



# Revista Colombiana de Hematología y Oncología

Biología molecular tumoral: de la básica a la clínica • Volumen 13 - Número 1  
Suplemento

REVISIÓN

## Bioinformática en cáncer

### Bioinformatics in cancer

»Leydi Tatiana Rojas-Ruco<sup>1,2</sup>



»Andrés Felipe Urrea-Orozco<sup>1,3</sup>

<sup>1</sup>Universidad del Valle, Cali, Colombia.

<sup>2</sup>Sección de Genética, Laboratorio de Genética Molecular Humana, Cali, Colombia.

<sup>3</sup>Sección de Genética, Grupo de Estudios Ecogenéticos y Biología Molecular, Cali, Colombia.

Recibido el 03 de septiembre de 2025. Aceptado el 15 de diciembre de 2025

<https://doi.org/10.51643/22562915.828>

### Resumen

**Introducción:** La bioinformática en el cáncer se ha convertido en una poderosa herramienta para la detección y el seguimiento de variantes moleculares asociadas a la enfermedad. Actualmente se manejan grandes cantidades de datos ómicos para desarrollar y aplicar nuevas herramientas que permitan analizar datos eficientemente. El objetivo de esta revisión fue describir el contexto en el que han surgido las nuevas tecnologías bioinformáticas y cómo estos avances están contribuyendo al estudio del cáncer.

**Métodos:** Se realizó una búsqueda en PubMed, Scopus, Google Scholar y ScienceDirect. Se expuso información basada en un conjunto de artículos disponibles en inglés y español, relacionados con el análisis de datos ( $n=66$ ). Además, se exponen las principales bases de datos y plataformas bioinformáticas para el estudio del cáncer.

**Resultados:** Los estudios revisados evidencian que la bioinformática actúa como un eje integrador en la investigación de la biología tumoral, facilitando la identificación de biomarcadores, clasificación molecular de tumores y estratificación de pacientes, mientras que el avance de la inteligencia artificial está revolucionando el análisis de datos.

**Discusión:** Las herramientas bioinformáticas han permitido una comprensión más profunda de los mecanismos moleculares del cáncer, apoyando el desarrollo de estrategias con potencial aplicación

\* Autor para correspondencia: Leydi Tatiana Rojas-Ruco, BSc. Sección de Genética, Laboratorio de Genética Molecular Humana, Cali, Colombia. Universidad del Valle, Cali, Colombia.

Dirección: Calle 13 # 100-00 Santiago de Cali, Valle del Cauca, Colombia.

Correo electrónico: rojas.leydi@correounalvalle.edu.co

<https://doi.org/10.51643/22562915.828>

Asociación Colombiana de Hematología y Oncología. Este es un artículo Open Access bajo la licencia CC BY-NC-ND. <http://creativecommons.org/licenses/by-nc-nd/4.0/>

clínica, así mismo este aumento de información requiere un control de calidad más estricto.

**Conclusión:** La bioinformática es una herramienta en expansión, necesaria en la investigación del cáncer, que puede facilitar la orientación adecuada en el diagnóstico y tratamiento, impactando la historia natural de la enfermedad.

**Palabras clave:** análisis de datos; biología computacional; biomarcadores; genómica; investigación biomédica; neoplasias.

## Abstract

**Introduction:** Bioinformatics in cancer has become a powerful tool for the detection and monitoring of molecular variants associated with the disease. Currently, large amounts of omics data are handled to develop and apply new tools that allow efficient data analysis. The aim of this review was to describe the context in which new bioinformatics technologies have emerged and how these advances are contributing to cancer research.

**Methods:** A search was conducted in PubMed, Scopus, Google Scholar, and ScienceDirect. Information was presented based on a set of articles available in English and Spanish related to data analysis ( $n=66$ ). In addition, the main databases and bioinformatics platforms for cancer research are presented.

**Results:** The reviewed studies show that bioinformatics acts as an integrative axis in tumor biology research, facilitating the identification of biomarkers, molecular classification of tumors, and patient stratification, while advances in artificial intelligence are revolutionizing data analysis.

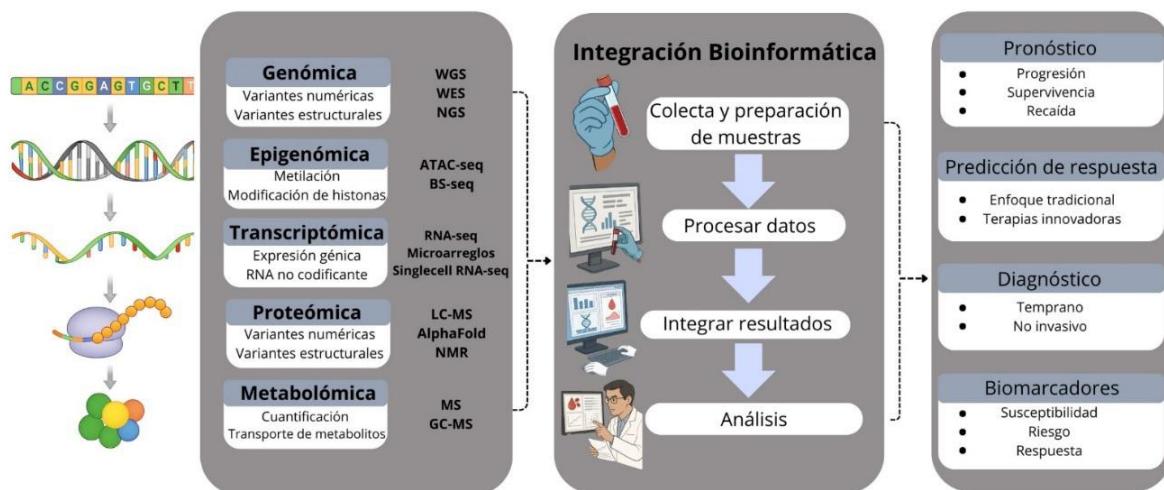
**Discussion:** Bioinformatics tools have enabled a deeper understanding of the molecular mechanisms of cancer, supporting the development of strategies with potential clinical application; likewise, this increase in information requires stricter quality control.

**Conclusion:** Bioinformatics is an expanding tool, necessary in cancer research, that can facilitate appropriate guidance in diagnosis and treatment, impacting the natural history of the disease.

**Keywords:** data analysis; computational biology; biomarkers; genomics; biomedical research; neoplasms.

## Resumen gráfico

### Bioinformática en el cáncer - Abstract Gráfico



#### Puntos clave

- Es una revisión narrativa donde se destaca el uso de la bioinformática desde la investigación básica a la clínica, centrándose en el rol diagnóstico y pronóstico de la bioinformática en el cáncer.
- Se enfatiza que, ante la heterogeneidad tumoral, el análisis multicapa permite identificar vías alteradas y posibles blancos terapéuticos; además, el creciente uso de la inteligencia artificial (IA) en la biomedicina está rompiendo barreras en el análisis de datos.
- Se resalta el impacto de los grandes consorcios como The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) y Pan-Cancer Analysis of Whole Genomes (PCAWG), que generaron redefiniciones moleculares de subtipos tumorales y hallazgos de mutaciones conductoras.
- El análisis computacional de los grandes datos ómicos mejora la identificación de alteraciones clave, el descubrimiento de biomarcadores accionables y la predicción de la respuesta, lo que permite estratificar a los pacientes y orientar los tratamientos.
- Los repositorios de datos se enriquecen continuamente, lo que permite su reutilización y establecer diferentes rutas de investigación para análisis integrales que articulen la investigación.

## Introducción

El cáncer es una de las principales causas de muerte en todo el mundo y ocupa una parte importante de la investigación biomédica actual; sin embargo, la detección de biomarcadores y el diagnóstico temprano han sido un desafío que ha requerido la integración de diversas áreas de la ciencia<sup>1</sup>. Esto ha hecho indispensable el uso de la bioinformática, una disciplina que se compone de estrategias computacionales, matemáticas y biológicas para analizar, visualizar y almacenar los enormes volúmenes de información generados por el creciente abordaje de las ciencias biológicas<sup>1,6</sup>. Los métodos de secuenciación genómica de alto rendimiento han enriquecido grandes bases de datos, que requieren estrategias para recopilar y organizar la información de manera coherente<sup>7</sup>. Además, la revolución de las ciencias ómicas (genómica, transcriptómica, epigenómica, proteómica y metabolómica) cambió el enfoque del estudio del cáncer al permitir caracterizar, de forma sistemática y a gran escala, las alteraciones moleculares que definen los tumores y sus subtipos clínicos<sup>8,9</sup>. Históricamente, el análisis molecular estaba limitado por la capacidad tecnológica; los estudios se centraban en genes aislados o en ensayos dirigidos. Actualmente, con la secuenciación masiva, la bioinformática dejó de ser una herramienta auxiliar para convertirse en el eje que conecta los datos moleculares con la clínica y la biología tumoral<sup>7,9</sup>. Asimismo, ha facilitado tanto la detección de biomarcadores como su validación en múltiples cohortes, así como la evaluación de su valor predictivo y la identificación de mecanismos biológicos subyacentes<sup>9</sup>. Lo que ha dado lugar a un cambio en la praxis clínica hacia una medicina de precisión más informada a nivel molecular. Este acercamiento integral entre la oncología y la biología ha permitido descomponer subpopulaciones celulares, rastrear clones tumorales y relacionar firmas moleculares con resistencias

terapéuticas o con el pronóstico; sin la bioinformática como área integradora, muchos de estos patrones habrían permanecido ocultos. El análisis computacional ha evolucionado de la mano de los avances tecnológicos en el estudio del cáncer<sup>7,9</sup>. El reconocimiento de la heterogeneidad tumoral ha orientado los análisis de datos hacia estrategias robustas y multicapa, con el fin de identificar vías biológicas frecuentemente alteradas en cada subtipo tumoral y genes blancos potenciales para terapias dirigidas. Por otro lado, la incorporación de la inteligencia artificial basada en aprendizaje automático a partir de datos heterogéneos y de gran volumen en la biología tumoral se perfila como un analizador de datos sistemático y rápido<sup>2</sup>. En consecuencia, la bioinformática proporciona las herramientas necesarias para almacenar, limpiar, integrar y analizar estas capas de información, y traducirlas en hipótesis biológicas aplicables, capaces de orientar el desarrollo de herramientas diagnósticas y terapéuticas más efectivas. En este trabajo se propuso revisar y sintetizar la contribución de la bioinformática al estudio del cáncer, desde las bases moleculares que pueden caracterizarse con métodos ómicos, pasando por la evolución de las estrategias de detección de biomarcadores, hasta el papel que juegan los grandes proyectos de datos y las técnicas de inteligencia artificial en la identificación de vías moleculares y genes blanco con relevancia terapéutica y pronóstica.

## Métodos

Esta revisión narrativa se basó en un proceso de búsqueda bibliográfica. El objetivo fue evaluar el papel que ha desempeñado la bioinformática en el cáncer, centrándose en su potencial diagnóstico y pronóstico, así como en las bases metodológicas de las principales plataformas bioinformáticas, bajo una lupa integral y actuali-

zada. La bioinformática es un área en constante crecimiento que abarca un amplio número de herramientas y bases de datos, así como una gran heterogeneidad en las formas de proceder. Por este motivo, las fases metodológicas de esta revisión narrativa (la selección bibliográfica, la extracción de datos y la síntesis) proporcionan una perspectiva más amplia.

### Búsqueda de literatura

La estrategia de búsqueda para identificar artículos elegibles se basó en el diseño experimental y conceptual de los estudios, limitando la búsqueda al período 2010-2025. Se incluyeron estudios publicados en inglés y en español. Se usaron diversas bases de datos utilizando las palabras clave “bioinformatics”, “cáncer driver genes”, “precision oncology”, “structural variation” y “omics”. Se consideraron artículos que exploran las diversas plataformas bioinformáticas y su papel en el diagnóstico y la detección de biomarcadores moleculares. La búsqueda se realizó en bases de datos como PubMed, Scopus, Google Scholar y ScienceDirect, que permitieron filtrar según el periodo de búsqueda indicado para obtener artículos con información actualizada y relevante en el área. Se consideraron estudios generales o específicos de acceso libre, así como revisiones narrativas y sistemáticas, ensayos clínicos y resúmenes de conferencias. La selección de los artículos se realizó mediante la revisión de títulos y resúmenes obtenidos a partir de la búsqueda y el filtro previo (n=66).

### Resultados

La oncología ha transitado de un abordaje terapéutico empírico a terapias enfocadas en la medicina de precisión. El conocimiento molecular del cáncer a partir de la investigación

clínica y molecular, en conjunto con la introducción de herramientas bioinformáticas, se integran para articular la investigación translacional que permite llevar a la práctica estos conocimientos<sup>3,4</sup>. La bioinformática en cáncer resulta fundamental para procesar e interpretar la cantidad de datos generados por las tecnologías modernas de secuenciación y perfilamiento molecular. Esto ha repercutido positivamente en la identificación de alteraciones genómicas clave en tumores individuales, además de favorecer el descubrimiento de biomarcadores accionables y la predicción de la respuesta a terapias específicas<sup>4,5</sup>. Diversos estudios y revisiones han subrayado que esta sinergia entre *big data* genómico y análisis computacional es la base para mejorar los resultados clínicos en oncología moderna, al permitir estratificar tumores y pacientes según sus vulnerabilidades moleculares específicas<sup>3,5</sup>.

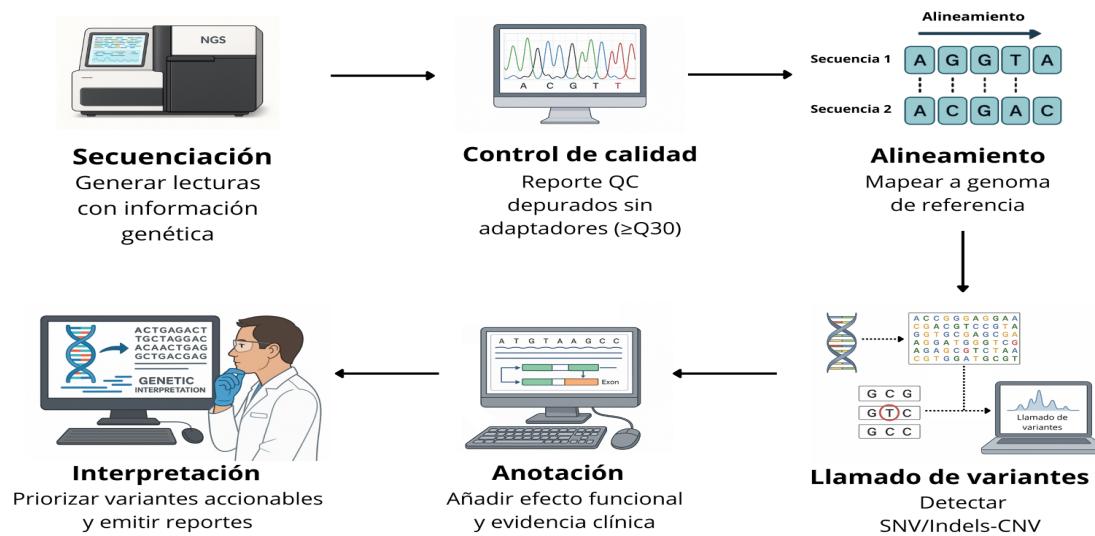
### Secuenciación

El progreso de la bioinformática en el cáncer ha venido de la mano de avances tecnológicos en la secuenciación de ADN y ARN, que han determinado la cantidad y la complejidad de los datos que deben utilizarse para un análisis completo. Para empezar, la secuenciación Sanger desarrollada en 1970 fue la técnica primordial durante décadas y permitió secuenciar el genoma humano por primera vez en el Proyecto Genoma Humano del 2003<sup>9</sup>. Sin embargo, el método Sanger, basado en terminadores dideoxinucleótidos y electroforesis capilar, es relativamente de baja productividad (~800 pb por corrida) y su aplicación al genoma completo requirió un esfuerzo internacional enorme. La llegada de la secuenciación de nueva generación (NGS, como sus siglas en inglés, Next-Generation Sequencing) marcó un punto de inflexión para la genómica del cáncer. Estas técnicas habilitaron el perfilamiento integral de genomas, exomas y transcriptomas tumorales, y con ello el desarrollo de pipelines

bioinformáticos estandarizados para el control de calidad, alineamiento, llamado y anotación de variantes (Figura 1)<sup>6</sup>.

**Figura 1.**

Flujo de trabajo para el análisis de secuencias.



estándar clínico e investigativo por medio de la amplificación por puente sobre celdas de flujo, el aumento de profundidad y cobertura de secuenciación, la baja tasa de error por sustitución y la operación masiva de secuencias, facilitaron análisis más complejos abriendo paso a las ciencias ómicas<sup>6,7</sup>. La secuenciación de tercera generación, como PacBio SMRT (Single Molecule, Real-Time), permite lecturas de longitudes mayores (~15–25 kb) con alta exactitud, y Oxford Nanopore ofrece lecturas ultralargas (~100 kb–1 Mb). Por lo tanto, ambas estrategias presentan mejoras sostenidas en la precisión y la longitud<sup>5</sup>. Estas tecnologías han mejorado la resolución estructural del genoma, facilitando la detección de reordenamientos complejos, inserciones/delecciones y haplotipos completos

que las lecturas cortas de NGS podrían fragmentar<sup>7</sup>.

Esta evolución tecnológica ha tenido un impacto directo en la investigación del cáncer, pues ha hecho posible que incluso laboratorios independientes secuencien genomas, ARN, células únicas o exomas tumorales con fines investigativos, permitiendo identificar variantes de resistencia emergentes o mutaciones germinales de predisposición en pacientes concretos<sup>6</sup>. Todo esto ha resultado en un crecimiento exponencial de los datos disponibles y, en consecuencia, en la necesidad de métodos bioinformáticos cada vez más eficientes y automatizados para almacenar, procesar e interpretar dichos datos<sup>6</sup>.

## Proyectos

Así, las tecnologías de secuenciación han dado lugar a proyectos internacionales de gran escala. Un precursor fue el proyecto del genoma humano (1990-2003), que proporcionó el primer genoma de referencia, facilitando las comparaciones posteriores entre genomas tumorales y sanos<sup>9</sup>. Tras este logro la biología molecular se volcó a la identificación de biomarcadores en tumores individuales a escala genómica, lo que potenció la detección de genes asociados con la oncogénesis y variantes somáticas recurrentes en ciertos tipos de cáncer, a su vez exigiendo avances experimentales como nuevas herramientas bioinformáticas para el ensamblaje, alineamiento y anotación de las variantes<sup>9,10</sup>.

Proyectos y estudios como los análisis integrales de cáncer de mama/colorrectal y glioblastoma han mostrado la eficacia de la secuenciación masiva combinada con análisis de expresión génica, para detectar montañas de genes que se encontraban frecuentemente mutados en ciertos tipos de cáncer (conductoras) y colinas de mutaciones raras (pasajeras) que, en conjunto, delinean el paisaje molecular de cada tumor<sup>12,13</sup>. La conceptualización moderna de mutaciones “conductoras” vs. “pasajeras” quedó sintetizada en revisiones clave, que estiman que un tumor típico acumula relativamente pocas mutaciones conductoras y muchas mutaciones pasajeras, sin efecto selectivo directo sobre el crecimiento tumoral<sup>10,11</sup>. Esta distinción motivó el desarrollo de métodos bioinformáticos para detectar señales de recurrencia de variantes en los datos somáticos y, de esta forma, priorizar candidatos de mutaciones conductoras<sup>14-16</sup>.

Algunas iniciativas de perfilamiento integral impulsaron la creación de grandes repositorios de datos genómicos del cáncer. Entre las bases de datos primarias se destaca The Cancer

Genome Atlas (TCGA) lanzado en 2006 por NCI, la cual marcó el inicio de la era de los “atlas” genómicos del cáncer. TCGA comenzó como un proyecto piloto en glioblastoma y cáncer de pulmón, pero eventualmente se extendió a decenas de tipos tumorales perfilando molecularmente más de 11.000 tumores de 33 tipos de cáncer distintos<sup>18</sup>. Los hallazgos de TCGA incluyeron la identificación de nuevas mutaciones conductoras, la redefinición molecular de subtipos y el descubrimiento de firmas moleculares en los tipos de cáncer estudiados. En paralelo, el International Cancer Genome Consortium (ICGC) coordinó proyectos en más de 15 países para secuenciar al menos 500 genomas completos de 50 tipos tumorales distintos, con énfasis en poblaciones diversas<sup>8</sup>. El ICGC, iniciado en 2010, complementó a TCGA extendiendo el alcance global del mapeo genómico del cáncer e integrando datos de regiones y etnias subrepresentadas<sup>8</sup>. Una colaboración entre TCGA-ICGC dio lugar al Pan-Cancer Analysis of Whole Genomes (PCAWG), publicado en 2020, que analizó el genoma completo de 2.658 tumores de 38 tipos diferentes<sup>17</sup>. Por su parte, PCAWG exploró regiones no codificantes y organizó las alteraciones estructurales a escala global. Entre sus hallazgos, se observó que cada tumor presenta, en promedio, 4 a 5 mutaciones conductoras, incluidas variantes estructurales mayores (como fusiones génicas o duplicaciones) con rol impulsor<sup>19</sup>. También identificó nuevos elementos no codificantes frecuentemente mutados e ilustró la ocurrencia de eventos catastróficos, como la cromotriplisis y la cromoplexia, que pueden impulsar la progresión tumoral<sup>19</sup>. Mientras que otro proyecto de alto alcance ha sido el Human Cell Atlas (HCA), iniciado en 2016, con la visión de mapear todos los tipos de células humanas a partir de sus perfiles moleculares individuales<sup>20</sup>. El HCA tiene repercusiones directas en el cáncer, pues sus avances en tecnologías y métodos de célula única han sido rápidamente aplicados en la investigación oncológica (Tabla

1). Los esfuerzos derivados están integrando la heterogeneidad intratumoral y la composición

de los microambientes tumorales en diversos tejidos<sup>21</sup>.

**Tabla 1.**

Síntesis de los principales proyectos de referencia para el análisis bioinformático del cáncer.

Proyecto	Avance	Impacto	Referencia
Human Genome Project (HGP)	Secuenciación del genoma de referencia; base para la anotación genómica.	Proporcionó el primer genoma de referencia, estandarizó los datos abiertos y habilitó los estudios genómicos en el cáncer.	Lander et al. 2001 <sup>9</sup>
The Cancer Genome Atlas (TCGA)	Perfiles multi-ómicos: WES/WGS, RNA-seq, metilación, CNV, clínica, imágenes.	Identificación de mutaciones conductoras, redefinición molecular de subtipos. Recurso estándar para el reanálisis.	Tomczak et al. 2013 <sup>18</sup>
International Cancer Genome Consortium (ICGC)	Estudios genómicos coordinados internacionalmente; WGS/WES y clínicos.	Amplió la representación poblacional de genomas tumorales y facilitó estudios comparativos internacionales.	Hudson et al. 2010 <sup>8</sup>
Pan-Cancer Analysis of Whole Genomes (PCAWG)	Análisis integrador de WGS.	Reveló el papel de las variantes estructurales y de las regiones no codificantes. Además de firmas mutacionales y procesos mutagénicos a escala.	Aaltonen et al. 2020 <sup>17</sup>
Human Cell Atlas (HCA)	Mapeo unicelular y espacial (scRNA-seq).	Referencias celulares que mejoran la interpretación de la heterogeneidad tumoral y del microambiente, con estándares unicelulares.	Regev et al. 2017 <sup>20</sup>

**Nota:** WES= whole exome sequencing, WGS= whole genome sequencing, CNV= copy number variants.

### Bases de datos

Las bases de datos se usan según corresponda al análisis que se realice. La información curada derivada de los proyectos se deposita en bases de datos secundarias o en plataformas de libre acceso, que facilitan la exploración e interpretación de los datos por investigadores de

todo el mundo. Una de las más utilizadas es cBioPortal for Cancer Genomics, una plataforma interactiva que permite consultar de forma integrada los perfiles genómicos de miles de tumores de TCGA y de otros estudios, y correlacionarse con datos clínicos<sup>22</sup>. Algunos archivos de secuencias disponibles permiten la reutilización de estudios publicados, como las

bases de datos ENA y SRA<sup>23,24</sup>. Por otro lado, los navegadores Ensembl, Genome Browser y los recursos de NCBI/GenBank proporcionan modelos génicos y bases de anotación con las respectivas isoformas que son necesarias para mapear las coordenadas genómicas<sup>25-27</sup>. Para priorizar variantes somáticas en cáncer y distinguir señales recurrentes, se utilizan sistemáticamente bases de datos especializadas, como COSMIC (variantes somáticas curadas) y ClinVar (interpretaciones clínicas), para la detección de biomarcadores<sup>28,29</sup>. La expresión génica y la validación mediante reanálisis se basan en repositorios como GEO<sup>30</sup>. La interpretación biológica de listados de genes y variantes se apoya en ontologías derivadas de plataformas como Reactome<sup>31</sup>, que representan vías y procesos biológicos y emplean un modelo de datos que funciona como una red de interacciones moleculares. Finalmente, para interpretar el posible efecto funcional de variantes a nivel de proteína, recursos como UniProt para evaluar estructuras experimentales y la base de predicciones de AlphaFold permiten realizar un modelado estructural que avale hipótesis mecanísticas<sup>36,37</sup>. Estas bases de datos y plataformas, en conjunto, permiten traducir una lista de variantes obtenidas mediante un abordaje biológico en hipótesis sobre rutas alteradas y posibles terapias efectivas, articulando la investigación traslacional.

### Ramas de estudio

Las ciencias ómicas estudian todos los componentes moleculares de un sistema biológico, y la bioinformática es fundamental para el análisis completo de estos datos ómicos<sup>54</sup>. Es por esto que las investigaciones oncológicas se han ramificado en diferentes caminos, que abordan distintos niveles de la complejidad tumoral<sup>17,32</sup>. Lo que se fundamenta en tres tipos de datos biológicos importantes: el ADN, el ARN y las secuencias de proteínas (Figura 2)<sup>19,33,34</sup>. A partir de este marco, los análisis genómicos

suelen ser el punto de partida natural<sup>16,17,35</sup>. Para realizar esto, se requieren herramientas y análisis más detallados; así surge el whole genome sequencing (WGS) o el whole exome sequencing (WES). La estrategia experimental típica consiste en: secuenciación, alineamiento, llamado de variantes, anotación funcional y priorización de mutaciones conductoras<sup>38,39</sup>. Los resultados se comparan mediante bases de datos para validar los patrones obtenidos en diferentes cohortes. Con estas herramientas se obtienen listas de variantes, como SNV/indels, que son variantes de uno o pocos nucleótidos, CNVs, que son variaciones en el número de copias, y SV/fusiones, que son variantes estructurales; y, entre ellas, se seleccionan candidatos accionables lo suficientemente robustos como para pasar a la verificación funcional<sup>40</sup>.

Otra rama de estudio es la transcriptómica, que articula las rutas, los estados transcripcionales activos y el contexto celular. Este campo de investigación estudia la suma de todas las moléculas de ARN<sup>41</sup>. Las plataformas que dominan este perfil son los microarreglos y el RNA-seq<sup>42</sup>, éstas demuestran que un evento genómico es operativo al evidenciar su transcripción y capturar firmas biológicas funcionales<sup>41,43</sup>. En el proceso de análisis transcriptómico, se alinean las lecturas, se normalizan los conteos y luego se prueban las hipótesis mediante el análisis de *fold change*. Algunos estudios transcriptómicos han permitido reconocer el dinamismo del ARN, dando lugar al reconocimiento de los microARN, pequeñas moléculas de ARN (20-22 pb) que pueden modular el control postranscripcional del ARNm y explicar por qué dos tumores con la misma variante pueden comportarse de manera distinta. Por ejemplo, el estudio de Lawrie destaca que en múltiples neoplasias se observaron miRNA supresores y oncogénicos simultáneamente, lo que destaca la importancia de la interacción biológica<sup>44,45</sup>. Por lo tanto, incorporar estos perfiles a los análisis bioinformáticos ayuda a clasificar pacien-

tes, anticipar la respuesta y proponer dianas reguladoras<sup>46-48</sup>. Por ello, su interpretación se vuelve más sólida cuando se integra con otras ciencias ómicas, como la epigenómica, para comprender cómo se habilitan o se bloquean los distintos mecanismos de transcripción<sup>49</sup>.

La epigenómica se caracteriza por estudiar las modificaciones que regulan la expresión génica sin alterar la secuencia, estos análisis revelan metilación del ADN, modificaciones de histonas y la accesibilidad de la cromatina<sup>50</sup>. En cáncer, el patrón típico es doble, Wilson y colaboradores señalaron que la hipometilación global desestabiliza el genoma y activa regiones repetitivas de genes de factores de crecimiento como el R-Ras y la familia de genes MAGE (melanoma-associated antigen genes)<sup>51</sup>, por otro lado, en una revisión Baylin explica que la hipermetilación local en promotores de genes supresores tales como p16, MLH1 y BRCA1 silencia y favorece la progresión tumoral<sup>52</sup>. En estos estudios la bioinformática resulta integrativa.

Otro de los enfoques existentes es la proteómica, que estudia de manera global el conjunto de proteínas y sus modificaciones posteriores a la traducción. Las proteínas contribuyen al desarrollo del cáncer debido a la activación aberrante de diversas vías de señalización proteica, por lo cual constituyen importantes dianas terapéuticas<sup>53,54</sup>. El trabajo de Slamon explica que cuando el oncogén ERBB2 está amplificado, la célula sobreexpresa la proteína HER2 en su superficie y en consecuencia, HER2 se autofosforila y activa rutas que impulsan proliferación y supervivencia celular<sup>55,56</sup>. Los avances más recientes en espectrometría de masas y en tecnologías de microarray de proteínas han situado a la proteómica como líder en la investigación oncológica. Estas tecnologías permiten cuantificar miles de proteínas y caracterizar redes proteicas tumorales. Esto, combinado con las plataformas

bioinformáticas, permite explorar las interacciones entre proteínas y mapear redes proteicas complejas. Además, la IA y el aprendizaje profundo han generado nuevas herramientas de predicción de proteínas, como AlphaFold, con las que es posible obtener información sobre características proteicas y su relación con posibles sitios de unión a fármacos<sup>37,54</sup>.

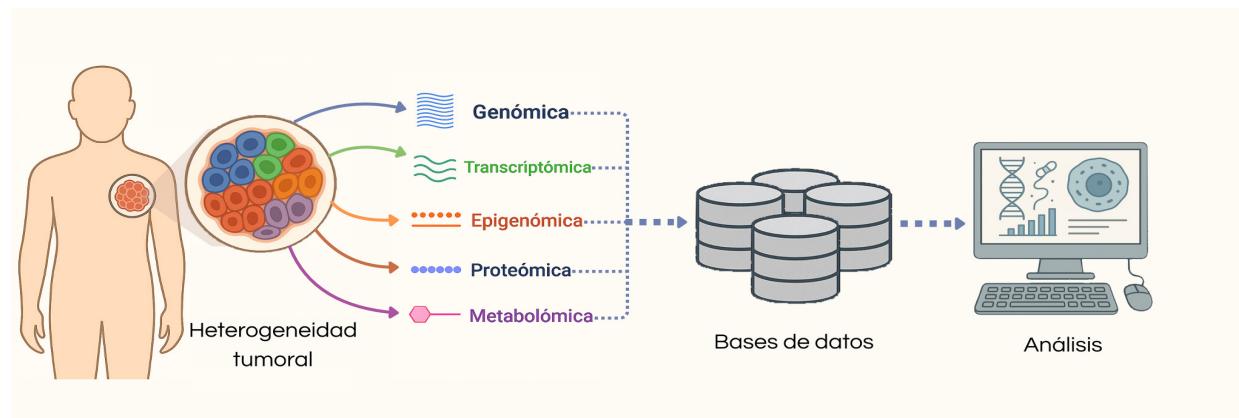
La metabolómica es el estudio y análisis científico de los metabolitos, que son el producto final de la actividad celular y, por tanto, reflejan directamente el estado funcional de las rutas y los fenotipos tumorales<sup>57</sup>. En el estudio metabolómico de Dang y colaboradores se identificó un oncometabolito denominado 2-hidroxiglutarato (2-HG) en tumores con alteraciones en IDH1/2, detectado mediante cromatografía líquida acoplada a espectrometría de masas (LC-MS); encontrando que la acumulación de 2-HG contribuye a la oncogénesis<sup>54,58</sup>. Otros estudios como el de Pollard y Selack han usado la técnica de cromatografía de gases acoplado a espectrometría de masas (GC-MS) para encontrar metabolitos que se acumulan en tumores y favorecen la metástasis y una mayor agresividad<sup>59,60</sup>. Además, los perfiles metabolómicos revelan dependencias de rutas bioquímicas que pueden focalizarse para ser utilizadas como dianas terapéuticas, como es el caso de la glutaminólisis cuando el oncogén MYC está activado<sup>61</sup>. En un estudio Wang y colaboradores observaron que los metabolitos circulantes pueden ser potenciales biomarcadores para la detección temprana y el diagnóstico del cáncer; además, señalaron que se han caracterizado metabolitos específicos con una gran capacidad para monitorear la progresión y la metástasis del cáncer<sup>62</sup>. Esto destaca la perfilación bioinformática de metabolitos presentes en tejidos tumorales o en fluidos mediante paneles metabólicos en biopsia líquida, proporcionando una firma metabólica del tumor<sup>63</sup>.

Sumado a esto, la inteligencia artificial ha tomado fuerza, actuando como una capa transversal que permite identificar patrones al integrar todas las ómicas<sup>64</sup>. Lo que resulta en un acoplamiento entre diferentes ramas medi-

ante la bioinformática convirtiendo los datos dispersos en conocimiento útil, contribuyendo al desarrollo de terapias dirigidas y de estrategias de tratamiento personalizadas.

**Figura 2.**

Flujo de datos de bioinformática a partir de muestras de tejido tumorales.



### Análisis multiómicos

Los estudios multiómicos combinan deliberadamente capas de información molecular con el objetivo de construir modelos integrados de la biología tumoral. Este enfoque va más allá de la adición de datos, pues permite validar hallazgos entre plataformas, reducir los falsos positivos y distinguir señales verdaderamente funcionales. Por ejemplo, en el cáncer de mama la clasificación de subtipos es crucial, así que en un estudio de Wang y colaboradores se aplicó una estrategia multiómica para tipificar tumores de mama, combinando datos de distintas plataformas<sup>65</sup>. Este enfoque reveló un nuevo subtipo molecular “híbrido” asociado a un pronóstico desfavorable. Los investigadores utilizaron biomarcadores derivados de datos de expresión diferencial entre subtipos para construir modelos pronósticos y clasificadores. Dichos modelos se validaron empleando

cohortes externas obtenidas de bases de datos públicas (GEO). Esta integración de datos de diferentes cohortes y tipos permite profundizar en la heterogeneidad del cáncer y en su traducción en terapias de precisión y pronóstico mejorado, usando métodos bioinformáticos que permitan articular datos de distintos orígenes<sup>65</sup>.

Otro estudio con un enfoque integrativo, donde basándose en la alta heterogeneidad del cáncer de pulmón, particularmente del adenocarcinoma de pulmón (LUAD), llevaron a cabo un análisis de 432 pacientes con LUAD de la base de datos de TCGA y 398 pacientes de los conjuntos de datos GEO<sup>66</sup>. Con el que integraron perfiles epigenéticos, transcriptómicos y clínicos. Luego, mediante algoritmos de aprendizaje automático, establecieron una nueva clasificación molecular a partir de estos perfiles y descubrieron subtipos basados en la regulación epigenética. Construyeron un modelo de

riesgo que predice el pronóstico, la infiltración inmune, la respuesta a la inmunoterapia y la sensibilidad a fármacos. Con lo cual lograron estratificar a pacientes con LUAD más allá de las mutaciones puntuales, lo cual puede orientar a nuevas formas de tratamiento en función de la firma epigenética-transcriptómica de cada tumor<sup>66</sup>. Estos estudios ilustran cómo la integración de capas ómicas ha permitido revelar relaciones causales y funcionales que serían difíciles de detectar con un único tipo de dato.

## Discusión

La bioinformática ha transformado profundamente el enfoque oncológico al integrar el estudio de grandes volúmenes de datos moleculares y clínicos, permitiendo avances sustanciales en la comprensión de la heterogeneidad del cáncer. Los estudios revisados destacan cómo la bioinformática no solo facilita la identificación de biomarcadores relevantes para diagnóstico y tratamiento<sup>13</sup>, sino que también optimiza el uso de tecnologías de secuenciación de alta capacidad, que ha impulsado la convergencia entre la biología molecular, las ciencias ómicas y el análisis computacional, consolidándose como un eje central de la oncología de precisión<sup>45</sup>. Todo esto, apoyándose en la evolución de las técnicas de secuenciación, desde Sanger hasta las tecnologías de NGS. Debido a que la capacidad de obtener datos masivos con alta resolución ha facilitado la identificación de mutaciones conductoras y pasajeras, redefiniendo la caracterización molecular de los tumores<sup>6,9</sup>.

En este contexto, proyectos como TCGA, ICGC y PCAWG han sido fundamentales para identificar patrones recurrentes de alteraciones genómicas mediante el análisis de grandes cohortes tumorales<sup>8,17,18</sup>. Estos consorcios han permitido

estratificar pacientes según su perfil molecular y mejorar la predicción de la respuesta a terapias específicas, fortaleciendo el vínculo entre la investigación genómica y la práctica clínica<sup>14</sup>. Los estudios de algunas neoplasias basándose en datos ómicos han mostrado cómo los enfoques integrativos de diferentes datos pueden revelar nuevos subtipos moleculares y mejorar la predicción pronóstica, apoyando a su vez el desarrollo de modelos terapéuticos personalizados<sup>8,65</sup>. Destacando el uso de plataformas y herramientas bioinformáticas para validación de datos, buscando mejorar la capacidad en la identificación de biomarcadores, dianas terapéuticas y la detección de firmas moleculares asociadas con la resistencia a terapias y la progresión tumoral<sup>9,30</sup>. A su vez, nuevas herramientas como la inteligencia artificial han tomado fuerza, mostrando que el aprendizaje automático de la IA puede ofrecer posibilidades en la medicina personalizada y la detección de patrones moleculares que han sido limitados con las herramientas tradicionales<sup>64</sup>. Acentuando la dirección integradora que ha tomado la investigación biomédica en el cáncer, con enfoques sistemáticos y predictivos, donde las nuevas posibilidades ofrecidas por la bioinformática pueden cambiar el curso natural de la enfermedad.

Adicionalmente, aunque la bioinformática ha demostrado un impacto decisivo en la investigación oncológica, su aplicación clínica enfrenta desafíos importantes. Entre ellos se encuentran el manejo y la calidad de los datos, la falta de estandarización entre plataformas, la reproducibilidad de los análisis y la necesidad de validación experimental de los hallazgos computacionales<sup>9</sup>. Asimismo, la integración multiómica requiere infraestructuras computacionales robustas y equipos interdisciplinarios capaces de interpretar los resultados en un contexto biológico y clínico. Abordar estas limitaciones será clave para consolidar el uso de la bioinformática como una herramienta rutinaria en la

oncología de precisión y maximizar su impacto en la toma de decisiones clínicas.

## Conclusiones

El cáncer representa un desafío continuo en la investigación biomédica, que requiere la acción conjunta de diversas áreas de la ciencia para lograr resultados eficientes en el estudio de la enfermedad. Para que los hallazgos ómicos lleguen a la clínica es imprescindible contar con

flujos de trabajo reproducibles; desde la calidad de las lecturas pasando por la anotación y priorización de variantes, para concretar en una interpretación clínica validada. Además, es fundamental consignar los datos en repositorios que permitan su reutilización y su validación independiente. Las herramientas bioinformáticas y el avance de la inteligencia artificial plantean formas novedosas y sistemáticas de análisis de datos, lo que pone sobre la mesa las limitaciones técnicas y éticas de estas tecnologías, destacando a su vez la importancia de la validación experimental.

### Financiamiento

Ninguno.

### Conflicto de interés

Los autores declaran no tener relaciones de interés comercial o personal en el marco de la investigación que condujo a la elaboración del manuscrito.

### Contribución de autoría

Ambos autores colaboraron desde la concepción y el diseño de la revisión hasta la redacción del manuscrito final y la revisión de su contenido.

## Referencias

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* [Internet]. 2021;71(3):209–49. Disponible en: <https://doi.org/10.3322/caac.21660>
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* [Internet]. 2019;25:44–56. Disponible en: <https://doi.org/10.1038/s41591-018-0300-7>
- Bode AM, Dong Z. Recent advances in precision oncology research. *NPJ Precis Oncol* [Internet]. 2018;2:11. Disponible en: <https://doi.org/10.1038/s41698-018-0055-0>
- Wang RC, Wang Z. Precision Medicine: Disease Subtyping and Tailored Treatment. *Cancers* [Internet]. 2023;15(15):3837. Disponible en: <https://doi.org/10.3390/cancers15153837>
- Molla G, Bitew M. Revolutionizing Personalized Medicine: Synergy with Multi-Omics Data Generation, Main Hurdles, and Future Perspectives. *Biomedicines* [Internet]. 2024;12(12):2750. Disponible en: <https://doi.org/10.3390/biomedicines12122750>
- Levy SE, Boone BE. Next-Generation sequencing Strategies. *Cold Spring Harb Perspect Med* [Internet]. 2018;9(7):a025791.

- Disponible en: <https://doi.org/10.1101/cshperspect.a025791>
7. Satam H, Joshi K, Mangrolia U, Waghoo S, Zaidi G, Rawool S, et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* [Internet]. 2023;12(7):997. Disponible en: <https://doi.org/10.3390/biology12070997>
  8. The International Cancer Genome Consortium. International network of cancer genome projects. *Nature* [Internet]. 2010;464(7291):993-8. Disponible en: <https://doi.org/10.1038/nature08987>
  9. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* [Internet]. 2001;409:860-921. Disponible en: <https://doi.org/10.1038/35087627>
  10. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. *Science* [Internet]. 2001;291(5507):1304-51. Disponible en: [https://doi.org/10.1016/s0002-9394\(01\)01077-7](https://doi.org/10.1016/s0002-9394(01)01077-7)
  11. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science* [Internet]. 2013;339(6127):1546-58. Disponible en: <https://doi.org/10.1126/science.1235122>
  12. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The Genomic Landscapes of Human Breast and Colorectal Cancers. *Science* [Internet]. 2007;318(5853):1108-13. Disponible en: <http://dx.doi.org/10.1126/science.1145720>
  13. Parsons DW, Jones S, Zhang X, Lin JCH, Leary RJ, Angenendt P, et al. An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science* [Internet]. 2008;321(5897):1807-12. Disponible en: <https://doi.org/10.1126/science.1164382>
  14. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* [Internet]. 2010;20(9):1297-303. Disponible en: <https://doi.org/10.1101/gr.107524.110>
  15. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* [Internet]. 2013;499(7457):214-8. Disponible en: <https://doi.org/10.1038/nature12213>
  16. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* [Internet]. 2018;173(2):371-385.e18. Disponible en: <https://doi.org/10.1016/j.cell.2018.02.060>
  17. Aaltonen LA, Abascal F, Abeshouse A, Aburatani H, Adams DJ, et al. Pan-cancer analysis of whole genomes. *Nature* [Internet]. 2020;578(7793):82-93. Disponible en: <https://doi.org/10.1038/s41586-020-1969-6>
  18. Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol* [Internet]. 2015;1A:68-77. Disponible en: <https://doi.org/10.5114/wo.2014.47136>

19. Cosenza MR, Rodriguez-Martin B, Korbel JO. Structural Variation in Cancer: Role, Prevalence, and Mechanisms. *Annu Rev Genom Hum Genet* [Internet]. 2022;23(1):123-52. Disponible en: <https://doi.org/10.1146/annurev-genom-120121-101149>
20. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. *eLife* [Internet]. 2017;6. Disponible en: <https://doi.org/10.7554/elife.27041>
21. Le J, Dian Y, Zhao D, Guo Z, Luo Z, Chen X, et al. Single-cell multi-omics in cancer immunotherapy: from tumor heterogeneity to personalized precision treatment. *Mol Cancer* [Internet]. 2025;24(1). Disponible en: <https://doi.org/10.1186/s12943-025-02426-3>
22. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* [Internet]. 2012;2(5):401-4. Disponible en: <https://doi.org/10.1158/2159-8290.cd-12-0095>
23. Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tarraga A, Cheng Y, et al. The European Nucleotide Archive. *Nucleic Acids Res* [Internet]. 2010;39:D28-31. Disponible en: <https://doi.org/10.1093/nar/gkq967>
24. Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res* [Internet]. 2010;39:D19-21. Disponible en: <https://doi.org/10.1093/nar/gkq1019>
25. Dyer SC, Austine-Orimoloye O, Azov AG, Barba M, Barnes I, Barrera-Enriquez VP, et al. Ensembl 2025. *Nucleic Acids Res* [Internet]. 2024;53(D1):D948-57. Disponible en: <https://doi.org/10.1093/nar/gkae1071>
26. Perez G, Barber GP, Benet-Pages A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2025 update. *Nucleic Acids Res* [Internet]. 2024;53(D1):D1243-9. Disponible en: <https://doi.org/10.1093/nar/gkae974>
27. O'Leary NA, Cox E, Holmes JB, Anderson WR, Falk R, Hem V, et al. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. *Sci. Data* [Internet]. 2024;11(1):732. Disponible en: <https://doi.org/10.1038/s41597-024-03571-y>
28. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* [Internet]. 2016;45(D1):D777-83. Disponible en: <https://doi.org/10.1093/nar/gkw1121>
29. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* [Internet]. 2014;42(1):D980-5. Disponible en: <https://doi.org/10.1093/nar/gkt1113>
30. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res* [Internet]. 2013;41:D991-5. Disponible en: <https://doi.org/10.1093/nar/gks1193>
31. Milacic M, Beavers D, Conley P, Gong C, Gillespie M, et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res* [Internet]. 2024. Disponible en: <https://doi.org/10.1093/nar/gkad1025>
32. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-Origin Patterns

- Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* [Internet]. 2018;173(2):291-304.e6. Disponible en: <https://doi.org/10.1016/j.cell.2018.03.022>
33. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol* [Internet]. 2017;18(1). Disponible en: <https://doi.org/10.1186/s13059-017-1215-1>
34. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* [Internet]. 2020;1(8):789-99. Disponible en: <https://doi.org/10.1038/s43018-020-0087-6>
35. Nair AS. Computational Biology & Bioinformatics: A Gentle Overview. Communications of the Computer Society of India [Internet]. 2007;1-12. Disponible en: [https://www.researchgate.net/publication/231337374\\_Computational\\_Biology\\_Bioinformatics\\_A\\_Gentle\\_Overview](https://www.researchgate.net/publication/231337374_Computational_Biology_Bioinformatics_A_Gentle_Overview)
36. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* [Internet]. 2022;51(D1):D523-31. Disponible en: <https://doi.org/10.1093/nar/gkac1052>
37. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* [Internet]. 2022;51(D1):D523-31. Disponible en: <https://doi.org/10.1093/nar/gkac1052>
38. Van Der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr Protoc Bioinformatics* [Internet]. 2013;43(1). Disponible en: <https://doi.org/10.1002/0471250953.bi110s43>
39. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods* [Internet]. 2013;10(11):1081-2. Disponible en: <https://doi.org/10.1038/nmeth.2642>
40. Koboldt DC. Best practices for variant calling in clinical sequencing. *Genome Med* [Internet]. 2020;12(1). Disponible en: <https://doi.org/10.1186/s13073-020-00791-w>
41. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, Jing X, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature* [Internet]. 2009;458(7234):97-101. Disponible en: <https://doi.org/10.1038/nature07638>
42. Kogenaru S, Yan Q, Guo Y, Wang N. RNA-seq and microarray complement each other in transcriptome profiling. *BMC Genomics* [Internet]. 2012;13(1). Disponible en: <https://doi.org/10.1186/1471-2164-13-629>
43. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* [Internet]. 2010;464(7289):773-7. Disponible en: <https://doi.org/10.1038/nature08903>
44. Lawrie CH. MicroRNA expression in lymphoma. *Expert Opin Biol Ther* [Internet]. 2007;7(9):1363-74. Disponible en: <https://doi.org/10.1517/14712598.7.9.1363>

45. Nelakurthi VM, Paul P, Reche A. Bioinformatics in Early Cancer Detection. *Cureus* [Internet]. 2023; Disponible en: <https://doi.org/10.7759/cureus.46931> doi.org/10.1016/j.bbcan.2006.08.007
46. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. *Nature* [Internet]. 2005;435(7043):834-8. Disponible en: <https://doi.org/10.1038/nature03702>
47. Van Roosbroeck K, Fanini F, Setoyama T, Ivan C, Rodriguez-Aguayo C, Fuentes-Mattei E, et al. Combining Anti-Mir-155 with Chemotherapy for the Treatment of Lung Cancers. *Clin Cancer Res* [Internet]. 2016;23(11):2891-904. Disponible en: <https://doi.org/10.1158/1078-0432.ccr-16-1025>
48. Buchberger E, Reis M, Lu TH, Posnien N. Cloudy with a Chance of Insights: Context Dependent Gene Regulation and Implications for Evolutionary Studies. *Genes* [Internet]. 2019;10(7):492. Disponible en: <https://doi.org/10.3390/genes10070492>
49. Xu M, Yuan L, Wang Y, Chen S, Zhang L, Zhang X. Integrative Analysis of DNA Methylation and Gene Expression Profiles Identifies Colorectal Cancer-Related Diagnostic Biomarkers. *Pathol Oncol Res* [Internet]. 2021;27. Disponible en: <https://doi.org/10.3389/pore.2021.1609784>
50. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis* [Internet]. 2009;31(1):27-36. Disponible en: <https://doi.org/10.1093/carcin/bgp220>
51. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochim Biophys Acta Rev Cancer* [Internet]. 2006;1775(1):138-62. Disponible en: <https://doi.org/10.1016/j.bbcan.2006.08.007>
52. Baylin SB. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* [Internet]. 2005;2(S1):S4-11. Disponible en: <https://doi.org/10.1038/ncponco354>
53. Manzoor Ahmad Mir, Hina Qayoom, Shazia Sofi, Nusrat Jan, Chapter 2 - Proteomics: A groundbreaking development in cancer biology, *Proteomics*, Academic Press, 2023, Pages 31-53, <https://doi.org/10.1016/B978-0-323-95072-5.00004-3>
54. Wolde T, Bhardwaj V, Pandey V. Current Bioinformatics Tools in Precision Oncology. *MedComm* [Internet]. 2025;6(7). Disponible en: <https://doi.org/10.1002/mco2.70243>
55. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2/neu Oncogene. *Science* [Internet]. 1987;235(4785):177-82. Disponible en: <https://doi.org/10.1126/science.3798106>
56. Swain SM, Shastry M, Hamilton E. Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Discov* [Internet]. 2022;22(2):101-26. Disponible en: <https://doi.org/10.1038/s41573-022-00579-0>
57. Qiu S, Cai Y, Yao H, Lin C, Xie Y, Tang S, et al. Small molecule metabolites: discovery of biomarkers and therapeutic targets. *Sig Transduct Target Ther* [Internet]. 2023;8(1). Disponible en: <https://doi.org/10.1038/s41392-023-01399-3>
58. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* [Internet].

- 2009;462(7274):739-44. Disponible en: <https://doi.org/10.1038/nature08617>
- Disponible en: <https://doi.org/10.1186/s40364-023-00507-3>
59. Selak MA, Armour SM, MacKenzie ED, Boulahbel H, Watson DG, Mansfield KD, et al. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF- $\alpha$  prolyl hydroxylase. *Cancer Cell* [Internet]. 2005;7(1):77-85. Disponible en: <https://doi.org/10.1016/j.ccr.2004.11.022>
60. Pollard PJ, Brière JJ, Alam NA, Barwell J, Barclay E, Wortham NC, et al. Accumulation of Krebs cycle intermediates and over-expression of HIF1 $\alpha$  in tumours which result from germline FH and SDH mutations. *Hum Mol Genet* [Internet]. 2005;14(15):2231-9. Disponible en: <https://doi.org/10.1093/hmg/ddi227>
61. Wise DR, DeBerardinis RJ, Mancuso A, Sayed N, Zhang XY, Pfeiffer HK, et al. Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc Natl Acad Sci U S A* [Internet]. 2008;105(48):18782-7. Disponible en: <https://doi.org/10.1073/pnas.0810199105>
62. Wang W, Rong Z, Wang G, Hou Y, Yang F, Qiu M. Cancer metabolites: promising biomarkers for cancer liquid biopsy. *Biomarker Res* [Internet]. 2023;11(1).
63. Wang G, Qiu M, Xing X, Zhou J, Yao H, Li M, et al. Lung cancer scRNA-seq and lipidomics reveal aberrant lipid metabolism for early-stage diagnosis. *Sci Transl Med* [Internet]. 2022;14(630). Disponible en: <https://doi.org/10.1126/scitranslmed.abk2756>
64. Ozaki Y, Broughton P, Abdollahi H, Valafar H, Blenda AV. Integrating Omics Data and AI for Cancer Diagnosis and Prognosis. *Cancers* [Internet]. 2024;16(13):2448. Disponible en: <https://doi.org/10.3390/cancers16132448>
65. Wang ZZ, Li XH, Wen XL, Wang N, Guo Y, Zhu X, et al. Integration of multi-omics data reveals a novel hybrid breast cancer subtype and its biomarkers. *Front Oncol* [Internet]. 2023;13. Disponible en: <https://doi.org/10.3389/fonc.2023.1130092>
66. Wang N, Li Y, Wang Y, Wang W. Integration of multi-omics profiling reveals an epigenetic-based molecular classification of lung adenocarcinoma: implications for drug sensitivity and immunotherapy response prediction. *Front Pharmacol* [Internet]. 2025;16. Disponible en: <https://doi.org/10.3389/fphar.2025.1540477>