



Universidad Internacional de La Rioja
Facultad de Ciencias de la Salud

Máster Universitario en Bioinformática

**Análisis bioinformático de variantes
genéticas asociadas a la anemia falciforme
mediante herramientas de predicción y
anotación de variantes**

Trabajo fin de Estudio presentado por:	Daniela Angie Simo Arias
Tipo de trabajo:	Propuesta Proyecto de investigación
Director/a:	Sergio Buenestado Serrano
Fecha:	16/07/2025

Resumen

La anemia falciforme (AF) es una hemoglobinopatía monogénica causada por una mutación puntual en el gen *HBB*. Aunque se origina por un único cambio genético, su expresión clínica varía ampliamente entre individuos debido a factores modificadores genéticos y ambientales. Este Trabajo Fin de Máster propone un estudio observacional bioinformático para identificar variantes genéticas asociadas a la AF y su relación con la severidad clínica. Utilizando datos simulados de pacientes y controles, se aplican tecnologías de secuenciación masiva (NGS) y herramientas como SIFT, PolyPhen-2, VEP o CADD para analizar variantes en genes como *HBB*, *BCL11A*, *HBS1L-MYB* y *KLF1*. El objetivo es priorizar mutaciones con impacto funcional y avanzar hacia una medicina personalizada. El estudio también considera aspectos éticos, criterios de inclusión/exclusión, recolección de datos clínicos, análisis estadístico e integración de resultados. Se espera identificar biomarcadores genéticos útiles para estratificación de pacientes y diseño de terapias avanzadas como la edición génica.

Palabras clave: Anemia falciforme, variantes genéticas y bioinformática

Abstract

Sickle cell anemia (SCA) is a monogenic hemoglobinopathy caused by a single point mutation in the *HBB* gene. Despite its monogenic origin, clinical expression varies significantly among individuals due to the influence of genetic modifiers and environmental factors. This Master's Thesis proposes a bioinformatic observational study aimed at identifying genetic variants associated with SCA and their correlation with clinical severity. Using simulated data from patients and healthy controls, the study applies next-generation sequencing (NGS) technologies and tools such as SIFT, PolyPhen-2, VEP, and CADD to analyze variants in genes including *HBB*, *BCL11A*, *HBS1L-MYB*, and *KLF1*. The goal is to prioritize functionally relevant mutations and support the development of personalized medicine. The research design includes ethical considerations, inclusion/exclusion criteria, clinical data collection, statistical analysis, and result integration. Expected outcomes include identifying relevant genetic biomarkers for patient stratification and contributing to the design of advanced therapies such as gene editing.

Keywords: Sickle cell anemia, genetic variants and bioinformatics

Índice de contenidos

1.	Introducción	8
1.1.	Objetivos	10
2.	Marco teórico	12
2.1.	Anemia Falciforme.....	12
2.2.	Fisiopatología de la anemia falciforme.....	13
2.3.	Genética y variantes asociadas a la anemia falciforme.....	14
2.4.	Aplicaciones de la bioinformática en el estudio de la anemia falciforme.....	17
2.5.	Epidemiología y situación global y en España	19
2.6.	Perspectivas futuras y avances en investigación	21
2.6.1.	Edición génica in vivo y nuevas tecnologías.....	21
2.6.2.	Desarrollo de nuevos fármacos y terapias combinadas	21
2.6.3.	Medicina personalizada y análisis de datos	22
2.7.	Aplicaciones de la bioinformática en la anemia falciforme.....	22
2.8.	Herramientas bioinformáticas específicas aplicadas a la anemia falciforme.....	24
2.8.1.	SIFT y PolyPhen-2: Predicción de impacto funcional.....	24
2.8.2.	Ensembl Variant Effect Predictor (VEP)	24
2.8.3.	UCSC Genome Browser y IGV	24
2.8.4.	Herramientas de diseño CRISPR	25
2.8.5.	Docking molecular y predicción estructural	25
2.8.6.	Plataformas de análisis de expresión génica	25
2.8.7.	Integración en plataformas multiómicas	25
3.	Marco metodológico.....	27
3.1.	Tipo de estudio.....	27
3.2.	Tamaño muestral.....	27

3.2.1.	Criterios de inclusión:	28
3.2.2.	Criterios de exclusión:.....	28
3.3.	Consideraciones éticas	29
3.4.	Variables del estudio	29
3.5.	Extracción de muestras y recolección de datos clínicos	31
3.5.1.	Reclutamiento y contacto inicial	31
3.5.2.	Recolección de datos clínicos	31
3.5.3.	Extracción de muestras biológicas.....	32
3.5.4.	Número de pacientes y análisis.....	32
3.6.	Procesamiento bioinformático de las variantes genéticas.....	33
3.6.1.	Extracción y secuenciación del ADN	33
3.6.2.	Control de calidad y alineamiento.....	33
3.6.3.	Llamada de variantes y anotación	33
3.6.4.	Herramientas de predicción de patogenicidad	34
3.6.5.	Interpretación y categorización de variantes.....	34
3.7.	Análisis estadístico.....	35
3.8.	Integración e interpretación de resultados	36
3.9.	Cronograma.....	37
3.10.	Presupuesto.....	37
4.	Resultados esperados	39
5.	Limitaciones y propuestas futuras.....	41
6.	Bibliografía.....	43
	Índice de acrónimos	56

Índice de figuras

Figura 1. Extensión de sangre periférica para encontrar células falciformes. Extraído de “Caracterización geométrica euclidiana y fractal de células falciformes” (4)	8
Figura 2. Distribución de los niveles de hemoglobina fetal según genotipos de SNP. Fuente Pubmed “DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease” Extraído del autor Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, et al. (32)	15
Figura 3. Cambios morfológicos durante el aumento de la dosis de hidroxíurea. Fuente Pubmed “How I use hydroxyurea to treat young patients with sickle cell anemia” Extraído del autor Ware RE (39).	17
Figura 4. Mapa de nacimientos, prevalencia y mortalidad de anemia falciforme. Extraído de “Prevalencia y carga de mortalidad mundial, regional y nacional de la anemia de células falciformes, 2000-2021” (57).....	20
Figura 5. Cronograma de proyecto de investigación. Elaboración propia con Python.	37

Índice de tablas

Tabla 1. Presupuesto detallado del estudio durante 5 años	37
---	----

1. Introducción

La anemia falciforme (AF) es una hemoglobinopatía hereditaria de transmisión autosómica recesiva que afecta a millones de personas a nivel mundial, especialmente en regiones donde la malaria ha sido endémica, como África subsahariana, Medio Oriente, la India, el Caribe y América del Sur (1). Esta enfermedad monogénica es causada por una mutación puntual en el gen *HBB*, responsable de codificar la cadena beta de la hemoglobina. En particular, dicha mutación produce la sustitución del ácido glutámico por una valina en la sexta posición de la cadena (Glu6Val), originando la hemoglobina S (HbS) (2).

La presencia de HbS provoca que, bajo condiciones de desoxigenación, las moléculas de hemoglobina se polimericen, lo cual distorsiona la forma de los eritrocitos, generando células con morfología en hoz. Estas células presentan menor deformabilidad, mayor fragilidad y una fuerte tendencia a la adhesión endotelial, favoreciendo fenómenos de vaso-oclusión y hemólisis intravascular. Como consecuencia, los pacientes pueden presentar dolor crónico, anemia, accidentes cerebrovasculares, daño multiorgánico, infecciones recurrentes y una reducción significativa de la esperanza de vida (3).

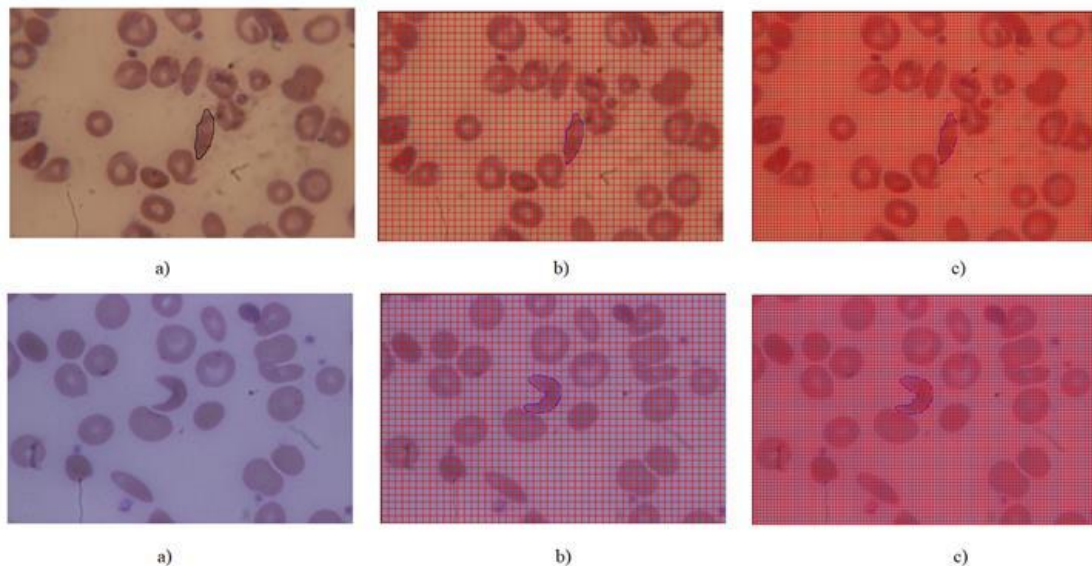


Figura 1. Extensión de sangre periférica para encontrar células falciformes. Extraído de “Caracterización geométrica euclidiana y fractal de células falciformes” (4)

Aunque la AF es producto de una única mutación puntual, la manifestación clínica de la enfermedad varía considerablemente entre los individuos. Esta heterogeneidad fenotípica se

debe tanto a factores ambientales como a la influencia de variantes genéticas adicionales, conocidas como genes modificadores. Algunas de estas variantes modulan la expresión de hemoglobina fetal (HbF), cuya mayor presencia inhibe la polimerización de la HbS y atenúa los síntomas (5). Otros genes modificadores están implicados en vías relacionadas con la inflamación, el estrés oxidativo o la adhesión celular.

La Organización Mundial de la Salud (OMS) reconoce a la AF como una de las enfermedades monogénicas más prevalentes, con más de 300.000 nacimientos afectados anualmente (6). En países en desarrollo, el limitado acceso a diagnóstico y tratamiento intensifica la carga de morbilidad. En contextos desarrollados, si bien terapias como la hidroxiurea, las transfusiones regulares y el trasplante de médula ósea han mejorado el pronóstico, persisten retos clínicos considerables.

En España, la prevalencia de la AF ha aumentado progresivamente como resultado de la inmigración desde regiones de alta incidencia. Se estima que más de mil personas padecen esta enfermedad en el territorio nacional, aunque la cifra real podría ser mayor debido a la baja notificación (7). Esta situación evidencia la necesidad de herramientas avanzadas que permitan caracterizar el perfil genético de los pacientes y optimizar las estrategias terapéuticas.

El análisis de variantes genéticas mediante herramientas bioinformáticas constituye una estrategia clave para identificar los factores que modulan la severidad clínica de la AF. Las tecnologías de secuenciación masiva (NGS) generan volúmenes de datos genómicos considerables, por lo que resulta fundamental emplear algoritmos de predicción que prioricen las variantes con mayor probabilidad de ser patogénicas. Herramientas como SIFT, PolyPhen-2, MutationTaster y CADD permiten evaluar el impacto funcional de las sustituciones de aminoácidos, considerando parámetros como la conservación evolutiva, la estructura proteica y la función bioquímica (8).

La integración de datos con bases como dbSNP, ClinVar, gnomAD y OMIM facilita la validación de variantes mediante la comparación con registros previos y frecuencias poblacionales, permitiendo diferenciar polimorfismos comunes de variantes raras potencialmente patogénicas (9). Este enfoque es especialmente relevante en enfermedades como la AF, donde la interacción de múltiples variantes puede influir de forma significativa en el fenotipo clínico.

Este Trabajo Fin de Máster se plantea como una propuesta de investigación que simula un ensayo clínico prospectivo con el objetivo de identificar y priorizar variantes genéticas implicadas en la anemia falciforme utilizando herramientas bioinformáticas. A partir de un conjunto de datos simulados que representan pacientes con AF y controles sanos, se llevará a cabo un análisis del genoma enfocado en genes relevantes como *HBB*, *BCL11A*, *HBS1L-MYB*, *KLF1*, entre otros.

La elección de este tema responde a una combinación de intereses personales, formativos y científicos. En primer lugar, la anemia falciforme representa un modelo paradigmático de enfermedad monogénica cuya complejidad clínica trasciende la simple presencia de una mutación puntual, lo que permite abordar el análisis genético desde una perspectiva integral. Desde el punto de vista académico, este trabajo permite aplicar y consolidar los conocimientos adquiridos durante el máster en áreas como la genómica, el análisis de variantes y la interpretación bioinformática de datos clínicos. Además, el uso de herramientas de predicción y anotación de variantes en enfermedades con gran impacto sanitario y social permite no solo profundizar en aspectos técnicos, sino también contribuir, aunque sea desde un enfoque teórico, a la mejora del diagnóstico y la atención personalizada. Finalmente, la creciente importancia de la medicina genómica y de precisión convierte este tipo de estudios en una vía prometedora para avanzar en la estratificación de pacientes y en la identificación de nuevos biomarcadores clínicamente relevantes.

1.1.Objetivos

Aunque la anemia falciforme se origina por una mutación puntual en el gen *HBB*, la variabilidad clínica observada entre pacientes se debe en parte a la presencia de variantes adicionales en genes modificadores. Estas variantes, detectables mediante técnicas de secuenciación de nueva generación (*NGS*) y análisis *in silico*, presentan correlaciones significativas con parámetros clínicos como la frecuencia de crisis vasooclusivas, los niveles de hemoglobina fetal (*HbF*), la tasa de hospitalización o la respuesta terapéutica a la hidroxiurea. Por tanto, se espera que determinados perfiles genéticos se asocien con un curso clínico más benigno, permitiendo establecer un modelo preliminar de estratificación genética del riesgo en pacientes con anemia falciforme.

Objetivo principal: Identificar y caracterizar variantes genéticas asociadas a la anemia falciforme mediante técnicas de secuenciación de nueva generación (NGS) y análisis *in silico*. Además de evaluar el impacto funcional de variantes en genes como *HBB*, *BCL11A*, *HBS1L-MYB* o *KLF1*, se integrarán modelos estadísticos para analizar la asociación entre el perfil genético y características clínicas relevantes, como la frecuencia de crisis vasooclusivas, el número de hospitalizaciones anuales, los niveles de HbF y la respuesta al tratamiento con hidroxiurea. Se busca avanzar hacia un modelo preliminar de estratificación genética del riesgo clínico en pacientes con anemia falciforme.

Objetivos específicos:

- Procesar y analizar datos genómicos de pacientes con anemia falciforme y controles sanos mediante técnicas de NGS, realizando control de calidad, alineamiento, llamada de variantes y anotación funcional con herramientas bioinformáticas especializadas.
- Priorizar variantes genéticas potencialmente patogénicas en genes candidatos mediante predicción *in silico* (SIFT, PolyPhen-2, CADD, VEP), considerando su impacto estructural, funcional y su frecuencia en bases de datos clínicas y poblacionales.
- Correlacionar variantes genéticas con variables clínicas, como la frecuencia de crisis vasooclusivas, número de hospitalizaciones, niveles de hemoglobina fetal (HbF) y respuesta terapéutica, utilizando modelos estadísticos multivariantes para identificar asociaciones significativas.
- Proponer un modelo preliminar de estratificación genética del riesgo clínico, basado en los perfiles genéticos identificados y su relación con la severidad del fenotipo, que contribuya a sentar las bases de un abordaje más personalizado en el manejo de la anemia falciforme.

2. Marco teórico

2.1. Anemia Falciforme

La anemia falciforme (AF), también conocida como drepanocitosis, es una hemoglobinopatía hereditaria de carácter autosómico recesivo que afecta a millones de personas en todo el mundo. Esta enfermedad se caracteriza por la presencia de hemoglobina S (HbS), una variante estructural de la hemoglobina normal, que provoca la deformación de los eritrocitos en forma de hoz bajo condiciones de desoxigenación (10). Esta alteración morfológica conlleva una serie de complicaciones clínicas, incluyendo anemia hemolítica crónica, episodios de dolor agudo, infecciones recurrentes y daño multiorgánico progresivo (11).

La mutación responsable de la AF consiste en una sustitución de adenina por timina en el codón 6 del gen **HBB**, que codifica la cadena beta de la hemoglobina. Esta mutación da lugar a la sustitución del aminoácido ácido glutámico por valina, lo que altera las propiedades fisicoquímicas de la hemoglobina y favorece su polimerización en condiciones de baja tensión de oxígeno (12). La polimerización de la HbS induce la rigidez y deformación de los eritrocitos, lo que compromete su capacidad para transitar por la microvasculatura y facilita su destrucción prematura (13).

La prevalencia de la AF es particularmente alta en regiones donde la malaria ha sido endémica, como África subsahariana, el Mediterráneo, Medio Oriente y la India. Esta distribución geográfica se explica por la **ventaja heterocigota**, ya que los portadores del rasgo falciforme (heterocigotos para la mutación) presentan una mayor resistencia a la infección por *Plasmodium falciparum*, el agente etiológico más virulento de la malaria (14). A nivel mundial, se estima que nacen aproximadamente 300,000 niños con AF cada año, siendo la mayoría en África (12).

En las últimas décadas, los avances en el diagnóstico y tratamiento de la AF han mejorado significativamente la calidad de vida y la esperanza de vida de los pacientes en países desarrollados. Sin embargo, en muchas regiones de bajos recursos, el acceso limitado a servicios de salud y terapias efectivas sigue siendo un desafío importante (10). Además, la expresión clínica de la enfermedad es altamente variable entre individuos, incluso entre aquellos con genotipos idénticos, lo que sugiere la influencia de factores genéticos modificadores y ambientales (15).

En este contexto, la bioinformática emerge como una herramienta poderosa para comprender la complejidad genética de la AF y desarrollar estrategias terapéuticas personalizadas. La integración de datos genómicos, transcriptómicos y clínicos mediante enfoques bioinformáticos permite identificar variantes genéticas asociadas con la severidad de la enfermedad, predecir la respuesta a tratamientos y descubrir nuevas dianas terapéuticas (14). Herramientas como **SIFT**, **PolyPhen-2**, **MutationTaster** y **CADD** se utilizan para evaluar el impacto funcional de las variantes genéticas, mientras que bases de datos como **dbSNP**, **ClinVar**, **gnomAD** y **OMIM** proporcionan información sobre la frecuencia y patogenicidad de dichas variantes (13).

2.2. Fisiopatología de la anemia falciforme

La fisiopatología de la anemia falciforme (AF) se origina a partir de una única mutación puntual en el gen **HBB**, localizado en el cromosoma 11, que da lugar a la sustitución del aminoácido ácido glutámico por valina en la posición 6 de la cadena beta de la hemoglobina (HbA) (16). Esta mutación provoca la síntesis de la **hemoglobina S (HbS)**, la cual tiene una tendencia anómala a polimerizarse cuando la tensión de oxígeno es baja, generando la deformación de los eritrocitos en forma de hoz o "falciforme" (17).

Bajo condiciones normales, la hemoglobina mantiene su conformación funcional durante el transporte de oxígeno. Sin embargo, la HbS desoxigenada se agrega en polímeros largos e insolubles que distorsionan la membrana celular del eritrocito, reducen su flexibilidad y promueven su fragilidad (18). Estas células deformadas tienen una vida media significativamente reducida (aproximadamente 10–20 días frente a los 120 días de un eritrocito normal), lo que conduce a una **anemia hemolítica crónica** (19).

Además de la hemólisis, la rigidez de los eritrocitos falciformes ocasiona episodios de **vaso-oclusión**, un fenómeno patognomónico de la enfermedad, donde los eritrocitos deformados se adhieren al endotelio vascular y bloquean el flujo sanguíneo capilar y postcapilar (20). Esta oclusión microvascular desencadena **episodios de isquemia y necrosis tisular**, responsables de los característicos **dolores agudos o "crisis falciformes"**, que pueden afectar a diversos órganos y tejidos (21).

La vaso-oclusión también está asociada a un estado inflamatorio crónico, promovido por la activación del endotelio vascular, la adhesión de leucocitos, la generación de especies

reactivas de oxígeno (ROS) y la liberación de citoquinas proinflamatorias como TNF- α e IL-6 (22). Este entorno inflamatorio contribuye al daño multiorgánico progresivo, afectando especialmente a órganos como el bazo, riñones, cerebro, pulmón y retina (23).

La **disfunción endotelial** y la activación de las plaquetas también participan en la fisiopatología de la enfermedad. Se ha observado un estado protrombótico en los pacientes con AF, con niveles elevados de trombina y fibrinógeno, lo que incrementa el riesgo de **accidentes cerebrovasculares** y trombosis venosa profunda, incluso en edades tempranas (24). De hecho, el infarto cerebral silente o clínico es una complicación frecuente en la infancia, con consecuencias neurocognitivas a largo plazo (25).

Otro aspecto relevante es la **hiperhemólisis**, que libera hemoglobina libre en el plasma, la cual capta óxido nítrico (NO), una molécula clave para la vasodilatación y la homeostasis vascular. La disminución de NO bioactivo contribuye a la **disfunción vascular** y al aumento de la presión arterial pulmonar, una complicación grave de mal pronóstico en la población adulta con AF (26).

La interacción de todos estos mecanismos (eritropoyesis ineficaz, hemólisis crónica, vaso-oclusión, inflamación y trombosis) conforman el complejo cuadro clínico y evolutivo de la anemia falciforme. La variabilidad en la expresión fenotípica de la enfermedad entre pacientes, incluso con el mismo genotipo, se atribuye a la influencia de **modificadores genéticos** como los niveles de hemoglobina fetal (HbF) y la coinherencia de otras hemoglobinopatías como la talasemia alfa (27).

Entender en profundidad estos mecanismos es fundamental para el desarrollo de estrategias terapéuticas eficaces, que van desde el uso de hidroxurea (capaz de incrementar los niveles de HbF) hasta intervenciones más recientes como la terapia génica, el trasplante de células madre hematopoyéticas y la edición genómica (28).

2.3. Genética y variantes asociadas a la anemia falciforme

La **anemia falciforme (AF)** es una hemoglobinopatía monogénica causada por una mutación puntual en el gen **HBB**, que codifica la cadena beta de la hemoglobina adulta. La mutación más frecuente consiste en un cambio de **adenina (A) por timina (T)** en el sexto codón del gen (GAG $\rightarrow$ GTG), lo que provoca la sustitución del **ácido glutámico por valina** en la posición 6 de la cadena beta-globina, originando la **hemoglobina S (HbS)** (29).

Esta variante HbS puede existir en **homocigosis (HbSS)**, que produce el fenotipo más severo de la enfermedad, o en **heterocigosis compuesta**, como ocurre en la coinherencia con otras hemoglobinopatías (HbSC, HbS/ β -talasemia), lo que da lugar a manifestaciones clínicas variables (30). La HbSC, por ejemplo, implica la herencia simultánea del alelo de HbS y del alelo de hemoglobina C (una sustitución de ácido glutámico por lisina en la misma posición), lo que resulta en una forma moderada de la enfermedad (31).

Además de estas variantes principales, estudios genéticos han identificado múltiples **modificadores genéticos** que influyen en el fenotipo clínico de la AF. Entre ellos destaca el nivel de **hemoglobina fetal (HbF)**, cuya persistencia en la edad adulta puede atenuar la gravedad de los síntomas, al interferir con la polimerización de HbS (32). Genes como **BCL11A**, **HBS1L-MYB** y el propio **locus β -globina** son determinantes clave en la regulación de la expresión de HbF (33). La presencia de ciertos polimorfismos de un solo nucleótido (**SNPs**) en estos genes se ha asociado a niveles más altos de HbF y mejor pronóstico clínico (34).

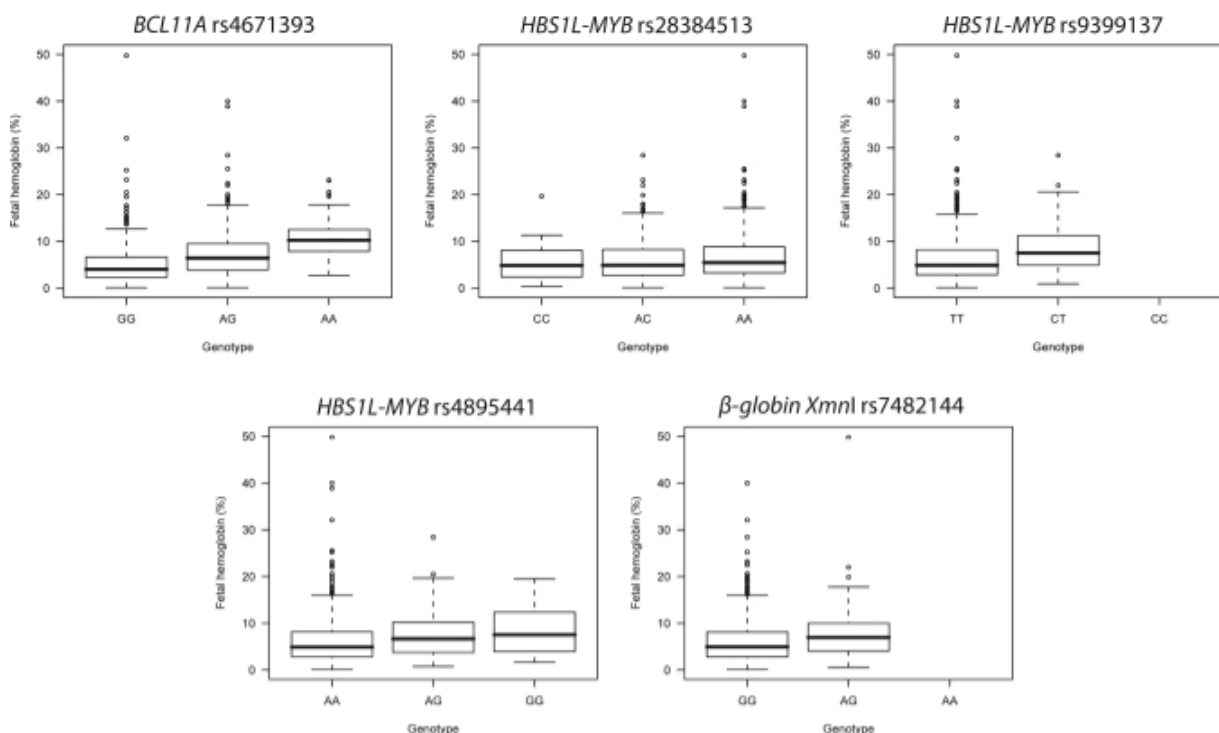


Figura 2. Distribución de los niveles de hemoglobina fetal según genotipos de SNP. Fuente Pubmed “DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease” Extraído del autor Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, et al. (32)

Asimismo, la coinherencia de α -talasemia, común en regiones endémicas, puede modificar el curso clínico de la enfermedad. La reducción en la síntesis de cadenas α disminuye la formación de HbS, lo que se traduce en una menor hemólisis y una mayor concentración de hematíes circulantes, aunque puede exacerbar la viscosidad sanguínea y la frecuencia de eventos vasooclusivos (35).

Los avances en la genómica han permitido también identificar variantes en genes relacionados con procesos inflamatorios, trombóticos y oxidativos, que podrían actuar como modificadores adicionales del fenotipo. Por ejemplo, variantes en genes como *TNF- α* , *IL-10*, *VCAM1* o *NOS3* han sido asociadas a un mayor riesgo de complicaciones como crisis dolorosas, síndrome torácico agudo o accidentes cerebrovasculares (36).

Las tecnologías de secuenciación masiva (NGS) han sido fundamentales para caracterizar estas variantes en estudios de asociación del genoma completo (GWAS) y estudios funcionales. Gracias a ellas, se ha podido describir la heterogeneidad genética de la enfermedad a nivel poblacional, especialmente en contextos africanos, afrodescendientes, asiáticos y mediterráneos (37). Por ejemplo, diferentes haplotipos del gen *β -globina* (Benin, Senegal, Camerún, Bantu, árabe-India) se han asociado con distinta severidad de la enfermedad y con niveles variables de HbF (38).

Desde una perspectiva de salud pública, el conocimiento de estas variantes ha permitido mejorar el cribado neonatal, la consejería genética y el diseño de estrategias personalizadas de tratamiento, como el uso de hidroxiaurea en pacientes con bajo nivel de HbF o el desarrollo de terapias génicas dirigidas (39). Además, la identificación de mutaciones en regiones reguladoras del genoma puede abrir nuevas vías terapéuticas mediante edición génica con CRISPR/Cas9, ya que algunas estrategias actuales buscan silenciar genes supresores de HbF como *BCL11A* (40).

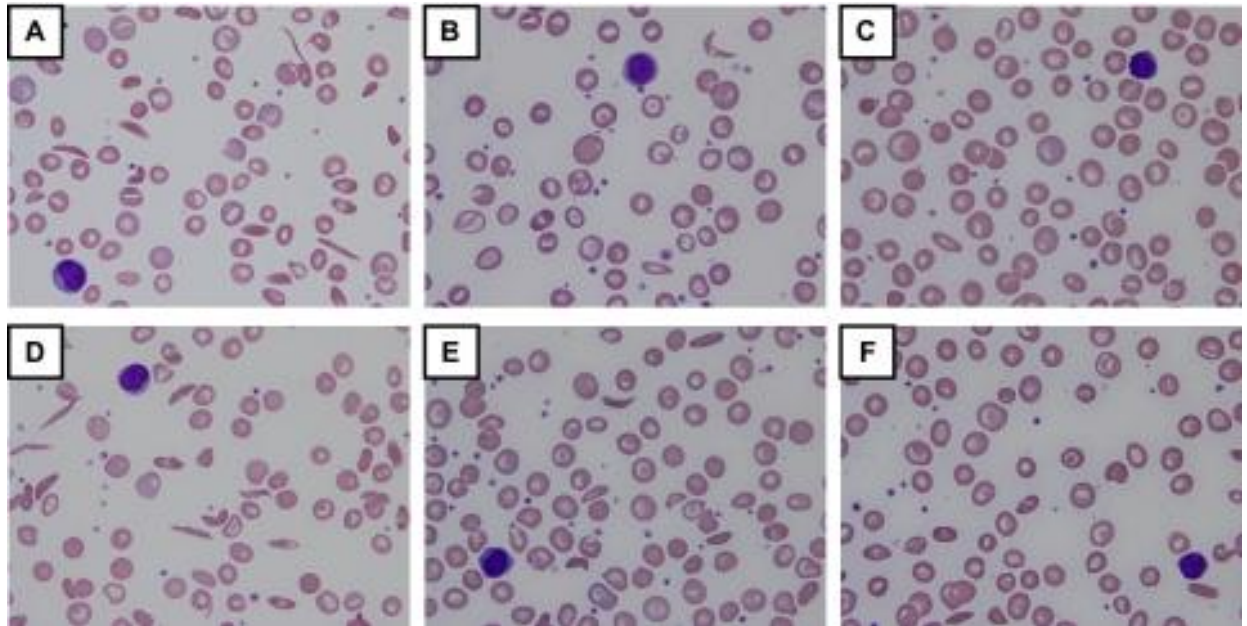


Figura 3. Cambios morfológicos durante el aumento de la dosis de hidroxurea. Fuente Pubmed “How I use hydroxyurea to treat young patients with sickle cell anemia” Extraído del autor Ware RE (39).

En conjunto, el estudio genético de la anemia falciforme ha evolucionado desde la identificación de una única mutación causal hacia una comprensión compleja del panorama genómico que modula el fenotipo. Esta transición ha sido posible gracias al desarrollo de herramientas bioinformáticas para el análisis de variantes, anotación funcional y modelado de interacciones génicas.

2.4. Aplicaciones de la bioinformática en el estudio de la anemia falciforme

La bioinformática ha revolucionado la investigación en anemia falciforme (AF), permitiendo abordar la complejidad genética y molecular de esta enfermedad a través de herramientas computacionales avanzadas. Estas tecnologías facilitan la gestión, análisis e interpretación de grandes volúmenes de datos genómicos, transcriptómicos y proteómicos derivados de pacientes afectados, con el objetivo de identificar variantes genéticas, biomarcadores y mecanismos patogénicos asociados a la enfermedad (41).

El análisis bioinformático de variantes ha sido fundamental para caracterizar las mutaciones y polimorfismos que modulan el fenotipo clínico de la AF. Herramientas de secuenciación de nueva generación (NGS) generan grandes cantidades de datos que requieren pipelines computacionales específicos para filtrar, anotar y priorizar variantes genéticas (42).

Plataformas como GATK, ANNOVAR y SnpEff permiten detectar mutaciones puntuales, inserciones, deleciones y variantes estructurales en el genoma de pacientes con AF (43).

Asimismo, la integración de bases de datos públicas (como dbSNP, ClinVar y 1000 Genomes) posibilita la interpretación funcional y clínica de estas variantes (44). Esto es crucial para distinguir mutaciones patogénicas de polimorfismos benignos y para comprender la heterogeneidad genética observada en diferentes poblaciones (45).

El estudio del perfil de expresión génica mediante tecnologías como RNA-Seq ofrece información sobre los mecanismos celulares alterados en la AF, incluyendo la regulación de genes implicados en la producción de hemoglobina, respuesta inflamatoria y estrés oxidativo. Los análisis bioinformáticos permiten identificar genes diferencialmente expresados entre pacientes y controles sanos, así como vías biológicas afectadas (46).

Por otra parte, la epigenética ha emergido como un campo clave para comprender la regulación de la expresión génica en la AF. Los datos de metilación del ADN y modificaciones de histonas pueden analizarse con herramientas bioinformáticas específicas para revelar patrones epigenéticos que modulan la expresión de HbF y otros genes relevantes (47).

La bioinformática facilita la predicción del impacto funcional de mutaciones mediante algoritmos de predicción de efectos en proteínas, como SIFT, PolyPhen-2 y MutationTaster. Estos programas evalúan la probabilidad de que un cambio aminoacídico altere la estructura o función proteica, apoyando la interpretación clínica (48).

Además, el modelado molecular y la simulación por dinámica molecular permiten estudiar la polimerización anómala de la HbS a nivel atómico, aportando conocimiento sobre la interacción entre moléculas y la formación de fibras que causan la deformación eritrocitaria (3).

El análisis bioinformático es esencial en el diseño de terapias avanzadas, como la edición génica con CRISPR/Cas9, para corregir mutaciones causantes de AF. Modelos computacionales ayudan a diseñar guías CRISPR específicas para el locus β -globina o para genes reguladores como *BCL11A*, optimizando la eficacia y minimizando efectos fuera del objetivo (49).

Además, la bioinformática apoya el desarrollo de fármacos mediante cribados virtuales y diseño de moléculas que interfieran con la polimerización de HbS o que aumenten la producción de HbF (50).

La integración de datos genómicos, clínicos y ambientales mediante enfoques bioinformáticos permite avanzar hacia una medicina personalizada en AF. Modelos de aprendizaje automático y análisis de *big data* pueden predecir la severidad clínica, la respuesta a tratamientos y el riesgo de complicaciones, facilitando la toma de decisiones terapéuticas individualizadas (51).

2.5. Epidemiología y situación global y en España

La anemia falciforme (AF) es una enfermedad genética de distribución global, cuya incidencia y prevalencia están determinadas en gran medida por factores étnico-geográficos. Tradicionalmente, se ha asociado con regiones donde la malaria ha sido endémica, como África Subsahariana, el subcontinente indio, el Mediterráneo y algunas zonas de América del Sur y Central. Sin embargo, en las últimas décadas, debido a los movimientos migratorios, su presencia se ha incrementado de manera significativa en países europeos como Francia, Reino Unido, Italia y España (52).

A nivel mundial, se estima que cada año nacen aproximadamente 300.000 niños con hemoglobinopatías graves, de los cuales más de 200.000 presentan anemia falciforme, principalmente en África (1). Según la OMS, hasta un 25% de la población en ciertas regiones africanas es portadora del alelo falciforme (HbS), y en países como Nigeria y la República Democrática del Congo, la enfermedad representa un grave problema de salud pública, siendo una de las principales causas de mortalidad infantil (53).

La prevalencia de la AF en otros continentes, como América del Norte y Europa, ha aumentado debido a la migración. En Estados Unidos, por ejemplo, afecta aproximadamente a 1 de cada 365 nacimientos afroamericanos, y más de 100.000 personas viven con la enfermedad (54). En Reino Unido, es la enfermedad genética más común, con aproximadamente 15.000 pacientes registrados y más de 240.000 portadores (55).

España ha sido testigo de un aumento sostenido de casos de anemia falciforme en los últimos 20 años, en gran parte como consecuencia de la inmigración procedente del África subsahariana y otras regiones endémicas. Aunque históricamente se consideraba una enfermedad rara en la península ibérica, en la actualidad está reconocida como una patología emergente dentro del sistema nacional de salud (56).

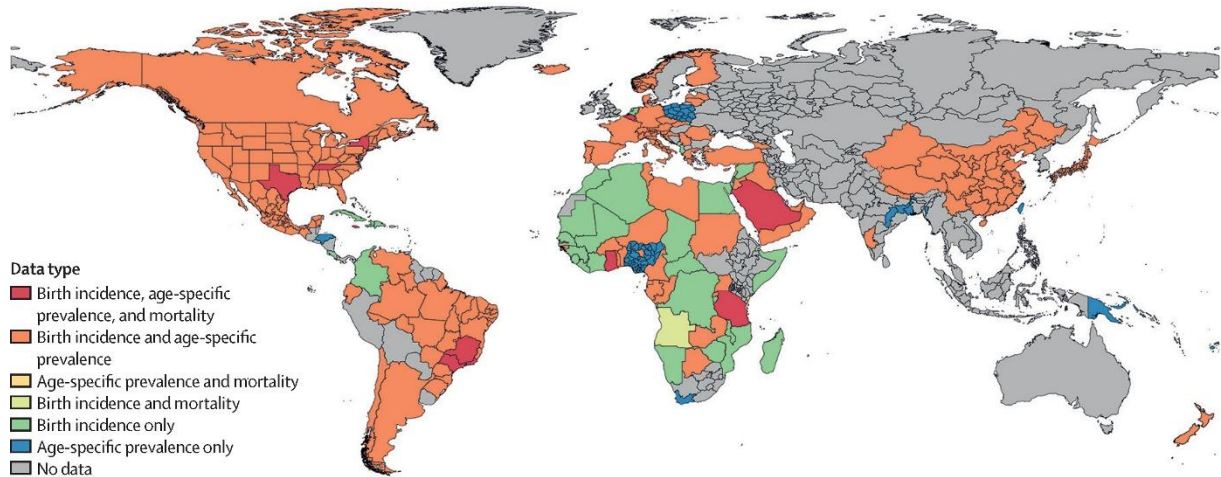


Figura 4. Mapa de nacimientos, prevalencia y mortalidad de anemia falciforme. Extraído de “Prevalencia y carga de mortalidad mundial, regional y nacional de la anemia de células falciformes, 2000-2021” (57)

Un estudio multicéntrico liderado por la Sociedad Española de Hematología y Hemoterapia (SEHH) y el Centro Nacional de Epidemiología (CNE) estimó que en España viven más de 1.000 personas con AF y que hay entre 15.000 y 20.000 portadores heterocigotos (rasgo falciforme) (58). Las comunidades autónomas con mayor prevalencia son Cataluña, Comunidad de Madrid, Andalucía y la Comunidad Valenciana, coincidiendo con las regiones con mayor proporción de población inmigrante procedente de África (59).

En la Comunidad de Madrid, diversos estudios epidemiológicos han evidenciado un aumento progresivo de los casos de hemoglobinopatías, especialmente de aquellas asociadas a la hemoglobina S (HbS). Este incremento se debe principalmente a los flujos migratorios procedentes de regiones de alta prevalencia como África subsahariana y el Caribe. En un análisis realizado entre 2007 y 2015 en centros hospitalarios madrileños, se observó que la mayoría de los nuevos diagnósticos correspondían a formas graves como la anemia falciforme, lo cual ha impulsado iniciativas de cribado neonatal y estrategias de seguimiento clínico en la comunidad autónoma (60).

El diagnóstico precoz de la anemia falciforme en recién nacidos es fundamental para iniciar tratamientos preventivos, reducir complicaciones y mejorar la calidad de vida. Algunas comunidades autónomas como Cataluña, Madrid, Castilla-La Mancha y el País Vasco han incorporado la detección de hemoglobinopatías en sus programas de cribado neonatal desde hace más de una década (61).

Sin embargo, no existe un programa nacional unificado, lo que genera desigualdades en el acceso al diagnóstico precoz según la región. Esta variabilidad ha sido señalada como una limitación significativa en la lucha contra las hemoglobinopatías en España (62). Además, la AF no siempre está incluida en los listados de enfermedades raras o prioritarias, lo que complica la financiación y continuidad de los programas de seguimiento (63).

A nivel hospitalario, se han desarrollado unidades especializadas en hemoglobinopatías que atienden a pacientes pediátricos y adultos, promoviendo la transición clínica entre servicios y facilitando el seguimiento a largo plazo. La colaboración entre hematólogos, pediatras, genetistas y trabajadores sociales es clave para ofrecer una atención integral a los pacientes con esta patología (64).

2.6. Perspectivas futuras y avances en investigación

La investigación en anemia falciforme (AF) ha experimentado avances significativos en los últimos años, abriendo nuevas posibilidades terapéuticas que podrían transformar el manejo de esta enfermedad. A continuación, se detallan las principales líneas de investigación y desarrollos recientes.

2.6.1. Edición génica in vivo y nuevas tecnologías

Más allá de las terapias *ex vivo*, se están explorando enfoques de edición génica *in vivo* que permitirían corregir la mutación responsable de la AF directamente en el organismo del paciente, sin necesidad de extracción y reinfusión de células madre (70). Estas estrategias podrían simplificar el tratamiento y hacerlo más accesible a nivel global.

Además, tecnologías emergentes como la edición génica de precisión (*prime editing*) han mostrado resultados prometedores en estudios preclínicos, logrando corregir eficientemente la mutación de la AF en células madre hematopoyéticas (71).

2.6.2. Desarrollo de nuevos fármacos y terapias combinadas

La investigación farmacológica continúa en busca de agentes que puedan aumentar la producción de HbF o mejorar la deformabilidad de los eritrocitos. Aunque algunos medicamentos como el voxelotor han sido retirados del mercado debido a preocupaciones sobre su eficacia y seguridad, otros compuestos están en fases avanzadas de desarrollo clínico (72).

Asimismo, se están evaluando terapias combinadas que integren tratamientos farmacológicos con intervenciones génicas o celulares, con el objetivo de maximizar los beneficios terapéuticos y minimizar los riesgos asociados (73).

2.6.3. Medicina personalizada y análisis de datos

La integración de datos genómicos, transcriptómicos y clínicos mediante herramientas de bioinformática y aprendizaje automático está permitiendo una comprensión más profunda de la heterogeneidad de la AF (74). Esto facilita la identificación de biomarcadores predictivos de respuesta al tratamiento y la estratificación de pacientes para terapias personalizadas.

Además, la implementación de programas de cribado neonatal y la mejora en los sistemas de vigilancia epidemiológica están contribuyendo a una detección más temprana y a un seguimiento más eficaz de los pacientes con AF (75).

2.7. Aplicaciones de la bioinformática en la anemia falciforme

La bioinformática ha emergido como una disciplina esencial en la investigación y comprensión de enfermedades genéticas como la anemia falciforme (AF). Al integrar herramientas computacionales y bases de datos genómicas, la bioinformática permite el análisis profundo de variantes genéticas, facilitando la identificación de mutaciones causales, la predicción de sus efectos funcionales y la exploración de posibles dianas terapéuticas (76).

La detección y anotación precisa de variantes genéticas es fundamental para comprender la etiología molecular de la AF. Herramientas como ANNOVAR y SnpEff permiten la anotación funcional de variantes, clasificándolas según su impacto en la estructura y función de las proteínas (76,77). Estas herramientas evalúan aspectos como la conservación evolutiva, la alteración de dominios funcionales y la posible afectación de sitios de *splicing*.

Además, bases de datos como ClinVar, dbSNP y gnomAD proporcionan información sobre la frecuencia alélica de variantes en diferentes poblaciones, así como su asociación con fenotipos clínicos (78). La integración de estos recursos permite priorizar variantes de interés y discernir entre polimorfismos benignos y mutaciones patogénicas.

La predicción *in silico* del efecto de variantes genéticas es crucial para anticipar su relevancia clínica. Herramientas como MutationTaster y Spliceman evalúan el potencial patogénico de

variantes mediante algoritmos que consideran factores como la alteración de sitios de *splicing*, la pérdida de dominios funcionales y la conservación filogenética (79,80).

Estas predicciones son especialmente útiles en la AF, donde variantes fuera del gen *HBB* pueden influir en la gravedad de la enfermedad al afectar la expresión de genes moduladores o la producción de hemoglobina fetal (80).

La bioinformática facilita la integración de datos transcriptómicos, epigenómicos y proteómicos para una comprensión holística de la AF. Por ejemplo, estudios de RNA-Seq han identificado patrones de expresión génica asociados con diferentes fenotipos clínicos de la AF, revelando posibles biomarcadores y nuevas dianas terapéuticas (81).

Herramientas como MAGMA permiten realizar análisis de asociación a nivel génico, integrando datos de GWAS para identificar genes cuya variación se correlaciona con características clínicas específicas de la AF (82).

La bioinformática desempeña un papel clave en el desarrollo de terapias génicas personalizadas para la AF. La identificación de variantes específicas en pacientes individuales permite diseñar estrategias de edición génica, como el uso de CRISPR/Cas9, para corregir mutaciones causales o modular la expresión de genes que afectan la severidad de la enfermedad (83). Además, permite simular y predecir los efectos de estas intervenciones, evaluando su eficacia y posibles efectos *off-target*, lo que es esencial para garantizar la seguridad y eficacia de las terapias génicas (83).

A pesar de los avances, la aplicación de la bioinformática en la AF presenta desafíos, como la necesidad de estandarizar *pipelines* de análisis, la interpretación de variantes de significado incierto y la integración de datos clínicos heterogéneos. Sin embargo, iniciativas como The Monarch Initiative buscan integrar datos genotípicos y fenotípicos de múltiples especies para mejorar la comprensión de las enfermedades genéticas y facilitar la traslación de hallazgos a la práctica clínica (84).

El futuro de la bioinformática en la AF se orienta hacia la implementación de enfoques de medicina de precisión, donde el análisis detallado del genoma de cada paciente guíe decisiones terapéuticas individualizadas, mejorando así los resultados clínicos y la calidad de vida de los afectados (84).

2.8. Herramientas bioinformáticas específicas aplicadas a la anemia falciforme

La anemia falciforme, como enfermedad monogénica, se beneficia especialmente del uso de herramientas bioinformáticas para el análisis de variantes puntuales, anotación funcional, predicción estructural y diseño de terapias génicas. A continuación, se presentan algunas de las herramientas más utilizadas en estudios recientes sobre esta patología, junto con sus aplicaciones específicas.

2.8.1. SIFT y PolyPhen-2: Predicción de impacto funcional

Las herramientas SIFT (*Sorting Intolerant From Tolerant*) y PolyPhen-2 (*Polymorphism Phenotyping v2*) se utilizan ampliamente para predecir el impacto funcional de variantes no sinónimas sobre la estructura y función de proteínas (85,86). Estas predicciones se basan en la conservación evolutiva de aminoácidos y en características fisicoquímicas.

En el contexto de la anemia falciforme, estas herramientas son especialmente útiles para evaluar variantes en genes modificadores, como *BCL11A*, *HBS1L-MYB* y *KLF1*, que influyen en la severidad clínica mediante la regulación de la hemoglobina fetal (32). Por ejemplo, el alelo rs1427407 del gen *BCL11A* ha sido analizado mediante PolyPhen-2 para estudiar su posible efecto sobre la represión de la hemoglobina fetal (87).

2.8.2. Ensembl Variant Effect Predictor (VEP)

El Variant Effect Predictor (VEP) de Ensembl permite la anotación de variantes genéticas con información detallada sobre localización, efecto potencial, y presencia en bases de datos como ClinVar, COSMIC o dbSNP. VEP se ha utilizado para identificar y clasificar mutaciones relacionadas con la anemia falciforme en estudios de asociación a gran escala.

Por ejemplo, variantes en promotores o regiones reguladoras del gen *HBB* pueden ser anotadas con VEP para predecir si interfieren con elementos de transcripción esenciales (88).

2.8.3. UCSC Genome Browser y IGV

El UCSC Genome Browser y el Integrative Genomics Viewer (IGV) permiten la visualización detallada de variantes dentro del contexto del genoma humano, junto con regiones reguladoras, histonas modificadas y patrones de metilación (89,90). Estas herramientas ayudan a investigadores a estudiar si ciertas variantes relacionadas con la AF se encuentran en zonas con accesibilidad de la cromatina o interacción con factores de transcripción.

Esto resulta útil, por ejemplo, al estudiar cómo ciertas mutaciones sin efectos codificantes aparentes podrían alterar la expresión génica a través de mecanismos epigenéticos (90).

2.8.4. Herramientas de diseño CRISPR

El diseño de estrategias terapéuticas basadas en edición génica se apoya fuertemente en bioinformática. Herramientas como CHOPCHOP, CRISPOR y Benchling permiten diseñar guías CRISPR con alta especificidad para dianas como el gen *BCL11A* o el locus de *HBB*, evaluando eficiencia y minimizando efectos *off-target* (91).

La predicción *in silico* de eficiencia de corte de Cas9 y la detección de sitios off-target ha sido clave en ensayos clínicos exitosos como el de CTX001, en el que se edita el *enhancer* de *BCL11A* (83).

2.8.5. Docking molecular y predicción estructural

Las herramientas de modelado estructural y docking molecular, como SWISS-MODEL, PyMOL, Phyre2 y AutoDock, permiten analizar cómo una mutación puntual afecta la estructura tridimensional de la proteína beta-globina (92,93).

Por ejemplo, la mutación Glu6Val responsable de la anemia falciforme puede modelarse para observar cómo induce agregación de hemoglobina desoxigenada mediante interacciones hidrofóbicas anómalas (93).

2.8.6. Plataformas de análisis de expresión génica

Herramientas como GEO2R, DESeq2 y edgeR se utilizan para analizar datos de expresión génica a partir de experimentos de microarrays o RNA-Seq. Han sido aplicadas para identificar genes diferencialmente expresados entre pacientes con diferentes fenotipos clínicos de AF, o para estudiar la expresión de genes relacionados con estrés oxidativo y adhesión celular (94).

Un análisis con DESeq2 identificó sobreexpresión de genes relacionados con citoquinas inflamatorias en pacientes con mayor frecuencia de crisis vasooclusivas (95).

2.8.7. Integración en plataformas multiómicas

Plataformas como ToppGene, StringDB y Cytoscape permiten integrar resultados de diferentes niveles ómicos (genómica, transcriptómica, proteómica) para construir redes de interacción génica y priorizar dianas terapéuticas (96).

Un estudio reciente integró expresión génica y variantes funcionales para identificar redes reguladoras centradas en *GATA1* y *BCL11A* como elementos clave de la fisiopatología en la AF.

3. MARCO METODOLÓGICO

3.1. Tipo de estudio

El presente estudio se plantea como un ensayo clínico observacional, analítico y prospectivo, con un enfoque bioinformático, cuyo principal objetivo es identificar y caracterizar variantes genéticas asociadas a la anemia falciforme mediante técnicas de secuenciación de nueva generación (NGS) y análisis *in silico*. El estudio se centra en el análisis comparativo entre pacientes diagnosticados con anemia falciforme y controles sanos, con el fin de encontrar correlaciones genotipo-fenotipo y evaluar el posible impacto funcional de las variantes halladas.

Aunque no se intervendrá en el tratamiento de los pacientes, este diseño permite observar asociaciones entre la presencia de variantes en el gen *HBB* (y genes relacionados) y la manifestación clínica de la enfermedad. El análisis posterior se realizará mediante el uso de herramientas bioinformáticas que permitan predecir la patogenicidad de las variantes halladas, su localización estructural y su impacto en la función proteica, alineándose con estudios similares ya publicados (97).

3.2. Tamaño muestral

La elección del tamaño muestral es un paso crucial en el diseño experimental, ya que determina la capacidad del estudio para detectar asociaciones significativas y evitar errores estadístico tipo I y II. En este proyecto, se ha seleccionado un total de 300 individuos, distribuidos en dos grupos:

- 150 pacientes con anemia falciforme (grupo caso)
- 150 individuos sanos (grupo control)

Esta decisión está basada en revisiones bibliográficas y estudios similares previos que han utilizado tamaños muestrales en este rango para estudios genéticos y genómicos en enfermedades monogénicas y multifactoriales (98). La potencia estadística estimada con este tamaño es del 80% para detectar asociaciones con un tamaño del efecto moderado (*odds ratio* entre 1.5 y 2.0) a un nivel de significancia del 5%, lo cual es adecuado para la exploración de variantes genéticas comunes y raras en la población estudiada.

Además, este número permite un análisis robusto considerando subgrupos por edad, sexo y otras variables clínicas, facilitando análisis estratificados y evitando la pérdida de poder estadístico por división de la muestra.

3.2.1. Criterios de inclusión:

- **Edad: 18-60 años:** Se ha establecido este rango de edad para limitar la variabilidad fenotípica derivada de procesos fisiológicos propios de la niñez o de la edad avanzada, los cuales pueden influir en los parámetros hematológicos y en la expresión clínica de la anemia falciforme (99). Los pacientes adultos suelen presentar una manifestación clínica más estable y representativa de la enfermedad crónica.
- **Diagnóstico confirmado de anemia falciforme mediante electroforesis de hemoglobina y pruebas moleculares:** Se garantizará que todos los casos cumplan con un diagnóstico confirmado para evitar contaminación del grupo con otros tipos de hemoglobinopatías o anemias hereditarias (100). Esto es fundamental para asegurar la validez del estudio, ya que la heterogeneidad diagnóstica puede sesgar los resultados genéticos.
- **Consentimiento informado escrito:** Cumpliendo con la Declaración de Helsinki y las normativas locales e internacionales sobre investigación en seres humanos, todos los participantes firmarán un consentimiento informado, que incluirá la explicación clara de los objetivos, procedimientos, riesgos y beneficios del estudio (101). Además, se asegurará la confidencialidad y anonimato de los datos.
- **Residencia en España:** Para controlar variables genéticas de origen poblacional y ambientales que puedan influir en la expresión clínica y genética, se incluirán solo pacientes residentes en España, con al menos tres generaciones en territorio español. Esto minimizará la heterogeneidad genética y facilitará la interpretación de los resultados en el contexto epidemiológico local.

3.2.2. Criterios de exclusión:

- **Coinfección con virus VIH, hepatitis B o C:** Estas infecciones virales pueden modificar la expresión clínica y hematológica de la anemia falciforme, afectando además la calidad y cantidad de ADN extraído (102). Se realizará cribado para descartar estas coinfecciones.

- **Otras hemoglobinopatías o trastornos hematológicos concomitantes:** Pacientes con talasemias, hemoglobinopatías dobles o enfermedades hematológicas malignas serán excluidos para evitar interferencias fenotípicas y genotípicas (103).
- **Tratamiento con transfusiones sanguíneas en los últimos 3 meses:** Las transfusiones recientes pueden alterar la proporción de células con ADN mutado y modificar los resultados de secuenciación y análisis genético (104).
- **Embarazo o lactancia:** Los cambios fisiológicos y hormonales propios del embarazo y la lactancia pueden afectar las variables hematológicas y la expresión genética (105), además de plantear consideraciones éticas adicionales.
- **Incapacidad para colaborar o cumplir con el protocolo del estudio:** Se excluyen sujetos con dificultades cognitivas, movilidad reducida o problemas para asistir a las visitas programadas, para garantizar la calidad y completitud de los datos (106).

El reclutamiento se realizará a través de hospitales asociados, unidades de hematología y centros de diagnóstico genético de ámbito nacional, siguiendo criterios éticos estrictos y aprobados por el Comité de Ética de Investigación correspondiente (107).

3.3. Consideraciones éticas

Todo el estudio se regirá por la Declaración de Helsinki y la legislación vigente en materia de investigación biomédica en España (Ley 14/2007 de Investigación Biomédica). Se garantizará el anonimato, la confidencialidad de los datos personales y el uso exclusivo de las muestras con fines científicos, previa firma del consentimiento informado. La recogida, almacenamiento y procesamiento de los datos estará supervisada por un delegado de protección de datos y se cumplirá con el Reglamento General de Protección de Datos (RGPD) de la UE (106).

3.4. Variables del estudio

El diseño contempla la recolección de variables clínicas, genéticas y sociodemográficas, clasificadas en:

Variables independientes:

- Estado de salud (paciente con anemia falciforme frente a control sano)
- Edad
- Sexo

- Etnia

Variables dependientes:

- Presencia/ausencia de variantes genéticas en *HBB* y otros genes modificadores (por ejemplo, *HBA1*, *BCL11A*, *MYB*)
- Tipo de mutación (*missense*, *nonsense*, silenciosa, *splicing*)
- Puntaje de patogenicidad (según predicción *in silico*)
- Ubicación estructural de la variante (región funcional de la proteína)

Variables clínicas asociadas:

El estudio recogerá un conjunto amplio de variables clínicas para poder correlacionar los hallazgos genéticos con el fenotipo y la gravedad de la enfermedad. Estas variables incluyen:

- Frecuencia anual de crisis vasooclusivas (dolorosas): Es la principal manifestación clínica de la anemia falciforme y está relacionada con la severidad y pronóstico. Estudios previos han demostrado que variantes en genes modificadores pueden influir en la frecuencia y gravedad de estas crisis (106,108,109).
- Número de hospitalizaciones por año: Indicador indirecto de la severidad clínica y la carga de la enfermedad en el paciente y el sistema sanitario (97).
- Niveles de hemoglobina total y fracciones (HbF, HbS): Los niveles de hemoglobina fetal (HbF) son un modificador conocido de la gravedad clínica, ya que su presencia inhibe la polimerización de la hemoglobina S (106,108,109). Además, se analizarán hemoglobina A y otras variantes para caracterizar el fenotipo hematológico.
- Complicaciones crónicas documentadas: Daño en órganos como riñón, pulmón, hígado, cerebro y retina, evaluado mediante historial clínico y pruebas complementarias, para evaluar la asociación con variantes genéticas implicadas en procesos de daño orgánico y fibrosis (102).
- Antecedentes genéticos: antecedentes familiares de anemia falciforme o hemoglobinopatías, que pueden influir en la expresión clínica y genética.
- Tratamientos farmacológicos y adherencia: Recogida de información sobre tratamientos con hidroxurea, analgésicos y otros, dado que pueden modificar el fenotipo y la expresión de la enfermedad.

3.5.Extracción de muestras y recolección de datos clínicos

El proceso de recolección de datos y obtención de muestras biológicas se llevará a cabo en el contexto de un estudio observacional prospectivo, cuyo objetivo es analizar la asociación entre variantes genéticas y la manifestación clínica de la anemia falciforme (AF). Este diseño permite establecer relaciones temporales y observar la evolución clínica de los pacientes a lo largo del tiempo (110).

3.5.1. Reclutamiento y contacto inicial

El reclutamiento de participantes se realizará en unidades clínicas especializadas en hemoglobinopatías, durante sus visitas programadas. Se priorizará el contacto directo con pacientes diagnosticados de AF que cumplan los criterios de inclusión y exclusión establecidos (ver sección 3.2.1 y 3.2.2). Una vez identificados, se les explicará verbalmente el objetivo del estudio, se resolverán dudas y, en caso de aceptación, firmarán un consentimiento informado conforme a las normas éticas internacionales, incluyendo la Declaración de Helsinki y el Reglamento (UE) 2016/679 de Protección de Datos (111).

3.5.2. Recolección de datos clínicos

En la primera consulta se administrará un cuestionario clínico estructurado, con ítems adaptados de instrumentos validados como el **SF-36 Health Survey** y escalas específicas para enfermedades crónicas hematológicas (112). Este incluirá:

- Historia clínica y comorbilidades.
- Edad al diagnóstico y tipo de mutación (si está disponible).
- Número de crisis vasooclusivas en el último año.
- Número de hospitalizaciones por complicaciones.
- Tratamiento actual (hidroxiurea, transfusiones, etc.).
- Adherencia terapéutica (escala visual analógica).
- Evaluación de calidad de vida.
- Nivel educativo y situación socioeconómica.

Todos los datos se almacenarán en una base de datos encriptada, con codificación alfanumérica que preserva el anonimato del participante, en conformidad con las guías del **CIOMS** y la normativa de la Agencia Española de Protección de Datos (111,113).

3.5.3. Extracción de muestras biológicas

En la misma consulta se procederá a la obtención de muestras sanguíneas. Esta recolección será realizada por personal de enfermería certificado, respetando las guías de buenas prácticas clínicas en la extracción de muestras humanas (114). Se extraerán:

- **Tubo EDTA (4–6 mL):** destinado al hemograma completo y posterior aislamiento de ADN genómico mediante kits comerciales (ej. Qiagen®).
- **Tubo con heparina sódica o litio:** útil en caso de estudios de expresión génica o análisis celulares adicionales.
- **Tubo sin anticoagulante (gel separador):** para obtención de suero en estudios bioquímicos.

Las muestras serán etiquetadas con un código único y transportadas bajo condiciones de cadena de frío (2–8 °C), siguiendo los protocolos ISO 20387 para biobancos. Posteriormente se almacenarán en congeladores a -80 °C en un biobanco universitario asociado, hasta su procesamiento genético (115).

3.5.4. Número de pacientes y análisis

Se estima reclutar un total de **300 participantes**: 150 con diagnóstico confirmado de anemia falciforme (mediante electroforesis de hemoglobina o secuenciación del gen *HBB*) y 150 controles sanos, emparejados por edad, sexo y origen étnico. Esta muestra permite alcanzar una potencia estadística del 80% para detectar asociaciones genotipo-fenotipo con un error alfa del 5% y tamaño del efecto moderado, según cálculos basados en estudios similares previos (116).

Una vez alcanzado el umbral mínimo de pacientes reclutados, se procederá al análisis de variantes mediante herramientas bioinformáticas, las cuales se detallan en el siguiente apartado (sección 3.6).

3.6. Procesamiento bioinformático de las variantes genéticas

El procesamiento bioinformático constituye una fase crítica del estudio, ya que permite identificar, predecir y clasificar las variantes genéticas encontradas en los pacientes con anemia falciforme. Este proceso incluye múltiples pasos secuenciales: desde la extracción del ADN, la obtención de datos genómicos mediante secuenciación, hasta la interpretación funcional mediante herramientas *in silico*.

3.6.1. Extracción y secuenciación del ADN

El ADN genómico se extraerá de muestras de sangre periférica previamente almacenadas, utilizando kits comerciales estandarizados como el **QIAamp DNA Blood Mini Kit** (Qiagen®), siguiendo las recomendaciones del fabricante. La calidad y concentración del ADN serán evaluadas con espectrofotometría (NanoDrop) y electroforesis en gel de agarosa. Solo las muestras con una relación A260/A280 entre 1.8 y 2.0 serán consideradas aptas para secuenciación.

La secuenciación del genoma o exoma se realizará mediante tecnología de **segunda generación (NGS)**, utilizando plataformas como **Illumina NovaSeq 6000**, que permiten una cobertura adecuada (>100x) del gen *HBB* y otros genes modificadores potenciales como *BCL11A*, *HBS1L-MYB* o *KLF1* (27).

3.6.2. Control de calidad y alineamiento

Los archivos generados (formato FASTQ) serán sometidos a control de calidad mediante **FastQC** y preprocesamiento con **Trimmomatic**, eliminando adaptadores y secuencias de baja calidad (117). Los datos limpios se alinearán al genoma de referencia humano **GRCh38** utilizando el algoritmo **BWA-MEM**. Posteriormente, se generarán archivos **BAM** ordenados con **SAMtools** y se eliminarán duplicados con **Picard**.

3.6.3. Llamada de variantes y anotación

La detección de variantes se realizará con el software **GATK (Genome Analysis Toolkit)**, que permite identificar variantes puntuales (SNVs) e indels. Las variantes serán filtradas por calidad (QUAL > 30) y profundidad (DP > 10), y se generará un archivo VCF consolidado.

Para la anotación funcional de las variantes se utilizarán las siguientes herramientas:

- **SnEff**: predice el impacto funcional de las variantes en regiones codificantes y no codificantes del genoma (77).
- **VEP (Variant Effect Predictor)**: desarrollado por Ensembl, permite acceder a datos de conservación evolutiva, frecuencia alélica en bases de datos como gnomAD, y vínculos con enfermedades descritas en ClinVar (88).
- **dbNSFP**: para integrar predicciones múltiples desde algoritmos como SIFT, PolyPhen-2, MutationTaster, entre otros (118).

3.6.4. Herramientas de predicción de patogenicidad

Las variantes no sinónimas identificadas se someterán a predicción funcional con herramientas especializadas:

- **SIFT** (*Sorting Intolerant From Tolerant*): evalúa el impacto de sustituciones aminoacídicas en función de la conservación evolutiva; valores <0.05 indican posible efecto perjudicial (85).
- **PolyPhen-2**: combina propiedades estructurales y evolución para estimar el impacto; clasifica las variantes como benignas, posiblemente dañinas o probablemente dañinas (86).
- **CADD (Combined Annotation Dependent Depletion)**: proporciona una puntuación global de deleteriedad combinando más de 60 anotaciones (119).

Además, se utilizarán recursos complementarios como **ClinVar**, **HGMD** y **OMIM** para explorar la relación conocida de ciertas variantes con fenotipos clínicos relacionados con la anemia falciforme o sus complicaciones (120).

3.6.5. Interpretación y categorización de variantes

La clasificación de las variantes se realizará siguiendo las directrices del **American College of Medical Genetics and Genomics (ACMG)**, que establece cinco categorías: benigna, probablemente benigna, incierta, probablemente patogénica y patogénica (121). Para esta clasificación se tendrá en cuenta:

- Evidencia funcional (predicciones *in silico*).
- Frecuencia en población general.

- Datos clínicos del paciente.
- Coincidencia con bases de datos clínicas.

Las variantes consideradas “de interés” (probablemente patogénicas o patogénicas) serán correlacionadas con variables clínicas (frecuencia de crisis, hospitalizaciones, respuesta a tratamiento, etc.) en la fase de análisis estadístico.

3.7. Análisis estadístico

Una vez obtenidos los datos genéticos y clínicos de los participantes, se procederá al análisis estadístico mediante software especializado como R o SPSS. En primer lugar, se realizará una descripción de las variables demográficas y clínicas (edad, sexo, región de origen, presencia de comorbilidades, niveles de hemoglobina, frecuencia de crisis vasooclusivas, etc.), utilizando medidas de tendencia central y dispersión para variables cuantitativas, y frecuencias absolutas y relativas para variables cualitativas (122).

Posteriormente, se llevará a cabo una comparación entre los grupos de estudio (pacientes con anemia falciforme y controles sanos) mediante pruebas de hipótesis: t de Student o U de Mann-Whitney para variables cuantitativas, y chi-cuadrado o test exacto de Fisher para las cualitativas, dependiendo del cumplimiento de los supuestos de normalidad y homogeneidad de varianzas (123). Para ello, se emplearán también pruebas como Shapiro-Wilk o Kolmogorov-Smirnov para la verificación de la normalidad, y el test de Levene para evaluar la igualdad de varianzas (124).

Con el objetivo de explorar asociaciones entre variantes genéticas y manifestaciones clínicas, se utilizarán modelos de regresión logística multivariante. En dichos modelos, las variables clínicas actuarán como variables dependientes, mientras que las variantes genéticas y las covariables relevantes (edad, sexo, origen geográfico, niveles de HbF, entre otras) se incluirán como variables independientes y de confusión (125). Este enfoque permitirá estimar *odds ratios* ajustados (ORa) y sus respectivos intervalos de confianza (IC95%) para valorar la fuerza de asociación entre los polimorfismos identificados y la severidad del fenotipo (126).

Además, para las variantes que muestren significación estadística, se realizará un análisis de interacción mediante regresión estratificada o términos de interacción en el modelo, con el fin de evaluar el posible efecto modificador de variables como el nivel de HbF o el uso de

hidroxiurea (127). Para todos los análisis, se considerará un nivel de significación estadística de $p < 0,05$.

3.8. Integración e interpretación de resultados

Una vez identificadas las variantes genéticas relevantes, se integrarán los resultados funcionales obtenidos mediante herramientas de predicción (como SIFT, PolyPhen-2, MutationTaster y CADD) con los datos clínicos y epidemiológicos recolectados (86). Esta triangulación de la información permitirá discriminar aquellas variantes que, además de ser estadísticamente significativas, tienen un impacto biológico demostrado sobre la función proteica o la regulación génica.

Los resultados se interpretarán en el contexto de las bases de datos de referencia como ClinVar, dbSNP, gnomAD y OMIM, con el objetivo de validar la patogenicidad de las variantes y clasificar su frecuencia poblacional (128). Esto permitirá diferenciar polimorfismos benignos de variantes raras con potencial efecto clínico, especialmente en genes como *BCL11A*, *HBS1L-MYB*, *KLF1* o *HMOX1*, que han demostrado modular la expresión de hemoglobina fetal y otros rasgos relacionados con la severidad de la AF (129).

Asimismo, los resultados obtenidos se contrastarán con estudios previos de cohortes internacionales, lo cual permitirá evaluar la consistencia de los hallazgos y determinar si las variantes identificadas coinciden con aquellas descritas en otras poblaciones de origen africano o latinoamericano (130). La integración de estos datos contribuirá no solo al conocimiento genético de la enfermedad en población española, sino también al diseño futuro de estrategias de medicina personalizada.

Finalmente, los hallazgos relevantes serán representados mediante gráficos de *forest plot*, diagramas de Manhattan y mapas de calor (*heatmaps*), facilitando así su visualización y comprensión. Se discutirá el potencial de las variantes más destacadas como biomarcadores de pronóstico o respuesta al tratamiento, planteando su utilidad clínica en el seguimiento individualizado de los pacientes (19).

3.9. Cronograma

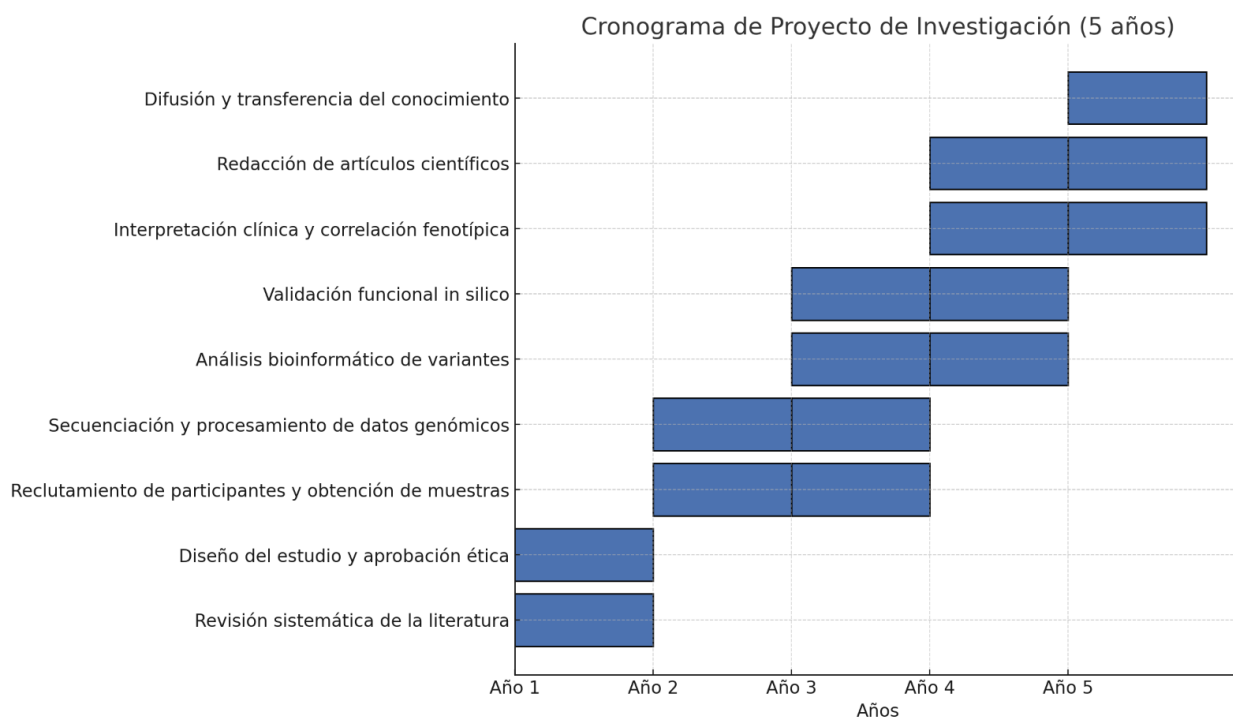


Figura 5. Cronograma de proyecto de investigación. Elaboración propia con Python.

3.10. Presupuesto

Tabla 1. Presupuesto detallado del estudio durante 5 años

Categoría	Detalle	Coste estimado (€)
1. Personal		
Bioinformático principal (50% JdP)	5 años x 25.000 €/año x 0.5	62.500
Investigador clínico asociado	5 años x 20.000 €/año x 0.5	50.000
Técnico de laboratorio / soporte	5 años x 15.000 €/año	75.000
Estudiantes de máster/doctorado	2 becas parciales durante 3 años	18.000
2. Infraestructura y equipamiento		
Servidores cloud de alto rendimiento	Análisis genómico (Amazon AWS/Google Cloud)	12.000

Software y licencias especializadas	ANNOVAR, CADD, herramientas de anotación	3.000
Estaciones de trabajo y mantenimiento	2 ordenadores potentes + renovación técnica	6.000
3. Adquisición y procesamiento de datos		
Compra de datos genómicos	Acceso a cohortes externas o consorcios	8.000
Secuenciación externa (si aplica)	Panel dirigido o WES de 50 pacientes	25.000
4. Validación y análisis funcional		
Herramientas in silico y pipelines	Coste de cálculo en plataformas bioinformáticas	5.000
Consultoría genética especializada	Validación funcional e interpretación clínica	4.000
5. Difusión y formación		
Publicaciones científicas (open access)	3 publicaciones en revistas relevantes	6.000
Asistencia a congresos científicos	2 personas/año x 3 años	9.000
Organización de taller de resultados	Evento final + material divulgativo	2.500
6. Gestión y gastos generales		
Coordinación administrativa	Informes, auditoría, seguimiento financiero	4.000
Imprevistos y emergencias	5% del total presupuestado	14.200
PRESUPUESTO TOTAL		298200

Elaboración propia

4. Resultados esperados

El presente estudio se plantea como una aproximación bioinformática integral para la caracterización de variantes genéticas asociadas a la anemia falciforme, con especial énfasis en el gen *HBB* y en genes modificadores como *BCL11A*, *KLF1* y el clúster *HBS1L-MYB*. A partir de las muestras genómicas obtenidas de pacientes y controles sanos, se espera identificar mutaciones de alto impacto funcional que permitan explicar tanto la aparición de la enfermedad como la heterogeneidad clínica observada entre los individuos afectados.

Se prevé encontrar, de forma consistente, la mutación puntual **Glu6Val (c.20A>T)** en el gen *HBB*, responsable directa de la sustitución del ácido glutámico por valina en la cadena beta de la hemoglobina. Esta variante patogénica, ampliamente descrita en la literatura (131), constituye el evento causal clásico de la hemoglobina S (HbS), y su detección mediante herramientas como ClinVar, SIFT y PolyPhen-2 será uno de los hallazgos esperados más relevantes. Además de confirmar su presencia, se anticipa que los algoritmos de predicción bioinformática la clasifiquen como una variante de **alto impacto funcional** y alta conservación evolutiva, reafirmando su rol crítico en la fisiopatología de la enfermedad.

Sin embargo, uno de los objetivos fundamentales de este trabajo es ir más allá del diagnóstico molecular primario y explorar **el efecto de variantes en genes modificadores**. Se espera que algunas de estas variantes se asocien con una **mayor producción de hemoglobina fetal (HbF)**, la cual tiene un efecto protector parcial al inhibir la polimerización de HbS y, por ende, reducir la severidad de las crisis vasooclusivas. Por ejemplo, se han descrito polimorfismos en regiones intrónicas del gen *BCL11A* que actúan como reguladores transcripcionales negativos de *HBG1/HBG2*, genes que codifican para las cadenas gamma de la HbF. Variantes que reduzcan la actividad represora de *BCL11A* podrían, por tanto, asociarse a un fenotipo más leve y constituir un hallazgo importante dentro de la cohorte estudiada (132).

De forma similar, se espera encontrar SNPs funcionales en el **clúster intergénico *HBS1L-MYB***, en el cromosoma 6q23, que también han sido vinculados con el incremento de HbF y con la modulación de la gravedad clínica en pacientes falciformes (133). Las herramientas utilizadas en el análisis, como Ensembl VEP y ANNOVAR, permitirán determinar la posición exacta, frecuencia alélica, efectos funcionales y posibles asociaciones fenotípicas de estas variantes. También se prevé identificar cambios en el gen *KLF1*, un factor de transcripción eritroide

esencial que puede alterar la expresión de múltiples genes implicados en la eritropoyesis y que podría influir indirectamente sobre los niveles de HbF (134).

Además de caracterizar variantes individuales, el análisis integrará modelos estadísticos para **evaluar correlaciones entre genotipo y fenotipo**, incluyendo parámetros como la frecuencia de crisis, el número de hospitalizaciones anuales, los niveles de HbF y la respuesta terapéutica a agentes como la hidroxiurea. Se espera que algunos perfiles genéticos se asocien de forma significativa con un curso clínico más benigno, permitiendo proponer un modelo preliminar de **estratificación genética de riesgo** en pacientes con anemia falciforme (135).

Por otra parte, uno de los resultados esperados más relevantes es demostrar la **utilidad de la bioinformática como herramienta de apoyo clínico**. La capacidad de integrar múltiples capas de información genómica (predicción funcional, frecuencia poblacional, efecto sobre la expresión) puede ayudar a los clínicos a tomar decisiones más informadas en cuanto a la personalización del tratamiento, la indicación de terapias modificadoras de HbF o incluso la inclusión de ciertos pacientes en ensayos clínicos dirigidos (136). Además, se espera sentar las bases para el diseño futuro de **paneles diagnósticos de variantes relevantes** que puedan implementarse en estudios poblacionales o cribados neonatales en contextos clínicos reales.

En conclusión, se anticipa que el estudio no solo confirmará las variantes causales conocidas, sino que identificará nuevas asociaciones genotipo-fenotipo relevantes para la práctica clínica. Estos hallazgos podrían contribuir significativamente a una mejor comprensión de la variabilidad en la expresión clínica de la enfermedad y abrir nuevas vías para la estratificación de pacientes y el desarrollo de terapias personalizadas.

5. Limitaciones y propuestas futuras

El análisis bioinformático propuesto representa una herramienta poderosa para la **identificación, caracterización y priorización de variantes genéticas asociadas a la anemia falciforme**, integrando datos funcionales, clínicos y poblacionales. Los resultados esperados refuerzan el conocimiento previamente establecido sobre el papel central de la mutación *Glu6Val* en el gen *HBB* como causa primaria de la enfermedad (19), pero también permiten destacar la relevancia de variantes en genes modificadores como *BCL11A*, *KLF1* y el clúster *HBS1L-MYB*, cuyos efectos reguladores sobre la expresión de hemoglobina fetal podrían explicar la variabilidad fenotípica observada entre pacientes con el mismo genotipo primario (132,134).

Uno de los aspectos más relevantes del presente trabajo es la **estratificación genética de los pacientes**. Este enfoque puede permitir, a medio plazo, la identificación de subgrupos con diferentes riesgos clínicos y respuestas terapéuticas, sentando las bases para una **medicina personalizada** en hemoglobinopatías (135). Esta perspectiva ya ha comenzado a implementarse en otros trastornos monogénicos como la fibrosis quística o algunas formas de cáncer hereditario, y su extensión a la anemia falciforme representa un paso natural en la evolución de la atención clínica (137).

Sin embargo, el enfoque planteado también presenta **diversas limitaciones**. En primer lugar, el estudio se basa en análisis *in silico*, lo que implica que **la funcionalidad real de muchas variantes predichas como dañinas o reguladoras no puede confirmarse experimentalmente** dentro del alcance del trabajo. Aunque herramientas como SIFT, PolyPhen-2 o CADD ofrecen una alta sensibilidad y especificidad, sus predicciones no sustituyen a la validación funcional en modelos celulares o animales (138). Además, el poder predictivo de estos algoritmos puede variar según la región genómica y el contexto del polimorfismo, por lo que se requiere cautela al interpretar sus resultados como evidencia definitiva.

Otra limitación relevante es la **falta de datos clínicos longitudinales emparejados con los genotipos**, lo que impide establecer correlaciones robustas entre las variantes y los desenlaces clínicos. Aunque se mencionan posibles asociaciones entre ciertos SNPs y la expresión de HbF, estas deben interpretarse como **hipótesis generadas por el análisis bioinformático** y no como conclusiones confirmadas. Idealmente, estas correlaciones

deberían validarse en cohortes independientes, con un seguimiento clínico prospectivo y homogéneo (95,139).

Asimismo, la representatividad poblacional es una cuestión crítica. Muchas bases de datos de variantes como gnomAD están sesgadas hacia poblaciones europeas o asiáticas, por lo que las **frecuencias alélicas observadas podrían no reflejar con precisión la realidad genética de poblaciones afrodescendientes**, donde la anemia falciforme es más prevalente. Este sesgo limita la capacidad de inferencia generalizada y refuerza la necesidad de incluir poblaciones subrepresentadas en estudios genómicos futuros (140).

En cuanto a las **propuestas futuras**, este estudio sienta las bases para varios desarrollos complementarios:

1. **Validación funcional** de variantes candidatas mediante ensayos moleculares in vitro (por ejemplo, mediante edición génica CRISPR/Cas9 para evaluar el efecto de SNPs en la regulación de HbF).
2. **Ampliación de la cohorte** y recopilación sistemática de variables clínicas, para permitir modelos genotipo-fenotipo más robustos.
3. **Desarrollo de paneles diagnósticos multigénicos**, que incluyan no solo *HBB* sino también genes modificadores, con utilidad potencial en cribado neonatal o asesoramiento genético.
4. **Evaluación de la utilidad clínica** de los perfiles genéticos en la toma de decisiones terapéuticas, especialmente en la indicación de tratamientos moduladores de HbF o en la selección de candidatos a terapias génicas emergentes (141).

Finalmente, este tipo de aproximaciones integradas pueden contribuir a reducir las inequidades sanitarias en poblaciones afectadas por hemoglobinopatías, proporcionando herramientas de diagnóstico más precisas y potenciando el acceso a tratamientos personalizados.

6. Bibliografía

1. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Dewi M, et al. Global epidemiology of Sickle haemoglobin in neonates: A contemporary geostatistical model-based map and population estimates. *The Lancet*. 2013;381:142–51.
2. Yawn BP, Buchanan GR, Afenyi-Annan AN, Ballas SK, Hassell KL, James AH, et al. Management of sickle cell disease: Summary of the 2014 evidence-based report by expert panel members. Vol. 312, *JAMA - Journal of the American Medical Association*. American Medical Association; 2014. p. 1033–48.
3. Eaton WA. Hemoglobin S polymerization and sickle cell disease: A retrospective on the occasion of the 70th anniversary of Pauling's Science paper. Vol. 95, *American Journal of Hematology*. Wiley-Liss Inc.; 2020. p. 205–11.
4. Rodríguez-Velásquez JO, Castillo-Bohórquez ML, Oliveros-Rozó AL, Soracipa-Muñoz MY, Prieto-Bohórquez SE. Caracterización geométrica euclidiana y fractal de células falciformes. *Nova*. 2020 Feb;18.
5. Steinberg MH, Sebastiani P. Genetic modifiers of sickle cell disease. Vol. 87, *American Journal of Hematology*. 2012. p. 795–803.
6. Sickle Cell Disease - ARISE Initiative. <https://www.ariseinitiative.org/arise-project/sickle-cell-disease/>.
7. Mañú Pereira M, Vives Corrons JL. Neonatal haemoglobinopathy screening in Spain. Vol. 62, *Journal of Clinical Pathology*. 2009. p. 22–5.
8. Garcia FA de O, Andrade ES de, Palmero EI. Insights on variant analysis in silico tools for pathogenicity prediction. Vol. 13, *Frontiers in Genetics*. Frontiers Media S.A.; 2022.
9. ClinVar. <https://www.ncbi.nlm.nih.gov/clinvar/>. ClinVar.
10. Kato GJ, Piel FB, Reid CD, Gaston MH, Ohene-Frempong K, Krishnamurti L, et al. Sickle cell disease. *Nature Reviews Disease Primers* 2018 4:1 [Internet]. 2018 Mar 15 [cited 2025 Apr 14];4(1):1–22. Available from: <https://www.nature.com/articles/nrdp201810>
11. <https://www.merckmanuals.com/es-us/hogar/trastornos-de-la-sangre/anemia/anemia-de-c%C3%A9lulas-falciformes-anemia-drepanoc%C3%ADtica->

12. Kavanagh PL, Fasipe TA, Wun T. Sickle Cell Disease: A Review. Vol. 328, JAMA. American Medical Association; 2022. p. 57–68.
13. Tebbi CK. Sickle Cell Disease, a Review. Vol. 3, Hemato. Multidisciplinary Digital Publishing Institute (MDPI); 2022. p. 341–66.
14. Brandow AM, Liem RI. Advances in the diagnosis and treatment of sickle cell disease. Vol. 15, Journal of Hematology and Oncology. BioMed Central Ltd; 2022.
15. <https://www.verywellhealth.com/sickle-cell-and-malaria-5323165> [Internet]. Sickle Cell and Malaria: What’s the Link?
16. Ingram VM. A specific chemical difference between the globins of normal human and sickle-cell anæmia hæmoglobin. Nature. 1956;178:792–4.
17. Everall IP, Lecouteur A. Firesetting in an adolescent boy with Asperger’s syndrome. British Journal of Psychiatry. 1990;157:284–7.
18. Ferrone FA. Polymerization and sickle cell disease: A molecular view. Microcirculation. 2004 Mar;11:115–28.
19. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. In: The Lancet. Elsevier B.V.; 2010. p. 2018–31.
20. Hebbel RP. Reconstructing sickle cell disease: A data-based analysis of the “hyperhemolysis paradigm” for pulmonary hypertension from the perspective of evidence-based medicine. Vol. 86, American Journal of Hematology. 2011. p. 123–54.
21. Brousse V, Buffet P, Rees D. The spleen and sickle cell disease: The sick(led) spleen. Vol. 166, British Journal of Haematology. Blackwell Publishing Ltd; 2014. p. 165–76.
22. Kau DK, Finnegan E, Barabino GA. Sickle red cell-endothelium interactions. Vol. 16, Microcirculation. 2009. p. 97–111.

23. Ware RE, de Montalembert M, Tshilolo L, Abboud MR. Sick cell disease. *The Lancet* [Internet]. 2017 Jul 15 [cited 2025 Apr 14];390(10091):311–23. Available from: <https://www.thelancet.com/action/showFullText?pii=S0140673617301939>
24. Ataga KI, Key NS. Hypercoagulability in sickle cell disease: new approaches to an old problem. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program.* 2007. p. 91–6.
25. Debaun MR, Kirkham FJ. Central nervous system complications and management in sickle cell disease. Vol. 127, *Blood. American Society of Hematology*; 2016. p. 829–38.
26. Kato GJ, Gladwin MT, Steinberg MH. Deconstructing sickle cell disease: Reappraisal of the role of hemolysis in the development of clinical subphenotypes. *Blood Rev.* 2007 Jan;21:37–47.
27. Steinberg MH. Genetic etiologies for phenotypic diversity in sickle cell anemia. Vol. 9, *TheScientificWorldJournal.* 2009. p. 46–67.
28. Ribeil JA, Hacein-Bey-Abina S, Payen E, Magnani A, Semeraro M, Magrin E, et al. Gene Therapy in a Patient with Sickle Cell Disease. *New England Journal of Medicine.* 2017 Mar;376:848–55.
29. Ingram VM. Gene mutations in human hæmoglobin: The chemical difference between normal and sickle cell hæmoglobin. *Nature.* 1957;180:326–8.
30. Serjeant GR. The natural history of sickle cell disease. *Cold Spring Harb Perspect Med.* 2013;3.
31. Nagel RL, Fabry ME, Steinberg MH. The paradox of hemoglobin SC disease. *Blood Rev.* 2003;17:167–78.
32. Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and β -globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A.* 2008 Aug;105:11869–74.
33. Menzel S, Garner C, Gut I, Matsuda F, Yamaguchi M, Heath S, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet.* 2007;39:1197–9.

34. Sedgewick AE, Timofeev N, Sebastiani P, So JCC, Ma ESK, Chan LC, et al. BCL11A is a major HbF quantitative trait locus in three different populations with β -hemoglobinopathies. *Blood Cells Mol Dis*. 2008 Nov;41:255–8.
35. Somervaille T. <https://pmc.ncbi.nlm.nih.gov/articles/PMC1282256/>. 2001. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management.
36. Makani J, Ofori-Acquah SF, Nnodu O, Wonkam A, Ohene-Frempong K. Sickle cell disease: New opportunities and challenges in Africa. *The Scientific World Journal*. 2013;2013.
37. Geard A, Pule GD, Chelo D, Bitoungui VJN, Wonkam A. Genetics of sickle cell-associated cardiovascular disease: An expert review with lessons learned in Africa. Vol. 20, *OMICS A Journal of Integrative Biology*. Mary Ann Liebert Inc.; 2016. p. 581–92.
38. Nagel RL, Fabry ME, Pagnier J, Zohoun I, Wajcman H, Baudin V, et al. Hematologically and Genetically Distinct Forms of Sickle Cell Anemia in Africa: The Senegal Type and the Benin Type. *N Engl J Med*. 1985;312:880–4.
39. Ware RE. How I use hydroxyurea to treat young patients with sickle cell anemia. Vol. 115, *Blood*. American Society of Hematology; 2010. p. 5300–11.
40. Canver MC, Orkin SH. Customizing the genome as therapy for the β -hemoglobinopathies. Vol. 127, *Blood*. American Society of Hematology; 2016. p. 2536–45.
41. Zhou H, Gul Y, Hameed Y, Alhomrani M, Alghamdi SA, Almalki AA, et al. Unveiling the unexplored novel signatures for osteoporosis via a detailed bioinformatics and molecular experiments based approach. *Am J Transl Res*. 2024;16:1306–21.
42. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. XThe next-generation sequencing revolution and its impact on genomics. Vol. 155, *Cell*. Elsevier B.V.; 2013. p. 27.
43. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010 Jul;38.

44. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018 Jan;46:D1062–7.
45. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, et al. The genetic structure and history of Africans and African Americans. *Science* (1979). 2009 May;324:1035–44.
46. Vona R, Sposi NM, Mattia L, Gambardella L, Straface E, Pietraforte D. Sickle cell disease: Role of oxidative stress and antioxidant therapy. Vol. 10, *Antioxidants*. MDPI; 2021. p. 1–27.
47. Starlard-Davenport A, Smith A, Vu L, Li B, Pace BS. MIR29B mediates epigenetic mechanisms of HBG gene activation. *Br J Haematol.* 2019 Jul;186:91–100.
48. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res.* 2001 May;11:863–74.
49. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, et al. CRISPR/Cas9 β -globin gene targeting in human haematopoietic stem cells. *Nature.* 2016;539:384–9.
50. Cetin B, Erendor F, Eksi YE, Sanlioglu AD, Sanlioglu S. Advancing CRISPR Genome Editing into Gene Therapy Clinical Trials: Progress and Future Prospects. *Expert Reviews in Molecular Medicine*. Cambridge University Press; 2025.
51. Yu I, Chen RK. A feasibility study of an extrusion-based fabrication process for personalized drugs. *J Pers Med.* 2020 Mar;10.
52. Modell B, Darlison M. Public health reviews Global epidemiology of haemoglobin disorders [Internet]. Vol. 86, *Bulletin of the World Health Organization*. 2008. Available from: www.chime.
53. Regional Committee for Africa 60. <https://iris.who.int/handle/10665/1682>. 2011. Sickle-Cell Disease: a strategy for the WHO African Region.
54. Hassell KL. Population Estimates of Sickle Cell Disease in the U.S. Vol. 38, *American Journal of Preventive Medicine*. 2010.
55. Streetly A, Latinovic R, Hall K, Henthorn J. Implementation of universal newborn bloodspot screening for sickle cell disease and other clinically significant

56. Aguilar Martinez P, Angastiniotis M, Eleftheriou A, Gulbis B, Mañú Pereira MDM, Petrova-Benedict R, et al. Haemoglobinopathies in Europe: Health & migration policy perspectives. Vol. 9, Orphanet Journal of Rare Diseases. BioMed Central Ltd.; 2014.
57. Thomson AM, McHugh TA, Oron AP, Teply C, Lonberg N, Vilchis Tella VM, et al. Global, regional, and national prevalence and mortality burden of sickle cell disease, 2000–2021: a systematic analysis from the Global Burden of Disease Study 2021. *Lancet Haematol*. 2023 Aug;10:e585–99.
58. <https://www.sehh.es/> [Internet]. Más del 30% de las consultas externas y hospitalarias de Hematología están relacionadas con la anemia.
59. Bardón Cancho EJ, García-Morín M, Beléndez C, Velasco P, Benéitez D, Ruiz-Llobet A, et al. Update of the Spanish registry of haemoglobinopathies in children and adults. *Med Clin (Barc)*. 2020 Aug;155:95–103.
60. Marco Sánchez JM, Bardón Cancho EJ, Benéitez D, Payán-Pernía S, Collado Gimbert A, Ruiz-Llobet A, et al. Haemoglobinopathies and other rare anemias in Spain: ten years of a nationwide registry (REHem-AR). *Ann Hematol*. 2024 Aug;103:2743–55.
61. Cribado neonatal de hemoglobinopatías [Internet]. Available from: <http://sergas.es/avalia-t>
62. Rodríguez-Moldes B, Carbajo AJ, Sanchez B, Fernandez M, Gari M, Fernandez MC, et al. Seguimiento en Atención Primaria de los recién nacidos con enfermedad falciforme detectados en el cribado neonatal de la Comunidad de Madrid. *An Pediatr (Engl Ed)*. 2015;82:222–7.
63. MINISTERIO DE SANIDAD. Situación de las Enfermedades Raras en 2010-2018 [Internet]. 2021 [cited 2025 Jun 11]. Available from: https://www.sanidad.gob.es/areas/calidadAsistencial/estrategias/enfermedadesRaras/docs/Informe_reer_reducido_2021.pdf
64. Española S, Hematología DE, Oncología Y. ENFERMEDAD DE CÉLULAS FALCIFORMES GUÍA DE PRÁCTICA CLÍNICA [Internet]. [cited 2025 Jun 11]. Available from:

65. FORWARDING OR COPYING IS A VIOLATION OF U.S. AND INTERNATIONAL COPYRIGHT LAWS The Medical Letter ® on Drugs and Therapeutics.
66. <https://www.yalemedicine.org/news/gene-therapies-sickle-cell-disease> [Internet]. Casgevy and Lyfgenia: Two Gene Therapies Approved for Sickle Cell Disease.
67. Klein J. <https://time.com/7094725/vertex-pharmaceuticals-and-crispr-therapeutics-casgevy/>. 2024. Vertex Pharmaceuticals and CRISPR Therapeutics Casgevy: the 200 Best Inventions of 2024.
68. Kanter J, Thompson AA, Pierciey FJ, Hsieh M, Uchida N, Leboulch P, et al. Lovo-cel gene therapy for sickle cell disease: Treatment process evolution and outcomes in the initial groups of the HGB-206 study. *Am J Hematol*. 2023 Jan;98:11–22.
69. <https://www.fda.gov/news-events/press-announcements/fda-approves-first-gene-therapies-treat-patients-sickle-cell-disease> [Internet]. FDA Approves First Gene Therapies to Treat Patients with Sickle Cell Disease.
70. Leonard A, Tisdale JF. Gene therapy for sickle cell disease. *Hematology Am Soc Hematol Educ Program*. 2023 Dec;2023:542–7.
71. <https://www.stjude.org/research/translational-innovation/sickle-cell-awareness.html> [Internet]. Advances in Sickle Cell Disease.
72. Joshi S. <https://www.delveinsight.com/blog/emerging-sickle-cell-disease-therapies>. Spotlight on 7 Emerging Sickle Cell Disease Therapies.
73. Kumar A, Bhattacharya S. Sickle cell disease: a comparative perspective on global and national initiatives. *Frontiers in Hematology*. 2024 Aug;3.
74. Barak M, Hu C, Matthews A, Fortenberry YM. Current and Future Therapeutics for Treating Patients with Sickle Cell Disease. Vol. 13, *Cells*. Multidisciplinary Digital Publishing Institute (MDPI); 2024.
75. Bell V, Varzakas T, Psaltopoulou T, Fernandes T. Sickle Cell Disease Update: New Treatments and Challenging Nutritional Interventions. Vol. 16, *Nutrients*. Multidisciplinary Digital Publishing Institute (MDPI); 2024.

76. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010 Jul;38.
77. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6:80–92.
78. Gudmundsson S, Singer-Berk M, Watts NA, Phu W, Goodrich JK, Solomonson M, et al. Variant interpretation using population databases: Lessons from gnomAD. Vol. 43, *Human Mutation*. John Wiley and Sons Inc; 2022. p. 1012–30.
79. Schwarz JM, Rödelberger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Vol. 7, *Nature Methods*. Nature Publishing Group; 2010. p. 575–6.
80. Lim KH, Ferraris L, Filloux ME, Raphael BJ, Fairbrother WG. Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proc Natl Acad Sci U S A*. 2011 Jul;108:11093–8.
81. Judy J, Wang X, Seifuddin F, Tumburu L, Pirooznia M, Thein SL. RNA Seq Profiles and Bioinformatics Validation in a Large Sample of Sickle Cell Disease Patients. *Blood*. 2020 Nov;136:13–4.
82. Tsukahara K, Chang X, Mentch F, Smith-Whitley K, Bhandari A, Norris C, et al. Identification of genetic variants associated with clinical features of sickle cell disease. *Sci Rep*. 2024 Dec;14.
83. Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, et al. CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia. *New England Journal of Medicine*. 2021 Jan;384:252–60.
84. Mungall CJ, McMurry JA, Kohler S, Balhoff JP, Borromeo C, Brush M, et al. The Monarch Initiative: An integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Res.* 2017 Jan;45:D712–22.
85. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012 Jul;40.

86. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nature Methods* 2010 7:4 [Internet]. 2010 Apr [cited 2025 Apr 16];7(4):248–9. Available from: <https://www.nature.com/articles/nmeth0410-248>
87. Liu N, Hargreaves V V., Zhu Q, Kurland J V., Hong J, Kim W, et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell*. 2018 Apr;173:430-442.e17.
88. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016 Jun;17.
89. James Kent W, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996–1006.
90. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief Bioinform*. 2013 Mar;14:178–92.
91. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*. 2019 Jul;47:W171–4.
92. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018 Jul;46:W296–303.
93. Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia, a molecular disease. *Science* (1979). 1949;110:543–8.
94. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec;15.
95. Conran N, Belcher JD. Inflammation in sickle cell disease. *Clin Hemorheol Microcirc*. 2018;68:263–99.
96. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37.

97. Thein SL. Genetic association studies in β -hemoglobinopathies. Vol. 2013, Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program. 2013. p. 354–61.
98. Milton JN, Gordeuk VR, Taylor JG, Gladwin MT, Steinberg MH, Sebastiani P. Prediction of fetal hemoglobin in sickle cell anemia using an ensemble of genetic risk prediction models. *Circ Cardiovasc Genet*. 2014;7:110–5.
99. Milton JN, Gordeuk VR, Taylor JG, Gladwin MT, Steinberg MH, Sebastiani P. Prediction of fetal hemoglobin in sickle cell anemia using an ensemble of genetic risk prediction models. *Circ Cardiovasc Genet*. 2014;7:110–5.
100. Kwiatkowski DP. How malaria has affected the human genome and what human genetics can teach us about malaria. Vol. 77, American Journal of Human Genetics. University of Chicago Press; 2005. p. 171–92.
101. Kaufman DP, Lappin SL. StatPearls. 2019. Physiology, Fetal Hemoglobin.
102. Mikobi TM, Kamuanya NC, Mikobi EKB, Kalela TI, Akilimali PZ, Lukusa PT. Sickle cell anemia and pregnancy: Profile of hemodynamic changes in sickle cell pregnant women in Kinshasa. *EJHaem*. 2023 Nov;4:977–83.
103. Kwiatkowski DP. How malaria has affected the human genome and what human genetics can teach us about malaria. Vol. 77, American Journal of Human Genetics. University of Chicago Press; 2005. p. 171–92.
104. Chou ST, Westhoff CM. Application of genomics for transfusion therapy in sickle cell anemia. *Blood Cells Mol Dis*. 2017 Sep;67:148–54.
105. Oteng-Ntim E, Meeks D, Seed PT, Webster L, Howard J, Doyle P, et al. Adverse maternal and perinatal outcomes in pregnant women with sickle cell disease: Systematic review and meta-analysis. Vol. 125, Blood. American Society of Hematology; 2015. p. 3316–25.
106. Kaufman DP, Lappin SL. Physiology, Fetal Hemoglobin. StatPearls. 2019.
107. Thein SL. Genetic association studies in β -hemoglobinopathies. Vol. 2013, Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program. 2013. p. 354–61.

108. Mtatiro SN, Makani J, Mmbando B, Thein SL, Menzel S, Cox SE. Genetic variants at HbF-modifier loci moderate anemia and leukocytosis in sickle cell disease in Tanzania. *Am J Hematol*. 2015 Jan;90:E1–4.
109. Bhatnagar P, Purvis S, Barron-Casella E, DeBaun MR, Casella JF, Arking DE, et al. Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. *J Hum Genet*. 2011 Apr;56:316–23.
110. Mann CJ. Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. Vol. 20, *Emergency Medicine Journal*. BMJ Publishing Group; 2003. p. 54–60.
111. REGLAMENTO (UE) 2016/ 679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO - de 27 de abril de 2016 - relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/ 46/ CE (Reglamento general de protección de datos).
112. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (Sf-36): I. conceptual framework and item selection. *Med Care*. 1992;30:473–83.
113. International Ethical Guidelines for Health-related Research involving Humans. 2016.
114. Organization WH. WHO guidelines on drawing blood: best practices in phlebotomy. <https://iris.who.int/handle/10665/44294>. 2010.
115. https://www.oecd.org/en/publications/oecd-best-practice-guidelines-for-biological-resource-centres_9789264128767-en.html [Internet]. OECD Best Practice Guidelines for Biological Resource Centres.
116. Piel FB, Steinberg MH, Rees DC. Sickle Cell Disease. *New England Journal of Medicine*. 2017 Apr;376:1561–73.
117. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014 Aug;30:2114–20.
118. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat*. 2016 Mar;37:235–41.

119. Kircher M, Witten DM, Jain P, O’roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46:310–5.
120. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018 Jan;46:D1062–7.
121. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine.* 2015 May;17:405–24.
122. Altman DG, Bland j. M. Statistics notes: The Normal distribution. *BMJ.* 1995 Feb;310:298.
123. Yim KH, Nahm FS, Han KA, Park SY. Analysis of statistical methods and errors in the articles published in the korean journal of pain. *Korean J Pain.* 2010 Mar;23:35–41.
124. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: A guide for non-statisticians. *Int J Endocrinol Metab.* 2012;10:486–9.
125. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. *Journal of Educational Research.* 2002;96:3–14.
126. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression: Third Edition.* Applied Logistic Regression: Third Edition. wiley; 2013. 1–510 p.
127. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol.* 2012 Apr;41:514–20.
128. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016;44:D862–8.
129. Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and β -globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A.* 2008 Aug;105:11869–74.

130. Sebastiani P, Ramoni MF, Nolan V, Baldwin CT, Steinberg MH. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Nat Genet.* 2005 Apr;37:435–40.
131. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. In: *The Lancet.* Elsevier B.V.; 2010. p. 2018–31.
132. Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, Van Handel B, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* (1979). 2008 Dec;322:1839–42.
133. Rodeghiero F. A critical appraisal of the evidence for the role of splenectomy in adults and children with ITP. Vol. 181, *British Journal of Haematology.* Blackwell Publishing Ltd; 2018. p. 183–95.
134. Perseu L, Satta S, Moi P, Demartis FR, Manunza L, Sollaino MC, et al. KLF1 gene mutations cause borderline HbA2. *Blood.* 2011 Oct;118:4454–8.
135. Sebastiani P, Ramoni MF, Nolan V, Baldwin CT, Steinberg MH. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Nat Genet.* 2005 Apr;37:435–40.
136. Giardine B, Borg J, Viennas E, Pavlidis C, Moradkhani K, Joly P, et al. Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. *Nucleic Acids Res.* 2014 Jan;42.
137. Collins FS, Varmus H. A New Initiative on Precision Medicine. *New England Journal of Medicine.* 2015 Feb;372:793–5.
138. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015 Jan;31:2745–7.
139. Chandrakasan S, Malik P. Gene therapy for hemoglobinopathies: The state of the field and the future. Vol. 28, *Hematology/Oncology Clinics of North America.* 2014. p. 199–216.
140. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Vol. 538, *Nature.* Nature Publishing Group; 2016. p. 161–4.

141. Esrick EB, Lehmann LE, Biffi A, Achebe M, Brendel C, Ciuculescu MF, et al. Post-Transcriptional Genetic Silencing of BCL11A to Treat Sickle Cell Disease . New England Journal of Medicine. 2021 Jan;384:205–15.

Índice de acrónimos

- ATC: Cáncer anaplásico de tiroides
- BAM: Formato binario de alineamiento (Binary Alignment Map)
- CADD: Depleción dependiente de anotación combinada (Combined Annotation Dependent Depletion)
- CNV: Variación en el número de copias (Copy Number Variation)
- CRISPR: Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (Clustered Regularly Interspaced Short Palindromic Repeats)
- DAPI: Tinte fluorescente que se une al ADN (4',6-diamidino-2-phenylindole)
- DAR: Regiones diferencialmente accesibles
- dbSNP: Base de datos de polimorfismos de un solo nucleótido
- DEG: Genes diferencialmente expresados
- DNA: Ácido desoxirribonucleico
- EGFR: Receptor del factor de crecimiento epidérmico
- EMT: Transición epitelio-mesénquima
- FASTQ: Formato para almacenar lecturas de secuenciación
- FBS: Suero fetal bovino (Fetal Bovine Serum)
- FDA: Administración de Alimentos y Medicamentos (Food and Drug Administration)
- FDR: Tasa de descubrimiento falso (False Discovery Rate)
- gDNA: ADN genómico
- GTF: Formato de transferencia de genes (Gene Transfer Format)
- GSEA: Análisis de enriquecimiento de conjuntos génicos
- GO: Ontología génica (Gene Ontology)
- gnomAD: Base de datos de agregación del genoma humano
- H&E: Tinción de hematoxilina y eosina
- ID: Identificador

- IFN- γ : Interferón gamma
- IGV: Visor Integrativo de Genómica (Integrative Genomics Viewer)
- INDEL: Inserción y delección
- KEGG: Enciclopedia de genes y genomas de Kyoto
- MAPK: Quinasa activada por mitógenos
- NGS: Secuenciación de nueva generación
- PCA: Análisis de componentes principales
- PBS: Solución salina tamponada con fosfato
- PCR: Reacción en cadena de la polimerasa
- PI3K: Fosfatidilinositol 3-cinasa
- PolyPhen: Fenotipado de polimorfismos
- QC: Control de calidad
- RNA: Ácido ribonucleico
- RPM: Lecturas por millón (Reads Per Million)
- RPMI: Medio de cultivo Roswell Park Memorial Institute
- RPKM: Lecturas por kilobase por millón de lecturas mapeadas
- RT: Temperatura ambiente (Room Temperature)
- scRNA-seq: Secuenciación de ARN de célula única
- SIFT: Clasificación de variantes dañinas o toleradas (Sorting Intolerant From Tolerant)
- SNV: Variante de un solo nucleótido
- snATAC-seq: Secuenciación de accesibilidad de cromatina de núcleos individuales
- SBS: Facultad de Ciencias Biomédicas (School of Biomedical Sciences)
- TCGA: Atlas del Genoma del Cáncer (The Cancer Genome Atlas)
- TF: Factor de transcripción
- TGF- β : Factor de crecimiento transformante beta
- TME: Microambiente tumoral
- TNF- α : Factor de necrosis tumoral alfa
- TSS: Sitio de inicio de transcripción
- TSV: Valores separados por tabulación
- UMAP: Aproximación uniforme y proyección a múltiples escalas
- UMI: Identificador molecular único
- UTR: Región no traducida (Untranslated Region)

- VCF: Formato de llamada de variantes (Variant Call Format)
- VDJ: Segmentos variables, de diversidad y de unión
- VEP: Predictor de efectos de variantes (Variant Effect Predictor)
- VST: Transformación estabilizadora de la varianza
- WNN: Vecino más cercano ponderado (Weighted Nearest Neighbor)