Instituto Nacional de Saúde Doutor Ricardo Jorge

NOVA FACULDADE DE CIÊNCIAS E TECNOLOGIA

Validation of next generation sequencing technology in the identification of complex and rare cases of hemoglobinopathies

Mariana Carmelo Antunes

MSc in Molecular Genetics and Biomedicine

Adviser: Paula Faustino (INSA)
Co-adviser: Alexandra Fernandes (FCT NOVA)

26th November 2025

INTRODUCTION

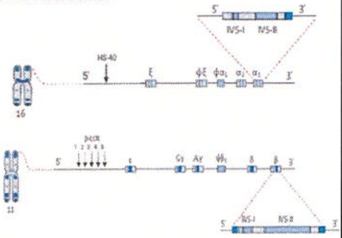
The Human Globin Gene Clusters

α-globin gene cluster (16p13.3)

- **ζ gene (*HBZ*)**, pseudogenes (ψζ1, ψα2, ψα1), **two functional α genes α2 and α1 (*HBA2* and *HBA1*)** and pseudogene θ (*HBQ1*)
- Regulatory element **HS-40** essential for α-globin expression

β-globin gene cluster (11p15.5)

- Functional genes **ε (*HBE*)**, **Gγ (*HBG2*)** and **Aγ (*HBG1*)**, **δ (*HBD*)** and **β (*HBB*)**, and pseudogene ψβ
- Regulatory element **locus control region (LCR)**, with five DNase I hypersensitive sites (HS-1 to HS-5)



INTRODUCTION

Developmental Regulation and Gene Switching

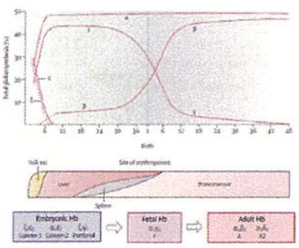
Globin gene expression changes throughout development:

- **Sequential activation** ζ → α and ε → γ → δ/β genes
- Coordinated by the **stage of human development** and **site of erythropoiesis**
- Promotes **differential synthesis** of globin chains
- **α : β chain ratio maintained** through all stages

Regulation coordinated by:

Cis-elements: HS-40, LCR

Trans-factors: *KLF1* (ch:19p13.13), *HBS1L-MYB* (ch:6q23.3), *BCL11A* (ch:2p16.1).



INTRODUCTION

Hemoglobinopathies: Molecular Defects and Clinical Diversity

Mutations or **deletions** in globin genes or their regulatory elements → **Abnormal structure (qualitative)** or **imbalanced synthesis (quantitative)** of globin chains → **Ineffective or unstable hemoglobin**

Structural variants

Thalassemia

HPFH

- **>900 variants described** → highly molecular heterogeneity
- **Clinical spectrum:** asymptomatic carriers → severe transfusion-dependent anemia
- Co-inheritance of different defects → **complex genotypes and phenotypes**

INTRODUCTION

Hemoglobinopathies in Portugal: Diagnostic Context and Workflow

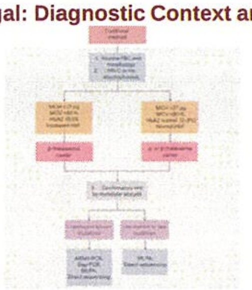
- Endemic in Mediterranean regions, **> prevalence in Portugal due to migration** from Africa, Asia, and the Middle East
- **PNCH, 1986**: improved diagnosis, counselling, and prevention
- **2021–2022 neonatal screening**: Sickle cell disease included

Diagnostic challenges:

- Traditional workflow is sequential, gene-specific and time-consuming
- May miss rare or complex genotypes

↓

Need for comprehensive, faster methods



OBJECTIVES

Objectives

Assess the applicability and reliability of next-generation sequencing (NGS) technology, using the Devyser Thalassemia NGS kit, in identifying genetic variants associated with haemoglobinopathies

1. Apply the **NGS methodology** to the analysis of complex ethiology and rare haemoglobinopathies cases
2. Validate the performance of the Devyser Thalassemia NGS kit, **comparing its results** with those previously obtained by traditional methods
3. Confirm the **variants detected by NGS**, when necessary
4. **Analyse the pathogenicity** of the variants identified

METHODS

Study Design and comparative approach

40 anonymised individuals with confirmed or suspected hemoglobinopathies

Traditional methods:
- Polymerase Chain Reaction (PCR)
- GAP-PCR
- Sanger sequencing
- Multiplex Ligation Probe Amplification (MLPA)

VS

NGS Methodology:
- Devyser Thalassemia NGS kit
- Analysis in the bioinformatic software (Devyser Amplicon Suite v3.5) workflow
- Re-evaluation of some complex cases through Devyser Thalassemia NGS kit v2

Comparative diagnostic performance analysis

7

METHODS

Traditional Molecular Methods

All samples were previously characterized through hematological and biochemical analyses

Provide a molecular starting point for genetic testing

Traditional methods provide **targeted**, **high-accuracy** results but require **prior suspicion** and **multiple assays**

PCR + Sanger Sequencing — Single nucleotide variations (SNVs)/ small indels

MLPA — Copy Number Variations (CNVs)-Large deletions/ duplications

8

METHODS

Devyser Thalassemia NGS

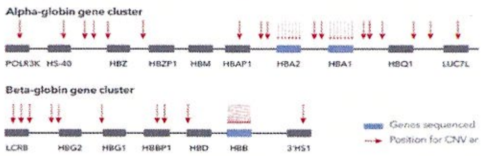
Generation of specific libraries for NGS analysis of SNVs and CNVs

Full coverage: *HBA1, HBA2, HBB* genes, including upstream and downstream regions

Partial coverage: *HBG1, HBG2, HBD* promoters + **distal regulatory regions**

Some common deletions in α and β are detected by direct detection (DD) method

Sequencing: MiSeq Illumina in paired-end (PE) mode (2 × 151 bp)

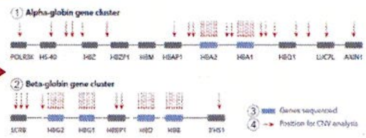


9

METHODS

Devyser Thalassemia NGS v2

Devyser Thalassemia v2: the improved new version of the kit



	Devyser Thalassemia v1	Devyser Thalassemia v2
Globin genes fully sequenced	*HBA1,HBA2,HBB*	*HBA1, HBA2, HBD, HBG1, HBG2, HBB*
Regulatory trans-factors	Not included	*KLF1, HBS1L-MYB, BCL11A*
α3.7 and α4.2 deletions	Through read-depth normalization	Long Range-PCR is employed

10

RESULTS AND DISCUSSION

Overview of Genetic Findings

The study population presented a total of **37 clinically relevant alterations**

Devyser Thalassemia NGS — 77 distinct alterations
- 70 SNVs
- 7 CNVs

31 clinically relevant alterations were found

VS

Traditional Methods — 45 distinct alterations
- 36 SNVs
- 9 CNVs

35 clinically relevant alterations were found

11

RESULTS AND DISCUSSION

Diagnostic Concordance Between Traditional Methods and NGS

Case diagnostic classification	Traditional Methods	Devyser Thalassemia NGS
Fully resolved	37	31
Cases with discrepancies	3	9
Incomplete characterization	3	7
Variants not previously found	3	--

The 9 cases re-evaluated with Devyser Thalassemia v2 were fully resolved

NGS v1 and traditional methods demonstrated substantial agreement (κ = 0.65) for SNVs detection

NGS v1 and traditional methods demonstrated near perfect agreement (κ = 0.95) for CNVs detection

12

RESULTS AND DISCUSSION

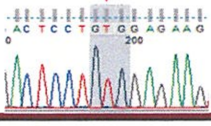
A rare and complex case of triple hemoglobinopathy mimicking Hb C-Harlem

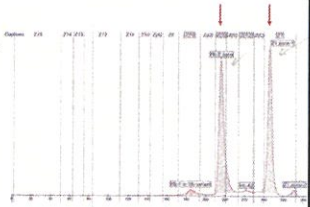
Clinical findings

- Female, anemia, microcytosis and hypochromia
- HPLC: no HbA peak + peaks Hb S (48 %) and near Hb C (42 %) → presumptive Hb S/Hb C-Harlem

Traditional methods

- *HBB*:c.20A>T (Hb S) homozygous
- No other alterations found



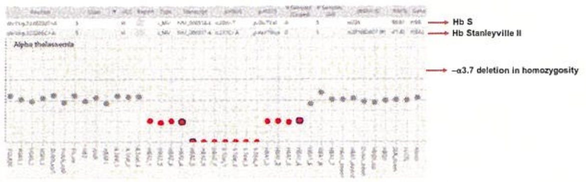


RESULTS AND DISCUSSION

A rare and complex case of triple hemoglobinopathy mimicking Hb C-Harlem

NGS findings

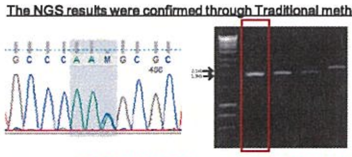
- Detected *HBA2*:c.237C>A (Hb Stanleyville II) in hemizygosity and *HBB*:c.20A>T (Hb S) in homozygosity
- −α3.7 deletion in homozygosity



RESULTS AND DISCUSSION

A rare and complex case of triple hemoglobinopathy mimicking Hb C-Harlem

The NGS results were confirmed through Traditional methods



Cases with complex genotypes could result in phenotypic and biochemical properties similar to other alterations

NGS resolved a misleading phenotype and prevented misclassification

RESULTS AND DISCUSSION

δ-globin variants: diagnostic relevance and NGS v2 improvement

Clinical relevance of delta variants

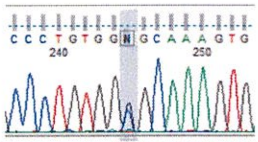
- Hb A_2 is composed of α and δ chains → routinely used to screen for β-thalassemia
- δ-thalassemia or δ-globin variants **alter Hb A_2 quantification → may mask β-thalassemia trait** → compromises initial triage for molecular testing

Several individuals carried ***HBD* alterations** (Hb A_2'/Hb B_2, Hb Babinga, δ-thalassemia) → Frequently co-inherited with **β-thalassemia** → complex phenotypic profiles

Traditional methods

Identified ***HBD*:c.49G>C (Hb A_2')** and ***HBD*:c.410G>A (Hb Babinga)**

Explained atypical Hb A_2 levels and confirmed compound genotypes



RESULTS AND DISCUSSION

δ-globin variants: diagnostic relevance and NGS v2 improvement

Clinical relevance of delta variants

Hb A_2 is composed of α and δ chains → routinely used to screen for β-thalassemia

δ-thalassemia or δ-globin variants **alter Hb A_2 quantification → may mimic or mask β-thalassemia** → compromises initial triage for molecular testing

NGS v1 limitations

Partial *HBD* coverage → **missed δ-variants** in multiple cases

Resulted in incomplete or discordant diagnoses

NGS v2 improvement

Full *HBD* and *HBG* coverage → correct detection of all δ-variants.

9 re-analysed cases = 100% concordance with traditional results

RESULTS AND DISCUSSION

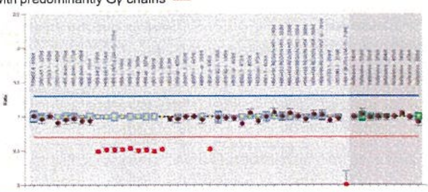
From incomplete to resolved: NGS v2 defines a large δβ deletion

Clinical findings

- Microcytosis and hypochromia
- Normal HbA
- High HbF with predominantly Gγ chains

Traditional methods

- Deletion on HBG1 exon 3 + HBD intron 2 to HBB exon 3

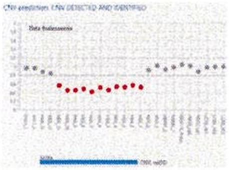


RESULTS AND DISCUSSION

From incomplete to resolved: NGS v2 defines a large δβ deletion

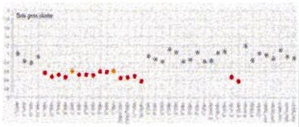
NGS v1 limitations

- Partial coverage → missed HBD and HBG deletion
- Not compatible with Sicilian deletion



NGS v2 improvement

- Full *HBD* and *HBG* coverage → complete characterization of deletions
- Allowed for comprehensive deletion delimitation



19

RESULTS AND DISCUSSION

Diagnostic Performance: Traditional Methods vs Devyser Thalassemia NGS

	Devyser Thalassemia NGS		Traditional Methods	
Variant Type	SNV/Indels	CNVs	SNV/Indels	CNVs
Sensitivity (95% CI)	0.86 (0.72–0.95)	0.89 (0.67–0.99)	0.98 (0.88–1.00)	0.84 (0.60–0.97)
Specificity (95% CI)	1.00 (0.54–1.00)	1.00 (0.85–1.00)	1.00 (0.54–1.00)	1.00 (0.85–1.00)

Both methods showed **perfect specificity**

NGS presents **higher sensitivity for CNV detection**

Traditional methods presented **higher sensitivity for SNV detections**

20

RESULTS AND DISCUSSION

Understanding discrepancies between traditional methods and NGS Performance

Nature of each method

- **Traditional methods** (PCR, Sanger, MLPA) are **targeted**, **high-accuracy**, but limited to suspected regions
- **NGS** performs **broad parallel screening** of multiple genes/regions, allowing **detection beyond presumptive diagnoses**.

Reasons for NGS discordant results

- NGS v1: partial coverage of *HBD* and *HBG* genes → **undetected δ-variants**.
- Some low-frequency CNVs near the detection threshold may need orthogonal confirmation

↓

Issues addressed in Devyser Thalassemia v2

RESULTS AND DISCUSSION

Advancements in Devyser Thalassemia NGS v2

Expanded Gene Coverage

- Complete sequencing of ***HBD***, ***HBG1***, and ***HBG2*** genes.
- Inclusion of **modifier genes** (*KLF1*, *BCL11A*, *HBS1L-MYB*) relevant to Hb F regulation.

Improved CNV Detection

- Added **long-range PCR (LR-PCR)** for direct detection of $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions.
- Enhanced probe density in α- and β-clusters → **greater sensitivity for complex CNVs**

↓

Better detection of δ- and γ-globin variants previously missed in v1

Reduced false negatives and ambiguous CNV calls

Broader clinical interpretation including disease modifiers

9 re-analysed samples → **100% concordance** with traditional methods

RESULTS AND DISCUSSION

Limitations of Traditional and NGS Methodologies

Traditional Methods

- Require a **priori suspicion**
- **Limited detection** of variants outside screened regions
- **Sequential workflow** = time-consuming and labor-intensive
- Potential for **delayed or incomplete diagnosis** in complex genotypes

NGS

- Requires **bioinformatics expertise** and careful variant interpretation
- **Homologous genes** → alignment ambiguity
- May produce **Variants of Uncertain Significance** (VUS)
- **High cost** and need for **specialized infrastructure**

CONCLUSIONS AND FUTURE PERSPECTIVES

Conclusions and Clinical Impact

Diagnostic validation

NGS proved highly reliable for **SNV, small indels, and CNV detection** in hemoglobinopathies.

Devyser v2 reached 100 % concordance with traditional methods, resolving previous blind spots

Technical and analytical advantages

Reduces need for multiple targeted tests and shortens turnaround time.

Detects **rare, complex, and modifier variants** missed by sequential assays.

Incorporation of modifier genes (*KLF1, BCL11A, HBS1L-MYB*) allows refined phenotype prediction.

Challenges and future directions

High sequence homology (*HBA1/HBA2, HBG1/HBG2*) still limits gene-level resolution.

Interpretation of VUS and benign variants increases analytic complexity.

Infrastructure and cost may restrict implementation in smaller diagnostic centres.

Traditional methods remain the gold standard for confirmation, but NGS offers speed and breadth.
Combined workflow = most efficient and balanced approach for routine diagnostics.

Acknowledgments

Instituto Nacional de Saúde Doutor Ricardo Jorge

REPÚBLICA PORTUGUESA | SAÚDE | SNS

NOVA FACULDADE DE CIÊNCIAS E TECNOLOGIA

Alexandra Fernandes

DGH

UID-PF:
Paula Faustino
Pedro Lopes

UTI:
Luís Vieira
Daniela Santos

UMO

DPS
Armandina Miranda
Cristina Vieira
Gisela Gaspar

25