Egypt. J. Chem. Vol. 67, SI: M. R. Mahran, pp. **1519 - 1531** (2024)

Egyptian Journal of Chemistry

http://ejchem.journals.ekb.eg/

Thalassemia: Next Generation (NGS) and Third Generation Sequencing (TGS) for the Diagnosis.

Moteab Abdulmohsen Althomari [1] , Ibrahim Taher Bohassan [2], Zahra Hajji Bohassan [3], Fayez Taher Alhajouji [4], Ebtihal Lafi M Alhejaili [5], Talal Jubayr Alharthi [6], Ahmed Mohammed Abdu Sofyani [7], Maram Saad Aljohani [8], Amal Faisal Khalaf Al Anzi [9], Hassan Mohammed Ali Hazazi [10], Adel Ahmed Harthi [10], Sajidah Saeed Abdrabuh Aljayzani [11], Ahmed Awadh Mohammed Albarqi [12], Hassan Mohammed Ali Nasseri [13], Safar Saad Alamoudh [14]

[1] KSA, Ministry of Health, King Abdullah Medical Complex Jeddah, Kingdom of Saudi Arabia
[2] KSA, Ministry of Health, King Fahad Hospital in Hofouf
[3] KSA, Ministry of Health, community laboratory
[4] KSA, Ministry of Health, Alyamamah hospital
[5] KSA, Ministry of Health, Alharam hospital
[6] KSA, Ministry of Health, Prince Salman bin Mohammed Hospital in Dalam
[7] KSA, Ministry of Health, Alardh General Hospital
[8] KSA, Ministry of Health, Almadinah regional laboratory
[9] KSA, Ministry of Health, Al Hamra Health Center
[10] KSA, Ministry of Health, AlHurrath General Hospital
[11] KSA, Ministry of Health, Prince Muhammad Bin Nasser Hospital in Jazan
[12] KSA, Ministry of Health, Mohayel General Hospital
[13] KSA, Ministry of Health, King Salman Bin Abdulaziz Hospital
[14] KSA, Ministry of Health, Prince Salman hospital.

Abstract

Background: Thalassemia is a genetic blood disorder caused by mutations in the α- or β-globin gene. Accurate and timely diagnosis is critical for appropriate management and genetic counseling. Traditional diagnostic techniques, such as PCR-based methods, often have limitations, including the inability to detect all genetic variants, long processing times, and high costs. Recent advances in sequencing technologies, specifically Next-Generation Sequencing (NGS) and Third-Generation Sequencing (TGS), offer potential improvements for thalassemia diagnosis by allowing comprehensive mutation detection, including single nucleotide variants (SNVs), insertions/deletions (indels), and copy number variations (CNVs).

Aim: This study aims to review the current diagnostic approaches for thalassemia and evaluate the advantages of NGS and TGS for identifying complex genetic variations associated with both α- and β-thalassemia.

Methods: A detailed review of conventional DNA analysis methods, such as Reverse Dot-Blot Analysis, Gap-PCR, ARMS-PCR, Sanger Sequencing, and MLPA, was conducted. Additionally, the latest advancements in NGS and TGS technologies were discussed, focusing on their ability to simultaneously detect SNVs, indels, and SVs. Case studies of thalassemia diagnoses using NGS and TGS technologies, including detection of rare and complex mutations, were analyzed.

Results: The study found that NGS and TGS technologies significantly improve the accuracy of thalassemia diagnosis by enabling the simultaneous detection of multiple mutation types. These technologies allow for faster and more comprehensive diagnoses, reduce the need for multiple diagnostic tests, and can detect complex mutations that are missed by conventional methods. Additionally, these methods require minimal DNA and offer reduced costs per sample through multiplexing.

Conclusion: NGS and TGS are emerging as valuable tools for the diagnosis of thalassemia, providing significant improvements in the detection of rare and complex mutations. While conventional methods remain important, the integration of NGS and TGS into routine diagnostic workflows will enhance the accuracy and efficiency of thalassemia diagnosis and genetic counseling.

Keywords: thalassemia, NGS, TGS, genetic mutations, diagnostic methods, α-globin, β-globin, genetic counseling, mutation detection..

1. Introduction

Pediatricians Thomas B. Cooley and Pearl Lee of Detroit gave the first clinical description of thalassemia when they discovered a severe form of anemia in children that was later dubbed Cooley's anemia and was marked by splenomegaly and bone abnormalities [1]. Cooley's anemia with erythroblastic anemia was referred to as "thalassemia" by Whipple and Bradford in 1936 [2]. Since most of the afflicted people were of Mediterranean ancestry, the word comes from Greek and means "sea" and "blood." The disease was later discovered to be widespread in the Middle East, Southeast Asia (SEA), and the Indian subcontinent.

Mutations that affect normal hemoglobin production—the main protein in red blood cells

*Corresponding author e-mail: malthomari@moh.gov.sa, (Moteab Abdulmohsen Althomari).
Receive Date: 15 November 2024, Revise Date: 06 December 2024, Accept Date: 10 December 2024
DOI: 10.21608/ejchem.2024.336711.10811

(RBCs) that carries oxygen from the lungs to tissues—cause thalassemia and other hereditary hemoglobinopathies. In order to maintain balanced production in erythropoietic cells, the synthesis of the two pairs of α- and β-globin chains that make up adult hemoglobin is closely coordinated [3]. An imbalance in α/β-globin occurs when one or both copies of the β-globin gene are unable to create functional β-globin, causing α-globin to continue producing normally. This imbalance causes free α-globin chains to accumulate in homozygous forms, where they produce hazardous inclusions [4]. There are two primary types of thalassemia: α-thalassemia and β-thalassemia. There are four phenotypes for the former: trait, silent carrier, HbH illness, and Hb Bart's disease. In contrast, β-thalassemia is divided into primary phenotypes, silent carriers (β++), carriers, and intermedia. People with a β-thalassemia mutation who have a normal β-globin allele and are clinically asymptomatic are referred to as having β-thalassemia mild or trait. Lifelong blood transfusions are necessary for survival in cases of severe anemia in infancy caused by β-thalassemia major [5]. This disorder causes intravascular hemolysis, erythroid precursor destruction, and inefficient erythropoiesis by aggregating extra α-globin chains into inclusion bodies that harm RBC membranes [6]. With clinical characteristics ranging from modest to serious [7], β-thalassemia intermedia is a milder variant that frequently affects homozygotes or compound heterozygotes for β-thalassemia [8]. Rarely, co-inheritance of genetic abnormalities, such as a segmental duplication of the α-globin gene, can cause symptoms in carriers of β-thalassemia. This can worsen the α/β-globin imbalance and result in a more severe phenotype [9,10].

Transfusion-dependent thalassemia (TDT) and non-transfusion-dependent thalassemia (NTDT) are new classifications that have been developed due to the complexity of phenotypic classification in thalassemia, especially in cases with concurrent α- and β-thalassemia mutations [11]. While NTDT comprises individuals who do not require continuous transfusions, although occasional or periodic transfusions may be necessary in specific clinical conditions [13], TDT refers to patients who need lifelong frequent blood transfusions for survival [12]. Hemoglobin E/β-thalassemia (mild and moderate variants), α-thalassemia intermedia (HbH illness), hemoglobin S/β-thalassemia, hemoglobin C thalassemia, and β-thalassemia intermedia are the five different types of NTDT [11]. α- and β-thalassemia are linked to mutations in the HBA and HBB genes, respectively. With more than a thousand different variations, including single nucleotide variations (SNVs), insertions and deletions (indels), segmental deletions and duplications, and inversions, these changes are extremely diverse [14,15,16]. While segmental deletions are prevalent in α-thalassemia, SNVs are the primary cause of β-thalassemia. In addition to being less commonly linked to inversions or donor insertions, deletions in the HBB cluster can cause diseases like high persistence of fetal hemoglobin (HPFH), delta beta (δβ-thalassemia), or β-thalassemia. Compared to the β-globin cluster, the α-globin cluster has a higher frequency of segmental duplications. Because of the disease's diverse phenotypic manifestation, molecular diagnostics is vital for verifying diagnoses, directing treatment, and offering genetic counseling, making it a critical component of thalassemia prevention and therapeutic management.

Conventional DNA Analysis

Before PCR was invented, Southern transfer and hybridization or Southern-blot analysis [19,20], linkage analysis, and restriction fragment length polymorphism (RFLP) [17,18] were used to diagnose β-thalassemia. Despite being initially used in 1971 [21], enzymatic DNA amplification was first used in clinical settings to diagnose sickle cell anemia in fetuses. Phosphorus-32 (^32P)-labeled oligonucleotide probes were used to increase the sensitivity of DNA hybridization using as little as 20 ng of DNA, or less than 100 times the typical amount [22]. Later, this enzymatic amplification method was dubbed the polymerase chain reaction (PCR) [23], which allowed for the use of non-radioactive probes for dot-blot analysis, such as oligonucleotides tagged with horseradish peroxidase [24, 25].

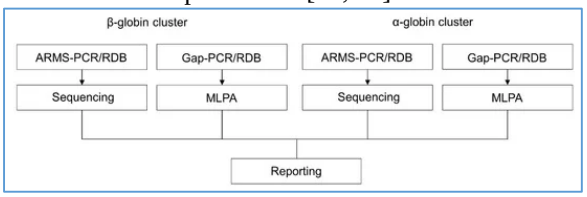

Figure 1: Conventional DNA Analysis.

Reverse Dot-Blot Analysis

The development of reverse dot-blot analysis followed, in which sequence-specific oligonucleotide probes labeled with horseradish peroxidase were attached to a nylon membrane to allow for simultaneous hybridization with several sequences [26]. Numerous populations have used this modified reverse dot-blot technique [27-34]. Three steps are usually involved in reverse dot-blot hybridization: first, allele-specific oligonucleotide probes are immobilized onto a nylon membrane; next, a biotinylated primer is used to amplify the target DNA region; and finally, streptavidin-alkaline phosphatase and colorimetric substrates are used to hybridize the biotinylated DNA with the probe-bound membrane.

Gap-PCR

In deletional types of thalassemia, the extra segment is amplified using gap-PCR. It is possible to estimate the precise deletional types based on the size of the amplicons produced. In multiplex detection of common α-thalassemia deletions [35], as well as αααanti 3.7 and αααanti 4.2 triplications [36], gap-

PCR is most frequently used. Additionally, HPFH, δβ-thalassemia, and β-thalassemia are examples of deletional variants of both α- and β-thalassemia that can be detected by gap-PCR [37,38,39].

Amplification Refractory Mutation System (ARMS) or Allele-Specific Polymerase Chain Reaction (ASPCR)

For the DNA investigation of sickle cell anemia, Wu and associates developed a straightforward allele-specific PCR (ASPCR) method that did not call for blot hybridization or enzyme digestion [40]. This idea was eventually developed into the amplification refractory mutation system (ARMS) [41], which matches the intended point mutation exactly by using allele-specific primers with a modified 3′ end. By stopping Taq polymerase from extending the primer, a mismatch at the fourth nucleotide at the 3′ end improves specificity. Primers that exactly match the wild-type sequence are used to determine homozygosity or heterozygosity, which allows ARMS to detect tiny indels and single nucleotide variants (SNVs). Allele-specific PCR is typically used in ARMS to amp up DNA fragments, and the amplified results are then subjected to gel electrophoresis. Many populations have used customized multiplex ARMS [33,41,42,43,44,45,46,47,48,49,50,51,52].

Sanger Sequencing

The introduction of Sanger sequencing in 1975 made it possible to identify certain DNA sequences [53]. Later, automated data capture [55] took the place of autoradiography, and fluorophore-labeled dNTPs covalently coupled to oligonucleotide primers for DNA sequencing [54] replaced isotope labeling. The development of PCR made it possible to identify uncommon β-thalassemia mutations by sequencing particular DNA regions [56,57]. A major development in sequencing technologies occurred in 1990 when capillary electrophoresis (CE) was used to distinguish fluorescently labeled DNA fragments [58]. Following targeted PCR, RFLP, and dot-blot investigations [29,33,50,51,59], this sequencing method emerged as the gold standard for identifying thalassemia mutations and examining copy number variations (CNVs).

Multiplex Ligation Probe-Dependent Analysis

First used to detect the Hepatitis C virus [60], ligation-dependent PCR was later patented in 1996 as multiplex ligation-dependent amplification (MLPA) [61]. Since then, MLPA has been used to identify a variety of illnesses, and MRC Holland (Amsterdam, The Netherlands) has increased its commercial uses in methylation status, tumor profiling, and inherited disorders. Two probes make up each MLPA probe set: one has the target-matching sequence linked to a universal primer, and the other has a stuffer sequence to produce a distinct amplicon following ligation, amplification, and hybridization. When probes are hybridized to target DNA, the probe signal can be amplified quantitatively, which is very helpful for identifying single nucleotide polymorphisms (SNPs). MLPA has been used to identify deletions in HBG1-HBG2 [64], εγδβ-thalassemia [70,71], significant segmental duplications in β-thalassemia carriers [65,66,67,68,69], and α-thalassemia [63,64].

Advanced Molecular Techniques in Single-Assay DNA Analysis

Two basic methods have been the mainstay of DNA analysis of Mendelian diseases, including thalassemia, for many years: finding rare mutations using Sanger sequencing and Multiplex Ligation-dependent Probe Amplification (MLPA) and detecting common mutations using a variety of targeted DNA analysis techniques. Sanger sequencing and MLPA, respectively, are used in comprehensive DNA analysis to detect variations in the HBA, HBB, and HBD genes as well as CNV. The HBB gene frequently exhibits indels, necessitating the use of extra sequencing reads in order to handle the frameshifted, heterozygous mixed reads in Sanger sequencing. Homozygous variations require extra testing to rule out possible deletions on the opposite allele since they must be considered. The HBA2 gene has been amplified [35] to identify deletional thalassemia, and cluster-spanning amplicons have been used [37,38,39] to verify that the other allele is normal. To rule out uncommon CNVs, homozygous deletions found by Gap-PCR must be validated by MLPA. In order to find CNVs, MLPA uses spanning probes to compare the sample to a known reference. Even while MLPA is usually easy to understand, its small number of probes prevents it from being a stand-alone method. Because the HBA and HBB gene clusters have a large number of accessible probes, long-range PCR can be used to validate deletions and duplications within these clusters. However, because there are so few MLPA probes outside of the HBA and HBB clusters, it is still difficult to identify breakpoints for bigger deletions and duplications.

Next-Generation Sequencing (NGS)

High-throughput sequencing was made possible by next-generation sequencing (NGS), a groundbreaking technology that surfaced two years after the Human Genome Project (HGP) was finished. Small DNA fragments (150–1000 bp) can be sequenced in parallel using NGS. These fragments can be utilized for specific techniques like methylation, RNA, or exome sequencing, or for whole-genome sequencing (WGS). Platform-specific sequencing chemistry varies; for example, Illumina sequencing employs a unique reversible terminator and clonal array creation for large-scale sequencing. This technique creates millions of DNA clusters by immobilizing adaptor and index-ligated DNA fragments on a flow cell and isothermally amplifying each fragment. Sequence by synthesis (SBS) is determined by illuminating these clusters and using the fluorescent signal that is released. While the emission wavelength and signal intensity aid in determining the base calls, the number of cycles

determines the read length. The origin of each sequencing read is ascertained by mapping the reads to a reference sequence prior to variant calling. The Scalable Nucleotide Alignment Program (SNAP) [76], Bowtie2 [74], minimap2 [75], Burrow-Wheeler Aligner-Maximal Exact Match (BWA-MEM) [73], and Genome Analysis Toolkit (GATK) [72] are popular mappers. By compressing, sorting, identifying duplicates, and indexing the output, SNAP streamlines the procedure. While some programs, like GATK [72], FreeBayes [77], and SAMtools [78], use Bayesian techniques to call variants and indels, others, such as DeepVariant, use deep neural networks [79], and Strelka2 uses a unique mixture-model-based approach [80].

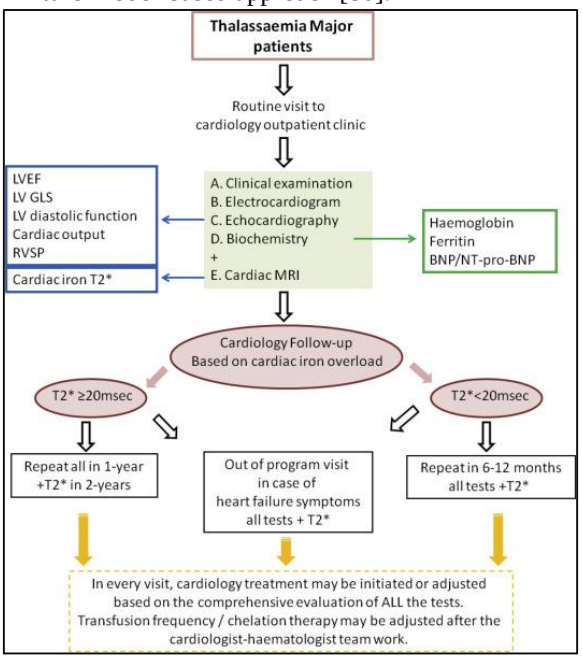

Figure 2: Biochemical and Imaging Markers of Thalassemia.

However, short lengths and GC-content biases make it difficult to identify significant rearrangements and CNVs using NGS data. There are a number of techniques to get around these restrictions, including CNVkit, which employs bias correction using rolling median and circular binary segmentation (CBS) for CNV calling [84], DELLY, which combines split-read analysis with short- and long-range paired-end mapping [83], and Control-FREEC, which uses a LASSO-based algorithm [82]. ExomeDepth, which combines likelihood values across multiple exons using a beta-binomial model and the hidden Markov model (HMM) [85], CoNIFER, which normalizes read depth using singular value decomposition (SVD) and generates z-scores for CNV detection [86], and FishingCNV, which uses principal component analysis (PCA) and CBS tests for CNV analysis [87], are tools specifically made for CNV analysis from exome sequencing data. A graphical software program called FishingCNV is also accessible.

The interpretation of output text formats, which frequently lack genetic or clinical annotations, presents another difficulty in CNV analysis. In response, a number of interactive and dynamic CNV visualization tools have been created. For example, SG-ADVISER provides a web server and annotation pipeline that exposes information on genomic variations, including CNVs, both known and anticipated [91]. Similar to this, inCNV is a web-based program that creates annotations from Ensembl, the Database of Genomic Variants (DGV), ClinVar, and OMIM [90] and combines CNV results from several techniques. Additional tools that offer interactive representations of CNV data, such reconCNV and CNVxplorer, enable functional evaluations of CNVs in a clinical diagnostic setting [92][93]. R tools that visualize copy number and beta allele frequency data include KaryoploteR [94], Gviz [95], and CopyNumberPlots [96]. Because the relevant clusters are tiny, targeted sequencing (TS) is a more economical method for analyzing thalassemia. In addition to detecting point mutations in the HBA, HBB, HBD, and HBG genes, a perfect TS strategy should incorporate uniform reads covering neighboring genes for breakpoint analysis and unusual CNV detection. It has been demonstrated that TS performs better than conventional DNA analysis techniques. For instance, simultaneous genotyping of globin genes and genetic modifiers utilizing the GATK variant calling pipeline and an internal CNV tool produced better results than standard screening techniques [97].

Furthermore, TS showed genotype agreement with traditional PCR results when employing Ion Torrent to genotype the HBB gene and common deletional α-thalassemia variants (-α3.7 and -α4.2) [98]. Compared to conventional screening methods based on MCV, MCH, and HbA2 levels, a combination of gap-PCR for α-thalassemia deletions and NGS for variant genotyping of the HBA and HBB genes produced a greater sensitivity [99]. Additionally, prioritized CNV genotyping in the α-globin cluster in conjunction with HBA and HBB variant analysis resulted in greater sensitivity and specificity [100,101]. RPKM analysis and a tailored targeted panel were used to characterize Inv-Del English V, a novel deletion in the HBB cluster [71]. Comparative genomic hybridization (CGH) and MLPA were used to identify a new variation of εγδβ-thalassemia that NGS had overlooked [102], while CNV analysis was not described. It is still difficult to map short reads in highly homologous regions, including the HBA1, HBA2, HBG1, and HBG2 genes, despite advancements in NGS. A specialized bioinformatics pipeline called NGS4THAL was created to realign ambiguously mapped reads from hemoglobin gene clusters by combining several CNV finding algorithms in order to increase genotyping sensitivity and specificity [104]. Furthermore, phasing

of the entire genome is made possible by the use of long-read sequencing technologies, such as link-read sequencing, which also yield useful haplotype information. Preimplantation genetic testing for embryo selection in carriers of α-thalassemia has shown this [106].

Third-Generation Sequencing (TGS)

Single-molecule sequencing (SMS), which is used in third-generation sequencing (TGS), avoids the need for previous clonal amplification of DNA by directly sequencing individual DNA or RNA strands inside a sample [107]. This method allows for continuous DNA synthesis by continuously incorporating fluorescently tagged deoxyribonucleoside triphosphates (dNTPs) into DNA polymerase [108]. By ligating adapters to double-stranded DNA and converting it into circular single-stranded DNA (SMRTbell), Pacific Biosciences (PacBio) uses single-molecule real-time (SMRT) isoform sequencing (Iso-Seq) to create a DNA library. A SMRT cell, a device with zero-mode waveguide (ZMW) wells, is then loaded with this library. A mounted polymerase in each ZMW well starts DNA replication and generates a fluorescent pulse that can be understood [109]. Circular Consensus Sequence (CCS) reads are produced by the Sequel II device.

By repeatedly sequencing the same molecule, HiFi reads—highly accurate long reads—are produced, improving the accuracy of single nucleotide variation (SNV) calling [110]. On the other hand, DNA molecules are translocated through tiny pores in nanopore-based SMS, which only permit single-stranded DNA to pass through in a rigid linear sequence [107]. As DNA fragments travel through protein nanopores implanted in a semi-synthetic insulating membrane, Oxford Nanopore Technologies (ONT) sequencers pick up variations in ionic current. A DNA fragment, an adapter-bound motor protein, and a tethering molecule that binds the DNA to the membrane and nanopore make up a library [111,112]. The motor protein directs the bases through the pore and regulates the rate of DNA translocation. MinKNOWTM software (version 20.10, Oxford, UK) records the "squiggle," a constant current change caused by the DNA as it passes through the nanopore. A neural network approach is then used to convert this squiggle into nucleotides in real time utilizing graphical processing units (GPUs) [113]. Because a single read can span across exons, genes, pseudogenes, highly duplicated regions, and copy number variants (CNVs), long-read sequencing is very useful for identifying CNVs. The greater indel error rate of this approach, however, presents difficulties [114,115]. To solve these problems, a number of cutting-edge mistake correction techniques are available and can be used while optimizing pipelines [116]. Either before or after genome assembly, error repair can be carried out. Before refining or fixing the assembly, assemblers like Flye [117], wtdbg2 [118], Shasta [119], and CONSENT [120] use the minimap2 pairwise aligner [75] to align the raw data. On the other hand, read errors are corrected before assembly using MECAT (an in-house aligner) [121], Canu [122] (using the MinHash Alignment Process (MHAP) aligner [123]), Falcon [124] (using the BLASR aligner [125]), and NECAT (an in-house assembly module) [126].

It has been demonstrated that variant calling for ONT data is improved using deep neural network (DNN) techniques such PEPPER-Margin-DeepVariant [127], NanoCaller [128], and Clair3-trio [129]. Similarly, GATK HaplotypeCaller [72,110], DeepVariant [79], and HELLO [130] have been shown to provide reliable variant calling for HiFi readings. For SMRT read analysis, PacBio's proprietary structural variant (SV) calling tool, pbsv, can be used. There are numerous CNV callers that support both ONT and SMRT reads. For CNV detection, Sniffles2 employs coverage-adaptive filtering, quick consensus sequence, and adaptive clustering (repeat-aware) [131]. While SVIM finds tandem duplications, interspersed duplications, and new element insertions [133], CuteSV2 uses heuristic signature purification and a specially made scanning line for CNV calling [132]. NanoVar's neural-network-based approach finds SVs in whole-genome sequencing (WGS) data with low read depth [134]. While library preparation, sequencing of the prepared library, quality assessment, and read trimming are the first steps in both NGS and TGS procedures, the assembly process is different. SNV, indel, and SV calling must come after reading polishing, which can be done either before or after mapping in TGS.

Targeted sequencing of certain amplicons to genotype common -thalassemia deletions (-3.7, -4.2, --SEA) and HBA and HBB single nucleotide variations (SNVs) using Circular Consensus Sequencing (CCS) showed perfect agreement with conventional PCR-based genotyping techniques [135]. Later, this method was changed to make it possible to genotype other deletional types of thalassemia. Targeted SMRT sequencing was used in this improved approach, called Comprehensive Analysis of Thalassemia Alleles (CATSA), which was evaluated on 1,759 samples and effectively identified copy number variations (CNVs), insertions/deletions (indels), and common and unusual thalassemia SNVs [136]. Later, 100 samples with aberrant hematological parameters that had not yielded any useful results from traditional genetic diagnostic techniques such reverse dot blot (RDB) and gap-PCR (genotyping for -α3.7, -α4.2, and --SEA) were genotyped using CATSA. In these situations, ten uncommon mutations were found [137]. More recently, Li et al. showed that different deletional and non-deletional α-thalassemia mutations can produce HbH disease. They also showed that α-thalassemia with point mutations and indels in β-thalassemia can also cause HbH disease [138]. A number of significant mutations were identified through the use of MLPA and SMRT sequencing. These included a deletion of

27,311 bp in the α-globin gene cluster (--27.3/αα), a deletion of 16,079 bp in the HS-40 region, a rearrangement of -α3.7α1α2 on one allele, a deletion of 4924 bp in the HBG1-HBG2 β-globin gene cluster, and a deletion of 15.8 kb in α-thalassemia [63,64].

In non-invasive prenatal testing (NIPT), 12 out of 13 fetal thalassemia statuses were successfully genotyped using a long-range 20 kb amplicon sequenced by Oxford Nanopore Technologies (ONT) and next-generation sequencing (NGS) to phase parental haplotypes for fetal inherited haplotypes via relative haplotype dosage (RHDO) analysis [139]. 100% agreement for HBB genotyping was found in a small-scale research in Tanzania that compared ONT with Sanger sequencing [140]. The viability of ONT sequencing for preimplantation genetic testing (PGT) was illustrated by Liu et al.'s successful genotyping of homozygous --SEA deletion embryos using straightforward read density plots throughout the HBA locus [141].

Discussion

DNA analysis has always been guided by differential diagnosis based on the patient's phenotypic and hematological characteristics. However, because of the limits of traditional DNA analysis tests and the variation in phenotypes, this approach may be unclear. Large deletions are found using gap-PCR or MLPA, whereas point mutations and insertions/deletions (indels) are usually found using ARMS-PCR or sequencing. ARMS-PCR and sequencing results for homozygous β-thalassemia may not always show a real homozygote, but rather a compound heterozygous state with a deletional or δβ-thalassemia mutation. To confirm the diagnosis in these situations, further testing with gap-PCR, MLPA, or cascade screening is necessary. Complex genotyping raises the risk of misdiagnosis, especially in situations like moderate β-thalassemia/δ-thalassemia with normal HbA2 levels. These numerous diagnostic techniques increase laboratory turnaround time (TAT) and are labor-intensive. Targeted genomic sequencing (TGS) and next-generation sequencing (NGS), on the other hand, have the benefit of concurrently identifying mutations in copy number variations (CNVs), indels, and single nucleotide variants (SNVs). This feature enhances the accuracy of DNA analysis and gives a better knowledge of the genotype–phenotype link by enabling more thorough genotyping of both α- and β-thalassemia and uncommon variations.

The fact that NGS and TGS can handle sample multiplexing and require less DNA input boosts throughput while lowering costs and turnaround times per sample. However, library preparation and sequencing, which need smaller amounts of DNA, are the main causes of the high costs connected with NGS and TGS. Notwithstanding its inability to match the resolution provided by NGS and TGS, conventional DNA analysis is more economical in comparison, mostly because it uses less expensive reagents (such as PCR master mix). The bioinformatics analysis, which might be technically complex, is one particular difficulty with these more recent technology. Haplotype phasing for short reads can be achieved using pangenomic mapping, even though TGS permits phasing during genome assembly. By combining whole-genome sequences from different people, a pangenome provides a more thorough depiction of genetic variation. The reference pangenome is anticipated to assist in mitigating the biases and inaccuracies associated with the single linear reference genome (GRCh38), which is being managed by the Human Pangenome Reference Consortium (HPRC) [142,143,144]. Graph-based pangenome mapping techniques and pipelines, however, are still developing and might not be widely available. Because NGS and TGS may concurrently detect SNVs, indels, and structural variations (SVs), they are superior to traditional DNA analysis. A single test needs uniform readings, long reads for haplotype phasing of homologous genes (like HBA and HBG), and breakpoint-spanning readings to directly detect common deletions and duplications because of the heterogeneity of mutations. The detection of heterogeneous thalassemia mutations will be further improved by advances in sequencing technologies, including Illumina's Complete Long-Read technology, the soon-to-be PacBio Revio long-read system, Onso short-read systems, and state-of-the-art error correcting techniques for TGS.

Notwithstanding these benefits, the complexity and usability of NGS and TGS technologies make them unlikely to displace traditional screening and PCR-based genotyping. It can be challenging to diagnose complex thalassemia cases using NGS, such as those with the HKαα allele (which has both -α3.7 and αααanti4.2 deletions). Additionally, traditional PCR is still a useful method for confirming TGS and NGS findings. By using blood test results before DNA analysis, traditional differential diagnosis enables genotype-phenotype correlation and reduces sample handling errors. The primary advantage of NGS and TGS is their capacity to genotype α- and β-thalassemia at the same time, offering a comprehensive diagnosis of thalassemia and its genetic modifiers—a crucial component of genetic counseling. Sequencing data can also be saved for later analysis if needed. Nevertheless, start-up labs may find the infrastructure needed for these technologies—which includes processing power, qualified bioinformatics technicians, well-established target sequencing (TS) methods, and efficient bioinformatics pipelines—to be expensive. These expenses should go down as sample throughput rises, increasing the technology's accessibility. A number of guidelines and suggestions have been established to enable the integration of these technologies into routine diagnostic practice in order to facilitate their clinical application [145-148].

Conclusion:

The rapid evolution of sequencing technologies has revolutionized the genetic diagnosis of thalassemia. Conventional DNA analysis methods, such as Reverse Dot-Blot Analysis, Gap-PCR, ARMS-PCR, Sanger sequencing, and Multiplex Ligation Probe-Dependent Analysis (MLPA), have been the cornerstone of thalassemia diagnosis for years. While these techniques have served their purpose, they have limitations, particularly in detecting complex mutations and variations like large deletions and rare variants, and they are often time-consuming and labor-intensive. In contrast, Next-Generation Sequencing (NGS) and Third-Generation Sequencing (TGS) have demonstrated significant potential in overcoming these limitations. NGS allows for the simultaneous detection of single nucleotide variants (SNVs), insertions and deletions (indels), and copy number variations (CNVs), providing a more comprehensive picture of the thalassemia genotype. The ability of TGS to generate long reads and facilitate haplotype phasing further enhances its potential to detect structural variations (SVs) and complex mutations, which are often missed by traditional methods. These advanced sequencing technologies offer several key advantages, including the ability to analyze multiple genetic markers simultaneously, reduce laboratory turn-around time (TAT), and minimize the DNA sample required. By increasing throughput and reducing costs per sample through multiplexing, NGS and TGS provide a more cost-effective and efficient approach to genetic testing. Furthermore, these technologies have the potential to detect rare and complex mutations associated with both α- and β-thalassemia, improving the overall accuracy of diagnosis and enabling more personalized treatment strategies. Despite their promising capabilities, NGS and TGS technologies are not without challenges. The bioinformatics analysis required for interpreting sequencing data can be complex and requires skilled technicians. Additionally, the infrastructure needed for these technologies can be costly for start-up laboratories. Nevertheless, as sequencing technology continues to improve and costs decrease, NGS and TGS are expected to become more accessible and integrated into clinical practice. In conclusion, NGS and TGS are transforming the diagnostic landscape for thalassemia, offering unprecedented accuracy, efficiency, and the ability to detect complex mutations. These technologies complement traditional methods and, when combined with conventional PCR-based genotyping, can provide a comprehensive approach to thalassemia diagnosis and genetic counseling, ultimately improving patient outcomes.

References:

1. Cooley, T.B.; Witwer, E.R.; Lee, P. Anemia in Children with Splenomegaly and Peculiar Changes in the Bones Report of Cases. *Am. J. Dis. Child.* **1927**, *34*, 347–363.
2. Whipple, G.H.; Bradford, W.L. Mediterranean Disease-Thalassemia (Erythroblastic Anemia of Cooley): Associated Pigment Abnormalities Simulating Hemochromatosis. *J. Pediatr.* **1936**, *9*, 279–311.
3. Ribeil, J.-A.; Arlet, J.-B.; Dussiot, M.; Cruz Moura, I.; Courtois, G.; Hermine, O. Ineffective Erythropoiesis in β-Thalassemia. *Sci. World J.* **2013**, *2013*, 394295.
4. Khandros, E.; Thom, C.S.; D'Souza, J.; Weiss, M.J. Integrated Protein Quality-Control Pathways Regulate Free α-Globin in Murine β-Thalassemia. *Blood* **2012**, *119*, 5265–5275.
5. Galanello, R.; Origa, R. Beta-Thalassemia. *Orphanet J. Rare Dis.* **2010**, *5*, 11.
6. Yuan, J.; Angelucci, E.; Lucarelli, G.; Aljurf, M.; Snyder, L.M.; Kiefer, C.R.; Ma, L.; Schrier, S.L. Accelerated Programmed Cell Death (Apoptosis) in Erythroid Precursors of Patients with Severe β-Thalassemia (Cooley's Anemia). *Blood* **1993**, *82*, 374–377.
7. Nathan, D.G.; Oski, F.A. *Hematology of Infancy and Childhood*, 4th ed.; W.B. Saunders: Philadelphia, PA, USA, 1993
8. Galanello, R.; Cao, A. Relationship between Genotype and Phenotype. *Ann. N. Y. Acad. Sci.* **1988**, *850*, 325–333.
9. Goossens, M.; Dozy, A.M.; Emburyt, S.H.; Zachariadest, Z.; Hadjiminast, M.G.; Stamatoyannopoulos, G.; Kan, Y.W.A.I. Triplicated A-Globin Loci. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 518–521.
10. Henni, T.; Belhani, M.; Morle, F.; Bachir, D.; Tabone, P.; Colonna, P.; Godet, J. Alpha Globin Gene Triplication in Severe Heterozygous Beta Thalassemia. *Acta Haematol.* **1985**, *74*, 236–239.
11. Weatherall, D.J. The Definition and Epidemiology of Non-Transfusion-Dependent Thalassemia. *Blood Rev.* **2012**, *26*, S3–S6.
12. Rachmilewitz, E.A.; Giardina, P.J. How I Treat Thalassemia. *Blood* **2011**, *118*, 3479–3488.
13. Musallam, K.M.; Rivella, S.; Vichinsky, E.; Rachmilewitz, E.A. Non-Transfusion-Dependent Thalassemias. *Haematologica* **2013**, *98*, 833–844.
14. Lederer, C.W.; Basak, A.N.; Aydinok, Y.; Christou, S.; El-Beshlawy, A.; Eleftheriou, A.; Fattoum, S.; Felice, A.E.; Fibach, E.; Galanello, R.; et al. An Electronic Infrastructure for Research and Treatment of the Thalassemias and Other Hemoglobinopathies: The Euro-Mediterranean ITHANET Project. *Hemoglobin* **2009**, *33*, 163–176.
15. Fokkema, I.F.A.C.; Taschner, P.E.M.; Schaafsma, G.C.P.; Celli, J.; Laros, J.F.J.; den Dunnen, J.T. LOVD v.2.0: The next Generation in Gene Variant Databases. *Hum. Mutat.* **2011**, *32*, 557–563.
16. Giardine, B.M.; Joly, P.; Pissard, S.; Wajcman, H.; Chui, D.H.K.; Hardison, R.C.; Patrinos, G.P. Clinically Relevant Updates of the HbVar

Database of Human Hemoglobin Variants and Thalassemia Mutations. *Nucleic Acids Res.* **2021**, *49*, D1192–D1196.
17. Kan, Y.W.; Lee, K.Y.; Furbetta, M.; Angius, A.; Cao, A. Polymorphism of DNA Sequence in the β-Globin Gene Region. *N. Engl. J. Med.* **1980**, *302*, 185–188.
18. Kazazian, H.H.; Phillips, J.A.; Boehm, C.D.; Vik, T.A.; Mahoney, M.J.; Ritchey, A.K. Prenatal Diagnosis of β-Thalassemias by Amniocentesis: Linkage Analysis Using Multiple Polymorphic Restriction Endonuclease Sites. *Blood* **1980**, *56*, 926–930.
19. Geever, R.F.; Wilson, L.B.; Nallaseth, F.S.; Milner, P.F.; Bittner, M. Direct Identification of Sickle Cell Anemia by Blot Hybridization. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 5081–5085.
20. Pirastu, M.; Kan, Y.W.; Cao, A.; Conner, B.J.; Teplitz, R.L.; Wallace, R.B. Prenatal Diagnosis of β-Thalassemia. *N. Engl. J. Med.* **1983**, *309*, 284–287.
21. Kleppe, K.; Ohtsuka, E.; Kleppe, R.; Molineux, I.; Khorana, H.G. Studies on Polynucleotides. XCVI. Repair Replication of Short Synthetic DNA's as Catalyzed by DNA Polymerases. *J. Mol. Biol.* **1971**, *56*, 341–361.
22. Saiki, R.K.; Scharf, S.; Faloona, F.; Mullis, K.B.; Horn, G.T.; Erlich, H.A.; Arnheim, N. Enzymatic Amplification of β-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science* **1985**, *230*, 1350–1354.
23. Mullis, K.B.; Faloona, F.A.B.T.-M. Specific Synthesis of DNA in Vitro via a Polymerase-Catalyzed Chain Reaction. In *Recombinant DNA Part F*; Academic Press: Cambridge, MA, USA, 1987; Volume 155, pp. 335–350. ISBN 0076-6879.
24. Amselem, S.; Nunes, V.; Vidaud, M.; Estivill, X.; Wong, C.; d'Auriol, L.; Vidaud, D.; Galibert, F.; Baiget, M.; Goossens, M. Determination of the Spectrum of β-Thalassemia Genes in Spain by Use of Dot-Blot Analysis of Amplified β-Globin DNA. *Am. J. Hum. Genet.* **1988**, *43*, 95–100.
25. Saiki, R.K.; Chang, C.-A.; Levenson, C.H.; Warren, T.C.; Boehm, C.D.; Kazazian, H.H.; Erlich, H.A. Diagnosis of Sickle Cell Anemia and β-Thalassemia with Enzymatically Amplified DNA and Nonradioactive Allele-Specific Oligonucleotide Probes. *N. Engl. J. Med.* **1988**, *319*, 537–541.
26. Saiki, R.K.; Walsh, P.S.; Levenson, C.H.; Erlich, H.A. Genetic Analysis of Amplified DNA with Immobilized Sequence-Specific Oligonucleotide Probes. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6230–6234.
27. Maggio, A.; Giambona, A.; Cai, S.P.; Wall, J.; Kan, Y.W.; Chehab, F.F. Rapid and Simultaneous Typing of Hemoglobin S, Hemoglobin C, and Seven Mediterranean Beta-Thalassemia Mutations by Covalent Reverse Dot-Blot Analysis: Application to Prenatal Diagnosis in Sicily. *Blood* **1993**, *81*, 239–242.
28. Cai, S.-P.; Wall, J.; Kan, Y.W.; Chehab, F.F. Reverse Dot Blot Probes for the Screening of β-Thalassernia Mutationsin Asians and American Blacks. *Hum. Mutat.* **1994**, *3*, 59–63.
29. Giambona, A.; Lo Gioco, P.; Marino, M.; Abate, I.; Di Marzo, R.; Renda, M.; Di Trapani, F.; Messana, F.; Siciliano, S.; Rigano, P. The Great Heterogeneity of Thalassemia Molecular Defects in Sicily. *Hum. Genet.* **1995**, *95*, 526–
30. Sutcharitchan, P.; Saiki, R.; Fucharoen, S.; Winichagoon, P.; Erlich, H.; Embury, S.H. Reverse Dot-Blot Detection of Thai β-Thalassaemia Mutations. *Br. J. Haematol.* **1995**, *90*, 809–816.
31. Sutcharitchan, P.; Saiki, R.; Huisman, T.H.; Kutlar, A.; McKie, V.; Erlich, H.; Embury, S.H. Reverse Dot-Blot Detection of the African-American Beta-Thalassemia Mutations. *Blood* **1995**, *86*, 1580–1585.
32. Chan, V.; Yam, I.; Chen, F.E.; Chan, T.K. A Reverse Dot-Blot Method for Rapid Detection of Non-Deletion α Thalassaemia. *Br. J. Haematol.* **1999**, *104*, 513–515.
33. Bashyam, M.D.; Bashyam, L.; Savithri, G.R.; Gopikrishna, M.; Sangal, V.; Devi, A.R.R. Molecular Genetic Analyses of β-Thalassemia in South India Reveals Rare Mutations in the β-Globin Gene. *J. Hum. Genet.* **2004**, *49*, 408–413.
34. Lin, M.; Zhu, J.-J.; Wang, Q.; Xie, L.-X.; Lu, M.; Wang, J.-L.; Wang, C.-F.; Zhong, T.-Y.; Zheng, L.; Pan, M.-C.; et al. Development and Evaluation of a Reverse Dot Blot Assay for the Simultaneous Detection of Common Alpha and Beta Thalassemia in Chinese. *Blood Cells Mol. Dis.* **2012**, *48*, 86–90.
35. Chong, S.S.; Boehm, C.D.; Higgs, D.R.; Cutting, G.R. Single-Tube Multiplex-PCR Screen for Common Deletional Determinants of α-Thalassemia. *Blood* **2000**, *95*, 360–362.
36. Wang, W.; Ma, E.S.; Chan, A.Y.; Prior, J.; Erber, W.N.; Chan, L.C.; Chui, D.H.; Chong, S.S. Single-Tube Multiplex-PCR Screen for Anti-3.7 and Anti-4.2 α-Globin Gene Triplications. *Clin. Chem.* **2003**, *49*, 1679–1682.
37. Craig, J.; Barnetson, R.; Prior, J.; Raven, J.; Thein, S. Rapid Detection of Deletions Causing Delta Beta Thalassemia and Hereditary Persistence of Fetal Hemoglobin by Enzymatic Amplification. *Blood* **1994**, *83*, 1673–1682.
38. Nussenzveig, R.H.; Vanhille, D.L.; Hussey, N.; Scott Reading, D.; Agarwal, A.M. Development of a Rapid Multiplex PCR Assay for Identification of the Three Common Hemoglobin-Lepore Variants (Boston-Washington, Baltimore,

and Hollandia) and Identification of a New Lepore Variant. *Am. J. Hematol.* **2012**, *87*, 74–75.
39. Tritipsombut, J.; Phylipsen, M.; Viprakasit, V.; Chalaow, N.; Sanchaisuriya, K.; Giordano, P.C.; Fucharoen, S.; Harteveld, C.L. A Single-Tube Multiplex Gap-Polymerase Chain Reaction for the Detection of Eight β-Globin Gene Cluster Deletions Common in Southeast Asia. *Hemoglobin* **2012**, *36*, 571–580.
40. Wu, D.Y.; Ugozzoli, L.; Pal, B.K.; Wallace, R.B. Allele-Specific Enzymatic Amplification of Beta-Globin Genomic DNA for Diagnosis of Sickle Cell Anemia. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2757–2760.
41. Newton, C.R.; Graham, A.; Heptinstall, L.E.; Powell, S.J.; Summers, C.; Kalsheker, N.; Smith, J.C.; Markham, A.F. Analysis of Any Point Mutation in DNA. The Amplification Refractory Mutation System (ARMS). *Nucleic Acids Res.* **1989**, *17*, 2503–2516.
42. Fortina, P.; Dotti, G.; Conant, R.; Monokian, G.; Parrella, T.; Hitchcock, W.; Rappaport, E.; Schwartz, E.; Surrey, S. Detection of the Most Common Mutations Causing Beta-Thalassemia in Mediterraneans Using a Multiplex Amplification Refractory Mutation System (MARMS). *PCR Methods Appl.* **1992**, *2*, 163–166.
43. Mirasena, S.; Shimbhu, D.; Sanguansermsri, M.; Sanguansermsri, T. Detection of β-Thalassemia Mutations Using a Multiplex Amplification Refractory Mutation System Assay. *Hemoglobin* **2008**, *32*, 403–409.
44. Hassan, S.; Ahmad, R.; Zakaria, Z.; Zulkafli, Z.; Abdullah, W.Z. Detection of β-Globin Gene Mutations among β-Thalassaemia Carriers and Patients in Malaysia: Application of Multiplex Amplification Refractory Mutation System-Polymerase Chain Reaction. *Malaysian J. Med. Sci.* **2013**, *20*, 13–20.
45. Tan, J.A.M.A.; Tay, J.S.H.; Lin, L.I.; Kham, S.K.Y.; Chia, J.N.; Chin, T.M.; Adb Aziz, N.B.; Wong, H.B. The Amplification Refractory Mutation System (ARMS): A Rapid and Direct Prenatal Diagnostic Technique for β-Thalassaemia in Singapore. *Prenat. Diagn.* **1994**, *14*, 1077–1082.
46. Chang, J.G.; Liu, H.J.; Huang, J.M.; Yang, T.Y.; Chang, C.P. Multiplex Mutagenically Separated PCR: Diagnosis of Beta-Thalassemia and Hemoglobin Variants. *Biotechniques* **1997**, *22*, 520–527.
47. Khateeb, B.; Moatter, T.; Shaghil, A.M.; Haroon, S.; Kakepoto, G.N. Genetic Diversity of Beta-Thalassemia Mutations in Pakistani Population. *J. Pakistan Med. Assoc.* **2000**, *50*, 293–296.
48. Eng, B.; Patterson, M.; Walker, L.; Chui, D.H.K.; Waye, J.S. Detection of Severe Nondeletional α-Thalassemia Mutations Using a Single-Tube Multiplex ARMS Assay. *Genet. Test.* **2001**, *5*, 327–329.
49. Bhardwaj, U.; Zhang, Y.-H.; Lorey, F.; McCabe, L.L.; McCabe, E.R.B. Molecular Genetic Confirmatory Testing from Newborn Screening Samples for the Common African-American, Asian Indian, Southeast Asian, and Chinese β-Thalassemia Mutations. *Am. J. Hematol.* **2005**, *78*, 249–255.
50. Darwish, H.M.; El-Khatib, F.F.; Ayesh, S. Spectrum of β-Globin Gene Mutations Among Thalassemia Patients in the West Bank Region of Palestine. *Hemoglobin* **2005**, *29*, 119–132.
51. El-Gawhary, S.; El-Shafie, S.; Niazi, M.; Aziz, M.; El-Beshlawy, A. Study of β-Thalassemia Mutations Using the Polymerase Chain Reaction-Amplification Refractory Mutation System and Direct DNA Sequencing Techniques in a Group of Egyptian Thalassemia Patients. *Hemoglobin* **2007**, *31*, 63–69.
52. Lacerra, G.; Musollino, G.; Di Noce, F.; Prezioso, R. Clementina Carestia Genotyping for Known Mediterranean α-Thalassemia Point Mutations Using a Multiplex Amplification Refractory Mutation System. *Haematologica* **2007**, *92*, 254–255.]
53. Sanger, F.; Coulson, A.R. A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase. *J. Mol. Biol.* **1975**, *94*, 441–448.
54. Smith, L.M.; Fung, S.; Hunkapiller, M.W.; Hunkapiller, T.J.; Hood, L.E. The Synthesis of Oligonucleotides Containing an Aliphatic Amino Group at the 5′ Terminus: Synthesis of Fluorescent DNA Primers for Use in DNA Sequence Analysis. *Nucleic Acids Res.* **1985**, *13*, 2399–2412.
55. Smith, L.M.; Sanders, J.Z.; Kaiser, R.J.; Hughes, P.; Dodd, C.; Connell, C.R.; Heiner, C.; Kent, S.B.; Hood, L.E. Fluorescence Detection in Automated DNA Sequence Analysis. *Nature* **1986**, *321*, 674–679.
56. Wong, C.; Dowling, C.E.; Saiki, R.K.; Higuchi, R.G.; Erlich, H.A.; Kazazian, H.H.J. Characterization of Beta-Thalassaemia Mutations Using Direct Genomic Sequencing of Amplified Single Copy DNA. *Nature* **1987**, *330*, 384–386.
57. Di Marzo, R.; Dowling, C.E.; Wong, C.; Maggio, A.; Kazazian, H.H. The Spectrum of β-Thalassaemia Mutations in Sicily. *Br. J. Haematol.* **1988**, *69*, 393–397.
58. Drossman, H.; Luckey, J.A.; Kostichka, A.J.; D'Cunha, J.; Smith, L.M. High-Speed Separations of DNA Sequencing Reactions by Capillary Electrophoresis. *Anal. Chem.* **1990**, *62*, 900–903.
59. Aulehla-Scholz, C.; Basaran, S.; Agaoglu, L.; Arcasoy, A.; Holzgreve, W.; Miny, P.; Ridolfi, F.; Horst, J. Molecular Basis of Beta-Thalassemia in Turkey: Detection of Rare Mutations by Direct Sequencing. *Hum. Genet.* **1990**, *84*, 195–197.

60. Hsuih, T.C.; Park, Y.N.; Zaretsky, C.; Wu, F.; Tyagi, S.; Kramer, F.R.; Sperling, R.; Zhang, D.Y. Novel, Ligation-Dependent PCR Assay for Detection of Hepatitis C in Serum. *J. Clin. Microbiol.* **1996**, *34*, 501–507.
61. Carrino, J.J. Multiplex Ligations-Dependent Amplification. 1996. Available online: **http://patent.google.com/patent/WO1996015271A1/en**
62. Schouten, J.P.; McElgunn, C.J.; Waaijer, R.; Zwijnenburg, D.; Diepvens, F.; Pals, G. Relative Quantification of 40 Nucleic Acid Sequences by Multiplex Ligation-Dependent Probe Amplification. *Nucleic Acids Res.* **2002**, *30*, e57.
63. Zhong, Z.; Zhong, G.; Guan, Z.; Chen, D.; Wu, Z.; Yang, K.; Chen, D.; Liu, Y.; Xu, R.; Chen, J. A Novel 15.8 Kb Deletion α-Thalassemia Confirmed by Long-Read Single-Molecule Real-Time Sequencing: Hematological Phenotypes and Molecular Characterization. *Clin. Biochem.* **2022**, *108*, 46–49.
64. Luo, S.; Chen, X.; Zeng, D.; Tang, N.; Yuan, D.; Liu, B.; Chen, L.; Zhong, Q.; Li, J.; Liu, Y.; et al. Detection of Four Rare Thalassemia Variants Using Single-Molecule Realtime Sequencing. *Front. Genet.* **2022**, *13*, 974999.
65. Harteveld, C.L.; Refaldi, C.; Cassinerio, E.; Cappellini, M.D.; Giordano, P.C. Segmental Duplications Involving the α-Globin Gene Cluster Are Causing β-Thalassemia Intermedia Phenotypes in β-Thalassemia Heterozygous Patients. *Blood Cells Mol. Dis.* **2008**, *40*, 312–316.
66. Sollaino, M.C.; Paglietti, M.E.; Perseu, L.; Giagu, N.; Loi, D.; Galanello, R. Association of a Globin Gene Quadruplication and Heterozygous β Thalassemia in Patients with Thalassemia Intermedia. *Haematologica* **2009**, *94*, 1445–1448.
67. Jiang, H.; Liu, S.; Zhang, Y.L.; Wan, J.H.; Li, R.; Li, D.Z. Association of an α-Globin Gene Cluster Duplication and Heterozygous β-Thalassemia in a Patient with a Severe Thalassemia Syndrome. *Hemoglobin* **2015**, *39*, 102–106.
68. Steinberg-Shemer, O.; Ulirsch, J.C.; Noy-Lotan, S.; Krasnov, T.; Attias, D.; Dgany, O.; Laor, R.; Sankaran, V.G.; Tamary, H. Whole-Exome Sequencing Identifies an α-Globin Cluster Triplication Resulting in Increased Clinical Severity of β-Thalassemia. *Cold Spring Harb. Mol. Case Stud.* **2017**, *3*, a001941.
69. Clark, B.; Shooter, C.; Smith, F.; Brawand, D.; Steedman, L.; Oakley, M.; Rushton, P.; Rooks, H.; Wang, X.; Drousiotou, A.; et al. Beta Thalassaemia Intermedia Due to Co-Inheritance of Three Unique Alpha Globin Cluster Duplications Characterised by next Generation Sequencing Analysis. *Br. J. Haematol.* **2018**, *180*, 160–164.
70. Kubiczkova Besse, L.; Sedlarikova, L.; Kryukov, F.; Nekvindova, J.; Radova, L.; Almasi, M.; Pelcova, J.; Minarik, J.; Pika, T.; Pikalova, Z.; et al. Combination of serum microRNA-320a and microRNA-320b as a marker for Waldenström macroglobulinemia. *Am. J. Hematol.* **2015**, *90*, E51–E52.
71. Shooter, C.; Rooks, H.; Thein, S.L.; Clark, B. Next Generation Sequencing Identifies a Novel Rearrangement in the HBB Cluster Permitting To-the-Base Characterization. *Hum. Mutat.* **2015**, *36*, 142–150.
72. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nat. Genet.* **2011**, *43*, 491–498.] [**CrossRef**]
73. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997v2.
74. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359.
75. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100.
76. Zaharia, M.; Bolosky, W.J.; Curtis, K.; Fox, A.; Patterson, D.A.; Shenker, S.; Stoica, I.; Karp, R.M.; Sittler, T. Faster and More Accurate Sequence Alignment with {SNAP}. *arXiv* **2011**, arXiv:1111.5572.
77. Garrison, E.; Marth, G. Haplotype-Based Variant Detection from Short-Read Sequencing. *arXiv* **2012**, arXiv:1207.3907.
78. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079.
79. Poplin, R.; Chang, P.-C.; Alexander, D.; Schwartz, S.; Colthurst, T.; Ku, A.; Newburger, D.; Dijamco, J.; Nguyen, N.; Afshar, P.T.; et al. A Universal SNP and Small-Indel Variant Caller Using Deep Neural Networks. *Nat. Biotechnol.* **2018**, *36*, 983–987.
80. Kim, S.; Scheffler, K.; Halpern, A.L.; Bekritsky, M.A.; Noh, E.; Källberg, M.; Chen, X.; Kim, Y.; Beyter, D.; Krusche, P.; et al. Strelka2: Fast and Accurate Calling of Germline and Somatic Variants. *Nat. Methods* **2018**, *15*, 591–594.
81. Teo, S.M.; Pawitan, Y.; Ku, C.S.; Chia, K.S.; Salim, A. Statistical Challenges Associated with Detecting Copy Number Variations with Next-Generation Sequencing. *Bioinformatics* **2012**, *28*, 2711–2718.
82. Boeva, V.; Popova, T.; Bleakley, K.; Chiche, P.; Cappo, J.; Schleiermacher, G.; Janoueix-Lerosey,

I.; Delattre, O.; Barillot, E. Control-FREEC: A Tool for Assessing Copy Number and Allelic Content Using next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 423–425.
83. Rausch, T.; Zichner, T.; Schlattl, A.; Stütz, A.M.; Benes, V.; Korbel, J.O. DELLY: Structural Variant Discovery by Integrated Paired-End and Split-Read Analysis. *Bioinformatics* **2012**, *28*, i333–i339.
84. Talevich, E.; Shain, A.H.; Botton, T.; Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **2016**, *12*, e1004873.
85. Plagnol, V.; Curtis, J.; Epstein, M.; Mok, K.Y.; Stebbings, E.; Grigoriadou, S.; Wood, N.W.; Hambleton, S.; Burns, S.O.; Thrasher, A.J.; et al. A Robust Model for Read Count Data in Exome Sequencing Experiments and Implications for Copy Number Variant Calling. *Bioinformatics* **2012**, *28*, 2747–2754.
86. Krumm, N.; Sudmant, P.H.; Ko, A.; O'Roak, B.J.; Malig, M.; Coe, B.P.; Project, N.E.S.; Quinlan, A.R.; Nickerson, D.A.; Eichler, E.E. Copy Number Variation Detection and Genotyping from Exome Sequence Data. *Genome Res.* **2012**, *22*, 1525–1532.
87. Shi, Y.; Majewski, J. FishingCNV: A Graphical Software Package for Detecting Rare Copy Number Variations in Exome-Sequencing Data. *Bioinformatics* **2013**, *29*, 1461–1462.
88. Zhang, Y.; Yu, Z.; Ban, R.; Zhang, H.; Iqbal, F.; Zhao, A.; Li, A.; Shi, Q. DeAnnCNV: A Tool for Online Detection and Annotation of Copy Number Variations from Whole-Exome Sequencing Data. *Nucleic Acids Res.* **2015**, *43*, W289–W294.
89. Derouault, P.; Chauzeix, J.; Rizzo, D.; Miressi, F.; Magdelaine, C.; Bourthoumieu, S.; Durand, K.; Dzugan, H.; Feuillard, J.; Sturtz, F.; et al. CovCopCan: An Efficient Tool to Detect Copy Number Variation from Amplicon Sequencing Data in Inherited Diseases and Cancer. *PLoS Comput. Biol.* **2020**, *16*, e1007503.
90. Chanwigoon, S.; Piwluang, S.; Wichadakul, D. InCNV: An Integrated Analysis Tool for Copy Number Variation on Whole Exome Sequencing. *Evol. Bioinforma.* **2020**, *16*, 1176934320956577.
91. Erikson, G.A.; Deshpande, N.; Kesavan, B.G.; Torkamani, A. SG-ADVISER CNV: Copy-Number Variant Annotation and Interpretation. *Genet. Med.* **2015**, *17*, 714–718.
92. Chandramohan, R.; Kakkar, N.; Roy, A.; Parsons, D.W. ReconCNV: Interactive Visualization of Copy Number Data from High-Throughput Sequencing. *Bioinformatics* **2021**, *37*, 1164–1167.
93. Requena, F.; Abdallah, H.H.; García, A.; Nitschké, P.; Romana, S.; Malan, V.; Rausell, A. CNVxplorer: A Web Tool to Assist Clinical Interpretation of CNVs in Rare Disease Patients. *Nucleic Acids Res.* **2021**, *49*, W93–W103.
94. Gel, B.; Serra, E. KaryoploteR: An R/Bioconductor Package to Plot Customizable Genomes Displaying Arbitrary Data. *Bioinformatics* **2017**, *33*, 3088–3090.
95. Hahne, F.; Ivanek, R. *Visualizing Genomic Data Using Gviz and Bioconductor BT—Statistical Genomics: Methods and Protocols*; Mathé, E., Davis, S., Eds.; Springer: New York, NY, USA, 2016; pp. 335–351. ISBN 978-1-4939-3578-9.
96. Gel, B.; Magallon, M. CopyNumberPlots: Create Copy-Number Plots Using KaryoploteR Functionality. 2021. Available online: **http://github.com/bernatgel/CopyNumberPlots**
97. Shang, X.; Peng, Z.; Ye, Y.; Asan; Zhang, X.; Chen, Y.; Zhu, B.; Cai, W.; Chen, S.; Cai, R.; et al. Rapid Targeted Next-Generation Sequencing Platform for Molecular Screening and Clinical Genotyping in Subjects with Hemoglobinopathies. *EBioMedicine* **2017**, *23*, 150–159.
98. Chen, P.; Yu, X.; Huang, H.; Zeng, W.; He, X.; Liu, M.; Huang, B. Evaluation of Ion Torrent Next-Generation Sequencing for Thalassemia Diagnosis. *J. Int. Med. Res.* **2020**, *48*, 0300060520967778.
99. Zhao, J.; Li, J.; Lai, Q.; Yu, Y. Combined Use of Gap-PCR and next-Generation Sequencing Improves Thalassaemia Carrier Screening among Premarital Adults in China. *J. Clin. Pathol.* **2020**, *73*, 488–492.
100. He, J.; Song, W.; Yang, J.; Lu, S.; Yuan, Y.; Guo, J.; Zhang, J.; Ye, K.; Yang, F.; Long, F.; et al. Next-Generation Sequencing Improves Thalassemia Carrier Screening among Premarital Adults in a High Prevalence Population: The Dai Nationality, China. *Genet. Med.* **2017**, *19*, 1022–1031.
101. Fan, D.M.; Yang, X.; Huang, L.M.; Ouyang, G.J.; Yang, X.X.; Li, M. Simultaneous Detection of Target CNVs and SNVs of Thalassemia by Multiplex PCR and Next-generation Sequencing. *Mol. Med. Rep.* **2019**, *19*, 2837–2848.
102. Zebisch, A.; Schulz, E.; Grosso, M.; Lombardo, B.; Acierno, G.; Sill, H.; Iolascon, A. Identification of a Novel Variant of Epsilon-Gamma-Delta-Beta Thalassemia Highlights Limitations of next Generation Sequencing. *Am. J. Hematol.* **2015**, *90*, E52–E54.
103. Fromer, M.; Moran, J.L.; Chambert, K.; Banks, E.; Bergen, S.E.; Ruderfer, D.M.; Handsaker, R.E.; McCarroll, S.A.; O'Donovan, M.C.; Owen, M.J.; et al. Discovery and Statistical Genotyping of Copy-Number Variation from Whole-Exome

Sequencing Depth. *Am. J. Hum. Genet.* **2012**, *91*, 597–607
104. Cao, Y.; Chan, G.C.-F. NGS4THAL, a One-Stop Molecular Diagnosis and Carrier Screening Tool for Thalassemia and Other Hemoglobinopathies by next-Generation Sequencing. *Res. Sq.* **2022**, *24*, 1089–1099.
105. Jiang, F.; Lyu, G.Z.; Zhang, V.W.; Li, D.Z. Identification of Thalassemia Gene Cluster Deletion by Long-Read Whole-Genome Sequencing (LR-WGS). *Int. J. Lab. Hematol.* **2021**, *43*, 859–865.
106. Li, Q.; Mao, Y.; Li, S.; Du, H.; He, W.; He, J.; Kong, L.; Zhang, J.; Liang, B.; Liu, J. Haplotyping by Linked-Read Sequencing (HLRS) of the Genetic Disease Carriers for Preimplantation Genetic Testing without a Proband or Relatives. *BMC Med. Genomics* **2020**, *13*, 117.
107. Korlach, J.; Turner, S.W. *Single-Molecule Sequencing BT—Encyclopedia of Biophysics*; Roberts, G.C.K., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 2344–2347. ISBN 978-3-642-16712-6.
108. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; et al. Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* **2009**, *323*, 133–138.
109. PacBio SMRT Sequencing—How It Works. Available online: **https://www.pacb.com/smrt-science/attachment/infographic_smrt-sequencing-how-it-works/%0Awww.pacb.com**
110. Wenger, A.M.; Peluso, P.; Rowell, W.J.; Chang, P.-C.; Hall, R.J.; Concepcion, G.T.; Ebler, J.; Fungtammasan, A.; Kolesnikov, A.; Olson, N.D.; et al. Accurate Circular Consensus Long-Read Sequencing Improves Variant Detection and Assembly of a Human Genome. *Nat. Biotechnol.* **2019**, *37*, 1155–1162.
111. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365.
112. Tomasz Dobrzycki Selecting the Right Library Prep Method for Your Experiment. Available online: **https://nanoporetech.com/resource-centre/video/lc22/selecting-the-right-library-prep-method-for-your-experiment**
113. Oxford Nanopore Technologies How It Works. Available online: **https://nanoporetech.com/how-it-works**
114. Carneiro, M.O.; Russ, C.; Ross, M.G.; Gabriel, S.B.; Nusbaum, C.; DePristo, M.A. Pacific Biosciences Sequencing Technology for Genotyping and Variation Discovery in Human Data. *BMC Genomics* **2012**, *13*, 375.
115. Cheng, S.H.; Jiang, P.; Sun, K.; Cheng, Y.K.Y.; Chan, K.C.A.; Leung, T.Y.; Chiu, R.W.K.; Lo, Y.M.D. Noninvasive Prenatal Testing by Nanopore Sequencing of Maternal Plasma DNA: Feasibility Assessment. *Clin. Chem.* **2015**, *61*, 1305–1306.
116. Zhang, H.; Jain, C.; Aluru, S. A Comprehensive Evaluation of Long Read Error Correction Methods. *BMC Genomics* **2020**, *21*, 889.
117. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546.
118. Ruan, J.; Li, H. Fast and Accurate Long-Read Assembly with Wtdbg2. *Nat. Methods* **2020**, *17*, 155–158.
119. Shafin, K.; Pesout, T.; Lorig-Roach, R.; Haukness, M.; Olsen, H.E.; Bosworth, C.; Armstrong, J.; Tigyi, K.; Maurer, N.; Koren, S.; et al. Nanopore Sequencing and the Shasta Toolkit Enable Efficient de Novo Assembly of Eleven Human Genomes. *Nat. Biotechnol.* **2020**, *38*, 1044–1053.
120. Morisse, P.; Marchet, C.; Limasset, A.; Lecroq, T.; Lefebvre, A. Scalable Long Read Self-Correction and Assembly Polishing with Multiple Sequence Alignment. *Sci. Rep.* **2021**, *11*, 761.
121. Xiao, C.-L.; Chen, Y.; Xie, S.-Q.; Chen, K.-N.; Wang, Y.; Han, Y.; Luo, F.; Xie, Z. MECAT: Fast Mapping, Error Correction, and de Novo Assembly for Single-Molecule Sequencing Reads. *Nat. Methods* **2017**, *14*, 1072–1074.
122. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation. *Genome Res.* **2017**, *27*, 722–736.
123. Berlin, K.; Koren, S.; Chin, C.-S.; Drake, J.P.; Landolin, J.M.; Phillippy, A.M. Assembling Large Genomes with Single-Molecule Sequencing and Locality-Sensitive Hashing. *Nat. Biotechnol.* **2015**, *33*, 623–630
124. Chin, C.-S.; Peluso, P.; Sedlazeck, F.J.; Nattestad, M.; Concepcion, G.T.; Clum, A.; Dunn, C.; O'Malley, R.; Figueroa-Balderas, R.; Morales-Cruz, A.; et al. Phased Diploid Genome Assembly with Single-Molecule Real-Time Sequencing. *Nat. Methods* **2016**, *13*, 1050–1054.
125. Chaisson, M.J.; Tesler, G. Mapping Single Molecule Sequencing Reads Using Basic Local Alignment with Successive Refinement (BLASR): Application and Theory. *BMC Bioinformatics* **2012**, *13*,
126. Chen, Y.; Nie, F.; Xie, S.-Q.; Zheng, Y.-F.; Dai, Q.; Bray, T.; Wang, Y.-X.; Xing, J.-F.; Huang, Z.-J.; Wang, D.-P.; et al. Efficient Assembly of Nanopore Reads via Highly Accurate and Intact Error Correction. *Nat. Commun.* **2021**, *12*, 60.

127. Shafin, K.; Pesout, T.; Chang, P.-C.; Nattestad, M.; Kolesnikov, A.; Goel, S.; Baid, G.; Kolmogorov, M.; Eizenga, J.M.; Miga, K.H.; et al. Haplotype-Aware Variant Calling with PEPPER-Margin-DeepVariant Enables High Accuracy in Nanopore Long-Reads. *Nat. Methods* **2021**, *18*, 1322–1332.
128. Ahsan, M.U.; Liu, Q.; Fang, L.; Wang, K. NanoCaller for Accurate Detection of SNPs and Indels in Difficult-to-Map Regions from Long-Read Sequencing by Haplotype-Aware Deep Neural Networks. *Genome Biol.* **2021**, *22*, 261.
129. Su, J.; Zheng, Z.; Ahmed, S.S.; Lam, T.-W.; Luo, R. Clair3-Trio: High-Performance Nanopore Long-Read Variant Calling in Family Trios with Trio-to-Trio Deep Neural Networks. *Brief. Bioinform.* **2022**, bbac301.
130. Ramachandran, A.; Lumetta, S.S.; Klee, E.W.; Chen, D. HELLO: Improved Neural Network Architectures and Methodologies for Small Variant Calling. *BMC Bioinformatics* **2021**, *22*, 404.
131. Smolka, M.; Paulin, L.F.; Grochowski, C.M.; Mahmoud, M.; Behera, S.; Gandhi, M.; Hong, K.; Pehlivan, D.; Scholz, S.W.; Carvalho, C.M.B.; et al. Comprehensive Structural Variant Detection: From Mosaic to Population-Level (Sniffles2). *bioRxiv* **2022**.
132. Cao, S.; Jiang, T.; Liu, Y.; Liu, S.; Wang, Y. Re-Genotyping Structural Variants through an Accurate Force-Calling Method. *bioRxiv* **2022**.
133. Heller, D.; Vingron, M. SVIM: Structural Variant Identification Using Mapped Long Reads. *Bioinformatics* **2019**, *35*, 2907–2915.
134. Tham, C.Y.; Tirado-Magallanes, R.; Goh, Y.; Fullwood, M.J.; Koh, B.T.H.; Wang, W.; Ng, C.H.; Chng, W.J.; Thiery, A.; Tenen, D.G.; et al. NanoVar: Accurate Characterization of Patients' Genomic Structural Variants Using Low-Depth Nanopore Sequencing. *Genome Biol.* **2020**, *21*,
135. Xu, L.; Mao, A.; Liu, H.; Gui, B.; Choy, K.W.; Huang, H.; Yu, Q.; Zhang, X.; Chen, M.; Lin, N.; et al. Long-Molecule Sequencing: A New Approach for Identification of Clinically Significant DNA Variants in α-Thalassemia and β-Thalassemia Carriers. *J. Mol. Diagnostics* **2020**, *22*, 1087–1095.
136. Liang, Q.; Gu, W.; Chen, P.; Li, Y.; Liu, Y.; Tian, M.; Zhou, Q.; Qi, H.; Zhang, Y.; He, J.; et al. A More Universal Approach to Comprehensive Analysis of Thalassemia Alleles (CATSA). *J. Mol. Diagnostics* **2021**, *23*, 1195–1204.
137. Peng, C.; Zhang, H.; Ren, J.; Chen, H.; Du, Z.; Zhao, T.; Mao, A.; Xu, R.; Lu, Y.; Wang, H.; et al. Analysis of Rare Thalassemia Genetic Variants Based on Third-Generation Sequencing. *Sci. Rep.* **2022**, *12*, 9907.
138. Li, Y.; Liang, L.; Qin, T.; Tian, M. Detection of Hemoglobin H Disease by Long Molecule Sequencing. *J. Clin. Lab. Anal.* **2022**, *36*, e24687.
139. Jiang, F.; Liu, W.; Zhang, L.; Guo, Y.; Chen, M.; Zeng, X.; Wang, Y.; Li, Y.; Xian, J.; Du, B.; et al. Noninvasive Prenatal Testing for β-Thalassemia by Targeted Nanopore Sequencing Combined with Relative Haplotype Dosage (RHDO): A Feasibility Study. *Sci. Rep.* **2021**, *11*, 5714.
140. Christopher, H.; Burns, A.; Josephat, E.; Makani, J.; Schuh, A.; Nkya, S. Using DNA Testing for the Precise, Definite, and Low-Cost Diagnosis of Sickle Cell Disease and Other Haemoglobinopathies: Findings from Tanzania. *BMC Genomics* **2021**, *22*, 902.
141. Liu, S.; Wang, H.; Leigh, D.; Cram, D.S.; Wang, L.; Yao, Y. Third-Generation Sequencing: Any Future Opportunities for PGT? *J. Assist. Reprod. Genet.* **2021**, *38*, 357–364.
142. Sirén, J.; Monlong, J.; Chang, X.; Novak, A.M.; Eizenga, J.M.; Markello, C.; Sibbesen, J.A.; Hickey, G.; Chang, P.-C.; Carroll, A.; et al. Pangenomics Enables Genotyping of Known Structural Variants in 5202 Diverse Genomes. *Science* **2021**, *374*, abg8871.
143. Li, H.; Feng, X.; Chu, C. The Design and Construction of Reference Pangenome Graphs with Minigraph. *Genome Biol.* **2020**, *21*, 265.
144. Wang, T.; Antonacci-Fulton, L.; Howe, K.; Lawson, H.A.; Lucas, J.K.; Phillippy, A.M.; Popejoy, A.B.; Asri, M.; Carson, C.; Chaisson, M.J.P.; et al. The Human Pangenome Project: A Global Resource to Map Genomic Diversity. *Nature* **2022**, *604*, 437–446.
145. Roy, S.; Coldren, C.; Karunamurthy, A.; Kip, N.S.; Klee, E.W.; Lincoln, S.E.; Leon, A.; Pullambhatla, M.; Temple-Smolkin, R.L.; Voelkerding, K.V.; et al. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J. Mol. Diagnostics* **2018**, *20*, 4–27.
146. Santani, A.; Simen, B.B.; Briggs, M.; Lebo, M.; Merker, J.D.; Nikiforova, M.; Vasalos, P.; Voelkerding, K.; Pfeifer, J.; Funke, B. Designing and Implementing NGS Tests for Inherited Disorders: A Practical Framework with Step-by-Step Guidance for Clinical Laboratories. *J. Mol. Diagnostics* **2019**, *21*, 369–374.
147. Roy, N.B.A.; Da Costa, L.; Russo, R.; Bianchi, P.; del Mar Mañú-Pereira, M.; Fermo, E.; Andolfo, I.; Clark, B.; Proven, M.; Sanchez, M.; et al. The Use of Next-Generation Sequencing in the Diagnosis of Rare Inherited Anaemias: A Joint BSH/EHA Good Practice Paper. *HemaSphere* **2022**, *6*, e739.
148. International Organization for Standardization [ISO] ISO 20397-1:2022 Biotechnology—Massively Parallel Sequencing. Available online: **https://www.iso.org/standard/74054.html**