

RESEARCH

Open Access



Ending diagnostic odyssey by reanalysis of whole exome sequencing data: reclassification of suspected Fanconi anemia cases to dyskeratosis congenita and Diamond-Blackfan anemia

Eudald Tejero¹, María José Ramírez de Haro^{1,2}, Roser Pujol^{1,2}, Massimo Bogliolo^{1,2,3} ,
Benjamín Rodríguez-Santiago^{1,2,4*} and Jordi Surrallés^{1,2,4*} 

Abstract

Background Initial Whole Exome Sequencing frequently fails to resolve rare disease cases. Bioinformatic reanalysis of existing genomic data utilizes advancing knowledge to enhance diagnosis without additional testing. This study investigated six patients with clinical features consistent with Fanconi Anemia but negative chromosomal breakage tests, whose initial genetic analyses were inconclusive.

Results Whole Exome Sequencing data from these patients (collected 2005–2009) underwent comprehensive reanalysis, including single nucleotide variants, insertions/deletions, and copy number variants across genes beyond those typically associated with Fanconi Anemia. Telomere length was assessed via monochrome multiplex quantitative PCR. Reanalysis identified clinically significant variants in two patients (33.3% yield): one harboured a heterozygous pathogenic loss-of-function variant in the Diamond-Blackfan anemia gene *RPL5*, while the second exhibited compound heterozygous variants in the *TERT* gene, indicative of dyskeratosis congenita.

Conclusions This study underscores the clinical value of reanalyzing existing genomic data in unresolved suspected genetic disorders, even when phenotype-specific assays are negative. The 33.3% diagnostic yield aligns with gains from larger reanalysis studies (10–25%). Systematic reassessment after sufficient time (24+ months) for genomic advancements offers a cost-effective diagnostic approach for long-undiagnosed cases, highlighting the dynamic nature of genomic interpretation as gene-disease understanding evolves.

Keywords WES Reanalysis, Rare Genetic Diseases, Fanconi Anemia, Diamond-Blackfan Anemia, Dyskeratosis Congenita, RPL5, TERT

*Correspondence:

Benjamín Rodríguez-Santiago

BRodriguezS@santpau.cat

Jordi Surrallés

Jordi.Surralles@uab.cat; JSurralles@santpau.cat

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Fanconi anemia (FA), dyskeratosis congenita (DC), and Diamond-Blackfan anemia (DBA) are inherited bone marrow failure (BMF) syndromes with high risks of bone marrow failure, leukemia, and solid tumors. The most common is FA, a rare hereditary condition characterized by defects in DNA repair mechanisms, leading to chromosomal instability and a diverse spectrum of clinical manifestations [1]. This disorder demonstrates complex inheritance patterns, with approximately 98% of cases inherited in an autosomal recessive manner, about 2% showing X-linked recessive inheritance (associated with the *FANCB* gene), and very rare instances of autosomal dominant transmission (related to the *FANCR* gene in <1% of cases). FA typically presents in early childhood with a constellation of symptoms including congenital anomalies, progressive bone marrow failure, and significantly elevated risk of hematologic and solid malignancies [2].

The diagnostic approach for FA primarily relies on the chromosomal breakage test, which serves as the distinctive diagnostic marker for the condition. This test involves exposing patient cells to DNA cross-linking agents such as diepoxybutane (DEB) or mitomycin C (MMC) to evaluate chromosome fragility [2]. While this test is the cornerstone of diagnosis, genetic testing plays a crucial complementary role in confirming the diagnosis and identifying the specific genetic variant responsible [3]. It is particularly valuable in cases of somatic mosaicism, where chromosomal breakage tests using blood cells may yield normal results, necessitating additional testing on fibroblasts [4, 5]. Despite thorough testing, some individuals exhibit clinical features strongly suggestive of FA but have negative chromosome fragility tests and no identifiable variants in known FA-related genes.

Next Generation Sequencing (NGS) is a powerful tool for the diagnosis of genetic disorders. In cases where initial genetic testing, such as Whole Exome Sequencing (WES), does not yield a diagnosis, reanalysis of the generated NGS data has emerged as a powerful strategy to increase the diagnostic yield without requiring additional sequencing [6, 7]. This approach holds particular relevance as genomic interpretation represents a dynamic process that evolves with the continuous discovery of new disease-causing genes and the development of improved bioinformatics tools and analytical methodologies. Multiple studies have demonstrated that routine reanalysis over time significantly improves diagnostic rates, with a recent systematic review and meta-analysis indicating an overall diagnostic yield of 10% in previously undiagnosed cases [8]. More specifically, research on dystonia patients has shown that genome sequencing

data reanalysis increased the diagnostic yield from 11.7% to 18.9%, with potential further extension up to 22.5% [9].

Several interconnected factors contribute to the increased diagnostic success achieved through reanalysis. Literature updates and newly established gene-disease associations represent primary drivers of new diagnoses, as the rapid pace of genetic discovery means that revisiting existing data allows for the identification of genes newly linked to disease after the initial analysis [10]. The constant expansion of genetic knowledge creates opportunities for reinterpretation that were simply unavailable at the time of initial testing. Concurrently, improvements in bioinformatics tools and analytical methods substantially enhance the ability to detect and interpret genetic variants [10]. Updated annotation databases and more sophisticated variant calling algorithms can reveal previously missed or misclassified variants, as exemplified by cases where pathogenic (P) variants were initially overlooked due to annotation to less relevant isoforms [10].

The accumulation of knowledge regarding genotype-phenotype correlations and the potential for phenotype expansion over time can significantly aid in the interpretation of molecular findings during reanalysis [11]. Detailed re-evaluation of the patient's clinical information becomes crucial in this context, as evolving understanding of genetic disorders may reveal that a patient's symptoms align with recently described manifestations of particular genetic condition [12]. Furthermore, collaborative research efforts and data sharing platforms facilitate the validation of candidate variants and the identification of novel gene-disease associations. International collaborations, such as those demonstrated in the European Solve-RD study, have yielded new diagnoses through combined expertise that would not have been possible through isolated analysis [12].

Given these substantial advantages, recent literature increasingly recommends the routine reanalysis of unsolved NGS data after a period of time [7]. Studies suggest that a timeframe of 24 months or more after the initial analysis may enhance effectiveness by allowing sufficient time for new discoveries and advancements [13]. Implementation of reanalysis programs also faces practical challenges, including resource allocation considerations, data storage requirements, and ethical questions regarding consent and communication of new findings [6]. Nevertheless, the potential benefits are substantial, particularly for patients with rare or complex presentations.

In the present study, a reanalysis of WES data was performed on six patients who presented clinically with a suspected FA phenotype. The investigation aimed to identify genetic variants that could explain their conditions by examining genes beyond those traditionally

associated with FA. This approach exemplifies the utility of reanalyzing genomic data in cases where initial genetic testing did not yield a diagnosis, potentially providing answers for patients who have long sought explanations for their medical conditions and opening avenues for more targeted clinical management strategies.

Materials and methods

Patients

The study included six patients, including one male (16.7%) and five females (83.4%), who were referred for genetic testing based on initial clinical suspicion of FA. The relevant clinical history for each patient is summarised as follows: Patient 1: Diagnosed with immune thrombocytopenic purpura (ITP) at age 10 (2001). Subsequently developed bilateral interstitial pneumonia with respiratory failure and recurrent aphthous ulcers of the lips and oral mucosa. Chronic diarrhoea led to a diagnosis of irritable bowel syndrome. In late 2003, hematological evaluation revealed total bone marrow aplasia. Patient 2: Diagnosed in 2010 with mild leukopenia-neutropenia anemia and a mean corpuscular volume (MCV) within the upper normal range for age. Patient 3: A sample from this patient was referred for FA exclusion in 2007. Due to a lack of available clinical data from the time of the fragility study and the inability to contact the physician, further clinical details are unavailable. Patient 4: Moderate thrombocytopenia was detected at age two. Additional clinical features included microcephaly, café-au-lait spots, epileptic episodes, and developmental delay. The mother presented with microcephaly, and there was a family history of epilepsy. The father was reportedly healthy. The patient's thrombocytopenia worsened, necessitating a transplant that was refused by the family. The patient died shortly thereafter. Patients 5 and 6 (Siblings): These siblings presented with macrocytic anemia, short stature, and a syndromic appearance. Their father had a childhood diagnosis of FA. Patient 6 was diagnosed with FA in 1988, while Patient 5's diagnosis remained uncertain. Patient 6 exhibited a more severe phenotype, including agenesis of the left kidney, bilateral cryptorchidism, hypertelorism, microcephaly, winged auricles, and psychomotor retardation. Additional clinical data for all patients is provided in Table 1.

Standard exome sequencing and analysis

Genomic DNA samples were isolated and obtained from whole blood following standard procedures. WES for four of these individuals (designated as Patients 1 to 4) was carried out at *Sistema Genómicos* (Valencia, Spain) during July 2013, following their standard protocols. Subsequently, in early 2014, samples from the remaining two

patients (Patients 5 and 6) were sent to the GATC Biotech sequencing center (Konstanz, Germany).

Sequencing data reanalysis

The six patients had inconclusive results from the initial analysis focused on known FA genes and related ones. Sequencing fastq files were recovered and underwent a systematic reanalysis as it is shown in Fig. 1. Briefly, sequence reads were aligned against the GRCh37 human reference genome assembly using the Burrows-Wheeler Aligner (BWA) [14] and Picard tool from Broad Institute (<http://broadinstitute.github.io/picard/>). Genome Analysis Toolkit (GATK) [15] was employed for the calling of single nucleotide variants (SNVs) and indels while functional annotation was performed with the use of ANNOVAR [16]. Copy number variants (CNVs) were analyzed using RStudio and following packages: ExomeDepth [17], Genomic Ranges and IRanges [18]. Sequencing quality was assessed using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Quality coverage metrics were obtained with Picard: the mean exome coverage was 86.1X, the median coverage was 70.3X and the percentage of accumulated coverage at least at 20X was 87%. A virtual panel of FA-related genes (Table S1) was applied to filter identified variants in the first genetic analysis. Two main groups of variants were considered: 1) variants in genes associated with OMIM diseases; 2) variants in genes not associated with OMIM diseases. Variants of the first group were analysed and in case no candidate variants were found, the second group was examined considering available literature to elucidate if a gene could explain the clinical phenotype of the patient. As the siblings (patients 5 and 6) had similar clinical features, variants shared in both patients were also deeply assessed. Variant prioritization to identify potential candidate disease variants was carried out by using different resources, including allele frequency from public databases gnomAD [19] and Kaviar [20] and our in-house database of individuals from the same geographical area; impact on protein (missense, in-frame, splicing, frameshift, nonsense, CNV) and inheritance model. For missense rare variants, in silico prediction REVEL program was used to assess potential impact on protein function. In the case of CNVs, variants with a Bayes factor (BF) below 10 were discarded. Only deletions with a reads ratio between 0.4 and 0.6 and duplications with a reads ratio equal or higher than 1.45 were considered. Reciprocal overlapping intervals between CNVs of the six patients and the siblings were analyzed employing BEDTools [21]. Final candidate variants were reviewed by an expert team of clinicians and geneticists to elucidate their involvement in the

Table 1 Demographic and chromosome fragility data and clinical features of suspected Fanconi anemia studied patients

Patient	Birth date	Sex	Ethnicity	Year of first diagnosis	Fragility study date	% Aberrant cells (DEB)	First clinical phenotype	Disease course	Family history
1	06/08/1991	F	Caucasian	2001	25/01/2006	0	Immune thrombocytopenic purpura	Bone marrow aplasia in 2003 (allogeneic transplant from a family donor), bilateral interstitial pneumonias with associated respiratory failure (lung transplant in 2021), aphthosis, irritable bowel syndrome, severe neutropenia in the most recent clinical report	-
2	04/04/2001	F	Caucasian	2010	01/03/2006	2	Mild leukopenia-neutropenia and anemia	Anemia was normalized, leukopenia maintained, medical discharge in 2015	-
3	26/12/2006	F	Caucasian	-	31/07/2007	4	Referred for "FA exclusion in 2007"	Not available	-
4	10/02/2001	F	Gypsy	2003	06/11/2007	9	Moderate thrombocytopenia, cafe-au-lait spots, microcephaly, epileptic episodes and mild mental retardation	Severe megakaryocyte thrombocytopenia in 2014 (transplant required, but the family refused, and the patient died shortly afterward)	Healthy father, mother with microcephaly and family history of epilepsy
5*	04/12/1979	F	Caucasian	1988	29/01/2009	11	Macrocytic anemia, short stature and a syndromic appearance	-	Paternal family history: FA diagnosed in childhood and subsequently ruled out
6*	21/06/1981	M	Caucasian	1988	29/01/2009	13	Macrocytic anemia, short stature and syndromic appearance, agenesis of the left kidney, bilateral cryptorchidism, hypertelorism, microcephaly, winged auricles, and psychomotor retardation	-	Paternal family history: FA diagnosed in childhood and subsequently ruled out

patient's phenotype. The ACMG/AMP guidelines were considered to give a final classification [22]. Candidate variants were confirmed using Sanger sequencing.

Topoisomerase-based cloning assay was used to identify variants in *trans* when no family relatives were available.

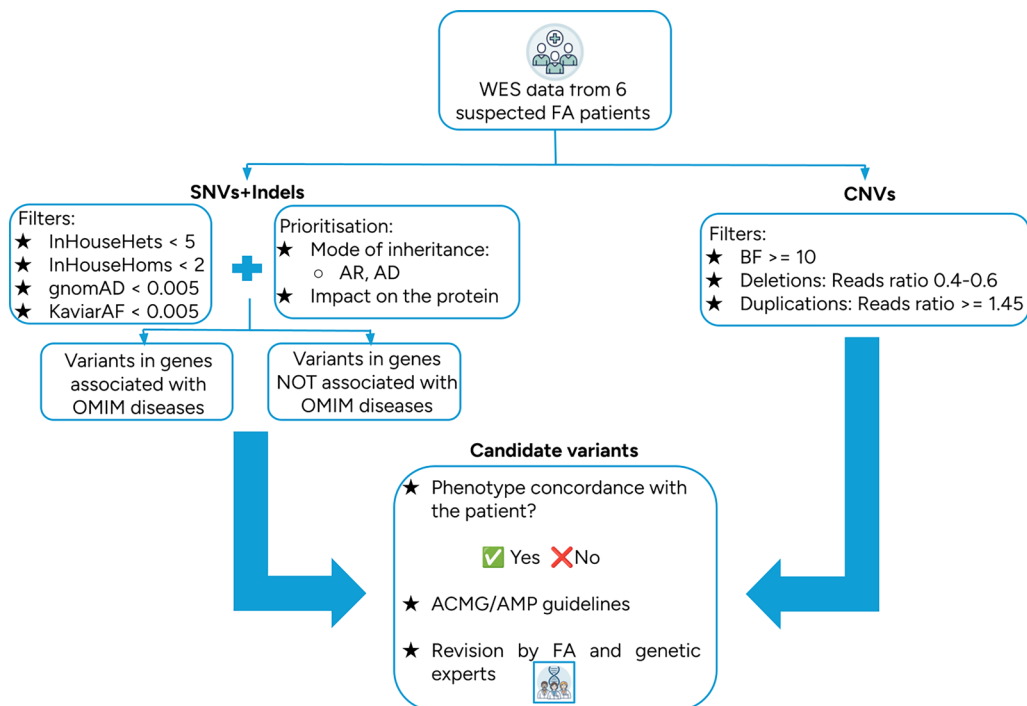


Fig. 1 Workflow of the reanalysis. Six patients with suspected FA clinical phenotype but with negative chromosome fragility test and no FA-related variants identified after WES were selected for reanalysis. SNVs and indels were filtered based on <0.5% allele frequency in variant databases as well as below 2 homozygotes or 5 heterozygotes individuals at our in-house database and prioritized by the disease mode of inheritance and effect on the encoded protein. CNVs were analyzed with ExomeDepth and filtered by BF and reads ratio. Finally, candidate variants were obtained and discussed in a committee with FA and genetic experts to elucidate their implication in their respective patients' condition. Variants were classified according to ACMG/AMP guidelines

Telomere length and chromosome fragility

Telomere length was measured as previously described by Cawthon (2009) [23] with some modifications. Briefly, genomic DNA was extracted from whole blood, purified using the DNeasy® PowerClean® Pro Cleanup Kit (Qiagen, 12,997–50), and quality controlled. Telomere length was measured using the monochrome multiplex qPCR on a 7900HT Applied Biosystems instrument. Reactions included telomere and albumin primers and PowerUp™ SYBR™ Green Master Mix (Applied Biosystems). Each sample was analyzed in triplicate wells on two separate plates. A five-point standard curve of control DNA was used to calculate the T/S ratio. Data analysis was performed using the Design and Analysis software (v.2.6.0; Thermo Fisher). Chromosome fragility was assessed in all patients using DEB induction, as previously described [24].

Results

Chromosome fragility and G2 arrest data (G2 data not shown) were incompatible with a classical FA phenotype (Fig. 2), disregarding FA as initial diagnosis. Patients

(1–6) showed values overlapping with those from non-Fanconi patients' group (Fig. 2).

Consistently with negative chromosome fragility test, no variants were identified in none of the known FA associated genes (Table S1) in the standard analysis. The performed reanalysis of sequencing data in other genes identified three clinically significant variants in two out of six patients (33.4%) (Table 2). Patient 1 showed two variants in the *TERT* gene: The first one was a nonsense variant creating a premature stop codon and a likely loss of function of the protein codified by the allele (NM_198253.3, c.337G>T:p.Glu113*). The second variant consisted in a missense change (c.193C>A:p.Pro65Thr). The variants were confirmed by Sanger sequencing and the topoisomerase-based cloning analysis showed that they were in *trans*. Given that *TERT* mutations are associated to telomere shortening, telomere length was measured in Patient 1 and the results were compatible with a telomeropathy (Fig. 3), exhibiting the most pronounced telomere shortening, with a z score considerably lower than the control mean, below percentile 1.

In Patient 2, a *RPL5* frameshift variant (NM_000969.5: c.175_176del; p.Asp59Tyrfs*53) was identified. This *RPL5*

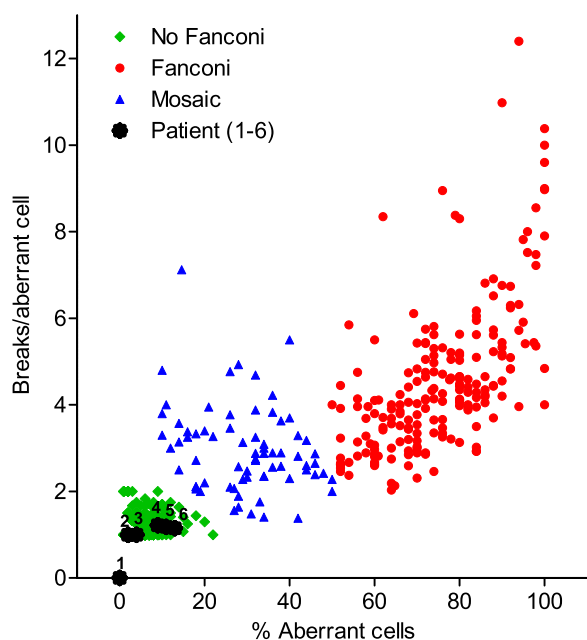


Fig. 2 Chromosome fragility results. The plot shows a graphical representation of the relative position of cases (suspected FA patients 1–6, with numbers over asterisk symbols in the plot) when considering the distribution of historical data from our laboratory in terms of percentage of aberrant cells and mean number of breaks per cell (all induced by DEB at 0,1 µg/ml). Green signals correspond to non-Fanconi patients, red are FA patients, blue represent mosaic FA patients, asterisk symbols correspond to analysed patients in this work. Mosaic patients are those FA patients with a percentage of aberrant cells below 50%

variant, which introduces a premature stop signal might cause loss of function, either by a shortened, non-functional protein or by triggering nonsense-mediated decay (NMD) of the mRNA. Telomere length analysis in Patient 2 revealed shortened telomeres with a z score below percentile 10 but above percentile 1 (Fig. 3); thus, telomeropathy in this patient based on this parameter was discarded.

In the remaining four patients, no clinically significant variants were identified in genes with known disease associations and concordant with their observed phenotype. Although patients 3 to 5 have short telomeres (below percentile 25) (Fig. 3), they exhibit a z score

above the 1st percentile, suggesting incompatibility with a telomeropathy. Additionally, the analysis of non-OMIM genes failed to reveal a novel candidate gene that could be associated with the phenotype, according to existing evidence and knowledge. Rare variants of uncertain significance (VUS) with an extreme low population frequency, predicted to be highly deleterious and/or identified in both siblings, and residing in genes associated to clinical phenotypes or functions partially overlapping patients' features were identified (Table S2), but without enough evidence to classify them as P or likely pathogenic (LP) following ACMG/AMP guidelines.

Discussion

Reanalysis of WES data led to the identification of disease-causing variants in patients 1 and 2. In Patient 1 a nonsense variant (NM_198253.3: c.337G>T; p.Glu113*) and a missense change (NM_198253.3: c.193C>A; p.Pro65Thr) were identified in the *TERT* gene. The nonsense variant shows an extremely low frequency in gnomAD and loss of function is a known mechanism of disease in the gene [25]. According to this evidence, the variant was classified as LP. The missense change also presents an extremely low frequency in gnomAD and the gene shows a low rate of benign missense mutations which are a common mechanism of disease. Although the variant was found in a patient with aplastic anemia also carrying a second missense change [26], it was classified as VUS as there was not enough available evidence to confirm its pathogenicity. *TERT* gene is associated with autosomal dominant or recessive DC, a telomere disease. While there is significant overlap in the clinical features of DC regardless of inheritance pattern, some nuances can exist: severe forms of DC, like Hoyeraal-Hreidarsson syndrome, which often involves cerebellar hypoplasia and developmental delay, can be associated with both autosomal recessive and dominant inheritance, depending on the specific gene involved. The age of onset and the rate of progression can vary between individuals and families, and this variation is not strictly tied to the inheritance pattern but more to the specific gene and mutation. The classic triad of nail dystrophy, lacy reticular pigmentation, and oral leukoplakia can be present in both autosomal recessive and dominant forms, but it is

Table 2 Clinically relevant variants identified after reanalyzing exome sequencing data

Patient	Sex of the patient	Gene	Reference transcript	Variant	Aminoacid change	Variant type	Variant classification	ACMG evidences
1	Female	<i>TERT</i>	NM_198253.3	c.337G>T	p.Glu113*	Nonsense	LP	PVS1, PM2
				c.193C>A	p.Pro65Thr	Missense	LP	PM2, PP2, PM3
2	Female	<i>RPL5</i>	NM_000969.5	c.175_176del	p.Asp59Tyrf5*5 3	Frameshift	P	PVS1, PM2, PS4, PP5

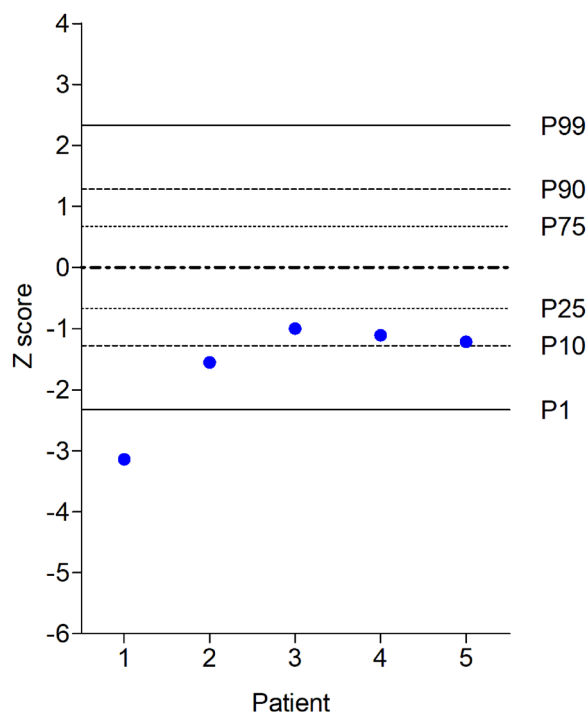


Fig. 3 Telomere length in suspected Fanconi anemia patients. Patients 1–5 with suspected FA represented in x-axis. Patient 6 not analyzed due to technical issues. Horizontal lines represent the mean (0), and the values -2.326, -1.282, -0.674, +0.674, +1.282, and +2.326 correspond to the standard deviations from the mean, equivalent to the 1st, 10th, 25th, 75th, 90th, and 99th percentiles (P1–P99) for control individuals

important to note that not all individuals with DC present with all three features. Both autosomal recessive and autosomal dominant forms of DC are characterized by shortened telomeres. The degree of telomere shortening can vary and may correlate somewhat with disease severity, but this isn't exclusively determined by the inheritance pattern. There is significant phenotypic variability in DC, even within families with the same mutation. Anticipation can occur in DC, particularly in autosomal dominant forms, where successive generations may exhibit earlier onset and more severe symptoms. The studied patient presents severe telomere shortening (Fig. 3) and developed a bone marrow aplasia which required an allogeneic transplant in 2003. In 2021 she was diagnosed with interstitial lung disease and received a lung transplant. The patient has also presented chronic diarrhea that was related to irritable bowel syndrome. In the most recent clinical report, she manifested severe neutropenia. The variants were confirmed by Sanger sequencing and the topoisomerase-based cloning analysis showed that they were in *trans*. This evidence gives the variant the ACMG/AMP PM3 evidence, that is a variant occurring in *trans* with a P variant for a recessive disorder, therefore

it can be classified as LP. The DC caused by *TERT* gene has been associated to dominant and recessive inheritance patterns, with the latter having additional characteristics such as leukoplakia, failure to thrive, cerebellar hypoplasia, microcephaly and developmental delay [27] not presented in our patient. The telomere length of this patient was 2,608 base pairs (bp) at 14.3 years of age, with a z score of -3.141, significantly below the 1st percentile (-2.326). This indicates severe telomere shortening, consistent with a telomere disease (Fig. 3).

In the case of the *RPL5* variant found in Patient 2 (NM_000969.5: c.175_176del; p.Asp59Tyrfs*53), it was classified as P based on the following evidence: loss-of-function variant in *RPL5*, a known mechanism in patients [28, 29], the deletion is extremely rare (no frequency at gnomAD) but it has been observed in individuals with DBA [28, 30, 31]. DBA is an inherited red blood cell aplasia that usually presents in the first year of life. The main features are normochromic macrocytic anemia, reticulocytopenia, and nearly absent erythroid progenitors in the bone marrow. Patients show growth retardation, and approximately 30 to 50% have craniofacial, upper limb, heart, and urinary system congenital malformations. Most patients have increased mean corpuscular volume, elevated erythrocyte adenosine deaminase activity, and persistence of hemoglobin F. However, some DBA patients do not exhibit these findings, and even in the same family, symptoms can vary between affected family members. The studied patient showed a mild disease phenotype as in 2010 she had anemia with a MCV at the normal-high limit for her age and a mild leukopenia-neutropenia. In the most recent clinical report, the anemia was normalized while she maintained the mild leukopenia. This can be explained by the clinical heterogeneity of the disease [32]: individuals with the same variant identified in our study exhibit different clinical characteristics [28]. In some cases, affected individuals within the same family may show mild anemia or even no evidence of clinical disease [28, 33]. In addition, this patient had shortened telomeres compared with controls (at 10.9 years of age was 7,116 bp, which is below the 10th percentile (-1.282) but above the 1st percentile (-2.326), with a z score of -1.557, although this was not as severe as in patients with DC. This finding is consistent with the diagnosis of DBA reported by Alter et al. (2015) [34].

Despite a comprehensive reanalysis of OMIM and non-OMIM genes, no further variants of interest were detected in the remaining patients. The absence of findings is likely multifactorial, with one key factor being the limitations of the short-read WES methodology used. Short-read sequencing is known to have reduced sensitivity for structural variants (SVs), due to the difficulty in assembling reads across complex genomic

rearrangements, especially within repetitive regions [35]. Furthermore, WES's exon-centric design hinders the detection of variants in critical non-coding regions [36], such as deep intronic and regulatory sequences. Other limitations, including uneven coverage in GC-rich regions [37] and the difficulty in detecting small CNVs [38], also contribute to the potential for missed variants. Long-read sequencing could overcome these limitations by effectively resolving complex and repetitive regions, leading to improved detection of SVs and other variant types, including SNVs [39–43]. Similarly, Whole Genome Sequencing (WGS) provides a more comprehensive genomic view, allowing for the identification of SVs and CNVs frequently missed by WES and enabling the exploration of potentially P variants in non-coding regions [44–46]. Another contributing factor may be the limited knowledge of many reviewed non-OMIM genes, which resulted in numerous rare VUS. In addition, there is also a possibility of a disease caused by the combined effects of multiple genes, each contributing a small risk. WES may identify some of these risk variants, but it is difficult to determine their overall contribution to the disease.

The findings of this study highlight the clinical benefit of reanalyzing WES data in patients with undiagnosed genetic conditions, particularly after initial, targeted analyses focusing on canonical FA genes have proven negative, while phenotype-specific assays also yield non-diagnostic results. Despite the modest number of studied patients, the percentage of new diagnoses identified through a subsequent, broader analysis of WES data, which included a more comprehensive panel of genes implicated in BMF syndromes, aligns with the diagnostic yields reported in larger cohorts (10–25%). This work underscores the importance of expanding the scope of genetic investigation beyond traditional FA genes in patients presenting suspected FA features but normal chromosome fragility. We propose that future studies implement an expanded gene panel in these phenotypes during the reanalysis of WES data. This strategy is expected to enhance diagnostic rates by detecting a wider array of genetic causes, ultimately leading to improved patient care.

Conclusions

Systematic reanalysis of WES data at appropriate intervals is a cost-effective and crucial approach for long-undiagnosed conditions. In this study we achieved a 33.3% diagnostic yield in a small cohort of six patients initially suspected of FA who had remained undiagnosed through conventional methods.

In our study, we diagnosed patients with DC and DBA, and our findings were corroborated by integrating them

with clinical and functional data, including telomere length measurements.

Our work highlights the importance of allowing sufficient time for clinical databases to expand, variant annotation to improve, and novel disease-gene relationships to be identified. In this context, WES reanalysis becomes an indispensable tool for maximizing diagnostic accuracy and delivering personalized care to patients who have long sought answers. Our study also recalls the very close clinical, cellular, and molecular links between these three syndromes as previously discussed [47].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13023-025-03928-5>.

Additional file 1.

Additional file 2.

Acknowledgements

Not applicable.

Author contributions

E.T.L. has participated in the acquisition, analysis, and interpretation of data. He has also contributed to the manuscript draft and revised the manuscript. He has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature. M.J.R. performed and analyzed the telomere data and prepared the corresponding figure. She also discussed and reviewed the manuscript. Also she has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature. R.P. performed chromosome fragility analysis and collected clinical data of the patients. Also she has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature. M.B. has participated in the NGS analysis of patients, the interpretation of data, drafted and revised the manuscript. Also, he has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in literature. B.R.S. has participated in the conception, design of the work, interpretation of data, drafted and revised the manuscript. Also, he has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in literature. J.S. designed experiments, analyzed and interpreted patients' data and contributed to financing the study and writing of the manuscript. Also, he has approved the submitted version and agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in literature.

Funding

E. T. L. is recipient of a *Formación Profesorado Universitario* (FPU) grant funded by the Ministry of Science, Innovation and Universities of the Government of Spain that supports his doctoral studies (grant code: FPU17/04141). Our research team is supported by the Spanish Ministry of Science, Innovation and Universities (PID2021-122411OB-I00/AEI, UE; Prueba de Concepto project PDC2022-133233-I00, financed by MICIU/AEI /10.13039/501100011033 and European Union Next Generation EU/PRTR; *Red de Investigación Colaborativa* RED2022-134164-T), the Autonomous Community of Catalonia (PI-G7180 HSP), the *Instituto de Salud Carlos III* (IC122/00076, IMPaCT *Genómica*- Ref. IMP/00009 and IMPaCT VUSCAN- Ref. PMP22/00064; co-financed with Next Generation EU Funds and the Recovery, Transformation and Resilience Plan), the ICREA-*Acadèmia* program (*Fundació Institut Català de Recerca i Estudis Avançats*), the *Agència de Gestió d'Ajuts Universitaris i de Recerca* (2021-SGR-00835, 2023–2025), the *Centro de Investigación Biomédica en Red en Enfermedades Raras* (CIBERER), the *Fanconi Cancer Foundation*, and *Fundació La Marató de TV3* (ref. 202040–33).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All relevant ethical regulations were applied and informed consent was obtained from all participants. This research has been performed in accordance with the Declaration of Helsinki.

Competing interests

J.S. has signed agreements in the last 3 years (advisory boards, service provision, research collaboration, or material transfer) with Boehringer Ingelheim, Glaxo Smith Kline, Moderna Therapeutics, Roche, Rocket Pharmaceuticals and Pfizer.

All the other authors declare that they have no conflict of interest.

Author details

¹Institut de Recerca Sant Pau (IR, SANT PAU) and Joint Research Unit on Genomic Medicine, Universitat Autònoma de Barcelona (UAB)-IR SANT PAU, Barcelona, Spain. ²Centro de Investigación Biomédica en Red de Enfermedades Raras, Instituto de Salud Carlos III (CIBERER, ISCIII), Madrid, Spain. ³Serra Hunter Fellow, Department of Genetics and Microbiology, Faculty of Biosciences, Universitat Autònoma de Barcelona, Barcelona, Spain. ⁴Hospital de La Santa Creu i Sant Pau, Barcelona, Spain.

Received: 6 June 2025 Accepted: 16 July 2025

Published online: 14 October 2025

References

1. Peake JD, Noguchi E. Fanconi anemia: current insights regarding epidemiology, cancer, and DNA repair. *Hum Genet.* 2022;141(12):1811–36.
2. Ramirez MJ, Pujol R, Minguillon J, Bogliolo M, Persico I, Caverio D, et al. Prognostic significance of mutation type and chromosome fragility in Fanconi anemia. *Am J Hematol.* 2025;100(2):272–84.
3. Bogliolo M, Pujol R, Aza-Carmona M, Munoz-Subirana N, Rodriguez-Santiago B, Casado JA, et al. Optimised molecular genetic diagnostics of Fanconi anaemia by whole exome sequencing and functional studies. *J Med Genet.* 2020;57(4):258–68.
4. Joshi G, Arthur NBJ, Geetha TS, Datar P, Modak K, Roy D, et al. Comprehensive laboratory diagnosis of Fanconi anaemia: comparison of cellular and molecular analysis. *J Med Genet.* 2023;60(8):801–9.
5. Ramirez MJ, Pujol R, Trujillo-Quintero JP, Minguillon J, Bogliolo M, Rio P, et al. Natural gene therapy by reverse mosaicism leads to improved hematology in Fanconi anemia patients. *Am J Hematol.* 2021;96(8):989–99.
6. van der Geest MA, Maekelberghe ELM, van Gijn ME, Lucassen AM, Swertz MA, van Langen IM, et al. Systematic reanalysis of genomic data by diagnostic laboratories: a scoping review of ethical, economic, legal and (psycho)social implications. *Eur J Hum Genet.* 2024;32(5):489–97.
7. Hiatt SM, Amaral MD, Bowling KM, Finnilla CR, Thompson ML, Gray DE, et al. Systematic reanalysis of genomic data improves quality of variant interpretation. *Clin Genet.* 2018;94(1):174–8.
8. Dai P, Honda A, Ewans L, McGaughan J, Burnett L, Law M, et al. Recommendations for next generation sequencing data reanalysis of unsolved cases with suspected Mendelian disorders: a systematic review and meta-analysis. *Genet Med.* 2022;24(8):1618–29.
9. Fellner A, Wali GM, Mahant N, Grosz BR, Ellis M, Narayanan RK, et al. Genome sequencing reanalysis increases the diagnostic yield in dystonia. *Parkinsonism Relat Disord.* 2024;124: 107010.
10. Schobers G, Schieving JH, Yntema HG, Pennings M, Pfundt R, Derks R, et al. Reanalysis of exome negative patients with rare disease: a pragmatic workflow for diagnostic applications. *Genome Med.* 2022;14(1):66.
11. Karaca E, Posey JE, Coban Akdemir Z, Pehlivan D, Harel T, Jhangiani SN, et al. Phenotypic expansion illuminates multilocus pathogenic variation. *Genet Med.* 2018;20(12):1528–37.
12. Stephenson KAJ, Zhu J, Dockery A, Whelan L, Burke T, Turner J, et al. Clinical and genetic re-evaluation of inherited retinal degeneration pedigrees following initial negative findings on panel-based next generation sequencing. *Int J Mol Sci.* 2022. <https://doi.org/10.3390/ijms23020995>.
13. Costain G, Jobling R, Walker S, Reuter MS, Snell M, Bowdin S, et al. Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. *Eur J Hum Genet.* 2018;26(5):740–4.
14. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–303.
16. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16): e164.
17. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics.* 2012;28(21):2747–54.
18. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol.* 2013;9(8): e1003118.
19. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434–43.
20. Glusman G, Caballero J, Mauldin DE, Hood L, Roach JC. Kaviar: an accessible system for testing SNV novelty. *Bioinformatics.* 2011;27(22):3216–7.
21. Quinlan AR, Hall IM. BEDtools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841–2.
22. 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. *Genet Med.* 2015;17(5):405–24.
23. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res.* 2009;37(3): e21.
24. Castella M, Pujol R, Callen E, Ramirez MJ, Casado JA, Talavera M, et al. Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact. *J Med Genet.* 2011;48(4):242–50.
25. Armanios M, Chen JL, Chang YP, Brodsky RA, Hawkins A, Griffin CA, et al. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci USA.* 2005;102(44):15960–4.
26. Arias-Salgado EG, Galvez E, Planas-Cerezales L, Pintado-Berninches L, Vallespin E, Martinez P, et al. Genetic analyses of aplastic anemia and idiopathic pulmonary fibrosis patients with short telomeres, possible implication of DNA-repair genes. *Orphanet J Rare Dis.* 2019;14(1):82.
27. Marrone A, Walne A, Tamarly H, Masunari Y, Kirwan M, Beswick R, et al. Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal-Hreidarsson syndrome. *Blood.* 2007;110(13):4198–205.

28. Gazda HT, Sheen MR, Vlachos A, Choessel V, O'Donohue MF, Schneider H, et al. Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am J Hum Genet.* 2008;83(6):769–80.
29. Quarello P, Garelli E, Carando A, Brusco A, Calabrese R, Dufour C, et al. Diamond-blackfan anemia: genotype-phenotype correlations in Italian patients with RPL5 and RPL11 mutations. *Haematologica.* 2010;95(2):206–13.
30. Gerrard G, Valganon M, Foong HE, Kasperaviciute D, Iskander D, Game L, et al. Target enrichment and high-throughput sequencing of 80 ribosomal protein genes to identify mutations associated with Diamond-Blackfan anaemia. *Br J Haematol.* 2013;162(4):530–6.
31. Tsangaris E, Klaassen R, Fernandez CV, Yanofsky R, Shereck E, Champagne J, et al. Genetic analysis of inherited bone marrow failure syndromes from one prospective, comprehensive and population-based cohort and identification of novel mutations. *J Med Genet.* 2011;48(9):618–28.
32. Landowski M, O'Donohue MF, Buros C, Ghazvinian R, Montel-Lehry N, Vlachos A, et al. Novel deletion of RPL15 identified by array-comparative genomic hybridization in Diamond-Blackfan anemia. *Hum Genet.* 2013;132(11):1265–74.
33. Errichiello E, Vetro A, Mina T, Wischmeijer A, Berrino E, Carella M, et al. Whole exome sequencing in the differential diagnosis of Diamond-Blackfan anemia: clinical and molecular study of three patients with novel RPL5 and mosaic RPS19 mutations. *Blood Cells Mol Dis.* 2017;64:38–44.
34. Alter BP, Giri N, Savage SA, Rosenberg PS. Telomere length in inherited bone marrow failure syndromes. *Haematologica.* 2015;100(1):49–54.
35. Kosugi S, Terao C. Comparative evaluation of SNVs, indels, and structural variations detected with short- and long-read sequencing data. *Hum Genome Var.* 2024;11(1):18.
36. Burdick KJ, Cogan JD, Rives LC, Robertson AK, Koziura ME, Brokamp E, et al. Limitations of exome sequencing in detecting rare and undiagnosed diseases. *Am J Med Genet A.* 2020;182(6):1400–6.
37. Wang Q, Shashikant CS, Jensen M, Altman NS, Girirajan S. Novel metrics to measure coverage in whole exome sequencing datasets reveal local and global non-uniformity. *Sci Rep.* 2017;7(1):885.
38. Yao R, Zhang C, Yu T, Li N, Hu X, Wang X, et al. Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. *Mol Cytogenet.* 2017;10:30.
39. Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. *Nat Rev Genet.* 2020;21(10):597–614.
40. Iyer SV, Goodwin S, McCombie WR. Leveraging the power of long reads for targeted sequencing. *Genome Res.* 2024;34(11):1701–18.
41. Warburton PE, Sebra RP. Long-read DNA sequencing: recent advances and remaining challenges. *Annu Rev Genomics Hum Genet.* 2023;24:109–32.
42. Wohlers I, Garg S, Hehir-Kwa JY. Editorial: Long-read sequencing-pitfalls, benefits and success stories. *Front Genet.* 2022;13: 1114542.
43. Olivucci G, Iovino E, Innella G, Turchetti D, Pippucci T, Magini P. Long read sequencing on its way to the routine diagnostics of genetic diseases. *Front Genet.* 2024;15:1374860.
44. Bowman P, Grimes H, Dallosso AR, Berry I, Mullin S, Rankin J, et al. Whole genome sequencing for copy number variant detection to improve diagnosis and management of rare diseases. *Dev Med Child Neurol.* 2025;67(1):126–31.
45. Nomakuchi TT, Teferedegny EY, Li D, Muirhead KJ, Dubbs H, Leonard J, et al. Utility of genome sequencing in exome-negative pediatric patients with neurodevelopmental phenotypes. *Am J Med Genet A.* 2024;194(12): e63817.
46. Brlek P, Bulic L, Bracic M, Projic P, Skaro V, Shah N, et al. Implementing whole genome sequencing (WGS) in clinical practice: advantages, challenges, and future perspectives. *Cells.* 2024;13(6):504.
47. Gueiderikh A, Maczkowiak-Chartois F, Rosselli F. A new frontier in Fanconi anemia: from DNA repair to ribosome biogenesis. *Blood Rev.* 2022;52: 100904.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.