Original Article

Short-Term Lymphocyte Culture Improves the Diagnostic Yield of Targeted RNA NGS in Cancer Predisposition Testing

(RNA / sequence capture / lymphocyte culture / NMDi / NGS / CZECANCA / hereditary cancer predisposition / germline genetic testing / alternative splicing / aberrant splicing / gene expression / reproducibility / FANCA)

MARTA ČERNÁ[1], KATEŘINA MATĚJKOVÁ[1,2], TAŤÁNA PTÁČKOVÁ[1], PETR NEHASIL[1,3,4], ROMANA MIHALOVÁ[5], KAMILA VESELÁ[5], MARKÉTA JANATOVÁ[1], JANA SOUKUPOVÁ[1], PETRA KLEIBLOVÁ[1,5]

[1]Institute of Medical Biochemistry and Laboratory Diagnostics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic
[2]Department of Genetics and Microbiology, Faculty of Science, Charles University, Prague, Czech Republic
[3]Institute of Pathological Physiology, First Faculty of Medicine, Charles University, Prague, Czech Republic
[4]Department of Paediatrics and Inherited Metabolic Disorders, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic
[5]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic

Abstract. Although next-generation RNA sequencing (RNA-seq) is increasingly incorporated into germline cancer predisposition testing, its diagnostic utility is often limited by low expression of many clinically relevant genes. To improve RNA yield and transcript representation for targeted RNA-seq, we optimized a simple protocol based on short-term lymphocyte culture prepared directly from whole blood collected in Li-heparin tubes. We systematically evaluated biological reproducibility and pre-analytical sample handling variability and demonstrated that whole blood can be stored at 4 °C for up to 5 days prior to lymphocyte cultivation without compromising RNA quality/gene expression. Gene expression was comparable for RNA isolated from K_2EDTA and Tempus tubes, whereas short-term lymphocyte culture resulted in a substantial increase in expression of clinically important genes including *BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *PALB2*, *CHEK2*, and multiple Fanconi anaemia genes otherwise low expressed in whole blood. Cultivation for 3–5 days did not significantly affect lymphocyte gene expression, providing flexibility for routine diagnostics. The protocol also enables inhibition of nonsense-mediated decay to facilitate analysis of variants causing premature termination. As a proof of principle, we characterized the splicing impact of *FANCA* c.2602-3C>G variant (located in intron 27) using cultured lymphocytes from its carrier. The variant causes deletion of six nucleotides in the mature transcript (ΔE28p(–6)/r.2602_2607del), resulting in an in-frame deletion (p. Gln869_Phe870del). This spliceogenic effect was reliably detectable only in cultured lymphocytes preferentially expressing the full-length *FANCA* transcript. Overall, short-term lymphocyte culture represents a simple and flexible RNA source that enhances variant interpretation in clinical RNA-seq analyses.

Received November 19, 2025. Accepted January 27, 2026.

This study was supported by the Ministry of Health of the Czech Republic grant projects (NU23-03-00150 and RVO-VFN 64165); Charles University research projects (SVV 260631 and Cooperatio); and the Ministry of Education, Youth and Sports of the Czech Republic project MULTIOMICS_CZ (Programme Johannes Amos Comenius, ID Project CZ.02.01.01/00/23_020/0008540) – co-funded by the European Union.

Corresponding author: Petra Kleiblová, Laboratory of Oncogenetics, Institute of Medical Biochemistry and Laboratory Diagnostics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Na Bojisti 1660/3, Prague 2, 128 00, Czech Republic. E-mail: pekleje@lf1.cuni.cz

Abbreviations: BAM – Binary Alignment Map, b(p) – base (pair), CZECANCA – CZEch CAncer paNel for Clinical Application, GTEx – Genotype-Tissue Expression, HBOC – hereditary breast and ovarian cancer, LCLs – lymphoblastoid cell lines, NGS – next-generation sequencing, NMD – nonsense-mediated decay, NMDi – NMD inhibition, PBS – phosphate-buffered saline, PCA – principal component analysis, RNA-seq – RNA-NGS, TPM – transcripts per million, VCF – variant call format. Abbreviations for genes in CZECANCA are provided in Supplementary Table S2

Copyright: © 2026 Author et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).

https://doi.org/10.14712/fb2026.0001

Introduction

Precise identification of germline pathogenic variants is crucial for personalized cancer prevention in asymptomatic carriers and for guiding treatment of patients with hereditary cancers (Stadler et al., 2021).

Over the past decade, next-generation DNA sequencing (NGS) has been routinely used to diagnose cancer predisposition. The majority of diagnostic protocols are based on DNA analysis of targeted gene panels or exomes (Soukupová et al., 2018; LaDuca et al., 2020; Kuzbari et al., 2023). In parallel, various NGS-based RNA-seq approaches have been adopted to analyse gene expression and splicing (Hong et al., 2020; Wen and Tang, 2025).

However, the clinical utility of RNA-seq analyses in cancer predisposition testing is often limited by low mRNA expression of many genes of interest in commonly available biological tissues. This limitation can be partially addressed by RNA target capture, enabling enrichment of transcripts of interest (Davy et al., 2017; Karam et al., 2020; Horton et al., 2022, 2024). We have recently described a diagnostic pipeline using an identical capture panel, CZECANCA, targeting 226 genes for parallel RNA and DNA analysis of germline cancer predisposition (Kleiblová et al., 2024). The addition of RNA-based NGS in clinical settings reduces hands-on time for variant confirmation, enables assessment of variant effect on pre-mRNA splicing, increases the accuracy of sample authentication and allows identification of tissue-specific alternative splicing. Nevertheless, the above-mentioned insufficient expression of genes of interest in the analysed tissue remains the principal limitation (Rowlands et al., 2022). Besides peripheral blood, which is the most frequent RNA source in routine clinical setting, nasopharyngeal epithelial cells are also easily accessible with minimal invasiveness (Kleiblová et al., 2024). Other specimens, including skin biopsies, skin fibroblasts or skeletal muscle samples, cannot be obtained routinely.

Another limitation of RNA-based evaluation of variant effect might be associated with nonsense-mediated decay (NMD). It eliminates aberrant mRNA transcripts containing premature termination codons and thereby hinders their detectability in the analysed samples. To overcome this limitation, inhibition of NMD (NMDi) could be applied; however, this process requires NMDi in viable, actively dividing cells. Although immortalized lymphoblastoid cell lines (LCLs) represent a gold standard for NMDi studies, their time-consuming and technically challenging nature prevents routine implementation (Walker et al., 2019; Mesman et al., 2020; Rowlands et al., 2022).

In this work, we described a simple protocol using short-term peripheral blood lymphocyte culture as an excellent source of high-quality biological material for RNA NGS-based analysis of cancer predisposition genes, including cultivation with NMDi to identify degraded transcripts. We also optimized the pre-analytical phase to define acceptable limits for sample transport and processing while maintaining the required RNA quality. Subsequently, we compared expression of genes of interest in cultured lymphocytes and whole blood. Finally, we demonstrated practical applicability of our approach by *FANCA* gene mRNA splicing analysis.

Material and Methods

Outline of optimization steps of the lymphocyte culture protocol

The short-term lymphocyte culture was based on a generic protocol for cultivation of peripheral blood lymphocytes for karyotyping (Ibraimov, 1983; Claussen et al., 2002).

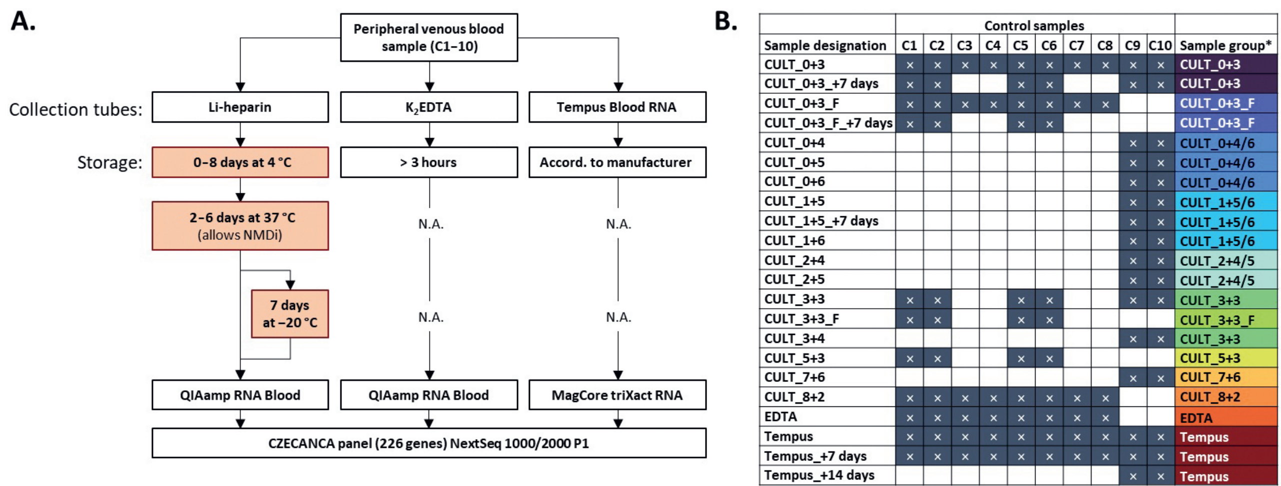
The aim of optimizing the pre-analytical phase of blood collected in Li-heparin tubes was to determine the effect of storage/transport (0–8 days at 4 °C) before lymphocyte cultivation and cultivation duration (2–6 days at 37 °C) on the results of the targeted RNA NGS analysis. In addition, we tested short-term storage of pelleted cultured lymphocytes (7 days at −20 °C) prior to RNA isolation to increase flexibility of downstream sample processing in routine clinical settings. We compared RNA isolated from cultured lymphocytes (allowing inclusion of a NMDi step if required) with RNA isolated from K_2EDTA blood (representing archival RNA sources from whole blood in many laboratories) and Tempus Blood RNA tubes (representing a standard RNA-preserving procedure) using CZECANCA panel RNA NGS-based gene expression analyses (Fig. 1A).

Patients and samples

To optimize the protocol, we obtained peripheral venous blood control samples from 10 healthy voluntary donors (C1–10; Fig. 1B). Paired RNA samples (cultured lymphocytes with and without NMDi) from seven patients undergoing RNA genetic testing for cancer predisposition were used as a real-sample set for expression analyses (P1–P7, details in Supplementary Table S1). All participants provided written informed consent with genetic testing approved by the Ethics Committee of General University Hospital in Prague, and the study was performed in accordance with the Declaration of Helsinki. All participants were Slavs of the Czech origin.

Lymphocyte culture

Peripheral blood collected in Li-heparin tubes was stored at 4 °C for 0–8 days. Approximately 0.5 ml of blood was inoculated into NUNC tissue culture flasks (Thermo Fisher Scientific, Waltham, MA) containing 5 ml of Chromosome Medium P (EuroClone, Pero MI, Italy). Consequently, samples were cultured at 37 °C for 2–6 days. Samples were designated as CULT_x+y, where x indicates the number of days of storage at 4 °C and y indicates the number of days of cultivation (Fig. 1B). To inhibit NMD, cultures were treated with cycloheximide (at a final concentration of 300 μg/ml, Santa Cruz Biotechnology, Dallas, TX) for 4 hours prior to RNA isolation.



	Control samples										
Sample designation	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Sample group*
CULT_0+3	×	×	×	×	×	×	×	×	×	×	CULT_0+3
CULT_0+3_+7 days	×	×			×	×			×	×	CULT_0+3
CULT_0+3_F	×	×	×	×	×	×	×	×			CULT_0+3_F
CULT_0+3_F_+7 days	×	×			×	×					CULT_0+3_F
CULT_0+4									×	×	CULT_0+4/6
CULT_0+5									×	×	CULT_0+4/6
CULT_0+6									×	×	CULT_0+4/6
CULT_1+5									×	×	CULT_1+5/6
CULT_1+5_+7 days									×	×	CULT_1+5/6
CULT_1+6									×	×	CULT_1+5/6
CULT_2+4									×	×	CULT_2+4/5
CULT_2+5									×	×	CULT_2+4/5
CULT_3+3	×	×			×	×			×	×	CULT_3+3
CULT_3+3_F	×	×			×	×					CULT_3+3_F
CULT_3+4									×	×	CULT_3+3
CULT_5+3	×	×			×	×					CULT_5+3
CULT_7+6									×	×	CULT_7+6
CULT_8+2	×	×	×	×	×	×	×	×			CULT_8+2
EDTA	×	×	×	×	×	×	×	×			EDTA
Tempus	×	×	×	×	×	×	×	×	×	×	Tempus
Tempus_+7 days	×	×	×	×	×	×	×	×	×	×	Tempus
Tempus_+14 days									×	×	Tempus

Fig. 1. **(A)** Blood samples were collected from voluntary donors into different collection tubes. Tempus Blood RNA and K_2EDTA tubes were used for direct RNA isolation from whole blood, while Li-heparin tubes were used for short-term lymphocyte culture followed by RNA isolation. Sample storage time prior to culture ranged from 0 to 8 days, and cultivation time ranged from 2 to 6 days. We also tested freezing the cultured lymphocytes prior to RNA isolation. Procedures optimized in the analysis are highlighted in red boxes. **(B)** In total, 100 RNA samples were processed as a part of the protocol optimization (detailed in Supplementary Table 1).
Note: The sample designation for cultured lymphocyte control samples (listed in section B) is defined as "CULT_x+y", where "x" denotes the number of days of storage at 4 °C and "y" denotes the number of days in culture. The underscore character indicates additional tested variables: F = freezing before RNA isolation; +7 days (+14 days) = sample obtained from the same individual at 7- or 14-day intervals.
** Colours in sample groups correspond to those shown in Fig. 2.*

RNA extraction

For RNA isolation from whole blood collected in Tempus Blood RNA Tubes (Thermo Fisher Scientific), the MagCore triXact RNA Kit was used on the MagCore Plus II instrument with a modified protocol #631T as described previously (Černá et al., 2025).

For RNA isolation from whole blood collected in K_2EDTA tubes, RNA was isolated from 1.5 ml of blood using the QIAamp RNA Blood Mini Kit (Qiagen, Venlo, NE) according to the manufacturer's instructions within 3 hours of collection.

RNA isolation from cultured lymphocytes was performed using the QIAamp RNA Blood Mini Kit (Qiagen). Cultured cells were pelleted in the culture flask at 300×*g* for 5 minutes in a precooled centrifuge at 4 °C. Supernatant was removed and the cell pellet was resuspended in 3 ml of ice-cold phosphate-buffered saline (PBS; Thermo Fisher Scientific). Following second centrifugation at 300×*g* for 5 minutes, supernatant was removed, and cells were resuspended in 1 ml of ice-cold PBS. This cellular suspension was used as an input for RNA isolation. The procedure was carried out either strictly according to the manufacturer's instructions or using a modified variant in which the lysed lymphocytes, resuspended in RLT buffer containing 1 % β-mercaptoethanol, were frozen at −20 °C for 7 days (designated as _F samples).

All RNA samples were stored at −80 °C until further use.

NGS library preparation and sequencing

Targeted RNA libraries were prepared using the Kapa RNA HyperPrep Kit (Roche, Base, Switzerland) following a previously described protocol (Kleiblová et al., 2024) with several modifications. Fragmentation time was shortened to 2 minutes at 94 °C, targeting a mean fragment length of 470 b. PCR amplifications were performed in 9 and 11 cycles, respectively. Forty-eight samples were equimolarly pooled and hybridized with custom-designed KAPA HyperChoice probes (CZECANCA v1.22; 226 cancer predisposition/candidate genes, 0.63 Mbp, Roche; Supplementary Table S2) (Soukupová et al., 2018) using the KAPA HyperCapture Reagent kit (Roche) and captured using the KAPA HyperCapture Bead kit (Roche). Final libraries (96 samples per run) were paired-end sequenced on a NextSeq2000 instrument using the NextSeq 1000/2000 P1 Reagents v 3.0 Kit (300 cycles; Illumina, San Diego, CA).

Bioinformatic analysis

CZECANCA panel RNA NGS: fastq files were mapped to the human reference genome (hg19) using the STAR aligner 2.5.2 with default settings (sjdbOverhang = 100) to generate Binary Aligment Map (BAM) files (Dobin et al., 2013). PCR duplicates were removed using Picard tools v1.129. The resulting BAM files were subsequently analysed using regtools (Cotto et al., 2023). Splice junction annotation was performed with the R package

SCANVIS (Agius et al., 2019) in R version 4.2. Splice junctions were included for alternative or aberrant splicing analysis if they were supported by at least three unique reads per sample. BAM files were further processed by GATK 4.3 (McKenna et al., 2010) with SplitNCigarReads tool, followed by variant calling with HaplotypeCaller to generate variant call format (VCF) files.

Gene expression analysis

Gene expression analysis was based on the maximum number of unique wild-type exon-exon junction counts for each sample and gene, as described previously (Kleiblová et al., 2024). The sum of all unique transcripts of targeted genes per sample represented the sequencing input for each sample. The absolute transcript counts were used to normalize gene expression to one million transcripts (TPM). The expression of three monoexonic genes (*CEBPA*, *FANCF*, *PTTG2)* was assessed using the average coverage and was not used for input quantity estimation.

Variability, reproducibility, and statistical analysis

The variability and reproducibility of different pre-analytical steps were based on the results of the gene expression analysis. Hereafter, blood collection, storage, cultivation and RNA isolation steps are collectively referred to as pre-analytical sample handling conditions. Their impact on the expression of genes of interest was evaluated within **sample handling variability** analysis. **Inter-sampling reproducibility** was assessed by comparing gene expression in RNA samples obtained from the same individual at 7- or 14-day intervals under selected sampling and culture conditions (Tempus tubes, CULT_0+3 and CULT_1+5). These parameters were analysed using principal component analysis and Spearman's rank correlation. Principal component analysis (PCA) was performed on normalized gene expression values. Genes with zero expression across samples were excluded. PCA was computed using the prcomp() function from the stats package in R (version 4.4.0) with scaling enabled (scale.=TRUE). The first two principal components (PC1 and PC2) were visualized using ggplot2 (Valero-Mora, 2010).

Correlations of gene expression between samples were estimated using Spearman's rank correlation, since the distribution of all the datasets was non-normal (Shapiro and Anderson-Darling normality tests were used). The results were plotted in R using the ggpubr package (Kassambara, 2025).

Wild-type junction analysis was used to more closely investigate changes in gene expression and transcript isoforms in RNA samples obtained under various pre-analytical sample-handling conditions. Normalized wild-type junction profiles for individual exon-exon junctions of selected genes (including *BRCA1*, *BRCA2*, *PALB2*, *CHEK2* and *ATM*) were manually inspected.

Results

Sample handling variability and inter-sampling reproducibility

The impact of different pre-analytical sample handling conditions on the gene expression profiles was assessed using targeted NGS with the CZECANCA panel. The sequencing input quantity (sum of all unique transcripts per sample) of the 100 samples used in the study optimization phase (Fig. 1B) ranged from 92,081 to 199,139 transcripts (mean 128,515; median 121,777) for K_2EDTA blood samples, and from 90,244 to 191,321 transcripts (mean 134,463; median 129,392) for Tempus blood RNA samples. For cultured samples stored at 4 °C for 0–5 days in Li-heparin tubes prior to lymphocyte culture, the sequencing input quantity ranged from 121,983 to 458,733 unique transcripts (mean 252,518; median 266,853). Samples stored for 7–8 days had a lower input compared to other culture conditions, ranging from 35,924 to 245,130 transcripts (mean 136,969; median 142,979). The input quantity of patient cultured samples ranged from 157,438 to 413,605 transcripts (mean 308,912; median 313,257) and was comparable to other cultured samples (Supplementary Table S3). Overall, considering that the panel targets 226 genes that are often lowly expressed, the sequencing input quantity per sample demonstrated robust enrichment across all tested pre-analytical sample handling conditions. RNA isolated from K_2EDTA and Tempus tubes provided adequate target coverage, which further increased with lymphocyte cultivation but was negatively affected by longer blood storage (at 4 °C in Li-heparin tubes for more than 5 days prior to setting up the lymphocyte culture).

Consequently, normalized gene expressions (in TPM) of targeted genes for each sample (Supplementary Table S4) were used for all analyses within the optimization phase, including principal component analysis (PCA) based on expression profiles of the 226 genes of interest. PCA separated whole-blood RNA (EDTA/Tempus) samples from RNA samples derived from cultured lymphocytes along PC1 (Fig. 2). There was no difference between whole-blood samples based on the collection/isolation method (K_2EDTA or Tempus tubes). Patient samples (all processed using lymphocyte culture) clustered together with cultured lymphocyte controls, which nicely reflects identical sample processing and supports comparability of expression and splicing analyses within the cultured lymphocyte workflow. Similarly, the analysis showed that freezing lymphocytes for 7 days at −20 °C prior to RNA isolation had no effect on gene expression. However, samples stored at 4 °C for 7 or 8 days prior to culture (CULT_7+6 and CULT_8+2) differed in their gene expression characteristics from other samples and were not used in further detailed gene expression analyses.

Overall, PCA indicated that RNA sources were the primary determinant of global expression variability, whereas storage duration and cultivation conditions had only minor secondary effects.

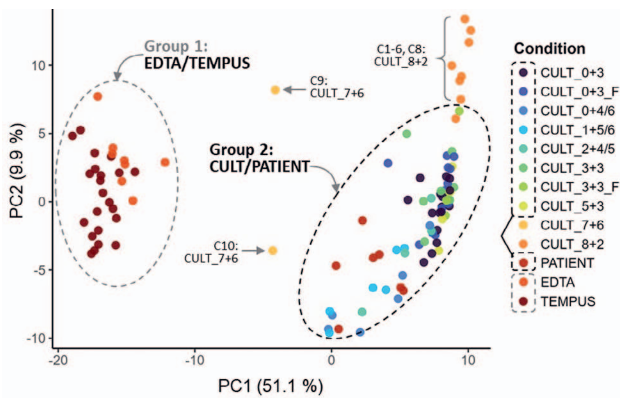


Fig. 2. **Principal component analysis** of all RNA samples separated whole-blood samples (Group 1) and all cultured samples (Group 2). The majority of cultured samples clustered to each other, with the exception of that stored 7 or 8 days at 4 °C prior to lymphocyte culture.

As a next step, we further evaluated the effect of blood collection, storage duration, and cultivation timing on gene expression using normalized expression values and correlation analyses (Fig. 3). The expression of CZECANCA panel-targeted genes was not affected by whether blood was sampled into K_2EDTA or Tempus tubes (Fig. 3A). Similarly, the sample storage time (4 °C for 0–5 days after collection) prior to lymphocytic culture did not affect gene expression (Fig. 3B), just as gene expression correlated in samples with/without freezing of cultured lymphocytes for 7 days at −20 °C (r = 0.98; Supplementary Table S4). In contrast, samples stored for 7–8 days prior to cultivation no longer showed sufficient correlation (r = 0.77 and r = 0.84, respectively) with the reference sample (CULT_0+3); however, they remained suitable for evaluation of mRNA splicing (data not shown). Extending cultivation from 3 to 6 days introduced moderate gene expression variability, reflected by progressively reduced correlation with 3-day cultures (Fig. 3C). Inter-sampling reproducibility was assessed under three selected collection and processing conditions repeated at a 7-day interval in the same voluntary donors. The RNA expression analyses of whole blood (Tempus) and lymphocyte cultures stored/cultivated for 0+3 or 1+5 days showed excellent correlation between these intra-individual samples (Fig. 3 D).

Therefore, the expression of our genes of interest in whole blood is not influenced by the type of collection tube used. Optimization steps have also shown that it is safe to transport and/or store blood samples at 4 °C for up to 5 days prior to lymphocyte culture and that the lymphocytes can be cultured for 3 to 5 days in culture without substantially affecting the gene expression profile. It all provides enough flexibility within diagnostic workflows.

Gene expression and splicing analyses in cultured lymphocytes

We next compared the expression of individual analysed genes between whole blood (Tempus) and cultured lymphocyte RNA samples (Fig. 4A, Supplementary Table S5; samples CULT_7+6 and CULT_8+2 were excluded). Targeted genes were divided into four categories according to their normalized expression levels in analysed RNA samples: high, moderate, low, and very low/no (reflecting their TPM > 5000, > 1500–5000, > 500–1500, and ≤ 500, respectively).

Looking specifically at the 41 genes with very low or no expression in whole-blood RNA, *CHEK1* moved to the high expression category in cultured lymphocytes. The expression of seven genes (*FANCG*, *RAD54B*, *XRCC3*, *RAD51AP1*, *EXO1*, *RAD51*, *RAD54L*) increased to the moderate category range, while the expression of another five additional genes (*FANCE*, *RAD51D*, *RECQL4*, *XRCC2, FANCC*) increased to the low expression category (Fig. 4A, Supplementary Table S5). The expression of the remaining 28 genes did not increase substantially and remained in the very low/no expression category, with eight of them not expressed at all (*PLA2G2A*, *DMBT1*, *KCNJ5*, *LMO1*, *WT1*, *HNF1A*, *HOXB13*, *ALK*, *PHOX2B*, *EGFR*), regardless of cultivation.

Overall, in comparison to Genotype-Tissue Expression Portal data (GTEx; assessed: 2025-10-06; https://gtexportal.org/), cultured lymphocytes resembled LCLs in the expression levels of evaluated cancer predisposition genes. Importantly, targeted RNA panel NGS provides ~ 100-fold higher coverage of unique transcripts and ~ 1000-fold higher yield when normalized to one million transcripts in a standard setup of 96 samples per NextSeq run in comparison to LCLs analysed by whole-transcriptome sequencing GTEx data (Fig. 4B, Supplementary Table S5).

Therefore, our analysis can also provide a more detailed view on individual splicing events at the level of wild-type exon-exon junctions. It can also identify other non-canonical splicing events. Wild-type junction analysis was used to assess whether altered expression under short-term lymphocyte cultivation affects all transcripts uniformly.

The counts of unique individual wild-type exon-exon junctions (Supplementary Table S6) and their normalized counts (Supplementary Table S7) across RNA samples enabled calculation of median normalized wild-type exon-exon junction coverage for individual RNA sample categories (EDTA, Tempus, cultured lymphocytes under various optimization conditions, and patient-cultured samples with and without NMDi; Supplementary Table S8).

Wild-type junction analysis of selected hereditary breast and ovarian (HBOC) predisposition genes showed that lymphocyte cultivation resulted in an approximately 10-fold increase in expression of *BRCA1*, *BRCA2* and *RAD51C* and a modest increase in expression of *PALB2*, *CHEK2* and *RAD51D*. Although the expression of *ATM*

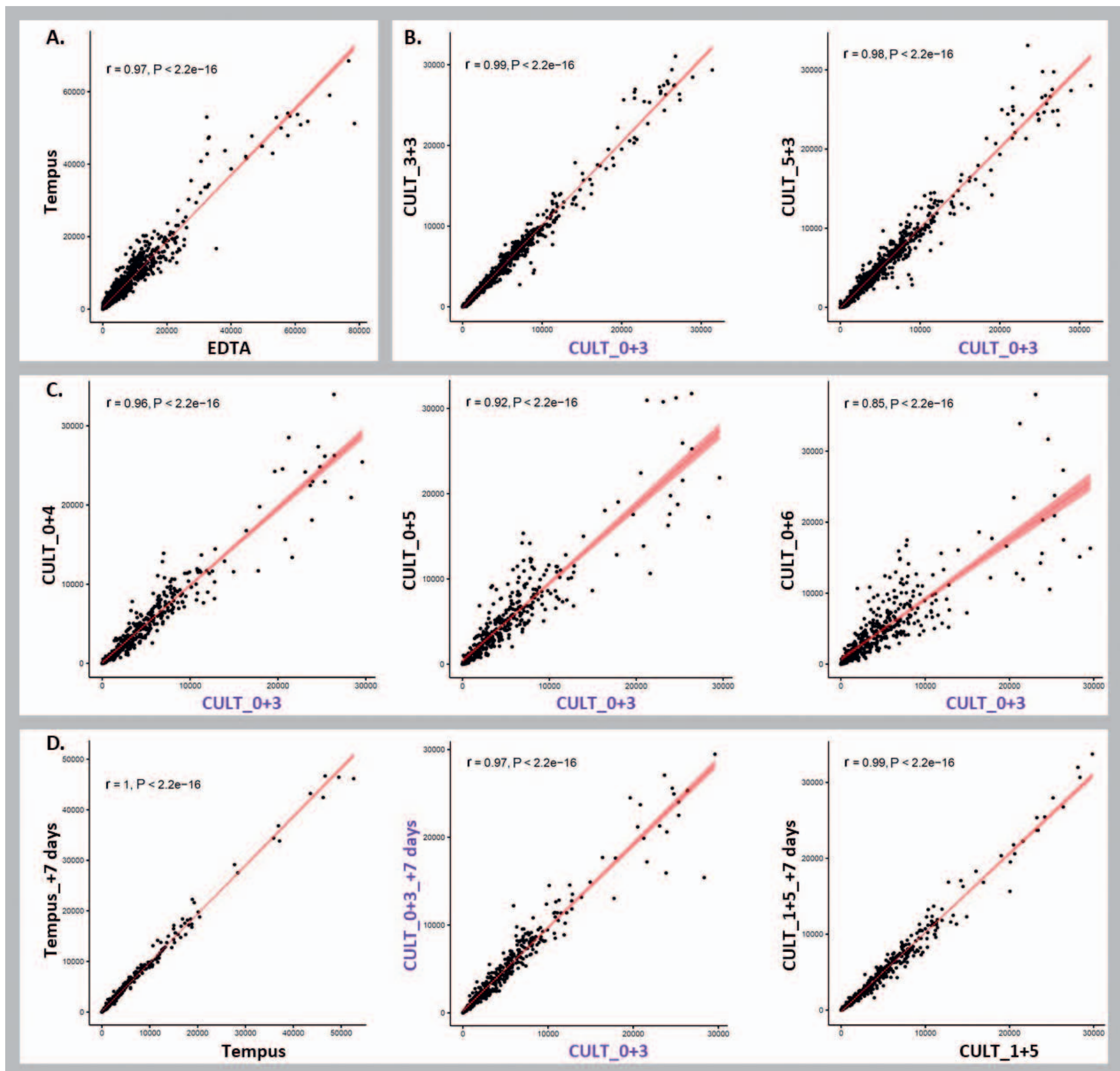


Fig. 3. **Sample handling variability** of RNA samples was compared in terms of gene expression, depending on the collection tubes **(A)**, storage length prior to cultivation **(B)** and culture length **(C)**, where samples CULT_0+3 served as a reference. **Inter-sampling reproducibility** in RNA samples at a 7-day interval from Tempus whole blood and from cultured lymphocytes (CULT_3+3 or CULT_1+5; **(D)**).
CULT_x+y," where "x" denotes the number of days of storage at 4 °C and "y" denotes the number of days in culture. The underscore character indicates additional tested variables: +7 days = sample obtained from the same voluntary donors at a 7-day interval.

was slightly reduced, it was still sufficient for splicing evaluation (Fig. 5). For all analysed genes, the coverage profiles of individual wild-type junctions were comparable.

Analysis of FANCA expression and spliceogenic variants

Compared to whole blood, the expression of genes in the Fanconi anaemia (FA) core complex increased markedly during lymphocyte cultivation (Fig. 5). Notably, *FANCA* showed a pronounced isoform shift: whereas the full-length (canonical) transcript (exons 1–43) predominated in cultured lymphocytes, whole blood (irrespective of K_2EDTA or Tempus collection) was dominated by shorter transcripts comprising exons 28–43 (Fig. 5; *FANCA*).

According to GTEx (https://gtexportal.org/home/gene/FANCA#gene-transcript-browser-block [Accessed 22.08.2025], *FANCA* expression across tissues does not exceed 5 TPM, with the exception of the testis, cultured

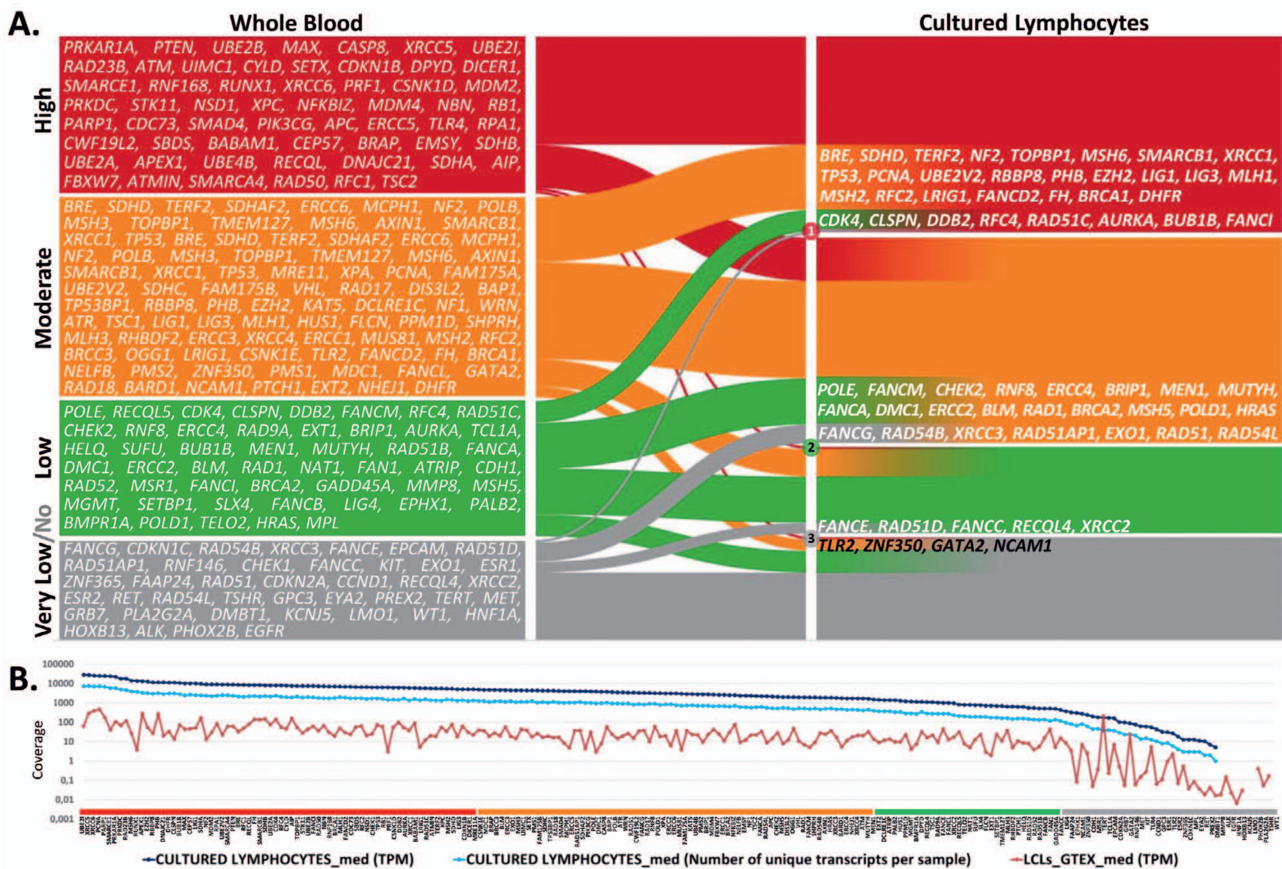


Fig. 4. **(A)** Sankey diagram showing the expression of 226 genes targeted by the CZECANCA panel in peripheral blood (left) and cultured lymphocytes (right). mRNA expression categories are high (red), moderate (orange), low (green) and very low/no (grey). All targeted genes are shown in blood RNA (left column). Only genes with increased expression category (white letters) or decreased across two categories (black letters) expression in cultured lymphocyte RNA (compared to blood) are shown in the right column; genes are listed according to their expression in whole-blood RNA. **(B)** CZECANCA panel capture-based expression data (medians) in absolute (light blue) and normalized (dark blue) unique transcript counts compared to GTEx median expression in blood RNA (red), predominantly generated by RNA-seq; genes are listed according to their expression in cultured lymphocyte RNA (Supplementary Table S5).
Note: Increase in CHEK1 (1) expression; decrease in expression of DPYD (2) and TLR4 (3).

fibroblasts and EBV-transformed lymphocytes. The canonical *FANCA* transcript (ENST00000389301.8; NM_000135.4; exons 1–43) predominates in the majority of tissues. The most highly expressed *FANCA* transcripts in whole blood are two shorter non-coding isoforms, ENST00000562424.1 (exons 40–43) and ENST00000305699.15 (exons 28–39), and one protein-coding isoform ENST00000564475.1 (exons 37–42; Fig. 6A, Supplementary Table S9); however, these isoforms cannot be directly distinguished using short-read sequencing. In whole blood, the canonical mRNA isoform is expressed at a very low level, which largely limits downstream analyses of potentially spliceogenic variants residing in the 5' part of the transcript. In this regard, cultured lymphocytes are a more informative source of RNA, expressing the canonical full FANCA isoform as a dominant transcript. Therefore, cultured lymphocytes could be used to analyse potentially spliceogenic variants located in exons and introns 1–27, including NMD inhibition if necessary. We documented the suitability of our approach by example analysis of *FANCA* germline variant of uncertain significance, c.2602-3C>G, identified in heterozygous status in patient 5 (P5). The variant causes deletion of six bases in a final mRNA transcript ΔE28p(-6) (r.2602_2607del) resulting in two amino acids deletion (p.Gln869_Phe870del). Aberrant splicing was also detectable in RNA from whole blood (Tempus), but the coverage of the evaluated region was insufficient (Fig. 6B).

Discussion

RNA-based NGS analyses are increasingly incorporated into germline cancer predisposition diagnostics primarily because they facilitate interpretation of sequence variants by supporting assessment of splicing and other RNA-level effects. However, the limited expression of genes of interest in routinely analysed tissue is a persistent technical limitation of bulk RNA-seq (Rowlands

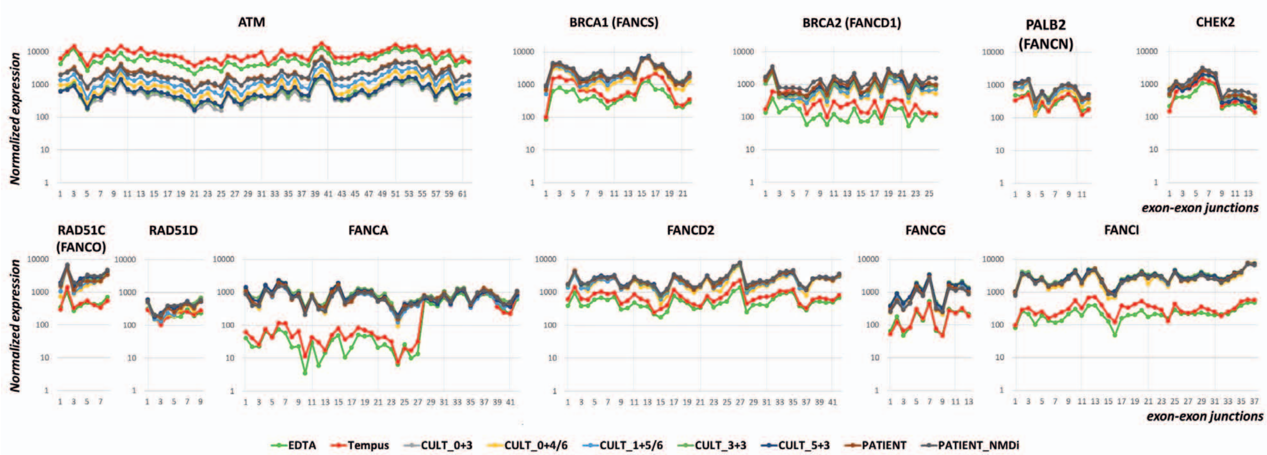


Fig. 5. **Normalized expression of wild-type exon-exon junction** of selected HBOC and FA predisposition genes in RNA from whole blood collected in K_2EDTA or Tempus tubes and under various lymphocyte cultivation conditions (Supplementary Table S1).
X-axis – individual exon-exon junctions (1 means a junction between exons 1 and 2); Y-axis – number of unique exon-exon junctions normalized to 1 million transcripts

et al., 2022). For example, the expression of *BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *PALB2* and *CHEK2* – the most frequently altered genes in individuals with hereditary breast and ovarian cancer syndrome – does not exceed 1.5 TPM in whole-blood RNA, according to the GTEx Portal.

Capture-based RNA NGS analyses can partially overcome this limitation by enriching transcripts of interest (Davy et al., 2017; Karam et al., 2020; Horton et al., 2022, 2024; Kleiblová et al., 2024).

Our data show that there are no significant differences in expression of the 226 capture-targeted genes between whole blood isolated from K_2EDTA and Tempus tubes. Therefore, K_2EDTA represents a cost-effective alternative for RNA-based analyses when RNA isolation can be performed within three hours after collection. As a next step, we focused on establishing a simple protocol for short-term lymphocyte culture to obtain RNA suitable for NGS-based analysis of cancer predisposition genes with minimal invasiveness, including the option of NMDi.

Recently, De Cock and colleagues published a minimally invasive RNA-seq protocol using short-term cultured peripheral blood mononuclear cells isolated from K_2EDTA blood tubes with the density gradient medium Lymphoprep applied to diagnostics of neurodevelopmental disorders (De Cock et al., 2025). In contrast, our approach represents an even simpler protocol that does not require density gradient centrifugation. The protocol is adapted from a standard lymphocyte culture method routinely used for karyotyping (Claussen et al., 2002).

Particular attention was given to evaluation of the preanalytical phase, including lymphocyte stability prior to cultivation, enabling transport of cooled blood samples for culture from collaborating laboratories worldwide. We demonstrated that storage of blood collected in Li-heparin tubes for up to five days at 4 °C does not affect gene expression profiles of consequently cultured lymphocytes. In addition, extending the standard cultivation period from three to five days did not compromise the clinical yield of subsequent RNA analyses, providing greater flexibility for routine workflows.

Short-term lymphocyte culture, very similarly to LCLs, resulted in increased expression of many targeted genes that are reported by GTEx as highly expressed in LCLs in comparison to whole blood (Supplementary Table S5), including frequently altered breast ovarian cancer predisposition genes *BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *PALB2* and *CHEK2* (Kleibl and Kristensen, 2016) or candidate predisposition genes from the FA family (Soukupová et al., 2024). Additionally, short-term lymphocyte culture might not only increase the expression level of genes of interest but might provide a more favourable profile of the detectable transcript variant. As an example, short-term lymphocyte culture did not substantially change overall *FANCA* expression levels, but it promoted expression of the full-length isoform in comparison to shorter ones detected in non-cultured lymphocytes (which contained only the C-terminal region encoded by exons 28–43).

FANCA encodes a key protein of the Fanconi anaemia (FA) pathway, which is essential for the repair of DNA inter-strand crosslinks and maintenance of genomic stability (Hodson and Walden, 2012; Ceccaldi et al., 2016; Yang et al., 2019). Biallelic germline pathogenic variants in *FANCA* cause FA complementation group A, a rare autosomal recessive disorder characterized by bone marrow failure, congenital abnormalities and elevated cancer risk (Kutler et al., 2003; Auerbach, 2009). *FANCA* is the most frequently mutated FA gene, accounting for two-thirds of FA cases (Kimble et al., 2018). Thus, lymphocyte

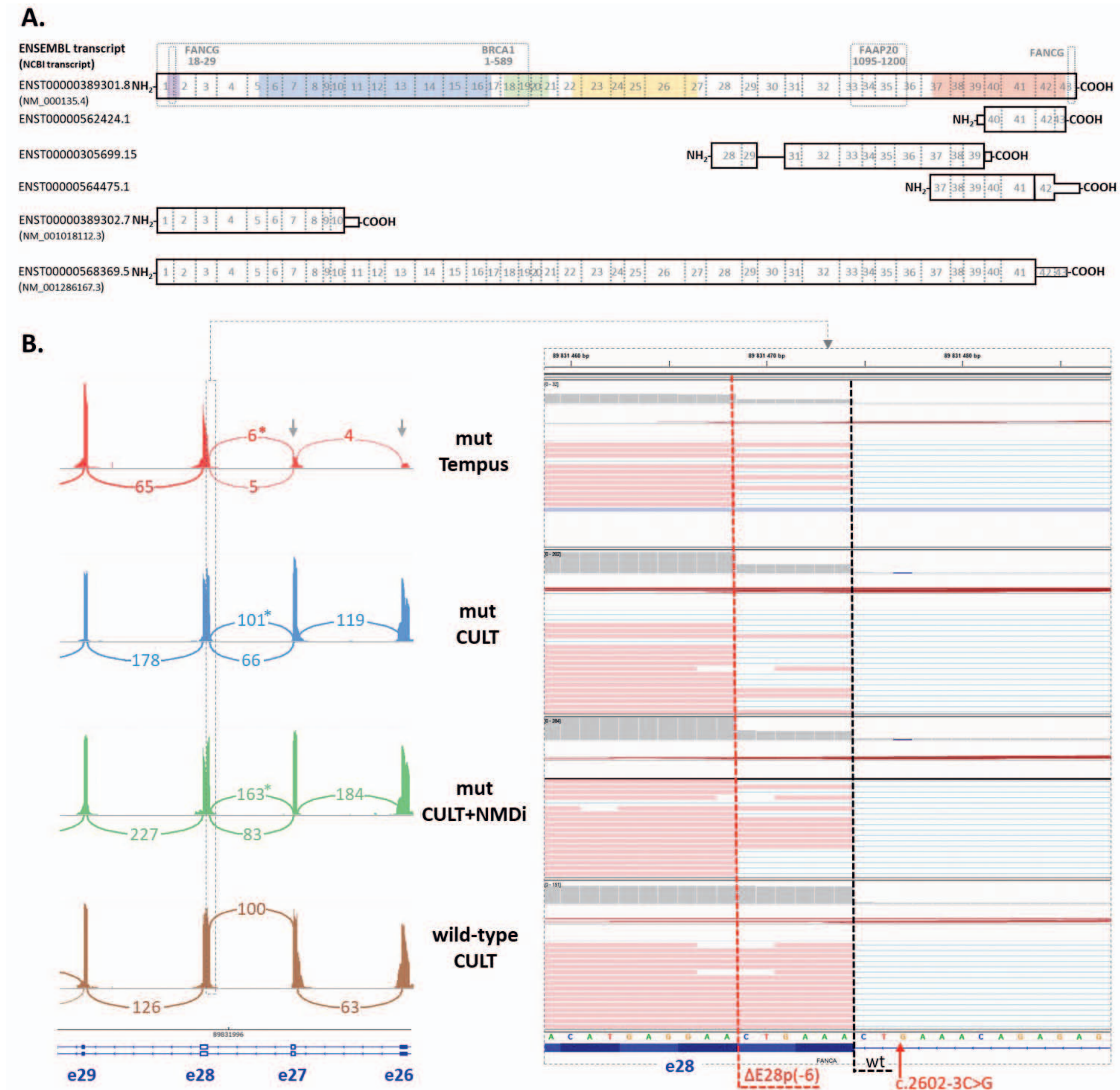


Fig. 6. **(A)** Schematic representation of **FANCA transcript isoforms and protein domains**. Isoform ENST00000389301.8 (canonical transcript) is shown with known structural and functional regions: bipartite nuclear localization signal (NLS; amino acids (aa) 18–34; purple); N-terminal region (aa 170–524; blue); helical domain (aa 544–625; green); arcN subdomain (aa 648–880; yellow); and C terminal domain (aa 1225–1431; red). Binding sites for FANCG (18–29 aa, C-terminus), BRCA1 (1–598 aa) and FAAP (aa1095–1200) are indicated. Information compiled from Shahid and Firasat (2019); GTExPortal, 2025; UniProt, 2025, https://www.uniprot.org/uniprotkb/O15360/entry [Accessed 22.08.2025]. **(B) IGV visualization of splicing impact of *FANCA* heterozygous germline variant c.2602-3C>G** localized in a region poorly expressed (grey arrows) in whole blood (mut Tempus) and cultured samples without (mut CULT) and with NMD inhibition (mut CULT+NMDi) and in a cultured control sample (wild-type CULT). The numbers indicate the number of unique exon-exon junctions, *indicates an aberrant junction.

culture has proved to be a useful tool for investigating the actual splicing impact of potentially spliceogenic variants identified in the exon 1–28 *FANCA*. We demonstrated the benefits of our approach using a germline *FANCA* variant c.2602-3C>G as an example. This variant causes a splice acceptor shift in exon 28, leading to in-frame deletion of six nucleotides in the mature mRNA transcript. It is located within the region of the transcript that is underrepresented in whole-blood RNA. Using Splice AI (data not shown), the variant was predicted to weaken the acceptor splice site and to generate aberrant transcripts ΔE28**,** ΔE28p(-6)**,** ΔE28p(-20) and

ΔE28p(-103). However, only ΔE28p(-6) resulting in p.Gln869_Phe870del was massively present in cultured lymphocytes in > 50 % of all exon-exon junctions in the region, regardless of the presence/absence of NMDi. Other predicted transcripts were minimally present in cultured lymphocytes, but not in Tempus whole-blood RNA, accounting together for less than 5 % of all transcripts.

Our approach also enabled us to address certain limitations of RNA analysis from whole blood, such as low expression of genes of interest and degradation of aberrant transcripts by NMD in carriers of variants causing premature termination of translation.

In conclusion, we demonstrated that combining short-term lymphocyte culture using a simple modified protocol originally developed for cytogenetic karyotype with capture-based NGS analysis of low-expression genes further improves the diagnostic yield of cancer predisposition RNA analyses.

Acknowledgements

We thank Pavel Just, Pavel Pešek and Olena Yakovchuk for excellent collaboration and technical support.

Conflict of interest

The authors declare no potential conflict of interest. They declare that the results summarized in this manuscript have not been published previously and have not been submitted for consideration to any other journal.

Author contribution

MC: Methodology; Investigation; Formal analysis and data curation; Writing – Original Draft. KM: Methodology; Formal analysis and data curation; Bioinformatics; Statistics; Writing – Editing. TP: Methodology; Investigation; Formal analysis and data curation; Bioinformatics; Statistics; Writing – Editing. PN: Formal analysis and data curation; Bioinformatics. RM: Methodology; Writing – Editing. KV: Resources, Writing – Editing. MJ: Writing – Editing. JS: Writing – Editing. PK: Conceptualization and methodology; Investigation; Formal analysis and data curation; Writing – Original Draft and Editing; Resources; Funding acquisition.

References

Agius, P., Geiger, H., Robine, N. (2019) SCANVIS: a tool for SCoring, ANnotating and VISualizing splice junctions. *Bioinformatics* **35**, 4843-4845.

Auerbach, A. D. (2009) Fanconi anemia and its diagnosis. *Mutat. Res.* **668**, 4-10.

Ceccaldi, R., Sarangi, P., D'Andrea, A. D. (2016) The Fanconi anaemia pathway: new players and new functions. *Nat. Rev. Mol. Cell Biol.* **17**, 337-349.

Černá, M., Šťastná, B., Pešek, P. et al. (2025) Semi-automated RNA isolation from Tempus Blood RNA Tubes using the MagCore plus II instrument. *Folia Biol. (Praha)* **71**, 88-94.

Claussen, U., Michel, S., Muhlig, P. et al. (2002) Demystifying chromosome preparation and the implications for the concept of chromosome condensation during mitosis. *Cytogenet. Genome Res.* **98**, 136-146.

Cotto, K. C., Feng, Y. Y., Ramu, A. et al. (2023) Integrated analysis of genomic and transcriptomic data for the discovery of splice-associated variants in cancer. *Nat. Commun.* **14**, 1589.

Davy, G., Rousselin, A., Goardon, N. et al. (2017) Detecting splicing patterns in genes involved in hereditary breast and ovarian cancer. *Eur. J. Hum. Genet.* **25**, 1147-1154.

De Cock, L., D'Haenens, E., Vantomme, L. et al. (2025) Cracking rare disorders: a new minimally invasive RNA-seq protocol. *NPJ Genom. Med.* **10**, 45.

Dobin, A., Davis, C. A., Schlesinger, F. et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Hodson, C., Walden, H. (2012) Towards a molecular understanding of the Fanconi anemia core complex. *Anemia* **2012**, 926787.

Hong, M., Tao, S., Zhang, L. et al. (2020) RNA sequencing: new technologies and applications in cancer research. *J. Hematol. Oncol.* **13**, 166.

Horton, C., Cass, A., Conner, B. R. et al. (2022) Mutational and splicing landscape in a cohort of 43,000 patients tested for hereditary cancer. *NPJ Genom. Med.* **7**, 49.

Horton, C., Hoang, L., Zimmermann, H. et al. (2024) Diagnostic outcomes of concurrent DNA and RNA sequencing in individuals undergoing hereditary cancer testing. *JAMA Oncol.* **10**, 212-219.

Ibraimov, A. I. (1983) Chromosome preparations of human whole blood lymphocytes: an improved technique. *Clin. Genet.* **24**, 240-242.

Karam, R., LaDuca, H., Richardson, M. E. et al. (2020) RNA-seq analysis is a useful tool in variant classification. *JCO Precis. Oncol.* **4**, 1226-1227.

Kimble, D. C., Lach, F. P., Gregg, S. Q. et al. (2018) A comprehensive approach to identification of pathogenic FANCA variants in Fanconi anemia patients and their families. *Hum. Mutat.* **39**, 237-254.

Kleibl, Z., Kristensen, V. N. (2016) Women at high risk of breast cancer: molecular characteristics, clinical presentation and management. *Breast* **28**, 136-144.

Kleiblová, P., Černá, M., Zemánková, P. et al. (2024) Parallel DNA/RNA NGS using an identical target enrichment panel in the analysis of hereditary cancer predisposition. *Folia Biol. (Praha)* **70**, 62-73.

Kutler, D. I., Singh, B., Satagopan, J. et al. (2003) A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood* **101**, 1249-1256.

Kuzbari, Z., Bandlamudi, C., Loveday, C. et al. (2023) Germline-focused analysis of tumour-detected variants in 49,264 cancer patients: ESMO Precision Medicine Working Group recommendations. *Ann. Oncol.* **34**, 215-227.

LaDuca, H., Polley, E. C., Yussuf, A. et al. (2020) A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. *Genet. Med.* **22**, 407-415.

McKenna, A., Hanna, M., Banks, E. et al. (2010) The genome analysis toolkit: a MapReduce framework for analyzing

next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303.

Mesman, R. L. S., Calleja, F., de la Hoya, M. et al. (2020) Alternative mRNA splicing can attenuate the pathogenicity of presumed loss-of-function variants in BRCA2. *Genet. Med.* **22**, 1355-1365.

Rowlands, C. F., Taylor, A., Rice, G. et al. (2022) MRSD: a quantitative approach for assessing suitability of RNA-seq in the investigation of mis-splicing in Mendelian disease. *Am. J. Hum. Genet.* **109**, 210-222.

Shahid, M., Firasat, S. (2019) FANCA and contribution of studies from Asian populations to the understanding of fanca mediated Fanconi anemia. *Genetika* **51**, 1197-1225.

Soukupová, J., Šťastna, B., Kanwal, M. et al. (2024) A comprehensive study evaluating germline FANCG variants in predisposition to breast and ovarian cancer. *Cancer Med.* **13**, e70103.

Soukupová, J., Zemánková, P., Lhotová, K. et al. (2018) Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes. *PLoS One* **13**, e0195761.

Stadler, Z. K., Maio, A., Chakravarty, D. et al. (2021) Therapeutic implications of germline testing in patients with advanced cancers. *J. Clin. Oncol.* **39**, 2698-2709.

Valero-Mora, P. M. (2010) ggplot2:Elegant Graphics for Data Analysis. *Journal of Statistical Software* **35 (Book Review 1)**, 1-3.

Walker, L. C., Lattimore, V. L., Kvist, A. et al. (2019) Comprehensive assessment of BARD1 messenger ribonucleic acid splicing with implications for variant classification. *Front. Genet.* **10**, 1139.

Wen, L., Tang, F. (2025) Single-cell omics sequencing technologies: the long-read generation. *Trends Genet*. **42**, 46-62.

Yang, X., Zhang, X., Jiao, J. et al. (2019) Rare variants in FANCA induce premature ovarian insufficiency. *Hum. Genet.* **138**, 1227-1236.