

Simultaneous Study of Circular RNAs and Messenger RNAs in Colorectal Cancer: The Unbalanced Fate of a Couple?

Corentin Levacher , [Joanna Delfosse](#) , Camille Charbonnier , Françoise Charbonnier , [Mathieu Viennot](#) , Edwige Kasper , Jacques Mauillon , Nathalie Parodi , Stéphanie Baert-Desurmont , Philippe Ruminy , [Claude Houdayer](#) *

Posted Date: 1 January 2026

doi: 10.20944/preprints202512.2762.v1

Keywords: circular RNAs; messenger RNAs; colorectal cancer



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Simultaneous Study of Circular RNAs and Messenger RNAs in Colorectal Cancer: The Unbalanced Fate of a Couple?

Corentin Levacher ¹, Joanna Delfosse ¹, Camille Charbonnier ², Françoise Charbonnier ¹, Mathieu Viennot ³, Edwige Kasper ⁴, Jacques Mauillon ⁵, Nathalie Parodi ⁵, Stéphanie Baert-Desurmont ⁴, Philippe Ruminy ³ and Claude Houdayer ^{4,*}

- ¹ Univ Rouen Normandie, Inserm U1245, Normandie Univ, F-76000 Rouen, France
- ² Univ Rouen Normandie, Inserm U1245, Normandie Univ, CHU Rouen, Department of Biostatistics, F-76000 Rouen, France
- ³ Univ Rouen Normandie, Inserm U1245, Normandie Univ, Centre Henri Becquerel, F-76000 Rouen, France
- ⁴ Univ Rouen Normandie, Inserm U1245, Normandie Univ, CHU Rouen, Department of Genetics, F-76000 Rouen, France
- ⁵ Univ Rouen Normandie, Normandie Univ, CHU Rouen, Department of Genetics, F-76000 Rouen, France
- * Correspondence: claud.houdayer@chu-rouen.fr; Tel.: +33 (0)2.32.88.57.83

Simple Summary

Circular RNAs (circRNAs) can compete with messenger RNAs (mRNAs) from their host gene, which can lead to the downregulation of mRNA. We investigated this potential mechanism of downregulation in colorectal cancer (CRC) predisposition in a large cohort of 712 patients suspected of having hereditary CRC and 249 healthy controls, with a focus on 23 genes associated with CRC. Using the novel SEALigHTS technique (Splice and Expression Analyses by Exon Ligation and High-Throughput Sequencing), we identified 220 circular RNA junctions, including 47 new ones. Patients exhibited a circRNA/mRNA ratio approximately twice that of controls, particularly for the *POLD1* gene, which produces a circRNA with potential cancer-promoting effects. The findings imply that the balance between circRNAs and mRNAs is disturbed in CRC, though not due to competition between the two. They also suggest that specific circRNAs play a role in CRC development, either as a cause or a consequence.

Abstract

Background/Objectives: Circular RNAs (circRNAs) are emerging players in human diseases, with function as part of competing endogenous networks. Given the importance of messenger RNA (mRNA) regulation in human diseases and the potential of circRNAs in this regulation, we studied the circRNA-mRNA couple in blood within a cohort of 712 patients suspected of having hereditary colorectal cancer (CRC) and 249 matched controls. **Methods:** The circRNA-mRNA couple was studied by SEALigHTS (Splice and Expression Analyses by exon Ligation and High-Throughput Sequencing) using a panel of 23 genes involved in CRC predisposition, i.e., 788 probes designed at exon ends, enabling the exploration of all exon-exon junctions. Following reverse transcription and probe hybridization on cDNA, nearby probes are ligated, and the number of ligations quantified using unique molecular identifiers and sequencing. **Results:** We described 220 circular junctions, including 47 novel ones. The circRNA/mRNA ratio was 1.93-fold higher in patients compared to controls ($p < 2 \times 10^{-16}$), irrespective of age of cancer onset. This increase was mainly driven by *POLD1* (fold change 3.11) and a single circPOLD1 with oncogenic potential. **Conclusions:** This study supports the idea of a physiological balance between circRNA and mRNA that can be disrupted under pathological conditions. It rules out a competitive mechanism between circular and linear transcripts in CRC predisposition and raises questions about the role of specific circRNAs in the development of CRC, either as a cause or a consequence.

Keywords: circular RNAs; messenger RNAs; colorectal cancer

1. Introduction

Known for decades [1,2], the messenger RNA (mRNA) is the centerpiece of the transcriptome. This intermediary between DNA and protein is regulated by various mechanisms, including transcription [3–5], splicing [6], stability [7], capping [8], transport [9] and translation [10]. Among the elements involved in this regulation, many types of RNA have been documented. MicroRNAs, but also pseudogene transcripts and long non-coding RNAs, have been shown to play a role in regulating mRNA [11]. These different elements are grouped together in a network known as competing endogenous RNA (ceRNA), and accumulating evidence points to its preponderant role in various pathologies such as cancer [12], neurodegenerative [13], cardiovascular [14] or autoimmune diseases [15]. Recently, with growing interest, circular RNAs (circRNAs) have been included in this ceRNA network [16].

circRNAs are single-stranded, closed-loop covalent structures, devoid of 5' caps nor 3' poly-A tails. They result from a process known as backsplicing, whereby a 5' splice site binds to an upstream 3' splice site [17]. Interest in circRNAs has increased greatly, mainly because the levels of certain circRNAs have proved important for better diagnosis, prognosis, disease monitoring in Alzheimer disease [18], cardiovascular disease [19] or cancer [20]. These observations are often attributed to their functions as microRNA and RNA binding protein sponges [21] but other mechanisms are emerging e.g. cap-independent translation that enables the generation of peptides involved in different tumorigenesis pathways [22,23] or transcription modulation. This last mechanism is particularly interesting, as it has been shown that circRNAs can regulate the levels of various mRNAs, including those of the parental gene [24,25] and a physiological balance between these two forms of transcript has been demonstrated [26,27]. Consequently, it may be hypothesized that a disruption of this physiological balance represents a novel, hidden mechanism in genetic diseases [28–31]. The question is particularly relevant in diseases where known mechanisms explain only a fraction of cases, the rest are referred to as “missing heritability”. Colorectal cancer (CRC) exemplifies this issue as the majority of non-polyposis familial and early onset microsatellite stable (MSS) cases have no identified genetic cause [32]. To investigate this hypothesis, we performed a simultaneous analysis of circRNA and mRNA expression in blood samples for 23 CRC predisposition genes in a cohort of 712 CRC patients (i.e. the “missing heritability” cohort) and 249 matched controls using a novel, dedicated technique named SEALigHTS (Splice and Expression Analyses by exon Ligation and High-Throughput Sequencing) [29,33]. We described the backsplicing landscape for these genes, studied the circRNA-mRNA balance between patients and controls and searched for a competitive mechanism between circRNA and mRNA expression.

2. Materials and Methods

Patients and controls

The collection of unexplained CRC cases was previously described [32] and includes 1029 patients with a personal and/or familial history suggestive of hereditary predisposition to CRC, Lynch syndrome and polyposis excluded (Table 1). All patients were recruited by the French network of Cancer Genetics Departments. In addition, 500 healthy volunteers without any personal or familial CRC history among their first-degree relatives were recruited by the Clinical Investigation Center of Rouen University Hospital. Among them, 712 patients and 500 controls had their blood sampled on PAXgene tubes for RNA extraction. To ensure quality, PAXgene tubes were transported within two days after sampling to the investigating center and extracted within 2 weeks. All of the 712 patients were included in the study. Among 500 healthy controls with available RNA samples, matched by age and sex to patients, 249 were randomly included. Written informed consent was obtained for all (ethics approval: DC-2013-1759).

Inclusion criteria

- CRC in two first-degree relatives, one being diagnosed before 61 years of age
- CRC diagnosed before 51 years of age or advanced colorectal adenoma (diameter over 1 cm, and/or tubulovillous or villous and/or with high-grade dysplasia) before 41 years of age
- Multiple primary CRCs in the same individual, the first one being diagnosed before 61 years of age

Exclusion criteria

- Lynch syndrome, as defined by the presence in the patient of a germline MMR gene pathogenic variant and/or a MSI tumour, in the context of a suggestive presentation (familial history, early age of onset)
- Adenomatous polyposis, as defined by more than 10 histologically proven adenomas
- Hamartomatous polyposis, as defined by the presence of histologically proven hamartomas

Table 1. Description of the criteria used to select patients for inclusion in the collection.

RNA extraction

RNA was extracted from peripheral blood with the PAXgene Blood RNA kit (PreAnalytiX, Switzerland) and concentration was determined with the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, MA, USA), with a mean of 88 ng/μL (min: 4; max: 446).

SEALigHTS

Principle

Initially used as a multiplex technique for fusion transcript detection [34] and for measuring gene expression [35], SEALigHTS has been adapted and validated to study and quantify splicing and backsplicing [29]. Briefly, SEALigHTS allows for the simultaneous exploration of all exon-exon junctions on a panel of genes of interest, thanks to probes designed at exon extremities. Following reverse transcription and probe hybridization on cDNA, nearby probes are ligated if splicing and/or backsplicing occurs and the number of ligations is quantified using unique molecular identifiers and high throughput sequencing. All possible combinations of exons, i.e. splicing and backsplicing, are detectable.

Protocol

Oligonucleotides probes contained (i) specific sequences for each exon (19 – 30 bases length to obtain an optimal melting temperature of 70°C), (ii) Unique Molecular Identifiers (UMI), consisting of 7 random bases, to count the number of ligations (iii) complementary sequences of universal PCR primers. In accordance with the transcripts present in Ensembl and/or GTex Portal and literature, 788 probes (Supplementary Table S1) were designed at exon boundaries for the 23 CRC predisposition genes listed by the French experts from the Groupe Génétique et Cancer (GGC) [36] (Table 2). For exon ends where a single nucleotide polymorphism (SNP) is present with a frequency higher than 1% in the Caucasian population, the corresponding probe was designed with a combination of the 2 nucleotides involved to avoid any hybridization problems. If the alternative and canonical splice sites to be explored are close, partial probe overlap occurs. In this case, and when the sequence overlap was greater than 50%, the alternative probe was not selected. To control for DNA contamination, intronic probes were designed for each gene and added in the probe mix (Supplementary Table S1). Following quantification, 14–500 ng of total RNA were converted into cDNA using a SuperScript™ VILO™ cDNA Synthesis Kit (Invitrogen, Carlsbad, CA, USA). cDNAs were incubated for 1 h at 60 °C with our mix of 788 oligonucleotide probes (Supplementary Table S1) in 1× SALSA MLPA buffer (MRC Holland, Amsterdam, The Netherlands). Following hybridization, neighboring probes were ligated using the thermostable SALSA DNA ligase (MRC Holland, Amsterdam, The Netherlands).

Among the 23 genes studied (Table 2), the gene expression given by GTex Portal is up to 10,000 times higher for some genes than for others. After an initial trial, it appeared that *PTEN* and *RPS20* high expression mandated a separate amplification mix in combination with the Q5® Hot Start High-Fidelity 2X Master Mix (NEB, Ipswich, MA, USA), which is why 2 independent PCRs were performed using different universal PCR primers. PCR products were purified using AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA). The library containing *PTEN* and *RPS20* was diluted 1:20

in the library with the other genes. Sequencing of amplicons was carried out using a NextSeq® system with 75 cycles (Illumina, San Diego, CA, USA).

Sequencing reads, comprising a maximum of 75 bases, including UMIs, left and right probes, and barcodes, were demultiplexed using the barcodes and aligned with probe sequences. Exon junctions were counted for quantification purposes, thanks to the 7 random bases of the UMI, allowing 16 384 different combinations of unique molecules. A minimum of 65% mapping of left and right probes was deemed necessary to avoid primer dimers and ensure sufficient ligation and reliable results. A custom Python script, available on request, generated schematic backsplicing and splicing profiles (Supplementary Figure S1).

Bioinformatics and statistics analyses

Ratio circRNAs/mRNAs

Linear mRNAs (splicing) were distinguished from circRNAs (backsplicing) thanks to the order of the probes. If a probe located at the 3' boundary of an exon is ligated with a probe located at the 5' boundary of the following exon, the transcript is linear. Conversely, if this 3' probe is ligated to a 5' probe of a preceding exon, the transcript is circular (Supplementary Figure S1). Each circular RNA is a unique molecule, so all circular junctions must be considered, unlike an mRNA molecule made up of successive linear junctions. Hence, for each gene, the ratio between circRNAs and mRNAs was obtained by dividing the total number of UMIs counts of all circular junctions by the median number of UMIs counts of all canonical linear junctions of that gene i.e. $\frac{UMI \text{ all circular junctions}}{\text{median UMI all canonical linear junctions}} \times 100$. Importantly, some of the UMIs attributed to linear junctions may actually stem from circRNAs. This occurs when a circRNA contains multiple exons, allowing probes to hybridize to internal exons. Therefore, the signal may be misinterpreted as linear junctions, impacting mRNA expression calculation. To mitigate this, only linear junction UMIs with no potential influence from circRNAs expressed at more than 10% of the corresponding mRNA were counted for mRNA calculation. A Student t-test was used to assess differences between patient and control groups, and p-values were corrected with a Bonferroni method to guarantee an overall control at 5% of type-I error rate.

mRNA and circRNAs expression

For each sample and junction, the UMI counts for each junction were normalized according to the library size i.e. $\frac{UMI \text{ junction}}{UMI \text{ all junctions}} = UMI \text{ normalized junction}$. Hence, for each gene and sample, mRNA expression was calculated as the mean of normalized "circular-free" canonical linear junctions. In contrast, as circRNAs represented distinct transcripts, expression per gene was obtained by summing the UMIs of normalized circular junctions.

Circular RNAs characterization

To characterize the full circRNA sequences, divergent primers were designed (Supplementary Figure S2A and B). From 500ng of RNA, cDNA was generated with the SuperScript™ VILO™ cDNA Synthesis kit (Invitrogen, Carlsbad, CA, USA), followed by a PCR step with the Q5® Hot Start High-Fidelity 2X Master Mix kit (NEB, Ipswich, MA, USA). Gel-purified PCR products were sequenced using PCR primers and the BigDye v3 kit (Applied Biosystems, South San Francisco, CA, USA) and analyzed by capillary electrophoresis on a "3500 Genetic Analyzer" (Applied Biosystems, South San Francisco, CA, USA).

3. Results

3.1. General considerations

Of the 961 samples analyzed, 657 (68.4%) i.e. 433 patients and 224 controls, met the quality criteria (65% mapping) with a mean UMI count of 1,183,735 ensuring homogeneous samples and results, as shown by principal component analysis (PCA) (Supplementary Figure S3).

GREM1 and *EPCAM* genes are poorly expressed in whole blood. As the number of UMI is gene expression-dependent we were not able to detect all the canonical junctions and therefore *GREM1* and *EPCAM* were excluded from further analyses.

For the 21 genes analyzed, 559 linear junctions i.e. canonical and alternatives were identified (Supplementary Table S1), with diversity in their numbers between genes (Figure 1 and Table 2). Although the number of linear junctions depends on the number of exons, the number of expressed alternative junctions differed between genes, e.g. *APC* and *MSH2* both have 16 exons but 43 and 25 total linear junctions were expressed in blood, respectively. As the probes hybridize on both *PMS2* and its pseudogene *PMS2CL*, their respective mRNAs and circRNAs cannot be distinguished. Therefore, *PMS2* was not considered in the analyses.

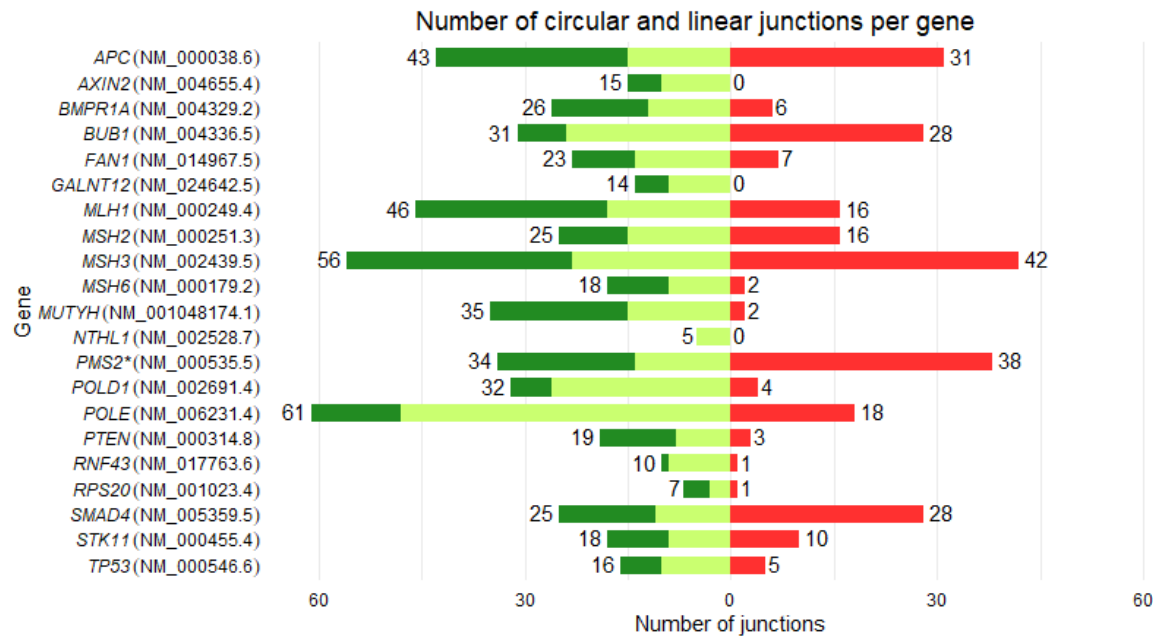


Figure 1. Number of circular and linear junctions per gene. Canonical (NM references) and alternative junctions are indicated, respectively, in light green and green. Circular junctions are indicated in red. The number of linear and circular junctions detected is shown in front of the bars. *PMS2* is indicated with an asterisk due to the probable detection of pseudogene junctions.

3.2. Backsplicing landscape

Of the 20 genes studied (*PMS2* excluded), SEALigHTS detected 220 circular junctions (Supplementary Table S1) of which 47 (21.36%) were novel (Table 2), not reported in RJunBase [37], a database compiling data on linear junctions, circular junctions, and fusion transcripts. A wide diversity in the number and level of expression of circRNAs (Table 2 and Figure 1) was found. Three genes (*AXIN2*, *GALNT12* and *NTHL1*) did not show circular junctions. For the remaining 17, the median number of circular junctions per gene was 7 and ranged from 1 to 42 for *RPS20* and *MSH3*, respectively. The number of circular junctions was not linked to the number of linear junctions e.g. *MUTYH* had 35 linear junctions and only 2 circular junctions, unlike *SMAD4* which has 25 and 28 respectively (Table 2 and Figure 1). While circRNAs are expressed at low levels for most of the genes (17/20), *POLD1*, *BUB1* and *SMAD4* had higher or similar levels of circRNA compared to mRNA molecules. The relative expression of circRNA to mRNA varied between genes and was not related to the number of circRNAs produced by the gene. For example, *POLD1* had the highest circRNA/mRNA expression level 547.28%, i.e. an relative abundance of 547 circRNAs for 100 mRNA molecules, with only 4 circRNAs produced, while *APC* had a ratio of 9.65% with 31 circRNAs produced (Table 2).

Table 2. Splicing and backsplicing summary for each gene. The number of exons is indicated according to the reference transcript. The number of linear junctions include canonical and alternative ones. The number of new circular RNAs is indicated with RJunBase as reference. *PMS2* is indicated with an asterisk due to pseudogene

issues (see text). *GREM1* and *EPCAM* genes are not included in the table, as they were excluded from the analysis due to lack of expression. Genes are listed by the ratio of circular to linear UMI.

Gene	Reference transcript	Number of exons	Number of linear junctions	Number of circular junctions	Number of new circular RNAs	Circular / linear junctions (%)	Circular / linear UMIs (%)
POLD1	NM_002691.4	27	32	4	3	12,50	547,28
BUB1	NM_004336.5	25	31	28	8	90,32	426,02
PMS2 *	NM_000535.7	15	34	38	17	111,76	134,81
SMAD4	NM_005359.6	12	25	28	6	112,00	96,61
POLE	NM_006231.4	49	61	18	6	29,51	40,91
MSH3	NM_002439.5	24	56	42	1	75,00	31,02
RNF43	NM_017763.6	10	10	1	0	10,00	18,37
FAN1	NM_014967.5	15	23	7	0	30,43	15,58
MLH1	NM_000249.4	19	46	16	6	34,78	12,58
MSH2	NM_000251.3	16	25	16	2	64,00	11,38
APC	NM_000038.5	16	43	31	3	72,09	9,65
MUTY H	NM_00104817.4.2	16	35	2	1	5,71	2,93
STK11	NM_000455.5	10	18	10	4	55,56	2,31
BMPRI1 A	NM_004329.3	13	26	6	2	23,08	1,74
MSH6	NM_000179.3	10	18	2	2	11,11	1,68
TP53	NM_000546.6	11	16	5	2	31,25	1,54
PTEN	NM_000314.8	9	19	3	0	15,79	0,17
RPS20	NM_001023.4	4	7	1	1	14,29	0,03
AXIN2	NM_004655.4	11	15	0	0	0,00	0,00
GALN T12	NM_024642.5	10	14	0	0	0,00	0,00
NTHL1	NM_002528.7	6	5	0	0	0,00	0,00

Following circRNA detection, we wanted to describe the entire sequence of the most frequent circRNAs i.e. *POLD1_circRNA_3-2*, *BUB1_circRNA_19-10*, *BUB1_circRNA_9-2*, *BUB1_circRNA_9-6*, *SMAD4_circRNA_8-6*, *SMAD4_circRNA_10-5*, *SMAD4_circRNA_8-5* and found alternative backsplicing events in 2 cases. Sequencing the circRNA joining exons 3 to 2 of *POLD1* (*POLD1_circRNA_3-2*) characterized a *POLD1_circRNA_3-2p* (Supplementary Figure S2C), i.e., with the use of an alternative splice site in exon 2, reported in RJunBase for both linear and circular transcripts. For the circRNA joining exons 10 to 5 of *SMAD4* (*SMAD4_circRNA_10-5*), sequence analysis showed a full-length backsplice transcript and an alternative one with exons 6 and 7 skipping (Supplementary Figure S2D). As these alternative splicing events are also present in mRNAs, this suggests the use of the same (back)splicing machinery to produce messenger and circular RNAs to increase circRNAs diversity.

3.3. Ratio circRNA/mRNA between CRC patients and controls

The ratio was calculated for the 17 genes harboring circular junctions. We observed a significant 1.93-fold increase in the mean circRNA/mRNA ratio in patients compared to controls (433 vs 224; $p<0.001$), (Figure 2A). This result was confirmed independently on *POLD1*, *BUB1* and *SMAD4*, which exhibited the highest circRNA/mRNA ratios. This increase was predominantly driven by *POLD1* (Figure 2B), with a fold change of 3.11, rather than *BUB1* and *SMAD4* (fold changes 1.36 and 1.33, respectively) (Figure 2C and D). We identified 31 outlier patients i.e. with a circRNA/mRNA ratio above mean ± 2 DS. Among these patients, 4 exhibited outlier ratios for all three genes (*POLD1*, *BUB1*, and *SMAD4*), 14 and 13 patients displayed outlier ratios for two and one of these genes,

respectively. To explore potential genotype-phenotype correlations, we investigated whether this subgroup of 31 patients was associated with specific clinical phenotypes. No significant association was found between this subgroup and the presence of adenocarcinoma ($p = 0.43$; Fisher's Exact test), adenoma ($p = 0.75$; Pearson's Chi-squared test), familial cancer ($p = 0.51$; Pearson's Chi-squared test), early sporadic cancer ($p = 0.41$; Fisher's exact test), or multiple primary tumors ($p = 0.64$; Fisher's exact test).

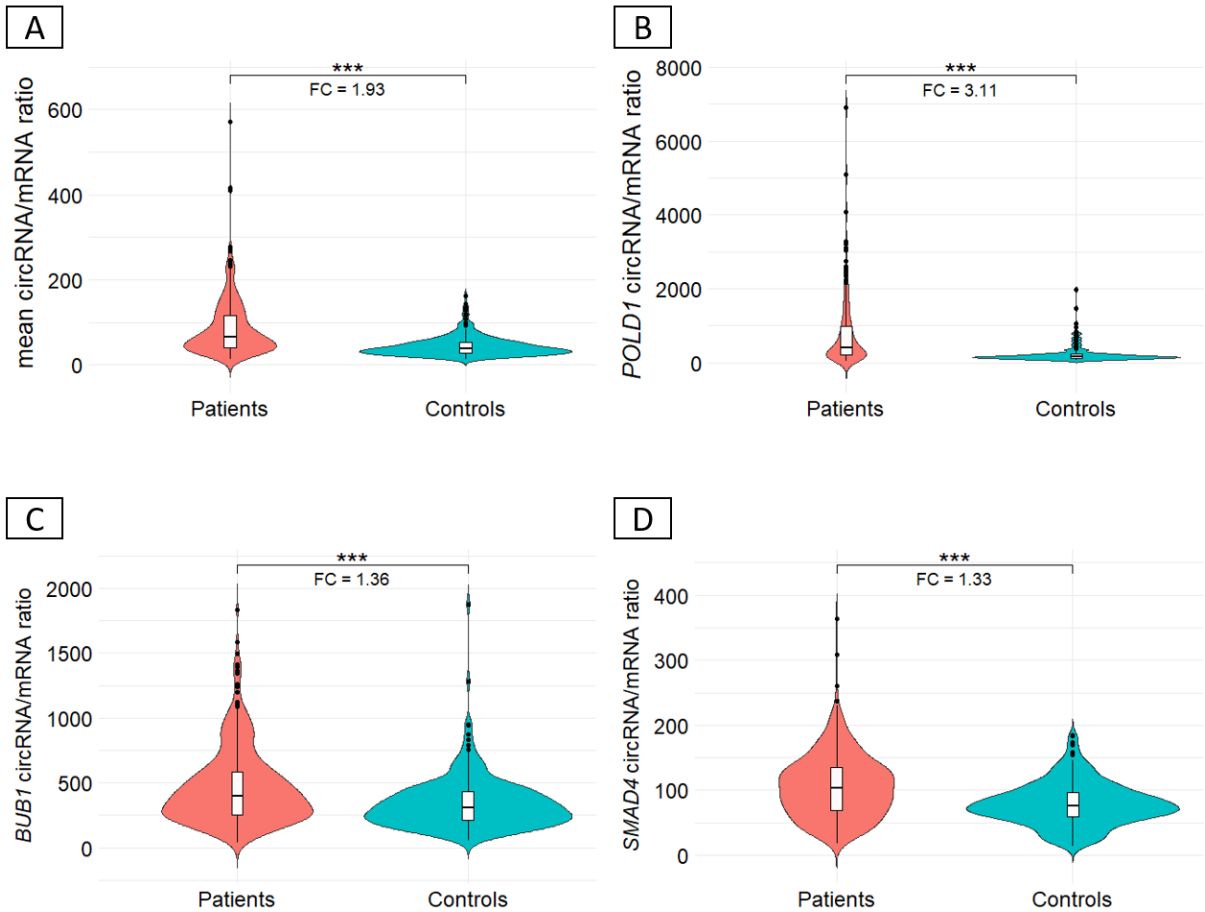


Figure 2. Ratio circRNA/mRNA between CRC patients and controls. (A) Mean ratio for the 17 genes is indicated for patients and controls, in a red and a turquoise violin plot, respectively. Mean ratio between patients and controls for (B) *POLD1*, (C) *BUB1*, (D) *SMAD4* violin plots ratios. Fold Changes (FC) are calculated with the mean ratio of patients and controls. (***: $p < 0.001$; Student's test). Black dots represented the outliers (mean $\pm$ 2DS).

To shed light on a possible mechanism, we then investigated whether this circRNA/mRNA increase was linked to a drop or an increased level of mRNA or circRNA, respectively. Regression analysis between circRNA and mRNA expression showed that this imbalance probably reflected an increase in circRNA production or accumulation rather than a lower *POLD1*, *BUB1* and *SMAD4* expression, thereby suggesting an independent regulation (Figure 3).

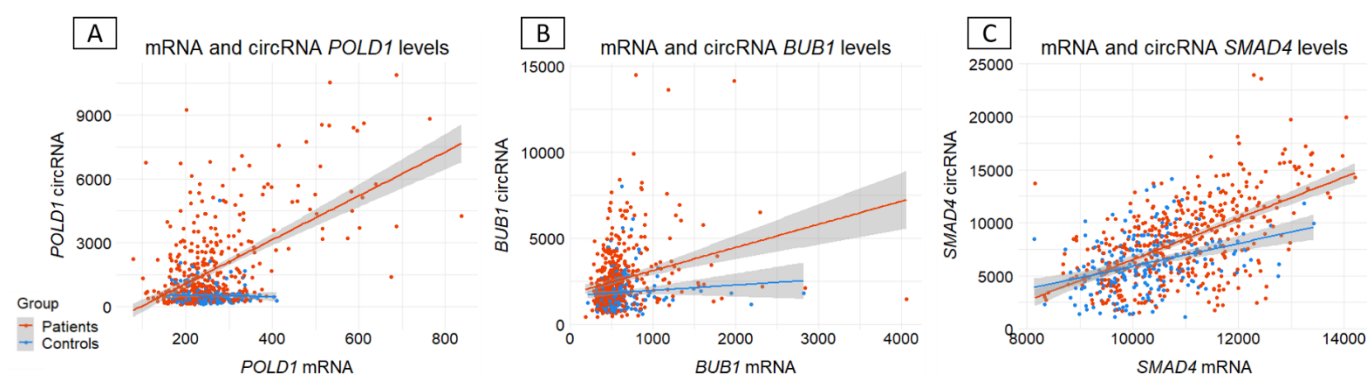


Figure 3. Linear regression analysis between circular RNAs (circRNAs) and messenger RNAs (mRNAs) levels for (A) *POLD1*, (B) *BUB1* and (C) *SMAD4*. X-axis: mRNA expression; Y-axis: circRNA expression. mRNA expression is calculated by averaging the unique molecular identifiers (UMI) of the canonical junctions. circRNA expression is the sum of circular junction UMI counts. Orange and blue points represent patients and controls, respectively. The linear regression is shown by lines of the same color and an area of uncertainty in gray.

4. Discussion

CircRNAs are now established as key players in human pathology, but original data remain scarce. We investigated whether disruption to the coregulation of mRNA–circRNA could represent a novel disease mechanism in a cohort of 712 unexplained CRC cases and 249 matched controls. While our data do not support a competitive regulatory mechanism, they reveal a significant imbalance in the mRNA–circRNA couple in blood between CRC patients and controls.

These results were obtained using SEALigHTS, a novel targeted high-throughput approach that, to our knowledge, is the only technology enabling the simultaneous quantitative analysis of mRNAs and circRNAs, including from FFPE-compatible material. By relying on probe hybridization at exon boundaries and UMI-based quantification prior to amplification, SEALigHTS circumvents the major limitations of conventional RNA-seq strategies for circRNA detection [38], without requiring RNA enrichment or specific sample treatment [39].

Our study provides new insights into i) the mechanism of backsplicing ii) the landscape of circRNAs for 20 CRC genes iii) the involvement of circRNAs from these 20 genes in CRC. These different aspects will be discussed successively. Linear splicing and backsplicing of exons are mediated by the spliceosome and occur simultaneously during the transcription of most human genes. Briefly, two main models for exonic circRNAs biogenesis coexist. Following the first model, known as "lariat-driven circularization" or "exon skipping", the skipping of one or more exons of the pre-mRNA results in a spliced mRNA and a lariat containing the exon(s) that will be circularized. In the second "Intron-Pairing-Driven Circularization" [17] model the presence of intronic repeated inverted complementary sequences, such as Alu sequences, enables the pairing of two introns that border exons that will be circularized. These sequences promote the spatial connection of the 5' splice site junction to that of the upstream 3' site [40] and leads to the formation of a circular transcript. Although both models may explain our results, we counted 220 circular junctions and 138 alternative linear splicing (i.e. exon skipping) junctions, not related to circRNAs for the most part. As an example, *BUB1* has 28 circular junctions but 7 alternative linear splicing junctions. Consequently, and based on these observations, the "Intron-Pairing-Driven Circularization" model is more likely used.

Albeit counterintuitive, the number of circRNAs doesn't increase with the number of exons/introns (Table 2). This might be explained by independent regulation and/or the fact that circRNA biogenesis also depends on several parameters such as the size of flanking introns and the presence of Alu sequences (see the "Intron-Pairing-Driven Circularization" model), the transcription speed [41] in particular through interaction with DNA forming circRNA:DNA hybrids (circR loops) [42], and the strength of splice sites [26]. *POLD1_circRNA_3-2*, *BUB1_circRNA_9-2* and *BUB1_circRNA_4-3* were more highly expressed than their parental mRNA, but the vast majority of

circRNAs (217/220, 94%) represented less than 10% of the corresponding mRNA (Supplementary Table S1).

These results confirm that backsplicing is less efficient than splicing [43] and suggest a transcriptional backsplicing background as already known for linear splicing. More surprisingly, at the gene level, the overall expression of circRNAs compared with the corresponding mRNAs is often (9/20 genes) greater than 10% (Table 2). The values observed for *POLD1*, *BUB1* and *SMAD4*, 5 times higher for the patients' outlier group for *POLD1* (Figure 2), can hardly be considered as background noise and could have a biological impact (Table 2). This suggests that each gene has its own circRNA repertoire with different, finely but uncoupled circRNA/mRNA expression levels [44,45]. Although a competitive mechanism has been ruled out, we demonstrate that the couple may be unbalanced in certain pathological situations [20,31].

Patients suspected of hereditary predisposition to colorectal cancer had a 1.93-fold higher circRNA/mRNA ratio as compared to controls, an imbalance especially driven by *POLD1* with a 3.11-fold change. No correlation with the phenotype could be made for the subgroup of outlier patients with the highest circRNA/mRNA ratio. Our results suggest an increased production or a lack of clearance of *POLD1* circRNA rather than a low *POLD1* expression, a finding in line with known *POLD1* contribution to hereditary CRC predisposition that is not related to decreased expression but to missense pathogenic variants in the exonuclease domain [46]. Interestingly, circPOLD1 itself has recently been shown to play a functional role in tumorigenesis. In a recent study, circPOLD1 expression was found to increase with lesion severity in cervical cancer and to promote oncogenic signaling through interaction with RNA-binding proteins such as YBX1, leading to activation of pathways involved in cell proliferation and tumor progression [47]. These data support the biological relevance of circPOLD1 and reinforce the interest of the imbalance observed in our study.

In any case, the circRNA/mRNA imbalance found in CRC patients suggests that the circRNAs studied play a role in some CRC, either as a cause or a consequence of other hidden defect(s). As circRNAs are expected to be finely regulated with cell-type dependent levels [48], this imbalance could also reflect cellular composition changes following tumorigenesis or chemotherapy, and may serve as a biomarker linked to the onset of cancer and progression [20]. Given the phenotype of the patients under study i.e. a cohort of patients suspected of hereditary predisposition to colon cancer, the small subgroup of 31 outlier patients is of special interest and requires further investigations e.g. cosegregation analyses of this imbalance in families, circRNA-mRNA tumor analyses and the search for newly formed peptide(s) to shed light on novel mechanism(s).

5. Conclusions

Overall, splicing and backsplicing analyses with SEALigHts highlight the RNA architecture of each gene, indicating no evidence of a competitive mechanism or coordinated regulation between circular and linear transcripts. The observed imbalance is largely driven by a single circRNA with oncogenic potential (circPOLD1), suggesting that specific circRNAs may play a pivotal role in disease susceptibility. These findings open new avenues for the study of circRNAs in blood as potential biomarkers or functional contributors to cancer predisposition.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/doi/s1>, Figure S1. Schematic representation of splicing and backsplicing counts. Figure S2. Characterization of circular RNAs (circRNAs) by RT-PCR and Sanger sequencing. Figure S3. Principal Component Analysis (PCA) plot demonstrating the homogeneity of results. Table S1. Probes name and sequences are listed by alphabetic order.

Author Contributions: Conceptualization, C.L. and C.H.; methodology, C.L. and P.R.; software, P.R.; validation, C.L., C.C. and C.H.; formal analysis, C.L., J.D. and C.C.; investigation, J.D., F.C., M.V. and P.R.; resources, E.K., J.M., N.P., S.B.D. and C.H.; data curation, C.L. and P.R.; writing—original draft preparation, C.L. and C.H.; writing—review and editing, C.C. and S.B.D.; visualization, C.L., J.D., C.C., P.R.; supervision, C.H.;

project administration, C.H.; funding acquisition, C.H.. All authors have read and agreed to the published version of the manuscript.

Funding: The work was financed by the Region Normandie and European Regional Development Fund (ERDF). Funding for open access charge: University of Rouen Normandy.

Institutional Review Board Statement: NA.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors thank all the clinical and molecular geneticists from the Groupe Génétique et Cancer.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brenner, S.; Jacob, F.; Meselson, M. An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature* **1961**, *190*, 576–581, doi:10.1038/190576a0.
2. Gros, F.; Hiatt, H.; Gilbert, W.; Kurland, C.G.; Risebrough, R.W.; Watson, J.D. Unstable Ribonucleic Acid Revealed by Pulse Labelling of Escherichia Coli. *Nature* **1961**, *190*, 581–585, doi:10.1038/190581a0.
3. Bentley, D.L. Coupling mRNA Processing with Transcription in Time and Space. *Nat Rev Genet* **2014**, *15*, 163–175, doi:10.1038/nrg3662.
4. Anvar, S.Y.; Allard, G.; Tseng, E.; Sheynkman, G.M.; de Klerk, E.; Vermaat, M.; Yin, R.H.; Johansson, H.E.; Ariyurek, Y.; den Dunnen, J.T.; et al. Full-Length mRNA Sequencing Uncovers a Widespread Coupling between Transcription Initiation and mRNA Processing. *Genome Biol* **2018**, *19*, 46, doi:10.1186/s13059-018-1418-0.
5. Rodríguez-Molina, J.B.; West, S.; Passmore, L.A. Knowing When to Stop: Transcription Termination on Protein-Coding Genes by Eukaryotic RNAPII. *Mol Cell* **2023**, *83*, 404–415, doi:10.1016/j.molcel.2022.12.021.
6. Mironov, A.; Petrova, M.; Margasyuk, S.; Vlasenok, M.; Mironov, A.A.; Skvortsov, D.; Pervouchine, D.D. Tissue-Specific Regulation of Gene Expression via Unproductive Splicing. *Nucleic Acids Res* **2023**, *51*, 3055–3066, doi:10.1093/nar/gkad161.
7. Passmore, L.A.; Collier, J. Roles of mRNA Poly(A) Tails in Regulation of Eukaryotic Gene Expression. *Nat Rev Mol Cell Biol* **2022**, *23*, 93–106, doi:10.1038/s41580-021-00417-y.
8. Aregger, M.; Cowling, V.H. Regulation of mRNA Capping in the Cell Cycle. *RNA Biol* **2017**, *14*, 11–14, doi:10.1080/15476286.2016.1251540.
9. Buccitelli, C.; Selbach, M. mRNAs, Proteins and the Emerging Principles of Gene Expression Control. *Nat Rev Genet* **2020**, *21*, 630–644, doi:10.1038/s41576-020-0258-4.
10. Das, A.; Rout, P.K.; Gorospe, M.; Panda, A.C. Rolling Circle cDNA Synthesis Uncovers Circular RNA Splice Variants. *IJMS* **2019**, *20*, 3988, doi:10.3390/ijms20163988.
11. Tay, Y.; Rinn, J.; Pandolfi, P.P. The Multilayered Complexity of ceRNA Crosstalk and Competition. *Nature* **2014**, *505*, 344–352, doi:10.1038/nature12986.
12. Anastasiadou, E.; Jacob, L.S.; Slack, F.J. Non-Coding RNA Networks in Cancer. *Nat Rev Cancer* **2018**, *18*, 5–18, doi:10.1038/nrc.2017.99.
13. Wu, Y.-Y.; Kuo, H.-C. Functional Roles and Networks of Non-Coding RNAs in the Pathogenesis of Neurodegenerative Diseases. *J Biomed Sci* **2020**, *27*, 49, doi:10.1186/s12929-020-00636-z.
14. Zheng, Y.; Zhang, Y.; Zhang, X.; Dang, Y.; Cheng, Y.; Hua, W.; Teng, M.; Wang, S.; Lu, X. Novel lncRNA-miRNA-mRNA Competing Endogenous RNA Triple Networks Associated Programmed Cell Death in Heart Failure. *Front Cardiovasc Med* **2021**, *8*, 747449, doi:10.3389/fcvm.2021.747449.
15. Huang, Z.; Kuang, N. Construction of a ceRNA Network Related to Rheumatoid Arthritis. *Genes (Basel)* **2022**, *13*, 647, doi:10.3390/genes13040647.

16. Qu, S.; Liu, Z.; Yang, X.; Zhou, J.; Yu, H.; Zhang, R.; Li, H. The Emerging Functions and Roles of Circular RNAs in Cancer. *Cancer Letters* **2018**, *414*, 301–309, doi:10.1016/j.canlet.2017.11.022.

17. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs Are Abundant, Conserved, and Associated with ALU Repeats. *RNA* **2013**, *19*, 141–157, doi:10.1261/rna.035667.112.

18. Wen, X.; Huang, C.; Xie, H.; Hu, D.; Luo, J.; Li, K. The Applications of CircRNA in the Diagnosis and Treatment of Alzheimer’s Disease. *Mol Neurobiol* **2024**, doi:10.1007/s12035-024-03977-0.

19. Min, X.; Liu, D.-L.; Xiong, X.-D. Circular RNAs as Competing Endogenous RNAs in Cardiovascular and Cerebrovascular Diseases: Molecular Mechanisms and Clinical Implications. *Front Cardiovasc Med* **2021**, *8*, 682357, doi:10.3389/fcvm.2021.682357.

20. Xu, C.; Jun, E.; Okugawa, Y.; Toiyama, Y.; Borazanci, E.; Bolton, J.; Taketomi, A.; Kim, S.C.; Shang, D.; Von Hoff, D.; et al. A Circulating Panel of circRNA Biomarkers for the Noninvasive and Early Detection of Pancreatic Ductal Adenocarcinoma. *Gastroenterology* **2024**, *166*, 178–190.e16, doi:10.1053/j.gastro.2023.09.050.

21. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA Circles Function as Efficient microRNA Sponges. *Nature* **2013**, *495*, 384–388, doi:10.1038/nature11993.

22. Huang, D.; Zhu, X.; Ye, S.; Zhang, J.; Liao, J.; Zhang, N.; Zeng, X.; Wang, J.; Yang, B.; Zhang, Y.; et al. Tumour Circular RNAs Elicit Anti-Tumour Immunity by Encoding Cryptic Peptides. *Nature* **2024**, *625*, 593–602, doi:10.1038/s41586-023-06834-7.

23. Zeng, K.; Peng, J.; Xing, Y.; Zhang, L.; Zeng, P.; Li, W.; Zhang, W.; Pan, Z.; Zhou, C.; Lin, J. A Positive Feedback Circuit Driven by m6A-Modified Circular RNA Facilitates Colorectal Cancer Liver Metastasis. *Mol Cancer* **2023**, *22*, 202, doi:10.1186/s12943-023-01848-1.

24. Li, Z.; Huang, C.; Bao, C.; Chen, L.; Lin, M.; Wang, X.; Zhong, G.; Yu, B.; Hu, W.; Dai, L.; et al. Exon-Intron Circular RNAs Regulate Transcription in the Nucleus. *Nat Struct Mol Biol* **2015**, *22*, 256–264, doi:10.1038/nsmb.2959.

25. Rossi, F.; Beltran, M.; Damizia, M.; Grelloni, C.; Colantoni, A.; Setti, A.; Di Timoteo, G.; Dattilo, D.; Centrón-Broco, A.; Nicoletti, C.; et al. Circular RNA ZNF609/CKAP5 mRNA Interaction Regulates Microtubule Dynamics and Tumorigenicity. *Mol Cell* **2022**, *82*, 75–89.e9, doi:10.1016/j.molcel.2021.11.032.

26. Ashwal-Fluss, R.; Meyer, M.; Pamudurti, N.R.; Ivanov, A.; Bartok, O.; Hanan, M.; Evantal, N.; Memczak, S.; Rajewsky, N.; Kadener, S. circRNA Biogenesis Competes with Pre-mRNA Splicing. *Molecular Cell* **2014**, *56*, 55–66, doi:10.1016/j.molcel.2014.08.019.

27. Singh, N.N.; Singh, R.N. How RNA Structure Dictates the Usage of a Critical Exon of Spinal Muscular Atrophy Gene. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2019**, *1862*, 194403, doi:10.1016/j.bbagr.2019.07.004.

28. Bachmayr-Heyda, A.; Reiner, A.T.; Auer, K.; Sukhbaatar, N.; Aust, S.; Bachleitner-Hofmann, T.; Mesteri, I.; Grunt, T.W.; Zeillinger, R.; Pils, D. Correlation of Circular RNA Abundance with Proliferation – Exemplified with Colorectal and Ovarian Cancer, Idiopathic Lung Fibrosis and Normal Human Tissues. *Sci Rep* **2015**, *5*, 8057, doi:10.1038/srep08057.

29. Levacher, C.; Viennot, M.; Drouet, A.; Beaussire, L.; Coutant, S.; Théry, J.-C.; Baert-Desurmont, S.; Laé, M.; Ruminy, P.; Houdayer, C. Disequilibrium between BRCA1 and BRCA2 Circular and Messenger RNAs Plays a Role in Breast Cancer. *Cancers (Basel)* **2023**, *15*, 2176, doi:10.3390/cancers15072176.

30. Trsova, I.; Hrustincova, A.; Krejcik, Z.; Kundrat, D.; Holoubek, A.; Staflova, K.; Janstova, L.; Vanikova, S.; Szikszai, K.; Klema, J.; et al. Expression of Circular RNAs in Myelodysplastic Neoplasms and Their Association with Mutations in the Splicing Factor Gene SF3B1. *Mol Oncol* **2023**, *17*, 2565–2583, doi:10.1002/1878-0261.13486.

31. Wedge, E.; Ahmadov, U.; Hansen, T.B.; Gao, Z.; Tulstrup, M.; Côme, C.; Nonavinkere Srivatsan, S.; Ahmed, T.; Jespersen, J.S.; Schlotmann, B.C.; et al. Impact of U2AF1 Mutations on Circular RNA Expression in Myelodysplastic Neoplasms. *Leukemia* **2023**, *37*, 1113–1125, doi:10.1038/s41375-023-01866-4.

32. Baert-Desurmont, S.; Charbonnier, F.; Houivet, E.; Ippolito, L.; Mauillon, J.; Bougeard, M.; Abadie, C.; Malka, D.; Duffour, J.; Desseigne, F.; et al. Clinical Relevance of 8q23, 15q13 and 18q21 SNP Genotyping to Evaluate Colorectal Cancer Risk. *Eur J Hum Genet* **2016**, *24*, 99–105, doi:10.1038/ejhg.2015.72.

33. Amiot, J.; Levacher, C.; Thibaut, L.M.; Lienard, G.; Vasseur, S.; Quenez, O.; Coutant, S.; Fourneaux, S.; Charbonnier, F.; Thorn, H.; et al. Bridging the Diagnostic Gap in Hereditary Cancers with Simple, Cost-Effective, High-Throughput RNA Splicing Analysis. *J Mol Diagn* **2025**, *27*, 954–968, doi:10.1016/j.jmoldx.2025.06.003.
34. Lanic, M.-D.; Le Loarer, F.; Rainville, V.; Sater, V.; Viennot, M.; Beaussire, L.; Viailly, P.-J.; Angot, E.; Hostein, I.; Jardin, F.; et al. Detection of Sarcoma Fusions by a Next-Generation Sequencing Based–Ligation-Dependent Multiplex RT-PCR Assay. *Modern Pathology* **2022**, *35*, 649–663, doi:10.1038/s41379-021-00980-x.
35. Bobée, V.; Drieux, F.; Marchand, V.; Sater, V.; Veresezan, L.; Picquenot, J.-M.; Viailly, P.-J.; Lanic, M.-D.; Viennot, M.; Bohers, E.; et al. Combining Gene Expression Profiling and Machine Learning to Diagnose B-Cell Non-Hodgkin Lymphoma. *Blood Cancer J.* **2020**, *10*, 59, doi:10.1038/s41408-020-0322-5.
36. Schubert, S.A.; Morreau, H.; van Wezel, T. The Missing Heritability of Familial Colorectal Cancer. **2020**, *35*, 11.
37. Li, Q.; Lai, H.; Li, Y.; Chen, B.; Chen, S.; Li, Y.; Huang, Z.; Meng, Z.; Wang, P.; Hu, Z.; et al. RJunBase: A Database of RNA Splice Junctions in Human Normal and Cancerous Tissues. *Nucleic Acids Res* **2021**, *49*, D201–D211, doi:10.1093/nar/gkaa1056.
38. Viscardi, M.J.; Arribere, J.A. Poly(a) Selection Introduces Bias and Undue Noise in Direct RNA-Sequencing. *BMC Genomics* **2022**, *23*, 530, doi:10.1186/s12864-022-08762-8.
39. Hou, L.; Zhang, J.; Zhao, F. Full-Length Circular RNA Profiling by Nanopore Sequencing with CIRI-Long. *Nat Protoc* **2023**, *18*, 1795–1813, doi:10.1038/s41596-023-00815-w.
40. Ivanov, A.; Memczak, S.; Wyler, E.; Torti, F.; Porath, H.T.; Orejuela, M.R.; Piechotta, M.; Levanon, E.Y.; Landthaler, M.; Dieterich, C.; et al. Analysis of Intron Sequences Reveals Hallmarks of Circular RNA Biogenesis in Animals. *Cell Reports* **2015**, *10*, 170–177, doi:10.1016/j.celrep.2014.12.019.
41. Liang, D.; Tatomer, D.C.; Luo, Z.; Wu, H.; Yang, L.; Chen, L.-L.; Cherry, S.; Wilusz, J.E. The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting. *Molecular Cell* **2017**, *68*, 940–954.e3, doi:10.1016/j.molcel.2017.10.034.
42. Conn, V.M.; Gabryelska, M.; Toubia, J.; Kirk, K.; Gantley, L.; Powell, J.A.; Cildir, G.; Marri, S.; Liu, R.; Stringer, B.W.; et al. Circular RNAs Drive Oncogenic Chromosomal Translocations within the MLL Recombinome in Leukemia. *Cancer Cell* **2023**, *41*, 1309–1326.e10, doi:10.1016/j.ccell.2023.05.002.
43. Zhang, Y.; Xue, W.; Li, X.; Zhang, J.; Chen, S.; Zhang, J.-L.; Yang, L.; Chen, L.-L. The Biogenesis of Nascent Circular RNAs. *Cell Rep* **2016**, *15*, 611–624, doi:10.1016/j.celrep.2016.03.058.
44. Kristensen, L.S.; Okholm, T.L.H.; Venø, M.T.; Kjems, J. Circular RNAs Are Abundantly Expressed and Upregulated during Human Epidermal Stem Cell Differentiation. *RNA Biol* **2018**, *15*, 280–291, doi:10.1080/15476286.2017.1409931.
45. Alonso-García, M.; Liaubet, L.; Faraut, T.; Dehais, P.; Kühn, C.; Demars, J.; Gutiérrez-Gil, B.; Robic, A. Decoupling Transcriptome Layers: The Distinct and Variable Nature of Circular RNAs. *BMC Biol* **2025**, *23*, 269, doi:10.1186/s12915-025-02375-9.
46. Palles, C.; Cazier, J.-B.; Howarth, K.M.; Domingo, E.; Jones, A.M.; Broderick, P.; Kemp, Z.; Spain, S.L.; Guarino, E.; Salguero, I.; et al. Germline Mutations Affecting the Proofreading Domains of POLE and POLD1 Predispose to Colorectal Adenomas and Carcinomas. *Nat Genet* **2013**, *45*, 136–144, doi:10.1038/ng.2503.
47. Zhao, L.; Chen, X.; Zhang, Y.; Cen, Y.; Zhu, T.; Wang, L.; Xia, L.; Li, Y.; Cheng, X.; Xie, X.; et al. The Biomarker Potential of circPOLD1 and Its Binding Protein YBX1 in Cervical Carcinogenesis. *J Transl Med* **2025**, *23*, 506, doi:10.1186/s12967-025-06494-3.
48. Nicolet, B.P.; Engels, S.; Agliarolo, F.; van den Akker, E.; von Lindern, M.; Wolkers, M.C. Circular RNA Expression in Human Hematopoietic Cells Is Widespread and Cell-Type Specific. *Nucleic Acids Res* **2018**, *46*, 8168–8180, doi:10.1093/nar/gky721.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.