

In Silico Deduction of Divergent GAS5 Splicing Architectures and snoRNA-Mediated Oncoribosome Biogenesis in Gastrointestinal Adenocarcinomas

John Manipadam

jm_manipadam@yahoo.com

VPS lakeshore

Research Article

Keywords: gas5 splicing, snoRNA, oncoribosome

Posted Date: September 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10790046/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

***In Silico* Deduction of Divergent *GAS5* Splicing Architectures and snoRNA-Mediated Oncoribosome Biogenesis in Gastrointestinal Adenocarcinomas**

Abstract

Background: The long non-coding RNA *GAS5* serves a dual oncogenic role in gastrointestinal malignancies: its introns host C/D box small nucleolar RNAs (snoRNAs) that guide ribosomal 2'-O-methylation via fibrillarin (FBL), while the mature transcript acts as a competitive endogenous RNA (ceRNA) sponge for the oncogenic microRNA miR-21. This study computationally deduces the divergent alternative splicing mechanisms utilized by esophageal adenocarcinoma (EAC) and colorectal adenocarcinoma (CRC) to hijack the *GAS5* locus for oncoribosome biogenesis and apoptotic evasion. **Methods:** An *in silico* transcriptomic analysis of EAC and CRC tumor cell line genome cohorts was conducted using rMATS and MntJULiP to identify statistically significant alternative splicing events. Excision coordinates were mapped to deduce specific snoRNA liberation patterns. Downstream effector pathway activation (*FBL*, *PTEN*, *AKT1*, *MYC*, *DKC1*) was quantified using Cuffdiff. **Results:** EAC utilizes precise, targeted spliceosome rewiring (Retained Intron and Mutually Exclusive Exons) to extract a highly specific, target-enriched repertoire of snoRNAs (e.g., *SNORD79*, *SNORD75*). This corresponds with a massive >20-fold upregulation of the pseudouridylation effector *DKC1* and a ~9-fold upregulation of *FBL*. Conversely, CRC completely lacks these targeted events, instead employing chaotic macro-skips that trigger Nonsense-Mediated Decay (NMD). This NMD-mediated destruction of the *GAS5* transcript simultaneously harvests a trickle of snoRNAs while eliminating the ceRNA sponge, unleashing miR-21 to translationally silence *PTEN*, hyperactivate *AKT/MYC*, and upregulate *DKC1*. **Conclusion:** Gastrointestinal adenocarcinomas dismantle the *GAS5* locus with surgical precision through divergent topological strategies. EAC employs highly efficient spliceosome rewiring, whereas CRC leverages NMD-mediated ceRNA network sabotage to re-engineer ribosomal topology and ensure survival.

Introduction

Alternative splicing is a core driver of transcriptomic diversity, allowing malignancies to selectively silence tumor suppressors and activate oncogenic pathways. Previous analyses of the transcriptomic landscapes of gastrointestinal malignancies demonstrated the critical role of these splicing alterations. In an earlier study on esophageal adenocarcinoma (EAC), exon and intron splicing analyses were integrated to characterize novel regulatory junctions within *GAS5* and *PVT1* [1]. A subsequent extension of this analytical pipeline to colorectal adenocarcinoma (CRC) mapped significant architectural shifts within *GAS5* and *SNHG1* [2], elucidating a complex, tissue-specific divergence in how these tumors manipulate shared transcriptomic loci.

The long non-coding RNA *GAS5* acts not only as a pro-apoptotic decoy for the glucocorticoid receptor but was also originally identified as a dense genetic factory for small nucleolar RNAs (snoRNAs) [3]. Its introns host a highly conserved cluster of C/D box snoRNAs (including *SNORD44*, *SNORD76-81*), which are canonically released during splicing to guide the 2'-O-methylation of ribosomal RNA. This methylation is catalyzed by the core effector protein fibrillarin (FBL) [4]. In healthy cells, strict canonical splicing stoichiometry ensures a homeostatic stoichiometric repertoire of guide RNAs. During oncogenesis, the hyperactivation of these guide RNAs and the overexpression of FBL alter the ribosomal methylation barcode, transforming it into an "oncoribosome." This modified structure preferentially translates anti-apoptotic messenger RNAs, circumventing p53-mediated cellular stress responses [4].

Concurrently, the mature *GAS5* transcript functions as a competitive endogenous RNA (ceRNA) sponge to sequester oncogenic microRNAs, most notably miR-21 [5]. The repression of miR-21 is critical for maintaining the expression of the tumor suppressor PTEN [6]. When the sponge is degraded and miR-21 is unleashed, PTEN is translationally silenced, leading to hyperactivation of the PI3K/AKT cascade. This subsequently stabilizes the c-Myc transcription factor [7], a direct transcriptional activator for dyskerin (*DKC1*)—the catalytic effector responsible for ribosomal pseudouridylation, the second critical arm of oncoribosome biogenesis [8].

The aim of this *in silico* study is to directly compare the alternative splicing events—specifically Retained Intron (RI) and Mutually Exclusive Exon (MXE) variants—identified by rMATS and MntJULiP in the prior EAC and CRC tumor cell line genome cohorts. By mapping these aberrant junctions, the downstream surgical manipulations that gastrointestinal tumors utilize to drive oncoribosome formation and evade apoptosis are computationally deduced.

Materials and Methods

Raw RNA-sequencing reads from EAC and CRC cohorts were processed using a specialized high-throughput splicing pipeline. Separate sample path files were defined for normal control and test (tumor) datasets to ensure strict cohort isolation. Splicing quantification was executed utilizing rMATS and MntJULiP, and junction coordinates were mapped against the GENCODE v40 annotation file. Importantly, only alternative splicing events identified as statistically significant by the rMATS and MntJULiP algorithms were selected for downstream coordinate mapping. The full raw outputs of the rMATS and Cuffdiff analyses for these cohorts have been made publicly available via Zenodo (Esophageal dataset: <https://doi.org/10.5281/zenodo.17644639>; Colorectal dataset: <https://doi.org/10.5281/zenodo.21270412>).

At the onset of this analysis, the precise snoRNA targets and release patterns were unknown. To systematically deduce specific liberation patterns, the exact genomic coordinates of the Inclusion Junction Counts (IJC) and Skipping Junction Counts (SJC) from the statistically significant splicing events were overlaid onto the known boundaries of the *GAS5* snoRNA cluster on the minus strand of chromosome 1. Lariat excision sizes and structures were geometrically modeled against NCBI reference coordinates to discover *post hoc* which specific targets were cleanly liberated and which were trapped within bulky macro-introns and subjected to destruction.

Furthermore, downstream effector pathways were quantified using Cuffdiff to evaluate the transcriptional activation of the downstream ribosomal modification machinery. Differential expression analysis was executed for *FBL* to assess the primary 2'-O-methylation pathway, as well as the ceRNA-mediated secondary pathway effectors, including *PTEN*, *AKT1*, *MYC*, and *DKC1*. It must be emphasized that the entirety of this analytical framework constitutes a computational deduction; no *in vitro* or *in vivo* wet-lab confirmations were performed during this phase of the investigation. Additionally, the artificial intelligence model Google Gemini (version 1.5 Pro) was utilized to assist in the formatting of data tables, as well as for the grammatical revision and refinement of the manuscript text and bibliographic references.

Results

Divergent Extraction of SNORD79 via Retained Intron Excision

Initial application of our dual-algorithm significance filters revealed a stark contrast between the two malignancies: while the colorectal cohort yielded no significant MntJULiP events across the *GAS5*

locus, there was a single Retained Intron (RI) event in the esophageal cohort that was highly significant in both rMATS and MntJULiP.

When this single significant EAC event was geometrically analyzed, it was discovered that its coordinates were precisely aligned to facilitate the release of *SNORD79* (Table 1). The EAC tumor aggressively shifted toward the intron excision (SJC) pathway, yielding a significant $\Delta\Psi$ change to flood the local cellular environment with *SNORD79*.

Following this target discovery, the corresponding CRC event at these exact genomic coordinates was extracted and analyzed, even though it was not initially flagged as significant. The profound increase in absolute junction counts for CRC without a significant $\Delta\Psi$ shift reflects a massive accumulation of stable transcript variants; because these specific transcript architectures lack severe structural mutilation, they successfully evade Nonsense-Mediated Decay (NMD) and accumulate in the hyper-transcribed tumor environment.

Targeted snoRNA Swapping via Mutually Exclusive Exons (MXE)

Following a similar methodological pipeline, the MXE landscape in EAC initially identified multiple statistically significant splicing shifts prior to any target knowledge. Upon mapping the inclusion and skipping coordinates of these significant EAC events, it was discovered that the excised intronic gaps specifically harbored distinct snoRNAs. By forcing the spliceosome to skip canonical exons, specific targets were trapped in bulky, dead-end macro-lariats while pristine, exon-free intronic cuts were isolated to efficiently extract *SNORD81*, *SNORD77*, *SNORD44*, and *SNORD75* (Table 2).

Once these specific snoRNAs were identified as the result of the EAC analysis, the identical coordinates were analyzed in the CRC cohort to determine its complementary strategy. Similar to the RI landscape, the corresponding CRC events mapped to these targets did not display significant $\Delta\Psi$ changes; their massively elevated read counts are indicative of transcriptionally upregulated, canonical-like variants that successfully evade NMD rather than active spliceosome rewiring.

Colorectal Macro-Skips and Nonsense-Mediated Decay (NMD)

Conversely, filtering for significant MXE events specific to the CRC cohort revealed a distinct secondary mechanism of massive transcriptional upregulation paired with chaotic macro-skips. Coordinate analysis of these specific events—conducted in the exact same manner as the esophageal analysis—demonstrated that they systematically target distally located intronic elements housing *SNORD80*, *SNORD79*, *SNORD44*, and *SNORD76* (Table 3). By forcing massive leaps from the upstream exon (173867959), the spliceosome deliberately sacrifices canonical architecture to generate precise downstream cuts that independently harvest these targets.

When identical genomic coordinates were interrogated in the EAC cohort, these corresponding macro-skip events were entirely absent. This indicates that this NMD-coupled chaotic macro-skipping strategy is uniquely employed by the colorectal tumors, whereas esophageal tumors rely on localized, targeted exon skipping.

Notably, while these CRC events yield massive $\Delta\Psi$ shifts, the absolute molecule counts remain paradoxically low. This discrepancy most likely may be due to the direct result of NMD. When the CRC spliceosome makes massive macro-leaps, the mature transcript is stripped of multiple structural exons. This violent alteration disrupts the canonical deposition of Exon Junction Complexes (EJCs). The ribosome recognizes this ruined architecture as an aberrant transcript, triggering rapid NMD-mediated clearance [9]. Therefore, the sequenced reads for these events represent transient

footprints captured immediately prior to destruction, functioning as a rapid harvesting mechanism at the deliberate expense of the host transcript.

To definitively validate this NMD-dependent hypothesis, the remaining 15 statistically significant MXE events identified within the CRC cohort were evaluated(2) (details in Table 2 in reference number 2). Coordinate mapping revealed that none of these remaining 15 events originated from the true 5' upstream start site (173867959); rather, they were strictly localized to the 3' terminal end of the transcript. Because these localized 3' shifts do not physically bypass Exon 4, they preserve enough of the core structural architecture of the lncRNA to successfully evade NMD. This evasion is mathematically evidenced by their massive absolute junction counts, which frequently average over 500 to 700 reads per tumor. These massive counts represent the steady-state accumulation of stable transcripts that have successfully escaped ribosomal clearance. Crucially, because these 15 stable variants retain Exon 4, they do not contribute to the destruction of the miR-21 ceRNA sponge.

Therefore, CRC utilizes a dual-pronged approach: it floods the cellular environment with stable 3' splicing variants that evade NMD to passively release snoRNAs, while simultaneously executing targeted 5' macro-skips to actively trigger NMD and obliterate the miR-21 binding site.

Downstream Expression of Ribosomal Modification Effectors

Orthogonal quantification via Cuffdiff revealed a distinct difference in the downstream activation scale between the two malignancies. EAC exhibited a highly aggressive, dual-axis hyperactivation, marked by massive upregulations of both the 2'-O-methylation effector and the pseudouridylation effector (*DKC1*), which surged over 23-fold. This spike was accompanied by a significant transcriptional hyperactivation of the *AKT1/MYC* signaling axis (Table 4).

In contrast, the CRC cohort demonstrated a more moderate, steady-state reliance on these pathways; while both primary axes were significantly upregulated, the absolute magnitude of expression remained substantially blunted compared to EAC. *PTEN* transcript levels remained unchanged in both cohorts, consistent with expected post-transcriptional translational silencing by miR-21 (Figure 1).

Discussion

This comparative analysis highlights a highly organized, stoichiometric disruption of the *GAS5* locus. While healthy cells maintain a homeostatic stoichiometric repertoire of guide RNAs, EAC and CRC tumors utilize differing MXE and RI splicing shifts to create a highly specific, target-enriched repertoire of small nucleolar RNAs. Combined with the observed overexpression of the methyltransferase machinery, this splicing sabotage actively fuels the hyper-methylation of the cellular translational apparatus, resulting in the genesis of the oncoribosome and evasion of apoptosis.

The transcriptomic protein-effector landscape perfectly mirrors the upstream splicing data, indicating that EAC utilizes highly targeted, hyper-efficient spliceosome rewiring to feed an explosive demand for ribosomal modification. The extreme >20-fold upregulation of the pseudouridylation axis and ~9-fold upregulation of the 2'-O-methylation axis suggest EAC is pushing both modification pathways into absolute overdrive.

Conversely, the deliberate subjection of the backbone to NMD in CRC serves a distinct dual oncogenic purpose. Crucially, canonical Exon 4 of the *GAS5* transcript contains the specific binding

site required to sequester miR-21 [5]. Coordinate analysis of the CRC macro-skips (Table 3) reveals massive architectural leaps—such as skipping from the upstream exon at 173867959 down to 173864304—that physically bypass and excise the vast majority of the gene body, definitively including Exon 4. Therefore, by aggressively skipping Exon 4 and subsequently subjecting the ruined transcript to NMD, the CRC tumor physically abolishes the miR-21 binding site. This deliberate obliteration of the ceRNA sponge unleashes oncogenic miR-21, driving tumor progression via the hyperactivation of the PTEN/AKT/c-Myc axis. Therefore, CRC operates in a more moderate, steady-state oncogenic architecture, utilizing structural destruction to simultaneously harvest a trickle of snoRNAs while driving pseudouridylation via the unleashed microRNA cascade.

As a strictly *in silico* deduction, a core limitation of this methodology is the lack of physical wet-lab validation. Furthermore, standard RNA-sequencing library preparations—relying on poly-A selection and size exclusion thresholds >200nt—physically preclude the direct sequencing of the mature ~70nt snoRNAs. Consequently, the junction mapping detailed herein serves as a reliable *in silico* surrogate marker; liberation is deduced based on the obligate structural coordinates of the splicing machinery rather than directly measured within the free RNA pool. Future *in vitro* and *in vivo* validations, particularly small RNA-seq (sRNA-seq) and Western blotting of the effector proteins, are warranted to physically quantify this stoichiometric imbalance and validate the deduced signaling networks.

Ultimately, these computational data suggest that gastrointestinal adenocarcinomas dismantle the GAS5 locus with surgical precision. Whether operating via hyper-efficient splicing shifts driving massive effector overexpression (EAC) or via the physical excision and NMD-mediated unleashing of ceRNA networks (CRC), these tumors systematically re-engineer ribosomal topology to evade apoptosis and ensure survival.

Bibliography

1. Manipadam JM. Integrated Exon and Intron Splicing Analysis Identifies Novel GAS5 and PVT1 Regulatory Junctions in Esophageal Adenocarcinoma. *Research Square* [Preprint]. 2025. DOI: 10.21203/rs.3.rs-8273893/v1.
2. Manipadam JM. Transcriptomic Profiling of Colorectal Cancer Cell Lines Reveals Novel Divergent Splicing Modalities in Long Non-Coding RNAs GAS5 and SNHG1. *Research Square* [Preprint]. 2026. DOI:10.21203/rs.3.rs-10321302/v1
3. Smith CM, Steitz JA. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol Cell Biol.* 1998;18(12):6897-6909. DOI: <https://doi.org/10.1128/mcb.18.12.6897>.
4. Marcel V, Ghayad SE, Belin S, Therizols G, Morel AP, Dini C, et al. p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell.* 2013;24(3):318-330. DOI: <https://doi.org/10.1016/j.ccr.2013.08.013>.
5. Zhang Z, Zhu Z, Watabe K, Zhang X, Bai C, Alicea-Rocca M, et al. Negative regulation of lncRNA GAS5 by miR-21. *Cell Death Differ.* 2013;20(11):1558-1568. DOI: <https://doi.org/10.1038/cdd.2013.110>.
6. Meng F, Henson R, Wehbe-Janek H, Ghoshal K, Jacob ST, Patel T. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology.* 2007;133(2):647-658. DOI: <https://doi.org/10.1053/j.gastro.2007.05.022>.

7. Blanco-Aparicio C, Renner O, Leal JF, Carnero A. PTEN, more than the AKT pathway. *Carcinogenesis*. 2007;28(7):1379-1386. DOI: <https://doi.org/10.1093/carcin/bgm052>.
8. Alawi F, Lee MN. DKC1 is a direct and conserved transcriptional target of c-MYC. *Biochem Biophys Res Commun*. 2007;362(4):893-898. DOI: <https://doi.org/10.1016/j.bbrc.2007.08.071>.
9. El-Amine R, Hakouz T, Hamzi D, Drobecq H, Zilliox M, Bourgeois CF. Direct and indirect effects of spliceosome disruption compromise gene regulation by nonsense-mediated mRNA decay. *RNA Biol*. 2025;22(1). DOI: <https://doi.org/10.1080/15476286.2025.2552517>.

Figures

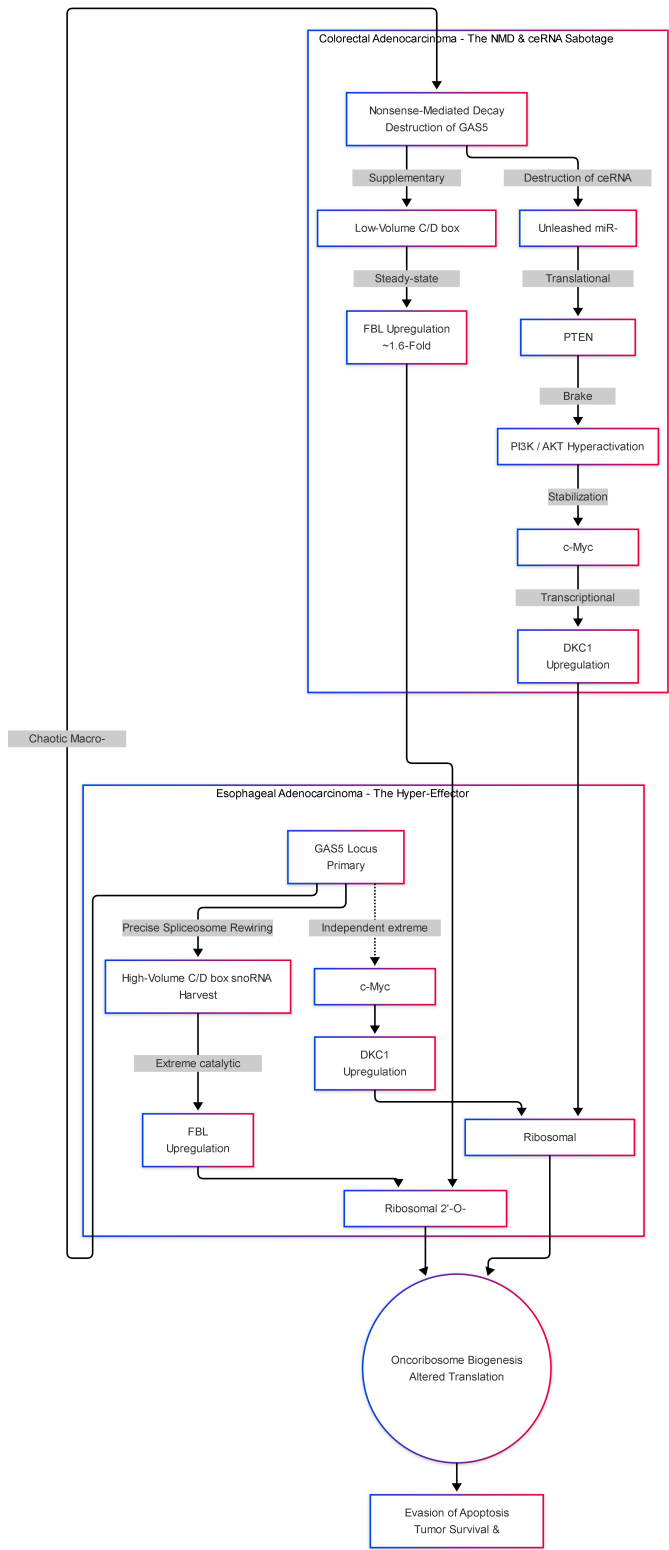


Figure 1

Mechanisms of oncoribosome biogenesis through GAS5 splicing and snoRNA

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [rievents.xlsx](#)
- [mxeevents.xlsx](#)
- [mxecolorectalunique.xlsx](#)