Differential Gene Expression Between Ectopic and Eutopic Endometrium Highlights Cancer-Linked Pathways in Endometriomas

Michael E. Gage

***Abstract*— Endometriomas are ovarian lesions formed from ectopic endometrial tissue and are clinically linked to an increased risk of epithelial ovarian cancers. However, the molecular mechanisms that underlie this malignant potential remain poorly defined. In this study, we performed transcriptomic profiling on ectopic ovarian endometrioma tissue (n=10) and paired eutopic endometrial biopsies (n=10) using publicly available RNA-sequencing data. Differential expression analysis identified 7,248 significantly altered genes (FDR < 0.05), with distinct transcriptional signatures observed between lesional and control samples. Functional enrichment revealed downregulation of core genes involved in cell cycle regulation and DNA damage repair, including TP53, BRCA1, CCNA1, and CCNE2. Interestingly, ATM was markedly upregulated, suggesting an uncoupling of DNA damage sensing from downstream checkpoint activation. Novel genes such as C9orf152 and RP11-6E9.4, previously unreported in endometriomas, were also significantly downregulated and may contribute to immune and epithelial dysregulation. Together, these findings point to early oncogenic shifts in endometrioma tissue and offer new insight into potential molecular drivers linking endometriosis to ovarian cancer progression.**

I. Introduction

Endometriosis is a chronic disease associated with pelvic pain, infertility, pain during menstruation, chronic inflammation, retrograde menstruation, and scar tissue buildup in the pelvic region. This disease often affects women from the onset of their first period until menopause [1]. Endometriosis affects 6-10% of women of reproductive age, and among women experiencing pelvic pain, infertility, or both, the prevalence of endometriosis is estimated to be between 35% and 50% [2].

Endometriomas, or OMAs, are pseudocyst-like growths of endometrial tissue in the ovaries. These ectopic lesions are associated with an increased risk of Type I ovarian cancers, including clear-cell and endometrioid cancer [3]. Aggressive Type II ovarian cancers like high-grade serous ovarian cancer or carcinosarcomas are also risk factors for patients with OMAs.

The development of ovarian cancer is heavily associated with genetic mutations. In 95% of patients with Type II ovarian cancer, TP53 is mutated [3]. Mutations in BRCA1 and BRCA2 are also associated with ovarian cancer in around 15% of cases, though these genes are more commonly associated with breast cancer [4]. TP53 and BRCA genes function as key tumor suppressors, and their mutation, suppression, or degradation—particularly when co-occurring—is strongly associated with the development of ovarian cancer [5].

Cyclin genes like CCNE and CCNA have also been established as biomarkers for a variety of cancers when overexpressed in tissues [6]. These genes are downstream from TP53 in the G1/S cell cycle checkpoint, and their dependence on the functionality of TP53 is essential in tumor suppression pathways. Increased expression of cyclin genes has been implicated in the enhancement of mitotic activity and the tumorigenesis of ovarian cancer [7].

The increased risk of malignant transformation in patients with ovarian endometriomas underscores the need to investigate the genetic mechanisms that may contribute to this progression. We hypothesize that ectopic endometrioma tissue exhibits gene expression profiles and pathway dysregulation similar to those observed in ovarian cancer. By comparing ectopic endometrioma lesions to histologically normal eutopic endometrium from patients with endometriosis, this study aims to identify transcriptional signatures indicative of early carcinogenic processes and to uncover molecular mechanisms that may underlie the elevated cancer risk associated with these lesions.

II. Materials and Methods

Paired-end RNA-seq FASTQ files were retrieved from the Sequence Read Archive (SRA) [8] under BioProject PRJNA1174758, corresponding to the study indexed in the Gene Expression Omnibus (GEO) under accession GSE279835 [9]. The samples selected include a cohort of 10 ectopic ovarian endometrioma biopsies (Stage III-IV; ENDO group) and 10 eutopic endometrial tissue biopsies (CTL group)(Table 1).

All samples were collected at Tartu University Hospital, Estonia, and disease states were confirmed by a pathologist according to ASRM guidelines [9]. Criteria for exclusion included infections, endocrine or metabolic disorders, anatomic abnormalities, autoimmune diseases, and carcinomas. Women included in the study were not taking hormonal medications within three months of biopsy collection [9].

SRA accession numbers for the relevant samples were submitted to the Galaxy platform to facilitate downstream quality assessment and alignment procedures [10].

Raw sequence reads underwent initial quality checks using FastQC [11], with outputs subsequently compiled into a unified summary via MultiQC [12]. To enhance read quality, adapters and low-quality ambiguous reads were removed using Trimmomatic with default parameters for paired end reads [13]. The GRCh38.p14 version [14] of the human genome reference sequence (FASTA) and its corresponding gene

annotation file (GTF) were downloaded from Ensembl [15] and uploaded to Galaxy. Trimmed reads were aligned to the reference genome using HISAT2, yielding BAM alignment files for each sample [16].

Aligned reads were quantified using HTSeq-count, pairing BAM files with the Ensembl GTF annotation to generate gene-level read counts [17]. The resulting count tables were extracted from Galaxy and imported into RStudio for statistical analysis [18,19]. Individual tables were merged into a single matrix for uniform processing.

Genes with fewer than 10 reads across all samples were excluded to minimize noise from low expression. Exploratory analysis was conducted via pairwise scatter plots (Fig. 1) and principal component analysis (Fig. 2) to evaluate sample similarity, detect outliers, and assess clustering patterns. Pairwise comparison across samples was further analyzed qualitatively to visualize correlation across groups (Fig. 3).

Normalized counts and differential expression testing were performed using the *DESeq2* package [20]. The Benjamini-Hochberg procedure for multiple testing correction was used to adjust p-values for false discovery rate (FDR). Comparisons were made between the ectopic endometrioma (ENDO) group and the eutopic endometrium (CTL) group. Significant results were visualized using MA plots for the pairwise contrast to display the relationship between expression strength and statistical significance.

Gene Ensembl IDs were annotated with gene symbols and supplemented with metadata. Volcano plots were generated using the *EnhancedVolcano* package [21] in conjunction with *org.Hs.eg.db* [22] to annotate gene functions and significance levels. Candidate gene sets were used to construct heatmaps using the *pheatmap* package in R [23]. Additional candidate genes were visualized in boxplots using *ggplot2* [24] and *reshape2* R [25] packages.

Differentially expressed gene lists were exported from the R environment and uploaded to the Database for Annotation, Visualization, and Integrated Discovery (DAVID) for functional enrichment analysis. This included Gene Ontology (GO) terms, KEGG pathway associations, Reactome pathway associations, and protein domain over-representation 26,27. Gene set enrichment analysis was conducted with the UC San Diego/Broad Institute's GSEA program [28] using a pre-ranked gene list and the *h.all.v2024.1.Hs.symbols.gmt* MSigDB gene set [29]. Results were cross validated with findings from previous literature.

The complete analysis workflow, including all tools and decision points, was visually summarized using BioRender (Fig. 4). The R analysis pipeline can be found as a supplementary RMarkdown script (.rmd) and PDF.

III. Results

Principal component analysis revealed distinct clustering between the ENDO and CTL sample groups, validating the unique transcriptomic expression of ectopic endometrioma and eutopic endometrial lesions (Fig. 2).

Differential expression analysis identified 7,248 genes significantly altered between ENDO and CTL groups (FDR < 0.05). Of these genes, 4,267 were upregulated and 2,981 were downregulated. These results suggest that gene expression profiles vary largely between ectopic and eutopic endometrial lesions (Fig. 5).

Among the genes showing the greatest upregulation in ectopic endometrioma tissue were PLA2G2A, DLK1, INMT, SCN7A, and SFRP2 (Table 2). Similarly, genes exhibiting the strongest downregulation included C9orf152, TRH, MMP26, RP11-6E9.4 (antisense), and SCGB1D2 (Table 3).

To visualize the overall distribution of gene expression changes, volcano plots were generated, highlighting genes with large effect sizes (log2 fold change) and high statistical significance (adjusted p-value) (Fig. 6).

Functional enrichment analysis in DAVID revealed significant expression patterns associated with cell division processes. Specifically, 49 genes were mapped to the KEGG pathway, "Cell cycle" (FDR <0.05) (Fig. 7). All genes associated with this pathway were downregulated apart from CDC14B, CCND3, ATM, NIPBL, ATRX, and STAG2, which were upregulated across ectopic endometrioma samples (Fig. 8).

The REACTOME pathway "p53-Dependent G1/S DNA Damage Checkpoint" (Fig. 9) was significantly enriched, with 22 genes meeting the FDR threshold of <0.05. Among these, ATM was the only gene upregulated in ectopic endometrioma (ENDO) samples, while all other genes in the pathway—including TP53, CCNA1, CCNA2, and CCNE2—were consistently downregulated (Fig. 10).

Boxplots were generated to visualize rlog-normalized expression levels of ATM, TP53, CCNA1, CCNA2, CCNE2, and BRCA1 across ectopic endometrioma (ENDO) and eutopic control (CTL) samples. The plot highlights clear downregulation of TP53 and cyclin genes in ENDO samples relative to CTL, alongside significant upregulation of ATM, reinforcing the patterns observed in pathway enrichment analyses (Fig. 11).

Gene Set Enrichment Analysis (GSEA) revealed additional pathway-level insights consistent with the observed differential gene expression patterns. Among the most enriched upregulated Hallmark pathways were several immune-related processes (Table 4). In contrast, downregulated pathways were predominantly associated with cell cycle progression, DNA replication, and genomic maintenance, including components of the DNA damage response (Table 5).

IV. Discussion

The top upregulated genes identified—PLA2G2A, DLK1, INMT, SCN7A, and SFRP2—are each involved in distinct biological processes that are often implicated in tissue remodeling or malignancy. For instance, overexpression of PLA2G2A plays a role in inflammatory lipid signaling and has been linked to cell proliferation and metastasis [30]. High DLK1 expression is also linked to metastasis and has been found in ovarian and adrenocortical carcinoma [31]. INMT expression is highly increased in a variety of cancers and is also linked to schizophrenia and psychoses [32]. Dysregulation of SCN7A has been linked to altered immune

signaling pathways and an antitumor response [33]. Increased SFRP2 levels are associated with oncogenesis and poor patient prognosis in high-grade serous ovarian carcinoma [34].

The top downregulated genes—C9orf152, TRH, MMP26, RP11-6E9.4, and SCGB1D2—also suggest biologically significant roles. C9orf152 function is not widely known, though it may be linked to tumor immunity and epithelial cell differentiation [35]. TRH and MMP26 are known target genes for ovarian steroid hormones in endometrial development [36]. RP11-6E0.4 has not been thoroughly researched, though it has been associated and co-expressed with inflammation-associated genes in a prior study [37]. SCGB1D2 is widely associated with gynecological cancer and estrogen receptor positive tumors [38]. The identification and significant downregulation of C9orf152 and RP11-6E9.4 are novel in context of ectopic endometrioma lesional samples.

The suppression of TP53 is particularly notable, as it is a central regulator of DNA damage response and frequently disrupted in early tumorigenesis. Although ATM, a known upstream activator of TP53 through phosphorylation, was upregulated—possibly reflecting a compensatory cellular response to genotoxic stress—this activation does not appear to result in TP53 expression at the transcript level. This dissociation may suggest post-transcriptional regulation, epigenetic silencing, or a mutation preventing TP53 activation, despite ATM signaling.

Furthermore, the downregulation of G1/S cyclins, including CCNA and CCNE family members, supports the notion that G1 arrest remains intact, at least partially, within the DNA damage checkpoint framework. However, the concurrent downregulation of TP53 suggests that this checkpoint control may be unstable or at risk of failure.

Collectively, these findings point toward a transcriptional environment in ectopic endometrioma tissue that mirrors early features of cellular transformation and may represent a pre-neoplastic state.

Notably, although BRCA1 was not mapped to any significantly enriched pathways in this analysis, it was found to be significantly downregulated in ectopic endometrioma tissue. Given BRCA1's critical role in homologous recombination repair and genomic stability, its suppression remains biologically relevant and may further support the presence of a pre-neoplastic transcriptomic profile and increased susceptibility to oncogenic transformation.

These findings align with enrichment results from DAVID and further support the hypothesis that ectopic endometrioma tissue may exhibit early molecular features associated with ovarian malignancy or elevated oncogenic potential.

Taken together, the results of this study are consistent with prior research demonstrating substantial transcriptomic divergence between ectopic and eutopic endometrial lesions, as well as the heightened risk of ovarian cancer associated with ectopic endometriomas. Notably, our analysis also uncovered several novel differentially expressed genes that warrant further investigation into their potential roles in gynecological oncogenic processes.

ACKNOWLEDGMENT

We acknowledge Frisendahl et al. for generating and publicly sharing the RNA-seq dataset GSE279835. We thank the developers of Galaxy, DESeq2, DAVID, GSEA, and the Bioconductor community for providing powerful open-access resources for bioinformatics research. This project was completed independently without external funding.

REFERENCES

[1] World Health Organization. Endometriosis [Internet]. Geneva: World Health Organization; 2023 [cited 2025 May 16]. Available from: https://www.who.int/news-room/fact-sheets/detail/endometriosis

[2] Burney RO, Giudice LC. Pathogenesis and pathophysiology of endometriosis. Fertil Steril. 2012 Sep;98(3):511–9.

[3] Grandi G, Toss A, Cortesi L, Botticelli L, Volpe A, Cagnacci A. The association between endometriomas and ovarian cancer: preventive effect of inhibiting ovulation and menstruation during reproductive life. Biomed Res Int. 2015;2015:751571.

[4] Pan Z, Xie X. BRCA mutations in the manifestation and treatment of ovarian cancer. Oncotarget. 2017 Dec;8(57):97657–70.

[5] Ghezelayagh TS, Pennington KP, Norquist BM, Khasnavis N, Radke MR, Kilgore MR, et al. Characterizing TP53 mutations in ovarian carcinomas with and without concurrent BRCA1 or BRCA2 mutations. Gynecol Oncol. 2020 Mar;160(3):786–92.

[6] Yang X, Zhou Y, Ge H, Tian Z, Li P, Zhao X. Identification of a transcription factor cyclin family genes network in lung adenocarcinoma through bioinformatics analysis and validation through RT qPCR. Exp Ther Med. 2022;25(1):11762.

[7] Lashen A, Algethami M, Alqahtani S, Shoqafi A, Sheha A, Jeyapalan JN, et al. The clinicopathological significance of the cyclin D1/E1–Cyclin-dependent kinase (CDK2/4/6)–Retinoblastoma (RB1/PRB1) pathway in epithelial ovarian cancers. Int J Mol Sci. 2024;25(7):4060.

[8] Leinonen R, Sugawara H, Shumway M. The Sequence Read Archive. Nucleic Acids Res. 2010 Jan;39(Database issue):D19–21.

[9] Frisendahl C, Tang Y, Boggavarapu NR, Peters M, Lalitkumar PG, Piltonen TT, et al. MIR-193B-5P and MIR-374b-5P are aberrantly expressed in endometriosis and suppress endometrial cell migration in vitro. Biomolecules. 2024;14(11):1400.

[10] Abueg AL, Afgan E, Allart O, Awan AH, Bacon WA, Baker D, et al. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res. 2024;52(W1).

[11] Andrews S. FastQC: A quality control tool for high throughput sequence data [Internet]. Babraham Bioinformatics; 2019. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[12] Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016 Oct;32(19):3047–8.

[13] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug;30(15):2114–20.

[14] Morales J, Pujar S, Loveland JE, Astashyn A, Bennett R, Berry A, et al. A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. Nature. 2022 Apr;604(7905):310–5.

[15] Harrison PW, Amode MR, Austine-Orimoloye O, Azov AG, Barba M, Barnes I, et al. Ensembl 2024. Nucleic Acids Res. 2023;52(D1):D891–9.

[16] Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015 Apr;12(4):357–60.

[17] Anders S, Pyl PT, Huber W. HTSeq: a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015 Jan;31(2):166–9.

[18] R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2023.

[19] RStudio Team. Posit: The Open-Source Data Science Company. Boston: Posit Software; 2023. Available from: https://posit.co/

[20] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

[21] Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. Bioconductor. 2025. Available from: https://bioconductor.org/packages/EnhancedVolcano
[22] Carlson M. org.Hs.eg.db: Genome wide annotation for Human (R package version 3.19.1). Bioconductor. 2024.
[23] Kolde R. pheatmap: Pretty Heatmaps [Internet]. 2019 [cited 2025 May 16]. Available from: https://cran.r-project.org/package=pheatmap
[24] Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag; 2016.
[25] Wickham H. Reshaping data with the reshape package. J Stat Softw. 2007;21(12).
[26] Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50(W1).
[27] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.
[28] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005 Oct;102(43):15545–50.
[29] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. 2015 Dec;1(6):417–25.
[30] Ge W, Yue M, Lin R, Zhou T, Xu H, Wang Y, et al. PLA2G2A+ cancer-associated fibroblasts mediate pancreatic cancer immune escape via impeding antitumor immune response of CD8+ cytotoxic T cells. Cancer Lett. 2023;558:216095.
[31] Pittaway JFH, Lipsos C, Mariniello K, Guasti L. The role of delta-like non-canonical Notch ligand 1 (DLK1) in cancer. Endocr Relat Cancer. 2021 Dec;28(12):R271–87.
[32] Zhong S, Jeong J, Huang C, Chen X, Dickinson SI, Dhillon J, et al. Targeting INMT and interrupting its methylation pathway for the treatment of castration resistant prostate cancer. J Exp Clin Cancer Res. 2021;40(1):1–13.
[33] Li W, Zhou K, Li M, Hu Q, Wei W, Liu L, et al. Identification of SCN7A as the key gene associated with tumor mutation burden in gastric cancer. BMC Gastroenterol. 2022;22(1):112.
[34] Monjé N, Dragomir MP, Sinn BV, Hoffmann I, Makhmut A, Simon T, et al. AHRR and SFRP2 in primary versus recurrent high-grade serous ovarian carcinoma and their prognostic implication. Br J Cancer. 2024;130(8):1249–60.
[35] Clark AJ, Singh R, Leonis RL, Stahlberg EA, Clark ZS, Lillard JW. Gene co-expression network analysis associated with endometrial cancer tumorigenesis and survival outcomes. Int J Mol Sci. 2024;25(22):12356.
[36] Tamm K, Rõõm M, Salumets A, Metsis M. Genes targeted by the estrogen and progesterone receptors in the human endometrial cell lines HEC1A and RL95-2. Reprod Biol Endocrinol. 2009;7:150.
[37] Lv D, Su C, Li Z, Chai X, Xu Z, Pang T. Expression of long non-coding RNAs in chondrocytes from proximal interphalangeal joints. Mol Med Rep. 2017;16(4):5175–80.
[38] Zafrakas M, Petschke B, Donner A, Fritzsche F, Kristiansen G, Knüchel R, et al. Expression analysis of mammaglobin A (SCGB2A2) and lipophilin B (SCGB1D2) in more than 300 human tumors and matching normal tissues reveals their co-expression in gynecologic malignancies. BMC Cancer. 2006;6:88.

TABLES AND FIGURES

TABLES

TABLE I. DESIGN MATRIX OF SELECTED SAMPLES FROM GSE279835

GEO Accession	Tissue	SRA ID	Sample ID
GSM8581746	ectopic ovarian endometrioma	SRR31039950	ENDO_50
GSM8581745	ectopic ovarian endometrioma	SRR31039951	ENDO_51
GSM8581744	ectopic ovarian endometrioma	SRR31039952	ENDO_52
GSM8581743	ectopic ovarian endometrioma	SRR31039953	ENDO_53
GSM8581742	ectopic ovarian endometrioma	SRR31039954	ENDO_54
GSM8581741	ectopic ovarian endometrioma	SRR31039955	ENDO_55
GSM8581740	ectopic ovarian endometrioma	SRR31039956	ENDO_56
GSM8581739	ectopic ovarian endometrioma	SRR31039957	ENDO_57
GSM8581738	ectopic ovarian endometrioma	SRR31039958	ENDO_58
GSM8581737	ectopic ovarian endometrioma	SRR31039959	ENDO_59
GSM8581736	eutopic endometrium	SRR31039960	CTL_60
GSM8581735	eutopic endometrium	SRR31039961	CTL_61
GSM8581734	eutopic endometrium	SRR31039962	CTL_62
GSM8581733	eutopic endometrium	SRR31039963	CTL_63
GSM8581732	eutopic endometrium	SRR31039964	CTL_64
GSM8581731	eutopic endometrium	SRR31039965	CTL_65
GSM8581730	eutopic endometrium	SRR31039966	CTL_66
GSM8581729	eutopic endometrium	SRR31039967	CTL_67
GSM8581728	eutopic endometrium	SRR31039968	CTL_68
GSM8581727	eutopic endometrium	SRR31039969	CTL_69

TABLE II. TOP TEN DIFFERENTIALLY EXPRESSED GENES RANKED BY HIGHEST LOG2 FOLD CHANGE (FDR <0.05)

Gene	Biotype	Log2FC	FDR
PLA2G2A	protein_coding	+11.09	1.55E-17
DLK1	protein_coding	+10.65	1.18E-08
INMT	protein_coding	+9.55	1.42E-18
SCN7A	protein_coding	+9.43	1.15E-12
SFRP2	protein_coding	+9.40	6.86E-07
FIBIN	protein_coding	+9.35	1.50E-29
IGHG3	IG_C_gene	+9.22	1.68E-06
PLA2G5	protein_coding	+9.00	1.40E-14
PNOC	protein_coding	+8.93	7.92E-05
MIR202HG	antisense	+8.73	5.14E-08

TABLE III. TOP TEN DIFFERENTIALLY EXPRESSED GENES RANKED BY LOWEST LOG2 FOLD CHANGE (FDR <0.05)

Gene	Biotype	Log2FC	FDR
C9orf152	protein_coding	-8.35	1.45E-18
TRH	protein_coding	-7.88	1.40E-10
MMP26	protein_coding	-7.75	4.74E-22
RP11-6E9.4	antisense	-7.62	1.03E-22
SCGB1D2	protein_coding	-7.61	7.11E-12
GP2	protein_coding	-7.56	9.68E-03
GJB1	protein_coding	-7.37	3.29E-15
OLFM4	protein_coding	-7.24	2.39E-08
SPDEF	protein_coding	-7.18	2.64E-24
PRR15L	protein_coding	-7.11	3.33E-12

TABLE IV. TOP UPREGULATED GSEA PATHWAYS

Hallmark Term	ES	NES	FDR
HALLMARK_ALLOGRAFT_REJECTION	0.53	1.79	0.00
HALLMARK_INFLAMMATORY_RESPONSE	0.51	1.72	0.00
HALLMARK_TNFA_SIGNALING_VIA_NFKB	0.50	1.69	0.00
HALLMARK_COMPLEMENT	0.44	1.48	0.04
HALLMARK_IL6_JAK_STAT3_SIGNALING	0.46	1.45	0.04
HALLMARK_INTERFERON_GAMMA_RESPONSE	0.41	1.42	0.06
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	0.39	1.35	0.12
HALLMARK_KRAS_SIGNALING_UP	0.39	1.34	0.12
HALLMARK_MYOGENESIS	0.39	1.32	0.13
HALLMARK_UV_RESPONSE_DN	0.39	1.31	0.13
HALLMARK_COAGULATION	0.36	1.20	0.36
HALLMARK_HYPOXIA	0.35	1.17	0.40
HALLMARK_APICAL_JUNCTION	0.35	1.17	0.37
HALLMARK_ANGIOGENESIS	0.42	1.16	0.37
HALLMARK_PANCREAS_BETA_CELLS	0.43	1.12	0.47
HALLMARK_INTERFERON_ALPHA_RESPONSE	0.34	1.10	0.52
HALLMARK_KRAS_SIGNALING_DN	0.33	1.09	0.50
HALLMARK_IL2_STAT5_SIGNALING	0.32	1.09	0.48
HALLMARK_APOPTOSIS	0.29	0.97	0.82
HALLMARK_HEME_METABOLISM	0.28	0.95	0.83
HALLMARK_APICAL_SURFACE	0.32	0.90	0.93
HALLMARK_XENOBIOTIC_METABOLISM	0.27	0.90	0.90
HALLMARK_TGF_BETA_SIGNALING	0.29	0.87	0.92
HALLMARK_UV_RESPONSE_UP	0.26	0.86	0.91
HALLMARK_P53_PATHWAY	0.25	0.85	0.88
HALLMARK_NOTCH_SIGNALING	0.27	0.75	0.98
HALLMARK_ADIPOGENESIS	0.16	0.53	1.00

TABLE V. TOP DOWNREGULATED GSEA PATHWAYS

Hallmark Term	ES	NES	FDR
HALLMARK_E2F_TARGETS	-0.69	-3.36	0.00
HALLMARK_G2M_CHECKPOINT	-0.58	-2.74	0.00
HALLMARK_MYC_TARGETS_V1	-0.56	-2.57	0.00
HALLMARK_MYC_TARGETS_V2	-0.62	-2.48	0.00
HALLMARK_OXIDATIVE_PHOSPHORYLATION	-0.49	-2.39	0.00
HALLMARK_ESTROGEN_RESPONSE_LATE	-0.49	-2.26	0.00
HALLMARK_DNA_REPAIR	-0.45	-2.09	0.00
HALLMARK_PEROXISOME	-0.42	-1.82	0.00
HALLMARK_GLYCOLYSIS	-0.33	-1.56	0.02
HALLMARK_MTORC1_SIGNALING	-0.33	-1.56	0.02
HALLMARK_SPERMATOGENESIS	-0.35	-1.55	0.02
HALLMARK_ESTROGEN_RESPONSE_EARLY	-0.32	-1.54	0.02
HALLMARK_ANDROGEN_RESPONSE	-0.31	-1.41	0.04
HALLMARK_MITOTIC_SPINDLE	-0.29	-1.40	0.04
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	-0.32	-1.38	0.04
HALLMARK_FATTY_ACID_METABOLISM	-0.29	-1.32	0.06
HALLMARK_BILE_ACID_METABOLISM	-0.25	-1.08	0.34
HALLMARK_HEDGEHOG_SIGNALING	-0.25	-0.91	0.83
HALLMARK_PI3K_AKT_MTOR_SIGNALING	-0.21	-0.90	0.83
HALLMARK_WNT_BETA_CATENIN_SIGNALING	-0.24	-0.88	0.84
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	-0.17	-0.63	1.00
HALLMARK_PROTEIN_SECRETION	-0.14	-0.60	1.00

FIGURES

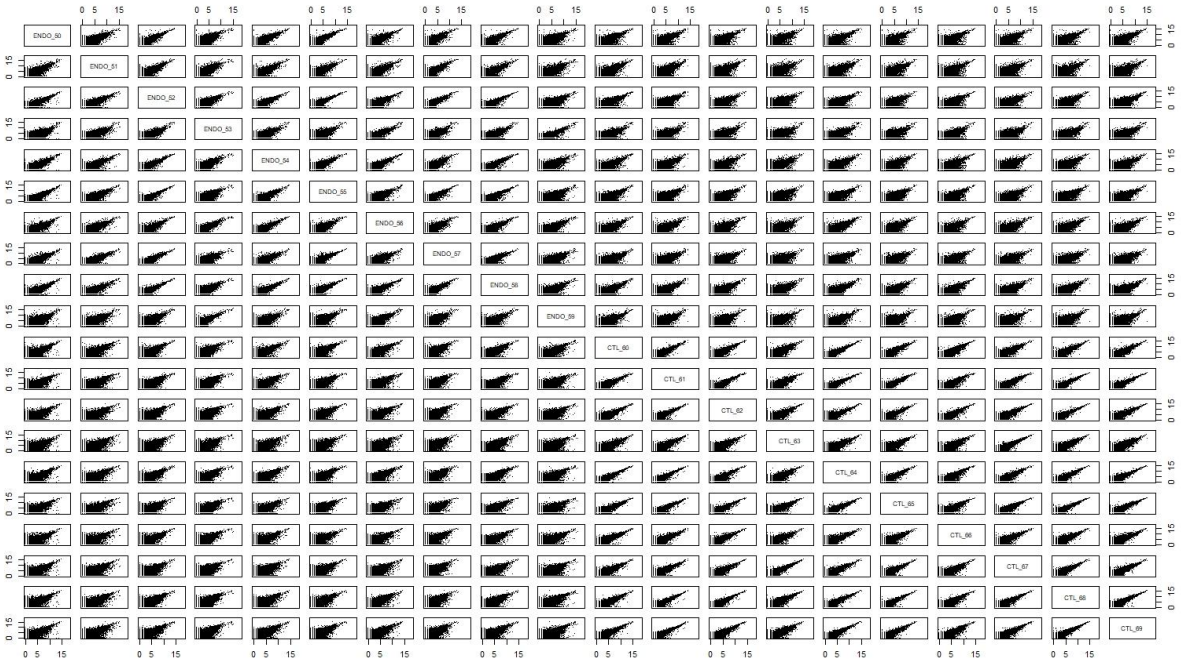

Figure 1. Pairwise scatter plot representing correlation between samples

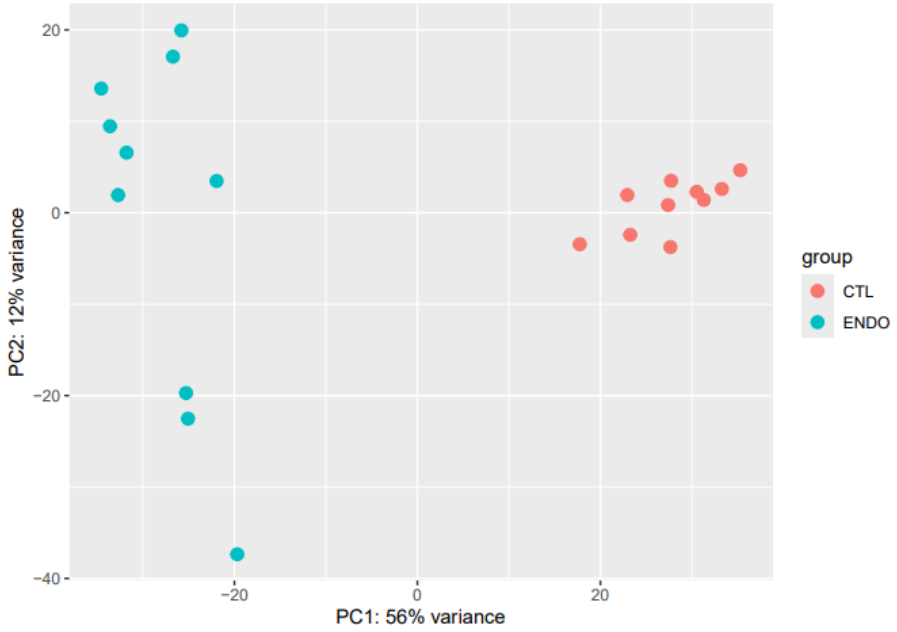

Figure 2. Principal component analysis of sample clustering

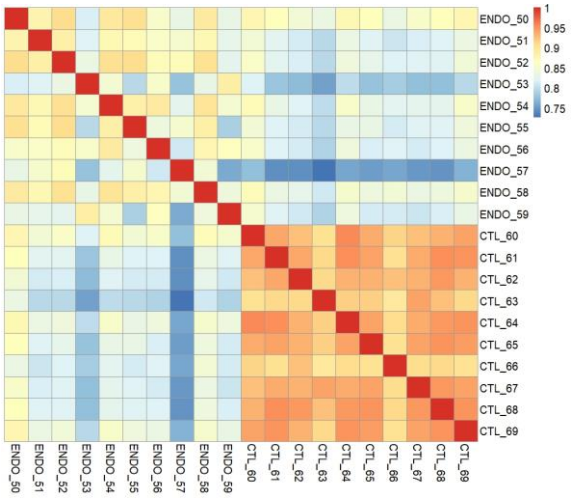

Figure 3. Pairwise heatmap of selected samples

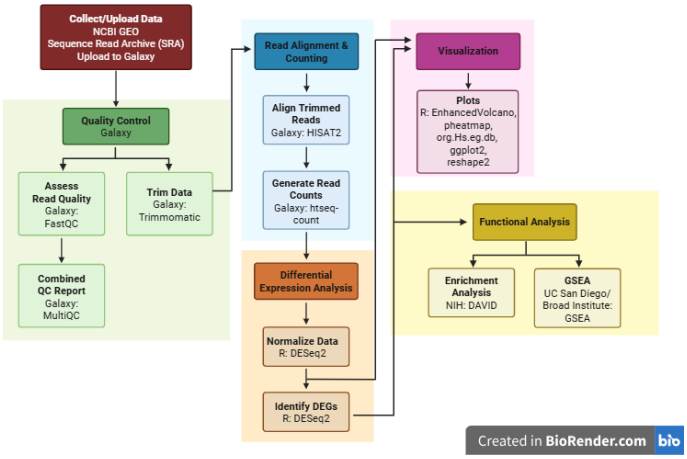

Figure 4. BioRender bioinformatics workflow

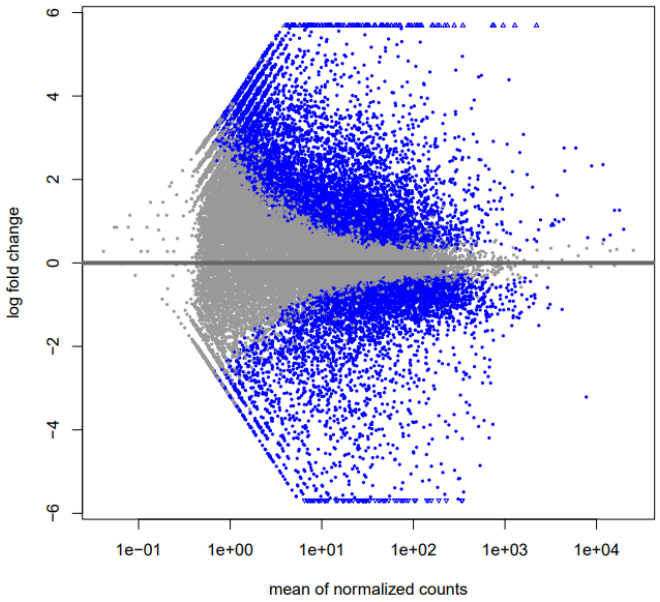

Figure 5. MA plot of differential gene expression between ectopic and eutopic lesions

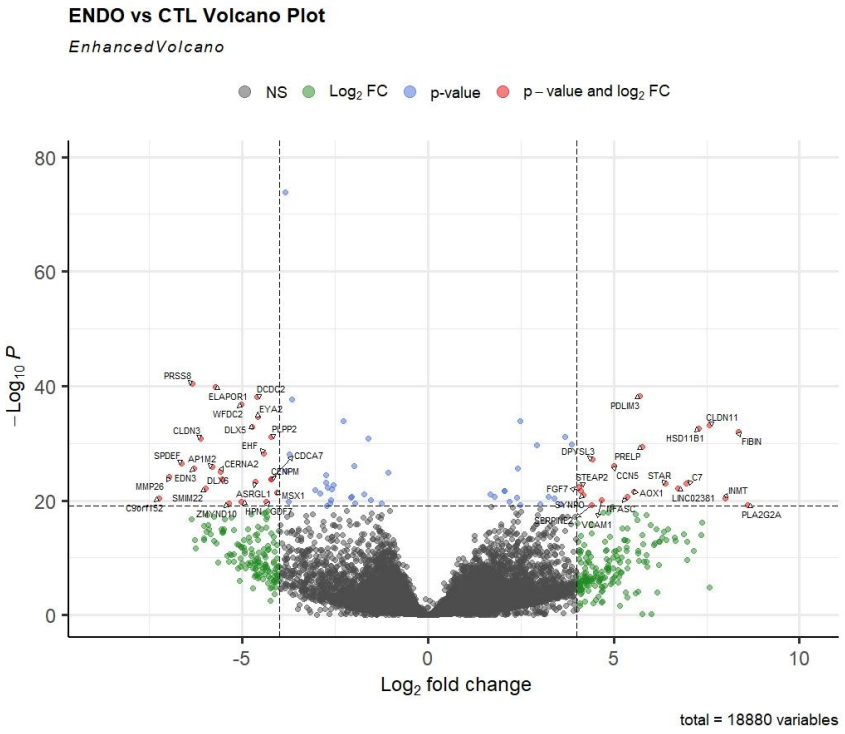

Figure 6. Volcano plot of differential gene expression between ectopic and eutopic lesions

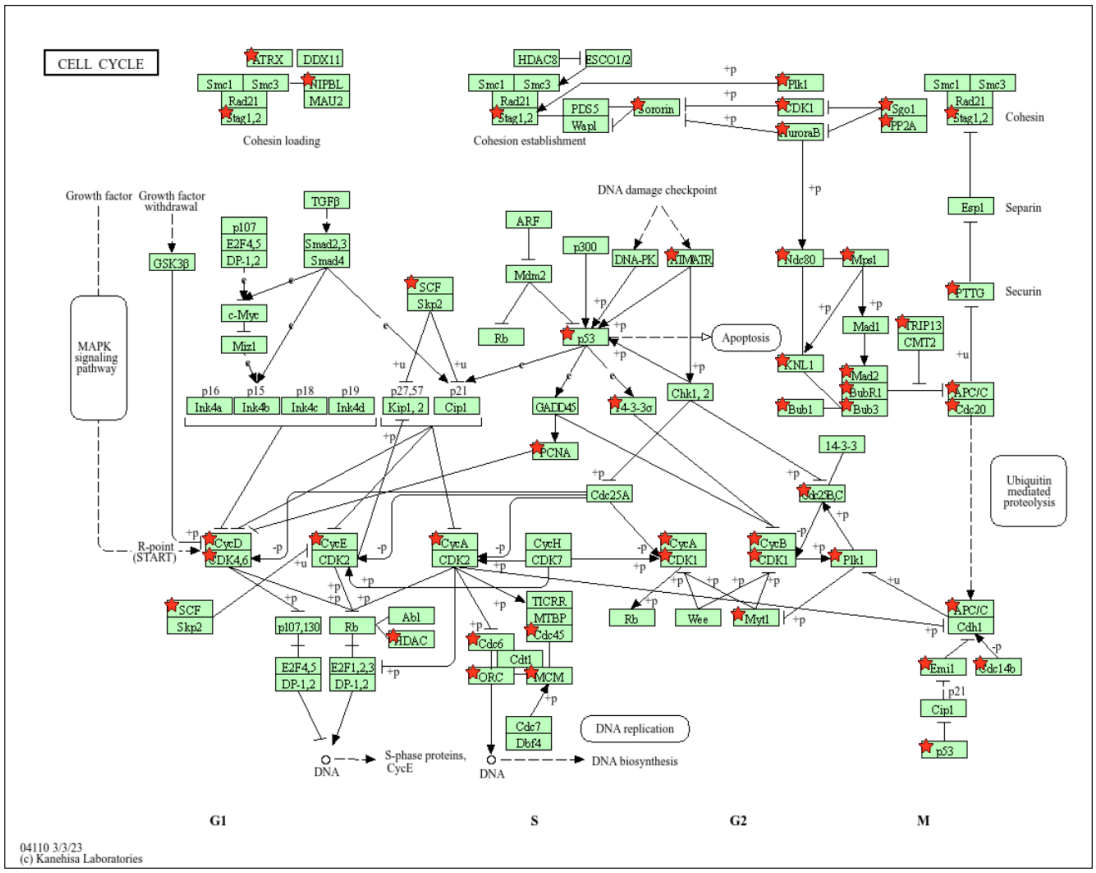

Figure 7. KEGG *Cell Cycle* pathway with mapped genes

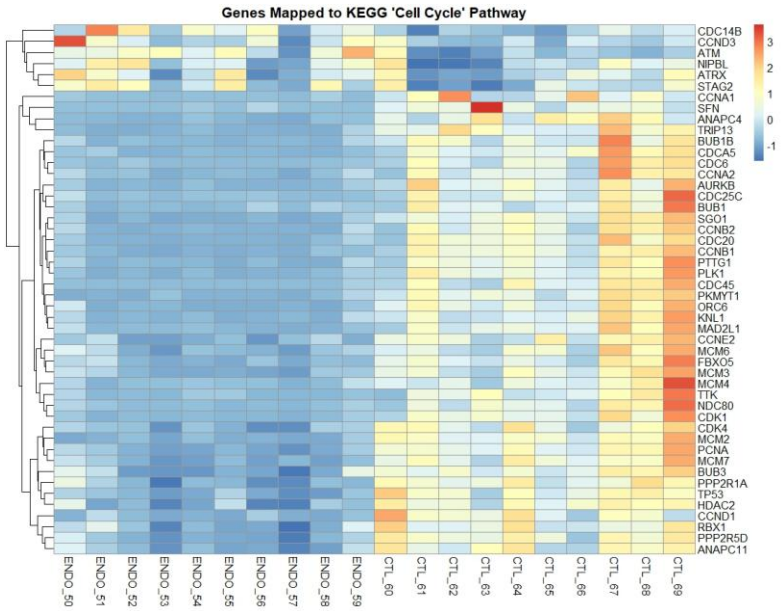

Figure 8. Heatmap of genes mapped to *Cell Cycle* KEGG pathway

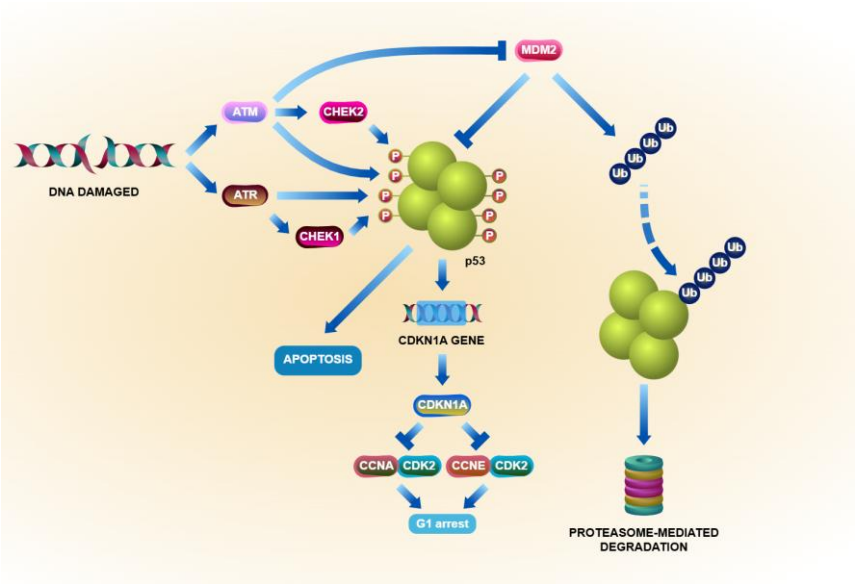

Figure 9. *p53-Dependent G1/S DNA Damage Checkpoint* Reactome pathway

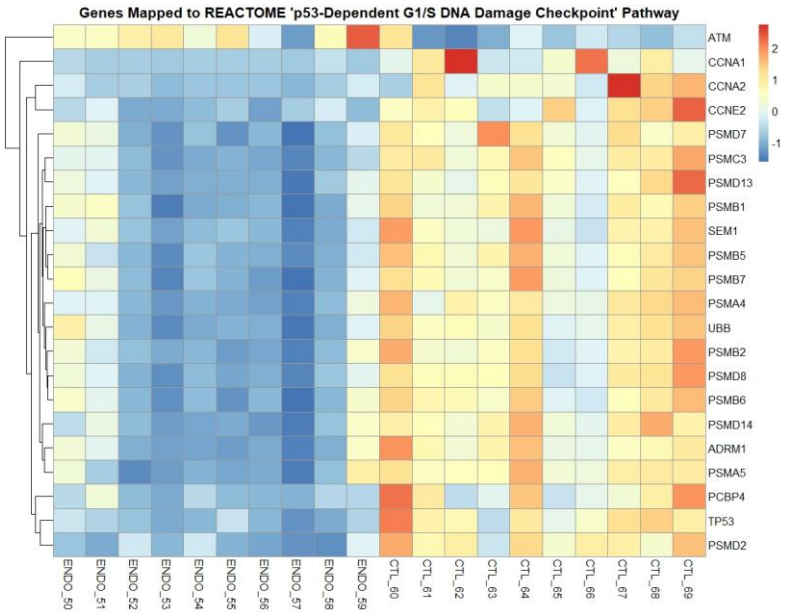

Figure 10. Heatmap of genes mapped to *p53-Dependent G1/S DNA Damage Checkpoint* Reactome pathway

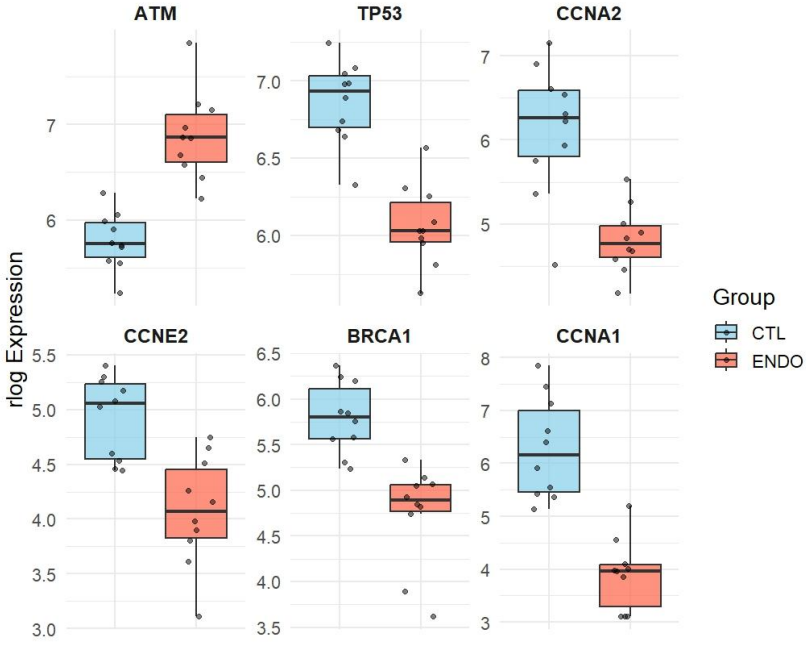

Figure 11. Boxplot of select genes in ectopic vs eutopic samples