Exploring Mitochondrial Gene Expression in Progression to Colorectal Cancer

Problem Statement

Despite advances in cancer research, uncertainty remains about the origin and evolution of normal tissue to cancerous tissue, especially with regards to the role that mitochondrial metabolic processes play. Thus, identifying mitochondrial genes that are dysregulated across different stages of colorectal tissue provides insight into the progression of metabolic dysfunction and development of disease.

Abstract

Colorectal cancer is one of the most common and deadliest cancers in the United States, making research into its origins and underlying causes particularly important. Cancer has long been considered a consequence of random genetic mutation, but in recent years, researchers have begun exploring the possibility that the origin of some cancers may be linked to mitochondrial metabolic dysfunction. This paper examines changes in mitochondrial gene expression during colorectal cancer progression. Analysis of gene expression levels of mitochondrial-related genes across normal, polyp, and cancerous tissue revealed a host of dysregulated genes, including *AKR1B10*, *SLC25* family genes, and *HIGD1A*. Each of these genes displayed dysregulation at the polyp stage in addition to the tumor stage. This finding indicates that mitochondrial dysfunction occurs prior to tumorigenesis and may contribute to the metabolic changes associated with tumor progression.

Literature Review

As of 2026, colorectal cancer (CRC) is the third most commonly diagnosed cancer in men and women and has the second highest rates of cancer-related deaths in the United States (Siegel et al., 2026). Although CRC incidence has seen a decline in older populations, recent studies have shown that incidence rates and mortality are increasing in younger populations (Siegel et al., 2026). Risk factors of CRC include smoking, alcohol consumption, unhealthy diets, and physical inactivity (Siegel et al., 2026). In light of the recent incidence and mortality trends, it is more crucial than ever to explore the etiology of CRC and cancer pathways more broadly.

The origin of cancer has long been attributed to acquired nuclear DNA mutations or inherited genetic abnormalities. This theory is sometimes referred to as the somatic mutation theory, and it suggests that random mutations can occur during the replication of DNA that cause noncancerous cells to turn cancerous (Seyfried, 2015). While this perspective of cancer development is key to understanding some aspects of the disease, it may not explain the whole picture. In recent years, researchers have renewed interest in theories that explore the possibility that the origin of some cancers is more directly linked to cellular bioenergetics and mitochondrial metabolic dysfunction.

Colloquially known as the “powerhouse of the cell,” the mitochondria is considered to be the center through which the cell produces energy, primarily in the form of ATP. Typically, the mitochondria uses processes like the tricarboxylic acid (TCA) cycle, the electron transport chain

(ETC), and oxidative phosphorylation (OXPHOS) to efficiently synthesize energy, mostly from glucose molecules and fatty acids. Due to the key role that oxygen and electrons play in these energetic systems, byproducts called reactive oxidative species (ROS) are produced. At physiological levels, ROS help promote cellular homeostasis and signalling.

However, at higher levels, ROS overwhelm the antioxidants human bodies use to keep them in check and act as highly toxic free radicals that can cause cellular damage (San-Millán, 2023). Because mitochondrial DNA is stored in close proximity to ROS production and is not wrapped around protective histones, it is vulnerable to damage and mutations (Danesova et al., 2025). Furthermore, mitochondrial DNA does not have the same repair mechanisms as nuclear DNA, making some damage irreversible (Danesova et al., 2025). Researchers are still uncertain whether ROS abundance triggers mitochondrial dysfunction or whether mitochondrial dysfunction leads to the overproduction of ROS (San-Millán, 2023).

Studies have shown that when nuclei from tumor cells are combined with cytoplasm from healthy, noncancerous cells tumorigenicity is suppressed (Seyfried, 2015). Conversely, research has also shown that introducing healthy nuclei to cytoplasm from a cancerous cell will stimulate tumor formation (Seyfried, 2015). This observation disrupts some of the ideas core to the model of the somatic mutation theory. Rather than support the idea that genomic mutations drive tumorigenesis, these findings demonstrate that cytoplasmic activities may be the key driver of certain malignancies.

Damage to mitochondrial DNA may not be the only precipitating factor in tumorigenesis. Carcinogenesis could potentially stem directly from metabolic dysregulation. A phenomenon known as the “Warburg Effect” describes the observation that tumor cells favor alternative forms of energy production (San-Millán, 2023, p. 13). Instead of relying primarily on oxidative phosphorylation to produce energy, the cancer cells exhibit an increase in glycolysis and high levels of lactate, a byproduct that is produced through an alternative energy producing process called fermentation (San-Millán, 2023, p. 13). These high levels of lactate can supply molecules for biosynthetic pathways rather than pathways that produce ATP, suggesting that tumors may take advantage of the alternative distribution of resources (Offei & Cluntun, 2026).

Normally, mitochondria plays a key role in the biosynthesis of an important array of molecules, including fatty acids, cholesterol, amino acids, and nucleotides (Wang & Patti, 2023). Neurons, a cell type which almost never replicate, use high numbers of mitochondria to sustain high ATP demand, rarely develop cancer, even though they likely produce high levels of ROS (Tomasetti & Vogelstein, 2015). However, cell types that replicate much more frequently, such as hepatocytes and colonic epithelial cells, are at significantly higher risk for developing cancer (Tomasetti & Vogelstein, 2015). Given these observations, cells that engage in increased division, thus utilizing mitochondrial biosynthesis pathways, appear to be more at risk for tumorigenesis. The disparity between tumor risk in non proliferating and proliferating cell lines could indicate that one significant origin of cancer may arise from a malfunction in mitochondrial resource allocation.

The accumulation of these observations indicate that the mitochondria plays a crucial and not yet fully understood role in cancer progression. Given the rising prevalence of CRC in young people and around the world, as well as the unique relationship the gastrointestinal tract has to the body's energy production, CRC is a useful cancer type to investigate.

Methods

This project used publicly available data from the Gene Expression Omnibus (GEO) database on the National Institutes of Health's (NIH) National Center for Biotechnology Information (NCBI) platform. Three different datasets—identified by the GEO Series accession numbers GSE89076 (Dataset 1), GSE128435 (Dataset 2), and GSE41657 (Dataset 3)—were chosen for analysis. Each dataset provided microarray gene expression data for normal, polyp, and cancerous tissue samples from human patients.

Dysregulated gene expression was evaluated across each of the three datasets using GEO2R. For each dataset, comparisons between different pathological stages were run. For Dataset 1 and Dataset 2, tumor tissue and normal tissue, polyp tissue and normal tissue, and tumor tissue and polyp tissue were compared. Dataset 3 had more specific distinctions between pathological categories. Consequently, the data was analyzed by comparing tumor tissue to normal tissue, tissue with high grade dysplasia to normal tissue, and tissue with low grade dysplasia to normal tissue. Low and high grade dysplasia tissues, as well as tissue with high grade dysplasia and tumor tissue were compared, but no significant changes in mitochondrial-related gene expression were found.

The results from each dataset were then screened for significance (p-adjusted values below 0.05 and log fold change (logFC) less than -2 or greater than 2). Next, the gene expression data that was not represented by a gene symbol was excluded from the analysis, given that the data would not be identifiable without a name. Using Microsoft Excel, the gene symbols of the data that met the significance requirements were compared to the MitoCarta, a list of 1137 genes related to the mitochondria and mitochondrial proteins.

After narrowing down the significant genes by mitochondrial relevance, the logFC, a measure of change between gene expression levels, was found for each significant mitochondrial-related gene. Genes that were consistently up or down regulated across subgroups as well as datasets were identified, and further research was conducted into their specific roles.

Findings

Figure 1 (a-c)

MitoCarta Genes with Significant Up- or Downregulation Across Tissue Type

(a) Dataset 1

Polyp vs. Normal	log FC	Tumor vs. Polyp	logFC	Tumor vs. Normal	log FC
ARMCX1	-2.32	BNIP3	3.33	ACADS	-2.23
BCL2A1	-2.32			ACSL6	3.02
BNIP3	-2.63			ADHFE1	-2.50
CPS1	4.97			AKR1B10	-3.74
EFHD1	-2.19			FABP1	-3.07
GLDC	-2.62			HIGD1A	-2.26
HIGD1A	-2.03			HMGCS2	-3.66
HPDL	2.00			LDHD	-3.01
MAOB	-2.86			MTHFD1L	2.24
MOCS1	-2.25			NEU4	-2.38
MSRB3	-2.60			PDK4	-2.62
PDE2A	-2.34			PYCR1	2.12
PYCR1	2.26			RECQL4	2.25
SLC25A34	-2.55			SLC25A34	-3.15

(b) Dataset 2

Polyp vs. Normal	log FC	Tumor vs. Polyp	logFC	Tumor vs. Normal	log FC
ACAA1	-2.35	ACAA1	2.23	ACACA	-2.63
ACACA	-2.21	MRPS22	-3.76	ACAD10	2.10
ADCY10	-2.12	NDUFB5	2.54	ACOT11	2.22
AKR1B10	-3.77	PHB	3.67	AKR1B10	-3.28
BID	2.93	SPR	-2.16	ARG2	2.05
COX5A	2.03	SYNJ2BP	2.23	BID	2.14
CPT2	-4.62	TRMT10C	2.36	CA5B	-3.20
DUS2	-2.41			CPT2	-4.58
FDXR	2.02			FKBP8	-3.31
FKBP8	-2.81			FOXRED1	-2.55
GTPBP3	-2.43			GTPBP3	-2.04
MRPL24	-2.31			IFI27	-2.24
MRPL54	2.10			MRPL23	2.03
MRPS35	-2.39			MRPS35	-2.63
MTERF4	-2.04			MSRA	-2.48
MYO19	-2.06			MYO19	-2.33
NDUFB5	-2.94			NDUFB6	-2.08
NDUFB6	-2.06			PIF1	-2.38
NDUFS5	2.02			PNPO	-3.41
NDUFV1	-2.07			SLC25A20	2.16
NIT1	-3.09			SOD2	-2.41
NLN	2.09			THEM4	-3.52
PDE2A	-2.23			TIMM13	2.00
PHB	-2.19				
PIF1	-2.80				
PNPO	-3.67				
SHMT2	-2.32				
SLC25A20	2.11				
SLC25A27	2.70				
SOD2	-2.57				
SPG7	-2.27				
SYNJ2BP	-2.31				
THEM4	-3.58				
TIMM21	-2.23				
TTC19	-3.12				

(c) Dataset 3

Low Grade vs. Normal	log FC	High Grade vs. Normal	log FC	High Grade vs. Low Grade	Adenocarcinoma vs. High	Adenocarcinoma vs. Normal	log FC
ACAT1	-2.55	ACAT1	-2.35	nil	nil	ACADS	-2.10
AGXT	-2.45	ACSL6	2.46			ACAT1	-2.19
BCL2A1	-2.24	AGXT	-2.52			ACSL6	**2.93**
BCL2L11	-2.07	AKR1B10	-2.46			AGXT	-2.21
BID	2.36	BCL2A1	-2.57			AKR1B10	-3.52
COX7A1	-3.77	BCL2L11	-2.09			ARG2	2.16
ECHDC3	-2.10	COQ10B	-2.18			BCL2	-2.51
HIGD1A	-2.25	COX7A1	-3.32			C15orf48	-2.54
MICU3	-2.92	CPS1	2.56			CKMT2	-4.10
MOCS1	-2.14	ECHDC3	-2.49			FKBP10	2.15
MSRB3	-2.16	HIGD1A	-2.77			GPT2	2.44
NEU4	-2.04	LDHD	-2.48			HIGD1A	-2.57
SLC25A21	2.14	MAOA	-2.34			HMGCS2	-2.52
		MICU3	-2.64			LDHD	-3.72
		MSRB3	-2.27			MAOA	-2.46
		MTHFD1L	2.00			MTHFD1L	2.36
		NEU4	-2.48			NEU4	-2.41
		PXMP2	-2.12			PINK1	-2.13
		SLC25A20	-2.15			PXMP2	-2.10
						RECQL4	2.23
						SCP2	-2.02

Note. Green indicates upregulated genes in both polyp and adenocarcinoma (tumor) stages. Red represents downregulated genes in both polyp and adenocarcinoma stages. Yellow depicts genes that have changed directionality in their dysregulation. Purple shows genes that were dysregulated in both low and high grade dysplasia stages, but not significantly in the tumor stage.

Dataset 1

In Dataset 1, 14 mitochondria-related genes were dysregulated when comparing normal tissue to polyp tissue. All but two genes were down regulated. When comparing polyp tissue to tumor tissue, only one MitoCarta gene, *BNIP3*, was dysregulated. Interestingly, while *BNIP3* was downregulated in the polyp stage, with a logFC of -2.63, it was upregulated in tumor tissue compared to polyp tissue, with a logFC of 3.33. Comparing tumor tissue to normal tissue yielded 14 dysregulated mitochondrial genes, with three of those genes—*PYCR1*, *HIGD1A*, and *SLC25A34*—in continued up- or downregulation from the polyp stage.

Dataset 2

Dataset 2 found 33 downregulated and 2 upregulated MitoCarta genes in polyp tissue compared to normal tissue. When comparing tumor tissue to polyp tissue, 7 MitoCarta genes were dysregulated. As in Dataset 1, there were four genes that appeared downregulated at the polyp stage, but were upregulated by similar magnitudes between polyp and tumor stages: *ACAA1*, *NDUFB5*, *PHB*, and *SYNJ2BP*. Like *BNIP3* in Dataset 1, this effect brought the net gene dysregulation closer to zero. Comparing tumor tissue to normal tissue, a total of 23 MitoCarta genes were identified as dysregulated. Of the 23 genes, 13 also appeared dysregulated

in the precancerous polyp tissue stage. This finding suggests that a majority of the dysfunctional mitochondrial genes in the tumor were already imbalanced at the polyp stage.

Dataset 3

In Dataset 3, tissue with low grade dysplasia had 13 significantly over or under expressed mitochondrial genes compared to normal tissue. Tissues with high grade dysplasia yielded 19 dysregulated MitoCarta genes when compared to normal tissue, with 10 of those genes already appearing dysregulated in the low grade dysplasia tissue. When comparing adenocarcinoma tissue to normal tissue, 21 genes appeared dysregulated. Ten of the dysregulated tumor tissue genes overlapped with genes that were dysregulated in high grade dysplasia tissue.

However, when comparing tissue of low and high grade dysplasia, as well as high grade dysplasia and adenocarcinoma tissue, there were no mitochondrial genes that were significantly up- or downregulated between these stages. This finding suggests that mitochondrial genes are dysregulated in initial stages of abnormality, but either cease to be dysregulated or do not significantly increase in dysregulation as the tissue approaches tumor status.

Discussion

The genes identified in the analysis are unique not only because they exhibit abnormal expression levels in colorectal cancer tissue, but also because they play vital or supporting roles in mitochondrial function. Although the small sample size limits the generalizability of the findings, they are nevertheless an important starting point. As more research is conducted surrounding the theory that some cancer may originate from metabolic mitochondrial dysfunction, these genes may provide insight into the earliest stages of bioenergetic dysregulation.

From the findings, some of the genes can be grouped by functionality or family. Although not all of the roles of the dysregulated genes were explored, certain genes stood out because of the continuity or frequency of their dysregulation. For example, *AKR1B10* appeared to be downregulated in all three datasets at the tumor stage, and displayed early underexpression in the polyp stage in Datasets 2 and 3. *AKR1B10* is a gene that is normally highly expressed in stomach and intestinal epithelial cells (Endo et al., 2021). It has been known to be downregulated in gastrointestinal cancers and upregulated in hepatocellular cancer (Endo et al., 2021). *AKR1B10* is normally involved in regulating metabolism and breaking down toxic compounds that can inflict cellular damage (Endo et al., 2021). The early downregulation of *AKR1B10* in polyps suggests that intestinal cells may be incurring more damage than usual. Whether that unchecked toxicity is the byproduct or source of the carcinogenesis is unclear.

Another gene family that appeared frequently dysregulated was the *SLC25* family. This group of genes is crucial to the transportation of multiple types of substrates across the mitochondrial inner membrane (Ruprecht & Kunji, 2020). Members of the *SLC25* family exhibited up- and downregulation in early polyp stages in all three datasets. Furthermore, two *SLC25* members continued to be under- and overexpressed in the tumor stage in Datasets 1 and 2 respectively. The dysregulation of this gene type indicates that there may be early changes to substrate traffic through the mitochondria, which likely impacts mitochondrial resource

allocation. The fact that more *SLC25* gene members were dysregulated in the polyp stage than the tumor stage may imply that altered substrate traffic through the mitochondria proceeds tumorigenesis. A deeper analysis should be conducted into the molecular dynamics of *SLC25* genes and their downstream effects.

Additionally, the *HIGD1A* gene appeared to be downregulated in all three datasets at the polyp stage. The gene was also underexpressed in Dataset 1 and Dataset 3 in the tumor stage. *HIGD1A* encodes for a mitochondrial inner membrane protein that is important in maintaining metabolic homeostasis and promoting cellular adaptation and survival under hypoxic conditions (Zhu et al., 2022). The protein is also involved in cell survival and anti-apoptotic processes (Zhu et al., 2022). Past research suggests that abnormally expressed *HIGD1A* behaves differently in various tumors (Xu et al., 2021). In colorectal cancer, the downregulation of *HIGD1A* contributes to increased cell proliferation and apoptosis inhibition (Xu et al., 2021). Thus, the finding that *HIGD1A* was underexpressed corroborates prior studies. Furthermore, the observed polyp stage downregulation indicates early changes to mitochondrial respiration, potentially stemming from shifts in oxygen consumption, disruptions in the electron transport chain, or altered ROS levels and energy metabolism.

In a different vein, comparing gene expression between studies may shed light on broader trends in diet and colorectal cancer. Dataset 1 and Dataset 3 shared more overlapping dysregulated gene expression. Notably, these two datasets used tissues collected from Asian populations—Japan and China—which historically have lower CRC rates (Wong et al., 2019). Dataset 2, however, was conducted using tissue samples from Finnish and Spanish populations. While it is not entirely clear what drove this discrepancy, one possible explanation could be differences in diet, due to the implications that food may have on mitochondrial functioning. Epidemiological research suggests that dietary factors such as red meat, whole grain, calcium, and fiber intake are associated with regional differences of CRC risk in Europe and Asia (Yeo et al., 2026). These dietary elements have diverse chemical makeups, and trends in diet may skew the kinds of molecules the mitochondria interacts with to make energy. Thus, diet and mitochondrial function are linked, with similarities in abnormal polyp stage gene expression as a potential indicator of this dynamic.

The findings of this analysis confirm the abnormal expression of previously observed oncogenic genes, but also shed new light on the early dysregulation of those genes in the polyp stage before tissue becomes cancerous. Furthermore, the findings may implicate the role that diet has on the dysregulation of specific genes. In the future, further investigation into the complete set of dysregulated genes, their biochemistry, and their role in the mitochondria should be conducted, and the connection between food and early stage gene disruption should be explored in greater depth.

References

Danesova, N., Horak, J., Valickova, A., Gil-Korilis, A., Ergui-Arbizu, J., Palek, R., Bruha, J., Levy, M., Skrobanek, P., Kral, J., Jungwirth, J., Neuzil, J., Vymetalkova, V., Vodicka, P., Vodenkova, S., & Tomasova, K. (2025). Dysregulated mitochondrial homeostasis and DNA repair in the progression from colon adenoma to cancer. *Molecular Medicine*, *31*(1), 341. https://doi.org/10.1186/s10020-025-01400-5

Endo, S., Matsunaga, T., & Nishinaka, T. (2021). The Role of AKR1B10 in Physiology and Pathophysiology. *Metabolites*, *11*(6), 332. https://doi.org/10.3390/metabo11060332

Esteban-Gil, A., Pérez-Sanz, F., García-Solano, J., Alburquerque-González, B., Parreño-González, M. A., del Carmen Legaz-García, M., Fernández-Breis, J. T., Rodriguez-Braun, E., Pimentel, P., Tuomisto, A., Mäkinen, M., Slaby, O., & Conesa-Zamora, P. (2019). *Differences in expression profiling between tumoral colorectal carcinomas subsets and normal tissue and polyps [mRNA]* [Data set]. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128435

Offei, N. A., & Cluntun, A. A. (2026). Lactate biology: Subcellular routing and chemical form define function. *Biochemistry, 65*(14), 2191–2201. https://doi.org/10.1021/acs.biochem.6c00251

Ruprecht, J. J., & Kunji, E. R. S. (2020). The SLC25 Mitochondrial Carrier Family: Structure and Mechanism. *Trends in biochemical sciences*, *45*(3), 244–258. https://doi.org/10.1016/j.tibs.2019.11.001

San-Millán, I. (2023). The Key Role of Mitochondrial Function in Health and Disease. *Antioxidants*, *12*(4), 782. https://doi.org/10.3390/antiox12040782

Satoh, K., Yachida, S., Sugimoto, M., Oshima, M., Nakagawa, T., Akamoto, S., Okano, K., Suzuki, Y., & Soga, T. (2017). *DNA microarray analysis of human colorectal cancer specimens* [Data set]. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89076

Seyfried T. N. (2015). Cancer as a mitochondrial metabolic disease. Frontiers in cell and developmental biology, 3, 43. https://doi.org/10.3389/fcell.2015.00043

Shi, X., Zhang, Y., Cao, B., Lu, N., Feng, L., Di, X., Han, N., Luo, C., Wang, G., Cheng, S., & Zhang, K. (2015). *Gene expression profiling of colorectal normal mucosa, adenoma and adenocarcinoma tissues* [Data set]. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41657

Siegel, R. L., Wagle, N. S., Star, J., Kratzer, T. B., Smith, R. A., & Jemal, A. (2026). Colorectal cancer statistics, 2026. CA: a cancer journal for clinicians, 76(2), e70067. https://doi.org/10.3322/caac.70067

Tomasetti, C., & Vogelstein, B. (2015). Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science (New York, N.Y.)*, *347*(6217), 78–81. https://doi.org/10.1126/science.1260825

Wang, Y., & Patti, G. J. (2023). The Warburg effect: a signature of mitochondrial overload. *Trends in cell biology*, *33*(12), 1014–1020. https://doi.org/10.1016/j.tcb.2023.03.013

Wong, M. C., Ding, H., Wang, J., Chan, P. S., & Huang, J. (2019). Prevalence and risk factors of colorectal cancer in Asia. *Intestinal research*, *17*(3), 317–329. https://doi.org/10.5217/ir.2019.00021

Xu, Z., Sun, J., Mao, Y., Chen, Y., Zhang, T., Qin, Y., & Hua, D. (2021). HIG1 domain family member 1A disrupts proliferation, migration, and invasion of colon adenocarcinoma cells. *Bioengineered*, *12*(2), 10501–10511. https://doi.org/10.1080/21655979.2021.1999368

Yeo, D., Hwang, S. H., Lee, S., Jeong, J., Hong, S., Kim, J., Kim, T. H., Hwang, J., Lee, H., Lee, J., Jung, J., & Yon, D. K. (2026). Global, regional, and national burden of colorectal cancer attributable to dietary factors, 1990 to 2021: A global burden of disease study 2021. *Medicine*, *105*(9), e47700. https://doi.org/10.1097/MD.0000000000047700

Zhu, J.-Y., Chen, M., Mu, W.-J., Luo, H.-Y., & Guo, L. (2022). The functional role of Higd1a in mitochondrial homeostasis and in multiple disease processes. Genes, 10(5), 1833–1845. https://doi.org/10.1016/j.gendis.2022.04.005