Impact of mini-driver genes in the prognosis and tumor features of colorectal cancer samples: a novel perspective to support current biomarkers

Anthony Vladimir Campos Segura[1,2,3], Mariana Belén Velásquez Sotomayor[3,4], Ana Isabel Flor Gutiérrez Román[1,2], César Alexander Ortiz Rojas[3,5] and Alexis Germán Murillo Carrasco[3,6]

[1] Biochemistry and Molecular Biology Research Laboratory. Faculty of Natural Sciences and Mathematics, Universidad Nacional Federico Villarreal, Lima, Peru
[2] Research Group in Biochemistry and Synthetic Biology (GIBBS-UNFV), Lima, Peru
[3] Research Group in Immunology and Cancer (IMMUCA), Lima, Peru
[4] School of Human Medicine, Faculty of Health Sciences, Universidad Científica del Sur, Lima, Peru
[5] Hematology Division, LIM31, Medical School, Universidade de São Paulo, Sao Paulo, Brazil
[6] Centro de Investigação Translacional em Oncologia (LIM24), Departamento de Radiologia e Oncologia, Faculdade de Medicina, Universidade de São Paulo, Sao Paulo, Brazil

ABSTRACT

Background: Colorectal cancer (CRC) is the second leading cause of cancer-related deaths, and its development is associated with the gains and/or losses of genetic material, which leads to the emergence of main driver genes with higher mutational frequency. In addition, there are other genes with mutations that have weak tumor-promoting effects, known as mini-drivers, which could aggravate the development of oncogenesis when they occur together. The aim of our work was to use computer analysis to explore the survival impact, frequency, and incidence of mutations of possible mini-driver genes to be used for the prognosis of CRC.

Methods: We retrieved data from three sources of CRC samples using the cBioPortal platform and analyzed the mutational frequency to exclude genes with driver features and those mutated in less than 5% of the original cohort. We also observed that the mutational profile of these mini-driver candidates is associated with variations in the expression levels. The candidate genes obtained were subjected to Kaplan–Meier curve analysis, making a comparison between mutated and wild-type samples for each gene using a p-value threshold of 0.01.

Results: After gene filtering by mutational frequency, we obtained 159 genes of which 60 were associated with a high accumulation of total somatic mutations with Log_2 (fold change) > 2 and p values < 10^{-5}. In addition, these genes were enriched to oncogenic pathways such as epithelium-mesenchymal transition, hsa-miR-218-5p downregulation, and extracellular matrix organization. Our analysis identified five genes with possible implications as mini-drivers: *DOCK3, FN1, PAPPA2, DNAH11*, and *FBN2*. Furthermore, we evaluated a combined classification where CRC patients with at least one mutation in any of these genes were separated from the main cohort obtaining a p-value < 0.001 in the evaluation of CRC prognosis.

Conclusion: Our study suggests that the identification and incorporation of mini-driver genes in addition to known driver genes could enhance the accuracy of prognostic biomarkers for CRC.

INTRODUCTION

Colorectal cancer (CRC) is a significant global health problem, ranking as the third most diagnosed neoplasia and the second leading cause of cancer-related death worldwide (*Bray et al., 2018*; *Sung et al., 2021*). In 2020, there were approximately 1.93 million new cases of CRC, representing 10% of all cancer cases (*Sung et al., 2021*). Despite advances in detection and treatment, CRC incidence and mortality rates have increased by 7.2% from 2018 to 2020. Therefore, it is essential to continue research into the underlying mechanisms driving CRC progression (*Bray et al., 2018*; *Sung et al., 2021*).

KRAS, *NRAS*, *BRAF*, *TP53*, and *APC* are considered driver genes for CRC since a few pathogenic mutations in any of these genes are sufficient to develop a tumor (*Thierry et al., 2014*). However, defining a single group of genes as the drivers of all tumors is challenging and only explains a small number of cancer cases (*Thierry et al., 2014*). This challenge arises because the description of any genetic variation as a pathogenic mutation depends on factors such as its impact on the translated protein (missense, nonsense, *etc.*) or the number of nucleotides affected (single-nucleotide, insertion, deletion, *etc.*).

To differentiate deleterious variations from polymorphisms, current genetic definitions consider mutations to be any genetic variations with a reduced frequency (<1%) in a healthy population (*Al-Koofee & Mubarak, 2020*). However, these definitions do not account for the possibility of new transcriptional switches that can support novel consequences of previously known polymorphisms, as recently described in CRC (*Abdi, Latifi-Navid & Latifi-Navid, 2022*).

The identification of coding genes, lncRNA, circRNA, and miRNA through transcriptomic studies has broadened the concept of mini-driver genes to genes or genomic regions that may collectively be associated with poor prognosis in cancer (*Yang et al., 2018*; *Wu et al., 2020*). Nevertheless, considering the large group of genes associated with cancer features and prognosis whose function is not elucidated, it is essential to propose a strategy that includes all types of genetic variations as mutations for analyzing candidate genes capable of supporting current driver genes. To this end, we adopt the concept of mini-driver genes, which refers to low-frequency genetic alterations with a relatively weak tumor-promoting effect (*Castro-Giner, Ratcliffe & Tomlinson, 2015*).

There are established criteria for identifying mutated genes that may be considered mini-driver genes. Firstly, individual mutations in a mini-driver gene should provide a growth advantage for cancer cells compared to normal cells, although this is not necessarily critical for tumor development. Secondly, mini-driver genes must be present in

a small proportion of tumors. Thirdly, they may be present in subclones because they have a relatively weak selective advantage, resulting in less probability of selective sweeping. Fourthly, mini-driver genes should show parallel or convergent evolution between cancer subclones and between cancers of the same type. Finally, they should be involved in processes such as gene expression regulation, mRNA stability, transcriptional changes, DNA methylation, and other non-coding genomic features (*Castro-Giner, Ratcliffe & Tomlinson, 2015*; *van Ginkel, Tomlinson & Soriano, 2023*).

Mini-driver genes play a significant role in tumor diversification, but the mechanisms underlying their effects are not well understood. One possible mechanism is that mini-drivers may help to maintain tumor homeostasis by counteracting the deleterious effects of some passenger mutations (*Li & Thirumalai, 2016*; *Cuykendall, Rubin & Khurana, 2017*). In some cases, non-coding mutations could be called "mini-drivers" because they alter transcriptional regulation, mRNA translation and stability, splicing control, and chromatin structure, leading to altered gene expression that favors tumor progression (*Elliott & Larsson, 2021*). Another possible mechanism by which mini-driver genes contribute to tumor development is through working together with driver genes/mutations. Large-scale genome-wide association studies (GWAS) have shown that even the most significant loci explain only a fraction of the predicted genetic variation for typical traits (*Boyle, Li & Pritchard, 2017*). Therefore, mini-driver genes may explain how polygenic effects provide a means by which heterogeneous mutation patterns can generate distinctive changes consistent with the phenotype observed in tumors (*Bennett et al., 2018*).

Clinical studies have identified mini-driver genes as prognostic biomarkers in cancer (*Bennett et al., 2018*) and our group evaluated mini-driver features in selected genes (*Campos Segura, 2022*) to focus on their contribution to CRC progression. In this study, we propose a strategy to identify potential mini-driver genes in CRC using Next Generation Sequencing (NGS) data and determine whether they could serve as prognostic markers, providing insights into their role in CRC progression.

MATERIALS AND METHODS

Database filtering and mutational frequency analysis

We retrieved data from Next Generation Sequencing (NGS) experiments and clinical information of colorectal cancer (CRC) patients using the cBioPortal platform (https://www.cbioportal.org/) ("cBioPortal for Cancer Genomics"; *Cerami et al., 2012*). Genomic data was only utilized to provide mutational information, while transcriptomic data, when available, was used to obtain mutational status, prediction of copy number alterations, and gene expression levels. Three cohorts of colorectal adenocarcinoma were selected for analysis, including the Dana-Farber Cancer Institute (DFCI) cohort ($n = 619$) (*Giannakis et al., 2016*), the Pan-Cancer Atlas from The Cancer Genome Atlas (TCGA) cohort ($n = 594$) (*Liu et al., 2018*), and the Memorial Sloan Kettering—Cancer Center (MSKCC) cohort ($n = 138$) (*Brannon et al., 2014*) datasets. All data sets were generated using Illumina HiSeq sequencers, including information for 18,215 genes. Clinical data are summarized in Table 1.

Table 1 Clinical and biological characteristics in colorectal cancer.

Variable		Study participants (*n* = 1,351)
Age (average ± standard deviation)		68.42 ± 11.14
Sex, *n* (%)	Female	660 (48.9)
	Male	551 (40.8)
	N/A	140 (10.4)
CIMP category, *n* (%)	CIMP-high	95 (7.0)
	CIMP-0/low	405 (30.0)
	N/A	851 (63.0)
MSI status, *n* (%)	MSI-high	91 (6.8)
	MSS	438 (32.4)
	N/A	822 (60.8)
Tumor stage, *n* (%)	I	152 (11.3)
	II	187 (13.8)
	III	159 (11.8)
	IV	65 (4.8)
	N/A	788 (58.3)
Tumor site, *n* (%)	Rectum	137 (10.2)
	Cecum	114 (8.4)
	Ascending to transverse colon	201 (14.8)
	Splenic flexure to sigmoid colon	166 (12.3)
	N/A	733 (54.3)

Note:
N/A, No Available; CIMP, CpG island methylator phenotype; MSI, Microsatellite Instability.

Characterization of somatic mutational profile in CRC

The cBioPortal platform was used to generate a report that displays the number of genetic variants detected in each gene, the number of patients with at least one variant per gene, and the total number of patients with available mutational information per gene. For this study, we included all somatic genetic variations, encompassing both SNPs (≥1% frequency in populations) and pathological mutations (<1% frequency in populations), as mutations. We then calculated the mean number of mutations per gene and patient (Table S1).

Using these values, we established four groups of genes. For that, we consider the power of the sample (number of patients), the percentage of participating genes, and the ranking of common driver genes in CRC. As result, we classified the genes as rarely mutated (<7%), low-mutated levels (7–10%), moderately mutated (11–50%), and highly mutated (>50%) genes. To visualize these results, we designed a scatterplot using the R software v.4.2.0 (*R Core Team, 2022*) with the ggplot2 package (*Wickham, 2011*).

Association between mutational status and tumor mutational burden

After filtering the prior data, we selected a putative group carrying mini-driver genes, consisting of low-mutated genes in patients (7% and 10%). Based on these genes, we compared the total number of variations (tumor mutational burden, TMB) per patient

according to their mutational status per gene. In our comparison, we considered any patient with at least one somatic variation in a specific gene as "mutated". To compare TMB, we utilized the Mann–Whitney test and calculated the fold change between mutated and wild-type groups per gene. We then adjusted the p-values using the Benjamini–Hochberg (BH) method (*Benjamini & Hochberg, 1995*). Finally, we established a mean-value threshold (10^{-5}) for selecting the top genes whose mutational status was associated with changes in TMB. To visualize our findings, we plotted the results using the Enhanced Volcano package in the R software ("EnhancedVolcano"; *Blighe, Rana & Lewis 2018*).

Association between mutational status and gene expression/copy number variations

We investigated the expression levels of genes with mutational status and copy number variations. To accomplish this, we used the TCGAretriever package ("TCGAretriever: Retrieve Genomic and Clinical Data from TCGA"; *Fantini, 2019*) to verify the normalized expression levels of genes and putative copy number alterations predicted by Gistic (*Mermel et al., 2011*) in colon and rectal adenocarcinoma samples. We applied the same criteria as in the previous step to dichotomize the patient population for each selected gene (mutated or wild-type) and compared the expression levels of genes or copy number alterations. We then used the Mann–Whitney test to compare these values, taking into account the distribution of expression levels. Finally, we represented all samples in boxplots for each pair of genes. The Y-axis displays expression levels or the number of altered copies of the first gene, while the X-axis shows patient categories based on the mutational status of the second gene.

Gene enrichment

Gene lists were obtained after applying the Enrichr tool (*Xie et al., 2021*). We used all available databases from Transcription (17 databases), Ontology (25 databases), and Pathway (eight pathways) modules. After loading gene lists, we only considered relevant pathways where more than three genes were associated with $p < 0.001$.

Survival analysis

To investigate the prognostic value of the mini-drivers, we employed Kaplan–Meier curves and Cox proportional hazard regression analysis (*Borgan Rnulf, 2001*). Then, we identified the mini-drivers that potentially play a role in disease progression, including *DOCK3*, *FN1*, *PAPPA2*, *DNAH11*, and *FBN2*. Then, we established a gene panel based on these genes. Cox regression analysis was performed, including age (as continuous), sex, and tumor stages (AJCC, American Joint Committee on Cancer) as potential confounding variables. To perform these analyses, we utilized the survival and survminer packages of R software version 4.2.0 (*R Core Team, 2022*).
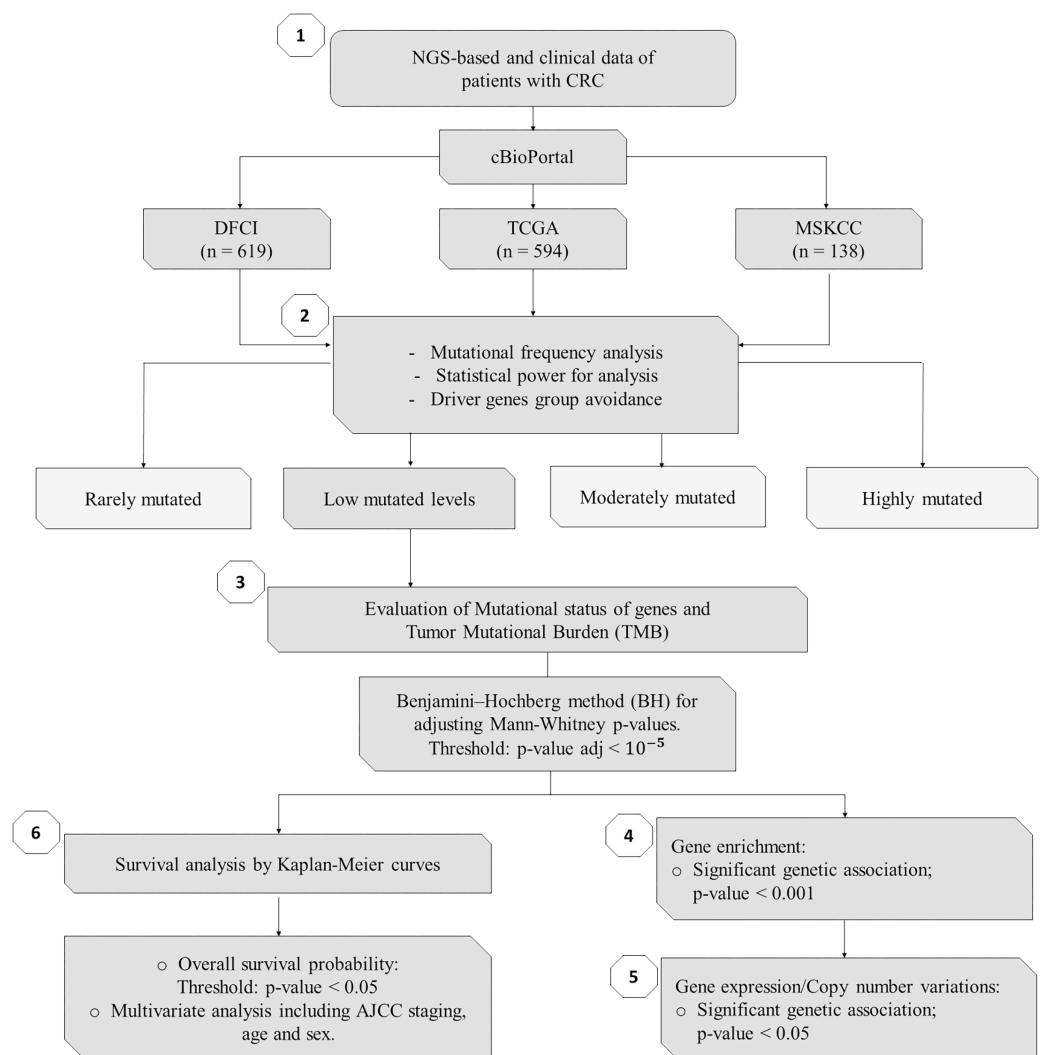
Figure 1 Flowchart of methods of this study. NGS, Next-Generation Sequencing; CRC, Colorectal Cancer; DFCI, Dana-Farber Cancer Institute; TCGA, The Cancer Genome Atlas; MSKCC, Memorial Sloan Kettering Cancer Center; *n*, number of samples.　Full-size 🔍 DOI: 10.7717/peerj.15410/fig-1

RESULTS

Identifying mini-drivers in CRC based on the mutational frequency

As summarized in Fig. 1, to identify potential mini-drivers in CRC, we analyzed the gene mutation frequency, and the statistical power. Then, we classified the genes into four groups (Table S1, Fig. 2): rarely mutated (≤7%, 17,993 genes), lowly mutated (7–10%, 159 genes), moderately mutated (11–50%, 62 genes), and highly mutated (>50%, two genes). The rarely mutated group includes genes with less than one variation in at least 59 patients, which represents less than 5% of the patients with available information for mutational profiling (980 patients from three cohorts with available information). To avoid non-representative results in further analyses, we excluded this group of genes.

The moderately and highly mutated groups include genes considered as drivers, such as *APC* with 66%, *TP53* with 56%, *KRAS* with 35%, *BRAF* with 16%, and *ARID1A* with 11%,
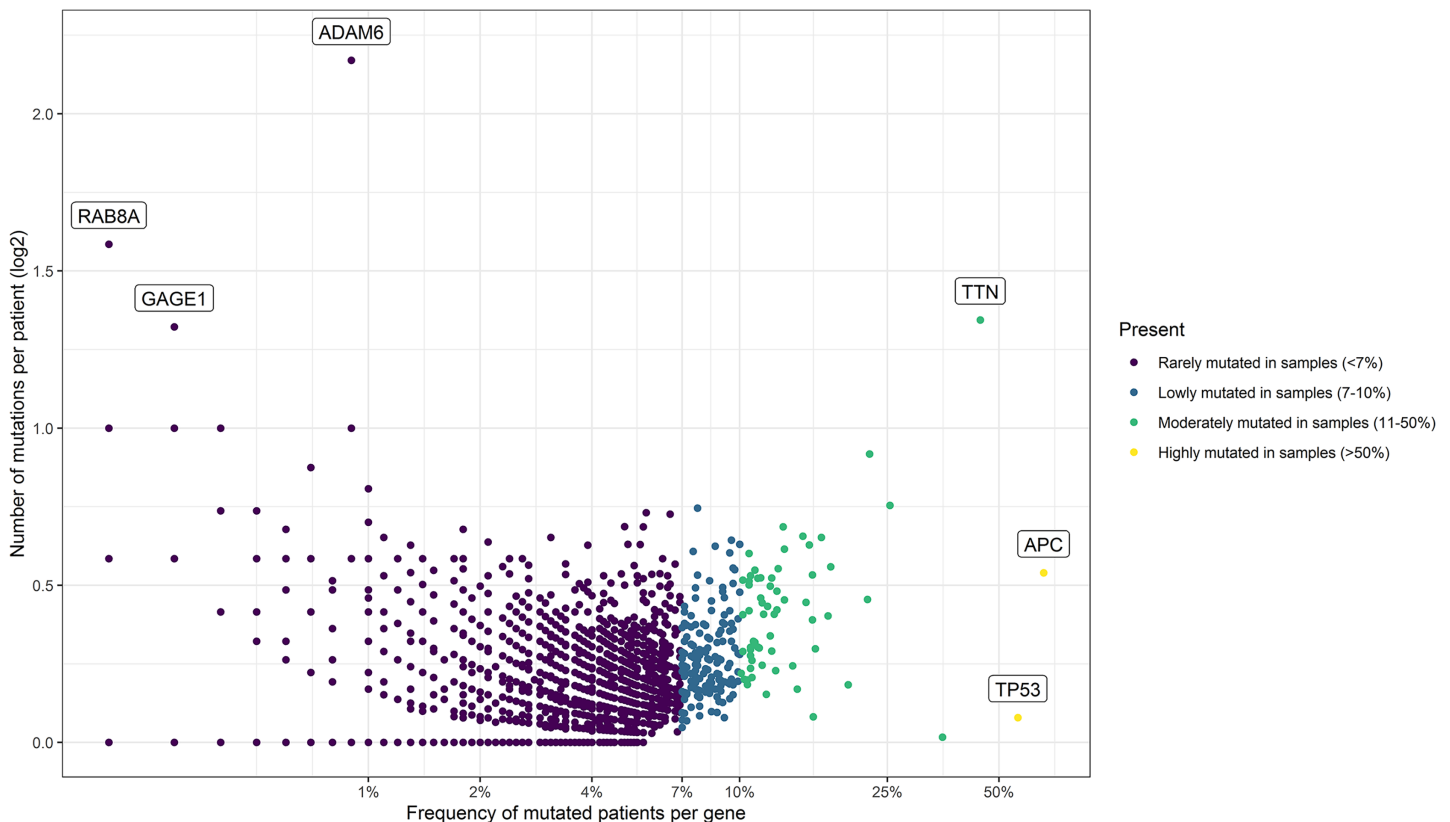
Figure 2 Mutational frequency present in the samples obtained from the three cohorts: DCFI, TCGA, and MSKCC. Based on the number of mutations per patient we classify the mutated genes into four groups: rarely mutated (<7%), low mutated levels (7–10%), moderately mutated (11–50%), and highly mutated (>50%). Name labels were added to recognize prominent genes with a high number of mutations per patient or a high percentage of affected patients. Full-size ◩ DOI: 10.7717/peerj.15410/fig-2

among others. We focused on exploring the genes in the lowly mutated group (159 genes) as potential mini-drivers. We analyzed these genes with the Enrichr database and found that they are mainly associated with hsa-miR-218-5p regulation (miRTarBase 2017 database, adjusted p-value = 1.9×10^{-4}), extracellular matrix organization (Reactome 2022 database, adjusted p-value = 9.1×10^{-11}), and epithelial-mesenchymal transition (EMT, MsigDB Hallmark 2020 database, adjusted p-value = 2.5×10^{-8}).

Putative mini-driver genes are associated with high mutation rates

In order to assess the impact of mini-drivers on tumoral progression, we analyzed the polygenic effect of 159 genes selected from the previous step. Our goal was to identify which of these genes were associated with high mutation rates. As shown in Fig. 3, we found that 60 genes had a significantly higher TMB when mutated compared to their respective wild-type group, with an increase ranging from 5.4–8.7-fold (p-value < 10^{-5}). Among the most statistically significant genes associated with high TMB were *MUC5B*, *DNAH7*, *DOCK3*, and *BMPR2*. Meanwhile, we identified *SYNE2*, *COL7A1*, *NOTCH3*, and *SPEG* as the genes with the highest mutation counts.
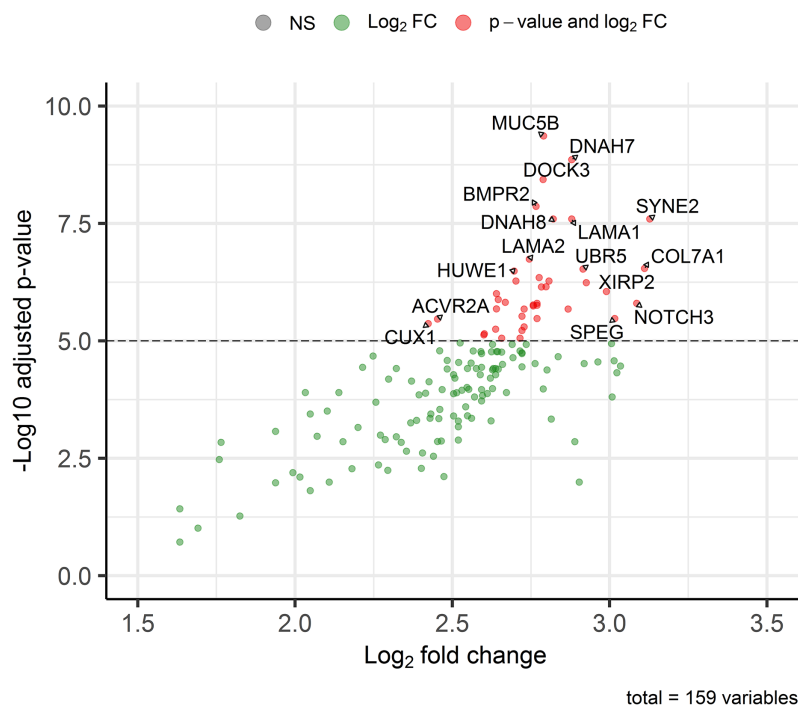
Figure 3 **Association of the mutational status of genes and the total mutation count in CRC patients.** Scatter plot showing $\log_2$Fold Change and adjusted p-values between the number of mutations in patients classified as mutated or wild-type for each gene. The p-value threshold was set at 10^{-5} to show strongly associated genes with high mutation rates in CRC patients. Full-size 🔺 DOI: 10.7717/peerj.15410/fig-3

Mini-drivers are associated with specific gene signatures in CRC

After identifying the genes associated with high TMB (Fig. 3), we compared the expression levels and copy number alterations with the mutational status of these genes. We discovered 46 genes whose mutational status was linked to changes in the expression levels of other genes ($p < 0.05$, Table S2). Remarkably, the mutated group of the *BMPR2* gene alters the expression of nine genes (*ACVR2A, DOCK3, MUC5B, FLNA, DNAH8, SYNE2, FBN2, ITPR3,* and *DLC1*), whereas mutated groups of *FLNB, ACVR2A, SIPA1L3,* and *CELSR1* genes altered the levels of other six genes. Table 2 provides a summary of the 43 relevant gene pairs ($\log_{10}(\text{FC}) > 0.4$, $p < 0.01$), while Fig. 4 displays a selection of these findings. Among them, we found that the subexpression of *MYH11* was related to its mutational profile ($p = 0.0043$), *UBR5* levels were decreased in patients with mutations in *FLNB* ($p = 0.0012$) or *COL7A1* ($p = 0.0075$), whereas *FBN3* mutated was associated with low expression of *FBN2* ($p = 0.0079$). Additionally, we observed reduced expression levels of *ZNF536* and *CUX1* genes in samples mutated for *FLNB* and *TMEM132D* genes, respectively ($p < 0.008$). Furthermore, we compared the mutational status of these 60 genes with copy number variations (CNVs). Specifically, we compared the number of altered copies (gains or losses) between mutated and WT groups for these genes. We identified 12 genes whose mutational status was linked to copy number alterations in other genes

Table 2 Comparison of expression level and mutational status for putative mini-driver genes.

Gene (expression levels)	Gene (mutational status)	p-value	FC (log$_{10}$)	Gene (expression levels)	Gene (mutational status)	p-value	FC (log$_{10}$)
CUX1	TMEM132D	0.0005	−0.406	ACVR2A	BMPR2	0.0028	0.567
MYH11	MYH11	0.0043	−0.44	DOCK3	XIRP2	0.0028	0.58
SYNE2	LAMA2	0.0056	0.429	MUC5B	COL7A1	0.0028	1.49
UBR5	COL7A1	0.006	−0.456	DOCK3	SZT2	0.0027	1.143
ZNF536	FLNB	0.0075	−0.484	MUC5B	LAMA2	0.0027	1.102
FLNA	ACVR2A	0.0075	0.463	MUC5B	ATP10A	0.0024	1.015
FBN2	FBN3	0.0079	−0.4	ACVR2A	ACVR2A	0.0022	0.517
DNAH7	PCNT	0.009	0.423	MUC5B	ACVR2A	0.0016	1.286
SYNE2	TRIOBP	0.009	0.402	MUC5B	ADGRV1	0.0013	1.003
SYNE2	SRRM2	0.0099	0.434	UBR5	FLNB	0.0012	−0.538
DOCK3	TCHH	0.0095	0.608	DOCK3	FLNB	0.0008	0.917
DOCK3	CELSR1	0.0085	0.846	MUC5B	WDFY3	0.0007	1.077
MUC5B	XIRP2	0.0084	1.107	DOCK3	MYO15A	0.0006	0.979
DOCK3	ADGRV1	0.0081	0.541	MUC5B	MYO15A	0.0005	1.442
DOCK3	DNAH1	0.0074	0.608	MUC5B	ASPM	0.0004	1.263
SYNE2	HSPG2	0.0074	0.611	MUC5B	TENM2	0.0003	1.271
MUC5B	RP1L1	0.0074	1.426	MUC5B	KIAA1109	0.0003	1.013
ACVR2A	SIPA1L3	0.0072	0.5	SYNE2	ASPM	0.0003	0.59
DOCK3	POLE	0.0066	0.726	MUC5B	FLNB	0.0002	1.51
MUC5B	BMPR2	0.004	1.559	MUC5B	DNAH7	0.0002	1.494
DOCK3	BMPR2	0.0039	0.903	MUC5B	POLE	0.0001	1.373
DOCK3	TRPM3	0.0039	0.515				

Note:
Genes shown log$_{10}$(FC) > 0.4 and p < 0.01. FC, Fold change.

(p < 0.05, Table S3). Interestingly, the mutated *SLITRK5* gene cluster was related to copy number alterations in five genes (*NOTCH3*, *FBN3*, *SIPA1L3*, *PTPRS*, and *CUX1*), particularly in *NOTCH3* and *FBN3*, where the *SLITRK5* mutated group showed large ratios (greater than 4) compared to WT patients (Fig. 5).

Overall, our results demonstrate that *MUC5B*, *TMEM132D*, *SYNE2*, *POLE*, *CUX1*, and *NOTCH3* play critical roles as effectors with altered expression levels (Table S2) or copy number variations (Table S3) in the context of mini-driver genes.

Mini-driver genes contribute to CRC prognosis

We used the 60 genes obtained to analyze their performance as biomarkers for survival in CRC (Fig. 3). We applied this analysis to 201 CRC patients that had complete information on overall survival rates and mutational profiles of the 60 genes. The mutational status of five genes, *DOCK3, FN1, PAPPA2, DNAH11*, and *FBN2*, were associated (p < 0.05) with poor survival in CRC patients (Fig. 6A). We observed a survival rate of ~25% or less for the mutated group, while the wild-type (WT) group retained more than 75% of survival after 5 years of follow-up. Then, when we constructed a panel with these genes, we observed an
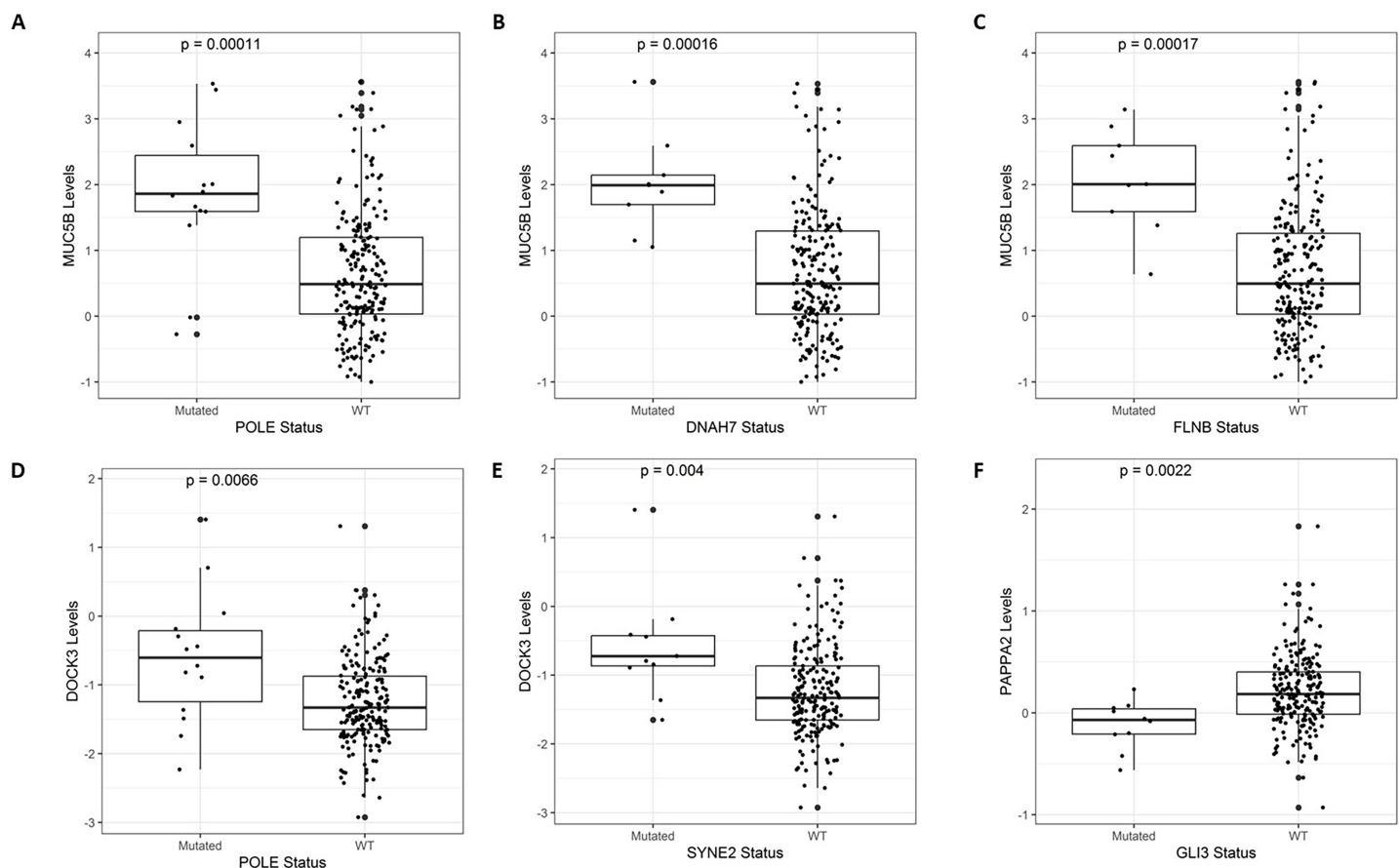
Figure 4 A selection of genes whose mutational status was associated with alterations in expression levels. (A) *MUC5B vs. POLE* ($p = 0.00011$). (B) *MUC5B vs. DNAH7* ($p = 0.00016$). (C) *MUC5B vs. FLNB* ($p = 0.00017$). (D) *DOCK3 vs. POLE* ($p = 0.0066$). (E) *DOCK3 vs. SYNE2* ($p = 0.0056$). (F) *PAPPA2 vs. GLI3* ($p = 0.0022$).

improvement in the survival rates differences between both groups (median $OS_{mutated} = 38.5$ months *vs.* median OS_{WT} = Not reached; Log-rank, $p < 0.0001$; Fig. 6B). Interestingly, we were able to identify a gene expression signature constituted of 16 genes that characterize the mutated group ($p < 0.05$ and absolute FC > 1.5). Between these genes, *DOCK3, FN1, ADAMTS2, AHNAK, AHNAK2, DNAH7, NBEA, SACS, SMAD4*, and *VWF*, were upregulated in the mutated group; whereas *AMER1, DIDO1, LRP1, LRP1B, RNF43*, and *TG*, were downregulated (Fig. S1). Finally, we tested the ability to predict the outcome of our mini-driver gene panel, in the presence of possible confounding variables (Fig. 6C). Our multivariate Cox regression demonstrated that our gene panel maintains its usefulness in predicting prognosis independent of age, sex, and AJCC tumoral staging (HR = 2.92, $p = 0.002$).

Finally, Fig. 7 presents a visual abstract to summarize all our findings in this study and how our strategy could propose mini-driver genes as an additional set of markers to support current driver genes in CRC.
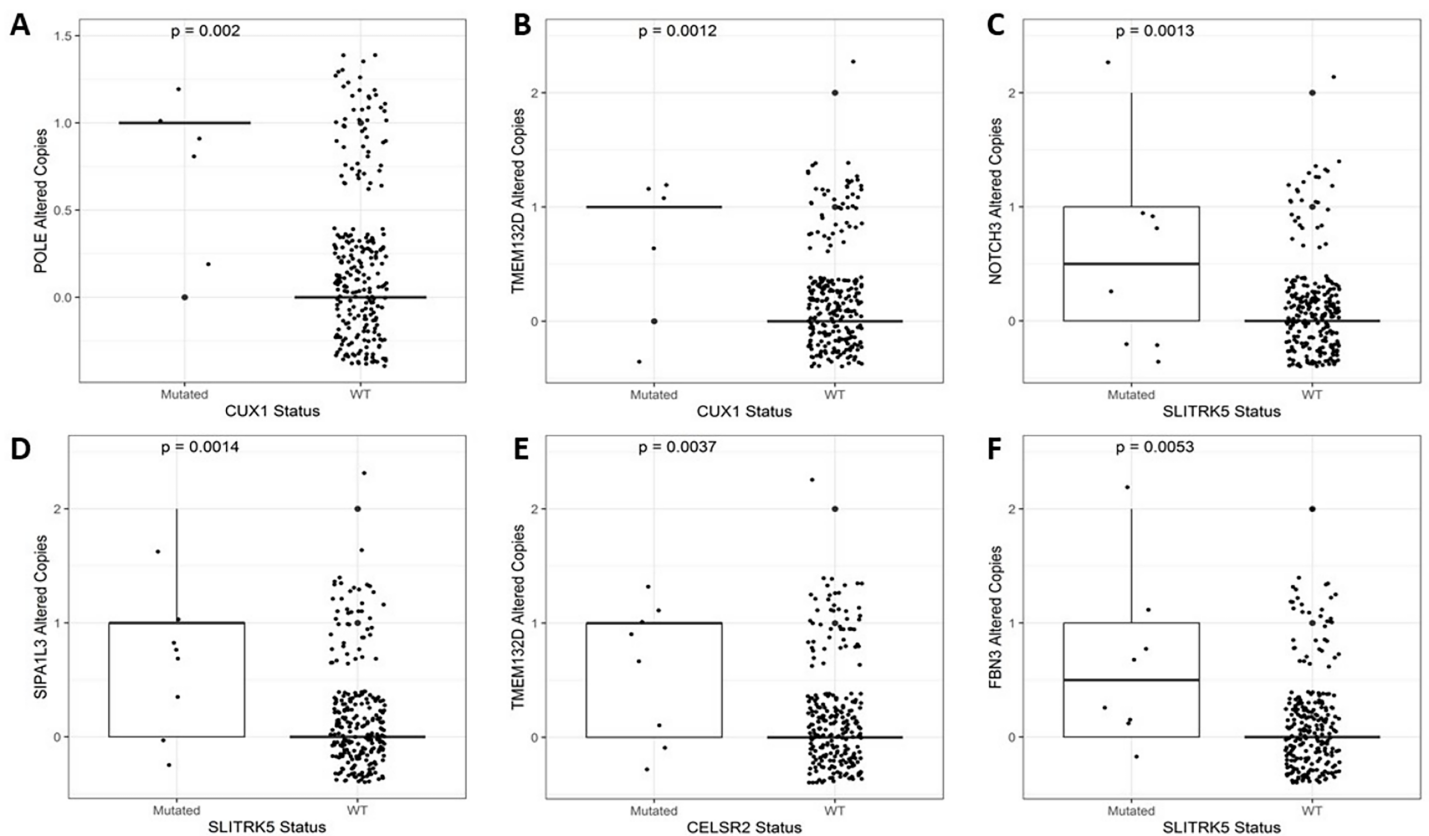
Figure 5 A selection of genes whose mutational status was associated with alterations in the number of copies. (A) *POLE vs. CUX1* ($p = 0.002$). (B) *TMEM132D vs. CUX1* ($p = 0.0012$). (C) *NOTCH3 vs. SLITRK5* ($p = 0.0013$). (D) *SIPA1L3 vs. SLITRK5* ($p = 0.0014$). (E) *TMEM132D vs. CELSR2* ($p = 0.0037$). (F) *FBN3 vs. SLITRK5* ($p = 0.0053$). Full-size ☑ DOI: 10.7717/peerj.15410/fig-5

DISCUSSION

In this study, we utilized three reliable databases of colorectal cancer patients to propose novel perspectives on the analysis of putative mini-driver genes. These studies provide crucial and representative information on colorectal cancer. For instance, DFCI's study focused on molecular characterization utilizing whole exome sequencing (WES) to gather tumor genomic data alongside detailed pathological and clinical information (*Giannakis et al., 2016*). TCGA's comprehensive analysis using around 10,000 samples representing 33 cancer types was employed for our interest in gene expression analysis corresponding to CRC samples (*Liu et al., 2018*). Finally, MSKCC's study analyzed inter- and intratumoral heterogeneity as evidence in the development of CRC (*Brannon et al., 2014*).

As mini-driver genes have a low mutational frequency, their impact alone is insufficient to generate a significant advantage for tumor cells. *Badr et al. (2022)* state that multiple accumulated weak mutations can combine to form a polygenic conductor (main driver) with enough impact to modify cellular function and patient prognosis (*Badr et al., 2022*), as depicted in Fig. 6. Our findings indicate that patients with mutations in *DOCK3*, *FN1*, *PAPPA2*, *DNAH11*, and *FBN2* genes had a shorter survival rate compared to patients without mutations. These results suggest that these alterations lead to cell dysregulation, as
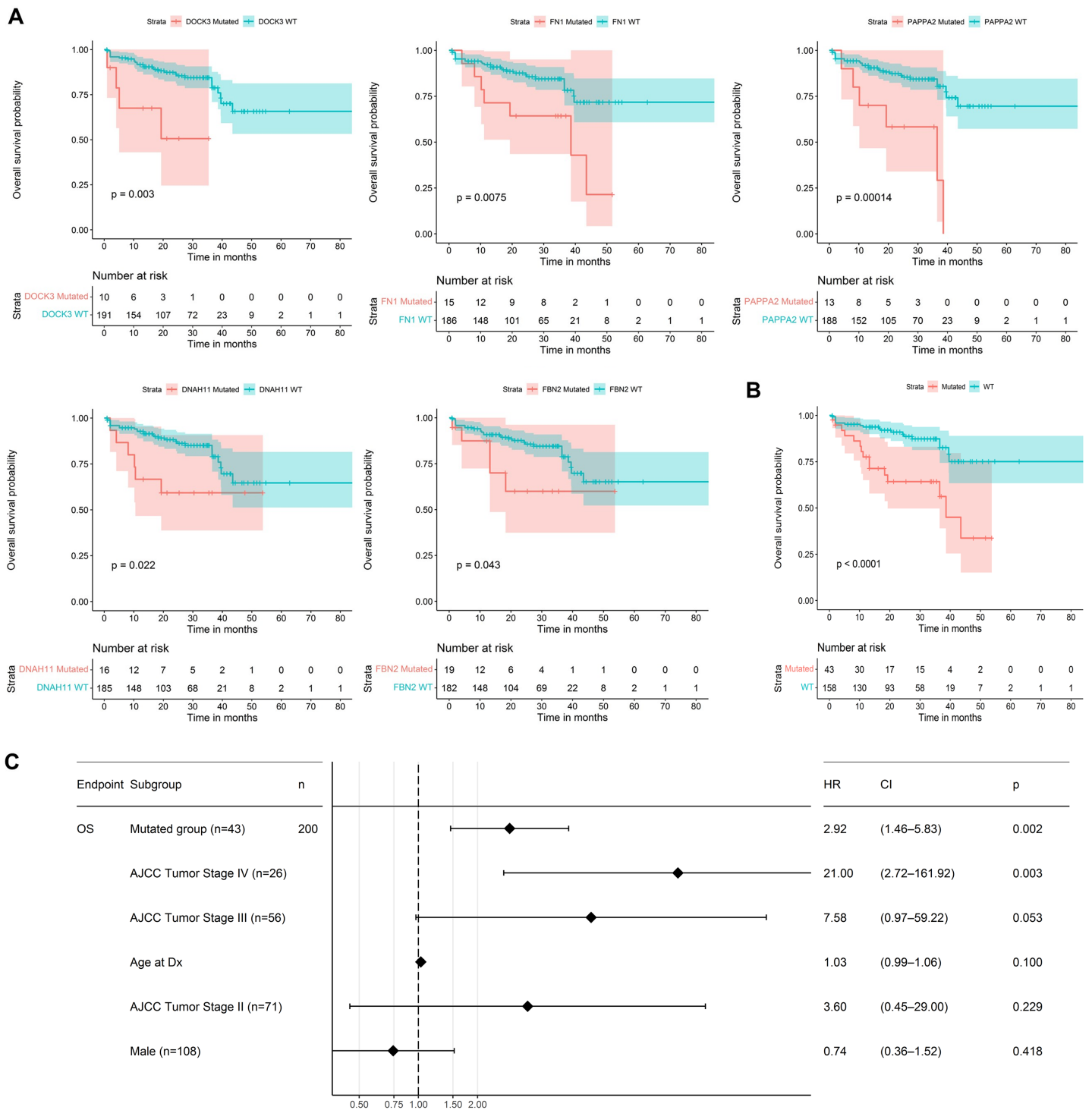
Figure 6 Prognosis analysis of five potential mini-drivers. (A) Kaplan–Meier curves analyses according to the mutational status of the *DOCK3, FN1, PAPPA2, DNAH11,* or *FBN2* genes. (B) Overall survival curve according to the combined mutational profile of *DOCK3, FN1, PAPPA2, DNAH11,* and *FBN2* genes. (C) Multivariate Cox regression demonstrating that the prognostic value of our gene panel is not included in variables commonly associated with the outcome. HR, Hazard ratio; CI, Confidence interval; *n*, number of samples.

Full-size ☑ DOI: 10.7717/peerj.15410/fig-6

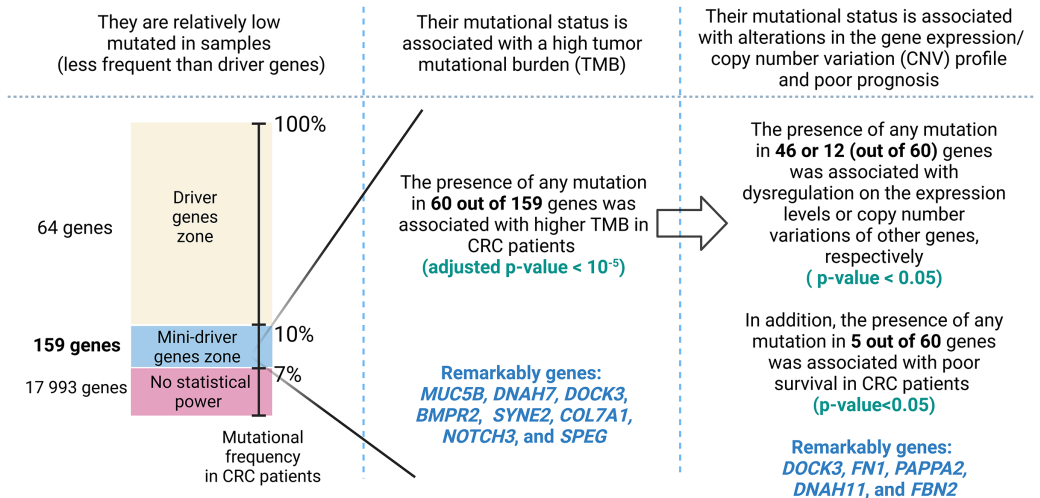**Features of Mini-driver genes in Colorectal Cancer Samples**

Figure 7 A visual abstract of the study. Since mutations in driver genes use canonical criteria to be characterized and could be not sufficient to explain all cancer cases, we evaluated a strategy for proposing additional genes using the mini-driver hypothesis. Full-size 🖼 DOI: 10.7717/peerj.15410/fig-7

seen in other cancer types (*Irmak-Yazicioglu, 2016*; *Wilk & Braun, 2018*; *Furuya et al., 2021*).

The *DOCK3* gene has been reported to participate in various processes related to invasion, migration, and metastasis in cancer cells (*Hofer et al., 2017*; *Kotelevets & Chastre, 2020*; *Lu et al., 2020*). However, there is limited research on the involvement of the *MUC5B* gene in CRC, with recent evidence showing high expression of this gene in elderly CRC patients, particularly in poorly differentiated tumors (*Iranmanesh et al., 2021*). Extracellular vesicles with high levels of phosphorylated and expressed *FN1* have been identified as potential prognostic factors and therapeutic targets in CRC (*Qi et al., 2020*; *Zheng et al., 2020*). It has been postulated that *FN1* could function as a promoter gene in non-canonical pathways of mini-driver genes and their mutations, with upregulation of *FN1* by the *HMGA2* gene contributing to a metastatic profile in CRC cells (*Wu et al., 2016*). Mutations in the *PAPPA2* have been associated with tumor progression and treatment of digestive tumors, although its role in CRC is not yet understood (*Miao et al., 2022*). Similarly, the rs2285947 polymorphism in the *DNAH11* gene has been linked to an increased risk of several cancer types, suggesting its potential contribution as a mini-driver gene in carcinogenesis (*Wang et al., 2015*).

The *FBN2* gene is hypermethylated in CRC tissue and serum samples from patients with CRC and liver metastases, and its expression is directly correlated with shorter survival rates in colon cancer patients, suggesting a possible role as a tumor suppressor gene (*Leygo et al., 2017*; *Wang et al., 2022*). Although methylation data for all genes in a representative number of CRC samples were not available for this study, we anticipate that upcoming omics data will provide more information on methylation levels in CRC patients, allowing

us to update this perspective and stimulate more NGS experiments for a more robust comparison.

Other genes, such as *NOTCH3* and *SLITRK5*, have clear contributions to tumor development. *NOTCH3* promotes tumor cell survival and proliferation, induces EMT and cancer stem cell (CSC) properties, and has been linked to various clinical and pathological features, including larger tumor size, advanced TNM stage, higher pathological grade, and tumor metastasis (*Pastò et al., 2014*; *Aburjania et al., 2018*; *Xiu et al., 2021*). Frequent genetic, epigenetic, and transcriptional changes have been observed in the *SLITRK5* gene in colorectal neoplasias (*Hesson et al., 2016*). Nevertheless, genes such as *CUX1* have controversial reports. It has been recently discovered that *CUX1* is a tumor suppressor paradoxically overexpressed in tumor cells (*Cancer Genome Atlas Network, 2012*; *Jo et al., 2017*; *Liu et al., 2020*). Overall, a mini-driver gene approach could be a useful tool to support further analysis (sense and antisense strands) of these controversial regions to understand their involvement in tumor growth.

According to *Dressler et al. (2022)*, when a driver gene mutates, it significantly impacts cancer cell growth, providing them with a fundamental advantage in their development (*Dressler et al., 2022*). However, our classification of patients as mutated, which includes at least one mutation in one of five genes, has resulted in decreased survival time. The universe of mutations present in these genes may have different pathways to favor tumor proliferation. Nonetheless, the classification of somatic mutations is affected by the initial analysis. Typically, researchers detect all somatic mutations but exclude those highly present in large populations (*Timmermann et al., 2010*; *Ma et al., 2020*). This can limit the analysis to a reduced number of targets (*Leedham & Tomlinson, 2012*; *Lee-Six et al., 2019*) based on the pathological effect related to variations rarely present in healthy individuals. However, in cancer, aberrant expression levels and unexpected pathways (*Hanahan, 2022*) may support new functions for these polymorphisms that are usually discounted when found as somatic variations.

Therefore, our study suggests analyzing all somatic mutations to assess cancer prognosis, combining traditionally evaluated driver genes and mutations with additional tumor-promoting regions (mini-drivers). We believe that even silent mutations may have additional functions related to the expression of non-coding RNA expressed from the same genomic locations. *Li & Thirumalai (2016)* argue that when the main drivers lose their mutagenic capacity, mini-drivers help normalize the physical condition advantage of the drivers. Additionally, mini-drivers could confer a physical fitness advantage on cancer cells, especially when they accumulate during tumor progression due to stochastic genetic mutations (*Li & Thirumalai, 2016*). To increase the number of therapy proposals and biomarkers discovered for colorectal cancer (CRC), we suggest exploring candidate mini-driver regions in a gene-panel strategy.

We used NGS-based information to analyze somatic mutational profiles from a broader perspective to evaluate a new strategy for predicting biomarkers for CRC. However, our study has limitations. We have no control over the data collection, information, and permissions declared by the patients. Nonetheless, the databases used are supported and validated by accredited public and private health institutions. Furthermore, many of the

proposed biomarkers for CRC have limited specificity and require further experimental research. Nevertheless, our methodology uses critical and rigorous settings to determine putative mini-driver genes, providing deeper insights into colorectal carcinogenesis to mitigate the risk of presenting biased information. Therefore, we expect that our findings will support future research to find prognostic markers for CRC using solid and liquid biopsies by testing panels of driver and mini-driver genes.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Anthony Vladimir Campos Segura conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Mariana Belén Velásquez Sotomayor analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Ana Isabel Flor Gutiérrez Román performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- César Alexander Ortiz Rojas analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Alexis Germán Murillo Carrasco conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:
The following datasets are available at cBioPortal (https://www.cbioportal.org/): Colorectal Adenocarcinoma (DFCI, Cell Reports 2016), Colorectal Adenocarcinoma (TCGA, PanCancer Atlas), and Colorectal Adenocarcinoma Triplets (MSK, Genome Biol 2014).
The mutations per gene and patient is available in the Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.15410#supplemental-information.

REFERENCES

Abdi E, Latifi-Navid S, Latifi-Navid H. 2022. Long noncoding RNA polymorphisms and colorectal cancer risk: progression and future perspectives. *Environmental and Molecular Mutagenesis* **63(2)**:98–112 DOI 10.1002/em.22477.

Aburjania Z, Jang S, Whitt J, Jaskula-Stzul R, Chen H, Bart Rose J. 2018. The role of *Notch3* in cancer. *The Oncologist* **23(8)**:900–911 DOI 10.1634/theoncologist.2017-0677.

Al-Koofee DAF, Mubarak SMH. 2020. Genetic polymorphisms. In: *The Recent Topics in Genetic Polymorphisms*. London: IntechOpen. *Available at https://www.intechopen.com/chapters/68164.*

Badr H, Blutrich R, Chan K, Tong J, Taylor P, Zhang W, Kafri R, Röst HL, Tsao M-S, Moran MF. 2022. Proteomic characterization of a candidate polygenic driver of metabolism in non-small cell lung cancer. *Journal of Molecular Biology* **434(13)**:167636 DOI 10.1016/j.jmb.2022.167636.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57(1)**:289–300 DOI 10.1111/j.2517-6161.1995.tb02031.x.

Bennett L, Howell M, Memon D, Smowton C, Zhou C, Miller CJ. 2018. Mutation pattern analysis reveals polygenic mini-drivers associated with relapse after surgery in lung adenocarcinoma. *Scientific Reports* **8(1)**:14830 DOI 10.1038/s41598-018-33276-3.

Blighe K, Rana S, Lewis M. 2018. EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. *Available at https://github.com/kevinblighe/EnhancedVolcano* (accessed 1 March 2023).

Borgan Rnulf. 2001. Modeling survival data: extending the cox model. In: Therneau TM, Grambsch PM, eds. *Statistics in Medicine.* Vol. 20. New York: Springer-Verlag, 2053–2054.

Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169(7)**:1177–1186 DOI 10.1016/j.cell.2017.05.038.

Brannon AR, Vakiani E, Sylvester BE, Scott SN, McDermott G, Shah RH, Kania K, Viale A, Oschwald DM, Vacic V, Emde A-K, Cercek A, Yaeger R, Kemeny NE, Saltz LB, Shia J, D'Angelica MI, Weiser MR, Solit DB, Berger MF. 2014. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biology* **15(8)**:454 DOI 10.1186/s13059-014-0454-7.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68(6)**:394–424 DOI 10.3322/caac.21492.

Campos Segura A. 2022. Análisis de mutaciones somáticas en SAMHD1, GNL1, POLE, MRE11 y ASXL2 como mini impulsores en adenocarcinoma colorrectal. Universidad Nacional Federico Villarreal. *Available at https://repositorio.unfv.edu.pe/handle/20.500.13084/6389.*

Cancer Genome Atlas Network. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487(7407)**:330–337 DOI 10.1038/nature11252.

Castro-Giner F, Ratcliffe P, Tomlinson I. 2015. The mini-driver model of polygenic cancer evolution. *Nature Reviews Cancer* **15(11)**:680–685 DOI 10.1038/nrc3999.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y. 2012. The cBio cancer genomics portal: an open platform for

exploring multidimensional cancer genomics data. *Cancer discovery* **2(5)**:401–404 DOI 10.1158/2159-8290.CD-12-0095.

Cuykendall TN, Rubin MA, Khurana E. 2017. Non-coding genetic variation in cancer. *Current Opinion in Systems Biology* **1**:9–15 DOI 10.1016/j.coisb.2016.12.017.

Dressler L, Bortolomeazzi M, Keddar MR, Misetic H, Sartini G, Acha-Sagredo A, Montorsi L, Wijewardhane N, Repana D, Nulsen J, Goldman J, Pollitt M, Davis P, Strange A, Ambrose K, Ciccarelli FD. 2022. Comparative assessment of genes driving cancer and somatic evolution in non-cancer tissues: an update of the Network of Cancer Genes (NCG) resource. *Genome Biology* **23(1)**:35 DOI 10.1186/s13059-022-02607-z.

Elliott K, Larsson E. 2021. Non-coding driver mutations in human cancer. *Nature Reviews Cancer* **21(8)**:500–509 DOI 10.1038/s41568-021-00371-z.

Fantini D. 2019. TCGAretriever: retrieve genomic and clinical data from TCGA. *Available at https://cran.r-project.org/web/packages/TCGAretriever/index.html* (accessed 1 March 2023).

Furuya TK, Murta CB, Murillo Carrasco AG, Uno M, Sichero L, Villa LL, Cardilli L, Coelho RF, Guglielmetti GB, Cordeiro MD, Leite KRM, Nahas WC, Chammas R, Pontes J Jr. 2021. Disruption of miRNA-mRNA networks defines novel molecular signatures for penile carcinogenesis. *Cancers* **13(19)**:4745 DOI 10.3390/cancers13194745.

Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, Bahl S, Cao Y, Amin-Mansour A, Yamauchi M, Sukawa Y, Stewart C, Rosenberg M, Mima K, Inamura K, Nosho K, Nowak JA, Lawrence MS, Giovannucci EL, Chan AT, Ng K, Meyerhardt JA, Van Allen EM, Getz G, Gabriel SB, Lander ES, Wu CJ, Fuchs CS, Ogino S, Garraway LA. 2016. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. *Cell Reports* **15(4)**:857–865 DOI 10.1016/j.celrep.2016.03.075.

Hanahan D. 2022. Hallmarks of cancer: new dimensions. *Cancer Discovery* **12(1)**:31–46 DOI 10.1158/2159-8290.CD-21-1059.

Hesson LB, Ng B, Zarzour P, Srivastava S, Kwok C-T, Packham D, Nunez AC, Beck D, Ryan R, Dower A, Ford CE, Pimanda JE, Sloane MA, Hawkins NJ, Bourke MJ, Wong JWH, Ward RL. 2016. Integrated genetic, epigenetic, and transcriptional profiling identifies molecular pathways in the development of laterally spreading tumors. *Molecular Cancer Research: MCR* **14(12)**:1217–1228 DOI 10.1158/1541-7786.MCR-16-0175.

Hofer P, Hagmann M, Brezina S, Dolejsi E, Mach K, Leeb G, Baierl A, Buch S, Sutterlüty-Fall H, Karner-Hanusch J, Bergmann MM, Bachleitner-Hofmann T, Stift A, Gerger A, Rötzer K, Karner J, Stättner S, Waldenberger M, Meitinger T, Strauch K, Linseisen J, Gieger C, Frommlet F, Gsur A. 2017. Bayesian and frequentist analysis of an Austrian genome-wide association study of colorectal cancer and advanced adenomas. *Oncotarget* **8(58)**:98623–98634 DOI 10.18632/oncotarget.21697.

Iranmanesh H, Majd A, Mojarad EN, Zali MR, Hashemi M. 2021. Investigating the relationship between the expression level of mucin gene cluster (MUC2, MUC5A, and MUC5B) and clinicopathological characterization of colorectal cancer. *Galen Medical Journal* **10**:e2030 DOI 10.31661/gmj.v10i0.2030.

Irmak-Yazicioglu MB. 2016. Mechanisms of microRNA deregulation and microRNA targets in gastric cancer. *Oncology Research and Treatment* **39(3)**:136–139 DOI 10.1159/000443224.

Jo YS, Kim MS, Yoo NJ, Lee SH. 2017. Intratumoral heterogeneity for inactivating frameshift mutation of CUX1 and SIRT1 genes in gastric and colorectal cancers. *Polish Journal of Pathology* **68(3)**:258–260 DOI 10.5114/pjp.2017.71534.

Kotelevets L, Chastre E. 2020. Rac1 signaling: from intestinal homeostasis to colorectal cancer metastasis. *Cancers* **12(3)**:665 DOI 10.3390/cancers12030665.

Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, Georgakopoulos N, Torrente F, Noorani A, Goddard M, Robinson P, Coorens THH, O'Neill L, Alder C, Wang J, Fitzgerald RC, Zilbauer M, Coleman N, Saeb-Parsy K, Martincorena I, Campbell PJ, Stratton MR. 2019. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574(7779)**:532–537 DOI 10.1038/s41586-019-1672-7.

Leedham S, Tomlinson I. 2012. The continuum model of selection in human tumors: general paradigm or niche product? *Cancer Research* **72(13)**:3131–3134 DOI 10.1158/0008-5472.CAN-12-1052.

Leygo C, Williams M, Jin HC, Chan MWY, Chu WK, Grusch M, Cheng YY. 2017. DNA methylation as a noninvasive epigenetic biomarker for the detection of cancer. *Disease Markers* **2017(4)**:3726595 DOI 10.1155/2017/3726595.

Li X, Thirumalai D. 2016. Interplay of driver, mini-driver, and deleterious passenger mutations on cancer progression. **49(17)**:4682 DOI 10.1101/084392.

Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, Seoane JA, Farshidfar F, Bowlby R, Islam M, Kim J, Chatila W, Akbani R, Kanchi RS, Rabkin CS, Willis JE, Wang KK, McCall SJ, Mishra L, Ojesina AI, Bullman S, Pedamallu CS, Lazar AJ, Sakai R, Cancer Genome Atlas Research Network, Thorsson V, Bass AJ, Laird PW. 2018. Comparative molecular analysis of gastrointestinal adenocarcinomas. *Cancer Cell* **33(4)**:721–735.e8 DOI 10.1016/j.ccell.2018.03.010.

Liu N, Sun Q, Wan L, Wang X, Feng Y, Luo J, Wu H. 2020. CUX1, a controversial player in tumor development. *Frontiers in Oncology* **10**:738 DOI 10.3389/fonc.2020.00738.

Lu Y, Wu S, Cui C, Yu M, Wang S, Yue Y, Liu M, Sun Z. 2020. Gene expression along with genomic copy number variation and mutational analysis were used to develop a 9-gene signature for estimating prognosis of COAD. *OncoTargets and Therapy* **13**:10393–10408 DOI 10.2147/OTT.S255590.

Ma R, Jing C, Zhang Y, Cao H, Liu S, Wang Z, Chen D, Zhang J, Wu Y, Wu J, Feng J. 2020. The somatic mutation landscape of Chinese Colorectal Cancer. *Journal of Cancer* **11(5)**:1038–1046 DOI 10.7150/jca.37017.

Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. 2011. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biology* **12(4)**:57 DOI 10.1186/gb-2011-12-4-r41.

Miao Y-R, Liu C-J, Hu H, Yang M, Guo A-Y. 2022. Biological pathway-derived TMB robustly predicts the outcome of immune checkpoint blockade therapy. *Cells* **11(18)**:2802 DOI 10.3390/cells11182802.

Pastò A, Serafin V, Pilotto G, Lago C, Bellio C, Trusolino L, Bertotti A, Hoey T, Plateroti M, Esposito G, Pinazza M, Agostini M, Nitti D, Amadori A, Indraccolo S. 2014. NOTCH3 signaling regulates MUSASHI-1 expression in metastatic colorectal cancer cells. *Cancer Research* **74(7)**:2106–2118 DOI 10.1158/0008-5472.CAN-13-2022.

Qi J, Hu Z, Liu S, Li F, Wang S, Wang W, Sheng X, Feng L. 2020. Comprehensively analyzed macrophage-regulated genes indicate that PSMA2 promotes colorectal cancer progression. *Frontiers in Oncology* **10**:618902 DOI 10.3389/fonc.2020.618902.

R Core Team. 2022. *R: A language and environment for statistical computing*. Version 4.2.0. Vienna: R Foundation for Statistical Computing. *Available at* https://www.r-project.org.

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **71(3)**:209–249 DOI 10.3322/caac.21660.

Thierry AR, Mouliere F, El Messaoudi S, Mollevi C, Lopez-Crapez E, Rolet F, Gillet B, Gongora C, Dechelotte P, Robert B, Del Rio M, Lamy P-J, Bibeau F, Nouaille M, Loriot V, Jarrousse A-S, Molina F, Mathonnet M, Pezet D, Ychou M. 2014. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. *Nature Medicine* **20(4)**:430–435 DOI 10.1038/nm.3511.

Timmermann B, Kerick M, Roehr C, Fischer A, Isau M, Boerno ST, Wunderlich A, Barmeyer C, Seemann P, Koenig J, Lappe M, Kuss AW, Garshasbi M, Bertram L, Trappe K, Werber M, Herrmann BG, Zatloukal K, Lehrach H, Schweiger MR. 2010. Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis. *PLOS ONE* **5**:e15661 DOI 10.1371/journal.pone.0015661.

van Ginkel J, Tomlinson I, Soriano I. 2023. The evolutionary landscape of colorectal tumorigenesis: recent paradigms, models, and hypotheses. *Gastroenterology* **164(5)**:841–846 DOI 10.1053/j.gastro.2022.11.049.

Wang J, Akter R, Shahriar MF, Uddin MN. 2022. Cancer-associated stromal fibroblast-derived transcriptomes predict poor clinical outcomes and immunosuppression in colon cancer. *Pathology Oncology Research: POR* **28**:1610350 DOI 10.3389/pore.2022.1610350.

Wang R, Zhu L, Zhang Y, Miao L, Ma H, Yuan H, Chen N. 2015. Genetic variants at 6p21.1 are associated with head and neck cancer in Chinese Han population. *Cancer Biomarkers: Section A of Disease Markers* **15(1)**:27–32 DOI 10.3233/CBM-140442.

Wickham H. 2011. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics* **3(2)**:180–185 DOI 10.1002/wics.147.

Wilk G, Braun R. 2018. Integrative analysis reveals disrupted pathways regulated by microRNAs in cancer. *Nucleic Acids Research* **46(3)**:1089–1101 DOI 10.1093/nar/gkx1250.

Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, Xiong F, Guo C, Wu X, Li Y, Li X, Li G, Zeng Z, Xiong W. 2020. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Molecular Cancer* **19(1)**:22 DOI 10.1186/s12943-020-1147-3.

Wu J, Wang Y, Xu X, Cao H, Sahengbieke S, Sheng H, Huang Q, Lai M. 2016. Transcriptional activation of FN1 and IL11 by HMGA2 promotes the malignant behavior of colorectal cancer. *Carcinogenesis* **37(5)**:511–521 DOI 10.1093/carcin/bgw029.

Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M, Ma'ayan A. 2021. Gene set knowledge discovery with enrichr. *Current Protocols* **1(3)**:e90 DOI 10.1002/cpz1.90.

Xiu M, Wang Y, Li B, Wang X, Xiao F, Chen S, Zhang L, Zhou B, Hua F. 2021. The role of Notch3 signaling in cancer stemness and chemoresistance: molecular mechanisms and targeting strategies. *Frontiers in Molecular Biosciences* **8**:694141 DOI 10.3389/fmolb.2021.69414.

Yang S, Sun Z, Zhou Q, Wang W, Wang G, Song J, Li Z, Zhang Z, Chang Y, Xia K, Liu J, Yuan W. 2018. MicroRNAs, long noncoding RNAs, and circular RNAs: potential tumor biomarkers and targets for colorectal cancer. *Cancer Management and Research* **10**:2249–2257 DOI 10.2147/CMAR.

Zheng X, Xu K, Zhou B, Chen T, Huang Y, Li Q, Wen F, Ge W, Wang J, Yu S, Sun L, Zhu L, Liu W, Gao H, Yue L, Cai X, Zhang Q, Ruan G, Zhu T, Wu Z, Zhu Y, Shao Y, Guo T, Zheng S. 2020. A circulating extracellular vesicles-based novel screening tool for colorectal cancer revealed by shotgun and data-independent acquisition mass spectrometry. *Journal of Extracellular Vesicles* **9(1)**:1750202 DOI 10.1080/20013078.2020.1750202.