Karadeniz Fen Bilimleri Dergisi, 15(1), 519-535, 2025. DOI: 10.31466/kfbd.1608816

KFBD

Karadeniz Fen Bilimleri Dergisi
The Black Sea Journal of Sciences
ISSN (Online): 2564-7377 https://dergipark.org.tr/tr/pub/kfbd

Araştırma Makalesi / Research Article

Proteomic Profiling in Colorectal Cancer: Identifying Druggable Biomarkers for Personalized Therapy

Abdulkadir ELMAS[1,2*]

Abstract

Colorectal cancer (CRC) remains a major global health challenge, with limited treatment options for advanced-stage patients. While genomic and transcriptomic analyses aid in target identification, proteomic alterations offer a more direct link to tumor biology and therapeutic opportunities. In this study, we analyzed mass spectrometry-based proteomics data from 102 primary CRC patients, including tumor and matched normal tissues, to systematically identify overexpressed, druggable therapeutic targets, with a particular focus on the patient kinome. Using the OPPTI approach, we discovered 31 kinases with notable overexpression, including 16 currently targetable by existing drugs, such as FGR, EPHA2, and PBK. Furthermore, we revealed 884 overexpressed non-kinase proteins, 253 of which are druggable, including ERAP2, FLG, and MT1H. Differential expression analysis identified 165 dysregulated kinases and 3,903 non-kinase proteins, with MET and STK3 emerging as potential candidates due to their substantial upregulation. Integrating differential expression and overexpression analyses, we highlighted a cohort of druggable targets, including EPHA2 and MET, whose inhibition has shown promising preclinical efficacy. This comprehensive proteomic study provides a resource for novel therapeutic target discovery in CRC, offering a framework for more personalized interventions through the identification of clinically actionable protein-level alterations.

Keywords: Colorectal cancer (CRC), Proteomics, Personalized therapy, Drug targets, Tumor biomarkers.

Kolorektal Kanserin Proteomik Profili: Kişiselleştirilmiş Tedavi için İlaçlanabilir Biyobelirteçlerin Keşfi

Öz

Kolorektal kanser (CRC), ileri evre hastalar için sınırlı tedavi seçenekleri nedeniyle küresel ölçekte önemli bir sağlık sorunu olmaya devam etmektedir. Genomik ve transkriptomik analizler hedef belirlemede değerli bir rol oynasa da, proteomik düzeydeki değişimler tümör biyolojisini daha doğrudan yansıtmakta ve terapötik açıdan daha uygulanabilir fırsatlar sunmaktadır. Bu çalışmada, 102 birincil CRC hastasına ait, tümör ve eşlenmiş normal dokuları içeren kütle spektrometresi proteomik verilerini analiz ederek, hasta kinomu üzerine odaklanarak sistematik bir şekilde aşırı ifade (eksprese) edilen, ilaçlanabilir terapötik hedefleri tanımladık. OPPTI metodunu kullanarak, mevcut ilaçlarla hedeflenebilir olan FGR, EPHA2 ve PBK gibi 16 kinazı içeren, toplamda 31 kinazın belirgin şekilde aşırı ifade edildiğini tespit ettik. Ayrıca, ERAP2, FLG ve MT1H gibi 253'ü ilaçlanabilir olan 884 kinaz dışı proteinin aşırı ifade edildiğini ortaya koyduk. Diferansiyel ifade analizi, 165 düzensiz kinaz ve 3,903 kinaz dışı proteini belirlerken, MET ve STK3, önemli derecede artış gösteren potansiyel adaylar olarak öne çıktı. Diferansiyel ifade ve aşırı ifade analizlerini birleştirerek, EPHA2 ve MET gibi ilaçlanabilir hedeflerden oluşan bir grup belirledik; bu hedeflerin inhibisyonu, umut verici preklinik etkinlik göstermiştir. Bu kapsamlı proteomik çalışma, CRC'de yeni terapötik hedeflerin keşfi için bir kaynak sunmakta ve klinik olarak uygulanabilir protein düzeyindeki değişikliklerin tanımlanması yoluyla daha kişiselleştirilmiş müdahalelere yönelik bir çerçeve sağlamaktadır.

Anahtar Kelimeler: Kolorektal kanser (CRC), Proteomik analiz, Kişiselleştirilmiş tedavi, İlaç hedefleri, Tümör biyobelirteçleri.

[1]Giresun University, Department of Electrical and Electronics Engineering, Giresun, Türkiye, abdulkadir.elmas@giresun.edu.tr
[2]Icahn School of Medicine at Mount Sinai, Dept. of Genetics and Genomic Sciences, New York, USA, abdulkadir.elmas@mssm.edu

[*]Sorumlu Yazar/Corresponding Author

Geliş/Received: 28.12.2024 **Kabul/Accepted:** 02.03.2025 **Yayın/Published:** 15.03.2025

Elmas, A. (2025). Proteomic Profiling in Colorectal Cancer: Identifying Druggable Biomarkers for Personalized Therapy. *Karadeniz Fen Bilimleri Dergisi*, 15(1), 519-535.

1. Introduction

Colorectal cancer (CRC) remains one of the leading causes of cancer-related morbidity and mortality worldwide (Siegel et al., 2023; Morgan et al., 2023). While early-stage CRC can be effectively treated through surgical resection and standard chemotherapy, advanced and metastatic CRC continues to pose a substantial clinical challenge due to limited treatment options and therapeutic resistance (Van Cutsem et al., 2016). A critical gap in improving CRC outcomes lies in the identification of reliable and actionable molecular targets that can drive the development of more effective, personalized therapeutic strategies (Rustgi, 2007; De Roock et al., 2010).

Over the past two decades, genomic and transcriptomic analyses have revolutionized cancer research, enabling the discovery of genetic alterations and aberrant transcriptional programs associated with CRC progression (Nunes et al., 2024). Large-scale initiatives such as The Cancer Genome Atlas (TCGA) have unveiled recurrent mutations in key oncogenes and tumor suppressors, including APC, KRAS, SMAD4, and TP53, as well as dysregulated signaling pathways such as WNT/β-catenin, PI3K/AKT, and MAPK (Muzny et al., 2012). However, these DNA- and RNA-level findings do not always translate into protein-level changes that are functionally relevant to tumor biology or therapeutically targetable (Huang et al., 2017). Proteins, as the ultimate effectors of cellular processes, provide a more direct representation of tumor phenotypes and actionable vulnerabilities (Mertins et al., 2014; Ruggles et al., 2017). Therefore, comprehensive proteomic profiling is essential to bridge the gap between genomic alterations and clinical application by identifying overexpressed, druggable proteins that drive CRC progression.

Kinases represent a particularly promising class of therapeutic targets due to their central roles in oncogenic signaling, proliferation, and survival (Zhang et al., 2009). Despite their success, the identification and clinical implementation of kinase targets in CRC remain limited, highlighting the need for systematic, proteome-wide assessments of dysregulated kinases in CRC tumors (Bhullar et al., 2018). Beyond kinases, non-kinase proteins may also contribute to CRC pathogenesis and therapeutic resistance, representing an underexplored reservoir of potential drug targets.

This work presents a comprehensive proteomic analysis of 95 primary CRC tumors and matched normal tissues (Vasaikar et al., 2019). Our goal was to systematically identify overexpressed, druggable proteins, with a particular emphasis on the CRC tumor kinome and actionable non-kinase proteins. Using a combination of the OPPTI algorithm for overexpression analysis and paired differential expression (DE) analysis, we identified a cohort of proteins that are significantly upregulated in CRC tumors and linked to existing or potential therapeutic strategies. This proteomic study provides an extensive resource for understanding CRC biology at the protein level, by offering

a framework for identifying clinically actionable targets in CRC tumors that can advance personalized therapeutic interventions.

2. Materials and Methods

2.1. Data Acquisition, Normalization, and Standardization

The proteomic datasets for the colorectal cancer (CRC) cohort were sourced from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) under the National Cancer Institute (Vasaikar et al., 2019). This dataset included 95 tumor samples and 94 paired normal controls, with a total of 7,412 unique proteins quantified, among which 333 were kinases. For robustness, the Median Absolute Deviation (MAD) was used to normalize each sample, aligning all datasets to a unit MAD for consistency. Additionally, protein markers with more than 20% missing values were excluded to maintain data quality and robustness.

2.2. Identification of Overexpressed Proteins

We utilized the OPPTI algorithm (Elmas et al., 2021) to identify markers showing significant overexpression. The expression level of each marker in a tumor sample is assessed using a weighted k-nearest neighbor (KNN) algorithm, with the weights determined by the relative abundance of other markers that are co-expressed at varying levels. To determine the statistical significance of overexpression, OPPTI conducts a permutation test. Specifically, dysregulation scores for proteins within each sample are randomly permuted, generating null overexpression events. This procedure is repeated multiple times to build a comprehensive permutation distribution based on the accumulated null overexpression scores across all iterations.

2.3. Identification of Differentially-Expressed Proteins

Differential protein expression in the CRC cohort was assessed through a paired analysis, where tumor samples were compared to their corresponding normal tissues using the limma R package (v3.42.2) (Ritchie et al., 2015). To account for potential confounders, adjustments were made for demographic factors, including age and gender. The Benjamini-Hochberg (BH) method was applied to correct p-values for multiple testing and control the false discovery rate (FDR) (Benjamini & Hochberg, 1995).

3. Findings and Discussion

3.1. CRC cohort

We acquired the mass spectrometry (MS) proteomics data from a cohort of 102 colorectal cancer patients, comprising 95 tumor samples and 94 matched normal tissues (Vasaikar et al., 2019) (Figure 1). To ensure the integrity and consistency of the data, a rigorous normalization pipeline and thorough quality control measures were applied (Methods). As a result, 7,418 proteins were reliably quantified for subsequent analyses. Additionally, we integrated the Drug-Gene Interaction Database (DGIdb) (Cotto et al., 2018) to identify druggable targets. Cross-referencing our quantified proteomic dataset with the DGIdb revealed 1,312 proteins that are currently associated with known therapeutic compounds.

We prioritized kinase proteins for further investigation, emphasizing those with known roles in driving oncogenesis and serving as validated therapeutic targets across various cancer types (Bhullar et al., 2018). Cross-referencing with the catalog of 683 human kinases compiled by (Manning et al., 2002), we identified 333 proteins that were robustly quantified in the CRC dataset, including 211 classified as druggable kinases (Figure 1). Additionally, these kinases were mapped to ten well-established oncogenic signaling pathways, as defined by the TCGA PanCanAtlas (Weinstein et al., 2013). These pathways encompass the Cell Cycle, HIPPO pathway, MYC regulatory network, NOTCH pathway, NRF2, PI3K axis, TGFβ pathway, receptor tyrosine kinase (RTK)/RAS/MAPK signaling cascade, TP53 pathway, and WNT/β-catenin signaling pathway (Sanchez-Vega et al., 2018) (Figure S1) (Table S1).

Study	Colorectal Cancer (CRC)
PMID	31031003
Sample Size (Patients/Tumors/Normals)	P: 102 T: 95 N: 100
Female Percentage	57.4%
Average Age at Onset	65.2
Tumor Stage Distribution	1: 10.2% 2: 40.6% 3: 41.1% 4: 8.1%
# of quantified proteins	7,418
# of quantified DGIdb-druggable proteins	1,312
# of quantified kinases	333
# of quantified DGIdb-druggable kinases	211

Figure 1. CRC cohort study overview. Summary of the proteomic dataset analyzed in this study for the human colorectal cancer cohort.

3.2. Overexpressed Proteins

To uncover potential targets for therapy from the proteins with elevated expression in colorectal cancer, we employed the OPPTI method (Elmas et al., 2021). This method is tailored to pinpoint protein overexpression within large-scale proteomic datasets generated by mass spectrometry, effectively handling variations and inconsistencies in quantitative data distributions. By applying OPPTI, we pinpointed proteins that could be effective therapeutic targets upon inhibition, independent of technical platform differences.

Our analysis revealed significant overexpression in 31 kinases (False discovery rate [FDR] < 0.05, OPPTI permutation test for marker overexpression), with 16 classified as druggable genes according to DGIdb. Notable examples include FGR (Protein overexpression rate [PRO] = 20.5, FDR = 1e-4), EPHA2 (PRO = 20, FDR = 1.2e-4), GUCY2C (PRO = 17.8, FDR = 0.0018), NME1 (PRO = 16.8, FDR = 0.0025), TNIK (PRO = 16.5, FDR = 0.005), and PBK (PRO = 16.1, FDR = 0.0056) (Figure 2A, 2B) (Table S2). Also, we found 884 non-kinase proteins with significant overexpression (FDR < 0.05), and 253 of them are currently-targeted by DGIdb, notably ERAP2 (PRO = 38.9, FDR = < 1e-100), FLG (PRO = 26.6, FDR = < 1e-100), and MT1H (PRO = 24.2, FDR = 2e-6) (Figure S2A, S2B) (Table S2).

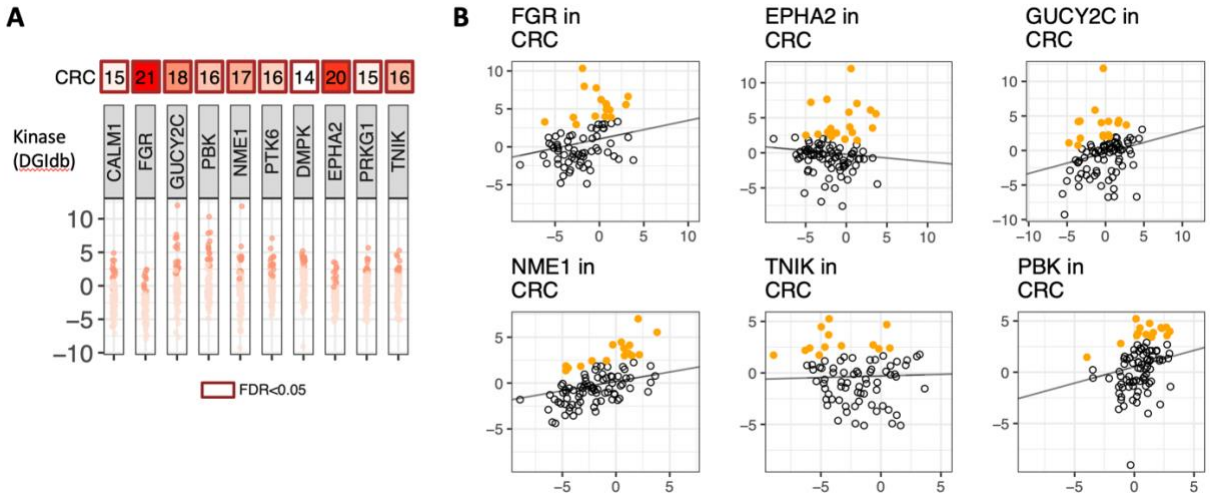

Figure 2. Kinase overexpression in the CRC cohort. (A) Druggable protein kinases identified as significantly overexpressed in the CRC cohort using the OPPTI method. (B) A sample-level assessment of kinase overexpression for the proteins highlighted in panel A. This OPPTI-based analysis illustrates how observed protein expression levels (y-axis) deviate from inferred background values (x-axis), with overexpression determined relative to a significance threshold (not explicitly shown).

3.3. Tumor-normal differentially-expressed proteins (DEPs)

To identify proteins that show tumor-specific expression, we carried out a paired differential expression (DE) analysis using tumors and matched-normal samples obtained from the same CRC patients. To deal with potential confounders of expression, we corrected the results for certain covariates including age, gender and ethnicity. For statistical robustness, false discovery rates (FDR) were calculated for the identified DEPs to control for type 1 error, and the t-statistics are moderated with an empirical Bayes method (Ritchie et al., 2015) (Methods). This analysis revealed a total of 165 kinases that were significantly differentially expressed at 5% FDR cutoff. Among them, 17 belonged to several oncogenic signaling pathways annotated, where two kinases showed more than 2-fold upregulation of expression in tumors compared to their matched normal tissues, namely MET (FC = 1.4, FDR = 0.012), and STK3 (FC = 1.1, FDR = 0.025) (Figure 3A, 3B) (Table S3).

Among the non-kinase proteins, we similarly identified 3,903 significant differentially expressed markers (FDR < 0.05), including 58 proteins acting in oncogenic signaling pathways. Of these, 18 proteins displayed at least 2-fold increase in expression among the tumor tissues compared to their matched normals, including THBS2 (FC = 6.1, FDR = 1e-7), SFRP4 (FC = 6, FDR = 4e-5), RB1 (FC = 3.2, FDR = 1e-7), CHD4 (FC = 2, FDR = 8e-11), RPS6 (FC = 1.7, FDR = 1e-8), and MFAP2 (FC = 2.5, FDR = 1e-5) (Figure S3A, S3B) (Table S3).

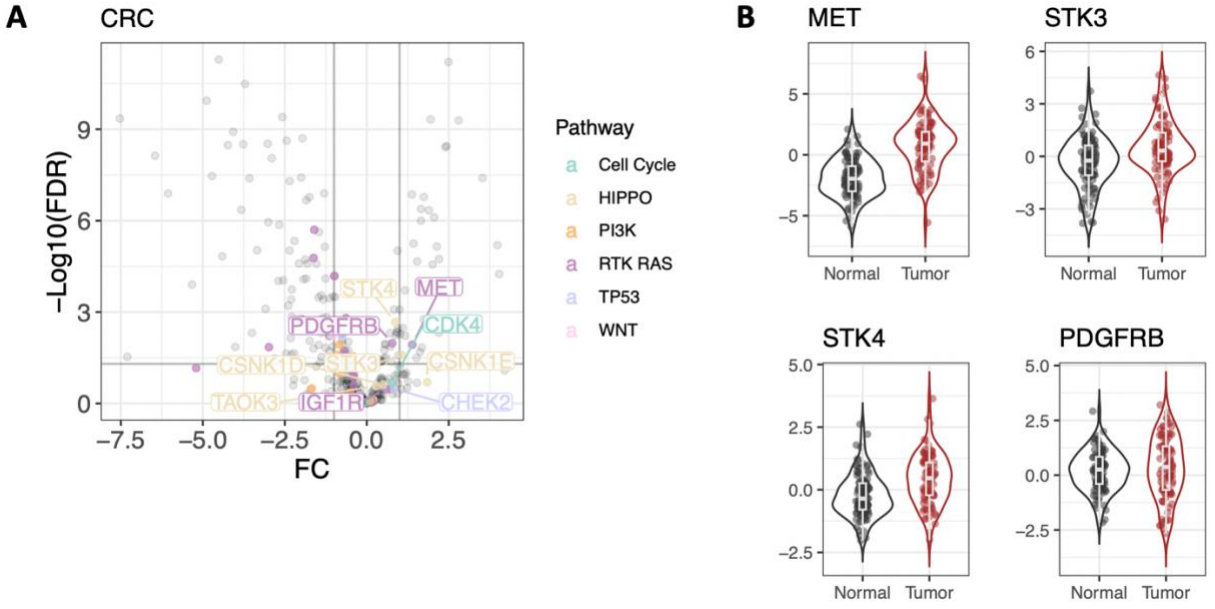

Figure 3. Differentially-expressed kinases in human CRC tumors. (A) The volcano plot illustrates kinase proteins that show differential expression between tumor vs normal tissue samples. Key kinases acting in oncogenic signaling pathways with significant changes are highlighted and labeled. (B) Violin plots of top kinases highlighted in panel A, displaying elevated levels of protein expressions in CRC tumors versus normal tissue samples. The gray box-plots illustrate the distribution, median, and interquartile range of protein expression across samples.

3.4. Cohort-wise overexpressed DEPs

Tumor-expressed protein markers (DEPs) may exhibit silent or inconsistent expression across patient cohorts. Therefore, we aimed to identify differentially expressed proteins with cohort-specific overexpression, as they represent promising candidates for potential therapeutic targeting. By cross-referencing the identified tumor DEPs with OPPTI-derived proteins, we identified 57 DGIdb kinases that are targetable by existing drugs and demonstrate elevated levels of differential expression (DE) and overexpression in the CRC cohort. Of these, two were significant across both analyses, namely FGR (FC = 2.1, FDR = 0.044; PRO = 20.5%, FDR = 0.0001), and PBK (FC = 2.2, FDR = 0.0047; PRO = 16.1%, FDR = 0.0056). Among the kinases not targeted by existing drugs, 26 were positive in both DE and overexpression, with 4 reaching statistical significance across the analyses, i.e., HKDC1 (FC = 4, FDR = 6e-5; PRO = 25.3%, FDR = 8e-7), HK3 (FC = 2.2, FDR = 2e-5; PRO = 14.7%, FDR = 0.013), PTK7 (FC = 2.1, FDR = 7e-6; PRO = 13.7%, FDR = 0.03), and FASTKD5 (FC = 1.5, FDR = 0.032; PRO = 17.2%, FDR = 0.0022) (Figure 4) (Table S4).

Furthermore, we analyzed the non-kinase proteins and identified 50 DGIdb proteins showing significant DE and overexpression in the CRC cohort, including FAP (FC = 6.9, FDR = 8.2e-9; PRO = 28.4%, FDR = < 1e-100), THBS2 (FC = 6.1, FDR = 1.1e-7; PRO = 33.7%, FDR = < 1e-100), THBS1 (FC = 4.6, FDR = 1.6e-7; PRO = 27.4%, FDR = < 1e-100), and MMP1 (FC = 5.1, FDR = 3.4e-6; PRO = 28.4%, FDR = < 1e-100). There were also 106 significant overexpressed DEPs not targeted by existing drug compounds, including GPRC5A (FC = 6.5, FDR = 4.2e-10; PRO = 27.4%, FDR = < 1e-100), TRIM29 (FC = 6.6, FDR = 1.7e-7; PRO = 34.7%, FDR = < 1e-100), S100P (FC = 6, FDR = 2e-7; PRO = 28.4%, FDR = < 1e-100), SERPINB5 (FC = 6, FDR = 9e-6; PRO = 33.7%, FDR = < 1e-100), and SFRP4 (FC = 6, FDR = 4e-5; PRO = 34.1%, FDR = < 1e-100) (Figure S4) (Table S4).

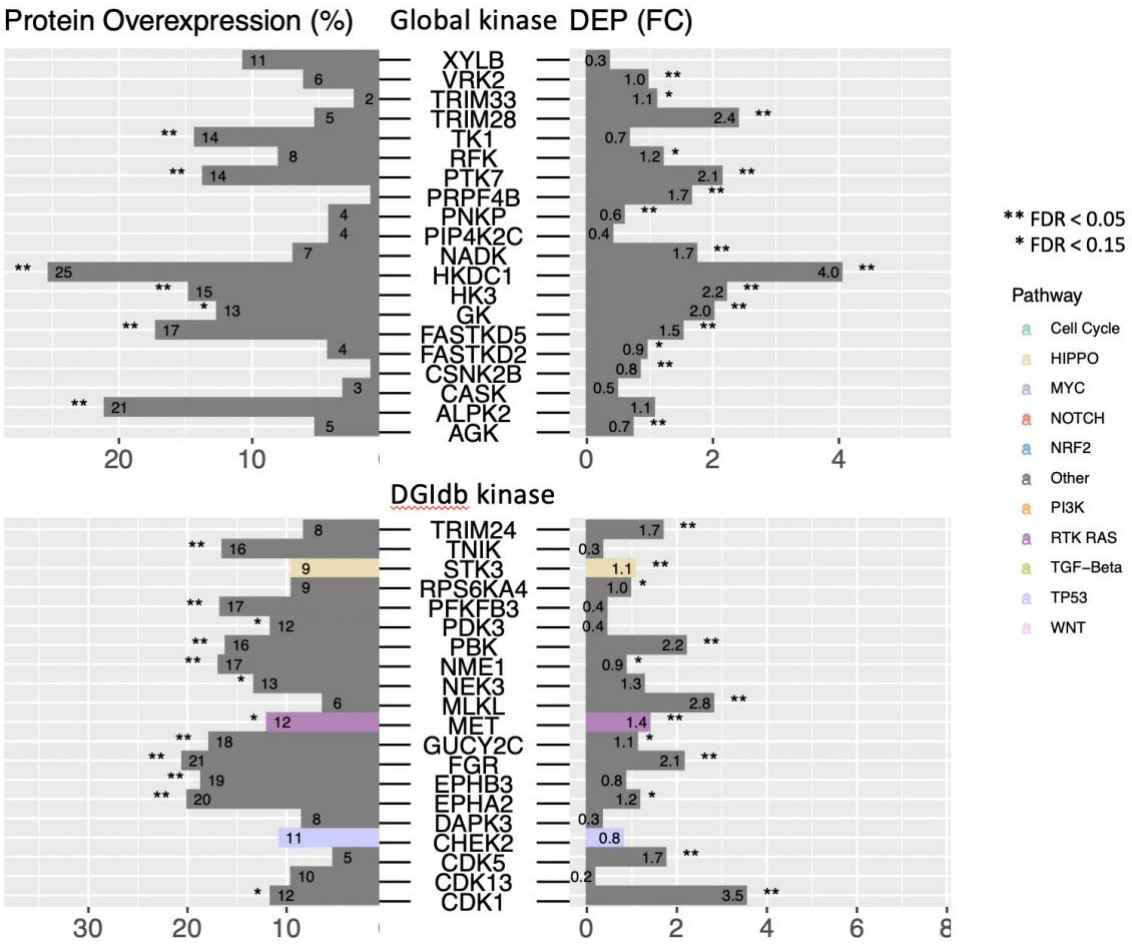

Figure 4. Potential CRC kinase targets demonstrating protein overexpression and differential expression. Top-20 global kinases (above) that exhibited significantly higher expression in tumors compared to normal tissues, as well as protein overexpression in the CRC cohort. The same analysis for the druggable kinases (with related drug compounds from DGIdb) is given below.

4. Conclusions and Recommendations

This study involved a comprehensive examination of protein expression profiles in 95 primary tumor samples from patients with colorectal cancer (CRC) using mass spectrometry proteomics data (Figure 1, S1). Our findings highlight significant proteomic dysregulation in CRC, revealing novel overexpressed kinases and non-kinase proteins with substantial therapeutic potential. These results address the existing gaps in current genomic-based target discovery strategies, which often overlook protein-level alterations critical for tumor progression and therapeutic intervention.

Kinases play pivotal roles in driving oncogenic signaling pathways. Our analyses revealed 31 significantly overexpressed kinases, 16 of which are already listed as druggable targets in the Drug Gene Interaction Database (DGIdb) (Figure 2). Among these, kinases such as FGR**,** EPHA2**,** and PBK stood out with consistent overexpression across both differential expression and OPPTI analyses, underscoring their potential as priority therapeutic candidates for CRC. EPHA2 has been

implicated in CRC progression through aberrant RTK signaling, promoting tumor cell survival and migration (Guo et al., 2024; Xiao et al., 2020). Preclinical studies have demonstrated that the EPHA2 inhibitor ALW-II-41-27 can significantly reduce tumor growth and improve the efficacy of Cetuximab by overcoming resistance (Tröster et al., 2023). EPHA2 inhibition has also shown promising effects on reducing CRC cell proliferation in vitro and tumor growth in xenograft mouse models, suggesting a dual role in directly suppressing tumor progression and restoring sensitivity to anti-EGFR therapies (Tröster et al., 2023). These findings further validate EPHA2 as a clinically relevant target, particularly for patients who develop resistance to standard therapies.

In addition to kinases, our analysis uncovered 884 significantly overexpressed non-kinase proteins, with 253 listed as druggable targets (Figure S2). Among these, proteins such as ERAP2, FLG, and MT1H displayed pronounced overexpression and therapeutic significance. Importantly, Endoplasmic Reticulum Aminopeptidase 2 (ERAP2) is a key enzyme involved in trimming peptides for presentation by HLA class I molecules, thereby shaping immune recognition. Its ability to alter antigen processing and presentation can either promote immune evasion in tumors or enhance immune recognition, suggesting its potential for developing targeted cancer immunotherapies (Lee, 2017; Fruci et al., 2008)**.**

Our tumor-normal differential expression analysis revealed 165 significantly dysregulated kinases and 3,903 non-kinase proteins, further reinforcing the proteomic heterogeneity of CRC (Figure 3, S3). Kinases such as MET and STK3 demonstrated notable upregulation, aligning with their established roles in CRC tumorigenesis. MET, has been widely recognized as a key driver of invasive growth and metastasis through activation of the HGF-MET axis. Targeting MET is promising due to its involvement in aggressive tumorigenesis and metastasis, with recent advances in MET inhibitors, degraders, and antibody-based therapies showing efficacy in addressing MET-driven cancers and overcoming resistance mechanisms (Gallo et al., 2024). Notably, the MET inhibitor Capmatinib has shown encouraging clinical activity when combined with the anti-EGFR monoclonal antibody Cetuximab. In MET-positive, KRAS wild-type metastatic CRC patients, Capmatinib in combination with Cetuximab demonstrated preliminary tumor shrinkage ranging from 29–44% and was well tolerated, providing evidence for its ability to overcome resistance to anti-EGFR therapy (Delord et al., 2020). STK3, a serine/threonine kinase in the Hippo pathway, plays a critical role in regulating apoptosis and cellular proliferation. Its promising potential as a target lies in its ability to act as a tumor suppressor, inhibiting tumor growth and metastasis through phosphorylation-dependent mechanisms, as demonstrated in multiple studies (Zhao et al., 2024). The identification of THBS2 and SFRP4 among highly upregulated non-kinase proteins further emphasizes their involvement in CRC through Wnt/β-catenin pathway. Thrombospondin-2 (THBS2), a member of the calcium-binding glycoprotein family, plays critical roles in extracellular matrix

remodeling, cell proliferation, and angiogenesis. In colorectal cancer, THBS2 emerges as a promising target due to its dual role in promoting metastasis through activation of the Wnt/β-catenin pathway and inducing M2 macrophage polarization via exosome secretion, thereby fostering an immunosuppressive tumor microenvironment (Liu et al., 2024; Qu et al., 2022). On the other hand, secreted frizzled-related protein 4 (SFRP4) is a class I antagonist of the Wnt signaling pathway, which regulates cell growth, differentiation, and apoptosis. SFRP4 shows promise as a therapeutic target in colorectal cancer due to its ability to inhibit the Wnt/β-catenin pathway, reduce tumor cell proliferation, and modulate chemoresistance, making it a potential candidate for gene therapy and improved treatment strategies (Liu et al., 2019).

To identify clinically actionable targets with consistent overexpression across the cohort, we integrated tumor-normal differential expression results with OPPTI-derived overexpression data. This approach enabled us to pinpoint 57 druggable kinases and 50 druggable non-kinase proteins exhibiting robust tumor-specific overexpression (Figure 4, S4). Among druggable kinases, **t**he tumor-specific overexpression of FGR and PBK kinases are associated with immune signaling and mitotic progression, respectively, further supporting their role in CRC pathogenesis. FGR, a member of the Src family kinases (SFKs), plays a role in inflammatory signaling and tumor progression by modulating immune cell responses within the tumor microenvironment. In colorectal cancer, FGR is particularly favorable as a therapeutic target due to its association with poor prognosis, reduced lymphocytic infiltration, and its potential to synergize with immunotherapies to improve treatment outcomes (Roseweir et al., 2019). PBK (PDZ-binding kinase) is a serine/threonine kinase involved in cellular proliferation and migration processes, and its inhibition has emerged as an effective strategy in colorectal cancer. PBK enhances tumor cell proliferation via Histone H3 phosphorylation and simultaneously suppresses migration and invasion through E-cadherin (CDH1) stabilization, suggesting its potential as a therapeutic target to reduce tumor growth and metastasis while improving CRC patient outcomes (Koshino et al., 2022). Finally, THBS2, FAP, and MMP1 represented non-kinase druggable targets with strong therapeutic relevance due to their involvement in extracellular matrix remodeling and tumor microenvironment regulation. Fibroblast activation protein-α (FAP) is a serine protease commonly overexpressed in tumor stromal fibroblasts and certain cancer epithelial cells. In colorectal cancer, FAP shows promise as a therapeutic target by promoting tumor migration, invasion, and metastasis via binding to Enolase 1 (ENO1) and activating the NF-κB signaling pathway, highlighting its potential for targeting metastatic progression (Yuan et al., 2021). Matrix metalloproteinase-1 (MMP1) is a collagenase involved in extracellular matrix degradation and tumor progression. In colorectal cancer, MMP1 emerges as a candidate therapeutic target due to its ability to promote tumor proliferation, migration, and invasion through the epithelial-mesenchymal

transition (EMT) and activation of the PI3K/Akt signaling pathway, highlighting its potential to mitigate tumor aggressiveness and metastasis (Wang et al., 2020).

Our findings highlight the importance of proteomic approaches for discovering actionable therapeutic targets in CRC. While genomic studies provide valuable insights into mutation-driven oncogenesis, they often fail to capture protein-level dysregulation that directly influences tumor biology and therapeutic vulnerabilities. By leveraging proteomic data, we identified several existing oncogenic drug targets and revealed many more novel druggable candidates, including both kinases and non-kinase proteins. Despite the strengths of our study, certain limitations warrant consideration. The reliance on proteomic data from a single cohort may introduce biases related to sample size and clinical heterogeneity. Future studies should validate these findings in independent CRC cohorts and explore functional mechanisms underlying the identified targets' roles in colorectal cancer. Furthermore, experimental validation of candidate proteins, such as FGR and PBK, will be critical to confirming their therapeutic potential. Integrating proteomic data with functional screens and drug response assays could present further evidence for the identified targets' therapeutic potential and accelerate the development of precision therapies for CRC.

In summary, our study highlights the effectiveness of mass spectrometry-based proteomics in uncovering new candidates for targeted treatment in colorectal cancer. Through the integration of advanced computational techniques and proteomic datasets, we identified a diverse set of overexpressed, druggable kinases and non-kinase proteins, including FGR, PBK, MET, EPHA2, THBS2, and ERAP2. These results reinforce the need for clinical investigations exploring the efficacy of established or novel targeted inhibitors to improve treatment outcomes and overcome drug resistance in CRC patients. Our results offer a critical foundation for the development of personalized therapeutic strategies in colorectal cancer and highlight the need for further investigation into the clinical application of these identified targets.

Data and Software Availability

Data Access

The supplementary tables referenced in this study are available at: https://www.columbia.edu/~ae2321/workspace/CRC/SupplementaryTables.xlsx. The CRC cohort data used in this study, originally reported by (Vasaikar et al., 2019) , is publicly accessible through resources provided by The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC). The data can be retrieved via the following links: https://proteomic.datacommons.cancer.gov/pdc/study/PDC000116 and https://pdc.cancer.gov/pdc/.

Software and Code

The OPPTI algorithm utilized in this study is open-source and can be accessed via GitHub at: https://github.com/Huang-lab/Oppti. All analyses were conducted using custom scripts developed in R (version 3.6.2).

Acknowledgements

The author extends gratitude to the colorectal cancer (CRC) cohort participants, along with their families, for their invaluable contributions to this study. The author would also like to thank the editor and the referees for their valuable contributions to the review of this manuscript. All scripts and textual content were thoroughly developed, reviewed, and validated by the author.

Authors' Contributions

AE conceptualized the study, designed the analytical framework, and developed the OPPTI software and tailored for the study. AE also carried out the bioinformatics analyses and was responsible for writing, revising, and approving the final version of the manuscript.

Statement of Conflict of Interests

The author confirms there are no conflicts of financial interest to disclose.

Statement of Research and Publication Ethics

The author declares that this study complies with Research and Publication Ethics.

Ethical Approval

This research utilized publicly available data from previously published sources. The dataset is open access and can be freely obtained from the CPTAC Data Portal (https://pdc.cancer.gov/pdc/). As such, no separate ethical approval was required for this study.

References

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: Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bhullar, K. S., Lagarón, N. O., McGowan, E. M., Parmar, I., Jha, A., Hubbard, B. P., & Rupasinghe, H. P. V. (2018). Kinase-targeted cancer therapies: Progress, challenges and future directions. In *Molecular Cancer* (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12943-018-0804-2

Cotto, K. C., Wagner, A. H., Feng, Y.-Y., Kiwala, S., Coffman, A. C., Spies, G., Wollam, A., Spies, N. C., Griffith, O. L., & Griffith, M. (2018). DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Research*, *46*(D1), D1068–D1073. https://doi.org/10.1093/nar/gkx1143

De Roock, W., Claes, B., Bernasconi, D., De Schutter, J., Biesmans, B., Fountzilas, G., Kalogeras, K. T., Kotoula, V., Papamichael, D., Laurent-Puig, P., Penault-Llorca, F., Rougier, P., Vincenzi, B., Santini, D., Tonini, G., Cappuzzo, F., Frattini, M., Molinari, F., Saletti, P., … Tejpar, S. (2010). Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *The Lancet Oncology*, *11*(8), 753–762. https://doi.org/10.1016/S1470-2045(10)70130-3

Delord, J.-P., Argilés, G., Fayette, J., Wirth, L., Kasper, S., Siena, S., Mesia, R., Berardi, R., Cervantes, A., Dekervel, J., Zhao, S., Sun, Y., Hao, H.-X., Tiedt, R., Vicente, S., Myers, A., & Siu, L. L. (2020). A phase 1b study of the MET inhibitor capmatinib combined with cetuximab in patients with MET-positive colorectal cancer who had progressed following anti-EGFR monoclonal antibody treatment. *Investigational New Drugs*, *38*(6), 1774–1783. https://doi.org/10.1007/s10637-020-00928-z

Elmas, A., Tharakan, S., Jaladanki, S., Galsky, M. D., Liu, T., & Huang, K.-L. (2021). Pan-cancer proteogenomic investigations identify post-transcriptional kinase targets. *Communications Biology*, *4*(1), 1112. https://doi.org/10.1038/s42003-021-02636-7

Fruci, D., Giacomini, P., Nicotra, M. R., Forloni, M., Fraioli, R., Saveanu, L., van Endert, P., & Natali, P. G. (2008). Altered expression of endoplasmic reticulum aminopeptidases ERAP1 and ERAP2 in transformed non-lymphoid human tissues. *Journal of Cellular Physiology*, *216*(3), 742–749. https://doi.org/10.1002/jcp.21454

Gallo, S., Folco, C. B., & Crepaldi, T. (2024). The MET Oncogene: An Update on Targeting Strategies. *Pharmaceuticals*, *17*(11), 1473. https://doi.org/10.3390/ph17111473

Guo, X., Yang, Y., Tang, J., & Xiang, J. (2024). Ephs in cancer progression: complexity and context-dependent nature in signaling, angiogenesis and immunity. In *Cell Communication and Signaling* (Vol. 22, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12964-024-01580-3

Huang, K.-L., Li, S., Mertins, P., Cao, S., Gunawardena, H. P., Ruggles, K. V, Mani, D. R., Clauser, K. R., Tanioka, M., Usary, J., Kavuri, S. M., Xie, L., Yoon, C., Qiao, J. W., Wrobel, J., Wyczalkowski, M. A., Erdmann-Gilmore, P., Snider, J. E., Hoog, J., … Ding, L. (2017). Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. *Nature Communications*, *8*, 14864. https://doi.org/10.1038/ncomms14864

Koshino, A., Nagano, A., Ota, A., Hyodo, T., Ueki, A., Komura, M., Sugimura-Nagata, A., Ebi, M., Ogasawara, N., Kasai, K., Hosokawa, Y., Kasugai, K., Takahashi, S., & Inaguma, S. (2022). PBK Enhances Cellular Proliferation With Histone H3 Phosphorylation and Suppresses Migration and Invasion With CDH1 Stabilization in Colorectal Cancer. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.772926

Lee, E. D. (2017). Endoplasmic Reticulum Aminopeptidase 2, a common immunological link to adverse pregnancy outcomes and cancer clearance? *Placenta*, *56*, 40–43. https://doi.org/10.1016/j.placenta.2017.03.012

Liu, Y., Li, J., & Qi, J. (2019). The role of the class I Wnt pathway antagonist sFRP4 in colorectal cancer. *Digestive Medicine Research*, *2*, 18–18. https://doi.org/10.21037/dmr.2019.08.01

Liu, Y., Lv, H., Liu, X., Xu, L., Li, T., Zhou, H., Zhu, H., Hao, C., Lin, C., & Zhang, Y. (2024). The RP11-417E7.1/THBS2 signaling pathway promotes colorectal cancer metastasis by activating the Wnt/β-catenin pathway and facilitating exosome-mediated M2 macrophage polarization. *Journal of Experimental and Clinical Cancer Research*, *43*(1). https://doi.org/10.1186/s13046-024-03107-7

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science (New York, N.Y.)*, *298*(5600), 1912–1934. https://doi.org/10.1126/science.1075762

Mertins, P., Yang, F., Liu, T., Mani, D. R., Petyuk, V. A., Gillette, M. A., Clauser, K. R., Qiao, J. W., Gritsenko, M. A., Moore, R. J., Levine, D. A., Townsend, R., Erdmann-Gilmore, P., Snider, J. E., Davies, S. R., Ruggles, K. V, Fenyo, D., Kitchens, R. T., Li, S., … Carr, S. A. (2014). Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Molecular & Cellular Proteomics : MCP*, *13*(7), 1690–1704. https://doi.org/10.1074/mcp.M113.036392

Morgan, E., Arnold, M., Gini, A., Lorenzoni, V., Cabasag, C. J., Laversanne, M., Vignat, J., Ferlay, J., Murphy, N., & Bray, F. (2023). Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut*, *72*(2), 338–344. https://doi.org/10.1136/gutjnl-2022-327736

Muzny, D. M., Bainbridge, M. N., Chang, K., Dinh, H. H., Drummond, J. A., Fowler, G., Kovar, C. L., Lewis, L. R., Morgan, M. B., Newsham, I. F., Reid, J. G., Santibanez, J., Shinbrot, E., Trevino, L. R., Wu, Y. Q., Wang, M., Gunaratne, P., Donehower, L. A., Creighton, C. J., … Thomson., E. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, *487*(7407), 330–337. https://doi.org/10.1038/nature11252

Nunes, L., Li, F., Wu, M., Luo, T., Hammarström, K., Torell, E., Ljuslinder, I., Mezheyeuski, A., Edqvist, P.-H., Löfgren-Burström, A., Zingmark, C., Edin, S., Larsson, C., Mathot, L., Osterman, E., Osterlund, E., Ljungström, V., Neves, I., Yacoub, N., … Sjöblom, T. (2024). Prognostic genome and transcriptome signatures in colorectal cancers. *Nature*, *633*(8028), 137–146. https://doi.org/10.1038/s41586-024-07769-3

Qu, H. L., Hasen, G. W., Hou, Y. Y., & Zhang, C. X. (2022). THBS2 promotes cell migration and invasion in colorectal cancer via modulating Wnt/β-catenin signaling pathway. *Kaohsiung Journal of Medical Sciences*, *38*(5), 469–478. https://doi.org/10.1002/kjm2.12528

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Roseweir, A. K., Powell, A. G. M. T., Horstman, S. L., Inthagard, J., Park, J. H., McMillan, D. C., Horgan, P. G., & Edwards, J. (2019). Src family kinases, HCK and FGR, associate with local inflammation and tumour progression in colorectal cancer. *Cellular Signalling*, *56*, 15–22. https://doi.org/10.1016/j.cellsig.2019.01.007

Ruggles, K. V., Krug, K., Wang, X., Clauser, K. R., Wang, J., Payne, S. H., Fenyö, D., Zhang, B., & Mani, D. R. (2017). Methods, Tools and Current Perspectives in Proteogenomics. *Molecular & Cellular Proteomics*, *16*(6), 959–981. https://doi.org/10.1074/mcp.MR117.000024

Rustgi, A. K. (2007). The genetics of hereditary colon cancer. In *Genes and Development* (Vol. 21, Issue 20, pp. 2525–2538). https://doi.org/10.1101/gad.1593107

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., Dimitriadoy, S., Liu, D. L., Kantheti, H. S., Saghafinia, S., Chakravarty, D., Daian, F., Gao, Q., Bailey, M. H., Liang, W.-W., Foltz, S. M., Shmulevich, I., Ding, L., Heins, Z., … Schultz, N. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell*, *173*(2), 321-337.e10. https://doi.org/10.1016/j.cell.2018.03.035

Siegel, R. L., Wagle, N. S., Cercek, A., Smith, R. A., & Jemal, A. (2023). Colorectal cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, *73*(3), 233–254. https://doi.org/10.3322/caac.21772

Tröster, A., Jores, N., Mineev, K. S., Sreeramulu, S., DiPrima, M., Tosato, G., & Schwalbe, H. (2023). Targeting EPHA2 with Kinase Inhibitors in Colorectal Cancer. *ChemMedChem*, *18*(23). https://doi.org/10.1002/cmdc.202300420

Van Cutsem, E., Cervantes, A., Adam, R., Sobrero, A., Van Krieken, J. H., Aderka, D., Aranda Aguilar, E., Bardelli, A., Benson, A., Bodoky, G., Ciardiello, F., D'Hoore, A., Diaz-Rubio, E., Douillard, J.-Y., Ducreux, M., Falcone, A., Grothey, A., Gruenberger, T., Haustermans, K., … Arnold, D. (2016). ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Annals of Oncology*, *27*(8), 1386–1422. https://doi.org/10.1093/annonc/mdw235

Vasaikar, S., Huang, C., Wang, X., Petyuk, V. A., Savage, S. R., Wen, B., Dou, Y., Zhang, Y., Shi, Z., Arshad, O. A., Gritsenko, M. A., Zimmerman, L. J., McDermott, J. E., Clauss, T. R., Moore, R. J., Zhao, R., Monroe, M. E., Wang, Y.-T., Chambers, M. C., … Clinical Proteomic Tumor Analysis Consortium. (2019). Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell*, *177*(4), 1035-1049.e19. https://doi.org/10.1016/j.cell.2019.03.030

Wang, K., Zheng, J., Yu, J., Wu, Y., Guo, J., Xu, Z., & Sun, X. (2020). Knockdown of MMP-1 inhibits the progression of colorectal cancer by suppressing the PI3K/Akt/c-myc signaling pathway and EMT. *Oncology Reports*, *43*(4), 1103–1112. https://doi.org/10.3892/or.2020.7490

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., & Stuart, J. M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics*, *45*(10), 1113–1120. https://doi.org/10.1038/ng.2764

Xiao, T., Xiao, Y., Wang, W., Tang, Y. Y., Xiao, Z., & Su, M. (2020). Targeting EphA2 in cancer. In *Journal of Hematology and Oncology* (Vol. 13, Issue 1). BioMed Central. https://doi.org/10.1186/s13045-020-00944-9

Yuan, Z., Hu, H., Zhu, Y., Zhang, W., Fang, Q., Qiao, T., Ma, T., Wang, M., Huang, R., Tang, Q., Gao, F., Zou, C., Gao, X., Wang, G., & Wang, X. (2021). Colorectal cancer cell intrinsic fibroblast activation protein alpha binds to Enolase1 and activates NF-κB pathway to promote metastasis. *Cell Death and Disease*, *12*(6). https://doi.org/10.1038/s41419-021-03823-4

Zhang, J., Yang, P. L., & Gray, N. S. (2009). Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer*, *9*(1), 28–39. https://doi.org/10.1038/nrc2559

Zhao, Z., Chu, Y., Feng, A., Zhang, S., Wu, H., Li, Z., Sun, M., Zhang, L., Chen, T., & Xu, M. (2024). STK3 kinase activation inhibits tumor proliferation through FOXO1-TP53INP1/P21 pathway in esophageal squamous cell carcinoma. *Cellular Oncology*, *47*(4), 1295–1314. https://doi.org/10.1007/s13402-024-00928-8

Supplementary Figures

Pathway	# genes	# DGIdb kinases
Cell Cycle	15	3
HIPPO	38	6
MYC	13	-
NOTCH	71	-
OSR/NRF2	3	-
PI3K	29	9
TGFβ	7	-
RTK/RAS/MAP	85	17
TP53	6	3
β-catenin/WNT	68	1

Figure S1. CRC kinome. (A) The number of available druggable kinases quantified in the CRC cohort, as annotated by the oncogenic signaling pathways they are involved.

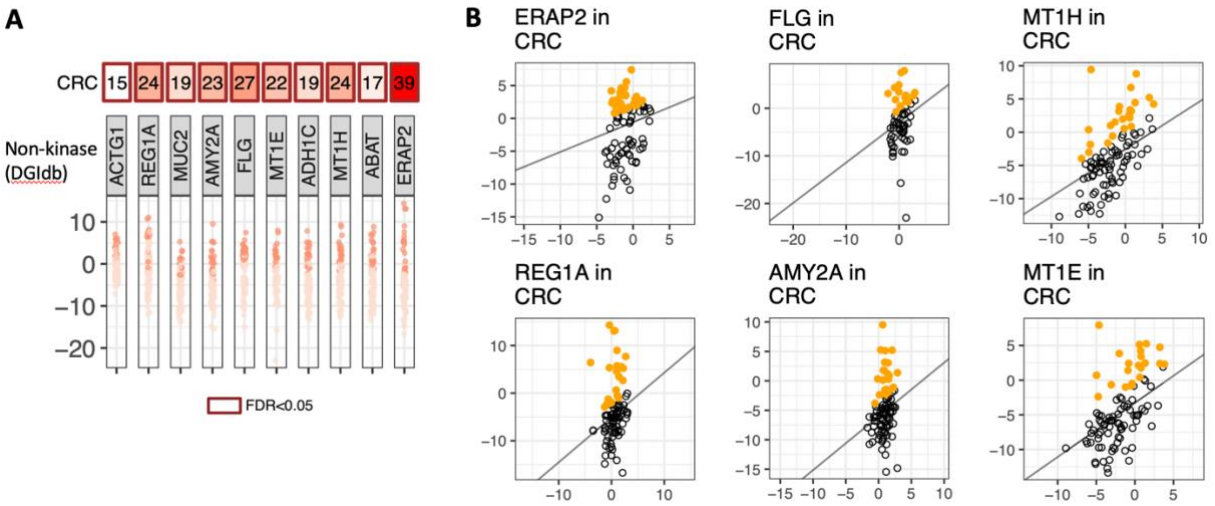

Figure S2. Overexpression of non-kinase proteins in the CRC cohort. (A) The most significantly overexpressed druggable non-kinase proteins were identified in the CRC tumor cohort using the OPPTI method. (B) Sample-level plots illustrate the overexpression patterns for the proteins listed in panel A, as determined by OPPTI.

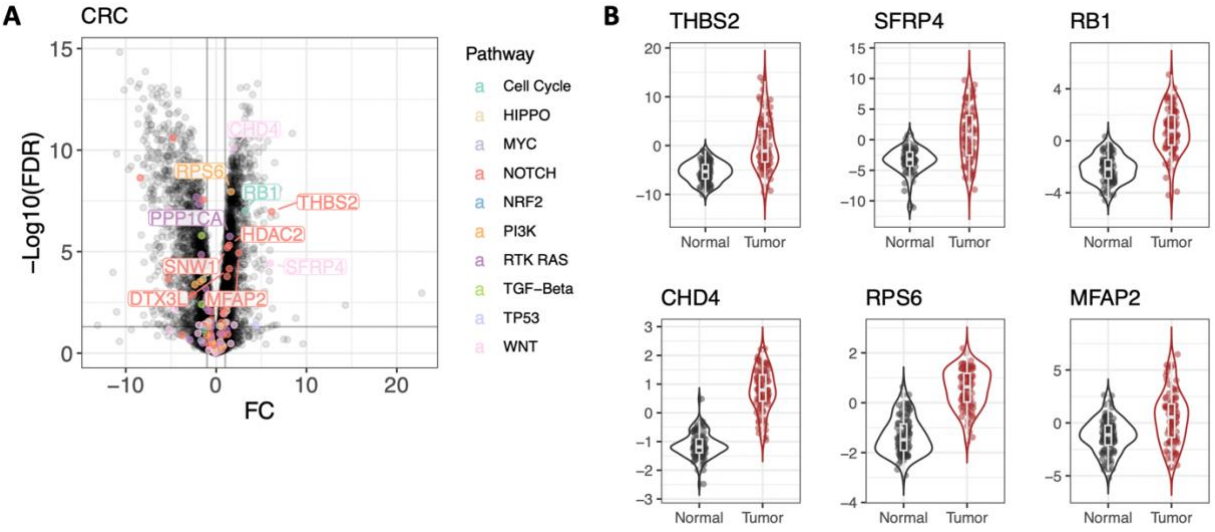

Figure S3. Differential expression analysis of proteins in human CRC tumors. (A) Volcano plot visualizes the non-kinase proteins showing significant expression differences between tumor and normal samples. Proteins with the most notable changes, particularly those involved in oncogenic signaling pathways, are annotated. (B) Violin plots provide a detailed view of the expression patterns for the key proteins highlighted in panel A, illustrating their higher abundance in CRC tumor samples compared to normal tissues.

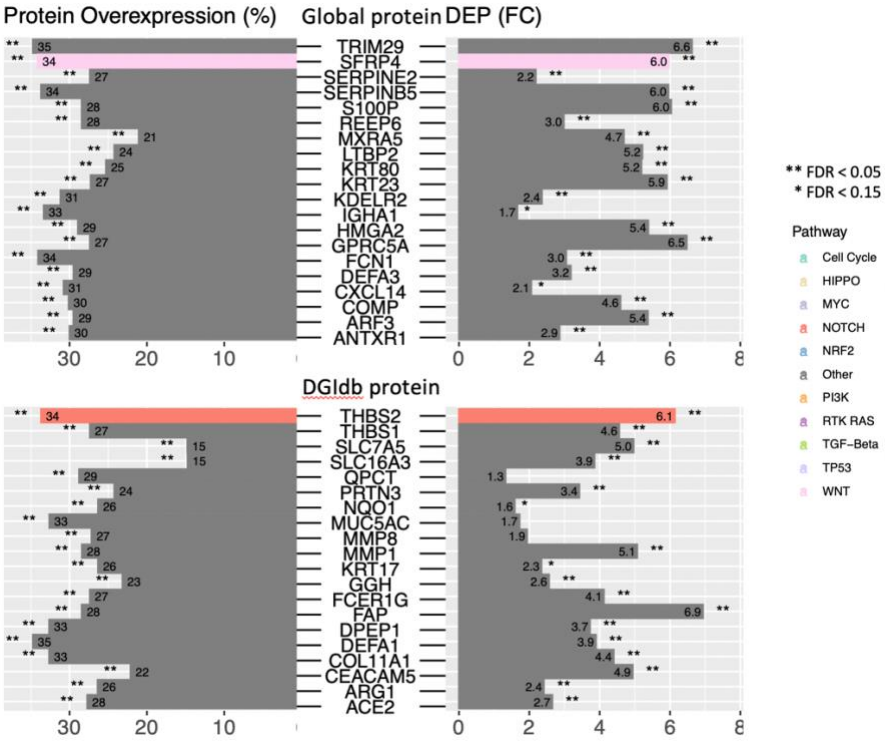

Figure S4. Potential CRC protein targets demonstrating protein overexpression and differential expression. The non-kinase proteins that showed notably higher expression in tumors relative to normal tissues and protein overexpression in the CRC cohort (above). The same analysis for the druggable proteins is given below.