

Title: Innate immune genes distinguish the immune microenvironment of early onset colorectal cancer

Authors:

Ivy H. Gardner MD¹, Ragavan Siddharthan MD¹, Katherine Watson MD¹, Elizabeth Dewey MS¹, Rebecca Ruhl BS^{2,3}, Xiangnan Guan PhD^{3,4,5}, Zheng Xia PhD^{3,4,5}, Liana V. Tsikitis MD, MCR, MBA^{1,3#}, and Sudarshan Anand PhD^{2,3,6#}

1. Department of Surgery,
 2. Department of Cell, Developmental and Cancer Biology,
 3. Knight Cancer Institute,
 4. Department of Molecular Microbiology & Immunology,
 5. Computational Biology Program,
 6. Department of Radiation Medicine,
- Oregon Health & Science University, 2720 S. Moody Avenue, Portland, OR

Corresponding Author:

Sudarshan Anand PhD

2720 S. Moody Avenue, KCRB 3004

Portland, OR 97201

503-494-8043

anands@ohsu.edu

These authors contributed equally

Conflict of interest: The authors have declared that no conflict of interest exists.

Abstract

Despite a decrease in the incidence of colorectal cancer (CRC) over the last 40 years, the incidence of CRC in people under 50 years old is increasing around the globe. Early onset (≤ 50 years old) and late onset (≥ 65 years old) CRC appear to have differences in their clinicopathological and genetic features, but it is unclear if there are differences in the tumor microenvironment. We hypothesized that the immune microenvironment of early onset CRC is distinct from late onset CRC and promotes tumor progression. We used Nanostring immune profiling to analyze mRNA expression of immune genes in FFPE surgical specimens from patients with early (N=40) and late onset (N=39) CRC. We found three genes, SAA1, C7, and CFD, have increased expression in early onset colorectal cancer and distinct immune signatures based on the tumor location. After adjusting for clinicopathological features, increased expression of CFD and SAA1 were associated with worse progression free survival and increased expression of C7 was associated with worse overall survival. Our data demonstrate that the immune microenvironment in early onset CRC is unique, location dependent, and associated with worse outcomes.

Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related death in the United States (1). Over the last 40 years, there has been an overall decrease in the incidence of CRC (1, 2) attributed to current screening guidelines (1, 2). Despite this decrease, there has been an increase in the incidence of CRC in young adults under the age of 50 over the same time period (2, 3). In particular, between 1980 and 2010, the incidence of rectal cancer in people aged 20-39 has quadrupled and it appears that the overall increase in early onset CRC (≤ 50 years old) is largely driven by an increase in rectal cancer in young patients (4). Based on these trends it is predicted that by the year 2030, the incidence of colon cancer will increase by 90% and rectal cancer will increase by 124% in people aged 20-34 (2). In addition, the rates of advanced disease manifested as localized, regional, and distant CRC have increased in patients 40-49 years old (5). This indicates that the increase in incidence is real and does not represent a shift in the age of diagnosis due to earlier detection with screening (5).

The etiology of the rise in early onset CRC is unclear, but it appears to be global (6–8). The proportion of CRC cases occurring in young patients has risen more in the highest income quartile, urban population, and Western states in the United States (3, 9). This rise in early onset CRC does not correlate with patterns of obesity or heavy alcohol use (9). Young patients with CRC are more likely to present with regional or metastatic disease, high grade, mucinous, and signet ring tumors (3, 7, 8, 10–16). Early onset is more commonly located in the rectum and left colon compared to late onset CRC (7, 14–16). Patients with early onset CRC appear to have similar or improved stage-specific and overall survival (7, 10, 11, 13, 16).

The majority of early onset CRC are sporadic with only 16% of early onset CRC cases having genetic mutations (17). About 10-21% of early onset CRCs are mismatch repair deficient with 6-8% having Lynch Syndrome and these cancers are usually right sided (17–19). In addition to distinct clinicopathological features, early onset CRC appears to have molecular differences to late onset CRC. Compared to late onset CRC, early onset CRC has lower rates of BRAF mutations and methylator phenotype (20, 21). Early onset microsatellite stable CRC has more mutations in TP53 and CTNNB1 compared to late onset microsatellite stable CRC (22). Several genes and key pathways in early onset CRC have been identified and include PEG10, the Wnt/beta catenin, MAP kinase, growth factor signaling, PI3KT-AKT, and TNFR1 pathways (20, 23–25).

The immune microenvironment in CRC is a key player in disease progression, therapy response, and overall survival (26–30). To our knowledge the role of the immune microenvironment in early onset CRC has not been well-studied. We set out to characterize the differences between the immune microenvironment in early onset (≤ 50 years old) and late onset (≥ 65 years old) CRC in a well-annotated cohort of patients at our tertiary cancer center.

Combined Results and Discussion

Our institutional cohort consisted of 40 early onset and 39 late onset CRC patients. A higher percentage of the early onset CRC group had metastatic disease and therefore underwent more neoadjuvant and adjuvant treatments (Figure 1A). Additionally, our early onset CRC group had less microsatellite instability compared to the late onset group (Figure 1B). We performed gene expression profiling for 770 immune response genes using the NanoString Immune panel from RNA extracted from FFPE tissues. We found 3 immune genes significantly (adjusted $p < 0.1$) upregulated in early onset CRC (SAA1, C7, CFD) and 7 genes upregulated in late onset CRC (CXCL3, CCL18, IL8, DMBT1, CEACAM6, CXCL5, CXCL10) (Figure 2A, 2E, and Supplemental Figure 1). Right colon, left colon, and rectal cancer have differentially expressed immune genes between early and late onset groups (Figure 2B-D). The increased expression of SAA1, C7, and CFD in early onset CRC is largely driven by an increase in rectal cancer, although C7 was also increased in early onset right colon cancers. We identified 10 genes of interest that were differentially expressed between early and late onset CRC in the whole cohort and at least one location specific group (SAA1, C7, CFD, CXCL12, DMBT1, CEACAM6, CCL18, IL8, CXCL3, CXCL10).

We validated the increased expression of one gene (CXCL3) in late onset colorectal cancer and another gene (SAA1) in early onset rectal cancer using qRT-PCR assays from the same patient samples (Figure 3A-B). We also found that SAA1 gene expression was higher in a group of liver metastases, but the significance of this finding is unclear because we did not have access to non-tumor liver samples within our colorectal cancer registry to compare to. Additionally, SAA1 expression was lower in a group non-tumor samples compared to early onset rectal cancer. We

analyzed a publicly available dataset (GSE87211) of gene expression microarray profiles from rectal cancer (N=105) and matched adjacent mucosa specimens (N=84) to further validate our findings (31). In this rectal cancer dataset, we found a difference in CXCL3 expression between tumor and mucosa in both groups, but no difference in gene expression between early and late onset rectal tumor specimens (Figure 3C). This is congruent with our findings since CXCL3 is not differentially expressed in rectal cancer in our cohort. We found that SAA1 levels were increased in the early onset tumors compared to late onset tumors, but there was no difference in expression between age groups in the adjacent mucosa (Figure 3D). This suggests that the upregulation of this gene may not simply be a product of immune aging, but more specifically due to the tumor immune microenvironment in early onset CRC.

We found several associations between the expression of our genes of interest and the clinical characteristics, pathological features, and outcomes of our patient population. Increased expression of C7 is associated with neoadjuvant treatment, right colon location, metastatic disease, angiolymphatic and perineural invasion (Supplemental Figure 2). Elevated CFD is only associated with neoadjuvant treatment and SAA1 does not appear to be associated with any clinical or pathological features (Supplemental Figure 2). After adjusting for age, metastatic disease, angiolymphatic invasion, perineural invasion, tumor grade, and neoadjuvant treatment, we found increased expression of SAA1 and CFD are associated with worse progression free survival (Figure 4A) and increased expression of C7 is associated with worse overall survival (Figure 4B). High expression of C7 and CFD was also associated with worse overall survival in The Cancer Genome Atlas Colon Adenocarcinoma (COAD) cancer cohort (Supplemental Figure

3). The combination of SAA1, C7, and CFD as an immune signature may have a role in predicting poor progression free and overall survival.

We found that elements of complement (C7, CFD) and SAA1 have increased expression in early onset CRC and are associated with worse outcomes. Analysis of gene set enrichment signatures from our profiles indicate complement and B-cell function is the most enriched pathway in early onset CRC (Supplementary Figures 4). C7 and CFD play a role in the terminal and alternative pathways of complement, but components of complement also have a non-canonical role that modulates the immune system and plays a role in the tumor immune microenvironment (32). Complement can play a pro-tumorigenic or anti-tumor effect in the tumor immune microenvironment and it is unclear what role C7 and CFD play in CRC (32). Complement activation in intestinal epithelial cells has been shown to be associated with inflammation; therefore the increase in C7 and CFD in early onset CRC may be related to an underlying inflammatory process (33). Similar to complement, SAA1 has also been shown to be overexpressed in patients with active inflammatory bowel disease (34, 35). SAA1 is an acute phase reactant produced by the liver, but is also present in multiple cancers and can be used as a biomarker for disease burden (36–38). SAA1 has been shown to direct pathogenic Th17 cell differentiation (34, 35). It is unclear if the increased expression of CFD, C7, and SAA1 in early onset CRC is the result of underlying immune dysregulation leading to a tumor-permissive microenvironment or an inappropriate immunological response to malignancy.

One of the hallmarks of immune aging is the phenomenon of “inflammaging” which is characterized by low levels of persistent inflammation both systemically and within the tissue

microenvironment (39). Although the majority of research on immune system aging has been on adaptive immunity, the phenomenon of inflammaging is now largely attributed to innate immune system dysfunction (39). Immune aging appears to be driven by a complex interaction between genetic and environmental factors and appears to exist along a trajectory towards an older immune system with people advancing across the trajectory at different speeds (40). Our findings paradoxically suggest that early onset CRC is associated with an inflammatory phenotype which is what would be expected in an older population. It may be that early onset CRC is associated with expedited inflammaging that leads to a tumor-permissive microenvironment. We found that SAA1, an acute phase protein, is elevated in early onset CRC which is akin to the elevation of another acute phase protein, C-reactive protein, in inflammaging. Studies have shown that older individuals have reduced microbiota diversity, decreased density of beneficial bacteria, and weakened integrity of the intestinal barrier (41). It is hypothesized that these factors allow leakage of proinflammatory microbial products into the systemic circulation and peripheral tissues which leads to the chronic inflammation associated with aging (41, 42). The microbiome and dysregulation of the intestinal barrier may play a role in the etiology of early onset CRC.

Our findings point to the innate immune system, specifically complement, as a component of the immune microenvironment that distinguishes early onset CRC. Although complement has been shown to have both protumor and antitumor effects, therapeutics targeting complement are currently being investigated. The majority of studies on complement inhibition in the treatment of cancer have been done in mice. Both C3a and C5a receptors may be potential targets for cancer immunotherapy and the inhibition of C5a in mice has been shown to decrease metastasis in colon cancer (32, 43, 44). Additionally, inhibition of the C5a receptor in mice reduces cancer

promoting inflammation that results from consumption of a high-fat diet (45). Of note, there is currently a phase I clinical trial for a C5aR1 monoclonal antibody in combination with durvalumab in advanced solid tumors (32). Our study suggests that early onset CRC may benefit from the addition of a complement inhibitor in addition to the standard use of chemotherapy and immunotherapy. Another possible therapeutic route in early onset CRC is manipulation of the microbiome. Mouse studies have shown that the microbiome appears to play a role in the effectiveness of chemotherapy and immunotherapy, including immune checkpoint inhibitors (46, 47). Manipulation of the microbiome by diet alterations, probiotics, antibiotics, or fecal transplant are being investigated as potential avenues to improve responsiveness to cancer treatments (47). Our findings suggest an inflammatory tumor promoting phenotype in early onset CRC and these treatments may be particularly useful in combating this immune microenvironment.

Our data demonstrate that the immune microenvironment is different between early and late onset CRC and these differences do not appear to solely be the result of age-related immune changes. Additionally, differences in the immune microenvironment between early and late onset CRC are location specific. *C7*, *CFD*, and *SAA1* are increased in early onset colorectal cancer and are associated with worse outcomes. These immune genes are associated with intestinal inflammation, suggesting that the immune microenvironment in early onset CRC may be pro-inflammatory and tumor permissive as a result of immune dysregulation.

Methods

The Oregon Colorectal Cancer Registry was queried to find patients with CRC who underwent surgical resection between January 2008 and October 2019 and were either under 50 or over 65 years old at the time of their diagnosis. In addition, six patients who underwent sigmoid colectomy for diverticulitis were identified to serve as a control group. The electronic medical record was used to obtain patient demographics, neoadjuvant treatment, pathology, genetic mutational status, adjuvant treatment, and outcomes.

Formalin-fixed paraffin-embedded (FFPE) slides were reviewed by pathology and only used if at least 70% of the tissue was tumor. The FFPE blocks were sectioned into 5 μ m-thick slides of tissue for use. Hematoxylin and Eosin stained slides were prepared for each sample to identify the tumor and were used as guides to sharply dissect the tumor from the unstained slides to be used for RNA extraction. RNA was isolated using the RNeasy FFPE kit (Qiagen). RNA concentration and purity was confirmed with the NanoDrop spectrophotometer (ThermoFisher Scientific).

NanoString immune profiling was performed using the human PanCancer Immune panel (cat#115000132) and run on the NanoString nCounter SPRINT profiler. We chose to use the NanoString platform since this platform had better tolerance for moderate RNA integrity scores typical of archival FFPE specimens. 100ng of RNA extracted from the FFPE tissue was used per the manufacturer's instructions.

Quantitative-PCR analysis was performed to validate expression of two genes identified as differentially expressed based on the NanoString analysis using TaqMan Gene Expression Assays (Applied Biosystems, Cat# 4448892, SAA1 Hs00761940_s1 and CXCL3 Hs00171061_m1) and the ViiA7 qRT-PCR machine (Applied Biosystems). Two housekeeping genes, GAPDH and ACTB (Applied Biosystems, Cat#433182 (Hs02758991_g1, Hs001060665_g1)), were used for all samples. Gene expression was reported as the change in the cycle threshold ($-2^{\Delta Ct}$, normalized to housekeeping genes).

Statistics

Early and late onset patient cohorts were compared using a two-tailed unpaired T-test for continuous variables and chi-squared test was used for categorical data. Raw NanoString immune gene expression counts were normalized to the best 10 housekeeping genes. The nSolver Advanced Analysis module was used to derive differential expression signatures according to NanoString recommendations. Adjusted P-values were obtained with a Benjamini-Hochberg post-hoc correction. The genes with an adjusted p value of less than 0.2 in the entire colorectal cancer cohort and at least one location specific cohort (right colon, left colon, or rectum) were designated genes of interest. This short list of 10 genes was then used for further analysis. Hazard ratios were calculated using multivariable Cox Proportional Hazards, and all estimates were adjusted for age, metastasis at diagnosis, neoadjuvant therapy, angiolymphatic invasion or perineural invasion, and tumor grade. The PH assumption was verified using time-dependent covariates. Comparisons between gene expression for qPCR results was done with a two-tailed unpaired T-test with Welch's correction. Statistical analysis was performed using

GraphPad Prism version 8.4.2 (GraphPad Software Inc., San Diego, CA), and JMP 14.3 (SAS Inc. Cary, NC).

Public Datasets

The GSE87211 rectal cancer gene expression microarray dataset was used to compare expression of our genes of interest in rectal tumor specimens and matched adjacent mucosa (31). This dataset was downloaded and divided into early (<50 years old) and late onset (>65 years old) groups. Tumor and mucosa specimens were compared between groups using a two-tailed unpaired T-test with Welch's corrections.

To further examine the correlation of the up-regulated genes with the clinical survival outcomes in the independent and large datasets, we investigated their expressions in bulk RNA-seq data from The Cancer Genome Atlas (TCGA). The processed gene expression data (fragments per kilobase per million fragments) as well as clinical data for primary colon adenocarcinoma (COAD) solid tumors were downloaded using the Bioconductor *TCGAbiolinks* package (version 2.14.0). Patients with overall survival less than 10 days were excluded. The remaining COAD tumor samples were stratified into the high and low groups based on the top quantile and bottom quantile of the expression of each individual gene. The overall survival curves of these two groups of patients were estimated by the Kaplan-Meier method with statistical significance calculated by the log-rank test.

Study Approval

This study was approved by the OHSU Institutional Review Board (study #00021278). Only patients who had undergone informed consent to be included in the Oregon Colorectal Cancer Registry were included in this study.

Author Contributions

IG- planning experiments, conducting experiments, analyzing data, writing manuscript

RS- planning experiments, conducting experiments, analyzing data, editing manuscript

KW- planning experiments, conducting experiments, writing manuscript

ED- analyzing data, editing manuscript

RR- planning experiments, conducting experiments, analyzing data, editing manuscript

XG- analyzing data, editing manuscript

ZX- analyzing data, editing manuscript

LVT- planning experiments, writing manuscript

SA- planning experiments, analyzing data, writing manuscript

Acknowledgements

This work was supported by funding from NHLBI to S.A. (R01 HL137779 and R01 HL143803).

We thank members of the Anand lab for useful discussions.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA. Cancer J. Clin.* 2020;70(1):7–30.
2. Bailey CE et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. *JAMA Surg.* 2015;150(1):17–22.
3. Virostko J, Capasso A, Yankeelov TE, Goodgame B. Recent trends in the age at diagnosis of colorectal cancer in the US National Cancer Data Base, 2004-2015. *Cancer* 2019;125(21):3828–3835.
4. Tawadros PS et al. Adenocarcinoma of the rectum in patients under age 40 is increasing: Impact of signet-ring cell histology. *Dis. Colon Rectum* 2015;58(5):474–478.
5. Meester RGS, Mannalithara A, Lansdorp-Vogelaar I, Ladabaum U. Trends in Incidence and Stage at Diagnosis of Colorectal Cancer in Adults Aged 40 Through 49 Years, 1975-2015. *JAMA* 2019;321(19):1933–1934.
6. Hofseth LJ et al. Early-onset colorectal cancer: initial clues and current views. *Nat. Rev. Gastroenterol. Hepatol.* 2020;1–13.
7. Wang MJ et al. The prognostic factors and multiple biomarkers in young patients with colorectal cancer. *Sci. Rep.* 2015;5:1–11.
8. Frostberg E, Rahr HB. Clinical characteristics and a rising incidence of early-onset colorectal cancer in a nationwide cohort of 521 patients aged 18-40 years. *Cancer Epidemiol.* 2020;66:101704.
9. Siegel RL, Medhanie GA, Fedewa SA, Jemal A. State variation in early-onset colorectal cancer in the United States, 1995-2015. *J. Natl. Cancer Inst.* 2019;111(10):1104–1106.
10. O’Connell JB, Maggard MA, Liu JH, Etzioni DA, Ko CY. Are survival rates different for young and older patients with rectal cancer?. *Dis. Colon Rectum* 2004;47(12):2064–2069.

11. Abdelsattar ZM et al. Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. *Cancer* 2016;122(6):929–934.
12. Dozois EJ et al. Young-onset Colorectal Cancer in Patients with No Known Genetic Predisposition: Can we increase early recognition and improve outcome?. *Medicine (Baltimore)*. 2008;87(5):259–263.
13. Dinaux AM, Leijssen LGJ, Bordeianou LG, Kunitake H, Berger DL. Rectal Cancer in Patients Under 50 Years of Age. *J. Gastrointest. Surg.* 2017;21:1898–1905.
14. Gausman V et al. Risk Factors Associated With Early-Onset Colorectal Cancer. *Clin. Gastroenterol. Hepatol.* 2019;in press.
15. Georgiou A et al. Outcomes of patients with early onset colorectal cancer treated in a UK specialist cancer center. *Cancers (Basel)*. 2019;11(10):1–15.
16. Burnett-Hartman AN et al. Treatment patterns and survival differ between early-onset and late-onset colorectal cancer patients: the patient outcomes to advance learning network. *Cancer Causes Control* 2019;30(7):747–755.
17. Pearlman R et al. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. *JAMA Oncol.* 2017;3(4):464–471.
18. Antelo M et al. Lynch-like syndrome is as frequent as Lynch syndrome in early-onset nonfamilial nonpolyposis colorectal cancer. *Int. J. Cancer* 2019;145(3):705–713.
19. Perea J et al. Classifying early-onset colorectal cancer according to tumor location: New potential subcategories to explore. *Am. J. Cancer Res.* 2015;5(7):2308–2313.
20. Kirzin S et al. Sporadic early-onset colorectal cancer is a specific sub-type of cancer: A morphological, molecular and genetics study. *PLoS One* 2014;9(8).
21. Willauer AN et al. Clinical and molecular characterization of early-onset colorectal cancer.

Cancer 2019;125(12):2002–2010.

22. Lieu CH et al. Comprehensive Genomic Landscapes in Early and Later Onset Colorectal Cancer. *Clin. Cancer Res.* 2019;25(19):5852–5858.

23. Jandova J, Xu W, Nfonsam V. Sporadic early-onset colon cancer expresses unique molecular features. *J. Surg. Res.* 2016;204(1):251–260.

24. Nfonsam V, Xu W, Koblinski J, Jandova J. Gene Expression Analysis of Sporadic Early-Onset Rectal Adenocarcinoma. *Gastrointest. Cancer* 2016;1(1).

25. Watson K et al. Differential expression of PEG10 contributes to aggressive disease in early vs late onset colorectal cancer. *Dis. Colon Rectum* 2020;In press.

26. Pérez-Ruiz E, Berraondo P. Immunological landscape and clinical management of rectal cancer. *Front. Immunol.* 2016;7:61.

27. Markman JL, Shiao SL. Impact of the immune system and immunotherapy in colorectal cancer. *J. Gastrointest. Oncol.* 2015;6(2):208–223.

28. Wang Y, Cardell SL. The Yin and Yang of invariant natural killer T cells in tumor immunity-suppression of tumor immunity in the intestine. *Front. Immunol.* 2018;8:1945.

29. Kather JN, Halama N, Jaeger D. Genomics and emerging biomarkers for immunotherapy of colorectal cancer. *Semin. Cancer Biol.* 2018.

30. Mlecnik B et al. Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity* 2016;44(3):698–711.

31. Hu Y et al. Colorectal cancer susceptibility loci as predictive markers of rectal cancer prognosis after surgery. *Genes Chromosom. Cancer* 2018;57(3):140–149.

32. Roumenina LT, Daugan M V., Petitprez F, Sautès-Fridman C, Fridman WH. Context-dependent roles of complement in cancer. *Nat. Rev. Cancer* 2019;19(12):698–715.

33. Sina C, Kemper C, Derer S. The intestinal complement system in inflammatory bowel disease: Shaping intestinal barrier function. *Semin. Immunol.* 2018;37:66–73.
34. Tang MS et al. Integrated Analysis of Biopsies from Inflammatory Bowel Disease Patients Identifies SAA1 as a Link between Mucosal Microbes with T H 17 and T H 22 Cells. *Inflamm. Bowel Dis.* 2017;23(9):1544–1554.
35. Lee J-Y et al. Serum Amyloid a Proteins Induce Pathogenic T H17 Cells and Promote Inflammatory Disease. *Cell* 2020;180:79–91.
36. Sun L, Ye RD. Serum amyloid A1: Structure, function and gene polymorphism. *Gene* 2016;583(1):48–57.
37. Yang M et al. Serum amyloid A expression in the breast cancer tissue is associated with poor prognosis. *Oncotarget* 2016;7(24):35843–35852.
38. Ignacio RMC et al. Serum amyloid A predisposes inflammatory tumor microenvironment in triple negative breast cancer. *Oncotarget* 2019;10(4):511–526.
39. Pinti M et al. Aging of the immune system- focus on inflammation and vaccination. *Eur J Immunol* 2016;46(10):2286–2301.
40. Alpert A et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat. Med.* 2019;25(3):487–495.
41. Biragyn A, Ferrucci L. Gut dysbiosis: a potential link between increased cancer risk in ageing and inflammaging. *Lancet Oncol.* 2018;19(6):e295–e304.
42. Fransen F et al. Aged gut microbiota contributes to systemical inflammaging after transfer to germ-free mice. *Front. Immunol.* 2017;8:1–12.
43. Piao C et al. Complement 5a enhances hepatic metastases of colon cancer via monocyte chemoattractant protein-1-mediated inflammatory cell infiltration. *J. Biol. Chem.*

2015;290(17):10667–10676.

44. Piao C et al. Complement 5a stimulates macrophage polarization and contributes to tumor metastases of colon cancer. *Exp. Cell Res.* 2018;366(2):127–138.

45. Doerner SK et al. High-fat diet-induced complement activation mediates intestinal inflammation and neoplasia, independent of obesity. *Mol. Cancer Res.* 2016;14(10):953–965.

46. Hendler R, Zhang Y. Probiotics in the Treatment of Colorectal Cancer. *Medicines* 2018;5(3):101.

47. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The Influence of the Gut Microbiome on Cancer, Immunity, and Cancer Immunotherapy. *Cancer Cell* 2018;33(4):570–580.

(A) Clinical Characteristics of OHSU Patient Cohort

	Early Onset N=40	Late Onset N=39	p-value
Female	22 (55%)	16 (41%)	0.2139
BMI (range)	26.7 (19-59)	27 (20-43)	0.8739
Location			
Right colon	9 (22.5%)	16 (41%)	0.0826
Left colon	18 (45%)	9 (23.1%)	
Rectum	13 (32.5%)	14 (35.9%)	
Metastasis at diagnosis			
Liver	10 (25%)	3 (7.7%)	0.0461
Lung	0	0	
Liver and lung	3 (7.5%)	0	
Carcinomatosis	1 (2.5%)	2 (5.1%)	
Recurrent colorectal cancer	0	2 (5.1%)	0.5268
Inflammatory Bowel Disease	3 (7.5%)	1 (2.6%)	0.3171
Lynch Syndrome	4 (10%)	0	0.1659
Polyposis	1 (2.5%)	1 (2.6%)	0.9855
Neoadjuvant Treatment			
Chemoradiation	10 (25%)	4 (10.3%)	0.0198
Chemotherapy	9 (22.5%)	3 (7.7%)	
Adjuvant chemotherapy			
Chemotherapy only	21 (52.5%)	10 (25.6%)	0.0015
Chemotherapy and immunotherapy	6 (15%)	1 (2.6%)	

(B) Pathological Characteristics of OHSU Patient Cohort

	Early Onset N=40	Late Onset N=39	p-value
Mucinous or signet ring	4 (10%)	7 (17.9%)	0.3076
Staging ^a			
T1	2 (5%)	4 (10.3%)	0.7400
T2	5 (12.5%)	5 (12.8%)	
T3	24 (60%)	24 (61.5%)	
T4	9 (22.5%)	6 (15.4%)	
N0	18 (45%)	21 (53.8%)	0.1550
N1	14 (35%)	11 (28.3%)	
N2	8 (20%)	2 (5.1%)	
Response to treatment			
None	6 (15%)	2 (5.1%)	0.8583
Minimal	3 (7.5%)	0	
Partial	8 (20%)	5 (12.8%)	
Complete	1 (2.5%)	0	
Positive margins	3 (7.5%)	3 (7.7%)	0.9742
High grade tumor	4 (10%)	5 (12.8%)	0.6932
Angiolymphatic Invasion	19 (47.5%)	18 (46.2%)	0.9046
Perineural Invasion	13 (32.5%)	6 (15.4%)	0.0752
Perforation	2 (5%)	1 (2.6%)	0.5712
Genetics ^b			
Kras mutant	9/20 (45%)	3/12 (25%)	0.2579
Braf mutant	2/20 (10%)	3/12 (25%)	0.2579
Microsatellite instability	2/36 (5.6%)	6/27 (22.2%)	0.0493

^a4 late onset patients underwent transanal excisions so no N stage was available; ^bNot all patients underwent genetic testing (denominator is number of patients who were tested)

Figure 1: Clinical and pathological characteristics of institutional cohort. Descriptive statistics of (A) clinical and (B) pathological characteristics of early (N=40) and late (N=39) onset colorectal cancer cohorts. Early and late onset patient cohorts were compared using a two-tailed unpaired T-test for continuous variables and chi-squared test was used for categorical data.

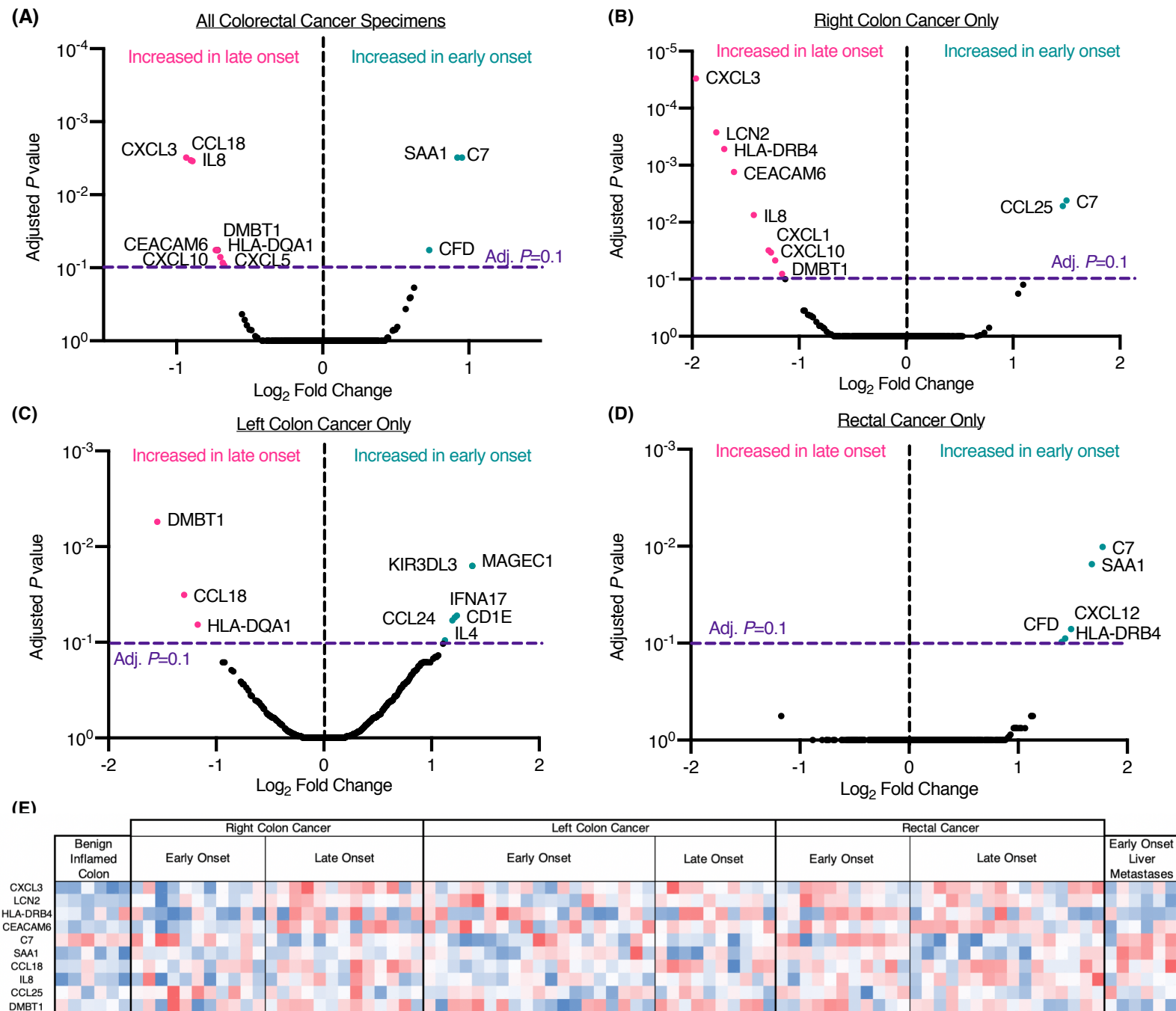


Figure 2: Immune microenvironment differences with age in colorectal cancer. Volcano plots depicts log₂ fold change in gene expression vs adjusted P-values of early vs late onset colorectal cancer patients. Nanostring Immune profiling panel (770 genes, human cancer immune panel) was used to generate expression profiles from RNA extracted from FFPE tissues slides. Raw counts were normalized to the best 10 housekeeping genes. The nSolver Advanced Analysis module was used to derive differential expression signatures according to Nanostring recommendations. Adjusted P-values were obtained with a Benjamini-Hochberg post-hoc correction. (A) Early (N=40) vs late onset (N=39) in all colorectal cancer specimens. (B) Right colon cancer specimens only (early N=10, late N=14). (C) Left colon cancer specimens only (early N=19, late N=10). (D) Rectal cancer specimens only (early N=11, late N=15). (E) Heatmap of the top 10 most differentially expressed immune genes with blue representing low relative gene expression and red representing high relative gene expression compared to other specimens.

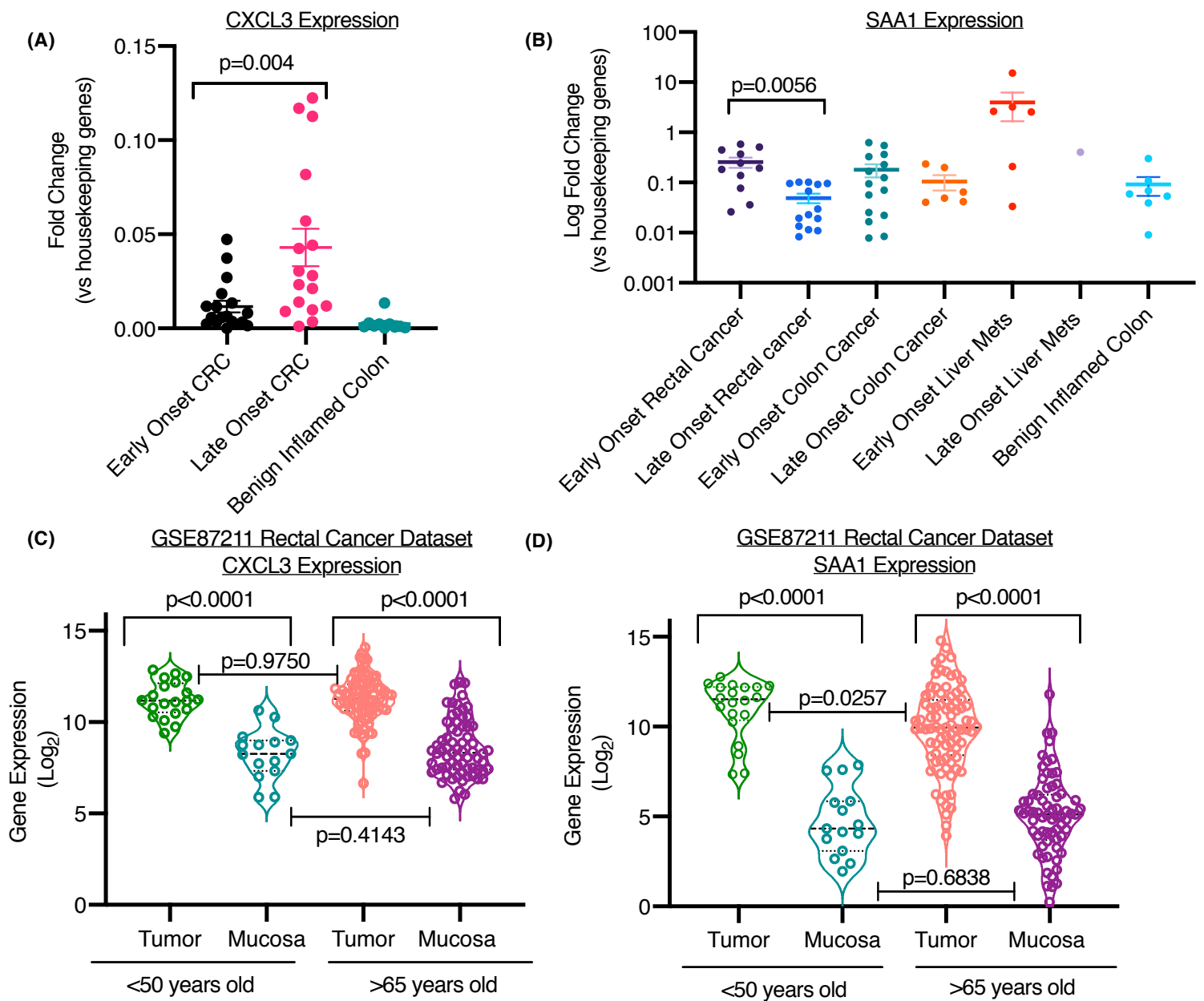


Figure 3: Validation of Nanostring data with independent qRT-PCR assays and independent datasets. (A) CXCL3 and (B) SAA1 expression of individual tumor RNA samples from FFPE slides measured using Taqman probes in a qRT-PCR assay. Change in gene expression ($2^{-\text{DCt}}$) are depicted as fold change over 2 housekeeping genes. Violin Plots depict (C) CXCL3 and (D) SAA1 gene expression from a publicly available Rectal Cancer dataset GSE87211. Values are log gene expression measured via microarrays in tumors and matched adjacent mucosal samples. P-values are from a two-tailed unpaired T-test with Welch's correction.

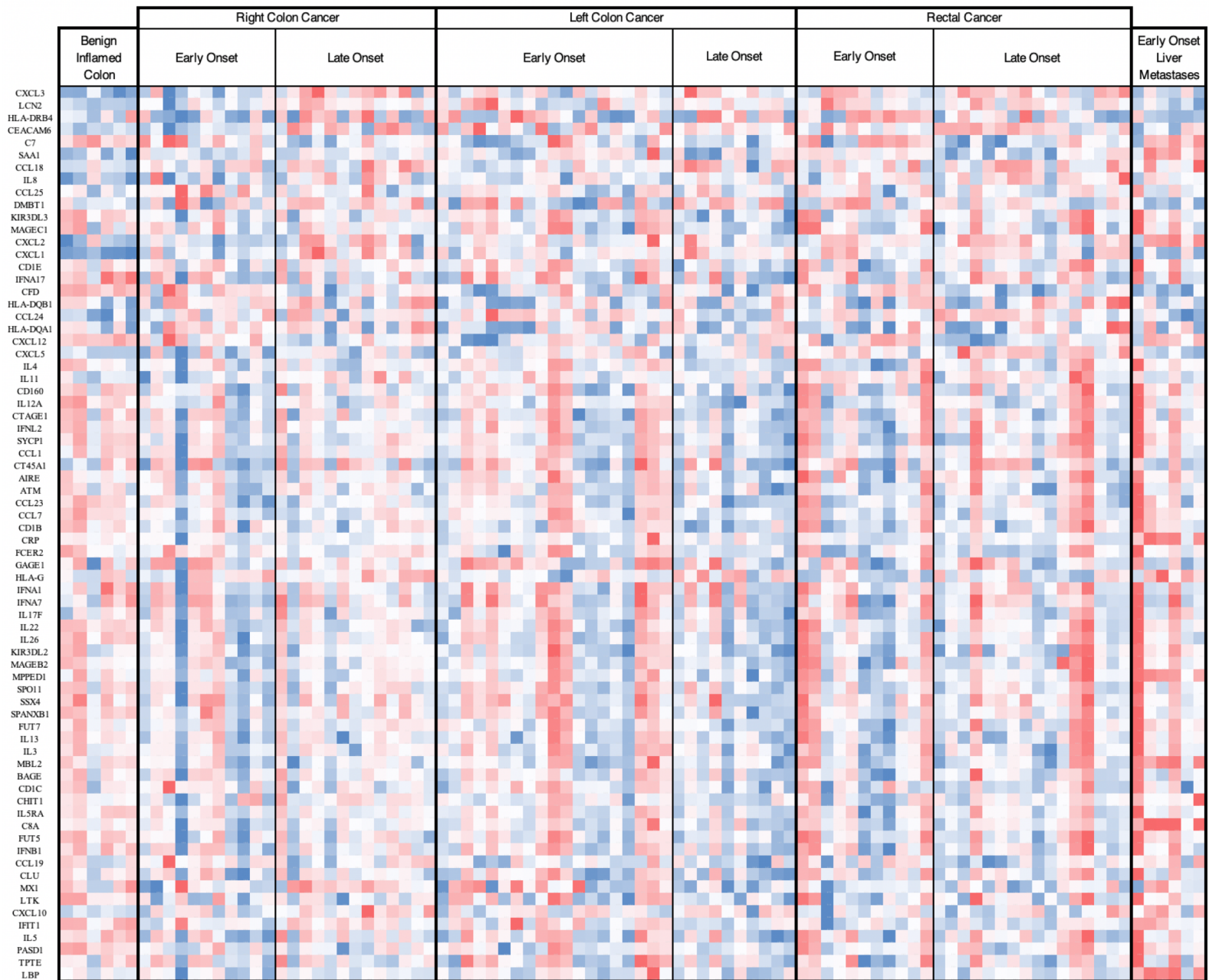
(A) Adjusted Hazard Ratios for Progression

Gene	HR	95% CI		p-value
		Lower	Upper	
SAA1	1.25	1.04	1.64	0.026
CFD	1.21	1.00	1.94	0.05
C7	1.21	0.97	1.56	0.08
C7+CFD+SAA1	1.33	1.05	1.32	0.009
IL8	0.97	0.77	1.16	0.77
DMBT1	0.99	0.85	1.13	0.89
CXCL12	1.02	0.78	1.28	0.84
CXCL3	0.97	0.70	1.22	0.76
CEACAM6	1.09	0.85	1.38	0.43
CXCL10	0.83	0.59	1.02	0.08
CCL18	1.01	0.74	1.27	0.90

(B) Adjusted Hazard Ratios for Death

Gene	HR	95% CI		p-value
		Lower	Upper	
SAA1	1.05	0.73	1.33	0.76
CFD	1.34	0.92	2.65	0.08
C7	1.65	1.23	3.55	0.01
C7+CFD+SAA1	1.42	1.00	1.46	0.051
IL8	0.91	0.62	1.19	0.58
DMBT1	1.00	0.77	1.23	0.99
CXCL12	1.12	0.74	1.58	0.49
CXCL3	0.90	0.55	1.23	0.55
CEACAM6	1.01	0.68	1.34	0.94
CXCL10	0.95	0.60	1.25	0.73
CCL18	0.97	0.60	1.29	0.85

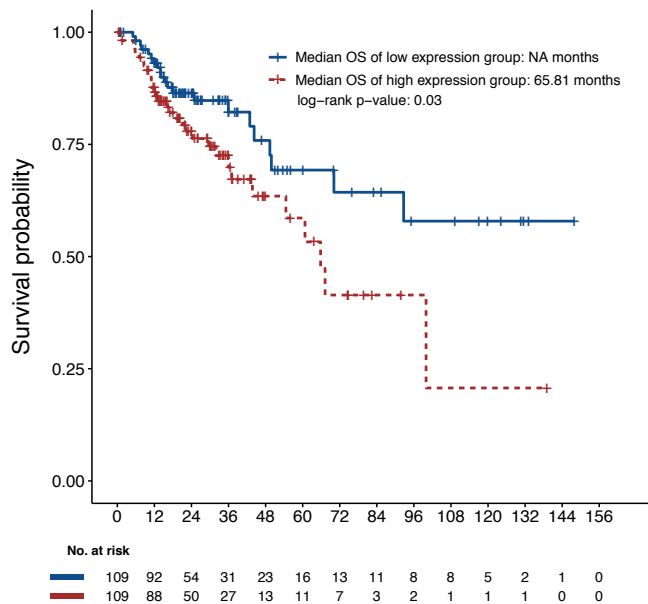
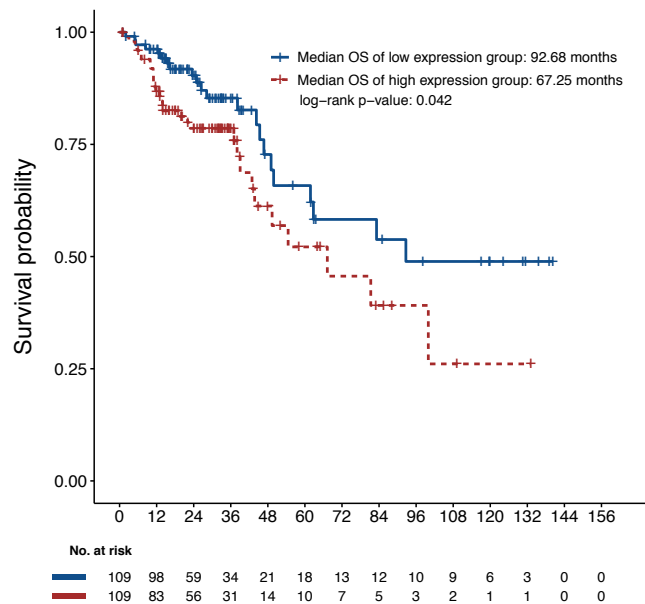
Figure 4: Adjusted Hazard Ratios. Hazard ratios for (A) progression and (B) death were calculated using multivariable Cox Proportional Hazards and are adjusted for age, metastasis at diagnosis, angiolymphatic invasion, perineural invasion, neoadjuvant therapy, and tumor grade. Hazard ratios are per half of a standard deviation from the mean of gene expression.



Supplemental Figure 1: Heatmap of differentially expressed immune genes between early and late onset colorectal cancer. Nanostring Immune profiling panel (770 genes, human immune panel) was used to generate expression profiles from RNA extracted from FFPE tissues slides. Raw counts were normalized to the best 10 housekeeping genes. The nSolver Advanced Analysis module was used to derive differential expression signatures according to Nanostring recommendations. Adjusted P-values were obtained with a Benjamini-Hochberg post-hoc correction. Genes with an adjusted p value >0.2 in all colorectal cancer, right colon cancer, left colon cancer, and rectal cancer specimens are included in the heatmap. The genes are ordered by p values with lowest p values at the top. Blue represents low relative gene expression and red represents high relative gene expression compared to other specimens.

↑ Gene Expression	Clinicopathologic Characteristics
C7	Neoadjuvant treatment (p=0.004), right colon location (p=0.034), metastatic disease (p=0.03), AVI/ PNI (p=0.048)
CFD	Neoadjuvant treatment (p=0.005)
SAA1	None
CCL8	None
CEACAM6	Low grade tumor (p=0.01)
CXCL10	No neoadjuvant treatment (p=0.001), high grade tumor (p=0.035), right colon location (p=0.008), no metastatic disease (p=0.028)
CXCL12	Neoadjuvant treatment (p=0.035)
CXCL3	No neoadjuvant treatment (p=0.001), no metastatic disease (p=0.006)
DMBT1	Low grade tumor (p=0.005)
IL8	No neoadjuvant treatment (p=0.008), no metastatic disease (p=0.007)

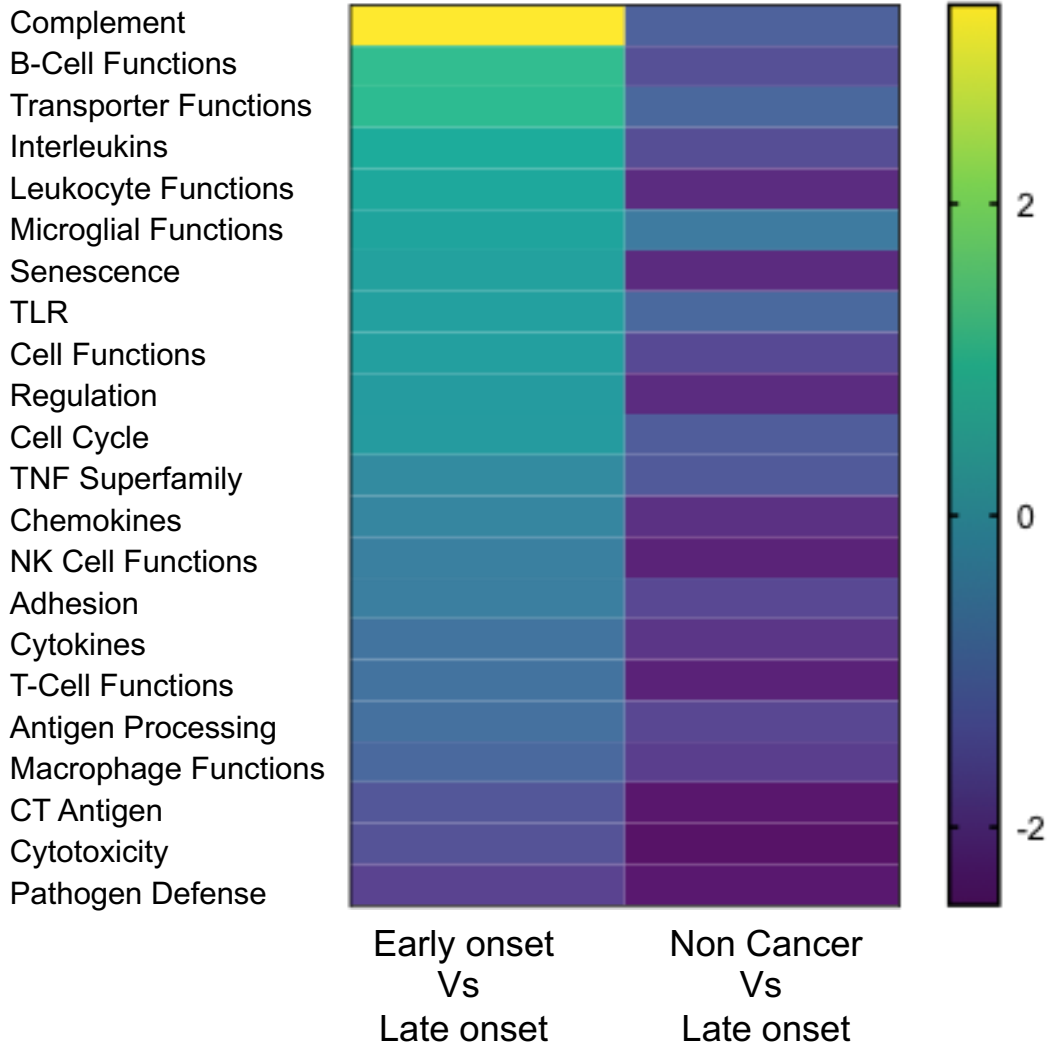
Supplemental Figure 2: Associations between immune gene expression and clinicopathologic characteristics. Associations between immune gene expression and neoadjuvant treatment, tumor grade, tumor location, metastatic disease, angiolymphatic invasion (AVI), and perineural invasion (PNI). Other clinicopathological features were not significantly associated with gene expression.

(A)**C7****(B)****CFD**

Supplemental Figure 5: Kaplan-Meier curves of colorectal adenocarcinoma patients from The Cancer Genome Atlas (TCGA).

Overall survival of TCGA colorectal adenocarcinoma patients in high (brown color line) and low (blue color line) expression groups of (A) C7 and (B) CFD. p-value was calculated by log-rank test. NA indicates that the median overall survival time is not reached. OS: overall survival.

Gene Set Enrichment Analysis



Supplemental Figure 4: Heatmap of global significance scores for gene set enrichment between early and late onset colorectal cancer. The global significance score is calculated as the square root of the mean squared t-statistic for the genes in a gene set, with t-statistics coming from the linear regression underlying our differential expression analysis.