Colorectal and Cancer Biology Research Literatures

Mark Herbert, PhD

39-06 Main Street, Flushing, Queens, New York 11354, USA, ma8080@gmail.com

Abstract: Cancer is the general name for a group of more than 100 diseases. Although there are many kinds of cancer, all cancers start because abnormal cells grow out of control. Untreated cancers can cause serious illness and death. The body is made up of trillions of living cells. Normal body cells grow, divide, and die in an orderly fashion. During the early years of a person's life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries. This article introduces recent research reports as references in the related studies.
[Mark H. **Colorectal and Cancer Biology Research Literatures.** *Academ Arena* 2019;11(7):56-92]. ISSN 1553-992X (print); ISSN 2158-771X (online). http://www.sciencepub.net/academia. 7. doi:10.7537/marsaaj110719.07.

Key words: cancer; life; research; literature; cell

1. Introduction

Cancer is the general name for a group of more than 100 diseases. Although there are many kinds of cancer, all cancers start because abnormal cells grow out of control. Untreated cancers can cause serious illness and death. The body is made up of trillions of living cells. Normal body cells grow, divide, and die in an orderly fashion. During the early years of a person's life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries. This article introduces recent research reports as references in the related studies.

The following introduces recent reports as references in the related studies.

Ali Hassan, N. Z., et al. (2014). "Integrated analysis of copy number variation and genome-wide expression profiling in colorectal cancer tissues." PLoS One **9**(4): e92553.

Integrative analyses of multiple genomic datasets for selected samples can provide better insight into the overall data and can enhance our knowledge of cancer. The objective of this study was to elucidate the association between copy number variation (CNV) and gene expression in colorectal cancer (CRC) samples and their corresponding non-cancerous tissues. Sixty-four paired CRC samples from the same patients were subjected to CNV profiling using the Illumina HumanOmni1-Quad assay, and validation was performed using multiplex ligation probe amplification method. Genome-wide expression profiling was performed on 15 paired samples from the same group of patients using the Affymetrix Human Gene 1.0 ST array. Significant genes obtained from both array results were then overlapped. To identify molecular pathways, the data were mapped to the KEGG database. Whole genome CNV analysis that compared primary tumor and non-cancerous epithelium revealed gains in 1638 genes and losses in 36 genes. Significant gains were mostly found in chromosome 20 at position 20q12 with a frequency of 45.31% in tumor samples. Examples of genes that were associated at this cytoband were PTPRT, EMILIN3 and CHD6. The highest number of losses was detected at chromosome 8, position 8p23.2 with 17.19% occurrence in all tumor samples. Among the genes found at this cytoband were CSMD1 and DLC1. Genome-wide expression profiling showed 709 genes to be up-regulated and 699 genes to be down-regulated in CRC compared to non-cancerous samples. Integration of these two datasets identified 56 overlapping genes, which were located in chromosomes 8, 20 and 22. MLPA confirmed that the CRC samples had the highest gains in chromosome 20 compared to the reference samples. Interpretation of the CNV data in the context of the transcriptome via integrative analyses may provide more in-depth knowledge of the genomic landscape of CRC.

Anderson, G. R., et al. (2001). "Intrachromosomal genomic instability in human sporadic colorectal cancer measured by genome-wide allelotyping and inter-(simple sequence repeat) PCR." Cancer Res **61**(22): 8274-8283.

We have used genome-wide allelotyping with 348 polymorphic autosomal markers spaced, on average, 10 cM apart to quantitate the extent of intrachromosomal instability in 59 human sporadic colorectal carcinomas. We have compared instability measured by this method with that measured by inter-(simple sequence repeat) PCR and microsatellite instability assays. Instability quantitated by fractional allelic loss rates was found to be independent of that detected by microsatellite instability analyses but was weakly associated with that measured by inter-(simple sequence repeat) PCR. A set of seven loci were identified that were most strongly associated with

elevated rates of fractional allelic loss and/or inter-(simple sequence repeat) PCR instability; these seven loci were on chromosomes 3, 8, 11, 13, 14, 18, and 20. A lesser association was seen with two loci flanking p53 on chromosome 17. Coordinate loss patterns for these loci suggest that at least two separate sets of cooperating loci exist for intrachromosomal genomic instability in human colorectal cancer.

Bajenova, O., et al. (2016). "The Genome-Wide Analysis of Carcinoembryonic Antigen Signaling by Colorectal Cancer Cells Using RNA Sequencing." PLoS One **11**(9): e0161256.

capital ES, Cyrillicarcinoembryonic antigen (CEA, CEACAM5, CD66) is a promoter of metastasis in epithelial cancers that is widely used as a prognostic clinical marker of metastasis. The aim of this study is to identify the network of genes that are associated with CEA-induced colorectal cancer liver metastasis. We compared the genome-wide transcriptomic profiles of CEA positive (MIP101 clone 8) and CEA negative (MIP 101) colorectal cancer cell lines with different metastatic potential in vivo. The CEA-producing cells displayed quantitative changes in the level of expression for 100 genes (over-expressed or down-regulated). They were confirmed by quantitative RT-PCR. The KEGG pathway analysis identified 4 significantly enriched pathways: cytokine-cytokine receptor interaction, MAPK signaling pathway, TGF-beta signaling pathway and pyrimidine metabolism. Our results suggest that CEA production by colorectal cancer cells triggers colorectal cancer progression by inducing the epithelial- mesenchymal transition, increasing tumor cell invasiveness into the surrounding tissues and suppressing stress and apoptotic signaling. The novel gene expression distinctions establish the relationships between the existing cancer markers and implicate new potential biomarkers for colorectal cancer hepatic metastasis.

Berg, M., et al. (2010). "Distinct high resolution genome profiles of early onset and late onset colorectal cancer integrated with gene expression data identify candidate susceptibility loci." Mol Cancer **9**: 100.

BACKGROUND: Estimates suggest that up to 30% of colorectal cancers (CRC) may develop due to an increased genetic risk. The mean age at diagnosis for CRC is about 70 years. Time of disease onset 20 years younger than the mean age is assumed to be indicative of genetic susceptibility. We have compared high resolution tumor genome copy number variation (CNV) (Roche NimbleGen, 385 000 oligo CGH array) in microsatellite stable (MSS) tumors from two age groups, including 23 young at onset patients without known hereditary syndromes and with a median age of 44 years (range: 28-53) and 17 elderly patients with median age 79 years (range: 69-87). Our aim was to identify differences in the tumor genomes between these groups and pinpoint potential susceptibility loci. Integration analysis of CNV and genome wide mRNA expression data, available for the same tumors, was performed to identify a restricted candidate gene list. RESULTS: The total fraction of the genome with aberrant copy number, the overall genomic profile and the TP53 mutation spectrum were similar between the two age groups. However, both the number of chromosomal aberrations and the number of breakpoints differed significantly between the groups. Gains of 2q35, 10q21.3-22.1, 10q22.3 and 19q13.2-13.31 and losses from 1p31.3, 1q21.1, 2q21.2, 4p16.1-q28.3, 10p11.1 and 19p12, positions that in total contain more than 500 genes, were found significantly more often in the early onset group as compared to the late onset group. Integration analysis revealed a covariation of DNA copy number at these sites and mRNA expression for 107 of the genes. Seven of these genes, CLC, EIF4E, LTBP4, PLA2G12A, PPAT, RG9MTD2, and ZNF574, had significantly different mRNA expression comparing median expression levels across the transcriptome between the two groups. CONCLUSIONS: Ten genomic loci, containing more than 500 protein coding genes, are identified as more often altered in tumors from early onset versus late onset CRC. Integration of genome and transcriptome data identifies seven novel candidate genes with the potential to identify an increased risk for CRC.

Bien, S. A., et al. (2017). "Enrichment of colorectal cancer associations in functional regions: Insight for using epigenomics data in the analysis of whole genome sequence-imputed GWAS data." PLoS One **12**(11): e0186518.

BACKGROUND: The evaluation of less frequent genetic variants and their effect on complex disease pose new challenges for genomic research. To investigate whether epigenetic data can be used to inform aggregate rare-variant association methods (RVAM), we assessed whether variants more significantly associated with colorectal cancer (CRC) were preferentially located in non-coding regulatory regions, and whether enrichment was specific to colorectal tissues. METHODS: Active regulatory elements (ARE) were mapped using data from 127 tissues and cell-types from NIH Roadmap Epigenomics and Encyclopedia of DNA Elements (ENCODE) projects. We investigated whether CRC association p-values were more significant for common variants inside versus outside AREs, or 2) inside colorectal (CR) AREs versus AREs of other tissues and cell-types. We employed an integrative epigenomic RVAM for variants with allele frequency

<1%. Gene sets were defined as ARE variants within 200 kilobases of a transcription start site (TSS) using either CR ARE or ARE from non-digestive tissues. CRC-set association p-values were used to evaluate enrichment of less frequent variant associations in CR ARE versus non-digestive ARE. RESULTS: ARE from 126/127 tissues and cell-types were significantly enriched for stronger CRC-variant associations. Strongest enrichment was observed for digestive tissues and immune cell types. CR-specific ARE were also enriched for stronger CRC-variant associations compared to ARE combined across non-digestive tissues (p-value = 9.6 x 10-4). Additionally, we found enrichment of stronger CRC association p-values for rare variant sets of CR ARE compared to non-digestive ARE (p-value = 0.029). CONCLUSIONS: Integrative epigenomic RVAM may enable discovery of less frequent variants associated with CRC, and ARE of digestive and immune tissues are most informative. Although distance-based aggregation of less frequent variants in CR ARE surrounding TSS showed modest enrichment, future association studies would likely benefit from joint analysis of transcriptomes and epigenomes to better link regulatory variation with target genes.

Broderick, P., et al. (2007). "A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk." Nat Genet **39**(11): 1315-1317.

To identify risk variants for colorectal cancer (CRC), we conducted a genome-wide association study, genotyping 550,163 tag SNPs in 940 individuals with familial colorectal tumor (627 CRC, 313 advanced adenomas) and 965 controls. We evaluated selected SNPs in three replication sample sets (7,473 cases, 5,984 controls) and identified three SNPs in SMAD7 (involved in TGF-beta and Wnt signaling) associated with CRC. Across the four sample sets, the association between rs4939827 and CRC was highly statistically significant (P (trend) = 1.0 x 10(-12)).

Crujeiras, A. B., et al. (2018). "Identification of an episignature of human colorectal cancer associated with obesity by genome-wide DNA methylation analysis." Int J Obes (Lond).

BACKGROUND: Obesity was established as a relevant modifiable risk factor in the onset and progression of colorectal cancer (CRC). This relationship could be mediated by an epigenetic regulation. OBJECTIVES: The current work aimed to explore the effects of excess body weight on the DNA methylation profile of CRC using a genome-wide DNA methylation approach and to identify an epigenetic signature of obesity-related CRC. METHODS: Fifty-six CRC-diagnosed patients (50 years) were included in the study and categorized according to their body mass index (BMI) as non-obese (BMI </= 25 kg/m (2)) or overweight/obese (BMI > 25 kg/m (2)). Data from Infinium 450k array-based methylomes of 28 CRC tumor samples were coupled with information on BMI categories. Additionally, DNA methylation results were validated in 28 CRC tumor samples. RESULTS: The analysis revealed statistically significant differences at 299 CpG sites, and they were mostly characterized as changes towards CpG hypermethylation occurring in the obese group. The 152 identified genes were involved in inflammatory and metabolic functional processes. Among these genes, novel genes were identified as epigenetically regulated in CRC depending on adiposity. ZNF397OS and ZNF543 represented the top scoring associated events that were further validated in an independent cohort and exhibited strong correlation with BMI and excellent and statistically significant efficiency in the discrimination of obese from non-obese CRC patients (area under the curve >0.80; p < 0.05). CONCLUSIONS: The present study identifies a potential epigenome mark of obesity-related CRC that could be useful for precision medicine in the management of this disease taking into account adiposity as a relevant risk factor.

Dehghanian, F., et al. (2018). "Reconstruction of the genome-scale co-expression network for the Hippo signaling pathway in colorectal cancer." Comput Biol Med **99**: 76-84.

The Hippo signaling pathway (HSP) has been identified as an essential and complex signaling pathway for tumor suppression that coordinates proliferation, differentiation, cell death, cell growth and stemness. In the present study, we conducted a genome-scale co-expression analysis to reconstruct the HSP in colorectal cancer (CRC). Five key modules were detected through network clustering, and a detailed discussion of two modules containing respectively 18 and 13 over and down-regulated members of HSP was provided. Our results suggest new potential regulatory factors in the HSP. The detected modules also suggest novel genes contributing to CRC. Moreover, differential expression analysis confirmed the differential expression pattern of HSP members and new suggested regulatory factors between tumor and normal samples. These findings can further reveal the importance of HSP in CRC.

Diep, C. B., et al. (2004). "Genome characteristics of primary carcinomas, local recurrences, carcinomatoses, and liver metastases from colorectal cancer patients." Mol Cancer **3**: 6.

BACKGROUND: Colorectal cancer (CRC) is one of the most common causes of cancer-related deaths in the Western world, and despite the fact that metastases are usually the ultimate cause of deaths, the knowledge of the genetics of advanced stages of this disease is limited. In order to identify potential genetic abnormalities underlying the development of local and distant metastases in CRC patients, we have, by comparative genomic hybridization, compared the DNA copy number profiles of 10 primary carcinomas, 14 local recurrences, 7 peritoneal carcinomatoses, and 42 liver metastases from 61 CRC patients. RESULTS: The median number of aberrations among the primary carcinomas, local recurrences, carcinomatoses, and liver metastases was 10, 6, 13, and 14, respectively. Several genetic imbalances, such as gains of 7, 8q, 13q, and 20, and losses of 4q, 8p, 17p, and 18, were common in all groups. In contrast, gains of 5p and 12p were more common in the carcinomatoses than in other stages of the disease. With hierarchical cluster analysis, liver metastases could be divided into two main subgroups according to clusters of chromosome changes. CONCLUSIONS: Each stage of CRC progression is characterized by a particular genetic profile, and both carcinomatoses and liver metastases are more genetically complex than local recurrences and primary carcinomas. This is the first genome profiling of local recurrences and carcinomatoses, and gains of 5p and 12p seem to be particularly important for the spread of the CRC cells within the peritoneal cavity.

Djureinovic, T., et al. (2006). "A genome wide linkage analysis in Swedish families with hereditary non-familial adenomatous polyposis/non-hereditary non-polyposis colorectal cancer." Gut **55**(3): 362-366.

BACKGROUND AND AIMS: Known colorectal cancer syndromes, such as familial adenomatous polyposis and hereditary non-polyposis colorectal cancer, have been identified in only a small proportion of cases with a family history of disease. In an attempt to identify loci harbouring novel predisposing genes, we have performed a genome wide linkage analysis in 18 colorectal cancer families recruited from the Department of Clinical Genetics at Karolinska Hospital, Sweden. METHODS: Multipoint parametric and non-parametric linkage analyses were performed using two affected status criteria, stringent and less stringent. Parametric analysis was performed under the assumption of locus homogeneity and locus heterogeneity. RESULTS: The initial scan performed using the less stringent affected status criteria revealed regions of interest on chromosome 11 (marker D11S1314: heterogeneity logarithm of odds (HLOD) score 1.96, non-parametric LOD (NPL) score 1.28; and marker D11S908: HLOD score 2.10, NPL score 2.16) and chromosome 14 (marker D14S258: HLOD score 2.61, NPL score 2.88). Using the stringent affected status criteria, a locus on chromosome 22 was suggested in the parametric analysis (marker D22S315: HLOD score 1.26). After finemapping of the regions on chromosomes 11 and 14, HLOD and NPL scores were reduced but still within the range of suggestive linkage. Haplotype analysis revealed overlapping regions between D11S987 and D11S4207 (proximal region), D11S4120 and D11S4090 (distal region), on chromosome 11, and between D14S1038 and D14S1069 on chromosome 14. CONCLUSION: Our study provides evidence of genetic heterogeneity among Swedish colorectal cancer families. Three novel regions were suggested to be of interest in a proportion of families analysed. Further studies are needed to confirm this result.

Du, M., et al. (2014). "No evidence of gene-calcium interactions from genome-wide analysis of colorectal cancer risk." Cancer Epidemiol Biomarkers Prev **23**(12): 2971-2976.

BACKGROUND: Calcium intake may reduce risk of colorectal cancer, but the mechanisms remain unclear. Studies of interaction between calcium intake and SNPs in calcium-related pathways have yielded inconsistent results. METHODS: To identify gene-calcium interactions, we tested interactions between approximately 2.7 million SNPs across the genome with self-reported calcium intake (from dietary or supplemental sources) in 9,006 colorectal cancer cases and 9,503 controls of European ancestry. To test for multiplicative interactions, we used multivariable logistic regression and defined statistical significance using the conventional genome-wide alpha = 5E-08. RESULTS: After accounting for multiple comparisons, there were no statistically significant SNP interactions with total, dietary, or supplemental calcium intake. CONCLUSIONS: We found no evidence of SNP interactions with calcium intake for colorectal cancer risk in a large population of 18,509 individuals. IMPACT: These results suggest that in genome-wide analysis common genetic variants do not strongly modify the association between calcium intake and colorectal cancer in European populations.

Dumenil, T. D., et al. (2014). "Genome-wide DNA methylation analysis of formalin-fixed paraffin embedded colorectal cancer tissue." Genes Chromosomes Cancer **53**(7): 537-548.

Formalin fixation and embedding of clinical tissue samples in paraffin is a common method for archiving biological material. These samples are often well annotated and provide an invaluable resource for research. However, this process of fixation and storage of tissue leads to DNA damage and fragmentation. The

use of DNA from formalin fixed, paraffin-embedded (FFPE) tissue to interrogate methylation levels on a genome-wide scale can pose challenges. We compared fresh and matched FFPE tissue DNA samples using the Illumina Infinium HD Human Methylation 450K BeadChip platform with a companion application for repair and "restoration" of DNA from FFPE tissue. Our results showed good correlation between fresh and FFPE sample data. FFPE DNA captured 99% of the CpG sites on the array on average. Significant cancer subgroups based on the CpG island methylator phenotype (CIMP) were clearly distinguished for both fresh and FFPE sample sets with cluster and scaling analysis. The DNA methylation status for the five standard CIMP panel genes which was evaluated for all samples by the MethyLight assay was correctly assigned in both fresh and FFPE samples by the array data. We conclude that the "restoration" method followed by assay on the Infinium HD Human Methylation 450K microarray can produce good quality data for DNA from FFPE samples.

Enroth, S., et al. (2011). "Cancer associated epigenetic transitions identified by genome-wide histone methylation binding profiles in human colorectal cancer samples and paired normal mucosa." BMC Cancer **11**: 450.

BACKGROUND: Despite their well-established functional roles, histone modifications have received less attention than DNA methylation in the cancer field. In order to evaluate their importance in colorectal cancer (CRC), we generated the first genome-wide histone modification profiles in paired normal colon mucosa and tumor samples. METHODS: Chromatin immunoprecipitation and microarray hybridization (ChIP-chip) was used to identify promoters enriched for histone H3 trimethylated on lysine 4 (H3K4me3) and lysine 27 (H3K27me3) in paired normal colon mucosa and tumor samples from two CRC patients and for the CRC cell line HT29. RESULTS: By comparing histone modification patterns in normal mucosa and tumors, we found that alterations predicted to have major functional consequences were quite rare. Furthermore, when normal or tumor tissue samples were compared to HT29, high similarities were observed for H3K4me3. However, the differences found for H3K27me3, which is important in determining cellular identity, indicates that cell lines do not represent optimal tissue models. Finally, using public expression data, we uncovered previously unknown changes in CRC expression patterns. Genes positive for H3K4me3 in normal and/or tumor samples, which are typically already active in normal mucosa, became hyperactivated in tumors, while genes with H3K27me3 in normal and/or tumor samples and which are expressed at low levels in normal mucosa, became hypersilenced in tumors. CONCLUSIONS: Genome wide histone modification profiles can be used to find epigenetic aberrations in genes associated with cancer. This strategy gives further insights into the epigenetic contribution to the oncogenic process and may identify new biomarkers.

Eskiocak, U., et al. (2011). "Functional parsing of driver mutations in the colorectal cancer genome reveals numerous suppressors of anchorage-independent growth." Cancer Res **71**(13): 4359-4365.

Landmark cancer genome resequencing efforts are leading to the identification of mutated genes in many types of cancer. The extreme diversity of mutations being detected presents significant challenges to subdivide causal from coincidental mutations to elucidate how disrupted regulatory networks drive cancer processes. Given that a common early perturbation in solid tumor initiation is bypass of matrix-dependent proliferation restraints, we sought to functionally interrogate colorectal cancer candidate genes (CAN-genes) to identify driver tumor suppressors. We have employed an isogenic human colonic epithelial cell (HCEC) model to identify suppressors of anchorage-independent growth by conducting a soft agar-based short hairpin RNA (shRNA) screen within the cohort of CAN-genes. Remarkably, depletion of 65 of the 151 CAN-genes tested collaborated with ectopic expression of K-RAS (V12) and/or TP53 knockdown to promote anchorage-independent proliferation of HCECs. In contrast, only 5 of 362 random shRNAs (1.4%) enhanced soft agar growth. We have identified additional members of an extensive gene network specifying matrix-dependent proliferation, by constructing an interaction map of these confirmed progression suppressors with approximately 700 mutated genes that were excluded from CAN-genes, and experimentally verifying soft agar growth enhancement in response to depletion of a subset of these genes. Collectively, this study revealed a profound diversity of nodes within a fundamental tumor suppressor network that are susceptible to perturbation leading to enhanced cell-autonomous anchorage-independent proliferative fitness. Tumor suppressor network fragility as a paradigm within this and other regulatory systems perturbed in cancer could, in large part, account for the heterogeneity of somatic mutations detected in tumors.

Fang, W. J., et al. (2012). "Genome-wide analysis of aberrant DNA methylation for identification of potential biomarkers in colorectal cancer patients." Asian Pac J Cancer Prev **13**(5): 1917-1921.

BACKGROUND: Colorectal cancer is one of the leading causes of mortality worldwide. Genome wide

analysis studies have identified sequence mutations causing loss-of-function that are associated with disease occurrence and severity. Epigenetic modifications, such DNA methylation, have also been implicated in many cancers but have yet to be examined in the East Asian population of colorectal cancer patients. METHODS: Biopsies of tumors and matched non-cancerous tissue types were obtained and genomic DNA was isolated and subjected to the bisulphite conversion method for comparative DNA methylation analysis on the Illumina Infinium HumanMethylation27 BeadChip. RESULTS: Totals of 258 and 74 genes were found to be hyper- and hypo-methylated as compared to the individual's matched control tissue. Interestingly, three genes that exhibited hypermethylation in their promoter regions, CMTM2, ECRG4, and SH3GL3, were shown to be significantly associated with colorectal cancer in previous studies. Using heatmap cluster analysis, eight hypermethylated and 10 hypomethylated genes were identified as significantly differentially methylated genes in the tumour tissues. CONCLUSIONS: Genome-wide methylation profiling facilitates rapid and simultaneous analysis of cancerous cells which may help to identify methylation markers with high sensitivity and specificity for diagnosis and prognosis. Our results show the promise of the microarray technology in identification of potential methylation biomarkers for colorectal cancers.

Fehringer, G., et al. (2016). "Cross-Cancer Genome-Wide Analysis of Lung, Ovary, Breast, Prostate, and Colorectal Cancer Reveals Novel Pleiotropic Associations." Cancer Res **76**(17): 5103-5114.

Identifying genetic variants with pleiotropic associations can uncover common pathways influencing multiple cancers. We took a two-stage approach to conduct genome-wide association studies for lung, ovary, breast, prostate, and colorectal cancer from the GAME-ON/GECCO Network (61,851 cases, 61,820 controls) to identify pleiotropic loci. Findings were replicated in independent association studies (55,789 cases, 330,490 controls). We identified a novel pleiotropic association at 1q22 involving breast and lung squamous cell carcinoma, with eQTL analysis showing an association with ADAM15/THBS3 gene expression in lung. We also identified a known breast cancer locus CASP8/ALS2CR12 associated with prostate cancer, a known cancer locus at CDKN2B-AS1 with different variants associated with lung adenocarcinoma and prostate cancer, and confirmed the associations of a breast BRCA2 locus with lung and serous ovarian cancer. This is the largest study to date examining pleiotropy across multiple cancer-associated loci, identifying common mechanisms of cancer development and progression. Cancer Res; 76(17); 5103-14. (c)2016 AACR.

Fernandez-Rozadilla, C., et al. (2013). "Pharmacogenomics in colorectal cancer: a genome-wide association study to predict toxicity after 5-fluorouracil or FOLFOX administration." Pharmacogenomics J **13**(3): 209-217.

The development of genotyping technologies has allowed for wider screening for inherited causes of variable outcomes following drug administration. We have performed a genome-wide association study (GWAS) on 221 colorectal cancer (CRC) patients that had been treated with 5-fluorouracil (5-FU), either alone or in combination with oxaliplatin (FOLFOX). A validation set of 791 patients was also studied. Seven SNPs (rs16857540, rs2465403, rs10876844, rs10784749, rs17626122, rs7325568 and rs4243761) showed evidence of association (pooled P-values 0.020, 9.426E-03, 0.010, 0.017, 0.042, 2.302E-04, 2.803E-03) with adverse drug reactions (ADRs). This is the first study to explore the genetic basis of inter-individual variation in toxicity responses to the administration of 5-FU or FOLFOX in CRC patients on a genome-wide scale.

Fernandez-Rozadilla, C., et al. (2014). "A genome-wide association study on copy-number variation identifies a 11q11 loss as a candidate susceptibility variant for colorectal cancer." Hum Genet **133**(5): 525-534.

Colorectal cancer (CRC) is a complex disease, and therefore its development is determined by the combination of both environmental factors and genetic variants. Although genome-wide association studies (GWAS) of SNP variation have conveniently identified 20 genetic variants so far, a significant proportion of the observed heritability is yet to be explained. Common copy-number variants (CNVs) are one of the most important genomic sources of variability, and hence a potential source to explain part of this missing genetic fraction. Therefore, we have performed a GWAS on CNVs to explore the relationship between common structural variation and CRC development. Phase 1 of the GWAS consisted of 881 cases and 667 controls from a Spanish cohort. Copy-number status was validated by quantitative PCR for each of those common CNVs potentially associated with CRC in phase I. Subsequently, SNPs were chosen as proxies for the validated CNVs for phase II replication (1,342 Spanish cases and 1,874 Spanish controls). Four common CNVs were found to be associated with CRC and were further replicated in Phase II. Finally, we found that SNP rs1944682, tagging a 11q11 CNV, was nominally associated with

CRC susceptibility (p value = 0.039; OR = 1.122). This locus has been previously related to extreme obesity phenotypes, which could suggest a relationship between body weight and CRC susceptibility.

Fernandez-Rozadilla, C., et al. (2013). "A colorectal cancer genome-wide association study in a Spanish cohort identifies two variants associated with colorectal cancer risk at 1p33 and 8p12." BMC Genomics **14**: 55.

BACKGROUND: Colorectal cancer (CRC) is a disease of complex aetiology, with much of the expected inherited risk being due to several common low risk variants. Genome-Wide Association Studies (GWAS) have identified 20 CRC risk variants. Nevertheless, these have only been able to explain part of the missing heritability. Moreover, these signals have only been inspected in populations of Northern European origin. RESULTS: Thus, we followed the same approach in a Spanish cohort of 881 cases and 667 controls. Sixty-four variants at 24 loci were found to be associated with CRC at p-values <10-5. We therefore evaluated the 24 loci in another Spanish replication cohort (1481 cases and 1850 controls). Two of these SNPs, rs12080929 at 1p33 (Preplication=0.042; Ppooled=5.523x10-03; OR (CI95%)=0.866(0.782-0.959)) and rs11987193 at 8p12 (Preplication=0.039; Ppooled=6.985x10-5; OR (CI95%)=0.786(0.705-0.878)) were replicated in the second Phase, although they did not reach genome-wide statistical significance. CONCLUSIONS: We have performed the first CRC GWAS in a Southern European population and by these means we were able to identify two new susceptibility variants at 1p33 and 8p12 loci. These two SNPs are located near the SLC5A9 and DUSP4 loci, respectively, which could be good functional candidates for the association signals. We therefore believe that these two markers constitute good candidates for CRC susceptibility loci and should be further evaluated in other larger datasets. Moreover, we highlight that were these two SNPs true susceptibility variants, they would constitute a decrease in the CRC missing heritability fraction.

Figueiredo, J. C., et al. (2014). "Genome-wide diet-gene interaction analyses for risk of colorectal cancer." PLoS Genet **10**(4): e1004228.

Dietary factors, including meat, fruits, vegetables and fiber, are associated with colorectal cancer; however, there is limited information as to whether these dietary factors interact with genetic variants to modify risk of colorectal cancer. We tested interactions between these dietary factors and approximately 2.7 million genetic variants for colorectal cancer risk among 9,287 cases and 9,117 controls from ten studies. We used logistic regression to investigate multiplicative gene-diet interactions, as well as our recently developed Cocktail method that involves a screening step based on marginal associations and gene-diet correlations and a testing step for multiplicative interactions, while correcting for multiple testing using weighted hypothesis testing. Per quartile increment in the intake of red and processed meat were associated with statistically significant increased risks of colorectal cancer and vegetable, fruit and fiber intake with lower risks. From the case-control analysis, we detected a significant interaction between rs4143094 (10p14/near GATA3) and processed meat consumption (OR = 1.17; p = 8.7E-09), which was consistently observed across studies (p heterogeneity = 0.78). The risk of colorectal cancer associated with processed meat was increased among individuals with the rs4143094-TG and -TT genotypes (OR = 1.20 and OR = 1.39, respectively) and null among those with the GG genotype (OR = 1.03). Our results identify a novel gene-diet interaction with processed meat for colorectal cancer, highlighting that diet may modify the effect of genetic variants on disease risk, which may have important implications for prevention.

Figueiredo, J. C., et al. (2011). "Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: results from a genome-wide association study." Cancer Epidemiol Biomarkers Prev **20**(5): 758-766.

BACKGROUND: Genome-wide association studies (GWAS) have led to the identification of a number of common susceptibility loci for colorectal cancer (CRC); however, none of these GWAS have considered gene-environment (G x E) interactions. Therefore, it is unclear whether current hits are modified by environmental exposures or whether there are additional hits whose effects are dependent on environmental exposures. METHODS: We conducted a systematic search for G x E interactions using genome wide data from the Colon Cancer Family Registry that included 1,191 cases of microsatellite stable (MSS) or microsatellite instability-low (MSI-L) CRC and 999 controls genotyped using either the Illumina Human1M or Human1M-Duo BeadChip. We tested for interactions between genotypes and 14 environmental factors using 3 methods: a traditional case-control test, a case-only test, and the recently proposed 2-step method by Murcray and colleagues. All potentially significant findings were replicated in the ARCTIC Study. RESULTS: No G x E interactions were identified that reached genome-wide significance by any of the 3 methods. When analyzing previously reported susceptibility loci, 7 significant G x E interactions were found at a 5% significance level. We investigated these 7 interactions in an independent

sample and none of the interactions were replicated. CONCLUSIONS: Identifying G x E interactions will present challenges in a GWAS setting. Our power calculations illustrate the need for larger sample sizes; however, as CRC is a heterogeneous disease, a tradeoff between increasing sample size and heterogeneity needs to be considered. IMPACT: The results from this first genome-wide analysis of G x E in CRC identify several challenges, which may be addressed by large consortium efforts.

Furuta, K., et al. (2012). "Integrated analysis of whole genome exon array and array-comparative genomic hybridization in gastric and colorectal cancer cells." Cancer Sci **103**(2): 221-227.

Whole genome-scale integrated analyses of exon array and array-comparative genomic hybridization are expected to enable the identification of unknown genetic features of cancer cells. Here, we evaluated this approach in 22 gastric and colorectal cancer cell lines, focusing on protein kinase genes and genes belonging to the cadherin-catenin family. Regarding alternative splicing patterns, several cancer cell lines predominantly expressed isoform 1 of protein kinase A catalytic subunit beta (PRKACB). Paired gastric cancer specimens demonstrated that isoform 1 of PRKACB was a novel cancer-related variant transcript in gastric cancers. In addition, the exon array analysis clearly identified exon 3 or exon 3-4 skipping in catenin beta 1, a short intron insertion with exon 9 skipping in CDH1, and a deletional transcript of CDH13. These abnormal transcripts were shown to have arisen from small genomic deletions. Meanwhile, an integrated analysis of 11 gastric cancer cell lines revealed that four cell lines amplified fibroblast growth factor receptor 2, with truncated forms observed in two of the cell lines. Gene amplification, and not the truncated form, was found to determine the sensitivity to a fibroblast growth factor receptor inhibitor, indicating that our cell line panel might be useful for cell-based evaluations of specific inhibitors. Using an integrated analysis, we identified several abnormal transcripts and genomic alterations in gastric and colorectal cancer cells. Our approach might enable genetic changes to be identified more efficiently, and the present results warrant further investigation using clinical samples and integrated analyses.

Gaiser, T., et al. (2011). "Genome and transcriptome profiles of CD133-positive colorectal cancer cells." Am J Pathol **178**(4): 1478-1488.

Colorectal carcinomas (CRC) might be organized hierarchically and contain a subpopulation of tumorigenic, putative cancer stem cells that are CD133 positive. We studied the biological and genetic characteristics of such cells in CRC cell lines and primary tumors. Three CRC cell lines were sorted in CD133 positive and negative fractions. The respective genetic aberration profiles were studied using array comparative genomic hybridization (aCGH) and expression profiling. Tumorigenicity for each cellular population was tested by injection into nude mice. Additionally, we compared CD133+ and CD133- cells of 12 primary colorectal tumors using laser capture microdissection and aCGH. Three of five CRC cell lines displayed both CD133+ and CD133- cells, but tumorigenicity of these subfractions did not differ significantly and aCGH revealed essentially identical genomic imbalances. However, 96 genes were differentially expressed between the two populations. Array comparative genomic hybridization analysis after laser capture microdissection of CD133+ and CD133- areas in primary colorectal tumors revealed genetic differences in 7 of 12 cases. The use of cell lines for studying genomic alterations that define cancer stem cell characteristics, therefore, seems questionable. In contrast, CD133+ cells in primary cancer samples showed a unique genomic aberration profile. In conclusion, our data suggest that CD133 positivity defines a genetically distinct cellular compartment in primary CRC, which potentially includes tumor initiating cells.

Gao, B., et al. (2016). "Weighted gene co-expression network analysis of colorectal cancer liver metastasis genome sequencing data and screening of anti-metastasis drugs." Int J Oncol **49**(3): 1108-1118.

Approximately 9% of cancer-related deaths are caused by colorectal cancer (CRC). CRC patients are prone to liver metastasis, which is the most important cause for the high CRC mortality rate. Understanding the molecular mechanism of CRC liver metastasis could help us to find novel targets for the effective treatment of this deadly disease. Using weighted gene co-expression network analysis on the sequencing data of CRC with and with metastasis, we identified 5 colorectal cancer liver metastasis related modules which were labeled as brown, blue, grey, yellow and turquoise. In the brown module, which represents the metastatic tumor in the liver, gene ontology (GO) analysis revealed functions including the G-protein coupled receptor protein signaling pathway, epithelial cell differentiation and cell surface receptor linked signal transduction. In the blue module, which represents the primary CRC that has metastasized, GO analysis showed that the genes were mainly enriched in GO terms including G-protein coupled receptor protein signaling pathway, cell surface receptor linked signal transduction, and negative regulation of cell differentiation. In the yellow and turquoise modules, which represent the primary non-metastatic CRC, 13 downregulated CRC liver metastasis-related candidate

miRNAs were identified (e.g. hsa-miR-204, hsa-miR-455, etc.).

Gao, B., et al. (2017). "[Corrigendum] Weighted gene co-expression network analysis of colorectal cancer liver metastasis genome sequencing data and screening of anti-metastasis drugs." Int J Oncol **50**(1): 339.

After the publication of the article, the authors noted that the affiliation for Dr Hani Choudhry is wrong. The correct affiliation should be as follows: Bo Gao1, Qin Shao2, Hani Choudhry3, Victoria Marcus2, Kung Dong5, Jiannis Ragoussis4 and Zu-Hua Gao2, 1Department of General Surgery, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150001, P.R. China; 2Department of Pathology, The Research Institute of McGill University Health Center, Montreal, Quebec H4A 3J1, Canada; 3Department of Biochemistry, Faculty of Science, Cancer and Mutagenesis Unit, King Fahd Center for Medical Research, Center of Innovation in Personalized Medicine, King Abdulaziz University, Jeddah, Saudi Arabia; 4McGill University and Genome Quebec Innovation Centre, Montreal, Quebec H3B 1S6, Canada; 5Department of Pathology, Beijing Youan Hospital, Capital Medical University, Beijing 100069, P.R. China. [the original article was published in the International Journal of Oncology 49: 1108-1118, 2016; DOI: 10.3892/ijo.2016.3591].

Ghanadi, K., et al. (2016). "Colorectal cancer and the KIR genes in the human genome: A meta-analysis." Genom Data **10**: 118-126.

Colorectal cancer is one of the most common types of inflammation-based cancers and is occurred due to growth and spread of cancer cells in colon and/or rectum. Previously genetic association of cell cycle genes, both proto-oncogenes and the tumor suppressors has been proved. But there were few studies about association of immune related genes such as killer-cell immunoglobulin-like receptors (KIRs). Thus we intend to perform a meta-analysis to find the association of different genes of KIR and susceptibility to be affected by colorectal cancer. The overall population of the four studies investigated in our meta-analysis was 953 individuals (470 individuals with colorectal cancer and 483 individuals in control groups). After the analyses, we concluded that colorectal cancer is affected by KIR2DS5 and also there were no protecting gene. This result shows the inflammatory basis of this cancer. In other words, in contrast to leukemia and blood cancers, colorectal cancers seem to be affected by hyper activity of natural killer-cells (NKs). Whys and therefore of this paradox, is suggested to be investigated further.

Giraldez, M. D., et al. (2013). "Circulating microRNAs as biomarkers of colorectal cancer: results from a genome-wide profiling and validation study." Clin Gastroenterol Hepatol **11**(6): 681-688 e683.

BACKGROUND & AIMS: Circulating microRNAs (miRNAs/miRs) might be used as biomarkers for the diagnosis of cancer and other diseases. Noninvasive approaches are needed to complement and improve upon current strategies for colorectal cancer (CRC) screening. We investigated whether plasma levels of miRNA can differentiate patients with CRC from healthy individuals. We also investigated whether plasma samples from patients with premalignant neoplastic lesions, such as advanced adenomas (AAs), also had a different expression pattern of miRNAs. METHODS: We analyzed 196 plasma samples from 123 patients newly diagnosed with sporadic colorectal neoplasia (63 with CRC and 60 with AAs) and 73 healthy individuals (controls) seen at 2 tertiary medical centers in Spain. An initial set of samples was analyzed using a genome-wide miRNA expression profiling assay (n = 61). Quantitative reverse-transcription PCR was used to validate the expression of selected miRNAs in an independent cohort (n = 135). RESULTS: Patients with CRC or AAs had plasma miRNA expression profiles that differed significantly from those of controls. We selected a group of 13 miRNAs for validation in an independent cohort of patients; 6 (miR18a, miR19a, miR19b, miR15b, miR29a, and miR335) were confirmed to be significantly up-regulated in patients with CRC, differentiating patients with CRC from controls with area under the receiver operating characteristic curve values ranging from 0.80 (95% confidence interval [CI], 0.71-0.89) to 0.70 (95% CI, 0.59-0.80). Only miR18a was confirmed to be significantly up-regulated in patients with AAs, compared with controls; the area under the receiver operating characteristic curve value was 0.64 (95% CI, 0.52-0.75). CONCLUSIONS: Patients with CRC have significantly different patterns of miRNA expression than healthy individuals. These patterns might be developed as biomarkers for CRC, although they have limited value in identifying patients with premalignant neoplastic lesions.

Gong, J., et al. (2016). "Genome-Wide Interaction Analyses between Genetic Variants and Alcohol Consumption and Smoking for Risk of Colorectal Cancer." PLoS Genet **12**(10): e1006296.

Genome-wide association studies (GWAS) have identified many genetic susceptibility loci for colorectal cancer (CRC). However, variants in these loci explain only a small proportion of familial aggregation, and there are likely additional variants that are associated with CRC susceptibility. Genome-

wide studies of gene-environment interactions may identify variants that are not detected in GWAS of marginal gene effects. To study this, we conducted a genome-wide analysis for interaction between genetic variants and alcohol consumption and cigarette smoking using data from the Colon Cancer Family Registry (CCFR) and the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO). Interactions were tested using logistic regression. We identified interaction between CRC risk and alcohol consumption and variants in the 9q22.32/HIATL1 (Pinteraction = 1.76x10-8; permuted p-value 3.51x10-8) region. Compared to non-/occasional drinking light to moderate alcohol consumption was associated with a lower risk of colorectal cancer among individuals with rs9409565 CT genotype (OR, 0.82 [95% CI, 0.74-0.91]; P = 2.1x10-4) and TT genotypes (OR,0.62 [95% CI, 0.51-0.75]; P = 1.3x10-6) but not associated among those with the CC genotype (p = 0.059). No genome-wide statistically significant interactions were observed for smoking. If replicated our suggestive finding of a genome-wide significant interaction between genetic variants and alcohol consumption might contribute to understanding colorectal cancer etiology and identifying subpopulations with differential susceptibility to the effect of alcohol on CRC risk.

Groothuis-Oudshoorn, C. G., et al. (2014). "Public stated preferences and predicted uptake for genome-based colorectal cancer screening." BMC Med Inform Decis Mak **14**: 18.

BACKGROUND: Emerging developments in nanomedicine allow the development of genome-based technologies for non-invasive and individualised screening for diseases such as colorectal cancer. The main objective of this study was to measure user preferences for colorectal cancer screening using a nanopill. METHODS: A discrete choice experiment was used to estimate the preferences for five competing diagnostic techniques including the nanopill and iFOBT. Alternative screening scenarios were described using five attributes namely: preparation involved, sensitivity, specificity, complication rate and testing frequency. Fourteen random and two fixed choice tasks, each consisting of three alternatives, were offered to 2225 individuals. Data were analysed using the McFadden conditional logit model. RESULTS: Thirteen hundred and fifty-six respondents completed the questionnaire. The most important attributes (and preferred levels) were the screening technique (nanopill), sensitivity (100%) and preparation (no preparation). Stated screening uptake for the nanopill was 79%, compared to 76% for iFOBT. In the case of screening with the nanopill, the percentage of people preferring not to be screened would be reduced from 19.2% (iFOBT) to 16.7%. CONCLUSIONS: Although the expected benefits of nanotechnology based colorectal cancer screening are improved screening uptake, assuming more accurate test results and less preparation involved, the relative preference of the nanopill is only slightly higher than the iFOBT. Estimating user preferences during the development of diagnostic technologies could be used to identify relative performance, including perceived benefits and harms compared to competitors allowing for significant changes to be made throughout the process of development.

Gundert, M., et al. (2017). "Genome-wide DNA methylation analysis reveals a prognostic classifier for non-metastatic colorectal cancer (ProMCol classifier)." Gut.

OBJECTIVE: Pathological staging used for the prediction of patient survival in colorectal cancer (CRC) provides only limited information. DESIGN: Here, a genome-wide study of DNA methylation was conducted for two cohorts of patients with non-metastatic CRC (screening cohort (n=572) and validation cohort (n=274)). A variable screening for prognostic CpG sites was performed in the screening cohort using marginal testing based on a Cox model and subsequent adjustment of the p-values via independent hypothesis weighting using the methylation difference between 34 pairs of tumour and normal mucosa tissue as auxiliary covariate. From the 1000 CpG sites with the smallest adjusted p-value, 20 CpG sites with the smallest Brier score for overall survival (OS) were selected. Applying principal component analysis, we derived a prognostic methylation-based classifier for patients with non-metastatic CRC (ProMCol classifier). RESULTS: This classifier was associated with OS in the screening (HR 0.51, 95% CI 0.41 to 0.63, p=6.2E-10) and the validation cohort (HR 0.61, 95% CI 0.45 to 0.82, p=0.001). The independent validation of the ProMCol classifier revealed a reduction of the prediction error for 3-year OS from 0.127, calculated only with standard clinical variables, to 0.120 combining the clinical variables with the classifier and for 4-year OS from 0.153 to 0.140. All results were confirmed for disease-specific survival. CONCLUSION: The ProMCol classifier could improve the prognostic accuracy for patients with non-metastatic CRC.

Habermann, J. K., et al. (2008). "From the genome to the proteome--biomarkers in colorectal cancer." Langenbecks Arch Surg **393**(1): 93-104.

BACKGROUND AND AIMS: Colorectal cancer is the second leading cause of cancer-related death. Current clinical practice in colorectal cancer screening (fecal occult blood test, FOBT; colonoscopy) has

contributed to a reduction of mortality. However, despite these screening programs, about 70% of carcinomas are detected at advanced tumor stages (UICC III/IV) presenting poor patient prognosis. Thus, innovative tools and methodologies for early cancer detection can directly result in improving patient survival rates. PATIENTS/METHODS: Biomedical research has advanced rapidly in recent years with the availability of technologies such as global gene and protein expression profiling. Comprehensive tumor profiling has become a field of intensive research aiming at identifying biomarkers relevant for improved diagnostics and therapeutics. RESULTS: In this paper, we report a comprehensive review of genomic, transcriptomic, and proteomic approaches for biomarker identification in tissue and blood with a main emphasis on two-dimensional gel-electrophoresis (2-DE) and mass spectrometry analyses. CONCLUSION: Proteomics-based technologies enable to distinguish the healthy patient from the tumor patient with high sensitivity and specificity and could greatly improve common classification systems and diagnostics. However, this progress has not yet been transferred from bench to bedside but could open the door to a more accurate and target specific personalized medicine with improved patient survival.

Habermann, J. K., et al. (2006). "From genome to proteome in tumor profiling: molecular events in colorectal cancer genesis." Adv Exp Med Biol **587**: 161-177.

Biomedical research has advanced rapidly in recent years with the sequencing of the human genome and the availability of technologies such as global gene and protein expression profiling using different chip platforms. However, this progress has not yet been transferred to the bedside. While detection of cancer at early stages is critical for curative treatment interventions, efficient diagnostic and therapeutic markers for the majority of malignancies still seem to be lacking. Comprehensive tumor profiling has therefore become a field of intensive research aiming at identifying biomarkers relevant for improved diagnostics and therapeutics. This chapter will demonstrate a genomic and proteomic approach while focusing on tumor profiling during colorectal cancer development.

Hatzis, P., et al. (2008). "Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells." Mol Cell Biol **28**(8): 2732-2744.

Wnt signaling activates gene expression through the induced formation of complexes between DNA-binding T-cell factors (TCFs) and the transcriptional coactivator beta-catenin. In colorectal cancer, activating Wnt pathway mutations transform epithelial cells through the inappropriate activation of a TCF7L2/TCF4 target gene program. Through a DNA array-based genome-wide analysis of TCF4 chromatin occupancy, we have identified 6,868 high-confidence TCF4-binding sites in the LS174T colorectal cancer cell line. Most TCF4-binding sites are located at large distances from transcription start sites, while target genes are frequently "decorated" by multiple binding sites. Motif discovery algorithms define the in vivo-occupied TCF4-binding site as evolutionarily conserved A-C/G-A/T-T-C-A-A-A-G motifs. The TCF4-binding regions significantly correlate with Wnt-responsive gene expression profiles derived from primary human adenomas and often behave as beta-catenin/TCF4-dependent enhancers in transient reporter assays.

Hinoue, T., et al. (2012). "Genome-scale analysis of aberrant DNA methylation in colorectal cancer." Genome Res **22**(2): 271-282.

Colorectal cancer (CRC) is a heterogeneous disease in which unique subtypes are characterized by distinct genetic and epigenetic alterations. Here we performed comprehensive genome-scale DNA methylation profiling of 125 colorectal tumors and 29 adjacent normal tissues. We identified four DNA methylation-based subgroups of CRC using model-based cluster analyses. Each subtype shows characteristic genetic and clinical features, indicating that they represent biologically distinct subgroups. A CIMP-high (CIMP-H) subgroup, which exhibits an exceptionally high frequency of cancer-specific DNA hypermethylation, is strongly associated with MLH1 DNA hypermethylation and the BRAF (V600E) mutation. A CIMP-low (CIMP-L) subgroup is enriched for KRAS mutations and characterized by DNA hypermethylation of a subset of CIMP-H-associated markers rather than a unique group of CpG islands. Non-CIMP tumors are separated into two distinct clusters. One non-CIMP subgroup is distinguished by a significantly higher frequency of TP53 mutations and frequent occurrence in the distal colon, while the tumors that belong to the fourth group exhibit a low frequency of both cancer-specific DNA hypermethylation and gene mutations and are significantly enriched for rectal tumors. Furthermore, we identified 112 genes that were down-regulated more than twofold in CIMP-H tumors together with promoter DNA hypermethylation. These represent approximately 7% of genes that acquired promoter DNA methylation in CIMP-H tumors. Intriguingly, 48/112 genes were also transcriptionally down-regulated in non-CIMP subgroups, but this was not attributable to promoter DNA hypermethylation. Together, we identified four distinct DNA methylation subgroups of CRC and provided novel insight

regarding the role of CIMP-specific DNA hypermethylation in gene silencing.

Hofer, P., et al. (2017). "Bayesian and frequentist analysis of an Austrian genome-wide association study of colorectal cancer and advanced adenomas." Oncotarget **8**(58): 98623-98634.

Most genome-wide association studies (GWAS) were analyzed using single marker tests in combination with stringent correction procedures for multiple testing. Thus, a substantial proportion of associated single nucleotide polymorphisms (SNPs) remained undetected and may account for missing heritability in complex traits. Model selection procedures present a powerful alternative to identify associated SNPs in high-dimensional settings. In this GWAS including 1060 colorectal cancer cases, 689 cases of advanced colorectal adenomas and 4367 controls we pursued a dual approach to investigate genome-wide associations with disease risk applying both, single marker analysis and model selection based on the modified Bayesian information criterion, mBIC2, implemented in the software package MOSGWA. For different case-control comparisons, we report models including between 1-14 candidate SNPs. A genome-wide significant association of rs17659990 (P=5.43x10(-9), DOCK3, chromosome 3p21.2) with colorectal cancer risk was observed. Furthermore, 56 SNPs known to influence susceptibility to colorectal cancer and advanced adenoma were tested in a hypothesis-driven approach and several of them were found to be relevant in our Austrian cohort. After correction for multiple testing (alpha=8.9x10(-4)), the most significant associations were observed for SNPs rs10505477 (P=6.08x10(-4)) and rs6983267 (P=7.35x10(-4)) of CASC8, rs3802842 (P=8.98x10(-5), COLCA1,2), and rs12953717 (P=4.64x10(-4), SMAD7). All previously unreported SNPs demand replication in additional samples. Reanalysis of existing GWAS datasets using model selection as tool to detect SNPs associated with a complex trait may present a promising resource to identify further genetic risk variants not only for colorectal cancer.

Hosono, S., et al. (2016). "A risk prediction model for colorectal cancer using genome-wide association study-identified polymorphisms and established risk factors among Japanese: results from two independent case-control studies." Eur J Cancer Prev **25**(6): 500-507.

Most genome-wide association studies of colorectal cancer (CRC) carried out to date have been in populations with European ancestry, and the extent to which the identified variants contribute as predictors of CRC among Japanese populations has not been clarified. We analyzed 23 genetic variants identified in previous genome-wide association studies in a derivation case-control study with 558 cases and 1116 age-matched and sex-matched controls. Six single nucleotide polymorphisms were selected for synthesis of the genetic risk score. A dose-dependent association was observed between CRC risk and genetic risk score, which is the aggregate number of alleles in six selected variants: 8q24 - rs6983267, 15q13 - rs4779584 and rs1696961, 14q22 - rs444435, 16q22 - rs9929218, and 3q26.2 - rs1093599. The c statistic for a model that included the genetic risk score and conventional risk factors was 0.7167, versus 0.7009 with the conventional risk factors only (P=0.0013). This model was evaluated in a replication study with 547 cases and 547 age-matched and sex-matched controls, and the corresponding c statistics were 0.6356 and 0.6391 with no statistical significance. When the two studies were combined, the corresponding c statistics were 0.6132 and 0.6198 (P=0.0126). We developed a risk model that incorporates a genetic risk score and established risk factors, but this model was not satisfactory in the replication study. The results in the combined study still encourage further attempts using a similar approach among individual countries.

Houlston, R. S., et al. (2010). "Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33." Nat Genet **42**(11): 973-977.

Genome-wide association studies (GWAS) have identified ten loci harboring common variants that influence risk of developing colorectal cancer (CRC). To enhance the power to identify additional CRC risk loci, we conducted a meta-analysis of three GWAS from the UK which included a total of 3,334 affected individuals (cases) and 4,628 controls followed by multiple validation analyses including a total of 18,095 cases and 20,197 controls. We identified associations at four new CRC risk loci: 1q41 (rs6691170, odds ratio (OR) = 1.06, P = 9.55 x 10(-) (1) (0) and rs6687758, OR = 1.09, P = 2.27 x 10(-) (9), 3q26.2 (rs10936599, OR = 0.93, P = 3.39 x 10(-) (8)), 12q13.13 (rs11169552, OR = 0.92, P = 1.89 x 10(-) (1) (0) and rs7136702, OR = 1.06, P = 4.02 x 10(-) (8)) and 20q13.33 (rs4925386, OR = 0.93, P = 1.89 x 10(-) (1) (0)). In addition to identifying new CRC risk loci, this analysis provides evidence that additional CRC-associated variants of similar effect size remain to be discovered.

Jasmine, F., et al. (2012). "A genome-wide study of cytogenetic changes in colorectal cancer using SNP microarrays: opportunities for future personalized treatment." PLoS One **7**(2): e31968.

In colorectal cancer (CRC), chromosomal instability (CIN) is typically studied using comparative-genomic hybridization (CGH) arrays. We studied paired (tumor and surrounding healthy) fresh frozen tissue from 86 CRC patients using Illumina's Infinium-based SNP array. This method allowed us to study CIN in CRC, with simultaneous analysis of copy number (CN) and B-allele frequency (BAF)--a representation of allelic composition. These data helped us to detect mono-allelic and bi-allelic amplifications/deletion, copy neutral loss of heterozygosity, and levels of mosaicism for mixed cell populations, some of which can not be assessed with other methods that do not measure BAF. We identified associations between CN abnormalities and different CRC phenotypes (histological diagnosis, location, tumor grade, stage, MSI and presence of lymph node metastasis). We showed commonalities between regions of CN change observed in CRC and the regions reported in previous studies of other solid cancers (e.g. amplifications of 20q, 13q, 8q, 5p and deletions of 18q, 17p and 8p). From Therapeutic Target Database, we identified relevant drugs, targeted to the genes located in these regions with CN changes, approved or in trials for other cancers and common diseases. These drugs may be considered for future therapeutic trials in CRC, based on personalized cytogenetic diagnosis. We also found many regions, harboring genes, which are not currently targeted by any relevant drugs that may be considered for future drug discovery studies. Our study shows the application of high density SNP arrays for cytogenetic study in CRC and its potential utility for personalized treatment.

Kandimalla, R., et al. (2018). "Genome-wide Discovery and Identification of a Novel miRNA Signature for Recurrence Prediction in Stage II and III Colorectal Cancer." Clin Cancer Res **24**(16): 3867-3877.

Purpose: The current tumor-node-metastasis (TNM) staging system is inadequate at identifying patients with high-risk colorectal cancer. Using a systematic and comprehensive biomarker discovery and validation approach, we aimed to identify an miRNA recurrence classifier (MRC) that can improve upon the current TNM staging as well as is superior to currently offered molecular assays.Experimental Design: Three independent genome-wide miRNA expression profiling datasets were used for biomarker discovery (N = 158) and in silico validation (N = 109 and N = 40) to identify an miRNA signature for predicting tumor recurrence in patients with colorectal cancer. Subsequently, this signature was analytically trained and validated in retrospectively collected independent patient cohorts of fresh-frozen (N = 127, cohort 1) and formalin-fixed paraffin-embedded (FFPE; N = 165, cohort 2 and N = 139, cohort 3) specimens.Results: We identified an 8-miRNA signature that significantly predicted recurrence-free interval (RFI) in the discovery (P = 0.002) and two independent publicly available datasets (P = 0.00006 and P = 0.002). The RT-PCR-based validation in independent clinical cohorts revealed that MRC-derived high-risk patients succumb to significantly poor RFI in patients with stage II and III colorectal cancer [cohort 1: hazard ratio (HR), 3.44 (1.56-7.45), P = 0.001; cohort 2: HR, 6.15 (3.33-11.35), P = 0.001; and cohort 3: HR, 4.23 (2.26-7.92), P = 0.0003]. In multivariate analyses, MRC emerged as an independent predictor of tumor recurrence and achieved superior predictive accuracy over the currently available molecular assays. The RT-PCR-based MRC risk score = (-0.1218 x miR-744) + (-3.7142 x miR-429) + (-2.2051 x miR-362) + (3.0564 x miR-200b) + (2.4997 x miR-191) + (-0.0065 x miR-30c2) + (2.2224 x miR-30b) + (-1.1162 x miR-33a). Conclusions: This novel MRC is superior to currently used clinicopathologic features, as well as National Comprehensive Cancer Network (NCCN) criteria, and works regardless of adjuvant chemotherapy status in identifying patients with high-risk stage II and III colorectal cancer. This can be readily deployed in clinical practice with FFPE specimens for decision-making pending further model testing and validation. Clin Cancer Res; 24(16); 3867-77. (c)2018 AACRSee related commentary by Rodriguez et al., p. 3787.

Kemp, Z., et al. (2006). "Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan." Hum Mol Genet **15**(19): 2903-2910.

To identify a novel susceptibility gene for colorectal cancer (CRC), we conducted a genome-wide linkage analysis of 69 pedigrees segregating colorectal neoplasia in which involvement of known loci had been excluded, using a high-density single nucleotide polymorphism (SNP) array containing 10,204 markers. Multipoint linkage analyses were undertaken using both non-parametric (model-free) and parametric (model-based) methods. After the removal of SNPs in strong linkage disequilibrium, we obtained a maximum non-parametric linkage statistic of 3.40 (P=0.0003) at chromosomal region 3q21-q24. The same genomic position also yielded the highest multipoint heterogeneity LOD (HLOD) score under a dominant model (HLOD=3.10, genome-wide P=0.038) with 62% of families linked to the locus. We provide evidence for a novel CRC susceptibility gene. Further studies are needed to confirm this localization and to evaluate the contribution of this locus to disease incidence.

Khamas, A., et al. (2012). "Genome-wide screening for methylation-silenced genes in colorectal cancer." Int J Oncol **41**(2): 490-496.

Identification of methylation-silenced genes in colorectal cancer (CRC) is of great importance. We employed oligonucleotide microarrays to identify differences in global gene expression of five CRC cell lines (HCT116, RKO, Colo320, SW480 and HT29) that were analyzed before and after treatment with 5-aza-2'-deoxycitidine. Selected candidates were subjected to methylation-specific PCR and real-time quantitative reverse transcription-PCR using 15 CRC cell lines and 23 paired tumor and normal samples from CRC patients. After 5-aza-2'-deoxycitidine treatment, 139 genes were re-expressed in all 5 CRC cell lines collectively with a fold change of more than 1.5 in at least one cell line. These genes include known methylated and silenced genes in CRC. After applying study selection criteria we identified 20 candidates. The GADD45B and THSD1 genes were selected for further analysis. Among 15 colon cancer cell lines, methylation was only identified in THSD1 (27%). THSD1 methylation was subsequently investigated in 23 colorectal tumors and methylation was detected in 9% of the analyzed samples; the observed promoter hypermethylation was cancer-specific. THSD1 mRNA down-regulation was observed in tumor tissues. This genome-wide screening led to the identification of genes putatively affected by methylation in CRC. The THSD1 gene may play a role in the tumorigenesis of CRC.

Ki, D. H., et al. (2007). "Whole genome analysis for liver metastasis gene signatures in colorectal cancer." Int J Cancer **121**(9): 2005-2012.

Liver metastasis is one of the major causes of death in colorectal cancer (CRC) patients. To understand this process, we investigated whether the gene expression profiling of matched colorectal carcinomas and liver metastases could reveal key molecular events involved in tumor progression and metastasis. We performed experiments using a cDNA microarray containing 17,104 genes with the following tissue samples: paired tissues of 25 normal colorectal mucosa, 27 primary colorectal tumors, 13 normal liver and 27 liver metastasis, and 20 primary colorectal tumors without liver metastasis. To remove the effect of normal cell contamination, we selected 4,583 organ-specific genes with a false discovery rate (FDR) of 0.0067% by comparing normal colon and liver tissues using significant analysis of microarray, and these genes were excluded from further analysis. We then identified and validated 46 liver metastasis-specific genes with an accuracy of 83.3% by comparing the expression of paired primary colorectal tumors and liver metastases using prediction analysis of microarray. The 46 selected genes contained several known oncogenes and 2 ESTs. To confirm that the results correlated with the microarray expression patterns, we performed RT-PCR with WNT5A and carbonic anhydrase II. Additionally, we observed that 21 of the 46 genes were differentially expressed (FDR=2.27%) in primary tumors with synchronous liver metastasis compared with primary tumors without liver metastasis. We scanned the human genome using a cDNA microarray and identified 46 genes that may play an important role in the progression of liver metastasis in CRC.

Kim, M. Y., et al. (2006). "Recurrent genomic alterations with impact on survival in colorectal cancer identified by genome-wide array comparative genomic hybridization." Gastroenterology **131**(6): 1913-1924.

BACKGROUND & AIMS: Although genetic aspects of tumorigenesis in colorectal cancer (CRC) have been well studied, reliable biomarkers predicting prognosis are scarce. We aimed to identify recurrently altered genomic regions (RAR) in CRC with high resolution, to investigate their implications on survival and to explore novel cancer-related genes in prognosis-associated RARs. METHODS: A 1-Mb resolution microarray-based comparative genomic hybridization (array CGH) was applied to 59 CRCs. RARs, defined as genomic alterations, detected in more than 10 cases were identified and analyzed for their association with survival. Expression levels of genes in prognosis-associated RARs were examined by real-time quantitative polymerase chain reaction. RESULTS: Twenty-seven RARs were identified. Eleven high-level amplifications and 2 homozygous deletions also were detected, but they were not as common as RARs. Multivariate analysis revealed RAR-L1 (loss on 1p36; hazard ratio = 8.15, P =.002) and RAR-L20 (loss on 21q22; hazard ratio = 3.53, P =.034) are independent indicators of poor prognosis. Expression of CAMTA1, located in RAR-L1, was reduced frequently in CRCs, and low CAMTA1 expression was associated significantly with poor prognosis, which indicates that CAMTA1 may play a role as a tumor suppressor in CRC. Five pairs of RARs were correlated significantly to each other and 3 pairs share genes involved in the same biological functions, suggesting possible collaborative roles in tumorigenesis. CONCLUSIONS: We identified recurrent genomic changes in 59 CRCs. RARs could be more important in sporadic tumors where the effect of genomic changes on tumorigenesis is relatively smaller than in familial cancer. Our results and analysis strategy will be helpful to elucidate pathogenesis of CRCs or to develop biomarkers for predicting prognosis.

Kundu, S., et al. (2018). "Linking FOXO3, NCOA3, and TCF7L2 to Ras pathway phenotypes through a genome-wide forward genetic screen in human colorectal cancer cells." Genome Med **10**(1): 2.

BACKGROUND: The Ras pathway genes KRAS, BRAF, or ERBBs have somatic mutations in ~ 60% of human colorectal carcinomas. At present, it is unknown whether the remaining cases lack mutations activating the Ras pathway or whether they have acquired mutations in genes hitherto unknown to belong to the pathway. METHODS: To address the second possibility and extend the compendium of Ras pathway genes, we used genome-wide transposon mutagenesis of two human colorectal cancer cell systems deprived of their activating KRAS or BRAF allele to identify genes enabling growth in low glucose, a Ras pathway phenotype, when targeted. RESULTS: Of the 163 recurrently targeted genes in the two different genetic backgrounds, one-third were known cancer genes and one-fifth had links to the EGFR/Ras/MAPK pathway. When compared to cancer genome sequencing datasets, nine genes also mutated in human colorectal cancers were identified. Among these, stable knockdown of FOXO3, NCOA3, and TCF7L2 restored growth in low glucose but reduced MEK/MAPK phosphorylation, reduced anchorage-independent growth, and modulated expressions of GLUT1 and Ras pathway related proteins. Knockdown of NCOA3 and FOXO3 significantly decreased the sensitivity to cetuximab of KRAS mutant but not wild-type cells. CONCLUSIONS: This work establishes a proof-of-concept that human cell-based genome-wide forward genetic screens can assign genes to pathways with clinical importance in human colorectal cancer.

Lagerstedt, K. K., et al. (2007). "Tumor genome wide DNA alterations assessed by array CGH in patients with poor and excellent survival following operation for colorectal cancer." Cancer Inform **3**: 341-355.

Genome wide DNA alterations were evaluated by array CGH in addition to RNA expression profiling in colorectal cancer from patients with excellent and poor survival following primary operations. DNA was used for CGH in BAC and cDNA arrays. Global RNA expression was determined by 44K arrays. DNA and RNA from tumor and normal colon were used from cancer patients grouped according to death, survival or Dukes A, B, C and D tumor stage. Confirmed DNA alterations in all Dukes A - D were judged relevant for carcinogenesis, while changes in Dukes C and D only were regarded relevant for tumor progression. Copy number gain was more common than loss in tumor tissue ($p < 0.01$). Major tumor DNA alterations occurred in chromosome 8, 13, 18 and 20, where short survival included gain in 8q and loss in 8p. Copy number gains related to tumor progression were most common on chromosome 7, 8, 19, 20, while corresponding major losses appeared in chromosome 8. Losses at chromosome 18 occurred in all Dukes stages. Normal colon tissue from cancer patients displayed gains in chromosome 19 and 20. Mathematical Vector analysis implied a number of BAC-clones in tumor DNA with genes of potential importance for death or survival. The genomic variation in colorectal cancer cells is tremendous and emphasizes that BAC array CGH is presently more powerful than available statistical models to discriminate DNA sequence information related to outcome. Present results suggest that a majority of DNA alterations observed in colorectal cancer are secondary to tumor progression. Therefore, it would require an immense work to distinguish primary from secondary DNA alterations behind colorectal cancer.

Lange, C. P., et al. (2012). "Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer." PLoS One **7**(11): e50266.

BACKGROUND: There is an increasing demand for accurate biomarkers for early non-invasive colorectal cancer detection. We employed a genome-scale marker discovery method to identify and verify candidate DNA methylation biomarkers for blood-based detection of colorectal cancer. METHODOLOGY/PRINCIPAL FINDINGS: We used DNA methylation data from 711 colorectal tumors, 53 matched adjacent-normal colonic tissue samples, 286 healthy blood samples and 4,201 tumor samples of 15 different cancer types. DNA methylation data were generated by the Illumina Infinium HumanMethylation27 and the HumanMethylation450 platforms, which determine the methylation status of 27,578 and 482,421 CpG sites respectively. We first performed a multistep marker selection to identify candidate markers with high methylation across all colorectal tumors while harboring low methylation in healthy samples and other cancer types. We then used pre-therapeutic plasma and serum samples from 107 colorectal cancer patients and 98 controls without colorectal cancer, confirmed by colonoscopy, to verify candidate markers. We selected two markers for further evaluation: methylated THBD (THBD-M) and methylated C9orf50 (C9orf50-M). When tested on clinical plasma and serum samples these markers outperformed carcinoembryonic antigen (CEA) serum measurement and resulted in a high sensitive and specific test performance for early colorectal cancer detection. CONCLUSIONS/SIGNIFICANCE: Our

systematic marker discovery and verification study for blood-based DNA methylation markers resulted in two novel colorectal cancer biomarkers, THBD-M and C9orf50-M. THBD-M in particular showed promising performance in clinical samples, justifying its further optimization and clinical testing.

Lascorz, J., et al. (2010). "Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility." Carcinogenesis **31**(9): 1612-1619.

Genetic susceptibility accounts for approximately 35% of all colorectal cancer (CRC). Ten common low-risk variants contributing to CRC risk have been identified through genome-wide association studies (GWASs). In our GWAS, 610 664 genotyped single-nucleotide polymorphisms (SNPs) passed the quality control filtering in 371 German familial CRC patients and 1263 controls, and replication studies were conducted in four additional case-control sets (4915 cases and 5607 controls). Known risk loci at 8q24.21 and 11q23 were confirmed, and a previously unreported association, rs12701937, located between the genes GLI3 (GLI family zinc finger 3) and INHBA (inhibin, beta A) [P = 1.1 x 10(-3), odds ratio (OR) 1.14, 95% confidence interval (CI) 1.05-1.23, dominant model in the combined cohort], was identified. The association was stronger in familial cases compared with unselected cases (P = 2.0 x 10(-4), OR 1.36, 95% CI 1.16-1.60, dominant model). Two other unreported SNPs, rs6038071, 40 kb upstream of CSNK2A1 (casein kinase 2, alpha 1 polypeptide) and an intronic marker in MYO3A (myosin IIIA), rs11014993, associated with CRC only in the familial CRC cases (P = 2.5 x 10(-3), recessive model, and P = 2.7 x 10(-4), dominant model). Three software tools successfully pointed to the overrepresentation of genes related to the mitogen-activated protein kinase (MAPK) signalling pathways among the 1340 most strongly associated markers from the GWAS (allelic P value < 10(-3)). The risk of CRC increased significantly with an increasing number of risk alleles in seven genes involved in MAPK signalling events (P (trend) = 2.2 x 10(-16), OR (per allele) = 1.34, 95% CI 1.11-1.61).

Le Marchand, L. (2009). "Genome-wide association studies and colorectal cancer." Surg Oncol Clin N Am **18**(4): 663-668.

Genome-wide association studies (GWAS) provide a powerful new approach to identify common, low-penetrance susceptibility loci without prior knowledge of biologic function. Results from three GWAS conducted in populations of European ancestry are available for colorectal cancer (CRC). These studies have identified 11 disease loci that, for the majority, were not previously suspected to be related to CRC. The proportions of the familial and population risks explained by these loci are small and they currently are not useful for risk prediction. However, the power of these studies was low, indicating that a number of other loci may be identified in new ongoing GWAS, and in pooled analyses. Thus, the risk prediction ability of susceptibility markers identified in GWAS for CRC may improve as more variants are discovered. This may, in turn, have important implications for targeting high-risk individuals for colonoscopy screening.

Lemire, M., et al. (2015). "A genome-wide association study for colorectal cancer identifies a risk locus in 14q23.1." Hum Genet **134**(11-12): 1249-1262.

Over 50 loci associated with colorectal cancer (CRC) have been uncovered by genome-wide association studies (GWAS). Identifying additional loci has the potential to help elucidate aspects of the underlying biological processes leading to better understanding of the pathogenesis of the disease. We re-evaluated a GWAS by excluding controls that have family history of CRC or personal history of colorectal polyps, as we hypothesized that their inclusion reduces power to detect associations. This is supported empirically and through simulations. Two-phase GWAS analysis was performed in a total of 16,517 cases and 14,487 controls. We identified rs17094983, a SNP associated with risk of CRC [p = 2.5 x 10(-10); odds ratio estimated by re-including all controls (OR) = 0.87, 95% confidence interval (CI) 0.83-0.91; minor allele frequency (MAF) = 13%]. Results were replicated in samples of African descent (1894 cases and 4703 controls; p = 0.01; OR = 0.86, 95% CI 0.77-0.97; MAF = 16 %). Gene expression data in 195 colon adenocarcinomas and 59 normal colon tissues from two different studies revealed that this locus has genotypes that are associated with RTN1 (Reticulon 1) expression (p = 0.001), a protein-coding gene involved in survival and proliferation of cancer cells which is highly expressed in normal colon tissues but has significantly reduced expression in tumor cells (p = 1.3 x 10(-8)).

Madhavan, S., et al. (2013). "Genome-wide multi-omics profiling of colorectal cancer identifies immune determinants strongly associated with relapse." Front Genet **4**: 236.

The use and benefit of adjuvant chemotherapy to treat stage II colorectal cancer (CRC) patients is not well understood since the majority of these patients are cured by surgery alone. Identification of biological markers of relapse is a critical challenge to effectively target treatments to the ~20% of patients destined to relapse. We have integrated molecular profiling results

of several "omics" data types to determine the most reliable prognostic biomarkers for relapse in CRC using data from 40 stage I and II CRC patients. We identified 31 multi-omics features that highly correlate with relapse. The data types were integrated using multi-step analytical approach with consecutive elimination of redundant molecular features. For each data type a systems biology analysis was performed to identify pathways biological processes and disease categories most affected in relapse. The biomarkers detected in tumors urine and blood of patients indicated a strong association with immune processes including aberrant regulation of T-cell and B-cell activation that could lead to overall differences in lymphocyte recruitment for tumor infiltration and markers indicating likelihood of future relapse. The immune response was the biologically most coherent signature that emerged from our analyses among several other biological processes and corroborates other studies showing a strong immune response in patients less likely to relapse.

McInnes, T., et al. (2017). "Genome-wide methylation analysis identifies a core set of hypermethylated genes in CIMP-H colorectal cancer." BMC Cancer **17**(1): 228.

BACKGROUND: Aberrant DNA methylation profiles are a characteristic of all known cancer types, epitomized by the CpG island methylator phenotype (CIMP) in colorectal cancer (CRC). Hypermethylation has been observed at CpG islands throughout the genome, but it is unclear which factors determine whether an individual island becomes methylated in cancer. METHODS: DNA methylation in CRC was analysed using the Illumina HumanMethylation450K array. Differentially methylated loci were identified using Significance Analysis of Microarrays (SAM) and the Wilcoxon Signed Rank (WSR) test. Unsupervised hierarchical clustering was used to identify methylation subtypes in CRC. RESULTS: In this study we characterized the DNA methylation profiles of 94 CRC tissues and their matched normal counterparts. Consistent with previous studies, unsupervized hierarchical clustering of genome-wide methylation data identified three subtypes within the tumour samples, designated CIMP-H, CIMP-L and CIMP-N, that showed high, low and very low methylation levels, respectively. Differential methylation between normal and tumour samples was analysed at the individual CpG level, and at the gene level. The distribution of hypermethylation in CIMP-N tumours showed high inter-tumour variability and appeared to be highly stochastic in nature, whereas CIMP-H tumours exhibited consistent hypermethylation at a subset of genes, in addition to a highly variable background of hypermethylated genes. EYA4, TFPI2 and TLX1 were hypermethylated in more than 90% of all tumours examined. One-hundred thirty-two genes were hypermethylated in 100% of CIMP-H tumours studied and these were highly enriched for functions relating to skeletal system development (Bonferroni adjusted p value =2.88E-15), segment specification (adjusted p value =9.62E-11), embryonic development (adjusted p value =1.52E-04), mesoderm development (adjusted p value =1.14E-20), and ectoderm development (adjusted p value =7.94E-16). CONCLUSIONS: Our genome-wide characterization of DNA methylation in colorectal cancer has identified 132 genes hypermethylated in 100% of CIMP-H samples. Three genes, EYA4, TLX1 and TFPI2 are hypermethylated in >90% of all tumour samples, regardless of CIMP subtype.

Naumov, V. A., et al. (2013). "Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips." Epigenetics **8**(9): 921-934.

Illumina's Infinium HumanMethylation450 BeadChip arrays were used to examine genome-wide DNA methylation profiles in 22 sample pairs from colorectal cancer (CRC) and adjacent tissues and 19 colon tissue samples from cancer-free donors. We show that the methylation profiles of tumors and healthy tissue samples can be clearly distinguished from one another and that the main source of methylation variability is associated with disease status. We used different statistical approaches to evaluate the methylation data. In general, at the CpG-site level, we found that common CRC-specific methylation patterns consist of at least 15,667 CpG sites that were significantly different from either adjacent healthy tissue or tissue from cancer-free subjects. Of these sites, 10,342 were hypermethylated in CRC, and 5,325 were hypomethylated. Hypermethylated sites were common in the maximum number of sample pairs and were mostly located in CpG islands, where they were significantly enriched for differentially methylated regions known to be cancer-specific. In contrast, hypomethylated sites were mostly located in CpG shores and were generally sample-specific. Despite the considerable variability in methylation data, we selected a panel of 14 highly robust candidates showing methylation marks in genes SND1, ADHFE1, OPLAH, TLX2, C1orf70, ZFP64, NR5A2, and COL4A. This set was successfully cross-validated using methylation data from 209 CRC samples and 38 healthy tissue samples from The Cancer Genome Atlas consortium (AUC = 0.981 [95% CI: 0.9677-0.9939], sensitivity = 100% and specificity = 82%). In summary, this study reports a large number of loci with novel differential methylation statuses,

some of which may serve as candidate markers for diagnostic purposes.

Neklason, D. W., et al. (2008). "Common familial colorectal cancer linked to chromosome 7q31: a genome-wide analysis." Cancer Res **68**(21): 8993-8997.

Present investigations suggest that approximately 30% of colorectal cancer cases arise on the basis of inherited factors. We hypothesize that the majority of inherited factors are moderately penetrant genes, common in the population. We use an affected sibling pair approach to identify genetic regions that are coinherited by siblings with colorectal cancer. Individuals from families with at least two siblings diagnosed with colorectal adenocarcinoma or high-grade dysplasia were enrolled. Known familial colorectal cancer syndromes were excluded. A genome-wide scan on 151 DNA samples from 70 kindreds was completed using deCODE 1100 short tandem repeat marker set at an average 4-cM density. Fine mapping on a total of 184 DNAs from 83 kindreds was done in regions suggesting linkage. Linkage analysis was accomplished with Merlin analysis package. Nonparametric linkage analysis revealed three genetic regions with logarithm of the odds (LOD) scores >or=2.0: Ch. 3q29, LOD 2.61 (P = 0.0003); Ch. 4q31.3, LOD 2.13 (P = 0.0009); and Ch. 7q31.31, LOD 3.08 (P = 0.00008). Affected siblings with increased sharing at the 7q31 locus have a 3.8-year (+/- 3.5) earlier age of colorectal cancer onset although this is not statistically significant (P = 0.11). No significant linkage was found near genes causing known syndromes or regions previously reported (8q24, 9q22, and 11q23). The chromosome 3q21-q24 region reported to be linked in colorectal cancer relative pairs is supported by our study, albeit a minor peak (LOD 0.9; P = 0.02). No known familial cancer genes reside in the 7q31 locus, and thus the identified region may contain a novel susceptibility gene responsible for common familial colorectal cancer.

Oh, T., et al. (2013). "Genome-wide identification and validation of a novel methylation biomarker, SDC2, for blood-based detection of colorectal cancer." J Mol Diagn **15**(4): 498-507.

Aberrant DNA methylation has shown promise as a biomarker for the early detection of cancer. To discover novel genes frequently methylated at an early stage in colorectal cancer (CRC), DNA microarray analysis coupled with enriched methylated DNA was performed in primary tumors and compared with adjacent nontumor tissues of 12 patients with CRC at stages I to IV. Stepwise filtering for candidate selection in microarray data analysis yielded a set of genes that are highly methylated across all CRC tumors and that can be used as a composite biomarker for CRC detection. Verification assay identified the SDC2 gene as a potential methylation biomarker for early CRC detection. In clinical validation in tissues from 139 CRC patients, a much higher level of aberrant SDC2 methylation was measured in most primary tumors (97.8%), compared with corresponding nontumor tissue of CRC patients, irrespective of clinical stage. Clinical validation of SDC2 methylation in serum DNA from CRC patients (n = 131) at stages I to IV and from healthy individuals (n = 125) by quantitative methylation-specific PCR demonstrated a high sensitivity of 87.0% (95% CI, 80.0% to 92.3%) in detecting cancers, with a specificity of 95.2% (95% CI, 89.8% to 98.2%). Importantly, sensitivity at stage I was 92.3%, indicating the potential of SDC2 methylation as a blood-based DNA test for early detection of CRC.

Pande, M., et al. (2018). "Genetic susceptibility markers for a breast-colorectal cancer phenotype: Exploratory results from genome-wide association studies." PLoS One **13**(4): e0196245.

BACKGROUND: Clustering of breast and colorectal cancer has been observed within some families and cannot be explained by chance or known high-risk mutations in major susceptibility genes. Potential shared genetic susceptibility between breast and colorectal cancer, not explained by high-penetrance genes, has been postulated. We hypothesized that yet undiscovered genetic variants predispose to a breast-colorectal cancer phenotype. METHODS: To identify variants associated with a breast-colorectal cancer phenotype, we analyzed genome-wide association study (GWAS) data from cases and controls that met the following criteria: cases (n = 985) were women with breast cancer who had one or more first- or second-degree relatives with colorectal cancer, men/women with colorectal cancer who had one or more first- or second-degree relatives with breast cancer, and women diagnosed with both breast and colorectal cancer. Controls (n = 1769), were unrelated, breast and colorectal cancer-free, and age- and sex- frequency-matched to cases. After imputation, 6,220,060 variants were analyzed using the discovery set and variants associated with the breast-colorectal cancer phenotype at P<5.0E-04 (n = 549, at 60 loci) were analyzed for replication (n = 293 cases and 2,103 controls). RESULTS: Multiple correlated SNPs in intron 1 of the ROBO1 gene were suggestively associated with the breast-colorectal cancer phenotype in the discovery and replication data (most significant; rs7430339, Pdiscovery = 1.2E-04; rs7429100, Preplication = 2.8E-03). In meta-analysis of the discovery and replication data, the most significant association remained at rs7429100 (P = 1.84E-06).

CONCLUSION: The results of this exploratory analysis did not find clear evidence for a susceptibility locus with a pleiotropic effect on hereditary breast and colorectal cancer risk, although the suggestive association of genetic variation in the region of ROBO1, a potential tumor suppressor gene, merits further investigation.

Pander, J., et al. (2015). "Genome Wide Association Study for Predictors of Progression Free Survival in Patients on Capecitabine, Oxaliplatin, Bevacizumab and Cetuximab in First-Line Therapy of Metastatic Colorectal Cancer." PLoS One **10**(7): e0131091.

PURPOSE: Despite expanding options for systemic treatment, survival for metastatic colorectal cancer (mCRC) remains limited and individual response is difficult to predict. In search of pre-treatment predictors, pharmacogenetic research has mainly used a candidate gene approach. Genome wide association (GWA) studies offer the benefit of simultaneously analyzing a large number of SNPs, in both known and still unidentified functional regions. Using a GWA approach, we searched for genetic markers affecting progression free survival (PFS) in mCRC patients treated with first-line capecitabine, oxaliplatin and bevacizumab (CAPOX-B), with or without cetuximab. PATIENTS AND METHODS: 755 patients were included in the CAIRO2-trial, a multicenter phase III trial, randomizing between first-line treatment with CAPOX-B versus CAPOX-B plus cetuximab. Germline DNA and complete clinical information was available from 553 patients and genome wide genotyping was performed, using Illumina's OmniExpress beadchip arrays, with 647,550 markers passing all quality checks. Another 2,202,473 markers were imputated by using HapMap2. Association with PFS was analysed using a Cox proportional hazards model. RESULTS: One marker, rs885036, associated significantly with PFS (P = 2.17x10(-8)) showing opposite effects on PFS depending on treatment arm. The minor allele was associated with increased PFS in patients receiving cetuximab. A cluster of markers located on chromosome 8 associated with PFS, irrespective of treatment arm (P-values of 2.30x10(-7) to 1.04x10(-6)). CONCLUSION: This is the first GWA study to find SNPs affecting PFS in mCRC patients treated with CAPOX-B, either with or without cetuximab. Rs885036 is a potential predictive marker for cetuximab efficacy. These markers need to be validated in independent treatment cohorts.

Peng, Z., et al. (2003). "Genome-wide search for loss of heterozygosity in Chinese patients with sporadic colorectal cancer." Int J Gastrointest Cancer **34**(1): 39-48.

In an attempt to integrally investigate the loss of tumor suppressor genes and search for putative suppressor loci associated with tumor occurrence and progression, we conducted a genome-wide loss of hetero zygosity (LOH) study of 83 tumor samples obtained from Chinese patients with sporadic colorectal cancer. We employed 400 fluorescence-labeled microsatellite marker primers to amplify the corresponding loci of the genomic DNA and then electrophoresed the polymerase chain reaction products and analyzed the fluorescent signals. The LOH frequencies were high (>35%) but were not associated with the tumor stage and progression in 20 loci, including the regions where TP53, E-cadherin, deleted in colorectal carcinoma (DCC), phosphatase and tensin homolog deleted on chromosome 10 (PTEN), mothers against decapentaplegic, Drosophila, homolog of 2 (MADH2) and mothers against decapentaplegic, Drosophila, homolog of 4 (MADH4) reside. Loss of other loci, including two narrow regions on chromosome 2, was found to relate to the tumor stage, suggesting that this genomic instability may contribute to tumor progression.

Penney, M. E., et al. (2018). "Associations of single nucleotide polymorphisms with mucinous colorectal cancer: genome-wide common variant and gene-based rare variant analyses." Biomark Res **6**: 17.

Background: Colorectal cancer has significant impact on individuals and healthcare systems. Many genes have been identified to influence its pathogenesis. However, the genetic basis of mucinous tumor histology, an aggressive subtype of colorectal cancer, is currently not well-known. This study aimed to identify common and rare genetic variations that are associated with the mucinous tumor phenotype. Methods: Genome-wide single nucleotide polymorphism (SNP) data was investigated in a colorectal cancer patient cohort (n = 505). Association analyses were performed for 729,373 common SNPs and 275,645 rare SNPs. Common SNP association analysis was performed using univariable and multivariable logistic regression under different genetic models. Rare-variant association analysis was performed using a multi-marker test. Results: No associations reached the traditional genome-wide significance. However, promising genetic associations were identified. The identified common SNPs significantly improved the discriminatory accuracy of the model for mucinous tumor phenotype. Specifically, the area under the receiver operating characteristic curve increased from 0.703 (95% CI: 0.634-0.773) to 0.916 (95% CI: 0.873-0.960) when considering the most significant SNPs. Additionally, the rare variant

analysis identified a number of genetic regions that potentially contain causal rare variants associated with the mucinous tumor phenotype. Conclusions: This is the first study applying both common and rare variant analyses to identify genetic associations with mucinous tumor phenotype using a genome-wide genotype data. Our results suggested novel associations with mucinous tumors. Once confirmed, these results will not only help us understand the biological basis of mucinous histology, but may also help develop targeted treatment options for mucinous tumors.

Peters, U., et al. (2012). "Meta-analysis of new genome-wide association studies of colorectal cancer risk." Hum Genet **131**(2): 217-234.

Colorectal cancer is the second leading cause of cancer death in developed countries. Genome-wide association studies (GWAS) have successfully identified novel susceptibility loci for colorectal cancer. To follow up on these findings, and try to identify novel colorectal cancer susceptibility loci, we present results for GWAS of colorectal cancer (2,906 cases, 3,416 controls) that have not previously published main associations. Specifically, we calculated odds ratios and 95% confidence intervals using log-additive models for each study. In order to improve our power to detect novel colorectal cancer susceptibility loci, we performed a meta-analysis combining the results across studies. We selected the most statistically significant single nucleotide polymorphisms (SNPs) for replication using ten independent studies (8,161 cases and 9,101 controls). We again used a meta-analysis to summarize results for the replication studies alone, and for a combined analysis of GWAS and replication studies. We measured ten SNPs previously identified in colorectal cancer susceptibility loci and found eight to be associated with colorectal cancer (p value range 0.02 to 1.8 x 10(-8)). When we excluded studies that have previously published on these SNPs, five SNPs remained significant at $p < 0.05$ in the combined analysis. No novel susceptibility loci were significant in the replication study after adjustment for multiple testing, and none reached genome-wide significance from a combined analysis of GWAS and replication. We observed marginally significant evidence for a second independent SNP in the BMP2 region at chromosomal location 20p12 (rs4813802; replication p value 0.03; combined p value 7.3 x 10(-5)). In a region on 5p33.15, which includes the coding regions of the TERT-CLPTM1L genes and has been identified in GWAS to be associated with susceptibility to at least seven other cancers, we observed a marginally significant association with rs2853668 (replication p value 0.03; combined p value 1.9 x 10(-4)). Our study suggests a complex nature of the contribution of common genetic variants to risk for colorectal cancer.

Phipps, A. I., et al. (2016). "Common genetic variation and survival after colorectal cancer diagnosis: a genome-wide analysis." Carcinogenesis **37**(1): 87-95.

Genome-wide association studies have identified several germline single nucleotide polymorphisms (SNPs) significantly associated with colorectal cancer (CRC) incidence. Common germline genetic variation may also be related to CRC survival. We used a discovery-based approach to identify SNPs related to survival outcomes after CRC diagnosis. Genome-wide genotyping arrays were conducted for 3494 individuals with invasive CRC enrolled in six prospective cohort studies (median study-specific follow-up = 4.2-8.1 years). In pooled analyses, we used Cox regression to assess SNP-specific associations with CRC-specific and overall survival, with additional analyses stratified by stage at diagnosis. Top findings were followed-up in independent studies. A P value threshold of $P < 5x10(-8)$ in analyses combining discovery and follow-up studies was required for genome-wide significance. Among individuals with distant-metastatic CRC, several SNPs at 6p12.1, nearest the ELOVL5 gene, were statistically significantly associated with poorer survival, with the strongest associations noted for rs209489 [hazard ratio (HR) = 1.8, $P = 7.6x10(-10)$ and HR = 1.8, $P = 3.7x10(-9)$ for CRC-specific and overall survival, respectively). No SNPs were statistically significantly associated with survival among all cases combined or in cases without distant-metastases. SNPs in 6p12.1/ELOVL5 were associated with survival outcomes in individuals with distant-metastatic CRC, and merit further follow-up for functional significance. Findings from this genome-wide association study highlight the potential importance of genetic variation in CRC prognosis and provide clues to genomic regions of potential interest.

Picelli, S., et al. (2008). "Genome-wide linkage scan for colorectal cancer susceptibility genes supports linkage to chromosome 3q." BMC Cancer **8**: 87.

BACKGROUND: Colorectal cancer is one of the most common causes of cancer-related mortality. The disease is clinically and genetically heterogeneous though a strong hereditary component has been identified. However, only a small proportion of the inherited susceptibility can be ascribed to dominant syndromes, such as Hereditary Non-Polyposis Colorectal Cancer (HNPCC) or Familial Adenomatous Polyposis (FAP). In an attempt to identify novel colorectal cancer predisposing genes, we have performed a genome-wide linkage analysis in 30 Swedish non-FAP/non-HNPCC families with a strong

family history of colorectal cancer. METHODS: Statistical analysis was performed using multipoint parametric and nonparametric linkage. RESULTS: Parametric analysis under the assumption of locus homogeneity excluded any common susceptibility regions harbouring a predisposing gene for colorectal cancer. However, several loci on chromosomes 2q, 3q, 6q, and 7q with suggestive linkage were detected in the parametric analysis under the assumption of locus heterogeneity as well as in the nonparametric analysis. Among these loci, the locus on chromosome 3q21.1-q26.2 was the most consistent finding providing positive results in both parametric and nonparametric analyses Heterogeneity LOD score (HLOD) = 1.90, alpha = 0.45, Non-Parametric LOD score (NPL) = 2.1). CONCLUSION: The strongest evidence of linkage was seen for the region on chromosome 3. Interestingly, the same region has recently been reported as the most significant finding in a genome-wide analysis performed with SNP arrays; thus our results independently support the finding on chromosome 3q.

Prasad, R. M., et al. (2018). "A genome-wide analysis of colorectal cancer in a child with Noonan syndrome." Pediatr Blood Cancer **65**(11): e27362.

Noonan syndrome (NS) is a developmental syndrome caused by germline mutations in the Ras signaling pathway. No association has been shown between NS and pediatric colorectal cancer (CRC). We report the case of CRC in a pediatric patient with NS. The patient underwent whole genome sequencing. A germline SOS1 mutation c.1310T>C (p. Ile437Thr) confirmed NS diagnosis. No known hereditary cancer syndromes were identified. Tumor analysis revealed two mutations: a TP53 missense mutation c.481G>A (p. Ala161Tyr) and NCOR1 nonsense mutation c.6052C>T (p. Arg2018*). This report highlights the complexity of Ras signaling and the interplay between developmental syndromes and cancer.

Quan, B., et al. (2015). "Pathway analysis of genome-wide association study and transcriptome data highlights new biological pathways in colorectal cancer." Mol Genet Genomics **290**(2): 603-610.

Colorectal cancer (CRC) is a common malignancy that meets the definition of a complex disease. Genome-wide association study (GWAS) has identified several loci of weak predictive value in CRC, however, these do not fully explain the occurrence risk. Recently, gene set analysis has allowed enhanced interpretation of GWAS data in CRC, identifying a number of metabolic pathways as important for disease pathogenesis. Whether there are other important pathways involved in CRC, however, remains unclear. We present a systems analysis of KEGG pathways in CRC using (1) a human CRC GWAS dataset and (2) a human whole transcriptome CRC case-control expression dataset. Analysis of the GWAS dataset revealed significantly enriched KEGG pathways related to metabolism, immune system and diseases, cellular processes, environmental information processing, genetic information processing, and neurodegenerative diseases. Altered gene expression was confirmed in these pathways using the transcriptome dataset. Taken together, these findings not only confirm previous work in this area, but also highlight new biological pathways whose deregulation is critical for CRC. These results contribute to our understanding of disease-causing mechanisms and will prove useful for future genetic and functional studies in CRC.

Real, L. M., et al. (2014). "A colorectal cancer susceptibility new variant at 4q26 in the Spanish population identified by genome-wide association analysis." PLoS One **9**(6): e101178.

BACKGROUND: Non-hereditary colorectal cancer (CRC) is a complex disorder resulting from the combination of genetic and non-genetic factors. Genome-wide association studies (GWAS) are useful for identifying such genetic susceptibility factors. However, the single loci so far associated with CRC only represent a fraction of the genetic risk for CRC development in the general population. Therefore, many other genetic risk variants alone and in combination must still remain to be discovered. The aim of this work was to search for genetic risk factors for CRC, by performing single-locus and two-locus GWAS in the Spanish population. RESULTS: A total of 801 controls and 500 CRC cases were included in the discovery GWAS dataset. 77 single nucleotide polymorphisms (SNP)s from single-locus and 243 SNPs from two-locus association analyses were selected for replication in 423 additional CRC cases and 1382 controls. In the meta-analysis, one SNP, rs3987 at 4q26, reached GWAS significant p-value (p = 4.02x10(-8)), and one SNP pair, rs1100508 CG and rs8111948 AA, showed a trend for two-locus association (p = 4.35x10(-11)). Additionally, our GWAS confirmed the previously reported association with CRC of five SNPs located at 3q36.2 (rs10936599), 8q24 (rs10505477), 8q24.21(rs6983267), 11q13.4 (rs3824999) and 14q22.2 (rs4444235). CONCLUSIONS: Our GWAS for CRC patients from Spain confirmed some previously reported associations for CRC and yielded a novel candidate risk SNP, located at 4q26. Epistasis analyses also yielded several novel candidate susceptibility pairs that need to be validated in independent analyses.

Risques, R. A., et al. (2003). "Genetic pathways and genome-wide determinants of clinical outcome in colorectal cancer." Cancer Res **63**(21): 7206-7214.

Various studies have suggested the existence of different pathways of tumor progression in colorectal cancer that associate with specific molecular, chromosomal, and clinicopathological features. We hypothesize that a comprehensive analysis of cumulated genomic damage in colorectal cancers would aid the characterization of different tumor progression pathways and identify the factors determining clinical outcome of tumors of each type. Genome-wide disruption was studied by DNA fingerprinting in a series of 129 sporadic colorectal carcinomas. These results, taken together with data for DNA ploidy, microsatellite instability, p53, and K-ras mutations and clinicopathological characteristics of the patients, have been used to classify colorectal carcinomas. The following five groups can be defined based on the type and level of cumulated genomic damage: (a) tumors with microsatellite instability, right location, and good prognosis; (b) diploid tumors lacking p53 mutations, left and right location, low subchromosomal damage, and bad prognosis; (c) diploid tumors with p53 mutations, left location, high levels of subchromosomal damage, and good prognosis; (d) high aneuploid tumors, p53 mutations, left location, high levels of numerical and structural chromosomal alterations, and bad prognosis; and finally (e) low aneuploid tumors, no p53 mutations, left and right location, low levels of structural chromosomal alterations, and good prognosis. We postulate that these groups represent alternative pathways of tumor progression, each with determinants of aggressiveness. This indicates a need for different prognostic assessments depending on which group the tumor belongs to.

Schumacher, F. R., et al. (2015). "Genome-wide association study of colorectal cancer identifies six new susceptibility loci." Nat Commun **6**: 7138.

Genetic susceptibility to colorectal cancer is caused by rare pathogenic mutations and common genetic variants that contribute to familial risk. Here we report the results of a two-stage association study with 18,299 cases of colorectal cancer and 19,656 controls, with follow-up of the most statistically significant genetic loci in 4,725 cases and 9,969 controls from two Asian consortia. We describe six new susceptibility loci reaching a genome-wide threshold of P<5.0E-08. These findings provide additional insight into the underlying biological mechanisms of colorectal cancer and demonstrate the scientific value of large consortia-based genetic epidemiology studies.

Shaik, A. P., et al. (2015). "Colorectal cancer: A review of the genome-wide association studies in the kingdom of Saudi Arabia." Saudi J Gastroenterol **21**(3): 123-128.

Genome-wise association studies (GWAS) identify risk variants and modifiers that can influence the pathophysiological processes involved in colorectal cancer (CRC) and thus are important to detect associations between disease phenotypes. Our literature review, performed as per PRISMA statement indicates a significant lack of GWAS functional studies in Saudi Arabia. Therefore, studies on sequencing and mapping are needed to identify gene variants that play a role in the pathophysiology of CRC in this specific population. Because it is not apt to generalize disease associations found in other racial and/or ethnic groups to the Arabic or Middle Eastern population, it is very important to conduct GWAS taking into account multiple ethnicities in this region. In addition, linkage studies and case-control studies that include the various confounding and epigenetic factors are needed for appropriate diagnosis of CRC. We recommend that studies in this region be conducted to understand the role of gene-environment interactions across the various ethnic groups, stages of cancer, tumor type, clinical variables, and the population risk to CRC.

Shanmugam, V., et al. (2014). "Whole genome sequencing reveals potential targets for therapy in patients with refractory KRAS mutated metastatic colorectal cancer." BMC Med Genomics **7**: 36.

BACKGROUND: The outcome of patients with metastatic colorectal carcinoma (mCRC) following first line therapy is poor, with median survival of less than one year. The purpose of this study was to identify candidate therapeutically targetable somatic events in mCRC patient samples by whole genome sequencing (WGS), so as to obtain targeted treatment strategies for individual patients. METHODS: Four patients were recruited, all of whom had received > 2 prior therapy regimens. Percutaneous needle biopsies of metastases were performed with whole blood collection for the extraction of constitutional DNA. One tumor was not included in this study as the quality of tumor tissue was not sufficient for further analysis. WGS was performed using Illumina paired end chemistry on HiSeq2000 sequencing systems, which yielded coverage of greater than 30X for all samples. NGS data were processed and analyzed to detect somatic genomic alterations including point mutations, indels, copy number alterations, translocations and rearrangements. RESULTS: All 3 tumor samples had KRAS mutations, while 2 tumors contained mutations in the APC gene and the PIK3CA gene. Although we did not identify a TCF7L2-VTI1A translocation, we

did detect a TCF7L2 mutation in one tumor. Among the other interesting mutated genes was INPPL1, an important gene involved in PI3 kinase signaling. Functional studies demonstrated that inhibition of INPPL1 reduced growth of CRC cells, suggesting that INPPL1 may promote growth in CRC. CONCLUSIONS: Our study further supports potential molecularly defined therapeutic contexts that might provide insights into treatment strategies for refractory mCRC. New insights into the role of INPPL1 in colon tumor cell growth have also been identified. Continued development of appropriate targeted agents towards specific events may be warranted to help improve outcomes in CRC.

Sharma, M. R., et al. (2017). "Exceptional Chemotherapy Response in Metastatic Colorectal Cancer Associated With Hyper-Indel-Hypermutated Cancer Genome and Comutation of POLD1 and MLH1." JCO Precis Oncol **2017**.

Purpose: A73-year-old woman with metastatic colon cancer experienced a complete response to chemotherapy with dose-intensified irinotecan that has been durable for 5 years. We sequenced her tumor and germ line DNA and looked for similar patterns in publicly available genomic data from patients with colorectal cancer. Patients and Methods: Tumor DNA was obtained from a biopsy before therapy, and germ line DNA was obtained from blood. Tumor and germline DNA were sequenced using a commercial panel with approximately 250 genes. Whole-genome amplification and exome sequencing were performed for POLE and POLD1. A POLD1 mutation was confirmed by Sanger sequencing. The somatic mutation and clinical annotation data files from the colon (n = 461) and rectal (n = 171) adenocarcinoma data sets were downloaded from The Cancer Genome Atlas data portal and analyzed for patterns of mutations and clinical outcomes in patients with POLE- and/or POLD1-mutated tumors. Results: The pattern of alterations included APC biallelic inactivation and microsatellite instability high (MSI-H) phenotype, with somatic inactivation of MLH1 and hypermutation (estimated mutation rate > 200 per megabase). The extremely high mutation rate led us to investigate additional mechanisms for hypermutation, including loss of function of POLE. POLE was unaltered, but a related gene not typically associated with somatic mutation in colon cancer, POLD1, had a somatic mutation c.2171G>A [p.Gly724Glu]. Additionally, we noted that the high mutation rate was largely composed of dinucleotide deletions. A similar pattern of hypermutation (dinucleotide deletions, POLD1 mutations, MSI-H) was found in tumors from The Cancer Genome Atlas. Conclusion: POLD1 mutation with associated MSI-H and hyper-indel-hypermutated cancer genome characterizes a previously unrecognized variant of colon cancer that was found in this patient with an exceptional response to chemotherapy.

Shia, J., et al. (2017). "Morphological characterization of colorectal cancers in The Cancer Genome Atlas reveals distinct morphology-molecular associations: clinical and biological implications." Mod Pathol **30**(4): 599-609.

The Cancer Genome Atlas data on colorectal carcinoma have provided a comprehensive view of the tumor's genomic alterations and their tumorigenic roles. Tumor morphology, however, has not been fully integrated into the analysis. The aim of this study was to explore relevant associations between tumor morphology and the newly characterized genomic alterations in colorectal carcinoma. Two hundred and seven colorectal carcinomas that had undergone whole-exome sequencing as part of The Cancer Genome Atlas project and had adequate virtual images in the cBioPortal for Cancer Genomics constituted our study population. Upon analysis, a tight association between 'microsatellite instability-high histology' and microsatellite instability-high ($P<0.001$) was readily detected and helped validate our image-based histology evaluation. Further, we showed, (1) among all histologies, the not otherwise specified type had the lowest overall mutation count ($P<0.001$ for entire cohort, $P<0.03$ for the microsatellite-instable group), and among the microsatellite-instable tumors, this type also correlated with fewer frameshift mutations in coding mononucleotide repeats of a defined set of relevant genes ($P<0.01$); (2) cytosine phosphate guanine island methylator phenotype-high colorectal cancers with or without microsatellite instability tended to have different histological patterns: the former more often mucinous and the latter more often not otherwise specified; (3) mucinous histology was associated with more frequent alterations in BRAF, PIK3CA, and the transforming growth factor-beta pathway when compared with non-mucinous histologies ($P<0.001$, $P=0.01$, and $P<0.001$, respectively); and (4) few colorectal cancers (<9%) exhibited upregulation of immune-inhibitory genes including major immune checkpoints; these tumors were primarily microsatellite-instable (up to 43%, vs <3% in microsatellite-stable group) and had distinctly non-mucinous histologies with a solid growth. These morphology-molecular associations are interesting and propose important clinical implications. The morphological patterns associated with alterations of immune checkpoint genes bear the potential to guide patient selection for clinical trials that target immune checkpoints in colorectal cancer, and provide directions for future studies.

Shiao, S. P. K., et al. (2018). "Genome wide DNA differential methylation regions in colorectal cancer patients in relation to blood related family members, obese and non-obese controls - a preliminary report." Oncotarget **9**(39): 25557-25571.

Despite evidences linking methylation changes in the cancer tissues, little is known about the methylation modification in the peripheral blood. With the current study, we identified differential methylation regions (DMRs) across human genome by collecting the blood samples of colorectal cancer (CRC) patients compared to that of their blood-related family who shared genetic inheritance and environmental influences, and unrelated obese and non-obese controls by accessing publicly available Gene Expression Omnibus data. We performed genome-wide analyses using the reduced representation bisulfite sequencing (RRBS) method covering about 25% of CpGs for whole human genome of the four groups (n = 5 each). In comparison to the non-obese controls, we observed significant DMRs in CRC for genes involved in tumorigenesis including MLH3, MSH2, MSH6, SEPT9, GNAS; and glucose transporter genes associated with obesity and diabetes including SLC2A1/GLUT1, and SLC2A3/GLUT3 that were reported on methylation being modified in cancer tissues. In addition, we observed significant DMRs in CRC for genes involved in the methylation pathways including PEMT, ALDH1L1, and DNMT3A. CRC and family members shared significant DMRs for genes of tumorigenesis including MSH2, SEPT9, GNAS, SLC2A1/GLUT1 and SLC2A3/GLUT3); and CAMK1, GLUT1/SLC2A1 and GLUT3/SLC2A3 genes involved in glucose and insulin metabolism that played vital role in development of obesity and diabetes. Our study provided evidences that these differentially methylated genes in the blood could potentially serve as candidate biomarkers for CRC diagnostic and may provide further understanding on CRC progression. Further studies are warranted to validate these methylation changes for diagnostic and prevention of CRC.

Siegert, S., et al. (2013). "Genome-wide investigation of gene-environment interactions in colorectal cancer." Hum Genet **132**(2): 219-231.

Colorectal cancer (CRC), one of the most frequent neoplasias worldwide, has both genetic and environmental causes. As yet, however, gene-environment (G x E) interactions in CRC have been studied mostly for a small number of candidate genes only. Therefore, we investigated the possible interaction, in CRC etiology, between single-nucleotide polymorphisms (SNPs) on the one hand, and overweight, smoking and alcohol consumption on the other, at a genome-wide level. To this end, we adopted a two-tiered approach comprising a case-only screening stage I (314 cases) and a case-control validation stage II (259 cases, 1,002 controls). Interactions with the smallest p value in stage I were verified in stage II using multiple logistic regression analysis adjusted for sex and age. In addition, we specifically studied known CRC-associated SNPs for possible G x E interactions. Upon adjustment for sex and age, and after allowing for multiple testing, however, only a single SNP (rs1944511) was found to be involved in a statistically significant interaction, namely with overweight (multiplicity-corrected p = 0.042 in stage II). Several other G x E interactions were nominally significant but failed correction for multiple testing, including a previously reported interaction between rs9929218 and alcohol consumption that also emerged in our candidate SNP study (nominal p = 0.008). Notably, none of the interactions identified in our genome-wide analysis was with a previously reported CRC-associated SNP. Our study therefore highlights the potential of an "agnostic" genome-wide approach to G x E analysis.

Sipos, F., et al. (2013). "Genome-wide screening for understanding the role of DNA methylation in colorectal cancer." Epigenomics **5**(5): 569-581.

DNA methylation analysis methods have undergone an impressive revolution over the past 15 years. Regarding colorectal cancer (CRC), the localization and distribution of several differently methylated genes have been determined by genome-wide DNA methylation assays. These genes do not just influence the pathogenesis of CRC, but can be used further as diagnostic or prognostic markers. Moreover, the identified four DNA methylation-based subgroups of CRC have important clinical and therapeutic merit. Since genome-wide DNA methylation analyzes result in a large amount of data, there is a need for complex bioinformatic and pathway analysis. Future challenges in epigenetic alterations of CRC include the demand for comprehensive identification and experimental validation of gene abnormalities. By introduction of genome-wide DNA methylation profiling into clinical practice not only the patients' risk stratification but development of targeted therapies will also be possible.

Skowronski, K., et al. (2014). "Genome-wide analysis in human colorectal cancer cells reveals ischemia-mediated expression of motility genes via DNA hypomethylation." PLoS One **9**(7): e103243.

DNA hypomethylation is an important epigenetic modification found to occur in many different cancer types, leading to the upregulation of previously silenced genes and loss of genomic stability. We

previously demonstrated that hypoxia and hypoglycaemia (ischemia), two common micro-environmental changes in solid tumours, decrease DNA methylation through the downregulation of DNMTs in human colorectal cancer cells. Here, we utilized a genome-wide cross-platform approach to identify genes hypomethylated and upregulated by ischemia. Following exposure to hypoxia or hypoglycaemia, methylated DNA from human colorectal cancer cells (HCT116) was immunoprecipitated and analysed with an Affymetrix promoter array. Additionally, RNA was isolated and analysed in parallel with an Affymetrix expression array. Ingenuity pathway analysis software revealed that a significant proportion of the genes hypomethylated and upregulated were involved in cellular movement, including PLAUR and CYR61. A Matrigel invasion assay revealed that indeed HCT116 cells grown in hypoxic or hypoglycaemic conditions have increased mobility capabilities. Confirmation of upregulated expression of cellular movement genes was performed with qPCR. The correlation between ischemia and metastasis is well established in cancer progression, but the molecular mechanisms responsible for this common observation have not been clearly identified. Our novel data suggests that hypoxia and hypoglycaemia may be driving changes in DNA methylation through downregulation of DNMTs. This is the first report to our knowledge that provides an explanation for the increased metastatic potential seen in ischemic cells; i.e. that ischemia could be driving DNA hypomethylation and increasing expression of cellular movement genes.

Sokolova, V., et al. (2016). "Integration of genome scale data for identifying new players in colorectal cancer." World J Gastroenterol **22**(2): 534-545.

Colorectal cancers (CRCs) display a wide variety of genomic aberrations that may be either causally linked to their development and progression, or might serve as biomarkers for their presence. Recent advances in rapid high-throughput genetic and genomic analysis have helped to identify a plethora of alterations that can potentially serve as new cancer biomarkers, and thus help to improve CRC diagnosis, prognosis, and treatment. Each distinct data type (copy number variations, gene and microRNAs expression, CpG island methylation) provides an investigator with a different, partially independent, and complementary view of the entire genome. However, elucidation of gene function will require more information than can be provided by analyzing a single type of data. The integration of knowledge obtained from different sources is becoming increasingly essential for obtaining an interdisciplinary view of large amounts of information, and also for cross-validating experimental results. The integration of numerous types of genetic and genomic data derived from public sources, and via the use of ad-hoc bioinformatics tools and statistical methods facilitates the discovery and validation of novel, informative biomarkers. This combinatory approach will also enable researchers to more accurately and comprehensively understand the associations between different biologic pathways, mechanisms, and phenomena, and gain new insights into the etiology of CRC.

Staub, E., et al. (2006). "A genome-wide map of aberrantly expressed chromosomal islands in colorectal cancer." Mol Cancer **5**: 37.

BACKGROUND: Cancer development is accompanied by genetic phenomena like deletion and amplification of chromosome parts or alterations of chromatin structure. It is expected that these mechanisms have a strong effect on regional gene expression. RESULTS: We investigated genome-wide gene expression in colorectal carcinoma (CRC) and normal epithelial tissues from 25 patients using oligonucleotide arrays. This allowed us to identify 81 distinct chromosomal islands with aberrant gene expression. Of these, 38 islands show a gain in expression and 43 a loss of expression. In total, 7.892 genes (25.3% of all human genes) are located in aberrantly expressed islands. Many chromosomal regions that are linked to hereditary colorectal cancer show deregulated expression. Also, many known tumor genes localize to chromosomal islands of misregulated expression in CRC. CONCLUSION: An extensive comparison with published CGH data suggests that chromosomal regions known for frequent deletions in colon cancer tend to show reduced expression. In contrast, regions that are often amplified in colorectal tumors exhibit heterogeneous expression patterns: even show a decrease of mRNA expression. Because for several islands of deregulated expression chromosomal aberrations have never been observed, we speculate that additional mechanisms (like abnormal states of regional chromatin) also have a substantial impact on the formation of co-expression islands in colorectal carcinoma.

Study, C., et al. (2008). "Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer." Nat Genet **40**(12): 1426-1435.

Genome-wide association (GWA) studies have identified multiple loci at which common variants modestly influence the risk of developing colorectal cancer (CRC). To enhance power to identify additional loci with similar effect sizes, we conducted a meta-analysis of two GWA studies, comprising 13,315

individuals genotyped for 38,710 common tagging SNPs. We undertook replication testing in up to eight independent case-control series comprising 27,418 subjects. We identified four previously unreported CRC risk loci at 14q22.2 (rs4444235, BMP4; P = 8.1 x 10(-10)), 16q22.1 (rs9929218, CDH1; P = 1.2 x 10(-8)), 19q13.1 (rs10411210, RHPN2; P = 4.6 x 10(-9)) and 20p12.3 (rs961253; P = 2.0 x 10(-10)). These findings underscore the value of large sample series for discovery and follow-up of genetic variants contributing to the etiology of CRC.

Suzuki, H., et al. (2011). "Genome-wide profiling of chromatin signatures reveals epigenetic regulation of MicroRNA genes in colorectal cancer." Cancer Res **71**(17): 5646-5658.

Altered expression of microRNAs (miRNA) occurs commonly in human cancer, but the mechanisms are generally poorly understood. In this study, we examined the contribution of epigenetic mechanisms to miRNA dysregulation in colorectal cancer by carrying out high-resolution ChIP-seq. Specifically, we conducted genome-wide profiling of trimethylated histone H3 lysine 4 (H3K4me3), trimethylated histone H3 lysine 27 (H3K27me3), and dimethylated histone H3 lysine 79 (H3K79me2) in colorectal cancer cell lines. Combining miRNA expression profiles with chromatin signatures enabled us to predict the active promoters of 233 miRNAs encoded in 174 putative primary transcription units. By then comparing miRNA expression and histone modification before and after DNA demethylation, we identified 47 miRNAs encoded in 37 primary transcription units as potential targets of epigenetic silencing. The promoters of 22 transcription units were associated with CpG islands (CGI), all of which were hypermethylated in colorectal cancer cells. DNA demethylation led to increased H3K4me3 marking at silenced miRNA genes, whereas no restoration of H3K79me2 was detected in CGI-methylated miRNA genes. DNA demethylation also led to upregulation of H3K4me3 and H3K27me3 in a number of CGI-methylated miRNA genes. Among the miRNAs we found to be dysregulated, many of which are implicated in human cancer, miR-1-1 was methylated frequently in early and advanced colorectal cancer in which it may act as a tumor suppressor. Our findings offer insight into the association between chromatin signatures and miRNA dysregulation in cancer, and they also suggest that miRNA reexpression may contribute to the effects of epigenetic therapy.

Takahashi, Y., et al. (2017). "Japanese genome-wide association study identifies a significant colorectal cancer susceptibility locus at chromosome 10p14." Cancer Sci **108**(11): 2239-2247.

Genome-wide association studies are a powerful tool for searching for disease susceptibility loci. Several studies identifying single nucleotide polymorphisms (SNP) connected intimately to the onset of colorectal cancer (CRC) have been published, but there are few reports of genome-wide association studies in Japan. To identify genetic variants that modify the risk of CRC oncogenesis, especially in the Japanese population, we performed a multi-stage genome-wide association study using a large number of samples: 1846 CRC cases and 2675 controls. We identified 4 SNP (rs7912831, rs4749812, rs7898455 and rs10905453) in chromosome region 10p14 associated with CRC; however, there are no coding or non-coding genes within this region of fairly extensive linkage disequilibrium (a 500-kb block) on 10p14. Our study revealed that the 10p14 locus is significantly correlated with susceptibility to CRC in the Japanese population, in accordance with the results of multiple studies in other races.

Tan, C., et al. (2016). "Risk of eighteen genome-wide association study-identified genetic variants for colorectal cancer and colorectal adenoma in Han Chinese." Oncotarget **7**(47): 77651-77663.

BACKGROUND: Recent genome-wide association studies (GWAS) identified eighteen single-nucleotide polymorphisms (SNPs) to be significantly associated with the risk of colorectal cancer (CRC). However, overall results of the following replications are inconsistent and little is known about whether these associations also exit in colorectal adenomas (CRA). METHODS: The SNP genotyping was performed using a Sequenom MassARRAY to investigate the association of these eighteen SNPs with colorectal neoplasm in a case-control study consisted of 1049 colorectal cancers, 283 adenomas, and 1030 controls. RESULTS: Two of these SNPs, rs10505477 and rs719725, showed evidence of an association in both CRC and CRA in our study population. Besides, seven SNPs (rs10808555, rs7014346, rs7837328, rs704017, rs11196172, rs4779584, and rs7229639) were significantly associated with CRC, and another one SNP rs11903757 was over-represented in CRA compared with controls. The strongest association was provided by rs11196172 (OR = 2.02, 95% CI = 1.66 - 2.46, P < 0.0001) and rs11903757 (OR = 1.96, 95% CI = 1.28 - 3.00, P = 0.0026). CONCLUSION: These results suggest that some previously reported SNP associations also have impact on CRC and CRA predispositions in the Han Chinese population. A part of genetic risk to CRC is possibly mediated by susceptibility to adenomas.

Tanami, H., et al. (2005). "Involvement of cyclin D3 in liver metastasis of colorectal cancer, revealed by

genome-wide copy-number analysis." Lab Invest **85**(9): 1118-1129.

The question of whether any genetic differences exist between primary and colorectal cancers (CRCs) and their metastatic foci is controversial. To look for genetic aberrations involved in metastasis of CRCs to the liver, we performed subtractive comparative genomic hybridization (CGH) experiments using paired samples from 20 CRC patients with primary tumors and synchronous or metachronous liver metastases. Relatively frequent gains in DNA copy number were detected at 6p, suggesting the presence of one or more metastasis-related genes in the region. Analysis of 11 CRC cell lines using array-based CGH (CGH-array) revealed one 6p candidate gene, CCND3. Quantitative reverse transcriptase-polymerase chain reaction experiments showed that CCND3 was significantly upregulated in liver-metastatic lesions compared with primary lesions (P<0.0152). In addition, immunohistochemical analysis of 120 primary CRC tumors demonstrated that cyclin D3 expression in the region of rolled edge was significantly associated with total recurrence, especially hematogenous recurrence (P=0.0307). The results implied involvement of cyclin D3 in liver metastasis of CRC, and the data may contribute to the development of a novel therapy or diagnostic agent for this currently intractable disease. Our experiments also confirmed the power of subtractive CGH and CGH-array analysis for identifying cancer-related genes.

Tanskanen, T., et al. (2018). "Genome-wide association study and meta-analysis in Northern European populations replicate multiple colorectal cancer risk loci." Int J Cancer **142**(3): 540-546.

Genome-wide association studies have been successful in elucidating the genetic basis of colorectal cancer (CRC), but there remains unexplained variability in genetic risk. To identify new risk variants and to confirm reported associations, we conducted a genome-wide association study in 1,701 CRC cases and 14,082 cancer-free controls from the Finnish population. A total of 9,068,015 genetic variants were imputed and tested, and 30 promising variants were studied in additional 11,647 cases and 12,356 controls of European ancestry. The previously reported association between the single-nucleotide polymorphism (SNP) rs992157 (2q35) and CRC was independently replicated (p = 2.08 x 10(-4); OR, 1.14; 95% CI, 1.06-1.23), and it was genome-wide significant in combined analysis (p = 1.50 x 10(-9); OR, 1.12; 95% CI, 1.08-1.16). Variants at 2q35, 6p21.2, 8q23.3, 8q24.21, 10q22.3, 10q24.2, 11q13.4, 11q23.1, 14q22.2, 15q13.3, 18q21.1, 20p12.3 and 20q13.33 were associated with CRC in the Finnish population (false discovery rate < 0.1), but new risk loci were not found. These results replicate the effects of multiple loci on the risk of CRC and identify shared risk alleles between the Finnish population isolate and outbred populations.

Tenesa, A. and M. G. Dunlop (2009). "New insights into the aetiology of colorectal cancer from genome-wide association studies." Nat Rev Genet **10**(6): 353-358.

Genome-wide association studies have recently identified ten common genetic variants associated with colorectal cancer susceptibility, several suggesting the involvement of components of the transforming growth factor beta (TGFbeta) superfamily signalling pathway. To date, no causal sequence variants have been identified, and risk seems to be mediated through effects on gene regulation. Several markers are located close to poorly characterized genes or in gene deserts, raising challenges for elucidating mechanisms of susceptibility. Disease-associated common genetic variation offers the potential to refine risk stratification within populations and enable more targeted disease prevention strategies.

Tenesa, A., et al. (2008). "Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21." Nat Genet **40**(5): 631-637.

In a genome-wide association study to identify loci associated with colorectal cancer (CRC) risk, we genotyped 555,510 SNPs in 1,012 early-onset Scottish CRC cases and 1,012 controls (phase 1). In phase 2, we genotyped the 15,008 highest-ranked SNPs in 2,057 Scottish cases and 2,111 controls. We then genotyped the five highest-ranked SNPs from the joint phase 1 and 2 analysis in 14,500 cases and 13,294 controls from seven populations, and identified a previously unreported association, rs3802842 on 11q23 (OR = 1.1; P = 5.8 x 10(-10)), showing population differences in risk. We also replicated and fine-mapped associations at 8q24 (rs7014346; OR = 1.19; P = 8.6 x 10(-26)) and 18q21 (rs4939827; OR = 1.2; P = 7.8 x 10(-28)). Risk was greater for rectal than for colon cancer for rs3802842 (P < 0.008) and rs4939827 (P < 0.009). Carrying all six possible risk alleles yielded OR = 2.6 (95% CI = 1.75-3.89) for CRC. These findings extend our understanding of the role of common genetic variation in CRC etiology.

Teo, A. S., et al. (2015). "Single-molecule optical genome mapping of a human HapMap and a colorectal cancer cell line." Gigascience **4**: 65.

BACKGROUND: Next-generation sequencing (NGS) technologies have changed our understanding of the variability of the human genome. However, the identification of genome structural variations based on

NGS approaches with read lengths of 35-300 bases remains a challenge. Single-molecule optical mapping technologies allow the analysis of DNA molecules of up to 2 Mb and as such are suitable for the identification of large-scale genome structural variations, and for de novo genome assemblies when combined with short-read NGS data. Here we present optical mapping data for two human genomes: the HapMap cell line GM12878 and the colorectal cancer cell line HCT116. FINDINGS: High molecular weight DNA was obtained by embedding GM12878 and HCT116 cells, respectively, in agarose plugs, followed by DNA extraction under mild conditions. Genomic DNA was digested with KpnI and 310,000 and 296,000 DNA molecules (>/= 150 kb and 10 restriction fragments), respectively, were analyzed per cell line using the Argus optical mapping system. Maps were aligned to the human reference by OPTIMA, a new glocal alignment method. Genome coverage of 6.8x and 5.7x was obtained, respectively; 2.9x and 1.7x more than the coverage obtained with previously available software. CONCLUSIONS: Optical mapping allows the resolution of large-scale structural variations of the genome, and the scaffold extension of NGS-based de novo assemblies. OPTIMA is an efficient new alignment method; our optical mapping data provide a resource for genome structure analyses of the human HapMap reference cell line GM12878, and the colorectal cancer cell line HCT116.

Thean, L. F., et al. (2010). "Genome-wide scan identifies a copy number variable region at 3q26 that regulates PPM1L in APC mutation-negative familial colorectal cancer patients." Genes Chromosomes Cancer **49**(2): 99-106.

Familial adenomatous polyposis (FAP) is an autosomal dominantly inherited form of colorectal cancer (CRC) caused by mutation in the adenomatous polyposis coli (APC) gene. However, APC mutations are not detected in 10-50% of FAP patients. We searched for a new cancer gene by performing genome-wide genotyping on members of an APC mutation-negative FAP variant family and ethnicity-matched healthy controls. No common copy number change was found in all affected members using the unaffected members and healthy controls as baseline. A 111 kb copy number variable (CNV) region at 3q26.1 was shown to have copy number loss in all eight polyps compared to matched lymphocytes of two affected members. A common region of loss in all polyps, which are precursors to CRC, is likely to harbor disease-causing gene in accordance to Knudsen's "two-hit" hypothesis. There is, however, no gene within the deleted region. A 2-Mb scan of the genomic region encompassing the deleted region identified PPM1L, coding for a novel serine-threonine phosphatase in the TGF-beta and BMP signaling pathways. Real-time PCR analyses indicate that the 3'UTR of PPM1L transcript was down-regulated more than two-folds in all six polyps and tumors compared to matched mucosa of the affected member. This down-regulation was not observed in APC mutation-positive FAP patients. Our results suggest that the CNV region at 3q26 harbors an element that regulates the expression of an upstream candidate tumor suppressor, PPM1L, thus providing a novel mechanism for colorectal tumorigenesis in APC mutation-negative familial CRC patients.

Thean, L. F., et al. (2018). "Genome-wide association study identified copy number variants associated with sporadic colorectal cancer risk." J Med Genet **55**(3): 181-188.

BACKGROUND: Multiple single nucleotide polymorphisms (SNPs) have been associated with colorectal cancer (CRC) risk. The role of structural or copy number variants (CNV) in CRC, however, remained unclear. We investigated the role of CNVs in patients with sporadic CRC. METHODS: A genome-wide association study (GWAS) was performed on 1000 Singapore Chinese patients aged 50 years or more with no family history of CRC and 1000 ethnicity-matched, age-matched and gender-matched healthy controls using the Affymetrix SNP 6 platform. After 16 principal component corrections, univariate and multivariate segmentations followed by association testing were performed on 1830 samples that passed quality assurance tests. RESULTS: A rare CNV region (CNVR) at chromosome 14q11 (OR=1.92 (95% CI 1.59 to 2.32), p=2.7e-12) encompassing CHD8, and common CNVR at chromosomes 3q13.12 (OR=1.54 (95% CI 1.33 to 1.77), p=2.9e-9) and 12p12.3 (OR=1.69 (95% CI 1.41 to 2.01), p=2.8e-9) encompassing CD47 and RERG/ARHGDIB, respectively, were significantly associated with CRC risk. CNV loci were validated in an independent replication panel using an optimised copy number assay. Whole-genome expression data in matched tumours of a subset of cases demonstrated that copy number loss at CHD8 was significantly associated with dysregulation of several genes that perturb the Wnt, TP53 and inflammatory pathways. CONCLUSIONS: A rare CNVR at 14q11 encompassing the chromatin modifier CHD8 was significantly associated with sporadic CRC risk. Copy number loss at CHD8 altered expressions of genes implicated in colorectal tumourigenesis.

Theodoratou, E., et al. (2018). "Genome-wide scan of the effect of common nsSNPs on colorectal cancer survival outcome." Br J Cancer.

BACKGROUND: We conducted a genome-wide scan to identify non-synonymous SNPs (nsSNPs) that might influence survival after a diagnosis of colorectal cancer (CRC). METHODS: We genotyped 7679 nsSNPs in 1939 Scottish patients from the Scottish Colorectal Cancer Study recruited soon after a CRC diagnosis and prospectively followed for survival outcomes. All-cause and CRC-specific survival analyses were conducted using Cox proportional hazard models adjusted for stage, age and sex for all cancer cases, after cancer type stratification and assuming additive and recessive models of inheritance. For all the SNPs that had a p-value < 0.10 a meta-analysis was performed combining the results of the discovery set and a replication set of 899 Scottish CRC patients. The p-value threshold of significance was set as at p < 10(-8). RESULTS: 897 and 894 nsSNPs were associated with all-cause and CRC-specific mortality, respectively, at a p-value level < 0.10 in the discovery set. Meta-analysis of the results from the discovery and replication sets was performed overall and for cancers of colon and rectum separately and none of the variants reached a p-value < 10(-8). CONCLUSIONS: This large scale well-powered analysis demonstrates that common nsSNPs are not associated with CRC prognosis overall.

Thorsen, K., et al. (2011). "Tumor-specific usage of alternative transcription start sites in colorectal cancer identified by genome-wide exon array analysis." BMC Genomics **12**: 505.

BACKGROUND: Approximately half of all human genes use alternative transcription start sites (TSSs) to control mRNA levels and broaden the transcriptional output in healthy tissues. Aberrant expression patterns promoting carcinogenesis, however, may arise from alternative promoter usage. RESULTS: By profiling 108 colorectal samples using exon arrays, we identified nine genes (TCF12, OSBPL1A, TRAK1, ANK3, CHEK1, UGP2, LMO7, ACSL5, and SCIN) showing tumor-specific alternative TSS usage in both adenoma and cancer samples relative to normal mucosa. Analysis of independent exon array data sets corroborated these findings. Additionally, we confirmed the observed patterns for selected mRNAs using quantitative real-time reverse-transcription PCR. Interestingly, for some of the genes, the tumor-specific TSS usage was not restricted to colorectal cancer. A comprehensive survey of the nine genes in lung, bladder, liver, prostate, gastric, and brain cancer revealed significantly altered mRNA isoform ratios for CHEK1, OSBPL1A, and TCF12 in a subset of these cancer types.To identify the mechanism responsible for the shift in alternative TSS usage, we antagonized the Wnt-signaling pathway in DLD1 and Ls174T colorectal cancer cell lines, which remarkably led to a shift in the preferred TSS for both OSBPL1A and TRAK1. This indicated a regulatory role of the Wnt pathway in selecting TSS, possibly also involving TP53 and SOX9, as their transcription binding sites were enriched in the promoters of the tumor preferred isoforms together with their mRNA levels being increased in tumor samples. Finally, to evaluate the prognostic impact of the altered TSS usage, immunohistochemistry was used to show deregulation of the total protein levels of both TCF12 and OSBPL1A, corresponding to the mRNA levels observed. Furthermore, the level of nuclear TCF12 had a significant correlation to progression free survival in a cohort of 248 stage II colorectal cancer samples. CONCLUSIONS: Alternative TSS usage in colorectal adenoma and cancer samples has been shown for nine genes, and OSBPL1A and TRAK1 were found to be regulated in vitro by Wnt signaling. TCF12 protein expression was upregulated in cancer samples and correlated with progression free survival.

Tomlinson, I., et al. (2007). "A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21." Nat Genet **39**(8): 984-988.

Much of the variation in inherited risk of colorectal cancer (CRC) is probably due to combinations of common low risk variants. We conducted a genome-wide association study of 550,000 tag SNPs in 930 familial colorectal tumor cases and 960 controls. The most strongly associated SNP (P = 1.72 x 10(-7), allelic test) was rs6983267 at 8q24.21. To validate this finding, we genotyped rs6983267 in three additional CRC case-control series (4,361 affected individuals and 3,752 controls; 1,901 affected individuals and 1,079 controls; 1,072 affected individuals and 415 controls) and replicated the association, providing P = 1.27 x 10(-14) (allelic test) overall, with odds ratios (ORs) of 1.27 (95% confidence interval (c.i.): 1.16-1.39) and 1.47 (95% c.i.: 1.34-1.62) for heterozygotes and rare homozygotes, respectively. Analyses based on 1,477 individuals with colorectal adenoma and 2,136 controls suggest that susceptibility to CRC is mediated through development of adenomas (OR = 1.21, 95% c.i.: 1.10-1.34; P = 6.89 x 10(-5)). These data show that common, low-penetrance susceptibility alleles predispose to colorectal neoplasia.

Tomlinson, I. P., et al. (2008). "A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3." Nat Genet **40**(5): 623-630.

To identify colorectal cancer (CRC) susceptibility alleles, we conducted a genome-wide association study. In phase 1, we genotyped 550,163

tagSNPs in 940 familial colorectal tumor cases (627 CRC, 313 high-risk adenoma) and 965 controls. In phase 2, we genotyped 42,708 selected SNPs in 2,873 CRC cases and 2,871 controls. In phase 3, we evaluated 11 SNPs showing association at P < 10(-4) in a joint analysis of phases 1 and 2 in 4,287 CRC cases and 3,743 controls. Two SNPs were taken forward to phase 4 genotyping (10,731 CRC cases and 10,961 controls from eight centers). In addition to the previously reported 8q24, 15q13 and 18q21 CRC risk loci, we identified two previously unreported associations: rs10795668, located at 10p14 (P = 2.5 x 10(-13) overall; P = 6.9 x 10(-12) replication), and rs16892766, at 8q23.3 (P = 3.3 x 10(-18) overall; P = 9.6 x 10(-17) replication), which tags a plausible causative gene, EIF3H. These data provide further evidence for the 'common-disease common-variant' model of CRC predisposition.

Uddin, S., et al. (2011). "Genome-wide expression analysis of Middle Eastern colorectal cancer reveals FOXM1 as a novel target for cancer therapy." Am J Pathol **178**(2): 537-547.

To identify genes potentially playing an important role in the progression of colorectal carcinoma (CRC), we screened global gene expression using cDNA expression array on 41 CRC tissue samples and 25 noncancerous colorectal tissue samples. Among the up-regulated genes, forkhead box M1 (FOXM1) has been shown to play a critical role in pathogenesis of various malignancies. Using immunohistochemistry on 448 Saudi CRC samples in tissue microarray format, FoxM1 protein overexpression was seen in 66% of CRC tissues and was significantly associated with poorly differentiated and highly proliferative tumors (P = 0.0200 and 0.0018, respectively). FoxM1 expression was also significantly associated with MMP-9 protein expression (P = 0.0002). In vitro data using CRC cell lines showed that inhibition of FoxM1 by thiostrepton resulted in inhibition of proliferation and induction of apoptosis in a dose-dependent manner. Overexpression of FoxM1 potentiated cell proliferation, cell transformation, and migration/invasion of CRC cells via up-regulation of FoxM1 target genes MMP2 and MMP9 and protected these cells from thiostrepton-mediated antiproliferative effects. Finally, in vivo, overexpression of FoxM1 promoted growth of CRC-cell line xenograft tumors in nude mice. Altogether, our data indicate that FoxM1 signaling contributes to aggressiveness in a subset of CRC and that the FOXM1 gene may serve as a useful molecular biomarker and potential therapeutic target.

Vishnubalaji, R., et al. (2015). "Genome-wide mRNA and miRNA expression profiling reveal multiple regulatory networks in colorectal cancer." Cell Death Dis **6**: e1614.

Despite recent advances in cancer management, colorectal cancer (CRC) remains the third most common cancer and a major health-care problem worldwide. MicroRNAs have recently emerged as key regulators of cancer development and progression by targeting multiple cancer-related genes; however, such regulatory networks are not well characterized in CRC. Thus, the aim of this study was to perform global messenger RNA (mRNA) and microRNA expression profiling in the same CRC samples and adjacent normal tissues and to identify potential miRNA-mRNA regulatory networks. Our data revealed 1273 significantly upregulated and 1902 downregulated genes in CRC. Pathway analysis revealed significant enrichment in cell cycle, integrated cancer, Wnt (wingless-type MMTV integration site family member), matrix metalloproteinase, and TGF-beta pathways in CRC. Pharmacological inhibition of Wnt (using XAV939 or IWP-2) or TGF-beta (using SB-431542) pathways led to dose- and time-dependent inhibition of CRC cell growth. Similarly, our data revealed up- (42) and downregulated (61) microRNAs in the same matched samples. Using target prediction and bioinformatics, ~77% of the upregulated genes were predicted to be targeted by microRNAs found to be downregulated in CRC. We subsequently focused on EZH2 (enhancer of zeste homolog 2), which was found to be regulated by hsa-miR-26a-5p and several members of the let-7 (lethal-7) family in CRC. Significant inverse correlation between EZH2 and hsa-miR-26a-5p (R (2)=0.56, P=0.0001) and hsa-let-7b-5p (R (2)=0.19, P=0.02) expression was observed in the same samples, corroborating the belief of EZH2 being a bona fide target for these two miRNAs in CRC. Pharmacological inhibition of EZH2 led to significant reduction in trimethylated histone H3 on lysine 27 (H3K27) methylation, marked reduction in cell proliferation, and migration in vitro. Concordantly, small interfering RNA-mediated knockdown of EZH2 led to similar effects on CRC cell growth in vitro. Therefore, our data have revealed several hundred potential miRNA-mRNA regulatory networks in CRC and suggest targeting relevant networks as potential therapeutic strategy for CRC.

Vogtmann, E., et al. (2016). "Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing." PLoS One **11**(5): e0155362.

Accumulating evidence indicates that the gut microbiota affects colorectal cancer development, but previous studies have varied in population, technical methods, and associations with cancer. Understanding these variations is needed for comparisons and for

potential pooling across studies. Therefore, we performed whole-genome shotgun sequencing on fecal samples from 52 pre-treatment colorectal cancer cases and 52 matched controls from Washington, DC. We compared findings from a previously published 16S rRNA study to the metagenomics-derived taxonomy within the same population. In addition, metagenome-predicted genes, modules, and pathways in the Washington, DC cases and controls were compared to cases and controls recruited in France whose specimens were processed using the same platform. Associations between the presence of fecal Fusobacteria, Fusobacterium, and Porphyromonas with colorectal cancer detected by 16S rRNA were reproduced by metagenomics, whereas higher relative abundance of Clostridia in cancer cases based on 16S rRNA was merely borderline based on metagenomics. This demonstrated that within the same sample set, most, but not all taxonomic associations were seen with both methods. Considering significant cancer associations with the relative abundance of genes, modules, and pathways in a recently published French metagenomics dataset, statistically significant associations in the Washington, DC population were detected for four out of 10 genes, three out of nine modules, and seven out of 17 pathways. In total, colorectal cancer status in the Washington, DC study was associated with 39% of the metagenome-predicted genes, modules, and pathways identified in the French study.

Vychytilova-Faltejskova, P., et al. (2016). "Genome-wide microRNA Expression Profiling in Primary Tumors and Matched Liver Metastasis of Patients with Colorectal Cancer." Cancer Genomics Proteomics **13**(4): 311-316.

BACKGROUND: Primary tumor spread to the liver is the major cause of disease progression and death in patients with colorectal cancer (CRC). MicroRNAs (miRNAs) are small non-coding RNAs that are involved in cancer development and progression, but their role in metastasis has not been extensively investigated. MATERIALS AND METHODS: Firstly, expression profiling of 752 miRNAs in 20 primary tumors and their corresponding liver metastases was performed. Secondly, validation of the results was carried out on an independent cohort of 66 patients with metastatic CRC using reverse transcription-quantitative polymerase chain (RT-qPCR) reaction. RESULTS: In total, 33 miRNAs were found to be significantly deregulated in liver metastases compared to their primary tumors. Fifteen miRNAs were chosen for subsequent validation, which confirmed significantly reduced expression of miR-143, miR 10b, and miR-28-5p, and increased expression of miR 122, miR-122*, and miR 885-5p in the tissue of liver metastases. CONCLUSION: These results indicate that miRNAs could serve as new therapeutic targets in patients with metastatic CRC.

Wang, H., et al. (2014). "Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A." Nat Commun **5**: 4613.

The genetic basis of sporadic colorectal cancer (CRC) is not well explained by known risk polymorphisms. Here we perform a meta-analysis of two genome-wide association studies in 2,627 cases and 3,797 controls of Japanese ancestry and 1,894 cases and 4,703 controls of African ancestry, to identify genetic variants that contribute to CRC susceptibility. We replicate genome-wide statistically significant associations (P<5 x 10(-8)) in 16,823 cases and 18,211 controls of European ancestry. This study reveals a new pan-ethnic CRC risk locus at 10q25 (rs12241008, intronic to VTI1A; P=1.4 x 10(-9)), providing additional insight into the aetiology of CRC and highlighting the value of association mapping in diverse populations.

Wang, H., et al. (2013). "Fine-mapping of genome-wide association study-identified risk loci for colorectal cancer in African Americans." Hum Mol Genet **22**(24): 5048-5055.

Genome-wide association studies of colorectal cancer (CRC) in Europeans and Asians have identified 21 risk susceptibility regions [29 index single-nucleotide polymorphisms (SNPs)]. Characterizing these risk regions in diverse racial groups with different linkage disequilibrium (LD) structure can help localize causal variants. We examined associations between CRC and all 29 index SNPs in 6597 African Americans (1894 cases and 4703 controls). Nine SNPs in eight regions (5q31.1, 6q26-q27, 8q23.3, 8q24.21, 11q13.4, 15q13.3, 18q21.1 and 20p12.3) formally replicated in our data with one-sided P-values <0.05 and the same risk directions as reported previously. We performed fine-mapping of the 21 risk regions (including 250 kb on both sides of the index SNPs) using genotyped and imputed markers at the density of the 1000 Genomes Project to search for additional or more predictive risk markers. Among the SNPs correlated with the index variants, two markers, rs12759486 (or rs7547751, a putative functional variant in perfect LD with it) in 1q41 and rs7252505 in 19q13.1, were more strongly and statistically significantly associated with CRC (P < 0.0006). The average per allele risk was improved using the replicated index variants and the two new markers (odds ratio = 1.14, P = 6.5 x 10(-16)) in African Americans, compared with using all index SNPs (odds ratio = 1.07, P = 3.4 x 10(-10)). The

contribution of the two new risk SNPs to CRC heritability was estimated to be 1.5% in African Americans. This study highlights the importance of fine-mapping in diverse populations.

Wang, H. M., et al. (2013). "A new method for post Genome-Wide Association Study (GWAS) analysis of colorectal cancer in Taiwan." Gene **518**(1): 107-113.

Recently, single nucleotide polymorphisms (SNPs) located in specific loci or genes have been identified associated with susceptibility to colorectal cancer (CRC) in Genome-Wide Association Studies (GWAS). However, in different ethnicities and regions, the genetic variations and the environmental factors can widely vary. Therefore, here we propose a post-GWAS analysis method to investigate the CRC susceptibility SNPs in Taiwan by conducting a replication analysis and bioinformatics analysis. One hundred and forty-four significant SNPs from published GWAS results were collected by a literature survey, and two hundred and eighteen CRC samples and 385 normal samples were collected for post-GWAS analysis. Finally, twenty-six significant SNPs were identified and reported as associated with susceptibility to colorectal cancer, other cancers, obesity, and celiac disease in a previous GWAS study. Functional analysis results of 26 SNPs indicate that most biological processes identified are involved in regulating immune responses and apoptosis. In addition, an efficient prediction model was constructed by applying Jackknife feature selection and ANOVA testing. As compared to another risk prediction model of CRC for European Caucasians population, which performs 0.616 of AUC by using 54 SNPs, the proposed model shows good performance in predicting CRC risk within the Taiwanese population, i.e., 0.724 AUC by using 16 SNPs. We believe that the proposed risk prediction model is highly promising for predicting CRC risk within the Taiwanese population. In addition, the functional analysis results could be helpful to explore the potential associated regulatory mechanisms that may be involved in CRC development.

Wang, T. L., et al. (2002). "Prevalence of somatic alterations in the colorectal cancer cell genome." Proc Natl Acad Sci U S A **99**(5): 3076-3080.

Although a small fraction of human cancers have increased rates of somatic mutation because of known deficiencies in DNA repair, little is known about the prevalence of somatic alterations in the vast majority of human cancers. To systematically assess nonsynonymous somatic alterations in colorectal neoplasia, we used DNA sequencing to analyze approximately 3.2 Mb of coding tumor DNA comprising 1,811 exons from 470 genes. In total, we identified only three distinct somatic mutations, comprising two missense changes and one 14-bp deletion, each in a different gene. The accumulation of approximately one nonsynonymous somatic change per Mb of tumor DNA is consistent with a rate of mutation in tumor cells that is similar to that of normal cells. These data suggest that most sporadic colorectal cancers do not display a mutator phenotype at the nucleotide level. They also have significant implications for the interpretation of somatic mutations in candidate tumor-suppressor genes.

Webb, E., et al. (2009). "A genome-wide scan of 10 000 gene-centric variants and colorectal cancer risk." Eur J Hum Genet **17**(11): 1507-1514.

Genome scans based on gene-centric single nucleotide polymorphisms (SNPs) have been proposed as an efficient approach to identify disease-causing variants that is complementary to scans based on tagging SNPs. Adopting this approach to identify low-penetrance susceptibility alleles for colorectal cancer (CRC) we analysed genotype data from 9109 gene-centric SNPs, 7014 of which were non-synonymous (nsSNPs), in 2873 cases and 2871 controls using Illumina iselect arrays. Overall the distribution of associations was not significantly different from the null. No SNP achieved globally significant association after correction for multiple testing (lowest P value 1.7 x 10(-4), rs727299). We then analysed the dataset incorporating information on the functional consequences of nsSNPs. We used results from the in silico algorithm PolyPhen as prior information to weight the association statistics, with weights estimated from the observed test statistics within predefined groups of SNPs. Incorporating this information did not, however, yield any further evidence of a specific association (lowest P value 2.2 x 10(-4), rs1133950). There was a strong relationship between effect size and SNPs predicted to be damaging (P=1.63 x 10(-5)), however, these variants which are most likely to impact on risk are rare (MAF<5%). Hence although the rationale for searching for low-penetrance cancer susceptibly alleles by conducting genome-wide scans of coding changes is strong, in practice it is likely that natural selection has rendered such alleles to be too rare to be detected by association studies of the size employed.

Weber, T. K., et al. (1999). "Genome-wide allelotyping indicates increased loss of heterozygosity on 9p and 14q in early age of onset colorectal cancer." Cytogenet Cell Genet **86**(2): 142-147.

Colorectal cancer remains a significant public health challenge, despite our increased understanding of the genetic mechanisms involved in the initiation

and progression of this disorder. It has become clear that multiple mechanisms lead to the tumorigenic phenotype, with familial predisposition syndromes accounting for less than 15% of all colorectal cancers. A genome-wide scan for loss of heterozygosity (LOH) was carried out with 150 highly polymorphic markers in an effort to identify additional loci involved in colorectal tumorigenesis in DNA samples from 42 colorectal cancer patients. The results confirm earlier observations that tumor DNAs from patients with hereditary nonpolyposis colon cancer (HNPCC) either maintain heterozygosity or exhibit altered or additional alleles. DNAs from patients with early onset colorectal carcinomas (diagnosed prior to age 50) revealed a higher overall degree of LOH than DNAs from patients with sporadic colorectal cancers diagnosed later in life (after age 50). While regions on 1p, 10q and 14q are suggestive, statistical analysis of LOH at these regions failed to reach significance. However, LOH at 9p did reveal a statistically significant increase in the early onset patient group, compared to the greater than age 50 group. LOH on 9p may involve inactivation of p16/CDKN2 through aberrant DNA methylation on the remaining chromosome, resulting in a situation analogous to a homozygous deletion of p16 and providing a selective growth advantage to these cells. This marker may prove to be a useful prognostic indicator for patient stratification in the design of therapy for early onset colorectal cancer patients.

Whiffin, N., et al. (2014). "Identification of susceptibility loci for colorectal cancer in a genome-wide meta-analysis." Hum Mol Genet **23**(17): 4729-4737.

To identify common variants influencing colorectal cancer (CRC) risk, we performed a meta-analysis of five genome-wide association studies, comprising 5626 cases and 7817 controls of European descent. We conducted replication of top ranked single nucleotide polymorphisms (SNPs) in additional series totalling 14 037 cases and 15 937 controls, identifying a new CRC risk locus at 10q24.2 [rs1035209; odds ratio (OR) = 1.13, P = 4.54 x 10(-11)]. We also performed meta-analysis of our studies, with previously published data, of several recently purported CRC risk loci. We failed to find convincing evidence for a previously reported genome-wide association at rs11903757 (2q32.3). Of the three additional loci for which evidence of an association in Europeans has been previously described we failed to show an association between rs59336 (12q24.21) and CRC risk. However, for the other two SNPs, our analyses demonstrated new, formally significant associations with CRC. These are rs3217810 intronic in CCND2 (12p13.32; OR = 1.19, P = 2.16 x 10(-10)) and rs10911251 near LAMC1 (1q25.3; OR = 1.09, P = 1.75 x 10(-8)). Additionally, we found some evidence to support a relationship between, rs647161, rs2423297 and rs10774214 and CRC risk originally identified in East Asians in our European datasets. Our findings provide further insights into the genetic and biological basis of inherited genetic susceptibility to CRC.

The above contents are the collected information from Internet and public resources to offer to the people for the convenient reading and information disseminating and sharing.

References

1. Ali Hassan, N. Z., et al. (2014). "Integrated analysis of copy number variation and genome-wide expression profiling in colorectal cancer tissues." PLoS One 9(4): e92553.
2. Anderson, G. R., et al. (2001). "Intrachromosomal genomic instability in human sporadic colorectal cancer measured by genome-wide allelotyping and inter-(simple sequence repeat) PCR." Cancer Res 61(22): 8274-8283.
3. Baidu. http://www.baidu.com. 2018.
4. Bajenova, O., et al. (2016). "The Genome-Wide Analysis of Carcinoembryonic Antigen Signaling by Colorectal Cancer Cells Using RNA Sequencing." PLoS One 11(9): e0161256.
5. Berg, M., et al. (2010). "Distinct high resolution genome profiles of early onset and late onset colorectal cancer integrated with gene expression data identify candidate susceptibility loci." Mol Cancer 9: 100.
6. Bien, S. A., et al. (2017). "Enrichment of colorectal cancer associations in functional regions: Insight for using epigenomics data in the analysis of whole genome sequence-imputed GWAS data." PLoS One 12(11): e0186518.
7. Broderick, P., et al. (2007). "A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk." Nat Genet 39(11): 1315-1317.
8. Crujeiras, A. B., et al. (2018). "Identification of an episignature of human colorectal cancer associated with obesity by genome-wide DNA methylation analysis." Int J Obes (Lond).
9. Dehghanian, F., et al. (2018). "Reconstruction of the genome-scale co-expression network for the Hippo signaling pathway in colorectal cancer." Comput Biol Med 99: 76-84.
10. Diep, C. B., et al. (2004). "Genome characteristics of primary carcinomas, local recurrences, carcinomatoses, and liver metastases from colorectal cancer patients." Mol Cancer 3: 6.

11. Djureinovic, T., et al. (2006). "A genome wide linkage analysis in Swedish families with hereditary non-familial adenomatous polyposis/non-hereditary non-polyposis colorectal cancer." Gut 55(3): 362-366.
12. Du, M., et al. (2014). "No evidence of gene-calcium interactions from genome-wide analysis of colorectal cancer risk." Cancer Epidemiol Biomarkers Prev 23(12): 2971-2976.
13. Dumenil, T. D., et al. (2014). "Genome-wide DNA methylation analysis of formalin-fixed paraffin embedded colorectal cancer tissue." Genes Chromosomes Cancer 53(7): 537-548.
14. Enroth, S., et al. (2011). "Cancer associated epigenetic transitions identified by genome-wide histone methylation binding profiles in human colorectal cancer samples and paired normal mucosa." BMC Cancer 11: 450.
15. Eskiocak, U., et al. (2011). "Functional parsing of driver mutations in the colorectal cancer genome reveals numerous suppressors of anchorage-independent growth." Cancer Res 71(13): 4359-4365.
16. Fang, W. J., et al. (2012). "Genome-wide analysis of aberrant DNA methylation for identification of potential biomarkers in colorectal cancer patients." Asian Pac J Cancer Prev 13(5): 1917-1921.
17. Fehringer, G., et al. (2016). "Cross-Cancer Genome-Wide Analysis of Lung, Ovary, Breast, Prostate, and Colorectal Cancer Reveals Novel Pleiotropic Associations." Cancer Res 76(17): 5103-5114.
18. Fernandez-Rozadilla, C., et al. (2013). "A colorectal cancer genome-wide association study in a Spanish cohort identifies two variants associated with colorectal cancer risk at 1p33 and 8p12." BMC Genomics 14: 55.
19. Fernandez-Rozadilla, C., et al. (2013). "Pharmacogenomics in colorectal cancer: a genome-wide association study to predict toxicity after 5-fluorouracil or FOLFOX administration." Pharmacogenomics J 13(3): 209-217.
20. Fernandez-Rozadilla, C., et al. (2014). "A genome-wide association study on copy-number variation identifies a 11q11 loss as a candidate susceptibility variant for colorectal cancer." Hum Genet 133(5): 525-534.
21. Figueiredo, J. C., et al. (2011). "Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: results from a genome-wide association study." Cancer Epidemiol Biomarkers Prev 20(5): 758-766.
22. Figueiredo, J. C., et al. (2014). "Genome-wide diet-gene interaction analyses for risk of colorectal cancer." PLoS Genet 10(4): e1004228.
23. Furuta, K., et al. (2012). "Integrated analysis of whole genome exon array and array-comparative genomic hybridization in gastric and colorectal cancer cells." Cancer Sci 103(2): 221-227.
24. Gaiser, T., et al. (2011). "Genome and transcriptome profiles of CD133-positive colorectal cancer cells." Am J Pathol 178(4): 1478-1488.
25. Gao, B., et al. (2016). "Weighted gene co-expression network analysis of colorectal cancer liver metastasis genome sequencing data and screening of anti-metastasis drugs." Int J Oncol 49(3): 1108-1118.
26. Gao, B., et al. (2017). "[Corrigendum] Weighted gene co-expression network analysis of colorectal cancer liver metastasis genome sequencing data and screening of anti-metastasis drugs." Int J Oncol 50(1): 339.
27. Ghanadi, K., et al. (2016). "Colorectal cancer and the KIR genes in the human genome: A meta-analysis." Genom Data 10: 118-126.
28. Giraldez, M. D., et al. (2013). "Circulating microRNAs as biomarkers of colorectal cancer: results from a genome-wide profiling and validation study." Clin Gastroenterol Hepatol 11(6): 681-688 e683.
29. Gong, J., et al. (2016). "Genome-Wide Interaction Analyses between Genetic Variants and Alcohol Consumption and Smoking for Risk of Colorectal Cancer." PLoS Genet 12(10): e1006296.
30. Google. http://www.google.com. 2018.
31. Groothuis-Oudshoorn, C. G., et al. (2014). "Public stated preferences and predicted uptake for genome-based colorectal cancer screening." BMC Med Inform Decis Mak 14: 18.
32. Gundert, M., et al. (2017). "Genome-wide DNA methylation analysis reveals a prognostic classifier for non-metastatic colorectal cancer (ProMCol classifier)." Gut.
33. Habermann, J. K., et al. (2006). "From genome to proteome in tumor profiling: molecular events in colorectal cancer genesis." Adv Exp Med Biol 587: 161-177.
34. Habermann, J. K., et al. (2008). "From the genome to the proteome--biomarkers in colorectal cancer." Langenbecks Arch Surg 393(1): 93-104.
35. Hatzis, P., et al. (2008). "Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells." Mol Cell Biol 28(8): 2732-2744.

36. Hinoue, T., et al. (2012). "Genome-scale analysis of aberrant DNA methylation in colorectal cancer." Genome Res 22(2): 271-282.
37. Hofer, P., et al. (2017). "Bayesian and frequentist analysis of an Austrian genome-wide association study of colorectal cancer and advanced adenomas." Oncotarget 8(58): 98623-98634.
38. Hosono, S., et al. (2016). "A risk prediction model for colorectal cancer using genome-wide association study-identified polymorphisms and established risk factors among Japanese: results from two independent case-control studies." Eur J Cancer Prev 25(6): 500-507.
39. Houlston, R. S., et al. (2010). "Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33." Nat Genet 42(11): 973-977.
40. Jasmine, F., et al. (2012). "A genome-wide study of cytogenetic changes in colorectal cancer using SNP microarrays: opportunities for future personalized treatment." PLoS One 7(2): e31968.
41. Kandimalla, R., et al. (2018). "Genome-wide Discovery and Identification of a Novel miRNA Signature for Recurrence Prediction in Stage II and III Colorectal Cancer." Clin Cancer Res 24(16): 3867-3877.
42. Kemp, Z., et al. (2006). "Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan." Hum Mol Genet 15(19): 2903-2910.
43. Khamas, A., et al. (2012). "Genome-wide screening for methylation-silenced genes in colorectal cancer." Int J Oncol 41(2): 490-496.
44. Ki, D. H., et al. (2007). "Whole genome analysis for liver metastasis gene signatures in colorectal cancer." Int J Cancer 121(9): 2005-2012.
45. Kim, M. Y., et al. (2006). "Recurrent genomic alterations with impact on survival in colorectal cancer identified by genome-wide array comparative genomic hybridization." Gastroenterology 131(6): 1913-1924.
46. Kundu, S., et al. (2018). "Linking FOXO3, NCOA3, and TCF7L2 to Ras pathway phenotypes through a genome-wide forward genetic screen in human colorectal cancer cells." Genome Med 10(1): 2.
47. Lagerstedt, K. K., et al. (2007). "Tumor genome wide DNA alterations assessed by array CGH in patients with poor and excellent survival following operation for colorectal cancer." Cancer Inform 3: 341-355.
48. Lange, C. P., et al. (2012). "Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer." PLoS One 7(11): e50266.
49. Lascorz, J., et al. (2010). "Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility." Carcinogenesis 31(9): 1612-1619.
50. Le Marchand, L. (2009). "Genome-wide association studies and colorectal cancer." Surg Oncol Clin N Am 18(4): 663-668.
51. Lemire, M., et al. (2015). "A genome-wide association study for colorectal cancer identifies a risk locus in 14q23.1." Hum Genet 134(11-12): 1249-1262.
52. Ma H, Chen G. Stem cell. The Journal of American Science 2005;1(2):90-92.
53. Ma H, Cherng S. Eternal Life and Stem Cell. Nature and Science. 2007;5(1):81-96.
54. Ma H, Cherng S. Nature of Life. Life Science Journal 2005;2(1):7-15.
55. Ma H, Yang Y. Turritopsis nutricula. Nature and Science 2010;8(2):15-20. http://www.sciencepub.net/nature/ns0802/03_1279_hongbao_turritopsis_ns0802_15_20.pdf.
56. Ma H. The Nature of Time and Space. Nature and science 2003;1(1):1-11. Nature and science 2007;5(1):81-96.
57. Madhavan, S., et al. (2013). "Genome-wide multi-omics profiling of colorectal cancer identifies immune determinants strongly associated with relapse." Front Genet 4: 236.
58. Marsland Press. http://www.sciencepub.net. 2018.
59. McInnes, T., et al. (2017). "Genome-wide methylation analysis identifies a core set of hypermethylated genes in CIMP-H colorectal cancer." BMC Cancer 17(1): 228.
60. National Center for Biotechnology Information, U.S. National Library of Medicine. http://www.ncbi.nlm.nih.gov/pubmed. 2018.
61. Naumov, V. A., et al. (2013). "Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips." Epigenetics 8(9): 921-934.
62. Neklason, D. W., et al. (2008). "Common familial colorectal cancer linked to chromosome 7q31: a genome-wide analysis." Cancer Res 68(21): 8993-8997.
63. Oh, T., et al. (2013). "Genome-wide identification and validation of a novel methylation biomarker, SDC2, for blood-based detection of colorectal cancer." J Mol Diagn 15(4): 498-507.
64. Pande, M., et al. (2018). "Genetic susceptibility markers for a breast-colorectal cancer phenotype: Exploratory results from genome-wide association studies." PLoS One 13(4): e0196245.

65. Pander, J., et al. (2015). "Genome Wide Association Study for Predictors of Progression Free Survival in Patients on Capecitabine, Oxaliplatin, Bevacizumab and Cetuximab in First-Line Therapy of Metastatic Colorectal Cancer." PLoS One 10(7): e0131091.
66. Peng, Z., et al. (2003). "Genome-wide search for loss of heterozygosity in Chinese patients with sporadic colorectal cancer." Int J Gastrointest Cancer 34(1): 39-48.
67. Penney, M. E., et al. (2018). "Associations of single nucleotide polymorphisms with mucinous colorectal cancer: genome-wide common variant and gene-based rare variant analyses." Biomark Res 6: 17.
68. Peters, U., et al. (2012). "Meta-analysis of new genome-wide association studies of colorectal cancer risk." Hum Genet 131(2): 217-234.
69. Phipps, A. I., et al. (2016). "Common genetic variation and survival after colorectal cancer diagnosis: a genome-wide analysis." Carcinogenesis 37(1): 87-95.
70. Picelli, S., et al. (2008). "Genome-wide linkage scan for colorectal cancer susceptibility genes supports linkage to chromosome 3q." BMC Cancer 8: 87.
71. Prasad, R. M., et al. (2018). "A genome-wide analysis of colorectal cancer in a child with Noonan syndrome." Pediatr Blood Cancer 65(11): e27362.
72. Quan, B., et al. (2015). "Pathway analysis of genome-wide association study and transcriptome data highlights new biological pathways in colorectal cancer." Mol Genet Genomics 290(2): 603-610.
73. Real, L. M., et al. (2014). "A colorectal cancer susceptibility new variant at 4q26 in the Spanish population identified by genome-wide association analysis." PLoS One 9(6): e101178.
74. Risques, R. A., et al. (2003). "Genetic pathways and genome-wide determinants of clinical outcome in colorectal cancer." Cancer Res 63(21): 7206-7214.
75. Schumacher, F. R., et al. (2015). "Genome-wide association study of colorectal cancer identifies six new susceptibility loci." Nat Commun 6: 7138.
76. Shaik, A. P., et al. (2015). "Colorectal cancer: A review of the genome-wide association studies in the kingdom of Saudi Arabia." Saudi J Gastroenterol 21(3): 123-128.
77. Shanmugam, V., et al. (2014). "Whole genome sequencing reveals potential targets for therapy in patients with refractory KRAS mutated metastatic colorectal cancer." BMC Med Genomics 7: 36.
78. Sharma, M. R., et al. (2017). "Exceptional Chemotherapy Response in Metastatic Colorectal Cancer Associated With Hyper-Indel-Hypermutated Cancer Genome and Comutation of POLD1 and MLH1." JCO Precis Oncol 2017.
79. Shia, J., et al. (2017). "Morphological characterization of colorectal cancers in The Cancer Genome Atlas reveals distinct morphology-molecular associations: clinical and biological implications." Mod Pathol 30(4): 599-609.
80. Shiao, S. P. K., et al. (2018). "Genome wide DNA differential methylation regions in colorectal cancer patients in relation to blood related family members, obese and non-obese controls - a preliminary report." Oncotarget 9(39): 25557-25571.
81. Siegert, S., et al. (2013). "Genome-wide investigation of gene-environment interactions in colorectal cancer." Hum Genet 132(2): 219-231.
82. Sipos, F., et al. (2013). "Genome-wide screening for understanding the role of DNA methylation in colorectal cancer." Epigenomics 5(5): 569-581.
83. Skowronski, K., et al. (2014). "Genome-wide analysis in human colorectal cancer cells reveals ischemia-mediated expression of motility genes via DNA hypomethylation." PLoS One 9(7): e103243.
84. Sokolova, V., et al. (2016). "Integration of genome scale data for identifying new players in colorectal cancer." World J Gastroenterol 22(2): 534-545.
85. Staub, E., et al. (2006). "A genome-wide map of aberrantly expressed chromosomal islands in colorectal cancer." Mol Cancer 5: 37.
86. Study, C., et al. (2008). "Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer." Nat Genet 40(12): 1426-1435.
87. Suzuki, H., et al. (2011). "Genome-wide profiling of chromatin signatures reveals epigenetic regulation of MicroRNA genes in colorectal cancer." Cancer Res 71(17): 5646-5658.
88. Takahashi, Y., et al. (2017). "Japanese genome-wide association study identifies a significant colorectal cancer susceptibility locus at chromosome 10p14." Cancer Sci 108(11): 2239-2247.
89. Tan, C., et al. (2016). "Risk of eighteen genome-wide association study-identified genetic variants for colorectal cancer and colorectal adenoma in Han Chinese." Oncotarget 7(47): 77651-77663.
90. Tanami, H., et al. (2005). "Involvement of cyclin D3 in liver metastasis of colorectal cancer, revealed by genome-wide copy-number analysis." Lab Invest 85(9): 1118-1129.

91. Tanskanen, T., et al. (2018). "Genome-wide association study and meta-analysis in Northern European populations replicate multiple colorectal cancer risk loci." Int J Cancer 142(3): 540-546.
92. Tenesa, A. and M. G. Dunlop (2009). "New insights into the aetiology of colorectal cancer from genome-wide association studies." Nat Rev Genet 10(6): 353-358.
93. Tenesa, A., et al. (2008). "Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21." Nat Genet 40(5): 631-637.
94. Teo, A. S., et al. (2015). "Single-molecule optical genome mapping of a human HapMap and a colorectal cancer cell line." Gigascience 4: 65.
95. Thean, L. F., et al. (2010). "Genome-wide scan identifies a copy number variable region at 3q26 that regulates PPM1L in APC mutation-negative familial colorectal cancer patients." Genes Chromosomes Cancer 49(2): 99-106.
96. Thean, L. F., et al. (2018). "Genome-wide association study identified copy number variants associated with sporadic colorectal cancer risk." J Med Genet 55(3): 181-188.
97. Theodoratou, E., et al. (2018). "Genome-wide scan of the effect of common nsSNPs on colorectal cancer survival outcome." Br J Cancer.
98. Thorsen, K., et al. (2011). "Tumor-specific usage of alternative transcription start sites in colorectal cancer identified by genome-wide exon array analysis." BMC Genomics 12: 505.
99. Tomlinson, I. P., et al. (2008). "A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3." Nat Genet 40(5): 623-630.
100. Tomlinson, I., et al. (2007). "A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21." Nat Genet 39(8): 984-988.
101. Uddin, S., et al. (2011). "Genome-wide expression analysis of Middle Eastern colorectal cancer reveals FOXM1 as a novel target for cancer therapy." Am J Pathol 178(2): 537-547.
102. Vishnubalaji, R., et al. (2015). "Genome-wide mRNA and miRNA expression profiling reveal multiple regulatory networks in colorectal cancer." Cell Death Dis 6: e1614.
103. Vogtmann, E., et al. (2016). "Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing." PLoS One 11(5): e0155362.
104. Vychytilova-Faltejskova, P., et al. (2016). "Genome-wide microRNA Expression Profiling in Primary Tumors and Matched Liver Metastasis of Patients with Colorectal Cancer." Cancer Genomics Proteomics 13(4): 311-316.
105. Wang, H. M., et al. (2013). "A new method for post Genome-Wide Association Study (GWAS) analysis of colorectal cancer in Taiwan." Gene 518(1): 107-113.
106. Wang, H., et al. (2013). "Fine-mapping of genome-wide association study-identified risk loci for colorectal cancer in African Americans." Hum Mol Genet 22(24): 5048-5055.
107. Wang, H., et al. (2014). "Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A." Nat Commun 5: 4613.
108. Wang, T. L., et al. (2002). "Prevalence of somatic alterations in the colorectal cancer cell genome." Proc Natl Acad Sci U S A 99(5): 3076-3080.
109. Webb, E., et al. (2009). "A genome-wide scan of 10 000 gene-centric variants and colorectal cancer risk." Eur J Hum Genet 17(11): 1507-1514.
110. Weber, T. K., et al. (1999). "Genome-wide allelotyping indicates increased loss of heterozygosity on 9p and 14q in early age of onset colorectal cancer." Cytogenet Cell Genet 86(2): 142-147.
111. Whiffin, N., et al. (2014). "Identification of susceptibility loci for colorectal cancer in a genome-wide meta-analysis." Hum Mol Genet 23(17): 4729-4737.
112. Wikipedia. The free encyclopedia. Cancer. https://en.wikipedia.org/wiki/Cancer. 2018.
113. Wikipedia. The free encyclopedia. http://en.wikipedia.org. 2018.
114. Wikipedia. The free encyclopedia. Stem cell. https://en.wikipedia.org/wiki/Stem_cell. 2018.