Final Paper

Genetics and Cancer Prevention: Focus on Colorectal Cancer

BMI258: Amikar Sehdev

Introduction

Colorectal cancer (CRC) has a high worldwide incidence and mortality. It is the second most commonly diagnosed cancer in females and third most commonly diagnosed cancer in males. Over 1.2 million incident cases and 608,700 deaths were recorded globally in 2008[1]. Recent census data have shown that incidence and mortality of CRC are decreasing in the United States. This is particularly due to CRC screening and detection of pre-neoplastic lesions[2,3]. The current knowledge about the risk factors of cancer is good to prevent at least 50% of cancer[4]. The modifiable risk factors for CRC have been well studied in epidemiological, preclinical and clinical studies and include physical inactivity, smoking, red and processed meat consumption, obesity, and excessive alcohol consumption[5,6]. The improvement in our understanding of the molecular mechanisms causing cancer makes it possible to use genetic testing to predict one's risk of cancer and implement strategies to overcome that risk. Genetic testing has already proven useful to identify such high-risk populations for hereditary cancers and is used routinely in clinic. It will save many lives if we are able to identify high-risk populations for non-hereditary cancers and target interventions to lower their risk.

We know that cancer cells undergo decades of genetic change acquiring multiple driver mutations (mutations that provide selective growth advantage)[7]. This long "incubation period" before the full-blown (and often hard to treat) cancer gives an excellent opportunity to identify people with high-risk mutations and intervene at the premalignant stage. Colorectal cancer serves as the perfect model of stepwise progression from normal epithelial cell growing into a premalignant adenoma and finally carcinoma acquiring mutations at each step (discussed in detail later in this review)[7]. Therefore, we will use the example of sporadic colorectal cancer in this

review to describe, i) molecular mechanisms driving colorectal cancer, ii) knowledge gained from genome wide association studies, iii) current landscape of colorectal cancer prevention, iv) application of genetic knowledge for cancer prevention, v) challenges in cancer prevention, and vi) conclusions.

I) Molecular Mechanisms Driving Sporadic Colorectal Cancer

The development of colorectal cancer is the result of loss of genomic stability. The mechanisms described to explain this loss of stability include chromosomal instability, microsatellite instability (MSI), and CpG island methylator phenotype (CIMP)[8]. About 80-85% of sporadic colorectal cancers are due to chromosomal instability resulting in either the loss of a wild-type copy of a tumor suppressor gene or a mutation in a gatekeeper gene (causing inactivation of the same), or another gene that regulates cell proliferation or cell survival. Microsatellite instability is the second most common form of genomic instability[8]. Although MSI is commonly known to cause Lynch syndrome (a hereditary form of colorectal cancer), it can also account for 15% cases of sporadic CRC[8]. The MSI arises through the inactivation of four mismatch repair genes (*MLH1*, *MSH2*, *PMS2*, and *MSH6*). The mismatch repair machinery works during DNA replication and recombination to sense, excise, and replace the mismatched bases[9,10]. Besides the inactivation, methylation of the *MLH1* gene alone can also lead to MSI and account for 15% of sporadic CRC cases[11,12]. Lastly, CIMP is a mechanism that leads to epigenetic silencing by aberrant DNA methylation of gene promoter region. The CIMP is responsible for about 15% of colon tumorigenesis, however it should be noted that CIMP and MSI pathways overlap. For instance, lack of *MLH1* function can result from its loss (MSI) or methylation of the *MLH1* promoter (CIMP) leading to the same phenotype[8].

Research over the last three decades has characterized the molecular details of the genetic alterations seen driving colorectal cancer initiation and progression from normal epithelium to cancer (Figure 1). This process requires at the minimum two to eight driver mutations in specific

genes leading to the formation of a malignant tumor[7]. The key pathway identified as the primary driver for CRC development is Wnt signaling. The activation of Wnt signaling marks the initiating event of colorectal cancer development[8]. The most common defect leading to activation of Wnt signaling is inactivation of *APC* gene. *APC* gene is the gatekeeper of Wnt signaling through regulation of β–catenin stability. About 85% of all sporadic cases have this defect[8,13]. Interestingly, inactivating germline mutations of *APC* gene can also lead to a hereditary form of colorectal cancer, familial adenomatous polyposis (FAP). CRC arising due to germline or sporadic mutations of APC has different ages of onset. Germline *APC* gene mutation usually causes tumor onset before the age of 40 years whereas sporadic *APC* gene mutation leads to tumor over the age of 49 years[14].

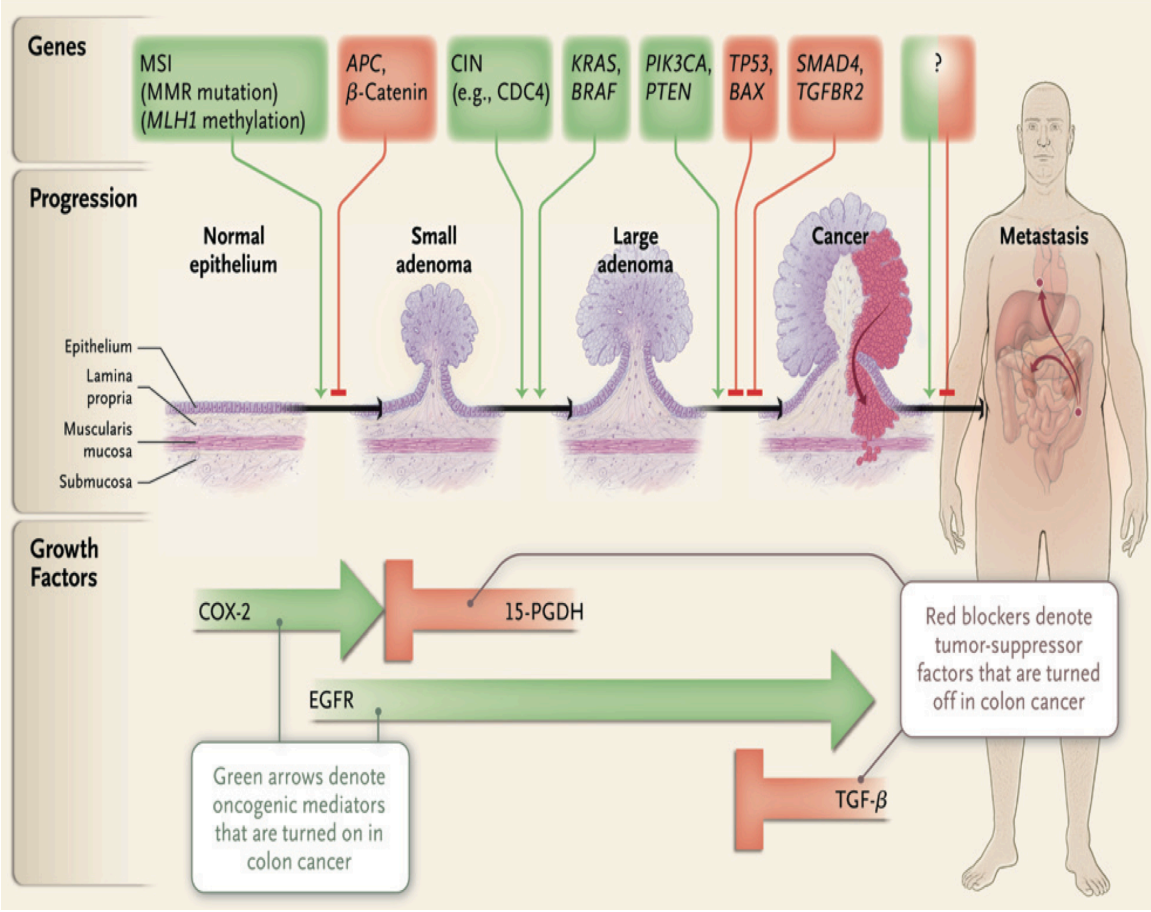

Figure 1: Shown above are genetic alterations at different stages of colorectal cancer progression[8].

The key pathways dysregulated in the transition from adenoma to carcinoma are primarily due to successive somatic mutations in p53, TGFβ, and the mitogen-activated protein kinase (MAPK) signaling pathways (Figure 1). The *TP53* gene is primarily inactivated in 35-55% of colorectal tumors by the loss of both alleles resulting from a combination of chromosomal deletion of 17p and missense mutation. The p53 regulates cell cycle arrest and cell death under cellular stress. The loss of p53 gives the cell a growth advantage by loss of both control of cell survival and genomic stability[8]. The TGFβ pathway plays an important role by controlling apoptosis and growth arrest and is inactivated in about 50% of all colorectal tumors. Mutations in either *TGFBR2* or one of the downstream signaling molecules – *SMAD4*, *SMAD2*, or *SMAD3* can result in the inactivation of this pathway. MSI is associated with mutations in *TGFBR2*[8]. Activation of the classical MAPK signaling pathway is also seen in about 50% of colorectal cancer[8]. The MAPK pathway regulates cell growth, cell differentiation, and apoptosis, thereby controlling tumor progression process[15]. Activating mutations in *KRAS* gene accounts for inactivation of MAPK pathway in 30-40% of all colorectal cancers[8,16]. Additionally, activating mutations in *BRAF* (8-12%) and *PIK3CA* (33%) are also known to inactivate MAPK pathway[8]. Therefore, mutations affecting p53, TGFβ, and the mitogen-activated protein kinase (MAPK) signaling pathways either directly or indirectly leads to the development and progression of colorectal cancer mainly by the loss of control of cell survival, growth, apoptosis and differentiation.

II) <u>Genome Wide Association Studies</u>

There were no common genetic variants known for colorectal cancer before 2007. The advances in cost effective, high throughput genotyping technologies and HapMap project (international, freely available, haplotype map of human genome) facilitated the GWAS studies. This lead to undertaking of genotyping tens of thousands of CRC patients resulting in reporting of many large GWAS studies in 2007-2008. Three most significant GWAS studies carried out of

England[17], Scotland[18,19], and Canada[19] have identified six associations which were recently reported in a series of publications[17-21]. However, the single nucleotide polymorphisms (SNP) associations found in these studies had modest effect (odds ratio ~ 1.2) in terms of risk of colorectal cancer. International collaborative efforts are required to conduct very large studies (>15,000 samples per locus identified) to test these associations with sufficient power to detect their true effect. Nonetheless, a meta-analysis of these GWA studies analyzing for 550,000 SNPs in 1,632 cases and 1,977 controls found four additional SNPs again with modest effect size[22]. All together, these ten loci (Table 1) can explain full-sibling relative risk of about 6% which corresponds to a phenotypic variance ~0.04 and ~1.26% on the liability and observed scale respectively[23].

Table 1: SNP loci identified in GWAS studies to be associated with increased risk of colorectal cancer[23].

Gene (or locus)	Chr.	SNP	Study Population	Sample size (cases/controls)		Effect size OR (95% CI)	Allele Frequency	PAR (%)
				GWAS	Total			
POU5F1P1, DQ515897, MYC	8	rs6983267	England	940/965	8,264/ 6,206	1.21 (1.15 - 1.27)	0.51	9.7
POU5F1P1, DQ515897, MYC		rs10505477	Canada	1,226/ 1,239	7,480/ 7,779	1.17 (1.12 -1.23)	0.50	7.8
POU5F1P1, DQ515897, MYC		rs7014346	Scotland	1,012/ 1,012	14,500/ 13,294	1.19 (1.14 -1.24)	0.37	6.6
SCG5, GREM1, FMN1	15	rs4779584	England	718/960	7,922/ 6,741	1.26 (1.19 – 1.34)	0.18	4.5
SMAD7 (intron 3)	18	rs4939827	England	940/965	8,413/ 6,949	1.18 (1.12-1.23)	0.52	8.6
SMAD7 (intron 3)		rs4939827	Scotland	1,012/ 1,012	14,500/ 13,294	1.20 (1.16 – 1.24)	0.51	9.2
LOC120376, FLJ45803, c11orf53, POU2AF1	11	rs3802842	Scotland	1,012/ 1,012	14,500/ 13,294	1.12 (1.07 – 1.17)	0.29	3.4
c8orf53, EIF3H	8	rs16892766	England	940/965	18,831/ 18,540	1.25 (1.19 – 1.32)	0.07	1.7
FLJ3802842	10	rs10795668	England	940/965	18,831/ 18,540	1.12 (1.10-1.16)	0.67	7.4
BMP4	14	rs4444235	United Kingdom	1,952/ 1,977	20,288/ 20,971	1.11 (1.08 – 1.15)	0.46	4.8
CDH1	16	rs9929218	United Kingdom	1,952/ 1,977	20,288/ 20,971	1.10 (1.06 – 1.12)	0.71	6.6
RHPN2	19	rs10411210	United Kingdom	1,952/ 1,977	20,288/ 20,971	1.15 (1.10- 1.20)	0.90	11.9
BMP2	20	rs961253	United Kingdom	1,952/ 1,977	20,288/ 20,971	1.12 (1.08 – 1.16)	0.35	4.0

The single nucleotide polymorphisms (SNP) identified in these GWAS studies did not establish any causal association and found a modest risk increased risk of CRC. However, they do add to the understanding of cancer biology and opens new venue of gene loci and protein that can be targeted for therapy in future. Importantly, five of the ten SNPs identified are linked to the TGFβ pathway of signaling explained above[23]. This suggests for the first time that perturbations of TGFβ signaling pathway may increase the colorectal cancer susceptibility. Further collaborative efforts are needed to overcome the limitations of current GWAS studies and to identify new loci with even smaller effects.

III) <u>Current Landscape of Colorectal Cancer Prevention</u>

Much of the preventive efforts in the past two decades have focused on targeting the modifiable risk factors, for example dietary interventions (low fat diet, high vitamin D diet, folate supplementation, etc.). However the evidence in support of such interventions is limited[4]. The prospective prevention trials are hard to conduct due to long follow up period and need for multicenter support. At present the preventive interventions that have shown to be of benefit and are in routine clinical use for colorectal cancer prevention include aspirin therapy and sigmoidoscopy/colonoscopy[4]. Aspirin use reduces the incidence and mortality of colorectal cancer by 20-40% however the benefit becomes apparent only after at least 3-5 years of use[24,25]. Sigmoidoscopy and colonoscopy has also shown to reduce the incidence of colorectal cancer in epidemiological[26] and a recent randomized controlled trial. A single screening flexible sigmoidoscopy between the age of 55 and 64 years was associated with 33% reduction in incidence and 43% reduction in mortality of colorectal cancer in this study[26]. At present, there are no clinically used tests to prevent colorectal cancer despite clear evidence of years of identifiable genetic changes before the development of cancer. Genetic tests are emerging for secondary prevention by assessing the risk of recurrence as well as for personalized therapy of chemoprevention drugs as discussed in the next section.

IV) Applications of Genetic Knowledge for Cancer Prevention

As described above we don't have many therapeutic options for colorectal prevention. In fact, besides screening modalities as described above, aspirin is the only proven therapeutic agents available for prevention of colorectal cancer[25,27,28]. Furthermore, there are no genetic tests that are available currently to predict risk of sporadic colorectal cancer. In this section, we will discuss role of genetic testing first by describing the two most common gene expression signature tests available for predicting the prognosis of CRC and then give an example of how genetic testing might be very useful for prescribing aspirin for chemoprevention.

i) Oncotype DX: The Oncotype DX Colon Cancer Assay (Genomic Health, Redwood City, CA) a tool for predicting the recurrence risk of CRC based on a 12-gene assay in patients with II disease[29,30]. There is no clear consensus for chemotherapy treatment for patients with stage II CRC. In fact, the Current National Comprehensive Cancer Network (NCCN) guidelines recommend chemotherapy only for high-risk stage II patients, where risk is determined by clinical presentation and pathologic parameters (MSI status). Therefore, additional tools which can help decide quantify the risk of a patient can be very helpful. The Oncotype DX Colon Cancer Assay is a reverse transcriptase-polymerase chain reaction (RT-PCR) assay in which standardized multigene expression analysis is done on formalin-fixed paraffin-embedded (FFPE) primary colon tumor tissue[29]. The recurrence score is calculated using the expression of seven genes normalized against five reference genes in a set of twelve total genes.

The continuous recurrence score predicted by Oncotype DX is significantly associated ($P = 0.004$) with recurrence risk and was validated in a retrospective analysis of phase III, randomized controlled QUASAR trial[30]. The association remained significant ($P = 0.008$) in a multivariate model besides tumor stage and MSI. The recurrence risk was estimated at 10% versus 25% for low and high recurrence score respectively[29].

ii) ColoPrint: The ColoPrint Assay (Genomic Health) is another available gene expression assay that predicts the risk of recurrence of stage II colorectal cancer[31]. Quantitative expression of 18 genes on FFPE tissue is used in this assay to estimate the risk of recurrence. The 5-year distant metastasis free survival was estimated at 80% and 95% for low-risk and high-risk patients respectively[31]. The assay was shown to predict development of distant metastasis based on the recurrence risk score with a hazard ratio (HR) of 4.3 (95% CI, 1.36–13.56; $P = 0.007$). In a more recent validation study of this signature the authors reported 60% of patients as low risk and 40% as high risk[32]. Multivariate model identified recurrence score as reported by ColoPrint to be the most significant predictor of five year relapse-free survival[32].

Although encouraging, it must be acknowledged that these tests need prospective validation and are not in routine clinical use at present.

iii) Role of PI3K mutation in Aspirin: Aspirin use has been shown to improve colorectal cancer survival[27,28]. Cyclooxygenase-2 (COX-2) is a rate-limiting enzyme that catalyzes prostaglandin synthesis and is blocked by aspirin[24,33]. COX-2 overexpression in CRC tumors (as detected by IHC) correlates with survival improvement if aspirin is used after chemotherapy treatment[27]. Therefore, an important potential target for aspirin in CRC is COX-2. Mutations in PIK3CA are present in approximately 15 to 20% of sporadic CRC and enhance prostaglandin synthesis, resulting in inhibition of apoptosis in colon-cancer cells[24]. Aspirin by inhibiting this PI3K pathway through downstream COX-2 may suppress cancer-cell growth and induce apoptosis. Liao *et al.* conducted a recent retrospective analysis to see if there is a differential response to aspirin therapy based on the PIK3CA mutation status in CRC[24]. They found a higher CRC-specific survival among patients with mutated-PIK3CA tumors as compared to wild type-PIK3CA tumors with regular use of aspirin after diagnosis (multivariate HR 0.18; 95% CI, 0.06 to 0.61; $P<0.001$ by log-rank)[24].

V) Challenges in Cancer Prevention

There have been many challenges in application of genetic testing for chemoprevention. Except for detection of hereditary cancer syndromes such testing is currently only done for research purposes. This is due to many different reasons. First, the most common obstacle is lack of prospectively validated gene expression signatures that can be used for primary as well as secondary prevention. Ongoing interdisciplinary research is required to develop genetic tests that can predict the risk of cancer development for an individual. Second, such are cost prohibitive for routine use. Although, there has significant decrease in the cost of gene sequencing (please see my final paper for BMI231), the total cost including the analysis is still quite high. Third, there are ethical and legal issues pertaining to routine genetic testing as it might lead to insurance and job discrimination based on results of such testing. Although US legal system passed a legislation (Genetic Information Nondiscrimination Act) to protect against health insurance discrimination based on the results of genetic testing, this law does not protect against life and disability insurance discrimination[34]. Issues like this need to addressed before wide spread implementation of genetic testing. Fourth, lack of infrastructure to do in house testing of these tests. At present only a few large academic centers and research institutes have the capability to conduct genetic tests. Fifth, the practicing physicians are not accustomed to interpreting these results and before these tests can be routinely used knowledge gaps need to be addressed universally in the medical community. Lastly, there needs to be a standardized way of reporting and analyzing these tests and Food and Drug Administration or another such organization will have to take the responsibility to examine the evidence in support of clinical use.

VI) Conclusions

The treatment of cancer has significantly improved in the last century however it is still difficult to treat metastatic cancer and cancer control rather than cure is the only possibility. Resources focused on early identification of cancer or premalignant lesion will be immensely

useful in preventing the morbidity and mortality of cancer. With the increasing understanding of molecular pathogenesis of cancer and faster comprehensive mutational analysis techniques, genetic testing holds tremendous potential for cancer prevention. Currently the United States has 13 million cancer survivors[35] who are at a very high risk of developing a second cancer and can benefit from secondary prevention strategies. There is promise in new screening tools for both primary and secondary prevention using genetic testing. However, there are many challenges in front of us before genetic testing for prevention purposes can be routinely used that need to be overcome with interdisciplinary and inter-institutional collaboration and public health education efforts.

Acknowledgment

I would like to thank Dr. Douglas Brutlag for providing me with a great opportunity to understand the basics of translational bioinformatics. This has been a huge learning curve for me that will help me in my professional aspirations.

References

1. Jemal A, Bray F, Center MM, et al: Global cancer statistics. CA Cancer J Clin 61:69-90, 2011
2. Edwards BK, Ward E, Kohler BA, et al: Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 116:544-73, 2010
3. Center MM, Jemal A, Smith RA, et al: Worldwide variations in colorectal cancer. CA Cancer J Clin 59:366-78, 2009
4. Colditz GA, Wolin KY, Gehlert S: Applying What We Know to Accelerate Cancer Prevention. Science Translational Medicine 4:127rv4-127rv4, 2012
5. Giovannucci EL: Cancers of the colon and rectum, Cancer Epidemiology and Prevention, New York: Oxford University Press, 2006, pp 809-829
6. Ferrari P, Jenab M, Norat T, et al: Lifetime and baseline alcohol intake and risk of colon and rectal cancers in the European prospective investigation into cancer and nutrition (EPIC). Int J Cancer 121:2065-72, 2007
7. Vogelstein B, Papadopoulos N, Velculescu VE, et al: Cancer Genome Landscapes. Science 339:1546-1558, 2013
8. Markowitz SD, Bertagnolli MM: Molecular Basis of Colorectal Cancer. The New England journal of medicine 361:2449-2460, 2009
9. Jascur T, Boland CR: Structure and function of the components of the human DNA mismatch repair system. International journal of cancer, 2006
10. Iyer RR, Pluciennik A, Burdett V: DNA mismatch repair: functions and mechanisms. Chem ..., 2006
11. Issa JP: CpG island methylator phenotype in cancer. Nat Rev Cancer 4:988-93, 2004
12. Kane MF, Loda M, Gaida GM, et al: Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 57:808-11, 1997
13. Kinzler KW, Vogelstein B: Lessons from Hereditary Review Colorectal Cancer. Cell, 1996
14. Lynch HT, de la Chapelle A: Hereditary colorectal cancer. The New England journal of medicine, 2003
15. Halilovic E, Solit DB: Therapeutic strategies for inhibiting oncogenic BRAF signaling. Current opinion in pharmacology, 2008
16. Roth AD, Tejpar S, Delorenzi M, et al: Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. Journal of Clinical ..., 2010

17. Tomlinson I, Webb E, Carvajal-Carmona L, et al: A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nature Genetics 39:984-988, 2007
18. Tenesa A, Farrington SM, Prendergast JGD, et al: Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nature Genetics 40:631-637, 2008
19. Zanke BW, Greenwood CM, Rangrej J, et al: Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nature Genetics 39:989-994, 2007
20. Broderick P, Carvajal-Carmona L, Pittman AM, et al: A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nature Genetics 39:1315-1317, 2007
21. Tomlinson IP, Webb E, Carvajal-Carmona L, et al: A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nature Genetics 40:623-630, 2008
22. Houlston RS, Webb E, Broderick P, et al: Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nature Genetics 40:1426-1435, 2008
23. Tenesa A, Dunlop MG: New insights into the aetiology of colorectal cancer from genome-wide association studies. Nature Reviews Genetics 10:353-358, 2009
24. Liao X, Lochhead P, Nishihara R: Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. ... England Journal of ..., 2012
25. Rothwell PM, Wilson M, Elwin CE, et al: Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. The Lancet, 2010
26. Atkin WS, Edwards R, Kralj-Hans I, et al: Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. The Lancet, 2010
27. Chan AT, Giovannucci EL, Meyerhardt JA: Long-term use of aspirin and nonsteroidal anti-inflammatory drugs and risk of colorectal cancer. JAMA: the journal of ..., 2005
28. Chan AT, Ogino S, Fuchs CS: Aspirin use and survival after diagnosis of colorectal cancer. JAMA: the journal of the American ..., 2009
29. Webber EM, Lin JS, Evelyn PW: Oncotype DX tumor gene expression profiling in stage II colon cancer. Application: prognostic, risk prediction. PLoS Curr 2, 2010
30. Catenacci DVT, Kozloff M, Kindler HL, et al: Personalized Colon Cancer Care in 2010. Seminars in oncology 38:284-308, 2011
31. R. Rosenberg MM, U. Nitsche, et al.: Independent validation of a prognostic genomic profile (ColoPrint) for stage II colon cancer (CC) patients. J Clin Oncol. 28:15s (suppl; abstr 3513).

32. Salazar R, Roepman P, Capella G, et al: Gene Expression Signature to Improve Prognosis Prediction of Stage II and III Colorectal Cancer. Journal of Clinical Oncology 29:17-24, 2010
33. Dougherty U, Sehdev A, Cerda S, et al: Epidermal growth factor receptor controls flat dysplastic aberrant crypt foci development and colon cancer progression in the rat azoxymethane model. Clin Cancer Res 14:2253-62, 2008
34. Terry SF: Genetic information nondiscrimination act insurance protections issued. Genet Test Mol Biomarkers 13:709-10, 2009
35. Jacobs: Care of the Adult Cancer Survivor.1-16, 2013