

**COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD,
CONSUMER PRODUCTS AND THE ENVIRONMENT**

Genome Wide Association Studies (GWAS) – Bladder Cancer and Colorectal Cancer

At the COC meeting in July 2010, the Committee discussed a paper entitled “Interaction Between Genotype And Chemicals In The Environment On The Induction Of Cancer In Risk Assessment – Scoping Paper for Update Review” (CC/2010/02). This detailed the major projects that have been initiated over the past decade in the fields of both genetics and exposure assessment and their contribution to understanding of gene-environment interactions and cancer risk. There have been substantial advances in the field of human genomics over the past ten years. At the meeting, the Committee indicated that they would like to review some examples of Genome-Wide Association Studies (GWAS), used to measure genome variation on a large scale, and cancer. The HPA Toxicology Unit at Imperial has produced this paper (CC/11/07) detailing the main GWAS of bladder cancer and colorectal cancer from the available literature. Members are invited to discuss the paper and respond to the questions below.

Questions for the Committee

In general, have the Committee any comments on GWAS?

Will information such as that described here be of help in assessments by the Committee and, if so, how?

Do the findings of these GWAS shed any light on the environmental causes of colorectal or bladder cancer?

Do members have any research or other recommendations with respect to GWAS?

Secretariat

March 2011

COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT

Genome Wide Association Studies (GWAS) – Bladder Cancer and Colorectal Cancer

GWAS

1. Genome-wide association studies (GWAS), also known as whole genome association studies (WGA studies), are an examination of genetic variation across a given genome, designed to identify genetic associations with observable traits. Completion of the Human Genome Project in 2003 has made it possible to assess genetic contributions to common diseases and analyse whole genome samples for genetic variation that contribute to the onset of these diseases. GWAS are genome-wide, non-hypothesis driven in nature. GWAS use high throughput genotyping technologies to assay hundreds of thousands of single nucleotide polymorphisms (SNPs) and relate these variants to disease. GWA studies build on lessons learned from candidate gene and family linkage studies, as well as the expanding knowledge of relationships among SNP variants generated by the International Hapmap Project, to work out the majority of common genetic differences among individuals and relate them to health and disease. At the time of writing this paper (March 2011), 708 GWA studies have been published (<http://www.gwascentral.org/>) with 85 GWA studies specifically addressing cancer. Most GWA studies yield at least one highly statistically significant association but almost all of the associations are weak [small odds ratios (ORs), in the order of 1.1–1.6]. According to Vineis et al. (2008), almost all results from cancer GWA studies pointed to novel loci, previously unidentified. Vineis et al. (2008) provide a list of robust cancer GWA studies with relative risks in the order of 1.06–2.23, but mostly ~ 1.2. These included studies on breast, prostate and colorectal cancers (Easton et al., 2007; Stacey et al., Hunter et al., 2007; Yeager et al., 2007; Gudmundsson et al., 2007; Tenesa et al., 2008 and Tomlison et al., 2008).

2. A typical GWAS has 4 parts; 1) Selection of a large number of individuals with the disease and a suitable comparison group, 2) DNA isolation, genotyping and data review to ensure high quality genotyping, 3) Statistical test for associations between the SNPs and the disease, 4) Replication of identified associations in an independent population sample (Pearson and Manolio, 2008).

Colorectal Cancer and GWAS

3. Cancer of the large bowel (also known as colorectal cancer - CRC) is a common form of malignancy in developed countries but occurs much less frequently in the developing world. Around 106 new cases of colorectal cancer are diagnosed each day in the UK and it is the third most common cancer after breast and lung. In 2007 there were 38,608 new cases of large bowel cancer registered in the UK: around two-thirds (24,274) in the colon and one-third (14,334) in the rectum. In 2008 there were 16,259 deaths from colorectal cancer in the UK, comprising 10,164 from colon and 6,095 from rectal cancer. Colorectal cancer is the second most common cause of death from cancer

in the UK after lung cancer. In the ten years between 1999 and 2008, the bowel cancer age-standardised mortality rates in the UK fell by 13%. This fall in mortality affected all age groups with the largest fall in the 40-54 age group for men and the 55-69 and 70-79 age groups in women (ONS, 2010; ISD, 2010; Welsh Cancer Intelligence and Surveillance Unit, 2010 and Northern Ireland Cancer Registry, 2010).

4. There is evidence of a genetic component to the aetiology of colorectal cancer. The literature suggests an increased incidence of colorectal cancer among persons with a family history of colorectal cancer and in families in which multiple family members are affected with colorectal cancer, in a pattern indicating autosomal dominant inheritance of cancer susceptibility (Burt and Petersen, 1996; Lynch and Smyrk, 1996; Utsunomiya and Lynch, 1990; Herrera, 1990). The majority of CRC cases (up to 80%) are sporadic (Cheah, 2009), indicating that both genetic and environmental factors contribute to the disease aetiology, with inherited susceptibility being responsible for ~35% of all CRC (Lichtenstein, 2000). Genetic mutations in hereditary high-risk germ line mutations of the *APC* gene, the DNA mismatch repair genes, *MUTYH*, and more rarely in the *SMAD4*, *BMPRIA* and *STK11/LKB1* genes have been identified as the cause of inherited cancer risk in some colon cancer-prone families; these mutations are estimated to account for only 5-6 % of colorectal cancer cases.

5. In brief, but described in more detail below, a number of GWASs for CRC have been published to date in England, Scotland and Canada (Tenesa et al., 2008; Tomlinson et al., 2007; Zanke et al., 2007). These GWAS resulted in the identification of six susceptibility loci (8q23.3, 8q24.21, 10p14, 11q23, 15q13.3 and 18q210), the reporting of which can be found in a number of publications (Tenesa et al., 2008; Tomlinson et al., 2007; Zanke et al., 2007; Broderick et al., 2007 Jaeger et al., 2008, Tomlinson et al., 2008). A subsequent meta-analysis of the UK GWASs (Houlston et al., 2007) yielded four novel risk loci (14q22.2, 16q22.1, 19p13.1 and 20p12.3), including confirmation of these associations in additional large sample sets. Several follow-up studies have replicated these associations through genotyping of thousands of individuals (Berndt et al., 2008; Curtin et al., 2009; Li et al., 2008; Matsuo et al., 2009; Middeldorp et al. 2009; Pittman et al., 2008; Poynter et al. 2007; Schafmayer et al., 2009; Slattery et al., 2010).

6. Zanke et al. (2007) published the first GWAS for CRC and adopted a multistage genetic association approach to identify markers in specific chromosomal regions associated with colorectal cancer. Stage I of the GWAS involved genotyping 99,632 SNPs in 1,257 affected individuals and 1,336 controls from Ontario. Stage 2 involved testing 1,143 SNPs selected from the stage 1 SNP panel in two case-control sets, one from Seattle and one from Newfoundland. Stage 3 of the GWAS investigated 76 putative associations identified in stages 1 and 2 in an early-onset colorectal cancer case-control set from Scotland. 9 SNPs showed evidence for replication in the Scottish stage 3 sample set. In the 4th stage of the GWAS, the 9 loci identified from stage 3 were tested for association in a further independent Scottish case-control series of older onset cases. Two of the associations identified in stage 3 were replicated further in the stage 4 sample set: rs10505477 (tagged by rs6983267 in stage 3, $P = 0.03$) and rs719725 (tagged by rs7857628 or rs206636213 in stage 3, $P = 0.0025$). Stage 4 associations detected for rs10505477 gave an OR of 1.16 (c.i. 1.11–1.21; $P = 0.001$), and for rs719725, OR = 1.10 (c.i. 1.05–1.15; $P = 0.037$) in an allelic model. Combining overall genotype data for stages 1–4 gave an OR of 1.19 ($P = 6.40 \times 10^{-9}$) for the 8q24 locus

(rs10505477 or rs6983267) and 1.13 ($P = 4.98 \times 10^{-5}$) for the 9p24 locus (rs719725 and rs7857826). The paper also describes how a consortium genotyped 20 of the 76 SNPs from stages 1 and 2 (including rs10505477 from 8q24) in 2,199 affected individuals and 2,401 controls derived from the European Prospective Investigation into Cancer and Nutrition (EPIC), Nantes and French Familial Case Control studies, resulting in validation of the association of rs10505477 with colorectal cancer (log-additive OR = 1.16; $P = 5.05 \times 10^{-4}$).

7. Tomlinson et al. (2007) conducted a GWAS of 530,163 tag SNPs in 970 individuals with familial colorectal neoplasia and 960 controls using the Illumina Infinium platform. Of the 547,647 polymorphic tag SNPs, 27,673 showed an association with the disease at $p < 0.05$. The strongest associations identified were with SNPs mapping to chromosome 8q24.21 with the rs6983267 SNP exhibiting the strongest association at a global significance of $p = 1.86 \times 10^{-7}$, allelic test. Three additional CRC case control series were genotyped to validate the rs6983267 result. In the first series, comprising of 4,361 individuals with CRC and 3,752 controls, the allelic-based test for disease association provided independent confirmation of an association between the variation at 8q24.21 and the risk of CRC, with a significant association for rs6983267 ($p = 5.02 \times 10^{-8}$). In the 2nd replication series, (1901 individuals with CRC and 1079 controls), the equivalent p value for rs6983267 was 3.41×10^{-4} . In the 3rd series, (1,072 individuals with CRC and 415 controls) association was apparent but nonsignificant ($p = 0.149$). When the authors pooled the data from the individuals with CRC and controls, unequivocal evidence was provided for a relationship between rs6983267 and the risk of CRC ($p = 1.27 \times 10^{-14}$, OR = 1.21, 95% CI 1.15-1.27). As a portion of the original panel of individuals had a family history of CRC but were themselves unaffected by colorectal adenomas, 2 additional panels of cancer-free adenoma cases were examined to assess the possibility that CRC risk associated with rs6983267 is mediated through adenoma susceptibility. Pooling data from these series with the adenoma cases from the original series provided evidence that rs6983267 is associated with an elevated risk of adenoma development (OR = 1.21, 95% CI 1.10 – 1.34, $p = 6.9 \times 10^{-5}$).

8. Tenesa et al. (2008) conducted a comprehensive phase designed GWAS in early onset Scottish CRC cases and controls. Stage I of the GWAS involved genotyping 555,510 SNPs in 1,012 CRC cases and 1,012 controls using Illumina HumanHap300 and Human-Hap240s arrays. From the analysis of stage 1, 15,008 SNPs were selected for further investigation in phase 2 of the GWAS, which involved genotyping these SNPs in 2,057 cases and 2,111 controls using the Illumina iSelect platform. Stage 3 of the GWAS involved genotyping the five highest-ranked SNPs (rs7014346 (8q24), rs4939827 (18q21), rs6533603 (4q25), rs3802842 (11q23.1) and rs9951602 (18q23)) from the joint phase 1 and 2 analysis in 14,500 cases and 13,294 controls from seven populations. This study identified a previously unreported locus on chromosome 11q23, tagged by rs3802842, which is associated with CRC. Extending previous observations by other researchers (Zanke et al., 2007; Tomlinson et al., 2007) at chromosome 8q24 locus and at the SMAD7 locus (Broderick et al., 2007), Tenesa et al. (2007) further replicated and fine mapped the associations at 8q24 (rs7014346, OR = 1.19, 95% CI 1.15- 1.23, P value 8.6×10^{-26}) and 18q21 (rs4939827, OR= 1.2 , 95% CI 1.16-1.24, P value 7.8×10^{-28}), showing consistent effects across multi-ethnic populations.

9. Broderick et al. (2007) conducted a genome-wide association study to identify risk variants for colorectal cancer. Stage I of this GWAS involved genotyping 547,647

polymorphic tagging SNPs in 940 individuals with familial colorectal tumours (627 CRC and 313 advanced adenomas) and 965 controls using an Illumina Hap550 BeadChip. The second stage of the GWAS involved genotyping the top 5 % of SNPs with the greatest *P* values from the initial GWAS. The most significant *P* values in the GWAS were found at two polymorphic sites, rs6983267 mapping to 8q24.21 and rs4939827 mapping to SMAD7. This study replicates the result for rs6983267 the first common susceptibility variant to be identified in GWAS for CRC (Zanke et al. 2007; Tomlinson et al. 2007). For the newly identified loci from this study, the risk of colorectal neoplasia associated with the SMAD7 SNP rs4939827 reached a *P* value of 3.07×10^{-7} . Two additional SNPs in SMAD7 (rs12953717 and rs4464148) had the most extreme *P* values from the unadjusted analysis. In three additional independent case-control analyses, the three SNPs associated with SMAD7 were evaluated by allele-specific PCR and these analyses provided further evidence of an association between genotype and risk. Pooling genotype data for CRC cases and controls from the four panels provided unequivocal evidence for a relationship between the three SNPs and CRC. For example, the pooled *P* value across all studies for rs4939827 was 1.0×10^{-12} .

Bladder cancer and GWAS

10. Bladder cancer is a common cancer in the UK, with 10,091 new cases diagnosed in 2007 (ONS, 2010; ISD, 2010; Welsh Cancer Intelligence and Surveillance Unit, 2010 and Northern Ireland Cancer Registry, 2010). It is the most frequently occurring tumour of the urinary system and accounts for around 1 in every 29 new cases of cancer each year in the UK. In the UK, bladder cancer is the fourth most common cancer in males, with 7,284 new cases diagnosed in 2007 (ONS, 2010; ISD, 2010; Welsh Cancer Intelligence and Surveillance Unit, 2010 and Northern Ireland Cancer Registry, 2010). This compares to 2,807 female cases, giving a male:female ratio of 5:2. In females, it is the eleventh most common cancer. In 2008 there were 5,002 deaths from bladder cancer in the UK (3,273 male deaths and 1,729 female deaths). The distribution of deaths by age shows a steep climb with age with almost 9 in 10 deaths occurring after the age of 65. Generally there has been little change in the female bladder cancer mortality rates although overall rates have fallen slightly from around 3.5 per 100,000 in the late 1970s/early 1980s to 2.8 per 100,000 in 2008. In contrast, the male rates have shown a consistent fall since 1992, from 12.2 to 7.9 per 100,000 in the year 2008, a fall of 35% (ONS, 2010; ISD, 2010; Welsh Cancer Intelligence and Surveillance Unit, 2010 and Northern Ireland Cancer Registry, 2010).

11. There is evidence of a genetic component to the aetiology of bladder cancer with familial clustering (Kiltie et al., 2010). Aben et al. (2002) reported an almost 2-fold increased risk among first-degree relatives of patients with urothelial cell carcinoma (UCC). Common variants in low-penetrance genes involved in the metabolism of environmental toxins (xenobiotic biotransforming enzymes such as *N*-acetyltransferase 2 (*NAT2*) and glutathione *S*-transferase M1 (*GSTM1*)) have been shown to modify individual susceptibility to bladder carcinogens. A recent meta-analysis by (García-Closas et al., 2005) of 31 case-control studies assessing the risk of bladder cancer conferred by *NAT2*-slow acetylating genetic variants and of 28 case-control studies assessing the risk of bladder cancer conferred by *GSTM1*-null variants estimated odds ratios (OR) of 1.4 [95% confidence interval (95% CI), 1.2-1.6] and 1.5 (95% CI, 1.3-1.6), respectively.

12. Kiemeny et al. (2008) performed the first GWAS analysis, on Icelandic and Dutch urinary bladder cancer cases and controls. They used an Illumina Infinium Whole genome genotyping microarray. In phase I of the GWAS, 1803 cases and 34,336 controls were genotyped. Over 300,000 SNPs were tested for association with Urinary Bladder cancer. The authors found that this sample size had an 88% power to achieve a genome-wide significant result for a variant with an OR of 1.3 and a minor allele frequency of 50%. In this phase, no single SNP reached the genome-wide significance threshold ($P < 1.6 \times 10^{-7}$, corresponding to 0.05/304140) in either the combined or the individual analysis of the Icelandic or Dutch GWAS sets. In phase II, the 10 most significant SNPs (all $P < 5 \times 10^{-5}$) were analysed in an additional 2,165 bladder cases and 3,800 controls from 7 follow-up groups, all of European ancestry.

13. The strongest association with urinary bladder cancer, reaching genome-wide significance in the overall analysis of the discovery and following group was observed for the T allele of RS9642880 at 8q24.21 (combined OR=1.22, 95% CI 1.15-1.29), with a frequency of the T allele among controls of 45% relative to non carriers. The OR for heterozygotes and homozygotes was 1.22 and 1.49, respectively. Results from all groups combined demonstrated that the association of RS9642880 (T) to bladder cancer did not deviate from the multiplicative model ($P = 0.76$).

14. The second strongest signal in the combined analysis was observed for RS710521 (A) on chromosome 3q28 (OR = 1.19, $P = 1.15 \times 10^{-7}$). The association of RS710521 (A) to bladder cancer did not deviate from the multiplicative model ($P = 0.70$) and the estimated population attributable risk of RS710521 (A) was 32%. The ORs for heterozygote and homozygote carriers of the risk allele were 1.19 and 1.41 respectively, relative to the risk in non-carriers.

15. One of the major risk factors for bladder cancer is smoking. The authors tested the reported smoking associated variant, RS1051730, in all 9 bladder case-control groups but did not find any association between the risk allele and disease (OR = 1.03, $P = 0.26$). The authors did not find any differences in frequency of RS9642880(A) on chromosome 8 and RS710521(A) on chromosome 3 between the ever smoking cases and the never smoking cases ($P = 0.47$ and $P = 0.55$, respectively). They also observed a nominally significant difference in the frequency of rs710521 (A) on chromosome 3 between male and female cases (OR = 1.14, $P = 0.049$), with frequency in males being slightly greater than females, no difference was observed for RS9642880 ($p = 0.13$).

16. Kiemeny et al. (2010) analysed an extended GWAS dataset composed of a total of 611 Icelandic affected individuals (cases) and 37,478 Icelandic controls, and 1,278 Dutch cases and 1,832 Dutch controls, genotyped on HumanHap300 or HumanCNV370 - Duo BeadChips. They found that no SNPs reached genome-wide significance threshold ($P < 1.6 \times 10^{-7}$) but that the most significant marker was the previously reported marker at 8q24.21, rs9642880 (odds ratio (OR) = 1.22, $P = 2.5 \times 10^{-7}$). In the previous GWAS, Kiemeny et al. (2008) replicated eight of the top 20 markers. In this follow up study, they attempted to genotype the remaining 12 markers in 5/9 additional UBC case-control sample set from Italy, the UK, Spain, Sweden, Belgium, Germany, Eastern Europe and Holland. They reported that one marker, rs798766, located on chromosome 4p16.3 replicated in the combined follow-up groups ($P = 8.5 \times 10^{-8}$), reaching genome-wide significance in the overall analysis of the discovery and follow-up groups (OR = 1.24, 95% CI 1.17–1.32, $P = 9.9 \times 10^{-12}$).

rs798766 is located in intron 5 of *TACC3*, a member of a family of *TACC* genes that play a role in regulating microtubule dynamics. Rs798766 is located 70 kb away from fibroblast growth factor receptor 3 (*FGFR3*), which often harbours activating somatic mutations in low-grade, noninvasive bladder cancer. In combined analysis, Kiemeny et al. (2010) found that the frequency of rs798766[T] was significantly higher in cases with low risk of progression than in those with high risk of progression (combined OR = 1.17, $P = 0.009$).

17. The second GWAS study on bladder cancer was performed by Wu et al. (2009), using case-control participants from an ongoing study in Texas. The GWAS was conducted using the illumina Human Hap610 BeadChip. Analysis was carried out on 556,429 SNPs in 969 cases and 957 controls. No SNPs reached genome-wide significance at this 1st stage of the GWS. Removing highly linked SNPs¹, the authors reported that three SNPs had a p value of $<10^{-5}$ and 50 SNPs showed a p value $<10^{-4}$. Using three more US datasets, the authors performed a fast tracked replication of the top 50 SNPs ($P < 10^{-4}$) and the top additional SNPs in 8q24 ($p < 5 \times 10^{-3}$), a region associated with genetic susceptibility to cancers such as breast, bladder, prostate and colorectal cancer.

18. The rs2294008 SNP was consistent across the US discovery and the replication sets ($P = 7.34 \times 10^{-4}$ and 3.53×10^{-5} , respectively). Using the 9 European populations to replicate this SNP, the overall P value for the combined European populations was 9.83×10^{-5} . Combining all the US and European subjects (6,667 cases and 39,590 controls), the P value was 2.14×10^{-10} and the allelic OR was 1.19 (95% CI, 1.10-1.20). Adjusting for age, gender and smoking status (5,038 cases and 9,363 controls) and using the multi-variable logistic regression status, the reported OR for individuals carrying one copy of the variant allele (T) was 1.30 (95% CI, 1.18-1.42) and for those carrying two copies it was 1.40 (95% CI 1.25-1.56).

19. RS2294008 is a missense SNP located in exon 1 of the PSCA (prostate stem cell antigen precursor) gene. Linkage disequilibrium (LD) analysis of all Hapmap SNPs in the vicinity of RS2294008 showed that it maps to an 11-kb LD block on chromosome 8q24. RS2294008 alters the start codon and is predicted to cause a truncation of 9 amino acids from the N-terminal signal sequence of the primary PSCA translation product.

20. Rothman et al. (2010) conducted the third GWAS of bladder cancer, involving 3,532 cases and 5,120 controls of self-described European descent in stage 1. Follow-up of the most significant signals in two more stages of replication involved a total of 8,382 cases and 48,275 controls. In stage 1, 591,637 SNPs were available for analysis (based on the common SNPs called from both the Illumina Human1M and Human610-Quad BeadChips). Their results from stage 1 of the GWAS confirmed the associations reported with tag SNPs in the four previously identified genomic regions on chromosomes 3q28 (rs710521) and 8q24.21 (rs9642880) (Kiemeny et al., 2008), 8q24.3 (rs2294008) (Wu et al., 2009) and 4p16.3 (rs798766) (Kiemeny et al., 2010). The SNPs, which were most strongly associated with tumours of low grade and a low risk of progression, were rs9642880 on 8q24.21 and rs798766 on 4p16 and these

¹ When multiple SNPs with linkage disequilibrium were found in the analysis, only the SNP with the highest P value was chosen.

findings are consistent with those of Kiemeny et al. (2008) and Wu et al. (2009). They also suggested a strong association with low grade and low risk disease for rs401681 on 5p15.33. They also found a signal below genome-wide significance for rs1495741, which tags the NAT2 gene previously reported as a bladder cancer susceptibility locus on 8p22 (Garcia-Closa et al., (2005) and Rothman et al., (2007)). This rs1495741 is located approximately 10 kb from the 3' end of the NAT2 gene. The rs1495741 AA genotype, which marks the slow acetylation phenotype, showed a highly significant ($P = 5.5 \times 10^{-7}$) association with increased bladder cancer risk that was limited to cigarette smokers (OR = 1.24, 95% CI 1.16–1.32, $P = 1.3 \times 10^{-10}$, P for interaction = 2.8×10^{-4}) compared to the combined GG and AG genotypes which correspond to the intermediate and rapid acetylation phenotypes, respectively.

21. Rothman et al. (2010) identified three new genomic regions on chromosomes 22q13.1, 19q12 and 2q37.1 that were associated with bladder cancer risk below the threshold ($P < 5 \times 10^{-7}$); for genome-wide significance; rs1014971, ($P = 8 \times 10^{-12}$) maps to a non-genetic region of chromosome 22q13.1, rs8102137 ($P = 2 \times 10^{-11}$) on 19q12 maps to *CCNE*, a member of the cyclin-dependent kinase (Cdk)-retinoblastoma protein (pRB) pathway which determines the rate of cell cycle transition from the G1 to the S phase and rs11892031 ($P = 1 \times 10^{-7}$) maps to an intronic region of the locus on uridine 5' diphospho-glucuronosyltransferase 1A gene (*UGT1A*), which encodes the UGT1A family of proteins UGT1A cluster on 2q37.1.

Conclusions

22. GWAS have only recently become practical after three key tools have become available: the discovery of more than 10 million SNPs in the human genome, the results from the HapMap project have been made available and the development of high-throughput genotyping platforms. This paper demonstrated the usefulness of GWAS in identifying important regions of genetic variation associated with increased risk of bladder and colorectal cancer. However, understanding the mechanisms by which GWAS loci contribute to diseases such as cancer will require considerable effort and time. For example, further studies on the regions identified in GWAS are needed to help identify candidate variants for functional studies which in turn should shed light on the biological mechanisms involved in the association.

COC Secretariat (March 2011)

References

- Aben KK, Witjes JA, Schoenberg MP, Hulsbergen-van de Kaa C, Verbeek AL, Kiemeny LA. Familial aggregation of urothelial cell carcinoma. *Int J Cancer*. 2002 Mar 10;98(2):274-8.
- Berndt SI, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk. *Hum. Mol. Genet*. 2008;17:2665-2672.
- Broderick P, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat. Genet*. 2007;39:1315-1317.
- Burt RW, Petersen GM: Familial colorectal cancer: diagnosis and management. In: Young GP, Rozen P, Levin B, eds.: *Prevention and Early Detection of Colorectal Cancer*. London, England: WB Saunders, 1996, pp 171-194.

Curtin K, et al. Meta association of colorectal cancer confirms risk alleles at 8q24 and 18q21. *Cancer Epidemiol. Biomarkers Prev.* 2009;18:616-621.

Cheah PY. Recent advances in colorectal cancer genetics and diagnostics. *Crit. Rev. Oncol. Hematol.* 2009;69:45-55.

Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* (2007) 447:1087–1093.

García-Closas, M. *et al.* NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 2005 **366**, 649–659.

Gudmundsson J, Sulem P, Manolescu A, et al. (2007). Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat. Genet.* (2007) 39:631–637.

Herrera L, ed.: *Familial Adenomatous Polyposis*. New York, NY: Alan R. Liss Inc, 1990.

Houlston RS, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat. Genet.* 2008;40:1426-1435.

Hunter DJ, Kraft P, Jacobs KB, et al. (2007). A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.* 39:870–874.

ISD Online. Cancer Incidence, Mortality and Survival data. Accessed 2010

Jaeger E, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat. Genet.* 2008;40:26-28.

Kiemeny LA, Thorlacius S, Sulem P, Geller F, Aben KK, Stacey SN, Gudmundsson J, Jakobsdottir M, Bergthorsson JT, Sigurdsson A, Blondal T, Witjes JA, Vermeulen SH, Hulsbergen-van de Kaa CA, Swinkels DW, Ploeg M, Cornel EB, Vergunst H, Thorgeirsson TE, Gudbjartsson D, Gudjonsson SA, Thorleifsson G, Kristinsson KT, Mouy M, Snorraddottir S, Placidi D, Campagna M, Arici C, Koppova K, Gurzau E, Rudnai P, Kellen E, Polidoro S, Guarrera S, Sacerdote C, Sanchez M, Saez B, Valdivia G, Ryk C, de Verdier P, Lindblom A, Golka K, Bishop DT, Knowles MA, Nikulasson S, Petursdottir V, Jonsson E, Geirsson G, Kristjansson B, Mayordomo JI, Steineck G, Porru S, Buntinx F, Zeegers MP, Fletcher T, Kumar R, Matullo G, Vineis P, Kiltie AE, Gulcher JR, Thorsteinsdottir U, Kong A, Rafnar T, Stefansson K. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. *Nat Genet.* 2008;40(11):1307-12.

Kiemeny LA, Sulem P, Besenbacher S, Vermeulen SH, Sigurdsson A, Thorleifsson G, Gudbjartsson DF, Stacey SN, Gudmundsson J, Zanon C, Kostic J, Masson G, Bjarnason H, Palsson ST, Skarphedinsson OB, Gudjonsson SA, Witjes JA, Grotenhuis AJ, Verhaegh GW, Bishop DT, Sak SC, Choudhury A, Elliott F, Barrett JH, Hurst CD, de Verdier PJ, Ryk C, Rudnai P, Gurzau E, Koppova K, Vineis P, Polidoro S, Guarrera S, Sacerdote C, Campagna M, Placidi D, Arici C, Zeegers MP, Kellen E, Gutierrez BS, Sanz-Velez JI, Sanchez-Zalabardo M, Valdivia G, Garcia-Prats MD, Hengstler JG, Blaszkewicz M, Dietrich H, Ophoff RA, van den Berg LH, Alexiusdottir K, Kristjansson K, Geirsson G, Nikulasson S, Petursdottir V, Kong A, Thorgeirsson T, Mungan NA, Lindblom A, van Es MA, Porru S, Buntinx F, Golka K, Mayordomo JI, Kumar R, Matullo G, Steineck G, Kiltie AE, Aben KK, Jonsson E, Thorsteinsdottir U, Knowles MA, Rafnar T, Stefansson K. A sequence variant at 4p16.3 confers susceptibility to urinary bladder cancer. *Nat Genet.* 2010;42(5):415-9

Kiltie AE. Common predisposition alleles for moderately common cancers: bladder cancer. *Curr Opin Genet Dev.* 2010 Jun;20(3):218-2

Li L, et al A common 8q24 variant and the risk of colon cancer: a population-based case-control study. *Cancer Epidemiol. Biomarkers Prev.* 2008;17:339-342.

Lichtenstein P, et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* 2000;343:78-85.

Lynch HT, Smyrk T: Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. *Cancer* 78 (6): 1149-67, 1996.

Matsuo K, et al. Association between an 8q24 locus and the risk of colorectal cancer in Japanese. *BMC Cancer* 2009;9:379.

Middeldorp A, et al. Enrichment of low penetrance susceptibility loci in a Dutch familial colorectal cancer cohort. *Cancer Epidemiol. Biomarkers Prev.* 2009;18:3062-3067.

Northern Ireland Cancer Registry. Cancer Incidence and Mortality. Accessed 2010

Office for National Statistics, Cancer Statistics registrations: registrations of cancer diagnosed in 2007, England. 2010

Pearson TA, Manolio TA. (2008). How to interpret a genome-wide association study. *JAMA.* 299(11):1335-44.

Pittman AM, et al. Refinement of the basis and impact of common 11q23.1 variation to the risk of developing colorectal cancer. *Hum. Mol. Genet.* 2008;17:3720-3727.

Poynter JN, et al. Variants on 9p24 and 8q24 are associated with risk of colorectal cancer: results from the Colon Cancer Family Registry. *Cancer Res.* 2007;67:11128-11132.

Rothman, N., Garcia-Closas, M. & Hein, D.W. Commentary: Reflections on G.M. Lower and colleagues' 1979 study associating slow acetylator phenotype with urinary bladder cancer: meta-analysis, historical refinements of the hypothesis, and lessons learned. *Int. J. Epidemiol.* 36, 23–28 (2007).

Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, Real FX, Van Den Berg D, Matullo G, Baris D, Thun M, Kiemeny LA, Vineis P, De Vivo I, Albanes D, Purdue MP, Rafnar T, Hildebrandt MA, Kiltie AE, Cussenot O, Golka K, Kumar R, Taylor JA, Mayordomo JI, Jacobs KB, Kogevinas M, Hutchinson A, Wang Z, Fu YP, Prokunina-Olsson L, Burdett L, Yeager M, Wheeler W, Tardón A, Serra C, Carrato A, García-Closas R, Lloreta J, Johnson A, Schwenn M, Karagas MR, Schned A, Andriole G Jr, Grubb R 3rd, Black A, Jacobs EJ, Diver WR, Gapstur SM, Weinstein SJ, Virtamo J, Cortessis VK, Gago-Dominguez M, Pike MC, Stern MC, Yuan JM, Hunter DJ, McGrath M, Dinney CP, Czerniak B, Chen M, Yang H, Vermeulen SH, Aben KK, Witjes JA, Makkinje RR, Sulem P, Besenbacher S, Stefansson K, Riboli E, Brennan P, Panico S, Navarro C, Allen NE, Bueno-de-Mesquita HB, Trichopoulos D, Caporaso N, Landi MT, Canzian F, Ljungberg B, Tjonneland A, Clavel-Chapelon F, Bishop DT, Teo MT, Knowles MA, Guarrera S, Polidoro S, Ricceri F, Sacerdote C, Allione A, Cancel-Tassin G, Selinski S, Hengstler JG, Dietrich H, Fletcher T, Rudnai P, Gurzau E, Koppova K, Bolick SC, Godfrey A, Xu Z, Sanz-Velez JI, D García-Prats M, Sanchez M, Valdivia G, Porru S, Benhamou S, Hoover RN, Fraumeni JF Jr, Silverman DT, Chanock SJ. (2010). A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nature Genetics* 42(11), 978-983

Schafmayer C, et al. Investigation of the colorectal cancer susceptibility region on chromosome 8q24.21 in a large German case-control sample. *Int. J. Cancer* 2009;124:75-80.

Slattery ML, et al. Increased risk of colon cancer associated with a genetic polymorphism of SMAD7. *Cancer Res.* 2010;70:1479-1485.

Stacey SN, Manolescu A, Sulem P, et al. (2007). Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* 39:865–869.

Tenesa A, et al. New insights into the aetiology of colorectal cancer from genome

wide association studies. *Nat. Rev. Genet.* 2009;10:353-358.

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

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

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

Utsunomiya J, Lynch HT, eds.: *Hereditary Colorectal Cancer: Proceedings of the Fourth International Symposium on Colorectal Cancer (ISCC-4) November 9-11, 1989, Kobe, Japan.* Tokyo, Japan: Springer-Verlag, 1990.

Vineis P, Brennan P, Canzian F, Ioannidis JP, Matullo G, Ritchie M, Stromberg U, Taioli E, Thompson J. (2008). Expectations and challenges stemming from genome-wide association studies. *Mutagenesis.* 23(6):439-44.

Welsh Cancer Intelligence and Surveillance Unit, *Cancer Incidence in Wales.* 2010.

Wu X, Ye Y, Kiemeny LA, Sulem P, Rafnar T, Matullo G, Seminara D, Yoshida T, Saeki N, Andrew AS, Dinney CP, Czerniak B, Zhang ZF, Kiltie AE, Bishop DT, Vineis P, Porru S, Buntinx F, Kellen E, Zeegers MP, Kumar R, Rudnai P, Gurzau E, Koppova K, Mayordomo JI, Sanchez M, Saez B, Lindblom A, de Verdier P, Steineck G, Mills GB, Schned A, Guarrera S, Polidoro S, Chang SC, Lin J, Chang DW, Hale KS, Majewski T, Grossman HB, Thorlacius S, Thorsteinsdottir U, Aben KK, Witjes JA, Stefansson K, Amos CI, Karagas MR, Gu J. Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer. *Nat Genet.* 2009 Sep;41(9):991-5.

Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat. Genet.* (2007) 39:645–649.

Zanke BW et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat. Genet.* 2007;39:989-994.