RESEARCH ARTICLE

# 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

Jan Pander[1☯], Lieke van Huis-Tanja[2☯], Stefan Böhringer[3], Tahar van der Straaten[1], Hans Gelderblom[2], Cornelis Punt[4], Henk-Jan Guchelaar[1]*

1 Department of Clinical Pharmacy & Toxicology, Leiden University Medical Center, The Netherlands; PO box 9600, 2300 RC Leiden, The Netherlands, 2 Department of Clinical Oncology, Leiden University Medical Center, The Netherlands; PO box 9600, 2300 RC Leiden, The Netherlands, 3 Department of Biostatistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands, 4 Department of Medical Oncology, Academic Medical Center Amsterdam, The Netherlands; PO Box 22660, 1100 DD Amsterdam, The Netherlands

☯ These authors contributed equally to this work.

* h.j.guchelaar@lumc.nl

## Abstract

### 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 imputed by using HapMap2. Association with PFS was analysed using a Cox proportional hazards model.

## Results

One marker, rs885036, associated significantly with PFS (P = $2.17 \times 10^{-8}$) 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 $2.30 \times 10^{-7}$ to $1.04 \times 10^{-6}$).

## 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.

## Background

Colorectal cancer is among the most prevalent forms of cancer worldwide, with the estimated number of new diagnoses in the United States for 2013 exceeding 142,000 cases.[1] Combination chemotherapy including a fluoropyrimidine plus the monoclonal agent bevacizumab and either oxaliplatin or irinotecan is generally recommended for first-line treatment of metastatic colorectal carcinoma (mCRC).[2] Unfortunately, even with these modern-day treatment options, median survival is limited and variation in response is largely unpredictable. Optimal selection of patients who will benefit from these extensive treatment schedules is warranted, both from the individual aspect of preventing needless burden of toxicity, and from a population-based aspect of optimal cost-effectiveness.

Germline genetic variation has been shown to predict differences in response to many chemotherapeutic drugs.[3] Until now, most research has focused on genetic alterations in genes encoding for known target or metabolic enzymes. A disadvantage of this candidate-gene approach is the limited knowledge of the exact mechanism of action for many drugs. Considering the immense amount of single nucleotide polymorphisms (SNPs) the human genome harbors, it is likely that many SNPs with potential effect on drug efficacy have not yet been detected. Circumventing the limitations of a candidate gene approach, genome wide association (GWA) studies offer the possibility of simultaneously analyzing a large number of SNPs, even in regions that have not previously been associated with the drug under investigation. This type of study has been applied to identify risk factors for multiple types of cancers, including colorectal carcinoma.[4,5] Moreover, in pharmacogenetic GWA studies, an increasing number of SNPs associated with treatment response in cancer are identified. A recent GWA study in colorectal cancer patients identified SNPs that were associated with adverse drug reactions in response to treatment with 5-FU and oxaliplatin.[6]

Here, we present the results of the first clinical GWA study to find SNPs that are associated with progression free survival (PFS) of first-line combination chemotherapy for mCRC with capecitabine, oxaliplatin and bevacizumab (CAPOX-B), either with or without cetuximab.

## Patients and Methods

Germline DNA was obtained from 553 of 755 previously untreated mCRC patients, who were recruited from the CAIRO2 trial, a multicenter phase III trial of the Dutch Colorectal Cancer Group (DCCG), which randomized between first-line treatment with CAPOX-B versus

CAPOX-B plus cetuximab.[7] All patients received three-weekly cycles of capecitabine 1000mg/m$^2$ b.i.d. orally on days 1–14 plus oxaliplatin 130mg/m$^2$ (maximum of 6 cycles) and bevacizumab 7.5mg/m2, both intravenously on day 1 of each treatment cycle. From cycle 7 the capecitabine dosage was increased to 1250mg/m$^2$. For patients randomized to treatment with CAPOX-B plus cetuximab, cetuximab was administered intravenously at a loading dose of 400mg/m$^2$ on the first treatment day, followed by 250mg/m$^2$ once weekly thereafter. Treatment was continued until disease progression, unacceptable toxicity or death, whichever occurred first. Patient eligibility criteria and guidelines for response assessment have been described in detail elsewhere.[7] The study was approved by the Committee on Research involving Human Subjects Arnhem-Nijmegen and by all local institutional ethics boards. All subjects gave written informed consent.

## Genotyping

Whole blood was collected at baseline and germline DNA was isolated from peripheral leukocytes using MagnaPure Compact (Roche diagnostics, Almere, The Netherlands). Genotyping was performed on Human OmniExpress v12 BeadChip arrays containing 733,202 markers (Illumina, San Diego, CA, USA). Genotype calls were set using GenomeStudio software (Illumina). All products were used according to manufacturer's prescriptions. The following cut-off values were used to filter out incorrectly called genotypes: GenCall $\geq$ 0.85; ClusterSep $\geq$ 0.3; CallFreq >0.85; AB T-mean 0.2–0.8, resulting in the exclusion of 3172 markers (0.43%).

## Statistical analysis

**Quality control.** Quality control was performed using R, version 3.0.1 (http://www.r-project.org/), and *plink*.[8] Markers were excluded based on a minor allele frequency (MAF) threshold of 0.01 (1636 excluded markers; 0.3%) and missingness threshold of 0.05 (no excluded markers). Hardy-Weinberg equilibrium (HWE) was evaluated per marker, using a $\chi^2$ goodness-of-fit statistic with a cut-off P-value of $\leq 10^{-7}$ (36 excluded markers). After these quality checks, 647,550 markers remained in the analysis.

Individuals were excluded based on missingness > 2% (no excluded individuals).

Multidimensional scaling (MDS) was used to investigate possible stratification and outliers. After removal of outliers MDS did not show clustering. Association analysis was performed with and without the first four MDS coordinates as covariates. Ranking and P-values of the top 30 SNPs were almost identical and only analyses without MDS coordinates are shown. In total 33 patients were excluded due to low genotyping quality. Complete information, including genotype, was available for 520 patients. The study flowchart is shown in Fig 1.

**Imputation.** Additionally, non-measured genotypes were imputated using HapMap2, to further enhance SNP density. A minimum imputation accuracy of R$^2$>0.40 was applied to select reliably imputed SNPs. In total 2,202,473 imputed markers were used in this analysis.

**Association model.** For each marker, the association with PFS was calculated with a Cox-proportional hazards model using *R* package *survival*. Age, gender, treatment arm were included as covariates. A first model contained each SNP as variable of interest (marginal model) and a second model tested the interaction SNP*arm as variable of interest while controlling for the main SNP effect (interaction model). Markers were evaluated using an additive genetic model in all analyses. The inflation factor for association models was calculated based on the $\chi^2$-quantiles for the P-values of the evaluated models. Formal significance for a marker was assumed for a two-sided P < 5x10$^{-8}$ to correct for multiple testing, as has been described by others.[9–11] Kaplan-Meier curves were estimated for the marker with the lowest P-value using *R*.
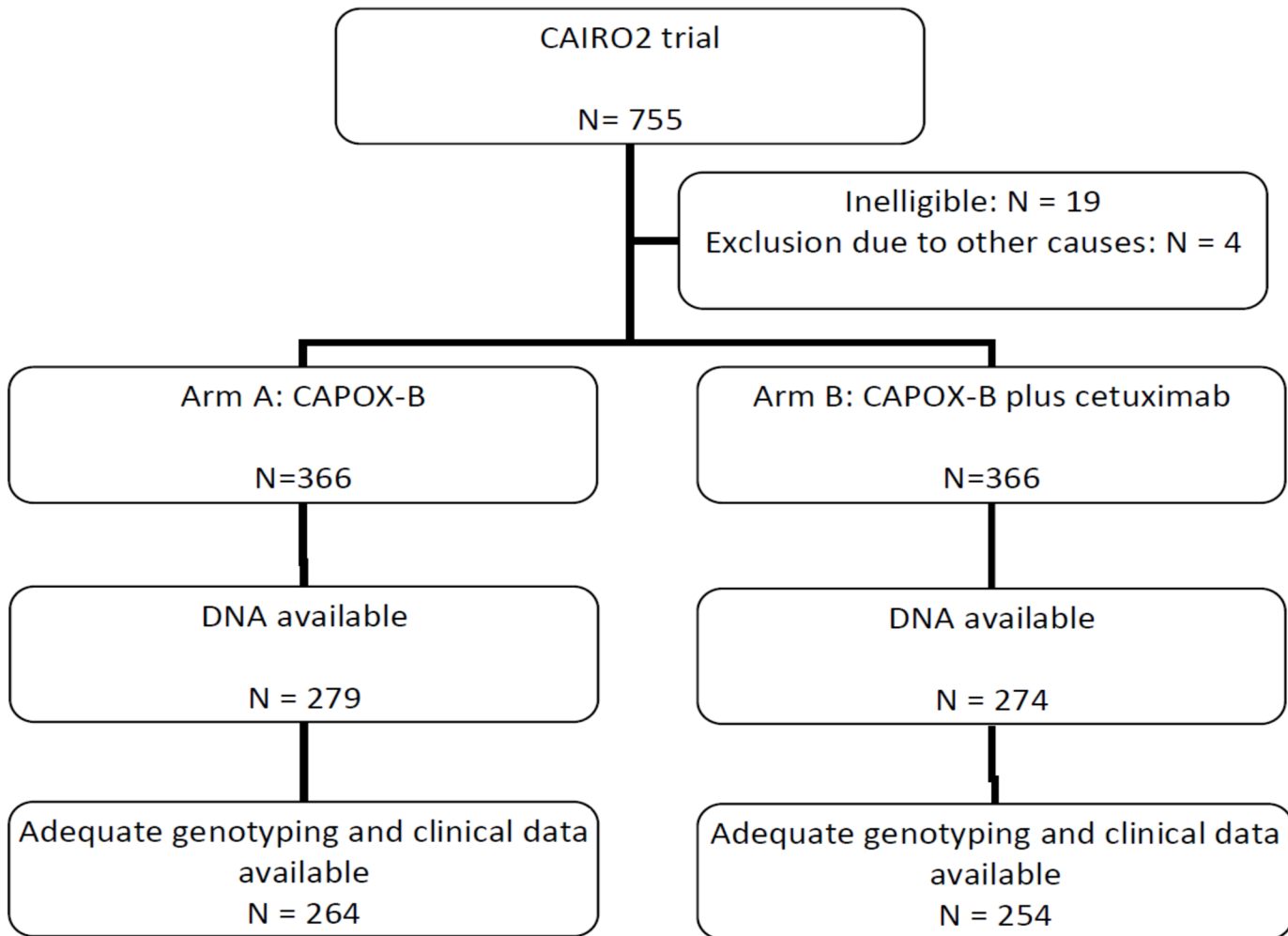
**Fig 1. Cairo2 study flow-chart.**

## Results

### Base-line characteristics

Baseline patient characteristics are described in Table 1 (online only). Median PFS was 10.3 months for all patients included in the analyses (range 0.1–44.7 months). At the time of analysis the primary endpoint PFS was reached by 487 of 520 patients (93.7%).

### Genome-wide association analysis including effect of treatment arm

Initially, analyses were performed including the interaction term of genetic markers with treatment arm as a covariate. The Manhattan plot for these analyses is shown in Fig 2 and details for the ten most significant markers are provided in Table 2. Markers with a differential effect according to treatment arm may actually reflect predictive markers for cetuximab efficacy. One marker (rs885036, position 98671226) on chromosome 2q12 showed a highly significant interaction with treatment arm ($P = 2.17 \times 10^{-8}$). In a stratified analysis, this SNP proved to have a contrasting effect in both treatment arms. Kaplan-Meyer survival curves for PFS according to genotype are shown in Fig 3. For patients treated with CAPOX-B, the G allele associated with

**Table 1. Baseline patient characteristics.**

| | All patients | CAPOX-B | CAPOX-B plus cetuximab |
|---|---|---|---|
| | (n = 520) | (n = 264) | (n = 256) |
| **Age in years** | | | |
| Median (range) | 63.2 (27.6–83.6) | 62.9 (27.6–83.6) | 63.9 (33.2–80.0) |
| **Sex–no (%)** | | | |
| Male | 316 (60.8) | 144 (54.5) | 172 (67.2) |
| Female | 204 (39.2) | 120 (45.5) | 84 (32.8) |
| **LDH at baseline–no (%)[1]** | | | |
| Normal | 307 (59.0) | 152 (57.6) | 155 (60.5) |
| Above upper limit of normal | 213 (41.0) | 112 (42.4) | 101 (39.5) |
| **Previous adjuvant chemotherapy–no (%)** | | | |
| Yes | 443 (85.2) | 223 (84.5) | 220 (86.6) |
| No | 71 (13.7) | 37 (14.0) | 34 (13.4) |
| Unknown | 6 (1.2) | 4 (1.5) | 2 (0.8) |
| **Progression free survival** | | | |
| Median (range) | 10.3 (0.1–44.7) | 10.6 (0.4–44.7) | 10.0 (0.1–40.8) |

[1] According to cut-off values of each individual center

CAPOX-B, capecitabine, oxaliplatin and bevacizumab; LDH, lactate dehydrogenase.

doi:10.1371/journal.pone.0131091.t001

shorter PFS (AA 13.47 months [95%CI 12.16–16.43]; AG 12.22 months [95%CI 9.17–14.32]; GG 9.00 months [95%CI 7.56–10.68]). For patients treated with CAPOX-B plus cetuximab, the G allele associated with increased PFS (AA 7.31 months [95%CI 6.53–10.05]; AG 10.05 months [95%CI 8.83–12.16]; GG 12.35 months [95%CI 10.35–15.44]).

## Overall genome-wide association analysis results

The Manhattan plot for the overall GWA analysis is shown in Fig 4. No marker reached the pre-determined formal significance level. A cluster of markers, located on chromosome 8,



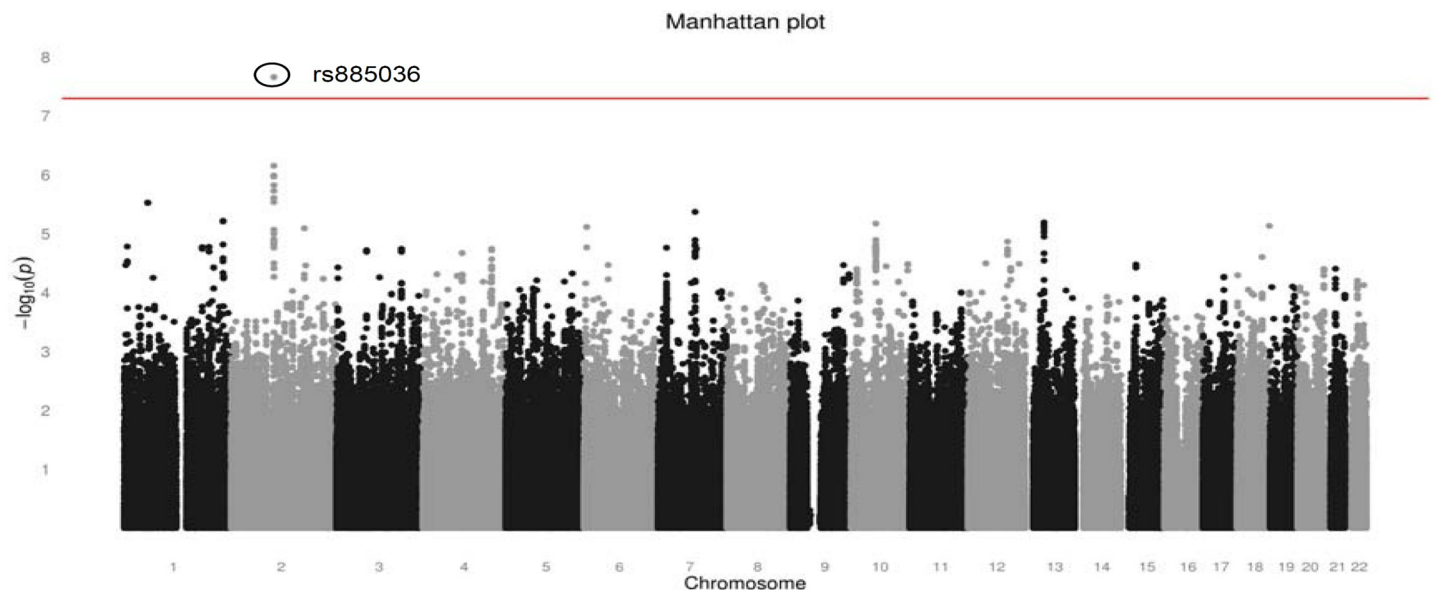**Fig 2. Manhattan plot of–log₁₀ (P-value) of the Cox-proportional hazards model.** Adjusted for age, gender and treatment arm and the interaction of marker with treatment arm. The horizontal line represents the formal genome-wide significance level of $5 \times 10^{-8}$.

doi:10.1371/journal.pone.0131091.g002

Table 2. Ten SNPs with lowest P-values for association with PFS.

| Marker | Chr | Position | Gene | A0 | A1 | Allele freq | $R^2$ | Allele HR* (95% CI) | P-value |
|--------|-----|----------|------|----|----|-------------|-------|---------------------|---------|
| rs885036 | 2 | 98671226 | GnT-IVa | A | G | 0.533 | 1.000 | 0.47 (0.36–0.61) | $2.17 \times 10^{-8}$ |
| rs3769689 | 2 | 98660129 | GnT-IVa | A | G | 0.327 | 0.975 | 0.49 (0.37–0.65) | $6.99 \times 10^{-7}$ |
| rs3769688 | 2 | 98660242 | GnT-IVa | A | G | 0.329 | 0.980 | 0.50 (0.38–0.66) | $1.02 \times 10^{-6}$ |
| rs17448420 | 2 | 98664315 | GnT-IVa | C | T | 0.329 | 0.983 | 0.50 (0.38–0.66) | $1.06 \times 10^{-6}$ |
| rs17448190 | 2 | 98656831 | GnT-IVa | C | T | 0.325 | 0.972 | 0.50 (0.38–0.66) | $1.06 \times 10^{-6}$ |
| rs17448211 | 2 | 98657628 | GnT-IVa | A | T | 0.325 | 0.973 | 0.50 (0.38–0.66) | $1.07 \times 10^{-6}$ |
| rs17448085 | 2 | 98653118 | GnT-IVa | C | T | 0.328 | 0.974 | 0.50 (0.38–0.67) | $1.49 \times 10^{-6}$ |
| rs17514013 | 2 | 98653226 | GnT-IVa | A | G | 0.672 | 0.974 | 1.99 (1.50–2.64) | $1.49 \times 10^{-6}$ |
| rs10199926 | 2 | 98658908 | GnT-IVa | C | T | 0.656 | 0.982 | 1.96 (1.49–2.60) | $1.87 \times 10^{-6}$ |
| rs2309434 | 2 | 98672936 | GnT-IVa | C | T | 0.340 | 0.987 | 0.51 (0.39–0.68) | $2.46 \times 10^{-6}$ |

Abbreviations: HR, Hazard ratio; Chr, chromosome; CI, confidence interval; GnT-IVa: mannosyl (alpha-1,3)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase isozyme A; R2: imputation accuracy. A0: reference allele according to hapmap2, A1: alternate allele.

* HR assuming an additive effect depending on dosage of the alternate alleles, obtained with multivariate Cox-proportionale hazards model, including age, gender, treatment arm, and the interaction treatment arm* marker as covariates.

doi:10.1371/journal.pone.0131091.t002
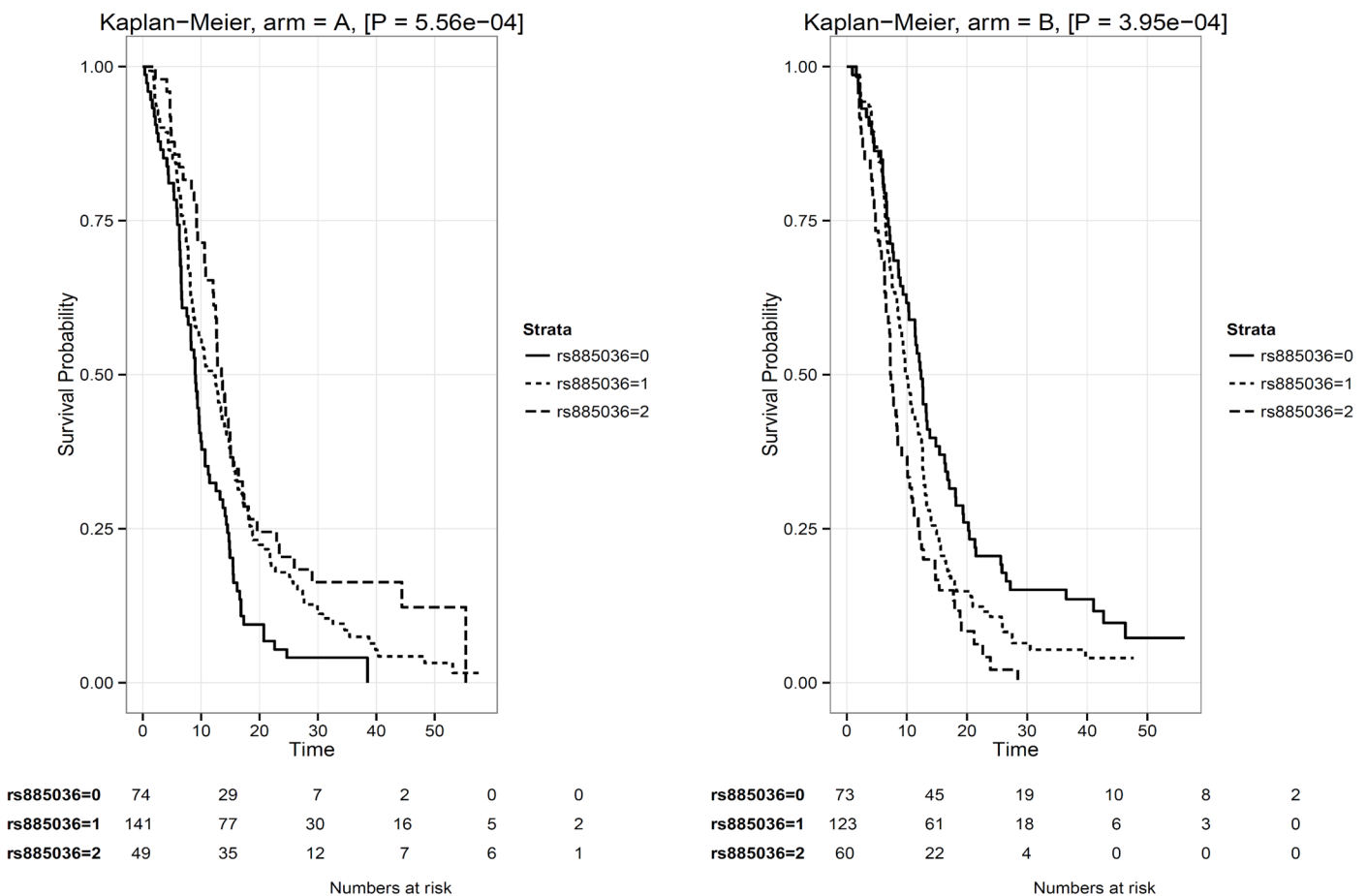


Fig 3. Kaplan-Meyer survival curves according to rs885036 genotype. A. Survival curves for patients in arm A, treated with CAPOX-B in first-line chemotherapy. B. Survival curves for patients in arm B, treated with CAPOX-B plus cetuximab in first-line chemotherapy.

doi:10.1371/journal.pone.0131091.g003
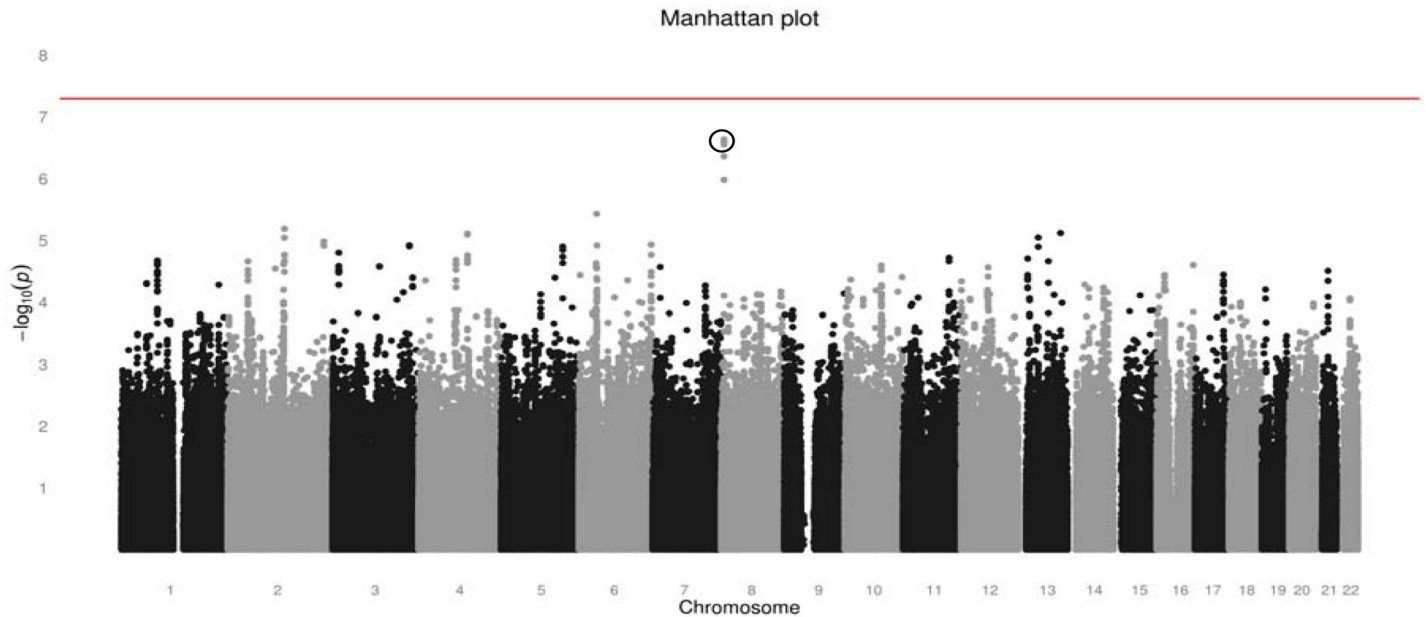
**Fig 4. Manhattan plot of–$\log_{10}$ (P-value) of the Cox-proprotional hazards model. Adjusted for age, gender and treatment arm.** The horizontal line represents the formal genome-wide significance level of $5 \times 10^{-8}$. The topmost significant markers are circled, i.e. rs2936519, rs2928608, rs2928609, rs2912024, rs2978926, rs2928607.

doi:10.1371/journal.pone.0131091.g004

showed the lowest P-values (P = $2.30 \times 10^{-7}$ to P = $1.04 \times 10^{-6}$, respectively). The inflation factor for this analysis was modest ($\lambda$ 1.02), indicating there was no evidence of population stratification or other bias in the analysis. The ten most significant markers are described in Table 3.

## Discussion

This is the first prospective GWA study to find SNPs that predict efficacy of first-line chemotherapy with CAPOX-B, with or without cetuximab, in mCRC patients in a clinical trial setting.

**Table 3. Ten SNPs with lowest P-values for overall association with PFS.**

| Marker | Chr | Position | Gene | A0 | A1 | Allele freq | R2* | Allele HR# (95% CI) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| rs2936519 | 8 | 6626650 | n.a. | A | G | 0.90 | 1.00 | 0.53 (0.43–0.67) | $2.30 \times 10^{-7}$ |
| rs2928608 | 8 | 6626434 | n.a. | C | T | 0.90 | 0.99 | 0.54 (0.43–0.67) | $2.33 \times 10^{-7}$ |
| rs2928609 | 8 | 6626528 | n.a. | C | T | 0.90 | 0.99 | 0.54 (0.43–0.67) | $2.36 \times 10^{-7}$ |
| rs2912024 | 8 | 6626309 | n.a. | C | T | 0.10 | 1.00 | 1.86 (1.49–2.32) | $2.60 \times 10^{-7}$ |
| rs2978926 | 8 | 6626835 | n.a. | A | G | 0.10 | 0.98 | 1.86 (1.49–2.33) | $2.67 \times 10^{-7}$ |
| rs2928607 | 8 | 6626120 | n.a. | C | G | 0.10 | 0.97 | 1.86 (1.49–2.33) | $2.79 \times 10^{-7}$ |
| rs11997869 | 8 | 6625874 | n.a. | C | G | 0.10 | 0.97 | 1.84 (1.47–2.30) | $4.32 \times 10^{-7}$ |
| rs2978931 | 8 | 6625491 | n.a. | A | C | 0.91 | 1.00 | 0.55 (0.44–0.69) | $1.04 \times 10^{-6}$ |
| rs2073016 | 6 | 41128900 | n.a. | C | T | 0.83 | 0.72 | 1.62 (1.31–2.00) | $3.70 \times 10^{-6}$ |
| rs4377367 | 2 | 131441375 | ARHGEF4 | C | T | 0.86 | 0.99 | 0.63 (0.52–0.76) | $6.47 \times 10^{-6}$ |

Abbreviations: HR, Hazard ratio; Chr, chromosome; CI, confidence interval; n.a., marker is not localized within a gene. ARHGEF4: Rho guanine nucleotide exchange factor 4.

* R2: imputation accuracy. A0: reference allele according to hapmap2, A1: alternate allele.

# HR assuming an additive effect depending on dosage of the alternate alleles, obtained with multivariate Cox-proportionale hazards model, including age, gender, treatment arm, and the interaction treatment arm* marker as covariates.

doi:10.1371/journal.pone.0131091.t003

One marker on chromosome 2, rs885036, was significantly associated with PFS, with a contrasting effect in both treatment arms. This SNP may therefore be a potential predictive marker for efficacy of cetuximab containing chemotherapy. Additionally, a cluster of SNPs on chromosome 8 influenced PFS with a similar effect in both treatment arms, almost reaching genome wide significance.

A genome-wide significant effect on PFS was found for rs885036, when treatment arm was taken into consideration. Interestingly, we found opposite effects of genotype on PFS in both treatment arms. This suggests that this polymorphism may have a deleterious prognostic effect on survival per se, but this effect is no longer seen after the addition of cetuximab to standard treatment with CAPOX-B.

Rs885036 is located in the gene encoding for mannosyl (alpha-1,3)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase isozyme A, *GnT-IVa* (also denominated MGAT4A), as are all ten most significant markers from these analyses. The top region of most significant markers exhibits long ranging linkage disequilibrium (LD), extending >50kb in both directions (HapMap3 data, data not shown), which also covers the candidate gene *GnT-IVa*. Of note, rs885036 is a directly measured SNP.

*GnT-IV* encodes for a glycosyltransferase, which is involved in the biosynthesis of oligosaccharides and has previously been associated with tumor behavior. In one study, *GnT-IV* mRNA was upregulated in metastatic colorectal tumors, as opposed to non-metastatic primary colorectal tumors.[12] Strong expression of *GnT-IVa* mRNA and enzyme activity was also found in drug-resistant hepatocellular carcinoma cells.[13]

The rs885036 SNP constitutes a C>T nucleotide change in the second intron of *GnT-IVa*, and is located within a predicted binding site for the highly conserved microRNA-34A (miR-34a).[14,15] MicroRNAs play an important role in post-transcriptional silencing of genes, as their interference with RNA leads to repressed translation or cleavage of RNA.[16] Expression of miR-34a itself is regulated by p53 expression.[17] Cytotoxic stress induced by chemotherapy or irradiation markedly increases miR-34a levels in the presence of p53[18], emphasizing its importance in response to anti-cancer treatment.

We believe that the C>T change leads to altered binding of miR-34a, resulting in differential translation efficacy of *GnT-IVa*. Although our analyses cannot clarify any further functional correlations, we speculate that the altered GnT-IVa protein level may then influence EGFR glycosylation, and thereby form a possible determinant of cetuximab efficacy.

Although no SNP in the overall analyses reached formal genome wide significance, a cluster of SNPs on chromosome 8p23.1 showed very low P-values for the association with PFS. Within the cluster, are two directly genotyped SPNs, including the most significant SNP (rs2936519).

These marers are not localized within any known gene, but the gene encoding for 1-acylglycerol-3-phosphate O-acyltransferase 5 (*APGAT5*), is found approximately 20 kbp upstream. [19] Again, the topmost significant SNPs are in long ranging LD, which covers this candidate gene (HapMap3 data, data not shown). AGPAT5 comes from a family of highly conserved enzymes that catalyze the acylation of lysophosphatidic acid (LPA) to phosphatidic acid, the second step in the *de novo* formation of glycerophospholipids.[20,21] Although the function of most AGPAT isoforms is largely unknown, multiple AGPAT isoforms have been linked to cancer. AGPAT9 and 11 are overexpressed in colon cancer tissue, as well as in other cancer types. [20,22] AGPAT2 inhibition was shown to induce *in vitro* growth arrest and cell death in different tumor types, including several colon cancer cell lines.[23]

To our knowledge, this is the first prospective GWA study investigating systemic treatment efficacy in Caucasian mCRC patients. Only two previous similar studies have been published worldwide, and none of the markers identified by those experiments overlap with our results. [24,25] These studies, both by the same Korean research group, followed a three step design,

validating results from GWA analysis in both an independent patient cohort and transfected colorectal cancer cell lines. They found two SNPs with putative influence on efficacy of cetuximab-containing regimens for colorectal carcinoma[25], as well as one SNP influencing disease recurrence after adjuvant chemotherapy with 5-FU for stage II or III colorectal carcinoma.[24] In both studies, the selection of markers to be carried forward to the validation studies was based on the minor allele frequency and location in a haplotype block in the Asian population, and only non-synonymous SNPs were included. Considerable differences in allele frequencies exist between Asian and European populations for the relevant genetic markers from both our and Kim's analyses, which in part may explain the lack of overlap between our results.[19] Also, considerably larger treatment heterogeneity was present in their population, thereby possibly diluting the effect of germline genetic markers on treatment efficacy.

Our study is subject to several limitations. Interpretation of GWA studies in general is hampered by the need for large patient populations, to ensure sufficient power to detect genotype effects. With 547 patients, our cohort used for this GWA study is larger than many other pharmacogenetic study populations. Non-significant results could therefore lead to the conclusion that indeed none of the tested SNPs are associated with PFS upon CAPOX-B treatment in mCRC. However, since the power to detect associations depends not only on population size, but also on the size of a SNP-effect and on genotype frequency. Despite a relatively large sample size, false negative results could arise for SNPs with low minor allele frequencies. In fact, minor allele frequencies of the three topmost significant SNPs from the overall analyses are only 0.10 in our population, which may well have led the results to be non-significant, despite an actual effect on PFS.

Obviously, it can be argued that rare variants or markers with a small effect on treatment efficacy are not clinically relevant, irrespective of their statistical significance level. However, genetic predisposition is unlikely to depend on only a few SNPs for most drug effects. Rather, multiple genetic markers could attribute in small amount and interact with each other.[26,27] Their combined effect, in conjunction with clinical and pathologic parameters, could be of value in predictive models. Such a model has recently been proposed for colorectal cancer risk assessment.[28] In future, similar models should be constructed for predicting treatment response as well as adverse events for diverse systemic treatment regimens. In this scenario, any SNP influencing the complex trait 'drug efficacy' is of value.

Another drawback of GWA analyses is that these studies only focus on the influence of genetic variation, without taking into consideration most patient and tumor characteristics. Whereas in the time the CAIRO2 trial was performed cetuximab was prescribed to unselected patients, regardless of tumor characteristics, it was shown shortly thereafter that patients carrying a somatic mutation in *Kras* codon 12 or 13 are resistant to treatment with cetuximab.[29–32] More recently, patients with mutations in *Nras* were also shown not to benefit from treatment with anti-epidermal growth factor antibodies.[33] More somatic mutations precluding EGFR inhibitor efficacy may be discovered.

At the time the study was conducted, EGFR inhibitor therapy was not restricted to Kras wild type tumors only. However, *Kras* mutation status has been analyzed restrospectively in those patients for whom tumor DNA was available. For this subgroup, we performed additional analyses, taking *Kras* codon 12 and codon 13 alleles into consideration. We found that rs885036 remained an independent predictor for PFS, with only a small change in HR (HR 0.466 without *Kras*, HR 0.472 with *Kras* as a covariate; Table 4).

Recently, it was found that patient survival in colorectal cancer is influenced by different molecular subtypes, including not only *Kras* mutation status, but several other genetic and epigenetic factors.[34] Because molecular subtype was not determined in the CAIRO2 population, we cannot completely exclude an effect of molecular subtype on the outcome of our analyses.

**Table 4. Ten SNPs with lowest P-values for association with PFS.**

| Marker | Chr | Position | A0 | Allele frequency | Allele HR* | P-value |
|---|---|---|---|---|---|---|
| Rs8180209 | 4 | 90863476 | G | 0.08 | 0.24 | $2.83 \times 10^{-01}$ |
| Rs530376 | 5 | 152819508 | G | 0.45 | 0.46 | $5.06 \times 10^{-01}$ |
| Rs11830274 | 12 | 90597138 | A | 0.02 | 0.03 | $7.28 \times 10^{-01}$ |
| Rs11837647 | 12 | 90624226 | A | 0.02 | 0.03 | $7.28 \times 10^{-01}$ |
| **Rs885036** | **2** | **98671226** | **A** | **0.46** | **2.12** | $6.87 \times 10^{-01}$ |
| Rs3118146 | 1 | 180533882 | C | 0.23 | 2.32 | $2.95 \times 10^{-01}$ |
| Rs2477185 | 1 | 180533140 | A | 0.23 | 2.32 | $3.19 \times 10^{-01}$ |
| Rs1776148 | 1 | 240109167 | A | 0.36 | 2.14 | $9.27 \times 10^{-02}$ |
| Rs1025159 | 1 | 163741907 | A | 0.15 | 0.37 | $7.67 \times 10^{-01}$ |
| Rs3025363 | 9 | 135471888 | A | 0.38 | 0.47 | $5.87 \times 10^{-01}$ |

Abbreviations: HR, Hazard ratio; Chr, chromosome; A0: reference allele according to hapmap2

* HR assuming an additive effect depending on dosage of the alternate alleles, obtained with multivariate Cox-proportionale hazards model, including age, gender, treatment arm, *Kras* mutation status and the interaction treatment arm* marker as covariates.

doi:10.1371/journal.pone.0131091.t004

However, there is no evidence in literature for an association between *GnT-IVa* genotype and colorectal cancer molecular subtype.

Contrary to practice in clinical trials, serum LDH level at baseline, although strongly correlated to PFS[7], was not included in our multivariate model. Confounding by the prognostic effect of LDH cannot be excluded, since there was a slight overrepresentation of patients with increased LDH levels in several genotype groups with shorter PFS (data not shown). LDH level in itself may be a predictive as well as a prognostic factor, and not completely independent of germline genotype. Serum LDH-levels have been associated with intratumoral gene expression of vascular endothelial growth factor type A (*VEGFA*) and vascular endothelial growth factor receptor type 1 (*VEGFR1*).[35] Both phenomena are thought to be the result of stimulation of hypoxia-inducible factor 1α (HIF-1α), due to the intratumoral aerobic glycolysis known as the Warburg effect.[36] Results of a recent study suggest that, although high serum LDH levels correlate with reduced survival in colorectal cancer patients, treatment with bevacizumab can improve PFS for patients with high pre-treatment LDH levels to that of patients with normal baseline LDH levels.[37] Because an interaction of LDH with genotype cannot be excluded, and inclusion of LDH in the experimental model could possibly have obscured genotype effects, we chose not to correct for the effect of LDH level at baseline in this study.

## Conclusion

In this GWA study, investigating the effect of genetic variation on PFS on first line CAPOX-B with or without cetuximab in mCRC, we found one marker with a significant effect on treatment efficacy, with opposing results in each treatment arm.

This SNP, rs885036, may be of predictive importance for patients treated with cetuximab-containing regimens, and should be further validated in patients with RAS wildtype tumors. Patients in our sample carrying the rs885036 GG genotype seem to respond better to CAPOX-B plus cetuximab, whereas patients with the AA genotype benefit more from CAPOX-B alone. Therefore, it is possible a subgroup of patients carrying the GG genotype may benefit from the addition of cetuximab to CAPOX-B, contrary to overall findings for the general study population of CAIRO2 and study with comparable design.[7,38] Whether this marker also predicts PFS in patients treated with cetuximab monotherapy should be subject to further research. We

formulated a hypothesis for the pathophysiological basis of this effect, which we are now testing in pre-clinical studies.

In addition, a cluster of SNPs on chromosome 8 with very low P-values was found, which may also have functional significance. For these SNPs, we cannot differentiate between a prognostic effect of these polymorphisms on PFS and a predictive effect in relation to chemotherapy. These results should also be validated in other populations.

## Acknowledgments

## Author Contributions

Conceived and designed the experiments: HJG HG CP. Performed the experiments: TvdS JP. Analyzed the data: JP LvH SB. Wrote the paper: JP LvH SB TvdS HG CP HJG.

## References

1. Siegel R, Naishadham D, Jemal A (2013) Cancer statistics, 2013. CA Cancer J Clin 63: 11–30. doi: 10.3322/caac.21166 PMID: 23335087

2. Edwards MS, Chadda SD, Zhao Z, Barber BL, Sykes DP (2012) A systematic review of treatment guidelines for metastatic colorectal cancer. Colorectal Dis 14: e31–e47. doi: 10.1111/j.1463-1318.2011.02765.x PMID: 21848897

3. Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D (2009) Genetic prognostic and predictive markers in colorectal cancer. Nat Rev Cancer 9: 489–499. nrc2645 [pii]; doi: 10.1038/nrc2645 PMID: 19536109

4. Stadler ZK, Thom P, Robson ME, Weitzel JN, Kauff ND, Hurley KE, et al. (2010) Genome-wide association studies of cancer. J Clin Oncol 28: 4255–4267. JCO.2009.25.7816 [pii]; doi: 10.1200/JCO.2009.25.7816 PMID: 20585100

5. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, et al. (2013) Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-Wide Meta-analysis. Gastroenterology 144: 799–807. S0016-5085(12)01846-X [pii]; doi: 10.1053/j.gastro.2012.12.020 PMID: 23266556

6. Fernandez-Rozadilla C, Cazier JB, Moreno V, Crous-Bou M, Guino E, Duran G, et al. (2013) Pharmacogenomics in colorectal cancer: a genome-wide association study to predict toxicity after 5-fluorouracil or FOLFOX administration. Pharmacogenomics J 13: 209–217. tpj20122 [pii]; doi: 10.1038/tpj.2012.2 PMID: 22310351

7. Tol J, Koopman M, Cats A, Rodenburg CJ, Creemers GJ, Schrama JG, et al. (2009) Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. N Engl J Med 360: 563–572. 360/6/563 [pii]; doi: 10.1056/NEJMoa0808268 PMID: 19196673

8. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575. S0002-9297(07)61352-4 [pii]; doi: 10.1086/519795 PMID: 17701901

9. Welcome Trust Case Controle Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678. nature05911 [pii]; doi: 10.1038/nature05911 PMID: 17554300

10. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. (2008) SLCO1B1 variants and statin-induced myopathy—a genomewide study. N Engl J Med 359: 789–799. NEJMoa0801936 [pii]; doi: 10.1056/NEJMoa0801936 PMID: 18650507

11. Dudbridge F, Gusnanto A (2008) Estimation of significance thresholds for genomewide association scans. Genet Epidemiol 32: 227–234. doi: 10.1002/gepi.20297 PMID: 18300295

12. D'Arrigo A, Belluco C, Ambrosi A, Digito M, Esposito G, Bertola A, et al. (2005) Metastatic transcriptional pattern revealed by gene expression profiling in primary colorectal carcinoma. Int J Cancer 115: 256–262. doi: 10.1002/ijc.20883 PMID: 15688387

13. Kudo T, Nakagawa H, Takahashi M, Hamaguchi J, Kamiyama N, Yokoo H, et al. (2007) N-glycan alterations are associated with drug resistance in human hepatocellular carcinoma. Mol Cancer 6: 32. 1476-4598-6-32 [pii]; doi: 10.1186/1476-4598-6-32 PMID: 17488527

14. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27: 91–105. S1097-2765(07) 00407-8 [pii]; doi: 10.1016/j.molcel.2007.06.017 PMID: 17612493

15. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15–20. S0092867404012607 [pii]; doi: 10.1016/j.cell.2004.12.035 PMID: 15652477

16. Ke XS, Liu CM, Liu DP, Liang CC (2003) MicroRNAs: key participants in gene regulatory networks. Curr Opin Chem Biol 7: 516–523. S1367593103000759 [pii]. PMID: 12941428

17. He L, He X, Lim LP, de SE, Xuan Z, Liang Y, et al. (2007) A microRNA component of the p53 tumour suppressor network. Nature 447: 1130–1134. nature05939 [pii]; doi: 10.1038/nature05939 PMID: 17554337

18. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, et al. (2007) p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Curr Biol 17: 1298–1307. S0960-9822(07)01638-7 [pii]; doi: 10.1016/j.cub.2007.06.068 PMID: 17656095

19. Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine.(dbSNP Build ID: 36.3). http://www.ncbi.nlm.nih.gov/SNP/.

20. Agarwal AK, Garg A (2010) Enzymatic activity of the human 1-acylglycerol-3-phosphate-O-acyltransferase isoform 11: upregulated in breast and cervical cancers. J Lipid Res 51: 2143–2152. jlr.M004762 [pii]; doi: 10.1194/jlr.M004762 PMID: 20363836

21. Shindou H, Shimizu T (2009) Acyl-CoA:lysophospholipid acyltransferases. J Biol Chem 284: 1–5. R800046200 [pii]; doi: 10.1074/jbc.R800046200 PMID: 18718904

22. Mansilla F, da Costa KA, Wang S, Kruhoffer M, Lewin TM, Orntoft TF, et al. (2009) Lysophosphatidylcholine acyltransferase 1 (LPCAT1) overexpression in human colorectal cancer. J Mol Med (Berl) 87: 85–97. doi: 10.1007/s00109-008-0409-0

23. Coon M, Ball A, Pound J, Ap S, Hollenback D, White T, et al. (2003) Inhibition of lysophosphatidic acid acyltransferase beta disrupts proliferative and survival signals in normal cells and induces apoptosis of tumor cells. Mol Cancer Ther 2: 1067–1078. PMID: 14578472

24. Kim JC, Kim SY, Cho DH, Roh SA, Choi EY, Jo YK, et al. (2010) Genome-wide identification of chemosensitive single nucleotide polymorphism markers in colorectal cancers. Cancer Sci 101: 1007–1013. CAS1461 [pii]; doi: 10.1111/j.1349-7006.2009.01461.x PMID: 20085586

25. Kim JC, Kim SY, Cho DH, Ha YJ, Choi EY, Kim CW, et al. (2011) Novel chemosensitive single-nucleotide polymorphism markers to targeted regimens in metastatic colorectal cancer. Clin Cancer Res 17: 1200–1209. 1078-0432.CCR-10-1907 [pii]; doi: 10.1158/1078-0432.CCR-10-1907 PMID: 21239504

26. Jiao S, Hsu L, Berndt S, Bezieau S, Brenner H, Buchanan D, et al. (2012) Genome-wide search for gene-gene interactions in colorectal cancer. PLoS One 7: e52535. doi: 10.1371/journal.pone.0052535 PONE-D-12-29750 [pii]. PMID: 23300701

27. Pander J, Wessels JA, Gelderblom H, van der Straaten T, Punt CJ, Guchelaar HJ (2011) Pharmacogenetic interaction analysis for the efficacy of systemic treatment in metastatic colorectal cancer. Ann Oncol 22: 1147–1153. mdq572 [pii]; doi: 10.1093/annonc/mdq572 PMID: 21048041

28. Yarnall JM, Crouch DJ, Lewis CM (2013) Incorporating non-genetic risk factors and behavioural modifications into risk prediction models for colorectal cancer. Cancer Epidemiol 37: 324–329. S1877-7821 (12)00175-0 [pii]; doi: 10.1016/j.canep.2012.12.008 PMID: 23375517

29. De Roock W, Piessevaux H, De Schutter J, Janssens M, De Herthogh G, Personeni N, et al. (2008) KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. Ann Oncol 19: 508–515. mdm496 [pii]; doi: 10.1093/annonc/mdm496 PMID: 17998284

30. Di Fiore F, Blanchard F, Charbonnier F, Le Pessot F, Lamy A, Galais MP, et al. (2007) Clinical relevance of KRAS mutation detection in metastatic colorectal cancer treated by Cetuximab plus chemotherapy. Br J Cancer 96: 1166–1169. 6603685 [pii]; doi: 10.1038/sj.bjc.6603685 PMID: 17375050

31. Khambata-Ford S, Garrett CR, Meropol NJ, Basik M, Harbison CT, Wu S, et al. (2007) Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. J Clin Oncol 25: 3230–3237. 25/22/3230 [pii]; doi: 10.1200/JCO.2006.10.5437 PMID: 17664471

32. Lievre A, Bachet JB, Boige V, Cayre A, Le Corre D, Buc E, et al. (2008) KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. J Clin Oncol 26: 374–379. 26/3/374 [pii]; doi: 10.1200/JCO.2007.12.5906 PMID: 18202412

33. Douillard JY, Oliner KS, Siena S, Tabernero J, Burkes R, Barugel M, et al. (2013) Panitumumab-FOL-FOX4 treatment and RAS mutations in colorectal cancer. N Engl J Med 369: 1023–1034. doi: 10.1056/NEJMoa1305275 PMID: 24024839

34. Phipps AI, Limburg PJ, Baron JA, Burnett-Hartman AN, Weisenberger DJ, Laird PW, et al. (2015) Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology 148: 77–87. S0016-5085(14)01192-5 [pii]; doi: 10.1053/j.gastro.2014.09.038 PMID: 25280443

35. Azuma M, Shi M, Danenberg KD, Gardner H, Barrett C, Jacques CJ, et al. (2007) Serum lactate dehydrogenase levels and glycolysis significantly correlate with tumor VEGFA and VEGFR expression in metastatic CRC patients. Pharmacogenomics 8: 1705–1713. doi: 10.2217/14622416.8.12.1705 PMID: 18086000

36. Lu H, Forbes RA, Verma A (2002) Hypoxia-inducible factor 1 activation by aerobic glycolysis implicates the Warburg effect in carcinogenesis. J Biol Chem 277: 23111–23115. doi: 10.1074/jbc.M202487200 M202487200 [pii]. PMID: 11943784

37. Scartozzi M, Giampieri R, Maccaroni E, Del PM, Faloppi L, Bianconi M, et al. (2012) Pre-treatment lactate dehydrogenase levels as predictor of efficacy of first-line bevacizumab-based therapy in metastatic colorectal cancer patients. Br J Cancer 106: 799–804. bjc201217 [pii]; doi: 10.1038/bjc.2012.17 PMID: 22315053

38. Saltz L, Badarinath S, Dakhil S, Bienvenu B, Harker WG, Birchfield G, et al. (2012) Phase III trial of cetuximab, bevacizumab, and 5-fluorouracil/leucovorin vs. FOLFOX-bevacizumab in colorectal cancer. Clin Colorectal Cancer 11: 101–111. S1533-0028(11)00063-6 [pii]; doi: 10.1016/j.clcc.2011.05.006 PMID: 22055112