Genome-wide association study of the efficacy of capecitabine, oxaliplatin and bevacizumab in metastatic colorectal cancer

Jan Pander • Stefan Böhringer • Tahar van der Straaten
Yavuz Ariyurek • Jeanine Houwing-Duistermaat • Hans Gelderblom
Cornelis Punt • Henk-Jan Guchelaar

This study was financially supported by a grant from the Dutch Cancer Society UL2009-4528
Abstract

Background
A more optimal selection of patients that will benefit from the frequently used first-line treatment of advanced colorectal cancer (ACC) consisting of the combination of capecitabine, oxaliplatin and bevacizumab (CAPOX-B) is warranted. We used a genome-wide association study to find single nucleotide polymorphisms (SNPs) that are associated with the efficacy of CAPOX-B.

Methods
Germline DNA was obtained from 547 previously untreated ACC patients in the randomized phase III CAIRO2 trial, in which patients were randomized between CAPOX-B or CAPOX-B plus cetuximab. Whole-genome genotyping was performed using 700 k Illumina OmniExpress BeadChip arrays. Associations between SNPs and progression-free survival (PFS) were tested using Cox-proportional hazard models. Associations were considered significant when \( P < 5 \times 10^{-8} \).

Results
Three SNPs located at 8p23.1 showed a trend toward significance for association with PFS (rs292936519, \( P = 1.24 \times 10^{-7} \); rs2912024, \( P = 1.38 \times 10^{-7} \) and rs2978931, \( P = 6.75 \times 10^{-7} \)). These SNPs are 20 kbp downstream of the AGPAT5 gene, which encodes a protein that is involved in phospholipid biosynthesis.

Conclusion
Even though these results possibly identify a novel genetic predictor for the efficacy of CAPOX-B, further analyses are required before definitive conclusions can be made based upon these data.

Background
A frequently used first-line treatment of advanced colorectal cancer (ACC) consists of the combination of capecitabine, oxaliplatin and bevacizumab (CAPOX-B). Even though this combination results in a prolongation of survival compared with no treatment, the one year progression-free survival (PFS) rate is below 50%. In order to reduce toxicity and costs, a more optimal selection of patients that will benefit from modern systemic treatment is warranted.

Heritable genetic variation has proven to predict variation in response to many therapeutics drugs. The basis of such research is currently limited to genetic variation in target or metabolic enzymes that have been selected using the candidate gene approach. The disadvantage of this approach is that it is limited to current knowledge of the mechanism of action of the investigated drugs. Since it is estimated that there are more than 10,000,000 single nucleotide polymorphisms (SNPs) in the human genome, it is very likely that many of these SNPs are not detected in the current approach of pathway based research.

Genome-wide association studies, in which the entire genome is characterized for SNPs, have been applied in the past years to identify risk factors for several types of cancers in large case-control series. Regarding outcome of systemic therapy, genome-wide association studies have identified SNPs associated with muscoskeletal adverse reactions to aromatase inhibitors, treatment response for childhood acute lymphoblastic lymphoma and pharmacokinetics of methotrexate. All of these studies are based upon a case-control design with \( \chi^2 \)-tests to test for associations, but survival could also be applied as an endpoint using Cox-proportional hazards models to test for associations.

Here we present the first results of a genome-wide association study to find SNPs that are associated with the efficacy of first-line CAPOX-B for ACC in a clinical trial setting with PFS as the primary endpoint.

Patients and Methods

Patients
Germline DNA was obtained from 547 of 736 previously untreated ACC patients who were randomized between treatment with CAPOX-B or CAPOX-B plus cetuximab in the multicenter randomized phase III CAIRO2 trial of the Dutch Colorectal Cancer Group (DCCG). Capecitabine 1000 mg/m\(^2\) (increased to 1250 mg/m\(^2\) from cycle 7) was administered orally twice daily on days 1–14 of each 3-week treatment cycle. Oxaliplatin 130 mg/m\(^2\) (maximum of six cycles) and bevacizumab 7.5 mg/kg were administered intravenously on day 1 of each treatment cycle. For patients randomized to the
CAPOX-B plus cetuximab arm, cetuximab was administered intravenously at a dose of 400 mg/m² on the first day, followed by 250 mg/m² weekly thereafter. Treatment was continued until disease progression, death or unacceptable toxicity, whichever occurred first. Patient eligibility criteria are described in detail elsewhere.9 The collection of a peripheral blood sample for pharmacogenetic research was pre-specified in the study protocol and required additional written informed consent. The protocol was approved by the local institutional review boards of all participating centers.

Genotype results
Three SNPs (rs2936519, rs2912024 and rs2978931) located on chromosome 8, cytogenic band 8p23.1, showed the lowest P-values (P = 1.24 x 10-7, P = 1.38 x 10-7 and P = 6.75 x 10-7, respectively). After these quality checks, 589,274 markers remained for the statistical analysis.

Statistical analysis
For each marker, a Cox proportional hazards model was calculated using R, which included age, gender and treatment arm as covariates. Since it is not known whether the effects of the markers are dominant, recessive or multiplicative, each marker was included in a multiplicative model (i.e. AA = 0, AB = 1 and BB = 2). Observed P-values were plotted against theoretical P-values (QQ-plot), and the inflation factor was calculated by \( \frac{\text{median}(T_1, \ldots, T_n)^2}{0.675} \), with \( T_1, \ldots, T_n \) being the square roots of the \( \chi^2 \) quantiles for the P-values of the markers. Formal significance for a marker was assumed for \( P < 5 \times 10^{-8} \). To check for effects that could be ascribed to the treatment arm, interaction between the marker and treatment arm was included in the model. The association was tested only in the CAPOX-B arm if the P-value of the marker*arm interaction term was < 0.001. Kaplan-Meier curves were estimated for the marker with the lowest P-value using SPSS version 17.0 (SPSS, Chicago, IL, USA).

Results

Patients
At the time of the analysis (December 2010), the primary endpoint PFS was reached in 459 patients (88.1%). Median PFS was 10.6 months (95% confidence interval [95%CI], 9.5 to 11.6 months). In the CAPOX-B and the CAPOX-B plus cetuximab arms, median PFS was 10.8 months (95%CI, 9.0 to 12.5 months) and 10.1 months (95%CI, 9.0 to 11.3 months), respectively. Baseline patient characteristics are described in Table 1.

Table 1  Baseline patient characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>63.1</th>
<th>27.6 - 83.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age - year</td>
<td>63.1</td>
<td>27.6 - 83.6</td>
</tr>
<tr>
<td>Sex - no (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>316 (60.7%)</td>
<td>205 (39.3%)</td>
</tr>
<tr>
<td>female</td>
<td>205 (39.3%)</td>
<td>316 (60.7%)</td>
</tr>
<tr>
<td>Arm - no (%)</td>
<td>264 (50.7%)</td>
<td>257 (49.3%)</td>
</tr>
<tr>
<td>CAPOX-B</td>
<td>264 (50.7%)</td>
<td>257 (49.3%)</td>
</tr>
<tr>
<td>CAPOX-B plus cetuximab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum lactate dehydrogenase level - no (%)</td>
<td>307 (58.9%)</td>
<td>213 (40.9%)</td>
</tr>
<tr>
<td>normal*</td>
<td>307 (58.9%)</td>
<td>213 (40.9%)</td>
</tr>
<tr>
<td>above normal*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* according to the cutoff values of each individual center
This gene encodes an integral membrane protein that converts lysophosphatidic acid (LPA) to phosphatidic acid (PA), the second step in de novo phospholipid biosynthesis, the major constituent of the cell membrane (http://www.ncbi.nlm.nih.gov). Additionally, LPA is a potent mitogen that has been linked to the development and progression of breast cancer. When these markers have been fully evaluated (i.e., functional analysis on gene-function or gene-expression level) and their associations have been confirmed in an independent cohort, they could be used to optimize selection of ACC patients for CAPOX-B treatment. This is the first pharmacogenomic genome-wide study on the efficacy of palliative therapy for ACC. Unfortunately, the associations between the SNPs and PFS did not reach formal statistical significance at the 5 x 10^{-8} level, but a trend toward significance was found for 3 SNPs. This could be the result of insufficient power, possibly in combination with very stringent correction for multiple testing. Otherwise, the results may simply be false positive findings based upon the large number of statistical tests. There are 47 more patients that have to be genotyped, and were therefore not genotype.

### Table 2

Top 10 SNPs with lowest P-values for association with PFS in a Cox-proportional hazards model with age, gender and treatment arm as covariates

<table>
<thead>
<tr>
<th>marker</th>
<th>chr</th>
<th>position</th>
<th>gene</th>
<th>allele frequency</th>
<th>P-value</th>
<th>allelic HR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2936519</td>
<td>8</td>
<td>6626650</td>
<td>n.a.</td>
<td>0.104</td>
<td>1.24 x 10^{-7}</td>
<td>0.545 (0.435 – 0.682)</td>
</tr>
<tr>
<td>rs2912024</td>
<td>8</td>
<td>6626309</td>
<td>n.a.</td>
<td>0.105</td>
<td>1.38 x 10^{-7}</td>
<td>0.547 (0.437 – 0.685)</td>
</tr>
<tr>
<td>rs2978931</td>
<td>8</td>
<td>6625491</td>
<td>n.a.</td>
<td>0.101</td>
<td>6.75 x 10^{-7}</td>
<td>0.561 (0.447 – 0.705)</td>
</tr>
<tr>
<td>rs4850159</td>
<td>2</td>
<td>131442241</td>
<td>ARHGEF4</td>
<td>0.136</td>
<td>2.58 x 10^{-6}</td>
<td>0.627 (0.516 – 0.762)</td>
</tr>
<tr>
<td>n6734725</td>
<td>2</td>
<td>46751074</td>
<td>n.a.</td>
<td>0.350</td>
<td>2.99 x 10^{-6}</td>
<td>0.713 (0.619 – 0.822)</td>
</tr>
<tr>
<td>rs17688362</td>
<td>18</td>
<td>39999678</td>
<td>n.a.</td>
<td>0.185</td>
<td>4.68 x 10^{-6}</td>
<td>0.657 (0.548 – 0.786)</td>
</tr>
<tr>
<td>n17444829</td>
<td>4</td>
<td>113593423</td>
<td>n.a.</td>
<td>0.133</td>
<td>5.49 x 10^{-6}</td>
<td>1.556 (1.285 – 1.884)</td>
</tr>
<tr>
<td>rs1730442</td>
<td>4</td>
<td>113581912</td>
<td>ALPK1</td>
<td>0.132</td>
<td>6.62 x 10^{-6}</td>
<td>1.554 (1.282 – 1.884)</td>
</tr>
<tr>
<td>rs10084940</td>
<td>8</td>
<td>92317629</td>
<td>n.a.</td>
<td>0.061</td>
<td>7.20 x 10^{-6}</td>
<td>1.849 (1.413 – 2.429)</td>
</tr>
<tr>
<td>rs17305016</td>
<td>4</td>
<td>86945264</td>
<td>ARHSAP24</td>
<td>0.402</td>
<td>7.61 x 10^{-6}</td>
<td>0.739 (0.647 – 0.844)</td>
</tr>
</tbody>
</table>

Abbreviations: ALPK1, α-kine 1; ARHGAP24, Rho GTPase activating protein 24; ARHGEF4, Rho guanine nucleotide exchange factor 4; chr, chromosome; 95%CI, 95% confidence interval; HR, hazards ratio; n.a., marker is not located within a gene

In Figure 2, the Kaplan-Meier curves are shown for the most significant SNP, rs2936519. Median PFS was 8.1 months (95%CI, 6.6 to 9.7 months) and 11.4 (95%CI, 10.4 to 12.4 months) for C/T and C/C genotypes, respectively. Only one patient was homozygous for the T-allele (this patient did contribute to the P-value, but is not shown in figure 2).

### Discussion

In this first analysis, three SNPs – that are in linkage – located at 8p23.1 showed a trend toward significance for association with PFS in ACC patients treated with CAPOX-B. The top three most significant SNPs are not located within a known gene, but are approximately 20 kbp downstream of the AGPAT5 gene (1-acylglycerol-3-phosphate O-acyltransferase 5, also known as lysophosphatidic acid acyltransferase, epsilon).

This gene encodes an integral membrane protein that converts lysophosphatidic acid (LPA) to phosphatidic acid (PA), the second step in de novo phospholipid biosynthesis, the major constituent of the cell membrane (http://www.ncbi.nlm.nih.gov). Additionally, LPA is a potent mitogen that has been linked to the development and progression of breast cancer. When these markers have been fully evaluated (i.e., functional analysis on gene-function or gene-expression level) and their associations have been confirmed in an independent cohort, they could be used to optimize selection of ACC patients for CAPOX-B treatment.

This is the first pharmacogenomic genome-wide study on the efficacy of palliative therapy for ACC. Unfortunately, the associations between the SNPs and PFS did not reach formal statistical significance at the 5 x 10^{-8} level, but a trend toward significance was found for 3 SNPs. This could be the result of insufficient power, possibly in combination with very stringent correction for multiple testing. Otherwise, the results might simply be false positive findings based upon the large number of statistical tests. There are 47 more patients that have to be genotyped, and were therefore not
Since all patients in our study were treated with CAPOX-B, no distinction could be made between prognostic (i.e. not related to treatment) and predictive (i.e. related to treatment) effects. If true significant effects would be found, the effects of the markers could therefore also be unrelated to therapy. However, it would be difficult to test whether the associations are predictive or prognostic, since a no-treatment control arm in the first-line treatment of ACC would be unethical. When the same associations would be found in a cohort of ACC patients that are treated with other agents as first-line therapy, the markers could then be regarded as prognostic rather than predictive. However, such a cohort is not feasible with fluoropyrimidines currently being the backbone of first-line ACC treatment.

The top three SNPs are in linkage, and are located near the AGPAT5 gene, which encodes a protein that converts LPA into PA, and is involved in phospholipid biosynthesis. It has to be elucidated whether these SNPs have an effect on the expression or function of this gene, or whether these SNPs are in linkage with a functional SNP in this gene. Possibly, fine-mapping or imputation in the region around the three significant SNPs could help finding the true causative SNP.

LPA has been linked to development and progression of breast cancer. Downstream of LPA receptor activation, the GTPase rho is activated. Two other genes with SNPs that are in the top 10 of most significant SNPs are possibly also involved in this signaling route (ARHGEF4 and ARHGAP24), suggesting that the LPA signaling pathway could be important for CAPOX-B efficacy or prognosis of ACC. Moreover, as phospholipids make up an essential component of the cell membrane, altered biosynthesis of phospholipids could have an effect on (tumor) cell division and therefore also efficacy of chemotherapy. However, such reasonings remain highly speculative, and the mechanism underlying the associations found in this study requires fundamental research.

Two SNPs in the top 10 are located in or near the gene encoding α-kinase 1 (ALPK1), which has been implicated in epithelial cell polarity and exocytic vesicular transport towards the apical plasma membrane. When other genetic information from the same genome-wide studies was included in a predictive model for human height, 45% of the variability could be explained. This strategy has also been applied for the risk of schizophrenia and bipolar disorder. Validation of such a predictive model in an independent cohort is very important because of the possibility of false positive findings due to the huge number of polymorphisms that are included, even though internal cross-validation can be used while developing the predictive model.

In conclusion, even though these results possibly identify a novel region that is associated with the efficacy of CAPOX-B, further analyses are required before firm conclusions can be made. A prediction model using the data from this study will probably better discriminate patients with long from short PFS than individual SNPs.
References

1. CBO guideline management of colorectal cancer (http://www.cbo.nl/product/richtlijnen/folder2002-1023121843/H_colonc_08.pdf)