Pharmacogenetics of advanced colorectal cancer treatment

Pander, J.

9

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 finacially 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).[1] 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%.[2,3] 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.[4] 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.[5] Regarding outcome of systemic therapy, genome-wide association studies have identified SNPs associated with muscoskeletal adverse reactions to aromatase inhibitors[6], treatment response for childhood acute lymphoblastic lymphoma[7] and pharmacokinetics of methotrexate[8]. All of these studies are based upon a case-control design with χ^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).[9] 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/m^2 on the first day, followed by 250 mg/m^2 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.

Genotyping and quality control

Germline DNA was isolated from peripheral white blood cells by the standard manual salting-out method. Genotyping was performed on Human OnmiExpress v12 BeadChip arrays containing 733,202 markers (Illumina, San Diego, CA, USA) using technical facilities at the Leiden Genome Technology Center (LGTC, Leiden, The Netherlands). Genotype calls were set using GenomeStudio software (Illumina). Patients with a call-rate of < 0.98 were excluded from further analysis. The following cut-off values were used to filter out incorrectly called genotypes: GenCall ≥ 0.85; ClusterSep ≥ 0.3; CallFreq > 0.85; AB T-mean 0.2 – 0.8, resulting in the exclusion of 3172 markers (0.43%).
Further quality control of the data was performed using R (http://www.r-project.org/). Ten patients (1.83%) were excluded based upon the sex check, and 16 patients (2.92%) were excluded based upon 4-dimensional multi-dimensional scaling (MDS) analysis to detect possible population stratification. In total, 26 patients were excluded, resulting in 521 evaluable patients.
For the markers, a minimum allele frequency for this analysis was set at 0.05, resulting in excluding 125,800 markers (17.2%). The call-rate cut-off per marker was set at 0.98, resulting in excluding 16,981 markers (2.3%). The distribution of the marker-missingness per chromosome showed no unexpected pattern. Hardy-Weinberg equilibrium (HWE) was evaluated per marker using a χ^2 goodness-of-fit statistic. Based upon the QQ-plot of observed P-values against expected P-values for HWE, 1168 markers (0.2%) were excluded with a HWE P-value of ≤ 5.0 x 10^{-7}. 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 (median(T_1,…,T_n)/0.675)^2 with T_1,…,T_n being the square roots of the χ^2

quantiles for the P-values of the markers. Formal significance for a marker was assumed for P < 5 x 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

Characteristic	
Age - year	
median	63.1
range	27.6 - 83.6
Sex - no (%)	
male	316 (60.7%)
female	205 (39.3%)
Arm - no (%)	
CAPOX-B	264 (50.7%)
CAPOX-B plus cetuximab	257 (49.3%)
Serum lactate dehydrogenase level - no (%)	
normal*	307 (58.9%)
above normal*	213 (40.9%)

* according to the cutoff values of each individual center

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; Table 2 and Figure 1), but did not reach the formal significance level of $P < 5 \times 10^{-8}$. These three SNPs were in linkage. None of ten most significant SNPs showed a significant interaction with treatment arm. The inflation factor for the analysis was 0.98, indicating that there was no population stratification or other bias in the analysis.

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

marker	chr	position	gene	allele frequency	P-value	allelic HR (95%CI)
rs2936519	8	6626650	*n.a.*	0.104	1.24×10^{-7}	0.545 (0.435 – 0.682)
rs2912024	8	6626309	*n.a.*	0.105	1.38×10^{-7}	0.547 (0.437 – 0.685)
rs2978931	8	6625491	*n.a.*	0.101	6.75×10^{-7}	0.561 (0.447 – 0.705)
rs4850159	2	131442241	*ARHGEF4*	0.136	2.58×10^{-6}	0.627 (0.516 – 0.762)
rs6734725	2	46751074	*n.a.*	0.350	2.99×10^{-6}	0.713 (0.619 – 0.822)
rs17688362	18	39999678	*n.a.*	0.185	4.68×10^{-6}	0.657 (0.548 – 0.786)
rs17444829	4	113593423	*n.a.*	0.133	5.49×10^{-6}	1.556 (1.285 – 1.884)
rs11730442	4	113581912	*ALPK1*	0.132	6.62×10^{-6}	1.554 (1.282 – 1.884)
rs10089490	8	92317629	*n.a.*	0.061	7.20×10^{-6}	1.849 (1.413 – 2.420)
rs17395916	4	86945264	*ARHGAP24*	0.402	7.61×10^{-6}	0.739 (0.647 – 0.844)

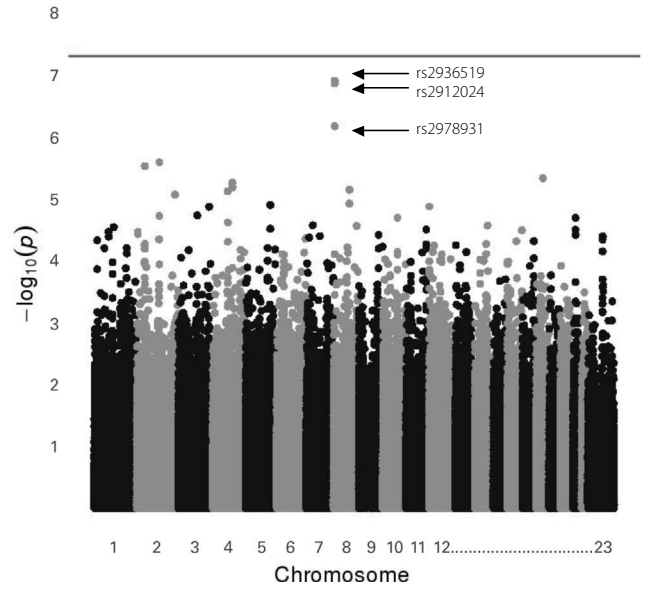
Abbreviations: ALPK1, α-kinase 1; ARHGAP24, Rho GTPase activating protein 24; ARHGEF4, Rho guanine nucleotide exchange factor 4; chr, chromosome; 95%CI, 95% confidence interval; HR, hazards ratio; n.a., marker is not located within a gene

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).

Figure 1 Manhattan plot of -$\log_{10}$(P-values) from the Cox-proportional hazards model adjusted for age, gender and treatment arm

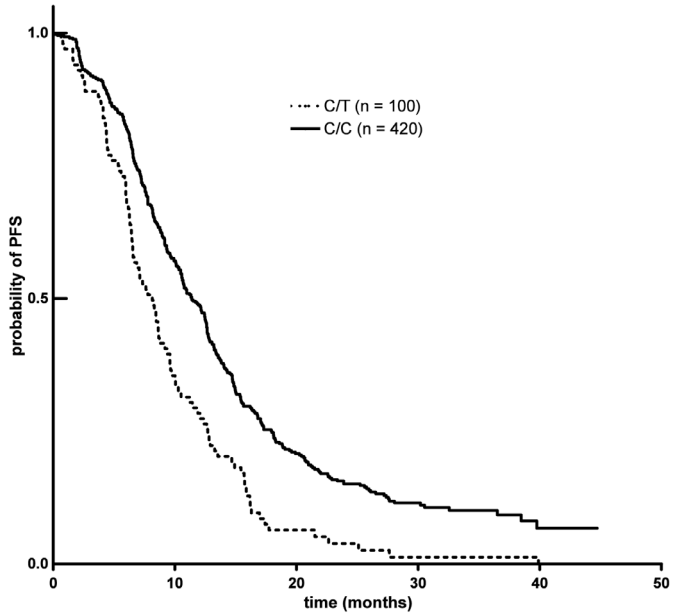


The horizontal line represents the formal genome-wide significance level of 5×10^{-8}

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.[10] 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 fist 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 5×10^{-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

Figure 2 Kaplan Meier curves of PFS for the rs2936519 SNP



included in this analysis. Possibly, including these patients would increase power enough for the results to become significant.

On the other hand, it is unlikely that a complex phenotype such as drug response depends on only a handful single SNPs. As with other complex traits, a series of polymorphisms could contribute to the phenotype. A predictive model can be built to assess the combined contribution of SNPs to the phenotype – which will be further evaluated in our study. That this could be a feasible approach is illustrated by a recent example on human height, in which it initially seemed that only a few SNPs were associated with this phenotype. However, the explained variability was only ~5%.[11-13] 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.[14] 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.[15]

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.[16] 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.[17] It is not clear how ALPK1 could be linked to (colorectal) cancer or the mechanism of action of CAPOX-B. Also, the consequence of the other SNPs in the top 10 is unknown.

For this study, we included patients who were treated with CAPOX-B as well as patients who were treated with CAPOX-B plus cetuximab. In our analysis, treatment allocation was included as a covariate. It is unlikely that the effects of the SNPs are linked to the efficacy of cetuximab, since none of the polymorphisms showed significant interaction with the treatment arm.

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/rl_colonc_08.pdf)
2. Punt CJ. New options and old dilemmas in the treatment of patients with advanced colorectal cancer. Ann Oncol 2004;15:1453-9.
3. Saltz LB, Clarke S, az-Rubio E, *et al.* Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. J Clin Oncol 2008; 26:2013-9.
4. Walther A, Johnstone E, Swanton C, *et al.* Genetic prognostic and predictive markers in colorectal cancer. Nat Rev Cancer 2009;9:489-99.
5. Stadler ZK, Thom P, Robson ME, *et al.* Genome-wide association studies of cancer. J Clin Oncol 2010; 28:4255-67.
6. Ingle JN, Schaid DJ, Goss PE, *et al.* Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. J Clin Oncol 2010;28:4674-82.
7. Yang JJ, Cheng C, Yang W, *et al.* Genome-wide interrogation of germline genetic variation associated with treatment response in childhood acute lymphoblastic leukemia. JAMA 2009;301:393-403.
8. Treviño LR, Shimasaki N, Yang W, *et al.* Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. J Clin Oncol 2009; 27:5972-8.
9. Tol J, Koopman M, Cats A, *et al.* Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. N Engl J Med 2009;360:563-72.
10. Panupinthu N, Lee HY, Mills GB. Lysophosphatidic acid production and action: critical new players in breast cancer initiation and progression. Br J Cancer 2010;102:941-6.
11. Gudbjartsson DF, Walters GB, Thorleifsson G, *et al.* Many sequence variants affecting diversity of adult human height. Nat Genet 2008;40:609-15.
12. Lettre G, Jackson AU, Gieger C, *et al.* Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 2008;40:584-91.
13. Weedon MN, Lango H, Lindgren CM, *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 2008;40:575-83.
14. Purcell SM, Wray NR, Stone JL, *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009;460:748-52.
15. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
16. Moolenaar WH, van Meeteren LA, Giepmans BN. The ins and outs of lysophosphatidic acid signaling. Bioessays 2004;26:870-81.
17. Heine M, Cramm-Behrens CI, Ansari A, *et al.* Alpha-kinase 1, a new component in apical protein transport. J Biol Chem 2005;280:25637-43.