Pharmacogenetic interaction analysis for the efficacy of systemic treatment in metastatic colorectal cancer

J. Pander\(^1\), J. A. M. Wessels\(^1\), H. Gelderblom\(^2\), T. van der Straaten\(^1\), C. J. A. Punt\(^3\) & H.-J. Guchelaar\(^1\)*

Departments of \(^1\)Clinical Pharmacy & Toxicology; \(^2\)Clinical Oncology, Leiden University Medical Center, Leiden; \(^3\)Department of Medical Oncology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Received 14 May 2010; revised 6 August 2010; accepted 17 August 2010

Background: Pharmacogenetic markers related to drug metabolism and mechanisms of action could help to better select patients with metastatic colorectal cancer (mCRC) for treatment. Genetic interaction analysis is used as a rational tool to study the contribution of polygenic variation in relation to drug response.

Patients and methods: A selection of 17 polymorphisms in genes encoding drug targets, pathway molecules and detoxification enzymes was analyzed in 279 previously untreated mCRC patients treated with capecitabine, oxaliplatin and bevacizumab (CAPOX-B). Multifactor dimensionality reduction analysis was used to identify a genetic interaction profile for progression-free survival (PFS).

Results: Median PFS was 10.9 [95% confidence interval (CI) 9.4–12.4] months. A genetic interaction profile consisting of the \(TYMS\) enhancer region and \(VEGF\) +405G>C polymorphisms was significantly associated with PFS. Median PFS was 13.3 (95% CI 11.4–15.3) and 9.7 (95% CI 7.6–11.8) months for the beneficial and unfavorable genetic profiles, respectively, corresponding to a hazards ratio for PFS of 1.58 (95% CI 1.14–2.19). None of the studied polymorphisms were individually associated with PFS.

Conclusions: Our results support a genetic interaction between the \(TYMS\) enhancer region and \(VEGF\) +405G>C polymorphisms as a predictor of the efficacy of CAPOX-B in mCRC patients.

Key words: metastatic colorectal cancer, multifactor dimensionality reduction, pharmacogenetics, predictive factor, systemic chemotherapy

Introduction

The combination of a fluoropyrimidine, such as 5-fluorouracil (5-FU) or capecitabine, oxaliplatin and the vascular endothelial growth factor (VEGF) blocking antibody bevacizumab (CAPOX-B) is a frequently used standard first-line treatment strategy for metastatic colorectal cancer (mCRC) [1, 2]. However, since not all patients respond to this regimen, better criteria to select patients for this treatment are warranted. For this purpose, pharmacogenetic studies have been carried out with germline polymorphisms in genes that encode metabolic enzymes and drug targets (Table 1). However, the findings from these studies are not consistent [43]. As a result, none of these polymorphisms are currently used in general practice to identify patients with an increased chance of response.

An explanation for these results could be that current analytical methods ignore or underestimate the complexity underlying drug response. Drug response involves many different proteins, such as therapeutic targets, molecules in the signaling pathway, metabolic enzymes or drug transporters. It may therefore be likely that the impact of polymorphisms in the corresponding genes exert their influence only in the presence of other polymorphisms. This concept is known as nonlinear interaction or epistasis [44]. Studying the interaction between polymorphisms could therefore provide more reliable information compared with separate analyses of associations between individual polymorphisms and response [45]. The resulting information can be transformed into genetic profiles that may have a prognostic and/or predictive value for mCRC patients.

The multifactor dimensionality reduction (MDR) methodology has been developed to study nonlinear patterns of interactions between genetic profiles and drug response [46]. In this study, we applied genetic interaction analysis using the MDR method to evaluate interaction between candidate polymorphisms in relation to the efficacy of CAPOX-B as first-line treatment in mCRC patients.

Materials and methods

Study population

Blood samples were collected from 279 of 368 previously untreated mCRC patients who were treated with CAPOX-B in the control arm of the
multicenter, prospective randomized phase III CAIRO2 study of the Dutch Colorectal Cancer Group [1]. Capecitabine 1000 mg/m² (increased to 1250 mg/m² from cycle 7) was administered orally twice daily on days 1–14 of each 3-week treatment cycle. Oxaliplatin 130 mg/m² (maximum of six cycles) and bevacizumab 7.5 mg/kg were administered i.v. on day 1 of each treatment cycle. Treatment was continued until disease progression, death or unacceptable toxicity, whichever occurred first. Patient eligibility criteria and further details of the study have been previously described [1]. The collection of a peripheral blood sample for pharmacogenetic research was prespecified in the study protocol and required additional written informed consent. The protocol was approved by the local institutional review boards of all participating centers.

Patients in the experimental cetuximab-containing study arm of the CAIRO2 study were not included in this pharmacogenetic study since the addition of cetuximab resulted in a decreased progression-free survival (PFS), the primary end point of the study [1].

**genotyping**

The studied genetic polymorphisms are shown in Table 1. These polymorphisms were selected primarily on the basis of the pharmacokinetics and pharmacodynamics of capecitabine, oxaliplatin and bevacizumab and on the known functional effects at the protein level [47]. Moreover, these polymorphisms have been included in previous pharmacogenetic association studies of 5-FU, capecitabine or oxaliplatin in mCRC [3–14, 20–28, 48]. Since results of only two pharmacogenetic studies for bevacizumab have been reported [33, 34], polymorphisms in VEGF and its receptor (kinase domain receptor) were selected [49]. Germline DNA was isolated from peripheral white blood cells by the standard manual salting-out method. Genotyping was carried out on a Biomark system (Fluidigm, South San Francisco, CA) according to the protocol provided by the manufacturer using predesigned TaqMan assays (Applied Biosystems, Foster City, CA).

The polymorphisms in the thymidylate synthase enhancer region (TSER) in the promoter of the TYMS gene (two or three 28-bp repeats including the C>C G polymorphism in the third repeat; TYMS-TSER) were analyzed by direct sequencing. The genotype was expressed as noncarriage of the 3G allele (2/2, 2/3C and 3C/3C genotypes) versus carriage of the 3G allele (3G/3G or 3G/2C genotypes) versus carriage of the 3G allele (3G/3G or 3G/2C genotypes). Fragments from all samples were sequenced in both directions.

Each assay was conducted with 10% duplicates, with water as negative control. The overall call rate was 0.948 (0.803–0.989) and none of the genotypes were deviated Hardy–Weinberg equilibrium ($P > 0.01$).

**statistical analysis**

Genotypes that are individually associated with drug response will usually end up in the best genetic profile in the genetic interaction analysis without providing substantial information gain. Therefore, the association between each individual polymorphism (treated as an ordinal variable, representing an additive model) and PFS as dependent variable was tested using a Cox proportional hazards model including serum LDH, age and gender as covariates. Polymorphisms significantly associated with PFS ($P < 0.05$) were
excluded from the subsequent interaction analysis but would be introduced in the final multivariate analysis (see below). Also, haplotypes disturb the selection of the best genetic profile because of over fitting the data due to the number of possible haplotype combinations and were therefore also not used in the interaction analysis. No haplotype was individually associated with PFS in our study (data not shown).

To study interaction between the polymorphisms in relation to response, the MDR software was used (version 2.0 beta 6; available on http://sourceforge.net/projects/mdr/) [46]. The software requires a complete dataset with no missing data. Therefore, missing data for polymorphisms with ≤5% missing data were imputed by genotypes based upon the genotype frequency of the polymorphism, taking the distribution of other polymorphisms in the same gene into account. Missing data for polymorphisms with >5% missing data were considered as a separate ‘missing genotype group’ in the genetic interaction analysis. If the genetic interaction analysis resulted in a combination consisting of a genotype with a missing genotype group, the procedure was repeated without this group and results were compared with the initial results.

Our study is designed to identify a subgroup of patients with increased PFS. The median PFS in our study population was 10.9 months. However, it is assumed that the patients with beneficial genetic profiles have a PFS much longer than the median, whereas patients with unfavorable genetic profiles have PFS much shorter than the median. We therefore included patients in the shortest and longest quartiles for PFS in the genetic interaction analysis in order to increase discriminating power [50–52]. The entire cohort was used in the final analysis of the genetic profile (see below).

Sensitivity analysis showed that when the patients with censored data before the 75% quartile cut-off point were included in the longest quartile or when tertiles were used instead of quartiles, results remained unchanged (data not shown), indicating that our choices regarding censoring and enrichment do not influence the results of the study.

In the genetic interaction analysis, the ratio between patients in the shortest quartile to patients in the longest quartile for each genotype combination is evaluated. Combinations with more patients in shortest quartile than in the longest quartile are considered high chance of short PFS and vice versa. This procedure was carried out across 10-fold cross-validation samples to avoid over fitting and was repeated for all possible combinations of two up to four polymorphisms. The genotype combination with the highest accuracy (fraction of correctly classified patients) in the validation sample was considered the combination that best predicts PFS and was selected for further analysis. A P value for the statistical significance of the accuracy was obtained using 1000-fold permutation testing (software available on https://sourceforge.net/projects/mdr/files/mdrpt/).

The genotype combination with the highest accuracy in the validation sample was recoded into a genetic profile predictive for PFS. This genetic profile was subsequently used for all 279 patients in the CAIRO2 study from whom a blood sample was available, including the patients from the intermediate PFS group, to estimate survival curves using the Kaplan–Meier method. The difference in PFS from the beneficial genetic profile versus the unfavorable genetic profile was estimated using the log-rank test. A Cox proportional hazards model including the genetic profile, age, gender, prior adjuvant chemotherapy (yes versus no), number of affected organs (1 versus >1), serum LDH and any polymorphisms that were individually associated with PFS was used to compute the adjusted hazards ratio (HR) and 95% confidence interval (95% CI). Given the exploratory nature of this study, no adjustment for multiple testing was carried out, and a P value of <0.05 was considered significant. The Kaplan–Meier and Cox proportional hazards analyses were carried out using SPSS version 17.0 (SPSS, Chicago, IL).

**Results**

At the time of analysis, the primary end point of PFS was reached in 225 of 279 eligible patients (80.6%). Median PFS was 10.9 months (95% CI 9.4–12.4 months). Two patients were censored in the shortest quartile and were excluded from the genetic interaction analysis since the actual PFS of these patients was unknown. Censored events in the longest quartile were not excluded since PFS for these patients was at least longer than the 75% quartile cut-off point. The shortest and longest quartiles for PFS were below 6.7 and above 15.5 months, respectively, each consisting of 70 patients.

None of the genetic polymorphisms were individually associated with PFS in the Cox proportional hazards analysis (Table 2). Therefore, all polymorphisms were included in the genetic interaction analysis with PFS.

The combination of the TYMS-TSER and VEGF +405G>C had the highest accuracy of 0.650 (P = 0.027, 1000-fold permutation testing; 0.624 after exclusion of missing data), meaning that 65% of the patients were correctly classified according to the genetic profile (Figure 1A). The distribution of patients in the shortest and longest PFS quartiles for the combination of TYMS-TSER and VEGF +405G>C genotypes is shown in Figure 1B. All other combinations of two, three and four polymorphisms each resulted in lower accuracies in the

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Allelic HR*</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTHFR 677C&gt;T</td>
<td>1.00</td>
<td>0.81–1.23</td>
<td>0.991</td>
</tr>
<tr>
<td>MTHFR 1298A&gt;C</td>
<td>0.91</td>
<td>0.74–1.13</td>
<td>0.393</td>
</tr>
<tr>
<td>TYMS 1494 ±6 bp</td>
<td>1.10</td>
<td>0.87–1.40</td>
<td>0.410</td>
</tr>
<tr>
<td>TYMS VNTR 2302G</td>
<td>1.02</td>
<td>0.77–1.36</td>
<td>0.884</td>
</tr>
<tr>
<td>ERCC1 1496C&gt;T</td>
<td>1.12</td>
<td>0.92–1.37</td>
<td>0.243</td>
</tr>
<tr>
<td>ERCC2 499C&gt;A</td>
<td>1.15</td>
<td>0.94–1.40</td>
<td>0.185</td>
</tr>
<tr>
<td>ERCC2 2251A&gt;C</td>
<td>1.00</td>
<td>0.82–1.21</td>
<td>0.968</td>
</tr>
<tr>
<td>ERCC2 965G&gt;A</td>
<td>0.80</td>
<td>0.63–1.01</td>
<td>0.058</td>
</tr>
<tr>
<td>XRCC1 1301G&gt;A</td>
<td>0.98</td>
<td>0.81–1.18</td>
<td>0.811</td>
</tr>
<tr>
<td>GSTP1 313A&gt;G</td>
<td>0.98</td>
<td>0.81–1.19</td>
<td>0.837</td>
</tr>
<tr>
<td>KDR 1719A&gt;T</td>
<td>1.08</td>
<td>0.88–1.33</td>
<td>0.465</td>
</tr>
<tr>
<td>KDR −604T&gt;C</td>
<td>1.03</td>
<td>0.86–1.24</td>
<td>0.738</td>
</tr>
<tr>
<td>VEGF = 1154G&gt;A</td>
<td>1.09</td>
<td>0.90–1.33</td>
<td>0.381</td>
</tr>
<tr>
<td>VEGF 405G&gt;C</td>
<td>0.97</td>
<td>0.81–1.18</td>
<td>0.785</td>
</tr>
<tr>
<td>VEGF 936C&gt;T</td>
<td>0.98</td>
<td>0.74–1.29</td>
<td>0.889</td>
</tr>
<tr>
<td>VEGF = 2578C&gt;A</td>
<td>1.03</td>
<td>0.86–1.23</td>
<td>0.763</td>
</tr>
<tr>
<td>VEGF = 460C&gt;T</td>
<td>1.00</td>
<td>0.84–1.20</td>
<td>0.990</td>
</tr>
</tbody>
</table>

*HR, 95% CI and P values were calculated for each polymorphism using a Cox proportional hazards model with age, gender and serum LDH as covariates.

CI, confidence intervals; ERCC1, excision repair cross-complementing group 1; ERCC2, excision repair cross-complementing group 2; GSTP1, glutathione s-transferase pi 1; HR, hazard ratios; KDR, kinase domain receptor (= vascular endothelial growth factor receptor 2); mCRC, metastatic colorectal cancer; MTHFR, methylentetrahydrofolate reductase; TYMS, thymidylate synthase; VEGF, vascular endothelial growth factor A; VNTR, variable number of tandem repeats; XRCC1, X-ray cross-complementing group 1.
genetic interaction analysis and were therefore not considered for further evaluation.

When all 246 patients with complete genotype data were used, 137 and 109 patients were in the beneficial and unfavorable profiles for PFS, respectively. In Figure 2, the frequency distribution of the genetic profile across the four quartiles for PFS is shown. Interestingly, the frequency of the unfavorable profile decreases for every quartile, even for the two middle quartiles ($P < 0.001$, $\chi^2$ test for trend).

The PFS curves for the genetic profile for all patients are shown in Figure 3. The median PFS was 13.3 (95% CI 11.4–15.3) and 9.7 (95% CI 7.6–11.8) months for the beneficial and unfavorable profiles, respectively ($P < 0.001$, log-rank test).

In the multivariate Cox proportional hazards model, including age, gender, prior adjuvant chemotherapy, number of affected organs and serum LDH, the HR for the genetic profile for PFS was 1.58 (95% CI 1.14–2.19, $P = 0.006$).

**discussion**

We showed that a genetic interaction profile consisting of the $VEGF +405G>C$ and $TYMS$-TSER polymorphisms correlates with PFS in mCRC patients treated with CAPOX-B. This approach provides a novel way to use pharmacogenetic variation to individualize treatment since individual polymorphisms were not associated with PFS.

To exclude profound individual associations with PFS that could interfere with the genetic interaction analysis, we first tested for associations of the individual polymorphisms with PFS. No significant associations were detected, analogous to the absence of associations for other individual molecular markers in mCRC [53].

The genetic interaction analysis takes the complexity of interacting polymorphisms in genes encoding drug targets, metabolic enzymes and detoxification enzymes into account. Our study shows that—in mCRC patients treated with first-line CAPOX-B—the $TYMS$-TSER polymorphism and the $VEGF +405G>C$ polymorphisms are dependent of each other in their impact on PFS.

$VEGF$ is the natural ligand for the $VEGF$ receptor, through which it induces angiogenesis. Bevacizumab neutralizes $VEGF$, resulting in decreased tumor angiogenesis, which in turn affects intratumoral hypoxia, nutrition status and/or disposition of concurrent chemotherapy [54]. The functional consequence of the $VEGF +405G>C$ polymorphism remains to be elucidated. One *in vitro* study reported increased $VEGF$ release by...
lipopolysaccharide-stimulated peripheral monoclonal blood cells with the VEGF +405G allele [38], but another study showed that VEGF serum levels were highest for healthy volunteers with the VEGF +405CC genotype [39]. The fluorodeoxyuridine monophosphate metabolite of capecitabine inhibits the TYMS enzyme and thereby induces DNA damage [55]. Previous in vitro experiments indicated that the TYMS-TSER 3G allele results in higher expression of TYMS [18, 19]. The finding by Marcuello et al. [9] that the TYMS-TSER 3G allele is associated with decreased efficacy of 5-FU-based chemotherapy in mCRC patients is therefore only present for VEGF +405C allele carriers in our study. However, further fundamental research should be undertaken to understand the exact biological mechanism of the genetic profile with regard to the efficacy of CAPOX-B.

For the genetic interaction analysis, the PFS end point initially was converted into a binary outcome. We assumed that the patients with beneficial genetic profiles have a PFS much longer than the median, whereas those with unfavorable genetic profiles have PFS much shorter than the median. By using the shortest and longest quartiles for PFS, we anticipated sufficient discriminating power, while keeping the groups reasonably large. This concept of enrichment is an accepted method in genetics [50–52]. Even though our choice of cut-off remains arbitrary, our sensitivity analysis showed similar results for the genetic profile, indicating that the results were not significantly influenced by the choice of quartiles for PFS.

There are some limitations to our findings. First, the genetic interaction analysis relies on data mining to identify the best model—or genetic profile—to fit the data [46], potentially leading to over fitting, with optimal results only in the initial test cohort. Although we used cross-validation to correct for over fitting, the genetic profile should be validated in an independent cohort to confirm our present finding and to assess its clinical utility [56].

Finally, without an untreated control group, it remains unclear whether the genetic profile is predictive for response to CAPOX-B in mCRC patients or prognostic for mCRC outcome regardless of treatment [57]. Also, given the many available salvage treatments, which were not part of the study protocol and were therefore not controlled, the assessment of a potential prognostic role was not feasible. However, the fact that both polymorphisms of the genetic profile are in the targets of two of the drugs suggests that the profile is predictive rather than prognostic.

In conclusion, we demonstrated a significant correlation between a genetic profile consisting of the TYMS-enhancer region and VEGF +405G>C polymorphisms and improved PFS. This genetic profile is a novel marker that may identify a subgroup of mCRC patients with increased probability of benefit to CAPOX-B. To our knowledge, this is the first study to explore the interaction between polymorphisms in relation to the efficacy of cancer chemotherapy. Testing for the interaction between polymorphisms is probably more rational than testing of each individual polymorphism since drug response is a complex phenomenon. If confirmed in independent studies, our results provide a novel tool to better select cancer patients for potentially toxic and expensive treatments.

**disclosure**

The authors declare no conflict of interest.
references


