Dissertation

Predictive and prognostic effects of germline polymorphisms in angiogenesis and cancer stem cell pathway genes and colorectal cancer

Armin Gerger, M.D., Priv.-Doz., Ass.-Prof.

Doktoratsstudium der Medizinischen Wissenschaft (O 790)

Division of Clinical Oncology
Research Unit Genetic Epidemiology and Pharmacogenetics
Medical University of Graz
Auenbruggerplatz 15, 8036 Graz, Austria

Under supervision of
Wilfried Renner, PhD, Univ.-Doz., Assoc.-Prof.
Hellmut Samonigg, M.D., Full Prof.
Michael Speicher, M.D., Full Prof.
Heinz-Josef Lenz, M.D., Full Prof.
(Cancer Genetics Laboratory, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, USA)

Graz, 20.2.2012

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 20.2.2012 *Unterschrift*

Danksagung

Mein ganz besonderer Dank geht an meine geliebte Ehefrau, Michaela, die mir mit unermüdlicher Geduld zur Seite gestanden hat.

Table of contents

Abstract…………………………………………………………………………………………………… 5

Background………………………………………………………………………………………………… 8

Introduction……………………………………………………………………………………………… 10

Patients and methods………………………………………………………………………………… 13

Eligible patients…………………………………………………………………………… 13

Candidate polymorphisms………………………………………………………………… 15

Genotyping………………………………………………………………………………… 15

Statistical analysis………………………………………………………………………… 16

Results……………………………………………………………………………………………………… 18

Metastatic setting………………………………………………………………………… 18

Adjuvant setting…………………………………………………………………………… 20

Discussion………………………………………………………………………………………………… 23

References………………………………………………………………………………………………… 31

Figures……………………………………………………………………………………………………… 36

Tables……………………………………………………………………………………………………… 41

Abstract

There is substantial germline genetic variability within angiogenesis and cancer stem cell (CSC) pathway genes, thereby causing inter-individual differences in angiogenic capacity and resistance to anti-angiogenesis therapy and tumor-recurrence capacity. In this study we investigated (a) germline polymorphisms in genes involved in VEGF-dependent and -independent angiogenesis pathways to predict clinical outcome and tumor response in metastatic colorectal cancer patients (mCRC) treated with bevacizumab (BV) and oxaliplatin-based chemotherapy and (b) germline polymorphisms in a comprehensive panel of CSC genes to predict time to tumor recurrence (TTR) in patients with stage III and high-risk stage II colon cancer. A total of 132 patients with mCRC treated with first-line BV and FOLFOX or XELOX and a total of 234 patients in adjuvant setting treated with 5-fluorouracil-based chemotherapy were included in this study. Whole blood samples were analyzed for germline polymorphisms in angiogenesis and CSC genes by PCR-RFLP or direct DNA-sequencing. In the metastatic setting the minor alleles of EGF rs444903 A>G and IGF-1 rs6220 A>G were associated with increased overall survival (OS) and remained significant in multivariate COX regression analysis (HR 0.52; 95%CI 0.31-0.87; adjusted-*P*=0.012 and HR 0.60; 95%CI 0.36-0.99; adjusted-*P*=0.046, respectively). CXCR1 rs2234671 G>C, CXCR2 rs2230054 T>C, EGFR rs2227983 G>A and VEGFR-2 rs2305948 C>T predicted tumor response, with CXCR1 rs2234671 G>C remaining significant in multiple testing (*P*act=0.003). In the adjuvant setting the minor alleles of CD44 rs8193 C>T, ALCAM rs1157 G>A and LGR5 rs17109924 T>C were significantly associated with increased TTR (9.4 vs 5.4 years; HR, 0.51; 95%CI, 0.35-0.93; *P*=0.022; 11.3 vs 5.7 years; HR, 0.56; 95%CI, 0.33-0.94; *P*=0.024 and 10.7 vs 5.7 years; HR, 0.33; 95%CI, 0.12-0.90; *P*=0.023, respectively) and remained significant in the multivariate analysis stratified by ethnicity. In recursive partitioning, a specific gene variant profile including LGR5 rs17109924, CD44 rs8193 and ALDH1A1 rs1342024 represented a high-risk subgroup with a median TTR of 1.7 years (HR, 6.71, 95%CI, 2.71-16.63, *P*<0.001). In this study we identified common germline variants in VEGF-dependent and -independent angiogenesis genes predicting clinical outcome and tumor response in patients with mCRC receiving first-line BV and oxaliplatin-based chemotherapy. Furthermore, we identified common germline variants in colon CSC genes as independent prognostic markers for stage III and high-risk stage II colon cancer patients.

Zusammenfassung

Es gibt eine große genetische Variabilität in Angiogenese- und Krebsstammzell-Pathways die zu einem inter-individuellen Unterschied in der Angiogenese Kapazität und Resistenz auf Anti-Angiogenese Therapie und Wiederauftreten des Tumors führen können. In dieser Studie untersuchten wir (a) Keimbahnpolymorphismen in VEGF-abhängigen und -unabhängigen Angiogenese Genen um die Prognose und das Tumoransprechen bei PatientInnen mit metastasiertem Kolorektalkarzinom und einer Therapie mit Bevacizumab und Oxaliplatin-basierender Chemotherapie vorherzusagen und (b) Keimbahnpolymorphismen in einem umfassenden Set aus Krebsstammzell-Genen um die Zeit bis zum Wiederauftreten des Tumors bei PatientInnen mit Stadium III and Hochrisiko-Stadium II Kolonkarzinom vorherzusagen. Wir haben insgesamt 132 PatientInnen, die als Erstlinientherapie im metastasierten Setting Bevacizumab und FOLFOX oder XELOX erhalten haben und insgesamt 234 PatientInnen in adjuvanten Stadium III und Hochrisiko-Stadium II, die mit 5-FU basierter Chemotherapie behandelt wurden in diese Studie inkludiert. Keimbahn-DNA wurde aus Blutproben isoliert und mittels PCR-RFLP oder direktem Sequenzieren genotypisiert. Im metastasiertem Setting war das seltene Allel von EGF rs444903 A>G und IGF-1 rs6220 A>G mit verbessertem Gesamtüberleben assoziiert und blieb in der multivariaten COX Analyse signifikant (HR 0.52; 95%CI 0.31-0.87; adjustiertes-*P*=0.012 and HR 0.60; 95%CI 0.36-0.99; adjustiertes-*P*=0.046). CXCR1 rs2234671 G>C, CXCR2 rs2230054 T>C, EGFR rs2227983 G>A und VEGFR-2 rs2305948 C>T konnten das Tumoransprechen vorhersagen, und CXCR1 rs2234671 G>C blieb bei multiplem Testen signifikant (*P*act=0.003). In der adjuvanten Situation war das seltene Allel von CD44 rs8193 C>T, ALCAM rs1157 G>A und LGR5 rs17109924 T>C signifikant mit der Zeit bis Wiederauftreten des Tumors assoziiert (9.4 vs 5.4 Jahre; HR, 0.51; 95%CI, 0.35-0.93; *P*=0.022; 11.3 vs 5.7 Jahre; HR, 0.56; 95%CI, 0.33-0.94; *P*=0.024 und 10.7 vs 5.7 Jahre; HR, 0.33; 95%CI, 0.12-0.90; *P*=0.023) und blieb in der multivariaten Analyse signifikant. Bei rekursiver Partitionierung repräsentierte ein spezifisches Genvarianten-Profil inklusive LGR5 rs17109924, CD44 rs8193 und ALDH1A1 rs1342024 eine Hochrisiko-Subgruppe mit einer medianen Zeit bis Wiederauftreten des Tumors von 1.7 Jahren (HR, 6.71, 95%CI, 2.71-16.63, *P*<0.001). In dieser Studie konnten wir erstmals Biomarker identifizieren, die Prognose und Tumoransprechen bei PatientInnen mit Kolorektalkarzinom und einer Therapie mit Bevacizumab und Oxaliplatin-basierender Chemotherapie vorhersagen können. Des

Weiteren konnten wir Genvarianten in Krebsstammzellen identifizieren, die unabhängige prognostische Marker für die Zeit des Wiederauftretens des Tumors bei PatientInnen mit Kolonkarzinom im Stadium III und Hochrisiko-Stadium II darstellen.

Background

Predictive biomarkers aim to identify patients who are likely to benefit from a specific therapy. In contrast, prognostic biomarkers are associated with the disease outcome and have significant application in the identification of patients at high risk for tumor recurrence.(1)

Colorectal cancer is a clinically heterogeneous disease, and it is generally accepted that the different clinical courses of patients with histologically similar tumors are due to molecular differences of the tumor and the host. Therefore, detailed molecular profiling could yield prognostic and predictive information. The continuous flow of molecular genetic information has cautiously started integrating into clinical practice changing the future landscape of the clinical management of colorectal cancer.(2) The TNM staging system and conventional clinicopathologic factors have led clinical decisions for decades. However, current staging methods remain suboptimal. Patients with colorectal cancer who are at low risk of tumor recurrence could be spared the toxicity of systemic treatment if clearly distinguished, while others at high risk of recurrence could get maximal benefit if therapy matched the genetic profile.(3) Areas of cancer research such as genetic polymorphisms keep on producing promising data which is believed to refine future preventive and early intervention strategies on an individual basis.(4)

The complexity of the metastatic process has made it difficult to gain full understanding of the origins of this most lethal aspect of cancer. Many factors probably have an important role, including somatic mutation, epigenetic modulations, interactions with normal stroma, and environmental stimuli.(5) Additionally, recent evidence implies a significant role for germline polymorphisms in colorectal cancer progression and chemoresistance.(6) The existence of inherited factors has potentially significant implications for models of metastasis, clinical outcome, chemoresistance and the development of tailored therapy. Further investigations into the inherited components of metastasis and chemoresistance might help resolve many of the questions that remain about colorectal cancer. Angiogenesis and cancer stem cell pathways play a key role in the metastatic process and chemoresistance. Genetic translational studies, which lead to a true association, are

expected to increase understanding of colorectal cancer progression and predict colorectal cancer outcome and chemoresistance.(7)

The impact of germline polymorphisms on cancer outcome is a burgeoning field of research.(8) Genetic polymorphisms are inherited variations in germline DNA that occur commonly. Germline DNA sequence variations may alter protein production and/or activity, thereby causing inter-individual differences (eg, for metastatic capacity and/or chemoresistance). With the completion of the human genome mapping project, there have been both increasing ability and interest in exploring common genetic variations as potential prognostic and/or predictive factors. Genotyping is the physical process of identifying which genetic variants are carried by a patient. Germline DNA may be obtained from a diverse array of biologic sources including leukocytes from blood and even adjacent normal tissue from the same organ as the cancer, rendering germline evaluation a practical and attractive field of research and one that can be translated into the clinical setting with relative ease. The stability of genomic DNA allows for easy handling and processing. In addition, SNP analysis is relatively robust between platforms and laboratories, which help reduce inter-laboratory variation.(9-11)

The goals of this translational research project were twofold:

1) **Investigation of germline polymorphisms in genes involved in VEGF-dependent and -independent angiogenesis pathways to predict clinical outcome and tumor response in metastatic colorectal cancer patients (mCRC) treated with bevacizumab (BV) and oxaliplatin-based chemotherapy.**
2) **Investigation of germline polymorphisms in a comprehensive panel of CSC genes to predict time to tumor recurrence (TTR) in patients with stage III and high-risk stage II colon cancer.**

Introduction

Angiogenesis is a universal requirement for the growth of solid tumors beyond the limits of oxygen diffusion from the existing vasculature.(12, 13) Inhibition of angiogenesis has proven to be beneficial in multiple types of malignancies, including colon cancer.(14) VEGF-A is one of the major regulators of both physiological and pathological angiogenesis. Rapid proliferation of tumor cells and poor blood flow suggest a hypoxia-conducive environment in different areas of tumors. Under hypoxic conditions, hypoxia inducible factor (HIF) binds to the hypoxia response element present in the VEGF-A gene, thus inducing the transcription of VEGF-A protein. Circulating VEGF-A binds to VEGF-receptor (VEGFR)-1 and VEGFR-2 stimulating the recruitment and proliferation of endothelial cells.(15-18) Bevacizumab (BV) is a humanized monoclonal antibody (mAB) directed against vascular endothelial growth factor (VEGF)-A.(19)

Treatment strategies incorporating BV have demonstrated efficacy in metastatic colorectal cancer (mCRC).(5, 20) The majority of previously untreated mCRC patients are treated with BV in combination with oxaliplatin and infusional 5-fluorouracil/leucovorin or capecitabine (FOLFOX or XELOX).(20) The addition of BV to first-line oxaliplatin-based chemotherapy in mCRC was shown to significantly prolong the median progression-free survival (PFS; from 8.0 to 9.4 months, corresponding to a hazard ratio (HR) of 0.83; *P*=0.0023). Median overall survival (OS) was 21.3 months in the BV group and 19.9 months in the placebo group (HR 0.89; *P*=0.077). The response rate (RR) was similar in both arms (38% vs 38%; odds ratio (OR) 1.00; *P*=0.99).(21)

The identification of biomarkers that may influence the efficacy of BV in combination with oxaliplatin-based chemotherapy is of considerable interest.(14) Despite the effects of BV in unselected mCRC patients, the ability to target therapy towards well selected subgroups of patients would increase the likelihood of benefit and would improve cost-effectiveness and therapeutic outcomes. Current evidence indicates potential predictive value for germline polymorphisms affecting genes of the VEGF-pathway in patients receiving BV.(22-25) This is not surprising because angiogenesis depends largely on the response of the host. There is substantial inherited genetic variability within angiogenesis pathway genes, thereby causing inter-individual differences in angiogenic capacity and resistance to BV. In a pioneering

study, Schneider *et al.* investigated five VEGF and two VEGFR-2 polymorphisms in a retrospective subset analyses of the E2100 trial cohort (paclitaxel±BV in metastatic breast cancer) and found two VEGF genotypes (VEGF 2578 A/A and VEGF 1154 A/A) predicting a superior OS for patients in the combination, but not in the control arm, thus indicating a predictive marker.(25)

Recent studies in several experimental models suggest that alternative angiogenic factors are potentially involved in resistance to anti-VEGF treatment.(18, 26, 27) Sustained tumor angiogenesis could occur through VEGF-independent mechanisms, thus indicating that these angiogenic factors may serve as predictors of BV efficacy. Schultheis et al. recently reported a functional germline polymorphism in interleukin (IL)-8 (251 T/A, A-allele associated with increased IL-8 protein levels), a potent VEGF-independent pro-angiogenic factor, significantly associated with lower RR in a phase II trial in patients with ovarian cancer treated with BV and cyclophosphamide.(23)

In the present study, we investigated germline polymorphisms in a comprehensive panel of angiogenesis genes to predict clinical outcome and tumor response in mCRC patients treated with first-line BV and oxaliplatin-based chemotherapy. We analyzed VEGF-dependent genes such as VEGF-A, VEGFR-2, HIF1α, aryl hydrocarbon receptor nuclear translocator (ARNT) and neuropilin-1 (NRP1) and VEGF-independent angiogenesis genes such as IL-1β, IL-6, IL-8, interleukin receptor-1/2 (CXCR1 and CXCR2), leptin, tissue factor (TF), endostatin (ES), fibroblast growth factor receptor (FGFR)-4, insulin like growth factor (IGF)-1/2, insulin like growth factor receptor (IGFR1), nuclear factor-κB (NF-κB), epidermal growth factor (EGF), epidermal growth factor receptor (EGFR), cyclooxygenase (COX)-2, tumor necrosis factor (TNF)-α and β, inter-cellular adhesion molecule (ICAM)-1 and matrix metalloproteinases (MMP)-2 and 7.

Adjuvant treatment is recommended for patients with stage III and high-risk stage II colon cancer. The risk of tumor recurrence can be significantly reduced by treating these patients with 5-fluorouracil (5-FU)-based chemotherapy. The addition of oxaliplatin to 5-FU-based chemotherapy is now a standard adjuvant treatment for colon cancer, with a higher 5-year disease-free survival (DFS) rate, compared with 5-FU-based treatment alone (73.3% vs

67.4%).(5) However, a considerable number of patients will relapse despite adjuvant treatment.(28)

Tumor recurrence after curative surgery remains a major obstacle for improving overall cancer survival, which may be in part due to the existence of cancer stem cells (CSC). Growing evidence suggests that human cancers are stem cell diseases and only a small subpopulation of cancer cells, endowed with stem cell-like features, might be responsible for tumor initiation, progression and chemoresistance.(29) Cancer cells with the properties of stem cells possess the ability to self-renew, to undergo multilineage differentiation, and to survive an adverse tissue microenvironment.

Putative CSC populations have been identified in colon cancer on the basis of the expression of specific markers and their functional properties; however, phenotypic characterization of colon CSCs is still a matter of debate and ongoing research studies.(30) Ideally, definitive markers should be gene products that are coupled to the function of the stem cell. CSC markers in colon cancer include CD133, CD44, and CD166.(31) More recently EpCAM, CD26, Msi-1, CD29, CD24, LGR5 and ALDH1A1 have been added to the list of putative stem cell markers for colon cancer.(32) These colon CSC markers are representative of a range of pathways including the Wnt-target genes, cell adhesion molecules, RNA-binding proteins, and detoxifying enzymes, and play distinct roles in a variety of processes including cell differentiation, proliferation, migration, apoptosis, adhesion, lymphocyte homing, angiogenesis and cellular response to chemotherapy.(33) Current therapies target populations of rapidly growing and differentiated tumor cells, but have been shown to lack activity against CSCs.(30) CSCs therefore may have an important role in tumor recurrence despite adjuvant chemotherapy. Thus far, pre-clinical studies in colon cancer have identified that CSCs are capable of initiating tumor development, however, little is known about the role of CSCs in colon cancer tumor recurrence.

There is substantial germline genetic variability within the genes used as markers to identify CSCs, including multiple single nucleotide polymorphisms (SNPs). These common DNA-sequence variations may alter the gene function and/or activity including transcription, translation or splicing, thereby causing inter-individual differences in relation to tumor

recurrence capacity and chemoresistance.(34) Winder et al. recently tested the impact of common gene variants in the cell surface glycoprotein, CD44, a gastric and colon CSC marker, on clinical outcome of patients with localized gastric adenocarcinoma and found that the minor allele of CD44 rs187116 was significantly associated with decreased time to tumor recurrence (TTR) and overall survival (OS), identifying a "high-risk" patient population based on a germline genetic variant.(35)

In the present study, we investigated 25 germline polymorphisms in a comprehensive panel of genes that have been previously associated with colon CSC to predict tumor recurrence in patients with stage III and high-risk stage II colon cancer. The analyzed CSC genes included CD44, Prominin-1 (CD133), dipeptidyl peptidase-4 (DPP4/CD26), epithelial cell adhesion molecule (EpCAM), activated leukocyte cell adhesion molecule (ALCAM/CD166), musashi homolog-1 (MSI-1), integrin beta-1 (ITGB1/CD29), CD24, leucine-rich repeat containing G protein-coupled receptor-5 (LGR5) and aldehyde dehydrogenase-1 family member A1 (ALDH1A1). To the best of our knowledge, this is the first study investigating common germline genetic variants in a comprehensive panel of colon CSC genes to predict tumor recurrence. This study was conducted adhering to the reporting recommendations for tumor marker prognostic studies.(36-38)

Patients and methods

Eligible patients

A total of 132 patients with histopathologically confirmed mCRC and first-line treatment with FOLFOX or XELOX and BV were included in this retrospective study. These patients received first-line treatment with FOLFOX or XELOX and BV (5mg/kg day 1 of a 2-week cycle when given with FOLFOX, 7.5mg/kg on day 1 of a 3-week cycle for XELOX) between April 2004 and October 2009 at the Norris Comprehensive Cancer Center/University of Southern California (NCCC/USC) or the Los Angeles County/USC Medical Center (LAC/USCMC) and the Division of Clinical Oncology, Medical University of Graz (MUG), Austria. Patients included in the study were required to be ≥18 years old, have present one or more unidimensionally measurable lesion, response data available during at least 2 cycles of BV plus FOLFOX or XELOX, and have not received prior systemic therapy for mCRC or previous treatment with monoclonal antibodies. At the time of treatment initiation, the following criteria were used

as contraindication for BV: brain metastases, high-dose NSAIDs, serious non-healing wound, prior pulmonary embolism or recent venous thromboembolic event, any arterial thromboembolic event, and/or baseline ≥ grade 2 proteinuria. Patient data were collected retrospectively through chart review by a medical oncologist. For quality control purposes all clinical data were independently reviewed by a second medical oncologist. Whole blood samples were collected at the time of diagnosis and stored at -80 degree Celsius. Blood samples from 119 patients were available for the current genetic analyses. This retrospective study was approved by the Institutional Review Boards of USC and MUG. All patients signed an informed consent for the analysis of molecular correlates. Baseline clinical examinations and staging CT-scans were performed within 4 weeks of starting treatment and repeated every 8 weeks until progression. The Response Evaluation Criteria in Solid Tumors (RECIST) were used to assess response.

For the adjuvant setting a total of 234 patients with stage III and high-risk stage II colon cancer were included in this study. High-risk stage II colon cancer patients were defined if they presented with at least one of the following features: lymph node sampling <12; poorly differentiated tumor; vascular, lymphatic or perineural invasion; tumor presentation with obstruction or perforation and pT4. All patients were treated with 5-FU-based adjuvant chemotherapy at the Norris Comprehensive Cancer Center/University of Southern California (NCCC/USC) or the Los Angeles County/USC-Medical Center (LAC/USCMC) from 1987 to 2007. All patients were included in the colon cancer surveillance program of NCCC/USC or LAC/USCMC, providing history and physical examination and CEA determination every 3 months for 3 years and every 6 months at years 4 and 5 after surgery, colonoscopy at year 1 and thereafter every 3-5 years and CT scans of chest and abdomen every 6 months for the first 3 years. Patient data were collected retrospectively through chart review. Whole blood was collected at the time of diagnosis and stored at -80 degrees Celsius. Blood samples from 216 patients were available for current genetic analyses. The study was approved by the Institutional Review Boards at USC and all study participants signed informed consent for the analysis of molecular correlates.

Candidate polymorphisms

Genes and polymorphisms known to modulate VEGF-dependent and –independent angiogenesis have been selected based on public literature resources and databases. Stringent and pre-defined criteria were used and included: (a) credible scientific basis to support a gene's involvement in angiogenesis signaling pathways; (b) polymorphism that could alter the function of the gene in a biologically relevant manner (either published data or predicted function using Functional-Single-Nucleotide-Polymorphism (F-SNP) database)(39, 40); (c) minor allele frequency ≥10% in Caucasians (for relative allelic frequencies of the polymorphisms in different ethnicities, we refer to the population genetics section in the Ensembl Genome Browser: http://uswest.ensembl.org/index.html). As it was not possible to select all angiogenesis signaling related genes and polymorphisms matching these criteria, this study focused on the most promising (Table 1 and 2).

Common and putatively functional polymorphisms within genes that have been previously associated with colon CSC were selected using public literature resources and databases including: NCBI-PubMed, dbSNP, Ensembl, GeneCards and the Pharmacogenomics-Knowledge-Base. Stringent and pre-defined selection criteria used were: (a) minor allele frequency ≥10% in Caucasians; (b) polymorphism that could alter the function of the gene in a biologically relevant manner (either published data or predicted function using Functional-Single-Nucleotide-Polymorphism (F-SNP) database, http://compbio.cs.queensu.ca/F-SNP; (c) published clinical associations (e.g. cancer risk/outcome or chemoresistance; Table 3).

Genotyping

Genomic DNA was extracted from peripheral blood using the QIAmp-kit (Qiagen). The majority of the samples were tested using PCR-based restriction fragment length polymorphism (PCR-RFLP) analysis. Briefly, forward and reverse primers were used for PCR amplification, PCR products were digested by restriction enzymes (New England Biolab), and alleles were separated on 4% NuSieve ethidium bromide stained agarose gel. If no matching restriction enzyme could be found, samples were analyzed by direct DNA- sequencing. The dinucleotide polymorphisms were determined by a 5′-end 33p γATP–labeled PCR protocol with a few modifications. In brief, DNA template, deoxynucleotide triphosphates, 5′-end 33p γATP-labeled primer, unlabeled complementary primer, Taq Polymerase (Perkin-Elmer), and

PCR Buffer were used in PCR. The reaction products were separated on a 6% denaturing polyacrylamid DNA-sequencing gel, which was vacuum blotted for 1 hour at 80 degree Celsius and exposed to XAR-film (Eastman-Kodak) overnight. For quality control purposes, a total of 5% PCR-RFLP analyzed samples were re-analyzed by direct DNA-sequencing. Patients' characteristics and clinical outcome were unknown to the investigator performing the genetic analyses.

Statistical analysis

In the metastatic setting the endpoints of the study were PFS, OS and RR. The PFS was calculated from the time of the first day of treatment until the first observation of disease progression or death from any cause. OS was defined as the time from the first day of treatment to death from any cause. If a patient was alive and had not progressed, PFS and OS were censored at the time of last follow-up. RR was categorized as two groups: complete or partial response and stable or progressive disease. With 119 patients, there was an 80% power to detect a minimum HR of 1.7 and 1.8, and 34% tumor response differences across the range of minor allele frequencies (0.1-0.5) using a dominant model for PFS, OS and RR, respectively. The NF-κB repeat polymorphism was analyzed by categorizing 3 groups: (a) carrying both alleles <24 repeats; (b) carrying one allele <24; and (c) carrying both alleles ≥24 repeats. For the VEGFR-2 repeat polymorphisms, we combined the 14 with the 13 CA repeats and categorized in 11/11, 11/13 and 13/13 CA repeats. Allelic distribution of the polymorphisms by ethnicity for deviation from Hardy-Weinberg equilibrium and the allelic frequencies of each polymorphism between different ethnic groups were tested using χ^2-test. The associations between polymorphisms and baseline demographic and clinical characteristics and RR were examined using contingency tables and the Fisher's exact test. The associations of polymorphisms with PFS and OS were analyzed using Kaplan-Meier curves and log-rank test. The true mode of inheritance of all polymorphisms tested is not established yet and we assumed a codominant, additive, dominant or recessive genetic model where appropriate. In the multivariate Cox regression analysis for PFS and OS and the logistic regression analysis for RR, the models were adjusted for sex, age, primary tumor site, histologic differentiation, metastatic site, number of metastatic sites, chemotherapy backbone (FOLFOX or XELOX) and study site (Hospital) stratified by ethnicity. *P*-values for all polymorphisms were adjusted for multiple testing using a modified test of Conneely and

Boehnke (*P*act) that was applied for the correlated tests due to linkage disequilibrium and different modes of inheritance considered.(41) Recursive partitioning (RP), including cross-validation, was used to explore and identify polymorphism interactions associated with RR, PFS and OS using the rPart-function in S-plus.(42, 43) Case-wise deletion for missing polymorphisms was used in univariate and multivariate analyses. In the RP-analysis, all patients with at least one polymorphism result available were included.

For the adjuvant setting the endpoint of the study was TTR. TTR was calculated from the date of diagnosis of colon cancer to the date of the first observation of tumor recurrence. TTR was censored at the time of death or at the last follow-up if the patient remained tumor recurrence-free at that time. With 216 patients, there was an 80% power to detect a minimum hazard ratio (HR) of 1.8 across the range of minor allele frequencies (0.2-0.5) for TTR using a dominant model. For the recessive model, the minimum HR was 3.1 when the allele frequency was 0.2 and approaches 1.8 when the allele frequency was 0.5. Allelic distribution of the polymorphisms by ethnicity was tested for deviation from Hardy-Weinberg equilibrium using χ^2-test. The distribution of polymorphisms across baseline demographic, clinical and pathological characteristics was examined using Fisher's exact test. The true mode of inheritance of all polymorphisms tested is not established yet and we assumed a codominant, additive, dominant or recessive genetic model where appropriate. The association of polymorphisms with TTR was analyzed using Kaplan-Meier curves and log-rank test. In the multivariate Cox-regression analysis, the model was adjusted for stage and type of adjuvant chemotherapy, and stratified by ethnicity. Interactions between polymorphisms and stage and gender on TTR were tested by comparing likelihood ratio statistics between the baseline and nested Cox proportional hazards models that include the multiplicative product term. *P*-values for all polymorphisms were adjusted for multiple testing using a modified test of Conneely and Boehnke that was applied for the correlated tests due to linkage disequilibrium and the different modes of inheritance considered. Recursive partitioning (RP), including cross-validation, was used to explore and identify polymorphism profiles associated with TTR using the rPart-function in S-plus. Case-wise deletion for missing polymorphisms was used in univariate and multivariate analyses. In the RP-analysis, all patients with at least one polymorphism result available were included. All analyses were performed using SAS 9.2 (SAS Institute Inc. NC, USA).

Results

Metastatic setting

Baseline patient characteristics and the association with clinical outcome and RR are summarized in Table 4. There were no statistically significant differences in age and sex between the patients from USC and MUG, however, more MUG patients had rectosigmoid tumors (28% vs 6%) and received FOLFOX (80% vs 43%), compared to USC patients. The median age at time of diagnosis was 56 years (range 28-81). The median follow-up time was 3.5 years (range 0.3-6.5) and the median PFS and OS was 11.2 months (95%CI 8.4-13.7) and 27.9 months (95%CI 22.4-32.4), respectively. Sixty-four patients (48.5%) showed complete or partial tumor response, 63 patients (47.7%) had stable or progressive disease, and for 5 patients (3.8%) response data were not available. The genotyping quality control by direct DNA-sequencing provided a genotype concordance of ≥99%. Genotyping was successful in at least 90% of cases in each polymorphism analyzed, except the NF-κB repeat polymorphism with 84%. The allelic frequencies for all polymorphisms were within the probability limits of Hardy-Weinberg equilibrium. VEGF-A rs2010963, rs699947, rs1570360 and rs833061 were in linkage disequilibrium (D' ranged from 0.79 to 1.0 and r^2 ranged from 0.13 to 0.81). VEGF-A rs3025039 was not in linkage disequilibrium with the other four VEGF-A gene variants (D'<0.16 and r^2<0.01). The VEGFR-2 polymorphisms were not in strong linkage disequilibrium (D'=0.47 and r^2=0.04). Genotype frequencies for all polymorphisms in each ethnicity are summarized in Supplementary Table 1.

In the univariate analysis, the minor allele of HIF1α rs11549465 C>T was significantly associated with increased PFS. Patients carrying at least one T allele showed a median PFS of 11.7 months. In contrast, patients homozygous C/C had a median PFS of 10.2 months (HR 0.55; 95%CI 0.31-0.99; *P*=0.038). None of the other tested polymorphisms were associated with PFS. In multivariate analysis stratified by ethnicity and multiple testing including all polymorphisms, HIF1α rs11549465 C>T did not remain significantly associated with PFS (HR 0.65; 95%CI 0.34-1.22; adjusted-*P*=0.18; *P*act=0.362; Table 5).

In the univariate analysis, the minor alleles of EGF rs444903 A>G and IGF-1 rs6220 A>G were significantly associated with increased OS. Patients carrying at least one G allele in EGF rs444903 A>G showed a median OS of 32.4 months. In contrast, patients homozygous A/A

had a median OS of 21.9 months (HR 0.54; 95%CI 0.33-0.88; *P*=0.011). Patients harboring the minor allele of IGF-1 rs6220 A>G showed a median OS of 32.4 months compared to 22.1 months harboring homozygous A/A (HR 0.51; 95%CI 0.32-0.83; *P*=0.005; Table 5; Figure 1). None of the other tested polymorphisms demonstrated a statistically significant association with OS. In the multivariate analysis stratified by ethnicity, the minor alleles of EGF rs444903 A>G and IGF-1 rs6220 A>G remained significantly associated with increased OS (HR 0.52; 95%CI 0.31-0.87; adjusted-*P*=0.012 and HR 0.60; 95%CI 0.36-0.99; adjusted-*P*=0.046, respectively). In multiple testing, EGF rs444903 A>G and IGF-1 rs6220 A>G did not remain significantly associated with OS (*P*act=0.664 and 0.780, respectively)

CXCR1 rs2234671 G>C, CXCR2 rs2230054 T>C, EGFR rs2227983 G>A and VEGFR-2 rs2305948 C>T were significantly associated with RR. Patients homozygous for the wild-type allele of CXCR1 rs2234671 G>C and VEGFR-2 rs2305948 C>T were more likely to show higher tumor response (71% and 57%, respectively), compared to patients carrying one (37% and 29%, respectively) or two minor alleles (17% and 33%; *P*<0.001 and *P*=0.024, respectively). In CXCR2 rs2230054 T>C and EGFR rs2227983 G>A, patients harboring the wild-type genotype had a significantly lower RR (38% and 43%, respectively), compared to patients heterozygous (56% and 55%, respectively) or homozygous for the minor allele (79% and 82%; *P*=0.008 and *P*=0.024, respectively; Table 5). None of the other analyzed polymorphisms predicted RR. In logistic regression analysis, CXCR1 rs2234671 G>C remained significantly associated with RR (*P*<0.001). In multiple testing including all polymorphisms, CXCR1 rs2234671 G>C remained statistically significant for RR (*P*act=0.003).

When RP was utilized to construct decision-trees as predictive models for PFS, OS and RR to classify patients based on the gene variants, high- and low-risk patient subgroups were identified. In the resultant tree for PFS, the most important factor that determined PFS in our study cohort was HIF1α rs11549465 C>T. Patients carrying the combination of HIF1α rs11549465 C>T wild-type and the minor allele of VEGF rs699947 C>A and EGFR rs2227983 G>A wild-type demonstrated a PFS of 7.8 months compared to 11.7 months in patients harboring at least one minor allele of HIF1α rs11549465 C>T (HR 2.66; 95%CI 1.30-5.42; *P*<0.001). The resultant tree for OS showed that patients heterozygous or homozygous for the minor alleles of IGF-1 rs6220 A>G and IL6 rs1800795 G>C demonstrated an OS of 60

months compared to 21.7 months in patients harboring the IGF-1 rs6220 A>G wild-type (HR 2.67; 95%CI 1.25-5.66; *P*<0.001). For RR, CXCR1 rs2234671 G>C was the main split criteria in the decision tree, but no other gene variants were shown to improve the prediction success.

When the allelic frequencies of each polymorphism, which was significantly associated with clinical outcome or tumor response were tested between ethnic groups, we found a significant difference for CXCR2 rs2230054 T>C, VEGFR-2 rs2305948 and EGF rs444903 A>G (*P*=0.017, *P*=0.018 and *P*=0.005, respectively). The minor allele of EGF rs444903 A>G remained significantly associated with increased OS (HR 0.47; 95%CI 0.24-0.93; *P*=0.021) in Caucasians, but not in Asians and Hispanics (HR 0.65; 95%CI 0.21-1.98; *P*=0.42 and HR 0.50; 95%CI 0.17-1.50; *P*=0.21, respectively). African Americans were excluded from sub-analyses because of the small number in our study cohort (n=5). CXCR2 rs2230054 T>C remained significantly associated with RR in Caucasians (wild-type: 35%, heterozygous for the minor allele: 60%, homozygous for the minor allele: 100%; *P*=0.035), but not in Asians and Hispanics (*P*=0.351 and *P*=0.191, respectively). Patients homozygous for the wild-type allele of VEGFR-2 rs2305948 C>T were more likely to show higher tumor response compared to patients carrying one or two minor alleles, but this effect did not remain significant in sub-analyses for ethnicity (Caucasian: *P*=0.17, Asian: *P*=0.055, Hispanic: *P*=0.86).

Adjuvant setting

The baseline characteristics of the 234 patients included in this analysis are summarized in Table 6. The median age at time of diagnosis was 59 years (range 22-87), with a median follow-up time of 4.4 years (range 0.4-16.8). Ninety (38.5%) patients showed tumor recurrence, with a stage III and high-risk stage II dependent probability of 3-year recurrence of 0.45±0.047 and 0.21±0.043, respectively. Median OS has not been reached yet. The genotyping quality control by direct DNA-sequencing provided a genotype concordance of ≥99%. Genotyping was successful in at least 90% of cases for each polymorphism analyzed, with the exception of CD44 rs8193 (88.4%). In failed cases, genotyping was not successful because of limited quantity and/or quality of extracted genomic DNA. The allelic frequencies for all polymorphisms were within the probability limits of Hardy-Weinberg equilibrium, with the exception of EpCAM rs17036526 (data not shown).

There were no significant associations between the polymorphisms and baseline demographic, clinical or pathological characteristics (data not shown). In the univariate analysis, the minor alleles of CD44 rs8193 C>T, ALCAM rs1157 G>A and LGR5 rs17109924 T>C were significantly associated with an increased TTR. Patients carrying at least one T allele in CD44 rs8193 showed a median TTR of 9.4 years. In contrast, patients with homozygous C/C had a median TTR of 5.4 years (HR, 0.51; 95%CI, 0.35-0.93; *P*=0.022). Patients harboring the minor allele of ALCAM rs1157 showed a median TTR of 11.3 years compared to 5.7 years for patients harboring the homozygous G/G (HR, 0.56; 95%CI, 0.33-0.94; *P*=0.024). Patients carrying one C allele in LGR5 rs17109924 had a median TTR of 10.7 years compared to 5.7 years for those patients carrying the homozygous T/T (HR, 0.33; 95%CI, 0.12-0.90; *P*=0.023; Figure 2). The other tested gene variants did not demonstrate any statistically significant associations with TTR in the univariate analyses.

In the multivariate analysis stratified by ethnicity, the minor alleles of CD44 rs8193 C>T, ALCAM rs1157 G>A and LGR5 rs17109924 T>C remained significantly associated with increased TTR (Table 7). There was no significant interaction between the polymorphisms and tumor stage or gender on TTR (*P*-values for interactions >0.05). In multiple testing including all polymorphisms analyzed, none of them remained significantly associated with TTR (adjusted-*P* for CD44 rs8193=0.142; adjusted-*P* for ALCAM rs1157=0.199; adjusted-*P* for LGR5 rs17109924=0.394)

When RP was utilized to construct a decision-tree as a predictive model to classify patients based on the gene variants, high- and low-risk patient subgroups were identified. In the resultant tree, the most important factor that determined the TTR in these patients was LGR5 rs17109924. Patients carrying the combination of LGR5 rs17109924 wild-type and at least one CD44 rs8193 wild-type allele and the mutant variant of ALDH1A1 rs1342024 demonstrated a TTR of 1.7 years (Node 5) compared to 10.7 years in patients with the minor allele of LGR5 rs17109924 (Node 1) or the combination of LGR5 rs17109924 wild-type and CD44 rs8193 mutant variant (Node 2; HR, 6.71, 95%CI, 2.71-16.63, *P*<0.001; Figure 3).

To evaluate if the high-risk patients identified from our gene variant profile benefit from combination chemotherapy (n=47) compared to 5-FU alone (n=66), the cases from node 4

and 5 of the decision tree were combined for further analysis. According to the treatment regimen, no significant difference in TTR were identified in this high-risk subgroup (P>0.05).

Discussion

Even though thousands of patients have been enrolled in randomized clinical trials of BV, only few insights are available about specific subgroups of patients who may actually benefit.(44) At the same time, no biomarkers are currently available to quantify the contribution of BV to the activity of cytotoxic drugs. Since the efficacy of BV may vary between different chemotherapy backbones, the current priority in translational research is focused on biomarkers to the response or resistance to combined therapies. In this investigation, we focused on host-related VEGF-dependent and –independent angiogenic biomarkers to predict survival and tumor response of mCRC patients treated with BV and oxaliplatin-based chemotherapy. This study has been performed to draw biological observations to be validated in biomarker-embedded trials.

In mCRC, *Loupakis et al.* found no association between VEGF and VEGFR-2 polymorphisms and clinical outcome in 57 patients enrolled in a phase II trial of FOLFOXIRI plus BV as first-line treatment.(45) One report by Formica *et al*. suggested that VEGF rs1570360 and VEGF rs2010963 genotypes predict PFS and RR in patients receiving BV and FOLFIRI.(24) A genetic interaction profile by Pander *et al.* including VEGF rs2010963 was associated with PFS in mCRC patients treated with XELOX plus BV.(46) While this study did not confirm these findings, we identified two other VEGF-dependent gene polymorphisms associated with PFS and RR in mCRC patients treated with BV and oxaliplatin-based chemotherapy. The minor allele of HIF1α rs11549465 C>T, which has been recently associated with higher HIF1α protein expression, predicted an increased PFS in the univariate analysis.(47) The VEGF-A promoter contains a hypoxic response element that can bind HIF1α, and initiate transcriptional activation of the VEGF-A gene.(48) Although the association found in our study is biologically plausible and in concordance with several studies showing that patients with VEGF-A activation more likely benefit from BV treatment, HIF1α rs11549465 C>T did not remain significant in the multivariate analysis.(49, 50) In VEGFR-2 rs2305948 C>T, the minor allele predicted a lower RR in the univariate analysis when analyzed in all study patients. However, this effect did not remain significant in sub-analyses for ethnicity. VEGFR-2 rs2305948 C>T is associated with microvessel density (MVD) in CRC with the C/T and T/T genotypes showing a significantly higher MVD compared to the wild-type.(51) Taking into account the functional effect of this polymorphism and the association found in our study,

we suggest that the variant allele in the VEGFR-2 gene may drive angiogenesis independent of VEGF-A ligand binding.

Since alternative angiogenic mechanisms are potentially involved in resistance to BV, we investigated a comprehensive panel of VEGF-independent gene polymorphisms. CXCR1 rs2234671 G>C represented the most promising polymorphism in our study, predicting RR in multivariate analysis and multiple testing, while CXCR2 rs2230054 C>T lost its significance in multivariate analysis and multiple testing. IL-8 exerts its angiogenic properties on endothelial cells through interaction with its cognate receptors CXCR1 and CXCR2.(52) Induction of IL-8 preserved the angiogenic response in HIF1α-deficient colon cancer cells, suggesting that IL-8 dependent angiogenesis is independent of VEGF.(53) The detailed molecular mechanisms involved in how the CXCR1 rs2234671 G>C and CXCR2 rs2230054 C>T polymorphisms exert effects on mCRC and predict RR in BV and FOLFOX or XELOX treated patients are unclear. We used the F-SNP database to predict the functional effects of these gene variants. F-SNP gathers computationally predicted functional information about polymorphisms, particularly aiming to facilitate identification of disease-related polymorphisms in association studies. Specifically, it provides information about potential deleterious effects of polymorphisms with respect to major molecular functions. When used for the CXCR gene variants, F-SNP predicted changes in splicing regulation and post-translation for CXCR1 rs2234671 G>C, and changes in splicing regulation and transcriptional regulation for CXCR2 rs2230054 C>T, thus supporting a biological function and the effects seen in our study.

In a recent study by Cascone *et al.* using mouse xenograft models of human lung adenocarcinomas, EGFR activation in stromal cells, but not in tumor cells has been shown to be involved in BV resistance, indicating a host-regulated process of VEGF-independent angiogenesis.(54) We found a functional germline polymorphism in EGFR associated with RR in univariate analysis in our study cohort. Patients harboring at least one minor allele of EGFR rs2227983 G>A were more likely to show a higher tumor response, compared to patients homozygous for the wild-type allele. The minor allele was previously found to be associated with attenuated EGFR ligand binding.(55) Our finding is therefore consistent with the functional effect of this polymorphism, as decreased EGFR-signaling may attenuate the VEGF-independent angiogenic capacity.

Another finding of our study was that EGF rs4444903 A>G and IGF-1 rs6220 A>G predict OS in the univariate and multivariate analysis. The minor allele of EGF rs4444903 A>G is transcriptionally more active than the wild-type and has been associated with higher EGF serum levels.(56) In a recent study by Pander *et al*., EGF rs4444903 A>G was not associated with PFS in mCRC patients treated with XELOX plus BV.(57) EGF-signaling contributes to cell proliferation and angiogenesis; therefore, the favorable effect of the EGF rs4444903 minor allele seen in our study may be counterintuitive. EGF-signaling promotes angiogenesis not only exerting direct effects on endothelial cells, but also by upregulating VEGF mRNA expression. This partly VEGF-dependent mechanism may be responsible for the increased OS survival in BV-treated patients harboring the minor allele. Similar potential mechanisms may also apply for IGF-1 rs6220 A>G, since the minor allele shows higher IGF-1 plasma levels, which is associated with increased VEGF mRNA expression.(58-60) On the other hand, it has been shown that at specific concentrations that vary between experimental model, EGF induces apoptosis and growth inhibition, rather than cell proliferation.(61-63) According to such findings and the fact that EGF rs4444903 A>G predicts OS but not PFS or RR, this polymorphism could represent a prognostic rather than a predictive biomarker, which is in concordance with several studies showing that the EGF upregulating genotype predicts significantly better survival rates.(64)

We performed multiple testing because of the large number of independent genetic variants evaluated. Application of a modified test of Conneely and Boehnke for correlated tests resulted in a significant *P*act-value only for CXCR1 rs2234671 G>C in predicting RR. Nevertheless, the biological plausibility and our translational findings hold promise for future investigations and warrant validation in a larger cohort. Investigating gene variant interactions using RP, we found that combinations of gene variants may improve the prediction success for PFS and OS, but not for RR in our study cohort. Interestingly, VEGF rs699947 C>A and IL6 rs1800795 G>C were included in the decision tree by RP for PFS and OS, respectively. The minor allele of VEGF rs699947 C>A is associated with lower VEGF serum levels and predicted the low PFS subgroup in RP analysis. This association is in concordance with several studies showing that patients with VEGF-A activation more likely benefit from BV treatment. The minor allele of IL6 rs1800795 G>C is associated with lower

IL6 plasma levels and may therefore contribute to the high OS subgroup predicted in RP analysis. The fact that allelic frequencies of polymorphisms can differ between ethnic groups and influence the clinical effect as seen in our study points out the importance of ethnicity specific analyses in future studies. Because of the combined treatment in our study and the lack of an appropriate control group, the results are not directly attributable to single-agent administration of BV, but should be referred to the combined treatment of BV with FOLFOX or XELOX.

This study provides evidence that functional germline polymorphisms in VEGF-dependent and –independent angiogenesis genes predict tumor response und survival in mCRC patients treated with BV and FOLFOX or XELOX. Biomarker-embedded translational trials are warranted to validate these findings.

In the adjuvant, we investigated germline polymorphisms in a comprehensive panel of genes that have been previously associated with colon CSC to predict tumor recurrence in patients with stage III and high-risk stage II colon cancer. The results indicate that common CSC gene variants in CD44, ALCAM, LGR5 and ALDH1A1 may be valuable to separate high-risk from low-risk colon cancer patients.

The detailed molecular mechanisms involved in how the CD44 rs8193, ALCAM rs1157, LGR5 rs17109924 and ALDH1A1 rs1342024 polymorphisms exert effects on colon cancer are unclear. Non-synonymous polymorphisms lead to amino acid changes and thus may affect the protein function.(65) 3'UTRs have been implicated in the modulation of gene regulation at the transcriptional or translational level and function as regulators mainly through control of mRNA stability and/or translational efficiency, and therefore play an important role in the overall fate of gene expression. Further, germline polymorphisms in the 3'UTRs have been shown to have functional effects on overall gene expression.(66) We used the F-SNP database to predict the functional effects of the analyzed polymorphisms. F-SNP gathers computationally predicted functional information about polymorphisms, particularly aiming to facilitate identification of disease-related polymorphisms in association studies. When used for this study set of polymorphisms, F-SNP predicted changes in transcription factor binding to the 3’UTR located CD44 rs8193 and ALCAM rs1157, and changes in splicing

regulation and protein coding for the non-synonymous LGR5 rs17109924, thus supporting the effects seen in our study. No prediction could be provided for the upstream located ALDH1A1 rs1342024 by the software.

CD44-signaling is crucial in cancer cell proliferation, motility and migration. As a Wnt-target gene, CD44 promotes cell proliferation via the phosphatidylinositol 3-kinase (PI3K)/Akt pathway. CD44 positive colon cancer cells have been reported to possess the capacity for self-renewal, longevity and multipotency.(67) ALCAM belongs to the immunoglobulin superfamily of cell adhesion molecules involved in cell-cell interactions. ALCAM may regulate through cytoskeletal anchoring and the integrity of the extracellular immunoglobulin-like domains complex cellular properties in regard to cell adhesion, migration and growth.(68) Isolation of ALCAM/CD44 double-positive cells from human colon cancer cells can recapitulate tumorigenesis when xenografted at low numbers into immune-deficient mice which represents a hallmark of CSCs.(31) Despite the potentially high clinical relevance of these CSC markers, little is known about their prognostic significance in colon cancer and contradictory findings have been reported. In a recent study based on 110 colorectal cancer (CRC) patients, membranous expression of CD44 and ALCAM did not predict survival in single-marker analyses, but gained significance when combined.(69) In contrast, Weichert *et al.* found a correlation between membranous ALCAM expression and decreased survival in 111 CRC patients.(70) Interestingly, loss, rather than overexpression, of membranous CD44 and ALCAM was correlated to outcome in an analysis including 101 CRC patients. The authors suggested that their results are in large part dependent on the cell adhesion function of CD44 and ALCAM with loss of cell adhesion representing a fundamental step underlying the initiation of the metastatic process.(71)

These conflicting results raised the question whether germline genetic variants putatively changing the gene's function rather than membranous evaluation of these proteins may predict colon cancer patient's outcome. We recently showed that the minor allele of CD44 rs187116 predicts decreased TTR and OS in patients with localized gastric adenocarcinoma.(35) More recently, Zhou *et al.* analyzed two polymorphisms in ALCAM investigating 1033 breast cancer patients and 1116 controls and found that individuals harboring the ALCAM rs6437585 C/T or T/T genotypes have an odds ratio (OR) of 1.38

(95%CI, 1.11-1.72) for developing breast cancer, compared to the C/C genotype. Additional experiments showed that the T allele was associated with a higher transcriptional activity of the ALCAM gene.(72) Both polymorphisms, CD44 rs187116 and ALCAM rs6437585, did not show any clinical associations in this study. However, these SNPs have been analyzed in other tumor entities and/or setting (risk assessment) and therefore might not exert their effects in colon cancer and tumor recurrence assessment. Since CD44 rs8193 and ALCAM rs1157 may change transcriptional factor binding to the polymorphic region of the gene as predicted by F-SNP, the findings in the present study that these gene variants affect TTR in colon cancer are biologically plausible. Taking into account the clinical association found in our study and the predicted function by F-SNP, we hypothesize that the wild-type genotype of CD44 rs8193 is associated with a higher transcriptional activity of the CD44 gene leading to a lower TTR in our study cohort. For ALCAM expression and clinical outcome in colon cancer conflicting results can be found in the literature, therefore we cannot assess the molecular mechanisms involved in how the ALCAM rs1157 polymorphism exerts its effects on colon cancer and studies elucidating the exact biological function of the ALCAM rs1157 polymorphism should be under future consideration.

Interestingly, the tree analysis provided LGR5 rs17109924 as the first split indicating the most important factor determining TTR in this patient cohort. LGR5 is a member of the G-protein-coupled receptor (GPCR) family comprising proteins with seven transmembrane domains. GPCRs function as receptors for various classes of ligands, including peptide hormones and chemokines; however, the ligand for and function of LGR5-related signaling remains unclear. LGR5, a Wnt-target gene, has been reported to be a marker for colon CSC, thus playing a putative role in the biological function of stem cells. In a recent study, high membranous LGR5 expression was shown to predict lower DFS in CRC patients.(73, 74) LGR5 rs17109924 represents a non-synonymous SNP and was predicted to affect splicing regulation and protein coding by F-SNP. In addition, LGR5 rs17109924 predicted TTR in both the univariate and multivariate analysis and was incorporated in the tree analysis, strongly indicating that this SNP has functional significance. Taking into account the clinical association found in our study and the predicted function by F-SNP, we hypothesize that the LGR5 rs17109924 wild-type genotype is associated with a higher protein expression of the LGR5 gene leading to a lower TTR in our study cohort.

A combination of gene variants in the tree analysis defined a high-risk subgroup with significantly lower TTR by incorporating ALDH1A1 rs1342024 when compared to single marker analysis. ALDH1 is a detoxifying enzyme that oxidizes intracellular aldehydes. This detoxification capacity may protect stem cells against oxidative insult.(75, 76) Taking into account the clinical association found in our study, we hypothesize that the mutant variant of ALDH1A1 rs1342024 is associated with a higher detoxification capacity of ALDH1 leading to a lower TTR in our study cohort when incorporated in the decision tree algorithm. In a recent study, membranous ALDH1 expression did not predict survival in CRC patients.(71) Although the mechanism of ALDH1A1 rs1342024 remains unclear, our data suggests that a multigenic approach, which assesses the combined effects of gene variants, may detect synergistic interactions between individual SNPs thus enhancing the predictive power of the model.

This study further utilized multiple testing due to the large number of independent genetic variants evaluated. Application of a modified test of Conneely and Boehnke resulted in a non-significant adjusted *P*-value for CD44 rs8193, ALCAM rs1157 and LGR5 rs17109924. Therefore, these data warrant further validation in a larger cohort. Nevertheless, the biological plausibility and our translational findings hold promise for further investigations in independent study populations. In a sub-analysis combining high-risk patients based on our gene variant profile, no benefit for the addition of oxaliplatin or irinotecan to 5-FU-based chemotherapy could be demonstrated. Since all patients included in this study represent stage III and high-risk stage II colon cancer treated with adjuvant therapy, it was not possible to correlate the genotypes with clinical outcome in an untreated control group. As a consequence, it could not be determined whether the high-risk patients, based on the gene variant profile, did not benefit from combination chemotherapy or from chemotherapy at all.

This study provides the first evidence that germline polymorphisms in CSC genes predict early tumor recurrence in patients with colon cancer. This may help to select subgroups of patients who may benefit from more aggressive treatment strategies or newly developed

stem cell targeting drugs. Future biomarker-embedded translational trials are warranted to validate these findings.

References

1. Wilson PM, Labonte MJ, Lenz HJ. Molecular markers in the treatment of metastatic colorectal cancer. Cancer J 2010;16: 262-72.
2. Benhaim L, Labonte MJ, Lenz HJ. Pharmacogenomics and metastatic colorectal cancer: Current knowledge and perspectives. Scand J Gastroenterol 2011;47: 325-39.
3. Bohanes P, LaBonte MJ, Winder T, Lenz HJ. Predictive molecular classifiers in colorectal cancer. Semin Oncol;38: 576-87.
4. Lenz HJ. Pharmacogenomics and colorectal cancer. Ann Oncol 2004;15 Suppl 4: iv173-7.
5. Cunningham D, Atkin W, Lenz HJ*, et al.* Colorectal cancer. Lancet 2010;375: 1030-47.
6. Wilson PM, Lenz HJ. Integrating biomarkers into clinical decision making for colorectal cancer. Clin Colorectal Cancer 2010;9 Suppl 1: S16-27.
7. Zhang W, Azuma M, Lurje G*, et al.* Molecular predictors of combination targeted therapies (cetuximab, bevacizumab) in irinotecan-refractory colorectal cancer (BOND-2 study). Anticancer Res 2010;30: 4209-17.
8. Zhang W, Stoehlmacher J, Park DJ*, et al.* Gene polymorphisms of epidermal growth factor receptor and its downstream effector, interleukin-8, predict oxaliplatin efficacy in patients with advanced colorectal cancer. Clin Colorectal Cancer 2005;5: 124-31.
9. Hsieh SM, Lintell NA, Hunter KW. Germline polymorphisms are potential metastasis risk and prognosis markers in breast cancer. Breast Dis 2006;26: 157-62.
10. Knechtel G, Hofmann G, Gerger A*, et al.* Analysis of common germline polymorphisms as prognostic factors in patients with lymph node-positive breast cancer. J Cancer Res Clin Oncol 2010;136: 1813-9.
11. Tan P. Germline polymorphisms as modulators of cancer phenotypes. BMC Med 2008;6: 27.
12. Folkman J, Shing Y. Angiogenesis. J Biol Chem 1992;267: 10931-4.
13. Carmeliet P. Angiogenesis in health and disease. Nat Med 2003;9: 653-60.
14. Gerger A, Labonte M, Lenz HJ. Molecular predictors of response to antiangiogenesis therapies. Cancer J 2011;17: 134-41.
15. Ferrara N. Vascular endothelial growth factor: basic science and clinical progress. Endocr Rev 2004;25: 581-611.
16. Ferrara N. VEGF-A: a critical regulator of blood vessel growth. Eur Cytokine Netw 2009;20: 158-63.
17. Ferrara N. Vascular endothelial growth factor. Arterioscler Thromb Vasc Biol 2009;29: 789-91.
18. Ferrara N. Pathways mediating VEGF-independent tumor angiogenesis. Cytokine Growth Factor Rev 2010;21: 21-6.
19. Jubb AM, Harris AL. Biomarkers to predict the clinical efficacy of bevacizumab in cancer. Lancet Oncol 2010;11: 1172-83.
20. Welch S, Spithoff K, Rumble RB, Maroun J. Bevacizumab combined with chemotherapy for patients with advanced colorectal cancer: a systematic review. Ann Oncol 2010;21: 1152-62.
21. Saltz LB, Clarke S, Diaz-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.

22. Steffensen KD, Waldstrom M, Brandslund I, Jakobsen A. The relationship of VEGF polymorphisms with serum VEGF levels and progression-free survival in patients with epithelial ovarian cancer. Gynecol Oncol 2010;117: 109-16.
23. Schultheis AM, Lurje G, Rhodes KE*, et al.* Polymorphisms and clinical outcome in recurrent ovarian cancer treated with cyclophosphamide and bevacizumab. Clin Cancer Res 2008;14: 7554-63.
24. Formica V, Palmirotta R, Del Monte G*, et al.* Predictive value of VEGF gene polymorphisms for metastatic colorectal cancer patients receiving first-line treatment including fluorouracil, irinotecan, and bevacizumab. Int J Colorectal Dis 2011;26: 143-51.
25. Schneider BP, Wang M, Radovich M*, et al.* Association of vascular endothelial growth factor and vascular endothelial growth factor receptor-2 genetic polymorphisms with outcome in a trial of paclitaxel compared with paclitaxel plus bevacizumab in advanced breast cancer: ECOG 2100. J Clin Oncol 2008;26: 4672-8.
26. Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. Clin Cancer Res 2008;14: 6735-41.
27. Cao Y, Zhong W, Sun Y. Improvement of antiangiogenic cancer therapy by understanding the mechanisms of angiogenic factor interplay and drug resistance. Semin Cancer Biol 2009;19: 338-43.
28. Tejpar S, Bertagnolli M, Bosman F*, et al.* Prognostic and predictive biomarkers in resected colon cancer: current status and future perspectives for integrating genomics into biomarker discovery. Oncologist 2010;15: 390-404.
29. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. Nat Rev Gastroenterol Hepatol 2011;8: 90-100.
30. Todaro M, Francipane MG, Medema JP, Stassi G. Colon cancer stem cells: promise of targeted therapy. Gastroenterology 2010;138: 2151-62.
31. Dalerba P, Dylla SJ, Park IK*, et al.* Phenotypic characterization of human colorectal cancer stem cells. Proc Natl Acad Sci U S A 2007;104: 10158-63.
32. Vermeulen L, Todaro M, de Sousa Mello F*, et al.* Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. Proc Natl Acad Sci U S A 2008;105: 13427-32.
33. Saif MW, Chu E. Biology of colorectal cancer. Cancer J 2010;16: 196-201.
34. Coate L, Cuffe S, Horgan A, Hung RJ, Christiani D, Liu G. Germline genetic variation, cancer outcome, and pharmacogenetics. J Clin Oncol 2009;28: 4029-37.
35. Winder T, Ning Y, Yang D*, et al.* Germline polymorphisms in genes involved in the CD44 signaling pathway are associated with clinical outcome in localized gastric adenocarcinoma (GA). Int J Cancer 2010.
36. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies. J Clin Oncol 2005;23: 9067-72.
37. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 2006;100: 229-35.
38. Alonzo TA. Standards for reporting prognostic tumor marker studies. J Clin Oncol 2005;23: 9053-4.
39. Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res 2008;36: D820-4.
40. Lee PH, Shatkay H. An integrative scoring system for ranking SNPs by their potential deleterious effects. Bioinformatics 2009;25: 1048-55.

41. Conneely KN, Boehnke M. So Many Correlated Tests, So Little Time! Rapid Adjustment of P Values for Multiple Correlated Tests. Am J Hum Genet 2007;81.
42. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. New York: Chapman&Hall (Wadsworth, Inc) 1984.
43. Therneau TM, Atkinson EJ. An Introduction to Recursive Partitioning Using the RPART Routines. Technical Report no 61, Mayo Clinic, Rochester, Mn 1997.
44. Hayes DF. Bevacizumab treatment for solid tumors: boon or bust? JAMA 2011;305: 506-8.
45. Loupakis F, Cremolini C, Fioravanti A*, et al.* Pharmacodynamic and pharmacogenetic angiogenesis-related markers of first-line FOLFOXIRI plus bevacizumab schedule in metastatic colorectal cancer. Br J Cancer 2010;104: 1262-9.
46. Pander J, Wessels JA, Gelderblom H, van der Straaten T, Punt CJ, Guchelaar HJ. Pharmacogenetic interaction analysis for the efficacy of systemic treatment in metastatic colorectal cancer. Ann Oncol 2011;22: 1147-53.
47. Kim HO, Jo YH, Lee J, Lee SS, Yoon KS. The C1772T genetic polymorphism in human HIF-1alpha gene associates with expression of HIF-1alpha protein in breast cancer. Oncol Rep 2008;20: 1181-7.
48. Pugh CW, Ratcliffe PJ. Regulation of angiogenesis by hypoxia: role of the HIF system. Nat Med 2003;9: 677-84.
49. Dowlati A, Gray R, Sandler AB, Schiller JH, Johnson DH. Cell adhesion molecules, vascular endothelial growth factor, and basic fibroblast growth factor in patients with non-small cell lung cancer treated with chemotherapy with or without bevacizumab--an Eastern Cooperative Oncology Group Study. Clin Cancer Res 2008;14: 1407-12.
50. Yang SX, Steinberg SM, Nguyen D, Wu TD, Modrusan Z, Swain SM. Gene expression profile and angiogenic marker correlates with response to neoadjuvant bevacizumab followed by bevacizumab plus chemotherapy in breast cancer. Clin Cancer Res 2008;14: 5893-9.
51. Hansen TF, Sorensen FB, Spindler KL*, et al.* Microvessel density and the association with single nucleotide polymorphisms of the vascular endothelial growth factor receptor 2 in patients with colorectal cancer. Virchows Arch 2010;456: 251-60.
52. Zeng Y, Sun HR, Yu C*, et al.* CXCR1 and CXCR2 are novel mechano-sensors mediating laminar shear stress-induced endothelial cell migration. Cytokine 2011;53: 42-51.
53. Mizukami Y, Jo WS, Duerr EM*, et al.* Induction of interleukin-8 preserves the angiogenic response in HIF-1alpha-deficient colon cancer cells. Nat Med 2005;11: 992-7.
54. Cascone T, Herynk MH, Xu L*, et al.* Upregulated stromal EGFR and vascular remodeling in mouse xenograft models of angiogenesis inhibitor-resistant human lung adenocarcinoma. J Clin Invest 2011.
55. Moriai T, Kobrin MS, Hope C, Speck L, Korc M. A variant epidermal growth factor receptor exhibits altered type alpha transforming growth factor binding and transmembrane signaling. Proc Natl Acad Sci U S A 1994;91: 10217-21.
56. Lanuti M, Liu G, Goodwin JM*, et al.* A functional epidermal growth factor (EGF) polymorphism, EGF serum levels, and esophageal adenocarcinoma risk and outcome. Clin Cancer Res 2008;14: 3216-22.
57. Pander J, Gelderblom H, Antonini NF*, et al.* Correlation of FCGR3A and EGFR germline polymorphisms with the efficacy of cetuximab in KRAS wild-type metastatic colorectal cancer. Eur J Cancer 2010;46: 1829-34.

58. van Cruijsen H, Giaccone G, Hoekman K. Epidermal growth factor receptor and angiogenesis: Opportunities for combined anticancer strategies. Int J Cancer 2005;117: 883-8.
59. Verheus M, McKay JD, Kaaks R*, et al.* Common genetic variation in the IGF-1 gene, serum IGF-I levels and breast density. Breast Cancer Res Treat 2008;112: 109-22.
60. Jiang ZY, He Z, King BL*, et al.* Characterization of multiple signaling pathways of insulin in the regulation of vascular endothelial growth factor expression in vascular cells and angiogenesis. J Biol Chem 2003;278: 31964-71.
61. Zhao X, Dai W, Zhu H*, et al.* Epidermal growth factor (EGF) induces apoptosis in a transfected cell line expressing EGF receptor on its membrane. Cell Biol Int 2006;30: 653-8.
62. Tikhomirov O, Carpenter G. Ligand-induced, p38-dependent apoptosis in cells expressing high levels of epidermal growth factor receptor and ErbB-2. J Biol Chem 2004;279: 12988-96.
63. Chiu B, Mirkin B, Madonna MB. Mitogenic and apoptotic actions of epidermal growth factor on neuroblastoma cells are concentration-dependent. J Surg Res 2006;135: 209-12.
64. Graziano F, Ruzzo A, Loupakis F*, et al.* Pharmacogenetic profiling for cetuximab plus irinotecan therapy in patients with refractory advanced colorectal cancer. J Clin Oncol 2008;26: 1427-34.
65. Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet 2006;7: 61-80.
66. Mandola MV, Stoehlmacher J, Zhang W*, et al.* A 6 bp polymorphism in the thymidylate synthase gene causes message instability and is associated with decreased intratumoral TS mRNA levels. Pharmacogenetics 2004;14: 319-27.
67. Du L, Wang H, He L*, et al.* CD44 is of functional importance for colorectal cancer stem cells. Clin Cancer Res 2008;14: 6751-60.
68. Levin TG, Powell AE, Davies PS*, et al.* Characterization of the intestinal cancer stem cell marker CD166 in the human and mouse gastrointestinal tract. Gastroenterology 2010;139: 2072-82 e5.
69. Horst D, Kriegl L, Engel J, Kirchner T, Jung A. Prognostic significance of the cancer stem cell markers CD133, CD44, and CD166 in colorectal cancer. Cancer Invest 2009;27: 844-50.
70. Weichert W, Knosel T, Bellach J, Dietel M, Kristiansen G. ALCAM/CD166 is overexpressed in colorectal carcinoma and correlates with shortened patient survival. J Clin Pathol 2004;57: 1160-4.
71. Lugli A, Iezzi G, Hostettler I*, et al.* Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD44s, EpCAM, and ALDH1 in colorectal cancer. Br J Cancer 2010;103: 382-90.
72. Zhou P, Du LF, Lv GQ*, et al.* Functional polymorphisms in CD166/ALCAM gene associated with increased risk for breast cancer in a Chinese population. Breast Cancer Res Treat 2011.
73. Uchida H, Yamazaki K, Fukuma M*, et al.* Overexpression of leucine-rich repeat-containing G protein-coupled receptor 5 in colorectal cancer. Cancer Sci 2010;101: 1731-7.
74. Takahashi H, Ishii H, Nishida N*, et al.* Significance of Lgr5(+ve) Cancer Stem Cells in the Colon and Rectum. Ann Surg Oncol 2010;18: 1166-74.
75. Deng S, Yang X, Lassus H*, et al.* Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers. PLoS One 2010;5: e10277.

76. Jiang L, Zhang C, Li Y*, et al.* A non-synonymous polymorphism Thr115Met in the EpCAM gene is associated with an increased risk of breast cancer in Chinese population. Breast Cancer Res Treat 2011;126: 487-95.

Figure 1

Overall survival by (A) EGF rs4444903 A>G and (B) IGF-1 rs6220 A>G

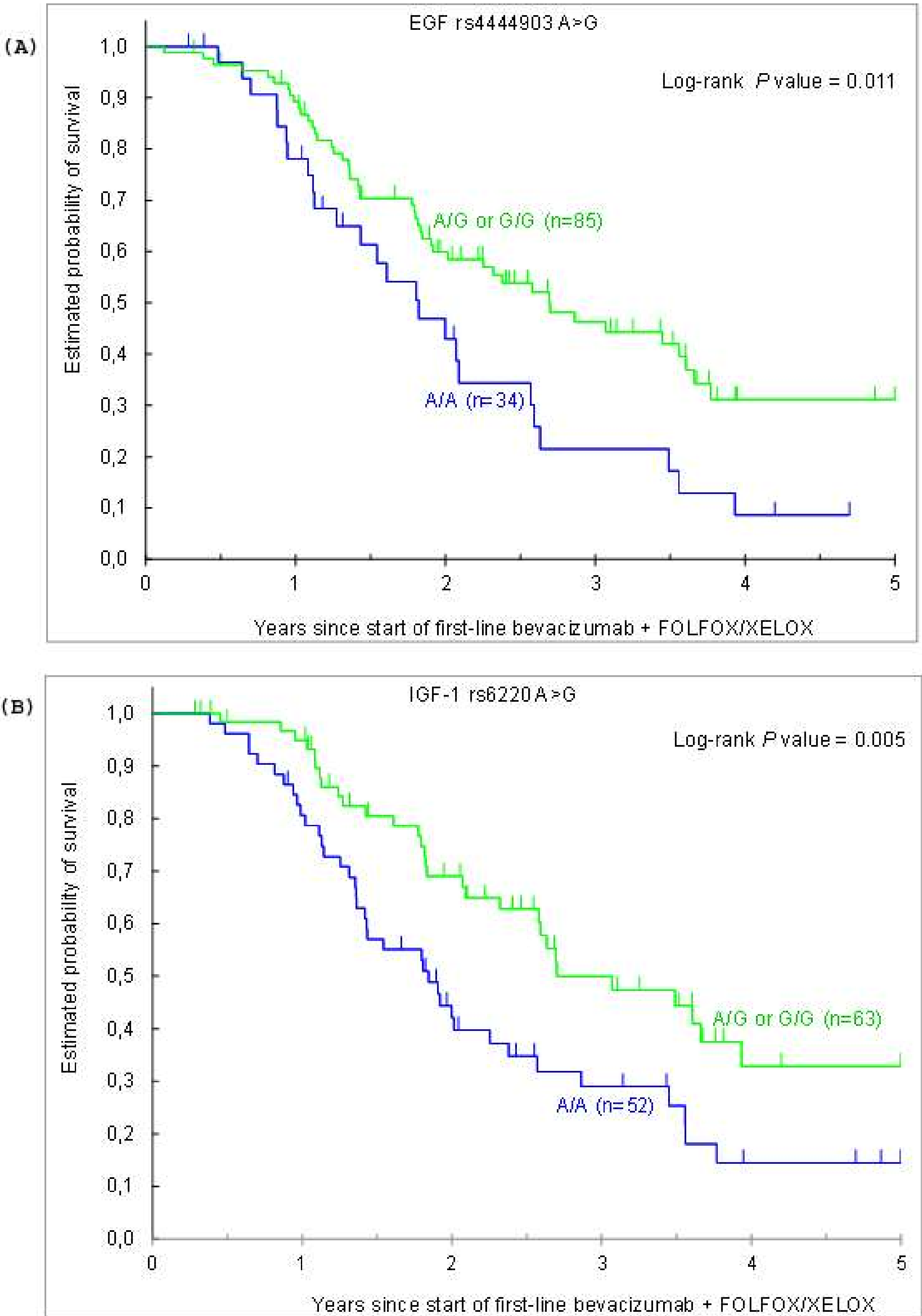
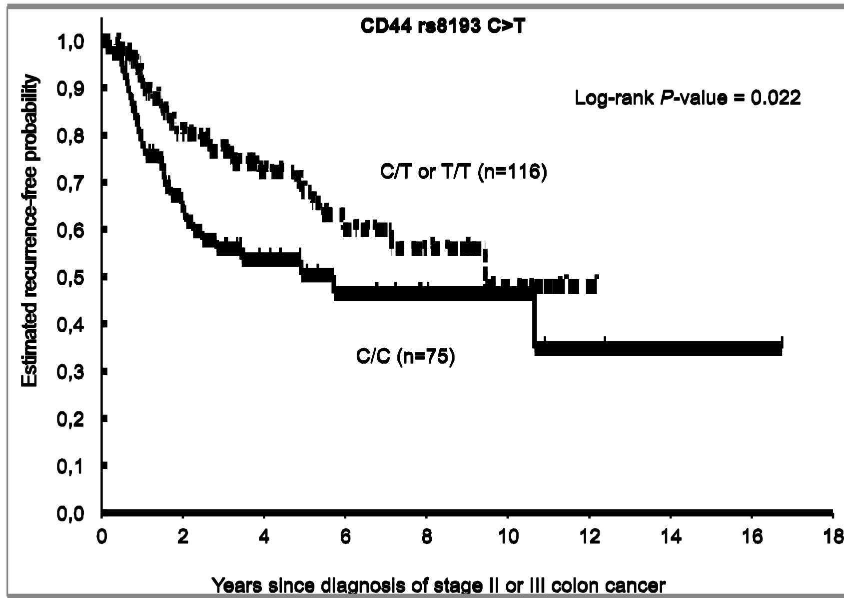


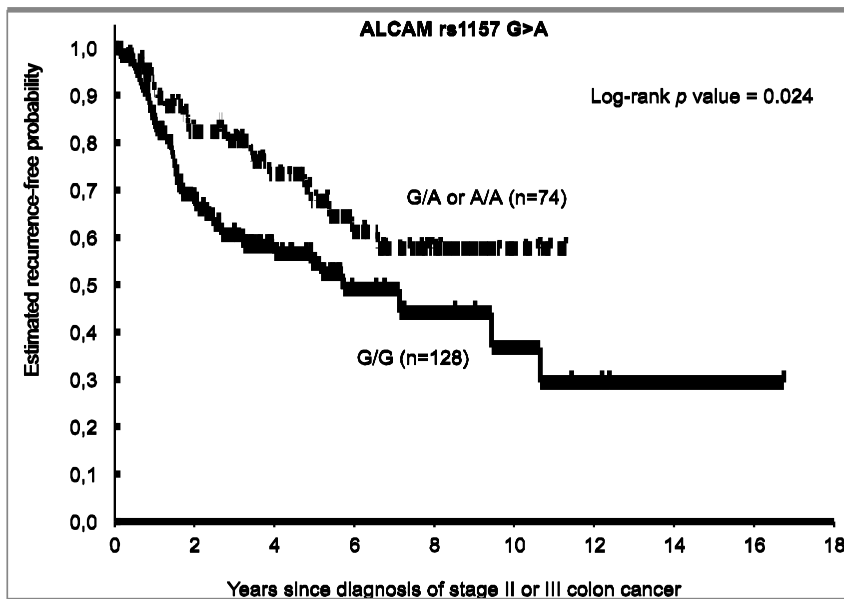
Figure 2

Time to tumor recurrence by (A) CD44 rs8193 C>T, (B) ALCAM rs1157 G>A and (C) Lgr5 rs17109924 T>C

A.



B.



C.

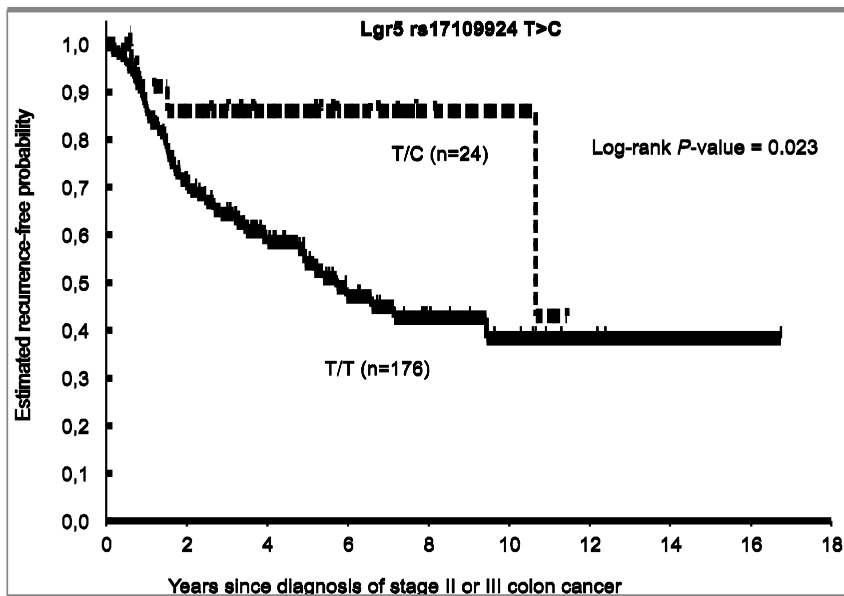
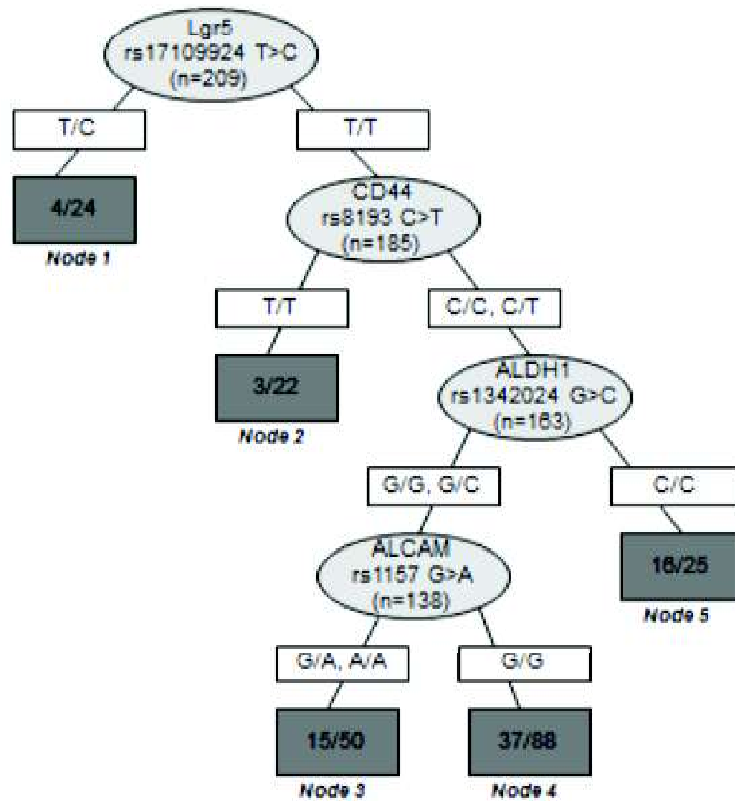


Figure 3

(A) RPart analysis of TTR. The end-nodes of the tree model represent subgroups of low- and high-risk patients based on either a single gene variant or combination of gene variants. Fractions within the end-nodes indicate patients who recurred/ total patients with this gene variant profile. (B) TTR by tree model defined subgroups. Node 5 represents a high-risk subgroup based on a specific gene variant profile including LGR5 rs17109924, CD44 rs8193 and ALDH1A1 rs1342024.

A.



B.

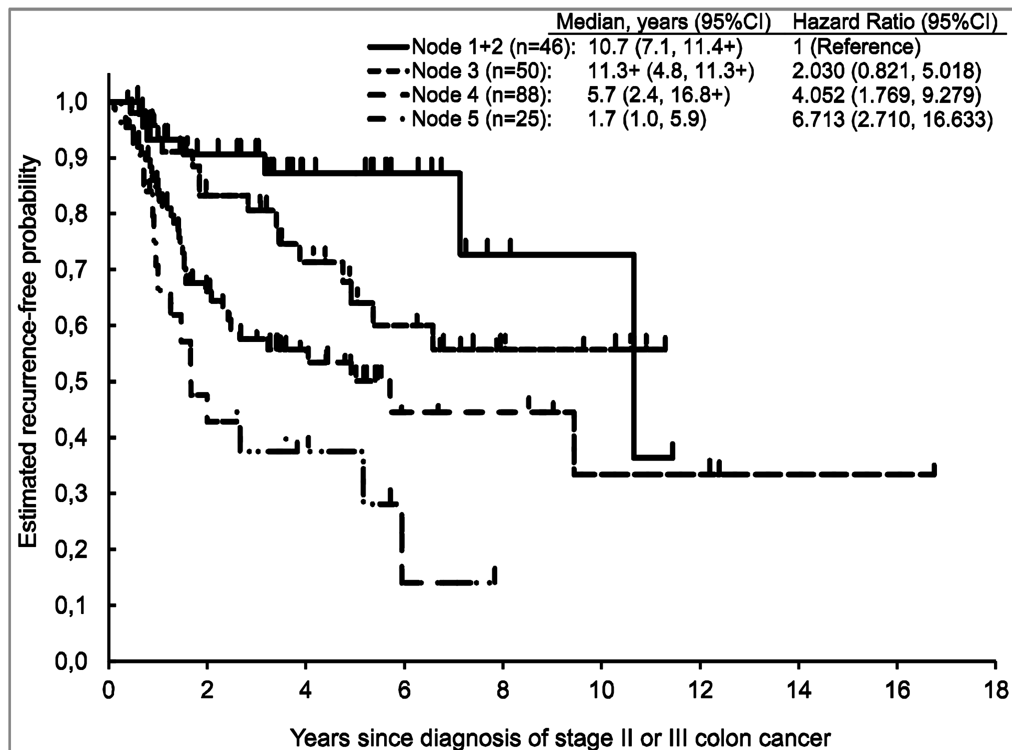


Table 1

Analyzed VEGF-dependent angiogenesis gene polymorphisms

Gene	rs-number	Base exchange	Function	Genotyping
VEGF-A	rs2010963	C>T	C lower VEGF plasma level	DS
	rs3025039	C>T	T lower VEGF plasma level	RE (NiaIII)
	rs1570360	G>A	A lower VEGF plasma level	DS
	rs833061	T>C	T lower BV associated hypertension	DS
	rs699947	C>A	C higher VEGF serum level	RE (BglII)
VEGFR-2	no rs-number	11-14 CA repeat	11/11 higher gene expression	γATP-labeled PCR
	rs2305948	C>T	T higher MVD	DS
	rs2071559	T>C	C higher MVD	RE (BsmI)
HIF1α	rs1154946	C>T	T higher HIF1α protein expression	RE (Tsp451)
ARNT	rs2228099	G>C	NA	DS
NRP1	rs3750733	C>T	NA	DS

Abbreviations: DS, direct DNA sequencing; RE, restriction enzyme; BC, breast cancer; MVD, microvessel density; NA, not analyzed; OS, overall survival; VEGF, vascular endothelial growth factor; VEGFR, VEGF-receptor; HIF, hypoxia inducible factor; ARNT, aryl hydrocarbon receptor nuclear translocator; NRP, neuropilin

Table 2

Analyzed VEGF-independent angiogenesis gene polymorphisms

Gene	rs-number	Base exchange	Function	Genotyping
IL-1β	rs16944	C>T	T higher IL-1β plasma level	RE (AvaI)
	rs1143634	C>T	T higher IL-1β plasma level	RE (TaqI)
IL-6	rs1800795	G>C	C lower IL-6 plasma level	RE (NiaIII)
IL-8	rs4073	T>A	A increased IL-8 plasma level	RE (MfeI)
CXCR1	rs2234671	G>C	NA	DS
CXCR2	rs2230054	T>C	NA	DS
Leptin	rs7799039	G>A	A higher Leptin serum level	RE (HhaI)
TF	rs1361600	A>G	G higher TF gene expression	RE (BstNI)
ES	rs12483377	G>A	A lower ES function	RE (MseI)
FGFR-4	rs351855	G>A	A higher FGFR4 gene expression	RE (BstNI)
IGF-1	rs6214	C>T	NA	NiaIII
	rs6220	A>G	G higher IGF plasma level	MnlI
IGF-2	rs10840452	G>A	NA	DS
IGFR1	rs2229765	G>A	A associated with lower IGF-1 plasma levels	RE (MnlI)
NF-κB	no rs-number	18-26 CA repeat	NA	γATP-labeled PCR

Cox-2	rs5275	T>C	C allele associated with lower promoter activity	DS
TNF-α	rs361525	G>A	A higher TNF-α plasma level	DS
TNF-β	rs909253	A>G	NA	RE (HinfI)
ICAM-1	rs5498	T>C	A lower ICAM-1 plasma level	RE (BstUI)
MMP-2	rs243865	C>T	T lower promoter activity	DS
MMP-7	rs1156881	A>G	G higher promoter activity	DS
EGF	rs4444903	A>G	G higher EGF serum levels	RE (AluI)
EGFR	rs2227983	G>A	A attenuated EGFR ligand binding	RE (BSTNI)

Abbreviations: BC, breast cancer; CRC, colorectal cancer; TTR, time to tumor recurrence; IL, interleukin; CXCR, interleukin receptor; TF, tissue factor; ES, endostatin; FGFR, fibroblast growth factor receptor; IGF, insulin like growth factor; IGFR, insulin like growth factor receptor; NF-κB, nuclear factor-κB; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; COX, cyclooxygenase; TNF, tumor necrosis factor; ICAM, inter-cellular adhesion molecule; MMP, matrix metalloproteinases

Table 3

Analyzed CSC gene polymorphisms

Gene	rs-number	Base exchange	Region	Genotyping
CD44	rs8193	C>T	3UTR	RE (BsrDI)
	rs187116	A>G	Intron	RE (MspI)
	rs4755392	T>A	3UTR	DS
	rs7116432	A>G	3UTR	RE (NlaIII)
Prominin-1	rs3130	A>G	3UTR	RE (EcorRI)
	rs2240688	A>C	3UTR	DS
	rs2286455	C>T	Splice Site	RE (MboI)
DPP4	rs2300757	G>C	Intron	RE (TfiI)
	rs1014444	A>G	Intron	RE (AluI)
	rs2268894	A>G	Intron	DS
EpCAM	rs17036526	G>C	Splice site	RE (DdeI)
	rs1126497	C>T	Non-synonymous coding	RE (NiaIII)
	rs1421	T>C	3UTR	DS
ALCAM	rs6437585	C>T	5UTR	RE (Cac8I)
	rs1044240	A>G	Non-synonymous coding	DS
	rs1044243	G>A	Non-synonymous coding	DS
	rs1157	G>A	3UTR	RE (MspI)
MSI-1	rs2522137	A>C	3UTR	DS
ITGB1	rs2153875	T>G	Splice site	RE (BfaI)
CD24	rs8734	C>T	Non-	RE (AciI)

			synonymous coding	
	rs3838646	-/CA	3UTR	RE (BsrI)
LGR5	rs17109924	T>C	Non-synonymous coding	DS
	rs17109926	G>A	3UTR	DS
ALDH1A1	rs13959	G>A	Synonymous coding	RE (Hypch4III)
	rs1342024	G>C	Upstream	DS

Abbreviations: DS, direct DNA sequencing; RE, restriction enzyme; UTR, untranslated region; DPP4, dipeptidyl peptidase-4; EpCAM, epithelial cell adhesion molecule; ALCAM, activated leukocyte cell adhesion molecule; MSI-1, musashi homolog-1; ITGB1, integrin beta-1; LGR5, leucine-rich repeat containing G protein-coupled receptor-5; ALDH1A1, aldehyde dehydrogenase-1 family member A1; nd, no data

Table 4

Baseline patient characteristics and the association with clinical outcome and RR (Ethnicity was self-defined by the patients)

	n	%	RR		PFS	OS
			CR/PR	**SD/PD**	**HR (95%CI)**	**HR (95%CI)**
Sex						
Female	57	43	23(41%)	33(59%)	1 (Reference)	1 (Reference)
Male	75	57	41 (58%)	30 (42%)	0.64 (0.44, 0.93)	0.78 (0.50, 1.22)
P value			0.075		0.016	0.27
Age, years						
<55	53	40	23 (44%)	29 (56%)	1 (Reference)	1 (Reference)
55-64	48	36	24 (55%)	20 (45%)	0.73 (0.47, 1.12)	0.80 (0.48, 1.33)
65+	31	23	17 (55%)	14 (45%)	0.90 (0.56, 1.46)	0.96 (0.55, 1.68)
P value			0.51		0.33	0.66
Ethnicity						
African American	5	4	1 (20%)	4 (80%)	1.41 (0.56, 3.55)	2.31 (0.80, 6.65)
Asian	27	20	14 (52%)	13 (48%)	1.04 (0.65, 1.68)	1.31 (0.76, 2.27)
Caucasian	65	49	29 (48%)	32 (52%)	1 (Reference)	1 (Reference)
Hispanic	35	27	20 (59%)	14 (41%)	0.83 (0.52, 1.30)	0.70 (0.39, 1.25)
P value			0.41		0.65	0.086
Karnofsky performance status%						
100	17	13	7 (41%)	10 (59%)	1 (Reference)	1 (Reference)
90	43	33	27 (64%)	15 (36%)	1.60 (0.84, 3.06)	1.33 (0.64, 2.80)
80	22	17	10 (45%)	12 (55%)	1.98 (0.97, 4.02)	2.38 (1.03, 5.53)
Missing	50	37				
P value			0.17		0.13	0.066
Primary tumor site						
Colon	87	66	45 (54%)	38 (46%)	1 (Reference)	1 (Reference)
Rectosigmoid	13	10	3 (25%)	9 (75%)	1.28 (0.71, 2.31)	0.96 (0.43, 2.11)
Rectum	32	24	16 (50%)	16 (50%)	1.37 (0.88, 2.11)	1.21 (0.73, 2.01)

P value			0.18		0.32	0.74
Histologic differentiation						
Well	2	1	2 (100%)	0 (0%)		
Moderate	94	72	47 (52%)	44 (48%)	1 (Reference)*	1 (Reference)*
Poor	31	23	14 (48%)	15 (52%)	1.07 (0.68, 1.68)	1.63 (0.98, 2.70)
Missing	5	4				
P value			0.60		0.76	0.056
Tumor stage at diagnosis						
Stage I	2	1	1 (50%)	1 (50%)		
Stage II	6	5	3 (50%)	3 (50%)		
Stage III	13	10	6 (46%)	7 (54%)	1 (Reference)*	1 (Reference)*
Stage IV	110	83	53 (50%)	52 (50%)	1.39 (0.82, 2.36)	2.11 (1.08, 4.10)
Missing	1	1				
P value			1.00		0.22	0.022
Metastatic site, n						
1 site	76	58	44 (58%)	31 (43%)	1 (Reference)	1 (Reference)
2+ sites	56	42	20 (38%)	32 (62%)	1.38 (0.95, 2.01)	1.72 (1.10, 2.70)
P value			0.031		0.085	0.013
Metastatic site, localization						
Liver	98	74	55 (57%)	42 (43%)	0.76 (0.50, 1.17)	0.94 (0.56, 1.58)
			0.012		0.21	0.82
Lung	31	23	8 (26%)	23 (74%)	1.42 (0.92, 2.17)	1.57 (0.98, 2.53)
			0.002		0.10	0.059
Peritoneum	23	17	7 (35%)	13 (65%)	1.18 (0.73, 1.90)	1.43 (0.81, 2.51)
			0.15		0.50	0.21
Other	32	24	10 (34%)	19 (66%)	0.95 (0.62, 1.47)	1.17 (0.70, 1.96)
P value			0.059		0.83	0.55
Chemotherapy backbone						
FOLFOX-4	66	50	23 (37%)	39 (63%)	1 (Reference)	1 (Reference)
XELOX	66	50	41 (63%)	24 (37%)	0.98 (0.68, 1.42)	1.07 (0.69, 1.67)
P value			0.005		0.91	0.75
Hospital						

USC		107	81	57 (55%)	46 (45%)	1 (Reference)	1 (Reference)
MUG		25	19	7 (29%)	17 (71%)	1.75 (1.09, 2.81)	0.84 (0.44, 1.58)
	P value			0.025		0.016	0.59

* Combined with the above categories for stable estimates.

Abbreviations: USC, University of Southern California; MUG: Medical University of Graz; n, number; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; RR, response rate; PFS, progression-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval.

Table 5

Univariate analysis of polymorphisms for PFS, OS and RR. PFS and OS were calculated for dominant models (combining heterozygous and homozygous minor variants)

	n	RR		PFS		OS	
		CR or PR	SD or PD	Median months (95%CI)	HR (95%CI)	Median months (95%CI)	HR (95%CI)
CXCR1 rs2234671							
G/G	51	35 (71%)	14 (29%)	13.9 (11.0, 16.0)	1 (Reference)	28.6 (21.5, 36.9)	1 (Reference)
G/C	62	22 (37%)	38 (63%)	8.1 (7.3, 10.4)	1.38 (0.93, 2.05)	27.9 (21.6, 41.9)	0.95 (0.60, 1.52)
C/C	6	1 (17%)	5 (83%)				
P			<0.001		0.10		0.83
CXCR2 rs2230054							
T/T	49	18 (38%)	29 (62%)	8.1 (7.1, 10.4)	1 (Reference)	24.2 (17.2, 34.4)	1 (Reference)
T/C	53	29 (56%)	23 (44%)	12.4 (10.1, 15.0)	0.89 (0.60, 1.33)	28.6 (22.0, 42.7)	0.88 (0.54, 1.42)
C/C	15	11 (79%)	3 (21%)				
P			0.008		0.57		0.59
EGFR rs222798							

G/G	64	27 (43%)	36 (57%)	8.3 (7.1, 11.2)	1 (Reference)	23.0 (19.3, 31.6)	1 (Reference)
G/A	43	22 (55%)	18 (45%)	12.4 (8.4, 17.0)	0.74 (0.50, 1.10)	32.4 (22.0, 43.3)	0.74 (0.46, 1.19)
A/A	11	9 (82%)	2 (18%)				
P			0.024		0.13		0.21
VEGFR-2 rs2305948							
C/C	92	50 (57%)	38 (43%)	10.1 (7.9, 12.4)	1 (Reference)	31.0 (22.9, 42.7)	1 (Reference)
C/T	24	7 (29%)	17 (71%)	10.8 (7.6, 16.0)	1.18 (0.74, 1.87)	22.0 (17.1, 32.4)	1.36 (0.79, 2.32)
T/T	3	1 (33%)	2 (67%)				
P			0.024		0.48		0.26
HIF1α rs11549465							
C/C	99	49 (52%)	46 (48%)	10.2 (7.9, 12.4)	1 (Reference)	27.9 (22.0, 34.4)	1 (Reference)
C/T	18	8 (44%)	10 (56%)	11.7 (6.3, 36.0+)	0.55 (0.31, 0.99)	32.4 (17.1, 60.0+)	0.81 (0.44, 1.52)
T/T	2	1 (50%)	1 (50%)				
P			0.68		0.038		0.51
EGF rs444903							
A/A	34	14	18	7.1 (5.7,	1	21.9	1

		(44%)	(56%)	13.2)	(Reference)	(13.5, 30.9)	(Reference)
A/G	55	32 (59%)	22 (41%)	11.1 (8.3, 13.7)	0.72 (0.47, 1.10)	32.4 (22.9, 43.3)	0.54 (0.33, 0.88)
G/G	30	12 (41%)	17 (59%)				
P			1.00		0.13		0.011
IGF-1 rs6220							
A/A	52	24 (46%)	28 (54%)	8.4 (7.4, 13.2)	1 (Reference)	22.1 (16.4, 28.6)	1 (Reference)
A/G	51	29 (60%)	19 (40%)	11.6 (8.1, 14.2)	0.92 (0.62, 1.37)	32.4 (27.9, 47.2)	0.51 (0.32, 0.83)
G/G	12	4 (33%)	8 (67%)				
P			1.00		0.68		0.005

Abbreviations: RR, response rate; PFS, progression-free survival; OS, overall survival; CXCR, interleukin receptor; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; VEGF, vascular endothelial growth factor; VEGFR, VEGF-receptor; IGF, insulin like growth factor; CR, complete response; PR, partial response; CI, confidence interval; HR, hazard ratio

Table 6

Baseline patient characteristics

	n	%
Sex		
Female	107	45.72
Male	127	54.28
Ethnicity		
Asian	34	14.53
African American	15	6.41
Caucasian	123	52.56
Hispanic	62	26.5
T		
T1	2	0.85
T2	14	5.98
T3	187	79.91
T4	27	11.54
Tx	4	1.72
Grade		
Well	11	2.18
Moderate	151	64.53
Poor/undifferentiated	54	23.08
Missing	18	10.21
N		
Negative	105	44.87
N1	72	30.77
N2	57	24.36
Stage		
High-risk II	105	44.87
III	129	55.13
N of resected lymph nodes		

≤12	70	29.91
>12	145	61.97
Missing	19	8.12
Tumor side		
Left	110	47.01
Right	115	49.15
Left and right	4	1.71
Missing	5	2.13
Adjuvant treatment		
5-FU	151	64.53
5-FU/LV/Oxaliplatin	60	25.64
5-FU/LV/Irinotecan	23	9.83

Table 7

Univariate and multivariate analysis of polymorphisms and TTR

	N	Univariate analysis				Multivariate analysis	
		Median TTR, yrs (95%CI)	Probability ± SE* of 3-year recurrence	Hazard Ratio (95%CI)	*P* value †	Hazard Ratio (95%CI)	*P* value ‡
CD44 rs8193					0.022		0.047
C/C	75	5.4 (2.1, 16.8)	0.44 ± 0.06	1 (Reference)		1 (Reference)	
C/T§	92	9.4 (5.9, 12.2)	0.23 ± 0.04	0.51 (0.35, 0.93)		0.60 (0.36, 0.99)	
T/T§	24						
CD44 rs187116					0.31		0.20
A/A	61	12.4 (9.4, 12.4)	0.25 ± 0.06	1 (Reference)		1 (Reference)	
A/G	94	10.7 (4.9, 11.3)	0.33 ± 0.05	1.40 (0.78, 2.51)		1.43 (0.79, 2.59)	
G/G	45	4.9 (3.5, 16.8)	0.34 ± 0.08	1.65 (0.85, 3.21)		1.92 (0.93, 3.95)	
CD44 rs4755392					0.94		0.74
T/T	48	9.4 (4.9, 16.8)	0.28 ± 0.07	1 (Reference)		1 (Reference)	
T/A	97	7.1 (4.9, 11.4)	0.33 ± 0.05	1.10 (0.61, 1.98)		1.22 (0.67, 2.21)	
A/A	57	12.4 (3.4, 12.4)	0.36 ± 0.07	1.04 (0.53, 2.02)		1.02 (0.51, 2.01)	
CD44 rs7116432					0.48		0.61
A/A	61	5.7 (2.5, 16.8)	0.40 ± 0.07	1 (Reference)		1 (Reference)	
A/G	87	9.4 (5.2, 9.6)	0.28 ± 0.05	0.75 (0.44, 1.30)		0.80 (0.45, 1.41)	
G/G	50	10.3 (4.9, 10.3)	0.23 ± 0.06	0.73 (0.38, 1.40)		0.72 (0.35, 1.47)	
Prominin-1 rs3130					0.59		0.40
A/A	124	7.1 (5.2, 12.2)	0.31 ± 0.05	1 (Reference)		1 (Reference)	
A/G	80	9.4 (3.4, 16.8)	0.34 ± 0.06	1.14 (0.71, 1.81)		1.24 (0.76, 2.03)	
Prominin-1 rs2240688					0.86		0.91
A/A	104	6.6 (4.9, 10.7)	0.27 ± 0.05	1 (Reference)		1 (Reference)	
A/C§	89	12.4 (3.2, 12.4)	0.38 ± 0.05	1.04 (0.66, 1.65)		0.97 (0.61, 1.56)	
C/C§	10						
Prominin-1 rs2286455					0.45		0.21
C/C	155	10.7 (5.7, 12.4)	0.32 ± 0.04	1 (Reference)		1 (Reference)	
C/T§	41	5.2 (3.9, 16.8)	0.33 ± 0.07	1.22 (0.73, 2.04)		1.41 (0.83, 2.40)	
T/T§	5						
DPP4 rs2300757					0.79		0.71
G/G	75	5.4 (3.4, 16.8)	0.35 ± 0.06	1 (Reference)		1 (Reference)	
G/C	96	10.7 (5.9, 12.2)	0.30 ± 0.05	0.85 (0.52, 1.41)		0.87 (0.52, 1.45)	
C/C	30	5.7 (2.4, 11.3)	0.33 ± 0.10	1.00 (0.50, 2.01)		0.75 (0.37, 1.54)	
DPP4 rs1014444					0.44		0.41
A/A	68	5.2 (3.4, 16.8)	0.35 ± 0.06	1 (Reference)		1 (Reference)	
A/G	92	10.7 (5.9, 12.2)	0.29 ± 0.05	0.72 (0.43, 1.20)		0.73 (0.43, 1.24)	

G/G	42	11.3 (2.4, 11.3)	0.37 ± 0.08	0.90 (0.48, 1.68)		0.70 (0.37, 1.35)	
DPP4 rs2268894					0.81		0.79
A/A	55	5.7 (2.4, 11.3)	0.37 ± 0.07	1 (Reference)		1 (Reference)	
A/G	98	10.7 (5.4, 12.2)	0.32 ± 0.05	0.84 (0.49, 1.46)		1.18 (0.67, 2.10)	
G/G	44	5.2 (3.4, 16.8)	0.30 ± 0.08	0.95 (0.51, 1.79)		1.22 (0.64, 2.33)	
EpCAM rs17036526					0.95		0.73
G/G	34	7.1 (2.5, 7.1)	0.32 ± 0.09	1 (Reference)		1 (Reference)	
G/C	152	9.4 (5.2, 16.8)	0.32 ± 0.04	1.03 (0.54, 1.97)		0.99 (0.50, 1.95)	
C/C	13	4.9 (1.7, 10.9)	0.38 ± 0.15	1.19 (0.41, 3.43)		1.45 (0.48, 4.34)	
EpCAM rs1126497					0.69		0.56
C/C	50	7.1 (3.2, 16.8)	0.32 ± 0.07	1 (Reference)		1 (Reference)	
C/T	104	5.9 (4.0, 10.7)	0.34 ± 0.05	1.00 (0.57, 1.75)		0.92 (0.52, 1.64)	
T/T	49	9.4 (4.9, 12.2)	0.30 ± 0.07	0.79 (0.40, 1.54)		0.70 (0.34, 1.41)	
EpCAM rs1421					0.10		0.24
T/T	157	6.6 (3.9, 16.8)	0.37 ± 0.04	1 (Reference)		1 (Reference)	
T/C§	45	7.1 (5.9, 12.4)	0.18 ± 0.06	0.60 (0.32, 1.12)		0.68 (0.36, 1.29)	
C/C§	2						
ALCAM rs6437585					0.080		0.16
C/C	175	9.4 (5.4, 16.8)	0.29 ± 0.04	1 (Reference)		1 (Reference)	
C/T	26	2.4 (1.5, 10.3)	0.50 ± 0.11	1.73 (0.93, 3.21)		1.62 (0.83, 3.17)	
ALCAM rs1044240					0.95		0.77
A/A	166	9.4 (5.4, 16.8)	0.32 ± 0.04	1 (Reference)		1 (Reference)	
A/G§	31	5.2 (3.2, 12.4)	0.31 ± 0.08	1.02 (0.57, 1.82)		0.91 (0.50, 1.68)	
G/G§	6						
ALCAM rs1044243					0.30		0.18
G/G	170	6.6 (4.9, 16.8)	0.34 ± 0.04	1 (Reference)		1 (Reference)	
G/A	32	10.9 (4.8, 10.9)	0.22 ± 0.08	0.70 (0.36, 1.38)		0.62 (0.31, 1.25)	
ALCAM rs1157					0.024		0.027
G/G	128	5.7 (3.2, 10.7)	0.39 ± 0.05	1 (Reference)		1 (Reference)	
G/A§	67	11.3 (5.9, 11.3)	0.20 ± 0.05	0.56 (0.33, 0.94)		0.55 (0.32, 0.93)	
A/A§	7						
MSI-1 rs2522137					0.59		0.41
A/A	65	5.7 (4.0, 12.4)	0.30 ± 0.06	1 (Reference)		1 (Reference)	
A/C	85	7.1 (3.2, 12.2)	0.35 ± 0.06	0.87 (0.51, 1.48)		0.80 (0.46, 1.37)	
C/C	51	16.8 (4.9, 16.8)	0.31 ± 0.07	0.73 (0.39, 1.34)		0.65 (0.35, 1.23)	
ITGB1 rs2153875					0.81		0.66
T/T	98	12.4 (5.2, 12.4)	0.32 ± 0.05	1 (Reference)		1 (Reference)	
T/G	89	6.6 (4.9, 10.7)	0.32 ± 0.05	1.14 (0.70, 1.85)		1.23 (0.74, 2.02)	
G/G	15	4.9 (2.7, 16.8)	0.30 ± 0.12	1.24 (0.55, 2.82)		1.34 (0.57, 3.15)	
CD24 rs8734					0.62		0.33
C/C	81	5.4 (3.9, 16.8)	0.33 ± 0.06	1 (Reference)		1 (Reference)	
C/T	62	6.6 (4.9, 10.9)	0.31 ± 0.06	0.85 (0.50, 1.46)		0.74 (0.43, 1.30)	
T/T	56	12.4 (5.2, 12.4)	0.33 ± 0.07	0.75 (0.42, 1.36)		0.65 (0.35, 1.19)	

CD24 rs3838646					0.39		0.53
(CA)1	182	9.4 (4.9, 16.8)	0.35 ± 0.04	1 (Reference)		1 (Reference)	
(CA)2§	16	7.1 (5.9, 7.2)	0.12 ± 0.08	0.67 (0.27, 1.67)		0.73 (0.28, 1.90)	
(CA)3§	1						
LGR5 rs17109924					0.023		0.035
T/T	176	5.7 (4.0, 16.8)	0.36 ± 0.04	1 (Reference)		1 (Reference)	
T/C	24	10.7 (10.7, 11.4)	0.14 ± 0.08	0.33 (0.12, 0.90)		0.33 (0.12, 0.93)	
LGR5 rs17109926					0.92		0.86
G/G	123	6.6 (5.2, 11.4)	0.31 ± 0.05	1 (Reference)		1 (Reference)	
G/A§	73	10.7 (3.4, 16.8)	0.35 ± 0.06	1.02 (0.64, 1.64)		1.05 (0.64, 1.70)	
A/A§	5						
ALDH1A1 rs13959					0.66		0.45
G/G	61	5.4 (4.8, 9.4)	0.25 ± 0.06	1 (Reference)		1 (Reference)	
G/A	98	10.7 (2.7, 16.8)	0.41 ± 0.06	0.97 (0.57, 1.63)		0.96 (0.56, 1.64)	
A/A	41	11.3 (3.4, 11.3)	0.25 ± 0.07	0.74 (0.38, 1.46)		0.65 (0.32, 1.33)	
ALDH1A1 rs1342024					0.18		0.37
G/G	79	9.4 (4.9, 11.4)	0.21 ± 0.05	1 (Reference)		1 (Reference)	
G/C	86	16.8 (3.2, 16.8)	0.38 ± 0.06	1.20 (0.71, 2.02)		1.04 (0.60, 1.78)	
C/C	35	5.2 (1.7, 7.8)	0.46 ± 0.09	1.79 (0.95, 3.34)		1.54 (0.81, 2.93)	

* Greenwood SE

† Based on log-rank test

‡ Based on Wald test within Cox proportional hazards model

§ In the dominant model

Abbreviations: TTR, time to tumor recurrence; DPP4, dipeptidyl peptidase-4; EpCAM, epithelial cell adhesion molecule; ALCAM, activated leukocyte cell adhesion molecule; MSI-1, musashi homolog-1; ITGB1, integrin beta-1; LGR5, leucine-rich repeat containing G protein-coupled receptor-5; ALDH1A1, aldehyde dehydrogenase-1 family member A1; SE, standard error; CI, confidence interval; yrs, years.

Supplementary Table 1

Genotype frequencies for all polymorphisms in each ethnicity.

	Caucasian	African American	Asian	Hispanic	Total
CXCR2 rs2230054					
T/T	33	1	4	11	49
T/C	26	3	11	13	53
C/C	3	1	6	5	15
Total	62	5	21	29	117
ES rs12483377					
G/G	55	5	21	28	109
G/A	9	0	0	1	10
A/A	0	0	0	0	0
Total	64	5	21	29	119
ICAM-1 rs5498					
T/T	22	5	8	7	42
T/C	29	0	10	11	50
C/C	13	0	3	11	27
Total	64	5	21	29	119
IGFR1 rs2229765					
G/G	19	2	9	6	36
G/A	27	3	7	18	55
A/A	15	0	2	1	18
Total	61	5	18	25	109
IL-6 rs1800795					
G/G	25	5	20	19	69
G/C	26	0	0	8	34
C/C	13	0	1	1	15
Total	64	5	21	28	118
VEGFR-2 rs2305948					
C/C	50	1	15	26	92
C/T	13	3	5	3	24
T/T	1	1	1	0	3
Total	64	5	21	29	119
MMP-2 rs243865					
C/C	30	4	15	15	64
C/T	28	1	4	12	45
T/T	5	0	2	2	9
Total	63	5	21	29	118
TNF-β rs909253					

A/A	23	3	4	17	47
A/G	29	2	12	9	52
G/G	8	0	5	3	16
Total	60	5	21	29	115
VEGF-A rs699947					
C/C	15	3	14	12	44
C/A	31	2	5	11	49
A/A	16	0	2	5	23
Total	62	5	21	28	116
IGF-1 rs6220					
A/A	28	3	7	14	52
A/G	25	1	10	15	51
G/G	8	1	3	0	12
Total	61	5	20	29	115
COX-2 rs5275					
T/T	30	2	11	11	54
T/C	28	1	9	15	53
C/C	6	2	1	3	12
Total	64	5	21	29	119
EGF rs4444903					
A/A	20	1	5	8	34
A/G	37	1	8	9	55
G/G	7	3	8	12	30
Total	64	5	21	29	119
FGFR-4 rs351855					
G/G	25	3	4	9	41
G/A	29	1	10	14	54
A/A	10	1	7	6	24
Total	64	5	21	29	119
IGF-1 rs6214					
C/C	13	1	8	11	33
C/T	41	3	10	14	68
T/T	8	1	3	4	16
Total	62	5	21	29	117
IL-1β rs1143634					
C/C	36	5	20	18	79
C/T	21	0	1	8	30
T/T	7	0	0	2	9
Total	64	5	21	28	118
IL-8 rs4073					
T/T	21	0	8	7	36
T/A	33	4	9	18	64
A/A	10	1	4	4	19

Total	64	5	21	29	119
VEGFR-2 rs2071559					
T/T	11	0	9	14	34
T/C	31	4	12	9	56
C/C	18	1	0	6	25
Total	60	5	21	29	115
MMP-7 rs11568818					
A/A	26	1	19	16	62
A/G	28	2	2	11	43
G/G	10	2	0	2	14
total	64	5	21	29	119
TF rs1361600					
A/A	17	0	12	8	37
A/G	29	2	8	14	53
G/G	18	3	1	7	29
Total	64	5	21	29	119
VEGF-A rs1570360					
G/G	29	4	17	21	71
G/A	32	1	3	5	41
A/A	3	0	1	1	5
Total	64	5	21	27	117
VEGF-A rs2010963					
C/C	26	2	12	9	49
C/T	31	3	4	16	54
T/T	7	0	5	4	16
Total	64	5	21	29	119
CXCR1 rs2234671					
G/G	26	2	13	10	51
G/C	35	2	8	17	62
C/C	3	1	0	2	6
Total	64	5	21	29	119
EGFR rs2227983					
G/G	36	4	9	15	64
G/A	24	1	6	12	43
A/A	3	0	6	2	11
Total	63	5	21	29	118
HIF1α rs1154946					
C/C	49	5	18	27	99
C/T	13	0	3	2	18
T/T	2	0	0	0	2
Total	64	5	21	29	119
IGF2 rs10840452					

G/G	26	3	16	8	53
G/A	30	2	5	13	50
A/A	7	0	0	7	14
Total	63	5	21	28	117
IL-1β rs16944					
C/C	30	1	8	4	43
C/T	26	3	6	11	46
T/T	8	1	7	13	29
Total	64	5	21	28	118
VEGFR-2 11-14 CA repeat					
11/11	31	2	14	15	62
11/13	23	2	7	8	40
13/13	5	0	0	5	10
Total	59	4	21	28	112
Leptin rs7799039					
G/G	27	2	13	6	48
G/A	24	2	3	13	42
A/A	13	1	3	8	25
Total	64	5	19	27	115
NF-kB 18-26 CA repeat					
<24/<24	20	3	7	12	42
<24/≥24	23	2	10	10	45
≥24/≥24	10	0	2	1	13
Total	53	5	19	23	100
TNF-α rs361525					
G/G	59	5	19	27	110
G/A	5	0	1	2	8
A/A	0	0	1	0	1
Total	64	5	21	29	119
VEGF-A rs833061					
T/T	17	1	14	11	43
T/C	30	3	4	12	49
C/C	17	1	3	6	27
Total	64	5	21	29	119
VEGF-A rs3025039					
C/C	47	5	14	14	80
C/T	16	0	7	9	32
T/T	1	0	0	6	7
Total	64	5	21	29	119
NRP1 rs3750733					
C/C	38	2	4	8	52
C/T	20	2	6	13	41
T/T	6	1	11	6	24

Total	64	5	21	27	117
ARNT rs2228099					
G/G	22	1	7	10	40
G/C	27	4	10	11	52
C/C	12	0	4	8	24
Total	61	5	21	29	116

Abbreviations: VEGF, vascular endothelial growth factor; VEGFR, VEGF-receptor; HIF, hypoxia inducible factor; ARNT, aryl hydrocarbon receptor nuclear translocator; NRP, neuropilin; IL, interleukin; CXCR, interleukin receptor; TF, tissue factor; ES, endostatin; FGFR, fibroblast growth factor receptor; IGF, insulin like growth factor; IGFR, insulin like growth factor receptor; NF-κB, nuclear factor-κB; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; COX, cyclooxygenase; TNF, tumor necrosis factor; ICAM, inter-cellular adhesion molecule; MMP, matrix metalloproteinases