Carcinogenesis, 2021, Vol. 42, No. 3, 378–394

doi:10.1093/carcin/bgaa136
Advance Access publication December 14, 2020
Original Article

ORIGINAL ARTICLE

Genetic variations in microRNA-binding sites of solute carrier transporter genes as predictors of clinical outcome in colorectal cancer

Petra Bendova[1,2,3], Barbara Pardini[4,5], Simona Susova[3,6], Jachym Rosendorf[3], Miloslav Levy[7], Pavel Skrobanek[8], Tomas Buchler[8], Jan Kral[9], Vaclav Liska[3], Ludmila Vodickova[1,2,3], Stefano Landi[10], Pavel Soucek[3,6], Alessio Naccarati[4,5], Pavel Vodicka[1,2,3] and Veronika Vymetalkova[1,2,3,*]

[1]Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Videnska 1083, 14200 Prague, Czech Republic [2]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Albertov 4, 128 00 Prague, Czech Republic [3]Biomedical Centre and Department of Surgery, Faculty of Medicine in Pilsen, Charles University, Alej Svobody 76, 323 00 Pilsen, Czech Republic [4]IIGM Italian Institute for Genomic Medicine, SP142, Km3.95, Candiolo, Italy [5]Candiolo Cancer Institute, FPO-IRCCS, SP142, km 3.95Candiolo, Italy [6]Toxicogenomics Unit, National Institute of Public Health, Srobarova 48, Prague 10, Czech Republic [7]Department of Surgery, Thomayer University Hospital, Videnska 800, 140 59, Prague, Czech Republic [8]Department of Oncology, Thomayer Hospital, Videnska 800, 140 59, Prague, Czech Republic [9]Institute for Clinical and Experimental Medicine, IKEM, Prague, Czech Republic [10]Department of Biology, University of Pisa, Via Derna 1, 56126 Pisa, Italy

*To whom correspondence should be addressed. Department of Molecular Biology of Cancer, Videnska 1083, 14200 Prague, Czech Republic. Tel: +420 2 296 4 2699; Fax: +420 2 410 6 2782;Email: veronika.vymetalkova@iem.cas.cz

Abstract

One of the principal mechanisms of chemotherapy resistance in highly frequent solid tumors, such as colorectal cancer (CRC), is the decreased activity of drug transport into tumor cells due to low expression of important membrane proteins, such as solute carrier (SLC) transporters. Sequence complementarity is a major determinant for target gene recognition by microRNAs (miRNAs). Single-nucleotide polymorphisms (SNPs) in target sequences transcribed into messenger RNA may therefore alter miRNA binding to these regions by either creating a new site or destroying an existing one. miRSNPs may explain the modulation of expression levels in association with increased/decreased susceptibility to common diseases as well as in chemoresistance and the consequent inter-individual variability in drug response. In the present study, we investigated whether miRSNPs in *SLC* transporter genes may modulate CRC susceptibility and patient's survival. Using an *in silico* approach for functional predictions, we analyzed 26 miRSNPs in 9 *SLC* genes in a cohort of 1368 CRC cases and 698 controls from the Czech Republic. After correcting for multiple tests, we found several miRSNPs significantly associated with patient's survival. SNPs in *SLCO3A1*, *SLC22A2* and *SLC22A3* genes were defined as prognostic factors in the classification and regression tree analysis. In contrast, we did not observe any significant association between miRSNPs and CRC risk. To the best of our knowledge, this is the first study investigating miRSNPs potentially affecting miRNA binding to *SLC* transporter genes and their impact on CRC susceptibility or patient's prognosis.

Received: August 12, 2020; **Revised:** December 1, 2020; **Accepted:** December 10, 2020

© The Author(s) 2020. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Abbreviations

5-FU	5-fluorouracil
ABC	ATP-binding cassette
BMI	body mass index
CART	classification and regression tree
CRC	colorectal cancer
miRNAs	microRNAs
OS	overall survival
RFS	recurrence-free survival
SLC	solute carrier
SNPs	single-nucleotide polymorphisms
TNM	tumor-node metastasis

Introduction

Colorectal cancer (CRC) ranks among the leading causes of cancer-related deaths worldwide (1–3). Survival rates of CRC patients largely depend on the disease stage at diagnosis: the 5-year survival for stage I is around 90%, whereas it drops dramatically to 7–14% for stage IV patients (4). 5-Fluorouracil (5-FU) is the most frequently used chemotherapeutic agent for treating patients with CRCs. However, the patients' benefit from 5-FU-based therapy is frequently compromised by the development of chemoresistance (5).

One of the most common mechanisms underlying drug resistance is associated with the low accumulation of the cytotoxic drugs in tumor cells caused by an increased efflux, a decreased influx, or both (6–8). The drug efflux is mediated mainly by transmembrane ATP-binding cassette (ABC) transporters [reviewed in (8)]. Another large family of proteins, the solute carrier (SLC) transporters, transfers into cells a wide range of substrates, including amino acids, lipids, inorganic ions, peptides, saccharides, metal ions, proteins and xenobiotics, including drugs [reviewed in (9–12)]. For example, nucleoside transporters, such as SLC28 and SLC29 subfamilies, accomplish the influx of gemcitabine and 5-FU (13,14). Moreover, several members of the SLC22A subfamily (SLC22A1/OCT1, SLC22A2/OCT2 and SLC22A3/OCT3) have been implicated in the transport of low-molecular weight drugs such as platinum compounds, although the available *in vivo* data on their role in platinum pharmacokinetics are rather controversial (15). Furthermore, human SLC22A7/OAT2 was found to have a very high affinity for 5-FU and is responsible for its hepatic uptake (16).

In CRC patients, inter-individual differences in the treatment response may be due to the unique genetic and epigenetic make-up of each individual as well as tumor and host characteristics. Interestingly, drug absorption, distribution, metabolism, excretion, and efficacy are also heavily influenced by epigenetic factors such as DNA/histone methylation and non-coding RNAs (especially microRNAs (miRNAs)) (17). Aberrant miRNA expression and/or function are frequently observed in many malignancies, including CRC (18,19). In addition, genetic variations in the 3′UTRs of target genes may affect miRNA binding, ultimately imposing additional variability into the differential mRNA and protein expressions (20). The increasing need for newer diagnostic strategies has led to the recognition of miRNAs as potential cancer biomarkers of new generation.

Our group already reported the association of single-nucleotide polymorphisms (SNPs) in miRNA target regions of several important genes with CRC risk and/or clinical outcome (21–24). In the present study, we hypothesized that SNPs in miRNA target regions (miRSNPs) of *SLC* transporter genes might modulate the efficiency of translation of corresponding proteins, thus affecting individual's susceptibility to CRC or patient's clinical outcome. We investigated 26 miRSNPs in 11 genes in a population of 1368 CRC patients and 698 controls from the Czech Rep.

Material and methods

Study population and data collection

Blood samples were collected from 1987 patients of Caucasian origins with histologically confirmed CRC recruited in several oncological departments in the Czech Republic between September 2003 and December 2016. The control group, comprising 1014 samples, was collected at the same time frame of cases recruitment. Details of CRC cases and controls have been reported previously (24–29). Study participants provided information on their lifestyle, body mass index (BMI) and family/personal history of cancer, using a structured questionnaire to determine basic demographic characteristics and potential risk factors for CRC.

All participants signed a written consent to participate in the study and approved the use of their biological samples for genetic analyses according to the Helsinki declaration. The design of the study was approved by the Ethics Committee of the Institute of Experimental Medicine, Prague, Czech Republic.

Follow-up of the patients

All CRC cases were monitored with a regular follow-up until December 31, 2016. For all subjects, clinical data at the time of diagnosis, including the location of the tumor, the International Union Against Cancer tumor-node-metastasis (TNM) stage system, grade and adjuvant or first line 5-fluorouracil (5-FU)-based chemotherapy treatment were assembled, along with information about distant metastasis, relapse and date of death.

One thousand nine hundred and eighty-seven CRC cases were considered in the follow-up. Four hundred and twenty-six CRC cases received 5-FU-based chemotherapy as adjuvant postoperative therapy. Metastatic CRC patients were administrated with palliative FOLFOX4 regimen. Seven hundred and fifty-seven subjects did not receive any adjuvant chemotherapy. In this study, the outcome variables measured were 5-FU-based chemotherapy (yes/no), overall survival (OS; described as the time from diagnosis until death or censorship), and recurrence-free survival (RFS; described as the time of surgery/end of chemotherapy until the date of relapse, death or censorship).

Selection of candidate genes and SNPs in miRNA target sites

A list of SLC transporters was retrieved from the Guide to Pharmacology (https://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=863). For each gene, miRSNPs within target binding sites for miRNA were identified by using freely available software MicroSNiPer (http://vm24141.virt.gwdg.de/services/microsniper/ (30)), miRSNP (http://bioinfo.bjmu.edu.cn/mirsnp/search/ (31)) and PolymiRTS (http://compbio.uthsc.edu/miRSNP/ (32)). The detected SNPs were further tested for minor allele frequency >5% in the Caucasian populations in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP) to reach an appropriate representation of all genotypes in our set of cases and controls. The information was primarily derived from EUR population; whenever this was not possible, other Caucasian reference populations were considered (i.e. 1000 genomes: phase 1, CEU population). MiRSNPs with the required minor allele frequency were further tested for the possibility to be in linkage disequilibrium using HaploView (v. 4.2.; https://www.broad.mit.edu/mpg/haploview), and Haploreg (v. 4.1; http://archive.broadinstitute.org/mammals/haploreg/haploreg.php).

DNA isolation and SNP genotyping

Genomic DNA was isolated from peripheral blood lymphocytes using phenol–chloroform procedures as described in (33). The selected 26 *SLC*

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

miRSNPs were genotyped using the KASP™ genotyping assay, a competitive allele-specific PCR SNP genotyping system (LGC Genomics, Hoddesdon, Herts, UK). For quality control purposes, an equal number of DNA samples from cases and controls were randomly placed on plates for simultaneous analysis. The results were regularly confirmed by random re-genotyping of more than 5% of the samples for each polymorphism and concordant results were yielded. Non-template controls (NTCs) were included in each plate. PCR reactions were carried out using 5 ng of purified DNA per reaction. The genotypes with unclear results were excluded from the study. The genotype correlation between the duplicate samples was >97%. Genotype call rate ranged between 94% and 99%.

Bioinformatics analysis

Assessment of Gibbs binding energy

For the selected SNPs, the algorithm RNAcofold (http://rna.tbi.univie.ac.at/cgi-bin/RNAcofold.cgi) was run to assess the Gibbs binding free energy (ΔG, expressed in KJ/mol), for both the common and the variant alleles. RNAcofold computes the hybridization energy and base-pairing pattern of two RNA sequences (20,21). The difference of the free energies between the two alleles was computed as "variation of ΔG" (i.e. |ΔΔG|). Since the neighboring sequence of each SNP can be a target for different miRNAs, we calculated the sum of the absolute values of ΔΔGs for each SNP (i.e. |ΔΔG|tot = Σ |ΔΔG|) (27). The |ΔΔG|tot should be considered as a sort of "disturbance index" predicting the likelihood for a given SNP to affect the function of the 3′UTR and it allows a ranking of SNPs for their relevance, as illustrated in previous studies (21,29).

The polymorphic and miRNA sequences used for the ΔΔG calculation were obtained from the MicroSniper tool (30). MicroSniper provides also an easy way to select miRNA target sites according to the binding to different alleles and a straightforward possibility to test binding free energy. For the miRNAs identified as binding to both common and variant alleles, ΔΔG calculation was performed. In detail, the sequences containing the original and variant allele around 50 nt long and miRNA sequences~22 nt long were obtained from MicroSniper and analyzed by RNAcofold tool.

Identification of candidate miRSNPs through eQTL analysis

The association between the miSNPs analyzed in the study and gene expression levels was obtained from the Genotype-Tissue Expression project (https://gtexportal.org/home/, version V6p) and SCAN database (http://www.scandb.org/newinterface/about.html). The Genotype-Tissue Expression project allows viewing and downloading computed eQTL results and aims to characterize variations in gene expression levels across individuals and diverse tissues of the human body, many of which are not easily accessible (34). The SCAN database provides the number of differentially transcribed genes for each SNP in lymphoblastoid cell lines from individuals of Caucasian origin.

Identification of miRNAs as potential pharmacogenomic biomarkers for anticancer drugs

To identify potential pharmacogenomic biomarkers characterized by miRNAs expression and discover the underlying mechanisms of anticancer drug responses mediated by miRNAs, the Small Molecule-miRNA Network-Based Inference (http://lmmd.ecust.edu.cn/database/smir-nbi/ (35–37)) model was used. The Small Molecule-miRNA Network-Based Inference model was built based on a heterogeneous network connecting drugs, miRNAs and genes (34) and predicts interactions between small molecules and miRNAs.

Statistical analysis

Pearson's chi-square test (1 degree of freedom), with a type I alfa error of 0.05, was used to verify whether the genotypes were in Hardy–Weinberg equilibrium in the control population. The association between the overall genotypes of all SNPs and the risk of CRC was estimated using a multivariate logistic regression. The covariates analyzed for inclusion in the multivariate model were: sex, age, smoking status (non-smokers versus smokers), BMI and familial history of CRC. The association between SNPs and CRC risk was calculated by estimation of the odd ratios, their 95% confidence intervals (CIs) and *P*-values ($P \leq 0.05$), adjusted for age and sex. For all SNPs with significant *P*-value *per genotype, the best model (recessive, dominant and* codominant) was calculated.

The association between patient survival and SNPs was assessed by the Log-rank test and the survival curves for OS and RFS were derived by the Kaplan–Meier method. The relative risk of death and recurrence was estimated by hazard ratio (HR) using Cox regression (R version 2.14-2, Survival package).

To assess the prognostic utility, the interactive effects of genotypes and clinicopathological parameters in association with 5-year OS and RFS were explained by classification and regression tree (CART) analysis (38). Twenty-six genotypes, together with 8 personal and clinicopathological parameters (age, sex, smoking status, positive family history of CRC, diagnosis, TNM stage, grade and therapy) were included in the multivariate model. Only patients with complete data for all parameters mentioned above were included in the analysis (1098 CRC patients). For both, OS and RFS, TNM stages as the initial split-up factor due to the variable with the optimal first split was selected. The CART analysis is an algorithm for predictive modeling machine learning. This analysis uses the decision tree (as a predictive model) to move from an observation about an item (represented in branches) to conclusions about the target value of the item (shown on the leaves). For survival analysis, it means that survival is better or worse in parallel with an accumulation of covariates and the presence of specific alleles in miRSNP.

All the analyses were adjusted for multiple testing corrections using the Benjamini–Hochberg False Discovery Rate measure.

Results

Selection of candidate genes and SNPs in miRNA-binding target sites

The SLC superfamily contains more than 400 transport proteins that mediate the influx and efflux of substances such as ions, nucleotides and sugars across biological membranes. Over 80 SLC transporters have been linked to human diseases. As the aim of the study was to analyze the outcomes for CRC, only those SLC transporters expressed in colon and liver (as a predominant CRC metastatic site) were selected into this study using the www.genecards.org data. Out of the 20 genes belonging to human SLC transporter family expressed in colon tissue, 11 had in total 52 miRSNPs predicted to bind miRNAs in their 3′UTRs. These miRSNPs were filtered for a minor allele frequency >5% in the European CEU population. At the resulting SNPs, an additional filter based on linkage disequilibrium analysis was applied using an r^2 threshold of 0.80. After these selection processes, 28 miRSNPs remained (rs1051298 in *SLC19A1*, rs694812, rs2450975, rs3127592 in *SLC22A2*, rs2076828, rs3088442 in *SLC22A3*, rs8025045 in *SLC28A1*, rs11140489 in *SLC28A3*, rs780680, rs1084004, rs2487068 in *SLC29A3*, rs10513202, rs10759637, rs10981707 in *SLC31A1*, rs739439, rs1128162 in *SLC46A1*, rs10735 in *SLC47A1*, rs875234, rs10841781, rs11045916, rs11045919 in *SLCO1A2*, rs8174, rs1053909, rs1060205, rs1060206, rs2108601, rs2270061, rs11074043 in *SLCO3A1*) and were analyzed in the present study (Supplementary Table 1, available at *Carcinogenesis* Online). Of these, two assays (rs10981707 in *SLC31A1*, and rs1060206 in *SLCO3A1*) did not pass quality control (non-specific amplification) and were excluded from further analyses. Finally, 26 SNPs were successfully genotyped.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

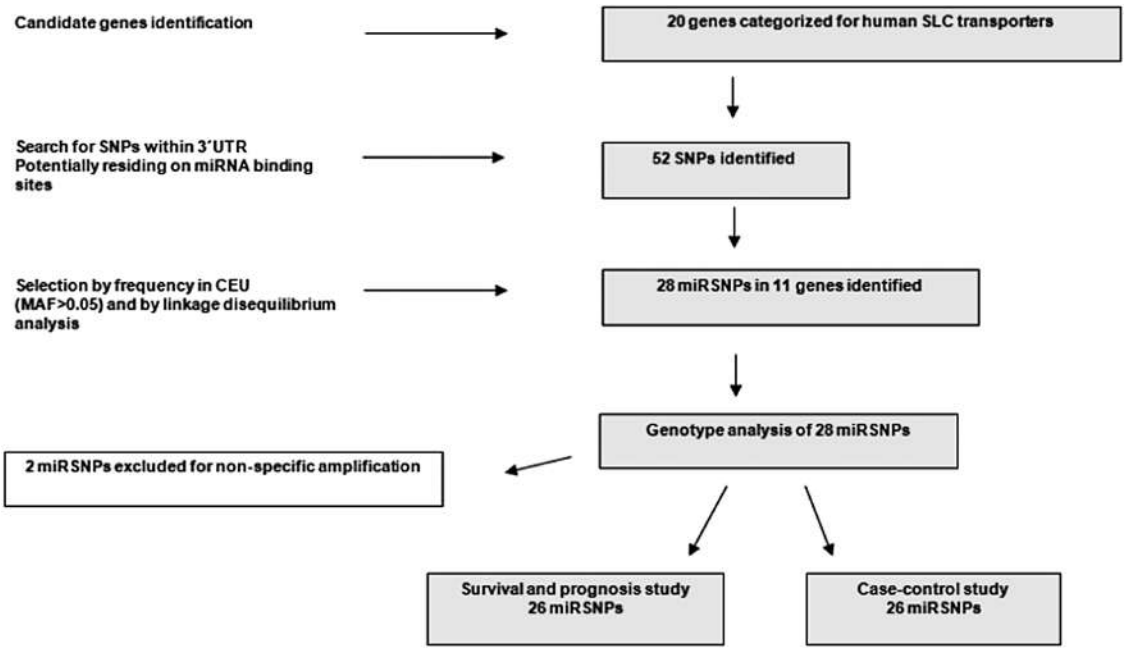

Figure 1. Workflow strategy for selection and analyses of miRSNPs in *SLC* transporter genes.

The workflow strategy used in the study for the selection and the analysis of miRSNPs is depicted in Figure 1.

Case–control study

Initially, the population in this study comprised 1987 sporadic CRC patients and 1014 controls. Due to significant differences in age and sex distribution between these two groups, an equivalence test to approximate the difference in age and sex between case and control groups was performed with a tolerance limit of ±5 years and ±10% representation of sex (Welch–Satterthwaite equation $P = 0.04$). After this test, the study included 1368 CRC (573 patients with a tumor in the colon, 279 in the sigmoideum and 516 in the rectum) and 698 controls. The characteristics of the study participants are given in Table 1.

The results of the associations between the investigated miRSNPs and CRC susceptibility are reported in Supplementary Table 2, available at *Carcinogenesis* Online. The distribution of genotypes within the studied SNPs in controls was in agreement with Hardy–Weinberg equilibrium.

We did not find any association between the investigated miRSNPs and CRC risk with the exception of the GA genotype in *SLCO3A1* rs1053909 that resulted associated with rectal cancer risk, but only when not-considering the correction for multiple testing.

Survival analysis

The mean OS and RFS for the studied population were 70.8 and 52.1 months, respectively. Several parameters, such as gender, age, smoking, alcohol intake, T, N, M status, CRC stage and chemotherapy treatment, were associated with OS and RFS in the univariate assessment of covariates (Table 2). Advanced age and stage, male gender and no alcohol consumption were significantly associated with shorter OS. Likewise, male gender and higher TNM status were associated with elevated risk of relapse or metastasis.

The strongest association with patient's survival was observed for *SLCO1A2* rs10841781 (Supplementary Table 3, available at *Carcinogenesis* Online). CRC patients carrying the heterozygous GA genotype displayed a shorter OS (HR 1.32; 95% CI 1.09–1.59; $P = 0.005$). The observed association became stronger after stratification for tumor site, particularly for patients with colon cancer (OS: HR 1.46; 95% CI 1.16–1.86; $P = 0.002$ and RFS: HR 1.37; 95% CI 1.11–1.70; $P = 0.004$, respectively).The reported associations remained significant also after adjusting for multiple testing corrections.

Survival and therapy

To examine the association of miRSNPs with the therapy outcome, we further stratified patients according to the treatment received into three distinct groups: (i) patients that did not underwent any treatment, (ii) patients receiving solely 5-FU regimen and (iii) patients receiving 5-FU in combination with oxaliplatin (Table 3). The univariate model for survival and therapy after false discovery rate correction showed several miRSNPs significantly associated with OS or RFS (detailed description in Table 3).

The second group of patients received a 5-FU-based adjuvant regimen as a postoperative therapy (based either on a Mayo, a simplified DeGramont, or a Xeloda regimen). The third group of subjects received adjuvant 5-FU treatment combined with oxaliplatin (based either on a FOLFOX or a XELOX regimen). Patients with palliative chemotherapy were stratified in CART analysis as TNM split (described in a separate section).

Group 1: ***CRC patients untreated***

Patients with colon cancer carrying homozygous AA genotype for *SLC22A2* rs2450975 displayed worse survival (OS: HR: 4.46; 95% CI 2.14–9.29; $P = 0.0001$; RFS: HR: 3.42; 95% CI 1.66–7.07; $P = 0.0009$, respectively) compared with carriers of the other genotypes. Colon cancer patients carrying heterozygous GA genotype for *SLC22A2* rs694812 also displayed higher recurrence risk (RFS: HR: 1.7; 95% CI 1.18–2.46; $P = 0.005$). Similar results were obtained for both miRSNPs in the dominant model.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 1. Characteristics of the study population

Variables		Controls (%)	Cases (%)	Odd ratios	95% CI	*p*-value
Total		698 (33.8)	1368 (66.2)			
Age (years)	<60	389 (18.8)	440 (21.3)	Ref.		
	60–65	93 (4.5)	255 (12.3)	2.42	1.84–3.19	**<0.0001**
	65–70	63 (3.1)	253 (12.3)	3.55	2.61–4.83	**<0.0001**
	>70	153 (21.9)	420 (20.3)	2.43	1.93–3.01	**<0.0001**
Sex	Females	301 (14.6)	510 (24.7)	Ref.		
	Males	397 (19.2)	858 (41.5)	1.28	1.06–1.54	**0.011**
BMI[a]	18.5–25	207(29.7)	344 (25.1)	Ref.		
	<18.5	7 (0.4)	9 (0.5)	1.38	0.48–3.97	0.55
	>30	136 (7.9)	265 (15.3)	1.12	0.88–1.42	0.35
Smoking status[a,b]	No	377 (19.1)	733 (37.1)	Ref.		
	Yes	263 (13.3)	603 (30.5)	1.24	1.02–1.51	**0.04**
Family history of cancer[a]	No	285 (15.1)	693 (36.6)	Ref.		
	Yes	341 (18.0)	573 (30.3)	0.75	0.61–0.91	**0.004**
Family history of CRC[a]	No	512 (27.7)	1055 (57.1)	Ref.		
	Yes	86 (4.7)	194 (10.5)	1.21	0.91–1.6	0.20
Alcohol intake[a]	No	245 (17.8)	286 (20.8)	Ref.		
	Yes	386 (28.1)	456 (33.2)	1.01	0.8–1.28	0.95

Significant results in bold.
[a]Numbers may not add up to 100% of available subjects because of missing data.
[b]Ex-smokers included in smokers.

Group 2: ***CRC patients receiving 5-FU regimen***
CRC carriers of the GC genotype of *SLC22A3* rs2076828 displayed longer RFS in codominant and dominant models (codominant model RFS: HR: 0.58; 95% CI 0.41–0.83; *P* = 0.003; dominant model RFS: HR: 0.58; 95% CI 0.42–0.81; *P* = 0.001). In CRC patients, carriers of at least one A allele for *SLC22A3* rs3088442 had higher risk of relapse or metastasis (dominant model RFS: HR: 1.52; 95% CI 1.10–2.10; *P* = 0.01). After stratification, the same effect was observed in rectal cancer patients (RFS: HR: 2.84; 95% CI 1.41–5.71; *P* = 0.003). The same trend was observed in the recessive model. Rectal cancer patients carrying homozygous genotype AA for *SLCO3A1* rs1053909 displayed worse OS and shorter RFS (recessive model OS: HR: 3.09, 95% CI 1.36–7.03; *P* = 0.007; RFS: HR: 2.38; 95% CI 1.15–4.91; *P* = 0.02).

Group 3: ***CRC patients receiving 5-FU regimen in combination with oxaliplatin***
All CRC patients carrying homozygous AA genotype for *SLCO3A1* rs8174 showed shorter RFS (recessive model RFS: HR: 1.68; 95% CI 1.13–2.50; *P* = 0.01).

CART survival analysis

To assess the prognostic utility of the studied miRSNPs, the interactive effects of genotypes and clinicopathological parameters in association with 5-year OS and RFS were analyzed using CART analysis. Only patients with complete data for all parameters described in the materials and methods (i.e. 1095 CRC cases) were included in the analysis.

Overall survival

In the 5-year OS analysis, four terminal nodes were created, and the structure of the tree included 14 following determining variables: *TNM, age, gender, diagnosis, grade*, *chemotherapy* and 7 *SNPs* (rs1053909, rs3088442, rs8025045, rs10735, rs739439, rs1051298, rs2108601 and rs3127592). The structure of the tree and corresponding survival curves from terminal nodes are presented in Figure 2A–C.**Figure 2.** Continued.

Among stage I CRC patients, the subsequent split showed interactions between *SLCO3A1* rs1053909, gender and age. Males older than 70 years of age were associated with worse prognosis (female 83.3% versus male 53.1%). However, carriers of GG+GA genotypes of rs1053909 and of <70 years old age displaying a better prognosis (GG+GA 95.0% versus AA 71.4%). In stage II, carriers of GA or GG genotypes of *SLC22A3*rs3088442 were associated with a better prognosis (GA+GG78.0% vs. AA 58.4%). In addition, AA genotype of the same miRSNP further interacted with GG genotype of *SLC28A1* rs8025045 and tumor localization (colon + sigmoideum 78.2% versus rectum 46.4%). Within stage III CRC patients, TC+CC genotypes of *SLC47A1* rs10735 were associated with better prognosis (TC+CC 72.1% versus TT 49.1%). However, the OS level within TT carriers rapidly decreased when being in combination with >75 age and AA+AG genotype in *SLC46A1*rs739439 (GG34.8% versus AA+AG0%). Nevertheless, patients younger than 75 years of age further interacted with lower grade (0 + 1+2), CC or TC genotypes in *SLC19A1* rs1051298, and GG genotype in *SLC28A1* rs8025045. The interaction was concluded by CC genotype of *SLCO3A1* rs2108601 (CC 97.1% versus CA+AA 63.8%). In stage IV, chemotherapy was the first split when patients with 5-FU+oxapliplatin-based therapy had a better prognosis when compared with those with no treatment (5-FU+oxapliplatin 39.7% versus no treatment 12.1%). Patients with no treatment had worse prognosis when associated with TC+CC genotype in *SLC22A2* rs3127592 (TT 15.1% versus TC+CC 0%). Among patients treated with 5-FU+oxaliplatin-based therapy, the presence of GG or GA genotypes of *SLC22A3* rs3088442 (AA 77.9% versus GA+GG 34.4%) was indicative of a worse prognosis.

Recurrence-free survival

Within 5-year RFS, the final tree structure contained six terminal nodes including 10 variables: TNM, diagnosis, gender and SNPs (rs8174, rs2270061, rs10841781, rs1051298, rs3088442, rs1084004 and rs10735). The structure of the tree and corresponding survival curves from terminal nodes are presented in Figure 3A and B.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 2. Clinical and anamnestic characteristics significantly affecting OS and RFS of CRC patients with complete follow-up (Cox regression)

			OS		RFS	
		N[b]	HR (95%CI)	P-value	HR (95%CI)	P-value
Sex	Female	556	Ref.		Ref.	
	Male	917	1.45 (1.2–1.8)	**0.0002**	1.28 (1.08–1.51)	**0.004**
Age (years)	≤60	440	Ref.		Ref.	
	60–65	255	1.32 (1.11–1.58)	**0.0022**	0.96 (0.81–1.14)	0.66
	65–70	253	1.42 (1.19–1.71)	**0.0001**	1.09 (0.93–1.27)	0.31
	>70	525	1.46 (1.17–1.8)	**0.0007**	1.17 (1–1.38)	**0.05**
Smoking status[a]	No	1215	Ref.		Ref.	
	Yes	222	1.01 (0.99–1.0)	0.2673	1.02 (1–1.04)	**0.037**
Alcohol consumption	No	306	Ref.		Ref.	
	Yes	475	0.75 (0.6–0.94)	**0.013**	0.87 (0.71–1.08)	0.20
pT	Tis + 0 + 1	59	Ref.		Ref.	
	2 + 3 + 4	1389	2.59 (1.38–4.85)	**0.003**	2.55 (1.47–4.42)	**0.0009**
pN	0	714	Ref		Ref	
	1 + 2 + 3	684	2.41 (1.99–2.92)	**< 0.001**	2.13 (1.81–2.52)	**< 0.0001**
pM	0	1136	Ref		Ref	
	1 + 2	273	3.17 (2.63–3.83)	**< 0.0001**	2.61 (2.2–3.1)	**< 0.0001**
Stage	0 + 1	194	Ref		Ref	
	2 + 3 + 4	1198	1.53 (1.13–2.06)	**0.006**	1.24 (0.97–1.59)	**0.093**
5FU-based chemotherapy	Yes	757	Ref		Ref	
	No	712	0.99 (0.83–1.18)	0.89	1.25 (1.06–1.46)	**0.006**

Bold value indicates significant results.

[a]Ex-smokers included in non-smokers.

[b]Numbers may not add up to 100% of available subjects because of missing information.

In stage I, CRC patients with GG genotype for *SLCO3A1* rs8174 displayed a better prognosis than patients carrying the other genotypes of the same miRSNP (GG 82.6% versus GA+AA 67.5%). However, GA or AA genotype for rs8174 in combination with TT genotype of *SLCO3A1* rs2270061, AA genotype of *SLCO1A2* rs10841781 and CC+TC genotypes of *SLC19A1* rs1051298 were further associated with better prognosis (CC+TC 100% versus TT 51.9%). Among individuals at stage II, those with rectal cancer and simultaneously carrying the AA genotype of *SLC22A3* rs3088442 were associated with worse prognosis (colon+sigmoideum 55.3% versus rectum 20.8%). GG or GA genotypes of the same miRSNP further in interaction with GA or AA genotypes in *SLCO3A1* rs8174, and female gender and displayed a better prognosis (female 85.2% versus male 59.5%). In stage III, carriers of TC+TT genotypes of *SLC29A3* rs1084004 showed better prognosis than CC carriers (TC+TT 60.5% versus CC 36.5%). On the other hand, carriers of the CC genotype in the same miRSNP showed worse prognosis in combination with TT genotype of *SLC47A1* rs10735 and AA genotype of *SLCO3A1* rs8174 (GG+GA 31.8% versus AA 12.1%).

Assessment of Gibbs binding energy

By calculating the Gibbs-free energy variation (ΔG), we could predict the impact of the allele change in affecting or impairing the binding with predicted miRNAs. There were several miRNAs that could bind at the same position in the presence of both alleles and for them we calculated ΔΔG (as reported in Table 4 and Supplementary Table 4, available at *Carcinogenesis* Online). The highest predicted difference of binding energy according to the allele was noticed for rs11045919 and miR-4487 (ΔΔG –4.4 kcal/mol) and rs2270061 and miR-654-5p (ΔΔG +5.8 kcal/mol). For rs11045919, the ΔG of binding to the C allele is less negative (=less favorable) than those with the more common A allele. This suggests a less efficient binding of miR-4487 in the presence of the C allele on the 3′UTR of *SLCO1A2* mRNA and implicates a potentially decreased post-transcriptional repression of *SLCO1A2* by this miRNA. Concerning the rs2270061, calculating the ΔΔG for the different alleles, the miRNA–mRNA binding is stronger with the presence of the variant T allele. This results in a stronger negative regulation of the target gene expression.

Identification of candidate miRSNPs through eQTL analysis

According to Genotype-Tissue Expression data, no significant eQTLs were found for the majority of miRSNP in any tissues (Supplementary Table 1, available at *Carcinogenesis* Online). In the SCAN database, we observed that only *SLC28A3* rs11140489 and *SLC46A1* rs739439 were associated with differentially expressed genes (26 and 15, respectively).

Identification of miRNAs as potential pharmacogenomic biomarkers for anticancer drugs

To identify those miRNAs that could work as potential pharmacogenomic biomarkers for anticancer drugs, we have performed additional analysis with the implementation of Small Molecule-miRNA Network-Based Inference model which predicts interactions between small molecules and miRNAs (Supplementary Table 5, available at *Carcinogenesis* Online). According to this *in silico* analysis, several miRNAs resulted regulated by many drugs, such as 5-FU, cisplatin and oxaliplatin. The most interesting results were observed for *SLC22A3* rs3088442. In presence of the common G allele at rs3088442, miR-147b is binding in the position of this miRSNP, while in presence of the variant A allele, another miRNA from the same family, miR-147a, is predicted to bind. miR-147a is predicted to be upregulated by 5-FU administration. This *in silico* outcome is in agreement with our results where CRC carriers of at least one A allele for *SLC22A3* rs3088442 had higher risk of relapse or metastasis after 5-FU-based therapy.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 3. miRSNPs associated with OS and RFS of patients stratified for tumor sites and therapy (Cox regression for adjusted estimates)

		All cancer patients				Colonsigmoideum cancer patients				Rectal cancer patients			
Gene SNP	Genotype	N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value
Overall survival													
Patients receiving no treatment													
SLC22A2 Rs694812	A:A	644	195	Ref.		378	114	Ref.		266	81	Ref.	
	G:A	105	40	1.20 (0.85–1.68)	0.31	60	30	1.53 (1.02–2.29)	**0.04**	45	10	0.73 (0.38–1.41)	0.35
	G:G	4	1	0.63 (0.09–4.50)	0.65	1	1	1.76 (0.25–12.6)	0.57	3	0	NA	0.98
	G:A+G:G	109	41	1.17 (0.83–1.63)	0.37	61	31	1.54 (1.03–2.29)	**0.03**	48	10	0.67 (0.34–1.29)	0.23
	A:A+G:A	749	235	Ref.		438	144	Ref.		311	91	Ref.	
	G:G	4	1	0.59 (0.08–4.23)	0.60	1	1	1.63 (0.23–11.7)	0.63	3	0	NA	0.98
SLC22A2 Rs2450975	C:C	517	155	Ref.		301	88	Ref.		216	67	Ref.	
	C:A	207	70	1.12 (0.84–1.48)	0.44	121	49	1.29 (0.91–1.84)	0.15	86	21	0.88 (0.54–1.45)	0.63
	A:A	22	11	1.83 (0.99–3.37)	**0.05**	12	8	4.46 (2.14–9.29)	**0.0001**	10	3	0.60 (0.19–1.92)	0.39
	C:A+A:A	229	81	1.18 (0.90–1.55)	0.22	133	57	1.44 (1.03–2.01)	**0.03**	96	24	0.83 (0.52–1.33)	0.45
	C:C+C:A	724	225	Ref.		422	137	Ref.		302	88	Ref.	
	A:A	22	11	1.77 (0.97–3.25)	0.06	12	8	4.17 (2.03–8.56)	**0.0001**	10	3	0.60 (0.19–1.92)	0.39
SLC29A3 Rs780680	G:G	597	178	Ref.		342	106	Ref.		255	72	Ref.	
	T:G	141	55	1.42 (1.05–1.92)	**0.02**	88	37	1.51 (1.04–2.19)	**0.03**	53	18	1.23 (0.73–2.06)	0.44
	T:T	9	3	1.01 (0.32–3.16)	0.99	6	2	0.81 (0.20–3.27)	0.76	3	1	1.93 (0.26–14.1)	0.52
	T:G+T:T	150	58	1.39 (1.03–1.87)	**0.03**	94	39	1.45 (1.00–2.09)	**0.05**	56	19	1.25 (0.75–2.08)	0.39
	G:G+T:G	738	233	Ref.		430	143	Ref.		308	90	Ref.	
	T:T	9	3	0.94 (0.3–2.94)	0.92	6	2	0.75 (0.19–3.02)	0.68	3	1	1.80 (0.25–13.01)	0.56
Patients receiving a 5-FU based therapy													
SLC22A3 rs2076828	C:C	129	47	Ref.		89	31	Ref.		40	16	Ref.	
	G:C	204	49	0.64 (0.43–0.95)	**0.03**	132	29	0.63 (0.38–1.04)	0.07	72	20	0.65 (0.34–1.27)	0.21
	G:G	84	23	0.69 (0.42–1.14)	0.15	54	16	0.89 (0.48–1.62)	0.69	30	7	0.42 (0.17–1.03)	0.06
	G:C+G:G	288	72	0.65 (0.45–0.94)	**0.02**	186	45	0.70 (0.44–1.10)	0.12	102	27	0.57 (0.31–1.06)	0.08
	C:C+G:C	333	96	Ref.		221	60	Ref.		112	36	Ref.	
	G:G	84	23	0.89 (0.57–1.41)	0.63	54	16	1.14 (0.66–1.99)	0.63	30	7	0.55 (0.25–1.24)	0.15
SLC22A3 Rs3088442	G:G	181	48	Ref.		117	30	Ref.		64	18	Ref.	
	G:A	173	46	1.33 (0.88–1.99)	0.17	117	31	1.33 (0.80–2.20)	0.27	56	15	1.32 (0.66–2.63)	0.43
	A:A	67	27	1.36 (0.85–2.18)	0.20	43	17	1.17 (0.65–2.13)	0.6	24	10	2.06 (0.95–4.47)	0.07
	G:A+A:A	240	73	1.35 (0.94–1.95)	0.10	160	48	1.30 (0.83–2.06)	0.25	80	25	1.48 (0.81–2.73)	0.20
	G:G+G:A	354	94	Ref.		234	61	Ref.		120	33	Ref.	
	A:A	67	27	1.20 (0.78–1.85)	0.40	43	17	1.03 (0.6–1.76)	0.92	24	10	1.79 (0.88–3.65)	0.11
SLC3A1 Rs1053909	G:G	188	59	Ref.		116	36	Ref.		72	23	Ref.	
	G:A	188	46	0.72 (0.49–1.06)	0.09	130	33	0.72 (0.45–1.16)	0.18	58	13	0.71 (0.36–1.4)	0.32
	A:A	46	15	0.99 (0.56–1.74)	0.96	32	8	0.65 (0.30–1.41)	0.28	14	7	2.35 (1.00–5.53)	**0.05**
	G:A+A:A	234	61	0.77 (0.54–1.10)	0.15	162	41	0.71 (0.45–1.11)	0.13	72	20	0.95 (0.52–1.72)	0.86
	G:G+G:A	376	105	Ref.		246	69	Ref.		130	36	Ref.	
	A:A	46	15	1.14 (0.66–1.95)	0.64	32	8	0.75 (0.36–1.57)	0.45	14	7	3.09 (1.36–7.03)	**0.007**
Patients receiving a 5-FU + oxaliplatin- based therapy													
SLC19A1 rs1051298	C:C	82	35	Ref.		57	23	Ref.		25	12	Ref.	
	T:C	140	71	1.35 (0.9–2.03)	0.15	105	57	1.57 (0.96–2.56)	0.07	35	14	1.01 (0.46–2.18)	0.99
	T:T	51	24	1.46 (0.86–2.48)	0.16	32	15	1.40 (0.72–2.74)	0.32	19	9	1.75 (0.73–4.21)	0.21
	T:C+T:T	191	95	1.37 (0.93–2.03)	0.11	137	72	1.54 (0.95–2.48)	0.08	54	23	1.21 (0.60–2.43)	0.60
	C:C+T:C	222	106	Ref.		162	80	Ref.		60	26	Ref.	

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 3. Continued

Gene SNP	Genotype	All cancer patients				Colonsigmoideum cancer patients				Rectal cancer patients			
		N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value
	T:T	51	24	1.22 (0.78–1.90)	0.39	32	15	1.04 (0.6–1.82)	0.88	19	9	1.77 (0.82–3.81)	0.15
SLC28A3 Rs11140489	T:T	199	97	Ref.		142	71	Ref.		57	26	Ref.	
	T:A	71	31	0.48 (0.15–1.60)	0.23	51	23	0.54 (0.07–4.06)	0.55	20	8	0.39 (0.08–1.97)	0.26
	A:A	4	3	0.44 (0.14–1.39)	0.16	1	1	0.37 (0.05–2.72)	0.33	3	2	0.53 (0.12–2.28)	0.39
	T:A+T:T	270	128	0.45 (0.14–1.42)	0.17	193	94	0.41 (0.06–2.98)	0.38	77	34	0.48 (0.12–2.04)	0.32
	A:A+T:A	75	34	Ref.		52	24	Ref.		23	10	Ref.	
	A:A	4	3	1.01 (0.68–1.50)	0.96	1	1	1.01 (0.64–1.61)	0.95	3	2	1.07 (0.51–2.23)	0.86
SLCO3A1 Rs8174	G:G	110	52	Ref.		77	36	Ref.		33	16	Ref.	
	G:A	123	56	0.81 (0.56–1.19)	0.28	85	40	0.80 (0.51–1.27)	0.35	38	16	0.81 (0.41–1.62)	0.55
	A:A	39	22	1.18 (0.72–1.95)	0.51	29	17	1.18 (0.66–2.12)	0.57	10	5	1.29 (0.47–3.56)	0.62
	G:A+A:A	162	78	0.90 (0.63–1.28)	0.55	114	57	0.89 (0.59–1.36)	0.61	48	21	0.88 (0.46–1.7)	0.71
	G:G+G:A	233	108	Ref.		162	76	Ref.		71	32	Ref.	
	A:A	39	22	1.29 (0.81–2.04)	0.28	29	17	1.23 (0.72–2.08)	0.45	10	5	1.45 (0.56–3.72)	0.45
SLC3A1 Rs2270061	A:A	29	18	Ref.		20	11	Ref.		36	17	Ref.	
	T:A	118	45	0.54 (0.31–0.93)	**0.03**	85	34	0.68 (0.34–1.34)	0.26	33	11	0.39 (0.15–1.02)	**0.05**
	T:T	119	60	0.65 (0.38–1.10)	0.11	83	43	0.77 (0.39–1.49)	0.43	9	7	0.45 (0.18–1.10)	0.08
	T:A+T:T	237	105	0.59 (0.36–0.97)	**0.04**	168	77	0.71 (0.38–1.34)	0.29	69	28	0.41 (0.18–0.96)	**0.04**
	A:A+T:A	147	63	Ref.		105	45	Ref.		42	18	Ref.	
	T:T	119	60	1.04 (0.73–1.48)	0.83	83	43	1.02 (0.67–1.56)	0.91	9	7	1.05 (0.54–2.05)	0.87
RFS													
Patients receiving no treatment													
SLC22A2 Rs694812	A:A	644	236	Ref.		378	134	Ref.		266	102	Ref.	
	G:A	105	48	1.24 (0.91–1.69)	0.17	60	36	1.70 (1.18–2.46)	**0.005**	45	12	0.67 (0.37–1.23)	0.20
	G:G	4	1	0.55 (0.08–3.90)	0.55	1	1	1.43 (0.20–10.20)	0.72	3	0	NA	0.98
	G:A+G:G	109	49	1.21 (0.89–1.64)	0.23	61	37	1.69 (1.18–2.44)	**0.005**	48	12	0.63 (0.35–1.15)	0.13
	A:A+G:A	749	284	Ref.		438	170	Ref.		311	114	Ref.	
	G:G	4	1	0.51 (0.07–3.64)	0.50	1	1	1.34 (0.19–9.55)	0.77	3	0	NA	0.98
SLC22A2 Rs2450975	C:C	517	190	Ref.		301	104	Ref.		216	86	Ref.	
	C:A	207	84	1.10 (0.85–1.42)	0.49	121	59	1.36 (0.99–1.87)	0.06	86	25	0.75 (0.48–1.18)	0.22
	A:A	22	11	1.62 (0.88–2.99)	0.12	12	8	3.42 (1.66–7.07)	**0.0009**	10	3	0.62 (0.19–1.96)	0.41
	C:A+A:A	229	95	1.14 (0.89–1.46)	0.30	133	67	1.47 (1.08–1.99)	**0.01**	96	28	0.74 (0.48–1.13)	0.17
	C:C+C:A	724	274	Ref.		422	163	Ref.		302	111	Ref.	
	A:A	22	11	1.58 (0.86–2.88)	0.14	12	8	3.21 (1.57–6.56)	**0.001**	10	3	0.66 (0.21–2.08)	0.48
SLC29A3 Rs780680	G:G	597	216	Ref.		342	126	Ref.		255	90	Ref.	
	T:G	141	64	1.41 (1.07–1.87)	**0.02**	88	43	1.59 (1.13–2.25)	**0.009**	53	21	1.14 (0.71–1.83)	0.60
	T:T	9	5	1.22 (0.5–2.97)	0.66	6	2	0.64 (0.16–2.58)	0.53	3	3	3.22 (1.01–10.2)	0.05
	T:G+T:T	150	69	1.40 (1.07–1.84)	**0.02**	94	45	1.50 (1.06–2.10)	**0.02**	56	24	1.24 (0.79–1.94)	0.36
	G:G+T:G	738	280	Ref.		430	169	Ref.		308	111	Ref.	
	T:T	9	5	1.15 (0.47–2.78)	0.76	6	2	0.59 (0.14–2.37)	0.45	3	3	3.09 (0.98–9.78)	0.05
Patients receiving a 5-FU based therapy													
SLC22A3 rs2076828	C:C	129	59	Ref.		89	40	Ref.		40	19	Ref.	
	G:C	204	65	0.58 (0.41–0.83)	**0.003**	132	39	0.56 (0.36–0.88)	**0.01**	72	26	0.63 (0.35–1.14)	0.12
	G:G	84	29	0.60 (0.39–0.94)	**0.03**	54	19	0.68 (0.39–1.17)	0.17	30	10	0.48 (0.22–1.04)	0.06

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 3. Continued

Gene SNP	Genotype	All cancer patients				Colonsigmoideum cancer patients				Rectal cancer patients			
		N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value	N[a]	Events	HR (95% CI)[b]	P-value
	G:C+G:G	288	94	0.58 (0.42–0.81)	**0.001**	186	58	0.59 (0.40–0.89)	**0.01**	102	36	0.57 (0.33–1.00)	**0.05**
	C:C+G:C	333	124	Ref.		221	79	Ref.		112	45	Ref.	
	G:G	84	29	0.82 (0.55–1.24)	0.35	54	19	0.93 (0.57–1.54)	0.79	30	10	0.65 (0.33–1.29)	0.22
SLC22A3 Rs3088442	G:G	181	59	Ref.		117	38	Ref.		64	21	Ref.	
	G:A	173	63	1.42 (1.00–2.03)	**0.05**	117	42	1.36 (0.88–2.11)	0.17	56	21	1.56 (0.85–2.88)	0.15
	A:A	67	33	1.73 (1.13–2.65)	**0.01**	43	20	1.39 (0.81–2.39)	0.23	24	13	2.84 (1.41–5.71)	**0.003**
	G:A+A:A	240	96	1.52 (1.10–2.10)	**0.01**	160	62	1.38 (0.92–2.06)	0.12	80	34	1.84 (1.06–3.19)	0.03
	G:G+G:A	354	122	Ref.		234	80	Ref.		120	42	Ref.	
	A:A	67	33	1.48 (1.00–2.17)	**0.05**	43	20	1.20 (0.73–1.96)	0.47	24	13	2.25 (1.21–4.21)	**0.01**
SLC3A1 Rs1053909	G:G	188	73	Ref.		116	44	Ref.		72	29	Ref.	
	G:A	188	57	0.71 (0.50–1.00)	**0.05**	130	40	0.74 (0.48–1.13)	0.16	58	17	0.65 (0.36–1.18)	0.16
	A:A	46	23	1.35 (0.84–2.15)	0.21	32	14	1.16 (0.63–2.11)	0.64	14	9	1.87 (0.87–3.99)	0.11
	G:A+A:A	234	80	0.82 (0.59–1.12)	0.21	162	54	0.81 (0.54–1.21)	0.30	72	26	0.84 (0.49–1.42)	0.51
	G:G+G:A	376	130	Ref.		246	84	Ref.		130	46	Ref.	
	A:A	46	23	1.59 (1.02–2.48)	**0.04**	32	14	1.34 (0.76–2.36)	0.31	14	9	2.38 (1.15–4.91)	**0.02**
Patients receiving a 5-FU + oxaliplatin-based therapy													
SLC19A1 rs1051298	C:C	82	53	Ref.		57	36	Ref.		25	17	Ref.	
	T:C	140	88	1.07 (0.76–1.50)	0.71	105	68	1.08 (0.72–1.62)	0.71	35	20	1.04 (0.54–2.00)	0.91
	T:T	51	35	1.28 (0.83–1.97)	0.26	32	20	1.01 (0.58–1.76)	0.96	19	15	2.05 (1.00–4.20)	**0.05**
	T:C+T:T	191	123	1.11 (0.81–1.54)	0.52	137	88	1.06 (0.71–1.56)	0.78	54	35	1.32 (0.74–2.38)	0.35
	C:C+T:C	222	141	Ref.		162	104	Ref.		60	37	Ref.	
	T:T	51	35	1.20 (0.83–1.74)	0.34	32	20	0.95 (0.59–1.54)	0.84	19	15	1.83 (1.00–2.37)	**0.05**
SLC28A3 Rs11140489	T:T	199	123	Ref.		142	87	Ref.		57	36	Ref.	
	T:A	71	49	0.45 (0.16–1.27)	0.13	51	35	0.97 (0.13–7.15)	0.98	20	14	0.25 (0.06–0.99)	**0.05**
	A:A	4	4	0.39 (0.14–1.05)	0.06	1	1	0.73 (0.10–5.29)	0.76	3	3	0.32 (0.09–1.05)	0.06
	T:A+T:T	270	172	0.40 (0.15–1.09)	0.07	193	122	0.79 (0.11–5.69)	0.82	77	50	0.29 (0.09–0.96)	**0.04**
	A:A+T:A	75	53	Ref.		52	36	Ref.		23	17	Ref.	
	A:A	4	4	0.85 (0.62–1.17)	0.32	1	1	0.83 (0.56–1.23)	0.36	3	3	0.92 (0.52–1.65)	0.79
SLCO3A1 Rs8174	G:G	110	69	Ref.		77	47	Ref.		33	22	Ref.	
	G:A	123	77	1.02 (0.74–1.41)	0.92	85	53	1.03 (0.69–1.53)	0.88	38	24	1.01 (0.57–1.81)	0.96
	A:A	39	30	1.73 (1.12–2.67)	**0.01**	29	22	1.67 (1.00–2.79)	**0.05**	10	8	1.96 (0.86–4.47)	0.11
	G:A+A:A	162	107	1.15 (0.85–1.55)	0.38	114	75	1.16 (0.81–1.67)	0.42	48	32	1.14 (0.66–1.97)	0.64
	G:G+G:A	233	146	Ref.		162	100	Ref.		71	46	Ref.	
	A:A	39	30	1.68 (1.13–2.5)	**0.01**	29	22	1.64 (1.03–2.61)	**0.04**	10	8	1.98 (0.92–4.23)	0.08
SLC3A1 Rs2270061	A:A	29	22	Ref.		20	13	Ref.		9	9	Ref.	
	T:A	118	65	0.68 (0.42–1.11)	0.12	85	48	0.77 (0.42–1.43)	0.41	33	17	0.57 (0.25–1.3)	0.18
	T:T	119	82	0.84 (0.52–1.34)	0.46	83	56	0.95 (0.52–1.73)	0.86	36	26	0.72 (0.33–1.56)	0.40
	T:A+T:T	237	147	0.76 (0.48–1.19)	0.23	168	104	0.86 (0.48–1.52)	0.60	69	43	0.64 (0.31–1.33)	0.24
	A:A+T:A	147	87	Ref.		105	61	Ref.		42	26	Ref.	
	T:T	119	82	1.15 (0.85–1.55)	0.37	83	56	1.15 (0.8–1.66)	0.44	36	26	1.12 (0.65–1.93)	0.69

NA, not applicable.

Results that passed the Benjamini–Hochberg False Discovery Rate test for multiple comparisons are in bold.

[a]Numbers may not add up to 100% of available subjects because of genotyping failure.

[b]Adjusted for age and sex.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

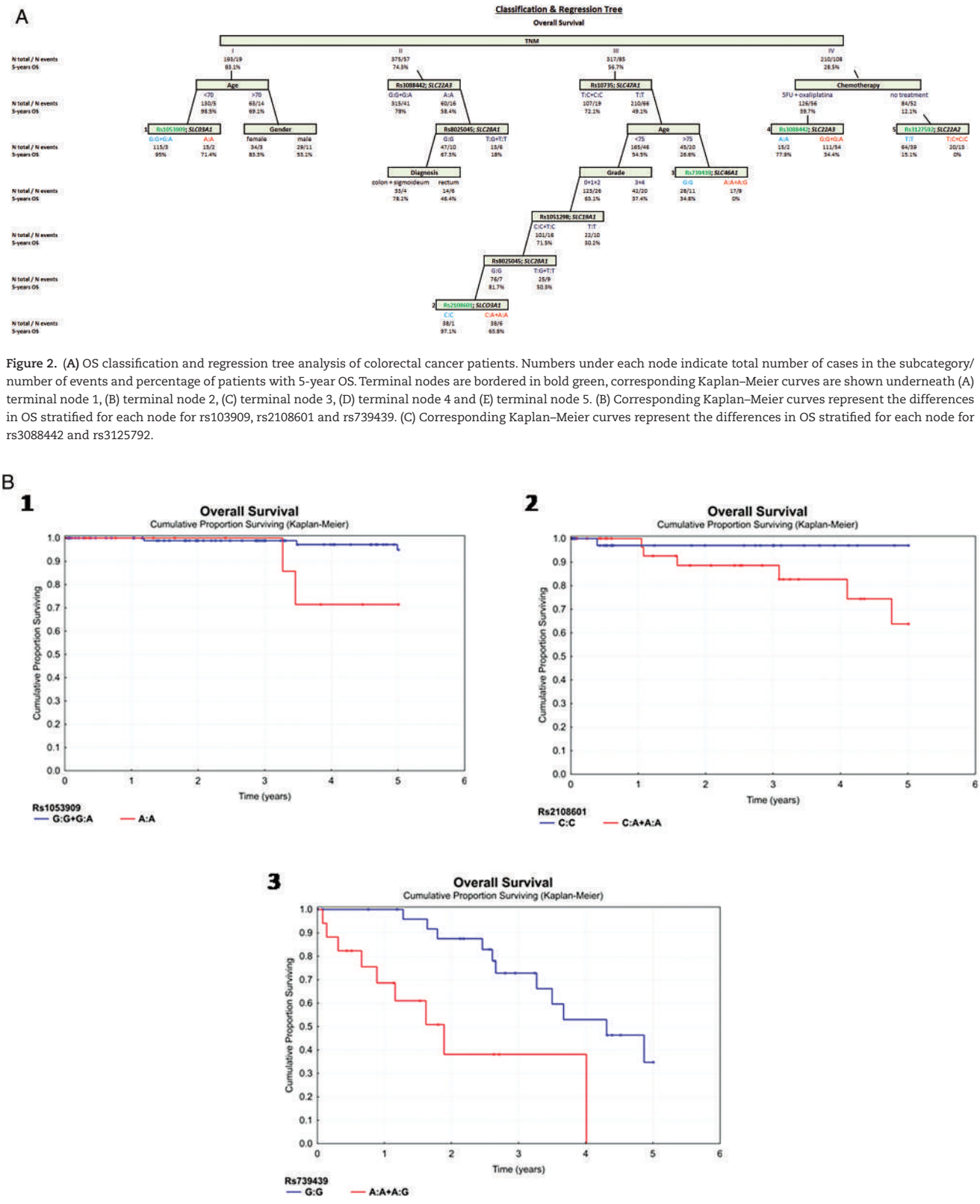

Figure 2. **(A)** OS classification and regression tree analysis of colorectal cancer patients. Numbers under each node indicate total number of cases in the subcategory/number of events and percentage of patients with 5-year OS. Terminal nodes are bordered in bold green, corresponding Kaplan–Meier curves are shown underneath (A) terminal node 1, (B) terminal node 2, (C) terminal node 3, (D) terminal node 4 and (E) terminal node 5. (B) Corresponding Kaplan–Meier curves represent the differences in OS stratified for each node for rs103909, rs2108601 and rs739439. (C) Corresponding Kaplan–Meier curves represent the differences in OS stratified for each node for rs3088442 and rs3125792.

Figure 2. Continued.

Discussion

Genetic variations affecting miRNA functions may have a pathogenic role in cancer (20,29) and miRSNPs have demonstrated a high penetrance for certain tumor phenotypes (39). Moreover, miRNA-binding site sequences in the majority of protein-coding genes are highly conserved (40). In the recent years, a number of studies have suggested the importance of inherited variants in miRNA target sites for human disease susceptibility and progression (20–22,24,27,28,41–44). It is well known that there are

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

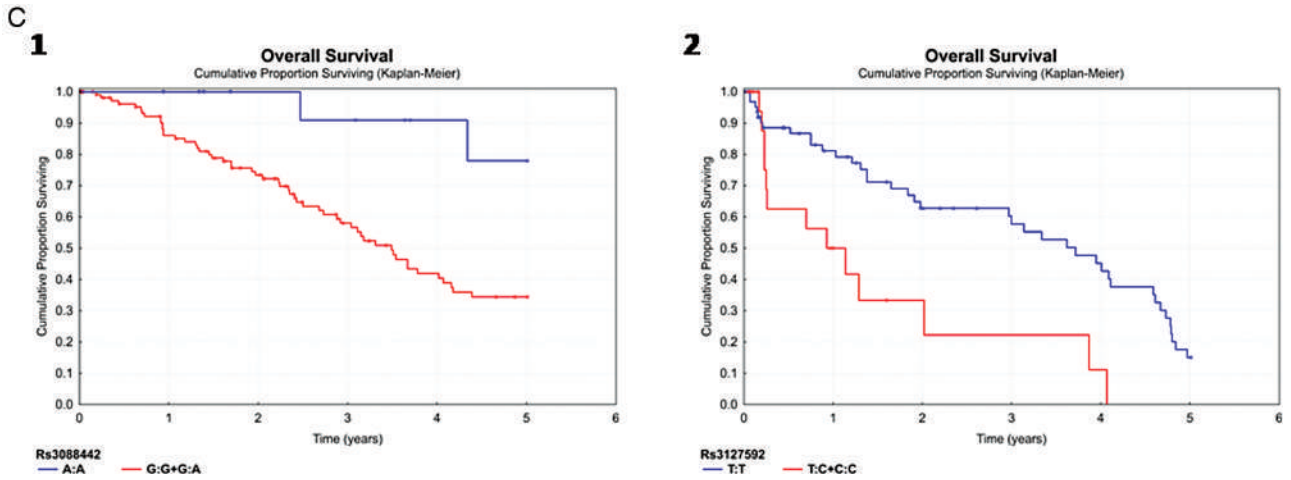

Figure 2. Continued.

considerable inter-individual differences in response to drug administrations and we assumed that the genetic variations may partially explain these differences. In the present study, we hypothesized that miRSNPs in target regions of *SLC* transporter genes might modulate the efficiency of translation of the corresponding proteins. This may have an impact on individual's susceptibility to CRC and affect the response to therapy. To our knowledge, this is the first study investigating the role of variations in SLC on CRC susceptibility or patient's prognosis.

In the present study, we performed the survival analyses both by traditional Cox regression and by survival tree (CART). The reason for using these two different approaches is that the results from survival tree complement the outcomes from traditional survival analysis (45). In fact, CART analyses can be beneficial to obtain additional information in data that is not captured by conventional survival analysis. Unlike logistic and linear regressions, CART does not create a prediction equation. Instead, the data are divided along the predictor axes into subsets with homogeneous values of the dependent variable—a process represented by a decision tree that can be used to make predictions from new observations. By comparing the results from both analyses, we observed significant overlaps and, these overlays will be discussed in details below.

In the traditional Cox regression, CRC and rectal cancer patients receiving a 5-FU-based therapy and carriers of at least one A allele for *SLC22A3* rs3088442 had higher risk of relapse or metastasis. Rectal cancer patients carrying homozygous genotype AA for *SLCO3A1* rs1053909 and also receiving the 5-FU based therapy displayed worse OS and shorter RFS. All CRC patients carrying homozygous AA genotype for *SLCO3A1* rs8174 and receiving 5-FU regimen in combination with oxaliplatin showed shorter RFS. All these miRSNPs were also crucial discriminant in CART analysis.

In the survival CART analyses, besides well-known prognostic factors, as TNM stage and gender, we observed significant associations of several miRSNPs with patients' survival and clinical outcomes. miRSNPs in *SLCO3A1* and *SLC22A3* genes appeared several times as the optimal split factors in the CART analyses (*SLCO3A1*: rs8174 three times and rs1053909 once; *SLC22A3*: rs3088442 three times). These outcomes suggest their potential relevance to patients' survival. It is interesting to observe that results from multivariate analysis overlapped with those from univariate analysis and the same miRSNPs evinced the same trends. In CART analysis, among CRC stage I, carriers of homozygous GG genotype of rs8174 in *SLCO3A1* evinced better prognosis when compared with A allele carriers.

The role of *SLCO3A1* gene in CRC pathogenesis is not fully understood yet, but there is evidence in other cancers that may be transferable to CRC. For instance, in breast cancer, estrone sulfate transporters, including *SLCO3A1* (*OATP3A1*), are primarily localized in the plasma membranes of malignant cells, and they have higher expression levels in hormone receptor-positive (HR+) tissues compared with HR tissues (46). On the other hand, the outcomes of prostate cancer patients might be potentially influenced by the overexpression of *SLCO3A1* in castration-resistant prostate cancer as observed by Wright et al. (47). Moreover, *SLCO3A1* was also upregulated in pancreatic and liver tumors (48). Hence, alterations in the mRNA levels of *SLCO3A1* in CRC tumors may be important contributors to the intra-tissue accumulation of various endogenous and exogenous compounds. Estrone sulfate transport mediated by *SLCO3A1* might facilitate colorectal synthesis of 17-b-estradiol that prevents CRC.

The functional prediction analysis showed that rs8174 in *SLCO3A1* gene is located in the region of H3K4me1 and H3K27ac: these histone marks are associated with the activity of enhancers and promote the transcription. Interestingly, regions marked by H3K27ac are considered as active promoters while regions occupied by H3K4me1 are considered as active enhancers (49). Therefore, rs8174 could affect the expression of *SLCO3A1* by influencing its promoter activity.

Let-7i-3p is predicted to bind to the common G allele in rs8174. Hur et al. (50) reported that low expression levels of let-7i are an independent predictor of distant metastasis in primary CRC tissues. Similarly, Zhang et al. (51) observed that the upregulation of let-7i correlated with decreased potential of metastasis in CRC cell lines. Recently, Coebergh et al. (52) identified two miRNAs (let-7i and miR-10b) whose expression could predict hepatic recurrence in CRC patients with stages I–II, but not the overall ability to develop distant metastasis. Further studies are needed to explain the role of let-7i in CRC and establish its clinical application for CRC therapy, diagnosis and prognosis.

SLC22A3, together with SLC22A1, SLC22A2, mediates the transport of a variety of structurally differing cations comprising both endogenous and exogenous compounds, e.g. neurotransmitters such as catecholamines and xenobiotics (including drugs), respectively (53,54). According to the Human protein Atlas (55), these proteins are expressed in CRC.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 4. List of miRSNPs included in the study and characteristics of their relative miRNAs binding to each polymorphism site

		miRNA-binding site		miRNA binding		Binding site			
Gene	miRSNP	Allele 1	Allele 2	Predicted[a]	Validated[b]	No change	Loss	New	Binding energy ΔΔG[c]
SLC22A2	*rs694812*	G	A	X					
		miR-34c-3p		X			x		
		miR-34b-3p		X			x		
		miR-492	miR-492	X		x			−2.00
		miR-640		X			x		
			miR-3159	X				x	
			miR-3920	x				x	
		miR-4533		x			x		
			miR-4517	x					
		miR-4720-3p	miR-4720-3p	x		x			1.80
		miR-4751	miR-4751	x		x			0.80
		miR-4772-5p		x			x		
		miR-4777-3p	miR-4777-3p	x		x			0
		miR-4802-5p		x			x		
		miR-5196-5p	miR-5196-5p	x		x			0
		miR-5583-3p	miR-5583-3p	x		x			−0.60
		miR-6072		x			x		
		miR-6079		x			x		
	rs2450975	A	C						
		miR-22-5p		x			x		
		miR-412		x			x		
		miR-876-5p		x			x		
		miR-3606-3p	miR-3606-3p	x		x			0
		miR-3924		x			x		
		miR-4495	miR-4495	x		x			0
		miR-5089-5p		x			x		
		miR-5096		x			x		
		miR-6515-3p		x			x		
SLC29A3	*rs780680*	G	T						
		miR-676-3p		x			x		
			miR-668	x				x	
		miR-3657		x			x		
		miR-3664-5p		x			x		
			miR-4480	x				x	
		miR-5691	miR-5691	x		x			−3.30
SLC22A3	*rs2076828*	C	G						
			miR-30d-5p	x				x	
			miR-200a-3p	x				x	
			miR-449b-3p	x				x	
		miR-575		x			x		
			miR-3120-3p	x				x	
			miR-3157-5p	x				x	
			miR-4672	x				x	
		miR-4797-5p		x			x		

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 4. Continued

Gene	miRSNP	miRNA-binding site		miRNA binding		Binding site			Binding energy ΔΔG[c]
		Allele 1	Allele 2	Predicted[a]	Validated[b]	No change	Loss	New	
	rs3088442	G	A						
		miR-130a-5p	miR-130a-5p	x		x			0.20
			miR-147a	x				x	
		miR-147b		x			x		
		miR-675-5p		x			x		
		miR-933		x			x		
			miR-3138	x				x	
			miR-3190-3p	x				x	
			miR-4317	x				x	
			miR-4455	x				x	
			miR-4657	x				x	
		miR-4761-5p		x			x		
		miR-6125		x			x		
SLC19A1	*rs1051298*	C	T						
		miR-634	miR-634	x		x			0
			miR-1227-5p	x				x	
			miR-1250	x				x	
			miR-3194-5p	x				x	
		miR-3648		x			x		
			miR-3922-5p	x				x	
			miR-4435	x				x	
		miR-4726-5p	miR-4726-5p	x		x			0
		miR-4786-3p	miR-4786-3p	x		x			−0.90
			miR-5092	x				x	
			miR-5704	x				x	
SLC28A3	*rs11140489*	G	T						
			miR-136-3p	x				x	
		miR-192-5p	miR-192-5p	x		x			0.80
		miR-221-5p	miR-221-5p	x		x			−1.10
		miR-202-5p	miR-202-5p	x		x			−2.10
			miR-204-3p	x				x	
		miR-215	miR-215	x		x			−0.20
			miR-337-3p	x				x	
		miR-487a	miR-487a	x		x			0.80
			miR-548g-3p	x				x	
		miR-551b-5p	miR-551b-5p	x		x			−0.20
		miR-649	miR-649	x		x			0.20
		miR-661		x			x		
		miR-767-3p		x			x		
		miR-1292-5p	miR-1292-5p	x		x			−2.60
		miR-3200-5p		x			x		
		miR-3153	miR-3153	x		x			−0.10
			miR-3187-5p	x				x	

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

Table 4. Continued

Gene	miRSNP	miRNA-binding site		miRNA binding		Binding site			Binding energy ΔΔG[c]
		Allele 1	Allele 2	Predicted[a]	Validated[b]	No change	Loss	New	
			miR-3192	x				x	
		miR-4471	miR-4471	x		x			−0.80
			miR-4646-5p	x				x	
		miR-4668-5p		x			x		
		miR-4727-5p		x			x		
			miR-4738-3p	x				x	
		miR-6071		x			x		
SLCO3A1	*rs8174*	T	C						
		let-7i-3p		x			x		
		miR-3648	miR-3648	x		x			1.30
		miR-4796-3p		x			x		
	rs2270061	T	A						
			miR-27a-5p	x				x	
			miR-185-3p	x				x	
		miR-323a-5p	miR-323a-5p	x		x			4.40
		miR-541-3p	miR-541-3p	x		x			1.70
		miR-608	miR-608	x		x			−0.10
		miR-654-5p	miR-654-5p	x		x			5.80
		miR-876-3p		x			x		
		miR-1293	miR-1293	x		x			−1.60
		miR-3151		x			x		
		miR-4296		x			x		
		miR-4417		x			x		
		miR-4437		x			x		
			miR-4446-3p	x				x	
		miR-4456		x			x		
		miR-4483		x			x		
			miR-4675	x				x	
			miR-5001-5p	x				x	
	rs1053909	G	A						
		miR-63		x			x		
			miR-155-5p	x				x	
		miR-579	miR-579	x		x			0
		miR-890		x			x		
		miR-4520b-3p	miR-4520b-3p	x		x			−0.10
			miR-4728-3p	x				x	
		miR-5696	miR-5696	x		x			0.30
		miR-6079		x			x		
		miR-6505-5p	miR-6505-5p	x		x			4.10

[a]According to MicroSniper tool.
[b]According to miRWalk.
[c]The expected ΔΔG for each given polymorphism is calculated as the difference of ΔG needed for the common allele and the ΔG needed for the rare allele.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

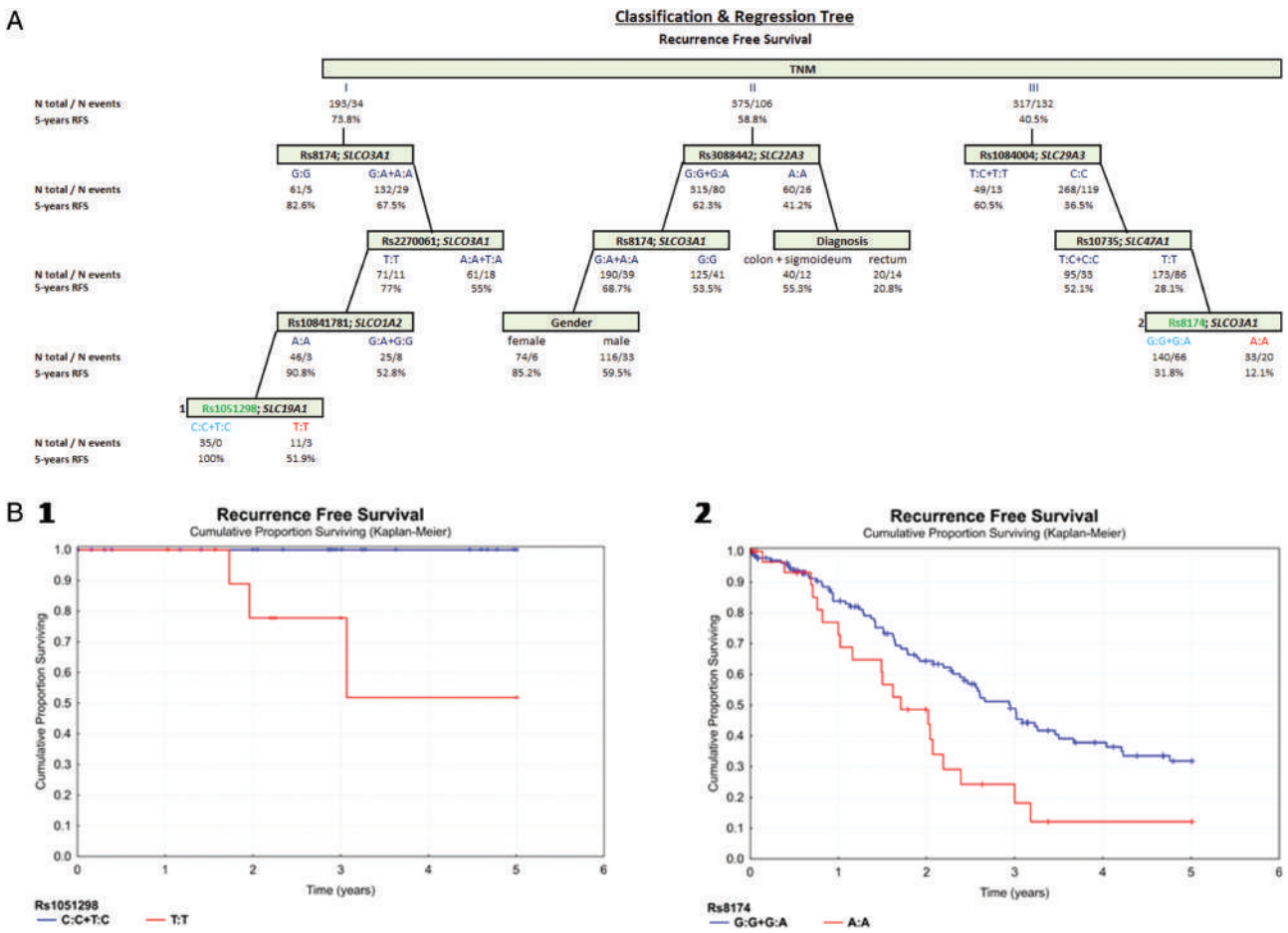

Figure 3. (A) Recurrence-free survival (RFS) classification and regression tree of colorectal cancer patients. Numbers under each node indicate total number of cases in the subcategory/number of events and percentage of patients with 5-year RFS. Terminal nodes are bordered in bold green, corresponding Kaplan–Meier curves are shown underneath (A) terminal node 1 and (B) terminal node 2. (B) Corresponding Kaplan–Meier curves represent the differences in RFS stratified for each node for rs1051298 and rs8174.

The results of the 5-year OS and RFS CART analysis showed that rs3088442 in *SLC22A3* was chosen three times as the optimal split and once even for the CRC stage IV patients receiving 5-FU+oxaliplatin-based chemotherapy. This finding indicates its possible use in personalized treatment strategies for identifying CRC stage IV patients who are likely to benefit from palliative therapy. In traditional Cox regression, CRC patients carrying at least one A allele for rs3088442 in *SLC22A3* had higher risk of relapse or metastasis. After stratification, the same effect was observed in rectal cancer patients. In the CART analysis, rectal cancer patients at stage II and with rs3088442 variant AA genotype had worse prognosis than those with colon cancer or with GG+GA genotypes. For this particular miRSNP, we can thus assume a rectal cancer-specific effect. Only one miRNA, miR-130a-5p, is predicted to bind to both alleles in the region where rs3088442 lies. Unfortunately, the presence of one or the other allele did not alter the binding free-energy necessary to this miRNA to bind to its target. Besides, in the presence of common G allele at rs3088442, miR-147b is also binding in the same position of this miRSNP. Conversely, miR-147a, which is predicted to be upregulated by 5-FU administration, binds more favorably to the variant A allele, as observed in (56). This *in silico* outcome is in agreement with our results where CRC carriers of at least one A allele for *SLC22A3* rs3088442 had higher risk of relapse or metastasis after 5-FU-based therapy. A few studies have previously described that miR-147 family is abnormally expressed in several types of cancers (57–60) and regulates several cellular events, including the inhibition of the stem cell-like traits (61), proliferation (62) and cell cycle arrest (63). Shen et al. (62) postulated that miR-147 family suppresses the proliferation and enhances the chemosensitivity of gastric cancer cells to 5-FU by promoting cell apoptosis via targeting directly PTEN and regulating the PI3K/AKT signalling pathway. It was not possible to retrieve which miR-147 was specifically used in the previous study. However, together with our results, these facts indicated that miR-147 family could represent promising therapeutic targets for CRC treatment. In the presence of both alleles of rs3088442 miRSNP in *SLC22A3* gene, there are several miRNAs predicted to bind in the region. However, for many of them, there is no evidence on their functionality or expression levels in colorectal tissues, although miR-675-5p has been identified as overexpressed in metastatic colon cancer cells (64).

Recently, Ren et al. (65) revealed a novel susceptible locus, rs420038 in *SLC22A3* gene, as involved in CRC development and progression. In particular, the carriers of at least one A allele were at a lower risk of CRC than those carrying the GG genotype. Other genetic variants in *SLC22A3* are associated with both disease risk and patients' survival for different tumor types, such as pancreas, prostate and CRC (66–70). None of these associated

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

SNPs were found in linkage disequilibrium with the miRSNPs examined here.

We are aware of some limitations of the present study. For instance, there was a different distribution in age and gender between cases and controls, as well as other parameters such as BMI. The original cohort included more subjects, but there was a strong difference in age and gender between cases and controls. To overcome this issue, we attempted to control the potential age and sex effect by approximating these differences between case and control groups with a tolerance limit of ±5 years and ±10% representation of sex using the equivalence test.

In summary, we provided observational and bioinformatics evidence that even subtle alterations in specific *SLC* transporters genes may contribute to the clinical outcome of CRC patients. In the future, it will be important to functionally characterize these miRSNPs and to find biological mechanisms underlying the associations to assess these variations as diagnostic, prognostic and/or predictive biomarkers in CRC.

Acknowledgement

The authors are thankful to all volunteers who contributed their biological material to the study and to all hospital employees who participated in sample collection.

Funding

This project was supported by the Grant Agency of the Ministry of Health of the Czech Republic (17-30920A) and the Grant Agency of the Czech Republic (18-09709S). We are thankful to Charles University Research Centre program UNCE/MED/006 "University Centre of Clinical and Experimental Liver Surgery" and INTER-COST project no. LTC19015 provided by the Ministry of Education Youth and Sports of the Czech Republic and by project No. CZ.02.1.01/0.0/0.0/16_019/0000787 "Fighting INfectious Diseases", awarded by the "MEYS CR", financed from EFRR.
This article is based upon work from COST Action CA17118, supported by COST (European Cooperation in Science and Technology) www.cost.eu.
Conflict of Interest Statement: None declared.

References

1. Haggar, F.A. *et al.* (2009) Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin. Colon Rectal Surg.*, 22, 191–197.
2. Siegel, R.L. *et al.* (2020) Cancer statistics, 2020. *CA Cancer J. Clin.*, 70, 7–30.
3. Keum, N. *et al.* (2019) Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.*, 16, 713–732.
4. Siegel, R.L. *et al.* (2020) Colorectal cancer statistics, 2020. *CA Cancer J. Clin.*, 70, 145–164.
5. Vodenkova, S. *et al.* (2020) 5-Fluorouracil and other fluoropyrimidines in colorectal cancer: past, present and future. *Pharmacol. Ther.*, 206, 107447.
6. Vymetalkova, V. *et al.* (2020) Expression quantitative trait loci in ABC transporters are associated with survival in 5-FU treated colorectal cancer patients. *Mutagenesis*, 35, 273–281.
7. Mohelnikova-Duchonova, B. *et al.* (2013) The association between the expression of solute carrier transporters and the prognosis of pancreatic cancer. *Cancer Chemother. Pharmacol.*, 72, 669–682.
8. Hlavata, I. *et al.* (2012) The role of ABC transporters in progression and clinical outcome of colorectal cancer. *Mutagenesis*, 27, 187–196.
9. Lemstrová, R. *et al.* (2014) Role of solute carrier transporters in pancreatic cancer: a review. *Pharmacogenomics*, 15, 1133–1145.
10. Huang, Y. *et al.* (2008) Genetic variations and gene expression of transporters in drug disposition and response. *Expert Opin. Drug Metab. Toxicol.*, 4, 237–254.
11. Huang, Y. *et al.* (2006) Membrane transporters and channels in chemoresistance and -sensitivity of tumor cells. *Cancer Lett.*, 239, 168–182.
12. Sissung, T.M. *et al.* (2012) Transporter pharmacogenetics: transporter polymorphisms affect normal physiology, diseases, and pharmacotherapy. *Discov. Med.*, 13, 19–34.
13. Mackey, J.R. *et al.* (1998) Functional nucleoside transporters are required for gemcitabine influx and manifestation of toxicity in cancer cell lines. *Cancer Res.*, 58, 4349–4357.
14. Baldwin, S.A. *et al.* (2004) The equilibrative nucleoside transporter family, SLC29. *Pflugers Arch.*, 447, 735–743.
15. Sprowl, J.A. *et al.* (2013) Polymorphic transporters and platinum pharmacodynamics. *Drug Metab. Pharmacokinet.*, 28, 19–27.
16. Kobayashi, Y. *et al.* (2005) Transport mechanism and substrate specificity of human organic anion transporter 2 (hOat2 [SLC22A7]). *J. Pharm. Pharmacol.*, 57, 573–578.
17. He, Y. *et al.* (2015) The effects of microRNA on the absorption, distribution, metabolism and excretion of drugs. *Br. J. Pharmacol.*, 172, 2733–2747.
18. Slattery, M.L. *et al.* (2011) MicroRNAs and colon and rectal cancer: differential expression by tumor location and subtype. *Genes. Chromosomes Cancer*, 50, 196–206.
19. Wei, L. *et al.* (2019) The emerging role of noncoding RNAs in colorectal cancer chemoresistance. *Cell. Oncol. (Dordr.)*, 42, 757–768.
20. Landi, D. *et al.* (2008) Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. *Carcinogenesis*, 29, 579–584.
21. Naccarati, A. *et al.* (2012) Polymorphisms in miRNA-binding sites of nucleotide excision repair genes and colorectal cancer risk. *Carcinogenesis*, 33, 1346–1351.
22. Vymetalkova, V. *et al.* (2014) Variations in mismatch repair genes and colorectal cancer risk and clinical outcome. *Mutagenesis*, 29, 259–265.
23. Pardini, B. *et al.* (2013) Variation within 3'-UTRs of base excision repair genes and response to therapy in colorectal cancer patients: a potential modulation of microRNAs binding. *Clin. Cancer Res.*, 19, 6044–6056.
24. Naccarati, A. *et al.* (2016) Double-strand break repair and colorectal cancer: gene variants within 3' UTRs and microRNAs binding as modulators of cancer risk and clinical outcome. *Oncotarget*, 7, 23156–23169.
25. Jiraskova, K., *et al.* (2018) Functional polymorphisms in DNA repair genes are associated with sporadic colorectal cancer susceptibility and clinical outcome. *Int. J. Mol. Sci.*, 20, 9.
26. Lu, S. *et al.* (2019) Single nucleotide polymorphisms within MUC4 are associated with colorectal cancer survival. *PLoS One*, 14, e0216666.
27. Pardini, B. *et al.* (2015) Polymorphisms in microRNA genes as predictors of clinical outcomes in colorectal cancer patients. *Carcinogenesis*, 36, 82–86.
28. Vymetalkova, V. *et al.* (2017) Polymorphisms in microRNA binding sites of mucin genes as predictors of clinical outcome in colorectal cancer patients. *Carcinogenesis*, 38, 28–39.
29. Schneiderova, M. *et al.* (2017) MicroRNA-binding site polymorphisms in genes involved in colorectal cancer etiopathogenesis and their impact on disease prognosis. *Mutagenesis*, 32, 533–542.
30. Barenboim, M. *et al.* (2010) MicroSNiPer: a web tool for prediction of SNP effects on putative microRNA targets. *Hum. Mutat.*, 31, 1223–1232.
31. Liu, C. *et al.* (2012) MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. *BMC Genomics*, 13, 661.
32. Ziebarth, J.D. *et al.* (2012) PolymiRTS Database 2.0: linking polymorphisms in microRNA target sites with human diseases and complex traits. *Nucleic Acids Res.*, 40, D216–D221.
33. Mustonen, R. *et al.* (1992) 7-Methylguanine levels in DNA of smokers' and non-smokers' total white blood cells, granulocytes and lymphocytes. *Carcinogenesis*, 13, 1951–1955.
34. Consortium, G.T. *et al.* (2017) Genetic effects on gene expression across human tissues. *Nature*, 550, 204–213.
35. Li, J. *et al.* (2016) Network-based identification of microRNAs as potential pharmacogenomic biomarkers for anticancer drugs. *Oncotarget*, 7, 45584–45596.
36. Cheng, F. *et al.* (2012) Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput. Biol.*, 8, e1002503.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022

37. Li, J. *et al.* (2014) Computational prediction of microRNA networks incorporating environmental toxicity and disease etiology. *Sci. Rep.,* 4, 5576.
38. Barlin, J.N. *et al.* (2013) Classification and regression tree (CART) analysis of endometrial carcinoma: seeing the forest for the trees. *Gynecol. Oncol.,* 130, 452–456.
39. Ryan, B.M. *et al.* (2015) Identification of a functional SNP in the 3'UTR of CXCR2 that is associated with reduced risk of lung cancer. *Cancer Res.,* 75, 566–575.
40. Ryan, B.M. *et al.* (2010) Genetic variation in microRNA networks: the implications for cancer research. *Nat. Rev. Cancer,* 10, 389–402.
41. Landi, D. *et al.* (2012) Identification of candidate genes carrying polymorphisms associated with the risk of colorectal cancer by analyzing the colorectal mutome and microRNAome. *Cancer,* 118, 4670–4680.
42. Landi, D. *et al.* (2011) Polymorphisms affecting micro-RNA regulation and associated with the risk of dietary-related cancers: a review from the literature and new evidence for a functional role of rs17281995 (CD86) and rs1051690 (INSR), previously associated with colorectal cancer. *Mutat. Res.,* 717, 109–115.
43. Vodicka, P. *et al.* (2016) Polymorphisms in non-coding RNA genes and their targets sites as risk factors of sporadic colorectal cancer. *Adv. Exp. Med. Biol.,* 937, 123–149.
44. Vymetalkova, V.P. *et al.* (2014) Molecular characteristics of mismatch repair genes in sporadic colorectal tumors in Czech patients. *BMC Med. Genet.,* 15, 17.
45. Zhou, Y. *et al.* (2015) Rationale and applications of survival tree and survival ensemble methods. *Psychometrika,* 80, 811–833.
46. Banerjee, N. *et al.* (2015) (125)I-Labelled 2-Iodoestrone-3-sulfate: synthesis, characterization and OATP mediated transport studies in hormone dependent and independent breast cancer cells. *Nucl. Med. Biol.,* 42, 274–282.
47. Wright, J.L. *et al.* (2011) Expression of SLCO transport genes in castration-resistant prostate cancer and impact of genetic variation in SLCO1B3 and SLCO2B1 on prostate cancer outcomes. *Cancer Epidemiol. Biomarkers Prev.,* 20, 619–627.
48. Hays, A. *et al.* (2013) Organic anion transporting polypeptides expressed in pancreatic cancer may serve as potential diagnostic markers and therapeutic targets for early stage adenocarcinomas. *Pharm. Res.,* 30, 2260–2269.
49. Nagarajan, S. *et al.* (2014) Bromodomain protein BRD4 is required for estrogen receptor-dependent enhancer activation and gene transcription. *Cell Rep.,* 8, 460–469.
50. Hur, K. *et al.* (2015) Identification of a metastasis-specific MicroRNA signature in human colorectal cancer. *J. Natl Cancer Inst.,* 107, 3.
51. Zhang, P. *et al.* (2012) Comprehensive gene and microRNA expression profiling reveals the crucial role of hsa-let-7i and its target genes in colorectal cancer metastasis. *Mol. Biol. Rep.,* 39, 1471–1478.
52. Coebergh van den Braak, R.R.J. *et al*; MATCH Study Group*. (2018) Confirmation of a metastasis-specific microRNA signature in primary colon cancer. *Sci. Rep.,* 8, 5242.
53. Cano-Soldado, P. *et al.* (2012) Transporters that translocate nucleosides and structural similar drugs: structural requirements for substrate recognition. *Med. Res. Rev.,* 32, 428–457.
54. Koepsell, H. *et al.* (2007) Polyspecific organic cation transporters: structure, function, physiological roles, and biopharmaceutical implications. *Pharm. Res.,* 24, 1227–1251.
55. Thul, P.J. *et al.* (2018) The human protein atlas: a spatial map of the human proteome. *Protein Sci.,* 27, 233–244.
56. Rossi, L. *et al.* (2007) Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil *in vitro*. *Pharmacol. Res.,* 56, 248–253.
57. Chu, G. *et al.* (2016) Serum level of microRNA-147 as diagnostic biomarker in human non-small cell lung cancer. *J. Drug Target.,* 24, 613–617.
58. Li, Z.Y. *et al.* (2018) The long noncoding RNA MEG3 and its target miR-147 regulate JAK/STAT pathway in advanced chronic Myeloid Leukemia. *EBioMedicine,* 34, 61–75.
59. Sui, C.J. *et al.* (2016) MicroRNA-147 suppresses human hepatocellular carcinoma proliferation migration and chemosensitivity by inhibiting HOXC6. *Am. J. Cancer Res.,* 6, 2787–2798.
60. Zhang, Y. *et al.* (2016) MicroRNA-147 suppresses proliferation, invasion and migration through the AKT/mTOR signaling pathway in breast cancer. *Oncol. Lett.,* 11, 405–410.
61. Ning, X. *et al.* (2019) Ectopic expression of miR-147 inhibits stem cell marker and epithelial-mesenchymal transition (EMT)-related protein expression in colon cancer cells. *Oncol. Res.,* 27, 399–406.
62. Shen, J. *et al.* (2018) Downregulation of MicroRNA-147 inhibits cell proliferation and increases the chemosensitivity of gastric cancer cells to 5-fluorouracil by directly targeting PTEN. *Oncol. Res.,* 26, 901–911.
63. Lee, C.G. *et al.* (2014) MicroRNA-147 induces a mesenchymal-to-epithelial transition (MET) and reverses EGFR inhibitor resistance. *PLoS One,* 9, e84597.
64. Costa, V. *et al.* (2017) MiR-675-5p supports hypoxia induced epithelial to mesenchymal transition in colon cancer cells. *Oncotarget,* 8, 24292–24302.
65. Ren, A. *et al.* (2019) Genetic variants in SLC22A3 contribute to the susceptibility to colorectal cancer. *Int. J. Cancer,* 145, 154–163.
66. Cui, R. *et al.* (2011) Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population. *Gut,* 60, 799–805.
67. Hao, Q. *et al.* (2016) Systematic meta-analyses of gene-specific genetic association studies in prostate cancer. *Oncotarget,* 7, 22271–22284.
68. Chen, H. *et al.* (2018) Genetic factors influencing prostate cancer risk in Norwegian men. *Prostate,* 78, 186–192.
69. Mohelnikova-Duchonova, B. *et al.* (2017) SLC22A3 polymorphisms do not modify pancreatic cancer risk, but may influence overall patient survival. *Sci. Rep.,* 7, 43812.
70. Zhu, L. *et al.* (2013) Genetic variant rs7758229 in 6q26-q27 is not associated with colorectal cancer risk in a Chinese population. *PLoS One,* 8, e59256.

Downloaded from https://academic.oup.com/carcin/article/42/3/378/6034128 by UCL Library Services user on 01 February 2022