CD44 SNP^{rs187115}: A novel biomarker signature that predicts survival in resectable pancreatic ductal adenocarcinoma

Giovanni Stracquadanio[1], Bart Vrugt[2], Renata Flury[3], Peter Schraml[2], Peter Würl[4], Thomas H. Müller[5], Uwe Knippschild[6], Doris Henne-Bruns[6], Stefan Breitenstein[7], Pierre-Alain Clavien[8], Rolf Graf[8], Gareth L. Bond[1], Lukasz F. Grochola[8]

Ludwig Institute for Cancer Research, University of Oxford, United Kingdom[1];
Institute for Surgical Pathology, University Hospital of Zurich, Switzerland[2];
Institute for Pathology, Cantonal Hospital of Winterthur, Switzerland[3];
Department of General and Visceral Surgery, Diakoniekrankenhaus Halle, Germany[4];
German Red Cross Blood Transfusion Service NSTOB, Springe, Germany[5];
Department of General and Visceral Surgery, Ulm University Hospital, Ulm, Germany[6];
Department of Visceral- and Thoracic Surgery, Cantonal Hospital of Winterthur, Switzerland[7];
Department of Visceral and Transplantation Surgery, University Hospital of Zurich, Switzerland[8].

Running title: CD44 SNP^{rs187115} predicts survival in PDAC

Keywords: Single nucleotide polymorphism, pancreatic ductal adenocarcinoma, SNP rs187115, survival, CD44

Financial support: This work was supported by the Department of Visceral- and Transplantation Surgery, University Hospital of Zurich, Switzerland.

Corresponding author: L.F. Grochola, Department of Visceral- and Transplantation Surgery, University Hospital of Zurich, Raemistrasse 100, 8091 Zurich, Switzerland; E-mail: lukaszfilip.grochola@usz.ch; Telephone: +41 787339411; Fax: +41 44255579

Conflict of interest: None

Translational relevance

The biomarker identified in this study, a single nucleotide polymorphism in the CD44 gene, robustly associates with over 2-fold allelic differences in risk for tumor-related death in 3 independent study cohorts that comprise a total of 348 patients with resectable pancreatic ductal adenocarcinoma. In contrast to other biomarkers currently in clinical use in this tumor type, CD44 SNP^{rs187115} can be determined before the initiation of any treatment by a simple blood test, which is readily available at the time of diagnosis. The data presented in this manuscript strongly suggest that CD44 SNP^{rs187115} could be utilized as a biomarker signature in order to identify individuals that are more likely to have a good oncologic outcome after surgical treatment and could therefore guide personalized treatment decisions. In addition, CD44 SNP^{rs187115} has the potential to affect the medical treatment of the patients with pancreatic ductal adenocarcinoma, particularly modern targeted therapies.

Abstract

Purpose: Although pancreatic ductal adenocarcinoma (PDAC) is an aggressive tumor, like other common cancers, it displays a wide range of biology. However, at present, there are no reliable tests to predict patients' cancer-specific outcomes and guide personalized treatment decisions. In this study, we aim to identify such biomarkers in resectable PDAC by studying single nucleotide polymorphisms (SNPs) in the CD44 gene, which drives the progression of pancreatic cancer.

Experimental Design: 348 PDAC patients from three independent cohorts (Switzerland, Germany, The Cancer Genome Atlas (TCGA)) who underwent pancreatic resection are included in the study. Information on the haplotype structure of the CD44 gene is obtained using 1000 Genomes Project data and the genotypes of the respective tagging SNPs are determined. Cox proportional hazards models are utilized to analyze the impact of SNP genotype on patients´ survival.

Results: We identify a SNP in the CD44 gene ($SNP^{rs187115}$) that independently associates with allelic differences in prognosis in all study cohorts. Specifically, in 121 Swiss patients, we observe an up-to 2.38-fold (p=0.020) difference in tumor-related death between the genotypes of $SNP^{rs187115}$. We validate those results in both the German (hazard ratio (HR)=2.32, p=0.044, 101 patients) and the TCGA cohort (HR=2.36, p=0.044, 126 patients).

Conclusions: CD44 $SNP^{rs187115}$ can serve as a novel biomarker readily available at the time of PDAC diagnosis that identifies patients at risk for faster tumor progression and guide personalized treatment decisions. It has the potential to significantly expand the pool of patients that would benefit from tumor resection.

Introduction

The late stage clinical presentation of pancreatic ductal adenocarcinoma (PDAC) in more than 80% of patients at the time of diagnosis, the high resistance to radiation- and chemotherapy as well as the limited effectiveness of current targeted therapeutic approaches result in an exceptionally poor overall prognosis (1-3). Surgery offers at present the only chance for a cure but carries a significant morbidity and mortality risk and results in varying oncologic outcomes with early recurrence and metastases in some individuals and long-term disease-free survival in others (4, 5). It is thought that the utilization of the knowledge of the likely oncologic course of an individual´s condition and the clinical response to any given therapy could guide a proper patient selection in a personalized fashion.

Such characterization of groups of individuals that are likely to benefit from individualized treatment decisions could be made possible by the understanding of human germline genetic variation (6, 7). Importantly, this type of genetic information about the patient is contained in the germline DNA and can be easily obtained by a simple blood or even saliva test before the initiation of any treatment and without the need for tumor biopsy. It has the potential to facilitate a personalized decision by an interdisciplinary team of oncologists, radiologists and surgeons about which treatment is most likely to benefit the patient. This type of personalized medicine is no longer a fantasy but a maturing reality. In other cancer types, such as acute lymphoblastic leukemia, colorectal, lung and breast cancer, the knowledge of an individual´s germline genetic code already affects clinical practice and the genotyping of certain alleles is

recommended by the US Food and Drug Administration (FDA) (8). However, despite emerging evidence that inherited genetic variation also affects the survival of pancreatic cancer patients and could influence the therapy of this dismal disease, no biomarker available at the time of diagnosis has been suggested for clinical use in PDAC to this date (9, 10). Therefore, in this study, we search for predictive biomarker signatures in PDAC patients that underwent tumor resection, which could help interdisciplinary teams make individualized treatment decisions.

To this end, we focus our study on single nucleotide polymorphisms (SNPs) in the CD44 gene. This transmembrane glycoprotein is involved in a vast range of cellular processes, such as regulation of growth, survival, differentiation, and motility (11). In pancreatic cancer, CD44 has been recently shown to affect the invasiveness, progression and the metastatic phenotype (12-14). The potential of germline genetic variation in the CD44 gene to affect human cancer progression, patient survival and therapy has been previously demonstrated (15). Hereby, a chemosensitivity screen has identified a SNP in the CD44 gene ($SNP^{rs187115}$) that associates with allelic differences in cellular responses to a large panel of cytotoxic chemotherapeutic agents in cells (on average 2.8-fold difference in drug sensitivity ($p=8.1\times10^{-24}$)) and survival of soft-tissue sarcoma patients (up-to 2.89 difference in tumor-related death, p=0.011, Cox multivariate analysis) (15). In this report, we build upon and expand those findings to search for clinically relevant single nucleotide polymorphisms in the CD44 gene that affect the survival of PDAC.

Materials and Methods

Swiss patient cohort. 154 patients with pancreatic cancer who were diagnosed in the years 2001-2013 at the Institutes of Pathology of the University Hospital of Zurich and Cantonal Hospital of Winterthur were included in a biobank-database (Table 1). From this database, 121 patients with pathologically confirmed PDAC were retrospectively included in the study. Approval from the local ethics committee was obtained.

German patient cohort. This prospective biobank-database comprises of 253 patients who underwent pancreatic surgery at the Department of General, Visceral and Transplantation Surgery (University of Ulm, Germany) in the years 2001 to 2007 (Table 1). From this database, 101 PDAC patients were included in the study. Approval from the local ethics committee was obtained.

The Cancer Genome Atlas (TCGA) cohort. This prospective database comprises of a total of 185 pancreatic cancer patients (years of initial pathological diagnosis 2007-2013) (Table 1). Thereof, 126 PDAC patients were included in the study. Approval from the NIH Data Access Committee was obtained.

Sample size. The sample size for the PDAC cohorts is based on, and limited by, the availability of tissue material for DNA extraction and analysis and the histopathological/clinical follow-up data of patients who underwent pancreatic resection. From all 3 PDAC cohorts (Swiss, German and TCGA), all patients with pathologically confirmed PDAC who underwent pancreatic resection and for whom tumor-free tissue or blood and a full clinical dataset was available were included in the study (Table 1). In addition, using data from our Swiss discovery cohort, we conducted a power analysis to

assess whether we can detect a 2-fold increase in risk (hazard ratio (HR)=2) for CD44 $SNP^{rs187115}$. Indeed, we estimated a power of 81.7% for a cohort of 125 patients, which is by approximation a common threshold (80%) utilized in clinical trials.

Control cohort. 498 blood donors (Germans of central European origin; 194 females and 304 males, ages 19-68 years; median 44.0 years) from whom samples were obtained at the German Red Cross Blood Transfusion Service NSTOB (Springe, Germany) were included in the study. Approval from the local ethics committee was obtained.

DNA extraction and sequence analysis. The germline allelic frequency of SNPs was determined from blood samples or normal pathologically confirmed tumor-free tissue adjacent to the resection specimen. Tissues (fresh-frozen sections or paraffin-embedded) and blood samples were stored at -80°C. Genomic DNA was extracted using the QIAamp DNA FFPE Tissue-Kit (Cat.no.56404, Qiagen, Hilden, Germany) and the Innuprep DNA/Blood DNA mini-kits (AJ Innuscreen, Berlin, Germany). The SNP genotypes (rs187115, rs353647, rs11033019, rs10768114, rs2065006, rs1547060, rs3794110, rs353630, rs353615, rs353623) were determined by PCR amplification and subsequent allelic discrimination using the C____779820_10, C___1622754_10, C__22273997_10, C___1622732_10, C__11648357_10, C___2143199_10, C__27484000_10, C____779802_10, C___1622768_10, C____779826_10 genotyping assays according to the manufacturers instructions (Applied Biosystems, Foster City, CA, USA) as described previously (15).

Haplotype analysis. For the 126 included patients, TCGA provides genotype data for matched normal tissue using the Affymetrix Genome-Wide Human SNP Array 6.0 with 906,000 probes. We performed a standard quality control step, by considering only

genotype calls with at least 95% confidence as reported by the Birdseed genotyping algorithm. We then selected SNPs with high-confidence genotype calls (SNP call-rate of at least 80%, MAF≥0.05, p-value≥1E-4 for the Hardy-Weinberg Equilibrium test). With these parameters, our final dataset consisted of forty-four SNPs in the CD44 gene.

Statistical analysis. Cox proportional hazards models were utilized to estimate the impact of SNP genotype on tumor-related death. The CD44 SNP^{rs187115} genotype was modeled as a categorical variable, with the T/T genotype being used as baseline. We also accounted for other prognostic factors that are relevant for PDAC, namely AJCC Stage (I, II, III, IV) and surgical margin status (R0, R1, R2); for these predictors, the baseline levels were set to II and R0, respectively. We then used the Wald-test to assess the statistical significance of the regression coefficients of the predictors in our model; predictors with $p<0.05$ were considered significant. Finally, for each model, we also tested that proportional hazards assumptions for Cox regression are verified, and found that all models verify these assumptions. Estimates for survival were calculated using Kaplan-Meier analysis and the log-rank test. The Likelihood ratio test was used to compare the goodness of fit of the tested models. The Fisher´s Exact Test and the Chi-Square test statistics were used for the cross-tabulation analysis. A Jonckeere-Terpstra test was utilized to assess the median differences in the age of PDAC diagnosis. Missing data were handled on a complete-case analysis basis. Values for $p<0.05$ were considered significant. All analyses were performed using SPSS 22.0 software (SPSS Inc., Chicago, IL, USA) and the Manual Survival-Package as provided in the R statistical software („A Package for Survival Analysis in S“, Terry Therneau, 2015, v2.38, {http://CRAN.R-project.org/package=survival}).

Results

CD44 SNP^{rs187115} and PDAC survival

First, we study the effects of CD44 SNP^{rs187115} (15) on the survival of patients with PDAC who underwent resection of their tumors and determine the genotype of the SNP^{rs187115} locus in 121 patients from our Swiss cohort. We start our analysis with a Kaplan-Meier survival estimate based only on the patient genotypes. At this stage of the analysis, other important clinical/histopathological factors that are known to affect the prognosis, i.e. tumor stage and completeness of surgical resection, are not included, in order to give a preliminary estimate of survival times for CD44 SNP^{rs187115}. In line with the previously reported results in the NCI60 anticancer drug screen and the clinical phenotype observed in soft-tissue sarcomas, we observe a trend whereby the C/C genotype of CD44 SNP^{rs187115} associates with the shortest estimated survival time (Table 2). Due to the expected inhomogeneity of the cohort regarding tumor stage and surgical margin status (R-status), which is inherent in most if not all clinical human datasets studying resectable PDAC, and the fact that Kaplan-Meier calculations can strongly under- or overestimate an effect of any single factor, we perform a multivariate survival analysis based on the Cox proportional hazards models. Indeed, in this analysis, which is adjusted for the known prognostic factors of PDAC, AJCC-stage and R-status and thus accounts for the heterogeneity of the cohort, the C/C genotype of CD44 SNP^{rs187115} associates with a significantly worse prognosis. Hereby, those PDAC patients with a C/C genotype have a 2.38-fold (p=0.020) increased risk for tumor-related death compared to those homozygous for the major T-allele (Figure 1A, Table 3).

Next, in order to validate these results, we include a second, independent PDAC cohort from Germany (University Hospital of Ulm) that comprises of a total of 101 patients who underwent pancreatic resection, the same cohort in which we have first described the gender-specific effects of a well-described polymorphisms in the MDM2 gene (MDM2 SNP309) on pancreatic cancer risk and prognosis (16, 17). Strikingly, in close similarity to the Swiss cohort, those patients from the German cohort homozygous for the C-allele of CD44 SNP^{rs187115} have a significantly worse prognosis compared to those T/T in genotype with a 2.32-fold (p=0.044) increased risk for tumor-related death (Cox multivariate analysis; Figure 1B and Table 3). Correspondingly, a Kaplan-Meier survival estimate demonstrates a bias towards the shortest survival time for patients with a C/C genotype (Table 2).

Subsequently, we are interested in validating the observations we have made in our German and Swiss cohorts, in a third and independent, publicly available dataset, The Cancer Genome Atlas (TCGA) pancreatic cancer database (https://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) (18). This database is supervised by the National Cancer Institute (funded by the US government) and comprises, among other tumor types, 126 PDAC patients who underwent surgical resection for their tumors and for whom normal tissue as well as a full clinical dataset is available (18). A thorough clinical and molecular characterization of the patients and their tumors has been done, including an extensive SNP-genotyping effort. Interestingly, again fully in line with all our previous observations, we note that the C/C genotype of CD44 SNP^{rs187115} correlates significantly with poorer survival in the TCGA cohort. Specifically, in those 126 PDAC patients from the TCGA database who underwent surgical resection of their primary

pancreatic tumors, patients with a C/C genotype associate with a 2.36-fold increased risk for tumor-related death compared to those with a T/T-genotype (p=0.044, Cox multivariate analysis, Figure 1C and Table 3). In line with the other cohorts, a Kaplan-Meier survival estimate demonstrates a trend whereby the shortest estimated survival time is noted in patients with a C/C genotype (Table 2). Those results were equally pronounced in those 104 Swiss, 52 German and 85 patients from the TCGA cohort who received radio-/chemotherapy, albeit possibly due to a reduced sample size not statistically significant (Supplementary Table S1).

In order to exclude that the observed effect of CD44 SNP^{rs187115} is caused by an enrichment of the C-allele in patients with advanced disease stages at the time of diagnosis or an incomplete tumor resection, we perform a cross-tabulation analysis in all three study cohorts utilizing a Fisher´s Exact test. We observe no significant enrichment of any of the genotypes in any AJCC stage or R-status of the tumors, further supporting the hypothesis that SNP^{rs187115} can serve as an independent predictor of survival (Supplementary Tables S2 and S3).

Lastly, we perform a direct comparison of the predictive power of the SNP and the two other clinical factors (R-status and AJCC stage). To do this in a way that would not be limited in statistical power, we perform a survival analysis after merging the three cohorts in which we have noted the significant association of CD44 SNP^{rs187115} with differential survival. Together, the resulting cohort consists of 348 patients with 206 tumor-related deaths. When we perform a Kaplan-Meier analysis using as predictors CD44 SNP^{rs187115}, R-status and AJCC status independently, we are able to determine that all three predictors are statistically significant (CD44 SNPrs187115 p-value: 8.95E-3; R-

status p-value: 1.34E-5; AJCC status: 2.05E-11; Table 4, Supplementary Figure S1). Analogously to the single cohort analysis, we perform a Cox regression to assess allelic differences in risk for CD44 $SNP^{rs187115}$, adjusting for R-status and AJCC stage. We find that the patients with C/C genotype have a 2.19-fold increase in risk of tumor-related death compared to patients with T/T genotype (p-value: 5.6E-4, Table 3, Figure 1D). To further assess the significance and the predictive power of CD44 $SNP^{rs187115}$, we perform a likelihood ratio test between our model, and a reduced model that considers only R-status and AJCC status (Supplementary Table S4). We observe that our model, which includes CD44 $SNP^{rs187115}$, fits significantly better than the reduced one with a p-value of 4.97E-3, suggesting that the inclusion of the genotypes at this locus together with the clinical predictors will have greater predictive power than the clinical information alone.

CD44 fine-mapping analysis

Next, we are interested whether any potential functional polymorphism in high linkage disequilibrium (LD) with $SNP^{rs187115}$ shows an even stronger association with PDAC survival. To this end, we first search for SNPs that are linked to $SNP^{rs187115}$ utilizing genotype data from European populations in 1000 Genomes Phase 1, and then retrieve their genotypes from the TCGA dataset to subsequently perform survival analyses. Of the forty-four polymorphisms that are located within the CD44 gene locus, we find $SNP^{rs187115}$ to be in high linkage disequilibrium (LD) with 4 other SNPs (rs353620, rs353615, rs353623 and rs353618) ($r^2 \geq 0.7$; Figure 2). We then perform a survival analysis independently for each SNP, assuming a dominant model for each of the

two alleles (allele A and allele B) in order to screen for SNPs that might serve as even better predictors of PDAC survival than rs187115. In this dominant model, the T-allele of $SNP^{rs187115}$ associates with a significant decrease in the risk of tumor-related death compared to the C/C genotype, which is consistent with the previous finding of C/C genotype being associated with higher risk (Cox multivariate analysis; HR: 0.441; p=0.043; Supplementary Table S5). In line with the linkage analysis, all SNPs from the haplotype tagged by rs187115 associate with allelic differences in risk for tumor-related death (Figure 2; Supplementary Table S5, up-to p=0.003, hazard ratio (HR)=0.288 (SNP rs353615)). Next, we determine the genotypes of the other SNPs from the rs187115 LD block, rs353615 and rs353623 that could add predictive value to the model, in our Swiss PDAC cohort. However, in contrast to $SNP^{rs187115}$, the slight improvement in predictive value observed in the TCGA cohort for some of those SNPs cannot be replicated in the Swiss patients (Supplementary Table S6; up-to p=0.037, HR=0.459 for SNP rs353623 as compared to HR=0.428, p=0.018 for rs187115), demonstrating that $SNP^{rs187115}$ indeed shows the strongest association with tumor-related death and the most consistent predictive value after PDAC resection within this haplotype.

Next, we expand the analysis to include SNPs within the entire genomic region of CD44 that tag other haplotypes than the rs187115 LD block ($r^2<0.7$). The remaining 39 tag SNPs cover all major haplotypes within the CD44 gene (LD block-size >3; $r^2\geq0.8$, Figure 2). Again assuming a dominant model for each of the two alleles, 11 of the 39 polymorphisms show allelic differences in the risk for tumor-related death (Figure 2; Supplementary Table S5; up-to p=0.005, HR=0.406 for SNP rs353647). Together, those 11 SNPs represent 8 different LD blocks within the CD44 gene (tagged by rs353630,

rs353647, rs1547060, rs11033019, rs2065006, rs10768114, rs353632 and rs3794110). We determine the genotypes of SNPs that tag 7 of those haplotype blocks in our Swiss cohort and repeat the survival analysis (one LD block tagged by rs353632 could not be determined due to the unavailability of an appropriate allelic discrimination assay). None of the effects of the SNPs can be validated in the Swiss cohort (Supplementary Table S6), suggesting that of all haplotypes included in the analysis, the LD block tagged by $SNP^{rs187115}$ has indeed the strongest predictive value after PDAC resection.

CD44 $SNP^{rs187115}$ and PDAC risk

Lastly, we test the hypothesis that the C-allele of CD44 $SNP^{rs187115}$, which has been previously shown to associate with age-dependent increased risk for soft-tissue sarcoma (15), also associates with an increased risk for PDAC. Hereby, no significant differences between the $SNP^{rs187115}$ genotypes are observed between 498 German blood donors and 101 German PDAC patients of the same ethnic background; no significant differences are noted for an age-dependent risk for PDAC (Supplementary Tables S7 and S8). In sum, the results do not support the hypothesis that $SNP^{rs187115}$ affects the risk for the development of PDAC.

Discussion

Taken together, these observations strongly suggest that CD44 $SNP^{rs187115}$ can serve as a predictive biomarker in patients with PDAC amenable to pancreatic resection.

Hereby, the C/C genotype of CD44 SNP^{rs187115} associates with an up-to 2.38-fold, significantly increased risk for tumor-related death in our Swiss discovery cohort. This observation is subsequently validated in two independent study cohorts, including the publicly available TCGA pancreatic cancer database supervised by the NCI.

The genotype of CD44 SNP^{rs187115} can be determined before the initiation of any treatment by a simple blood test, which is readily available at the time of diagnosis and utilized to identify patients who are at higher risk for faster tumor progression. In contrast to CD44 SNP^{rs187115}, other important prognostic factors currently in clinical use, such as completeness of resection and lymph node involvement, or those suggested to be useful in a clinical setting, such as postoperative CA19-9 serum levels (19), can only be accurately determined after surgical resection, and have therefore limited impact on therapeutic decisions.

Our data strongly suggest that patients with a C/C genotype at the CD44 SNP^{rs187115} locus with PDAC are less likely to have a good oncologic outcome after surgical treatment even in a resectable stage, and therefore should be considered for alternative systemic therapy or chemo-/radiotherapy protocols. Indeed, future adjuvant/neoadjuvant chemotherapy trials in resectable PDAC should stratify patients by the CD44 SNP^{rs187115} genotype and/or possibly exclude patients homozygous for the C-allele.

Importantly, CD44 SNP^{rs187115} also has a strong potential to serve as a biomarker signature that identifies patients with borderline resectable/locally advanced tumors, which could truly profit from established, more radical surgical therapies. These therapies aim to fully remove the tumors and therefore offer the only potential for cure in cases of

pancreatic cancer with a relatively favorable prognosis despite their locally advanced presentation. This is based on the fact that individuals who receive a resection of their tumors have an average life expectancy of approximately 23 months as compared to 11 months for those with locally advanced, non-metastatic disease who may be considered for potentially curative tumor resection and instead undergo palliative treatment (4, 20, 21). In fact, 50% of the tumors are classified as localized, and a large majority of these non-metastatic tumors (35% of all PDAC patients) are diagnosed in a borderline resectable or locally advanced stage. Despite technical resectability in many of those cases, they are only rarely considered for surgical resection. Therefore, curative treatment is currently offered only to the remaining 15% of the non-metastatic patients (2, 4). Such a restrictive surgical strategy is mainly caused by the observation that extended types of resection, which involve major visceral vessel resection and reconstruction, can lead to an increased perioperative morbidity and mortality (5, 22-24). However, due to the fact that extended surgical procedures offer the only chance for a cure and are technically feasible, they have been strongly advocated in the literature (5, 22-24). It is thought that these procedures are particularly useful in a subset of patients that are expected to have a relatively good long-term prognosis, such as individuals who carry the T-allele of CD44 $SNP^{rs187115}$, and who might truly benefit from a potentially curative therapy, as the expected long-term benefit would outweigh the increased perioperative risk. It will be important to study whether the findings of this analysis, which is limited to patients without locally advanced disease, can be repeated in a clinical trial in individuals with locally advanced/borderline resectable PDAC who undergo resection in curative intent.

In contrast to our previous study performed in the NCI60 cell line panel (which

does not include pancreatic cancer cells) and soft-tissue sarcoma patients (15), we observe that the association of CD44 SNP^{rs187115} and survival in PDAC is equally pronounced in patients who received DNA-damaging treatments and those who only underwent resection of their tumors. However, this observation is possibly due to differences in the biology of the tumors and their susceptibility to chemo- and/or radiotherapy. Hereby, in contrast to soft-tissue sarcomas, which are known to respond well to DNA-damaging treatments (25-28), the effect of such therapies is only very limited in resectable PDAC and most clinical studies have shown no or only a weak response to the agents utilized in the treatment of the patients included in our study (29-32).

In addition, besides the above-described impact on surgical strategy, CD44 SNP also has the potential to affect the medical treatment of the patients. This is based on the recent observation that CD44 can be successfully directly targeted in pancreatic adenocarcinoma with monoclonal antibodies to treat pancreatic xenograft tumors in mice (33, 34). Based on those findings, clinical trials in humans utilizing this approach are already under way (33, 34). Moreover, it has been shown that CD44 interacts with the EGFR receptor, which is already an FDA-approved therapeutic target in PDAC, to promote cell motility and invasiveness (35, 36). Therefore, it is tempting to speculate that the identified SNP in the CD44 gene could affect the response to current and future targeted therapies, such as direct anti-CD44 targeting or EGFR inhibition. It will be interesting to test this hypothesis in future clinical trials that utilize such novel therapeutic approaches. In addition, it will be important to elucidate the precise molecular mechanisms by which this SNP leads to the observed clinical phenotypes and to test its

potential to affect cellular signaling pathways that can be directly targeted by modern therapeutics.

Acknowledgements

We thank Mrs. Theresia Graf Reding and Mrs. Elisabeth Bond for their extraordinary support in the lab. We would also like to thank Dr. K. Slankamenac, Dr. A. Thürlimann and Dr. E. Schadde for their valuable help with the data collection for the Swiss patient cohort and Nadia Braga Grochola for her important assistance with the data management.

References

1. Lennon AM, Wolfgang CL, Canto MI, Klein AP, Herman JM, Goggins M, et al. The early detection of pancreatic cancer: what will it take to diagnose and treat curable pancreatic neoplasia? Cancer research. 2014;74:3381-9.

2. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. The New England journal of medicine. 2014;371:1039-49.

3. Moore MJ, Goldstein D, Hamm J, Figer A, Hecht JR, Gallinger S, et al. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:1960-6.

4. Heestand GM, Murphy JD, Lowy AM. Approach to patients with pancreatic cancer without detectable metastases. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33:1770-8.

5. Yekebas EF, Bogoevski D, Cataldegirmen G, Kunze C, Marx A, Vashist YK, et al. En bloc vascular resection for locally advanced pancreatic malignancies infiltrating major blood vessels: perioperative outcome and long-term survival in 136 patients. Annals of surgery. 2008;247:300-9.

6. Vazquez A, Bond EE, Levine AJ, Bond GL. The genetics of the p53 pathway, apoptosis and cancer therapy. Nature reviews Drug discovery. 2008;7:979-87.

7. Grochola LF, Zeron-Medina J, Meriaux S, Bond GL. Single-nucleotide polymorphisms in the p53 signaling pathway. Cold Spring Harbor perspectives in biology. 2010;2:a001032.

8. Wheeler HE, Maitland ML, Dolan ME, Cox NJ, Ratain MJ. Cancer pharmacogenomics: strategies and challenges. Nature reviews Genetics. 2013;14:23-34.

9. Innocenti F, Owzar K, Cox NL, Evans P, Kubo M, Zembutsu H, et al. A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in CALGB 80303. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18:577-84.

10. Wu C, Kraft P, Stolzenberg-Solomon R, Steplowski E, Brotzman M, Xu M, et al. Genome-wide association study of survival in patients with pancreatic adenocarcinoma. Gut. 2014;63:152-60.

11. Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol. 2003;4:33-45.

12. Zoller M. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? Nature reviews Cancer. 2011;11:254-67.

13. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, et al. Identification of pancreatic cancer stem cells. Cancer research. 2007;67:1030-7.

14. Jiang W, Zhang Y, Kane KT, Collins MA, Simeone DM, di Magliano MP, et al. CD44 regulates pancreatic cancer invasion through MT1-MMP. Mol Cancer Res. 2015;13:9-15.

15. Vazquez A, Grochola LF, Bond EE, Levine AJ, Taubert H, Muller TH, et al. Chemosensitivity profiles identify polymorphisms in the p53 network genes 14-3-3tau and CD44 that affect sarcoma incidence and survival. Cancer research. 2010;70:172-80.

16. Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell. 2004;119:591-602.

17. Grochola LF, Muller TH, Bond GL, Taubert H, Udelnow A, Wurl P. MDM2 SNP309 associates with accelerated pancreatic adenocarcinoma formation. Pancreas. 2010;39:76-80.

18. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455:1061-8.

19. Berger AC, Garcia M, Jr., Hoffman JP, Regine WF, Abrams RA, Safran H, et al. Postresection CA 19-9 predicts overall survival in patients with pancreatic cancer treated with adjuvant chemoradiation: a prospective validation by RTOG 9704. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26:5918-22.

20. Huguet F, Andre T, Hammel P, Artru P, Balosso J, Selle F, et al. Impact of chemoradiotherapy after disease control with chemotherapy in locally advanced pancreatic adenocarcinoma in GERCOR phase II and III studies. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:326-31.

21. Neoptolemos JP, Stocken DD, Bassi C, Ghaneh P, Cunningham D, Goldstein D, et al. Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial. Jama. 2010;304:1073-81.

22. Mollberg N, Rahbari NN, Koch M, Hartwig W, Hoeger Y, Buchler MW, et al. Arterial resection during pancreatectomy for pancreatic cancer: a systematic review and meta-analysis. Annals of surgery. 2011;254:882-93.

23. Fuhrman GM, Leach SD, Staley CA, Cusack JC, Charnsangavej C, Cleary KR, et al. Rationale for en bloc vein resection in the treatment of pancreatic adenocarcinoma adherent to the superior mesenteric-portal vein confluence. Pancreatic Tumor Study Group. Annals of surgery. 1996;223:154-62.

24. Martin RC, 2nd, Scoggins CR, Egnatashvili V, Staley CA, McMasters KM, Kooby DA. Arterial and venous resection for pancreatic adenocarcinoma: operative and long-term outcomes. Archives of surgery. 2009;144:154-9.

25. Pisters PW, Harrison LB, Leung DH, Woodruff JM, Casper ES, Brennan MF. Long-term results of a prospective randomized trial of adjuvant brachytherapy in soft tissue sarcoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1996;14:859-68.

26. Yang JC, Chang AE, Baker AR, Sindelar WF, Danforth DN, Topalian SL, et al. Randomized prospective study of the benefit of adjuvant radiation therapy in the treatment of soft tissue sarcomas of the extremity. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1998;16:197-203.

27. Kraybill WG, Harris J, Spiro IJ, Ettinger DS, DeLaney TF, Blum RH, et al. Long-term results of a phase 2 study of neoadjuvant chemotherapy and radiotherapy in the management of high-risk, high-grade, soft tissue sarcomas of the extremities and body wall: Radiation Therapy Oncology Group Trial 9514. Cancer. 2010;116:4613-21.

28. Kraybill WG, Harris J, Spiro IJ, Ettinger DS, DeLaney TF, Blum RH, et al. Phase II study of neoadjuvant chemotherapy and radiation therapy in the management of high-risk, high-grade, soft tissue sarcomas of the extremities and body wall: Radiation Therapy Oncology Group Trial 9514. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006;24:619-25.

29. Kosuge T, Kiuchi T, Mukai K, Kakizoe T, Japanese Study Group of Adjuvant Therapy for Pancreatic C. A multicenter randomized controlled trial to evaluate the effect of adjuvant cisplatin and 5-fluorouracil therapy after curative resection in cases of pancreatic cancer. Jpn J Clin Oncol. 2006;36:159-65.

30. Neoptolemos JP, Stocken DD, Friess H, Bassi C, Dunn JA, Hickey H, et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. The New England journal of medicine. 2004;350:1200-10.

31. Oettle H, Neuhaus P, Hochhaus A, Hartmann JT, Gellert K, Ridwelski K, et al. Adjuvant chemotherapy with gemcitabine and long-term outcomes among patients with resected pancreatic cancer: the CONKO-001 randomized trial. Jama. 2013;310:1473-81.

32. Smeenk HG, van Eijck CH, Hop WC, Erdmann J, Tran KC, Debois M, et al. Long-term survival and metastatic pattern of pancreatic and periampullary cancer after adjuvant chemoradiation or observation: long-term results of EORTC trial 40891. Annals of surgery. 2007;246:734-40.

33. Wood NJ. Pancreatic cancer: pancreatic tumor formation and recurrence after radiotherapy are blocked by targeting CD44. Nature reviews Gastroenterology & hepatology. 2014;11:73.

34. Li L, Hao X, Qin J, Tang W, He F, Smith A, et al. Antibody against CD44s inhibits pancreatic tumor initiation and postradiation recurrence in mice. Gastroenterology. 2014;146:1108-18.

35. Kim Y, Lee YS, Choe J, Lee H, Kim YM, Jeoung D. CD44-epidermal growth factor receptor interaction mediates hyaluronic acid-promoted cell motility by activating protein kinase C signaling involving Akt, Rac1, Phox, reactive oxygen species, focal adhesion kinase, and MMP-2. The Journal of biological chemistry. 2008;283:22513-28.

36. Grass GD, Tolliver LB, Bratoeva M, Toole BP. CD147, CD44, and the epidermal growth factor receptor (EGFR) signaling pathway cooperate to regulate breast epithelial cell invasiveness. The Journal of biological chemistry. 2013;288:26089-104.

Figure legends

Figure 1. Cox multivariate analysis of tumor-specific survival after resection of pancreatic ductal adenocarcinoma (PDAC). The C/C genotype of the CD44 SNP, which has been previously shown to associate with a weaker apoptotic response to chemotherapeutic treatment in the NCI60 cell lines and with a poorer survival in soft-tissue sarcoma patients (15) also correlates significantly with poorer survival in PDAC patients. **(A**) The Figure shows the results of the Cox analysis in 121 PDAC patients from Switzerland who underwent a surgical resection of their primary pancreatic tumors. Patients with a C/C genotype associated with the shortest survival time compared to those with a T/T-genotype (HR=2.38 for tumor-related death). **(B**) Results of the Cox analysis in 101 PDAC patients from Germany who underwent a surgical resection of their tumors at the University Hospital of Ulm. Patients with a C/C genotype had the shortest survival time compared to those with a T/T-genotype (HR=2.32). **(C)** The graph displays the survival curves of 126 PDAC patients from the publicly available TCGA database who underwent resection of their tumors. Patients with a C/C genotype associate with a 2.36-fold increased risk for tumor-related death compared to those T/T in genotype (Cox multivariate regression analysis, adjusted for the known prognostic factors of PDAC AJCC stage and resection margin). **(D)** Survival analysis in the merged dataset (all 3 PDAC cohorts: Swiss, German, TCGA) with a total of 348 patients.

Figure 2. CD44 fine-mapping: Cox survival analysis of CD44 SNPs in the TCGA PDAC cohort. For each SNP, we plot the best p-value of the Cox regression on the dominant models for the A and B alleles; the black dashed line is the p-value ≤ 0.05 significance threshold. Each SNP is plotted at its genomic locus, according with the hg19 human genome assembly. The rs187115 LD Block ($r2 \geq 0.7$, 5 SNPs) is the one showing the most significant associations with differential survival. All of the SNPs from the rs187115 LD Block are located within intron 1 of the CD44 gene.

Tables

Table 1. Clinical and histopathological data.

	Swiss cohort	German cohort	TCGA cohort	Merged (all 3 cohorts)
Median age*	66 yrs (19-87)	65 yrs (32-82)	65.5 yrs (35-85)	65 yrs (19-87)
Median observation time*	20.2 mo (1-89)	12.0 mo (1-85)	12.4 mo (0-66)	13 mo (0-89)
Frequency of cases assessed for inclusion	154	253	185	592
Frequency of included cases (%)				
Total	121	101	126	348
Females	63 (52.1)	40 (39.6)	56 (44.4)	159 (45.7)
Males	58 (47.9)	61 (60.4)	70 (55.6)	189 (54.3)
Frequency of excluded cases (%)				
Total excluded	33	152	59	244
Missing follow-up	2 (6.1)	0 (0)	5 (8.5)	7 (2.9)
No normal tissue and/or blood	2 (6.1)	14 (9.2)	18 (30.5)	34 (13.9)
Missing histopathologic data	0 (0)	1 (0.7)	9 (15.2)	10 (4.1)
Histology other than PDAC	0 (0)	137 (90.1)	27 (45.8)	164 (67.2)
No pancreatic resection	29 (87.9)	0 (0)	0 (0)	29 (11.9)
Status at last follow up (%)				
Died of tumour-related causes	71 (58.7)	77 (76.2)	59 (46.8)	207 (59.5)
Alive	36 (29.8)	20 (19.8)	58 (46.0)	114 (32.8)
Died of non tumour-related causes	4 (3.2)	0 (0)	7 (5.6)	11 (3.2)
In-hospital death	10 (8.3)	4 (4.0)	2 (1.6)	16 (4.6)
Chemo- and/or radiotherapy (%)**	104 (86.0)	52 (51.5)	85 (67.5)	241 (69.3)
Stage (%)				
I	4 (3.3)	7 (6.9)	10 (7.9)	21 (6.0)
II	103 (85.1)	82 (81.2)	108 (85.7)	293 (84.2)
III	6 (5.0)	4 (4.0)	4 (3.2)	14 (4.0)
IV	8 (6.6)	8 (7.9)	4 (3.2)	20 (5.7)
Tumour resection (%)***				
Radical (R0)	66 (54.5)	73 (72.3)	82 (65.1)	221 (63.7)
Not radical (R1)	48 (39.7)	17 (16.8)	42 (33.3)	107 (30.8)
Not radical (R2)	7 (5.8)	10 (9.9)	2 (1.6)	19 (5.5)

yrs=years; mo=months

* Included cases only; the range is given in brackets

** The chemotherapeutic agents included Gemcitabine, Oxaliplatin, Capecitabine,Irinotecan, 5-FU, Leucovirin (Swiss cohort), Gemicitabine (German cohort) and Gemcitabine, Oxaliplatin, Capecitabine, Irinotecan, 5-FU, Paclitaxel, Doxorubicin, Cyclophosphamide, Leucovirin (TCGA)

*** The R-Status of one patient (German cohort) could not be determined

Table 2. Medians for survival time based on Kaplan Meier estimates (not adjusted to AJCC Stage and R-Status)

	CD44 rs187115	n	Number of tumour-related deaths	Medians				p-value (log-rank)
				Estimate (months)	Std. Error	95% Confidence Interval		
						Lower Bound	Upper Bound	
Swiss	T/T	67	39	26.5	6.2	14.3	38.7	0.124
	T/C	39	22	20.0	1.9	16.3	23.7	
	C/C	15	10	17.0	3.8	9.5	24.5	
German	T/T	43	28	15.0	1.4	12.2	17.8	0.287
	T/C	47	39	14.0	1.8	10.5	17.5	
	C/C	11	9	12.0	NA	NA	NA	
TCGA	T/T	58	24	20.0	1.4	17.3	22.6	0.325
	T/C	54	27	20.2	1.8	16.7	23.6	
	C/C	14	8	15.8	5.9	4.2	27.4	

Table 3. Cox multivariate regression analysis.

Cohort	Adjustment variable	n	Hazard Ratio	95% Confidence Interval	P-value
Swiss	CD44 SNP rs187115				
	T/T	67	1	-	-
	C/T	39	1.043	0.604 - 1.802	>0.1
	C/C	15	2.380	1.143 - 4.940	0.020
	R-status				
	R0	66	1	-	-
	R1	48	1.378	0.820 - 2.320	>0.1
	R2	7	13.462	3.186 - 56.972	0.000
	AJCC Stage				
	I	4	0.571	0.079 - 4.146	>0.1
	II	103	1	-	-
	III	6	0.401	0.077 - 2.078	>0.1
	IV	8	0.976	0.331 - 2.870	>0.1
German	CD44 SNP rs187115				
	T/T	43	1	-	-
	C/T	47	1.516	0.891 - 2.581	>0.1
	C/C	11	2.320	1.022 - 5.270	0.044
	R-status				
	R0	73	1	-	-
	R1	17	1.106	0.572 - 2.138	>0.1
	R2	10	1.724	0.374 - 7.941	>0.1
	AJCC Stage				
	I	7	0.626	0.181 - 2.159	>0.1
	II	82	1	-	-
	III	4	6.040	1.786 - 20.429	0.004
	IV	8	4.004	0.777 - 20.630	0.097
TCGA	CD44 SNP rs187115				
	T/T	58	1	-	-
	C/T	54	1.094	0.608 - 1.970	>0.1
	C/C	14	2.361	1.023 - 5.449	0.044
	R-status				
	R0	82	1	-	-
	R1	42	2.025	1.124 - 3.647	0.019
	R2	2	NA	NA	>0.1
	AJCC Stage				
	I	10	0.580	0.171 - 1.972	>0.1
	II	108	1	-	-
	III	4	2.315	0.536 - 10.003	>0.1
	IV	4	1.506	0.458 - 4.953	>0.1
Merged (all 3 cohorts)	CD44 SNP rs187115				
	T/T	168	1	-	-
	C/T	140	1.213	0.902 - 1.631	>0.1
	C/C	40	2.188	1.402 - 3.415	0.001
	R-status				
	R0	221	1	-	-
	R1	107	1.327	0.969 - 1.818	0.078
	R2	19	4.056	2.001 - 8.220	0.000
	AJCC Stage				
	I	21	0.610	0.284 - 1.308	>0.1
	II	293	1	-	-
	III	14	1.501	0.689 - 3.269	>0.1
	IV	20	1.536	0.858 - 2.751	>0.1

Table 4. Survival time analysis in the merged database including all 3 PDAC cohorts based on Kaplan Meier estimates

	Group	n	Number of tumour-related deaths	Medians				p-value (log-rank)
				Estimate (months)	Std. Error	95% Confidence Interval		
						Lower Bound	Upper Bound	
CD44 rs187115	T/T	168	91	20.6	1.8	17.1	24.2	0.009
	T/C	140	88	18.0	1.6	14.8	21.2	
	C/C	40	27	15.8	2.6	10.7	20.8	
AJCC	1	21	7	NA	NA	NA	NA	0.000
	2	293	172	20.0	1.5	17.1	22.8	
	3	14	10	13.0	0.8	11.4	14.5	
	4	20	17	12.0	4.7	2.9	21.1	
R-status	0	221	126	20.2	1.7	16.9	23.4	0.000
	1	107	64	18.0	1.7	14.7	21.3	
	2	19	16	5.0	1.0	3.0	7.0	

Figure 1.

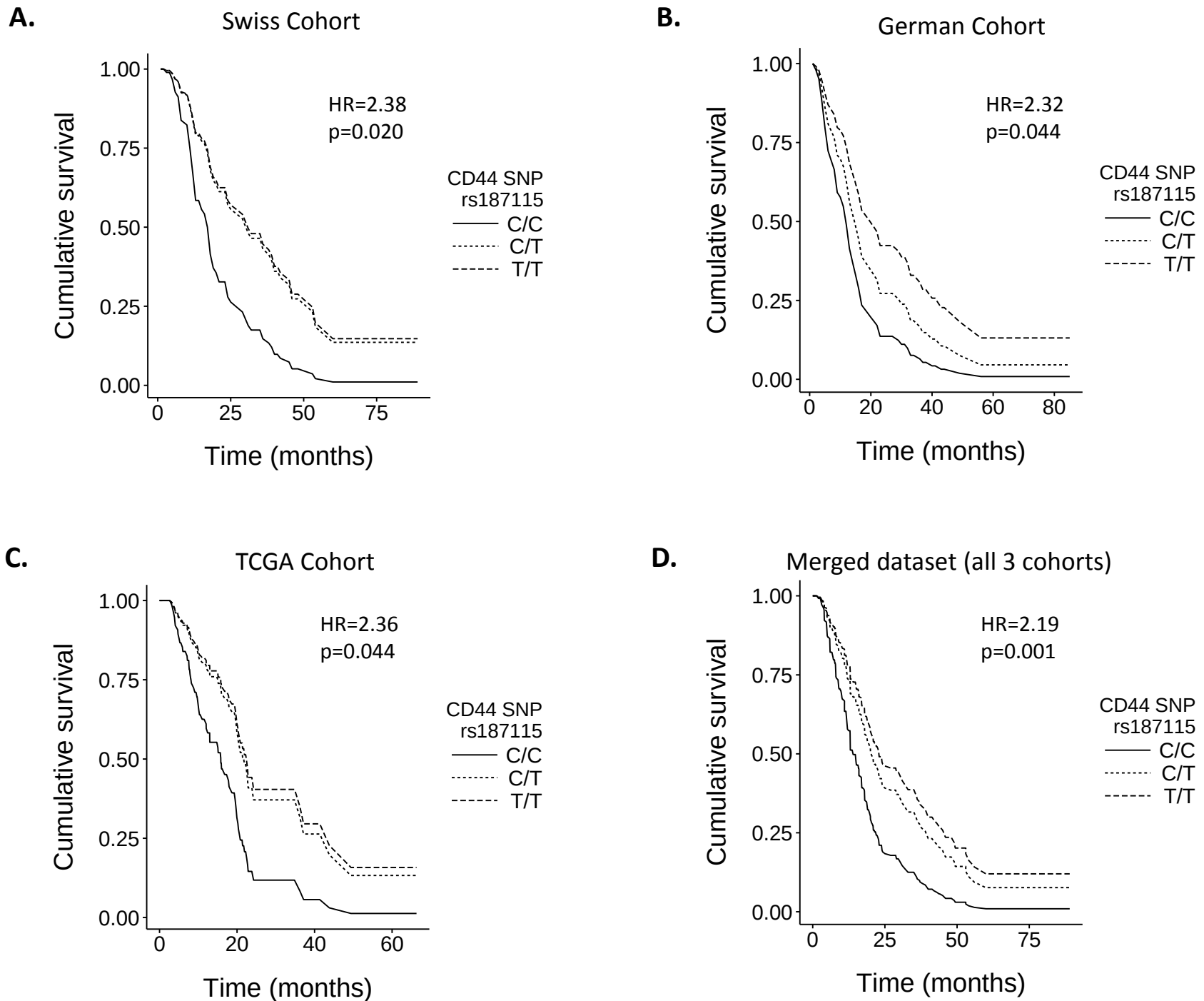
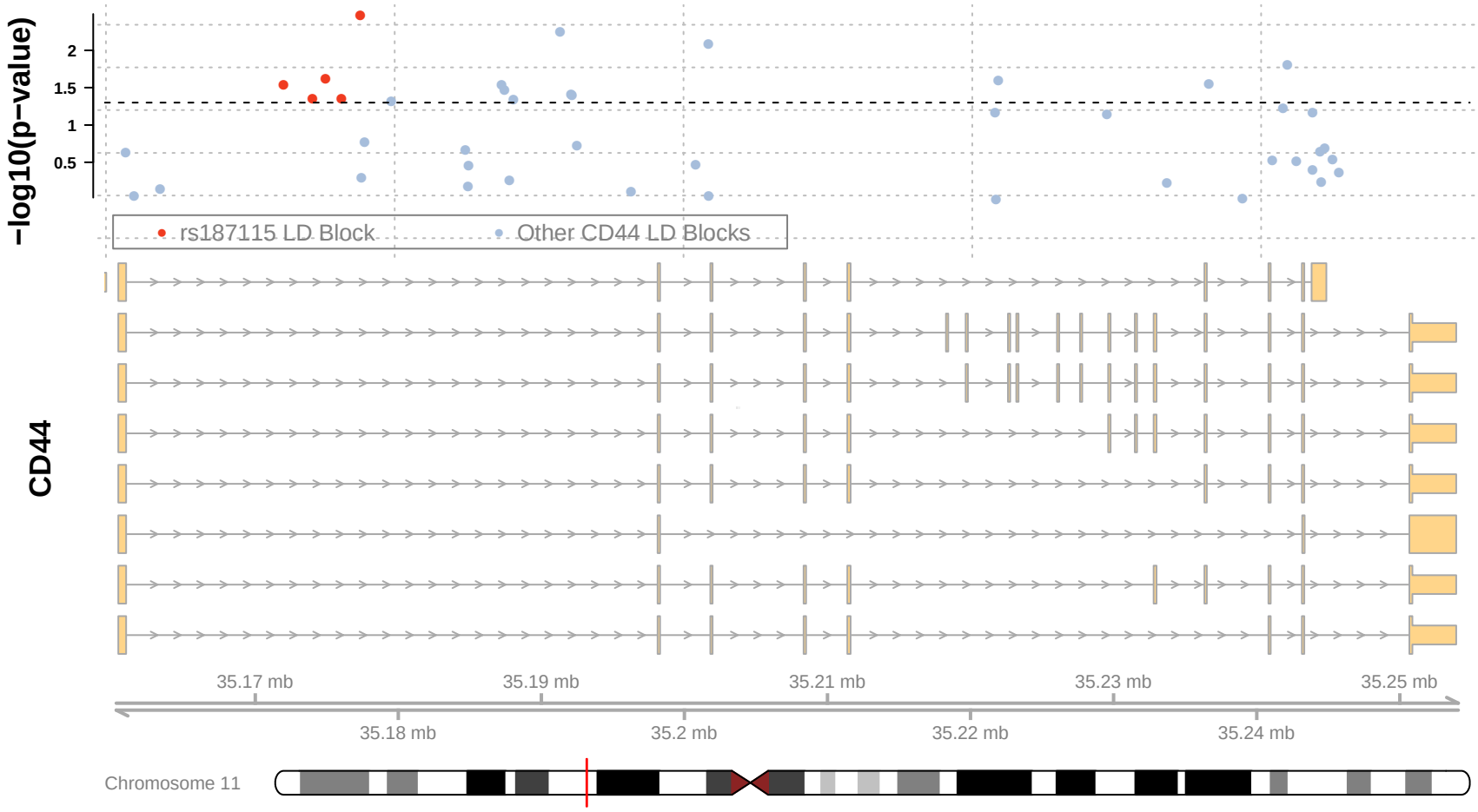

A.
Swiss Cohort
HR=2.38
p=0.020
Cumulative survival
Time (months)
CD44 SNP
rs187115
C/C
C/T
T/T
B.
German Cohort
HR=2.32
p=0.044
C.
TCGA Cohort
HR=2.36
p=0.044
D.
Merged dataset (all 3 cohorts)
HR=2.19
p=0.001

Figure 2.


−log10(p−value)
2
1.5
1
0.5
rs187115 LD Block
Other CD44 LD Blocks
CD44
35.17 mb
35.19 mb
35.21 mb
35.23 mb
35.25 mb
35.18 mb
35.2 mb
35.22 mb
35.24 mb
Chromosome 11