Published in final edited form as:

Lancet Oncol. 2013 January ; 14(1): 81–87. doi:10.1016/S1470-2045(12)70517-X.

Prospective Analysis of Genetic Polymorphisms and Risk of Recurrence in Renal Cell Cancer

Fabio A. B. Schutz, MD¹, Mark M. Pomerantz, MD^{1,2,*}, Kathryn P. Gray, PhD^{5,*}, Prof Michael B. Atkins, MD^{3,4}, Jonathan E. Rosenberg, MD^{1,2}, Michelle S. Hirsch, MD/PhD², David F. McDermott, MD^{3,4}, Megan E. Lampron, BS⁴, Gwo-Shu Mary Lee, MD/PhD¹, Sabina Signoretti, MD^{2,4}, Prof Philip W. Kantoff, MD^{1,2}, Matthew L. Freedman, MD/PhD¹, and Toni K. Choueiri, MD^{1,2,4}

¹Dana-Farber Cancer Institute and Harvard Medical School, Boston MA

²Brigham and Women's Hospital and Harvard Medical School, Boston, MA

³Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA

⁴Kidney Cancer Program Dana-Farber/Harvard Cancer Center, Boston, MA

⁵Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, MA

Summary

Background—Germline genetic polymorphisms may affect the risk of recurrence in patients with localized renal cell carcinoma (RCC). Our aim was to investigate the association of genetic polymorphisms with RCC recurrence.

Patients and Methods—We analyzed germline DNA samples extracted from 554 (discovery cohort of 403 and an independent validation cohort of 151) patients with localized RCC treated at Dana-Farber/Harvard Cancer Center (DF/HCC) and of European-American ancestry (Caucasians). The discovery cohort was selected from a prospective database at Dana-Farber/Harvard Cancer Center and the validation cohort was identified from the Brigham and Women's Hospital surgery and pathology department records. Single nucleotide polymorphisms (SNPs) residing in 70 genes involved in RCC pathogenesis including the VHL/HIF/VEGF, PI3K/AKT/mTOR pathways, and genes involved in immune regulation and metabolism were genotyped for the discovery cohort (total 285 SNPs successfully genotyped and assessable for analysis). The analyses of genotype associations with recurrence free survival (RFS) were assessed using Cox proportional hazards model, Kaplan-Meier method and logrank test. False discovery rate (FDR) q-value was used to

Corresponding Author: Toni K. Choueiri, MD, Dana-Farber Cancer Institute, 450 Brookline Ave. (Dana 1230), Boston MA 02115, Phone: 617-632-5456; Fax: 617-632-2165, Toni_Choueiri@DFCI.harvard.edu.

*Both authors contributed equally

Contributions of authors:

Concept and design: TKC

Search and collection of data: FABS, TKC, MEL, MMP, MLF

Analysis of data and interpretation: TKC, FABS, KPG, MBA, JER, MSH, DFM, MGL, MEL, SS, PWK, MMP and MLF

Tables and Figures: TKC, FABS, KPG

Writing the manuscript and review: TKC, FABS, KPG, MBA, JER, MSH, DFM, MGL, MEL, SS, PWK, MMP and MLF

Potential Conflicts of Interest:

TKC: Research funding: Pfizer. Advisory board: Novartis, Pfizer, Aveo, GSK.

MBA: Advisory Board: Genentech, BMS, Novartis, Prometheus, Amgen, Merck, Aveo

JER: Advisory Board/Consultant: Genentech, Abbott, Novartis. Speakers Bureau: Novartis. Research

Funding: Glaxo-Smithkline, Sanofi-Aventis, Novartis.

FABS, KPG, MSH, DFM, MGL, SS, PWK, MMP and MLF: declared no conflicts of interest

adjust for multiple comparisons in selecting potential SNPs with RFS association. The finding from the discovery cohort was validated in an external independent cohort.

Findings—We report the significant association between genotype variants of SNP rs11762213 (c.144G>A; p.Ala48Ala, located in exon two c-MET) and primary analysis endpoint of RFS using both univariate and multivariable analysis. Specifically, patients carrying one or two copies of the minor (risk) allele had an increased risk of recurrence or death (hazard ratio (HR) =1.86, 95% confidence interval (CI), 1.17,2.95; p=0.0084) in the multivariate analysis adjusted for clinical and pathological factors. The median RFS for carriers of the risk allele was 19 months (95%CI: 9,*) compared to 50 months (95%CI: 37,75) for homozygotes of the non-risk allele. The significant association was validated using data from the validation cohort with a HR of 2.45 (95%CI: 1.01,5.95; p=0.048), although of borderline significance. The rs11762213 results in a synonymous aminoacid change in cMET gene. * *unable to estimate due to small sample*.

Interpretation—Patients with localized RCC and c-MET polymorphism (rs11762213) may have an increased risk of recurrence after nephrectomy. If these results are further validated, it may be incorporated in future prognostic tools, potentially aiding in the design of adjuvant clinical trials with c-MET inhibitors, and clinical management.

Funding—This project is funded by the Conquer Cancer Foundation and ASCO under a Career Development Award (CDA) for Dr. Choueiri, The Trust Family Research for Kidney cancer for Dr. Choueiri and the NIH/NCI Kidney cancer SPORC.

Keywords

localized renal cell cancer; nephrectomy; recurrence free interval; genetic polymorphisms; single nucleotide polymorphisms; MET; VEGF

Introduction

A large number of patients with localized RCC treated with a curative intent will eventually experience disease recurrence; usually leading to incurable disease.¹ The risk of recurrence is highly associated with clinical and pathological factors, such as TNM stage, performance status and tumor Fuhrman grade. Still, there is significant heterogeneity in outcomes among patients with similar clinical and pathological features. Improved predictors of RCC recurrence are needed.^{2,3}

While several molecular markers for disease progression have been proposed, no biomarkers have been well established to assess recurrence risk. Germline DNA polymorphisms are particularly attractive biomarkers since they are present at the time of diagnosis and are not influenced by the state of the disease or the timing of diagnosis. Single nucleotide polymorphisms (SNPs) are inherited germline DNA sequence variants. These variants occur throughout the entire genome, in both coding and non-coding regions, and they may produce modifications in biologic pathways.⁴

Recently a GWAS for RCC identified SNP in HIF2a associated with high risk for renal cell carcinoma and a complex genetic architecture for association with risk for RCC.^{5–8} A follow-up study showed that a variant in 11q13.3 remodulates the binding and function of hypoxia-inducible factor (HIF) at a previously unrecognized transcriptional enhancer of CCND1. The protective haplotype impairs binding of HIF-2, resulting in an allelic imbalance in cyclin D1 expression.⁹ However, to our knowledge there are no large studies evaluating SNPs and RCC recurrence and survival after resection. Positive associations between particular germline polymorphisms and outcome in prostate cancer, breast cancer, lymphoid neoplasm and nasopharyngeal cancer (after initial treatment) have been

described.^{10–13} Therefore, we sought to evaluate the association of SNPs in genes implicated in RCC biology with the likelihood of RCC recurrence after localized therapy.

Patients and Methods

Discovery cohort

Patients were initially identified from the Dana-Farber/Harvard Cancer Center (DF/HCC) Kidney Cancer Program Clinical Research Information System (CRIS) database. This database, initiated in 2001, is designed to prospectively follow patients with a diagnosis of localized RCC submitted to a curative intent procedure. The first patient in our analytic cohort enrolled on January 18th, 2002 and last patient enrolled on May 6th, 2010. Patients provide an informed consent for baseline and prospective clinical and pathological data collection, as well as for collection of blood samples for DNA extraction for research purposes. The study was approved by the Dana-Farber/Harvard Cancer Center (DF/HCC) institutional review board. All included patients had histologically proven RCC, and were reviewed by genitourinary (GU) pathologists at DF/HCC. Patients' baseline clinical and pathological data, including age, gender, performance status, nephrectomy date, TNM stage, histology and tumor Fuhrman grade were available so were the disease outcome information, including date of recurrence, date of death or date of last follow-up. Date of recurrence was considered the date of first diagnosis of local or distant recurrence, or diagnosis of contra-lateral renal cell cancer by the treating physician recorded in patients charts. Patients were followed prospectively at a routine follow-up schedule devised by the treating physician. Patients included in the analysis were of European-American ancestry (Caucasians) to ensure homogeneity of allelic frequencies. The ancestry information was gathered from patients' charts.

Validation cohort

An independent cohort of 151 RCC patients with localized disease was identified from the Brigham and Women's Hospital pathology department records. Patients' clinical information was ascertained by medical records. The enrollment time for this cohort was from 28 August, 1991 to 10 June, 2010. With the help of 2 GU pathologists (MSH and SS), DNA was extracted from non-cancerous renal parenchyma tissue present in the paraffin blocks from each patient. This study was approved by the independent institutional review board (IRB) under a different protocol.

Genes and SNP selection

Several signaling pathways and key processes known to play an important role in the pathogenesis and progression of RCC were chosen for analysis in the discovery cohort. A total of 70 genes were selected, including genes from the VHL/HIF/VEGF, PI3K/AKT/mTOR pathways, and genes involved in immune regulation and metabolism/detoxification processes.^{14–19}

SNPs were selected across selected genes, including regions 50kb upstream and 50kb downstream of each gene with the goal of capturing common genetic variation. The European-American ancestry population (CEU) of the HapMap database (release 21) was used for the selection of the SNPs. A minimum allele frequency (MAF) of at least 5% and a minimum pairwise correlation index (r^2) of 80% were required for each SNP.²⁰ Additionally, SNPs from previously published studies that were found to be significantly associated with risk of developing RCC were also selected for genotyping. All non-synonymous coding SNPs in the 70 selected genes were included for analysis. In the validation cohort, a limited number of SNPs were selected, including the top 10 SNPs associated with RFS in the testing cohort with a p-value<0.05.

DNA extraction and genotyping

The DNA samples from the discovery cohort were extracted from peripheral blood mononuclear cell (PBMC) by using QIAamp DNABlood mini kit (QIAGEN Inc., Valencia, CA), while the DNA from validation cohort was extracted from the normal kidney parenchyma tissue in the formalin fixed paraffin embedded (FFPE) blocks using DNeasy 96 Blood & Tissue Kit (QIAGEN Inc., Valencia, CA). Genotyping was performed with staff blinded to patient status using the Sequenom iPLEX Gold platform with matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry. All SNP assays were combined into 12 multiplex pools design, and all reactions were carried out in 384-well format. Approximately 5% of duplicate samples were randomly selected and interspersed among plates for quality control. Concordance rate for duplicate genotyping was 100%.

Analysis was restricted to SNPs passing quality filters, excluding SNPs with a genotyping success rate <85% or with a significant deviation from Hardy-Weinberg equilibrium (HWE) to ensure adequate genotyping quality and proper number for analysis. In the discovery cohort, 290 out of 368 SNPs were successfully genotyped at an average rate of 97%.

Statistical Analysis

Patient clinical and pathological characteristics were summarized as number and percentage of patients for categorical factors or median and range of values for continuous factors. The primary analysis endpoint was recurrence free survival (RFS), defined as time from surgery for a curative intent to recurrence or death or censored at date last known alive. Other analysis endpoints included recurrence free interval (RFI), defined as time from surgery to renal cancer recurrence or censored at death prior to recurrence or last follow-up time and time from the surgery to all cause of death or overall survival (OS). Test for association between RFS and clinical factors used the log-rank test from a Cox regression model.

In the discovery cohort analysis, each SNP was treated as a categorical variable with levels of a common homozygote, a rare homozygote, and a heterozygote. Rare homozygotes were combined with heterozygotes if the rare homozygote count was very low (used cut-off of 10 or frequency of 2.5% for a variant group) to ensure that the regression model can run properly to obtain the coefficient estimates. The analysis of identifying significant association between RFS and genotypes used a logrank test. Multiple comparison adjustment used the false-discovery rate (FDR) measure q-value with a cutoff <10%²¹ to determine a significant association. The q-value represents the expected proportion of false-positive results when declaring a test significant. We choose to use a strict FDR threshold of 10% based on the suggestion from REMARK guideline [<http://www.nature.com/bjc/journal/v93/n4/abs/6602678a.html>] that a very stringent criteria is to be employed if we aim to identify biomarker that holds up in future studies.

Once a particular polymorphism was selected on the basis of log-rank test and false-discovery rate, the distribution of RFS stratified by allelic status was estimated using the Kaplan-Meier method. The associations between the polymorphism and the outcome endpoint (RFS, RFI and OS) used the Cox regression models adjusted for clinical and pathological factors including ECOG PS, clinical stage, tumor size, tumor Fuhrman grade, and histology (clear cell vs. non clear cell).²² The Cox model also estimated hazard ratio (HR) and 95% confidence interval (CI) comparing the SNP variants. The associations between the selected SNP variants and the clinical and pathological factors used Chi-square test for a categorical variable and Wilcoxon (two levels of SNP variant) or Kruskal-Wallis (3 levels of SNP variants) tests for a continuous variable. Similar analyses methods were used to confirm the finding in validation cohort.

All reported p values were two sided. Q-values were calculated using R package: qvalue (<http://cran.r-project.org/web/packages/qvalue>) on the basis of the observed distribution of P values from the log-rank test for individual SNPs. The analyses were carried out using SAS version 9.2 (SAS Institute Inc., Cary, NC) and R 2.10.1 (The R Foundation for Statistical Computing, Vienna, Austria).

Role of the funding source

The funding sources did not have any role in the study design, collection, analysis, interpretation or writing the manuscript. All authors had access to the raw data. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

Patient characteristics

A total of 554 patients were included in the analysis. The analytic cohort for the discovery had a total of 403 patients with localized RCC at presentation who underwent a surgery, while a similar and independent validation cohort of 151 patients were identified. The baseline demographic and pathological characteristics of the study cohorts are presented in Table 1. The majority of patients in discovery cohort were male, had an ECOG PS of 0, stage I/II and clear-cell histology, with a median follow-up time of 43 months (interquartile range (IQR): 26,64) since surgery. The validation cohort had more females and ECOG PS 1 and a median follow-up of 66 months (IQR: 23,106). The remaining clinical characteristics, such as age, tumor size, clinical staging, Fuhrman grade and histology, were comparable between the two cohorts.

Disease characteristics

For the discovery cohort, the RFS event rate was 46% (184/403), of which 167 patients had recurrence and 104 patients died for the follow-up period. The validation cohort had a RFS event rates of 29% (44/151), 22 patients had recurrence and 35 patients died.

The risks of recurrence according to important clinical and pathological characteristics are presented in Table 2. Consistent with previous reports in the setting of localized RCC, patients with higher stage and higher tumor grade had a higher risk of RCC recurrence in both cohorts.³

Association of SNPs with cancer recurrence

A total of 290 SNPs were successfully genotyped, among them, 5 were excluded from analysis due to having only single variant level. The analytic cohort had 285 SNPs with assessable data for analysis in the discovery cohort. Using genotypes as categorical variable, twelve SNPs were shown to be potentially associated with RFS (logrank-test p-value <0.05) and their distribution and properties are summarized in Table 3. Using a pre-specified false discovery rate (FDR) q-value cutoff of 10% for adjustment of multiple comparisons, one synonymous coding SNP (rs11762213) (c.144G>A; p.Ala48Ala) located in exon two of c-MET was shown to be significantly associated with RFS outcome (univariate log-rank $p < 0.0001$; and q-value: 0.027). The distribution of the SNP rs11762213 variants was summarized in Table 3, and the minor allele frequency (MAF) of the SNP rs11762213 (G>A) was 5.3%, consistent with the frequency from the CEU population from the HapMap database. No statistically significant differences ($p > 0.05$) were observed between this SNP in relation to baseline clinical and pathological characteristics (data shown in Supplementary Table 1.A and 1.B).

Further analysis of the genotype association between SNP rs11762213 and outcomes RFS, RFI and OS using univariate and multivariable regression model that adjusted for ECOG PS, stage, tumor size, tumor Fuhrman grade, and histology were performed and the results summarized in Table 4. Patients with one or two copies of the risk allele A (AG and AA genotypes versus GG, respectively) had a significantly increased risk of recurrence or death with a median time to RFS of 19 months (AA,AG) versus 50 months (GG), this corresponds to a univariate HR= 2.22 (95% CI: 1.47, 3.35; p=0.0001) and a multivariable HR=1.86 (95% CI: 1.17, 2.95; p=0.008). The complete multivariable model results, including the clinical and pathological factors, were included in Supplementary Table 2.A and 2.B.

The 5-year RFS event rate was estimated to be 19% (95% CI:9,44) for AA, AG variant group, compared to that of the event rate of 46% (95% CI: 40,53) for GG group. The association between RFI endpoints and the polymorphism was consistent with that of the association for RFS, with risk allele (A) being associated with the increased risk of recurrence. The association of OS with the polymorphism was in the same direction, though not statistically significant (Table 4).

Findings from the discovery cohort were validated in the validation cohort. Patients with risk allele A had an increased risk of recurrence or death with a multivariable HR=2.45 (95%CI: 1.01, 5.95; p=0.048), note that the p-value of 0.048 was marginally significant with a relatively wide 95% CI, this was likely due to the small sample size and a low count in the variant comparison group (16 for AA,AG) versus 132 for GG). The results for the other endpoints RFI and OS were also consistent with those from the discovery cohort (Supplementary Figures 1 and 2). We also explored the incremental value in risk prediction and discrimination, such as c-index and integrated discrimination improvement (IDI)²³ of the c-MET polymorphism rs11762213 over clinical and pathological factors of age, tumor size, clinical stage, Fuhrman grade and histology with respect to disease recurrence. Considering a risk prediction model for recurrence that included the c-MET polymorphism rs11762213 and the clinical and pathological predictors, the net improvement in risk discrimination of the model was quantified using the IDI, which estimated the change in difference in the mean predicted probabilities of the recurrence between those with and without the recurrence after the addition of c-MET polymorphism to the clinical and pathological factors. The improvement in the mean predicted probabilities of recurrence or IDI were from 4%(0.04 with 95% CI: 0.01, 0.08) at 3 years (short-term) and 13% (0.13 with 95%CI: -0.03, 0.3) at 8 years (long-term) (Supplementary Table 3). Figure 1 shows the estimated RFS distribution (Kaplan-Meier curve) by genotype variants for both discovery and validation cohort.

Discussion

In this study we identified patients with one or two copies of the minor risk allele A of SNP rs11762213 in exon 2 of c-MET (c.144G>A) to have a HR of 1.86 for the risk of RCC recurrence in our discovery cohort, and a HR of 2.45 in the validation cohort. The multivariable analysis of OS with minor risk allele was also statistically significant in the validation cohort, although it did not reach statistical significance level in the discovery set, a fact possibly influenced by the longer follow up and more mature survival in the validation set, as well as the possibility of subsequent targeted therapy in recurrent disease in the more recent discovery cohort. To our knowledge, this is the first study to report the association of a germline SNP with an increased risk of recurrence in patients with localized RCC.

c-MET is an intriguing candidate gene. c-MET is the transmembrane tyrosine kinase receptor for the hepatocyte growth factor (HGF). Notably, activating c-MET mutations in the tyrosine kinase domain have been identified in patients with type-1 hereditary papillary

RCC.²⁴ An increased expression and signaling of c-MET in tumor tissues have been observed in several malignancies, including sporadic RCCs.²⁵ In addition, the c-MET gene plays a significant role in RCC pathogenesis, usually cooperating with VEGFR to promote tumor growth.²⁶ Additionally, it has been shown that c-MET might act as a sensor of an adverse microenvironment (i.e. hypoxia) driving cell invasion and metastasis through the activation of certain genes involved in blood coagulation.²⁷ Furthermore, MET polymorphisms have been associated with autism. Recently it has been reported genetic and molecular evidence pointing to a common functional promoter variant in the MET gene as a contributing risk factor for autism susceptibility.²⁸

Fine mapping will be necessary to identify the true causal allele, as rs11762213 may be correlated with other alleles. It is possible that the true causal SNP in linkage disequilibrium with rs11762213 does have more obvious direct downstream effects. If the causal allele is synonymous or non-coding, it presumably does not directly alter the downstream protein product. However, synonymous or noncoding SNPs can affect protein abundance, structure, and/or function through various means: 1) changing mRNA structure and stability; 2) changing kinetics of transcription/translation; or, 3) alternate splicing.²⁹ In addition functional analysis is required to confirm that c-MET is a target gene driving disease aggressiveness and determine exactly how genotype influences this risk. Ultimately, if indeed inherited variation within MET increases risk of RCC recurrence, strategies aiming at using MET inhibitors in high risk patients may prove beneficial. Drugs targeting c-MET in many advanced solid tumors, including metastatic RCC are currently in clinical trials.²⁶

Our results are strengthened by several facts. First, we have evaluated a well-defined cohort in the current study with patients from European-American ancestry, making this the largest study of its kind. Second, all patients in the discovery cohort were followed prospectively after treatment of localized RCC, and their clinical and pathological characteristics are consistent with similar RCC cohorts. Third, we included a broad range of relevant genes that were previously highlighted in the literature. Fourth, our results were adjusted for multiple comparisons, limiting potential false discovery. Fifth, our results remained significant after adjusting for clinical and pathological covariates. Finally, we were able to reproduce our findings in an independent, albeit smaller, cohort of localized RCC patients with DNA extracted from normal tissue of FFPE samples. DNA extracted from normal FFPE tissue has been shown to provide excellent quality and quantity genome-wide genotyping data representing germline DNA, sufficient for both linkage and association analyses.³⁰

Of note, there is an observed deviation from HWE for rs11762213 in our discovery cohort due to the higher observed minor homozygote rate than expected, even though the observed minor homozygote rate was only 1% in the discovery cohort). There are a number of reasons that may contribute to this deviation including selection bias, since we only included patients with RCC but without control group. We acknowledge that the lack of the HWE as a limitation of study, and the results from our study should be regarded with caution, recognize that it may not be applicable for the direct comparisons with other population. In addition, it is important to point out that the observed frequency of A allele in our cohort was consistent with that of the expected frequency in HapMap database (CEU allele frequency of 5%) and the 1000 genome project (minor allele frequency of 4.85% in the CEU cohort).

The vast majority of previous germline polymorphism studies involving RCC evaluated the association of SNPs with the risk of developing cancer, rather than its recurrence.⁵⁻⁷ The majority of these previously reported SNPs were included in our analysis but were not associated with the risk of recurrence in our cohort. We are not aware of previous germline studies reporting on c-MET polymorphisms and cancer risk or prognosis.

In conclusion, our study suggests that a germline genetic polymorphism in c-MET (rs11762213) is independently associated with risk of recurrence in patients with localized RCC. A potential limitation of our findings is that it may not be applicable for other populations from different ancestry origin, as we included only patients from European-American ancestry. Replication in other populations and larger cohort are needed to further validate this finding. After further validation, the risk status at the polymorphism may be incorporated in future prognostic tools, potentially aiding in the design of adjuvant clinical trials with c-MET inhibitors, and clinical management. Additionally, functional studies are warranted in order to identify biological mechanisms involved in this association.

Research in context

Systematic review

We searched PubMed from 1966 to September 2012 for germline studies in renal cell carcinoma (RCC) and other malignancies, focused on recurrence and/or survival. We reviewed studies that assessed survival, recurrence free survival or cancer specific survival in patients who underwent nephrectomy. We have chosen genes involved in RCC pathogenesis including the VHL/HIF/VEGF, PI3K/AKT/mTOR pathways, and genes involved in immune regulation and metabolism. We used our large prospectively-collected database at Dana-Farber/Harvard Cancer Center, first as a training set. The top SNPs were tested in a separate independent cohort of RCC patients that were subjected to a curative-intent surgery. Our goal was to assess how germline genetic polymorphisms may affect the risk of recurrence in patients with localized renal cell carcinoma, taking into account known clinical prognostic factors such as stage and tumor grade.

Interpretation

Our findings show that 1 SNP in the cMET gene (The rs11762213, located in an exon) is independently associated with the primary endpoint of recurrence-free survival in the training and the validation cohorts. One possibility is that the SNP may be incorporated in future prognostic tools, in addition to the well known clinical prognostic factors in localized disease. One further future “validation” step would be to embark on fine mapping to identify the true causal allele.

With the availability of MET inhibitors, it is not unconceivable that a population with high-risk disease (based on clinical and germline risks) be subject to an adjuvant clinical trial with a MET inhibitor.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to acknowledge Meredith Regan, ScD for her help in results interpretation

References

1. Brugarolas J. Renal-cell carcinoma--molecular pathways and therapies. *N Engl J Med*. 2007; 356:185–7. [PubMed: 17215538]
2. Hupertan V, Roupret M, Poisson JF, et al. Low predictive accuracy of the Kattan postoperative nomogram for renal cell carcinoma recurrence in a population of French patients. *Cancer*. 2006; 107:2604–8. [PubMed: 17075871]

3. Cindolo L, Patard JJ, Chiodini P, et al. Comparison of predictive accuracy of four prognostic models for nonmetastatic renal cell carcinoma after nephrectomy: a multicenter European study. *Cancer*. 2005; 104:1362–71. [PubMed: 16116599]
4. Coate L, Cuffe S, Horgan A, Hung RJ, Christiani D, Liu G. Germline genetic variation, cancer outcome, and pharmacogenetics. *J Clin Oncol*. 2010; 28:4029–37. [PubMed: 20679599]
5. Andreotti G, Boffetta P, Rosenberg PS, et al. Variants in blood pressure genes and the risk of renal cell carcinoma. *Carcinogenesis*. 2010; 31:614–20. [PubMed: 20047954]
6. Moore LE, Brennan P, Karami S, et al. Apolipoprotein E/C1 locus variants modify renal cell carcinoma risk. *Cancer Res*. 2009; 69:8001–8. [PubMed: 19808960]
7. Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. *Nat Genet*. 2011; 43:60–5. [PubMed: 21131975]
8. Han SS, Yeager M, Moore LE, et al. The chromosome 2p21 region harbors a complex genetic architecture for association with risk for renal cell carcinoma. *Human molecular genetics*. 2012; 21:1190–200. [PubMed: 22113997]
9. Schodel J, Bardella C, Sciesielski LK, et al. Common genetic variants at the 11q13.3 renal cancer susceptibility locus influence binding of HIF to an enhancer of cyclin D1 expression. *Nature genetics*. 2012; 44:420–5. S1–2. [PubMed: 22406644]
10. Holt SK, Kwon EM, Koopmeiners JS, et al. Vitamin D pathway gene variants and prostate cancer prognosis. *Prostate*. 2010; 70:1448–60. [PubMed: 20687218]
11. Decock J, Long JR, Laxton RC, et al. Association of matrix metalloproteinase-8 gene variation with breast cancer prognosis. *Cancer Res*. 2007; 67:10214–21. [PubMed: 17974962]
12. Domingo-Domenech E, Benavente Y, Gonzalez-Barca E, et al. Impact of interleukin-10 polymorphisms (-1082 and -3575) on the survival of patients with lymphoid neoplasms. *Haematologica*. 2007; 92:1475–81. [PubMed: 18024395]
13. Pomerantz MM, Werner L, Xie W, et al. Association of prostate cancer risk Loci with disease aggressiveness and prostate cancer-specific mortality. *Cancer Prev Res (Phila)*. 2011; 4:719–28. [PubMed: 21367958]
14. Li L, Kaelin WG Jr. New insights into the biology of renal cell carcinoma. *Hematology/oncology clinics of North America*. 2011; 25:667–86. [PubMed: 21763962]
15. Mignogna C, Staibano S, Altieri V, et al. Prognostic significance of multidrug-resistance protein (MDR-1) in renal clear cell carcinomas: a five year follow-up analysis. *BMC cancer*. 2006; 6:293. [PubMed: 17177989]
16. Siegmund M, Brinkmann U, Schaffeler E, et al. Association of the P-glycoprotein transporter MDR1(C3435T) polymorphism with the susceptibility to renal epithelial tumors. *Journal of the American Society of Nephrology: JASN*. 2002; 13:1847–54. [PubMed: 12089380]
17. Thompson RH, Dong H, Kwon ED. Implications of B7-H1 expression in clear cell carcinoma of the kidney for prognostication and therapy. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2007; 13:709s–15s. [PubMed: 17255298]
18. Linehan WM, Srinivasan R, Schmidt LS. The genetic basis of kidney cancer: a metabolic disease. *Nature reviews Urology*. 2010; 7:277–85.
19. Frigola X, Inman BA, Lohse CM, et al. Identification of a soluble form of B7-H1 that retains immunosuppressive activity and is associated with aggressive renal cell carcinoma. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2011; 17:1915–23. [PubMed: 21355078]
20. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007; 449:851–61. [PubMed: 17943122]
21. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*. 2003; 100:9440–5. [PubMed: 12883005]
22. Cox DR. Regression models and life-tables. *JR Stat Soc [B]*. 1972; 34:187–220.
23. Chambless LE, Cummiskey CP, Cui G. Several methods to assess improvement in risk prediction models: Extension to survival analysis. *Statistics in Medicine*. 2011; 30(1):202, 38.

24. Schmidt L, Duh FM, Chen F, et al. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nat Genet.* 1997; 16:68–73. [PubMed: 9140397]
25. Peruzzi B, Bottaro DP. Targeting the c-Met signaling pathway in cancer. *Clin Cancer Res.* 2006; 12:3657–60. [PubMed: 16778093]
26. Eder JP, Vande Woude GF, Boerner SA, LoRusso PM. Novel therapeutic inhibitors of the c-Met signaling pathway in cancer. *Clin Cancer Res.* 2009; 15:2207–14. [PubMed: 19318488]
27. Boccaccio C, Comoglio PM. Invasive growth: a MET-driven genetic programme for cancer and stem cells. *Nat Rev Cancer.* 2006; 6:637–45. [PubMed: 16862193]
28. Campbell DB, Sutcliffe JS, Ebert PJ, et al. A genetic variant that disrupts MET transcription is associated with autism. *Proceedings of the National Academy of Sciences of the United States of America.* 2006; 103:16834–9. [PubMed: 17053076]
29. Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM. Silent polymorphisms speak: how they affect pharmacogenomics and the treatment of cancer. *Cancer Res.* 2007; 67:9609–12. [PubMed: 17942888]
30. Cannon-Albright LA, Cooper KG, Georgelas A, Bernard PS. High quality and quantity Genome-wide germline genotypes from FFPE normal tissue. *BMC research notes.* 2011; 4:159. [PubMed: 21615924]

cMET:rs11762213(Validation set)

cMET:rs11762213(Discovery set) **Figure 1.**

Kaplan Meier curve of RFS in the discovery and validation sets for patients with the dominant genotype (GG) in the c-MET polymorphism (rs11762213) versus patients with 1 or 2 copies of the risk allele A (genotypes AG and AA, respectively).

Table 1

Clinical and pathological characteristics of study cohorts.

Characteristics		Discovery cohort (n=403)	Validation cohort (n=151)
Follow-up (month)	Median (IQR [*])	43(26,64)	66(23,106)
Age (years)	Median (range)	58.0 (21.8 – 91.4)	61.2 (34.3, 86.5)
Tumor size (cm)	Median (range)	6.0 (0.8 – 20.0)	6.0 (1.3, 21.0)
	Missing	19	0
Gender: N (%)	Female	134 (33%)	81 (54%)
	Male	269 (67%)	70 (46%)
ECOG PS ^{**} : N (%)	0	321 (81%)	86 (57%)
	1	77 (19%)	64 (43%)
	Missing	5	1
Clinical Stage: N (%)	I/II	236 (62%)	109 (72%)
	III/IV	147 (38%)	42 (28%)
	Missing	20	0
Furhman Grade: N (%)	1–2	190 (50%)	90 (60%)
	3–4	188 (50%)	59 (40%)
	Missing	25	2
Histology: N (%)	Clear cell	310 (78%)	111 (75%)
	Non-clear cell ^{***}	86 (22%)	37 (25%)
	Papillary	43(50%)	19(51%)
	Chromo	31(36%)	18(49%)
	Mixed	8(9.3%)	-
	Others	4(4.7%)	-
	Missing	7	3

^{*} IQR: interquartile range

^{**} ECOG PS: Eastern Oncology Group Performance Status

^{***} includes: papillary, chromophobe, collecting duct, mixed, others.

Table 2

Association of recurrence free survival (RFS) according to clinical and pathological prognostic factors assessed using univariate Cox regression model.

Characteristics	Discovery cohort					Validation cohort				
	Estimate (coef)	se(coef)	z-statistic	HR (95% CI)	P-value	Estimate (coef)	se(coef)	z-statistic	HR (95% CI)	P-value
ECOG PS: 1 vs. 0(ref)	0.95	0.16	5.8	2.6 (1.88–3.59)	<0.0001	0.47	0.3	1.57	1.6 (0.89–2.88)	0.12
T size (cm)	0.15	0.02	9.07	1.17 (1.13–1.21)	<0.0001	0.03	0.05	0.55	1.03 (0.93–1.13)	0.59
Clinical stage: III/IV vs. I/II (ref)	1.43	0.16	9.06	4.2 (3.08–5.72)	<0.0001	1.34	0.3	4.45	3.83 (2.12–6.93)	<0.0001
Fuhrman grade: 3/4 vs. 1/2 (ref)	1.23	0.17	7.32	3.43 (2.47–4.77)	<0.0001	0.73	0.31	2.39	2.08 (1.14–3.8)	0.017
Histology: Clear-cell vs. Non-clear cell (ref)	0.61	0.2	3	1.84 (1.24–2.74)	0.0027	0.21	0.36	0.58	1.23 (0.61–2.49)	0.56

HR: hazard ratio; CI: confidence interval

ref: reference group in comparison

Table 3

Properties of SNPs identified to be potentially associated with RFS using categorical variant groups* with univariate log-rank p-value less than 0.05 and the corresponding false discovery rate (q-value) in the discovery cohort.

Discovery cohort									
Gene	SNP	Assessable No. ^a	Minor allele (%)	Homozygous(%)	Heterozygous (%)	Wild-type (%)	HWE** p-value	Log-rank p-value	q-value
cMET	rs11762213(G>A)	393	5.3	1.3	8.1	90.6	0.0028	9.40E-05	0.027
GLUT1	rs3820546(A>G)	387	46.8	23.8	46	30.2	0.15	0.0019	0.27
cMET	rs38846(T>C)	389	18.6	3.9	29.6	66.6	0.62	0.0093	0.73
VEGFR2	rs1531290(A>G)	397	46.3	22.4	47.9	29.7	0.48	0.01	0.73
MMP9	rs2236416(A>G)	398	13.8	2.5	22.6	74.9	0.29	0.023	0.89
cMET	rs38845(G>A)	390	45.5	19.7	51.5	28.7	0.48	0.029	0.89
AGT	rs1326889(T>C)	362	48.3	25.4	45.9	28.7	0.12	0.031	0.89
TNF	rs3093662(A>G)	386	8.2	0.5	15.3	84.2	1	0.032	0.89
TNF	rs361525(G>A)	398	4.8	0.3	9	90.7	0.6	0.033	0.89
ABCB1	rs10267099(A>G)	347	23.3	6.3	34	59.7	0.37	0.034	0.89
VHL	rs779805(A>G)	399	32.3	9.8	45.1	45.1	0.57	0.035	0.89
cMET	rs10271561(T>C)	391	10.4	0.8	19.2	80.1	0.78	0.037	0.89
Validation cohort									
Gene	SNP	Assessable No. ^a	Minor allele (%)	Homozygous(%)	Heterozygous (%)	Wild-type (%)	HWE** p-value	Log-rank p-value	
cMET	rs11762213(G>A)	148	5.4	0	10.8	89.2	1	0.042	
GLUT1	rs3820546(T>C)	148	47.3	21.6	51.4	27	0.87	0.064	

Note: C-MET: rs11762213, Chr:116126518 with CEU frequency for allele A=0.05

* categorical variable with levels of a common homozygote, a rare homozygote, and a heterozygote. Rare homozygotes were combined with heterozygotes if the rare homozygote count was very low (used cut-off of 10 or frequency of 2.5% for a variant group).

** HWE: Hardy-Weinberg Equilibrium (p-value reflects the exact test for HWE)

^a patients who had failed genotyping results were excluded from analysis.

Association of disease outcomes of RFS, RFI and OS by genotype variants of c-MET SNP (rs11762213 G>A) assessed using univariate and multivariable Cox regression model.

Table 4

Primary Endpoint		Discovery cohort (N=403)				Validation cohort (N=151)			
		a n(event)	5 year event rate (%) (95% CI)	Univariate HR(95%CI), p-value	Multivariable* HR(95%CI), p-value	a n(event)	5 year event rate (%) (95% CI)	Univariate HR(95%CI), p-value	Multivariable* HR(95%CI), p-value
RFS	Overall								
	cohort	393(184)	43 (37,50)			148(44)	74 (66,82)		
	GG (ref**)	356(157)	46 (40,53)			132(37)	75 (68,84)		
	AA,AG	37(27)	19 (9,44)	2.22 (1.47,3.35), 0.00014	1.86(1.17,2.95), 0.0084	16(7)	58 (36,92)	2.29 (1.01,5.18), 0.048	2.45(1.01,5.95), 0.048
Secondary Endpoints									
	RFI								
	Overall								
	cohort	393(167)	48 (42,55)			148(22)	85 (79,92)		
	GG (ref**)	356(143)	50 (44,57)			132(17)	87 (81,94)		
	AA,AG	37(24)	25 (12,52)	2.18 (1.41,3.37), 0.00043	1.69(1.03,2.76), 0.039	16(5)	68 (48,96)	2.93 (1.08,7.99), 0.035	2.99(0.97,9.2), 0.056
OS	Overall cohort	393(104)	61 (54,68)			148(35)	80 (73,88)		
	GG (ref**)	356(88)	62 (55,70)			132(29)	82 (75,90)		
	AA,AG	37(16)	52 (35,75)	1.68 (0.98,2.87), 0.061	1.24(0.67,2.27), 0.5	16(6)	62 (40,96)	3.06 (1.24,7.56), 0.015	3.52(1.32,9.38), 0.012

* Multivariable analysis adjusted for clinical prognostic factors: ECOG performance status, clinical stage, Fuhrman grade, histology(clear cell vs. all others)

** ref: Reference group

^a 10 patients from discovery cohort and 3 from validation cohort had failed genotyping results so excluded from analysis;

RFS: recurrence-free survival, RFI: recurrence-free interval, OS: overall survival, HR: hazard ratio