

A genome-wide association study identifies a novel susceptibility locus for renal cell carcinoma on 12p11.23

Xifeng Wu^{1,*}, Ghislaine Scelo^{3,†}, Mark P. Purdue^{4,†}, Nathaniel Rothman⁴, Mattias Johansson³, Yuanqing Ye¹, Zhaoming Wang^{4,6}, Diana Zelenika⁷, Lee E. Moore⁴, Christopher G. Wood², Egor Prokhortchouk⁸, Valerie Gaborieau³, Kevin B. Jacobs^{4,6}, Wong-Ho Chow⁴, Jorge R. Toro⁴, David Zaridze⁹, Jie Lin¹, Jan Lubinski¹⁰, Joanna Trubicka¹⁰, Neonilia Szeszenia-Dabrowska¹¹, Jolanta Lissowska¹², Peter Rudnai¹³, Eleonora Fabianova¹⁴, Dana Mates¹⁵, Viorel Jinga¹⁶, Vladimir Bencko¹⁷, Alena Slamova¹⁷, Ivana Holcatova¹⁷, Marie Navratilova¹⁸, Vladimir Janout¹⁹, Paolo Boffetta^{20,21}, Joanne S. Colt⁴, Faith G. Davis²², Kendra L. Schwartz²³, Rosamonde E. Banks²⁴, Peter J. Selby²⁴, Patricia Harnden²⁵, Christine D. Berg⁵, Ann W. Hsing⁴, Robert L. Grubb III²⁶, Heiner Boeing²⁷, Paolo Vineis^{26,29,30}, Françoise Clavel-Chapelon^{31,32}, Domenico Palli³³, Rosario Tumino³⁴, Vittorio Krogh³⁵, Salvatore Panico³⁶, Eric J. Duell³⁷, José Ramón Quirós³⁸, Maria-José Sanchez^{39,40}, Carmen Navarro^{40,41}, Eva Ardanaz^{40,42}, Miren Dorronsoro^{40,43}, Kay-Tee Khaw⁴⁴, Naomi E. Allen⁴⁵, H. Bas Bueno-de-Mesquita⁴⁶, Petra H.M. Peeters^{28,47}, Dimitrios Trichopoulos^{48,49}, Jakob Linseisen^{50,51}, Börje Ljungberg⁵², Kim Overvad⁵³, Anne Tjønneland⁵⁴, Isabelle Romieu³, Elio Riboli²⁸, Victoria L Stevens⁵⁵, Michael J Thun⁵⁵, W. Ryan Diver⁵⁵, Susan M. Gapstur⁵⁵, Paul D. Pharoah^{56,57}, Douglas F. Easton^{56,57}, Demetrius Albanes⁴, Jarmo Virtamo⁵⁸, Lars Vatten⁵⁹, Kristian Hveem⁵⁹, Tony Fletcher⁶⁰, Kvetoslava Koppova¹⁴, Olivier Cussenot⁶¹, Geraldine Cancel-Tassin⁶¹, Simone Benhamou^{62,63}, Michelle A. Hildebrandt¹, Xia Pu¹, Mario Foglio⁷, Doris Lechner⁷, Amy Hutchinson^{4,6}, Meredith Yeager^{4,6}, Joseph F. Fraumeni Jr⁴, Mark Lathrop⁷, Konstantin G. Skryabin⁶⁴, James D. McKay³, Jian Gu^{1,‡}, Paul Brennan^{3,‡} and Stephen J. Chanock^{4,*},[‡]

¹Department of Epidemiology, Division of Cancer Prevention and Population Sciences and ²Department of Urology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA, ³International Agency for Research on Cancer (IARC), Lyon 69008, France, ⁴Division of Cancer Epidemiology and Genetics, National Cancer Institute and ⁵Department of Health and Human Services, Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, ⁶Core Genotyping Facility, SAIC-Frederick Inc., National Cancer Institute-Frederick, Frederick, MD, USA, ⁷Commissariat à l'énergie Atomique, Institut Genomique, Centre National de Genotypage, Evry 91000, France, ⁸Center "Bioengineering" of Russian Academy of Sciences, 117312 Prospekt 60 letiya Oktyabrya 7-1, Moscow, Russian Federation, ⁹Russian N.N.Blokhin Cancer Research Centre, Kashirskoye shosse 24, Moscow 115478, Russian Federation, ¹⁰Department of Genetics and Pathomorphology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland, ¹¹Department of Epidemiology, Institute of Occupational Medicine, Lodz 90950, Poland, ¹²M. Sklodowska-Curie Memorial Cancer Center and Institute of

*To whom correspondence should be addressed. Tel: +1 7137452485; Fax: +1 7137924657; Email: xwu@mdanderson.org (X.W.) or chanocks@mail.nih.gov (S.J.C.)

[†]Co-first author.

[‡]Co-last author.

Oncology, Warsaw 02781, Poland, ¹³Department of Environmental Epidemiology, National Institute of Environmental Health, Gyali út 2-6, Budapest 1097, Hungary, ¹⁴Regional Authority of Public Health in Banska Bystrica, Banska Bystrica 97556, Slovakia, ¹⁵National Institute of Public Health, Bucharest 050463, Romania, ¹⁶“T. Burghel” Clinic of Urology, Bucharest, Romania, ¹⁷Institute of Hygiene and Epidemiology, First Faculty of Medicine, Charles University in Prague, Prague 2 12800, Czech Republic, ¹⁸Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno 65653, Czech Republic, ¹⁹Palacky University, Olomouc 77515, Czech Republic, ²⁰Mount Sinai School of Medicine, The Tisch Cancer Institute, One Gustave L. Levy Place, New York, NY 10029, USA, ²¹International Prevention Research Institute, Lyon 69006, France, ²²Division of Epidemiology/Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL 60612, USA, ²³Karmanos Cancer Institute and Department of Family Medicine, Wayne State University, Detroit, MI 48201, USA, ²⁴Cancer Research UK Centre, Leeds Institute of Molecular Medicine and ²⁵Department of Pathology, St James’s University Hospital, Leeds LS9 7TF, UK, ²⁶Division of Urologic Surgery, Washington University School of Medicine, St. Louis, MO 63110, USA, ²⁷Department of Epidemiology, German Institute of Human Nutrition, Potsdam-Rehbruecke, Nuthetal 14558, Germany, ²⁸School of Public Health and ²⁹MRC/HPA Centre for Environment and Health, Imperial College, London SW7 2AZ, UK, ³⁰HuGeF Foundation, Torino 10100, Italy, ³¹INSERM, Centre for Research in Epidemiology and Population Health, U1018, Institut Gustave Roussy, Villejuif F-94805, France, ³²Paris South University, UMRS 1018, Villejuif F-94805, France, ³³Molecular and Nutritional Epidemiology Unit Cancer Research and Prevention Institute–ISPO, Florence, Italy, ³⁴Cancer Registry and Histopathology Unit, “Civile M.P.Arezzo” Hospital, ASP 7, Ragusa, Italy, ³⁵Nutritional Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milano 20133, Italy, ³⁶Department of Clinical and Experimental Medicine, Federico II University, Naples 80131, Italy, ³⁷Unit of Nutrition, Environment and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona 08907, Spain, ³⁸Public Health and Health Planning Directorate, Asturias, Spain, ³⁹Andalusian School of Public Health, Granada 18011, Spain, ⁴⁰Epidemiología y Salud Pública, CIBER, (CIBERESP), Madrid 28029, Spain, ⁴¹Department of Epidemiology, Regional Council of Health and Consumer Affairs, Murcia 30008, Spain, ⁴²Public Health Institute of Navarra, Pamplona, Spain, ⁴³Basque Regional Health Department, Public Health Division of Gipuzkoa, San Sebastian 20113, Spain, ⁴⁴Department of Gerontology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 0XY, UK, ⁴⁵Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7LF, UK, ⁴⁶National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, ⁴⁷Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht 3508 GA, The Netherlands, ⁴⁸Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA, ⁴⁹Bureau of Epidemiologic Research, Academy of Athens, Greece, ⁵⁰Division of Cancer Epidemiology, German Cancer Research Centre, Heidelberg 69120, Germany, ⁵¹Institute of Epidemiology I, Helmholtz Centre Munich, Neuherberg 85764, Germany, ⁵²Department of Surgical and Perioperative Sciences, Urology and Andrology, Umeå University, Umeå 90185, Sweden, ⁵³Department of Epidemiology, School of Public Health, Aarhus University, Aarhus DK-8000, Denmark, ⁵⁴The Danish Cancer Society, Institute of Cancer Epidemiology, Copenhagen 2100, Denmark, ⁵⁵Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA, ⁵⁶Department of Oncology and ⁵⁷Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK, ⁵⁸Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, FIN-00300, Finland, ⁵⁹Department of Public Health and General Practice, Norwegian University of Science and Technology, Trondheim 7489, Norway, ⁶⁰London School of Hygiene and Tropical Medicine, University of London, London WC1H 9SH, UK, ⁶¹CeRePP, Tenon Hospital APHP (ER2-University Paris 6), Paris 75020, France, ⁶²INSERM U946, Paris 75010, France, ⁶³CNRS UMR8200, Institut Gustave Roussy, Villejuif 94805, France, ⁶⁴Kurchatov Scientific Center, 123182 Kurchatov sq 1, Moscow, Russian Federation

Renal cell carcinoma (RCC) is the most lethal urologic cancer. Only two common susceptibility loci for RCC have been confirmed to date. To identify additional RCC common susceptibility loci, we conducted an independent genome-wide association study (GWAS). We analyzed 533 191 single nucleotide polymorphisms (SNPs) for association with RCC in 894 cases and 1516 controls of European descent recruited from MD Anderson Cancer Center in the primary scan, and validated the top 500 SNPs *in silico* in 3772 cases and

8505 controls of European descent involved in the only published GWAS of RCC. We identified two common variants in linkage disequilibrium, rs718314 and rs1049380 ($r^2 = 0.64$, $D' = 0.84$), in the inositol 1,4,5-triphosphate receptor, type 2 (*ITPR2*) gene on 12p11.23 as novel susceptibility loci for RCC ($P = 8.89 \times 10^{-10}$ and $P = 6.07 \times 10^{-9}$, respectively, in meta-analysis) with an allelic odds ratio of 1.19 [95% confidence interval (CI): 1.13–1.26] for rs718314 and 1.18 (95% CI: 1.12–1.25) for rs1049380. It has been recently identified that rs718314 in *ITPR2* is associated with waist–hip ratio (WHR) phenotype. To our knowledge, this is the first genetic locus associated with both cancer risk and WHR.

INTRODUCTION

Kidney cancer is the eighth leading cancer in the United States, with an estimated 58 240 new cases and 13 040 deaths in 2010 (1). The incidence rates for kidney cancer have been steadily increasing in most high-income countries in recent decades until recently in some European countries where kidney cancer incidence rates have decreased or stabilized (2,3). In over 85% of kidney cancers, the pathological subtype is renal cell carcinoma (RCC). Established modifiable risk factors for RCC include cigarette smoking, obesity and hypertension. Other risk factors may include physical inactivity, occupational exposure to trichloroethylene and a history of diabetes mellitus (3).

There is compelling evidence for genetic predisposition to RCC, both in rare hereditary syndromes and in the general population (4,5). Rare high-penetrance susceptibility genes include *VHL* (von Hippel-Lindau syndrome), *MET* (hereditary papillary renal carcinoma), *BHD* (Birt-Hogg-Dube syndrome) and *FH* (hereditary leiomyomatosis and RCC), which are also involved in certain subtypes of sporadic RCC through somatic mutations (4). Genetic predisposition to sporadic RCC has been demonstrated in a meta-analysis of published epidemiological studies in which family history of kidney cancer conferred a 2.2-fold (95% CI: 1.6–2.9) increased risk (6). A recent genome-wide association study (GWAS) comprising individuals of European descent identified two common low-penetrance variants on 2p21 and 11q13.3 that were associated with the risk of RCC (5). A candidate gene on 2p21 is *EPAS1*, which encodes the hypoxia-inducible factor-2 α , central to the VHL–HIF pathway strongly implicated in RCC pathogenesis (5). These two variants are the only confirmed loci for RCC to date and explain a small fraction of the familial risk of RCC. To identify additional novel RCC susceptibility loci, we conducted an independent primary scan of RCC followed by validation of the top 500 single nucleotide polymorphisms (SNPs) in an *in silico* replication and meta-analysis using previously published GWAS results (5).

RESULTS

The primary scan was performed using the Illumina Infinium HumanHap660W BeadChip in an RCC case–control study consisting of 910 cases and 566 controls of European descent. Cases were newly diagnosed and histologically confirmed RCC patients from MD Anderson Cancer Center, and controls were recruited through random digital dialing in Texas and matched to cases on age, gender and residence (6). We also included 972 controls of Texas residence and

European descent who were genotyped using HumanHap610 in a recently published GWAS of bladder cancer (7). After applying strict quality control criteria (Materials and Methods), we restricted the analysis to genotyped SNPs common to the HumanHap610 and 660W Beadchips; accordingly, we analyzed 533 191 SNPs for association with RCC risk for 894 cases and 1516 controls in stage one. A quantile–quantile (Q–Q) plot of observed versus expected χ^2 -test statistics showed no evidence for inflation of χ^2 -tests, suggesting no evidence for differences in population substructure (inflation factor $\lambda = 1.037$; Supplementary Material, Fig. S1). We adjusted for genomic control in the analysis and there was little difference in the P -values (data not shown). While none of the SNPs reached genome-wide significance in stage one (Supplementary Material, Fig. S2), seven SNPs had a P -value $< 10^{-5}$ and 58 SNPs showed a P -value $< 10^{-4}$, all analyzed *in silico* in stage 2. The two GWAS-identified RCC susceptibility SNPs (5) were confirmed in our primary scan, with allelic ORs of 1.14 (95% CI, 1.01–1.28, $P = 0.039$) for rs11894252 on *EPAS1* (2p21), and 0.75 (95% CI, 0.59–0.95, $P = 0.017$) for rs7105934 on 11q13.3, respectively (Supplementary Material, Table S1). The corresponding results in the published GWAS were 1.14 (95% CI, 1.09–1.20, $P = 1.8 \times 10^{-8}$) and 0.69 (0.62–0.76, $P = 7.8 \times 10^{-14}$), respectively (5).

We performed a meta-analysis of the 500 top SNPs from our primary scan data with that of the recently published GWAS of RCC, which had a total of 4092 cases and 8991 controls of European descent undergoing primary scans using various Illumina HumanHap BeadChips (HapMap 300, 500, 610 and 660W) (5). The study design, population characteristics and genotyping platforms for that study were previously described (5). To ensure consistency of genotyping, we only selected the top 500 SNPs that are common across the different BeadChips and did not use imputed data for this meta-analysis. After quality control procedures, 3772 cases and 8505 controls were used for meta-analysis. Combining our primary scan and this GWAS population, there were 4666 RCC cases and 10 021 controls for the meta-analysis.

Two SNPs (rs718314 and rs1049380) on 12p11.23 reached genome-wide significance in meta-analysis (Fig. 1). These two SNPs were in high linkage disequilibrium (LD; $r^2 = 0.64$, $D' = 0.84$). The P -values for the discovery phase were 2.43×10^{-4} and 1.19×10^{-3} and for the validation phase were 4.28×10^{-7} and 9.19×10^{-7} , respectively. The P -values reached 8.89×10^{-10} for rs718314 and 6.07×10^{-9} for rs1049380 in meta-analysis of two scans using a fixed effects model (Supplementary Material, Table S2). There was no significant heterogeneity between the ORs

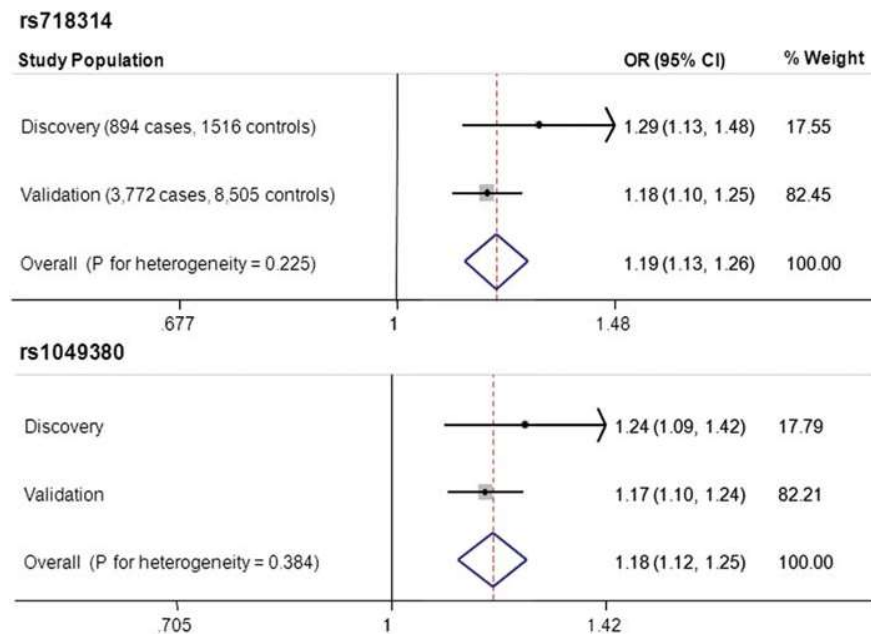


Figure 1. Forest plot showing the association of rs718314 and rs1049380 with RCC risk in discovery and validation populations and in meta-analysis.

Table 1. Association between rs718314 and the risk of RCC

| Analysis ^a | Discovery OR (95% CI) | P-value | Validation OR (95% CI) | P-value | Meta-analysis OR (95% CI) | P-value | P for heterogeneity |
|--|--------------------------|------------|---------------------------|------------|------------------------------|------------|------------------------|
| Univariate | 1.28 (1.12–1.46) | 2.65E – 04 | 1.17 (1.10–1.25) | 6.34E – 07 | 1.19 (1.13–1.26) | 1.39E – 09 | 0.240 |
| Adjusted for age and sex | 1.29 (1.13–1.48) | 2.43E – 04 | 1.18 (1.10–1.25) | 4.28E – 07 | 1.19 (1.13–1.26) | 8.89E – 10 | 0.225 |
| Adjusted for age, sex and BMI | 1.27 (1.09–1.47) | 2.23E – 03 | 1.17 (1.09–1.23) | 1.09E – 06 | 1.18 (1.12–1.26) | 1.34E – 08 | 0.349 |
| Adjusted for age, sex and eigenvector | 1.28 (1.12–1.47) | 3.09E – 04 | 1.18 (1.10–1.25) | 4.28E – 07 | 1.19 (1.13–1.26) | 1.05E – 09 | 0.246 |
| Adjusted for age, sex, BMI and eigenvector | 1.26 (1.08–1.46) | 3.12E – 03 | 1.17 (1.10–1.25) | 1.09E – 06 | 1.18 (1.12–1.25) | 1.69E – 08 | 0.397 |
| Stratified analysis ^b | | | | | | | |
| BMI < 25 | 1.41 (1.04–1.91) | 2.52E – 02 | 1.16 (1.01–1.33) | 3.14E – 02 | 1.20 (1.06–1.35) | 4.04E – 03 | 0.242 |
| BMI ≥ 25 and BMI < 30 | 1.12 (0.87–1.44) | 3.75E – 01 | 1.12 (1.01–1.25) | 3.79E – 02 | 1.12 (1.02–1.24) | 2.40E – 02 | 0.995 |
| BMI ≥ 30 and BMI < 35 | 1.31 (0.92–1.86) | 1.33E – 01 | 1.21 (1.02–1.44) | 3.03E – 02 | 1.23 (1.05–1.43) | 9.13E – 03 | 0.693 |
| BMI ≥ 35 | 1.15 (0.70–1.87) | 5.82E – 01 | 1.25 (0.94–1.66) | 1.20E – 01 | 1.22 (0.96–1.56) | 1.05E – 01 | 0.763 |

^aThe specified analyses were done in discovery set. In the validation set, study site was a covariate in all the analysis, the two eigenvectors were covariates in all analysis except univariate analysis, and age was not included as a covariate.

^bStratified analyses were adjusted for age and sex in the discovery set, and for sex, study site and the two eigenvectors in the validation set.

estimated for the two stages. None of the other SNPs reached genome-wide significance in meta-analysis (Supplementary Material, Table S2). We performed principal component analysis and found one borderline significant ($P < 0.1$) eigenvector. When we included this eigenvector in the logistic regression model, the results were similar to the model without this eigenvector (Table 1).

Since obesity is a risk factor for RCC, and more interestingly, rs718314 has recently been associated with waist–hip ratio [WHR (8)], we added body mass index (BMI) as a covariate in logistic regression analysis and also performed stratified analyses by BMI (Table 1). The ORs were similar with or without BMI adjustment and there was no meaningful difference in the allelic ORs across different BMI strata either. Similarly, we examined the results stratified by the other major risk factors

for RCC, namely smoking status, hypertension and family history of kidney cancer, and observed no differences across strata (data not shown).

We next imputed genotypes [IMPUTE version 2 (9)] within 1 Mb of these two SNPs using the 1000 Genomes Project data (Build 36) combined with HapMap 3 (release 2) in our primary scan population (Fig. 2A). LD analysis of genotyped and imputed SNPs showed that these two SNPs reside within an LD block on 12p11.23 containing only the *ITPR2* gene (Fig. 2B).

DISCUSSION

We identified two highly linked common variants in *ITPR2* on 12p11.23 that point to a novel susceptibility locus for RCC

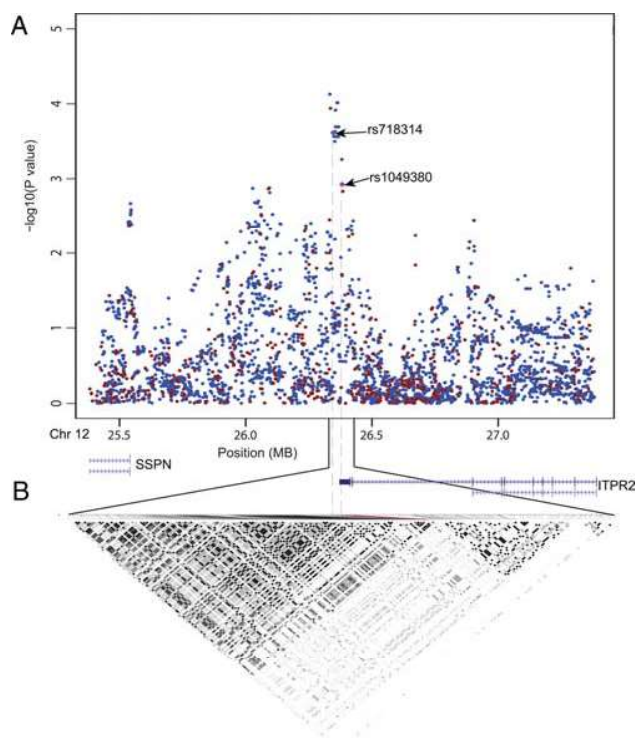


Figure 2. (A) Results of SNP association from primary scan: observed results from genotyped SNPs are in red and imputed results are in black. All known genes in this region are also shown. (B) LD structure across this region based on genotyped and imputed SNPs. Shown in each box is the square of the correlation coefficient (r^2) derived from genotyped and imputed SNPs in Haploview software. Darker shading box indicates greater extent of LD between two SNPs.

through a GWAS. Interestingly, rs718314 was recently unequivocally identified as a locus for WHR in a meta-analysis of 32 GWAS followed by candidate validation in 29 studies ($P = 1.14 \times 10^{-17}$ with a sample size of nearly 185 000) (8). Rs1049380 is located in the 3'-untranslated region of the *ITPR2* gene, 2786 nucleotides from the stop codon. Rs718314 lies ~ 36 kb distal to rs1049380 and resides just outside of the *ITPR2* gene.

There is biological plausibility that *ITPR2* could directly affect RCC susceptibility. *ITPR2* belongs to the family of inositol 1,4,5-triphosphate (IP3) receptors which are intracellular Ca^{2+} -release channels residing in the endoplasmic reticulum (ER) and which mediate Ca^{2+} mobilization from ER to cytoplasm in response to the binding of IP3, a second messenger. IP3-induced Ca^{2+} release is triggered by many additional external stimuli and plays important physiological roles (10). There are three subtypes of IP3 receptors in mammals—type 1 (*ITPR1*), type 2 (*ITPR2*) and type 3 receptor (*ITPR3*). These IP3-mediated receptors are expressed in a cell-specific manner and play important physiological functions. In the kidney, the vasopressin receptor, the endothelin receptor and the angiotensin II receptor are coupled to phosphatidylinositol responses. A previous study showed that *ITPR2* was the most abundantly expressed subtype in murine kidney tissues, higher than *ITPR1*, whereas *ITPR3* was not

detectable (11). Moreover, exposing mice to a nephrotoxic agent, uranyl acetate (UA), caused significantly increased mRNA and protein levels of *ITPR2* in kidney, but did not affect the levels of *ITPR1* (11). These data suggest a more prominent role of *ITPR2* than other two subtypes in kidney function and *ITPR2* may participate in UA-induced nephrotoxicity (11). In another microarray gene-expression study, *ITPR2* was also shown as among the most significant genes that were downregulated in RCC tumors when compared with adjacent normal kidney tissues (12). These studies provide biological plausibility that *ITPR2* could be involved in RCC tumorigenesis, although direct functional evidence supporting the link between *ITPR2* and RCC is still lacking.

It is also intriguing whether *ITPR2* affects RCC risk at least partially through obesity-related pathways. The most significant SNP in our study, rs718314, was unequivocally associated with WHR (8). Previous GWAS have identified a few common SNPs for BMI and certain cancers, not including RCC (13). Rs718314 is the first common SNP associated with both WHR and cancer. WHR has been suggested as a better measure than BMI for obesity. The positive association between BMI and RCC risk is well established (2,3). The positive association between WHR and RCC risk has been fairly consistent in several epidemiological studies and the association appeared to be stronger in women than in men (14–17). Biologically, a previous study showed that mice with double-knockout of both *ITPR2* and *ITPR3* exhibited exocrine dysfunction that caused difficulties in nutrient digestion. Despite a normal caloric intake, the double mutants were hypoglycemic and lean (18). *ITPR2* exhibited significantly differential expression between abdominal and gluteal subcutaneous adipose tissue (8). Unfortunately, we only have BMI data but do not have WHR data in our studies. Inclusion of BMI as a covariate in the logistic regression analysis did not significantly alter the OR, neither did we observe meaningful difference in ORs across different BMI strata in stratified analysis, suggesting that the association of rs718314 with RCC risk may be independent of obesity. However, future RCC association studies with WHR data are warranted to clarify the relationship between rs718314, WHR and RCC risk. Finally, we could not rule out that a separate gene located elsewhere may be tagged by the two significant SNPs and is the causal gene. Much more additional functional studies are needed to delineate the biological mechanism underlying the link between the SNPs on 12q11.23 and RCC risk.

In summary, we identified conclusively a new RCC susceptibility locus through a meta-analysis of GWAS. The common genetic variants identified in our study implicate a shared locus for WHR and RCC risk and also point towards a plausible candidate gene, *ITPR2*, for investigating the biological basis of this association.

MATERIALS AND METHODS

Study population

MD Anderson discovery population. Cases and controls for the primary scan were derived from an ongoing RCC case-control study starting 2002. The study design was described previously (6). Briefly, all cases were newly diagnosed,

histopathologically confirmed and previously untreated RCC patients recruited from the University of Texas MD Anderson Cancer Center. There were no age, gender, ethnicity and cancer-stage restrictions. The controls were healthy individuals with no prior history of cancer (except non-melanoma skin cancer) recruited through random digit dialing (RDD). The controls must have lived for at least 1 year in the same county or in socio-economically matched surrounding counties where the case resided. The controls were frequency matched to the cases by age (± 5 years), gender, ethnicity and county of residence. The overall response rate for RDD screening was 51% and, among those who agreed to participate, the response rate was 88%. The response rate for the eligible cases was 87%. We also included controls from an ongoing bladder cancer case-control study who were involved in a previously published GWAS of bladder cancer (7). These controls were recruited from Kelsey Seybold Clinic, the largest multi-specialty, managed-care physician group in the Houston metropolitan area and all these controls were also residents of Texas. We only included self-reported Caucasian individuals for this study. These studies were approved by the Institutional Review Boards of participating institutions and informed consent was obtained from all study participants before the collection of epidemiological data and blood samples by trained MD Anderson staff interviewers.

Validation populations. We used RCC cases and controls from a completed GWAS as our validation populations (5). There were two primary scans performed in that GWAS, one was coordinated by the International Agency for Research on Cancer (IARC) and the Centre National de Génotypage (CNG) consisting of 2639 RCC cases and 5392 controls of European background drawn from seven studies conducted in Europe, and the other was led by the US National Cancer Institute (NCI) based on 1453 RCC cases and 3531 controls of European background from four US studies. The study design of each participating study and population characteristics were previously described (5). After quality control, 3772 cases and 8505 controls were used in meta-analysis.

Genotyping. The primary scan for the discovery population of RCC case-control study (910 cases and 566 controls) was performed at MD Anderson Cancer Center using the Illumina HumanHap660W Beadchips. Genotyping of the additional 972 controls of the bladder cancer control study was performed using the HumanHap610 Beadchips (7). Cases and controls were excluded from analysis if they had genotyping call rates less than 95% ($n = 4$), were found on review to be duplicated samples or known relatives to another sample ($n = 12$), were found to have reported a gender that did not match with X chromosome heterozygosity ($n = 10$). We excluded 13 samples that deviated by more than 4 SD from other study subjects using similarity in genotypes implemented in PLINK. After all quality control procedures were completed, 2410 samples including 894 cases and 1516 controls were available for analyses. We restricted analysis to genotyped SNPs common to both the HumanHap610 and 660W Beadchips. There were 561 882 overlapping SNPs between these two chips. After removing SNPs that were copy number variation markers, did not yield genotype and had

minor allele frequency < 0.01 , call rate $< 95\%$, Hardy-Weinberg equilibrium $P < 0.0001$ in controls, or $P < 0.001$ in the comparison of RCC controls and bladder cancer controls, we finally had 533 191 SNPs in the final analysis. For the published GWAS, the genotyping platform for each study was described previously (5). Briefly, the primary scans of the European studies were performed with the HumanHap 300 or 610 Beadchips. The primary scans of the four US studies were performed at the NCI Core Genotyping Facility with the HumanHap 500, 610 or 660w BeadChips.

Statistical analyses. Associations between each SNP and the risk of RCC in the primary scan was estimated using multivariate unconditional logistic regression by the OR and 95% CI assuming an additive genetic model with 1 degree of freedom. The genotype was coded as 0, 1, 2 (the count of minor alleles), except for SNPs in chromosome X among men in which the genotype was coded 0 for major allele and 2 for the minor allele. Q-Q plots of the χ^2 -test statistics was used to assess possible population stratification, adequacy of the case-control matching, and the possibility of differential genotyping in cases and controls. We performed principal component analysis implemented in SVS (Golden Helix, Inc) in our primary scan. Logistic regression analysis adjusting for age and gender was used to assess the significance of the eigenvectors and revealed one borderline significant ($P < 0.1$) eigenvector. When we included this eigenvector in the model, the results were similar to the model without this eigenvector. The final logistic regression analysis was adjusted by age and gender in the discovery stage and adjusted by gender, study center and the two eigenvectors in the replication studies. We performed fixed effects meta-analysis using allelic OR and 95% CI derived from the unconditional logistic regression to estimate the combined ORs and P -value for discovery and validation phases and Cochran's Q -statistic for heterogeneity implemented in STATA. We used the Haploview software (v.4.2) to infer the LD structure (19). We used IMPUTE (version 2) (9) to impute SNPs within 1 Mb of rs1049380 based on the public reference data of the 1000 Genomes Project (Build 36) combined with HapMap 3 (release 2).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank the genotyping personnel, study coordinators, and interviewers for performing experiments and recruiting participants. We are especially thankful for all the study participants who made the research possible.

Conflict of Interest statement. None declared.

FUNDING

The study was partially supported by the US National Institutes of Health (NIH) grant R01CA 98897 (X.W.), MD

Anderson Research Trust (X.W.), French Institut National du Cancer (INCa) for those studies coordinated by IARC/CNG and the intramural research program of the National Cancer Institute (NCI), NIH for those studies coordinated by the NCI.

REFERENCES

- Jemal, A., Siegel, R., Xu, J. and Ward, E. (2010) Cancer statistics, 2010. *CA Cancer J. Clin.*, **60**, 277–300.
- Chow, W.H. and Devesa, S.S. (2008) Contemporary epidemiology of renal cell cancer. *Cancer J.*, **14**, 288–301.
- Chow, W.H., Dong, L.M. and Devesa, S.S. (2010) Epidemiology and risk factors for kidney cancer. *Nat. Rev. Urol.*, **7**, 245–257.
- Linehan, W.M., Pinto, P.A., Bratslavsky, G., Pfaffenroth, E., Merino, M., Vocke, C.D., Toro, J.R., Bottaro, D., Neckers, L., Schmidt, L.S. *et al.* (2009) Hereditary kidney cancer: unique opportunity for disease-based therapy. *Cancer*, **115**, 2252–2261.
- Purdue, M.P., Johansson, M., Zelenika, D., Toro, J.R., Scelo, G., Moore, L.E., Prokhorchouk, E., Wu, X., Kiemeny, L.A., Gaborieau, V. *et al.* (2011) Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. *Nat. Genet.*, **43**, 60–65.
- Clague, J., Lin, J., Cassidy, A., Matin, S., Tannir, N.M., Tamboli, P., Wood, C.G. and Wu, X. (2009) Family history and risk of renal cell carcinoma: results from a case-control study and systematic meta-analysis. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 801–807.
- Wu, X., Ye, Y., Kiemeny, L.A., Sulem, P., Rafnar, T., Matullo, G., Seminara, D., Yoshida, T., Saeki, N., Andrew, A.S. *et al.* (2009) Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer. *Nat. Genet.*, **41**, 991–995.
- Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M.C., Speliotes, E.K., Magi, R. *et al.* (2010) Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.*, **42**, 949–960.
- Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.*, **5**, e1000529.
- Mikoshiba, K. (2007) The IP3 receptor/Ca2+ channel and its cellular function. *Biochem. Soc. Symp.*, **74**, 9–22.
- Ondrias, K., Sirova, M., Kubovcakova, L. and Krizanova, O. (2008) Uranyl acetate modulates gene expression and protein levels of the type 2, but not type 1 inositol 1,4,5-trisphosphate receptors in mouse kidney. *Gen. Physiol. Biophys.*, **27**, 187–193.
- Dalgin, G.S., Holloway, D.T., Liou, L.S. and DeLisi, C. (2007) Identification and characterization of renal cell carcinoma gene markers. *Cancer Inform.*, **3**, 65–92.
- Roberts, D.L., Dive, C. and Renehan, A.G. (2010) Biological mechanisms linking obesity and cancer risk: new perspectives. *Annu. Rev. Med.*, **61**, 301–316.
- Adams, K.F., Leitzmann, M.F., Albanes, D., Kipnis, V., Moore, S.C., Schatzkin, A. and Chow, W.H. (2008) Body size and renal cell cancer incidence in a large US cohort study. *Am. J. Epidemiol.*, **168**, 268–277.
- Dal Maso, L., Zucchetto, A., Tavani, A., Montella, M., Ramazzotti, V., Talamini, R., Canzonieri, V., Garboglio, A., Negri, E., Tonini, A. *et al.* (2007) Renal cell cancer and body size at different ages: an Italian multicenter case-control study. *Am. J. Epidemiol.*, **166**, 582–591.
- Luo, J., Margolis, K.L., Adami, H.O., Lopez, A.M., Lessin, L. and Ye, W. (2007) Body size, weight cycling, and risk of renal cell carcinoma among postmenopausal women: the Women's Health Initiative (United States). *Am. J. Epidemiol.*, **166**, 752–759.
- Prineas, R.J., Folsom, A.R., Zhang, Z.M., Sellers, T.A. and Potter, J. (1997) Nutrition and other risk factors for renal cell carcinoma in postmenopausal women. *Epidemiology*, **8**, 31–36.
- Futatsugi, A., Nakamura, T., Yamada, M.K., Ebisui, E., Nakamura, K., Uchida, K., Kitaguchi, T., Takahashi-Iwanaga, H., Noda, T., Aruga, J. *et al.* (2005) IP3 receptor types 2 and 3 mediate exocrine secretion underlying energy metabolism. *Science*, **309**, 2232–2234.
- Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.