Published in final edited form as:
Nat Genet. 2010 May ; 42(5): 415–419. doi:10.1038/ng.558.

A sequence variant at 4p16.3 confers susceptibility to urinary bladder cancer

Lambertus A Kiemeney[1,2,3,44,*], **Patrick Sulem**[4,44], **Soren Besenbacher**[4], **Sita H Vermeulen**[1], **Asgeir Sigurdsson**[4], **Gudmar Thorleifsson**[4], **Daniel F Gudbjartsson**[4], **Simon N Stacey**[4], **Julius Gudmundsson**[4], **Carlo Zanon**[4], **Jelena Kostic**[4], **Gisli Masson**[4], **Hjordis Bjarnason**[4], **Stefan T Palsson**[4], **Oskar B Skarphedinsson**[4], **Sigurjon A Gudjonsson**[4], **J Alfred Witjes**[2], **Anne J Grotenhuis**[1], **Gerald W Verhaegh**[2], **D Timothy Bishop**[5], **Sei Chung Sak**[6], **Ananya Choudhury**[7], **Faye Elliott**[5], **Jennifer H Barrett**[5], **Carolyn D Hurst**[6], **Petra J de Verdier**[8], **Charlotta Ryk**[8], **Peter Rudnai**[9], **Eugene Gurzau**[10], **Kvetoslava Koppova**[11], **Paolo Vineis**[12,13], **Silvia Polidoro**[12,14], **Simonetta Guarrera**[12,14], **Carlotta Sacerdote**[15,16], **Marcello Campagna**[17], **Donatella Placidi**[17], **Cecilia Arici**[17], **Maurice P Zeegers**[18,19], **Eliane Kellen**[20], **Berta Saez Gutierrez**[21], **José I Sanz-Velez**[22], **Manuel Sanchez-Zalabardo**[23], **Gabriel Valdivia**[24], **Maria D Garcia-Prats**[25], **Jan G Hengstler**[26], **Meinolf Blaszkewicz**[26], **Holger Dietrich**[27], **Roel A Ophoff**[28,29], **Leonard H van den Berg**[30], **Kristin Alexiusdottir**[31], **Kristleifur Kristjansson**[4], **Gudmundur Geirsson**[32], **Sigfus Nikulasson**[33], **Vigdis Petursdottir**[33], **Augustine Kong**[4], **Thorgeir Thorgeirsson**[4], **N Aydin Mungan**[34], **Annika Lindblom**[35], **Michael A van Es**[30], **Stefano Porru**[17], **Frank Buntinx**[36,37], **Klaus Golka**[26], **José I Mayordomo**[38], **Rajiv Kumar**[39], **Giuseppe Matullo**[12,14], **Gunnar Steineck**[40,41], **Anne E Kiltie**[42], **Katja K H Aben**[1,3], **Eirikur Jonsson**[32], **Unnur Thorsteinsdottir**[4,43], **Margaret A Knowles**[6], **Thorunn Rafnar**[4], and **Kari Stefansson**[4,43]

[1]Department of Epidemiology, Biostatistics and Health Technology Assessment, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [2]Department of Urology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [3]Comprehensive Cancer Center East, Nijmegen, The Netherlands. [4]deCODE Genetics, Reykjavik, Iceland. [5]Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, St. James's University Hospital, Leeds, UK. [6]Section of Experimental Oncology, Leeds Institute of Molecular

Correspondence should be addressed to L.A.K. (b.kiemeney@ebh.umcn.nl) or K.S. (kstefans@decode.is)..
[44]These authors contributed equally to this work.
[*]A full list of author affiliations appears at the end of the paper.

AUTHOR CONTRIBUTIONS The study was designed and results were interpreted by L.A.K., P.S., U.T., M.A.K., T.R. and K.S. Statistical analysis was carried out by L.A.K., P.S., S.B., G.T., D.F.G., G. Masson and A.K. Subject ascertainment, recruitment, biological material collection and collection of clinical and lifestyle information was organized and carried out by S.H.V., J.A.W., A.J.G., G.W.V., D.T.B., S.C.S., A.C., F.E., J.H.B., C.D.H., P.J.d.V., C.R., P.R., E.G., K. Koppova, P.V., S. Polidoro, S.G., C.S., M.C., D.P., C.A., M.P.Z., E.K., B.S.G., J.I.S.-V., M.S.-Z., G.V., M.D.G.-P., J.G.H., M.B., H.D., R.A.O., L.H.v.d.B., K.A., K. Kristjansson, G.G., S.N., V.P., N.A.M., A.L., M.A.v.E., S. Porru, F.B., K.G., J.I.M., R.K., G. Matullo, G.S., A.E.K., K.K.H.A., T.T., E.J. and M.A.K. Principal investigators for the UBC follow-up populations were A.E.K. (UK), G. Matullo and P.V. (Torino), S. Porru (Brescia), M.P.Z. and F.B. (Belgium), R.K. (Eastern Europe), J.I.M. (Spain), G.S. (Sweden), K.G. (Germany) and L.A.K. (The Netherlands, group 2). Genotyping and laboratory experiments were carried out by A.S., S.N.S., J.G., J.K., H.B., S.T.P., O.B.S. and C.D.H. Bioinformatics analysis was carried out by P.S., A.S., G.T., C.Z. and S.A.G. L.A.K., P.S., U.T., M.A.K., T.R. and K.S. drafted the manuscript. All authors contributed to the final version of the paper.

METHODS Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

COMPETING FINANCIAL INTERESTS The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics/.

Europe PMC Funders Author Manuscripts

Medicine, St. James's University Hospital, Leeds, UK. [7]Christie Hospital National Health Service Foundation Trust, Manchester, UK. [8]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [9]National Institute of Environmental Health, Budapest, Hungary. [10]Environmental Health Centre, Cluj-Napoca, Romania. [11]State Health Institute, Banska Bystrica, Slovakia. [12]Institute for Scientific Interchange (ISI) Foundation, Torino, Italy. [13]Department of Epidemiology and Public Health, Imperial College, London, UK. [14]Department of Genetics, Biology and Biochemistry, University of Torino, Torino, Italy. [15]Unit of Cancer Epidemiology, University of Torino, Torino, Italy. [16]Centre for Cancer Epidemiology and Prevention (CPO Piemonte), Torino, Italy. [17]Department of Experimental and Applied Medicine, Section of Occupational Medicine and Industrial Hygiene, University of Brescia, Brescia, Italy. [18]Unit of Genetic Epidemiology, Department of Public Health and Epidemiology, University of Birmingham, Birmingham, UK. [19]Department of Complex Genetics, Cluster of Genetics and Cell Biology, Nutrition and Toxicology Research Institute, Maastricht University, Maastricht, The Netherlands. [20]Leuven University Centre for Cancer Prevention (LUCK), Leuven, Belgium. [21]Department of Medicine, University of Zaragoza, Zaragoza, Spain. [22]Division of Urology, San Jorge Hospital, Huesca, Spain. [23]Division of Urology, Hospital Clinico, Zaragoza, Spain. [24]Department of Urology, University of Zaragoza School of Medicine, Zaragoza, Spain. [25]Division of Surgical Pathology, San Jorge Hospital, Huesca, Spain. [26]Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany. [27]Department of Urology, Paul Gerhardt Foundation, Lutherstadt Wittenberg, Germany. [28]Department of Medical Genetics, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands. [29]University of California Los Angeles Center for Neurobehavioral Genetics, Los Angeles, USA. [30]Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands. [31]Department of Medical Oncology, Landspitali–University Hospital, Reykjavik, Iceland. [32]Department of Urology, Landspitali–University Hospital, Reykjavik, Iceland. [33]Department of Pathology, Landspitali–University Hospital, Reykjavik, Iceland. [34]Department of Urology, Zonguldak Karaelmas University, Zonguldak, Turkey. [35]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [36]Department of General Practice, Catholic University of Leuven, Leuven, Belgium. [37]Department of General Practice, Maastricht University, Maastricht, The Netherlands. [38]Division of Medical Oncology, University of Zaragoza, Zaragoza, Spain. [39]Division of Molecular Genetic Epidemiology, German Cancer Research Centre, Heidelberg, Germany. [40]Section of Clinical Cancer Epidemiology, Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden. [41]Division of Clinical Cancer Epidemiology, Department of Oncology, Institute of Clinical Sciences, The Sahlgrenska Academy, Gothenburg, Sweden. [42]Gray Institute for Radiation Oncology and Biology, University of Oxford, Oxford, UK. [43]Faculty of Medicine, University of Iceland, Reykjavik, Iceland.

Abstract

Previously, we reported germline DNA variants associated with risk of urinary bladder cancer (UBC) in Dutch and Icelandic subjects. Here we expanded the Icelandic sample set and tested the top 20 markers from the combined analysis in several European case-control sample sets, with a total of 4,739 cases and 45,549 controls. The T allele of rs798766 on 4p16.3 was found to associate with UBC (odds ratio = 1.24, $P = 9.9 \times 10^{-12}$). rs798766 is located in an intron of *TACC3*, 70 kb from *FGFR3*, which often harbors activating somatic mutations in low-grade, noninvasive UBC. Notably, rs798766[T] shows stronger association with low-grade and low-stage UBC than with more aggressive forms of the disease and is associated with higher risk of recurrence in low-grade stage Ta tumors. The frequency of rs798766[T] is higher in Ta tumors that carry an activating mutation in *FGFR3* than in Ta tumors with wild-type *FGFR3*. Our results show a link between germline variants, somatic mutations of *FGFR3* and risk of UBC.

UBC is estimated to be the 9th most common cancer world wide and the 13th most common cause of death from cancer1. About 90% of UBCs arise in the urothelium, and clinical and molecular evidence suggests that the disease can develop along at least two distinct pathways. About half of all UBCs are low-grade tumors characterized by activating somatic mutations in *FGFR3* (encoding fibroblast growth factor receptor 3) and by a favorable prognosis despite a high risk of local recurrence2. The other half of all UBCs present with high grade and/or high stage at diagnosis and progress to muscle invasion or metastases in up to 50% of cases.

Although cigarette smoking and occupational exposure to specific carcinogens are the strongest known risk factors for UBC, familial clustering of UBC cases suggests that there is also a genetic component to risk3,4. Recently, genome-wide association (GWA) studies by us and others have led to the discovery of four common sequence variants that associate with UBC[5–7]. These variants are located at 8q24.21 (30 kb upstream from *MYC*), 3q28 (near *TP63*), 8q24.3 (in *PSCA*) and 5p15.33 (at a locus containing *TERT*).

To search for additional variants affecting risk of UBC, we analyzed an extended GWA dataset composed of a total of 611 Icelandic affected individuals (cases) and 37,478 Icelandic controls, and 1,278 Dutch cases and 1,832 Dutch controls, genotyped on HumanHap300 or HumanCNV370-Duo BeadChips. After removing SNPs that failed quality-control checks, we tested 304,073 SNPs for association with UBC. No SNPs reached our genome-wide significance threshold ($P < 1.6 \times 10^{-7}$; calculated by 0.05/304,073 total SNPs). However, the most significant marker in this analysis was the previously reported marker at 8q24.21, rs9642880 (odds ratio (OR) = 1.22, $P = 2.5 \times 10^{-7}$)5. The results for the top 20 SNPs in the combined Dutch and Icelandic analysis are presented in Supplementary Table 1. The previously reported variant on 3q28 was also among the top 20 markers, but those on 8q24.3 or 5p15.33 were not.

Of the top 20 markers, we previously attempted to replicate eight of them5. We proceeded to genotype the remaining 12 markers in at least five of nine additional UBC case-control sample sets from Italy, the UK, Spain, Sweden, Belgium, Germany, Eastern Europe and The Netherlands (Supplementary Tables 2 and 3). Of these 12 follow-up markers, one, rs798766 on chromosome 4p16.3 (Table 1), replicated in the combined follow-up groups ($P = 8.5 \times 10^{-8}$) and reached genome-wide significance in the overall analysis of the discovery and follow-up groups (OR = 1.24, 95% CI 1.17–1.32, $P = 9.9 \times 10^{-12}$). To further investigate the association signal, we imputed all HapMap CEU (phase II) SNPs in a 1-Mb region centered on rs798766 for the Icelandic and Dutch discovery samples. Although rs798766 is strongly correlated ($r^2 > 0.8$) with 20 other HapMap CEU SNPs in the region (Supplementary Table 4), none of those SNPs showed stronger association with UBC than rs798766. After conditioning on rs798766 and taking into account the number of tests performed, no further SNPs were significantly associated with UBC. When we compared the full model to the multiplicative model, we could not reject the multiplicative model (combined $P = 0.91$). Relative to noncarriers, the ORs for heterozygous and homozygous carriers of the T risk allele were estimated to be 1.24 and 1.54, respectively. The estimated population attributable risk was 9%.

rs798766 is located in intron 5 of *TACC3* (encoding transforming, acidic coiled-coil containing protein 3), a member of a family of *TACC* genes that play a role in regulating microtubule dynamics. *TACC3* shares a linkage disequilibrium (LD) block with two other genes, *TMEM129* (encoding transmembrane protein 129 isoform b) and *SLBP* (encoding stem-loop (histone) binding protein) (Fig. 1). Neither *TMEM129*, *SLBP* nor *TACC3* has been previously implicated in UBC. In the adjacent LD block, centromeric to *TACC3* and 70 kb away from rs798766, is *FGFR3*, which encodes fibroblast growth factor receptor 3.

This gene has a strong link with UBC, as activating somatic mutations in *FGFR3* are the most common mutations in low-grade bladder tumors that do not invade the submucosa (stage Ta, grade G1 (TaG1))8,9, where they have been found in 84% of cases and in 74% of stage-Ta UBC tumors overall10. In TaG1 tumors, *FGFR3* mutations are associated with an increased risk of local recurrence, and in Ta and T1 tumors overall, *FGFR3* mutations are associated with a reduced risk of progression11. This strong link with a particular clinical behavior of UBC tumors prompted us to investigate whether the frequency of rs798766[T] differs between clinical subtypes of UBC. For nine of the sample sets, we had information on stage and grade at diagnosis, which we used to classify the UBC cases into those with a predicted 'low risk' of progression (tumors confined to the bladder mucosa and not poorly differentiated) or those with a predicted 'high risk' of progression (tumor invasion in or beyond the lamina propria or poorly differentiated) (see Online Methods). In the combined analysis, the frequency of rs798766[T] was significantly higher in cases with low risk of progression than in those with high risk of progression (combined OR = 1.17, $P = 0.009$; Supplementary Table 5). By regressing on the age at diagnosis, we found that rs798766[T] associated significantly with a younger age at diagnosis of UBC (effect = −0.81 years per allele, $P = 0.0036$; Supplementary Table 6).

Detailed clinical follow-up information was available for the Dutch subjects, allowing us to evaluate the effect of rs798766 on the observed risk of recurrence in 1,109 cases with non–muscle invasive bladder cancer. The risk of recurrence was higher among cases with one copy (hazard ratio (HR) = 1.20, $P = 0.046$) or two copies of rs798766[T] (HR = 1.12, $P = 0.58$) than in noncarriers of the allele (Supplementary Table 7). The rs798766[T] allelic HR was estimated to be 1.13 ($P = 0.090$). When stratified by the diagnostic criteria used to classify low or high risk of progression, a significant difference in risk of recurrence was observed based on rs798766 genotypes among cases with predicted low risk of progression (log-rank test, $P = 0.044$) (Supplementary Fig. 1a and Supplementary Table 7). The corresponding HR per allele among cases with low risk of progression was 1.23 ($P = 0.019$). No significant difference in risk of recurrence was observed based on rs798766 genotypes among cases with predicted high risk of progression (log-rank test: $P = 0.882$; Supplementary Fig. 1b).

Because smoking is a strong risk factor for UBC and because variants in the sequence of the genome that affect smoking behavior have been previously identified, we tested whether rs798766[T] was associated with smoking behavior by using genotypes and information on smoking from Icelandic subjects (15,310 ever smokers and 6,077 never smokers)12. No association was seen between rs798766 and smoking initiation ($P = 0.24$) or smoking quantity as measured by number of cigarettes smoked per day ($P = 0.57$). To search for potential interactions between smoking and rs798766, we compared the frequency of rs798766[T] in smokers and nonsmokers among UBC cases (2,505 ever smokers and 415 never smokers). We did not observe a significant difference between the two groups (combined OR = 1.15, 95% CI 0.95–1.40, $P = 0.16$). Notably, previous studies have shown that neither smoking nor occupational exposure to polycyclic aromatic hydrocarbons seem to influence either the frequency or the pattern of *FGFR3* mutations13,14.

To test if rs798766 was tagging a mutation in the coding or regulatory parts of *TACC3* and *FGFR3*, we sequenced all the exons, flanking regions and promoters of both *TACC3* and *FGFR3* in 184 UBC cases and 184 controls. We identified ten missense variants that occurred more than twice: five in *TACC3* and five in *FGFR3* (Supplementary Table 8). The genotypes of these variants were imputed into 561 UBC cases and 35,684 controls genotyped in the Icelandic study group by using a method of long-range phasing and haplotyping15,16. None of these variants showed a significant association with UBC in Iceland, and rs798766 remained significant after adjustment for each of the ten missense

variants (all $P < 0.005$). Because rs798766 resides in a noncoding region, it is possible that it may be associated with differential expression of nearby genes. Although it would be feasible to assess expression in normal urothelial biopsies, these are currently not available in the large numbers that would be needed to detect small differences. We assessed *FGFR3* expression in 27 low-passage normal human urothelial cell strains using quantitative real-time RT-PCR. No significant association between rs798766 and *FGFR3* expression was detected. However, our previous experiments using cultured normal urothelial cells showed that levels of *FGFR3* mRNA fall markedly when cells are placed in culture but rise again at confluence and with senescence17. Therefore, these cells may not be suitable to assess potential association between genotype and expression. We also examined 47 bladder-cancer cell lines and found no relationship between rs798766 genotype and *FGFR3* expression or mutation status. However, all of these 47 lines were derived from muscle-invasive tumors, and currently, there are no cell lines available that are representative of the papillary noninvasive bladder tumor group that shows the strongest association with rs798766 genotype.

As we did not have expression data from normal urothelial tissue, we attempted to search for correlation between rs798766 and expression of neighboring genes measured in high-quality samples from normal (healthy) blood ($n = 747$) and adipose tissue ($n = 606$)18. First, we examined whether rs798766 is associated with expression of any of the 15 transcripts located in a 1-Mb window centered on rs798766 (Supplementary Table 9). In adipose tissue, we observed a significant correlation between rs798766[T] and increased expression of *FGFR3*, *TACC3* and *TMEM129* (Fig. 2 and Supplementary Table 10). The expression of *FGFR3* in adipose tissue was increased by an estimated 22.4% with each T allele carried, whereas the expression of *TACC3* and *TMEM129* were increased by an estimated 9.1% and 3.6%, respectively (Fig. 2). Expression of *TACC3* was also significantly associated with rs798766[T] in blood ($P = 1.0 \times 10^{-11}$), but the expression of *FGFR3* in blood was too low to be reliably measured. When compared to the other 449 variants (directly typed or imputed) in the 1-Mb region centered on rs798766, rs798766 showed the strongest association with *FGFR3* expression in adipose tissue, whereas rs2236786 (r^2 with rs798766 in HapMap CEU = 1.00) had the strongest correlation with *TACC3* expression (Supplementary Table 10).

We used immunohistochemistry (IHC) to assess the level of FGFR3 protein in Ta tumors, scoring the tumors as having low (score 0 or 1) or high (score 2 or 3) levels of FGFR3 (Supplementary Fig. 2)19. Cases with high levels of FGFR3 in their tumors ($n = 28$) had a higher frequency of rs798766[T] than cases with low levels ($n = 33$); the frequency of the T-allele in tumors with high levels of FGFR3 was 0.30 compared to 0.17 in those with low levels (OR 2.18, $P = 0.073$). When the correlation between the number of rs798766[T] alleles and IHC class (0, 1, 2 or 3) was tested, we observed an increase of 0.42 in protein score per copy of rs798766[T] ($P = 0.026$).

Activating mutations in *FGFR3* are present in about 30% of all bladder cancers20 and have the highest frequency in Ta tumors (74%)10. To determine if there is an association between the risk variant and somatic mutation in *FGFR3*, we tested the frequency of rs798766[T] in stage-Ta tumors with ($n = 57$) or without ($n = 33$) mutation in the gene (Supplementary Table 11). The frequency of the T allele of rs798766 was significantly higher in Ta tumors with *FGFR3* mutations (frequency = 0.31) than in those without (frequency = 0.14) (allelic OR = 2.81, 95% CI 1.21–6.51, $P = 0.016$).

In summary, we have discovered an association between rs798766[T] at the *TACC3-FGFR3* locus and increased risk of UBC and UBC recurrence. The variant shows a stronger association with tumors that have a low risk of progression than with tumors that have a

higher risk of progression. Notably, the variant also associates with increased expression of *FGFR3* and *TACC3* in adipose tissue and FGFR3 protein level and somatic mutations of the *FGFR3* gene in Ta bladder tumors. Intriguingly, recent studies of UBC showed that some microRNAs that are commonly downregulated in low-risk tumors are negative regulators of *FGFR3*, and that the resulting upregulation of *FGFR3* expression appears to precede mutations in the gene21. Our observations are compatible with this model. Thus, it can be speculated that increased expression of *FGFR3*, caused either by germline variation or by downregulation of microRNAs targeting *FGFR3* RNA, may lead to carcinogenesis in at least two ways. First, increased production of FGFR3 protein may increase the rate of urothelial proliferation and thus facilitate the accumulation of mutations that contribute to UBC development. Second, increased transcription may increase the chance of mutation in the *FGFR3* gene itself. Further work will be needed to elucidate the link between germline variation at the *TACC3-FGFR3* locus and urinary bladder carcinogenesis.

ONLINE METHODs

Subjects

Detailed information on all case-control sample sets are found in the Supplementary Note.

Genotyping

Samples from Iceland and The Netherlands were assayed with either Infinium HumanHap300 or HumanCNV370-Duo SNP chips (Illumina). The analysis was restricted to 304,073 SNPs that passed quality filters and were deemed usable due to yield, Hardy-Weinberg equilibrium and consistency in genotype frequencies between the two arrays. All samples had call rates above 98%. All follow-up genotyping was carried out applying single-track Centaurus assays (Nanogen). The quality of the Centaurus SNP assays was evaluated by genotyping each assay in the CEU HapMap samples and comparing the results with the publicly released HapMap data. Assays with >1.5% mismatch rate were not used, and an LD test was used for SNPs known to be in LD. The concordance rate of genotypes derived from the two genotyping platforms was >99.5%.

Statistical analysis

A likelihood procedure described in a previous publication and implemented in the NEMO software developed at deCODE genetics was used for the association analyses22. We tested the association of each allele with UBC using a standard likelihood-ratio statistic that, if the subjects were unrelated, would have asymptotically a χ^2 distribution with one degree of freedom under the null hypothesis. Allelic frequencies, rather than carrier frequencies, are presented for the markers in the main text. Allele-specific ORs and associated *P* values were calculated assuming a multiplicative model for the two chromosomes of an individual23. Results from multiple case-control groups were combined using a Mantel-Haenszel model in which the groups were allowed to have different population frequencies for alleles and genotypes but were assumed to have common relative risks24. To adjust for possible population stratification and relatedness among individuals, we divided the χ^2 test statistics from the individual scans using the method of genomic control25, that is, the 304,073 χ^2 test statistics were divided by their means, which were 1.040 and 1.075 for Iceland and The Netherlands, respectively. Tests of heterogeneity were performed by comparing the null hypothesis of the effect being the same in all populations to the alternative hypothesis of one or more populations having different effects using a likelihood-ratio test. I^2 takes values between 0% and 100% and is the proportion of the total variation in the estimates that is due to heterogeneity.

Analysis of RNA expression

Expression levels of 15 transcripts in a 1-Mb region at the 4p16.3 locus were analyzed in whole blood and adipose tissue from 747 and 606 unaffected Icelandic individuals, respectively, and correlated with rs798766 genotype status. Isolation of RNA and hybridization to Agilent Technologies Human 25K microarrays has been described previously18. Expression differences between two samples were quantified as the MLR compared to a reference pool RNA sample. A list of the transcripts and probes is provided in Supplementary Table 7. The hybridizations went through the standard quality-control process, which included assessment of signal-to-noise ratio, reproducibility and accuracy at spike-in compounds. The correlation between genotype and MLR for the transcripts was tested by regressing the MLRs on the number of copies of the at-risk T allele of rs798766, adjusting for age, sex, the interaction of age × sex and, for whole blood, the differential blood cell count. All *P* values were adjusted for relatedness of the individuals by simulating genotypes through Icelandic genealogy as previously described26. We also tested the association between expression of *FGFR3*, *TACC3* and *TMEM129* and the other 449 SNPs in the 1-Mb region at 4p16.3 (directly typed or imputed) to test if any of them showed a stronger correlation to gene expression than rs798766. The array probes for *FGFR3* (NM_000142), *TACC3* (NM_006342) and *TMEM129* (NM_138385) were in the 3′ untranslated regions of the genes. Expression of *FGFR3* in low passage normal human urothelial cell strains was assessed by quantitative real-time RT-PCR as described17.

Classification of low risk and high risk of progression

Based on cancer stage and grade information, all cases were classified with regard to risk of progression. Subjects with low risk of progression were defined as those having TNM stage pTa in combination with World Health Organization (WHO) 1973 differentiation grade 1 or 2 or WHO/International Society of Urological Pathology (ISUP) 2004 low grade27. All other patients were classified as having tumors with high risk of progression (stage pTis or ≥ pT1 or WHO 1973 grade 3 or WHO/ISUP 2004 high grade).

Association between genotype and recurrence-free survival

Data on bladder cancer recurrence in the discovery group from The Netherlands were collected using the automated national Dutch archive of pathology data (PALGA). The date of first recurrence was defined as the first date after transurethral resection of tumor (TURT) of the primary tumor at which another TURT or biopsy material revealed urothelial malignancy. However, if the first TURT or biopsy during follow-up took place within 3 months after primary diagnosis, a positive result was considered residual disease instead of recurrence. In case of primary carcinoma in situ (CIS), the date of first recurrence was defined as the first date at which TURT or biopsy material revealed urothelial malignancy after at least one negative cytology or histology evaluation. In the case of no recurrence, follow up was censored at the last date of negative biopsies, cytology or bladder wash. The median follow-time until recurrence was 2.6 years. Kaplan-Meier survival analyses and univariable proportional hazard regression were performed and log-rank tests were calculated to study the association between rs798766 SNP genotype and recurrence-free survival.

Tumor samples and DNA isolation

The study was approved by the Leeds research ethics committee (Leeds East project number 156/99), and informed consent was obtained from all subjects. Cold-cup biopsies of urothelial carcinoma were collected at the time of transurethral resection, snap frozen and stored in liquid nitrogen. The remainder of the tumor sample was embedded in paraffin for diagnostic assessment. Samples (Supplementary Table 10) were graded and staged using the

1973 WHO and TNM criteria, respectively8,9. All samples were transitional cell carcinoma. Frozen sections were inspected for at least 70% tumor purity. DNA was extracted using the QIAamp DNA mini kit (Qiagen). Venous blood samples were collected in EDTA tubes, and DNA was extracted using a Nucleon DNA extraction kit (Nucleon Biosciences) or by a salt precipitation method.

FGFR3 mutation analysis

Blood samples from 90 subjects with Ta tumors were analyzed for rs798766 genotype using a TaqMan assay (Applied Biosystems). DNA extracted from fresh-frozen biopsies of tumors was assessed for *FGFR3* mutation status using SNaPshot analysis to detect all common hotspot mutations28. Exons 7, 10 and 15 of *FGFR3* were amplified in a multiplex reaction; the primer sequences we used are listed in Supplementary Table 12. PCR products were checked for quality and yield by agarose gel electrophoresis and the remainder of the PCR products were treated with 3 units of shrimp alkaline phosphatase and 2 units of exonuclease I. SNaPshot analysis was performed using an Applied Biosystems SNaPshot Multiplex Kit. Reactions (9 μl total volume) contained 2.5 μl of SNaPshot Ready Multiplex Ready Reaction Mix, 1 × BigDye sequencing buffer, 1 μl of probe mix (final concentrations determined empirically for each probe batch) and 1 μl of multiplex PCR product. Extension reactions consisted of 35 cycles of denaturation at 95 °C for 10 s followed by annealing and extension at 58.5 °C for 40 s. Labeled extension products were treated with shrimp alkaline phosphatase (1 unit per sample), diluted 1 in 10, and 1 μl of the diluted product mixed with 9.8 μl of highly deionized (HiDi) formamide and 0.2 μl of Genescan-120 LIZ size standard. Products were denatured and separated using an ABI PRISM 3100 Genetic Analyzer and analysis on the final products was performed using GeneMapper 3.7 software (Applied Biosystems).

IHC

FGFR3 protein expression was assessed by IHC on 61 tumors. Five-micrometer deparaffinized and rehydrated sections of the tumors were treated with 3% hydrogen peroxide (Sigma), microwaved for 20 min and blocked with an Avidin Biotin blocking kit (Vector Laboratories). A primary antibody (FGFR3 B9, Santa Cruz) was applied to the sample for 1 h at room temperature (16–24 °C), and its presence was then detected with a biotinylated secondary antibody and 3,3-diaminobenzidine. Slides were counterstained with hematoxylin, dehydrated and mounted. Antibody specificity was confirmed using sections from tumors with known high or low RNA expression levels of *FGFR3* measured previously by real-time RT-PCR. A semiquantitative scoring system was adopted with the following designations: 0, all tumor cells negative; 1, faint but detectable positivity in some or all tumor cells; 2, weak but extensive tumor positivity; 3, strong tumor positivity (regardless of extent)19. A representative example of IHC staining is provided in Supplementary Figure 2. In addition to negative controls, sections of ureter (score 1) and a tumor with known high-level expression (score 3) were included as reference sections.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the individuals who participated in the study and whose contribution made this work possible. We also thank the nurses at deCODE's recruitment center and the personnel at the deCODE core facilities. We acknowledge the Icelandic Cancer Registry for assistance in the ascertainment of the Icelandic UBC cases. C.Z. and S.B. are funded by a FP7-MC-IAPP Grant agreement no. 218071 (CancerGene). Collection of samples and data in Iceland and The Netherlands was funded in part by the European Commission (POLYGENE: LSHC-CT-2005) and a

research investment grant of the Radboud University Nijmegen Medical Centre. Control samples for the Dutch follow-up group were genotyped with generous support from the 'Prinses Beatrix Fonds', VSB Fonds, H. Kersten and M. Kersten (Kersten Foundation), The Netherlands ALS Foundation, J.R. van Dijk and the Adessium foundation. The controls from the Dutch Schizophrenia GWA study were genotyped with the support of the US National Institute of Mental Health (R.A.O.). The Leeds Bladder Cancer Study was funded by Cancer Research UK and Yorkshire Cancer Research. The Torino Bladder Cancer Case Control Study was supported by a grant to ECNIS (Environmental Cancer Risk, Nutrition and Individual Susceptibility), a network of excellence operating within the European Union Sixth Framework Program, Priority 5: 'Food Quality and Safety' (Contract No. 513943), and by grants of the Compagnia di San Paolo, of the Italian Association for Cancer Research and of the Piedmont Region Progetti de Ricerca Sanitaria Finalizzata, Italy. The Belgian case-control study on bladder cancer risk was supported by a grant of the Flemish government, the government of the Belgian province of Limburg and the Limburg Cancer Fund.

References

1. Parkin DM. The global burden of urinary bladder cancer. Scand. J. Urol. Nephrol. Suppl. 2008:12–20. [PubMed: 19054893]
2. Knowles MA. Molecular subtypes of bladder cancer: Jekyll and Hyde or chalk and cheese? Carcinogenesis. 2006; 27:361–373. [PubMed: 16352616]
3. Murta-Nascimento C, et al. Risk of bladder cancer associated with family history of cancer: do low-penetrance polymorphisms account for the increase in risk? Cancer Epidemiol. Biomarkers Prev. 2007; 16:1595–1600. [PubMed: 17684133]
4. Aben KK, et al. Familial aggregation of urothelial cell carcinoma. Int. J. Cancer. 2002; 98:274–278. [PubMed: 11857419]
5. Kiemeney LA, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat. Genet. 2008; 40:1307–1312. [PubMed: 18794855]
6. Wu X, et al. Genetic variation in the prostate stem cell antigen gene *PSCA* confers susceptibility to urinary bladder cancer. Nat. Genet. 2009; 41:991–995. [PubMed: 19648920]
7. Rafnar T, et al. Sequence variants at the *TERT-CLPTM1L* locus associate with many cancer types. Nat. Genet. 2009; 41:221–227. [PubMed: 19151717]
8. Mostofi, FK. Histological Typing of Urinary Bladder Tumours. World Health Organization; Geneva: 1973.
9. Union Internationale Contre le Cancer (UICC). TNM Classification of Malignant Tumors. 3rd edn. UICC; Geneva: 1978. Bladder; p. 113-117.
10. Billerey C, et al. Frequent *FGFR3* mutations in papillary non-invasive bladder (pTa) tumors. Am. J. Pathol. 2001; 158:1955–1959. [PubMed: 11395371]
11. Hernandez S, et al. Prospective study of *FGFR3* mutations as a prognostic factor in nonmuscle invasive urothelial bladder carcinomas. J. Clin. Oncol. 2006; 24:3664–3671. [PubMed: 16877735]
12. Thorgeirsson TE, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008; 452:638–642. [PubMed: 18385739]
13. Bakkar AA, et al. Occupational exposure to polycyclic aromatic hydrocarbons influenced neither the frequency nor the spectrum of *FGFR3* mutations in bladder urothelial carcinoma. Mol. Carcinog. 2010; 49:25–31. [PubMed: 19722178]
14. Wallerand H, et al. Mutations in *TP53*, but not *FGFR3*, in urothelial cell carcinoma of the bladder are influenced by smoking: contribution of exogenous versus endogenous carcinogens. Carcinogenesis. 2005; 26:177–184. [PubMed: 15347601]
15. Kong A, et al. Detection of sharing by descent, long-range phasing and haplotype imputation. Nat. Genet. 2008; 40:1068–1075. [PubMed: 19165921]
16. Kong A, et al. Parental origin of sequence variants associated with complex diseases. Nature. 2009; 462:868–874. [PubMed: 20016592]
17. Tomlinson DC, L'Hote CG, Kennedy W, Pitt E, Knowles MA. Alternative splicing of fibroblast growth factor receptor 3 produces a secreted isoform that inhibits fibroblast growth factor-induced proliferation and is repressed in urothelial carcinoma cell lines. Cancer Res. 2005; 65:10441–10449. [PubMed: 16288035]
18. Emilsson V, et al. Genetics of gene expression and its effect on disease. Nature. 2008; 452:423–428. [PubMed: 18344981]

19. Tomlinson DC, Baldo O, Harnden P, Knowles MA. FGFR3 protein expression and its relationship to mutation status and prognostic variables in bladder cancer. J. Pathol. 2007; 213:91–98. [PubMed: 17668422]
20. Cappellen D, et al. Frequent activating mutations of *FGFR3* in human bladder and cervix carcinomas. Nat. Genet. 1999; 23:18–20. [PubMed: 10471491]
21. Catto JW, et al. Distinct microRNA alterations characterize high- and low-grade bladder cancer. Cancer Res. 2009; 69:8472–8481. [PubMed: 19843843]
22. Gretarsdottir S, et al. The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. Nat. Genet. 2003; 35:131–138. [PubMed: 14517540]
23. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann. Hum. Genet. 1987; 51:227–233. [PubMed: 3500674]
24. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. 1959; 22:719–748. [PubMed: 13655060]
25. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]
26. Stefansson H, et al. A common inversion under selection in Europeans. Nat. Genet. 2005; 37:129–137. [PubMed: 15654335]
27. Epstein JI, Amin MB, Reuter VR, Mostofi FK. The World Health Organization/International Society of Urological Pathology consensus classification of urothelial (transitional cell) neoplasms of the urinary bladder. Bladder Consensus Conference Committee. Am. J. Surg. Pathol. 1998; 22:1435–1448. [PubMed: 9850170]
28. van Oers JM, et al. A simple and fast method for the simultaneous detection of nine fibroblast growth factor receptor 3 mutations in bladder cancer and voided urine. Clin. Cancer Res. 2005; 11:7743–7748. [PubMed: 16278395]

Europe PMC Funders Author Manuscripts

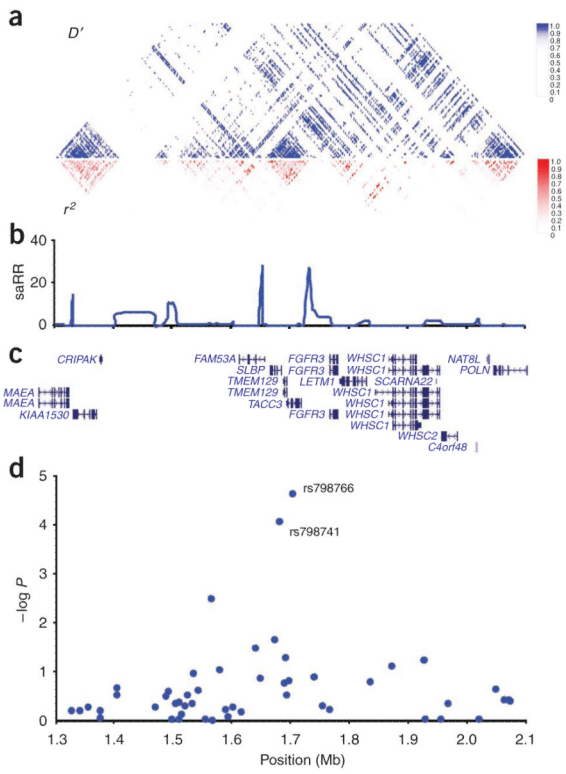

Figure 1.

Schematic view of the structure and association results in the UBC-associated region on chromosome 4p16.3. (**a**) Pairwise correlation structure in an 800-kb interval (1.3–2.1 Mb, NCBI B35) on chromosome 4p16.3. The upper plot shows pairwise D' for 292 common SNPs (with minor allele frequency > 5%) from the HapMap (v21) CEU dataset. The lower plot shows the corresponding r^2 values. (**b**) Estimated recombination rates (saRR) in cM per Mb from the HapMap (v21) Phase II data. (**c**) Location of known genes in the region. (**d**) Schematic view of the association with bladder cancer for all SNPs tested in the region for the initial scan (Iceland and The Netherlands). The y axis shows the $-\log_{10} P$ value.

Europe PMC Funders Author Manuscripts

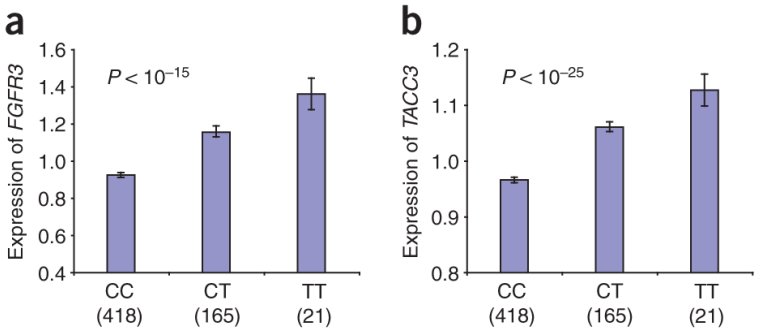

Figure 2.

Correlation between genotypes of rs798766 and the expression of *FGFR3* and *TACC3* in adipose tissue from 604 individuals. (**a**) Expression of *FGFR3*. (**b**) Expression of *TACC3*. Expression is shown as $10^{(\text{average MLR})}$, where MLR is the mean $\log_{10}$ expression ratio and the average is taken over individuals with a particular genotype. Regression of the MLR values on the number of copies of the UBC risk variant rs798766[T] shows that the expression of *FGFR3* is increased by an estimated 22.4% with each T allele carried, whereas the expression of *TACC3* is increased by an estimated 9.1% with each allele. The effect of age and sex is taken into account in the regression by including the variables age, sex and the interaction of age × sex among the explanatory variables.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Association of rs798766-T on 4p16.3 with UBC among individuals from 11 European study centers

				Cases		Controls			
Population	**OR**	**95% CI**	**P**	**n**	**Freq**	**n**	**Freq**	P_{het}	I^2
GWAS									
The Netherlands, discovery group[a]	1.19	1.06–1.35	4.8×10^{-3}	1,278	0.23	1,832	0.20		
Iceland[a]	1.26	1.09–1.45	1.4×10^{-3}	611	0.22	37,478	0.18		
GWAS combined	1.22	1.11–1.34	2.4×10^{-5}	1,889		39,310	0.19	0.56	0
Replication									
Belgium	1.06	0.77–1.46	0.73	183	0.19	375	0.18		
Germany	1.30	0.94–1.82	0.12	209	0.25	195	0.20		
Eastern Europe	1.55	1.19–2.02	1.3×10^{-3}	209	0.28	506	0.20		
Italy, Brescia	0.99	0.68–1.43	0.95	158	0.22	170	0.22		
Italy, Torino	1.22	0.94–1.57	0.14	303	0.24	369	0.21		
Spain	1.35	1.06–1.73	0.02	244	0.23	890	0.18		
Sweden	1.41	1.15–1.72	1.1×10^{-3}	344	0.25	1,194	0.19		
The United Kingdom	1.17	0.95–1.44	0.13	707	0.20	539	0.17		
The Netherlands, group 2	1.17	0.95–1.45	0.14	334	0.22	1,721	0.19		
Replication combined	1.26	1.16–1.37	8.5×10^{-8}	2,691		5,959	0.19	0.47	0
All combined[b]	1.24	1.17–1.32	9.9×10^{-12}	4,580		45,269	0.19	0.62	0

All *P* values shown are two sided. Shown are the allelic ORs with *P* values based on the multiplicative model, the corresponding numbers of cases and controls (*n*) and the allelic frequencies (Freq) of variants in affected and control individuals (freq).

[a] Results presented for Iceland and for The Netherlands discovery group were individually adjusted by the method of genomic control (see Online Methods).

[b] For the combined study populations, the reported control frequency was the average, unweighted control frequency of the individual populations, whereas the OR and the *P* value were estimated using the Mantel-Haenszel model. P_{het}, *P* value for heterogeneity.