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

Lambertus A. Kiemeney, Patrick Sulem, Soren Besenbacher, Sita H. Vermeulen, Asgeir Sigurdsson, Gudmar Thorleifsson, Daniel F Gudbjartsson, Simon N. Stacey, Julius Gudmundsson, Carlo Zanon, Jelena Kostic, Gisli Masson, Hjordis Bjarnason, Stefan T. Palsson, Oskar B. Skarphedinsson, Sigurjon A. Gudjonsson, J. Alfred Witjes, Anne J. Grotenhuis, Gerald W. Verhaegh, D. Timothy Bishop, Sei Chung Sak, Ananya Choudhury, Faye Elliott, Jennifer H. Barrett, Carolyn D. Hurst, Petra J. de Verdier, Charlotta Ryk, Peter Rudnai, Eugene Gurzau, Kvetoslava Koppova, Paolo Vineis, Silvia Polidoro, Simonetta Guarrera, Carlotta Sacerdote, Marcello Campagna, Donatella Placidi, Cecilia Arici, Maurice P. Zeegers, Eliane Kellen, Berta Saez Gutierrez, José I. Sanz-Velez, Manuel Sanchez-Zalabardo, Gabriel Valdivia, Maria D. Garcia-Prats, Jan G. Hengstler, Meinolf Blaszkewicz, Holger Dietrich, Roel A. Ophoff, Leonard H. van den Berg, Kristin Alexiusdottir, Kristleifur Kristjansson, Gudmundur Geirsson, Sigfus Nikulasson, Vigdis Petursdottir, Augustine Kong, Thorgeir Thorgeirsson, N. Aydin Mungan, Annika Lindblom, Michael A. van Es, Stefano Porru, Frank Buntinx, Klaus Golka, José I. Mayordomo, Rajiv Kumar, Giuseppe Matullo, Gunnar Steineck, Anne E. Kiltie, Katja K.H. Aben, Eirikur Jonsson, Unnur Thorsteinsdottir, Margaret A. Knowles, Thorunn Rafnar, Kari Stefansson.

Supplementary Note:

Case control sample sets used in the study:

Icelandic study population. Records of all urinary bladder cancer diagnoses were obtained from the Icelandic Cancer Registry (ICR) (http://www.krabbameinsskra.is). The ICR contains all cancer diagnoses in Iceland from January 1, 1955. The ICR contained records of 1,777 Icelandic UBC patients diagnosed until December 31, 2008, and all prevalent cases were eligible to participate. The participation rate for newly diagnosed cases was 65%. Patients were recruited by trained nurses on behalf of the patients' treating physicians, through special recruitment clinics. Participants in the study donated a blood sample and answered a lifestyle questionnaire.

Nature Genetics: doi:10.1038/ng.558

A total of 611 patients (76% males; diagnosed from December 1974 to December 2008) were included in a genome-wide SNP genotyping effort, using the Infinium II assay method and either the Sentrix HumanHap 300 or HumanCNV370-duo BeadChip (Illumina). The median age at diagnosis for all consenting cases was 68 years (range 22-95 years) as compared to 68.5 years for all UBC patients in the ICR. The 37,478 controls (41% males; mean age 61 years; SD = 21) used in this study consisted of individuals from other ongoing genome-wide association studies at deCODE and represent over 15% of the adult population of Iceland. No individual disease group is represented by more than 10% of the total control group. Cancer patients (prostate, breast, colorectal and lung) were analyzed separately, and the frequency of the sequence variants studied did not differ from other controls. Samples from prostate, breast, colorectal and lung cancer patients as well as individuals used for the analysis of smoking variables come from other ongoing project at deCODE Genetics. The study was approved by the Data Protection Authority of Iceland and the National Bioethics Committee. Written informed consent was obtained from all patients, relatives and controls. Personal identifiers associated with medical information and blood samples were encrypted with a third-party encryption system in which the Data Protection Authority maintains the code.

Dutch population

Two groups from the Netherlands were included in this study, the discovery population which was genotyped by chip and group 2 which was used for replication.

The Netherlands, discovery population. The Dutch patients were recruited for the Nijmegen Bladder Cancer Study (see http://dceg.cancer.gov/icbc/membership.html). The Nijmegen Bladder Cancer Study identified patients through the population-based regional cancer registry held by the Comprehensive Cancer Centre East, Nijmegen that serves a region of 1.3 million inhabitants

in the eastern part of the Netherlands (www.ikcnet.nl). Patients diagnosed between 1995 and 2006 under the age of 75 years were selected and their vital status and current addresses updated through the hospital information systems of the 7 community hospitals and one university hospital (Radboud University Nijmegen Medical Centre, RUNMC) that are covered by the cancer registry. All patients still alive on August 1, 2007 were invited to the study by the Comprehensive Cancer Center on behalf of the patients' treating physicians. In case of consent, patients were sent a lifestyle questionnaire to fill out and blood samples were collected by Thrombosis Service centers which hold offices in all the communities in the region. 1,651 patients were invited to participate. Of all the invitees, 1,082 gave informed consent (66%): 992 filled out the questionnaire (60%) and 1016 (62%) provided a blood sample. The number of participating patients was increased with a non-overlapping series of 376 bladder cancer patients who were recruited previously for a study on gene-environment interactions in three hospitals (RUNMC, Canisius Wilhelmina Hospital, Nijmegen, and Streekziekenhuis Midden-Twente, Hengelo, the Netherlands). Ultimately, completed questionnaires and blood samples were available for 1,276 and 1,392 patients, respectively. All the patients that were selected for the analyses (N=1,278) were of self-reported European descent. The median age at diagnosis was 62 (range 25-93) years. 82% of the participants were males. Data on tumor stage and grade were obtained through the cancer registry. The 1,832 control individuals (46% males) were cancer free and frequency-matched for age with the cases. They were recruited within a project entitled "Nijmegen Biomedical Study". The details of this study were reported previously [1]. Briefly, this is a population-based survey conducted by the Department of Epidemiology and Biostatistics and the Department of Clinical Chemistry of the Radboud University Nijmegen Medical Centre (RUNMC), in which 9,371 individuals participated from a total of 22,500 age and sex stratified, randomly selected inhabitants of Nijmegen. Control individuals from the Nijmegen Biomedical

Study were invited to participate in a study on gene-environment interactions in multifactorial diseases such as cancer. All the 1,832 participants in the present study are of self-reported European descent and were fully informed about the goals and the procedures of the study. The study protocols of the Nijmegen Bladder Cancer Study and the Nijmegen Biomedical Study were approved by the Institutional Review Board of the RUNMC and all study subjects gave written informed consent.

The Netherlands, group 2. The second group of Dutch cases (Netherlands group 2) consisted of 334 individuals from the Nijmegen Bladder Cancer Study who were diagnosed after 2006 and had not been genotyped by chip. The distributions of age at diagnosis and gender and response rates were comparable to those of the discovery group. Controls for group 2 consisted of samples that had been chip genotyped in conjunction with other studies and have been described in a previous publication[2]. Briefly, these controls were recruited from 3 sources; 1) 450 control individuals were unrelated, healthy volunteers who accompanied non-ALS patients to the UMC Utrecht neurology outpatient clinic, 2) 603 controls were recruited from an ongoing, prospective population-based study on ALS in The Netherlands and 3) 677 control individuals were included from a genome-wide association study on schizophrenia[3]. The controls were volunteers and were free of any psychiatric history. All were of Dutch descent, with at least three out of four grandparents of Dutch ancestry.

Leeds Bladder Cancer Study, United Kingdom. Details of the Leeds Bladder Cancer Study have been reported previously[4]. In brief, patients from the urology department of St James's University Hospital, Leeds were recruited from August 2002 to March 2006. All those patients attending for cystoscopy or transurethral resection of a bladder tumor (TURBT) who had previously been found, or were subsequently shown, to have urothelial cell carcinoma of the

bladder were included. Exclusion criteria were significant mental impairment or a blood transfusion in the past month. All non-Caucasians were excluded from the study leaving 764 patients. The median age at diagnosis of the patients was 73 years (range 30-101). 71% of the patients were male and 36% of all the patients had a low risk tumor (pTaG1/2). The controls were recruited from the otolaryngology outpatients and ophthalmology inpatient and outpatient departments at St James's Hospital, Leeds, from August 2002 to March 2006. All controls of appropriate age for frequency matching with the cases were approached and recruited if they gave their informed consent. As for the cases, exclusion criteria for the controls were significant mental impairment or a blood transfusion in the past month. Also, controls were excluded if they had symptoms suggestive of bladder cancer, such as haematuria. 2.8% of the controls were non-Caucasian leaving 530 Caucasian controls for the study. 71% of the controls were male. Data were collected by a health questionnaire on smoking habits and smoking history (non- ex- or current smoker, smoking dose in pack-years), occupational exposure history (to plastics, rubber, laboratories, printing, dyes and paints, diesel fumes), family history of bladder cancer, ethnicity and place of birth, and places of birth of parents. The response rate of cases was approximately 99%, that among the controls approximately 80%. Ethical approval for the study was obtained from Leeds (East) Local Research Ethics Committee, project number 02/192.

Torino Bladder Cancer Case Control Study, Italy. The source of cases for the Torino bladder cancer study are two urology departments of the main hospital in Torino, the San Giovanni Battista Hospital [5]. Cases are all Caucasian men, aged 40 to 75 years (median 63 years) and living in the Torino metropolitan area. They were newly diagnosed between 1994 and 2006 with a histologically confirmed, invasive or in situ, bladder cancer. Of all the patients with information on stage and grade, 56% were low risk (pTaG1/2). The sources of controls are

urology, medical and surgical departments of the same hospital in Torino. All controls are Caucasian men resident in the Torino metropolitan area. They were diagnosed and treated between 1994 and 2006 for benign diseases (such as prostatic hyperplasia, cystitis, hernias, heart failure, asthma, and benign ear diseases). Controls with cancer, liver or renal diseases and smoking related conditions were excluded. The median age of the controls was 57 years (range 40 to 74). Data were collected by a professional interviewer who used a structured questionnaire to interview both cases and controls face-to-face. Data collected included demographics (age, sex, ethnicity, region and education) and smoking. For cases, additional data were collected on tumor histology, tumor site, size, stage, grade, and treatment of the primary tumor. The response rates were 90% for cases and 75% for controls resulting in 328 cases and 389 controls. Ethical approval for the study was obtained from Comitato Etico Interaziendale, A.O.U. San Giovanni Batista – A.). C.T.O./ Maria Adelaide.

The Brescia bladder cancer study, Italy. The Brescia bladder cancer study is a hospital-based case-control study. The study was reported in detail previously[6]. In short, the catchment area of the cases and controls was the Province of Brescia, a highly industrialized area in Northern Italy (mainly metal and mechanical industry, construction, transport, textiles) but also with relevant agricultural areas. Cases and controls were enrolled in 1997 to 2000 from the two main city hospitals. The total number of eligible subjects was 216 cases and 220 controls. The response rate (enrolled/eligible) was 93% (N=201) for cases and 97% (N=214) for controls. Only males were included. All cases and controls had Italian nationality and were of Caucasian ethnicity. All cases had to be residents of the Province of Brescia, aged between 20 and 80, and newly diagnosed with histologically confirmed bladder cancer. The median age of the patients was 63 years (range 22-80). 29% of all the patients with known stage and grade had a low risk tumor

(pTaG1/2). Controls were patients admitted for various urological non-neoplastic diseases and were frequency matched to cases on age, hospital and period of admission. The study was formally approved by the ethical committee of the hospital where the majority of subjects were recruited. A written informed consent was obtained from all participants. Data were collected from clinical charts (tumor histology, site, grade, stage, treatments, etc.) and by means of face-to-face interviews during hospital admission, using a standardized semi-structured questionnaire. The questionnaire included data on demographics (age, ethnicity, region, education, residence, etc.), and smoking. ISCO and ISIC codes and expert assessments were used for occupational coding. Blood samples were collected from cases and controls for genotyping and DNA adducts analyses.

The Belgian Case Control Study of Bladder Cancer. The Belgian study has been reported in detail[7]. In brief, cases were selected from the Limburg Cancer Registry (LIKAR) and were approached through urologists and general practitioners. All cases were diagnosed with histologically confirmed urothelial cell carcinoma of the bladder between 1999 and 2004, and were Caucasian inhabitants of the Belgian province of Limburg. The median age of the patients was 68 years. 86% of all the patients were males. For the recruitment of controls, a request was made to the "Kruispuntbank" of the social security for simple random sampling, stratified by municipality and socio-economic status, among all citizens above 50 years of age of the province. The median age of the controls was 64 years; 59% of the controls were males. Three trained interviewers visited cases and controls at home. Information was collected through a structured interview and a standardized food frequency questionnaire. In addition, biological samples were collected. Data collected included medical history, lifetime smoking history, family history of bladder cancer and a lifetime occupational history. Informed consent was obtained from all

participants and the study was approved by the ethical review board of the Medical School of the Catholic University of Leuven, Belgium.

The Eastern Europe study population. The details of this study have been described previously[8]. Cases and controls were recruited as part of a study designed to evaluate the risk of various cancers due to environmental arsenic exposure in Hungary, Romania and Slovakia between 2002 and 2004. The recruitment was carried out in the counties of Bacs, Bekes, Csongrad and Jasz-Nagykun-Szolnok in Hungary; Bihor and Arad in Romania; and Banska Bystrica and Nitra in Slovakia. The cases (N=214) and controls (N=533) selected were of Hungarian, Romanian and Slovak nationalities. Bladder cancer patients were invited on the basis of histopathological examinations by pathologists. Hospital-based controls were included in the study, subject to fulfillment of a set of criteria. All general hospitals in the study areas were involved in the process of control recruitment. The controls were frequency matched with cases for age, gender, country of residence and ethnicity. Controls included general surgery, orthopedic and trauma patients aged 30–79 years. Patients with malignant tumors, diabetes and cardiovascular diseases were excluded as controls. The median age for the bladder cancer patients was 65 years (range 36-90). 83% of the patients were males. The median age for the controls was 61 years (range 28-83). 51% of the controls were males. The response rates among cases and controls were ~70%. Of all the patients with known stage and grade information, 28% had a low risk tumor (pTaG1/2). Clinicians took venous blood and other biological samples from cases and controls after consent forms had been signed. Cases and controls recruited to the study were interviewed by trained personnel and completed a general lifestyle questionnaire. Ethnic background for cases and controls was recorded along with other characteristics of the study population. Local ethical boards approved the study plan and design.

The Swedish Bladder Cancer Study. The Swedish patients come from a population-based study of urinary bladder cancer patients diagnosed in the Stockholm region in 1995-1996[9]. Blood samples from 352 patients were available out of a collection of 538 patients with primary urothelial carcinoma of the bladder. The average age at onset for these patients is 69 years (range 32-97 years) and 67% of the patients are males. Clinical data, including age at onset, grade and stage of tumor, were prospectively obtained from hospitals and urology units in the region. The control samples came from blood donors in the Stockholm region and were from cancer free individuals of both genders. The regional ethical committee approved of the study and all participants gave informed consent.

The Spanish bladder cancer sample set. The Spanish study patients were recruited from the Urology and Oncology Departments of Zaragoza Hospital between September 2007 and June 2009. 246 patients with histologically-proven urothelial cell carcinoma of the bladder were enrolled (response 77%). Clinical information including age at onset, grade and stage was obtained from medical records. The median age at diagnosis for the patients was 65 years (range 27 to 94) and 87% were males. The 890 Spanish control individuals were part of a larger collection of control samples obtained from individuals that had attended the University Hospital in Zaragoza, Spain, for diseases other than cancer between November 2001 and May 2007. The controls were of both genders and median age was 52 years (range 11-87). Controls were questioned to rule out prior cancers before drawing the blood sample. All patients and controls were of self-reported European descent. Study protocols were approved by the Institutional Review Board of Zaragoza University Hospital. All subjects gave written informed consent.

Lutherstadt Wittenberg bladder cancer study, Germany.

Details of the bladder cancer cases of this study have been reported previously[10,11]. In brief, 221 patients with a confirmed bladder cancer from the Department of Urology, Paul Gerhardt Foundation, Lutherstadt Wittenberg, Germany, were included. Patients were enrolled from December 1995 to January 1999. Exclusion criterion was a missing written informed consent into the study. The median age of the patients was 65 years (range 20-91); 86% of the patients were males. A total of 214 controls were from the same department of urology, but were admitted for treatment of benign urological diseases. Exclusion criteria were malignant disease in the medical history or a missing written informed consent. The median age of the controls was 68 years (range 29-91); 84% of the controls were males. Data were collected from July 2000 to May 2005. All cases and controls were Caucasians, which was confirmed by questionnaire-based documentation of nationality. Cases and controls were matched for age. Data collected in cases and controls include age, gender, a complete documentation of occupational activities performed at least for 6 months, documentation of work places with known bladder cancer risk over the entire working life, exposures to known or suspected occupational bladder carcinogens, lifetime smoking habits, family history of bladder cancer, numbers of urinary infections treated by drugs during the previous 10 years, place of birth and places of residency for more than 10 years. For bladder cancer cases, data on tumor staging, grading and treatment were taken from the records. First diagnosis of bladder cancer was recorded from July 1979 to January 1999. The local ethics committee approved the study plan and design.

REFERENCES

1. Wetzels, J.F., Kiemeney, L.A., Swinkels, D.W., Willems, H.L. & den Heijer, M. Age- and gender-specific reference values of estimated GFR in Caucasians: the Nijmegen Biomedical Study. *Kidney Int* **72**, 632-7 (2007).
2. van Es, M.A. *et al.* Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* **41**, 1083-7 (2009).
3. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744-7 (2009).
4. Sak, S.C., Barrett, J.H., Paul, A.B., Bishop, D.T. & Kiltie, A.E. The polyAT, intronic IVS11-6 and Lys939Gln XPC polymorphisms are not associated with transitional cell carcinoma of the bladder. *Br J Cancer* **92**, 2262-5 (2005).
5. Matullo, G. *et al.* Polymorphisms/haplotypes in DNA repair genes and smoking: a bladder cancer case-control study. *Cancer Epidemiol Biomarkers Prev* **14**, 2569-78 (2005).
6. Shen, M. *et al.* Polymorphisms of the DNA repair genes XRCC1, XRCC3, XPD, interaction with environmental exposures, and bladder cancer risk in a case-control study in northern Italy. *Cancer Epidemiol Biomarkers Prev* **12**, 1234-40 (2003).
7. Kellen, E., Zeegers, M., Paulussen, A., Van Dongen, M. & Buntinx, F. Fruit consumption reduces the effect of smoking on bladder cancer risk. The Belgian case control study on bladder cancer. *Int J Cancer* **118**, 2572-8 (2006).
8. Thirumaran, R.K. *et al.* Single nucleotide polymorphisms in DNA repair genes and basal cell carcinoma of skin. *Carcinogenesis* **27**, 1676-81 (2006).
9. Larsson, P. *et al.* A population-based study of 538 patients with newly detected urinary bladder neoplasms followed during 5 years. *Scand J Urol Nephrol* **37**, 195-201 (2003).
10. Golka, K. *et al.* The influence of polymorphisms of glutathione S-transferases M1 and M3 on the development of human urothelial cancer. *J Toxicol Environ Health A* **71**, 881-6 (2008).
11. Golka, K. *et al.* Susceptibility to urinary bladder cancer: relevance of rs9642880[T], GSTM1 0/0 and occupational exposure. *Pharmacogenet Genomics* (2009).

SUPPLEMENTARY TABLES

Supplementary table 1 - Top 20 markers from Genome wide association analysis of bladder cancer in the Netherlands and Iceland.

			The Netherlands discovery group						Iceland						Combined					
Chrom.	Position	SNP Allele	OR	P	Cases	Freq	Controls	Freq	OR	P	Cases	Freq	Controls	Freq	OR	CI 95	P	Phet[a]	I^{2} [b]	Prev. Rep.[c]
2	84,927,324	rs2162567 A	1.19	6.6E-03	1,277	0.22	1,832	0.20	1.28	4.6E-04	611	0.24	37,468	0.20	1.23	(1.12,1.35)	1.3E-05	0.42	0	n
3	191,079,549	rs1515490 T	1.17	1.2E-02	1,276	0.79	1,832	0.76	1.34	1.0E-04	611	0.83	37,455	0.78	1.23	(1.12,1.35)	1.0E-05	0.16	50	n
3	191,128,627	rs710521 A	1.24	3.0E-04	1,277	0.76	1,832	0.72	1.24	1.5E-03	611	0.76	37,466	0.72	1.24	(1.13,1.35)	1.5E-06	0.99	0	y
4	1,704,037	rs798766 T	1.19	4.8E-03	1,277	0.23	1,832	0.20	1.26	1.4E-03	611	0.22	37,478	0.18	1.22	(1.11,1.34)	2.4E-05	0.56	0	n
4	175,588,362	rs7664475 A	1.15	5.6E-03	1,278	0.47	1,832	0.43	1.21	1.2E-03	610	0.57	37,423	0.52	1.18	(1.09,1.27)	2.5E-05	0.53	0	n
5	82,115,972	rs1549262 A	1.15	3.4E-02	1,273	0.22	1,822	0.19	1.36	6.4E-06	602	0.28	37,064	0.22	1.24	(1.13,1.36)	3.4E-06	0.069	70	n
6	26,249,354	rs806970 T	1.33	6.7E-03	1,278	0.07	1,832	0.06	1.40	9.4E-04	610	0.10	37,471	0.07	1.37	(1.18,1.58)	2.1E-05	0.7	0	n
7	131,715,919	rs10240737 A	1.37	7.4E-05	1,278	0.90	1,832	0.86	1.20	7.4E-02	608	0.91	37,209	0.89	1.30	(1.15,1.47)	2.5E-05	0.28	13	y
8	128,787,250	rs9642880 T	1.22	1.5E-04	1,269	0.53	1,824	0.48	1.23	4.6E-04	604	0.53	37,329	0.48	1.22	(1.13,1.32)	2.5E-07	0.88	0	y
9	21,747,803	rs1335510 G	1.25	1.8E-05	1,277	0.47	1,832	0.42	1.15	2.0E-02	611	0.50	37,367	0.47	1.20	(1.11,1.30)	2.0E-06	0.28	16	n
9	21,761,241	rs1341866 G	1.24	2.6E-05	1,278	0.47	1,832	0.42	1.14	2.7E-02	611	0.50	37,477	0.46	1.20	(1.11,1.29)	3.8E-06	0.27	19	n
9	21,796,564	rs10757257 A	1.24	4.3E-05	1,278	0.47	1,831	0.41	1.13	3.6E-02	610	0.49	37,480	0.46	1.19	(1.10,1.28)	8.4E-06	0.26	22	n
10	80,536,170	rs1092116 C	1.29	2.6E-03	1,275	0.91	1,832	0.88	1.34	1.4E-03	611	0.89	37,454	0.86	1.31	(1.16,1.48)	1.2E-05	0.74	0	y
12	111,515,857	rs233722 T	1.21	2.3E-04	1,277	0.63	1,830	0.58	1.23	1.3E-03	610	0.73	37,443	0.69	1.22	(1.13,1.32)	1.0E-06	0.85	0	y
12	111,524,326	rs233716 A	1.21	3.3E-04	1,277	0.63	1,828	0.58	1.24	1.2E-03	610	0.73	37,417	0.68	1.22	(1.13,1.32)	1.4E-06	0.79	0	y
13	85,735,283	rs12584999 A	1.18	8.7E-03	1,277	0.24	1,832	0.21	1.29	3.0E-04	610	0.25	37,469	0.20	1.22	(1.12,1.34)	1.3E-05	0.33	0	y
17	16,468,520	rs8065506 C	1.17	5.0E-03	1,277	0.73	1,829	0.70	1.27	8.5E-04	608	0.80	37,315	0.76	1.21	(1.11,1.32)	1.9E-05	0.39	0	n
18	41,570,268	rs692899 G	1.16	5.7E-03	1,278	0.40	1,832	0.36	1.23	7.6E-04	609	0.38	37,440	0.33	1.19	(1.10,1.29)	1.8E-05	0.46	0	n
19	40,960,311	rs12982672 G	1.34	5.9E-04	1,277	0.91	1,831	0.88	1.28	1.1E-02	608	0.90	37,291	0.88	1.31	(1.16,1.49)	2.1E-05	0.72	0	y
X	56,444,998	rs5913935 G	1.37	2.6E-05	1,278	0.16	1,831	0.12	1.23	4.9E-02	611	0.12	37,449	0.10	1.32	(1.17,1.48)	5.1E-06	0.39	0	n

[a]Phet denotes the tests of heterogeneity performed by comparing the null hypothesis of the effect being the same in all populations to the alternative hypothesis of each population having a different effect using a likelihood ratio test.

[b]I^{2} takes values between 0% and 100% and describes the proportion of the total variation in estimates that is due to heterogeneity

[c]Prev. Rep. Indicates whether the markers had been a part of a previous replication effort described in Kiemeney et al. 2008

Supplementary table 2. Case control groups used in the study

Study Group	#cases	#controls	Average age at diagnosis (range)	% males (cases)	Study type
Discovery groups (GWA)					
Iceland	611	37,478	68 (20-95)	76	Population based
The Netherlands, discovery group	1,278	1,832	62 (25-93)	81	Population based
Follow up groups					
UK, Leeds	771	574	73 (30-101)	71	Hospital-based
Italy, Torino	332	391	63 (40-75)	100	Hospital-based
Italy, Brescia	183	193	63 (22-80)	100	Hospital-based
Belgium, Leuven	201	385	68 (40-93)	86	Population based
Eastern Europe (Hungary, Romania, Slovakia)	214	533	65 (36-90)	83	Hospital-based
Sweden, Stockholm	352	1,350	69 (32-97)	67	Population based
Spain, Zaragoza	246	890	65 (27-94)	87	Hospital-based
Germany, Lutherstadt Wittenberg	217	202	65 (20-91)	86	Hospital-based
The Netherlands, group 2	334	1,721	64 (30-91)	83	Population-based
Total	4,739	45,549			

Supplementary table 3 - Association of the top 20 markers from GWA of UBC among 9 follow-up study groups from Europe and among the discovery and follow-up groups combined.

			FOLLOW-UP COMBINED										GWA and FOLLOW UP COMBINED				
Chrom.	Position	SNP Allele	OR	95% CI	P	Phet[a]	I^{2} [b]	Cases	Freq	Controls	Freq	Centers	OR	CI 95	P[c]	Phet[a]	I^{2} [b]
2	84,927,324	rs2162567 A	1.02	(0.93,1.13)	0.66	0.71	0	2,170	0.19	3,616	0.19	7	1.13	(1.05,1.21)	0.0005	0.18	30
3	191,079,549	rs1515490 T	1.11	(1.01,1.23)	0.033	0.044	54	2,121	0.79	3,640	0.78	7	1.17	(1.10,1.26)	3.2×10^{-6}	0.028	54
3	191,128,627	rs710521 A	1.14	(1.04,1.25)	0.005	0.52	0	2,128	0.76	3,837	0.73	7	1.19	(1.12,1.27)	**6.1×10^{-8}**	0.55	0
4	1,704,037	rs798766 T	1.26	(1.16,1.37)	8.5×10^{-8}	0.47	0	2,691	0.23	5,959	0.19	9	1.24	(1.17,1.32)	**9.9×10^{-12}**	0.62	0
4	175,588,362	rs7664475 A	0.93	(0.85,1.03)	0.15	0.55	0	1,623	0.47	1,955	0.50	5	1.08	(1.01,1.14)	0.017	0.0072	66
5	82,115,972	rs1549262 A	0.93	(0.83,1.05)	0.25	0.43	0	1,594	0.18	1,998	0.19	5	1.12	(1.04,1.20)	0.0024	0.002	71
6	26,249,354	rs806970 T	1.00	(0.85,1.17)	0.96	0.1	43	2,169	0.07	3,627	0.07	7	1.18	(1.06,1.32)	0.0018	0.014	58
7	131,715,919	rs10240737 A	0.96	(0.85,1.08)	0.5	1	0	2,128	0.88	3,818	0.89	7	1.11	(1.02,1.21)	0.014	0.082	43
8	128,787,250	rs9642880 T	1.20	(1.11,1.30)	5.7×10^{-6}	0.36	9	2,091	0.49	3,780	0.44	7	1.21	(1.15,1.28)	**6.7×10^{-12}**	0.57	0
9	21,747,803	rs1335510 G	1.04	(0.96,1.12)	0.38	0.92	0	2,386	0.41	2,919	0.41	8	1.12	(1.06,1.12)	5.5×10^{-5}	0.28	18
9	21,761,241	rs1341866 G	1.04	(0.96,1.12)	0.34	0.89	0	2,401	0.41	2,938	0.41	8	1.12	(1.06,1.12)	6.9×10^{-5}	0.31	15
9	21,796,564	rs10757257 A	1.05	(0.95,1.16)	0.33	0.91	0	1,591	0.40	2,021	0.39	5	1.13	(1.07,1.20)	4.0×10^{-5}	0.4	4
10	80,536,170	rs1092116 C	0.98	(0.87,1.11)	0.81	0.52	0	2,105	0.88	3,780	0.89	7	1.14	(1.05,1.24)	0.0031	0.045	49
12	111,515,857	rs233722 T	0.97	(0.89,1.05)	0.4	0.15	37	1,994	0.57	3,822	0.59	7	1.09	(1.03,1.15)	0.0035	0.0012	69
12	111,524,326	rs233716 A	0.96	(0.89,1.04)	0.35	0.39	5	2,135	0.57	3,847	0.59	7	1.08	(1.02,1.15)	0.0062	0.0032	65
13	85,735,283	rs12584999 A	0.99	(0.90,1.10)	0.87	0.98	0	2,018	0.20	3,768	0.21	7	1.11	(1.04,1.19)	0.0017	0.19	29
17	16,468,520	rs8065506 C	0.99	(0.90,1.09)	0.86	0.11	41	2,064	0.73	2,994	0.74	7	1.10	(1.03,1.18)	0.0027	0.0092	61
18	41,570,268	rs692899 G	1.03	(0.93,1.14)	0.54	0.084	51	1,579	0.36	1,986	0.35	5	1.13	(1.06,1.20)	0.00017	0.036	56
19	40,960,311	rs12982672 G	1.06	(0.92,1.22)	0.43	0.9	0	2,076	0.92	3,769	0.91	7	1.19	(1.09,1.31)	0.00022	0.49	0
X	56,444,998	rs5913935 G	1.03	(0.92,1.17)	0.58	0.39	4	1,651	0.21	1,985	0.20	5	1.17	(1.08,1.27)	0.00026	0.048	53

[a]Phet denotes the tests of heterogeneity performed by comparing the null hypothesis of the effect being the same in all populations to the alternative hypothesis of each population having a different effect using a likelihood ratio test.

[b]I^{2} takes values between 0% and 100% and describes the proportion of the total variation in estimates that is due to heterogeneity

[c]P values that reach genome-wide significance are bolded

Supplementary Table 4. LD between rs798766 and 20 HapMap CEU SNPs with the strongest correlation in the Icelandic and Dutch populations

Marker 1	Marker 2	Location	D'	R^2
rs798766	rs798727	1654935	0.94732	0.852544
rs798766	rs798726	1655009	0.947917	0.898546
rs798766	rs811316	1659175	1	0.948454
rs798766	rs798719	1667981	1	1
rs798766	rs798756	1677245	1	1
rs798766	rs798741	1682211	1	1
rs798766	rs798744	1684482	1	1
rs798766	rs2236786	1689092	1	1
rs798766	rs798751	1689370	1	1
rs798766	rs798754	1690555	1	1
rs798766	rs798755	1690622	1	1
rs798766	rs3099555	1695449	1	1
rs798766	rs1665364	1697004	1	1
rs798766	rs1665366	1698179	1	0.948454
rs798766	rs798766	1704037	NA	NA
rs798766	rs3752749	1707060	1	1
rs798766	rs4865463	1712198	1	1
rs798766	rs2166580	1712967	1	1
rs798766	rs8389	1716642	1	1
rs798766	rs744658	1724872	1	0.861111
rs798766	rs732754	1732085	1	1

Supplementary table 5 - Association of rs798766-T with low risk versus high risk tumor characteristics among bladder cancer patients from 9 European centers.

Center	OR	95% CI	P	N Low risk cases	Freq	N High risk cases	Freq	Phet	I2
The Netherlands, discovery group	1.27	(1.05-1.54)	0.011	696	0.24	612	0.20		
Iceland	1.84	(1.04-3.23)	0.035	119	0.25	58	0.16		
Germany	0.87	(0.5-1.52)	0.573	69	0.23	140	0.26		
Eastern Europe	1.17	(0.58-2.33)	0.593	38	0.32	99	0.28		
Italy, Brescia	1.31	(0.73-2.33)	0.363	45	0.26	113	0.21		
Italy, Torino	1.10	(0.69-1.75)	0.638	146	0.25	112	0.23		
Spain	1.02	(0.50-2.06)	0.965	28	0.21	137	0.21		
Sweden	0.83	(0.57-1.20)	0.293	147	0.23	194	0.26		
The United Kingdom	1.22	(0.93-1.61)	0.151	271	0.22	410	0.18		
All Combined	1.17	(1.04-1.31)	0.009	1,559	0.25	1,875	0.22	0.35	10

Supplementary table 6 - Correlation of rs798766-T with age at diagnosis of bladder cancer among patients of 10 centers.

Center	N	Effect (years)	95%CI	Freq	P	Phet	I2
Belgium	183	0.65	(-1.98,3.29)	0.19	0.626		
Germany	209	-1.39	(-3.93,1.16)	0.25	0.285		
Eastern Europe	209	-1.84	(-4.15,0.48)	0.28	0.120		
The Netherlands	1,266	-0.80	(-1.74,0.14)	0.23	0.094		
Iceland	612	-1.06	(-2.86,0.75)	0.22	0.251		
Italy, Brescia	158	-3.66	(-6.76,-0.56)	0.22	0.021		
Italy, Torino	303	-0.97	(-2.39,0.45)	0.24	0.181		
Spain	220	-0.66	(-3.51-2.19)	0.23	0.650		
Sweden	344	0.57	(-1.47,2.61)	0.25	0.582		
The United Kingdom	707	-0.48	(-1.91,0.95)	0.20	0.510		
All Combined	4,211	-0.81	(-1.35,-0.26)	0.23	3.60E-03	0.59	0

Supplementary table 7 - Results of Cox proportional hazards regression analysis for risk of recurrence.

Group		rs798766	Hazard Ratio	95% CI	P value	Recurrence event	
Median recurrence-free survival time						no	yes
All NMIBC	Genotype	CC vs.CC	1	NA	NA	351	289
2.6 years	Genotype	CT vs. CC	1.198	1.003-1.432	0.046	198	212
	Genotype	TT vs. CC	1.118	0.753-1.658	0.580	30	27
	Allele	T vs. C	1.128	0.982-1.296	0.090		
Low risk NMIBC	Genotype	CC vs.CC	1	NA	NA	222	169
2.8 years	Genotype	CT vs. CC	1.309	1.043-1.642	0.020	136	136
	Genotype	TT vs. CC	1.397	0.848-2.301	0.190	17	17
	Allele	T vs. C	1.234	1.035-1.471	0.019		
High risk NMIBC	Genotype	CC vs.CC	1	NA	NA	125	111
2.3 years	Genotype	CT vs. CC	1.019	0.745-1.393	0.906	61	63
	Genotype	TT vs. CC	0.861	0.450-1.644	0.649	12	10
	Allele	T vs. C	0.971	0.760-1.241	0.816		

Supplementary table 8 - Missense mutations in *TACC3* and *FGFR3*

Pos	SNP name	Gene	Exon	aa change	P	case (n)	case (f)[a]	control (n)	control (f)[a]	r2 with rs798766	Homozygous common[b]	Heterozygous[b]	Homozygous rare[b]
Chr4_1699354	rs34205238	TACC3	4	E143K	0.32	561	0.1631	35,675	0.1748	0.0461	246	83	7
Chr4_1699751	rs17132047	TACC3	4	C275Y	0.57	561	0.2255	35,676	0.2331	0.0662	208	108	21
Chr4_1699786	rs1063743	TACC3	4	G287S	0.59	561	0.2273	35,674	0.2343	0.0667	203	109	22
Chr4_1700177		TACC3	4	P417L	0.14	561	0.0027	35,676	0.0011	0.0003	337	3	0
Chr4_1702776	rs17680881	TACC3	6	G514E	0.52	561	0.2246	35,674	0.2331	0.0663	200	113	23
Chr4_1764874	rs17880408	FGFR3	1	G15R	0.07	561	0.0348	35,667	0.0259	0.0040	317	18	1
Chr4_1764928		FGFR3	1	A33T	0.33	561	0.0053	35,667	0.0038	0.0008	336	4	0
Chr4_1772970		FGFR3	5	R175H	0.06	561	0.0027	35,684	0.0079	0.0017	350	3	0
Chr4_1776427		FGFR3	10	P449S	1.00	561	0.0009	35,684	0.0017	0.0004	340	4	0
Chr4_1778684		FGFR3	18	G661D	0.86	561	0.0080	35,684	0.0078	0.0017	344	7	0

a Frequency estimated from the imputed data

b Genotype counts observed among sequenced individuals

Supplementary table 9 - Transcripts located in a 1 MB region centered on rs798766 on chr4p16.3

Probe	Transcript	Chr	Start	End
NM_052861	MGC21675	chr4	1236201	1236261
NM_005882	MAEA	chr4	1323819	1323878
NM_175918	FLJ34443	chr4	1379707	1379767
NM_006527	SLBP	chr4	1664454	1664514
NM_138385	TMEM129	chr4	1687477	1687537
NM_006342	TACC3	chr4	1716581	1716641
NM_000142	FGFR3	chr4	1780330	1780390
NM_012318	LETM1	chr4	1784579	1784640
NM_133334	WHSC1	chr4	1920283	1920343
NM_005663	WHSC2	chr4	1954250	1954310
Contig22095_RC	WHSC2	chr4	2008040	2008100
ENST00000290995	(null)	chr4	2012859	2012919
NM_178557	FLJ37478	chr4	2037373	2037433
NM_181808	POLN	chr4	2043558	2043618
NM_024511	C4orf15	chr4	2203369	2203429

Supplementary table 10 - Correlation between rs798766-T and expression of nearby transcripts.

					Best *cis* SNP		
Gene (Transcript)	**Tissue**	**Effect**	***P***	P_{adj}	**SNP (r^2)**	***P***	P_{adj}
FGFR3 (NM_000142)	Adipose	22.40%	8.9×10^{-16}		rs798766		
TACC3 (NM_006342)	Adipose	9.10%	1.8×10^{-26}	0.48	rs2236786 (1.00)	7.2×10^{-29}	0.0022
TACC3 (NM_006342)	Blood	7.80%	1.0×10^{-11}	0.007	rs2854915 (0.16)	2.8×10^{-32}	1.3×10^{-22}
TMEM129 (NM_138385)	Adipose	3.60%	5.5×10^{-6}	0.0054	rs1374468 (0.03)	4.3×10^{-27}	9.3×10^{-24}

The table includes the correlation between the UBC risk variant rs798766-T and the expression levels of transcripts measured in adipose tissue from 606 individuals and in whole blood from 747 individuals. The correlation is tested for 15 transcript located in a 1 Mb window centred on rs798766 by regressing the MLR values (the mean log10 expression ratio) on the number of copies of rs798766[T] an individual carries. The effect of age and sex (and differential cell count for whole blood) is taken into account by including the corresponding terms in the regression analysis. The table includes the gene and the transcript tested, the tissue, the effected measure as percentage change in the relative expression, and the P value. Only correlation tests with $P < 0.001$ are included. For each transcript the table also includes the variant, out of 449 tested variants in the 1 Mb region, that shows the most significant correlation with the expression (cisSNP), the correlation (r2) between that cisSNP and rs798766, and adjusted P values when the correlation with rs798766 is conditional on the cisSNP and vice verse. All P values have been adjusted for relatedness of the individual by the method of genomic control.

Supplementary table 11 - Characteristics of tumors used for somatic mutation analysis.

Patient	Stage	Grade	rs798766	FGFR3 mutation
1	pTa	1	CC	WT
2	pTa	1	CC	WT
3	pTa	2	CC	S249C
4	pTa	2	CC	S249C
5	pTa	2	CC	Y375C
6	pTa	2	CC	S249C
7	pTa	2	CC	S249C
8	pTa	2	CC	WT
9	pTa	2	CC	G372C
10	pTa	2	CC	WT
11	pTa	2	CC	S249C
12	pTa	2	CC	K652E
13	pTa	2	CC	S249C
14	pTa	2	CC	S249C
15	pTa	2	CC	G372C
16	pTa	2	CC	G372C
17	pTa	2	CC	S249C
18	pTa	2	CC	WT
19	pTa	2	CC	R248C
20	pTa	2	CC	S249C
21	pTa	2	CC	WT
22	pTa	2	CC	S249C
23	pTa	2	CC	S249C
24	pTa	2	CC	S249C
25	pTa	2	CC	S249C
26	pTa	2	CC	WT
27	pTa	2	CC	WT
28	pTa	2	CC	S249C
29	pTa	2	CC	WT
30	pTa	2	CC	WT
31	pTa	2	CC	S249C
32	pTa	2	CC	WT
33	pTa	3	CC	WT
34	pTa	3	CC	WT
35	pTa	3	CC	WT
36	pTa	3	CC	WT
37	pTa	3	CC	WT
38	pTa	3	CC	WT
39	pTa	3	CC	WT
40	pTa	3	CC	WT
41	pTa	3	CC	S249C
42	pTa	3	CC	WT
43	pTa	1	CC	S249C
44	pTa	1	CC	S249C
45	pTa	1	CC	WT

46	pTa	1	CC	WT
47	pTa	1	CC	WT
48	pTa	1	CC	S249C
49	pTa	1	CC	WT
50	pTa	1	CC	R248C
51	pTa	2	CC	WT
52	pTa	1	CT	S249C
53	pTa	2	CT	WT
54	pTa	2	CT	S249C
55	pTa	2	CT	S249C
56	pTa	2	CT	WT
57	pTa	2	CT	WT
58	pTa	2	CT	S249C
59	pTa	2	CT	S249C
60	pTa	2	CT	WT
61	pTa	2	CT	S249C
62	pTa	2	CT	S249C
63	pTa	2	CT	Y375C
64	pTa	2	CT	WT
65	pTa	2	CT	WT
66	pTa	2	CT	S249C
67	pTa	2	CT	S249C
68	pTa	2	CT	S249C
69	pTa	2	CT	S249C
70	pTa	2	CT	S249C
71	pTa	2	CT	Y375C
72	pTa	2	CT	S249C
73	pTa	2	CT	S249C
74	pTa	2	CT	S249C
75	pTa	3	CT	S249C
76	pTa	3	CT	S249C
77	pTa	3	CT	Y375C
78	pTa	3	CT	Y375C
79	pTa	3	CT	S249C
80	pTa	3	CT	R248C
81	pTa	1	CT	G372C
82	pTa	1	CT	S249C
83	pTa	1	CT	WT
84	pTa	2	CT	S373C
85	pTa	2	CT	R248C
86	pTa	1	TT	S249C
87	pTa	2	TT	WT
88	pTa	2	TT	S249C
89	pTa	1	TT	S249C
90	pTa	2	TT	S249C

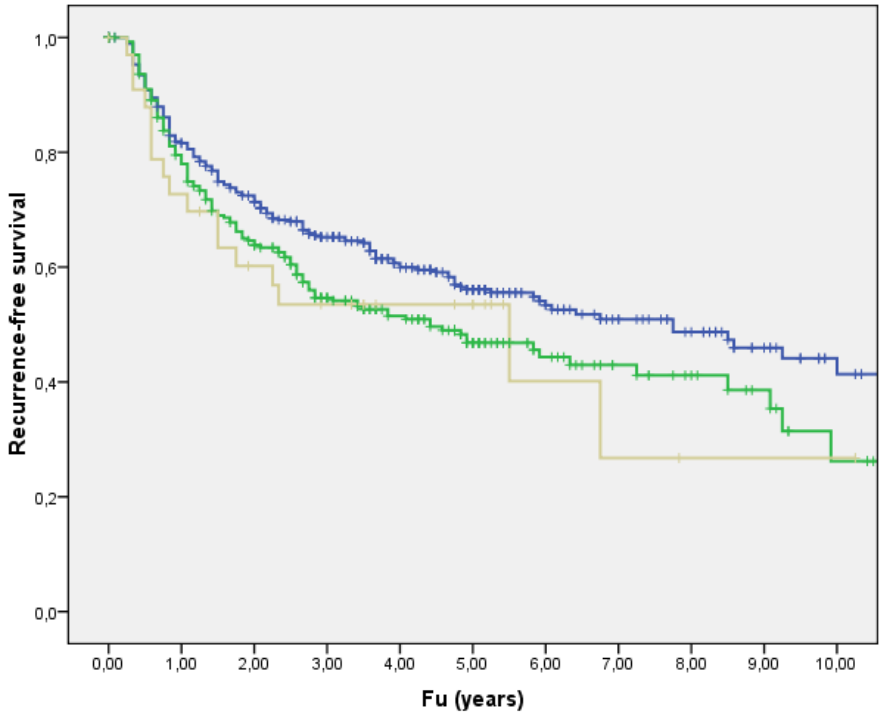
Supplementary Table 12 - Primers and probes used for *FGFR3* mutation analysis

Primers

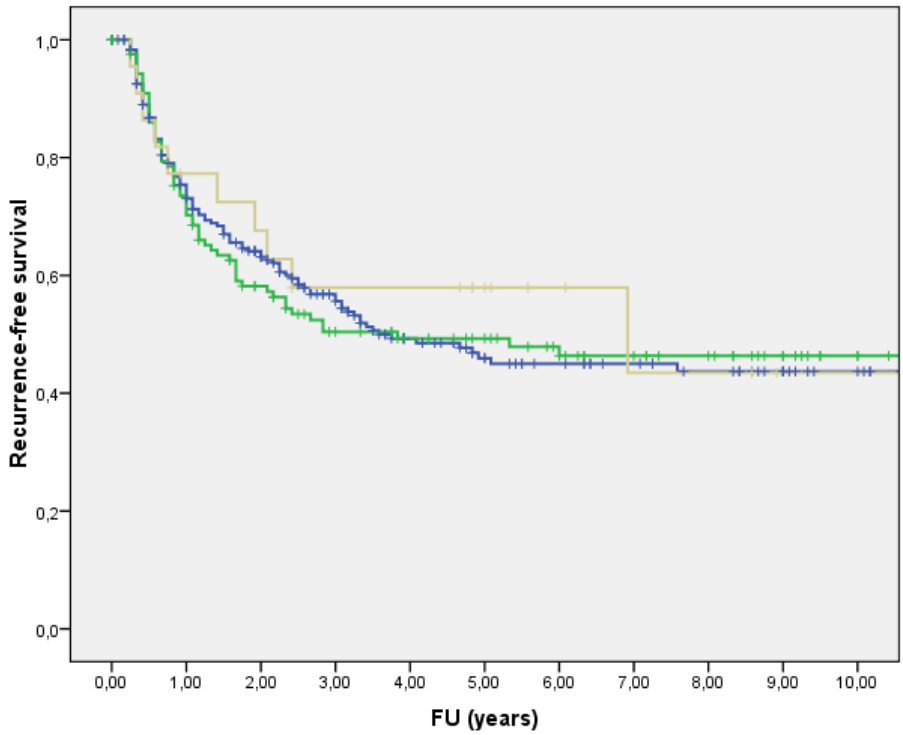
Exon	Forward primer	Reverse primer
7	5′-AGTGGCGGTGGTGGTGAGGGAG-3′	5′-GCACCGCCGTCTGGTTGG-3′
10	5′-CAACGCCCATGTCTTTGCAG-3′	5′-AGGCGGCAGAGCGTCACAG-3′;
15	5′-GACCGAGGACAACGTGATG-3′	5′-GTGTGGGAAGGCGGTGTTG-3′

Probes

Name		Probe sequence
R248C	T46	CGT CAT CTG CCC CCA CAG AG
S249C	T36	TCT GCC CCC ACA GAG CGC T
G372C	T29	GGT GGA GGC TGA CGA GGC G
Y375C	T43	ACG AGG CGG GCA GTG TGT
A393E	T34	CCT GTT CAT CCT GGT GGT GG
K652M/T	T20	CAC AAC CTC GAC TAC TAC AAG A
K652E/Q	T50	GCA CAA CCT CGA CTA CTA CAA G
S373C	T19	GAG GAT GCC TGC ATA CAC AC
G382R	T56	GAA CAG GAA GAA GCC CAC CC



(a) Log-rank test: p=0.044



(b) Log rank test: p=0.882

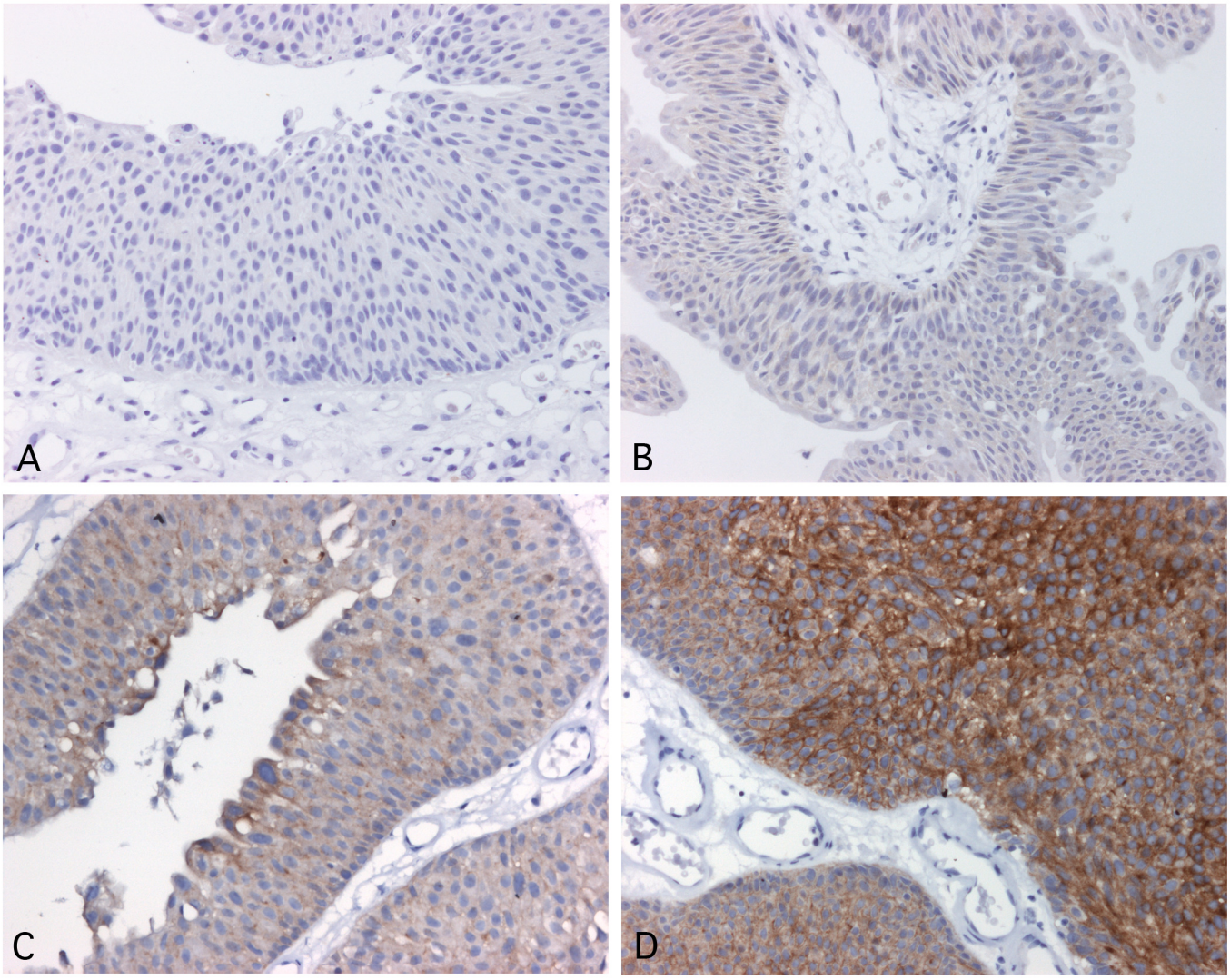
Supplementary Figure 1. Risk of recurrence of UBC by rs798766 genotype among patients with non-muscle invasive bladder cancer, stratified by predicted (a) low or (b) high risk of progression at diagnosis. Blue line represents rs798766-CC; green line represents rs798766-CT, beige line represents rs798766-TT. The corresponding recurrence event counts are displayed in the tables below.

rs798766 genotype and recurrence for patients with low predicted risk of progression

		Recurrence event		
		0 (=no)	1 (=yes)	Total
rs798766	CC	222	169	391
	CT	136	136	272
	TT	17	17	34
Total		375	322	697

rs798766 genotype and recurrence for patients with high predicted risk of progression

		Recurrence event		
		0 (=no)	1 (=yes)	Total
rs798766	CC	125	111	236
	CT	61	63	124
	TT	12	10	22
Total		198	184	382



Supplementary Figure 2. Patterns of FGFR3 expression in bladder tumors illustrating criteria used for scoring immunohistochemical staining. A. score 0; B. score 1; C. score 2; D. score 3.