A genome-wide association study identifies susceptibility loci for Wilms tumor

Clare Turnbull1, Elizabeth R Perdeaux1, David Pernet1, Arlene Naranjo2, Anthony Renwick1, Sheila Seal1, Rosa Maria Munoz-Xicola1, Sandra Hanks1, Ingrid Slade1, Anna Zachariou1, Margaret Warren-Perry1, Elise Ruark1, Mary Gerrard3, Juliet Hale4, Martin Hewitt5, Janice Kohler6, Sheila Lane7, Gill Levitt8, Mabrook Madi9, Bruce Morland10, Veronica Neefjes11, James Nicholson12, Susan Picton13, Barry Pizer14, Milind Ronghe15, Michael Stevens16, Heidi Traunecker17, Charles A Stiller18, Kathy Pritchard-Jones19, Jeffrey Dome20, Paul Grundy21,22 & Nazneen Rahman1

Wilms tumor is the most common renal malignancy of childhood. To identify common variants that confer susceptibility to Wilms tumor, we conducted a genome-wide association study in 757 individuals with Wilms tumor (cases) and 1,879 controls. We evaluated ten SNPs in regions significantly associated at \( P < 5 \times 10^{-5} \) in two independent replication series from the UK (769 cases and 2,814 controls) and the United States (719 cases and 1,037 controls). We identified clear significant associations at 2p24 (rs3755132, \( P = 1.03 \times 10^{-14} \); rs807624, \( P = 1.32 \times 10^{-14} \)) and 11q14 (rs790356, \( P = 4.25 \times 10^{-15} \)). Both regions contain genes that are plausibly related to Wilms tumorigenesis. We also identified candidate association signals at 5q14, 22q12 and Xp22.

Wilms tumor is a childhood embryonal kidney cancer that affects approximately 1 in 10,000 children in Western populations. The median age of diagnosis is between 3 and 4 years, and both kidneys are affected in ~5% of children. Histologically, Wilms tumor mirrors the development of the normal kidney and classically contains the three cell types present in the embryonic kidney: blastema, epithelia and stroma.

There is strong evidence of a genetic contribution to Wilms tumor. First, there is substantial variation in incidence that is attributable to ancestry rather than geographical location. Second, approximately 2% of cases are familial. Third, there are several predisposition syndromes associated with an increased risk of Wilms tumor, the most common of which are conditions that result from mutations in WT1 or epigenetic defects at 11p15 (ref. 3). Fewer than 5% of Wilms tumor cases are attributable to known causes, and the underlying basis of most Wilms tumor cases is unknown.

We performed a genome-wide association study (GWAS) to identify variant that confer susceptibility to Wilms tumor, using cases recruited through oncology clinics in North America. We compared the genotype frequencies in cases to those in data from North American controls obtained from the Database of Genotypes and Phenotypes (dbGaP). After quality control exclusions and removal of samples with non-European ancestry, we analyzed data on 599,255 SNPs in 757 cases and 1,879 controls. We compared genotype frequencies in cases and controls, primarily using the 1-degree-of-freedom Cochran-Armitage trend test (Supplementary Fig. 1). There was evidence of only modest inflation of the test statistics (\( \lambda = 1.08 \)), indicating that the extent to which population stratification contributed the results was limited (Supplementary Fig. 2).

We next considered 20 SNPs in 9 regions of linkage disequilibrium (LD) that were significantly associated with Wilms tumor at \( P < 5 \times 10^{-5} \). After eliminating SNPs that were strongly correlated with each other, we further evaluated these putative associations by genotyping ten SNPs at nine loci in two independent replication series. These series included 769 cases and 2,814 controls from the UK and 719 additional cases from North America. We did not have samples from US control individuals to directly genotype, so we analyzed the US control individuals to directly genotype, so we analyzed the

Received 18 November 2011; accepted 19 March 2012; published online 29 April 2012; corrected after print 6 June 2012; corrected after print 10 July 2013; doi:10.1038/ng.2251

1Division of Genetics & Epidemiology, Institute of Cancer Research, Sutton, UK. 2Children’s Oncology Group Statistics and Data Center, University of Florida, Gainesville, Florida, USA. 3Sheffield Children’s National Health Service (NHS) Trust, Sheffield, UK. 4Department of Paediatric Oncology, Royal Victoria Infirmary, Newcastle, UK. 5Department of Paediatric Oncology, University Hospital Nottingham, Nottingham, UK. 6Regional Paediatric Oncology Centre, Southampton General Hospital, Southampton, UK. 7Department of Paediatric Oncology, Oxford Children’s Hospital, John Radcliffe Hospital, Oxford, UK. 8Department of Paediatric Oncology, Great Ormond Street Hospital for Children NHS Trust, London, UK. 9Department of Paediatric Oncology, Leicester Royal Infirmary, Leicester, UK. 10Department of Paediatric Oncology, Birmingham Children’s Hospital, Birmingham, UK. 11Department of Paediatric Oncology, Royal Aberdeen Children’s Hospital, Aberdeen, UK. 12Department of Paediatric Oncology, Cambridge University Hospitals NHS Foundation Trust, Addenbrookes Hospital, Cambridge, UK. 13Paediatric Oncology Department, Leeds General Infirmary, Leeds, UK. 14Department of Paediatric Oncology, Alder Hey Children’s NHS Foundation Trust, Liverpool, UK. 15Department of Paediatric Oncology, Royal Hospital for Sick Children, Glasgow, UK. 16Department of Paediatric Oncology, Bristol Royal Hospital for Children, Bristol, UK. 17Paediatric Oncology Unit, Children’s Hospital for Wales, Cardiff, UK. 18Childhood Cancer Research Group, University of Oxford, Oxford, UK. 19Molecular Haematology and Cancer Biology Unit, University College London Institute of Child Health, London, UK. 20Division of Oncology, Children's National Medical Center, Washington, DC, USA. 21Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada. 22Department of Oncology, University of Alberta, Edmonton, Alberta, Canada. Correspondence should be addressed to N.R. (nazneen.rahman@icr.ac.uk).
Two SNPs on chromosomes 2p24 and one SNP on chromosome 11q14 showed clear evidence of replication in each of the replication series separately ($P = 4.1 \times 10^{-4}$ or better in the same direction as in the GWAS) and reached genome-wide significance levels over both stages combined of $P = 1.03 \times 10^{-14}$ and $1.32 \times 10^{-14}$ at 2p24 and $P = 4.25 \times 10^{-15}$ at 11q14 (Fig. 1, Table 1 and Supplementary Tables 1 and 2).

We also identified three low-frequency SNPs at 5q14, 22q12 and Xp22 for which associations achieved genome-wide significance with clear replication in the US series but with only weaker evidence of replication in the UK series (Table 2 and Supplementary Tables 1 and 2). Accordingly, further replication studies are required to confirm these associations, as low-frequency variants in particular are susceptible to signal artifact when using different genotyping platforms for cases and controls.

We next used imputation to evaluate whether a more strongly associated variant was present at the loci, by estimating the genotype probabilities at additional SNPs using HapMap 3 and 1000 Genomes Project data as a framework. At four of the five loci, imputation identified more strongly associated variants than the original SNP (Fig. 1 and Supplementary Table 3). In all regions, the imputed SNPs gave similar risk frequency profiles as the genotyped SNPs. Direct genotyping of these imputed SNPs in cases and controls would be of interest to confirm these associations. We also estimated the odds ratios associated with haplotypes of SNPs in each of the five genomic regions (Supplementary Table 4). At each locus, either the association was present on more than one haplotype carrying the risk allele, suggesting that the association was unlikely to be driven by a single rarer, higher penetrance variant, or the association was only present on a single haplotype of similar frequency to the index SNP, thus affording no additional information. There was no evidence of departure from a log-additive model for any SNP, meaning that the odds ratios for rare homozygotes did not differ significantly from the square of the odds ratios for heterozygotes. There was also no evidence of statistical gene–gene interaction between loci; the combined effects of SNPs were consistent with multiplicative (log-additive) combination of effects.

We investigated whether the loci were associated with different risks in subgroups of Wilms tumor cases characterized by specific phenotypic features or risk factors, including sex, age at diagnosis, family history of Wilms tumor, whether the tumor was bilateral or unilateral, histology, disease stage and whether there was relapse after treatment (Supplementary Tables 5–7). rs790356 at 11q14 showed evidence of a stronger effect in females than males ($P = 0.0017$) and a trend in strength of association with increasing age at diagnosis ($P = 0.0036$). Genotyping of rs790356 in additional series would be of interest to further evaluate these associations. There was no significant difference in risk at any other SNP in any subgroup (Supplementary Tables 6 and 7).

The strongest evidence of association ($P = 4.25 \times 10^{-15}$) was found at rs790356, which lies in a 68-kb LD block at 11q14.1 containing DLG2 (encoding discs, large homolog 2) and neighboring genes. DLG2 is a member of the membrane-associated guanylate kinase protein family, and its interaction with Scribbled and Lgl in the planar cell polarity pathway is essential for correct tissue morphogenesis during development, and its disruption has been implicated in oncogenesis.

Of note, the human homolog of Scribbled (encoded by SCRIB) is a recognized target of WT1, the Wilms tumor 1 gene product.

Table 1 Results for three confirmed Wilms tumor susceptibility SNPs

| Locus   | Chromosome | Position | Allele | Stage       | Controls | Cases | Controls | Cases | Per-allele OR (95% CI) | By stage | Combined |
|---------|------------|----------|--------|-------------|----------|-------|----------|-------|------------------------|----------|----------|
| rs3755132 | 2p24       | 15647271 | G/T    | GWAS        | 756      | 1,878 | 0.14     | 0.20  | 1.48 (1.25–1.74)        | 3.75 × 10⁻⁶ |          |
|         |            |          |        | UK replication | 753    | 2,715 | 0.14     | 0.19  | 1.41 (1.20–1.66)        | 2.12 × 10⁻⁵ | 0.34     |
|         |            |          |        | US replication | 702    | 1,034 | 0.14     | 0.21  | 1.50 (1.26–1.79)        | 5.45 × 10⁻⁶ |          |
| rs807624 | 2p24       | 15699922 | A/C    | GWAS        | 757      | 1,879 | 0.33     | 0.40  | 1.33 (1.17–1.51)        | 8.04 × 10⁻⁶ |          |
|         |            |          |        | UK replication | 747    | 2,708 | 0.36     | 0.42  | 1.29 (1.14–1.45)        | 2.83 × 10⁻⁶ | 1.32     |
|         |            |          |        | US replication | 709    | 1,037 | 0.34     | 0.42  | 1.39 (1.21–1.60)        | 3.95 × 10⁻⁶ |          |
| rs790356 | 11q14      | 83298435 | G/A    | GWAS        | 755      | 1,876 | 0.49     | 0.58  | 1.43 (1.27–1.61)        | 5.50 × 10⁻⁹ |          |
|         |            |          |        | UK replication | 761    | 2,725 | 0.49     | 0.55  | 1.27 (1.14–1.43)        | 3.18 × 10⁻⁶ | 4.25     |
|         |            |          |        | US replication | 697    | 1,037 | 0.51     | 0.57  | 1.28 (1.12–1.47)        | 0.000414  |          |

RAF, risk allele frequency.

*Chromosome and Build 36 position. †Risk/non-risk alleles. ‡Cochran-Armitage test for trend.
In mice, Scrib and Wt1 show coincident expression in the developing kidney. Moreover, in both mouse and human kidney cell lines, Wt1 has been shown to bind the SCRIB promoter and activate SCRIB expression. Thus, although the interactions of the DLG2 protein in humans have not been well characterized, it is plausible that susceptibility to Wilms tumor at the 11q14 locus is mediated through DLG2- and Wt1-related pathways.

We identified two SNPs associated with Wilms tumor susceptibility in a 109-kb LD block at 2p24.3. The effects of rs3755132 and rs807624 were maintained when corrected for each other (odds ratio (OR) = 1.25; 95% confidence interval (CI) = 1.03–1.53 and OR = 1.21, 95% CI = 1.04–1.41, respectively; \( P = 0.01 \)). Furthermore, the correlation between rs3755132 and rs807624 is weak (\( D' = 1 \) and \( r^2 = 0.328 \) in HapMap 3 Utah residents of Northern and Western European ancestry (CEU) samples; \( D' = 0.979 \) and \( r^2 = 0.328 \) in data from our GWAS controls). These data suggest that rs3755132 and rs807624 are independently associated with Wilms tumor. Because neither SNP alone can fully account for the association in this region, it is possible that a unique causal variant may exist that is in LD with rs3755132 and rs807624 and captures the effects of both of these SNPs. However, although imputation identified more strongly associated SNPs at this locus (Supplementary Table 3), none fully captured the associations of both rs3755132 and rs807624 and captures the effects of both of these SNPs. However, although imputation identified more strongly associated SNPs at this locus (Supplementary Table 3), none fully captured the associations of both rs3755132 and rs807624 and captures the effects of both of these SNPs. However, although imputation identified more strongly associated SNPs at this locus (Supplementary Table 3), none fully captured the associations of both rs3755132 and rs807624 and captures the effects of both of these SNPs. However, although imputation identified more strongly associated SNPs. 

| Locus | Chromosome | Position | Allele | Stage | Cases | Controls | Controls | Cases | Per-allele OR (95% CI) | By stage | Combined |
|-------|------------|----------|--------|-------|-------|----------|----------|-------|---------------------|----------|----------|
| rs1027643 | 5q14 | 91919548 | T/C | GWAS | 757 | 1,878 | 0.036 | 0.073 | 2.24 (1.69–2.97) | 1.94 x 10^{-4} |
| rs2283873 | 22q12 | 29343296 | A/G | GWAS | 757 | 1,879 | 0.015 | 0.036 | 2.65 (1.75–4.02) | 3.90 x 10^{-6} |
| rs5955543 | Xp22 | 17608318 | G/A | GWAS | 757 | 1,879 | 0.014 | 0.034 | 3.42 (1.91–6.12) | 3.69 x 10^{-5} |

**METHODS**

Methods and any associated references are available in the online version of the paper.

**ACKNOWLEDGMENTS**

We thank the families and the physicians and nurses that recruited them for their participation in this study, which was funded by the Wellcome Trust Case Control Consortium 3 (WTCCC3) initiative (grant reference 088804/Z/09/Z). We thank P. Donnelly for statistical advice throughout the design and execution of the study. We thank D. Dudukia, J. Bull, R. Linger, B. Ebbs, D. Hughes from Institute of Cancer Research (ICR) and Y. Mistry (from the Children’s Cancer and Leukaemia Group (CCLG) tumor bank) for assistance in sample collection, DNA extraction and genotyping. The UK samples and data were provided by the Children’s Oncology Group (study ARENO96B1) supported by the Chair’s grant U10 CA98543, SDC grant U10 CA98413 and Human Specimen Banking grant U24 CA114766 from the National Cancer Institute at the US National Institutes of Health. The UK samples were collected through the Factors Associated with Childhood Tumors (FACT) study, which is a CCLG Study (Multicentre Research Ethics Committee (MREC) 05/MRE02/17) and is supported by Cancer Research UK (grant references C8620/A9024 and C8620/A8857). A full list of collaborators is given in the Supplementary Note. The CCRG receives funding from the UK Department of Health, the National Cancer Intelligence Network, the Scottish Government and Children with Cancer UK. The views expressed in this publication are those of the authors and not necessarily of any of these organizations.

**URLs**

Database of Genotypes and Phenotypes (dbGaP), http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap; PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/; The 1000 Genomes Project, http://www.1000genomes.org/; haplview, http://www.broadinstitute.org/haplview/haplview; CaTS–Power Calculator, http://www.sph.umich.edu/csg/abecasis/CaTS/; IMPUTE2, http://mathgen.stats.ox.ac.uk/impute/impute_v2.html; 1958 Birth Cohort, http://www.cls.ioe.ac.uk/nccds.

**Note**

Supplementary information is available in the online version of the paper.

© 2012 Nature America, Inc. All rights reserved.

© 2012 Nature America, Inc. All rights reserved.
AUTHOR CONTRIBUTIONS
N.R. and C.T. designed the study and obtained financial support. M.G., J.H.,
M.H., J.K., S.L., G.L., M.M., B.M., V.N., J.N., S.P., B.P., M.R., M.S., H.T.
and N.R. undertook sample and data collection of UK cases, which was coordinated
by A.Z., M.W.-P., K.P.-J., C.A.S. and N.R. A.N., J.D. and P.G. coordinated the
transfer of US samples. E.R.P., S.S., R.M.M.-X., S.H., I.S. and A.R. coordinated
sample management and replication genotyping and sequencing. D.P. coordinated
data transfer and management. C.T. conducted statistical analyses with assistance
from D.P. and E.R. C.T., E.R.P. and N.R. wrote the manuscript. N.R. and C.T.
oversaw and managed all aspects of the study.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Stiller, C.A. & Parkin, D.M. International variations in the incidence of childhood
renal tumours. Br. J. Cancer 62, 1026–1030 (1990).
2. Breslow, N.E., Beckwith, J.B., Perlman, E.J. & Reeve, A.E. Age distributions, birth
weights, nephrogenic rests, and heterogeneity in the pathogenesis of Wilms tumor.
Pediatr. Blood Cancer 47, 260–267 (2006).
3. Scott, R.H., Stiller, C.A., Walker, L. & Rahman, N. Syndromes and constitutional
chromosomal abnormalities associated with Wilms tumour. J. Med. Genet. 43,
705–715 (2006).
4. Scott, R.H. et al. Constitutional 11p15 abnormalities, including heritable imprinting
center mutations, cause nonsyndromic Wilms tumor. Nat. Genet. 40, 1329–1334
(2008).
5. Little, S.E. et al. Frequency and heritability of WTI mutations in nonsyndromic
Wilms’ tumor patients: a UK Children’s Cancer Study Group Study. J. Clin. Oncol. 22,
4140–4146 (2004).
6. Mailman, M.D. et al. The NCBI dbGaP database of genotypes and phenotypes.
Nat. Genet. 39, 1181–1186 (2007).
7. Kim, E., Cho, K.O., Rothschild, A. & Sheng, M. Heteromultimerization and NMDA
receptor–clustering activity of Chapsyn-110, a member of the PSD-95 family of
proteins. Neuron 17, 103–113 (1996).
8. Humbert, P.D. et al. Control of tumourigenesis by the Scribble/Dlg/Lgl polarity
module. Oncogene 27, 6888–6907 (2008).
9. Bilder, D., Li, M. & Perrimon, N. Cooperative regulation of cell polarity and growth
by Drosophila tumor suppressors. Science 289, 113–116 (2000).
10. Wells, J., Rivera, M.N., Kim, W.J., Starbuck, K. & Haber, D.A. The predominant
WTI isoform (+KTS) encodes a DNA-binding protein targeting the planar cell
polarity gene Scribble in renal podocytes. Mol. Cancer Res. 8, 975–985 (2010).
11. Ll. L., Monckton, E.A. & Godbout, R. A role for DEAD box 1 at DNA double-strand
breaks. Mol. Cell. Biol. 28, 6413–6425 (2008).
12. Schaub, R. et al. Array comparative genomic hybridization reveals unbalanced gain
of the MYCN region in Wilms tumors. Cancer Genet. Cytogenet. 172, 61–65 (2007).
13. Squire, J.A. et al. Co-amplification of MYCN and a DEAD box gene (DIO1) in
primary neuroblastoma. Oncogene 10, 1417–1422 (1995).
ONLINE METHODS

Samples. Wilms tumor cases from the UK were recruited from Paediatric Oncology centers through the Factors Associated with Childhood Tumors (FACT) study and the Children’s Cancer and Leukemia Group (CCLG) (Supplementary Note). From 1985 onward, 97% of children in the UK with Wilms tumor have been registered with the CCLG (formerly called the UKCCSG). Phenotypic data were obtained from the referring center and/or the National Registry of Childhood Tumors (NRCT). Wilms tumor cases from North America were recruited through the National Wilms Tumor Study Group (NWTSG) and the Children’s Oncology Group (COG). Since 1969, the NWTSG has registered over 80% of incident cases of Wilms tumor.14 Cases were unrelated and self-reported to be of non-Hispanic, European ancestry. Information regarding ancestry, gender, age at diagnosis, tumor stage, histopathology, relapse, bilaterality and family history of Wilms tumor was obtained for as many cases as possible.

DNA samples from the UK controls were from the 1958 Birth Cohort, an ongoing study of persons born in the UK in one week in 1958. In the UK replication study, we used 2,814 population controls representative of geographical regions across the UK. We did not have samples from North American controls. Instead, we used data from dbGaP. In the GWAS, we used data from individuals recruited as controls for a study of Parkinson’s disease at the NeuroGenetics Research Consortium (NGRC)-affiliated movement disorder clinics in Oregon, Washington, Georgia and New York. Individuals were >20 years of age, of self-reported white ethnicity and free of neurological diseases. In the US replication study, we used controls recruited through being friends or spouses of attendees of patients at the MD Anderson Cancer Center. Individuals were described as non-Hispanic, Caucasian and had not had cancer.

The study had approval from the appropriate ethics and study committees in the UK (London Multicentre Ethics Committee, 05/MRE02/17) and the United States (Children’s Oncology Group, AREN09B1). Informed consent was obtained from affected individuals and/or their parents, as appropriate.

Genotyping. For the GWAS, cases were genotyped using the Illumina HumanOmniExpress-12v1_A SNP array at the Broad Institute Genetic Analysis Platform. Each 96-well plate contained a HapMap CEU control. Samples were clustered and called from intensity data using BeadStudio version 3.1.3.0 with genotyping module version 3.2.32 (Illumina). Genotypes for the 1,879 controls used in the GWAS were generated using the Illumina Omni 1M array as part of the Center for Inherited Disease Research (CIDR): NGRC Parkinson’s Disease Study.

For the replication studies, we used 5 exome capture assays (TaqMan, Applied Biosystems), following the manufacturer’s protocols. We combined cases with controls on 384-well plates, each of which included at least two negative controls and 1–2% duplicates (intra-platform concordance was 98.86%). We performed cross-platform validation, genotyping by TaqMan on all cases that had been run on the Illumina array (cross-platform concordance was 99.11%). Genotypes for US replication samples were compared to those available from 1,037 US controls, which had been genotyped on the Illumina Omni 1M array as part of the High Density SNP Association Analysis of Melanoma study.

For rs2283873 at 22q12, the genotype frequencies were in Hardy-Weinberg equilibrium (HWE) in the GWAS, but there was some deviation from HWE in the replication analyses (Supplementary Table 2). We therefore undertook direct Sanger sequencing of rs2283873 to validate the TaqMan data and to confirm genotypes of 257 samples, including all those heterozygous or homozygous for the minor allele or that had failed TaqMan replication. There was full concordance between genotype calls determined by TaqMan and sequencing, thereby confirming the validity of the replication genotyping.

Inclusion criteria for analyses. We only included cases with a definite histological diagnosis of Wilms tumor. We excluded individuals with histological diagnoses of nephroblastomatosis, renal cell carcinoma, renal sarcoma, rhabdoid tumor and peripheral primitive neuroectodermal tumor (PNET). We restricted analyses to individuals that were called on >97% of successfully genotyped SNPs. We computed identity-by-state (IBS) probabilities for all pairs in order to identify cryptic duplicates and close relatives (IBS > 0.80). For each pair identified, the sample with the highest genotype call rate was retained. We thus eliminated two cases, one from a pair of close relatives (IBS = 0.82) and one from a pair of duplicated samples (IBS > 0.99). Using a subset of 65,045 uncorrelated SNPs (r2 < 0.1), we estimated the average IBS between all participants together with Phase 2 HapMap samples (90 CEU, 90 Ibadan from Yoruba, Nigeria (YRI), 44 Japanese from Tokyo (JPT) and 45 Han Chinese from Beijing (CHB)). Using multi-dimensional scaling, we identified and removed 42 cases with >15% non-European ancestry (27 with African ancestry and 15 with Asian ancestry). Furthermore, we excluded samples with abnormal heterozygosity (>5 s.d. from the mean heterozygosity; one case sample eliminated). Following these exclusions, 757 cases and 1,879 controls were available for the genome-wide analysis.

For the GWAS, we included the 696,780 SNPs that were successfully genotyped in cases and controls. We filtered out all SNPs in either cases or controls with (i) a minor allele frequency (MAF) of <1%, (ii) a call rate of <95% in cases or controls or (iii) a MAF of 1–5% with a call rate of <99%. We also excluded SNPs whose genotyped frequency departed from HWE at P < 1 × 10−12 in cases or P < 0.00001 in controls (male samples were excluded for evaluation of HWE for SNPs on the X chromosome). We excluded synonymous (G−G and A−T) SNPs. Following these exclusions, we analyzed 599,255 SNPs genotyped in cases and controls. Cluster plots were inspected manually for all SNPs considered for replication.

For the replication, call rates of at least 97% per 384-well plate were required, and cluster plots were visually examined for each plate. Genotype distributions for each SNP were evaluated for deviation from HWE.

Statistical methods. For the GWAS, we assessed associations between each SNP and disease primarily using the 1-degree-of-freedom Cochran-Armitage trend test; as a secondary measure of association, we performed the 1-degree-of-freedom allelic association test. Inflation in the χ2 statistic was assessed using the genomic control approach. We derived an inflation factor (λ) by dividing the median of the lowest 90% of the 1-degree-of-freedom statistics by the 45% percentile of a 1-degree-of-freedom χ2 distribution (0.357). Because λ was small, we chose to present P values uncorrected for λ, as this made little difference to the significance levels and preserved consistency with the replication analysis.

We selected 20 SNPs that showed association at a significance level of P < 5 × 10−5 based on the 1-degree-of-freedom trend test in the GWAS for the replication phase. Where two or more SNPs were selected from the same region, we used multiple logistic regression to determine a minimal set of SNPs that showed evidence of association after adjustment for other SNPs.

We performed 1-degree-of-freedom tests of association for the GWAS and the US and UK replication analyses separately and then for the combined data stratified by stage (GWAS or replication) and study group (US or UK). For the principal association analysis, we have emphasized the OR estimates from the combined replication analysis (stratified by study) to minimize the effect of winner’s curse. We assessed each SNP for dose response by comparing 1-degree-of-freedom and 2-degree-of-freedom logistic regression models, adjusting for stage, using a likelihood ratio test (P < 0.01), and examined the combined effects of multiple SNPs by evaluating the effect on the model of adding an interaction term, using a likelihood ratio test (P < 0.01) and adjusting for stage.

Modification of the per-allele ORs by covariate phenotype and/or risk factors was assessed using analysis of cases against all controls to generate strata-specific ORs and using case-only analysis to look for heterogeneity of effect (Supplementary Table 6). Tumors defined as stage 3 or 4 were categorized as advanced, and tumors defined as stage 1 or 2 were categorized as non-advanced; bilateral tumors, which are usually termed stage 5, were excluded from this analysis. Tumors with histology described as high-risk blastemal, high-risk diffuse anaplasia or unfavorable were categorized as high risk, and tumors with histology described as blastemal, cystic, intermediate, triphasic, favorable, intermediate-focal anaplasia or intermediate non-anaplastic were categorized as standard risk. Age of diagnosis was analyzed in three categories: <2 years, 2–4 years and >4 years, and the effect of age at diagnosis on SNP genotype in the cases was analyzed using ordered polytomous regression.

Imputation was performed using IMPUTE2 using data from HapMap 3 (~1.5 million SNPs, 1,184 individuals) and the 1000 Genomes Project low-coverage
pilot haplotypes (179 individuals) as a framework. Imputed SNPs were retained only if they had an information score of >0.8 and a certainty score of >0.95 in both cases and controls. Haplotypes were evaluated using SNPs in each region that were significantly associated with Wilms tumor at $P < 0.001$, after eliminating perfectly correlated SNPs. For haplotypes of frequency >1% (in controls), the $P$ value for association for the haplotype was calculated. CaTS–Power Calculator software was used to estimate the power to detect each of the associations found. LD matrices between SNPs reported in HapMap were based on Data Release 27/phase 2 and 3 Feb09 on NCBI B36 assembly, dbSNP b126, viewed using Haploview software (v4.2). LD blocks were evaluated using the HapMap recombination rate (cM/Mb) and defined using the Oxford recombination hotspots

All genomic references are based on NCBI Build 36. Analyses were performed using R (v2.6), Stata10 (State College) and PLINK (v1.07) software.

14. D’Angio, G.J. The National Wilms Tumor Study: a 40 year perspective. Lifetime Data Anal. 13, 463–470 (2007).
15. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. Science 310, 321–324 (2005).
Corrigendum: A genome-wide association study identifies susceptibility loci for Wilms tumor

Clare Turnbull, Elizabeth R Perdeaux, David Pernet, Arlene Naranjo, Anthony Renwick, Sheila Seal, Rosa Maria Munoz-Xicola, Sandra Hanks, Ingrid Slade, Anna Zachariou, Margaret Warren-Perry, Elise Ruark, Mary Gerrard, Juliet Hale, Martin Hewitt, Janice Kohler, Sheila Lane, Gill Levitt, Mabrook Madi, Bruce Morland, Veronica Neefjes, James Nicholdson, Susan Picton, Barry Pizer, Milind Ronghe, Michael Stevens, Heidi Traunecker, Charles A Stiller, Kathy Pritchard-Jones, Jeffrey Dome, Paul Grundy & Nazneen Rahman

*Nat. Genet.* 44, 681–684 (2012); published online 29 April 2012; corrected after print 6 June 2012

In the version of this article initially published, the name of one of the authors was incorrectly listed as James Nicholdson. The correct name is James Nicholson. The error has been corrected in the HTML and PDF versions of the article.
In the version of this article initially published, the following statement was omitted from the Acknowledgments: “We acknowledge use of data from the database of Genotypes and Phenotypes (dbGaP) from the NCBI, US National Library of Medicine. Research support to collect data and develop an application to support High Density SNP Association Analysis of Melanoma (phs000187) was provided by grants 3P50CA093459, 5P50CA097007, 5R01ES011740 and 5R01CA133996 from the US National Institutes of Health. This study also used, in part, data from the National Institute of Neurological Disorders and Stroke (NINDS) dbGaP database from the CIDR:NGRC Parkinson’s Disease Study (phs000196).” The error has been corrected in the HTML and PDF versions of the article.