Published in final edited form as:

Science. 2019 December 06; 366(6470): 1247–1251. doi:10.1126/science.aax1323.

Embryonal precursors of Wilms tumor

Tim HH Coorens¹, Taryn D Treger^{1,2,3}, Reem Al-Saadi^{4,5}, Luiza Moore^{1,2}, Maxine GB Tran^{6,7}, Thomas J Mitchell^{1,2,8}, Suzanne Tugnait⁴, Christine Thevanesan⁴, Matthew D Young¹, Thomas RW Oliver^{1,2,9}, Minou Oostveen^{4,5}, Grace Collord^{1,2,3}, Patrick S Tarpey², Alex Cagan¹, Yvette Hooks¹, Mark Brougham¹⁰, Ben C Reynolds¹¹, Giuseppe Barone⁵, John Anderson^{4,5}, Mette Jorgensen⁵, GA Amos Burke^{2,3}, Johannes Visser², James C Nicholson^{2,3}, Naima Smeulders⁵, Imran Mushtaq⁵, Grant D Stewart^{2,8}, Peter J Campbell¹, David C Wedge^{12,13}, Iñigo Martincorena¹, Dyanne Rampling⁵, Liz Hook^{2,9}, Anne Y Warren^{2,9}, Nicholas Coleman^{2,9}, Tanzina Chowdhury⁵, Neil Sebire^{4,5}, Jarno Drost¹⁴, Kourosh Saeb-Parsy^{2,8}, Michael R Stratton¹, Karin Straathof^{4,5}, Kathy Pritchard-Jones^{4,5}, Sam Behjati^{1,2,3,*}

¹Wellcome Sanger Institute, Hinxton, CB10 1SA, UK ²Cambridge University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK ³Department of Paediatrics, University of Cambridge, Cambridge, CB2 0QQ, UK ⁴UCL Great Ormond Street Institute of Child Health, London, WC1N 1EH, UK ⁵Great Ormond Street Hospital for Children NHS Foundation Trust, London, WC1N 3JH, UK ⁶UCL Division of Surgery and Interventional Science, Royal Free Hospital, London, NW3 2PS, UK ⁷Specialist Centre for Kidney Cancer, Royal Free Hospital, London, NW3 2PS, UK ⁸Department of Surgery, University of Cambridge, Cambridge, CB2 0QQ, UK ⁹Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK ¹⁰Department of Haematology and Oncology, Royal Hospital for Sick Children, Edinburgh, EH9 1LF, UK ¹¹Department of Paediatric Nephrology, Royal Hospital for Children, Glasgow, G51 4TF, UK ¹²Big Data Institute, University of Oxford, Oxford, OX3 7LF, UK ¹³Oxford NIHR Biomedical Research Centre, John Radcliffe Hospital, Oxford, OX3 9DU, UK ¹⁴Princess Máxima Center for Pediatric Oncology, Oncode Institute, 3584 CT Utrecht, The Netherlands

Abstract

Some adult cancers arise from pre-malignant clonal expansions. It is unknown whether childhood tumors are also preceded by such expansions, or arise in isolation. Here, we investigated whether Wilms tumor (nephroblastoma) develops from a pre-cancerous background by examining the

*Corresponding author. sb31@sanger.ac.uk.

Author contributions: S.B. conceived of the experiment. T.C., T.D.T. and S.B. analyzed data. Statistical expertise was provided by M.D.Y., D.C.W., and In.M. M.D.Y., T.J.M., G.C., P.S.T., P.J.C., M.R.S., In.M and K.P-J contributed to discussion. L.M. and M.G.B.T. performed laser capture microscopy experiments. Samples were curated and/or experiments were performed by R.A., C.T., S.T., M.O., Y.H., M.B., B.C.R., G.B., J.A., M.J., G.A.A.B., J.V., J.C.N., N.S., K.S-P, G.D.S., K.S., T.Ch., I.M, J.D., and K.P-J. Pathological expertise was provided by L.H., T.R.W.O., D.R., A.Y.W., N.C., and N.S. A.C. created kidney illustrations. T.C., T.D.T., and S.B. wrote the manuscript. S.B. directed the study.

Competing Interests: GDS is a paid consultant for Pfizer, Merck, EUSA Pharma and Cambridge Medical Robotics. All other authors declare no competing interests.

Data and materials availability: Raw sequencing data have been deposited in the European Genome-phenome Archive (EGA) under study ID EGAD00001004774.

phylogenetic relation between tumors (n=66) and corresponding normal tissues (n=163). We found pre-malignant clonal expansions in morphologically normal kidney tissues that preceded tumors in 14 of 23 cases (61%). Somatic mutations, shared between tumor and normal tissue but absent from blood, defined these clonal expansions. We found hypermethylation of *H19*, a *bonafide* driver of Wilms tumor, in 58% of expansions. Phylogenetic analyses of bilateral tumors indicated that clonal expansions can evolve before left and right kidney primordia diverge. Our findings reveal embryonal precursors from which unilateral and multifocal cancers develop.

Adult cancers typically arise as a consequence of aging and mutagen exposure, at times via the generation of pre-cancerous clonal expansions such as Barrett's esophagus, clonal hematopoiesis, or colonic polyps. It is unknown whether childhood tumors, thought to result from aberrant fetal development, form from such precursors (Fig. 1A). We investigated whether Wilms tumor (nephroblastoma), the most common kidney cancer of childhood, may evolve from pre-cancerous clonal expansions. Wilms tumor is a prototypical embryonal malignancy of infants and young children (1). It arises from abnormal fetal nephrogenesis, which it resembles morphologically (1) and transcriptionally (2). It occurs sporadically, or in the context of bilateral tumors, multifocal lesions, urogenital developmental disorders or overgrowth syndromes (1).

To discover possible precursors of Wilms tumor, we used somatic mutations to infer the phylogenetic relationship between cancers and corresponding normal tissues (kidney and blood). We analyzed 229 whole genome sequences obtained from 54 individuals: 23 children with Wilms tumor, 16 parents of affected children, three children with non-Wilms tumor kidney tumors (congenital mesoblastic nephroma, malignant rhabdoid tumor), ten adults with clear cell renal cell carcinoma (ccRCC), and two adults without renal tumors (one kidney transplant and one kidney obtained at autopsy; table S1). We called base substitutions against the reference human genome and extracted mosaic mutations from each set of donor related tissues. We validated the method for calling mosaic mutations by sequencing parental germline DNA, re-sequencing tissues and inspection of raw data (fig. S1). Based on the variant allele frequency (VAF) and distribution of mutations across related tissues, we built phylogenetic trees of tumor development. We supplemented DNA data with analyses of RNA sequences and genome wide methylation patterns (table S1).

Our discovery cohort consisted of three children with unilateral Wilms tumor. We sampled tumors, blood, and histologically normal kidney tissues from the same individuals (table S1 and fig. S2). As expected, whole genome sequences revealed mosaic mutations attributable to the first cell divisions of the fertilized egg (fig. S3, table S2) (3, 4). In two cases, we also detected mosaic mutations in normal kidneys that were present in the corresponding cancer, but absent from blood (Fig. 1, B to D and table S2), indicating that the tumors had arisen from that particular normal kidney tissue.

Several features of these mutations showed that they defined clonal expansions in normal kidney tissue, as illustrated by case PD37272 (Fig. 1, B to D). The VAFs of mutations in the normal tissue of this kidney, variants #3-#5 (Fig. 1C), were as high as 44%, suggesting that the mutation was present in 88% of all cells in the biopsy. Mutations #3-#5 were present in the two parenchymal biopsies (i.e. cortex and medulla) whilst absent from blood DNA,

deeply sequenced to 106X genome-wide (fig. S4). Similarly, mutations #3-#5 were undetectable in renal pelvis, which is embryologically derived from a different lineage than kidney parenchyma(5). Furthermore, the VAF of early embryonic mutations, variants #1 and #2, was inflated in parenchyma and in tumors (Fig. 1D). Such inflation of early embryonic mutations is a feature of tissues that contain a clonal expansion of a single cell (fig. S3). By contrast, in tissues devoid of a major clone, such as renal pelvis and blood, the VAFs of early embryonic mutations were not inflated (Fig. 1D). Thus, these normal tissue variants #3-#5 demonstrate the presence of clonal expansions within kidney parenchyma, which we termed clonal nephrogenesis, accounting for up to 88% of cells sampled in the cortex.

To further study and validate our discovery of clonal nephrogenesis as an antecedent of Wilms tumor, we studied another 20 cases: 15 unilateral tumors with normal tissue biopsies curated through a British childhood renal tumor study (IMPORT); four cases of bilateral Wilms tumor; and one tumor with ten normal tissue biopsies (table S1 and fig. S5). Considering the entire cohort of 23 children, we found evidence of clonal nephrogenesis in ten of 19 (53%) children with unilateral disease and in all four children with bilateral cancers (Fig. 2, A to B). The presence of clonal nephrogenesis was further substantiated by the significant ($p<0.01$; Wilcoxon signed rank test) inflation of VAFs of early embryonic variants (fig. S6). There were no copy number changes detected in normal tissues by three different methods.

An alternative explanation for these findings could be tumor infiltration into normal tissue, not visible histologically (fig. S2), or cross-contamination of DNAs. This explanation is implausible, as contamination would manifest as shared variants at a low VAF, rather than select mutations at a high VAF. We statistically excluded possible contribution from tumor infiltration and contamination by employing a binomial mixture model on the observed base counts of tumor mutations in the normal samples (fig. S7).

Next, we investigated whether clonal nephrogenesis represents the normal clonal architecture of human nephrons by three approaches. First, using laser capture microscopy (LCM), we excised glomeruli ($n=7$), proximal and distal tubules ($n=15$) from kidneys of three individuals (Fig. 2, C to D). We subjected these LCM cuts to whole genome sequencing using an established method for generating low input DNA libraries (6, 7). Kidney tissues were obtained at autopsy (one case) or from normal portions of kidneys affected elsewhere by ccRCC (two cases). We analyzed somatic mutations across these LCM cuts which revealed a VAF distribution (fig. 2E) that is inconsistent with the monoclonal organization seen, for example, in endometrial glands or colonic crypts (6, 7). Second, we assessed whether mutations were commonly shared between renal tumors and surrounding normal kidney tissue. We studied childhood congenital mesoblastic nephroma (two tumors, six normal kidney samples), childhood malignant rhabdoid tumor (one cancer, one normal kidney sample) and adult ccRCC (eight cancers including one bilateral case, 15 normal tissues). Applying the same analysis pipeline, we identified early embryonic mutations shared between tumor, normal kidney tissues, and blood (fig. S8). However, we did not find mutations shared only between tumor and normal tissue (Fig. 2B), showing that such mutations were specific to, and enriched in, Wilms tumor ($p<0.001$; Fisher's exact test). Of particular relevance were renal cell carcinomas which, like Wilms tumor, are

derived from nephrons. If normal embryological clonal dynamics typically had generated clonal expansions, we would have expected to find clonal nephrogenesis in ccRCC cases. Third, we examined all developmental mutations of normal kidney tissues listed thus far, supplemented by an additional 18 biopsies obtained from bilateral kidneys declined for transplantation (Fig. 2B). We analysed somatic mutations present in kidney tissue and absent from non-renal tissue, irrespective of whether they were shared with tumors. Collectively, these analyses of 77 normal kidney biopsies revealed that variants of tissues without clonal nephrogenesis have a significantly lower VAF distribution than clonal nephrogenesis mutations ($p < 0.001$, Wilcoxon rank sum; Fig. 2B). Therefore our findings indicate that clonal nephrogenesis represents aberrant kidney development.

A central question that our findings raise is whether cancer-causing (driver) events underpin clonal nephrogenesis. However, clonal nephrogenesis mutations ($n=66$) (table S2 and fig. S8) were non-coding (64 of 66) or did not generate plausible oncogenic events. We searched further for driver events amongst germline and somatic DNA mutations, in transcriptomes and, where available, methylation patterns. We found hypermethylation of *H19* in seven of 12 (58%) normal kidney tissues with clonal nephrogenesis. It was absent from normal kidney tissues without clonal nephrogenesis and from blood (Fig. 2, A and F), bar the blood of a child (PD40738) with Beckwith-Wiedemann syndrome (fig. S9). *H19* hypermethylation is an established driver event of Wilms tumor, thought to operate by disrupting the epigenetic regulation of growth promoting genes that reside on 11p15 (8–10). The degree of hypermethylation of *H19* correlated with the VAF of clonal nephrogenesis, indicating that hypermethylation pervaded the entire clone (Fig. 2G). In the five wildtype samples, hypermethylation of *H19* may therefore have evaded discovery due to a small clone size (Fig. 2A). Methylation of the KvDMR1 locus was unchanged in clonal nephrogenesis. Hypomethylation of KvDMR1 underlies overgrowth syndromes that confer only a minimal predisposition to Wilms tumor (11). We did not find any further driver events accounting for clonal nephrogenesis, despite re-interrogation by exome sequencing of 15 of 17 tissues with clonal nephrogenesis. Gene expression profiles, including utilisation of fetal transcripts, did not differ between normal renal tissues that did, or did not, contain clonal nephrogenesis (fig. S9). Similarly, global methylation patterns did not differ between these two groups (fig. S9).

The timing of the emergence of clonal nephrogenesis during development could be defined in three children from whom we obtained bilateral tumors. In cases PD40735 and PD36159, left and right tumors were derived from the same trunk of clonal nephrogenesis (Fig. 3, A to D). In the third child, PD40378, all five left tumors, but not the right neoplasms, were related to clonal nephrogenesis on the left (Fig. 3E). Taken together, these findings indicate that in two children with bilateral cancers, clonal nephrogenesis must have arisen before left and right kidney primordia diverged, early in embryogenesis (5). In unilateral tumors, we cannot definitively comment on the timing of the occurrence of clonal nephrogenesis. It may have evolved before the kidney was formed or thereafter, followed by a “clonal sweep” of clonal nephrogenesis replacing kidney tissue.

In five cases, we sampled multiple neoplasms of the same kidney, which revealed two configurations of tumor development (Fig. 3). Tumors were either derived from a shared

trunk that had emerged from clonal nephrogenesis (Fig. 3, A to E). Or, they arose independently and successively from clonal nephrogenesis, alluding to a sustained potential for the latter to spawn tumors (Fig. 3, F to G). For example, PD36165 presented with two lesions at opposing poles of the kidney, a nephrogenic rest and a Wilms tumor. Cancer and rest had emerged from the same ancestral clone, yet at different time points, followed by clonal diversification within each (Fig. 3G).

In a final analysis we compared the somatic changes of Wilms tumor with, and without clonal nephrogenesis, which revealed a near mutual exclusivity of loss of heterozygosity (LOH) of 11p15 and clonal nephrogenesis. That is, in tumors with clonal nephrogenesis, LOH of 11p15 was mostly absent ($p=0.009$, Fisher's exact test) (fig. S10). This indicates that there may be two distinct pathways for Wilms tumor generation, which both utilise dysregulation of 11p15 genes as a driver (Fig. 1A). Accordingly, cancers may either arise directly in isolation through LOH of 11p15, or indirectly via clonal nephrogenesis with perturbation of 11p15 by hypermethylation of *H19*.

Overall, we identified clonal expansions in histologically normal tissue as an atypical outcome of renal development that commonly antedates Wilms tumor. We demonstrate a direct phylogenetic link between clonal expansions, *H19* hypermethylation, and the formation of cancer, thus identifying clonal nephrogenesis as an epigenetic progenitor of cancer, comprised of “neoplasia-ready cells” (12). In contrast to precursors of adult cancer, clonal nephrogenesis generated histologically and functionally normal tissues, that in the most pronounced cases occupied the bulk of renal tissues. It may be possible that the extent of clonal nephrogenesis informs on the malignant potential of kidneys and recurrence risk, which could be utilised to guide treatment intensity and surveillance schedules of Wilms tumor. Moreover, it is conceivable that *H19* hypermethylation may be amenable to therapeutic intervention, thus enabling targeting the cancer root. One may even envision that prevention of Wilms tumor may become feasible, if we were able to manipulate the neoplastic potential of clonal nephrogenesis. Collectively our findings portray Wilms tumor as an insurrection on the background of a premalignant tissue bed, rather than a clearly demarcated neoplasm in an otherwise normal polyclonal kidney. We speculate that embryonal clonal expansions, perhaps also driven epigenetically, may be a common phenomenon in childhood cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank research nursing staff at Cambridge University Hospitals, the Royal Hospital for Sick Children (Edinburgh), and the Royal Hospital for Children (Glasgow) as well as all IMPORT investigators. We thank Dr Moritz Gerstung for critical review of the manuscript. We are indebted to our little and older patients and their families for participating in our research.

Funding: This experiment was principally funded by The Little Princess Trust, the St. Baldrick's Foundation (Robert J. Arceci International Award to S.B.) and Wellcome (Fellowship to S.B.; Sanger core funding). Additional funding was received from CRUK (IMPORT study; fellowship to T.J.M.; Cambridge Centre), NIHR (Biomedical Research Centre Great Ormond Street; Cambridge Human Research Tissue Bank; Oxford Biomedical Research

Centre; Fellowship to T.R.W.O.), The Royal College of Surgeons of England (Fellowship to T.J.M.), Wellcome (Fellowship to T.C. and K.S.), Great Ormond Street Hospital Children's Charity (R.A., K.P.J.), Li Ka Shing foundation (D.C.W.). J.D. acknowledges funding from the Alpe d'HuZes foundation / KWF Dutch Cancer Society Bas Mulder Award (#10218) and the Oncode Institute.

References

1. Treger TD, Chowdhury T, Pritchard-Jones K, Behjati S. The genetic changes of Wilms tumour. *Nat Rev Nephrol.* 2019; 15:240–251. [PubMed: 30705419]
2. Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, Collord G, Botting RA, Popescu D-M, Loudon KW, Vento-Tormo R, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science.* 2018; 361:594–599. [PubMed: 30093597]
3. Behjati S, Huch M, van Boxtel R, Karthaus W, Wedge DC, Tamuri AU, Martincorena I, Petljak M, Alexandrov LB, Gundem G, Tarpey PS, et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature.* 2014; 513:422–425. [PubMed: 25043003]
4. Ju YS, Martincorena I, Gerstung M, Petljak M, Alexandrov LB, Rahbari R, Wedge DC, Davies HR, Ramakrishna M, Fullam A, Martin S, et al. Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature.* 2017; 543:714–718. [PubMed: 28329761]
5. Short KM, Smyth IM. The contribution of branching morphogenesis to kidney development and disease. *Nat Rev Nephrol.* 2016; 12:754–767. [PubMed: 27818506]
6. Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, Georgakopoulos N, Torrente F, Noorani A, Goddard M, Robinson P, et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature.* 2019; 574:532–537. [PubMed: 31645730]
7. Moore L, Leongamornlert D, Coorens THH, Sanders MA, Ellis P, Dawson K, Maura F, Nangalia J, Tarpey PS, Brunner SF, Lee-Six H, et al. The mutational landscape of normal human endometrial epithelium. *bioRxiv.* 2018
8. Moulton T, Crenshaw T, Hao Y, Moosikasuwan J, Lin N, Dembitzer F, Hensle T, Weiss L, McMorro L, Loew T, Kraus W, et al. Epigenetic lesions at the H19 locus in Wilms' tumour patients. *Nat Genet.* 1994; 7:440–447. [PubMed: 7920666]
9. Okamoto K, Morison IM, Taniguchi T, Reeve AE. Epigenetic changes at the insulin-like growth factor II/H19 locus in developing kidney is an early event in Wilms tumorigenesis. *Proc Natl Acad Sci U S A.* 1997; 94:5367–5371. [PubMed: 9144243]
10. Charlton J, Williams RD, Sebire NJ, Popov S, Vujanic G, Chagtai T, Alcaide-German M, Morris T, Butcher LM, Guilhamon P, Beck S, et al. Comparative methylome analysis identifies new tumour subtypes and biomarkers for transformation of nephrogenic rests into Wilms tumour. *Genome Med.* 2015; 7
11. Cooper WN, Luharia A, Evans GA, Raza H, Haire AC, Grundy R, Bowdin SC, Riccio A, Sebastio G, Blik J, Schofield PN, et al. Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur J Hum Genet.* 2005; 13:1025–1032. [PubMed: 15999116]
12. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet.* 2006; 7:21–33. [PubMed: 16369569]
13. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009; 25(14):1754–1760. [PubMed: 19451168]
14. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature.* 2016; 534(7605)
15. Jones D, Raine KM, Davies H, Tarpey PS, Butler AS, Teague JW, Nik-Zainal S, Campbell PJ. cgpCaVEManWrapper: simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Current protocols in bioinformatics.* 2016; 56(1):15–10.
16. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics.* 2009; 25(21):2865–2871. [PubMed: 19561018]

17. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, Weigman VJ, Marynen P, Zetterberg A, Naume B, Perou CM, et al. Allele-specific copy number analysis of tumors. *Proceedings of the National Academy of Sciences*. 2010; 107(39):16910–16915.
18. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, Raine K, Jones D, Marshall J, Ramakrishna M, Shlien A, et al. The life history of 21 breast cancers. *Cell*. 2012; 149:994–1007. [PubMed: 22608083]
19. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*. 1995:289–300.
20. Gerstung M, Papaemmanuil E, Campbell PJ. Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics*. 2014; 30(9):1198–1204. [PubMed: 24443148]
21. Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, Goodstein DM, Elisk CG, Lewis SE, Stein L, Holmes IH. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome biology*. 2016; 17(1)
22. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*. 1934; 26(4):404–413.
23. Behjati S, Tarpey PS, Haase K, Ye H, Young MD, Alexandrov LB, Farndon SJ, Collord G, Wedge DC, Martincorena I, Cooke SL, et al. Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. *Nature communications*. 2017; 8
24. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Yin Kok C, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic acids research*. 2014; 43(D1):D805–D811. [PubMed: 25355519]
25. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014; 30(10):1363–1369. [PubMed: 24478339]
26. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29(1):15–21. [PubMed: 23104886]
27. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31(2):166–169. [PubMed: 25260700]
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 2014; 15(12)
29. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015; 348(6235):648–660. [PubMed: 25954001]
30. Popescu D-M, Botting RA, Stephenson E, Green K, Webb S, Jardine L, Calderbank EF, Polanski K, Goh I, Efremova M, Acres M, et al. Decoding the development of the blood and immune systems during human fetal liver haematopoiesis. *Nature*. 2019; 574:365–371. [PubMed: 31597962]

One sentence summary

Clonal expansions within normal kidney tissue, some of which harbor driver variants, often precede the most common type of childhood kidney cancer, Wilms tumor (nephroblastoma).

Figure 1. Pre-cursor clonal expansions in normal human kidneys.

(A) Wilms tumor arises during embryogenesis on the background of an otherwise normal kidney, as traditionally thought, or from precursor lesions residing in normal tissues, as found here.

(B) Overview of tissue sampling in the kidney of PD37272.

(C) Somatic mutations can be timed used their VAF across the corresponding normal tissues, which is higher for earlier mutations. If the mutation is present in tumor, kidney, and blood, it is an early embryonic mutation (#1-#2). If it is present in kidney samples and tumor only, it is clonal nephrogenic (mutations #3-#5, marked by asterisk). If it is only in the tumor, it is labelled as such. White and black circles indicate whether the observed VAF is insignificant

(white) or significant (black), $p < 0.001$ (test of presence using beta-binomial overdispersion).

(D) The VAF for the last embryonic mutation in kidney samples and tumor compared with blood.

Figure 2. Clonal nephrogenesis and *H19* hypermethylation

(A) Sizes of nephrogenic clones as predicted by twice the VAF of the most prominent nephrogenic mutation alongside corresponding methylation values of the *H19* locus. Green symbols indicate significant deviation ($p < 0.05$, Wilcoxon rank-sum test) from the background methylation distribution (grey shaded area) as obtained from normal kidney samples without clonal nephrogenesis.

(B) Mutations present in samples obtained from normal kidneys but absent in matched blood. Only in Wilms tumor were some of these mutations shared with the corresponding

tumor. In the presence of clonal nephrogenesis, the VAF distribution of these mutations was significantly elevated (*** $p < 0.001$, Wilcoxon rank-sum test).

(C) Histology images showing components (arrow heads) of the human nephron excised by laser capture microscopy. Tenfold magnification with a 250 micrometre scale bar.

(D) VAF simulations to derive expected distributions depending on clonality of a tissue; monoclonal origin ($\sim \text{Bin}(n \sim \text{Pois}(cov), p = 0.5)$), oligoclonal origin ($\sim \text{Bin}(n \sim \text{Pois}(cov), p = 0.3)$), or polyclonal origin ($\sim \text{Bin}(n \sim \text{Pois}(cov), p = 0.1)$), where cov = coverage = 40X.

(E) VAF distributions for 22 microdissected samples (ten proximal tubules, five distal tubules, seven glomeruli) from three patients, one rapid autopsy donor and two ccRCC patients. Color indicates the underlying maximum likelihood VAF as predicted by a truncated binomial mixture model.

(F) Group-level methylation beta values of *H19* (* $p < 0.05$, Wilcoxon rank-sum test).

(G) Relationship between predicted clone sizes from nephrogenic mutation (see Fig. 2A) and the methylation level of *H19*. The black dot represents PD40738g, which is affected by germline *H19* hypermethylation (omitted from correlation and linear regression).

Figure 3. Phylogenies of bilateral and multifocal Wilms tumor.

(A,C,E,G) For each tumor, the phylogeny of shared mutations is shown including *de novo* germline mutations, embryonic mutations, mutations demarcating clonal nephrogenesis, and tumor mutations. Numbers refer to the number of substitutions defining each developmental trunk. Truncal driver events are detailed.

(B) Heatmap showing the contribution of a mutation to the sample PD40735 shown in (A) (as per legend). The pattern of shared mutations reveals a split between left and right kidney, which is obeyed by both tumor and normal samples.

- (D)** As revealed by the pattern of shared mutations, the left tumor is more closely related to the right branch of clonal nephrogenesis than to the left in PD36159.
- (F)** Two mutations indicate the independent emergence of tumors at different time points from the nephrogenic clone in PD40641.
- (G)** Tumor and nephrogenic rest in PD36165 both originated from clonal nephrogenesis despite being situated at opposing kidney poles.