

## Normal tissue clones preceding childhood kidney cancer

Tim HH Coorens<sup>1</sup>, Taryn D Treger<sup>1,2,3</sup>, Reem Al-Saadi<sup>4</sup>, Thomas J Mitchell<sup>1,2,5</sup>, Suzanne Tugnait<sup>4</sup>, Matthew D Young<sup>1</sup>, Minou Oostveen<sup>4,7</sup>, Patrick S Tarpey<sup>2</sup>, Thomas RW Oliver<sup>1,2,6</sup>, Alex Cagan<sup>1</sup>, Grace Collord<sup>1,2,3</sup>, Yvette Hooks<sup>1</sup>, Mark Brougham<sup>8</sup>, Ben Reynolds<sup>9</sup>, Giuseppe Barone<sup>7</sup>, John Anderson<sup>4,7</sup>, Amos Burke<sup>2,3</sup>, Johann Visser<sup>2</sup>, James D Nicholson<sup>2</sup>, Naima Smeulders<sup>7</sup>, Imran Mushtaq<sup>7</sup>, Kouros Saeb-Parsy<sup>2,5</sup>, Grant D Stewart<sup>2,5</sup>, Peter J Campbell<sup>1</sup>, Michael R Stratton<sup>1</sup>, Iñigo Martincorena<sup>1</sup>, Thomas Jacques<sup>4,7</sup>, Dyanne Rampling<sup>7</sup>, Anne Warren<sup>2,6</sup>, Nicholas Coleman<sup>2,6</sup>, Neil Sebire<sup>4,7</sup>, Karin Straathof<sup>4,7</sup>, Tanzina Chowdhury<sup>4,7</sup>, Jarno Drost<sup>10</sup>, Kathy Pritchard-Jones<sup>4,7</sup>, Sam Behjati<sup>1,2,3\*</sup>

<sup>1</sup>Wellcome Sanger Institute, Hinxton CB10 1SA, UK.

<sup>2</sup>Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK.

<sup>3</sup>Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK.

<sup>4</sup>UCL Great Ormond Street Hospital Institute of Child Health, London WC1N 1E, UK.

<sup>5</sup>Department of Surgery, University of Cambridge, Cambridge CB2 0QQ, UK.

<sup>6</sup>Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK.

<sup>7</sup>Great Ormond Street Hospital for Children NHS Foundation Trust, London WC1N 3JH, UK.

<sup>8</sup>Department of Haematology and Oncology, Royal Hospital for Sick Children, Edinburgh EH9 1LF, UK.

<sup>9</sup>Department of Paediatric Nephrology, Royal Hospital for Children, Glasgow G51 4TF, UK.

<sup>10</sup>Princess Máxima Center for Pediatric Oncology, 3584 CT Utrecht, The Netherlands.

\*Corresponding author

## Normal tissue clones preceding childhood kidney cancer

Some tumours arise from pre-cancerous clonal expansions, which are defined by somatic mutations shared between tumour and precursor lesion. Here, we studied the phylogenetic relation of nephroblastoma (Wilms tumour), the most common kidney cancer of childhood, and paired normal renal tissue. Based on analyses of 172 whole genome sequences derived from 62 tumours and 110 normal tissues, we discovered clones in morphologically normal kidney tissue. These clones were discernible through somatic mutations present in tumour and normal renal tissue whilst absent from blood. Normal tissue clones were found in an initial discovery cohort (2/3 children) and validated in an independent series of children with nephroblastoma (11/20). The emergence of clones, termed clonal nephrogenesis, preceded cancer development. Timing clonal nephrogenesis, from phylogenetic analyses in bilateral tumours, indicated that it can evolve early in embryonic organ development before left and right kidney primordia diverged. Normal tissue clones were not observed in 30 normal tissue biopsies, sampled in the context of clear cell renal cell carcinoma (8 individuals), suggesting that clonal nephrogenesis may be pathological. Our findings reveal normal tissue clones from which nephroblastoma emerges. It is conceivable that such clones form the substrate of recurrent nephroblastoma and of solid childhood tumours more generally.

We set out to define from somatic mutations the relationship between nephroblastoma, the most common childhood kidney cancer, and surrounding normal kidney tissue. Nephroblastoma is a prototypical embryonal cancer that mainly affects infants and children younger than five years of age<sup>1</sup>. It is thought to arise from aberrant embryogenesis, as it resembles fetal nephrogenesis morphologically<sup>1</sup> and transcriptionally<sup>2</sup>. It occurs sporadically in isolation, or in the context of bilateral tumours, multifocal lesions, urogenital developmental disorders or overgrowth syndromes<sup>1</sup>.

We analysed a total of 172 whole genome sequences obtained from 47 individuals: 23 children with nephroblastoma; 16 parents of affected children; and 8 adults with clear cell renal cell carcinoma (ccRCC; **Supplementary Table 1**). We called base substitutions against the reference human genome. From each set of donor related tissues, we extracted mosaic mutations, that is mutations present in two or more tissues at a variant allele frequency (VAF) of less than 50% in at least one tissue (**Methods**). We validated the method for calling mosaic mutations by sequencing parental normal tissue DNA, re-sequencing tissues, and inspection of raw data (**Methods, Supplementary Figure 1**). Based on VAF and distribution of mutations across related tissues, we built phylogenetic trees of tumour development (**Methods**).

The starting point of our analyses were whole genome sequences from three unilateral nephroblastomas, along with blood and morphologically normal kidney tissue (**Supplementary Table 1**). As expected, we found mosaic mutations describing the first few cell divisions of the fertilised egg (**Figures 1a-d**), i.e. variants present in all normal tissues at VAFs of less than 50% (**Supplementary Table 2**)<sup>3,4</sup>. Unexpectedly, however, we also detected mosaic mutations in normal kidneys that were absent in blood, in two cases (**Figure 1e-g, Supplementary Table 2**). These mutations were shared between normal kidney tissue and tumours.

Several features of these mutations indicated that they defined clones in normal kidney tissue, as illustrated by case PD37272 (**Figure 1e-g**). The VAFs of mutations in the normal tissue of this kidney, variants 3-5 (**Figure 1f**), were as high as 44%, suggesting that the mutation was significantly present in 88% of all cells in the biopsy. Mutations 3-5 were significantly present ( $p < 0.001$ ; see **Methods**) in the two parenchymal biopsies (i.e. cortex and medulla) whilst absent from blood DNA, deeply sequenced to 106X genome-wide (**Figure 1g**). Similarly, mutations 3-5 were undetectable in renal pelvis, which is embryologically derived from a different lineage than kidney parenchyma<sup>5</sup>. Furthermore, the VAF of early embryonic mutations, variants 1 and 2, was inflated in parenchyma and in tumours (**Figure 1f,h**). Such inflation of early embryonic mutations are a feature of tissues that contain a sizeable clone derived from a single cell (**Figure 1c,d**). By contrast, in tissues devoid of a major clone, such as renal pelvis and blood, the VAFs of early embryonic mutations are not inflated (**Figure 1b,h**). Thus, these normal tissue variants 3-5 delineate a clonal zone within kidney parenchyma, which we termed clonal nephrogenesis, accounting for up to 88% of cells sampled in the cortex.

An alternative explanation for these findings could be infiltration of normal tissue by tumour cells or cross-contamination of DNAs. This explanation is, however, implausible. Contamination would manifest as shared variants at a low VAF, rather than select mutations at a high VAF. We formally excluded possible contribution from tumour infiltration and contamination by employing a binomial mixture model on the observed base counts of tumour mutations in the normal samples (**Supplementary Figure 2, Methods**).

Next, we investigated whether clonal nephrogenesis might represent the normal clonal architecture of human nephrons. If nephrons were normally derived from only one or a few cells that are distant from the blood lineage, there would always be shared mutations between parenchymal kidney biopsies and nephron-derived tumours. To test this possibility, we analysed newly generated and previously published<sup>2,6</sup> whole genome sequences from another nephron-derived tumour, adult ccRCC. We studied a total of 49 samples, 19 tumours and 30 normal tissues. Included were one case of bilateral ccRCC (PD24242) and one resection specimen from which we sampled 11 normal tissues (PD41697). Applying the same analysis strategy did not detect clones in normal kidney tissue (**Figure 2a**). This suggests that the physiological clonal dynamics of fetal kidney development do not commonly generate major clones in kidney parenchyma. The normal tissue clones we observed in nephroblastoma may therefore represent abnormal clonal expansions associated with the development of nephroblastoma. It is, however, possible that clonal nephrogenesis is a common feature of normal paediatric kidneys that vanishes with age. This would require that potentially large areas of kidney involute and regenerate during childhood, which is difficult to entertain given that new glomeruli do not form in the human kidney after birth<sup>7</sup>.

To further investigate and validate our discovery of clonal nephrogenesis, we studied another 20 cases of nephroblastoma: 15 unilateral tumours with normal tissue biopsies curated through a British nephroblastoma study, four cases of bilateral nephroblastoma, and one extensively sampled tumour with ten normal tissue biopsies (**Figure 2b-e**). Considering the entire cohort of 23 children, we found evidence of clonal nephrogenesis in 9/19 children with unilateral nephroblastoma and in 4/4 children with bilateral tumours (**Methods, Figure 2b-e**).

The presence of clonal nephrogenesis in these tumours was further substantiated by the inflation of VAFs of early embryonic variants (**Figure 2f**).

The number of substitutions defining clonal nephrogenesis ranged from 1 to 11 variants per case (median 4; **Figure 2g**). Mutations of clonal nephrogenesis were predominantly C>T transversions in a sequence context attributable to ubiquitous endogenous mutational processes (**Supplementary Figure 3**)<sup>8,9</sup>. The VAFs of these nephrogenic mutations varied from 0.04 to 0.44 (median 0.13). A remarkable clone size was found in PD40735, where a single nephrogenic clone accounted for 75% and 85% of left and right normal kidney samples, respectively (**Figure 2d**). No copy number changes were detected in normal kidney tissues.

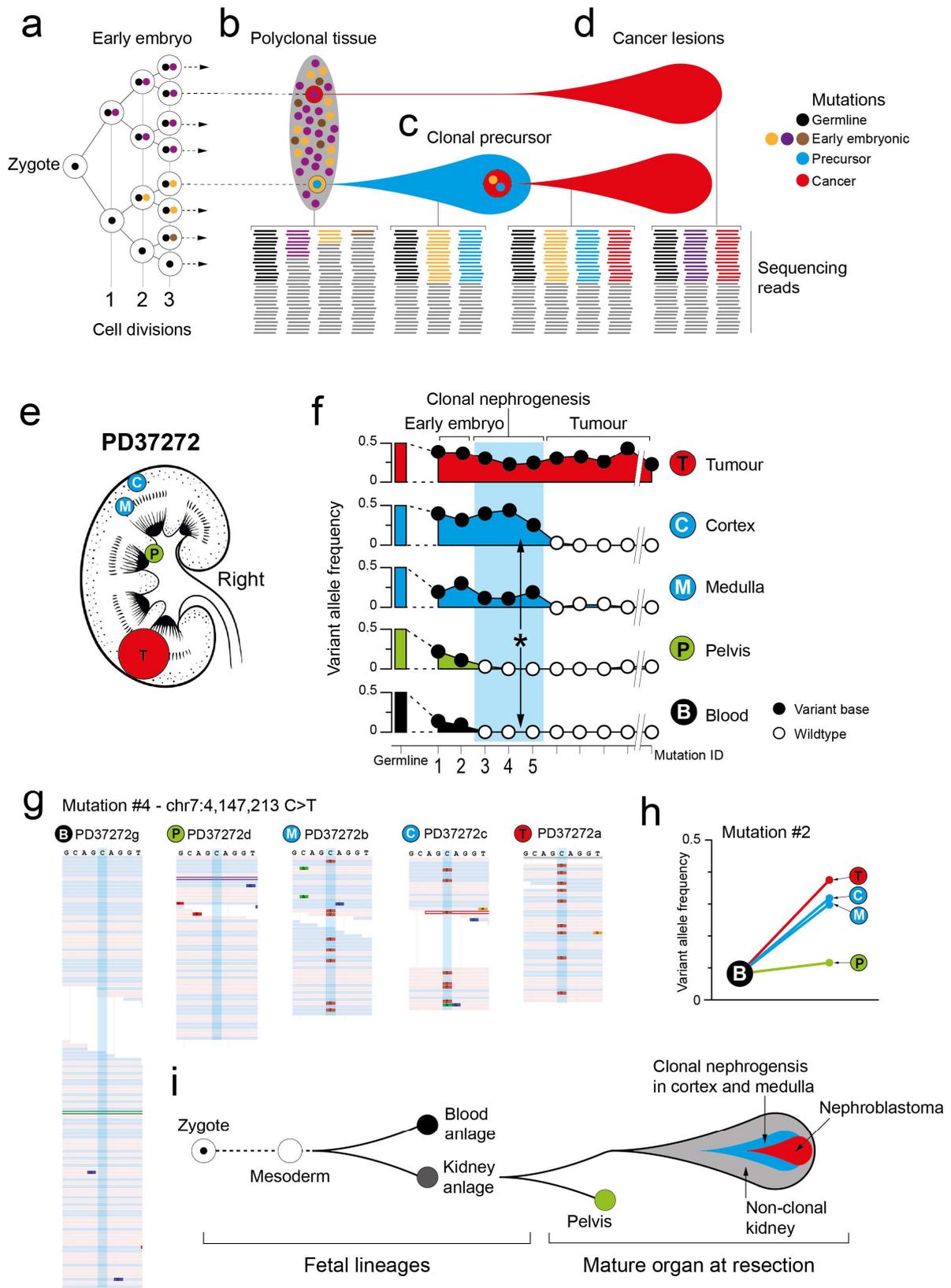
The timing of the emergence of clonal nephrogenesis during human development could be defined in three children from whom we obtained bilateral tumours. In two cases, left and right tumours were derived from the same trunk of clonal nephrogenesis (**Figure 3a-e**). Likewise, clonal nephrogenesis found in normal tissues bilaterally in these two children originated from a common stem (**Figure 3a,c**). PD36159 is a particularly striking example (**Figure 3c-e**). Here, three left tumour samples were phylogenetically more closely related to the right normal kidney than to the left. In the third child, PD40378, all four left tumour samples were related to clonal nephrogenesis on the ipsilateral side (**Figure 3f**). The right neoplasms, a tumour and a nephrogenic rest, were not derived from clonal nephrogenesis on the left. No normal tissue of the right side was available for analysis. Taken together, these findings indicate that in two children with bilateral tumours, clonal nephrogenesis must have arisen before left and right kidney primordia diverged, early in embryogenesis<sup>5</sup>. In unilateral tumours, 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” in which clonal nephrogenesis replaced normal kidney tissue.

In the five cases from which we sampled multiple neoplasms of the same kidney, two configurations of tumour development were seen (**Figure 3**). Tumours were either derived from a shared trunk that had emerged from clonal nephrogenesis (e.g., PD40735, **Figure 3a**). Or, tumours arose independently and successively from clonal nephrogenesis, alluding to a sustained potential for tumour budding from normal tissue clones. For example, PD36165 presented with two kidney lesions, a nephrogenic rest and a nephroblastoma. Tumour and rest, situated at opposing poles of the kidney, had emerged from the same nephrogenic clone, yet at different time points, followed by clonal diversification (**Figure 3h,i**) within tumour and rest.

A central question that our findings raise is whether distinct cancer causing (driver) mutations underpin clonal nephrogenesis. We annotated germline mutations, mosaic mutations, and variants of clonal nephrogenesis to identify plausible driver events, in addition to reviewing clinical genetics reports of children with a suspected predisposition to nephroblastoma. We detected pathogenic germline events (**Supplementary Table 3**) in two cases, one child with, and one child without clonal nephrogenesis. In the remainder of cases we did not find plausible drivers events explaining clonal nephrogenesis. It is possible that cryptic DNA mutations, epigenetic events or a stochastic clonal drift underlie clonal nephrogenesis.

We discovered sizeable clones in normal kidney tissue that preceded the development of unilateral and multifocal nephroblastoma. Our findings portray nephroblastoma as an insurrection on the background of a premalignant tissue bed, rather than a clearly demarcated neoplasm in an otherwise normal, polyclonal kidney. Our observations indicate that clonal nephrogenesis may be the source of recurrence in nephroblastoma treated with partial nephrectomy. We speculate that embryonal expansion of normal tissue clones may be a common phenomenon in childhood cancer.

**Figure 1**



## Figure 1. Clones in normal human kidneys.

**(a)** Schematic tree of early embryogenesis showing that somatic mutations are acquired in post-zygotic divisions and retained by the progeny of the cell.

**(b)** These early embryonic mutations are present in an exponentially decreasing variant allele frequency in polyclonal tissue. The ability to detect these mutations in bulk tissue therefore is limited by the sequencing depth. For example, at a typical depth of 30X genome wide, theoretically only the first four cell divisions can be determined, unless a clonal expansion occurs **(c)**.

**(d)** Tumours, which are clonal outgrowths, can either evolve directly from a single cell within polyclonal tissue or emerge from an expanded clonal precursor.

**(e)** Overview of tissue sampling in the kidney of PD37272.

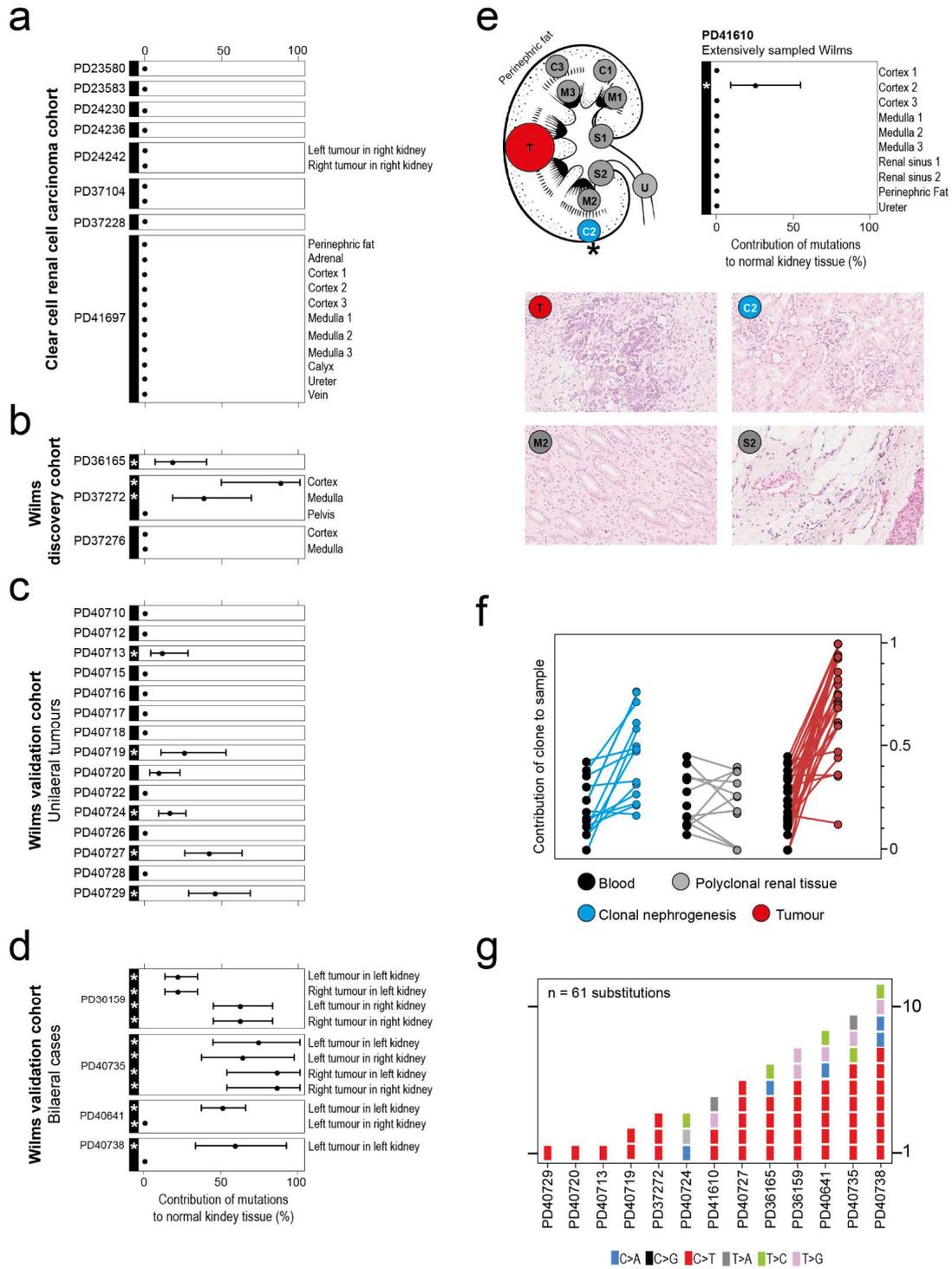
**(f)** Variants in the trunk of the tumour can be detected in normal tissues. If the mutation is present in tumour, kidney, and blood, it is classified as early embryonic. If it is present in kidney samples and tumour only, it is clonal nephrogenic. If it is only in the tumour, it is labelled as such. \* $p < 0.001$  (test of presence using beta-binomial overdispersion, **Methods**) White and black circles indicate whether the observed VAF is insignificant (white) or significant (black).

**(g)** Raw reads showing mutant locus across samples. Blue highlight shows variant locus.

**(h)** The VAF for the last embryonic mutation in kidney samples and tumour compared with blood.

**(i)** Schematic showing the development lineage of blood and kidney, and subsequently the pelvis and kidney parenchyma during early embryogenesis. Afterwards, clonal nephrogenesis occurred within the corticomedullary region from which the nephroblastoma tumour arose.

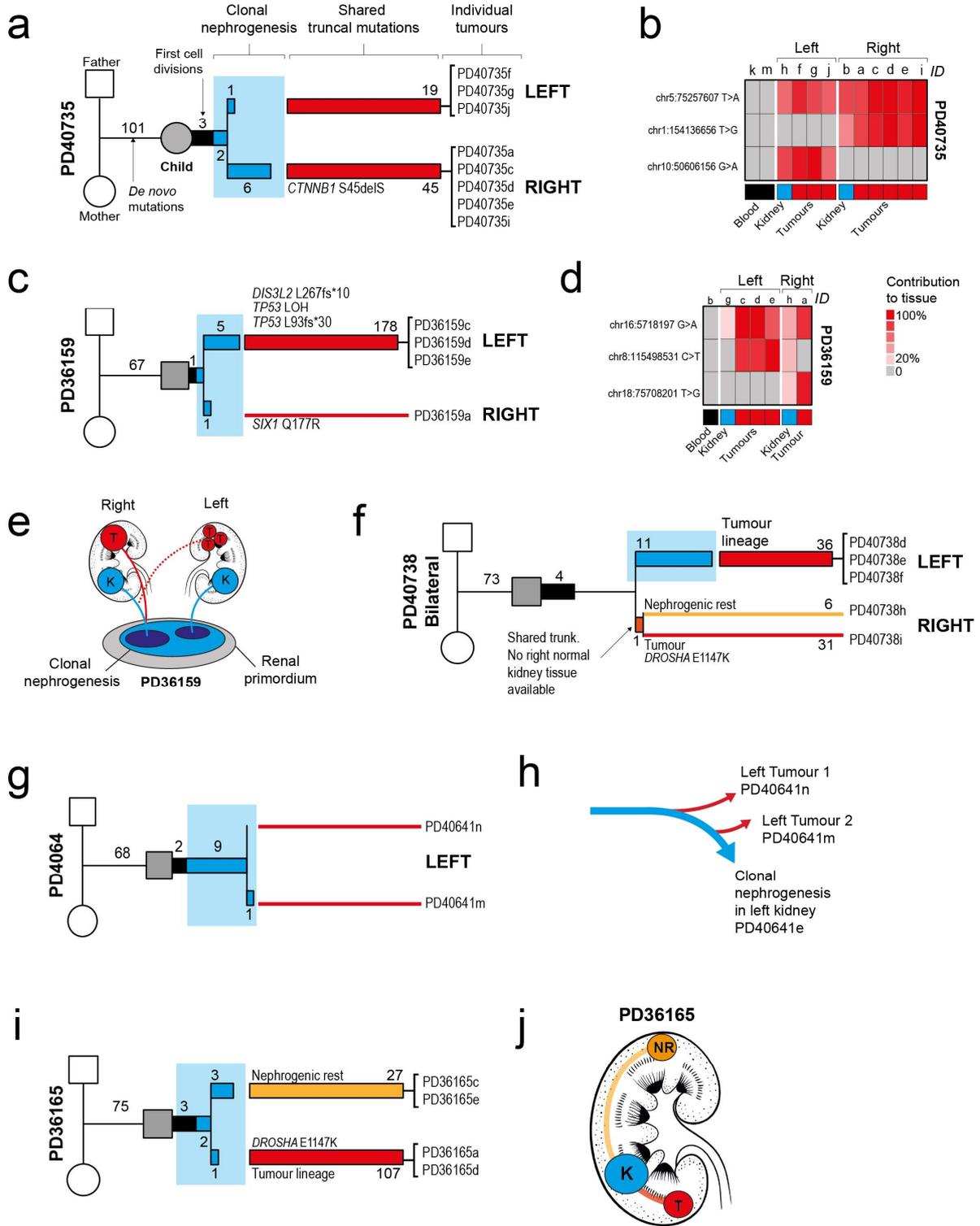
**Figure 2**



## Figure 2. Clonal nephrogenesis across cohorts.

- (a-d)** Presence of clonal nephrogenesis in kidney samples of various cohorts. Kidneys in which nephrogenic mutations were significantly present are highlighted with an asterisk (see **Methods** for statistical assessment). From the VAF of the estimated earliest nephrogenic mutation in these kidney samples we approximated the contribution of this ancestral tumour clone to the kidney sample (with 95% confidence interval, **Methods**).
- (e)** Overview of the extensive spatial sampling of kidney PD41610 with histology for selected sampling sites.
- (f)** Clones can also be detected using the VAF of the last embryonic variant. This plot shows the contribution of the last embryonic mutation in tumour and in samples with and without clonal nephrogenesis, alongside the contribution to blood.
- (g)** Mutation burden of clonal nephrogenesis by sample and mutation type.

**Figure 3**



### Figure 3. Bilateral and multifocal nephroblastoma.

For each tumour the phylogeny of shared mutations is shown including *de novo* mutations, embryonic mutations, clonal nephrogenesis, and tumour trunk (where available). Numbers refer to the number of substitutions defining each developmental segment. Driver mutation in tumour trunk are detailed. Driver mutations present in individual tumours, rather than the trunk, are not shown here. Heatmaps **(b and d)** show the contribution of a mutation to the sample in question (as per legend). **(e)** The left tumour is more closely related to the right branch of clonal nephrogenesis than to the right in PD36159. **(f)** Phylogenetic tree of tumours and rest in PD40738. Note that the right tumour and rest do not come from the nephrogenic clone that spawned the left tumour. **(g and h)** The emergence of tumours at different time points from clonal nephrogenesis in PD40641. **(i)** Tumour and nephrogenic rest in PD36165 both originated from clonal nephrogenesis despite being situated at opposing kidney poles **(j)**.

## References

1. Treger, T. D., Chowdhury, T., Pritchard-Jones, K. & Behjati, S. The genetic changes of Wilms tumour. *Nat. Rev. Nephrol.* 1 (2019). doi:10.1038/s41581-019-0112-0
2. Young, M. D. *et al.* Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science (80-. ).* **361**, 594–599 (2018).
3. Behjati, S. *et al.* Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
4. Ju, Y. S. *et al.* Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature* **543**, 714–718. (2017)
5. Short, K. M. & Smyth, I. M. The contribution of branching morphogenesis to kidney development and disease. *Nat. Rev. Nephrol.* **12**, 754–767 (2016).
6. Mitchell, T. J. *et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**, 611–623.e17 (2018).
7. Saxén, L., & Sariola, H. Early organogenesis of the kidney. *Pediatric nephrology*, **1**(3), 385-392 (1987).
8. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–21 (2013).
9. Alexandrov, L. B. *et al.* Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015).
10. Li, H., & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**(14), 1754-1760 (2009).
11. Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*, **534**(7605), 47054. (2016).
12. Jones, D. *et al.* cgpCaVEManWrapper: simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Current protocols in bioinformatics*, **56**(1), 15-10. (2016).
13. Ye, K. *et al.* Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, **25**(21), 2865-2871 (2009).
14. Van Loo, P. *et al.* Allele-specific copy number analysis of tumors. *Proceedings of the National Academy of Sciences* **107**(39), 16910-16915 (2010)

15. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994-1007 (2012)
16. 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)*, 289-300 (1995).
17. Gerstung M. *et al.* Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics*, **30**(9), 1198-1204 (2014).
18. Buels, R. *et al.* JBrowse: a dynamic web platform for genome visualization and analysis. *Genome biology*, **17**(1), 66 (2016).
19. Clopper, C. J., & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, **26**(4), 404-413 (1934).
20. Behjati, S. *et al.* Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. *Nature communications*, **8**, 15936 (2017).
21. Forbes, S. A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic acids research*, **43**(D1), D805-D811. (2014).
22. Alexandrov, L. *et al.* (2018). The Repertoire of Mutational Signatures in Human Cancer. *bioRxiv*, **322859**.
23. Rosenthal, R. *et al.* DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome biology*, **17**(1), 31 (2016).

## ACKNOWLEDGEMENTS

We thank for their help research nursing staff at Cambridge University Hospitals, the Royal Hospital for Sick Children, Edinburgh, and the Royal Hospital for Children, Glasgow. We thank IMPORT investigators for patient recruitment and 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.). Additional funding was received from CRUK (IMPORT study) and NIHR (BCR at Great Ormond Street).

**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 and I.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. Samples were curated and/or experiments were performed by R.A., S.T., M.O., Y.H., M.B., B.R., G.B., J.A., A.B., J.V., J.D.N., N.S., Im.M., K.S-P., G.D.S., K.S., T.Ch., J.D., and K.P-J. Pathological expertise was provided by T.RW.O., T.J., D.R., A.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.

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