© 2017. Published by The Company of Biologists Ltd | Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

RESEARCH ARTICLE

# Branching morphogenesis in the developing kidney is governed by rules that pattern the ureteric tree

**James G. Lefevre[1,*], Kieran M. Short[2,3,*], Timothy O. Lamberton[1,*], Odyssé Michos[4], Daniel Graf[5], Ian M. Smyth[2,3,6,‡] and Nicholas A. Hamilton[1,‡]**

## ABSTRACT

Metanephric kidney development is orchestrated by the iterative branching morphogenesis of the ureteric bud. We describe an underlying patterning associated with the ramification of this structure and show that this pattern is conserved between developing kidneys, in different parts of the organ and across developmental time. This regularity is associated with a highly reproducible branching asymmetry that is consistent with locally operative growth mechanisms. We then develop a class of tip state models to represent elaboration of the ureteric tree and describe rules for 'half-delay' branching morphogenesis that describe almost perfectly the patterning of this structure. Spatial analysis suggests that the observed asymmetry may arise from mutual suppression of bifurcation, but not extension, between the growing ureteric tips, and demonstrates that disruption of patterning occurs in mouse mutants in which the distribution of tips on the surface of the kidney is altered. These findings demonstrate that kidney development occurs by way of a highly conserved reiterative pattern of asymmetric bifurcation that is governed by intrinsic and locally operative mechanisms.

**KEY WORDS: Branching morphogenesis, Ureteric tree, Mathematical modelling**

## INTRODUCTION

The increase in organ size associated with the evolution of complex body plans presents challenges for the organism in mediating the efficient exchange of nutrients and the removal of waste. To effectively 'plumb' tissues, animals have evolved a developmental mechanism known as branching morphogenesis that serves to break up large tissues to maintain their effective surface area to volume ratio. The branching process typically initiates through the formation of bud-like organ anlage, which invades (and ramifies within) an associated mesenchymal field to form complex networks of epithelial tubes that ultimately allow for cellular exchange. Branching morphogenesis is pervasive in higher eukaryotes and is crucial for the development of organs such as the kidney, lung and mammary gland. In the metanephric kidney, organogenesis begins with the formation of the ureteric bud (UB), which appears as an outgrowth of the nephric duct at E10.5 in mice and during the fifth week of gestation in humans (Saxen, 1987). The UB then grows to contact and invade a surrounding field of cells known as the metanephric mesenchyme, which in later development condenses around the tips of the ureteric tree to form discrete cellular fields known as the cap mesenchyme (CM). The UB tips require the presence of CM cells to undergo ongoing branching (Self et al., 2006). In turn, the CM responds to tip signals to control the self-renewal and commitment of precursor populations to form the nephrons (Kreidberg et al., 1993), which will ultimately integrate with the ureteric tree to form the urine collecting system of the adult organ. Reciprocal epithelial-mesenchymal exchange of this nature is a feature of almost all developmental branching morphogenesis and is central to the functional differentiation of cells in both compartments.

A central question in the study of branching morphogenesis in the kidney is the degree to which stereotypy characterizes elaboration of the epithelium. Do epithelial branching events occur at predetermined places and at defined developmental stages to form a morphologically distinct structure? A landmark study of the developing mouse lung describes a highly predictable program that features three morphologically distinct branch types that are reproducibly employed in a program which establishes a precise pulmonary airway architecture (Metzger et al., 2008). Although this suggests a hard-wired genetic control of the branching process, other reports have argued for a more adaptive mechanism emerging from pulmonary epithelial-mesenchyme interactions (Blanc et al., 2012). Mathematical models have been developed based on such known interactions to explain the process of tip bifurcation in the lung (Menshykau et al., 2012) and kidney (Menshykau and Iber, 2013), and the induction of the ureteric bud (Lawson and Flegg, 2016). These models, which are based on a Schnakenburg-like Turing patterning arising from competition for mesenchymal signal, have so far not been able to explain the broad patterning of the epithelial trees. In the kidney, analysis of epithelial ureteric branching in culture suggests self-avoidance in the growing tips – observations that support an adaptive branching mechanism operating in the organ (Davies et al., 2014). This is consistent with our previous studies that have shown that although branching of the ureteric tree is not spatially stereotypical, its elaboration (based on hierarchy) is indeed highly reproducible (Short et al., 2014). That is to say, while specific branches do not form in the same 3D space at the same developmental time, the extent of ramification of the ureteric tree is highly similar between one organ and the next. These observations suggest that a constrained program of patterning drives the elaboration of the UB.

[1]Division of Genomics and Development of Disease, Institute for Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia. [2]Biomedicine Discovery Institute, Monash University, Clayton, Melbourne, Victoria 3800, Australia. [3]Department of Anatomy and Developmental Biology, Monash University, Clayton, Melbourne, Victoria 3800, Australia. [4]Department of Biosystems, Science and Engineering (D-BSSE), ETH Zurich, Basel 4058, Switzerland. [5]School of Dentistry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta T6G 1C9, Canada. [6]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Melbourne, Victoria 3800, Australia.
*These authors contributed equally to this work

‡Authors for correspondence (ian.smyth@monash.edu; n.hamilton@imb.uq.edu.au)

J.G.L., 0000-0002-5945-9575; K.M.S., 0000-0001-9137-008X; O.M., 0000-0002-0002-4315; D.G., 0000-0003-1163-8117; I.M.S., 0000-0002-1727-7829; N.A.H., 0000-0003-0331-3427

**Received 28 April 2017; Accepted 5 October 2017**

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

In order to understand how mouse ureteric tree patterning is repeatable, we have combined three-dimensional imaging, spatial quantification and mathematical modelling to develop a class of state-based mathematical models that describe the bifurcation and elaboration of ureteric trees. We then compared our models that best describe normal branching against ureteric trees from mice carrying mutations in genes thought to be crucial to shaping the branching process. Although highly stereotypical, the self-similar nature of the branching pattern we describe indicates that branching occurs under the influence of local rules, which is consistent with an emergent pattern arising from interaction between neighbouring epithelial and mesenchyme niches. We also show that the characteristic asymmetry of branching in both wild-type and mutant kidneys is linked to the relative positioning of the tips, which is consistent with inter-tip suppression of bifurcation. Taken together, these findings show that elaboration of the ureteric tree occurs by way of a previously unrecognized but highly reproducible pattern of branching that contributes to the form and structure of the kidney.

## RESULTS

### Branching asymmetry is consistent across location and age

We imaged ureteric trees from embryonic mice using optical projection tomography and quantified branching using Tree Surveyor software (Short et al., 2013). We analysed wild-type patterning in 32 C57BL/6J embryonic ureteric trees from E11.5 to E15.5. In observing the ureteric tree pattern, we noticed that each bifurcation led to significantly more tips from one child branch than the other, in a ratio that was broadly invariant. To quantify this, we defined the weight of a bifurcating branch point as the number of tips descending from it and defined its balance as the ratio between the larger and smaller child weights (Fig. 1A). We have previously proposed that the initial branches that establish the basic six-clade structure of the organ are distinct from further branching within clades (Fig. 1B) (Short et al., 2014). To examine this, we used our measures of weight and balance to study the root node as well as root branch points that establish the five ancestor and six-clade bifurcations that define the basic structure of the kidney. We refer to these as the 11 primary branch points and consider them separately from the other intra-clade branches that make up the vast majority of the ureteric tree (Fig. 1B). Branching at the ureter root node is significantly more symmetrical than in later generations (balance=1.18±0.18, mean±s.e.m.), with no anterior or posterior bias ($P$=0.495, paired $t$-test on log weight of anterior and posterior), whereas the next branches are substantially more asymmetrical than later generations (balance=2.21±1.20), principally because they define the clade structure of the organ (Fig. 1C). The remaining primary branch points have very similar balance (1.49±0.46, Fig. 1D). With the exception of the earliest developmental stage, balance at each of the primary branch points remains the same across developmental time ($P$>0.24 for linear regression against age as continuous or nominal variable, controlling for clade).

We next examined the balance of branches within individual clades, which make up the majority of the mature ureteric tree. In contrast to the variation we saw between many of the primary branches, the balance within clades was remarkably similar when comparing clades from all organs as a single dataset (balance=1.25±0.17, Fig. 1E): i.e. on average, the balance of any branch point within any clade is no different to any other branch point in another clade ($P$=0.98, ANOVA). The extensive reorganisation and branch resorption associated with the formation of the renal pelvis (Short et al., 2013) precludes the production of a ‘normally’ arborized ureteric after E15.5 (i.e. early branch generations are thereafter missing). However, to examine whether balance is maintained later in development, we investigated peripheral branching in four E16.5 kidneys. These organs have, on average, 1292 tips, which means that they are approximately one generation away from the cessation of branching. Using these data, we quantified balance in the three most distal branching generations (inverse generation 2,3,4 by our schema) and found no significant difference compared with younger organs (mean balance=1.252, $P$=0.1867 for a $t$-test on kidney means). This indicates that branching within clades occurs through the establishment of a constrained, asymmetrical and self-similar pattern of elaboration at least until the very latest stage of branching morphogenesis. The conservation of balance over time, irrespective of location within the organ, suggests that the elaboration of the ureteric tree occurs in a highly controlled manner after the initial establishment of basic organ pattern (through the primary branch points).

### Tip state models

The conserved branching asymmetry we observe suggests that a consistent, locally operative mechanism defines a consistent pattern of branching morphogenesis that drives renal development. Although other biological systems such as the mammary gland have been successfully modelled using stochastic approaches (Scheele et al., 2017), after consideration these were found to not be appropriate in the case of the kidney, for reasons outlined later in the Discussion. Instead, we developed theoretical ‘tip state’ models to represent the evolution of the branching tree over time, with a view to describing our biological data. These models posit that a given tip exists in a state governed by rules that define whether it either matures or bifurcates as development progresses and what state it next assumes. The simplest such model is the unique non-trivial ‘single state’ model giving a family of trees with uniform bifurcation, which we call perfect. Because by definition this cannot model the systematic asymmetry observed in the analysis of our experimental models, we then considered ‘two-’ and ‘three-state’ models. There are two distinct two-state models and 34 distinct three-state models (see supplementary Materials and Methods for details and a classification of all models with up to three states).

We assessed all of these and found only three that have a single bifurcating state with non-terminating child states. Such generalized delay models have the desired property that balance converges to a single asymptotic value as weight increases. Two of these were considered plausible candidates for modelling the ureteric tree: a two-state model, which we term the ‘Fibonacci’, in which the weight of all trees and sub-trees are Fibonacci numbers (Fig. 2A); and a three-state model that we term the ‘half delay’, the behaviour of which is outlined in Fig. 2B. The Fibonacci model has asymptotic balance ≈1.618 (the golden ratio), while the half-delay model has asymptotic balance ≈1.325. The third delay model has asymptotic balance ≈2.148, which is outside the observed range in ureteric trees. In all cases, the smaller trees produced by these models have balance values that differ from the asymptote and from each other, owing to a discrete weight effect.

### The half-delay model provides best global fit to ureteric trees

To define how these models of branching describe the behaviour of the ureteric bud *in vivo*, we compared the Fibonacci and half-delay models with our biological dataset using an optimal overlay algorithm (Lamberton et al., 2015), using the perfect model as a baseline. This approach computes a matching score between 0 and 1 for each pair of tree structures (Fig. 3A,B). We applied this method by clade, such that for every ureteric tree each of the six clades was

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

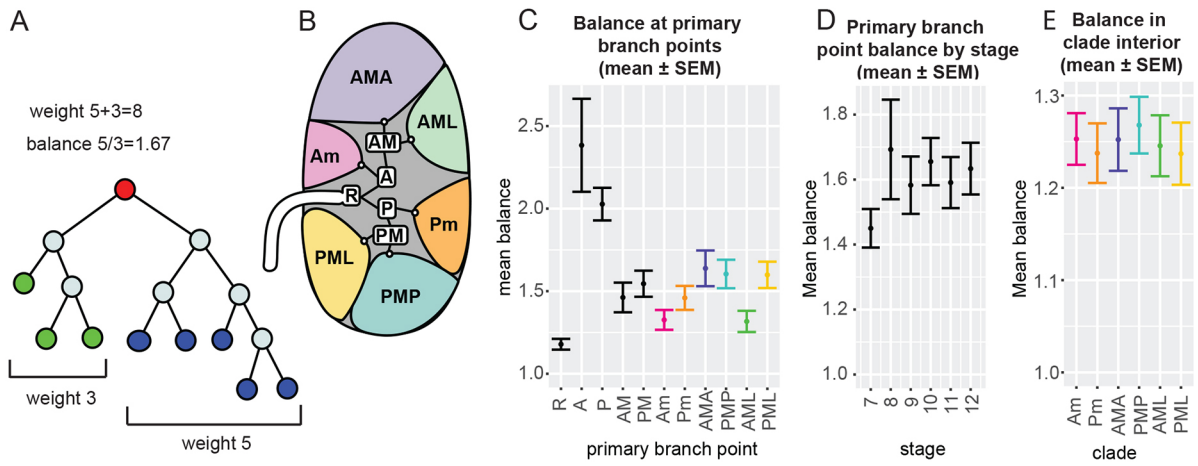



**Fig. 1. Quantification of branching asymmetry in developing kidneys.** (A) Branching asymmetry is quantified at each binary branch point (red) as the ratio between the larger (blue) and smaller (green) child sub-tree weight (number of tips). (B) Balance is calculated for the 11 primary branch points that define the six-clade structure. (C) Primary branch point balance ($n$=32,32,32,32,32,30,31,31,30,30,29 for groups in order listed). (D) Average primary branch point balance by limb stage ($n$=56,54,77,55,55,44 for stages 7-12, approximate ages E11.5, E12.5, E13.25, E13.75, E14.5, E15.5). (E) Average balance within each clade ($n$=27,29,30,29,26,28 for clades in order listed). Data are mean±s.e.m. calculated across kidneys. P, posterior; A, anterior; M, major; m, minor; L, lateral. In C and E, data are pooled over stages 7-12, including each branch point or clade in which branching has occurred.

overlaid with the closest tree by size from the selected model. We then calculated the overall matching score for the whole tree. When such a comparison was performed using our wild-type datasets, the best observed fit was the half-delay model that had a matching score between 0.74 and 1.0 depending on developmental stage; Fig. 3C shows an example of a subtree with an exact matching to the half-delay model. To understand the nature of the small amount of discordance between the model and *in vivo* branching of the ureteric bud, we undertook a program of pairwise matching between the model and wild-type data. For each developmental stage and clade type, we isolated the corresponding set of clades then identified the half-delay tree with best average matching score and added this tree to the set. We asked which member of this augmented set was best matched to all of the other members and then termed this the centroid (Lamberton et al., 2015). Using this approach we found that of the 36 cases (six clades and six developmental stages), the half-delay model was the centroid in 29 (80.6%), was the second ranked in six cases and third in only one (Fig. S1). This indicates that the model trees are 'in the middle' of the experimental ones that vary around them and that the half-delay model almost perfectly represents the majority of branching morphogenesis in the ureteric tree.

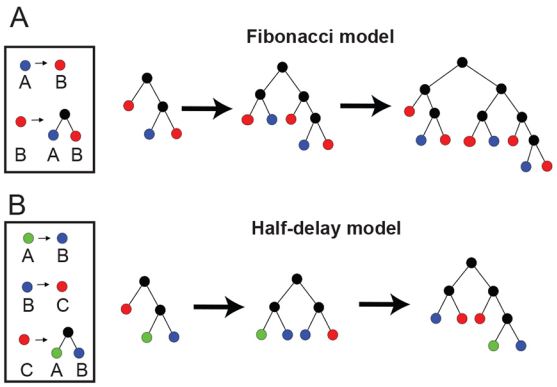



**Fig. 2. Tip state models.** (A) Tip state rules and initial sequence for the Fibonacci model. (B) Tip state rules and initial sequence for the half-delay model. Colours represent tip states; branch points have no state and are shown in black. See supplementary Materials and Methods for formal definitions.

### Sub-tree analysis provides a detailed developmental pathway of branching morphogenesis

To define the normal developmental program that gives rise to the mature ureteric tree, we took advantage of our datasets to capture snapshots of the average branch pattern. The frequency of each subtree type was calculated across the entire dataset and the centroid for each weight identified as the single most representative structure (Fig. 4, top row). Although multiple tree structures occur at all weights above three, in each case the distribution is dominated by either one or two patterns (Fig. 4, for weights up to 12). In the latter instance, this always represents only a transient divergence. Strikingly, for the subtree weights that can be described by a half-delay tree model, the half-delay model dominates the distribution for that weight (see Fig. 4 weights 3, 4, 5, 7, 9 and 12). Similar calculations at weights 16 and 21 (the other half-delay examples) also found that that the centroid trees are the half-delay models of those weights (Fig. S2). This indicates a definable sequence of growth in the kidney, with the addition of a bifurcation at each step (green in Fig. 4) that is centred on the half-delay model.

### Geometry of asymmetrical branching

One explanation for the asymmetry we have observed in whole clade and subtree analyses is that bifurcation of a given tip may be locally inhibited by its neighbours. If this were the case, we would expect the most closely related tips in the tree structure to provide the greatest local influence on growth and bifurcation, and hence expect a relationship between sibling node imbalance and the position of the next most closely related tips. To test this hypothesis, we analysed all pairs of sibling nodes with unequal weight (32 kidneys, 1963 pairs) and identified the common parent node p and its sibling, p′ (Fig. 5A,B). In the original pair of sibling nodes, we call the node that is spatially closest to p′ the inner sibling (b), which corresponds to the angle $\theta_1$ being greater than $\theta_2$ (Fig. 5A,B). If inter-tip suppression is driving asymmetry, we would expect the inner sibling to face stronger inhibitory signals and to have a lower weight (i.e. less branching). Analysis of these sibling pairs

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

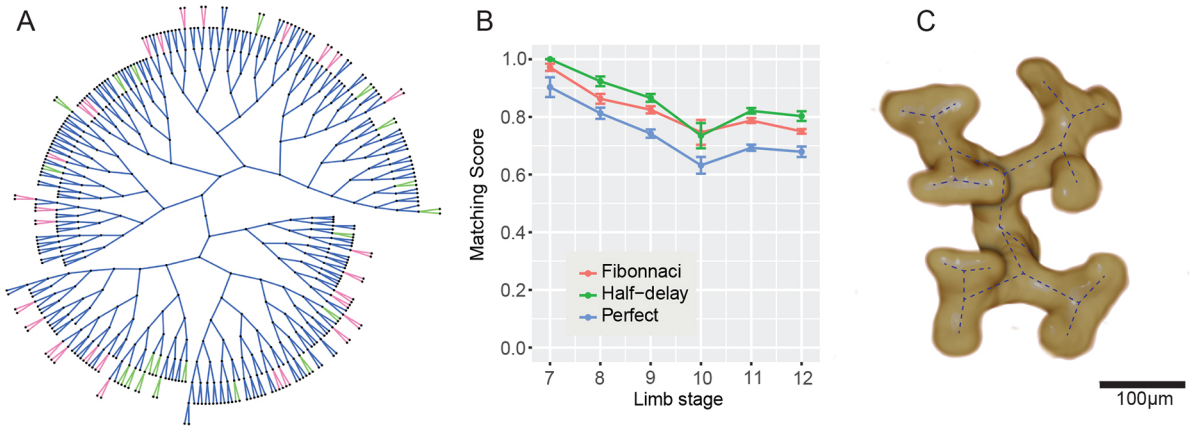



**Fig. 3. Overlay analysis supports half-delay model.** (A) Example optimal overlay between E14.5 ureteric tree and half-delay model (by clade); overlap is shown in blue (536 edges), model-only edges in green (40), real tree-only edges in red (58). Matching score is 536/(536+40+58)=0.845. (B) Overlay analysis of biological data against models shows the best match between ureteric tree structure and the half-delay model, out-performing the Fibonacci and perfect models ($n$=6,5,7,5,5,4). (C) Sample weight 12 subtree from stage 11/E14.5 wild-type data with half-delay structure. This structure is shown schematically in Fig. 4 (weight 12, far right).

described such a scenario in 69.9% of cases ($n$=1963), supporting this hypothesis. In addition, nodes are not generally co-planar, as depicted in Fig. 5, meaning that a and b could be almost equidistant from p′ (Fig. 5B). To correct for this we defined near-planar branches where the inner sibling is more strongly 'between' p′ and a as those where the $\theta_1/\theta_2$ difference is above 45°. In these cases, the proportion in which the inner sibling weight is lower rises to 83.5% ($n$=571), a relationship that is apparent at all of the developmental stages we examined (>66% at each stage).

To test whether the delayed bifurcation of the inner sibling is just the result of generally slowed growth, we compared the branch lengths in pairs of near-planar sibling nodes (in Fig. 5A,B we compare the lengths from p to a and p to b). In cases where both siblings have bifurcated, the inner sibling is longer in 65.5% of cases (median difference 14.6 µm or 19.5%), indicating that a failure to branch resulted in continued lengthening. To further clarify this observation, we considered node pairs in which neither sibling has branched (i.e. a and b in Fig. 5A are both tips). Doing so allowed us to consider tip extension in isolation from bifurcation and in such cases little difference in length was seen: the inner child is instead shorter in close to half of cases (56.9%) and the median difference is much smaller than in siblings that have bifurcated (only 3 µm or 5%). Based on these findings, we conclude that the local cues that we propose to inhibit tip bifurcation have little or no impact on the rate of tip extension.

### Divergence from normal patterning in mutant phenotypes

We next wished to address how the pattern of branching morphogenesis we observed during kidney development is established, with a view to understanding its genetic basis and robustness in the face of perturbation. On the basis that local cues are likely to influence patterning, we analysed mutants in which the environment of tips on the organ surface are altered. The first selected was bone morphogenetic protein 7 (*Bmp7*$^{-/-}$), an inhibitor of nephron progenitor cell differentiation (Tomita et al., 2013; Muthukrishnan et al., 2015) produced by cells in the UB epithelia

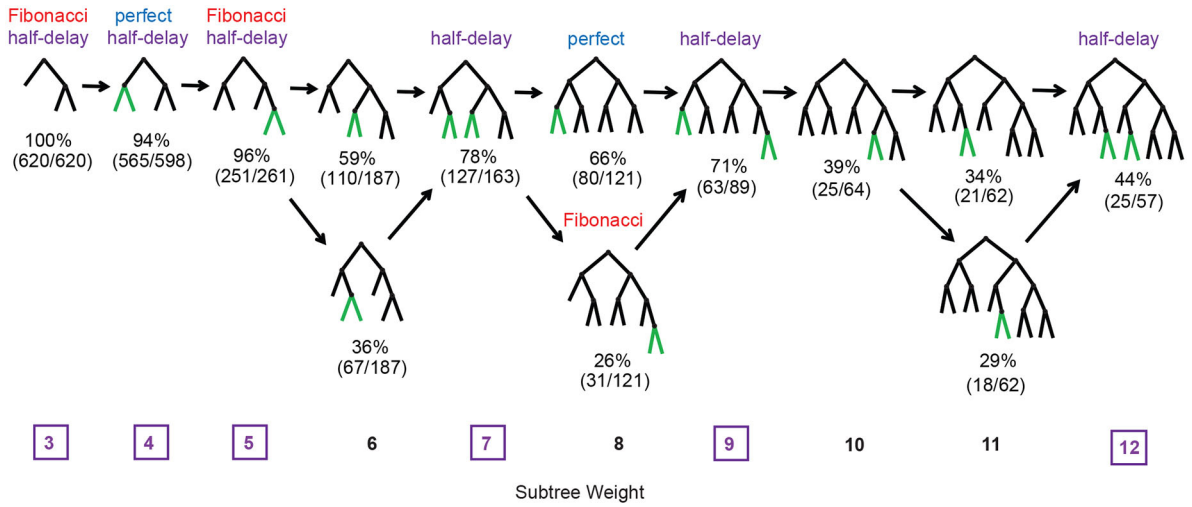



**Fig. 4. Sub-tree analysis of wild-type kidney sub-trees indicates that stereotypical growth closely matches half-delay model.** Within each weight group (columns), tree structures with relative frequency >10% are shown, annotated with the relative frequency and whether they are produced by the perfect, Fibonacci or half-delay model (some structures are produced by more than one model). Weights for which a half-delay model exist are marked in purple. Centroid trees are on the top row. Arrows indicate potential developmental pathways. New edges at each stage of this pathway are shown in green.

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

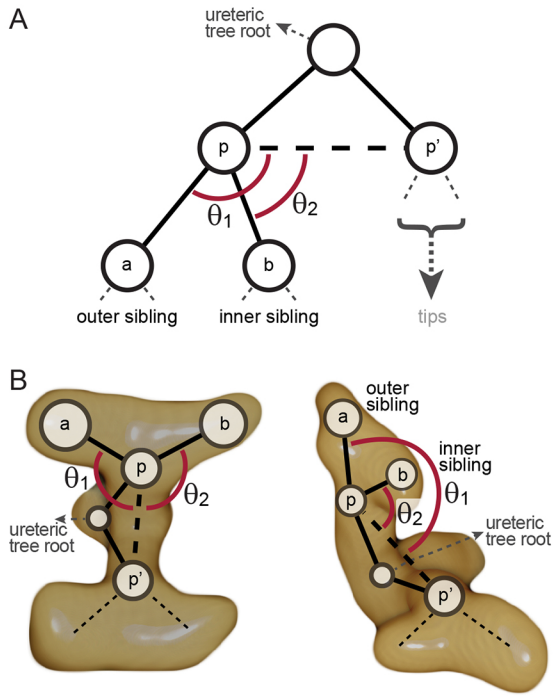



**Fig. 5. Geometrical analysis of branching asymmetry.** (A) For each pair of sibling nodes, the inner sibling occupies the interior position with respect to the 'aunt' node sibling p′, determined by the smaller angle θ. If $\theta_1-\theta_2>45°$, then we say that the branching at p is near planar. (B) Illustrative examples of branching geometry in wild-type ureteric trees. In the first example (left), there is little distinction between inner and outer siblings, whereas in the second example (right), the branching is near planar and the distinction is clear.

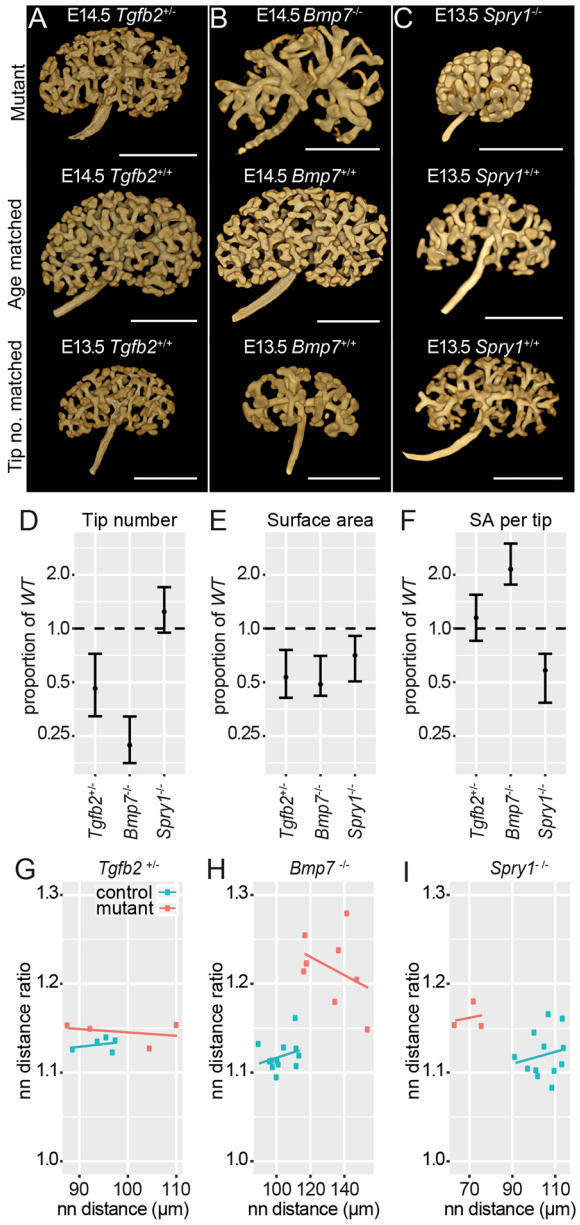



**Fig. 6. Mutant phenotypes.** (A-C) Sample mutant kidneys (top row) with wild-type controls (middle row) and alternative wild-type comparison kidneys matched by tip number (bottom row). (A) *Tgfb2*$^{+/-}$ (E14.5); (B) *Bmp7*$^{-/-}$ (E14.5); (C) *Spry1*$^{-/-}$ (E13.5). Scale bars: 500 µm. (D-F) Changes in (D) tip number, (E) surface area and (F) surface area per tip. Each plot shows estimated median proportional change from wild-type controls (dashed line) with 95% confidence interval, using a Wilcoxon rank sum test on the log transformed metrics (displayed in log scale, marked values are linear scale). (G-I) Ratio between nearest non-sibling neighbour (nn) distances in pairs of sibling tips, averaged by kidney and plotted against mean non-sibling nearest-neighbour distance over the same set of tips for *Tgfb2*$^{+/-}$ (G), *Bmp7*$^{-/-}$ (H) and *Spry1*$^{-/-}$ (I). For each pair of tips, the distance from each sibling to the nearest other tip (excluding each other) is determined; the nearest neighbour distance ratio is the larger of these two distances divided by the smaller. Control/test group sizes are: *Tgfb2*$^{+/-}$ , 5/4; *Bmp7*$^{-/-}$, 12/9; *Spry1*$^{-/-}$, 12/3.

and associated mesenchyme, and whose germline inactivation leads to renal hypoplasia (Dudley et al., 1995). Knockdown of expression in cultured organs results in fusion of ureteric tips and hence the molecule has been proposed to be essential in a tip-avoidance mechanism in the developing organ (Davies et al., 2014). The second was sprouty 1 (*Spry1*$^{-/-}$), a receptor tyrosine kinase inhibitor expressed in the ureteric bud tips (Zhang et al., 2001) that plays a central role in modulating cell signalling downstream of the Ret, a co-receptor for GDNF (Basson et al., 2005). Gene knockout results in precocious UB development and accelerated branching morphogenesis (Basson et al., 2005). Finally, we studied mice carrying heterozygous loss-of-function mutations in transforming growth factor β 2 (*Tgfb2*$^{+/-}$), a molecule expressed predominantly in the ureteric epithelium (Pelton et al., 1991). *Tgfb2* haploinsufficiency results in hypoplasia characterized by reduced branching and developmental delay (Short et al., 2010).

To examine the relationship between tip spacing/environment and developmental patterning of the branching ureteric tree, we collected foetal kidneys from *Bmp7*$^{-/-}$ (E14.5), *Spry1*$^{-/-}$ (E13.5) and *Tgfb2*$^{+/-}$ (E14.5) embryos and subjected them to OPT imaging and Tree Surveyor analysis as per the wild-type sample cohort used in our analysis of branch patterning (see supplementary Materials and Methods for basic metrics). Mutant kidneys of all three genotypes were found to be smaller than controls (Fig. 6A-C, Movie 1). To examine how this related to the physical environment in which the tips reside, we computed a surface area for each ureteric tree using a convex hull around the organ using the tip coordinates to establish the tree boundary (see supplementary Materials and Methods), which confirmed the hypoplasty in these organs (Fig. 6E). When tip environment was expressed as a surface area per tip 'occupancy', we observed no change in *Tgfb2* kidneys (12,059±1453 µm$^2$ versus 10,396±351 µm$^2$, $P$=0.56), an increase in *Bmp7* kidney (27,292±2844 µm$^2$ versus 11,783±489 µm$^2$, $P=6.8\times10^{-6}$) and a decrease in *Spry1* kidney (6761±1038 µm$^2$ versus 12,192±445 µm$^2$, $P$=0.0044) (Fig. 6F). These observations

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

are consistent with the respective hypoplasty and increased branching previously reported for these genotypes (Dudley et al., 1995; Basson et al., 2005). Importantly, differences in occupancy were still observed when mutant kidneys were instead compared with wild-type organs with comparable tip numbers (Fig. 6A-C, Movie 1), suggesting that such changes are not simple products of developmental delay but instead represent genetically encoded mediators of tree morphology (Fig. S3A). Notably, we observed no evidence for tip fusion or 'collision' in *Bmp7* mutant kidneys, as has been reported in cultured organs (Davies et al., 2014). In fact, the surface area occupied per tip in the *Bmp7* samples was considerably greater than in wild-type samples (see Fig. 6). Consistent with this observation, the average distance between tips was increased by 25.17 µm when compared with age-matched controls [95% CI (18.36,32.08), $P=6.8\times10^{-6}$] and by 17.42 µm when compared with wild-type kidneys with equivalent tip numbers [95%CI (9.94,24.84), $P=2.0\times10^{-4}$]. Differences in branching morphogenesis between cultured kidneys and organs dissected from mutant embryos are a feature we have noted before in several other mutants (Short et al., 2013). These may reflect artefactual products of anlage culture systems derived from tissue compression that interrupt the normal spatial arrangement of tissues and cells in the developing kidney.

Although the average distance is greater in *Bmp7*, we also considered how uniformly distributed the tips on the surface of each of the kidneys are for all genotypes. To determine this, we analysed the relative distance ratio of the closest and furthest neighbour tips from pairs of sibling tips. While this measure was unchanged when comparing *Tgfb2* kidneys and controls (1.145684 versus 1.131798, $P=0.1111$, Fig. 6G), we noted significant differences between *Spry1* and *Bmp7*, and age-matched controls (1.161907 versus 1.120212, $P=0.0485$ and 1.218161 versus 1.122488, $P=1.361\times10^{-5}$, Fig. 6H,I). We interpret this to mean that the distribution of tips on the surface of these organs is less even.

Having identified mutants with altered tip spacing and distribution, we employed overlay metrics to determine whether these changes in the tip environment and in branching dynamics resulted from, or were the consequence of, perturbation of the half-delay patterning we had described in wild-type organs. The average balance was calculated for each kidney (Fig. S4) and each mutant group was compared with the age-matched wild-type control group (Fig. 7A). The *Tgfb2* mutants show no detectable change in balance, consistent with our previous overlay analysis (Lamberton et al., 2015), but increases in branching asymmetry were observed in the *Bmp7* and *Spry1* mutants. Comparisons were repeated using size-matched controls (by tip number), with similar results (Fig. S3B). We then repeated the centroid analysis previously employed to characterise the wild-type dataset, examining the *Bmp7* and *Spry1* mutants for each of the six clades. For each mutant group, the selected clade was isolated from each kidney, and the resultant set of clades was augmented with the closest matching model tree (in this case either half delay or Fibonacci) before finding the centroid. In both cases, the half-delay model was the centroid in only 33% of clades (two out of six) as opposed to >80% of wild-type organs at the same developmental stage (29 of 36 cases, Fig. S1). In contrast, when the Fibonacci model was used to classify the clades, it was the centroid in 50% of cases for Spry1 (three out of six) and ~83% of cases for Bmp7 (five out of six). This modification in renal patterning is also reflected in the distribution of branch balances for different sub-tree sizes (Fig. 7B), where a greater imbalance in the mutant trees is evident. Hence, in both the *Spry1* and *Bmp7* kidneys, where there is a significant change in the spacing of tips on the organ surface, there is a definite shift from the wild-type half-delay model towards the Fibonacci model. This itself represents a change in the relationship between the extension and bifurcation rates in the mutant trees.

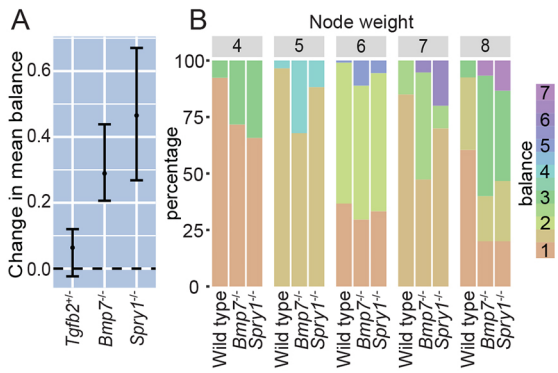



**Fig. 7. Balance change in mutant phenotypes versus wild type.** (A) Estimated change in balance compared with wild-type controls. Median with 95% confidence interval using Wilcoxon rank sum test on kidney means. The kidney mean balance values used are shown in Fig. S4. Control/test group sizes are: $Tgfb2^{+/-}$ (E14.5), 5/4; $Bmp7^{-/-}$ (E14.5), 12/9; $Spry1^{-/-}$ (E13.5), 12/3. (B) Distribution of node balance for each weight 4-8; weights 2 and 3 are excluded as only one balance value is possible in each case. Data are pooled across kidneys: 21 wild type (E14.5), nine $Bmp7^{-/-}$ (E14.5) and three $Spry1^{-/-}$ (E13.5). The distribution of balance in the mutant phenotypes is distinct from wild type across multiple node weights.

## DISCUSSION

Branching morphogenesis is a commonly employed developmental mechanism that establishes the airways of the lungs, the ductal network of glandular organs and the collecting duct network of the kidney. In all cases, the epithelial networks present in these organs are established by way of the arborisation of an initial epithelial bud, usually within a supporting population of mesenchymal cells. The factors that mediate this elaboration and the manner in which it occurs are only partially defined. In the lung, it has been proposed that a rigid and genetically hardwired program of three different branch types combine to form the airways. However, this pattern of growth appears to differ fundamentally from the kidney, as there is no recognizable and reproducible structure of the branched ureteric tree (Short et al., 2014). Our previous studies have identified similarity between different kidneys in the form of a higher order clade structure and also in the extent of branching (Short et al., 2014). In this study, we have sought to understand whether (and to what extent) this reflects a deeper underlying patterning that governs the development of the organ.

We have established a mathematical growth model of ureteric tree branching in which a stereotypical early structure is succeeded by a branching pattern that is locally similar regardless of the position of a branch within the ureteric tree and across developmental stages. This pattern is characterized by a conserved level of branching asymmetry, suggesting a common growth process acting at a local level. This is consistent with our previous quantitative analysis of branch sizes and geometry (Short et al., 2014), which similarly shows substantial underlying regularity, with minor local variation. The family of generative tip-state models that we have developed comprise tip populations that either bifurcate or mature solely according to their current state. These can be conceived more generally as delay models (see supplementary Materials and Methods), in which there is a characteristic difference in maturity or time to bifurcation between the two child nodes formed by each

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

bifurcation. Consequently, they possess the feature that the balance at each branch point will converge to a fixed asymptotic value as growth continues, and are well suited to modelling branching under a common locally operative growth process. We have shown that a three-state delay model (which we term 'half delay') recapitulates, with a very high degree of accuracy, the pattern of branching established by imaging biological samples. This encompasses measures of balance that are shared between and within the clades of the organs. Importantly this model accurately describes the structure, both on the global (Fig. 3) and local (Fig. 4) scale.

Our analysis suggests that an initial idiosyncratic phase of stereotypic UB branching occurs early in development and establishes the clades of the organs. The organization and balance of these branches is clearly distinct from later bifurcation events. At this point, the UB invades a discrete and relatively large and uniform region of $SIX2^+$ GDNF-producing mesenchyme. We propose that the early branching pattern and the establishment of clades occurs in response to the pre-existing shape and relative uniformity of this field of cells. However, by E12 the $SIX2^+$ population begins to condense around the growing tips to form discrete cap mesenchyme populations, which are then a feature of organ development until the termination of nephrogenesis at P4 in the mouse. During this period of development, the tips are confined to, and evenly arranged within, a thin zone on the surface of the growing kidney. We propose that this environment is conducive to the establishment of locally operant cell signalling systems that dictate the asymmetric pattern of branching morphogenesis we observe in our analysis. Although the surface area and tip number increase exponentially, the spacing between tips declines only modestly in synchrony with the reduction in the size of the tip and cap niches (Short et al., 2014). The fact that tip spacing is relatively constant indicates coordination between tip extension and branching: the two primary ureteric tree growth processes. A natural hypothesis is that bifurcation of the growing tips is suppressed to some extent by neighbouring tips and we propose that such a mechanism further contributes to the patterned branching we observe. The hypothesized mutual suppression of bifurcation by the tips is supported by our geometrical analysis. Bifurcation is delayed (but growth continues) in tips located between the most closely related branches in the tree structure (node b in Fig. 5), which we expect to have, on average, more and closer neighbouring tips than their siblings (node a in Fig. 5). In this way relatedness in the tree structure provides a useful proxy for tip spacing. Furthermore, this comparison of sibling branches demonstrates that the mutual suppression of bifurcation by tips may explain the systematic and conserved branching asymmetry that marks the branching structure.

Recent studies have demonstrated that the hierarchical patterning of the mouse mammary gland can be modelled with a locally operative, time invariant, stochastic growth process, in which growing tips bifurcate or cease growth with near-equal probability (Scheele et al., 2017). This report further suggested that a similar process may drive renal branching morphogenesis. We find little support for such a proposal. For example, there is no evidence of stochastic termination of bifurcation or extension – branches do not terminate 'within' the kidney and tips are all, by and large, bifurcating on the organ surface at any given time (see Fig. 6 and Movie 1). Furthermore, cell proliferation in all of the tips of the branching UB is actively maintained until the cessation of nephrogenesis at postnatal day 4 (Short et al., 2014), in a manner that relies on signals from the cap mesenchyme (Cebrian et al., 2014). The possibility that the observed asymmetry may arise from a random process was also considered. The clearest evidence against this hypothesis (and on derived models in which the timing of random branching is shaped by a specified probability distribution) is found in the geometrical analysis of sibling subtree clusters (see Figs 3 and 4). As we can predict which sibling subtree is most likely to be larger based on their positions relative to related tips, a purely random process can be excluded. However, ultimately the strongest argument for this approach is that we were able to provide an excellent fit to the centre of the distribution of real data (see centroid analysis). In these analyses the half-delay model trees were generally a better fit to the experimental data than were any of the individual experimentally determined trees. The remaining option for a stochastic model would be to add some degree of randomness to the half-delay model. Although this might be an accurate representation of the observed biological variation around the identified stereotypical pattern, the lack of any systemic difference between our deterministic model and the data indicated that fitting such a model would not further improve the description.

Taken together, these observations are entirely inconsistent with a model based on stochastic cessation of branch growth and/or bifurcation. Instead, we propose that there are much broader similarities between branching morphogenesis in the kidney and lung. Unlike the mammary gland, branching morphogenesis in these organs is particularly characterized by the involvement of highly specialized mesenchymal cell populations that produce factors that drive cell proliferation and branching of their respective epithelial networks. In support of this notion, we have found that the tip state models developed in this study are able to generate hierarchical structures like the domain branching seen in lungs (see supplementary Materials and Methods: three state models). Although further modelling and comparison with biological samples will be required to determine the full extent to which such models can describe pulmonary development, our findings suggest a commonality between branching in these organs and highlight considerable differences with the mammary gland.

Our analyses of mutant kidneys say much about the nature of patterning that governs the normal development of the kidney. They imply that alterations in the rate of organ growth do not necessarily perturb patterning. While haplo-insufficiency for *Tgfb2* significantly abrogates kidney growth, by ∼24 h (Lamberton et al., 2015), we observe no differences in the pattern of systematic asymmetry in branching identified in this study. In most respects, these organs, although small, are perfectly formed – suggesting that genetic control of organ growth is not intrinsically linked to patterning of the branching UB. In contrast, changes in *Bmp7* and *Spry1* mutants result in changes in branch patterning from a half delay to a pattern more closely approximating a Fibonacci model. Although the surface areas per tip of each model are larger (*Bmp7*) and smaller (*Spry1*), both models share the feature of more uneven tip distribution on the organ surface. This suggests that the normal factors governing tip distribution are somehow ameliorated in both organs. This might be through a reduction in suppressive actions of a protein that prevents the incursion of one tip into the field of another (as may be the case for *Bmp7*) or in the loss of factors that normally control the unchecked bifurcation of tips (as may be the case for *Spry1*). In both cases, a change in the uniformity of tip spacing is associated with altered branch patterning. At this point, we cannot say with certainty whether the alteration in patterning is a cause or a consequence of the change in the uniformity of tip distribution. However, given that the patterning is effectively determined by the decision of a tip to bifurcate or grow (which is likely locally determined), we favour the former model in

Development (2017) 144, 4377-4385 doi:10.1242/dev.153874

which the uniform distribution of tips on the organ surface dictates the systematic asymmetry we observe in the ureteric tree.

What biological factors might contribute to the half-delay models that describe patterning of the elaborating ureteric tree? The different 'nodes' in this scheme effectively represent points at which a given branch bifurcates or grows. Such states effectively represent different degrees of developmental delay (see supplementary Materials and Methods for approaches to modelling different delay to bifurcation time ratios). It is likely that both biophysical and molecular cues contribute to the decision of a given tip to bifurcate or to extend. The size of the organ surface effectively establishes the field in which such tip-tip interactions can occur and although it might be expected that this is defined by the extent of ureteric branching, our findings demonstrate that this is not the case: increased and decreased branching do not necessarily imply increased or decreased organ size. At the level of the tip, the bounding basement membrane likely constitutes a structure that is able to sterically impact adjacent tip environments. It is tempting to attribute similar features to the surrounding cap mesenchyme that (at least in section) appears as a ramified buffer between neighbouring tips. However, recent studies suggest that this cell population is dynamic and motile, which may reduce its steric impact on neighbouring tips (Combes et al., 2016). Equally, the role of the stromal cells – which sit between cap/tip populations – has been relatively poorly studied, although they are known to regulate branching morphogenesis in a non-cell-autonomous manner (Paroly et al., 2013; Mukherjee et al., 2017). Studies to assess the *in vivo* forces and tensions created by juxtaposed cap/tip/stromal niches and their respective cells are required to assess their biophysical impact on patterning. Considerable recent conceptual and experimental advances have extended our understanding of the cellular events that underlie bifurcation in branching epithelia (Ihermann-Hella et al., 2014; Menshykau et al., 2014; Huebner et al., 2016; Lin et al., 2017). Linking these to the emerging picture of shared and organ-specific growth factor production in the tips of branching epithelial networks (Rutledge et al., 2017) will therefore be important for integrating how these combine to direct the exquisite patterning we describe in the developing kidney at a whole organ level.

## MATERIALS AND METHODS

### Sample collection, staining and imaging

Embryonic mice from heterozygous *Tgfb2*$^{tm1Doe}$, homozygous *Bmp7*$^{tm1.2Dgra}$, homozygous *Spry1*$^{tm1Jdli}$ and wild-type C57BL/6J kidneys were collected, stained and imaged as previously described (Combes et al., 2014) with ages and sample numbers outlined in Table S1. Limb staging was used to broadly confirm age in most of the embryos and also to subclassify them based upon developmental progression (Wanek et al., 1989). Kidneys were stained using anti pan-cytokeratin (Abcam, AB11213; 1:300) and anti Trop2 (R&D Systems, AF1122; 1:100) with Alexa Fluor secondary antibodies (Life Technologies), followed by OPT imaging (Bioptonics). All animal procedures complied with standards set under Australian, US and Canadian federal and state guidelines for animal welfare, and experiments were approved by the relevant Animal Ethics Panels of Monash University, Columbia University and the University of Alberta.

### Sample data analysis and representation

Reconstructed ureteric tree datasets were mapped and quantified in 3D using Tree Surveyor software (Short et al., 2013). Tree Surveyor software is available to academic and not-for-profit researchers. GraphML data was exported, providing a representation of each tree in the form of nodes (branch points and tips) and edges between them (branch segments), annotated with information on size and position. The root node of the graph is the point where the ureter first branches, and defines an orientation of each edge from a start node, which is closer to the root, to an end node, which is further away. The start node is the parent of the end node, and the end node is the child of the start node. Each of the 11 primary branch points (Fig. 1B) were annotated on the ureteric tree graphs based on their relative position in the kidney the graph was derived from, also identifying the six-clade structure previously described (Short et al., 2014).

### Computational analysis

Graph overlays and pairwise comparisons were performed as previously described (Lamberton et al., 2015) using MATLAB and R code. All other analyses described were performed in R, including generation of tip-state model tree structures used in overlay analysis with real data. See supplementary Materials and Methods for details of tip state model definitions, properties and enumeration. Example model trees used in overlay analysis were generated in R.

#### Acknowledgements

We acknowledge the contributions of Frank Costantini and Jonathan Licht in providing the *Sprouty 1* mouse embryos. We thank the Monash Micro Imaging and Monash Animal Research Platforms for assistance with imaging and animal husbandry, respectively. The authors also thank Professor Melissa Little and Dr Alexander Combes for interesting and useful discussions during the preparation of this paper.

#### Competing interests

The authors declare no competing or financial interests.

#### Author contributions

Conceptualization: J.G.L., K.M.S., T.O.L., I.M.S., N.A.H.; Methodology: J.G.L., K.M.S., T.O.L., I.M.S., N.A.H.; Software: J.G.L., K.M.S., T.O.L.; Formal analysis: J.G.L., T.O.L., N.A.H.; Investigation: J.G.L., K.M.S., T.O.L., O.M., D.G.; Resources: O.M., D.G., I.M.S.; Data curation: J.G.L., K.M.S., T.O.L.; Writing - original draft: J.G.L., K.M.S., T.O.L., I.M.S., N.A.H.; Writing - review & editing: J.G.L., K.M.S., T.O.L., O.M., D.G., I.M.S., N.A.H.; Visualization: J.G.L., K.M.S., T.O.L., N.A.H.; Supervision: I.M.S., N.A.H.; Project administration: I.M.S., N.A.H.; Funding acquisition: I.M.S., N.A.H.

#### Funding

This work was supported by project grants to I.M.S. and N.A.H. from the Australian Research Council (DP160103100 and DP130100886) and to K.M.S. from the Monash University Strategic Grants scheme. I.M.S. acknowledges the support of an Australian Research Council Future Fellowship (FT100100620) and a National Health and Medical Research Council Senior Research Fellowship (APP1106516).

#### Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.153874.supplemental

#### References

**Basson, M. A., Akbulut, S., Watson-Johnson, J., Simon, R., Carroll, T. J., Shakya, R., Gross, I., Martin, G. R., Lufkin, T., McMahon, A. P. et al.** (2005). Sprouty1 is a critical regulator of GDNF/RET-mediated kidney induction. *Dev. Cell* **8**, 229-239.

**Blanc, P., Coste, K., Pouchin, P., Azaïs, J.-M., Blanchon, L., Gallot, D. and Sapin, V.** (2012). A role for mesenchyme dynamics in mouse lung branching morphogenesis. *PLoS ONE* **7**, e41643.

**Cebrian, C., Asai, N., D'Agati, V. and Costantini, F.** (2014). The number of fetal nephron progenitor cells limits ureteric branching and adult nephron endowment. *Cell Rep.* **7**, 127-137.

**Combes, A. N., Short, K. M., Lefevre, J., Hamilton, N. A., Little, M. H. and Smyth, I. M.** (2014). An integrated pipeline for the multidimensional analysis of branching morphogenesis. *Nat. Protoc.* **9**, 2859-2879.

**Combes, A. N., Lefevre, J. G., Wilson, S., Hamilton, N. A. and Little, M. H.** (2016). Cap mesenchyme cell swarming during kidney development is influenced by attraction, repulsion, and adhesion to the ureteric tip. *Dev. Biol.* **418**, 297-306.

**Davies, J. A., Hohenstein, P., Chang, C.-H. and Berry, R.** (2014). A self-avoidance mechanism in patterning of the urinary collecting duct tree. *BMC Dev. Biol.* **14**, 35.

**Dudley, A. T., Lyons, K. M. and Robertson, E. J.** (1995). A requirement for bone morphogenetic protein-7 during development of the mammalian kidney and eye. *Genes Dev.* **9**, 2795-2807.

**Huebner, R. J., Neumann, N. M. and Ewald, A. J.** (2016). Mammary epithelial tubes elongate through MAPK-dependent coordination of cell migration. *Development* **143**, 983-993.

**Ihermann-Hella, A., Lume, M., Miinalainen, I. J., Pirttiniemi, A., Gui, Y., Peränen, J., Charron, J., Saarma, M., Costantini, F. and Kuure, S.** (2014). Mitogen-activated protein kinase (MAPK) pathway regulates branching by remodeling epithelial cell adhesion. *PLoS Genet.* **10**, e1004193.

**Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D. and Jaenisch, R.** (1993). WT-1 is required for early kidney development. *Cell* **74**, 679-691.

**Lamberton, T. O., Lefevre, J., Short, K. M., Smyth, I. M. and Hamilton, N. A.** (2015). Comparing and distinguishing the structure of biological branching. *J. Theor. Biol.* **365**, 226-237.

**Lawson, B. A. and Flegg, M. B.** (2016). A mathematical model for the induction of the mammalian ureteric bud. *J. Theor. Biol.* **394**, 43-56.

**Lin, C., Yao, E., Zhang, K., Jiang, X., Croll, S., Thompson-Peer, K. and Chuang, P. T.** (2017). YAP is essential for mechanical force production and epithelial cell proliferation during lung branching morphogenesis. *eLife* **6**, e21130.

**Menshykau, D. and Iber, D.** (2013). Kidney branching morphogenesis under the control of a ligand-receptor-based Turing mechanism. *Phys. Biol.* **10**, 046003.

**Menshykau, D., Kraemer, C. and Iber, D.** (2012). Branch mode selection during early lung development. *PLoS Comput. Biol.* **8**, e1002377.

**Menshykau, D., Blanc, P., Unal, E., Sapin, V. and Iber, D.** (2014). An interplay of geometry and signaling enables robust lung branching morphogenesis. *Development* **141**, 4526-4536.

**Metzger, R. J., Klein, O. D., Martin, G. R. and Krasnow, M. A.** (2008). The branching programme of mouse lung development. *Nature* **453**, 745-750.

**Mukherjee, E., K. V. Maringer, E. Papke, D. S. Bushnell, C. M. Schaefer, R. Kramann, J. Ho, B. D. Humphreys, C. M. Bates, and S. and Sims-Lucas**, (2017). Endothelial markers expressing stromal cells are critical for kidney formation. *Am. J. Physiol. Renal. Physiol.* **313**, F611-F620.

**Muthukrishnan, S. D., Yang, X., Friesel, R. and Oxburgh, L.** (2015). Concurrent BMP7 and FGF9 signalling governs AP-1 function to promote self-renewal of nephron progenitor cells. *Nat. Commun.* **6**, 10027.

**Paroly, S. S., Wang, F., Spraggon, L., Merregaert, J., Batourina, E., Tycko, B., Schmidt-Ott, K. M., Grimmond, S., Little, M. and Mendelsohn, C.** (2013). Stromal protein Ecm1 regulates ureteric bud patterning and branching. *PLoS ONE* **8**, e84155.

**Pelton, R. W., Saxena, B., Jones, M., Moses, H. L. and Gold, L. I.** (1991). Immunohistochemical localization of TGF beta 1, TGF beta 2, and TGF beta 3 in the mouse embryo: expression patterns suggest multiple roles during embryonic development. *J. Cell Biol.* **115**, 1091-1105.

**Rutledge, E. A., Benazet, J. D., and McMahon, A. P.** (2017). Cellular heterogeneity in the ureteric progenitor niche and distinct profiles of branching morphogenesis in organ development. *Development* **114**, 3177-3188.

**Saxen, L.** (1987). *Organogenesis of the Kidney*. Cambridge, Cambridge University Press.

**Scheele, C. L., Hannezo, E., Muraro, M. J., Zomer, A., Langedijk, N. S., van Oudenaarden, A., Simons, B. D. and van Rheenen, J.** (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* **542**, 313-317.

**Self, M., Lagutin, O. V., Bowling, B., Hendrix, J., Cai, Y., Dressler, G. R. and Oliver, G.** (2006). Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. *EMBO J.* **25**, 5214-5228.

**Short, K. M., Hodson, M. J. and Smyth, I. M.** (2010). Tomographic quantification of branching morphogenesis and renal development. *Kidney Int.* **77**, 1132-1139.

**Short, K., Hodson, M. and Smyth, I.** (2013). Spatial mapping and quantification of developmental branching morphogenesis. *Development* **140**, 471-478.

**Short, K. M., Combes, A. N., Lefevre, J., Ju, A. L., Georgas, K. M., Lamberton, T., Cairncross, O., Rumballe, B. A., McMahon, A. P., Hamilton, N. A. et al.** (2014). Global quantification of tissue dynamics in the developing mouse kidney. *Dev. Cell* **29**, 188-202.

**Tomita, M., Asada, M., Asada, N., Nakamura, J., Oguchi, A., Higashi, A. Y., Endo, S., Robertson, E., Kimura, T., Kita, T. et al.** (2013). Bmp7 maintains undifferentiated kidney progenitor population and determines nephron numbers at birth. *PLoS ONE* **8**, e73554.

**Wanek, N., Muneoka, K., Holler-Dinsmore, G., Burton, R. and Bryant, S. V.** (1989). A staging system for mouse limb development. *J. Exp. Zool.* **249**, 41-49.

**Zhang, S., Lin, Y., Itäranta, P., Yagi, A. and Vainio, S.** (2001). Expression of Sprouty genes 1, 2 and 4 during mouse organogenesis. *Mech. Dev.* **109**, 367-370.

# SUPPLEMENTARY MATERIAL

## 1. Data

Analysis is based on a data set of 32 wild type C57BL/6J fetal mouse kidneys of embryonic age E11.5 to E15.5 and an additional 28 mutant and control samples (all on an inbred C57BL/6J background, see Table S1).

***Table S1. Mouse ureteric tree data used.***

| Group | Approximate Age (embryonic day) | Limb Stage | Samples | tip number (mean $\pm$ stdev) |
|---|---|---|---|---|
| Wild type time series | 11.5 | 7 | 6 | $13.5 \pm 5.4$ |
| Wild type time series | 12.5 | 8 | 5 | $25.0 \pm 2.9$ |
| Wild type time series | 13.25 | 9 | 7 | $85.7 \pm 18.6$ |
| Wild type time series | 13.75 | 10 | 5 | $116.2 \pm 15.8$ |
| Wild type time series | 14.5 | 11 | 5 | $224.8 \pm 45.1$ |
| Wild type time series | 15.5 | 12 | 4 | $621.3 \pm 35.8$ |
| Wild type time series | total | | 32 | |
| *Tgfb2+/-* mutant | 14.5 | 11 | 4 | $106.0 \pm 23.2$ |
| *Bmp7+/+* control | 14.5 | 11 | 12 | $247.9 \pm 71.4$ |
| *Bmp7-/-* mutant | 14.5 | 11 | 9 | $58.3 \pm 19.0$ |
| *Spry1-/-* mutant | 13.5 | 9/10 | 3 | $121.3 \pm 12.6$ |
| Mutant / control total | | | 28 | |

## 2. Statistical notes

Statistical analysis was performed using the R programming language and RStudio environment. All t-tests were two-sided and unpaired, using the Welch adjustment for possibly unequal variance (using the t.test function in R with default parameters). Linear regression and ANOVA was performed with the lm function in R (using the f-statistic).

**Age and size matched controls**

Dedicated control groups were available for the *Bmp7* mutants, while the *Spry1* and *Tgfb2* mutants were compared to age matched data from the wild type time series data. For *Spry1*, the combined E13.25 and E13.75 wild type data was used.

For some analyses, size matched alternative controls (based on tip number) were also used, to test whether observed differences in the mutant groups could be explained as the result of a general delay in development. The E14.5 *Tgfb2* mutant was compared to the combined E13.25 and E13.75 wild type data, and the *Bmp7* mutant to the combined E12.5 and E13.25 wild type data. The *Spry1* mutant did not differ substantially in tip number from the age matched control group, so no alternative control was used.

**Mutant comparisons**

Mutant phenotypes were quantified by a comparison of kidney level metrics between mutant and control groups. The large variation between samples in some mutant groups did not appear consistent with a normal distribution, so a Wilcoxon rank sum test was performed using the R function wilcox.exact from package exactRankTests. This function produced the estimate and 95% confidence interval for the median difference, used in Fig. 6, Fig. 7 and Fig. S3.

Surface area was calculated for a convex hull on the set of tip points, using the R function convhulln from the geometry package, which interfaces the Qhull library (Barber, Dobkin et al. 1996).

For the number of tips per unit surface area and balance metrics, we are interested in difference compared to both age and size (tip number) matched wild type kidneys. In Fig. 6 and 7, the standard (age matched) controls were used; the alternative comparisons using size matched controls are shown in Fig. S3.

## 3. Tip state models

A tip state model is defined by a finite set of *tip states*, a specified initial state, and a transition rule which deterministically maps each state to a set of one or more new states. The transition rule is applied to produce a sequence of *rooted trees* (Epp 2004) in which each tip has an associated state from the allowed set; the first tree in the sequence consists of a single tip with the specified initial state. At each iteration, the rule is applied independently and simultaneously to each tip. If the current state of a tip is mapped to a single new state, then only the state is changed. If the tip state is mapped to more than one state, then child nodes are added to the tip with the assigned states; the original node is no longer a tip and does not have an associated state. The order of child nodes is not important, as we are ultimately concerned with distinct structures only up to graph isomorphism (Epp 2004). A state which is mapped to more than one new state is a *branching* state.

Tip state is an internal parameter which is used to generate the trees but is ultimately discarded, as it is unobservable in real trees; we say that two trees are isomorphic (written $\equiv$) if they are graph isomorphic, with tip state ignored if either or both trees possess it. Another potential source of ambiguity is that two models may produce the same tree structure at different positions in their generated sequences. Hence we define the family of rooted trees associated with a model to be the set of trees that occur in the generated sequence, with tip state ignored, which are distinct up to graph isomorphism. Tip state models which produce the same family of trees (up to isomorphism) are considered *equivalent*, and a model which is equivalent to a model with fewer states is *degenerate*, and excluded from the enumeration.

It is possible that a tip state model will produce only a finite number of trees. If any branching state occurs which gives rise after some number of iterations to at least one tip with the same (branching) state, then the tree sequence produced is infinite. If there are $k$ states, there can be at most $k-1$ iterations of the model in which branching occurs before such a loop arises. Thus models which fail to give an infinite sequence of trees give at most $k$ distinct trees. We call such models *trivial*, and as with degenerate models we exclude them from enumeration.

**Classification of binary tip state models**

Since we are modelling bifurcating tree structures, from here on we only consider binary models, in which each state is mapped to either one or two states; a *bifurcating* state is one which is mapped to two new states. In the following we develop some properties of these models, in particularly classifying all non-equivalent binary tip state models with up to three states, and proving the asymptotic balance property of a class of tip state models.

**Notation**

We denote a tree consisting of a single tip of state $a$ by just writing $a$, relying on context to distinguish from the state $a$, and let $(T_1, T_2)$ denote a tree with more than one tip, in which the two children of the root node are the roots of subtrees of type $T_1$ and $T_2$. Note that this notation can be used recursively to specify any rooted binary tree together with tip states. We also use this notation with no state specified to denote tree structures where the tips do not have names or states; so for example $()$ is the trivial tree, and $((,),(,))$ is the symmetric tree with 4 tips. This is known as Newick notation (Krane 2003).

For a given tip state model, the transition rule is formally an automorphism on the set of rooted trees with labelled tips. When the model in question is clear we denote this associated function $f$, so $f^n(T)$ is the result of $n$ consecutive transformations of an initial (tip labelled) tree T. We write $T_1 \mapsto T_2$ as a shorthand for $f(T_1) = T_2$.

For a tree $T$ generated from a tip state model with $k$ states, the number of tips in each state can be represented by a *state vector* $\boldsymbol{s}(T) \in \mathbb{Z}^k$. This requires an ordering of the states; we typically denote states with lowercase Latin letters and assume alphabetical ordering (of course reordering states will not fundamentally change the model). Since the result of transforming a tip is deterministic and depends only on its initial state, the state vector of $f(T)$ is a linear transformation of the state vector of $T$. That is:

**Lemma 1:** For any binary tree state model, there is a unique $k \times k$ matrix $M$ such that $\boldsymbol{s}\big(f(T)\big) = M\boldsymbol{s}(T)$.

We refer to $M$ as the state transition matrix. Another immediate consequence of the transition function acting separately on each function is:

**Lemma 2:** $f\big((T_1, T_2)\big) = (f(T_1), f(T_2))$

We now introduce the first non-trivial family of rooted binary trees.

**Definition:** The *n*th *perfect* tree, denoted $P_T(n)$, is the completely balanced binary tree formed by $n-1$ consecutive bifurcations of every tip, starting with a single node. Recursively, $P_T(1) = ()$ and $P_T(n+1) = (P_T(n), P_T(n))$. $P_T(n)$ has $2^{n-1}$ tips and $2^n$-1 nodes in total.

**One-state models**

We begin the enumeration of binary tip state models by supposing there is only one possible tip state. In this case the only non-trivial model produces the perfect trees.

Observe that any model with no bifurcating state will produce only the trivial tree, while a model where all states bifurcate will give $f^n(a) = P_T(n)$. This immediately resolves the one-state case, and also gives

the following lemma. Recall that a degenerate model is one that is equivalent to (produces the same trees as) a model with fewer tip states, and we disregard such models:

**Lemma 3:** Given any binary tip state model with more than one state, then if all states are bifurcating the model is degenerate, while if no states are bifurcating then it is trivial.

**Two-state models**

We show that there are two distinct models with two tip states (excluding trivial and degenerate models). One is the Fibonacci model (see Fig. 2A); the other produces trees consisting of a single "trunk" and a series of tips branching from it, which we call the singleton domain model. This "domain" branching is named for the pattern of trunk and offshoot branches seen in the lung (Metzger, Klein et al. 2008) but it is not capable of modelling ureteric tree structures.

By Lemma 3, we can assume there is one bifurcating and one non-branching state. Up to isomorphism we have cases

1. $a \mapsto a$ and $b \mapsto (a, a)$
2. $a \mapsto a$ and $b \mapsto (b, b)$
3. $a \mapsto a$ and $b \mapsto (a, b)$
4. $a \mapsto b$ and $b \mapsto (a, a)$
5. $a \mapsto b$ and $b \mapsto (b, b)$
6. $a \mapsto b$ and $b \mapsto (a, b)$

In theory each of these cases corresponds to two models, depending on the choice of start state. But if the start state is *a* then it will either be mapped to itself, giving a trivial model, or it will be mapped to *b*, producing the same sequence of trees with a delay of one time step as the model with start state *b*, and is thus equivalent to that model. So we need only consider start state *b*. In fact, we can assume for any model (not restricted to bifurcating models) that the start state is a branching state. Although in general it may take more than one iteration for an initial non- branching state to be mapped to a branching state, it will always be equivalent to the model starting at that branching state.

**Lemma 4:** A tip state model with a non-branching initial state is either trivial or is equivalent to a model which has a branching initial state and is otherwise the same.

Thus when enumerating distinct models, we assume that the start state is branching.

Returning to two-state models and the six cases above, case 1 is trivial, while cases 2, 4, and 5 are degenerate (in cases 2 and 5 $f^n(b) \equiv P_T(n)$, while in case 4 $f^n(b) \equiv P_T(\lfloor (n+2)/2 \rfloor)$). Case 3 produces a new family with trees of the form $(a, \big(a, (\dots (a, (a, b)) \dots)\big))$. This pattern of repeated offshoots from a trunk is seen in the branching of organs such as the lung, and we refer to it as *domain* branching. But our 2-state system cannot model further development of the offshoots. We call this minimal model *singleton* domain branching.

This leaves Case 6, which we call the Fibonacci model.

**Definition:** The *n*th *Fibonacci* tree, denoted $F_T(n)$, is defined recursively by $F_T(1) = F_T(2) = ()$ and $F_T(n+2) = (F_T(n), F_T(n+1))$.

Case 6 produces exactly the set of Fibonacci trees: $f^n(b) \equiv F_T(n+1)$. From the recursive definition of the Fibonacci numbers, the following follows immediately:

**Lemma 5**: $F_T(n)$ has $F(n)$ tips, where $F(n)$ is the *n*th Fibonacci number.

We have completed the 2 state classification:

**Theorem 1:** The only non-trivial and non-degenerate 2-state binary tip state models are the Fibonacci and singleton domain branching models.

**Generalised delay models - single branching state**

Before we examine three-state models in detail we first consider a class which can be regarded as a generalization of the Fibonacci model. The Fibonacci model has a single branching state, with one child tip that branches at the next model iteration and the other child branching with the delay of one additional model iteration. Similar models with additional states allow variable delays before bifurcation.

Suppose there is a single branching state $c$, and $f(c) = (a, b)$, where $a, b, c$ are not necessarily distinct. Then we have 3 cases up to generality:

1. $f^p(a) = f^q(b) = c$, for some $p \geq q \geq 0$.
2. $f^q(b) = c$ for some $q \geq 0$ but there is no $p$ such that $f^p(a) = c$
3. There is no $p$ or $q$ such that $f^p(a) = c$ or $f^q(b) = c$.

Case 2 will produce the singleton domain branching trees (branching will be delayed but the tree structures produced are the same), and hence will be degenerate for more than 2 states, while case 3 will produce only the trees () and (,). We are left with case 1, which we define as the generalized $(p, q)$ delay model.

**Definition:** For integers $p \geq q \geq 0$, the $(p, q)$ delay model is a model with a single bifurcating state $c$, with $f(c) = (a, b)$ and $f^p(a) = f^q(b) = c$. The generalized delay tree $D_T^{p,q}(n)$ is the tree structure $f^n(c)$, produced by the $(p, q)$ delay model.

**Lemma 6:** The (non-degenerate) $(p, q)$ delay model will have $p + 1$ states $a, f(a), f^2(a), \ldots, f^p(a) = c$, with $b = f^{p-q}(a)$.
Proof: The listed states must exist and be distinct. Since this is sufficient to produce the delay trees, a model with more states is degenerate.

Note that $D_T^{1,0}(n) = F_T(n+1)$; the difference by one in the index is because the Fibonacci trees were defined to match the Fibonacci numbers (Lemma 5). The recursive definition of $F_T(n)$ generalises as follows:

**Lemma 7:** $D_T^{p,q}(m+p+1) = (D_T^{p,q}(m), D_T^{p,q}(m+p-q))$ for $m \geq 0$

Proof: From the definition and Lemma 2,

$$D_T^{p,q}(m+p+1) = f^{m+p+1}(c) = f^{m+p}\big((a,b)\big) = \big(f^{m+p}(a), f^{m+p}(b)\big) = \big(f^m(c), f^{m+p-q}(c)\big)$$

The result follows. □

The perfect trees can also be regarded as the base case of the generalized delay model: $D_T^{0,0}(n) = P_T(n)$. Thus we have categorised all models with a single branching state:

**Theorem 2:** Any non-trivial binary tip state model with exactly one branching state will produce either singleton domain trees or a family of generalized delay trees.

We now examine balance in the delay model.

**Theorem 3:** The asymptotic balance at the branching points of $D_T^{p,q}(n)$ as the weight tends to infinity is $\lambda^{p-q}$, where $\lambda$ is the (unique) positive real solution of $\lambda^{p+1} - \lambda^{p-q} - 1 = 0$.

> Proof: We consider the root node of $D_T^{p,q}(n)$ as $n \to \infty$. By Lemma 7, the balance at this node is the ratio between the weight of two smaller $D_T^{p,q}$ trees, one of which is $(p-q)$ iterations advanced from the other. Consider the $(p+1) \times (p+1)$ transition matrix M for the $(p,q)$ delay model. Assuming states are ordered $a, f(a), f^2(a), \ldots, f^p(a)$, M has ones in position $(i+1, i)$, $i = 1,2,\ldots,p$, plus positions $(1, p+1)$ and $(p-q, p+1)$ in the last column representing the bifurcating state. This matrix has characteristic equation $\lambda^{p+1} - \lambda^{p-q} - 1 = 0$. Since $p+1 > p-q \geq 0$, *M* will thus have exactly one positive real eigenvalue $\lambda$, which will be greater than 1. Therefore in the asymptote, the number of tips increases by a factor of $\lambda$ in each model iteration (the corresponding eigenvector gives the asymptotic tip state proportions). Thus by Lemma 7, the asymptotic root balance is $\lambda^{p-q}$.

The subtree associated with each branch point is itself a $D_T^{p,q}$ trees, so this asymptotic property generalizes from the root to each branch point.

Since $F_T(n) = D_T^{1,0}(n)$, we have the following result:

**Corollary 1**: The Fibonacci trees have asymptotic balance $(\sqrt{5}+1)/2 \cong 1.61$

**Three-state models**

In this section we show that there are 34 distinct three-state bifurcating tip models, including one which is potentially useful for modelling ureteric trees, which we call the half-delay model. By Lemma 3, the non-degenerate 3 state models have exactly one or two bifurcating states, so we consider these two cases in turn.

We have shown that a single branching state implies that the only non-degenerate models are $(p,q)$ delay models. By Lemma 6 we have $p = 2$, and hence $q = 0,1,2$. The $(2,2)$ delay model will give perfect trees so is degenerate. Hence we have two three-state models with a single branching state, the $(2,0)$ and $(2,1)$ delay models.

By Theorem 2, the $(2,1)$ delay model will have asymptotic balance $\cong 1.325$, while the $(2,0)$ delay model will have asymptotic balance $\cong 2.148$ (the square of the asymptotic growth rate $\cong 1.466$). Hence the $(2,1)$ delay model is potentially useful for modelling ureteric trees, but the $(2,0)$ delay model is not since it gives balance values outside the observed range.

We call the $(2,1)$ delay model the *half-delay*; see Fig. 2B.

In the remainder of this section we consider models with 2 branching states. Assume that states b and c branch, a does not. The cases are listed below up to generality, with each distinct, non-trivial and non-degenerate model noted in bold; we confirmed that these 32 models are distinct by generating all trees with up to 200 tips for each model, and comparing the tip numbers and in some cases tree structures until we were able to distinguish all models. Two main groups are domain structures, where one branching state is strictly downstream from the other, and nested structures, where each branching state will give rise to the other (possibly with delay). Note that by Lemma 4, in each case listed below we have two starting starts to consider, b and c; we use ss to refer to the initial state.

1. Both b and c branch symmetrically: this will give a perfect tree (note that all tips will have the same state at each step). Thus this case is degenerate.
2. Exactly one of b and c branch symmetrically. Without loss of generality assume it is c.
    a. $c \mapsto (c, c)$; we consider only initial state b since c gives perfect trees (degenerate).
        i. $b \mapsto (b, c)$: Degenerate, produces perfect trees.
        ii. $b \mapsto (a, b)$:
            1. $a \mapsto a$: Degenerate, produces singleton domain trees.
            2. $a \mapsto b$: Degenerate, produces Fibonacci trees.
            3. **$a \mapsto c$: Domain branching with a perfect tree at offshoots after single step delay; equivalent to a perfect tree with one pair of tips advanced beyond the rest.**
        iii. $b \mapsto (a, c)$:
            1. **$a \mapsto a$: Produces trees of the form $(a, P_T(n-1))$, with a single tip plus a perfect tree arising directly from the root node.**
            2. **$a \mapsto b$: Produces a family trees similar to domain branching except that the offshoots are more advanced than the trunk. We can write this family recursively as $T_1 = ()$, $T_2 = (,)$, $T_{n+2} = (T_n, P_T(n+1))$**
            3. **$a \mapsto c$: Produces trees of the form $(P_T(n), P_T(n+1))$; both sides are perfect trees, but one is one step more advanced.**
    b. $c \mapsto (b, b)$
        i. $b \mapsto (b, c)$: Degenerate, produces perfect trees.
        ii. $b \mapsto (a, b)$:
            1. **$a \mapsto a$: Gives trees consisting of two equal singleton domain trees descending from the root node, if ss=c.** Degenerate if ss=b (produces singleton domain trees).
            2. **$a \mapsto b$: Gives trees consisting of two equal Fibonacci trees descending from the root node, if ss=c.** Degenerate if ss=b (produces Fibonacci trees).
            3. **$a \mapsto c$: 2 Nested models**
        iii. $b \mapsto (a, c)$:
            1. **$a \mapsto a$: 2 distinct models: Resemble perfect trees with the addition of singleton offshoots off every node between each main branching generation; models vary on whether this occurs at odd or even generations.**
            2. **$a \mapsto b$: 2 distinct models: ss=b gives nested "3-cherries"; recursively, $T(n) = \Big(T(n-2), \big(T(n-2), T(n-2)\big)\Big)$; ss=c gives pairs of these trees descending from the root.**
            3. **$a \mapsto c$: 2 Nested models.**

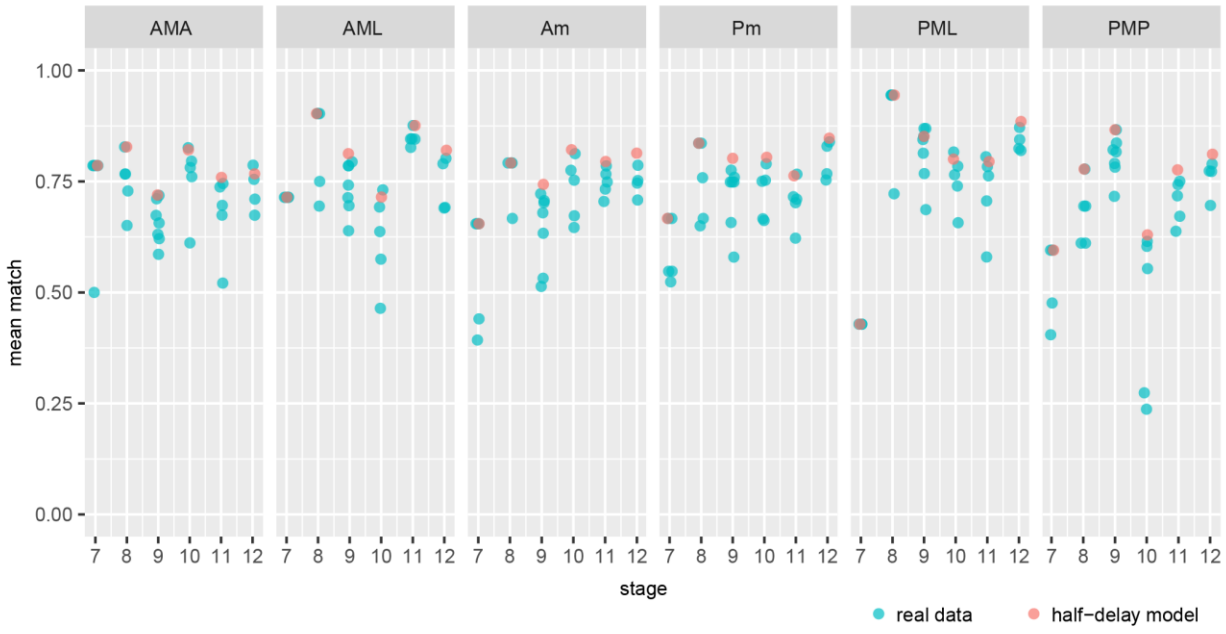
c. $c \mapsto (a, a)$
   i. $b \mapsto (b, c)$:
      1. **$a \mapsto a$: ss=b gives domain branching with $P(2)$ domains;** ss=c trivial
      2. **$a \mapsto b$: 2 Nested models.**
      3. **$a \mapsto c$: ss=b gives domain branching with perfect domain subtrees; compared to 2a(ii)3 ($c \mapsto (c, c)$, $b \mapsto (a, b)$, $a \mapsto c$) there is extra delay step, so more asymmetric;** ss=c degenerate, gives perfect trees.
   ii. $b \mapsto (a, b)$:
      1. $a \mapsto a$: Degenerate (domain branching).
      2. $a \mapsto b$: Equivalent to 2b(ii)2: Fibonacci, or each half is Fibonacci.
      3. **$a \mapsto c$: ss=b gives domain model with "slow growing" perfect tree offshoots;** ss=c degenerate, gives perfect trees.
   iii. $b \mapsto (a, c)$:
      1. $a \mapsto a$: Trivial, produces $\{(\,), (,), (,(\,))\}$
      2. **$a \mapsto b$: 2 Nested models.**
      3. **$a \mapsto c$: ss=b gives both perfect trees and trees of the form $(P_T(n), P_T(n+1))$;** ss=c degenerate, gives perfect trees only

3. Both b and c branch asymmetrically.
   a. $b \mapsto (a, x)$ and $c \mapsto (a, y)$, for $x, y \in \{b, c\}$: Degenerate; gives singleton domain branching (states b and c are redundant).
   b. $b \mapsto (w, x)$ and $c \mapsto (y, z)$, for $w, x, y, z \in \{b, c\}$: Degenerate; gives perfect trees.
   c. $c \mapsto (b, c)$
      i. $b \mapsto (a, b)$:
         1. **$a \mapsto a$: ss=c gives domain branching where every domain is singleton domain tree**; ss=b degenerate (singleton domain trees).
         2. **$a \mapsto b$: ss=c gives domain branching where each domain is Fibonacci**; ss=b degenerate (Fibonacci trees).
         3. **$a \mapsto c$: 2 Nested models.**
      ii. $b \mapsto (a, c)$:
         1. **$a \mapsto a$: 2 Nested structures with repeated singleton offshoots; for ss=b tip numbers equal Fibonacci, but structure is distinct.**
         2. **$a \mapsto b$: 2 Nested models.**
         3. **$a \mapsto c$: 2 Nested models.**

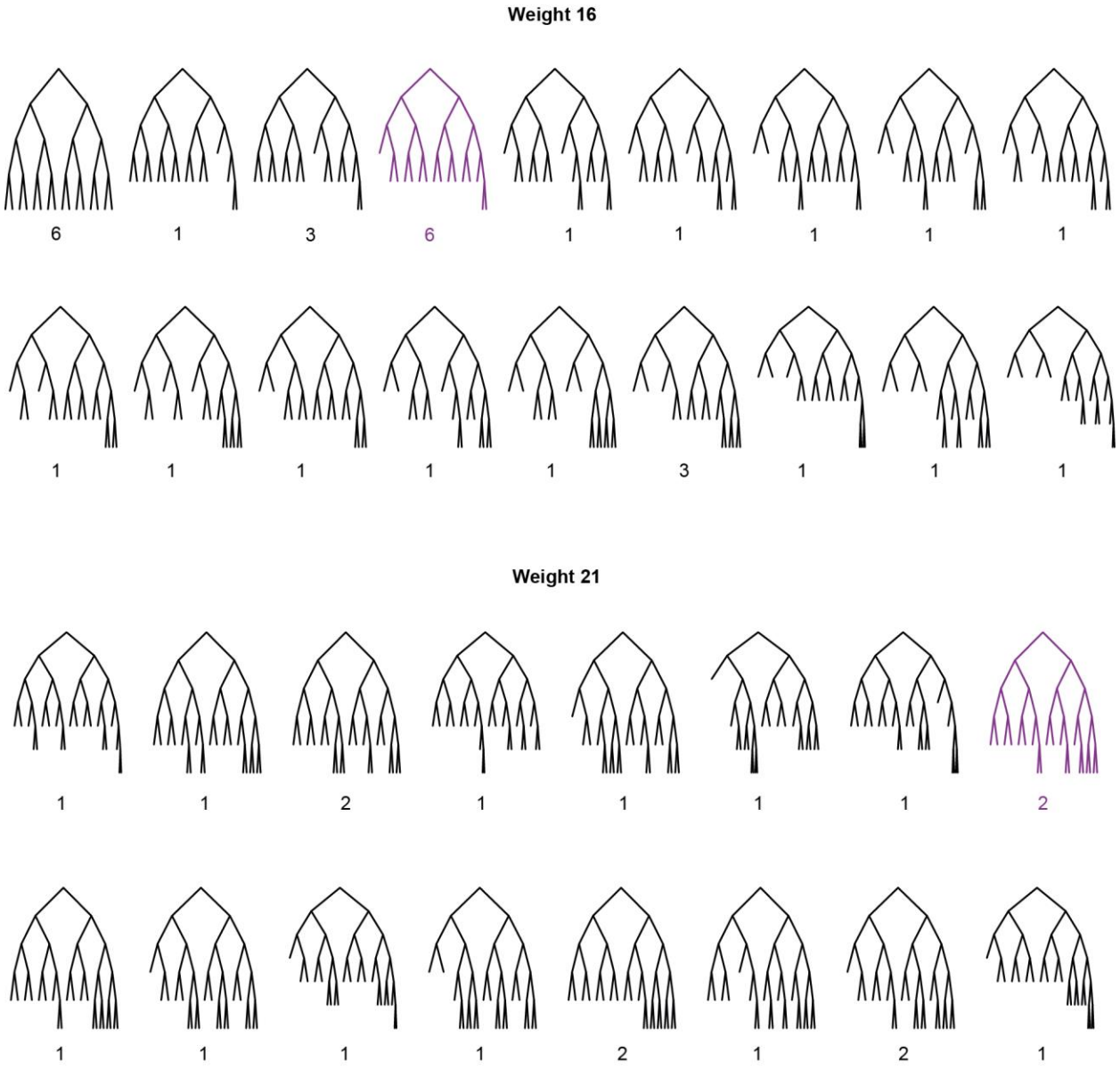
In total there are 32 distinct, non-degenerate 3 state models with 2 branching states. Of these, 20 are nested (these models are in pairs, with distinct models being produced depending on the initial state) and 7 are domain branching structures with infinite or finite (but non-trivial) offshoots. The remaining 5 trees are equivalent to 2 state models with additional (finite) structure around the root.

None of the models with two branching states had features that were observed in the ureteric tree data, although domain branching models may be of interest for modelling other structures such as the lung.

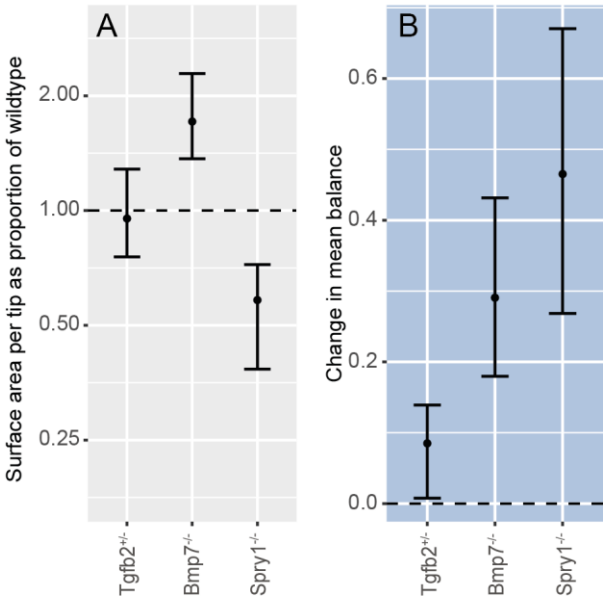
## SUPPLEMENTARY FIGURES



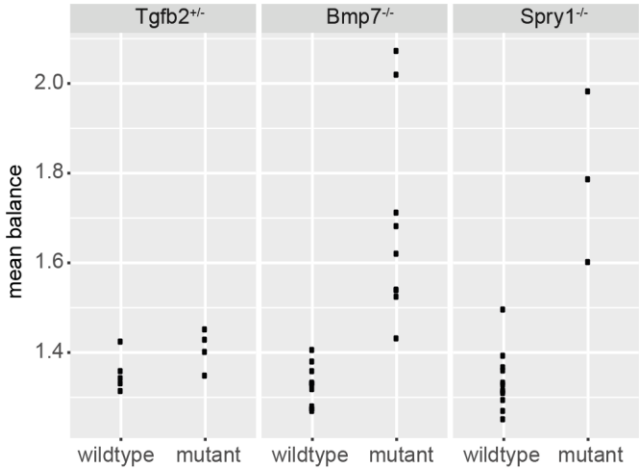

**Supplementary Figure 1:** Comparison of the best half delay model fit against variation between samples, within each stage and clade, showing that the model lies at the center of the natural variation. For each of the six developmental stages and six clade types, the corresponding set of clades was isolated from the wild type ureteric tree data. Clades which had not developed beyond a single tip were discarded (10 of 42 at stage 7 and 1 of 36 at stage 8). For each such set of clades, the half-delay model tree with the best average matching to the set was added (see Fig. 3A for definition of matching score). Then for each real or model tree, the average matching score with the other members of its set was calculated, and this metric was used to compare the model (red) against the real samples (blue). In 29 of 36 cases the half delay model has highest or equal highest average matching score, indicating that the model is closer to the sample clades on average than are any of the individual samples to the others; in the remaining 7 cases the model is surpassed by only small margins. A small amount of random horizontal permutation has been added to points in the plot to enable points with equal matching score to be distinguished.



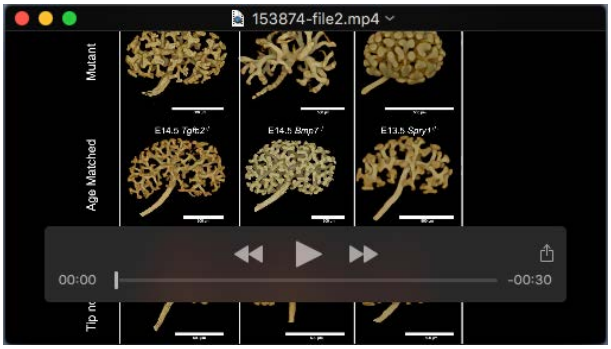

**Supplementary Figure 2:** Weight 16 and 21 subtree frequency in wild type time series dataset. In each case the highlighted tree is the centroid of the set (most representative structure, determined by maximum mean matching score with the other members of the set), and is also the half-delay tree of the given weight.




**Supplementary Figure 3**: Mutant phenotypes compared to tip-number matched wild type controls in terms of (A) surface area per tip; (B) balance, averaged over the binary branch points in each tree. Compare to Fig. 6F, Fig. 7A. Alternative control groups are to show that mutant differences are not an artefact of generally delayed development. Each plot shows estimated median change from wild type (dashed line) with 95% confidence interval, using a Wilcoxon rank sum test on the kidney level metrics. For A, values were first log transformed, and the estimates presented as proportional change. Control / test group sizes are: *Tgfb2*$^{+/-}$ 12/4; *Bmp7*$^{-/-}$ 12/9; *Spry1*$^{-/-}$ 12/3.




**Supplementary Figure 4**: **Mean balance of kidneys in mutant and control groups.** Control / test group sizes are: *Tgfb2*$^{+/-}$ (E14.5) 5/4; *Bmp7*$^{-/-}$ (E14.5) 12/9; *Spry1*$^{-/-}$ (E13.5) 12/3. Comparison is to same-age wildtype controls.




**Movie 1**: Ureteric trees from mutant *Tgfb2*$^{+/-}$ (A), *Bmp7*$^{-/-}$ (B) and *Spry1*$^{-/-}$ (C) embryos (top row) compared with wild type controls (middle row) and alternative wild type comparison kidneys matched by tip number (bottom row).

**References**

Barber, C. B., D. P. Dobkin and H. Huhdanpaa (1996). "The quickhull algorithm for convex hulls." ACM Trans. Math. Software **22**(4): 469--483.

Epp, S. S. (2004). Discrete Mathematics with Applications, Thomson Brooks/Cole.

Krane, D. E. (2003). Fundamental concepts of bioinformatics, Pearson Education India.

Metzger, R. J., O. D. Klein, G. R. Martin and M. A. Krasnow (2008). "The branching programme of mouse lung development." Nature **453**(7196): 745-750.