Journal of Theoretical Biology 365 (2015) 226–237

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Comparing and distinguishing the structure of biological branching

Timothy O. Lamberton [a], James Lefevre [a], Kieran M. Short [b], Ian M. Smyth [b,c], Nicholas A. Hamilton [a,d,*]

[a] Division of Genomics of Development and Disease, Institute for Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia
[b] Department of Biochemistry and Molecular Biology, Monash University, Clayton, Melbourne, Victoria, Australia
[c] Department of Anatomy and Developmental Biology, Monash University, Clayton, Melbourne, Victoria, Australia
[d] Division of Cell Biology and Molecular Medicine, Institute for Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia

HIGHLIGHTS

- A graph alignment metric for quantifying similarity in branching structures is developed.
- It shows increased ability to structurally distinguish ureteric trees across development.
- We developed a consensus metric to align multiple trees simultaneously.
- Its use to create an atlas of stages and across stages of development is shown.
- Strict stereotypic branching across development in control and mutant kidneys is shown.

ARTICLE INFO

Article history:
Received 27 May 2014
Received in revised form
29 September 2014
Accepted 1 October 2014
Available online 13 October 2014

Keywords:
Kidney
Development
Graph alignment
Algorithms

ABSTRACT

Bifurcating developmental branching morphogenesis gives rise to complex organs such as the lung and the ureteric tree of the kidney. However, a few quantitative methods or tools exist to compare and distinguish, at a structural level, the critical features of these important biological systems. Here we develop novel graph alignment techniques to quantify the structural differences of rooted bifurcating trees and demonstrate their application in the analysis of developing kidneys from in normal and mutant mice. We have developed two graph based metrics: *graph discordance*, which measures how well the graphs representing the branching structures of distinct trees graphs can be aligned or overlayed; and *graph inclusion*, which measures the degree of containment of a tree graph within another. To demonstrate the application of these approaches we first benchmark the discordance metric on a data set of 32 normal and 28*Tgfβ*$^{+/-}$ mutant mouse ureteric trees. We find that the discordance metric better distinguishes control and mutant mouse kidneys than alternative metrics based on graph size and *fingerprints* – the distribution of tip depths. Using this metric we then show that the structure of the mutant trees follows the same pattern as the normal kidneys, but undergo a major delay in elaboration at later stages. Analysis of both controls and mutants using the inclusion metric gives strong support to the hypothesis that ureteric tree growth is stereotypic. Additionally, we present a new generalised multi-tree alignment algorithm that minimises the sum of pairwise graph discordance and which can be used to generate maximum consensus trees that represent the archetype for fixed developmental stages. These tools represent an advance in the analysis and quantification of branching patterns and will be invaluable in gaining a deeper understanding of the mechanisms that drive development. All code is being made available with documentation and example data with this publication.

1. Introduction

1.1. Biology and software

The effective delivery and removal of biologically important fluids and gases is paramount to homoeostasis in multicellular organisms. Tubular branched networks are one structure by which this occurs, and branching morphogenesis is a fundamental process underlying the development of many organs of animals such as the vascular system, lung, kidney and mammary gland. In organs such as the kidney and lung, repeated branching of epithelial tissue efficiently fills a confined space while creating a large functional surface area at specialised tips to allow exchange of waste or oxygen. Recent reviews of the roles of branching morphogenesis in organ development can be found in Affolter et al. (2009) and Costantini and Kopan (2010).

* Corresponding author. Tel.: +61 7 3346 2033; fax: +61 7 3346 2101.
E-mail address: n.hamilton@imb.uq.edu.au (N.A. Hamilton).

http://dx.doi.org/10.1016/j.jtbi.2014.10.001
0022-5193/

Branching morphogenesis in developing mouse lungs has been suggested to follow a remarkable predictable programme with three distinct modes, which suggests hard-wired, genetic control of the branching process (Metzger et al., 2008). Similarly, the kidney is a heavily branched organ and shows genetically controlled branching and patterning mechanisms (Short et al., 2014). As such, branching morphogenesis of the kidney has been an area of considerable study with the cellular mechanisms of control have being of particular interest (Dressler, 2006; Kopan et al., 2007; Quaggin and Kreidberg, 2008). Much early kidney work focused on in vitro studies, and looked at development modes (al-Awqati and Goldberg, 1998) and quantifying spatial data (Cullen-McEwen et al., 2002; Watanabe and Costantini, 2004). The majority of kidney branching events are bifurcations, with trifurcations occurring relatively rarely (Short et al., 2014). Often, as the internal branches grow, the trifurcated branches will reform into a pair of bifurcated branches, with a pair of the trifurcated branches growing away from the third (Watanabe and Costantini, 2004). This suggests that trifurcations may be cases of a bifurcation followed by a rapid successive bifurcation (al-Awqati and Goldberg, 1998), which cannot be distinguished until the branches extend over time. More recent work in mouse kidneys suggests that a self-avoidance mechanism of branching structures could explain the growth modes seen in the lung and other studies (Davies et al., 2014).

The use of optical projection tomography (OPT) combined with quantitative imaging methods (Short et al., 2010) have demonstrated that ex vivo culture methods often poorly reflect in vivo development. Further refinements to this approach (Short et al., 2013) have seen the development of software packages which are able to extract and analyse skeletons that describe the branching ureteric tree, and reducing the 3D branching volumes of the ureteric tree of the kidney to a set of points and edges joining them. These methods and others were then applied to large scale data sets in Short et al. (2014) to give a comprehensive, quantitative, multiscale analysis of mammalian kidney development.

An important aim of in vivo studies is to objectively and quantitatively assess kidney development under a range of genetic conditions and determine individual genetic contributions. For instance, it has been demonstrated that mouse embryos heterozygous for a mutation in the gene encoding transforming growth factor-beta 2 ($Tgf\beta 2^{+/-}$) exhibit decreased branching, a significant developmental delay, changes to the branching programme, and a significantly smaller renal pelvis compared to control embryos (Short et al., 2013, 2010). At early developmental stages, $Tgf\beta 2^{+/-}$ and control kidneys are similar as measured by number of branch points and (terminal) tips (Fig. 1). However, as development progresses the increase in complexity of the $Tgf\beta 2^{+/-}$ ureteric tree (as measured by tip number) kidney growth lags the controls.

While tip number is a simple metric that can distinguish later stage mutant kidneys from controls it provides no information about the architecture of the distinct organs. It is therefore unclear as to whether the smaller later stage $Tgf\beta 2^{+/-}$ kidneys are similar in structure to the earlier stage controls but have retarded growth, or if their branch patterning is fundamentally different. Although earlier stage 7–10 controls and mutant kidneys have the same number of tips, the branching pattern could ostensibly be different, precipitating a later stage tip deficit.

A broader question is whether a fundamental set pattern exists which controls ureteric tree formation. For instance, if two ureteric trees are examined that have the same number of tips will their branching structures be identical or is the bifurcation of branches to some extent random in direction but at a constant rate giving rise to trees of the same size but with very different structures? Moreover, it is unknown whether the growth of the ureteric tree is

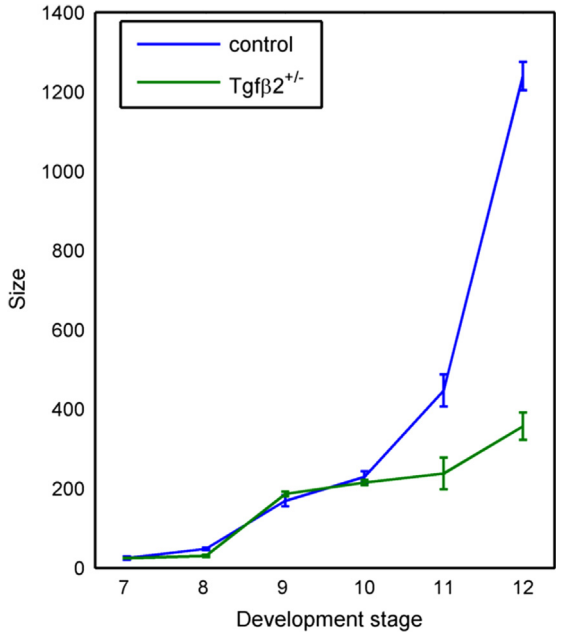

Fig. 1. Detecting difference between ureteric trees of control and $Tgf\beta 2^{+/-}$ kidneys across early development. Shown are the mean numbers of the size (number of branch points and tips) for ureteric trees of control and mutant ($Tgf\beta^{+/-}$) mouse kidneys at distinct developmental stages, adapted from Short et al. (2010) plot showing number of tips. While between stages 7 (12 dpc) to 10 (13.75 dpc) the controls and mutants are statistically indistinguishable, by stages 11 (14.5 dpc) and 12 (15.5 dpc), $Tgf\beta 2^{+/-}$ kidneys show notably smaller size than control kidneys at the same stage. Error bars are standard errors of the mean (SEM).

structurally stereotypical to the point where the branching pattern of any later stage normal kidney can be seen to be an extension of the structure of earlier stage kidneys.

In order to address these questions, metrics that are refined to be structurally aware are necessary for the comparison of the branching patterns. Here we describe the development of a number of such metrics for quantifying, structurally comparing and distinguishing rooted bifurcating structures. We evaluate the utility of these metrics and compare their performance against readily available metrics such as tip number. The new metrics are then used to detect and understand the structural nature of the $Tgf\beta 2^{+/-}$ mutant kidney phenotype. Finally, we give a detailed analysis of the patterning of the normal mouse kidney from stages from 12 dpc to 15.5 dpc (stages 7 to 12) and show that the growth is highly controlled and stereotypic.

Supporting software implementing the algorithms and methods, as well as example data, are available as Supplementary material. In order to improve readability, mathematical proofs and some details are included in Section 4.

Note: In the following, the term *tree* is used in the sense of a connected, acyclic, undirected graph T with a vertex set $V(T)$ and an edge set $E(T)$, derived from the physical kidney ureteric tree.

1.2. Quantitative tree metrics

This work considers four distance metrics which provide measures of the similarity between trees. As they are distance metrics they may also be used in other analyses such as hierarchical clustering (for example).

1.2.1. Graph size difference

The simplest metric for tree comparison, which we include as a benchmark for the more sophisticated methods, is to compare the