

# Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution

Romain Fernandez<sup>1,2,5</sup>, Pradeep Das<sup>3,5</sup>, Vincent Mirabet<sup>3</sup>, Eric Moscardi<sup>1</sup>, Jan Traas<sup>3</sup>, Jean-Luc Verdeil<sup>4</sup>, Grégoire Malandain<sup>2</sup> & Christophe Godin<sup>1</sup>

**Quantitative information on growing organs is required to better understand morphogenesis in both plants and animals. However, detailed analyses of growth patterns at cellular resolution have remained elusive. We developed an approach, multiangle image acquisition, three-dimensional reconstruction and cell segmentation–automated lineage tracking (MARS-ALT), in which we imaged whole organs from multiple angles, computationally merged and segmented these images to provide accurate cell identification in three dimensions and automatically tracked cell lineages through multiple rounds of cell division during development. Using these methods, we quantitatively analyzed *Arabidopsis thaliana* flower development at cell resolution, which revealed differential growth patterns of key regions during early stages of floral morphogenesis. Lastly, using rice roots, we demonstrated that this approach is both generic and scalable.**

The control of morphogenesis during plant and animal development is a major question in developmental biology. Although several studies have provided profound insight into the molecular regulatory networks that act during development, the effects of such networks on shape transformations are often only described qualitatively. Indeed, describing shape and shape change as a geometrical output of gene activity requires the quantification of growth patterns with cellular resolution. Obtaining accurate geometric information about cell positions and shapes will be essential to develop quantitative growth models<sup>1–6</sup> and to accurately test their predictions. Although several recent methods, mostly based on nuclear tracking in animal cell populations<sup>7–10</sup>, have addressed cell positioning and tracking, they cannot provide information on three-dimensional (3D) cell geometry.

We are interested in characterizing and quantifying growth in plant meristems, which are small groups of pluripotent cells that give rise to all organs in both the shoot and the root. Three types of meristems can be identified: the root apical meristem, the shoot apical meristem and the floral meristem. Meristem size may

vary from a few cells to several thousand cells, depending on the species and the meristem type. In *Arabidopsis thaliana*, an initiating floral meristem measures 20–35  $\mu\text{m}$  along each axis and contains 30–50 cells, each  $\sim 150\text{--}250\ \mu\text{m}^3$ . With divisions every 19–24 h on average<sup>11</sup>, the flower quickly grows into an object measuring 80–100  $\mu\text{m}$  along each axis and containing several hundred cells, even before the onset of differentiation and organ formation. Rice root meristems, in contrast, may contain many thousands of cells<sup>12</sup>. Plant cells are typically separated from each other by cell walls and cell membranes of less than 0.5  $\mu\text{m}$  in thickness, whose correct identification is essential to accurately determine cell shapes.

Different methods have been used to image and reconstruct tissues at cellular resolution<sup>7–10,13–15</sup>. Confocal laser-scanning microscopy has been used to image fixed roots and to semiautomatically identify individual cells<sup>16</sup>. In living plant tissues, most protocols have been restricted to surface reconstructions<sup>17–19</sup>. Confocal microscopy has the advantage of providing access to the inner parts of the tissue, but the fluorescence signal fades in the inner parts of thick tissue, making it difficult to identify cell walls in the deeper layers. Additionally, cell walls perpendicular to the focal axis of the microscope are very often not seen. These drawbacks severely limit the use of automatic processing to extract additional information from confocal images. Here we present a method to generate 3D digitized tissues at cell resolution and to automatically track cell lineage during growth. To create a digitized tissue that can be used to quantitatively analyze growth in four dimensions, we developed an experimental pipeline comprising two key steps: multiangle image acquisition, 3D reconstruction and cell segmentation (MARS) and automated lineage tracking (ALT). We applied this pipeline to analyze the developmental dynamics of young floral meristems over 70 h.

## RESULTS

### Multi-angle, real-time imaging via confocal microscopy

To overcome limitations owing to either tissue thickness or microscope resolution anisotropy (Supplementary Fig. 1), we

<sup>1</sup>Equipe-projet Virtual Plants, Institut National de la Recherche en Informatique et en Automatique, Sophia Antipolis, France and Unité Mixte de Recherche Développement et Adaptation des Plantes, Montpellier, France. <sup>2</sup>Equipe-projet Asclepios, Institut National de la Recherche en Informatique et en Automatique, Sophia Antipolis, France. <sup>3</sup>Laboratoire Reproduction et Développement des Plantes, Ecole Normale Supérieure de Lyon et Institut National de la Recherche Agronomique, Lyon, France. <sup>4</sup>Centre de Coopération Internationale en Recherche Agronomique pour le Développement et Unité Mixte de Recherche Développement et Adaptation des Plantes, Montpellier, France. <sup>5</sup>These authors contributed equally to this work. Correspondence should be addressed to C.G. (christophe.godin@inria.fr).

devised MARS based on the idea of visualizing a single sample from multiple angles (Fig. 1a). We stained shoot apical meristems bearing only flower buds between stages 1 and 3 with the vital dye FM4-64 and imaged them under a confocal laser-scanning microscope (Online Methods). First we imaged the flower from the top ('top-view stack'), and then we manually tilted it by 30–40° and reimaged it ('tilt-view stack'). We repeated this process such that we imaged every flower from at least three different angles (Fig. 1b–d). We avoided rotation angles greater than 50° to maintain sufficient common context between each image.

### 3D reconstruction and segmentation of volumetric images

We computed 3D cell-segmented images from multiangle scans in two steps. First, we fused the images acquired under different orientations to enhance the quality of the cell outlines (Supplementary Note 1), which required co-registering every image stack with a reference stack. We used a hierarchical strategy wherein, for convenience, we chose the top-view stack as the reference stack and in an incremental procedure registered the other (floating) tilt-view stacks onto this reference. Based on at least four landmarks that we manually identified in both the reference and the floating stacks (Fig. 1b–d and Supplementary Fig. 2) using surface reconstructions<sup>19</sup>, we computed an initial rigid transformation (which consisted of rotation and translation; Fig. 1e). This first step usually yielded only an approximate matching (Fig. 1f), possibly because of local shape changes resulting from plant growth or from changes in the mechanical or physiological properties of the cells. Then, a more robust, rigid transformation was automatically computed using all available image information<sup>20</sup>, followed by a refined nonlinear transformation computed using a block-based pyramidal algorithm<sup>21</sup> (Fig. 1g,h). This process substantially restored membrane-to-cytoplasm contrast

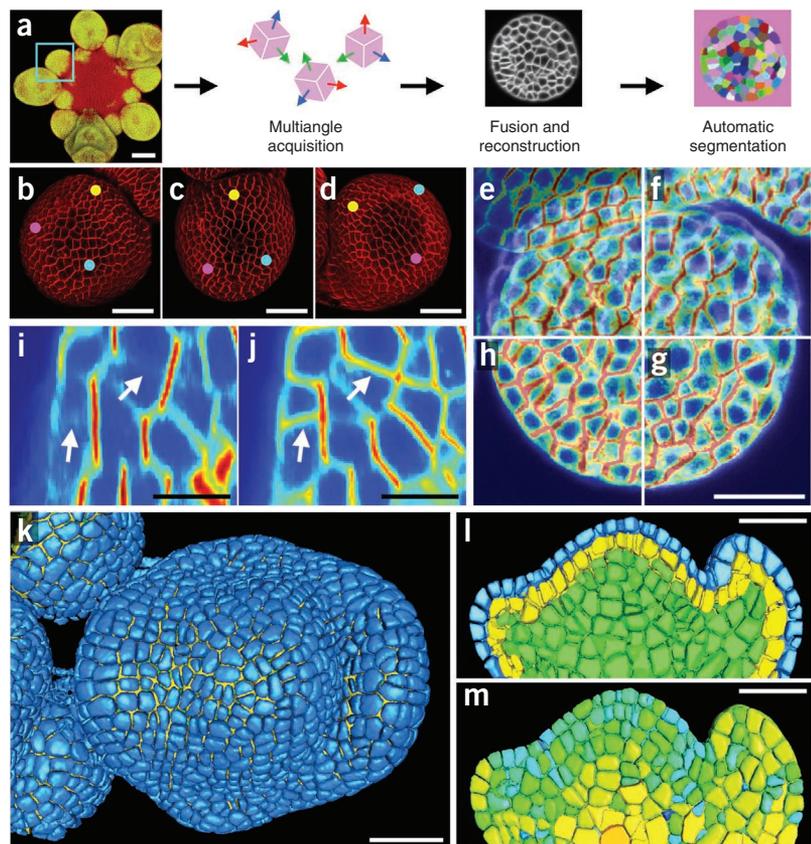
in the entire 3D structure and also typically provided a much more homogeneous contrast distribution in the tissue (Fig. 1i,j and Supplementary Fig. 3).

Second, we segmented the resultant 3D images using a 3D watershed algorithm<sup>22</sup> to identify the cells as individual 3D objects (Fig. 1k–m and Supplementary Note 2). Such an algorithm is well suited for processing our images, in which the objects to segment were dark areas (cell interiors) surrounded by bright and thin boundaries (cell walls). To initialize the procedure, every cell must have a unique marker. We defined these markers as the main local minima (computed using the *h*-min operator<sup>22</sup>) of the noise-filtered image.

We next assessed whether MARS was generic and scalable by applying it to rice roots, which are large (200–450 μm) and have a complex cellular organization (with about 30 cell layers)<sup>23</sup>. We observed the root meristem under four azimuthal angles spaced approximately 90 degrees from one another. The cell recognition was of better quality in the four-fused-view reconstructions than in the one-view reconstructions: the algorithm could segment 16,400 cells in the rice root tip for the four fused views and for one view with an estimated cell recognition rate of 88% and 61% in the root center, and of 93% and 87% in the cortical region, respectively (Online Methods, Supplementary Figs. 4 and 5 and Supplementary Note 3).

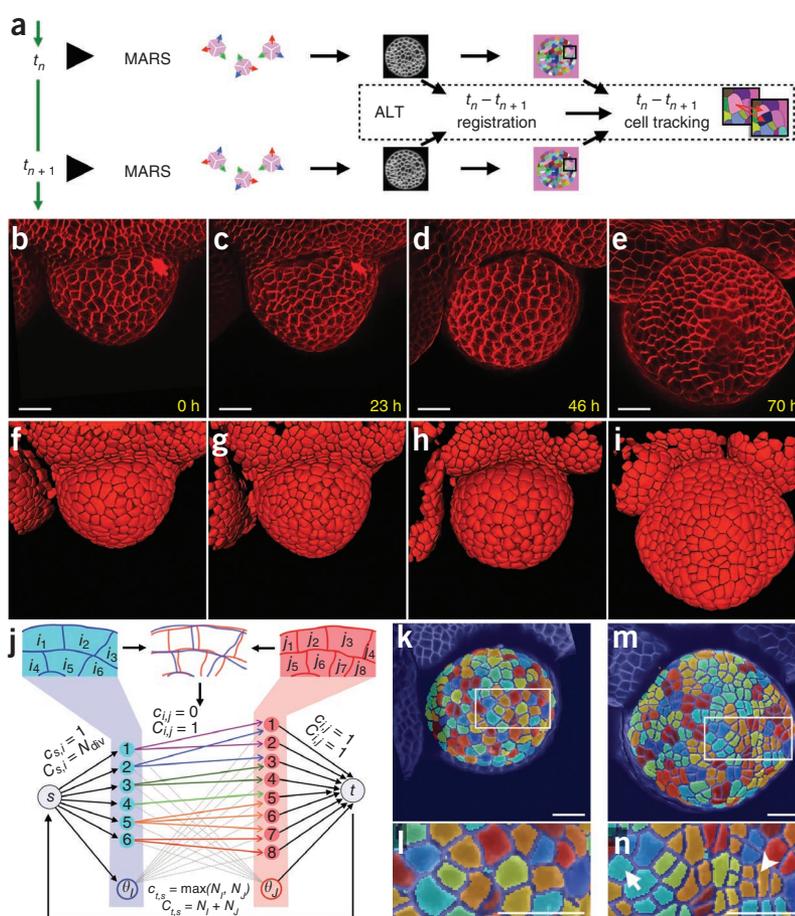
### Automatic lineage tracking during growth

To identify cell lineages during floral growth, we developed a second software pipeline called ALT (Fig. 2a). First, we used MARS to track the growth of young flower meristems (primordia) by imaging them from multiple angles every 24 h for up to 4 d and generating 3D cell-segmented images at each time point (Fig. 2b–i



**Figure 1** | MARS. (a) Pipeline for segmented 3D tissue reconstructions. From an image of inflorescence expressing a flower-specific *GFP* marker (*pLEAFY::ER-GFP*), one flower (blue box) is chosen for further study (left). For each flower, image stacks are acquired from multiple angles and computationally fused into a single 3D reconstruction, which is then automatically segmented. (b–d) Example confocal images of a flower acquired from three angles. After image acquisition, three to eight common landmark vertices (colored dots) are identified in each image stack. (e–h) These image stacks are registered using a hierarchical process with increasing precision and then fused as follows: superposition of the raw images from two different views (e), a manual, rigid registration (f), an automatic, linear registration (g) and a dense, nonlinear registration (h). (i,j) Images showing details of single-angle (i) and fused multiangle (j) acquisitions. Arrows highlight signal for cell outlines and interiors that is low or absent in single-angle images. (k–m) After automatic segmentation, tissue was visualized with a full organ reconstruction (k) or with virtual sections using color codes for cell layer (l) or cell volume (m). Scale bars, 50 μm (a), 25 μm (b–h,k–m) and ~10 μm (i,j).

**Figure 2 | ALT.** (a) ALT pipeline. Reconstructed and segmented data from serial time points ( $t_n$  and  $t_{n+1}$ ) are compared to automatically determine cell lineages. (b–e) Confocal image surface projections of the top view of a wild-type flower collected at the indicated times. (f–i) Segmented 3D reconstructions of each time point (corresponding to images in b–e, respectively), with many cells belonging to the shoot apex and to neighboring flowers manually discarded. (j) ALT is modeled as a problem of flow in which plausible lineage correspondences are expressed in a valued graph. Edges connect parent cells of the source image ( $i_{1-6}$ ) to candidate daughter cells of the target image ( $j_{1-8}$ ).  $\theta_s$ ,  $\theta_j$ ,  $s$  and  $t$  are additional nodes representing, respectively, void source cell, void target cell and source and target nodes emitting and collecting the flow. Each edge may carry units of flow that are bounded by the edge minimal and maximal capacities,  $c_{ij}$  and  $C_{ij}$ , respectively, and is associated with a cost per unit flow.  $N_i$  and  $N_j$ , the number of source and target cells, respectively;  $N_{div}$  hypothesized maximum number of daughters that can be associated with a given parent. (k–n) Application of the ALT algorithm to floral growth: cells at an earlier time point (k,l) and corresponding daughter cells (m,n) were colored to show lineages. Detailed views (L,n) show cells that have undergone up to two rounds of cell division (arrowhead). Also visible is a tracking error (cell marked with an arrow should be red). Scale bars, 10  $\mu$ m.



and **Supplementary Fig. 6**). Next, we manually identified an initial set of surface (L1 layer) cell lineages for every pair of consecutive segmented images from sequential time points (**Supplementary Note 4**). We used these ‘known’ reference lineages to initialize the set of high-confidence lineages of the ALT pipeline. ALT begins by estimating a rigid transformation between two images by minimizing the total square distance between parent cells and their descendants in high-confidence lineages. It then estimates new high-confidence cell lineages iteratively by performing the following steps until no candidate cell lineages remain (**Supplementary Note 4**).

In the first step, ALT computes a deformation field based on the high-confidence lineages. The residual positioning error is used to define a nonlinear transformation of the image voxels as a dense vector field (a vector per voxel). For this, each high-confidence lineage is transformed into one vector that links the center of mass of the parent cell to the center of mass of all its descendants, and the dense vector field is computed by interpolating between these vectors.

In the second step, the deformation field is refined. The initial nonlinear transformation is refined by an automated nonlinear registration algorithm that matches the voxel intensities of the two images. This defines a more precise nonlinear registration that makes it possible to overlap the two images and compare their segmentations.

In the third step, ALT builds lineage hypotheses and solves the lineage problem. Pairwise registration of sequential images permits candidate daughter cells for each parent cell to be

hypothesized, based on their spatial proximity in a common image frame. ALT formalizes the lineage problem as the search for a mapping between cells in sequential images that globally minimizes the total distances between the mapped cells. To solve this problem, ALT treats it as an optimal flow problem (**Fig. 2j**, **Supplementary Note 4** and Online Methods).

In the final step, ALT updates the set of high-confidence lineages. ALT automatically scans the list of lineages between two images to identify those lineages that comply with additional criteria (**Supplementary Note 4**) by applying plausibility tests based on geometric and topological rules that are not used to guide the algorithm. If a cell lineage passes the test, it is then added to the list of high-confidence lineages.

We used the ALT algorithm to identify lineages during flower growth over 70 h (**Fig. 2k–n**). ALT provided 84–100% accurate cell lineage information, depending on the extent of growth (**Table 1** and Online Methods).

### Validation of the MARS-ALT pipeline

We assessed the quality of the segmentation algorithm by an exhaustive visual examination of the results obtained from the automatic segmentations of *Arabidopsis* floral meristems. We examined eight MARS-analyzed floral meristems cell by cell using MARS exploratory tools (**Supplementary Note 5**). This included one flower (‘flower A’) imaged at an optimal resolution (varying from  $576 \times 576$  to  $680 \times 680$  pixels) at four time points separated by about 24 h each, and two flowers (‘flower 2’ and ‘flower 3’) imaged at a higher resolution ( $1,024 \times 1,024$  pixels)

**Table 1** | Results of the MARS-ALT pipeline applied to flower development

	$T_0$ (0 h)	$T_1$ (23 h)	$T_2$ (46 h)	$T_3$ (70 h)
Segmentation (MARS)				
Total number of cells imaged and segmented	716	798	706	1,236
Mean cell volume ( $\mu\text{m}^3$ )	212	189	250	299
Number of epidermal (L1 layer) cells	344	405	253	379
Number of subepidermal (L2 layer) cells	254	276	251	351
Percentage of correct MARS-segmented cells	97.8	96.2	98.6	98.1
Lineage tracking (ALT) <sup>a</sup>				
Total number of cells tracked by ALT	357	444	582	1236
Increase in total volume of cells tracked during $T_0$ - $T_3$ <sup>b</sup>	NA	1.04-fold	2.26-fold	2.50-fold
S.d. of parent cell volume increase in a $T_0$ - $T_3$ transition	NA	0.29	0.94	1.38
Reference lineages provided for ALT initialization	4	26	33	NA
Cells with one daughter at the next time point (no division)	252	194	141	NA
Cells with two daughters (1 division)	94	154	240	NA
Cells with three daughters (2 divisions)	1	16	109	NA
Cells with four daughters (3 divisions)	0	4	51	NA
Cell lineages identified manually for ALT evaluation <sup>c</sup>	98	129	184	NA
Percentage of correctly detected mother-to-daughter lineages <sup>d</sup>	100	89	84	NA
Percentage of correctly detected daughter-to-mother lineages <sup>e</sup>	NA	100	94	94

All data presented are from the 'flower-A' time-course experiment (Fig. 2a). NA, not applicable.

<sup>a</sup>For clarity, only the ancestors of cells present at  $T_3$  were retained for analysis. <sup>b</sup>At  $T_0$ , the total volume of the 357  $T_0$ - $T_3$ -tracked cells was  $77.5 \times 10^3 \mu\text{m}^3$ . The fold-increase values in the table are based on the automatically tracked cells. As a consequence, they may be slightly affected by lineage errors. <sup>c</sup>A manual lineage identification was performed to determine the ancestors of all epidermal and subepidermal cells in the upper part of the flower bud (up to and including the sepals) at  $T_3$ . <sup>d</sup>This is the most strict definition of a lineage wherein, for a given parent cell, ALT provides the identical set of daughters that had been identified manually. <sup>e</sup>In this definition of a correct lineage, for any given daughter cell, ALT provides the same parent cell that had been identified manually.

at two time points each. We created a reference segmentation for each of these multiangle images by manually correcting an automatic segmentation. To assess the performances of different segmentation algorithms or the effect of varying their parameters, we compared, for a given image, the automatic segmentation produced by these algorithms to the reference segmentation of this image. As two segmentations, each containing hundreds of cells in three dimensions, cannot be manually compared routinely, we designed a special algorithm to automate this comparison process (Supplementary Note 6). We considered cells correctly identified if they could be associated one-to-one with a manually corrected cell by the validation mapping algorithm.

Merging increasing numbers of view stacks made it possible to substantially increase the percentage of correctly segmented cells. Thus, when a view stack alone was processed, 91.2% of cells were correctly identified (Fig. 3a). With the addition of one and two tilt-view stacks, identification improved to 95.7% and 98.5%, respectively. Additional analysis revealed that in one- and two-angle segmentations, only about 60.9% and 81.5% of cells, respectively, had the correct volume relative to the reference segmentations (Fig. 3b). Here we considered cells correctly identified if, in addition to the above criterion, the volume of the mapped cells differed by a maximum of 10%. Image resolution had little effect on the efficiency of the algorithm, so that the results obtained for images acquired at optimal resolution or with oversampling differed by only about 2% (data not shown). We then analyzed how the segmentation errors were distributed in the data for floral meristem. As expected, in one-angle segmentations, the percentage of correctly segmented cells decreased steadily with the depth of the tissue (Fig. 3c). The fusion of one or two additional view stacks made it possible to substantially improve segmentation quality (Fig. 3c) such that at depths greater than 40  $\mu\text{m}$ , up to 7% more cells were correctly identified.

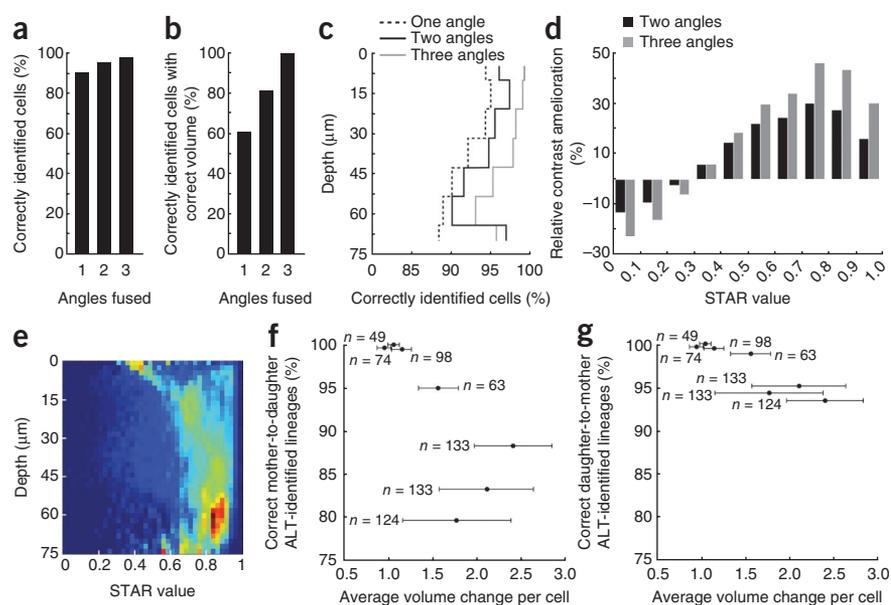
We analyzed how the cell wall signal depended on both its orientation with respect to the focal axis of the microscope and on

its depth in the tissue. Cell walls with different orientations have different projected silhouettes in the focal plane. We defined the ratio between the silhouette area of a cell wall projected in the focal plane and the total actual area of the wall as the silhouette to total area ratio (STAR), which has values between 0 and 1. Values close to 1 are reported for cell walls parallel to the focal plane, and values close to 0 are reported for cell walls perpendicular to it. The use of multiangle acquisitions enhanced the contrast for the different categories of cell-wall orientations (Fig. 3d), with particular improvement in regions with high STAR (that is, walls nearly parallel to the focal plane) in the reference image stack. Distribution of cell-wall orientations was not homogeneous in the tissue (Fig. 3e). At the meristem surface, cell walls had a horizontal bias (STAR > 0.5), consistent with our observation that segmentation quality for surface cells was lower in single-angle segmentations (Fig. 3c). The use of multiple angles efficiently corrected this effect of the wall orientation distribution.

To assess the quality of the automatic cell-lineage tracking procedure, we used the four time points ( $T_0$  to  $T_3$ ) from the flower-A data (Fig. 2b–i) and validated the MARS-processed segmentations with the automated procedure described above (Table 1). We then manually identified a large number,  $k_i$ , of reference lineages between parent cells (at  $T_{i-1}$ ) and daughter cells ( $T_i$ ) for each time transition ( $i = 1, 2$  or  $3$ ,  $k_1 = 98$ ,  $k_2 = 129$  and  $k_3 = 184$  lineages) in superficial as well as internal cells. We compared the results of the ALT pipeline to these manually identified reference lineages (Table 1). About half the cells at time  $T_0$  yielded daughter cells at time  $T_3$ , which corresponded to an average of 3.5 daughter cells at  $T_3$  per parent cell at  $T_0$ , for a 4.8-fold volume increase of the corresponding growing region. Our data showed a burst of growth between  $T_2$  and  $T_3$ , during which both the volume and the cell division rate increased. When the change in volume was small ( $T_0$  to  $T_1$ ), ALT correctly identified all of the reference lineages. The performance was slightly worse (89% of cell lineages identified) with a twofold increase in volume. With even bigger

**Figure 3** | Validation of MARS-ALT results.

(a–d) Comparisons of eight individual flowers (representing 6,038 cells) that were segmented using reconstructions from one, two or three image stacks. Plotted are data for correctly identified cells (a) and correctly identified cells with correct volume (b) relative to the manually curated three-angle segmentation. Percentage of correctly segmented cells as a function of depth in the tissue in reconstructions compared to the manual 3-angle segmentation was plotted (c; depth was measured along the focal axis of the reference image relative to the meristem surface). Relative amelioration of contrast (d) between cell membrane signal and cell interiors (that is, the quality of cell outlines) for various cell membrane orientations relative to one-angle reconstructions was graphed. STAR is a measure of cell-membrane orientation. (e) Heat map of cell membrane orientation distribution as a function of depth in the flower. Red color, high values of the distribution, and blue, low values. (f,g) ALT results from time-course experiments covering different developmental stages from different flowers, showing percentage of mother cells with fully correct daughter cells (f) or percentage of daughter cells with correct mother cells (g) for average volume change per cell. Horizontal lines show s.d. ( $n = 49$ –184 lineages, depending on the point in the diagram). Points clustered at 100% were manually separated for clarity.



increases of volume (2.5-fold between  $T_2$  and  $T_3$ ), only 84% of the cell lineages were perfectly identified. But this corresponded to an incorrect detection of a parent cell for only 6% of daughter cells. By extrapolation to the entire dataset, ALT correctly detected the parent cell for 1,162 of the 1,236 cells at  $T_3$ . We observed similar trends in other time-course experiments (Fig. 3f,g).

The ALT algorithm identifies not only the cellular lineages but also the precise geometrical deformations that occur during growth. Based on this four-dimensional reconstruction, it is possible to estimate the intermediate 3D images by interpolation between serial time steps. This virtual 3D representation of continuous growth at cellular resolution is a powerful way to observe organ-growth dynamics (Supplementary Note 7 and Supplementary Videos 1 and 2).

### Analysis of early floral growth at cellular resolution

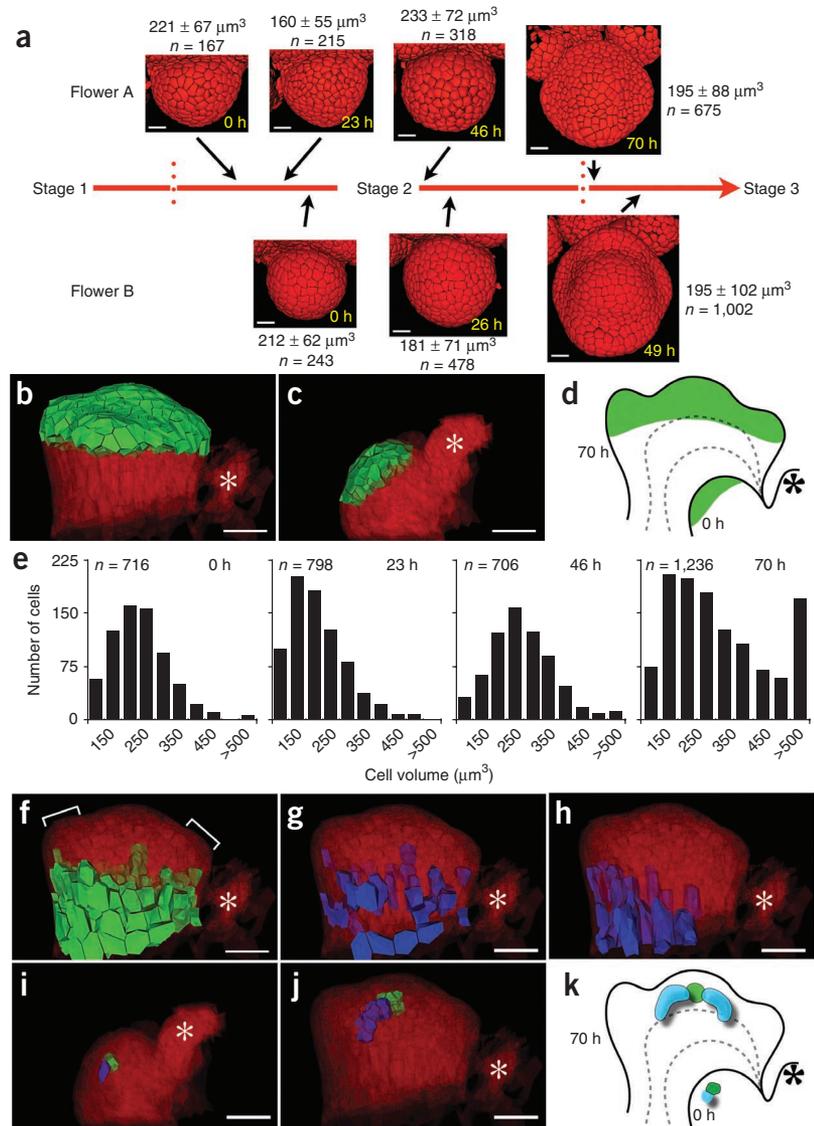
The *Arabidopsis* flower initiates on the flanks of the shoot apical meristem as a stage-1 bud. During stage 2, the bud becomes separated from the shoot apical meristem and goes through a long growth phase. The first organs form at stage 3, when the flower bud comprises four sepal primordia surrounding a dome of meristematic tissue in the center, which gives rise to all other floral organs. We used MARS-ALT to analyze cell lineage trajectories and cell morphometrics of this entire zone in two independent time-course experiments (Fig. 4a).

First, we used the digital reconstructions of the last time points in both time courses, selected all the cells of the sepals and the floral dome (the ‘floral organ zone’) and used ALT to identify these cells’ ancestors at the earliest time points in the two experiments. We then mapped the locations of those ancestors in the flower and observed that at mid-stage 2, these cells were located not at the vertical summit of the flower, but abaxially, in the region most distal to the shoot apical meristem (Fig. 4b,c). This suggests that as the flower develops, it undergoes unequal growth along the medial (proximo-distal) axis that pushes the abaxial cells toward the

summit (Fig. 4d). We next investigated whether such an unequal tissue growth was linked to particular cellular characteristics, such as cell size. We analyzed cell volumes during floral growth (Fig. 4e and Supplementary Fig. 7). We observed that at stage 3 in both time-course experiments, the floral organ zone cells had significantly similar mean volumes ( $195 \pm 88 \mu\text{m}^3$  for flower A ( $n = 674$  cells);  $195 \pm 102 \mu\text{m}^3$  for flower B ( $n = 611$  cells);  $P = 0.9833$  in a Welch two-sample  $t$ -test). However, during the transition from stage 2 to 3, many cells (12%, 70 h of flower A time course) grew to volumes of over  $500 \mu\text{m}^3$  (Fig. 4e). We observed that all of these cells were located below the floral organ zone and likely belonged either to the emerging vascular strands or to the elongating pedicel (Fig. 4f). Whereas the smallest cells in this group ( $500$ – $600 \mu\text{m}^3$ ) were distributed uniformly around the base of the flower (Fig. 4g), the largest cells ( $1,000$ – $2,200 \mu\text{m}^3$ ) were located primarily along the abaxial side of the flower (Fig. 4h). Thus, in contrast to the majority of cells, these cells had not divided in this 24 h period. These data suggest that changes in cell division rules are an important part of floral growth and specification and that these rules may be altered locally to control morphogenesis.

To explore cellular growth and dynamics in the floral organ zone, we examined the stem cells and the stem cell organizing center, which are crucial for proper flower development. We localized the putative stem cells and organizing center over time using a stem cell-specific reporter line. Consistent with our findings for the floral organ zone, the putative stem cells and organizing center ancestors were also located in an abaxial region at early stages (Supplementary Fig. 8). It has long been suggested that organizing center cells are large and slow dividing<sup>24</sup>, but it has never been possible to properly examine their morphologies or dynamics in four dimensions. We found that at early mid-stage 2, the uppermost (L3 layer) putative organizing center cell had an exceptionally large volume ( $402 \mu\text{m}^3$  in flower-A data) that was almost twice the mean volume of all floral cells ( $212 \mu\text{m}^3$ ; Table 1). Descendants of this cell formed a tight cluster at stage 3,

**Figure 4** | Application of MARS-ALT to flower development. **(a)** Reconstructions from two MARS-ALT time-course experiments (flower A and flower B) tracking wild-type floral growth. For comparison, stages 1–3 are positioned on a common timeline. Times of image acquisition (which are unrelated between the two experiments) and mean volumes ( $\pm$ s.d.) of cells in the floral organ zone are indicated. **(b–k)** Image renderings and schematics are side views of reconstructed flower A at early mid-stage 2 (**c,i**) or early stage 3 (**b,f–h,j**). The flower and a part of the shoot apical meristem (asterisk) are shown in red, with cells of interest in green or blue. Side view at 70 h with all floral organ zone cells selected (**b**) and at 0 h showing the ancestors of the cells in **b** (**c**). The schematic in **d** shows growth along the proximo-distal axis with the first and last time points drawn as solid lines and intermediate stages represented as dashed lines. Frequency distributions (**e**) of cell volumes from flower A during growth. Rendering at 70 h with 143 large cells marked (**f**; green). The smallest (**g**;  $500\text{--}600\ \mu\text{m}^3$ ) and the largest (**h**;  $>1,000\ \mu\text{m}^3$ ) of the marked cells in **f** are shown. Rendering in **i** shows a single stem cell organizing center cell (green) at 0 h and one immediate neighbor (blue). Daughter cells of both cells in **i** at 70 h are shown in **j**. **(k)** Schematic contrasting growth of the organizing center cell and that of its neighbor is shown in **k** (as in **d**). Scale bars, 10  $\mu\text{m}$  (**a**), and  $\sim 25\ \mu\text{m}$  (**b,c,f–j**).



whereas descendants of its immediate neighbors tended to form cell files that extended away from the organizing center cluster (Fig. 4i–k), reflecting local anisotropic cell growth that presumably contributes to the emergence of floral shape at stage 3.

By enabling examination of specific cells or groups of cells, MARS-ALT allowed us to analyze floral growth emergence. A powerful extension of this analysis will be to combine such studies with cellular identities (for instance, using cell-fate reporters) to clarify the temporal link between local growth properties and differentiation events.

## DISCUSSION

Our pipeline has some key advantages over existing methods. In contrast to published *ad hoc* methods of fusing multiple images<sup>25</sup>, MARS requires neither dedicated hardware nor precise external landmarks to guide the fusion. It works at cellular resolution, can efficiently segment internal cells deep in living tissues, can efficiently segment very large numbers of cells and can be applied to images with different markers (Supplementary Fig. 9). In contrast to methods that track nuclei in growing tissues at very small time intervals (typically one image per minute)<sup>7</sup>, the combination of MARS and ALT allowed us to describe growth deformations in the entire cell and over long developmental time periods. Thus, we can begin to describe floral morphogenesis in terms of the morphological changes in component cells. As our lineage

tracking algorithm relies on a global optimization procedure, it is very robust to local segmentation or tracking errors. Furthermore, as we designed the pipeline with no plant-specific dependencies, it should be applicable to other biological systems, with adaptation to address cell movement and/or cell death in animal tissues. Our algorithm already accommodates the loss of cells (for example, those that leave the field of view), and should thus be able to ‘accept’ cell death. The fact that the topology should not change in vegetal cells during growth is only used in ALT to select the most probable cell lineages (Supplementary Note 4). This test may be removed or weakened for use with animal tissues. It may be possible to address the issue of cell movement just by generating sufficiently clustered time point data.

One limitation of our algorithms is that the error rate for lineage tracking tends to increase with increasing extent of change (deformation and/or cell divisions) that the tissue undergoes during growth (Fig. 3f,g and Supplementary Fig. 7). Thus, during periods of rapid growth in a tissue, shortening the time interval between two acquisitions (and thus limiting the extent of change) may be required to obtain highly accurate ALT results. For 90% accurate daughter-to-mother cell lineage recognition, time intervals will have to be adjusted to maintain the s.d. of

volume expansion (which reflects nonhomogeneous tissue deformation) below 0.8 (**Supplementary Fig. 7**). A second limitation is that because our current floral analyses have been focused on the early stages, our experiments have not used time windows of more than a few days. It is possible that the biological protocols may require modification for much longer time course experiments. It will be necessary to determine the optimal time window to accurately observe certain developmental phase changes, as floral growth is not a temporally linear process. Furthermore, because the vital stain appears to have some toxic effects upon prolonged exposure, other fluorescent membrane markers (such as GFP fusions) may be needed. Finally, in the long term, scaling up the pipeline to improve throughput may present potential problems from an algorithmic point of view, such as in automatically identifying input reference points for MARS and reference lineages for ALT. These are challenging, highly combinatorial operations that will require additional development and testing to reach complete (or near-complete) automation.

Our work has many potential applications. First, MARS-ALT provides the opportunity for developmental biologists to quantitatively examine organ growth and track growth rate, anisotropy and growth direction in both surface and internal cells. Incorporation of reporter expression could enable linking quantitative growth analyses to gene activity and cellular identity. Second, our digitized floral growth template provides an opportunity to use 3D computational modeling to quantitatively test biological hypotheses. Third, we can now address a longstanding question on the precise nature of cell division rules in growing tissues and organs. Finally, our approaches may be used to examine and model organ growth in mutants with the aim of quantifying affected cellular behaviors and generating hypotheses on mechanisms that control them. This will serve to link the activities of specific genes to their morphogenetic outputs at cellular resolution.

## METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturemethods/>.

*Note: Supplementary information is available on the Nature Methods website.*

## ACKNOWLEDGMENTS

We thank A. Berger for help with establishing plant growth and imaging protocols; A. Lacroix for help with plant growth; and C. Lionnet for help with imaging. This work was funded by grants from the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (to J.-L.V.), the Region Languedoc-Roussillon (to R.F.) and from the Institut National de la Recherche en Informatique et en Automatique (to C.G.). P.D. was funded by a European Union Marie Curie Incoming International Fellowship grant (IIF-022002). The work of J.T. and C.G. was also funded by grants from the Agence Nationale de Recherche (Virtual Carpel, GeneShape and Flower model) and the European Union (Morphex).

## AUTHOR CONTRIBUTIONS

P.D., J.T., J.-L.V., G.M. and C.G. conceived the experiments; P.D. carried out the *Arabidopsis* experiments; J.-L.V. carried out the rice experiments; R.F., G.M. and C.G. conceived the software pipeline; R.F. wrote the software; R.F., E.M. and V.M. carried out the software experiments; E.M. and R.F. wrote the MARS-ALT documentation; P.D. and V.M. analyzed the output; P.D., G.M. and C.G. wrote the paper with input from the other authors.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturemethods/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Coen, E., Rolland-Lagan, A.G., Matthews, M., Bangham, J.A. & Prusinkiewicz, P. The genetics of geometry. *Proc. Natl. Acad. Sci. USA* **101**, 4728–4735 (2004).
2. Smith, R. *et al.* A plausible model of phyllotaxis. *Proc. Natl. Acad. Sci. USA* **103**, 1301–1306 (2006).
3. Jonsson, H., Heisler, M.G., Shapiro, B.E., Meyerowitz, E.M. & Mjolsness, E. An auxin-driven polarized transport model for phyllotaxis. *Proc. Natl. Acad. Sci. USA* **103**, 1633–1638 (2006).
4. Jonsson, H. *et al.* Modeling the organization of the WUSCHEL expression domain in the shoot apical meristem. *Bioinformatics* **21** (Suppl. 1), i232–i240 (2005).
5. Barbier de Reuille, P. *et al.* Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **103**, 1627–1632 (2006).
6. Hamant, O. *et al.* Developmental patterning by mechanical signals in *Arabidopsis*. *Science* **322**, 1650–1655 (2008).
7. Bao, Z. *et al.* Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **103**, 2707–2712 (2006).
8. Verveer, P.J. *et al.* High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy. *Nat. Methods* **4**, 311–313 (2007).
9. Jaqaman, K. *et al.* Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**, 695–702 (2008).
10. Murray, J.I. *et al.* Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat. Methods* **5**, 703–709 (2008).
11. Reddy, G.V., Heisler, M.G., Ehrhardt, D.W. & Meyerowitz, E.M. Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana*. *Development* **131**, 4225–4237 (2004).
12. Verdeil, J., Alemanno, L., Niemenak, N. & Tranbarger, T. Pluripotent versus totipotent plant stem cells: dependence versus autonomy? *Trends Plant Sci.* **12**, 245–252 (2007).
13. Boot, M.J. *et al.* *In vitro* whole-organ imaging: 4D quantification of growing mouse limb buds. *Nat. Methods* **5**, 609–612 (2008).
14. Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J. & Stelzer, E.H. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* **305**, 1007–1009 (2004).
15. Long, F., Peng, H., Liu, X., Kim, S.K. & Myers, E. A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Methods* **6**, 667–672 (2009).
16. Moreno, N., Bougourd, S., Haseloff, J. & Fejjo, J. in *Handbook of Biological Confocal Microscopy*, 3rd edn. (ed., J.B. Pawley) (Springer Science+Business Media, New York, 2006).
17. Kwiatkowska, D. Surface growth at the reproductive shoot apex of *Arabidopsis thaliana* pin-formed 1 and wild type. *J. Exp. Bot.* **55**, 1021–1032 (2004).
18. Grandjean, O. *et al.* *In vivo* analysis of cell division, cell growth, and differentiation at the shoot apical meristem in *Arabidopsis*. *Plant Cell* **16**, 74–87 (2004).
19. Barbier de Reuille, P., Bohn-Courseau, I., Godin, C. & Traas, J. A protocol to analyse cellular dynamics during plant development. *Plant J.* **44**, 1045–1053 (2005).
20. Ourselin, S., Roche, A., Prima, S. & Ayache, N. Block matching: a general framework to improve robustness of rigid registration of medical images. in *Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery* (eds., DiGioia, A.M. and Delp, S.) 557–566 (Springer, 2000).
21. Commowick, O. & Malandain, G. From statistical atlases to personalized models. in *Proceedings of the SA2PM Workshop* (2006).
22. Soille, P. *Morphological Image Analysis: Principles and Applications* 89–125 (Springer-Verlag, Berlin, 1999).
23. Rebouillat, J. *et al.* Molecular genetics of rice root development. *Rice* **2**, 15–34 (2009).
24. Steeves, S. & Sussex, I.M. *Patterns in Plant Development* (Cambridge University Press, 1989).
25. Preibisch, S., Saalfeld, S., Rohlfing, T. & Tomancak, P. Bead-based mosaicing of single plane illumination microscopy images using geometric local descriptor matching. *Proc. SPIE* **72592S**, 1–10 (2009).

## ONLINE METHODS

**Plant growth and imaging.** *Arabidopsis* plants were grown as described previously<sup>26</sup>. Approximately 2 cm of the upper part of the plant, including the shoot apical meristem and the flowers, were placed in growth medium. All flower buds older than stage 3 were then dissected away to facilitate observations of early stage flowers. During time course experiments, plants were maintained in growth chambers between acquisitions, and, if necessary, older flower buds were removed. In general, only one flower from any given inflorescence meristem was tracked over time, to limit damage from handling. Just before imaging, inflorescences were treated with about 2  $\mu\text{l}$  (of a 330  $\mu\text{g ml}^{-1}$  stock) of the water-soluble lipophilic dye FM 4-64 (Invitrogen), which labels cell membranes. The treated inflorescences were then immersed in water and observed on a Zeiss LSM 510 confocal microscope with a 63 $\times$  water immersion achroplan lens with a 2 mm working distance. Images were acquired at a resolution of 0.2  $\mu\text{m}$  along  $x$  and  $y$  axes and 1  $\mu\text{m}$  along the  $z$  axis. The meristems were manually rotated between angle scans, a process that took 2–5 min, which equates to a total acquisition time of 45–60 min per flower at each time point. Each set of multiangle scans took 45–60 min to acquire. In experiments in which the flowers were imaged multiple times, the inflorescence meristems were returned to a culture chamber after each observation and relabeled with FM4-64 before subsequent imaging.

All flowers presented here were either in the Columbia (Col) or the Wassilewskija (Ws) ecotypes. ‘Flower 3’ and ‘flower 2’ were observed at high resolution (1,024  $\times$  1,024 pixels) over two time points. ‘Flower A’ and ‘flower B’ were imaged at ‘optimal’ resolution (using the option and specifications provided by Zeiss LSM Browser software) at four and three time points, respectively. Every flower was observed from three angles.

Rice (*Oryza sativa* L., variety Nippon bare) crown roots were collected from young seedlings (three developed leaves) and mounted (coverslip grade 0) in a PBS buffer solution (pH 7). Imaging was carried out on an inverted Zeiss 510 META NLO multiphoton microscope in the nondescanned mode using a chameleon ultra 140 fs pulsed Ti-sapphire laser (Coherent). For cell-wall autofluorescence measurement, the laser was scanned from 690 to 770 nm range, with a peak performance at  $\sim$ 710 nm (which is roughly equivalent to 355 nm in single-photon excitation with a continuous wavelength laser system). Images were collected with a c-apochromat  $\times$ 40, 1.2 numerical aperture (NA) (Korr uv-vis-ir) water-immersion objective.

**Scalability of the MARS pipeline.** We applied our MARS protocol to rice roots using a multiphoton microscope. We first acquired images of the root apex from multiple lateral views (Supplementary Fig. 5a); the first image stack was defined as the ‘reference view’ (X), one other view (X’) was diametrically opposite to this and two others were at either end of the orthogonal diameter (Y and Y’). We segmented either the reference view alone (which we called ‘one-view’) or the 3D reconstructions generated by fusing the following combinations of image stacks: X plus X’ (two-opp-view); X plus Y or Y’ (two-ort-view); X plus X’ plus Y or Y’ (three-view) or all four (four-view). The results were evaluated on a sample of about 800 manually segmented cells chosen in either the superficial cortical layer or the central meristematic zone (Supplementary Note 5 and Supplementary Video 3).

**Lineage tracking as an optimization problem.** After each registration step of the ALT algorithm, we used the obtained common reference system to detect potential cell lineages. For every cell  $i$  of a given image in the sequence (source image), we look for daughter candidates  $j_1, j_2, \dots, j_k$  in the next image in the sequence (target image). Daughter candidates of  $i$  were identified based on their spatial proximity to  $i$  in the common reference system. Here these were defined as the cells  $j$  whose center of mass are at a distance less than a threshold distance,  $d_{\text{max}}$ , from the surface of  $i$  (the distance of a cell  $j$  to the surface of  $i$  is defined as the shortest distance between the center of  $j$  and any point of the surface of  $i$ ). This defined a set of lineage hypotheses at this step (note that the cells belonging to the set of already determined high-confidence lineages are not taken into consideration here).

To solve the lineage problem, we then formalized the search for a set of consistent lineages between the source and target images as a combinatorial optimization problem. Here we define a valid mapping between the two images as a list of pairs of cells  $(i, j)$ ,  $i$  in the source image  $I$  and  $j$  in the target image  $J$  such that  $j$  is one of the candidate cells for cell  $i$ , and such that  $j$  appears only in one pair of the mapping (a cell of the target image cannot be a descendant of several cells in the source image). If a parent cell does not appear in a valid mapping  $M$ , it has no descendant in the target image by  $M$ . The set of such cells and the set of daughter cells that have no parent in the source image are denoted

$$\bar{I}_M \text{ and } \bar{J}_M,$$

respectively. We then attach a cost  $\gamma_{ij}$  corresponding to the normalized distance between  $i$  and  $j$ . We also define constant costs,  $\gamma_I$  and  $\gamma_J$  that a cell  $i$  has no daughters, and a cell  $j$  has no parent, respectively (Supplementary Note 4). Then, the cost of a valid mapping  $M$  between the source and the target image can be defined as the sum of the local cost of the pairs of cells in  $M$ :

$$\Gamma(M) = \sum_M \gamma_{ij} + \sum_{\bar{I}_M} \gamma_I + \sum_{\bar{J}_M} \gamma_J$$

The search for a valid lineage between the two images can then be formulated as an optimization problem: among all the possible valid mappings, we look for a single  $M^*$  with minimal cost, that is, that globally minimizes the distance between the mapped cells:

$$M^* = \underset{M \text{ valid}}{\text{argmin}} \Gamma(M)$$

To solve this optimization problem, we modeled it as a problem of flow (Fig. 2j and Supplementary Note 3), which was shown to be tractable in polynomial time<sup>27,28</sup>.

**Assessing the effect of tissue deformation between consecutive acquisitions in ALT.** The ALT pipeline accuracy is critically related to the amount of change in the tissue (deformation and/or cell divisions), rather than to the time intervals between acquisitions. We tested this using serial time points where the deformation was limited (less than twofold), such as for ‘plant A’ at  $T_0$ ,  $T_1$  and  $T_2$ , and performed a direct tracking of cell lineages between  $T_0$  (considered as the source image) and  $T_2$  (as the target image), as if the image at  $T_1$  did not exist. We found that the tracking results were markedly better between  $T_0$ – $T_1$  (100%) and  $T_1$ – $T_2$  (94%), than between  $T_0$ – $T_2$  (87%), hence showing that keeping time intervals short in order to limit deformation increases ALT

accuracy. In addition, this also decreases the required manual annotation (for example,  $T_0-T_1$  is 100% accurate regardless of whether 4 or 22 reference lineages are provided). However, systematically decreasing time intervals from 24 h to 12 h may not be necessary as different growth transitions require different time intervals (**Fig. 3** and **Supplementary Fig. 7**).

**Software.** The MARS-ALT software pipeline is available as **Supplementary Software** and at <http://openalea.gforge.inria.fr/dokuwiki/doku.php?id=packages:packages/>. Software details

are described in **Supplementary Note 5**. Time estimates for the different phases of the MARS-ALT pipeline are given in **Supplementary Table 1**.

26. Das, P. *et al.* Floral stem cell termination involves the direct regulation of *AGAMOUS* by *PERIANTHIA*. *Development* **136**, 1605–1611 (2009).
27. Edmonds, J. & Karp, R.M. Theoretical improvements in algorithmic efficiency for network flow problems. *J. Association Computing Machinery* **19**, 248–264 (1972).
28. Tarjan, R. *Data structures and Network Algorithms*. (Society for Industrial and Applied Mathematics, 1983).

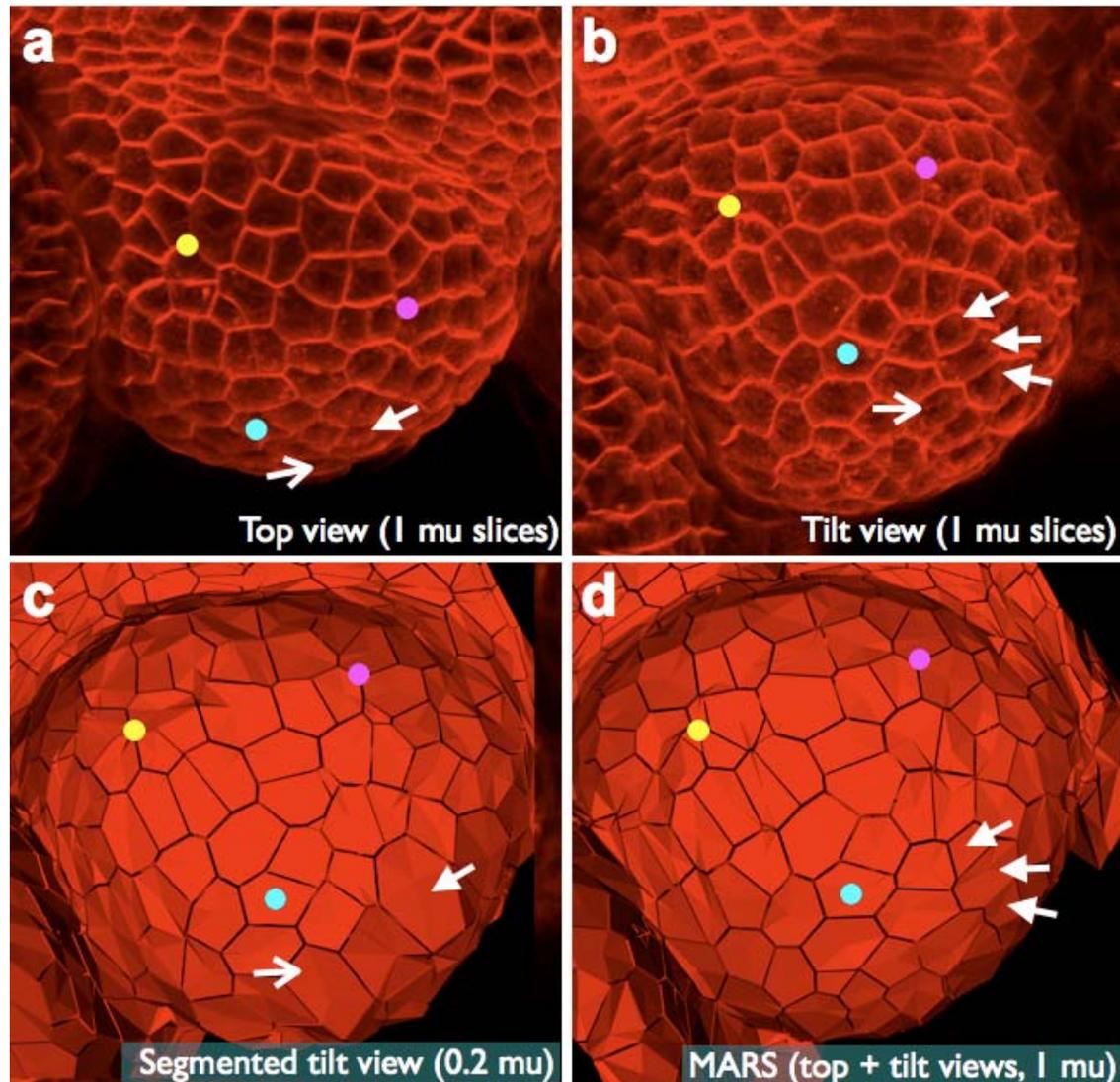
# Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution

Romain Fernandez, Pradeep Das, Vincent Mirabet, Eric Moscardi, Jan Traas, Jean-Luc Verdeil, Grégoire Malandain & Christophe Godin

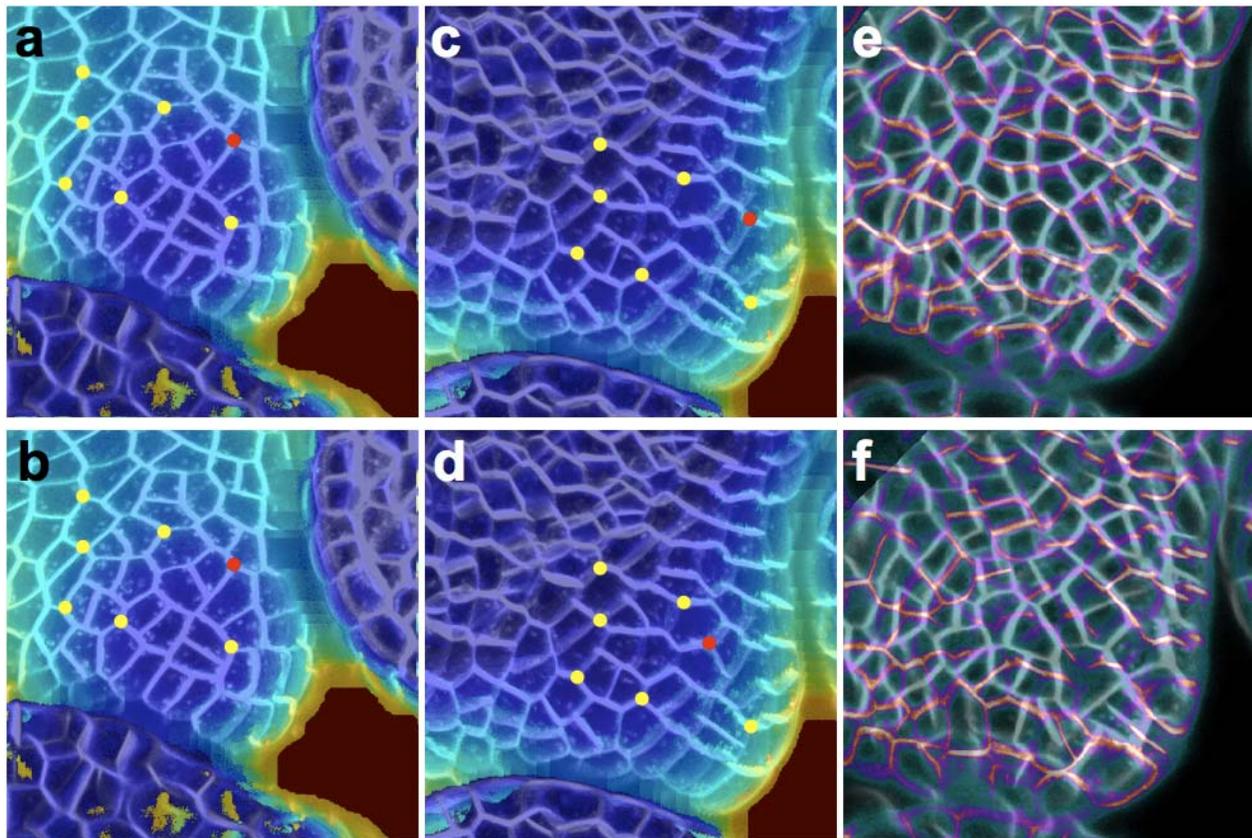
Supplementary figures and text:

<b>Supplementary Figure 1</b>	Oversampling in the Z-axis does not improve data quality
<b>Supplementary Figure 2</b>	Incorrectly positioned fusion landmarks generate obvious errors
<b>Supplementary Figure 3</b>	Effect of multiangle reconstructions on image quality
<b>Supplementary Figure 4</b>	Scalability: rice root segmentation with MARS
<b>Supplementary Figure 5</b>	Scalability of MARS assessed on rice roots
<b>Supplementary Figure 6</b>	Time course experiments to track flower development
<b>Supplementary Figure 7</b>	Quality of the cell lineage automatic tracking
<b>Supplementary Figure 8</b>	Characterizing the floral stem cell zone
<b>Supplementary Figure 9</b>	Automatic segmentation of cell fate identity reporters
<b>Supplementary Table 1</b>	Time estimate for the MARS-ALT pipeline
<b>Supplementary Note 1</b>	Reconstruction
<b>Supplementary Note 2</b>	Segmentation
<b>Supplementary Note 3</b>	Scalability of the MARS pipeline
<b>Supplementary Note 4</b>	ALT
<b>Supplementary Note 5</b>	Software tools
<b>Supplementary Note 6</b>	Detection of segmentation errors
<b>Supplementary Note 7</b>	Generation of the movie showing continuous growth

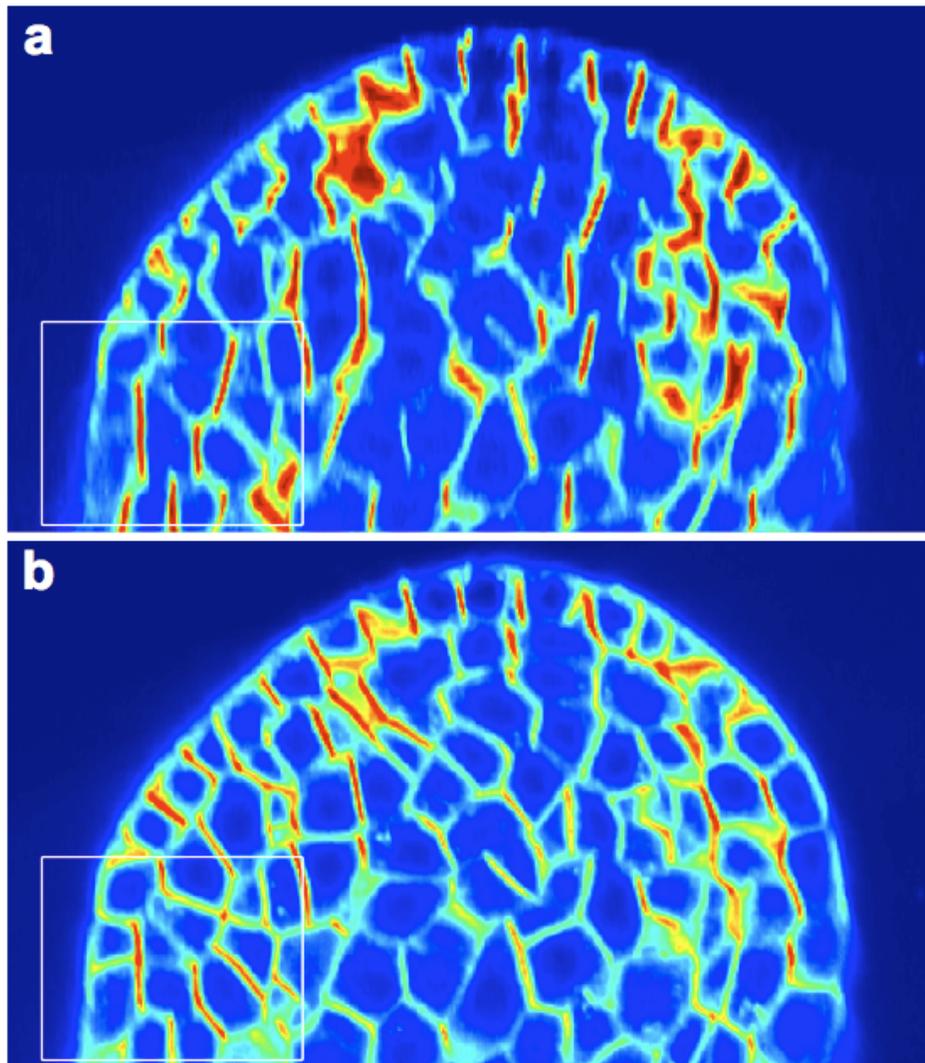
*Note: Supplementary Videos 1–3 and Supplementary Software are available on the Nature Methods website.*

**Supplementary Figure 1**

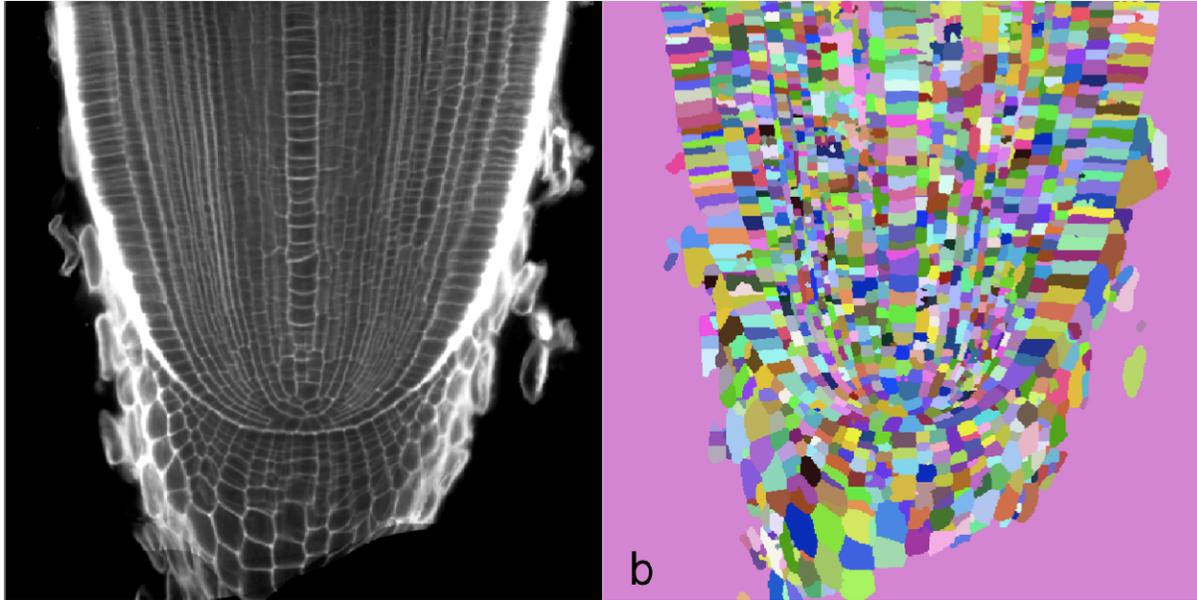
**Supplementary Figure 1 | Oversampling in the Z-axis does not improve data quality.** A single early-mid stage two flower was imaged in different ways. (a-b) Surfacic projections of the top (a) or tilted (b) views acquired every 1 μm along the Z axis. (c) Segmentation rendering of a top view acquired every 0.3 μm and rotated to match the tilt view in (b). (d) Rendering of a MARS-processed segmentation where the top and tilt views (a-b) were fused. Colored dots mark identical positions, while arrows and arrowheads indicate examples of cells that were not correctly identified with a higher Z resolution.

**Supplementary Figure 2**

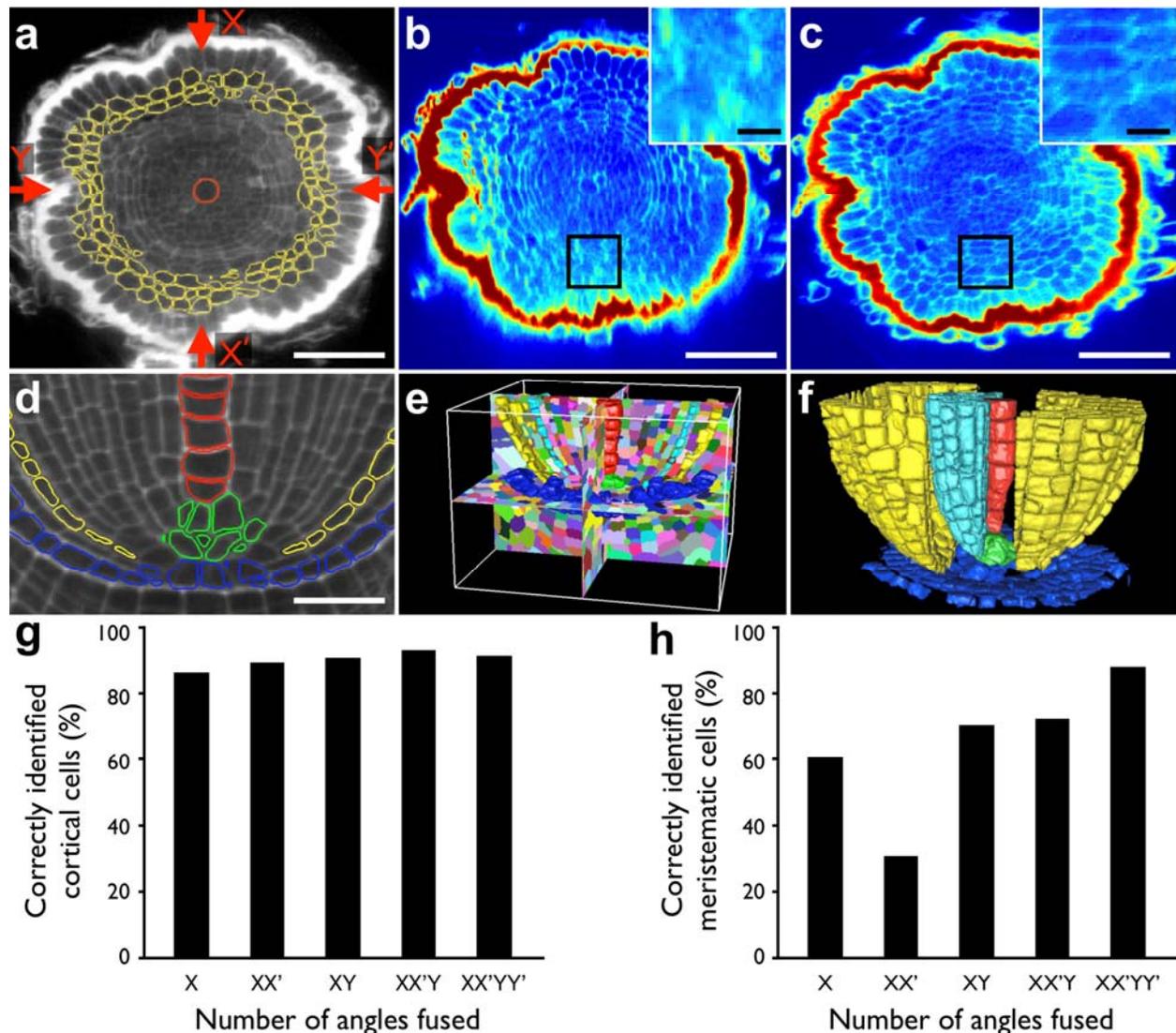
**Supplementary Figure 2 | Incorrectly positioned fusion landmarks generate obvious errors.** MARS-processing of a stage 1 flower using a top view stack (**a**, **b**) and one tilt view stack (**c**, **d**). Yellow dots indicate the landmarks provided using a custom-designed interface, while the red dots indicate landmarks that are either correctly (**a-c**) or incorrectly (**b-d**) identified. (**e-f**) Results of the rigid transformation step, with cell outlines from the top view shown in white and those from the tilt stack in purple. When all landmarks provided are correct, the outlines of the same cells from the top and tilt views being in close proximity (**e**), whereas if even one landmark is incorrectly positioned by just one cell width, the overlaps of the cell outlines are significantly perturbed (**f**).

**Supplementary Figure 3**

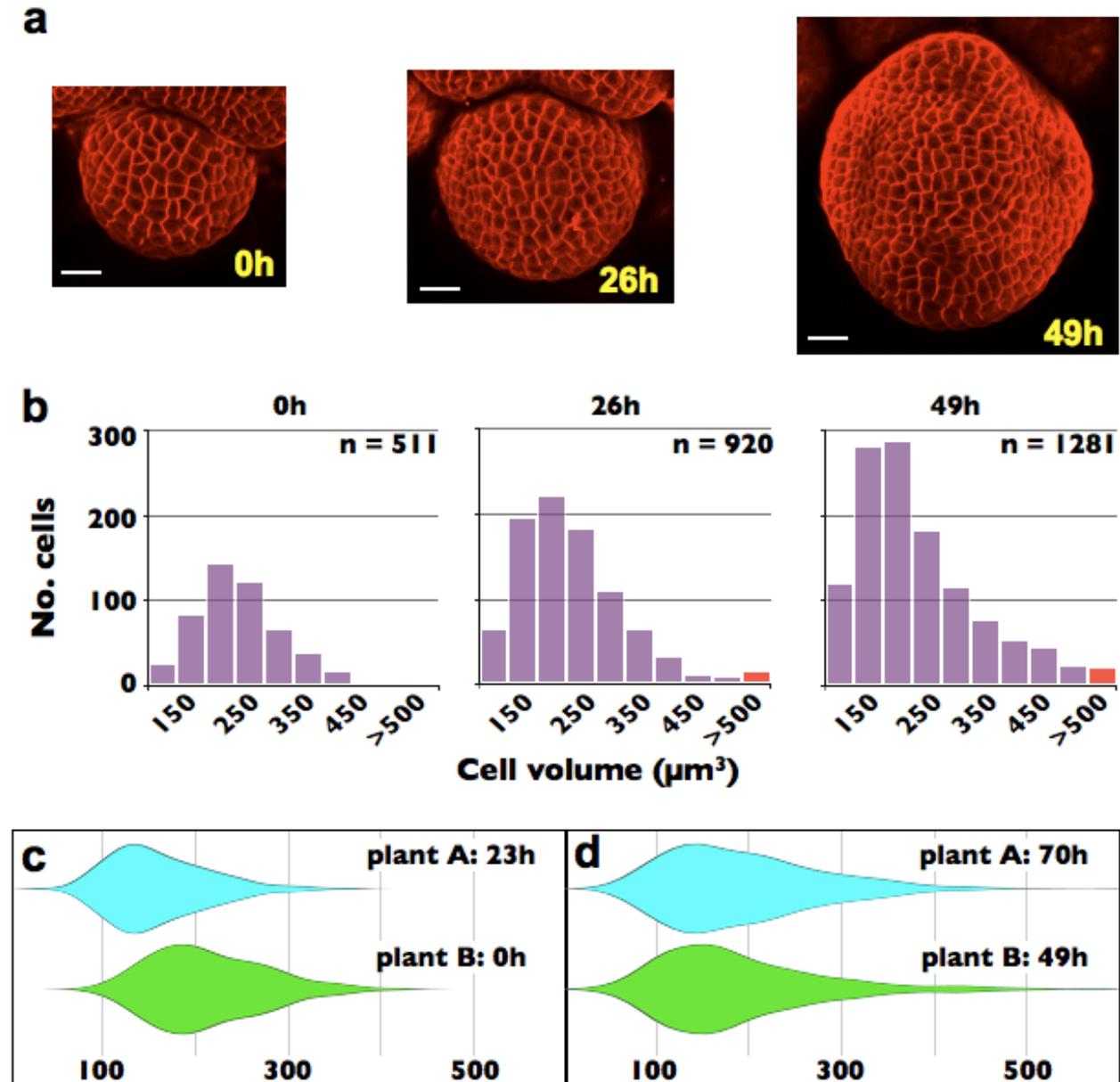
**Supplementary Figure 3 | Effect of multi-angle reconstructions on image quality.** An optical longitudinal section through a flower with (a) only the 'reference view' or (b) after the fusion of three view stacks. The boxed area is shown in Figure 1 (d-d') in the main text.

**Supplementary Figure 4**

**Supplementary Figure 4 | Scalability: rice root segmentation with MARS.** (a) Transversal image of the root observed with a multi-photon microscope (b) MARS segmented image showing the automatically identified individual cells. A total of 16,400 cells were segmented.

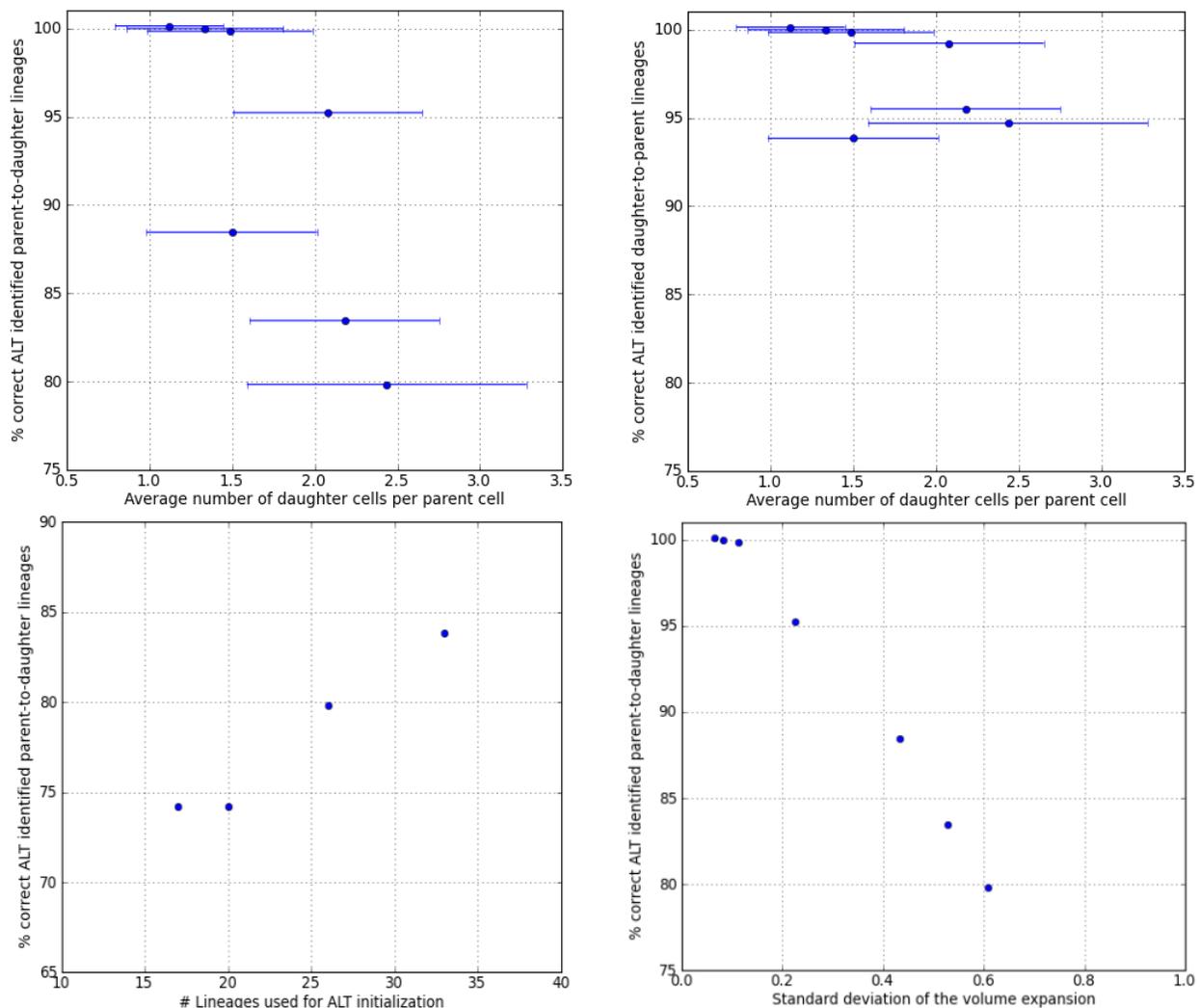
**Supplementary Figure 5**

**Supplementary Figure 5 | Scalability of MARS assessed on rice roots.** (a-d) Transversal views of a coronary rice root. (a) The different view angles are indicated with arrows (X is the reference view, X' is diametrically opposite X, Y is orthogonal to X and Y' opposite to Y). Colored regions indicate the various groups of cells that were manually segmented: the cortical cells (in yellow) and the central meristematic cells, including the metaxylem (red) and the quiescent centre (green). (b-c) Prior to the fusion of the various image stacks, many areas of the tissue have ill-defined cell outlines (b), whereas after fusion, many of these are clearly restored (c). Insets show magnified views of the boxed areas. (e-f) The segmented reconstructions may be easily rendered according to volume (e) or to cell type (f). (g-h) Percentage of correctly segmented cells as a function of the number of the fused view stacks in either the cortex (g) or the central meristematic region (h). The scale bars indicate 70  $\mu\text{m}$  in a-c, 25  $\mu\text{m}$  in the inset boxes in b-c, and 50  $\mu\text{m}$  in d.

**Supplementary Figure 6**

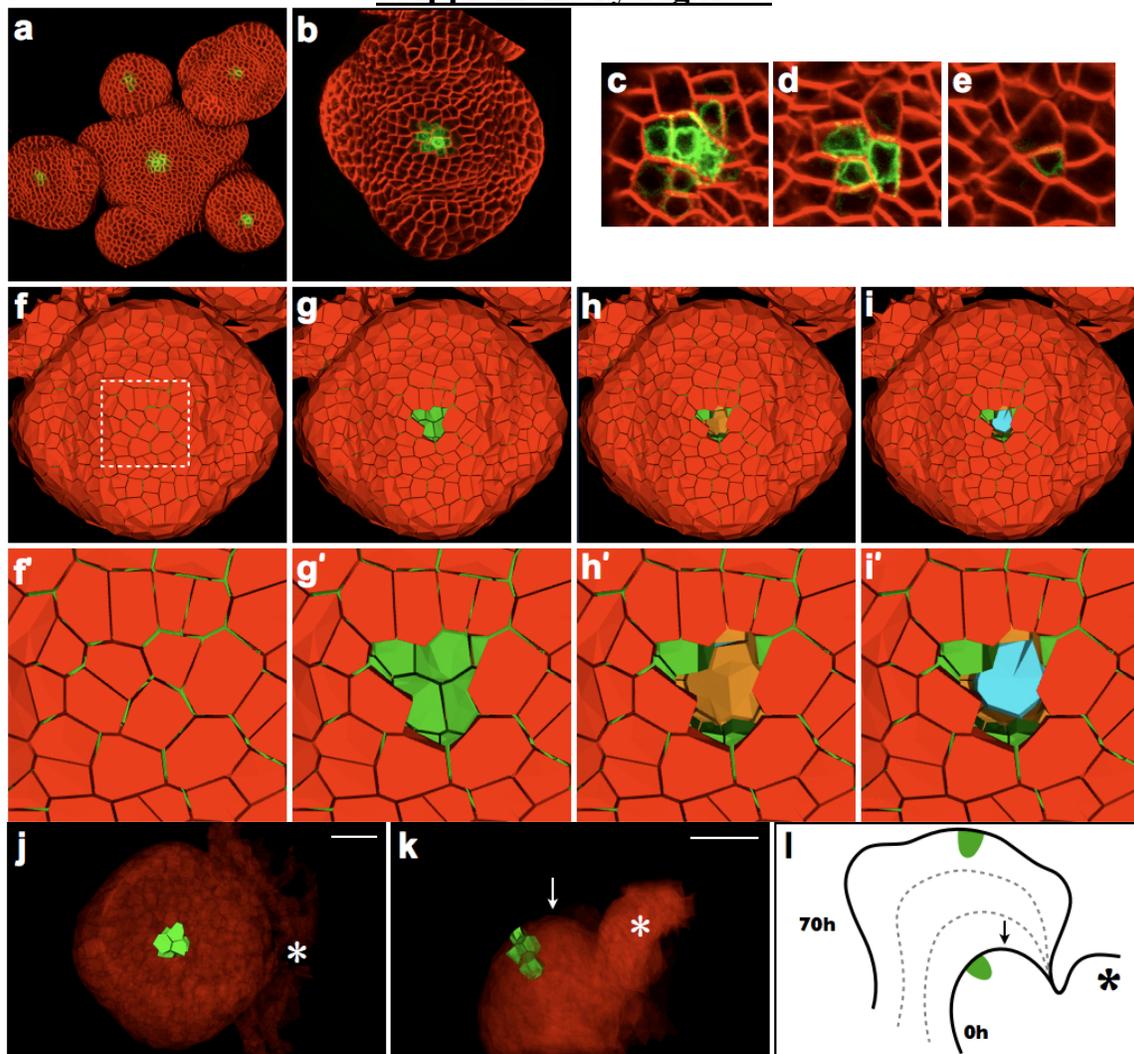
**Supplementary Figure 6 | Time-course experiments to track flower development.** (a) Surface projections of a time course experiment (called 'flower-B'; 'flower-A' is presented in Fig. 2 in the main text). Time intervals between acquisitions are indicated in hours. (b) Frequency distributions of cell volumes during growth (the distributions for flower-A can be seen in Fig. 5 in the main text). (c-d) Violin plots directly comparing volume distribution data from flower-A (blue) and flower-B (green) shows the dissimilarities at stage 2 and similarities at stage 3 (both are statistically significant values as tested using the Welch *t*-test).

### Supplementary Figure 7

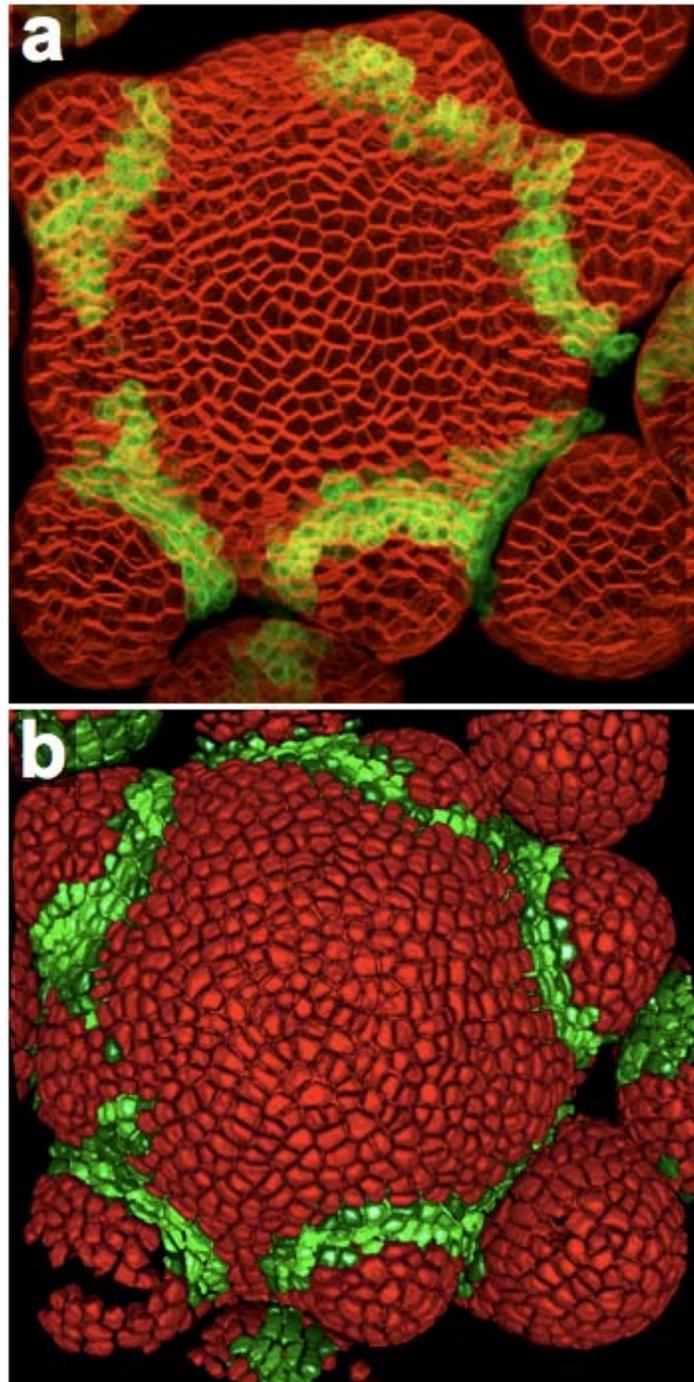


**Supplementary Figure 7 | Quality of the cell lineage automatic tracking.** (a) Percentage of parent cells with perfect ALT-identified daughter cell lineages as a function of the average number of cell divisions per parent cell. Each point corresponds to unique pairs of serial time points that cover different flowers and developmental stages. Horizontal bars show the standard deviation of the cell division number per parent cell. All tracking has been initialized with 26 expert lineages (b) Percentage of daughter cells with correct ALT-identified parent cells as a function of the average number of cell divisions per parent cell. (c) Percentage of parent cells with perfect ALT-identified daughter cell lineages as a function of the number of manually identified lineages used for ALT initialization in the case of the T2-T3 transition in flower-A. (d) Percentage of parent cells with perfect ALT-identified daughter cell lineages for varying volume expansions and sorted by increasing standard deviation of cellular volume expansion. This shows the strong correlation between the accuracy of the output of the ALT algorithm and the extent of non-homogeneity in the volume deformation of the cells.

## Supplementary Figure 8



**Supplementary Figure 8 | Characterizing the floral stem cell zone.** (a-b) Expression patterns of the well-studied stem cell marker line *pCLAVATA3::GFP* in (a) an inflorescence meristem and (b) an early stage 3 flower where sepal outgrowth has just commenced. Cells were counter-stained with the vital stain FM 4-64 to visualize cell membranes. (c-e) Magnified view of the *CLV*-expressing zone. Serial optical sections through the epidermal L1 layer (c), the sub-epidermal L2 layer (d) and the L3 layer just below (e) showing high expression in a total of seven cells. (f-i') Rendering of a stage 3 flower at the last time point from the flower-A time course experiment (f-i), and a magnified view of its center (f'-i') upon which a similar *CLV3* expression zone was modeled. The putative *CLV3* cells were marked and removed from the L1 (g, g'), L2 (h, h') and L3 (i, i'). (j-l) All data and images are similar to those shown in **Figure 4** of the main text. (j) A group of seven cells at the exact center of the flower at 70 h, in the putative expression domain of the stem cell identity reporter. (k) Their six lineage-tracked ancestors at 0 h are located abaxially with respect to the vertical summit of the flower (arrow). (l) The schematic shows the approximate positions of the stem cells (green) with respect to the vertical summit at 0 h and 70 h.

**Supplementary Figure 9**

**Supplementary Figure 9 | Automatic segmentation of cell fate identity reporters.** (a) An inflorescence meristem expressing a boundary marker gene (*pBNDRY*>>*GFP*). GFP expression is induced using the ethanol-inducible system, and observed under the confocal along with the vital stain FM4-64. (b) This image is then directly fed into the segmentation protocol and the output is rendered according to the amount of FP expressed in the cell.

**Supplementary Table 1**

<b>Supplementary Table 1. Time estimate for the MARS-ALT pipeline<sup>a</sup></b>		
Task	Time required (minutes)	
	Sample A (small)	Sample B (large)
Segmentation (MARS) <sup>b</sup>		
1. Imaging (per angle)	15	20
2. Manual positioning of anchor points (registration)	10	15
3. Automated reconstruction (rigid + dense registration)	60	90
4. Segmentation	15	15
5. Manual segmentation verification <sup>c</sup>	10-30	30-180
Lineage Tracking (ALT)		
6. Identifying lineage references	30-45/transition <sup>d</sup>	
7. Automatic lineage tracking	60-150/transition	

<sup>a</sup> The various tasks were performed on ordinary desktop computers with slightly different configurations. Tasks 2-5: Four 2.27GHz dual-core processors running Debian Linux 5.0.4, 17GB RAM and an NVIDIA Quadro FX 1800 Graphics card with 768MB RAM; Tasks 7: One 2.4GHz dual-core processor running Ubuntu Linux 9.4, 2GB RAM and an NVIDIA Quadro NVS 140M Graphics card with 512MB RAM.

<sup>b</sup> Examples are provided of reconstructions of a small stage 1 flower (Sample A) or of a larger stage 3 flower (Sample B). Both involve image stacks acquired from three angles at a resolution of 512 x 512 pixels. Sample A: 65.81  $\mu\text{m}$  in the X and Y axes; 33-37 Z slices at 1  $\mu\text{m}$  intervals; 940 cells total, of which about 225 belong to the sampled flower. Sample B: 153.56  $\mu\text{m}$  in the X and Y axes; 94-105 Z slices at 1  $\mu\text{m}$  intervals; about 3050 cells total, of which about 2000 belong to the sampled flower.

<sup>c</sup> The time estimate for this task depends both on the developmental stage of the flower (i.e. number of cells) & on the data resolution (i.e. image quality). Thus samples with fewer cells and/or better image quality may be validated faster than those with more cells or noisier images.

<sup>d</sup> The length of this task depends on the extent of growth and volumetric increase between serial time points.

## **Supplementary Note 1: Reconstruction**

This step aims at reconstructing an isotropic 3D image (i.e. with voxels of the same dimension along the 3 axis) by fusing image stacks acquired at different angles. Before fusion, these image stacks required being co-registered, and this is achieved in a hierarchical way. The complete pipeline consists then in the following steps:

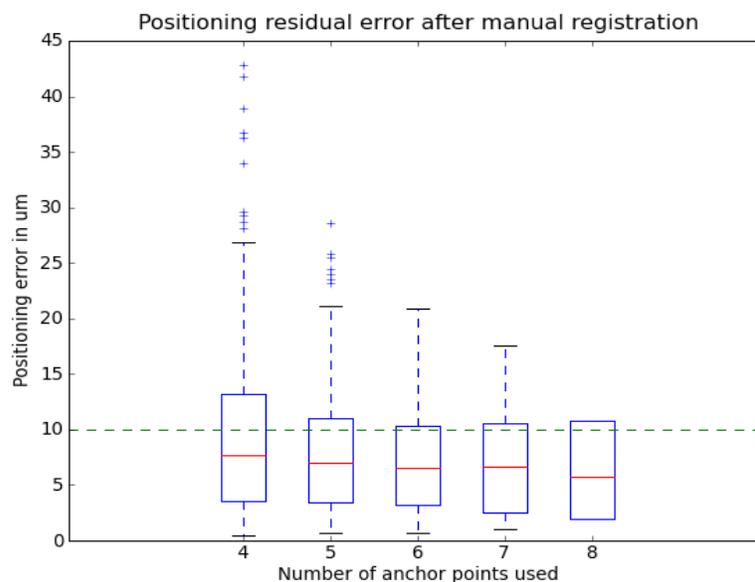
1. user-assisted linear co-registration
2. automated linear co-registration, and
3. automated non-linear co-registration and fusion

### **1.1. User-assisted linear co-registration**

Automated registration may fail to retrieve large transformations (for instance, rotation with an angle greater than 45 degrees). Therefore, we used a first transformation computed with user assistance as initialization for further automated registration.

For each couple of images to be registered, the user identified correspondences by pairing a few anchor points (note that this is the only manual operation of the MARS pipeline).

To ease this operation for the user, we reconstructed the surface of the meristem from each image stack<sup>1</sup>. Based on a projected view of this surface, anchor points were manually positioned on each 2D image using a dedicated user-interface and the 3D coordinates of the points were inferred from it. If the points are too close to each other on the meristem surface, the registration error can be high. We thus chose points homogeneously distributed over the whole surface and that appear in all the views. A reference image is then chosen (usually the top-view image of the meristem) and an initial rigid transformation is computed for all other images by minimizing the mean squared error between the anchor points in these images and those of the reference image<sup>2</sup>.



**Fig. A. Registration error (in  $\mu\text{m}$ ) between the manual rigid transformation and the optimal rigid transformation as a function of number of anchor points. The horizontal green dashed line indicates the average diameter of a cell ( $10 \mu\text{m}$ ). Using eight anchor points (or more) make it unlikely to have a final positioning error larger than one cell size.**

Since anchor points have to be entered manually by the user, we tested options to minimize the effort involved. For this, we studied how the initial rigid transformation was altered by decreasing numbers of anchor points. We considered the rigid transformation obtained at the end of the automated rigid registration (see below) as a reference and compared it to initial rigid transformations obtained from varying the number of anchor points. The error between two rigid transformations was defined as the distance between the transformed image domain centers obtained by these two rigid transformations respectively. Results are presented in Fig. A. To keep the error smaller than the radius of a cell, we found that 8 points were a satisfactory compromise between manual input and positioning accuracy.

### 1.2. Automated linear co-registration

An automated linear registration method is used to refine the previously computed transformation. We chose a block-matching based method<sup>3</sup>. The principle of this method is to pair sub-images (or blocks) between both images to be registered and then to compute a transformation that explains the obtained pairings. Pairing is achieved by maximizing the correlation coefficient between blocks of the two images, while the rigid transformation that best superimposes them is computed by minimizing the squared distances between the paired blocks' centers. This is iterated until convergence, and embedded into a multi-scale strategy.

#### *Parameters:*

In our experiments, the size of blocks is set to 4 x 4 x 4 voxels, block spacing is set to 3 voxels along each dimension, and only the 75% blocks of highest variance are kept (to avoid considering blocks of homogeneous intensity). Block similarity is assessed with the correlation coefficient. The transformation estimation is conducted with a least trimmed squares scheme, with a cut value of 0.75. The multi-scale pyramid starts with images of size 32 x 32 x 32 and ends at full resolution.

### 1.3. Automated non-linear co-registration and fusion

After linear co-registration, differences may still exist between image stacks that are due to slight deformations (**Fig. 1h** in the main text). To compensate for them, we used a non-linear registration method, initialized by the previously computed rigid transformations, that enables to compute a deformation field between images. The fusion of all images was done by averaging the voxel intensities after resampling them within a common geometry. To cancel the effect of voxel geometric anisotropy, the reference image was itself resampled at a higher resolution with smaller cubic voxels, while the other (floating) images were resampled with the combination of all computed transformations in the same geometry. However, the choice of a particular image stack as reference may bias the reconstruction result, since this image may be deformed too. Then, to reduce the influence of the reference image choice, the non-linear registration is embedded into an average image computation scheme<sup>4</sup>. As a non-linear registration method, we chose an extension to deformation field of the block-matching<sup>5</sup>.

#### *Parameters:*

In our experiments, the size of blocks is set to 7 x 7 x 7 voxels, and only the 90% blocks of highest variance are kept (to avoid considering blocks of homogeneous intensity). Block similarity is assessed with the correlation coefficient. The multi-scale pyramid comprises four levels.

## **Supplementary Note 2: Segmentation**

In the literature, cell segmentation has been achieved using a wide range of image processing methods, from deformable models<sup>6</sup> and high-sensitivity contour detection<sup>7</sup>, to performance-based thresholding<sup>8</sup>. In some cases, watershed algorithms are used on two- or three-dimensional images<sup>1, 9</sup>. Here, we have used the latter technique, which is well-suited to our images, in which the objects to segment are dark areas (cell interiors) surrounded by bright and thin boundaries (cell walls). To apply the watershed algorithm, we first extract seeds that will serve as sources for the catchment basins. We then run the watershed method, and subsequently correct automatically obvious over-segmentation errors.

### **2.1. Seed extraction**

Generally, sources for watershed methods are the local minima of the input image. However, this is known that such a strategy leads to over-segmentation. We first denoise the input image to enhance the signal/noise ratio and eliminate high frequency noise while preserving the main structural properties of the image. It appeared that the denoising method has to be adapted to the type of processed images. Then the main local minima are extracted.

- For the floral buds images, acquired with a single photon laser scanning microscope, the Alternate Sequential Filter (ASF) is used as denoising method<sup>10</sup>. This is a succession of morphological opening and closing operations with structuring elements of increasing size. For structuring elements, we chose discretized Euclidean sphere of radii ranging from 1 to  $N$ ,  $N$  being equal to 4 for the high resolution images (e.g. f2), and to 2 or 3 for the normal resolution images (f3 and p60).
- For the root meristem images, acquired with a multiphoton microscope, a Gaussian filtering (i.e. a convolution with a 3D Gaussian function) is preferred. The typical value of the standard deviation is 0.5  $\mu\text{m}$ .

The seeds are then extracted by computing the  $h$ -minima from the denoised image. The parameter  $h$  allows controlling the pertinence of extracted minima: two neighboring basins will be merged if they are separated by a “mountain” whose minimal height (with respect to the higher basin) is less than  $h$ . The parameter  $h$  is set for each series of image, whereas it varies between 3 and 5 (in intensity units) for the floral buds, and between 4 and 6 for the roots. These seeds serve as markers for the foreground. Eventually, an additional marker is added for the background. To that end, we threshold the image and keep points whose intensity value is less than the threshold (a typical value is 20). The largest connected component is the background marker.

### **2.2. Watershed transformation and over-segmentation correction**

The principle of the watershed transformation is to consider the image to be segmented as an elevation map, and to flood water from different sources (the markers) to extract the catchment basins that will be separated from the watershed ridges<sup>11</sup>. The extracted cells are then subsequently analyzed. Cell with too small volumes (typically less than 40  $\mu\text{m}^3$ ) are obvious errors. Corresponding markers are removed from the seeds, and the watershed transformation is computed again until convergence.

### **Supplementary Note 3: Scalability of the MARS pipeline**

To assess whether our method is generic (applicable to tissues other than inflorescences) and scalable (applicable to much larger tissues), we applied the MARS protocol to rice roots, which are both thicker (200-450  $\mu\text{m}$ ) and have a more complex cellular organization (with up to 10 cortical cell layers, **Supplementary Fig. 4a**), as compared with *Arabidopsis* roots (that are 80-100  $\mu\text{m}$  thick and have a single cortical layer).

We applied our multi-angle image acquisition protocol using a multi-photon microscope. We first acquired images of the root apex from multiple lateral views (**Supplementary Fig. 5a**); the first image stack is defined as the 'reference view' (**X**), one other view (**X'**) is diametrically opposite to this and two others are at either end of the orthogonal diameter (**Y** and **Y'**). We then segmented either the reference view alone (which we call '1-view'), or the 3-D reconstructions generated by fusing the following combinations of image stacks: X plus X' (2-Opp-view); X plus Y or Y' (2-Ort-view); X plus X' plus Y or Y' (3-view) or all four (4-view). The results were evaluated on a sample of about 800 manually segmented cells chosen in either the superficial cortical layer or the central meristematic zone (**Supplementary Fig. 5a,d**; **Supplementary Video 3** online). Whereas 86.5% of all tested cortical cells are accurately segmented in the 1-view, in the 2-view (2-Opp-view and 2-Ort-view combined) and the 3-view, this increases to 91.1% and 93.4%, respectively (**Supplementary Fig. 5g**). However, the fusion of all four views did not further improve the results.

The advantages of using multi-angle reconstruction are even more striking for the most inaccessible cells of the root, the central meristematic cells, located about 170  $\mu\text{m}$  under the root surface (**Supplementary Fig. 5h**). Of these cells, only 60.8% are accurately segmented in the 1-view, whereas in the 4-view, this number is 88.2% (**Supplementary Fig. 5h**). Not surprisingly, we found that for these cells, the two types of 2-angle reconstructions are not equivalent. In the 2-Ort-view, 70.6% of cells are correctly segmented, as compared to only 30.8% in the 2-Opp-view (**Supplementary Fig. 5h**) for which the wall anisotropy problem is not corrected. In this latter case, the walls that are clearly visible (parallel to the focal axis) are the same in both images and similarly, walls that are not clearly visible (perpendicular to the focal axis) are also the same both images. This means that two views in opposite directions amplify the discrepancy between clearly visible walls and those that are not, leading to a poorer segmentation result compared to segmentations from a single image. However, as soon as a third perpendicular angle is used (**XX'Y**), the effect is counterbalanced positively. In contrast to our results for the cortical cells, the fusion of all 4 views was essential to achieve good segmentation rates in the central meristematic zone. It is probable that the walls of these cells have different properties from those of the cortical layers, perhaps making them harder to visualize.

In summary, the MARS algorithm was able to automatically segment 16,400 cells in the root tip with an estimated 88% (for central meristematic cells) to 93% (for cortical cells) of cells correctly segmented. These results demonstrate the scalability of MARS (see **Supplementary Fig. 4-5**).

## **Supplementary Note 4: Automatic Lineage Tracking (ALT)**

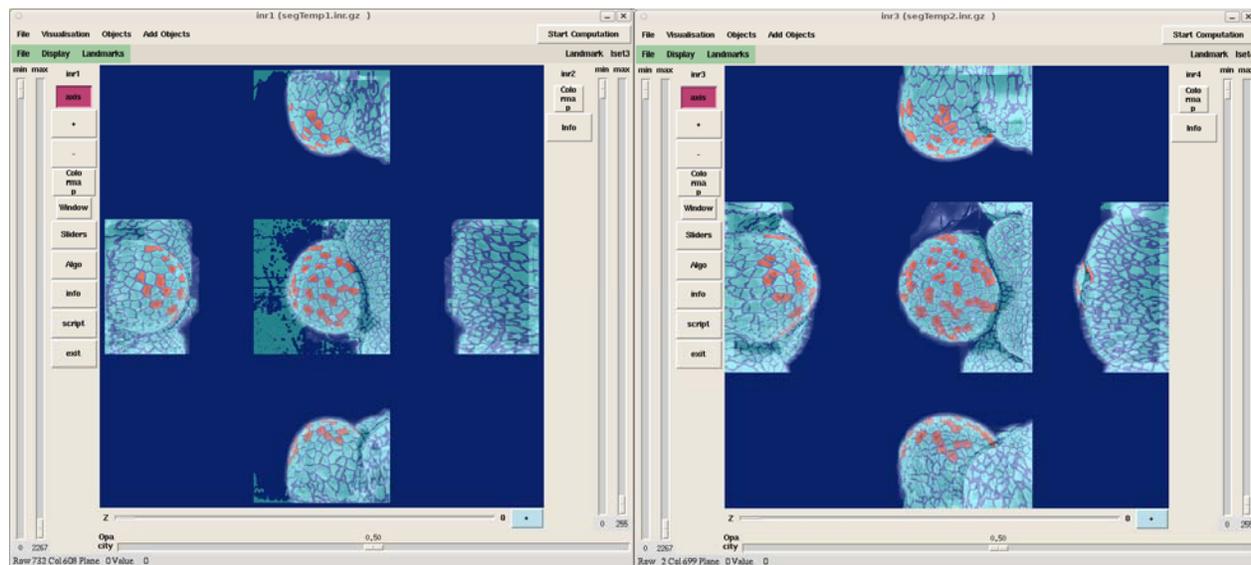
The computation of cell lineage from the target image  $I$  and the source image  $J$  consists of the following steps:

1. Initialization of the algorithm (initialization of the set of high confidence lineages)
2. Computation of a deformation field based on a set of high confidence lineages,  $L$ .
3. Refining of the deformation field with an automated non-linear registration algorithm.
4. Building of lineage hypotheses and solving the lineage problem.
5. Update of the set of high confidence lineages,  $L$

As explained in the main text, steps 1-3 are performed once, while steps 4 and 5 are called iteratively until stability is reached.

### **3.1. Initialization of the algorithm**

The algorithm requires defining an initial set of cell lineages to guide the computation of the deformation. The required number of initial cell lineages ultimately depends on the amount of deformation between two consecutive time frames and on the desired cell lineage detection accuracy. This part of the pipeline is extremely difficult to automate due to its highly combinatorial nature (initially, almost any cell of the first image can be paired with any cell of the second image, etc.) and to the critical impact this initial pairing on the quality of the algorithm results. In our approach, we chose to optimize the overall pipeline robustness by selecting manually a number of lineages on the surface of the meristem between each time frame (Fig. B). We made test benches to estimate the minimal number of lineages required for particular deformation amplitudes. Results are shown in **Supplementary Fig. 7**.



**Fig. B.** Expert definition of cell lineages on the meristem surface (L1) to initialize the ALT algorithm, illustrated on flower-B between time frames  $T_0$  and  $T_0+26h$ . The user clicks on a parent cell in the first image (dark bleu cell in the left-hand side image) and identifies a series of daughter cells in the second image (dark bleu cells in the right-hand side image).

### 3.2. Computation of a deformation field based on a set of high confidence lineages, $L$

A first rigid transformation is estimated between the two images by minimizing the total square distance between each parent cell in the set of high confidence lineages and its descendants. After this rigid registration, the residual positioning error of the high confidence lineages is used to define non-linear transformation as a dense vector field (*i.e.* a vector per voxel). For this, each high confidence lineage is transformed into one vector that links the center of mass of the parent cell to the center of mass of all its descendants, and the dense vector field is computed by interpolating between these vectors.

More precisely, a lineage  $l_n$  links a parent cell at time  $t_0$  to its daughter cell at time  $t_1$  (if there are several daughter cells, *i.e.* if the parent cell has experienced some division(s), we group all of them into one (virtual) single cell). The vector from the center of mass  $m_n$  of the parent cell to the center of mass  $d_n$  of the daughter cell,  $\overline{m_n d_n}$ , indicates the local displacement from time  $t_0$  to time  $t_1$ . Given a set of vectors  $\overline{m_n d_n}$ , a dense displacement can be interpolated. We choose here a Gaussian weighting but other interpolation schemes can be used as well. At voxel  $v_i$  of frame at time  $t_0$ , the displacement vector is then computed by

$$\mathbf{v}(v_i) = \frac{\sum_n G(\|v_i - m_n\|) \overline{m_n d_n}}{\sum_n G(\|v_i - m_n\|)}$$

where  $G(\cdot)$  denotes the Gaussian function. Note that due to the weighting coefficients, this interpolation scheme makes it possible to estimate vectors corresponding to voxels outside the convex hull defined by the known parent cells.

### 3.3. Refining of the deformation field with an automated non-linear registration algorithm

Then, based on this initial residual field, an automated non-linear registration algorithm is used to refine this initial transformation by matching the voxel intensities of the two images. This refined registration, computed as described in section 1.3 of the supplementary data, makes it possible to overlap the two images, and compare their segmentations.

### 3.4. Building of lineage hypothesis and solving the lineage problem

*Building the set of lineage hypotheses.* In this common reference system, we hypothesize potential daughter candidates for each parent cell, based on their spatial distance. A cell  $j$  from the target image  $J$  (resampled after non-linear registration) is a potential daughter cell of a cell  $i$  of the source image  $I$  (*i.e.* is a likely lineage hypothesis) if the distance  $d(i,j)$  between cells  $i$  and  $j$  is below a specified threshold  $d_{\max}$ . This distance is 0 if the center of mass  $c_j$  of  $j$  is inside cell  $i$ , else a 3D line is drawn between the centers of mass of cells  $i$  and  $j$ , and the  $d(i,j)$  is defined as the distance between  $c_j$  and the first point of cell  $i$  encountered along this line.

*Building a bipartite graph representing the set of hypotheses.* From the identification of potential lineage candidates, we construct a graph containing two types of nodes. Nodes labeled  $i$  correspond to cells of image  $I$  and nodes  $j$  correspond to cells of image  $J$ . An edge is created between a pair of nodes  $(i,j)$  if  $j$  is identified as a potential daughter candidate of  $i$ . To express the possibility that cells in either of the images may not be mapped, two special nodes are added to the graph:  $\theta_i$  and  $\theta_j$  representing “void” nodes, *i.e.* nodes that are used to map cells that are not

part of a valid mapping in both images.  $\theta_I$  is connected by an edge to every node of  $J$  and likewise  $\theta_J$  is connected to every node of  $I$ .

*Defining a flow on the graph.* On each edge  $(i,j)$  of the graph, a flow of 1 unit means that the two nodes are mapped. The flow is null otherwise. On edges, flows can only take integer values and are bounded by minimum and maximal values, called capacities. Here the maximal capacity  $C_{i,j}$  for each edge is 1 while the minimal capacity  $c_{i,j}$  is 0, except for edges involving the “void” nodes  $\theta_I$  and  $\theta_J$  that have an infinite maximal capacity and a null minimal capacity.

To sum up the flow going through all the nodes of an image, two other nodes are added,  $s$  and  $t$ , called source and target nodes. The source node is connected to every node of  $I$  and to  $\theta_I$  while each node of  $J$  and node  $\theta_J$  are connected to the target node. The minimal capacities  $c_{s,j}$  and  $c_{i,t}$  of all these edges is 1 (all the cells must be mapped either on another cell or on a void node). To allow only valid flows to circulate in the network different types of capacities are defined on these edges. The maximal capacity of the edges originating from the source node,  $C_{s,j}$  corresponds to the maximal number of divisions,  $NDiv$ , that a cell may have between two consecutive time-frames ( $NDiv$  is set to 8, i.e. a cell has a maximum of 8 daughter cells). The maximal capacity of the edges arriving on the target node,  $C_{i,t}$ , is 1 since a cell from the target image may at most have one parent cell.

To express the fact that, up to cell division, the number of cells in both images should be the same, a conservative law must be respected at each node of the network expressing that the total flow entering a node must be equal to the total flow leaving this node. Finally, an edge is added from the target to the source to express that the total input flow entering the system equals the output flow.

*Cost of a flow.* The cost  $\gamma_{ij}$  of the edge  $(i,j)$  is defined by  $d(i,j)/d_{\max}$  and ranges therefore from 0 to 1. To penalize links to void cells  $\theta_I$  and  $\theta_J$ , the costs of edges  $(i,\theta_I)$  and  $(\theta_J,j)$  are set to 2 so that it is always preferred to map a cell  $i$  to  $j$  rather than to remove both  $i$  and  $j$ . Other edges (to and/or from source and target vertices  $s$  and  $t$ ) have a cost of 0.

A possible lineage mapping between cells of the two images is represented as a set of binary values attached to each edge: 1 for a lineage relationship and 0 otherwise. The resulting list of binary values for all edges defines a flow between the two images. Not all flows correspond to possible lineage mappings. A flow is considered as being ‘valid’ if every cell in the source image is mapped at least once and if every cell in the target image is mapped to at most one parent cell in the source image. A cost is associated with each valid flow  $F$  by summing up all the elementary costs  $\gamma_{ij}$  of mapping cells  $(i,j)$  that are defined by edges with flow value  $f_{ij} = 1$  in  $F$ :

$$\Gamma(F) = \sum_F \gamma_{ij} f_{ij} + \sum_{I_F} \gamma_i f_i + \sum_{J_F} \gamma_j f_j$$

The resolution of equation (2) thus amounts to finding a valid flow  $F^*$  in the above network with minimum cost  $\Gamma(F^*)$ .

### 3.5. Update of the set of high confidence lineages, $L$

From the optimal mapping computed above, we extend the set of high confidence lineages. A couple (parent cell, descendant cells)  $= (p, \{d_1, \dots, d_p\})$  is added to the set if the following conditions are verified.

- The parent cell  $p$  is adjacent to a parent cell already in  $L$  (lineages are added by propagation).
- The descendant cells of a parent cell adjacent to  $p \in L$  must be adjacent to the descendants  $\{d_1, \dots, d_p\}$  of  $p$  (adjacencies are preserved).
- The descendant cells must be adjacent together, i.e. they form one single connected component (if they are split in several components, it is probably an erroneous lineage).
- The volume of all descendant cells should be larger than the one of the parent cell. To account for potential errors, we check that this total volume is at least 90 % of the one of the parent cell.

### **Supplementary Note 5: Software tools**

We used a collection of online binary executables, each of them being dedicated to a specific task, that are gathered using a script language to yield high-level procedures. In addition, a 3D image viewer allows visual inspection of images/results and to select anchor points for the user-guided registration procedures. These components are developed on a linux platform, and are available online. The different high-level tasks are listed below.

- a) Preprocessing
  - Conversion of input data (usually a series of 2D image file) into an in-house 3D image format, namely inrimage.
- b) Registration and fusion
  - Enhancement of the meristematic surface.
  - Selection of anchor points onto the surface thanks to a 3D image viewer.
  - Pairing of points for each couple of images, and computation of initial rigid transformations.
  - Iterative fusion (this involves all the needed registration procedures).
- c) Segmentation
  - Automated cell segmentation (a visual analysis of the results, thanks to the 3D image viewer, helps tuning the segmentation parameters).
  - User-guided segmentation correction: the user indicates over-segmented cells (one of the underlying watershed seed will then be suppressed) and under-segmented ones (and add a watershed seed at the correct location).
- d) Segmentation analysis
  - Global statistics computation (number of cells, cell volumes, etc).
  - User-guided selection of cells of interest for analysis.
  - 3D rendering of cell (possibly with superimposition of the 3D image) for visualization purposes thanks to the 3D image viewer.
- e) Lineage computation.
  - Selection of anchor points thanks to a 3D image viewer.
  - Pairing of points between the two images, and computation of the initial non-linear transformations.
  - Iterative lineage computation and non-linear registration
- f) Lineage validation.
  - Graphical analysis of the segmented 4-D time series was carried out using custom built interface based on the Visualisation Toolkit ([www.vtk.org/](http://www.vtk.org/)) and the mayavi1.5 ivtk librarie (<http://mayavi.sourceforge.net/>). The software imports segmented images into delaunay tessellations and allows the user to manipulate the lineages in 3-D.

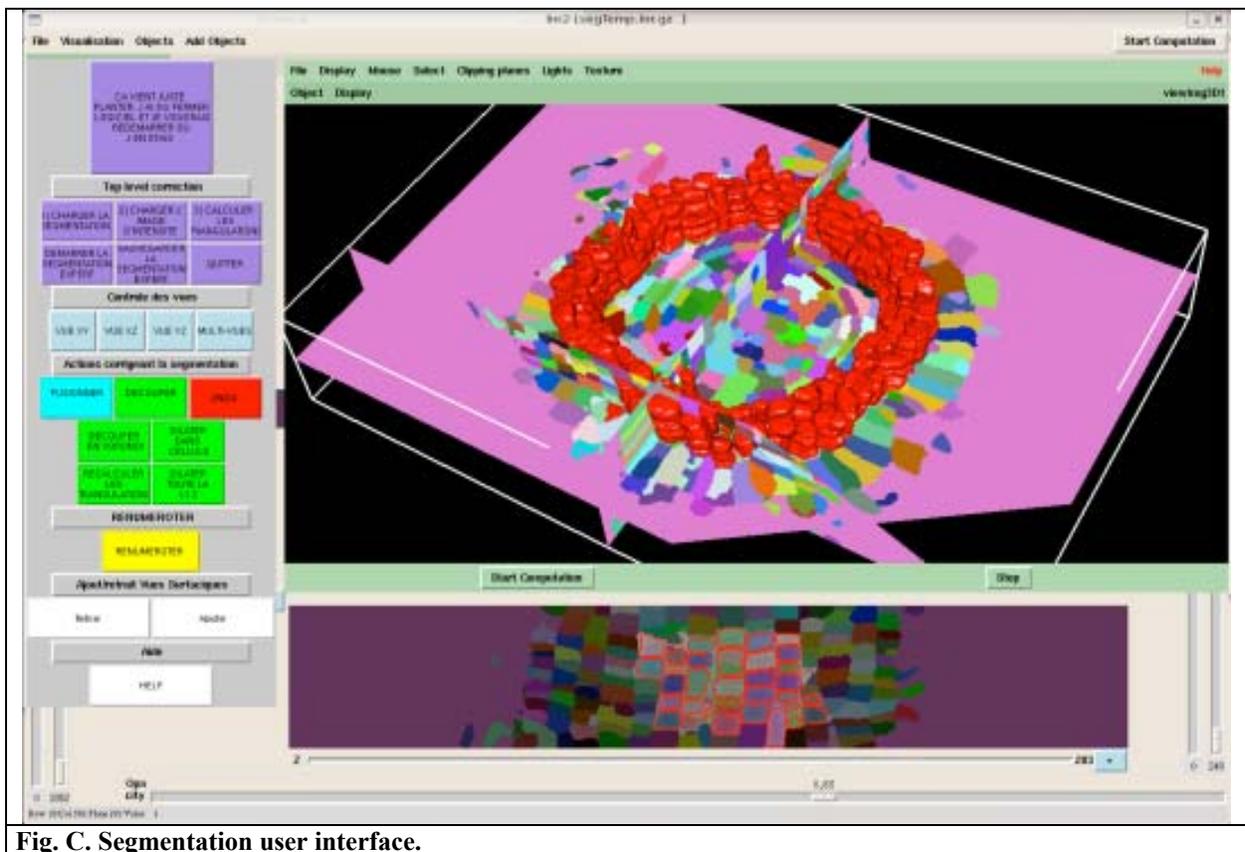


Fig. C. Segmentation user interface.

### **Supplementary Note 6: Detection of segmentation errors**

The segmentation errors are detected by comparing the output of the automated segmentation MARS method against a ground truth, i.e. a segmentation validated by experts. A cell is considered as being well segmented by the automated method if it can be associated without ambiguity to a unique cell of the ground truth segmentation. These associations can be viewed as lineages from the automated segmentation to the ground truth, with the following specificities:

- there is no geometric distortion between the two images, then no requirement of a (non-)linear registration,
- in case of oversegmentation, two “parent” cells can be linked to a single “daughter” cell.

Therefore, we reused the ALT algorithm, with the exception that the test and the reference segmentation did not required to be co-aligned. This also results in a few changes in the graph construction that are detailed below.

A cell  $j$  from the reference image  $J$  can be linked to a cell  $i$  of the test image  $I$  if the cells intersect. The cost  $\gamma_{ij}$  of the edge  $(i,j)$  is defined by the Dice’s coefficient between the two cells. It is defined by  $2\text{vol}(i \cap j) / (\text{vol}(i) + \text{vol}(j))$  and ranges therefore from 0 (no intersection) to 1 (cells are identical). To penalize links to void cells  $\theta_I$  and  $\theta_J$ , the costs of edges  $(i,\theta_I)$  and  $(\theta_J,j)$  are set to 2. Other edges (to and/or from source and target vertices  $s$  and  $t$ ) have a cost of 0. The minimal and maximal capacities of edges  $(i,j)$ ,  $(i,\theta_I)$  and  $(\theta_J,j)$  are respectively set to 0 and 1, this allows the cells to be linked or not. The minimal and maximal capacities of edges  $(s,i)$  and  $(j,t)$  are all set to 1, while those of edge  $(s,\theta_I)$  are respectively set to 0 and  $\max(N_J, N_I)$ .

The maximal feasible flow is computed using Ford-Fulkerson algorithm. Then, the cells from the test image that are linked to the void cell  $\theta_J$  are considered again and linked with the reference cell of smallest cost.

The well-identified cells are those for which there exists a one-to-one mapping. The oversegmentation is assessed by the reference cells that are linked to more one test cell. The undersegmentation is assessed by the test cells that are linked to void cell  $\theta_I$ .

### Supplementary Note 7: Generation of the movie showing continuous growth

After MARS-ALT methods, the pairs of successive time point acquisitions have been non-linearly co-registered. Each of these transformations can be decomposed into a rigid transformation and residual deformations. By combining the rigid transformations, all volumes and subsequently all MARS segmentations can be resampled into the same reference pose, e.g. the first volume.

Consider now two successive volumes  $I_i$  and  $I_j$  and the residual transformation  $T_{ij}$  that goes from  $I_i$  towards  $I_j$  (and then allows to warp  $I_j$  onto  $I_i$ ). To produce a movie, we have to generate intermediary images between  $I_i$  and  $I_j$ . Consider now that the time goes from 0 ( $I_i$ ) to 1 ( $I_j$ ), we generate warped images of both  $I_i$  and  $I_j$  at an intermediary time step  $t$ , and then blend them.

The transformation  $T_{ij}$  can be rewritten as  $T_{ij} = Id + U_{ij}$  where  $U_{ij}$  is a vector field depicting the deformation. Therefore, if we assume the linearity of the deformation with respect to time, the transformation that goes from time 0 to time  $t$  is  $Id + t \cdot U_{ij}$  and we denote it by  $t \cdot T_{ij}$ . To resample  $I_i$  at time  $t$ , we compute its inverse, i.e.  $(t \cdot T_{ij})^{-1}$ . Similarly, the transformation  $-(t-1) \cdot T_{ij}$  allows to resample  $I_j$  at time  $t$ .

Using the above transformation, the segmented cells of both  $I_i$  and  $I_j$  can be warped at the intermediary time  $t$ . Surface rendering are computed for both, resulting in two 2D views  $S_{i,t}$  and  $S_{j,t}$ . The virtual intermediary image  $S_t$  of the movie is computed by blending the 2D views, i.e.  $S_t = (1-t)S_{i,t} + tS_{j,t}$ . Note that, with such a procedure, all cell divisions seem to occur simultaneously, at time  $t=0.5$ .

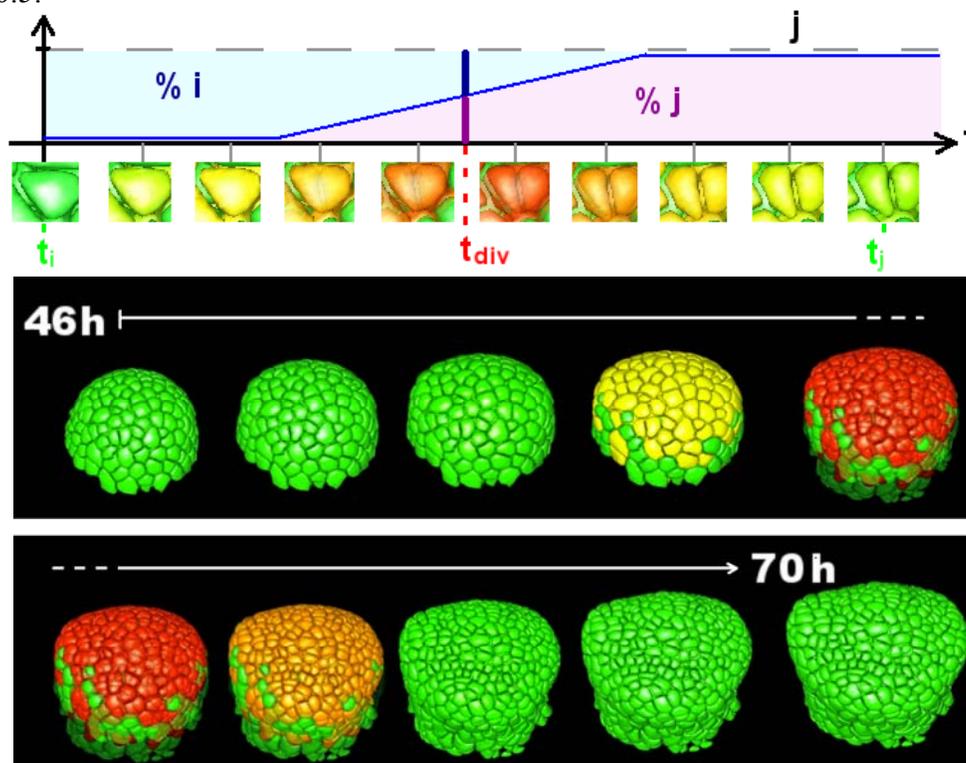


Fig. D. 3-D Morphing computation between two time frames (46h and 70h). The upper diagram shows how the two images are mixed at any time  $t$  (blue curve). In the bottom image, colors are emphasizing cells that are dividing between the two time frames.

### **Supplementary Data References**

1. Barbier de Reuille, P., Bohn-Courseau, I., Godin, C. & Traas, J. A protocol to analyse cellular dynamics during plant development. *Plant J.* 44, 1045-1053 (2005).
2. Faugeras, O. & Hebert, M. The representation, recognition, and locating of 3-D objects. *The International Journal of Robotics Research* 5, 27-49 (1986).
3. Ourselin, S., Roche, A., Prima, S. & Ayache, N. in Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI 2000). (ed. A.M.D.a.S. Delp) 557-566 (Springer, 2000).
4. Guimond, A., Meunier, J. & Thirion, J. Average Brain Models: A Convergence Study. *Computer Vision and Image Understanding* 77, 192-210 (2000).
5. Commowick, O. & Malandain, G. From Statistical Atlases to Personalized Models. In Proceedings of the SA2PM Workshop, Copenhagen, held in conjunction with MICCAI 2006; (2006).
6. Mischler, D., Romaniuk, B., Benassarou, A., Bittar, E. Robust 4D segmentation of cells in confocal images. *Machine Graphics & Vision International Journal* 15, 515-524 (2006).
7. Kasson, P., Huppa, J., Davis, M. & Brunger, A. in Bioinformatics2005).
8. Beaver, W., Kosman, D., Tedeschi, G. & Bier, E. in 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2007).
9. Moreno, N., Bougourd, S., Haseloff, J. & Feijo, J. in Handbook of Biological Confocal Microscopy (2006).
10. Serra, J. Image analysis and mathematical morphology: theoretical advances. *Book 2* (1988).
11. Vincent, L. & Soille, P. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 583-598 (1991).