Abstract

Arabidopsis thaliana is a conventional plant adopted for various studies on plants. On Arabidopsis thaliana leaf there are trichomes whose initiation and distribution are of interest to biologists. Partial differential equations models for studying pattern formations are quite complicated. Moreover, usage of these models for the leaf in the initial stage of growth is unclear. In our work, we developed a novel algebraic model, which is easier for and more acceptable to biologists, for close-to-real life prediction of initiation and distribution of trichomes on Arabidopsis thaliana. Though our work is limited to tessellating a leaf surface with squares, other shapes (i.e equilateral triangles and regular hexagons) are mentioned and numerical results are explored.
1 Introduction

Biology is becoming an important application area of mathematics, as there are growing needs of, or desires for, applied mathematical models in biology. However, in order to gain the advantages of a mathematical model, and in order to construct an efficient model, a trade-off between the level of mathematics used and the prerequisites in biological knowledge has to be considered. Though this may present an obstacle (from both 'sides'), it is apparent that a well-constructed model would be advantageous to understand and control all kinds of mechanisms and phenomena in biological sciences.

The *Arabidopsis thaliana* plant has a short life-cycle and is therefore a good specimen to use in plant gene laboratories. Because of the short time span between its initiation, reproduction phase and death, the time that lapses from gene input to (visible) output will be short and make for quicker (and often cheaper!) experiments.

On the *Arabidopsis thaliana* leaf there are single-celled epidermal hairs called trichomes. A common function of trichomes is to produce various classes of secondary metabolites. When a trichome is ruptured these molecules are secreted onto the leaf in the form of a resin or oil. This gives plants their scent as well as protecting the surface of the leaf. Many trichome-secreted compounds are used commercially as flavourings and medicines. Hence there is an interest in being able to understand and, if possible, also to control the effect of genes on the qualities of *Arabidopsis thaliana* trichomes.

In this report, we present a mathematical model which, given a legitimate input, representing some alteration of a gene, yields a modeled leaf with trichomes and distribution as output. The genes responsible for initiation of trichomes are classified as activators and inhibitors. In the model, size of the tile corresponds to the amount of activator and the threshold to the amount of inhibitor. Output trichome distribution should approximate real life samples and should correspond to known data. In other words, the output should be reasonable with respect to what is known about trichome formation, and distribution.

2 Biology of trichomes in *Arabidopsis thaliana*

2.1 Introduction

According to Wikipedia, *Arabidopsis thaliana* is a small flowering plant native to Europe, Asia and northwestern Africa. Moreso, it is a spring annual with a relatively short life cycle, and it is popular as a model tool for understanding the molecular biology of many plant traits, including flower development and light sensing.

Developmental biology is the study of the process by which organisms grow and develop. Modern developmental biology studies the genetic control of cell growth (the process by which a cell irreversibly increases in size over time by accretion and biosynthesis production of matter similar to that already present), differentiation (the process by which a less specialised cell becomes a more specialised cell type), and "morphogenesis" (which is concerned with the shapes of tissues, organs and entire organisms and positions of the various specialised cell types). And since mathematics is a concise language, many scientists are daily using it as a tool towards the understanding of molecular biology.
2.2 Trichomes in *Arabidopsis thaliana*

According to Wikipedia, trichomes are fine out-growths or appendages on plants and certain protists. These are of diverse structure and function. Examples are hairs, glandular hairs, scales, and papillae.

In *Arabidopsis thaliana*, trichomes are single and very large cells that are regularly spaced on the leaves. Trichome initiation is an early event in leaf organogenesis taking place while neighbour epidermal cells continue dividing.

2.3 Distribution of the *Arabidopsis thaliana* root epidermal cells

The Arabidopsis root epidermis is composed of two cell types. Root hair or trichoblast cells develop over underlying anticlinal walls of the cortex cells, while the remaining epidermal cells develop as atrichoblast or non-hair cells. According to Haseloff (2007), current models for root hair differentiation suggest that the fate of an epidermal cell is determined by whether or not it expresses a genetic pathway that inhibits hair cell differentiation. So, if a cell does not express this inhibitory pathway then it develops as a hair (H) cell; if it does express this pathway, then it develops as a non-hair (N) cell.

Moreso, the expression of this pathway appears to be partly controlled by the relative abundance of two MYB-type transcription factors, WEREWOLF (WER) and CAPRICE (CPC). And each of these is proposed to interact with the same bHlH (basic helix-loop-helix) transcription factor and the TRANSPARENT TESTA GLABRA (TTG) protein, with the WER protein able to generate an active transcription complex and the CPC protein generating an inactive complex.

2.4 Trichome Distribution

According to Haseloff (2007), the plant epidermis is a multifunctional tissue playing important roles in water relations, defence and pollinator attraction. This range of function is performed by a number of different types of specialised cells which differentiate from the early undifferentiated epidermis in adaptively significant patterns and frequencies.

In the words of Haseloff (2007), recent work on the Arabidopsis trichome suggests that interactions between neighbouring cells reinforce initial differences, possibly in levels of gene expression or cell cycle stage, to commit cells to different developmental programmes. According to Haseloff (2007), Arabidopsis trichomes are large unicellular structures that develop on the surface of most shoot-derived organs. In leaves, the number, spacing and shape of trichomes is tightly regulated, and this process has been used as an experimental system to study the control of cell fate and pattern formation.

Haseloff (2007) highlighted the following steps as the general approach to understanding trichome development:

- Description of the cellular events that take place during trichome initiation and differentiation.
- Screening for Arabidopsis mutants that affect this process.
- Isolation of the gene sequences.
• Promoter fusion and in situ hybridisation experiments, to determine in which cells the genes are transcribed.
• Gene misexpression and protein interaction studies.

The following three results of Larkin et al. (1996) are also used in our work:
• The trichome distribution on Arabidopsis leaves is not random.
• Cell lineage does not play a role in trichome spacing.
• Trichomes are not produced until leaf primordia are approximately 100µm long.

3 Differential Equation Models for Pattern Formation

According to Pereverzyev and Anderseen (2008), at the macroscopic level, ordinary and partial differential equations have been successfully used to simulate observed pattern formation process and at the cellular level, rewriting, implemented using L-systems methodologies, is a popular tool. The problem with models based on ordinary and partial differential equation is that they are usually too complicated for a non-mathematician to understand. This point, with other issues, are illustrated with the following (popular) models for pattern formation.

3.1 Reaction-diffusion model

Reaction-diffusion systems describe how two processes (local chemical reactions and diffusion) influence the concentration of one or more substances distributed in space. Pattern formation is often modeled using such systems.

Reaction-diffusion systems can be represented in the general form as
\[ \partial_t \vec{q} = D \Delta \vec{q} + R(\vec{q}) , \] (1)

where each component of the vector \( \vec{q}(x, t) \) is the concentration of a substance, \( D \) is a diagonal matrix of diffusion coefficients, \( \Delta \) denotes the Laplace operator and \( R \) accounts for all local reactions.

3.1.1 Two-component reaction-diffusion equations

According to Jupp et al. (2007), Turing suggested that chemicals can react and diffuse to produce patterns. An important idea that was first proposed by Alan Turing is that a state that is stable in the local system becomes unstable in the presence of diffusion. Based on this idea, pattern formation is often modeled by a series of reaction-diffusion equations for the interacting components.

A linear stability analysis however shows that when linearising the general two-component system
\[ \begin{pmatrix} \partial_t u \\ \partial_t v \end{pmatrix} = \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix} \ast \begin{pmatrix} \partial_{xx} u \\ \partial_{xx} v \end{pmatrix} + \begin{pmatrix} F(u,v) \\ G(u,v) \end{pmatrix} , \] (2)

and perturbing the system against plane waves
\[ \vec{q}_k(x, t) = \begin{pmatrix} \tilde{u}_k(t) \\ \tilde{v}_k(t) \end{pmatrix} \ast \exp(ik \cdot x) \] (3)

4
close to a stationary homogenous solution, one finds

\[
\begin{pmatrix}
\frac{\partial \tilde{u}_k(t)}{\partial t} \\
\frac{\partial \tilde{v}_k(t)}{\partial t}
\end{pmatrix} = -k^2 \ast \begin{pmatrix}
D_u \tilde{u}_k(t) \\
D_v \tilde{v}_k(t)
\end{pmatrix} + R' \ast \begin{pmatrix}
\tilde{u}_k(t) \\
\tilde{v}_k(t)
\end{pmatrix}.
\] (4)

Turing’s idea can only be realised in four equivalent classes of systems characterised by the signs of the Jacobian R of the reaction function. This class of systems is named activator-inhibitor system after its first representative. Though this theory seems to fit, in many cases, the appearance of trichomes in Arabidopsis, the following problems are identified with this approach:

- A novice in mathematics is outrightly put off with the mathematical jargons.
- Full specification of a reaction-diffusion problem is realised with the differential equations, some initial conditions and boundary conditions. As one does not know the exact initiation state of trichomes, prescribing close-to-right initial conditions is a difficult task. Moreso, as a plant cell does not necessarily live in isolation, influence of the neighbouring cells on the boundary of a particular cell needs to be taken into consideration in formulating the boundary conditions. This is again a difficult task.
- The domain (i.e Arabidopsis leaf) grows with time and this poses further complications.

3.2 Incompressible Navier-Stokes Equation

Plant leaves growth modelling using the Navier-Stokes equations is rare. The continuum concept is difficult to apply to a plant tissue, since the relative size of the molecules to a fluid is smaller than that of the cells to a plant tissue. But according to Wang et al. (2004), it is frequently found that the tissue or organ of interest consists of rather large number of cells. Wang et al. (2004) also gave the growing region of a primary root of a Zea seedling, consisting of more than 250 000 cells, as an example.

Matter is not conserved in leaf as cell division leads to creation of matter. And since Navier-Stokes equations assume conservation of mass, this becomes an issue. As a result, modification of the continuity equation is essential to accommodate the change in mass due to growth. Wang et al. (2004) gave modified Navier-Stokes equations as

\[
u_t + (u \cdot \nabla)u + Lu = -\nabla p + \frac{1}{3 \ast Re} \nabla L + \frac{1}{Re} \Delta u ,
\] (5)

\[
\nabla \cdot u = L(x,t)
\] (6)

where \( u \) is the fluid velocity vector, \( p \) is the pressure, and \( Re \) is the Reynolds number which represent the viscosity of the fluid. For the plant growth, in particular, Wang et al. (2004) claimed that \( u \) is identified with the growth velocity, \( p \) is the pressure exerted by the growing cells, and \( L(x,t) \) is the relative growth rate.

Even if the boundary conditions in Wang et al.(2004) correctly describe what happens in trichomes initiation in Arabidopsis, the system of equations is far from adequately describing trichomes appearance or initiation, as it says nothing about the cell differentiation, plant morphogenesis, and other key processes involved in cell growth.

Hence these shortcomings of PDE models call for another approach for modelling pattern formation, in particular trichomes appearance in \textit{Arabidopsis thaliana}. This necessitates our work, using an algebraic approach rather than the differential equation approach.
4 Mathematical Model of Leaf Growth and its Numerical Simulation

The initial growth starts from the meristem until the petiole forms. At this stage, the only significant expansion that occurs is that occurring at the meristem boundary between the tissue being formed, that will subsequently become the cells of the leaf, and the cells of the meristem. It is this initial growth that determines the basic shape and structure within a leaf. (It appears to play a homologous role similar to that of embryo formation in other organisms.)

For this stage, the initial growth from the meristem occurs in the following manner:

(i) The basic shape and internal structure of a leaf are determined at the boundary between the cells that have fully formed as leaf cells and the cells within the meristem controlling the formation of the tissue layer that will next become the next layer of leaf cells.

(ii) The positional information of the current tissue layer growing out from the meristem is determined by the signalling and communication occurring between the cells that have fully formed and the cells within the meristem controlling the formation of the current tissue layer.

We define the growth of leaf under four stages. Each cell is defined as a square. Other possibilities to model the cell are discussed in Section 5.2.

First Stage:
The "mother cell" grows and divides to produce "daughter cells". Thereafter, the expansion on the x-axis begins only on the boundary of the initial leaf and after a while the expansion terminates.

Second Stage:
The expansion on the x-axis ends and the expansion on the y-axis continues.
Third Stage:
The shrinkage on x-axis starts, so the shape becomes narrow.

Fourth Stage:
The expansion is again only on the y-axis and this makes the final shape.

In the next section, we present our model for appearance and distribution of trichomes on the simulated leaf.
5  Algebraic Approach for Trichomes initiation based on Young’s Model

5.1  Recursive Algebraic Modelling Approach

Recursive algebraic modelling approach was introduced by J.P.W. Young in 1983 which he applied to pea leaf morphogenesis. The pea leaf is made up of three types of structure which are leaflet, tendril and rachis.

The idea of the Young’s model is to consider an initial primordium, which is a part of the plant that produces new organs. Starting from this primordium, determination of the leaf structure proceeds by the growth and sub-division. Whether a primordium becomes a leaflet, tendril or rachis element depends on its size at a critical stage. The initial primordium has size $M_0$. The primordium first grows and then splits into three new primordia, which is called *trifurcation*. Two of these primordia are considered as lateral primordia with size $S_1$, while the third one is the central primordium of size $M_1$ (see Figure 1).

![Figure 1: Young’s Model](image)

Due to a threshold condition on the size of the new primordium it either becomes a leaflet, tendril or rachis element. Primordia that are smaller than the first threshold $T_1$ become committed as leaflets, or if smaller than a second, lower threshold $T_2$, as tendrils. Morphogenesis ceases when every primordium is either committed as leaflet or tendril.

This is summarised as follows:

**Rule 1.** The size of the next primordia to form is determined, in a dynamical manner, by given algebraic formulas defined only in terms of the size of the primordia from which they have just formed.

**Rule 2.** The disjoint dynamic stages in the subsequent development of primordia just formed are determined by the following size threshold constraints:
(a) $S_1 > T_1$, (or $M_1 > T_1$): the initiation repeats with $S_1$ (or $M_1$) as the new values for $M_0$.
(b) $T_2 < S_1$, $M_1 < T_1$: leaflet forms, and the process stops.
(c) $S_1 < T_2$ or $M_1 < T_2$: tendril forms, and the process stops.

5.2 Tessellation

A tessellation or tiling of the plane is a collection of plane figures that fills the plane with no overlaps and no gaps. One may also speak of tessellations of the parts of the plane or of other surfaces. Generalizations to higher dimensions are also possible.

We will now give a mathematical reasoning that supports our simulation of leaf growth with squares as presented in the previous section.

Proposition: There are only three regular polygons with which the plane can be tessellated. These are equilateral triangles, squares and regular hexagons.

Proof. It is obvious that the plane can be covered by these three shapes. We want to show that there are no other possibilities.

Let $n$ be the number of vertices of the regular polygon and let $\alpha$ be the interior angle of the polygon. We first show that $n \leq 6$.

When considering plane tiling, it is clear that in each vertex there are at least three edges because we need that $\alpha < 180^\circ$ in order to have a polygon. From this we can conclude that $\alpha \leq 120^\circ$ because the sum of all angles in one vertex has to be $360^\circ$.

With the formula

$$\alpha = \frac{n - 2}{n} \cdot 180^\circ$$

for computing $\alpha$, we finally get

$$\frac{n - 2}{n} \cdot 180^\circ \leq 120^\circ,$$

$$\frac{n - 2}{n} \leq \frac{2}{3},$$

$$n \leq 6.$$ 

What is left to show is that $n \neq 5$, i.e. it is not possible to use pentagon. For $n = 5$ we get $\alpha = 108^\circ$. In each vertex a number of equal angles, namely of size $\alpha$, sum up to $360^\circ$ which means that $\alpha$ must divide $360^\circ$. This is not the case for $\alpha = 108^\circ$. Therefore $n$ can only be equal to 3, 4 or 6 and it is clear that with these three shapes plane tiling is possible.

5.3 Our model

We develop our model based on Young’s ideas. The leaf is covered by squares of the same size. Each square represents a tile which is composed of a group of cells. There is a meristem level from which tiles appear in levels. (see Section 4.)
In accordance with Young’s model we assign a specific size to each tile. Whether the trichome initiation starts in a tile is determined by the threshold condition: If a tile has size $x$ such that $x \geq T$ (with some threshold $T$), then trichome initiation occurs, otherwise it does not.

There has to be some function for determining the tile size. In order to find a suitable function we have to take biological background into account. One of the facts is that there are no trichomes on the leaf boundary. Therefore we decided to set both the initial and all boundary values to $s_0 < T$. Furthermore, there are usually no neighbouring trichomes. In terms of modelling this means that if we want to compute the size of a tile, it must depend on the size of the neighbouring tiles. In order to simplify matters in our model, the tile size $x$ depends only on the sizes $x_1$, $x_2$ and $x_3$ of the neighbouring tiles from the level above (see Figure 2).

![Figure 2: Model Description](image)

This means that in our model $x$ is a function depending on $x_1$, $x_2$ and $x_3$, i.e. $x = f(x_1, x_2, x_3)$. The idea now is that if one of these neighbouring tiles has a trichome, i.e. $\max(x_1, x_2, x_3) \geq T$, then we need $x < T$ in order to avoid neighbouring trichomes. In the other case where the neighbouring tiles have no trichomes, i.e. $\max(x_1, x_2, x_3) < T$, $x$ should at least be greater than $\max(x_1, x_2, x_3)$, making it more likely that a trichome occurs. One possible function one can use is

$$x = \sum \omega_i x_i > \max(x_1, x_2, x_3).$$

Altogether we get the following algorithm for computing the tile size:

If $\max(x_1, x_2, x_3) \geq T$

then $x = s_0 < T$

else $x = \sum \omega_i x_i > \max(x_1, x_2, x_3)$

In our simulations we used $\omega_i = 1$ for $i = 1, 2, 3$. Of course this function is only one possible way to fulfil biological requirements.
In the next section we discuss our numerical results and possible improvements.

6 Results and Outlook

6.1 Numerical Results

Numerical results of our model are very promising in both biological and mathematical sense. One can generate a wide range of different trichome distribution patterns by varying the threshold. Figures 3–4 show the simulation of trichomes for $T = 20, 50, 80$ and $100$, where number of trichomes decreases with increase in $T$. Figures 5–6 show the typical simulation for $T = 10^5, 10^6$, $5 \times 10^6$ and $10 \times 10^6$. In Figures 5–6, number of trichomes again decreases with increase in $T$.

In each of the three sub-figures in Figure 7, number of trichomes is plotted as a function of threshold. Each sub-figure shows that as threshold increases the number of trichomes decreases. This validates the fact that as the concentration of inhibitor increases trichomes initiation rate is reduced. However, in the last two sub-figures, after initial rapid decrease, the number of trichomes seems to level out to a constant value. Another observation is that the curve is not monotonous in each sub-figure. Moreo, the model does not really break down even when the threshold is extremely high ($10^{50}$) (see Figures 5–7a).

Another point to consider is scaling of the model. In Figure 7 number of trichomes seems to scale linearly (in logarithmic scale) in the middle part of the curve and so it seems reasonable to fit a line of the type $y = a \cdot \log(x) + b$ to the data (upper surface of the ”curve”), where $x$ is the threshold, $y$ is the number of trichomes, and $a$ and $b$ are constants to be determined. This could actually be a benefit because in biology and chemistry scaling is sometimes logarithmic (for example ph-scale). Therefore it is not inconceivable that the effect of activators and inhibitors determining the trichome growth might also depend on the logarithm of the amount of growth compound(s).

To validate the patterns one needs to obtain some experimental data that can be compared to this model. Apart from comparing patterns of trichomes one could also compare some numerical values produced by the model with corresponding numerical values obtained from experiments. Such values could be, for example, trichome density, mean distance between the trichomes or collection of trichomes (such as, line, clusters, etc.).

6.2 Possible improvements

In the Matlab code the size of the tile is a double-type variable. This leads to unnecessary numerical errors. It is fairly easy to implement a 64 bit unsigned integer variable in Matlab and this would be an obvious way to improve the computational part of the model. However, even with this improvement it would not be possible to handle integers that have more than 19 digits.

In section 5, we proposed and proved that there are three regular polygons (i.e. equilateral triangles, squares and regular hexagons) that tessellate a plane without gaps or overlaps. Using triangles as tiles would lead to several problems. First of all one would need two kinds of tiles to construct a leaf (triangles with the base on top and ones with the base at the bottom). Secondly one would get a sawtooth pattern at the rim of the leaf.

Though we have used squares with varying results, there are two main problems in using squares as tiles. The first problem is that all neighbouring tiles are not equal. This can be seen in Figure
Figure 3: Typical Distributions for Small Thresholds; [a]: $T = 20$, [b]: $T = 50$. 
Figure 4: Typical Distributions for Medium Thresholds; [a]: $T = 80$, [b]: $T = 100$. 
Figure 5: Typical Distributions for Big Thresholds; [a]: $T = 10^5$, [b]: $T = 10^6$. 
Figure 6: Typical Distributions for Large Thresholds; [a]: $T = 5 \times 10^6$ and [b]: $T = 10 \times 10^6$. 


Figure 7: Number of trichomes in dependence on the value of the threshold
2, Section 5.3, where tiles $x_1$ and $x_3$ are connected to tile $x$ only by corners of the tiles whereas tile $x_2$ is connected to tile $x$ by one whole side. The second problem arises at the meristem when a tile is created. In our model it is taken as an experimental fact that there should not be trichomes in neighbouring tiles under any circumstances. Therefore the size of a tile depends not only its neighbours on layer above ($x_1$, $x_2$, $x_3$ in Figure 8) but also on its neighbours on the same layer (A and B in Figure 8) and hence the distribution of trichomes depends also on the order in which the tiles are created.

![Figure 8: Simulated Cells](image)

In deterministic numerical models one expects, at least in some sense, symmetrical outcome if the initial state of simulation is symmetric. Therefore in the meristem we first created a tile on the center line (mirror symmetry line) of the leaf and worked out from there symmetrically to left and right. First, this approach demands a symmetric leaf because there needs to exist a mirror symmetry line. It is an open question if the leaf should actually be symmetric on the length scale used in the simulation but in any case it would be beneficial to be able to explore non-symmetrical leaf forms. Moreso, this gives a preference to the tiles near the mirror symmetry line.

However one could try to overcome some of the problems resulting from tessellating squares by giving different tiles different weights in the sum. For example, $w_1 = \frac{1}{\sqrt{2}} = w_3$ and $w_2 = 1$ would make a sense if one assumes that the interaction between the cells is inversely proportional to the distance between the centers of tiles.

Hexagon is an ideal choice because all the neighbouring tiles are connected to the center tile by one whole side and in this sense all neighbouring tiles are equal. Also, with hexagons there are no problems in choosing the order in which the tiles are updated (or created) in the meristem because geometry (see Figure 9) can be chosen in such a way that there are no neighbouring tiles in the meristem.

7 Conclusion

Initiation and distribution of trichomes on *Arabidopsis thaliana* have been considered both from the biological and mathematical perspectives. A novel algorithm was used for this process. A leaf tessellated with squares has been used all through.

We do not claim our model is perfect (this is clear with close observation on Figures 3-6, as it only approximates real life scenario) as, in the words of George Box, all models are wrong and some are useful, but we hope that the model will be validated in the near future with experimental data for possible improvement. We also hope that a model based on leaf tessellation with regular hexagons will be developed in the future.
Figure 9: Hexagon tiles: Blue tiles are in the meristem, red tiles are in the layer just above the meristem and other existing tiles are green.

References


