Challenging *Arabidopsis thaliana* as the Ultimate Model Species: Can Seed Germination be the Achilles’ Heel?

Emilio Cervantes*

Instituto de Recursos Naturales y Agrobiología. Consejo Superior de Investigaciones Científicas (IRNASA-CSIC), Apartado 257, Salamanca, Spain

Correspondence: *ecervant@usal.es

ABSTRACT

*Arabidopsis thaliana* has reached maturity as a model system in plants. In this short text first, the meaning of the word model is discussed in relation to the ease with which developmental processes can be analyzed and also in relation to the influence that the model itself has on experimentation. Although the efficiency of *Arabidopsis* as a plant model has been demonstrated, and is beyond the scope of this discussion, the description of ecological strategies in plants indicates that *Arabidopsis* is an example of an r-type strategy species. Plants with the r-type strategy, and, in particular, *Arabidopsis*, although structurally simple, may support complex regulatory processes. This could involve complex genome dynamics, in particular in response to given environmental conditions during situations such as, for example, seed germination.

**Keywords:** complexity, development, evolution, genome, regulation

**ARABIDOPSIS AS A MATURE MODEL FOR PLANT BIOLOGY**

*Arabidopsis thaliana* has been the model species in Plant Biology for some time. Considering the initial publication of the *Arabidopsis* Information Service in 1964, the model is almost 45 years old; thus, we may say it has reached the age of maturity. Eleven years after the publication of the *Arabidopsis* Information Service, George Rédel’s article entitled “*Arabidopsis* as a Genetic Tool”, appeared in the Annual Review of Genetics (1975), contributing to extend the work in progress with *Arabidopsis* and to establish this species as the model plant. The model was slowly consolidating throughout the 1980’s and 1990’s. The 4th International Congress of Plant Molecular Biology, held in Amsterdam (1996), had a full session devoted to *Arabidopsis* that was entitled “Arabidopsis: a model plant species”. The model that was then developing, has now reached full maturity, and, as an example of this, a single search in Google for “Arabidopsis”, done in March, 2008, yielded 6.630.000 internet quotes (Table 1). These data give an idea about the actual success of *Arabidopsis* in research and invite us to consider whether the model satisfies all requirements or, if alternatively, it may present any particular aspects in which this species may not be conforming correctly as would be expected for such a mature and broadly applied model system.

The agreement to concentrate efforts on *Arabidopsis* originated from early genetic studies (Rédel 1975), and was further supported by the simplicity to grow this plant and the ease with which a large amounts of seeds could be obtained. Results from molecular biology contributed to the extension of *Arabidopsis* as a model showing that cells had a small genome with relatively low content in heterochromatin and a high proportion of coding sequences (The *Arabidopsis* Genome Initiative 2000). A drawback during plant and animal sequencing projects consists in the presence of repeated sequences in the centromeric regions and other locations representing a real difficulty to obtain the complete genomic sequence (Copenhaver *et al.* 1999; Hall *et al.* 2003). Centromeric regions of variable complexity, rich in repeated sequences, are characteristic of almost all eukaryotic species, excluding only the simplest yeast species like baker’s yeast (*Saccharomyces cerevisiae*) and its close relatives (Hegemann and Fleig 1993). Being one of the plants with a low amount of repeated DNA supported *Arabidopsis* as a model species.

| Table 1 Results of search in Internet (Google) and in Scopus (Elsevier’s publications database). |
| --- | --- | --- | --- |
| PubMed | Scopus | Google |
| www.pubmed.org | www.scopus.com | www.google.com |
| Arabidopsis | 4152 | 23857 | 6642 |
| Arabidopsis development | 876 | 7935 | 7533 |
| Arabidopsis flowering | 206 | 1377 | 1720 |
| Arabidopsis germination | 117 | 990 | 1176 |
| Arabidopsis microarray | 87 | 627 | 1053 |
| Arabidopsis MPSS | 10 | 26 | 30 |
| Arabidopsis cDNA-AFLP | 10 | 25 | 33 |

*Search done in March, 2008.*

Both aspects, maintenance of information through division cycles and regulation of development, are interlinked by microRNA molecules involved in both the structural organisation of the centromere (Reinhart and Bartel 2002) and the regulation of transcription factors, keys for the regulation of development (Yekta *et al.* 2004; Rinn *et al.* 2007). The maintenance of coding capacity for species-specific characteristics and conserved development are the two sides of the same coin, because the main characteristic of a given species is how it controls development.

Importantly, concepts from the field of ecology appear to indicate that when one wishes to concentrate on certain

Received: 7 February, 2008. Accepted: 13 April, 2008.
regulatory aspects of development, Arabidopsis may not be the simple system that we expected. However, before delving into the explanation of these concepts, allow me to concentrate on the definition of the characteristics that a model needs to fulfil.

WHAT IS A MODEL (PLANT)? DEFINITION AND CHARACTERISTICS

The concept of a model, in general, is an old and important one in science. According to specialists in the history of science, such as Duhem (1978) and Rossi (1986), models first appear in the English scientific tradition as the elements introduced to give connexion to a theory, to make it concrete and understandable. For example, Rossi (1986) describes how Faraday imagined a series of elastic strings to represent his model of electrostatic attraction, and this model was highly esteemed and admired by Maxwell.

We also often hear of mathematic models as the ideal mathematic representations of complex situations in nature. A summary of the application of mathematical models to plant development has been given by Alvarez-Buylla et al. (2007). In addition to authors quoted in this review, the books “Stabilité structurelle et morphogénèse” (Structural stability and morphogenesis), by René Thom (1972), and “Mathematical Biology”, by JD Murray (2002) are classic texts about mathematic models in the interpretation of nature.

But, more recently in biology, the concept of model species has gained importance as a collective working strategy, based in the concentration of work in a single, representative species. In this sense, a biological model is the species of any taxonomic group of living organisms that was chosen as a reference. The strategy consists in the concentration of efforts with the idea that gaining knowledge in the analysis of the model will help to understand all the elements better. The model is selected because it is representative (i.e., it has characteristics common to all elements), and because it has features that make it easier to analyse any given process. Implicit in the idea of a model is the concept of ease, but ease may be readily confused with “simplicity”. The more work is done with the model, the better the model becomes established, and, in general, analyses of processes become easier, because much work has been done before, thus channelling further work and resulting in the generation of resources needed to simplify further analysis. Nevertheless, this does not mean that all the developmental processes or aspects will be simpler in the model than in other species.

Among plants, diverse models where proposed in response to particular research interests. Thus, for monocots, maize, wheat and rice where considered models for some time. Now, the prevalent model for monocots is rice, mainly due to its smaller genome (Goff 1999; Goff et al, 2005), cDNA-AFLP (de Diego et al, 2006) or Massive Parallel Sequence Signature (Hoth et al, 2002).

Today, working in seed germination with Arabidopsis offers several advantages. It is feasible, for example, to identify genes up- and down-regulated during germination by diverse methods such as microarrays (Ogawa et al, 2003; Li et al, 2005), cDNA-AFLP (de Diego et al, 2006) or Massive Parallel Sequence Signature (Hoth et al, 2002).

The first method analyses the expression of sequences that were previously cloned and ordered in the array; while the second and third methods, departing from cDNA corresponding to RNA populations, allow the identification of novel, previously un-cloned sequences that were indeed expressed during this process. cDNA-AFLP (Bacher et al, 1996) and MPSS (Brenner et al, 2000) allow the cloning of cDNAs corresponding to RNA molecules expressed in particular tissue or environmental conditions. MPSS allows the cloning of cDNAs corresponding to molecules present in very low amounts in tissues. Both methods, MPSS and cDNA-AFLP, differ from microarrays in that no previous information is needed to be available on the sequences analysed. Although the three methods allow the identification of changes in mRNA populations that occur in the seed during germination, the first and more popular of them (microarrays) is based on previously cloned sequences; thus the use of microarrays, the method predominant in Arabidopsis (Table 1), is more frequent than others offering the possibility of cloning new sequences. In the same direction but in a more general sense, the identification of sequences induced during germination is a “subtle” aspect in the analysis of this process, having preference over other key aspects of the process of germination. For example, the analysis of variation at the time of germination.

Variation in time is a key aspect of germination. If all seeds of a seed lot germinate at the same time, this results in the compromise of the corresponding genotype. Thus mechanisms must exist in seeds by which diversity in germination is guaranteed. Two main hypotheses may explain variation at the time of germination. First, variation may be
Table 2 A comparison between the main characteristics of two main reproductive strategies, r- and K-strategies of plants and animals.

<table>
<thead>
<tr>
<th>r-selected</th>
<th>K-selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>mature rapidly</td>
<td>mature slowly</td>
</tr>
<tr>
<td>short-lived; most die before they reproduce</td>
<td>tend to live long lives: low juvenile mortality rate</td>
</tr>
<tr>
<td>have many offspring - tend to overproduce</td>
<td>compete well for resources</td>
</tr>
<tr>
<td>invest little in individual youngsters</td>
<td>have few offspring at a time</td>
</tr>
<tr>
<td>most pest species are r-selected</td>
<td>most endangered species are K-selected</td>
</tr>
<tr>
<td>population not regulated by density: boom and bust population figures</td>
<td>population stabilizes near carrying capacity</td>
</tr>
<tr>
<td>opportunistic – invade new areas</td>
<td>maintain numbers in stable ecosystems</td>
</tr>
</tbody>
</table>

Adapted from Begon et al. (2006)

considered as the additive result of small differences in multiple processes occurring during germination. Second, variation may be the consequence of the activity of regulatory processes directed to this particular end.

If variation in time of germination was just the consequence of an additive result of small differences in processes, then there would not be differences in seed germination through species, but the consideration of some principles of plant ecology indicates that differences must exist between species and that the second hypothesis may be true.

ECOLOGICAL STRATEGIES IN PLANTS

Plant and animal ecologists distinguish two types of reproductive strategies, r- and K- (See for example Begon et al. 2006; Table 2). r- stands for rapid, thus organisms with an r-strategy will invest in rapid reproduction, whereas K-species will invest in prolonged development and a long life. Life cycles are shorter and morphological complexity is reduced in r-strategy species. Arabidopsis is clearly an example of an r-strategy species.

Seeds of r-strategy species are also single-shaped. They are produced in large quantities, released from the mother plant and do not have any sophisticated structural characteristics allowing them to be distributed by specific means such as air, water or animals; thus, in contrast to seeds of plants of the K-strategy, r-strategy seeds remain close to their site of production near the mother plant. Whereas K-strategy is based on a spatial distribution of seeds, r-strategy is related with seed germination scattered through time. Arabidopsis seeds do not move far away from the mother plant; instead, their strategy depends on a prolonged germination in time. Broadly, morphologically simpler seeds may have complex germination regulation. This can be seen when considering seed dormancy in relation to different life cycles, as for example in Table 4.2 of Baskin and Baskin (2001). In general, non dormant seeds belong to biannual or perennial plants (K-type strategy), whereas dormant seeds correspond more frequently to annual plants (r-type strategy). The complexity in the regulation of germination in Arabidopsis seeds is shown in the results of germination tests done through the years in our laboratory under standard germination conditions.

VARIATION IN SEED GERMINATION IN ARABIDOPSIS

In a seed population, there will be always some variation in germination speed. This is due to internal differences (in seed structure or composition) that can be the product of differences in genotype, in growth conditions, or in both. Variation is a necessary characteristic for adaptation to a changing environment as well as for evolution, because if all seeds were identical in their responses to environmental stimuli, they would all be unable to germinate under given conditions, this being negative for the maintenance of the species. Through evolution, plants need to invest in energy to maintaining diversity. The key to variation lies in the plasticity of the genome by dynamic processes still unknown.

In general, seeds of similar genotypes may respond variably to the same environmental conditions. For example, in lots of seeds sharing overall the same genetic composition (i.e., lots of isogenic lines or even in lots of seeds from the same plant), there will be variation observed in the analysis of seed germination. Variation is observed not only from experiment to experiment, when the environmental conditions may change, but also in repetitions of the same experiment, where only small variations occur: Seeds of the same lot present variable germination rates under only slightly varying experimental conditions. One of the classical problems in seed companies consists in obtaining high percentages of germination rates in seed lots. Even the best lots optimised for high germination rates will present variation and (almost) never could be guaranteed 100% germination in a seed lot. Unfortunately, the reports of germination tests have concentrated on the mean values and often do not include an analysis of the variation.

Multiple factors may influence seed germination, such as variations in temperature, humidity, light conditions and other types of radiation, and the presence in the germinating medium of salts, metals or nutrients, just to mention some factors, but it is important to remark that probably seeds perceive environmental variables in an integrated fashion and that some other factors, important for seeds may be still unknown or poorly defined.

Standard conditions for germination in our laboratory involve a sterilization process (washing 2 min. in ethanol and 7 min. in sodium hypochlorite, followed by rinsing four times in distilled water). After sterilization, the seeds are placed in Petri dishes and following a cold-incubation for three days, they are exposed to optimum light and temperature (20°C, about 300 µE m–2) in a day-night cycle (16 h/8 h). Under these conditions, the germination of Arabidopsis thaliana var. Columbia is highly variable (Table 3). Not
only are mean values of germination highly variable, but also deviations from the mean are variable and reach high values (Table 3). High variability in germination rates and deviations from the mean correlates with the complex regulation of germination characteristic of the r-type strategy.

Seeds belonging to plant species having an r-type strategy may present more variation, whereas those having a K-type strategy may germinate more evenly. Plants that follow an r-type strategy have simpler morphology, but on the other hand, their regulation of germination may be more complex. Although in the studies of seed germination in diverse species it is not frequent to analyze the deviation from the mean, our results in Arabidopsis show that germination under standard laboratory conditions results in high variation rates.

AN ANCIENT EVOLUTIONARY THEORY
In his book entitled ‘Les transformations du monde animal’, published in Paris in 1929, Charles Dépéret mentions what he calls “Non-specialisation doctrine” and explains that this is an original concept derived from Edward Cope, who considered this as one of the main principles in evolution. Following Cope, the stages of advanced specialisation that are characteristic of apparently complex life forms, represent indeed powerful adaptations, but on the other hand, their specialisation results in their inaptitude in the face of changing conditions. Complex, sophisticated structures would be the result of simple (specifically adapted) regulatory processes, whereas simple structures could support more complex regulatory processes resulting in new adaptive developments in the case where environmental changes require new adaptations. Clearly, plants with strategies of the r-type in general, and Arabidopsis in particular, are simpler in form, but are able to support complex regulatory processes. This may involve genome dynamics, in particular in response to given environmental conditions, and in consequence, a higher variation rate in seed germination. In fact our data from germination tests done through the years indicate that seed germination is in Arabidopsis, a highly variable process.

CONCLUSION: AN HYPOTHESIS
Based on three independent arguments we propose that the simple morphology that characterizes the model plant Arabidopsis thaliana is accompanied by complex regulatory mechanisms during seed germination. The arguments are: 1) From the field of ecology, showing that Arabidopsis is an example of an r-strategy plant; 2) From our own germination data, showing high standard deviation in germination tests and 3) Finally, the evolutionary argument first established by Cope and described by Dépéret (1929) and called the “Non-specialisation doctrine” may support the hypothesis that Arabidopsis complexity in seed germination is a characteristic example of how simple structures may support more complex regulatory processes resulting in new adaptive developments in the case of environmental changes. We are currently investigating in the laboratory the physiological basis of the high variation rates observed during A. thaliana seed germination.

ACKNOWLEDGEMENTS
I thank Juan José Aladasoro (Real Jardín Botánico CSIC, Madrid) for sharing ideas through an interesting discussion one morning in winter, 2006.

REFERENCES
Duhem P (1978) La Teoría Fisica, Il Mulino, Bolonia, Italy, 391 pp