

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

Table 1 Results of search in Internet (Google) and in Scopus (Elsevier's publications database).

	Science direct www.sciencedirect.com	PubMed www.pubmed.org	Scopus www.scopus.com	Google www.google.es
<i>Arabidopsis</i>	4152	23857	6642	6.630.000
<i>Arabidopsis</i> development	876	7935	7533	426.000
<i>Arabidopsis</i> flowering	206	1377	1720	230.000
<i>Arabidopsis</i> germination	117	990	1176	539.000
<i>Arabidopsis</i> microarray	87	627	1053	234.000
<i>Arabidopsis</i> MPSS	10	26	30	63.700
<i>Arabidopsis</i> cDNA-AFLP	10	25	33	9580

Search done in March, 2008.

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.

Genome organisation and the complexity of the centromeric region may be related to two aspects of interest in biology that are particularly refractory to analysis:

- First, the need to maintain constant DNA information through mitosis and meiosis.
- Second, regulation of development.

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.

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

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 were proposed in response to particular research interests. Thus, for monocots, maize, wheat and rice were considered models for some time. Now, the prevalent model for monocots is rice, mainly due to its smaller genome (Goff 1999; Goff *et al.* 2002). In legumes, pea used to be the historic model, because of broad agricultural application, relatively small genome and, being the species of Mendel experiments, but in the age of genomics, *Lotus japonicus* (Asamizu *et al.* 2004) and *Medicago truncatula* (Bell *et al.* 2001) became the species of choice due to a smaller genome than pea. Other plant models selected with different purposes were: soybean, tomato, spinach and *Antirrhinum*, but finally *Arabidopsis* was imposed as the main model in plants. The properties that made *Arabidopsis* a general model were detailed by the Multinational *Arabidopsis* Steering Committee (www.nsf.gov/pubs/2002/bio0202/model.htm).

Simplicity and ease, both characteristics highly desired in a model, are present in *Arabidopsis* and this may have an additional undesired effect that makes things look simpler than they really are. Let's first discuss the meaning of the words simple and easy, after which we may be in a better condition to see why work with a model system can make us tend to see things simpler than they are.

The word Simple has two meanings: one is clear, the other more complicated. The first refers to composition and

is used for example in botany. In botany, “simple” means something that has no parts, in contrast to “complex” that means something having parts, but this is not the unique meaning of “simple”. The second meaning of “simple” refers to level of complexity. “Simple” in this sense means lacking complexity. But in this case, the meaning is relative and subjective, i.e., it refers to our personal view of something. In this sense, the objects or the processes of the natural world are not simple or complex in an objective way: their complexity is relative and depends upon our point of view. We can perceive something as being “simple”, but if we change our point of view, our criteria or our method of approximation, this may become complex. A drop of water may be seen as an extremely complex entity when we observe it under a particular type of chemical considerations, whereas the moon may be seen as a simple object under a poor optics telescope.

The meaning of Easy is related to the second meaning of Simple and when we say, for example, that the analysis of flowering in *Arabidopsis* is simpler than in other plant, we mean that it would take us less effort to obtain a certain amount of information, results or to reach new conclusions in *Arabidopsis*. This is based on the fact that *Arabidopsis* offers facilities and resources not available in other plants. But it is important to notice here that not only our results but also our conclusions will be dependent upon our resources. Also the hypotheses forming our point of departure will be strongly influenced by our resources. Working with a species consolidated as a model plant, as is *Arabidopsis*, forces us to work on its aspects that present us with a more straightforward or consolidated way to develop a successful project, and not with the ones that could be considered more interesting before or in the absence of such a model. We may see this better with an example.

CHALLENGING THE MODEL: THE ANALYSIS OF SEED GERMINATION IN *ARABIDOPSIS*

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 are induced during this process. cDNA-AFLP (Bachem *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 need 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 “preferential” 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.

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

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

Table 3 Seed germination in *Arabidopsis thaliana* ecotype Columbia. Germination was under standard conditions (cold storage of imbibed seeds during three days and then exposure to 20°C with a 16/8 light-dark cycle) Summary results of experiments done through the years 1999, 2001 and 2004.

Date	% Germination	Mean	Std Dev.
04/05/1999	59	46	18.3
	33		
05/05/1999	81	85.5	6.4
	90		
05/05/1999	78	82.5	6.4
	87		
21/02/2001	73.9	66.5	10.5
	59.1		
26/02/2001	52.1	41	15.6
	30		
26/06/2001	33.3	36.6	4.7
	45.3		
23/06/2004	31.1	42.7	16.4
	54.3		
23/06/2004	35.3	59.9	30.5
	78.5		
TOTAL		57.58	13.6

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 $\mu\text{E m}^{-2}$) 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 Depéret 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 Depéret (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 Aldasoro (Real Jardín Botánico CSIC, Madrid) for sharing ideas through an interesting discussion one morning in winter, 2006.

REFERENCES

- Alvarez-Buylla ER, Benítez M, Balleza Dávila E, Chaos A, Espinosa-Soto C Padilla-Longoria P (2007) Gene regulatory network models for plant development. *Current Opinion in Plant Biology* 10, 83-91
- Asamizu E, Nakamura Y, Sato S, Tabata S (2004) Characteristics of the *Lotus japonicus* gene repertoire deduced from large-scale expressed sequence tag (EST) analysis. *Plant Molecular Biology* 54, 405-414
- Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF (1996) Visualisation of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during tuber development. *The Plant Journal* 9, 745-753
- Baskin CC, Baskin JM (2001) *Seeds. Ecology, Biogeography and Evolution of Dormancy and Germination*, Academic Press, San Diego, CA, USA, 666 pp
- Begon M, Townsend CA, Harper JL (2006) *Ecology: From Individuals to Ecosystems* (4th Edn), Blackwell, UK, 738 pp
- Bell CJ, Dixon RA, Farmer AD, Flores R, Inman J, Gonzales RA, Harrison MJ, Paiva NL, Scott AD, Weller JW, May GD (2001) The *Medicago* Genome Initiative: a model legume database. *Nucleic Acids Research* 29, 114-117
- Brenner S, Williams SR, Vermaas EH, Storck T, Moon K, McCollum C, Mao JI, Luo S, Kirchner JJ, Eletr S, DuBridge RB, Burcham T, Albrecht G (2000) *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proceedings of the National Academy of Science USA* 97, 1665-1670
- Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S, Lin X, Bevan M, Murphy G, Harris B, Parnell LD, McCombie WR, Martienssen RA, Marra M, Preuss D (1999) Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* 286, 2468-2474
- Depéret Ch (1929) *Les Transformations du Monde Animal*, Flammarion, Paris, France, 360 pp
- De Diego JG, Rodríguez FD, Rodríguez Lorenzo JL, Grappin P, Cervantes E (2006) cDNA-AFLP analysis of seed germination in *Arabidopsis thaliana* identifies transposons and new genomic sequences. *Journal of Plant Physiology* 163, 452-462
- Duhem P (1978) *La Teoria Fisica*, Il Mulino, Bologna, Italy, 391 pp
- Goff SA (1999) Rice as a model for cereal genomics. *Current Opinion in Plant Biology* 2, 86-89
- Goff SA, Ricke D, Lan T, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, MB Lange, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W-I, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92-100
- Hall SE, Kettler G, Preuss D (2003) Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains. *Genome Research* 13, 195-205
- Hegemann JH, Fleig UN (1993) The centromere of budding yeast. *Bioessays* 15, 451-460
- Hoth S, Morgante M, Sanchez JP, Hanafey MK, Tingey SV, Chua NH (2002) Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the *abi-1* mutant. *Journal of Cell Science* 115, 4891-900
- Li F, Wu X, Tsang E, Cutler AJ (2005) Transcriptional profiling of imbibed *Brassica napus* seed. *Genomics* 86, 718-730
- Murray JD (2002) *Mathematical Biology I: An Introduction* (3rd Edn), Springer, New York, 586 pp
- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S (2003) Gibberellin biosynthesis and response during *Arabidopsis* seed germination. *Plant Cell* 15, 1591-1604
- Rédél G (1975) *Arabidopsis* as a genetic tool. *Annual Review of Genetics* 9, 111-127
- Reinhart BJ, Bartel DP (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science* 297, 1831
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Good-nough LH, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311-1323
- Rossi P (1986) *I Ragni e le Formiche: Un'Apologia della Storia della Scienza*, Il Mulino, Bologna, Italy, 292 pp
- The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815
- Thom R (1977) *Stabilité Structurale et Morphogénèse* (2nd Edn), InterEditions, Paris, 351 pp
- Yekta S, Shih I, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304, 594-596