

Why is *Arabidopsis* so Special?

Jaime A. Teixeira da Silva

Faculty of Agriculture and Graduate School of Agriculture, Kagawa University, Miki-Cho, Ikenobe, 2393, Kagawa-Ken, 761-0795, Japan

Corresponding author: * jaimetex@yahoo.com

ABSTRACT

The correct form of representing the Latin name of a plant is in italics. For example, the Latin name for potato is *Solanum tuberosum*. Therefore, *Arabidopsis thaliana* would be italicized. However, while the common name of most plant species is written with a lower-case letter, such as potato, spinach or lettuce, *Arabidopsis* is not written as *arabidopsis* in many if not most *arabidopsis* papers. This is most likely because the common name for *A. thaliana* is not *Arabidopsis*, it is thale cress or mouse-ear cress (in English). This begs the question, where did the “technical jargon” or “informal short-hand” come from? However, a very recent (2012) publication uses the term *Arabidopsis* to describe three *Arabidopsis* species (*A. thaliana*, *A. suecica* and *A. arenosa*), even if only in the title, throwing this theory into disarray. Who introduced it into the scientific literature, and for what reasons? This short paper aims to discover a little more about a little-known and explored issue of this tiny model species, which rose from weed status to star power.

Keywords: *Arabidopsis thaliana*, thale cress

INTRODUCTION

In the past decade or so, I have personally seen such a wide ranging-use of the terms *Arabidopsis*, *Arabidopsis* and occasionally *arabidopsis* within the literature. Initially, I made the innocent supposition that in fact there was an error in the literature, since the common name of all plant species was always written in lower-case letters while the Latin, botanical name was always written in italics, for example, wheat for *Triticum aestivum*, maize (UK) or corn (US) for *Zea mays*, potato for *Solanum tuberosum*, etc. Then I realized my own error, but one may be made by many similarly astute plant scientists, i.e., that the common name for *Arabidopsis thaliana* is thale cress or mouse-ear cress. This then left me with very few final pieces of the puzzle which were not matching the rest of the 1000-piece puzzle. Why was I observing *Arabidopsis* throughout the literature (using data-bases such as Elsevier’s Scopus and sciencedirect.com, SpringerLink, Taylor and Francis, Wiley-Blackwell, Oxford University Press, and PubMed) and not *arabidopsis* (i.e., why was the A being capitalized and not written as lower-case a)? Interestingly, if you try to write *arabidopsis* in Window’s Microsoft Word (2003, 2007 or 2010 versions), it is automatically converted to *Arabidopsis*. This fact, in itself, is fascinating. For two reasons: 1) it indicates that this word is formally recognized as being a regular lexicon in the English language (at least according to Microsoft); 2) the formalization of the use of *Arabidopsis* (neither as a common name nor as the Latin genus name *Arabidopsis*) was prior to 2003.

Seeing this odd fact, I decided to explore further. Initially, I contacted about 50 scientists who had published work on *Arabidopsis* (broadly) in leading plant science and related molecular biology journals, because the assumption was that such individuals would surely know such basic information about their test plants. Although I did receive a few fairly confident and convincing arguments, I did nonetheless, note in at least 80% of respondents, some hesitation in the response. What did emerge was the fact that the term *Arabidopsis* emerged as a “technical jargon”, an “informal short-hand”, or “nick-name” for *A. thaliana*. That is fine. This would then leave one less question left to answer. But one paper caught my eye that made me realize that maybe the rule was not that uniform. In a paper recently published

in 2012 (Pontvianne *et al.* 2012), the term *Arabidopsis* was used in the title to describe three *Arabidopsis* species (*A. thaliana*, *A. suecica* and *A. arenosa*). I accepted this to be a highly respectable decision because we were indeed dealing with a legendary plant science group led by Prof. Craig Pikaard, thus this decision was likely to be true and valid. I approached the corresponding author about this issue, and the response I received was “It is a good question. In my personal case, I use *Arabidopsis* with an A when I am talking about the genus *Arabidopsis*. For example, in my last article [(i.e., Pontvianne *et al.* 2012)], big A is used here because I am talking about different members of the genus *Arabidopsis* (*A. thaliana*, *A. suecica* and *A. arenosa*). However, I agree that in that case I should have used “within the *arabidopsis* genus” instead of “in *Arabidopsis*”, but title should be as short as possible. It another reason using *Arabidopsis* instead of *Arabidopsis thaliana* could be use as a shortcut...” The author did however clearly state that he was not absolutely sure of these facts, and did state that the use of *Arabidopsis* to describe three *Arabidopsis* species was limited strictly to the title. However, this latter use and interpretation of the use of *Arabidopsis* for describing several *Arabidopsis* spp. would go against the more commonly attributed, and exclusive, use to *A. thaliana*. Could a new “error” have been accidentally introduced into the literature, and could the Pontvianne *et al.* (2012) paper be an interesting milestone in the history of the *Arabidopsis* literature? I use the term “error” because it seems as if the actual consequences of this publication were not even apparent to the authors themselves. With this publication, in fact, a “dangerous” precedent has been set in which future manuscripts related to any *Arabidopsis* spp. could in fact refer to any or all of them as simply *Arabidopsis* (which was originally destined for *A. thaliana*). Currently there are two queries arise about the original introduction of the term *Arabidopsis* (capitalized A and no italics) to only describe *A. thaliana*: a) WHO introduced this term for the first time into the literature and why was the term allowed to be published (assuming that the journal was peer reviewed)? b) WHY was this term introduced and given such red-carpet and exclusive treatment – at least more so than other major model crops such as maize, wheat, tomato or potato – when common names exist for *A. thaliana*? Certainly these issues lead to a conspiracy theory, stimulating the ensuing questions:

- Was a “movement” started years ago to popularize *A. thaliana* and to create some exclusivity within the scientific community, i.e. was a biased system intentionally introduced?
- Why is *Zea mays* not referred to as *Zea*, *Triticum aestivum* as *Triticum*, *Spinacia oleracea* as *Spinacia*, or *Solanum tuberosum* as *Solanum*? Is this because maize, wheat, spinach and potato researchers did not have as much initiative (or power/clout) as the *Arabidopsis* scientific community?
- *A. thaliana* was the first plant to have its genome sequenced (The *Arabidopsis* Genome Initiative 2000). Does this have something to do with the “superior”, No. 1 status of *A. thaliana* or with the genome sequencing project? According to Bevan *et al.* (2001), the *Arabidopsis* Multinational Steering Committee was established in 1990, with the objective of concentrating international effort on exploring this plant based on several landmark research results that had been found with this plant.

Prof. Randy Scholl, who started working with *Arabidopsis* in the 1970's, and who always referred to it as *A. thaliana* in scientific papers, believes that “in the 80s, with the influx of molecular biologists into the field, referring the specie[s] as *Arabidopsis* definitely became common usage”, suggesting a biased influence on the use of this popularized term.

I decided to explore further.

Some typical examples of:

- Case 1 (most common) *A. thaliana* abbreviated as *Arabidopsis* (capitalized and italicized): Tominaga-Wada *et al.* (2011). This use goes back at least as far as 1997 (Delseny *et al.* 1997) and the use may have been set in stone with the high-profile and historic sequencing of the *A. thaliana* genome (The *Arabidopsis* Genome Initiative 2000);
- Case 2 (next most common) *A. thaliana* abbreviated as *Arabidopsis* (capitalized and non-italicized): Louvet *et al.* (2011), Zhang *et al.* (2011). Interestingly, the first two cases were published in the same journal, suggesting that the use of *Arabidopsis* and not *Arabidopsis*, may have been encouraged by the editor board.

Due to the massive data-base on *A. thaliana* and related species, I was unable to quantify the number of cases 1 and 2, but several important conclusions have emerged:

- there is no uniformity in the use of *Arabidopsis* or *Arabidopsis* when describing *A. thaliana*;
- it is unclear whether in fact *Arabidopsis* or *Arabidopsis* should be used to exclusively describe *A. thaliana* or also other *Arabidopsis* species.

Several authors appear to simply “follow the leader”, and do as preceding authors have done, without question (see *verbatim* response 5 in **Appendix**).

CONCLUSIONS

Certainly *Arabidopsis*, referring to *A. thaliana*, is getting special treatment relative to other plants. Whether this treatment has come about as a result of common usage, or as a result of this plant having become an important model plant for molecular genetics as a result of concerted efforts made several decades ago, remains uncertain. Even the *Arabidopsis* Biological Resource Center at Ohio State University, USA, is unable to, or unwilling to, provide a formal response to what would appear to be a most basic question, leaving this issue and query mysteriously unresolved. Fundamentally, what I have not yet been able to determine is WHO exactly and WHEN exactly that first “colloquial”

usage came into being. Perhaps this inquiry will spur others to find the answer to this enigmatic case.

ACKNOWLEDGEMENTS

I wish to thank the following individuals for their thought-provoking responses: Prof. Dr. Marcus A. Koch (Botanic Garden and Herbarium Heidelberg (HEID), University of Heidelberg, Germany), Prof. Rainer Breitling (Manchester University, UK), Dr. Rogelio Rodríguez-Sotres (Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, México), Prof. Maria-Jesus Marcote (Universidad de Valencia, Spain), Dr. Frédéric Pontvianne (Perpignan University, France), Dr. Fikret Mamedov, Department of Chemistry, Uppsala University, Sweden), Dr. James Mann and Dr. Luz Rivero (*Arabidopsis* Biological Resource Center, Ohio State University, USA), and Prof. Randy Scholl (retired). I also thank others (but not by name) who responded, but who did not actually provide new insight.

REFERENCES

- Bevan M, Mayer K, White O, Eisen JA, Preuss D, Bureau T, Salzberg SL, Mewes H-W (2001) Sequence and analysis of the *Arabidopsis* genome. *Current Opinion in Plant Biology* 4 (2), 105-110
- Delseny M, Cooke R, Comella P, Wu H-J, Raynal M, Grellet F (1997) The *Arabidopsis thaliana* genome project. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie* 320 (8), 589-599
- Louvet R, Rayon C, Domon J-M, Rusterucci C, Fournet F, Leautic A, Crépeau M-J, Ralet M-C, Rihouey C, Bardor M, Lerouge P, Gillet F, Pelloux J (2011) Major changes in the cell wall during silique development in *Arabidopsis thaliana*. *Phytochemistry* 72 (1), 59-67
- Pontvianne F, Blevins T, Chandrasekhara C, Feng W, Stroud H, Jacobsen SE, Michaels SD, Pikaard CS (2012) Histone methyltransferases regulating rRNA gene dose and dosage control in *Arabidopsis*. *Genes and Development* 26, 945-957
- The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815
- Tominaga-Wada R, Ishida T, Wada T (2011) New insights into the mechanism of development of *Arabidopsis* root hairs and trichomes. *International Review of Cell and Molecular Biology* 286, 67-106
- Zhang J-X, Sun X-M, Zhang Z-P, Ni Y-W, Zhang Q, Liang X-M, Xiao H-B, Chen J-P, Tokuhisa JG (2011) Metabolite profiling of *Arabidopsis* seedlings in response to exogenous sinalbin and sulfur deficiency. *Phytochemistry* 72 (14-15), 1767-1778

Appendix

Some of the *verbatim* responses I received (actual respondents' names not identified for “privacy” reasons).

- “because it is used frequently as a NAME (e.g. accepted by many journals), not as an OBJECT (potato). This means “*Arabidopsis*” is simply the name. Taxonomically, only the version in italics is correct.”
- “*Arabidopsis* is not a common name (the common name of *A. thaliana* is thale cress), it is technical jargon, an informal short-hand. Even when it is not used as a formal scientific (“Latin”) name and therefore not written in italics, the initial capital still remains. Capital letters are also used in many botany books for the common names of plants, where these are meant to refer to individual well-defined species (i.e. are used as the English language equivalent of a scientific name). A comparable case is, by the way, found in zoology: *Drosophila* is also written in capitals, even if it is not used as a scientific name and not in italics (and it is also not a common name).”
- “I am definitively not an expert on the issue, what I am going to offer you is entirely an opinion and it may be wrong, you decide if it helps you. For the scientific community, *Arabidopsis* has become a proper noun (i.e. a name), very much like “El Niño”. The common noun of this plant would be mouse-ear cress (at least in some communities, where the plant was recognized as a wild plant). This noun is not capitalized. The noun *Arabidopsis* was probably coined as a result of its repeated used in the Scientific literature, and somehow became a sort of tradition. It is clear to me that if not capitalized, it may create confusion, as it could refer to other species of the same genus, such as *A. lyrata*. Yet I cannot tell you where this tradition started. On the other hand, the rules for scientific designation of biological species are clear and in that sense “*Arabidopsis thaliana*” is no exception.”
- “I don't really know the reason. I always thought that the name *Arabidopsis* when is not in italics plays the role of the common name as there was no common name of *Arabidopsis* in English or Spanish etc. Then, they made up the common name in capital letter, as the latin name in short, and of course not in

italics. Tobacco, pea, lettuce are common names that already exist in different languages and they are already in lower-case letter.”

5. “We decided to use *Arabidopsis* as a common name in our papers because: 1. It is shorter and more convenient than full Latin name and if it is used many times in the paper, a shorter form is needed. 2. It starts with upper case I think because the first word in the Latin name is the same and also starts with big A. For example another important plant in our research is spinach and it is used like that in the papers but full name is *Spinacia oleracea*, so it is different. And of course it is more common that *Arabidopsis*. 3. We also saw that other groups use *Arabidopsis* as a common name in their papers.”

6. “My opinion on the use of the term *Arabidopsis* as if it were the common name is as follows: The common names of thale cress or mouse ear cress are fairly generic and nonspecific. The common names are also not well known and are not often used. *Arabidopsis* is not a plant discussed or written about very often in non scientific circles. Generally speaking a person that sees it would only term it as a weed. If you mention thale cress or mouse ear cress most people will have no idea what you are talking about. So the name of the plant is pretty much only discussed in scientific circles where the Latin name would be used in any professional writing or discussion. The common name not being particularly interesting, specific, well known or possibly even liked may have led the people working with this plant to begin to refer to it even in casual circumstances by its shortened Latin name. This shorthand probably propagated fairly quickly through the community and then through the literature. As far as this use being extended to the other members of the genus it is probably because the same logical process that leads one to call *A. thaliana* by *Arabidopsis* is equally true for any other member of the genus and so is equally valid. This unfortunately could lead to some confusion but *A. thaliana* is by far the most discussed member of the genus and you would be generally safe to assume that if someone is talking about *Arabidopsis* they mean *A. thaliana*. It is my opinion that “*Arabidopsis*” is the proper common name for the plant even though it also happens to be the genus name for the plant we are referring to. It is certainly more often called by that name then it is ever called thale cress or mouse ear cress.”

7. “There are guidelines for biological nomenclature that enables people to communicate about plants and animals without confusion. The terms, taxonomy and nomenclature are often confused, but have quite distinct meanings. Taxonomy is the science of classifying, describing and characterising different groups (taxa) of living organisms. Nomenclature, on the other hand, is about giving names to those different entities or groups. Scientific names of plants follow internationally agreed rules which are published as their respective “Codes of Nomenclature”. These rules are largely the same for the different groups of organisms. Each scientific name is tied to a type specimen and are essentially ‘binomials’ consisting of the name of a genus (*Arabidopsis*) followed by the name of the species (*thaliana*). This system of naming plants and animals has remained largely unchanged since Linnaeus developed it in the mid 18th

Century. The convention is that scientific names are written in italics with an initial upper-case letter for the genus and all lower case letters for the species name (*Arabidopsis thaliana*). A genus name should be spelt out in full the first time it is used and then may be abbreviated to an initial letter and full stop when it is unambiguous to do so. For example, *Arabidopsis thaliana* may be abbreviated to *A. thaliana*. The authors of a species name may be included, but more often than not. For example, *Arabidopsis thaliana* (L.) Heynh. is called by this name since 1842. The *thaliana* honors Johannes Thal, a XVI century German physician who described the plant in 1577; but Linnaeus used other names for this plant.”

8. “George P. Rédei could have been one of the pioneers who introduced the generalized term into the literature.”

9. “... *Arabidopsis thaliana* and other species of *Arabidopsis* may be the best model system for basic research in the biology of all multicellular eukaryotes. *Arabidopsis* has been the organism of choice for many plant biochemists, physiologists, developmental biologists and geneticists for several decades, since then a great deal of knowledge has been gained about the biology of this flowering plant. We now have in hand the sequence of its genome. Ongoing research within the community has resulted in working knowledge of many of the biochemical, physiological, and developmental processes of *Arabidopsis*. The availability of a broad base of knowledge about *Arabidopsis* and the previously developed research invites scientists to establish new techniques, develop new approaches, and test new concepts in *Arabidopsis* prior to their application in other species. The novel technologies made available in this way not only continually increase the efficiency of research done in *Arabidopsis*, but expose researchers to the most up-to-date methods in plant research. Although *Arabidopsis* is closely related to economically important crop plants as turnip, cabbage, broccoli, and canola, *Arabidopsis* is not an economically important plant. Despite this, it has been the focus of intense genetic, biochemical and physiological study for over 40 years because of several traits that make it very desirable for laboratory study. All together, these traits make *Arabidopsis* an ideal model organism for biological research and the species of choice for a large and growing community of scientists studying complex, advanced multicellular organisms. Why *Arabidopsis*? Why not concentrate our research efforts and resources on a species that will actually provide food for our world or useful products for industrial uses? In order to make the strides necessary to increase crop production in a relatively short time, we have to be able to move forward quickly and spend the available human and financial resources as efficiently as possible. This is the advantage of a model system: an organism that is easily manipulated, genetically tractable, and about which much is already known. By studying the biology of *Arabidopsis*, the model plant, we can gain comprehensive knowledge of a complete plant. In the laboratory, *Arabidopsis* offers the ability to test hypotheses quickly and efficiently. With the knowledge we gain from the model plant thus established as a reference system, we can move forward with research and rapidly initiate improvements in plants of economic and cultural importance.”