Review

New Phytologist

Tansley insight

Archaic lineages broaden our view on the history of *Arabidopsis thaliana*

Andrea Fulgione and Angela M. Hancock

Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany

Author for correspondence:
Andrea Fulgione
Tel: +49 162 7706340
Email: fulgione@mpipz.mpg.de

Received: *28 February 2018*
Accepted: *25 April 2018*

Contents

New Phytologist (2018) **219:** 1194–1198
doi: 10.1111/nph.15244

Key words: Africa, *Arabidopsis thaliana*, demography, genome, natural variation, population history, relict, world-wide.

Summary

Natural variation in *Arabidopsis thaliana* has contributed to discoveries in diverse areas of plant biology. While *A. thaliana* has typically been considered a weed associated primarily with human-mediated environments, including agricultural and urban sites and railways, it has recently been shown that it is also native in remote natural areas, including high altitude sites in Eurasia and Africa, from the Atlas mountains in Morocco to the afro-alpine regions in Eastern and South Africa to Yunnan in China, the Himalayas and the Tibetan Plateau. This finding suggests that while *A. thaliana* has been extensively studied in Europe and Western Asia there are still many open questions about its population history, genotype–phenotype relationships and mechanisms of adaptation.

I. Introduction

For early Arabidopsis researchers, natural variation held a prominent role because it was recognized as a useful tool for understanding how perturbations at the biochemical level could lead to phenotypic change (Alonso-Blanco *et al.*, 2009; Weigel, 2012). Collections of diverse natural accessions assembled over the past 75 years have been a vital resource for molecular discovery and evolutionary analyses (Provart *et al.*, 2016). For example, our knowledge of the functional and ecological roles of many genetic loci is based on work done in natural accessions (e.g. for flowering time and seed dormancy; Alonso-Blanco *et al.*, 1998, 2003; Wilczek *et al.*, 2009; Burghardt *et al.*, 2016). Reconstructing the history of the species provides crucial information about how to effectively leverage natural variation to address biological questions. Knowing the history of migrations, adaptation to climate and other environmental factors informs choices of accessions for studying specific traits, as well as for mapping studies. For this reason, extensive work over the past 20 years has clarified the history of *Arabidopsis thaliana* in Europe, Western Asia and North America. However, the origin and early history of the species still remained unclear. Here, we review recent work on the population history of *A. thaliana* population history and discuss how novel data from relict lineages reshapes our view on the origin and history of the species.

II. Origin of the *A. thaliana* species

Arabidopsis thaliana differs from other *Arabidopsis* species in several respects. It has an annual life cycle, predominantly self-fertilizes its flowers, exhibits reduced floral morphology and has reduced genome size compared to other species in the genus. The reduced petal size and pollen number found in *A. thaliana*, combined with the re-orientation of the stamen towards the stigma, are likely downstream effects of the transition to predominant selfing. This combination of features is part of the 'selfing syndrome', which has been observed across many species that transitioned to predominate selfing (Sicard & Lenhard, 2011). Furthermore, in *A. thaliana* the haploid chromosome number is reduced from the ancestral set of eight to five, and genome size is reduced from greater than 200 Mbp to *c.* 120 Mbp, through major genomic rearrangements including chromosome fusions (Kuittinen *et al.*, 2004). This reduced genome size has been hypothesized to result from the shift to predominate selfing (Oyama *et al.*, 2008; Hu *et al.*, 2011), consistent both with empirical observations (Wright *et al.*, 2008) and theoretical predictions (Charlesworth, 1992). These genomic changes render *A. thaliana* incompatible with its closest relatives and thereby create a species boundary (Nasrallah *et al.*, 2000).

From studies that have examined the genetic basis of selfing in *A. thaliana* (Nasrallah *et al.*, 2002), we know that two genes in the S-locus region are responsible for self-incompatibility in outcrossing *Arabidopsis* species, where many highly divergent S-haplogroups are maintained by balancing selection. In *A. thaliana*, three S-haplogroups have made it through the speciation bottleneck, all of which include independent loss-of-function mutations (Vekemans *et al.*, 2014; Shimizu & Tsuchimatsu, 2015). These S-haplogroups are prevalent in different regions in Eurasia and they were never found to co-occur all together. To account for the independent loss-of-function mutations and the disjunct geographic distribution of the S-haplogroups, the transition to selfing in *A. thaliana* was hypothesised to have evolved multiple times independently in different geographic regions (Shimizu *et al.*, 2008; Boggs *et al.*, 2009). However, this is difficult to reconcile with the massive genome size reduction (and associated speciation event), which likely followed the transition to predominant selfing.

Possible locations of the original speciation event were originally proposed based on species distributions in other members of the family and genus or intraspecific variation. These hypotheses revolved around Central Asia (Berger, 1965; Price *et al.*, 1994) and Europe (Hoffmann, 2002), and more recently the Caucasus (Beck *et al.*, 2008), the Balkans and the Carpathian Mountains (Koch & Matschinger, 2007; Koch *et al.*, 2008).

III. The classic model of the history of *A. thaliana*

Numerous population genetic studies over the past 20 years identified some robust biogeographic signals. In particular, Sharbel *et al.* (2000), found signals of isolation by distance across the species range, identified the Iberian Peninsula and Asia as centers of diversity, and described an east-west clinal distribution in genetic variation. These patterns suggested post-glacial expansions from Pleistocene refugia in Asia and the Iberian Peninsula towards a secondary contact zone in Central Europe (Sharbel *et al.*, 2000). Subsequent studies with much more dense genetic data and larger sample sizes remain remarkably consistent with these early findings (Nordborg *et al.*, 2005; Schmid *et al.*, 2006; Beck *et al.*, 2008; François *et al.*, 2008; Picó *et al.*, 2008). More generally, a recent large-scale survey of Eurasian genomic variation provided hints that there may be other pockets of highly diverged accessions along the Southern edge of the Eurasian range (1001 Genomes Consortium, 2016).

These patterns of genomic variation suggested a general model of the history of *A. thaliana* in Eurasia (outlined in Fig. 1). The hypothesised origin of the species revolved around Eastern Europe, and in particular the Caucasus (Beck *et al.*, 2008). The spread of *A. thaliana* from this region was dated to *c.* 120 thousand years ago (ka). Several studies have also hypothesized a second wave of expansion from the Caucasus following human agriculture, *c.* 8 ka (François *et al.*, 2008; 1001 Genomes Consortium, 2016; Lee *et al.*, 2017). Based on a small set of ascertained SNPs, a shared variation between North-Western Africa and the Iberian Peninsula (Brennan *et al.*, 2014) suggested a recent colonization event from Iberia. The rest of the African continent and Eastern Asia were thought to be outside the native range of the species (Hoffmann, 2002), and the few accessions found there, to have been recently introduced by human migrations.

IV. New genomic data from outside Eurasia challenge our view of *A. thaliana* history

In spite of extensive efforts to survey natural variation in *A. thaliana* over the years, beginning with Laibach (1943), there has historically been a strong Eurocentric focus in *A. thaliana* sampling and population genetic analysis, which reflects the accepted native range of the species: temperate regions in Eurasia (Hoffmann, 2002). Moreover, genotyping studies have often relied on Eurasian and North American accessions for variant discovery, so that non-European variation was under-represented, even in studies where these samples were included. These two factors limited the scope of the conclusions that could be drawn. However, several recent studies examined complete genome sequences of accessions from outside Eurasia, including Africa (Durvasula *et al.*, 2017), Madeira

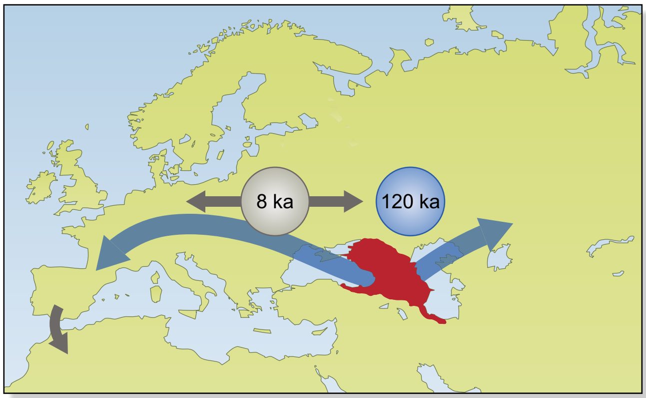

Fig. 1 The classic model of the history of *Arabidopsis thaliana*. In red, the hypothesised origin of the species. In blue, the expansion in Eurasia estimated at *c.* 120 thousand years ago (ka) (Beck *et al.*, 2008). In grey, the hypothesised recent expansion following human agriculture *c.* 8 ka (1001 Genomes Consortium, 2016; Lee *et al.*, 2017), and the colonization of Morocco from the Iberian peninsula.

(Fulgione *et al.*, 2018) and Central Asia (Zeng *et al.*, 2017; Zou *et al.*, 2017) and compared them to results from across Eurasia (1001 Genomes Consortium, 2016). These studies reveal several surprises and helped to reconcile previous inconsistencies, as outlined below.

Deep divergence among major continental groups

Contrary to the idea that the center of *A. thaliana* diversity lies within Eurasia, the joint analysis of variation across continents revealed evidence for deep divergence among several relict lineages (1001 Genomes Consortium, 2016; Durvasula *et al.*, 2017; Fulgione *et al.*, 2018; Zeng *et al.*, 2017; Zou *et al.*, 2017). Particularly surprising was the finding that all African populations, including South Africans and Tanzanians, are native and that all sampled African individuals represent relicts when compared to the major Eurasian clade (Durvasula *et al.*, 2017). Other notable findings are that the island of Madeira harbors the most extremely diverged individuals sampled to date (Fulgione *et al.*, 2018) and that diverged samples from Central Asia (Tibet, Zeng *et al.*, 2017; Yunnan, Zou *et al.*, 2017; Western Himalayas, Tyagi *et al.*, 2015; the Altai Mountains, Yin *et al.*, 2010) suggest additional refugia.

African populations best represent the early history of the species

Genome-wide diversity levels are higher in Africa than elsewhere, and Africa harbours extensive genetic variation that is not represented in sequenced Eurasian accessions (Durvasula *et al.*, 2017). Furthermore, historical population sizes in Africa appear to have been larger than those in Eurasia, suggesting an ancient history of the species on this continent. Considering that high levels of private variation and ancestral population sizes have been used to infer the origin of the species in previous studies (Beck *et al.*, 2008), and in other species (e.g. in humans), these results are consistent with an African origin of *A. thaliana*.

The origin of the *A. thaliana* lineage is likely tied to the evolution of selfing. The conundrum in the classic view of this transition was that the disjunct distribution of S-haplogroups in Eurasia suggested multiple independent events. However, the genomic shrinkage and rearrangements, which likely followed the transition to selfing and created a species boundary, are shared species-wide. African data help solve this enigma. All three S-locus haplogroups as well as multiple novel haplotypes are found together in the Atlas Mountains in Morocco for the first time in the world-wide sample (Durvasula *et al.*, 2017). The regional hypothesis for the evolution of selfing now appears unnecessary; rather, the transition to predominant selfing likely happened in an isolated population best represented today within Morocco.

Population splits roughly correspond to major climatic events

Data from new relict samples provide additional information about the timing of population separations and expansions. The deepest population splits appear to have occurred between East and West Africa and the Levant (a region roughly comprising North-eastern Africa and the northwest of the Arabian plate) at *c.* 90 ka (Durvasula *et al.*, 2017). This timing corresponds to the end of a major pluvial period in Africa and interglacials across the Northern Hemisphere. In Eurasia there were likely at least two major expansions, the most ancient of which is represented by pockets of relicts in Iberia, Italy (1001 Genomes Consortium, 2016; Günther *et al.*, 2016) and Central Asia (Zeng *et al.*, 2017; Zou *et al.*, 2017). Relicts in Western Europe may represent a separate migration from Africa (possibly through Gibraltar (Fulgione *et al.*, 2018)) from that which populated Central Asia (likely through the Levant). More recent splits between South Africa and Tanzania and between the currently widespread European and Asian clades are inferred at around 45 ka based on joint analyses of Eurasian, African and East Asian data sets (Durvasula *et al.*, 2017; Zou *et al.*, 2018). This suggests that non-relicts spread through Eurasia well before the current interglacial, but a third post-glacial expansion, as hypothesized in Lee *et al.* (2017), is also possible.

An updated model of *A. thaliana* population history

Taken together, these new data suggest several revisions to the classic model of *A. thaliana* history (outlined in Fig. 2). In this model, a subpopulation became geographically isolated from the outcrossing ancestor by 1.2–0.8 Ma. This period corresponds to the Middle Pleistocene Transition, a shift to more arid climates,

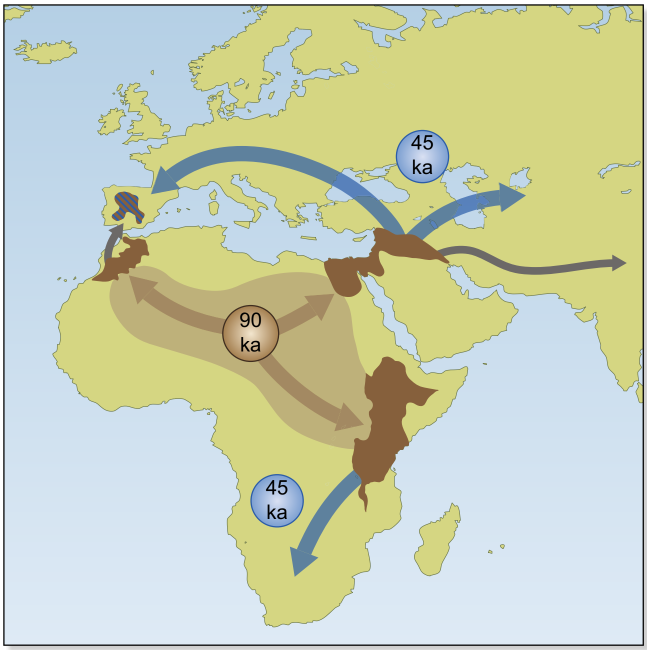

Fig. 2 The updated model of the history of *Arabidopsis thaliana*. The faint brown area represents the wide spread of *A. thaliana* in Africa during the Abassia pluvial (120–90 thousand years ago (ka)). In brown, the three main clades that split *c.* 90 ka (Durvasula *et al.*, 2017). In grey, the expansion of relict lineages in Eurasia (Iberia and Central Asia). Timing of these events have not been inferred to date. In blue, the migration from East- into South-Africa and the expansion of the currently widespread Eurasian clade, both dated at *c.* 45 ka (Durvasula *et al.*, 2017; Zou *et al.*, 2017). This expanding lineage mixed with archaics in Iberia, resulting in the admixed Iberian relicts (Fulgione *et al.*, 2018), represented as brown and blue stripes.

and more open habitats in Africa (woodlands to grasslands) (Cerling & Hay, 1986; deMenocal, 2004). This range expansion and consequent bottlenecks may have predisposed the marginal population toward the evolution of selfing, as exemplified in North American *Arabidopsis lyrata* (Mable *et al.*, 2017). Consistent with previous timing estimates of this event (Bechsgaard *et al.*, 2006; Tang *et al.*, 2007), at *c.* 1 Ma all *A. thaliana* populations show evidence for an ancient severe bottleneck. Based on current sampling and S-locus diversity, the location of the speciation event is best represented within Morocco, but may have occurred elsewhere, e.g. within the Levant, where current sampling is extremely limited. These regions are distant from the center of diversity of other *Arabidopsis* species (Koch & Matschinger, 2007; Koch *et al.*, 2008), consistent with the origin of *A. thaliana* at the edge of a range expansion.

There is evidence that *A. thaliana* was widespread across Africa during the Abassia pluvial (an extended wet and rainy pluvial period in North-Africa that lasted from *c.* 120 to 90 ka) and the last interglacial in Eurasia (*c.* 120–90 ka). During this period, the Sahara was scattered with lakes and wet corridors, and vegetation was abundant across Africa (Dupont, 2011) and Eurasia (Kukla *et al.*, 2002; Kaspar *et al.*, 2005). This may also have been when the ancestors of current 'Iberian relicts' migrated out of Morocco (Fulgione *et al.*, 2018), and relicts currently in Central Asia reached this region. The end of this humid period *c.* 90 ka led to a general loss of migration corridors and population subdivision in Africa and the Levant. This also corresponds to a time when modern humans are thought to have first migrated out of Africa (Osborne *et al.*, 2008; Timmermann & Friedrich, 2016).

More recently, at *c.* 45 ka, there is evidence that the currently widespread Eurasian clade expanded from the Caucasus, resulting in divergence among European and Asian sub-clusters (Durvasula *et al.*, 2017; Zou *et al.*, 2017). Although evidence of an interglacial period at this time is more controversial, analogous migration patterns are well-supported in human populations (Voight *et al.*, 2005; Henn *et al.*, 2012; Mallick *et al.*, 2016). This is also consistent with evidence that the Caucasian clade mixed with the African-derived ancestor of the current Iberian relicts 20–40 ka (Fulgione *et al.*, 2018).

V. Conclusions

Although many investigations over the past two decades have assayed and used natural variation in *A. thaliana*, recent work revealed that these studies focused on only a fraction of the total variation in the species. Just when we thought we had gained a comprehensive survey of variation in the species, we are finding that there is immense variation left to explore. To address this, additional sampling is needed. The African continent itself remains largely unexplored, especially across east and south Africa. Additional sampling in the Levant and Arabia will be crucial to resolve details of the timing and direction of gene flow. Further, relicts found in Italy, Central Asia, and offshore African islands suggest there may be several more unexplored refugia, which will open the way for future studies into adaptation to climate extremes. Moreover, the finding that historical migrations of *A. thaliana* populations were associated with climatic events suggests that future climate change could influence the distribution and adaptive potential of this species.

One of the most enticing aspects of recent findings is the enormous amount of untapped genetic and phenotypic diversity. Just as natural variation in *A. thaliana* has served as an important tool for biological discovery in the past, this new variation provides the raw material for novel research questions and findings in the future.

Acknowledgements

We thank Maarten Koornneef for useful comments on a previous version of the manuscript. This work was supported by the Max Planck Gesselschaft and ERC Grant CVI_ADAPT 638810 (A.M.H.).

References

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D, Koornneef M. 2009. What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell* **21**: 1877–1896.

Alonso-Blanco C, Bentsink L, Hanhart CJ, Blankestijn-de Vries H, Koornneef M. 2003. Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics* **164**: 711–729.

Alonso-Blanco C, El-Assal SE, Coupland G, Koornneef M. 1998. Analysis of natural allelic variation at owering time loci in the Landsberg erecta and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. *Genetics* **149**: 749–764.

Bechsgaard JS, Castric V, Charlesworth D, Vekemans X, Schierup MH. 2006. The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr. *Molecular Biology and Evolution* **23**: 1741–1750.

Beck JB, Schmuths H, Schaal BA. 2008. Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reacts Pleistocene glacial dynamics. *Molecular Ecology* **17**: 902–915.

Berger B. 1965. The taxonomic confusion within Arabidopsis and allied genera. In: Robbelen G, ed. *Arabidopsis research, report of an international symposium*. Gottingen, Germany: University of Göttingen, 19–25.

Boggs NA, Nasrallah JB, Nasrallah ME. 2009. Independent S-locus mutations caused self-fertility in *Arabidopsis thaliana*. *PLoS Genetics* **5**: e1000426.

Brennan AC, Mendez-Vigo B, Haddioui A, Martinez-Zapater JM, Pico FX, Alonso-Blanco C. 2014. The genetic structure of *Arabidopsis thaliana* in the south-western Mediterranean range reveals a shared history between North Africa and southern Europe. *BMC Plant Biology* **14**: 17.

Burghardt LT, Edwards BR, Donohue K. 2016. Multiple paths to similar germination behavior in *Arabidopsis thaliana*. *New Phytologist* **209**: 1301–1312.

Cerling TE, Hay RL. 1986. An isotopic study of paleosol carbonates from Olduvai Gorge. *Quaternary Research* **25**: 63–78.

Charlesworth B. 1992. Evolutionary rates in partially self-fertilizing species. *American Naturalist* **140**: 126–148.

Dupont L. 2011. Orbital scale vegetation change in Africa. *Quaternary Science Reviews* **30**: 3589–3602.

Durvasula A, Fulgione A, Gutaker RM, Alacakaptan SI, Flood PJ, Neto C, Tsuchimatsu T, Burbano HA, Pico FX, Alonso-Blanco C *et al.* 2017. African genomes illuminate the early history and transition to selfing in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **114**: 5213–5218.

François O, Blum MG, Jakobsson M, Rosenberg NA. 2008. Demographic history of European populations of *Arabidopsis thaliana*. *PLoS Genetics* **300**: e1000075.

Fulgione A, Koornneef M, Roux F, Hermisson J, Hancock A. 2018. Madeiran *Arabidopsis thaliana* reveals ancient long-range colonization and clarifies demography in Eurasia. *Molecular Biology and Evolution* **35**: 564–574.

1001 Genomes Consortium. 2016. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**: 481–491.

Günther T, Lampei C, Barilar I, Schmid KJ. 2016. Genomic and phenotypic differentiation of *Arabidopsis thaliana* along altitudinal gradients in the North Italian Alps. *Molecular Ecology* **25**: 3574–3592.

Henn BM, Cavalli-Sforza LL, Feldman MW. 2012. The great human expansion. *Proceedings of the National Academy of Sciences, USA* **109**: 17758–17764.

Hoffmann MH. 2002. Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). *Journal of Biogeography* **29**: 125–134.

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H *et al.* 2011. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics* **43**: 476–481.

Kaspar F, Kühl N, Cubasch U, Litt T. 2005. A model-data comparison of European temperatures in the Eemian interglacial. *Geophysical Research Letters* **32**: L11703.

Koch MA, Matschinger M. 2007. Evolution and genetic differentiation among relatives of *Arabidopsis thaliana. Proceedings of the National Academy of Sciences, USA* **104**: 6272–6277.

Koch MA, Wernisch M, Schmickl R. 2008. *Arabidopsis thaliana*'s wild relatives: an updated overview on systematics, taxonomy and evolution. *Taxon* **57**: 933–943.

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley CH, Savolainen O. 2004. Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana. Genetics* **168**: 1575–1584.

Kukla GJ, Bender ML, Beaulieu J-Ld, Bond G, Broecker WS, Cleveringa P, Gavin JE, Herbert TD, Imbrie J, Jouzel J *et al.* 2002. Last interglacial climates. *Quaternary Research* **58**: 2–13.

Laibach F. 1943. *Arabidopsis thaliana* (L.) Heynh. als Object für genetischeund entwicklungs-physiologische Untersuchungen. *Botanisches Archiv* **44**: 439–455.

Lee CR, Svardal H, Farlow A, Exposito-Alonso M, Ding W, Novikova P, Alonso-Blanco C, Weigel D, Nordborg M. 2017. On the post-glacial spread of human commensal *Arabidopsis thaliana. Nature Communications* **8**: 14458.

Mable BK, Hagmann J, Kim ST, Adam A, Kilbride E, Weigel D, Stift M. 2017. What causes mating system shifts in plants? *Arabidopsis lyrata* as a case study. *Heredity (Edinburgh)* **118**: 110.

Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A *et al.* 2016. The Simons Genome Diversity Project: 300 genomes from diverse populations. *Nature* **538**: 201–206.

deMenocal PB. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220**: 3–24.

Nasrallah ME, Liu P, Nasrallah JB. 2002. Generation of self-incompatible *Arabidopsis thaliana* by transfer of two S locus genes from *A. lyrata. Science* **297**: 247–249.

Nasrallah ME, Yogeeswaran K, Snyder S, Nasrallah JB. 2000. Arabidopsis species hybrids in the study of species differences and evolution of amphiploidy in plants. *Plant Physiology* **124**: 1605–1614.

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R *et al.* 2005. The pattern of polymorphism in *Arabidopsis thaliana. PLoS Biology* **3**: e196.

Osborne AH, Vance D, Rohling EJ, Barton N, Rogerson M, Fello N. 2008. A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. *Proceedings of the National Academy of Sciences, USA* **105**: 16444–16447.

Oyama RK, Clauss MJ, Formanová N, Kroymann J, Schmid KJ, Vogel H, Weniger K, Windsor AJ, Mitchell-Olds T. 2008. The shrunken genome of *Arabidopsis thaliana. Plant Systematics and Evolution* **273**: 257–271.

Picó FX, Mendez-Vigo B, Martinez-Zapater JM, Alonso-Blanco C. 2008. Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian peninsula. *Genetics* **180**: 1009–1021.

Price RA, Palmer JD, Al-Shehbaz IA. 1994. Systematic relationships of Arabidopsis: a molecular and morphological perspective. In: Somerville C, Meyerowitz E, eds. *Arabidopsis*. Cold Spring Harbour, NY, USA: Cold Spring Harbour Laboratory Press, 7–19.

Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S *et al.* 2016. 50 years of Arabidopsis research: highlights and future directions. *New Phytologist* **209**: 921–944.

Schmid KJ, Torjek O, Meyer R, Schmuths H, Hoffmann MH, Altmann T. 2006. Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide single nucleotide polymorphism markers. *Theoretical and Applied Genetics* **112**: 1104–1114.

Sharbel TF, Haubold B, Mitchell-Olds T. 2000. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Molecular Ecology* **9**: 2109–2118.

Shimizu KK, Shimizu-Inatsugi R, Tsuchimatsu T, Purugganan MD. 2008. Independent origins of self-compatibility in *Arabidopsis thaliana. Molecular Ecology* **17**: 704–714.

Shimizu KK, Tsuchimatsu T. 2015. Evolution of selfing: recurrent patterns in molecular adaptation. *Annual Review of Ecology Evolution and Systematics* **46**: 593–622.

Sicard A, Lenhard M. 2011. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany* **107**: 1433–1443.

Tang C, Toomajian C, Sherman-Broyles S, Plagnol V, Guo YL, Hu TT, Clark RM, Nasrallah JB, Weigel D, Nordborg M. 2007. The evolution of selfing in *Arabidopsis thaliana. Science* **317**: 1070–1072.

Timmermann A, Friedrich T. 2016. Late Pleistocene climate drivers of early human migration. *Nature* **538**: 92–95.

Tyagi A, Singh S, Mishra P, Singh A, Tripathi AM, Jena SN, Roy S. 2015. Genetic diversity and population structure of *Arabidopsis thaliana* along an altitudinal gradient. *AoB Plants* **8**: plv145.

Vekemans X, Poux C, Goubet PM, Castric V. 2014. The evolution of selfing from outcrossing ancestors in Brassicaceae: what have we learned from variation at the S-locus? *Journal of Evolutionary Biology* **27**: 1372–1385.

Voight BF, Adams AM, Frisse LA, Qian Y, Hudson RR, Di Rienzo A. 2005. Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. *Proceedings of the National Academy of Sciences, USA* **102**: 18508–18513.

Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology* **158**: 2–22.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J *et al.* 2009. Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**: 930–934.

Wright SI, Ness RW, Foxe JP, Barrett SCH. 2008. Genomic consequences of outcrossing and selfing in plants. *International Journal of Plant Sciences* **169**: 105–118.

Yin P, Kang J, He F, Qu LJ, Gu H. 2010. The origin of populations of *Arabidopsis thaliana* in China, based on the chloroplast DNA sequences. *BMC Plant Biology* **10**: 22.

Zeng L, Gu Z, Xu M, Zhao N, Zhu W, Yonezawa T, Liu T, Qiong L, Tersing T, Xu L *et al.* 2017. Discovery of a high-altitude ecotype and ancient lineage of *Arabidopsis thaliana* from Tibet. *Science Bulletin* **62**: 1628.

Zou YP, Hou XH, Wu Q, Chen JF, Li ZW, Han TS, Niu XM, Yang L, Xu YC, Zhang J *et al.* 2017. Adaptation of *Arabidopsis thaliana* to the Yangtze River basin. *Genome Biology* **18**: 239.