Enhanced resistance to bacterial and oomycete pathogens by short tandem target mimic RNAs in tomato

Alex Canto-Pastor¹,², Bruno A. M. C. Santos³,², Adrian A. Valli³,², William Summers³, Sebastian Schornack¹,², and David C. Baulcombe¹,²

¹Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom; and ²Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, United Kingdom

Edited by Xuemei Chen, University of California, Riverside, CA, and approved December 20, 2018 (received for review August 22, 2018)

Nucleotide binding site leucine-rich repeat (NLR) proteins of the plant innate immune system are negatively regulated by the miR482/2118 family miRNAs that are in a distinct 22-nt class of miRNAs with a double mode of action. First, they cleave the target RNA, as with the canonical 21-nt miRNAs, and second, they trigger secondary siRNA production using the target RNA as a template. Here, we address the extent to which the miR482/2118 family affects expression of NLR mRNAs and disease resistance. We show that structural differences of miR482/2118 family members in tomato (Solanum lycopersicum) are functionally significant. The predicted target of the miR482 subfamily is a conserved motif in multiple NLR mRNAs, whereas for miR2118b, it is a noncoding RNA target formed by rearrangement of several different NLR genes. From RNA sequencing and degradome data in lines expressing short tandem target mimic (STTM) RNAs of miR482/2118, we confirm the different targets of these miRNAs. The effect on NLR mRNA accumulation is slight, but nevertheless, the tomato STTM lines display enhanced resistance to infection with the oomycete and bacterial pathogens. These data implicate an RNA cascade of miRNAs and secondary siRNAs in the regulation of NLR RNAs and show that the encoded NLR proteins have a role in quantitative disease resistance in addition to dominant gene resistance that has been well characterized elsewhere. We also illustrate the use of STTM RNA in a biotechnological approach for enhancing quantitative disease resistance in highly bred cultivars.

Nucleotide binding site leucine-rich repeat (NLR) proteins of plants are central components of the innate immune system that protects against pests and pathogens. These proteins are encoded in multigene families and they regulate signal transduction pathways leading to disease resistance (1). There are several negative regulators of these resistance pathways that, presumably, reduce the likelihood that disease resistance is activated in the absence of a pest or pathogen. Such pathogen-independence induction could be damaging to the host because the disease resistance mechanisms may be associated with programmed cell death called hypersensitive response, changes to the cell wall, production of active oxygen species, and strong activation of pathogenesis-related genes that compromise reproductive fitness. Transgenic lines overexpressing NLRs (2–4) and gain-of-function mutations, like suppressor of npr1-1, constitutive 1 (5) or suppressor of salicylic acid insensitive 4 (6), illustrate how inappropriate activation of disease resistance can damage the plant. In normal conditions the negative regulators may mitigate the potential cost of NLR gene diversification and they could be particularly beneficial to plants with many NLR genes (7).

The importance of these negative regulators is reflected in their diversity. There are, for example, suppressor proteins affecting either the NLR proteins themselves or components of the downstream signal transduction pathways (8, 9). There are also miRNAs and siRNAs that function as negative regulators of NLR mRNA (10–13). These small RNAs (sRNAs) bind to their target mRNA by Watson–Crick base pairing, and they silence the expression of the mRNA-encoded protein through various mechanisms affecting RNA stability or translation in which the effectors are proteins of the Argonaute family (14).

Among the miRNA regulators of NLRs (7), the miR482/2118 family is the most diverse. This family is present in modern seed plants, although, in some instances, the apparent conservation may reflect repeated rounds of convergent evolution (15). The miR482/2118 family members are all 22 nt rather than the more usual 21 nt in length (10–12), and the additional nucleotide is functionally significant because it influences the fate of the targeted miRNA: targets of 21-nt miRNA are simply degraded by the Argonaute nuclease activity (14), whereas 22-nt miRNA targets are converted into a dsRNA by the RNA-dependent RNA polymerase 6. The dsRNA is then cleaved by a DCL protein to generate an array of 21-nt secondary siRNAs that may be phased

Significance

Plants, like animals, have complex disease resistance systems in which receptors interact directly or indirectly with effectors of disease produced by pests and pathogens. To minimize the fitness cost of these systems to the plant, there are miRNAs that target the mRNAs of a family of receptor proteins required for disease resistance. Target site mimics of these miRNAs confer enhanced quantitative resistance in tomato against an oomycete and a bacterium. These findings are consistent with a role of the receptor proteins in quantitative disease resistance and show how blocking these miRNAs could be a useful approach in crop protection.

Author contributions: A.C.-P., S.S., and D.C.B. designed research; A.C.-P. and B.A.M.C.S. performed research; A.A.V. and W.S. contributed new reagents/analytic tools; A.C.-P. analyzed data; A.A.V. assisted with writing and analyses; and A.C.-P., S.S., and D.C.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the NCBI BioProject (accession no. PRJNAS505207).

¹Present address: Department of Plant Biology and Genome Center, University of California, Davis, CA 95616.
²Present address: Department of Crop Biotechnology, National Institute of Agricultural Botany, Cambridge CB3 0LE, United Kingdom.
³Present address: Department of Plant Molecular Genetics, Spanish National Center for Biotechnology, Madrid 28049, Spain.
⁴To whom correspondence should be addressed. Email: dcb40@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1814380116/-/DCSupplemental.

Published online January 24, 2019.
with respect to the binding site of the miRNA (16, 17). As a result of this process there is the potential for 22-nt miRNAs to establish regulatory cascades in which mRNAs are targeted by both primary miRNAs and secondary sRNAs.

Here we focus on the miR482/2118 family in tomato and we set out to establish the extent to which they influence NLR RNAs and disease resistance. Our approach involved the use of short tandem target mimic (STTM) RNAs that would inactivate the miR482/2118 family members and prevent primary or secondary regulation of NLR mRNAs. Our findings confirm that, although the miR482 and miR2118 subfamilies have similar sequences, they are functionally distinct in tomato. The miR482 subfamily targets NLR mRNAs as described (10–12) and triggers secondary siRNA production. It is possible, therefore, that NLR RNAs could be both primary and secondary targets of the miR482 family. One isoform of miR2118—miR2118b—is likely to influence NLRs through secondary sRNAs. Its main primary target is not an NLR but a long noncoding RNA—TAS5—from which secondary sRNAs are produced that could target NLR coding sequence RNAs. This primary and secondary negative regulation of NLR RNAs by miR482/2118b has an effect on quantitative disease resistance because the STTM lines are less susceptible to an oomycete and a bacterial pathogen than control lines. This enhanced resistance was achieved without a large effect on growth and development of the plants, and it may be useful to protect tomato and other species against pests and diseases.

Results
Revisiting the miR482/2118 Family and Their Targets in Tomato.
There are numerous families of miRNAs that negatively regulate defense by targeting conserved motifs in NLR RNAs. The miR482/2118 family is the most extensive of these families and it is present in most lineages of seed plants due to either conservatism or convergent evolution (7). This analysis focused on the five members of this family (SI Appendix, Table S1) that (i) align to the tomato genome and (ii) feature in our sRNA datasets from leaves of 1-mo-old tomato plants.

The sequences of these five miRNAs (Fig. 1A) are complementary to the RNA representation of the conserved P-loop motif (GMGGVGKT) in NLR proteins with sequence variation at 5/6 variable sites corresponding to wobble positions. This pattern indicates that this small family of miRNAs could target a larger number of NLR RNAs with synonymous coding sequence variation (12). Using refined prediction algorithms (18) we estimate that individual miR482/2118 species could potentially target between 15 and 55 NLR RNAs (Fig. 1B and SI Appendix, Table S2).

The miR482/2118 family members also have the potential to trigger secondary siRNA production on their mRNA targets (16) because they are 22 nt in length (10–12); correspondingly, there were 32 NLR genomic regions with 10 or more 21- to 22-nt siRNA normalized reads per million (nRPM) mapping (Fig. 1C) from our sRNA datasets. Other studies in different species refer to these NLRs with overlapping sRNAs as phas-siRNAs (10, 15, 19), although the extent of phasing may be low. We use the term sNL here referring to NLR mRNAs with secondary sRNAs that may or may not be phased. Most of these tomato sNLS in our datasets are predicted targets of miR482s with 30 of 32 of them encoding nucleotide binding site leucine-rich repeats with coiled coil domains (CNLs) rather than nucleotide binding site leucine-rich repeat with Toll/interleukin-1 receptor homology domains (TNLs) at the amino terminus (Fig. 1B and SI Appendix, Table S3).

Compared with miR482 predictions, potential targets of miR2118s were less enriched for NLRs (Fig. 1B and SI Appendix, Table S2) or secondary sRNAs. One of the two isoforms miR2118b—like its close homologue in other species (20, 21)—was much less abundant in leaf tissues than in flower and fruit (Fig. 1D) and it may not have a large effect on NLR mRNAs. In contrast, miR2118b, like miR482 isoforms, was abundant in leaves and reproductive tissues (Fig. 1D) (12) and it can potentially target several RNAs, including seven sNLS (Fig. 1B and SI Appendix, Table S2) and an RNA species that had been identified previously as TAS5 (22).

Our analysis reveals that TAS5 is atypical of NLR RNAs, because it has sequence similarity to both CNL and TNL types of NLR RNA on both the sense and antisense orientations (SI Appendix, Fig. S1). It is unlikely to be translated into a functional protein [calculated coding potential score of −1.146 (23, 24)], and a more likely interpretation is that TAS5 is a long noncoding RNA composed of rearranged and degenerated sequences from multiple NLR genes (SI Appendix, Fig. S1). It has two miR2118b target sites that were detected bioinformatically and

---

Fig. 1. The miR482/2118 family in tomato. (A) Nucleotide sequence alignment of mature miR482/2118 members in tomato. The consensus sequence of each position in the alignment vs. the P-loop motif is shown at the bottom. (B) Dot plot representing the sum of 21-nt sRNA nRPM aligning to an individual NLR. A total of 32 NLRs presented >10 nRPM counts and were defined as sNLS (purple). TAS5 (red) is added as a reference. (C) Summary of target prediction of all miR482/2118 members. (D) RNA gel blot analysis of tomato miR482/2118 members in various tissues of plant development. Column 4 shows the same blot hybridized with U6 as a loading control. Anth, anthesis; EI, early influorescence; GF, green fruit; LI, late influorescence.
by degradome analysis (SI Appendix, Fig. S2 and Tables S2 and S4), and the siRNAs are predominantly in a phased register. This phasing pattern starts at or adjacent to the more 5′ miR2118b target sites (SI Appendix, Fig. S2).

These findings indicate that there are both structural and functional differences between the miR482 and miR2118 subfamilies. The miR482 subfamily (12) targets the CNL NLRs directly and may also have an indirect effect on NLR RNAs via secondary siRNAs. The miR2118 subfamily, in contrast, is different in that its alignment with the P-loop coding motif is shifted by two nucleotides and is only associated with secondary siRNAs (Fig. 1C) in a few instances. The miR2118a variant is likely to have little if any effect on NLRs. The miR2118b, in contrast, may act via TASS secondary siRNAs on both TNL- and CNL-type NLRs.

Target Mimics Block Production of Secondary NLR siRNAs. To further investigate the function of miRNAs from the miR482/2118 family in tomato, we generated STTM transgenes. They were expressed from the strong 35S promoter and were designed so that their transcripts would inactivate the miRNAs 482 or 2118b. We did not test miR2118a, as it is expressed predominantly in reproductive tissue (Fig. 1D). The design of the STTM constructs was based on the IPS-1 natural target mimic of miR399 (25) with two tandem target sites (Fig. 2A) to enhance efficiency (26). These target mimic sequences had a 3′-nt bulge in the miRNA/target RNA duplex structure that would prevent cleavage of the target mimic RNA and channel the miRNAs away from their natural targets.

Lines transformed with the STTM482 and STTM2118b constructs were screened for transgene expression and effects on target miRNAs. Of 16 stable lines (8 per construct), one-half showed reduced levels of the corresponding miRNAs (SI Appendix, Fig. S3). From these lines, for subsequent detailed analysis, the top inactivating lines for STTM482 (lines 1 and 3 in SI Appendix, Fig. S3) and STTM2118b (lines 5 and 7 in SI Appendix, Fig. S3) were carried to the further generations. In 1-mo-old T2 plants, high-throughput sequence analysis of miRNAs in STTM482.1 revealed a specific 10-fold reduction over the WT in all miR482 subfamily members, whereas in STTM2118b.5 the effect was exclusively on miR2118b (Fig. 2B and SI Appendix, Fig. S4). All other miRNAs were similarly abundant in the STTM lines and WT plants. The effect of the STTMs, therefore, was specific for the cognate miRNA.

There was an additional effect of the STTMs on sRNAs other than miRNAs. To detect this effect, we aligned sRNA reads to the tomato genome and identified differential small RNA loci (DSLs) using the segmentSeq and baySeq packages (27). DSLs in STTM482.1 were enriched for sRNAs, and most of them (10 of 14) overlapped with predicted targets of miR482 members, whereas in STTM2118b.5, the only DSL was TASS (SI Appendix, Table S6). The siRNA levels at 14 DSL sRNAs were at least twofold higher in the WT than in STTM482.1 (SI Appendix, Table S7). At LRR1 (Solyc02g036270) and LRR2 (Solyc04g005540), two well-studied CNL RNAs targeted by miR482 (12), there were abundant siRNAs aligned to the 3′-side of the miRNA cleavage sites (Fig. 3B) that were fourfold reduced in the STTM482.1 but not in STTM2118b.5. Conversely, there was a fivefold reduction in TASS-derived siRNAs in STTM2118b.5 but not in STTM482.1 (Fig. 3B and SI Appendix, Table S7).

Target Mimic Effects on NLR mRNA Accumulation. We further investigated the effect of miR482/2118 species on the NLR mRNAs using RNAseq data of WT and T2 STTM lines. We used young and old leaves and detached leaves inoculated with zoospore droplets of Phytophthora infestans 88069. The analysis of these datasets revealed that the overall effects of the STTM RNAAs were small in all of the conditions tested. The expressed sRNAs were more affected than other NLRs or other genes, but only in the infected plants was the trend toward elevated expression (Fig. 4A). The previously described miR482 targets LRR1 and LRR2 and a few other sRNAs and NLR RNAs showed consistent small up-regulation (1.5- to 2.5-fold changes) in all conditions in STTM482 but not STTM2118b lines, and TASS was consistently up-regulated in STTM2118b (SI Appendix, Table S7) but not in the STTM482 lines. These observations were validated using qRT-PCR (SI Appendix, Fig. S5).

To find out whether the small up-regulation effect was due directly to reduced targeting by the respective miRNAs, we focused on the P. infestans-inoculated samples and we compared the targeting score of either miR482 or miR2118b with the degree of overexpression for each of the SNLs and expressed NLRs (SI Appendix, Table S2). There was no correlation (Fig. 4B); we conclude from these data (Figs. 3 and 4) that the miR482/2118 species are regulators of NLR expression, but to account for the lack of a simple correlation of targeting score and mRNA accumulation, there may be additional mechanisms involved as discussed below.

Inactivation of miR482 and miR2118b Enhances Resistance to P. infestans and Pseudomonas syringae. To investigate the effect of miR482/2118 on disease resistance, we inoculated detached leaves of T2 STTM lines and WT plants with zoospore droplets of the P. infestans 88069. The tomato cultivar M82 (LA3475), from which all of the plants in this study are derived, is highly susceptible to P. infestans (28). We measured the size of necrotic lesions to measure progression of the disease, and according to this criterion, the STTM482.1 and STTM2118b.5 were less susceptible than the nontransgenic control plants with significantly smaller lesion size at 3 dpi in both T2 STTM482.1 and STTM2118b.5 than WT and STTM171.1 lines.
A Canto-Pastor et al. vs. the WT. Dotted (blue) and TAS5 STTM482.1 In this paper, we describe definitive evidence that was added as a reference. In all instances, RNA abundance (SI). One exception is a predicted target of miR2118b locus miR482- and miR2118b-mediated silencing is relieved in STTM482 P. infestans (32) in a separate study, SI Appendix LRR2 Verticillium and TAS5 P. infestans TAS5 and A and SI Appendix Fusarium oxysporium loci. Positions Ps. syringae (30) and its encoded B www.pnas.org/cgi/doi/10.1073/pnas.1814380116

Overall NLR expression is increased in STTM482 and STTM2118b (Fig. 5C and SI Appendix, Fig. S4). This effect indicates that the primary or secondary targets of miR482s and miR2118b have a role in antibacterial immunity.

Defense activation may be at the expense of the plant’s fitness (29). In the STTM lines, however, general growth was similar in the control and STTM lines used in the disease resistance tests other than a small decrease in shoot length in STTM482.1 (SI Appendix, Fig. S6). These tests may not be sensitive to other slight differences in fitness but they indicate that enhanced disease resistance in the STTM lines was not associated with gross effects on growth and development.

Discussion miR482/2118b as Regulators of NLRs and Quantitative Disease Resistance. In this paper, we describe definitive evidence that miR482/2118b family members are negative regulators of quantitative disease resistance in tomato: suppression of miR482 and miR2118b using STTMs enhanced resistance against both P. infestans and Ps. syringae (Fig. 5). These findings are consistent with hypersusceptibility to Verticillium and P. infestans caused by overexpression of miR482 family members (30, 31). Similarly, higher expression levels of miR482/2118 members correlated with hypersusceptibility to P. infestans (32) in a separate study, and their expression level correlated inversely with resistance against Fusarium oxysporium (33). Our findings also confirm the very recent report that STTM482b confers enhanced resistance against P. infestans (31).

To explain the quantitative disease resistance in the STTM lines (Fig. 5) there could be, in principle, mechanisms that either involve NLRs as miR482/2118b targets or are NLR independent. We do not favor the NLR-independent mechanism, however, because known defense-related non-NLR mRNAs do not undergo large increases in abundance in the STTM lines (SI Appendix, Fig. S8 and Table S8) and there are very few predicted miR482/2118b targets that are not NLR RNAs (SI Appendix, Table S2). One exception is a predicted target of miR2118b encoding a DnaJ-type chaperone protein (Solyc01g105340) (SI Appendix, Table S2), but this RNA is not differentially expressed in the STTM lines (SI Appendix, Table S8) and its encoded protein is not easily accommodated in models of quantitative disease resistance. For these reasons, we favor quantitative disease resistance in the STTM lines that is mediated by NLRs, although we are aware that additional tests are required to block NLR signaling pathways or to knock out parts of the TAAS locus and sNL loci.

How many NLRs might affect the quantitative resistance in the STTM plants? With miR482, there are at least 32 NLRs generating secondary siRNAs (Fig. 1) and, in principle, there could be many more implicated in a miR482/2118b cascade. The first layer in this cascade would involve NLR RNAs that are direct targets of miR482 and a second layer might involve NLR targets of secondary siRNAs. With miR2118b, however,
the main primary target generating secondary siRNA is a non-coding RNA—TAS5. It is unlikely that this TAS5 RNA affects resistance directly, and it is likely that NLRs are targeted by the secondary siRNAs that are reduced in the STTM lines. (Fig. 3 and SI Appendix, Fig. S2 and Tables S4–S7).

Well-described RNA silencing mechanisms involving siRNA and miRNA result in cleavage and degradation of the target RNAs (14), and we had expected that the sNLRs and other NLR RNA targets of miR482/2118b family would increase in abundance in the STTM lines. There was, however, only a small increase (Fig. 4 and SI Appendix, Fig. S6 and Table S7) even for the sNLR RNAs with reduced secondary siRNAs in the STTM lines (Fig. 3). To reconcile these RNAseq results (Fig. 4 and SI Appendix, Fig. S6 and Table S7) with NLR-mediated quantitative resistance (Fig. 5), we hypothesize that there is a translational effect in the STTM plants that is not associated with major changes in the abundance of the targeted NLR RNAs. An effect in tomato of miR482 family members on NLR translation is also supported by an independent report (33).

NLR genes are normally associated with dominant gene resistance and with mechanisms that confine the pathogen to the site of initial infection. This dominant gene resistance was the first described phenotype of NLRs (34), and it is normally specific for races or strains of the pathogen. The NLRs in this race-specific resistance mediate direct or indirect recognition of effectors that are transported into the infected cell (1). This “effector-triggered immunity” then triggers metabolic and molecular changes that hinder the progression of disease (35). Our tests were not, however, with race-specific resistance. The plant genotypes were fully susceptible to both pathogens, and the STTM-mediated effect was manifested as quantitative resistance rather than complete suppression of the pathogen (Fig. 5).

To explain this quantitative resistance in the STTM plants, we propose that there is a low level of recognition of pathogen effectors, even in susceptible plants, but that any triggering of resistance would be too slow or too weak to prevent accumulation and spread of P. infestans and Ps. syringae. In the STTM plants, however, the reduced silencing of NLRs may allow stronger recognition and more rapid activation of defense. Alternatively, any increased expression of NLRs in the STTM lines could cause autoactivation of disease resistance in the absence of elicitor recognition, analogous to the effects of overexpressing NLRs in noninfected plants (2–4).

Evolutionary Dynamics of NLR Regulation by miRNA. The miRNA regulation of NLRs is highly variable in angiosperm species. In addition to the miR482/2118 family, there are numerous other miRNAs that target NLRs, including miR5300, miR6024, miR6026, miR2118a and others, with the potential to target NLRs (36). All of these miRNAs are variably present depending on the species, indicating a dynamic evolutionary process in which RNA silencing of NLR RNAs is lost and gained (7, 37).

The miR2118 family is a special case of NLR regulation that is dependent on two genome events. One of these changes in the Solanaceae involves functional diversification of miR2118 members, with one of them (miR2118b) being expressed in the vegetative phase including leaves and another (miR2118a) having the pattern that is typical of most monocot and dicot species (Fig. 1) in reproductive phases (20, 21). Consistent with a previous report (38) we were unable to find homologues of miR2118b outside of the Solanum clade (SI Appendix, Fig. S7). A second Solanaceae-specific genome change (SI Appendix, Figs. S1 and S7) would have been a rearrangement of TNL and CNL NLR genes, resulting in the TASS locus. These changes illustrate how the evolutionary dynamics of NLR regulation are not only dependent on loss or gain of the miRNA genes. There can also be neofunctionalization of miRNA genes, as with miR2118b, and the addition of noncoding RNAs into the RNA silencing cascade, as with TASS.

We have speculated previously that the miR482-mediated down-regulation of NLRs is a process that allows the plant to trade off the costs and benefits of NLRs (37). Accumulation of NLR proteins has a cost that is presumably related to metabolic changes associated with disease resistance and a benefit due to protection against pests and pathogens. The balance of costs and benefits may depend on the ecological niche occupied by the plant and on the effectiveness of other defense systems.

A component of our earlier speculation was that pathogen-derived suppressors of RNA silencing would relieve the miR482-mediated suppression of NLRs in infected plants (12). We originally envisioned that this process would operate in virus-infected plants (14) but with the identification of such suppressors encoded by other pathogens, including P. infestans (39), it could be more general. The effect of a suppressor of silencing could explain the increase of NLRs in P. infestans-infected plants (Fig. 44).

To explain the apparent minimal cost of STTMs on growth and development of the plants, we suggest that the protected glasshouse environment of our tomato plants allowed them to tolerate the increased expression of NLRs without effects on growth and development of the plant (SI Appendix, Fig. S6). There may also be other changes associated with the domestication and breeding of cultivated tomato that protect the
plant against the costs of increased NLRs. Additional testing will be required to establish whether there is a cryptic growth phenotype associated with STTM expression and also, whether the basal immunity can be enhanced further. It could be that a combination of the two STTMs would have a stronger effect than the individual species. Alternatively, the enhanced basal immunity would be useful if it is part of an integrated management strategy to protect crops against disease.

Materials and Methods

Plant strains, plasmid constructs, Northern blots, bioinformatic analyses, and pathogen tests are described in SI Appendix. The tomato (Solanum lycopersicum) cultivar M82 was used for all experiments. STTM transgenic lines were obtained via Agrobacterium tumefaciens-mediated stable transformation. Image analysis was performed on ImageJ. Statistical and bioinformatics analyses were performed using R.

ACKNOWLEDGMENTS. We thank James Barlow, Mel Steer, Antonia Yarur, and Pawel Baster for technical and horticultural assistance; Stuart Fawke for maintaining and providing P. infestans 88069; Zhengming Wang for the STTM171.1 seeds; and Thomas Hardcastle for bioinformatics advice. This work was supported by the Biotechnology and Biological Research Council Advanced Investigator Grant ERC-2013-AdG 340642. S.S. is funded by Gatsby Foundation Fellowship GAT3395/GLD and Royal Society University Research Fellowship UF160413, D.C.B. is the Royal Society Edward Penley Abraham Research Professor. The funding bodies had no roles in the design of the study; the collection, analysis, or interpretation of data; or the writing of the manuscript.