

# **Title: Convergent evolution of plant prickles by repeated gene co-option over deep time**

**Authors:** James W. Satterlee<sup>1,2</sup>, David Alonso<sup>3</sup>, Pietro Gramazio<sup>3</sup>, Katharine M. Jenike<sup>4,5</sup>, Jia He<sup>1,2</sup>, Andrea Arrones<sup>3</sup>, Gloria Villanueva<sup>3</sup>, Mariola Plazas<sup>3</sup>, Srividya Ramakrishnan<sup>4,5</sup>, Matthias Benoit<sup>6</sup>, Iacopo Gentile<sup>7</sup>, Anat Hendelman<sup>1,2</sup>, Hagai Shohat<sup>1</sup>, Blaine Fitzgerald<sup>1,2</sup>, Gina M. Robitaille<sup>1,2</sup>, Yumi Green<sup>8</sup>, Kerry Swartwood<sup>8</sup>, Michael J. Passalacqua<sup>7</sup>, Edeline Gagnon<sup>9,10</sup>, Rebecca Hilgenhof<sup>10</sup>, Trevis D. Huggins<sup>11</sup>, Georgia C. Eizenga<sup>11</sup>, Amit Gur<sup>12</sup>, Twan Rutten<sup>13</sup>, Nils Stein<sup>13,14</sup>, Shengrui Yao<sup>15,16</sup>, Adrien Poncet<sup>17</sup>, Clement Bellot<sup>17</sup>, Amy Frary<sup>18</sup>, Sandra Knapp<sup>19</sup>, Mohammed Bendahmane<sup>17</sup>, Tiina Särkinen<sup>10</sup>, Jesse Gillis<sup>1,20</sup>, Joyce Van Eck<sup>8,21</sup>, Michael C. Schatz<sup>4,5</sup>, Yuval Eshed<sup>22</sup>, Jaime Prohens<sup>3</sup>, Santiago Vilanova<sup>3</sup>, Zachary B. Lippman<sup>1,2,7\*</sup>

## **Affiliations:**

<sup>1</sup> Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

<sup>2</sup> Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

<sup>3</sup> Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain.

<sup>4</sup> Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA.

<sup>5</sup> Department of Genetic Medicine, Johns Hopkins University, Baltimore, MD, USA.

<sup>6</sup> French National Institute for Agriculture, Food, and Environment, Laboratory of Plant-Microbe Interactions, Toulouse, France.

<sup>7</sup> School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

<sup>8</sup> Boyce Thompson Institute, Ithaca, New York, USA.

<sup>9</sup> Department of Integrative Biology, University of Guelph, Ontario, Canada.

<sup>10</sup> Royal Botanic Garden Edinburgh, Edinburgh, UK.

<sup>11</sup> USDA-ARS, Dale Bumpers National Rice Research Center, Stuttgart, AR, USA.

<sup>12</sup> Cucurbits Section, Department of Vegetable Sciences, Agricultural Research Organization, Newe Ya'ar Research Center, Ramat Yishay, Israel.

<sup>13</sup> Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany.

<sup>14</sup> Crop Plant Genetics, Martin Luther University of Halle-Wittenberg, Halle (Saale), Germany.

<sup>15</sup> Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM, USA.

<sup>16</sup> Sustainable Agriculture Sciences Center, New Mexico State University, Alcalde, NM, USA.

<sup>17</sup> Laboratoire Reproduction et Développement des Plantes, INRAE, CNRS, Université Lyon, Ecole Normale Supérieure de Lyon, Lyon, France.

<sup>18</sup> Department of Biological Sciences, Mount Holyoke College, South Hadley, MA, USA.

<sup>19</sup> Natural History Museum, London, UK.

<sup>20</sup> Physiology Department and Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada.

<sup>21</sup> Plant Breeding and Genetics Section, Cornell University, Ithaca, NY USA.

5 <sup>22</sup> Department of Plant and Environmental Sciences, The Weizmann Institute of Science, Rehovot, Israel.

\*Corresponding author. Email: lippman@cshl.edu

10 **Abstract:** An enduring question in evolutionary biology concerns the degree to which episodes of convergent trait evolution depend on the same genetic programs, particularly over long timescales. Here we genetically dissected repeated origins and losses of prickles, sharp epidermal projections, that convergently evolved in numerous plant lineages. Mutations in a cytokinin hormone biosynthetic gene caused at least 16 independent losses of prickles in eggplants and wild relatives  
15 in the genus *Solanum*. Homologs underlie prickle formation across angiosperms that collectively diverged over 150 million years ago, including rice and rose. By developing new *Solanum* genetic systems, we leveraged this discovery to eliminate prickles in a wild species and an indigenously foraged berry. Our findings implicate a shared hormone-activation genetic program underlying evolutionarily widespread and recurrent instances of plant morphological innovation.

20 **One-Sentence Summary:** Uncovering a shared genetic basis for the convergent evolution of prickles facilitates their elimination in crop improvement.

**Main Text:** Trait convergence, defined as the emergence of analogous traits in distantly related organisms, was a key observation made by Darwin in support of his theory of evolution. He recognized that similar selective pressures could lead to similar yet independently derived adaptations across species. However, the extent to which phenotypic convergence is driven by corresponding convergence in underlying genetic programs is poorly understood. Within a species, adaptive traits may arise from selection acting on standing genetic variation within and among populations, making phenotype-genotype convergence more likely (1, 2). At higher taxonomic levels and with increasing evolutionary divergence, phenotype-genotype convergence is posited to decline due to variation in allelic diversity, genomic background, and developmental mechanisms (3, 4). However, opportunities to dissect convergence at these timescales are scarce; finding convergent traits across wide evolutionary spans that are genetically tractable and well-supported by genomic data has remained a significant challenge.

In plants, sharp epidermal projections known as prickles convergently evolved at least 28 times over more than 400 million years of evolution (5) (**Fig. 1A and Table S1**). Prickles serve adaptive functions in herbivore deterrence, climbing growth, plant competition, and water retention (6–9). Rose (*Rosa* spp.) is a widely recognized taxon bearing prickles, though these prickles are vernacularly called thorns. True thorns, which are found on the trees of citrus (*Citrus* spp.) and honey locusts (*Gleditsia* spp.), for example, develop from axillary branches, whereas prickles originate from the epidermis or cortex, typically in association with hair-like structures known as trichomes (6). Despite their diverse adaptive roles and the broad phylogenetic diversity of their origins, prickles exhibit remarkable morphological similarity (**Fig. S1A-C**). Moreover, prickles have been lost or suppressed in numerous lineages. Therefore, prickle formation is an attractive system to determine whether episodes of repeated trait evolution rely on the same genetic programs over both short and long evolutionary timescales.

In the genus *Solanum*, which includes the major crops eggplant, potato, and tomato, prickles emerged in the common ancestor of the so-called “spiny *Solanums*” around 6 million years ago (Mya) (10, 11). This lineage includes the large Leptostemonum clade, which comprises hundreds of globally distributed species, including all cultivated eggplants and their wild progenitors (**Fig. 1B**). Prickle morphologies across the clade range from broad at the base (broad-based), or narrow-based and needle-like. Prickles occur on stems, along the vasculature of leaves, and on calyces, the outer whorl of floral organs. Several spiny *Solanum* species underwent human-driven selection for losses or suppression of prickles (12, 13), facilitating comparisons of prickled and prickless sister species, crop species, and wild relatives (**Fig. 1C and Table S2**). An agriculturally significant instance of prickle loss occurred during the domestication of the widely cultivated Brinjal eggplant (*S. melongena*); however, prickle losses have also been observed in wild *Solanum* species without history of domestication (**Fig. 1D**). The specific genes controlling prickle development are unknown.

### **Repeated losses of prickles in cultivated eggplants are caused by *LOG* gene mutations**

Previous mapping studies in Brinjal eggplant showed that the loss of prickles is inherited as a single Mendelian locus designated *prickleless* (*pl*) and localized to a genomic interval on chromosome 6 (14). Using a recurrent backcross-derived mapping population between Brinjal eggplant and its prickled wild progenitor *S. insanum*, we confirmed this result and further fine-mapped *pl* to a ~100 kb interval containing 10 annotated genes (**Fig. 2A**). Just outside this interval is the previously proposed *pl* candidate gene *SmelARF18*, a putative auxin hormone response

transcription factor (15). However, we did not find conspicuous coding region loss-of-function mutations in this gene or in any other gene in the interval. Instead, we identified a probable splice-site mutation in a gene encoding a LONELY GUY (LOG)-family cytokinin biosynthetic enzyme. LOG family members catalyze the final step in the biosynthesis of bioactive cytokinin, a hormone with roles in plant cell proliferation and differentiation (16). In a collection of 23 re-sequenced eggplant accessions (17), we found that this splice-site mutation was consistently associated with the prickless phenotype, except in one accession, which harbored a 474 bp deletion in exon 6 of the *LOG* gene (Fig. S2A and Table S3).

The discovery of two independent mutations in the *LOG* candidate gene suggested that the loss of prickles occurred at least twice in Brinjal eggplant or its wild relatives. It also raised the possibility that mutations in orthologous genes may have caused parallel prickle losses in two other independently domesticated African eggplant species, the Scarlet eggplant (*S. aethiopicum*) and the Gboma eggplant (*S. macrocarpon*). Genomic resources for these indigenous crop species are limited. We therefore sequenced and assembled high-quality (QV  $\geq$  51, completeness > 99) chromosome-scale genomes and generated gene annotations for both species (Fig. 1C, Tables S4 and S5). Using these resources, we found that synteny within the *pl* locus was retained across all three cultivated eggplant species (Fig. 2B), and that prickless Scarlet eggplant and Gboma eggplant each harbored different loss-of-function mutations in their respective *LOG* orthologs (Fig. 2C). Scarlet eggplant carries an indel mutation leading to a frameshift in the coding sequence and a prematurely terminated protein product, while Gboma eggplant carries a splice-site mutation. Reverse-transcriptase polymerase chain reaction (RT-PCR) on cDNA revealed lower expression and multiple mis-spliced transcripts in Brinjal eggplant and a mis-spliced isoform with a retained intron in Gboma eggplant (Fig. 2D). PCR sequencing revealed these transcripts were non-functional (Fig. S2B-D).

To further validate that these independent mutations explain the prickless phenotypes, we next performed co-segregation analysis in F2 populations derived from intraspecific crosses between prickled and prickless parents (Fig. 1C and S2). In Scarlet eggplant, homozygosity of the *LOG* mutant allele co-segregated with the prickless phenotype in a Mendelian recessive fashion in all examined individuals ( $\chi^2 = 0.52$ , df = 1,  $p = 0.47$ ). In Gboma eggplant we observed segregation patterns that indicated the presence of another unlinked recessive variant independently contributing to prickle loss ( $\chi^2 = 14.8$ , df = 1,  $p < 0.001$ ). Leveraging our newly developed genomic resources, we used a mapping-by-sequencing approach to identify a second large interval associated with the loss of prickles on chromosome 4, which we designated *pl2* (Fig. 2E). Importantly, all segregating homozygous mutant individuals at *pl* on chromosome 6 carried the *LOG* gene splice-site mutation, although this genotype class was represented at lower-than-expected frequency, likely owing to segregation distortion (Fig. S2G). Finally, we modified existing plant regeneration and transformation protocols to engineer loss-of-function *PL* alleles using CRISPR-Cas9 genome editing in a prickled accession of *S. aethiopicum*. Analysis of three independently edited multiallelic transformants revealed suppression of prickle development due to numerous frameshift mutations resulting in *PL* loss-of-function (Fig. 2F and S2F). Transformants lacking mutations retained prickles. Taken together, these results indicate that *PL* is the *LOG* candidate gene, and that at least four independent mutations in this gene enabled repeated selection for losses of prickles in cultivated eggplant species.

## Mutations in *PL* are found in prickless wild and cultivated species across the *Solanum*



The clade encompassing all three of the cultivated eggplants diverged ~2 Mya, but prickles in *Solanum* are more ancient, having emerged over ~6 Mya, and 31 independent losses of prickles have been documented, including in additional domesticated and wild species (11). We tested whether mutations in *PL* underlie these repeated instances of prickle loss across this broader evolutionary timescale by sampling DNA from additional prickleless species and their prickled close relatives. Because many wild *Solanum* species are too rare or geographically inaccessible for live-tissue sampling, we used a combination of PCR-amplified exon sequencing from herbarium tissue samples and whole-gene sequencing from available live tissue samples to detect *PL* mutations (**Fig. S3 and Table S6**).

Along with the four *PL* mutations identified in our analysis of prickleless eggplants, we identified an additional 12 allelic mutations predicted to deleteriously affect *PL* function at the pan-genus level across the spiny *Solanum* (**Fig. 3**). These mutations, together with those detected by mapping, were associated with 14 out of the 31 recorded losses of prickles across the genus at the species level (**Fig. 1B and Table S7**). We then confirmed that these mutations were not found in prickled species from closely-related lineages (**Fig. S3**). In some cases, we detected the same, although not necessarily ancestrally derived, alleles in separate species. For example, the same splice-site mutation found in prickleless Gboma eggplant, native to and cultivated almost exclusively in Africa, was also identified in the wild species *S. donianum*, whose native range is in Central America and the Caribbean. Likewise, an identical splice-site mutation was found in both the wild species *S. lanzae*, from western Africa, and the foraged and sometimes cultivated species *S. stramonifolium*, native to northern South America. Such genetic convergence at the allelic level may reflect the high penetrance of *PL* splicing defects, which can be conferred by mutationally accessible single nucleotide variants (18). Together, our results suggest that *PL* had an important and repeated genetic role in the convergent losses of prickles across *Solanum* in the wild and in cultivation. However, loci other than *PL* may explain prickles losses in lineages for which mutations were not identified.

### Repeated co-option of *LOG* homologs underlies prickle convergent evolution

The finding of recurrent mutations in *PL* orthologs across the spiny *Solanums* suggested that co-option of cytokinin biosynthetic gene function was critical to prickle evolution. This spurred us to ask whether genetic convergence through *LOG* gene co-option extends to other prickled species across flowering plants. We searched the literature for studies associating instances of loss or suppression of sharp outgrowths with specific genomic loci or genes. Strikingly, we found that in the grass family (Poaceae) independent mutant alleles in a *LOG* homolog from rice (*Oryza sativa*) and barley (*Hordeum vulgare*) conferred near complete suppression of epidermally-derived sharp projections commonly called “barbs” but botanically classified as prickles (19, 20). In contrast to the conspicuous, multi-cellular, and lignified prickles found in the *Solanum* (**Fig. S1**), grass prickles are homologous structures made of silicified single-cells that develop on awns (**Fig. 4A**), an outer-whorl structure of the grass flower involved in seed dispersal, along with leaves and spikelets.

Mining additional genomic data for *LOG* mutations co-occurring with losses of prickles in other eudicot lineages, we found that the fruit-bearing tree crop jujube, commonly known as Chinese date (*Ziziphus jujuba*), in the Rhamnaceae family, carried two independent mutations (a 1 bp deletion and an exonic insertion) in a *LOG* homolog in two cultivars with suppressed prickles (also known as stipular spines) (**Fig. 4B**) (21–23). Importantly, neither mutation was found in Sour

jujube (*Z. jujuba* var. *spinosa*), the prickled wild progenitor. We also detected an exonic insertion in a *LOG* homolog of the prickle suppressed ‘Purple Queen’ cultivar of giant spider flower (*Teranaya hassleriana*), a widely cultivated ornamental plant in the Cleomaceae, a small family within the Brassicales closely related to Arabidopsis (**Fig. 4C and Table S6**) (24). Finally, in rose, which is a commercially important cultivated cut flower, previous mapping for loci conferring “thornlessness” identified two major effect loci (9), as we found in *S. macrocarpon*. One of these was a ~2.5 Mb interval containing 156 annotated genes (**Fig. 4D**), which we found includes a *LOG* homolog. Though there were no obvious coding or splicing mutations in this *LOG*, we found that its expression was substantially reduced in the leaves of the mapping parent cultivar ‘Bayse’s thornless’ (*Rosa wichuraiana*) compared to the prickled parent *Rosa chinensis* (**Fig. 4D**). To determine whether this candidate *LOG* has a role in rose prickle development we used a virus induced gene silencing (VIGS) approach to reduce *LOG* function (25). In 2/14 rose plants infected with the tobacco rattle virus (TRV) expressing an inverted repeat of *LOG* RNA, strong suppression of prickle development was observed, while wild-type plants of the same background showed normal prickle development (**Fig. 4D**).

Taken together, these findings suggested that *LOG* gene reuse was critical in the independent acquisition of prickles in numerous plant lineages that last shared a common ancestor ~150 Mya. Most sequenced seed plants (angiosperms and gymnosperms) retain multiple *LOG* gene copies within their genomes. In these taxa, the mean number of annotated *LOG* genes is 15, inflated by recent polyploid lineages, while the median and mode copy numbers are 12 and 10, respectively ( $N = 160$ ). To understand the phylogenetic context of *LOG* co-option and to ask whether repeated co-option occurs in a specific clade of *LOG* gene family members, we conducted an analysis of *LOG* family proteins from prickled and prickless species across the angiosperms (**Fig. 4E**). Most of the prickle co-option associated *LOG*s occurred within a specific subclade of the *LOG* family, suggesting that co-option was more favorable in certain *LOG* family subclades, particularly those with lineage-specific duplications. However, the *LOG* homolog co-opted in barley is derived from an earlier diverging subclade (20), indicating that despite a subclade bias, co-option of other *LOG* family members in different clades may also be associated with prickle evolution.

### ***LOG* gene diversification preceded *PL* co-option in *Solanum***

Given the recurrent co-option of *LOG* genes against a backdrop of paralogous gene family members, we sought to better understand the phylogenetic and genomic context that facilitated *LOG* co-option in *Solanum*. We examined the conservation of the *PL* locus, comparing the region across *Solanum*, including Brinjal eggplant and two additional spiny *Solanum* species, with tomato (*S. lycopersicum*), an ancestrally prickless species that diverged prior to the evolution of the spiny *Solanums*. We first constructed high-quality chromosome-scale genome assemblies for *S. prinophyllum* (Forest nightshade; QV = 51.6, completeness > 99) and *S. cleistogamum* [Desert raisin; QV = 49.8, completeness > 99 (**Tables S4 and S5**)]. Forest nightshade is endemic to southeastern Australia whereas Desert raisin is native to the arid center of Australia and has been foraged by First Nations people for thousands of years for their sweet, dried berries (26) (**Fig. 5A**). In our screen for *PL* mutations across the spiny *Solanum* we did not identify any naturally occurring *PL* mutations in the Australian *Solanum* lineages to which these species belong (**Fig. S3**). Indeed, neither species has been domesticated and are distinct lineages from the cultivated eggplants.

Leveraging these newly developed genomic resources, we found that synteny at the *PL* locus was conserved across the *Solanum*, suggesting that *PL* was co-opted from a standing syntenic ortholog that existed at least since the divergence of tomato and the spiny *Solanums* ~14 Mya (**Fig. 5B**). To better understand ancestral *PL* function across eudicots, we performed a meta-analysis of gene expression data from Arabidopsis (3154 samples) and tomato (5491 samples), reasoning that shared expression profiles reflect the degree of shared inter-species function (27). We assessed each member of the *LOG* family for its ability to predict co-expression in every other member of the *LOG* family in the other species. An area under the receiver operating characteristic (AUROC) curve statistic of 0.93 indicated that *SlycPL* in tomato is co-expressed with nearly identical genes to that of *AthaLOG1* in Arabidopsis, pointing to a conserved function. Likewise, three other tomato *LOG* gene family members also exhibited strongly conserved co-expression with *AthaLOG1* (**Fig. 5C**). The *LOG1* clade, to which *PL* belongs, has therefore maintained signatures of functional conservation across ~120 My.

Tissue-specific knockdown of *AthaLOG1* in the Arabidopsis floral meristem has been shown to impair floral organ initiation, suggesting that *AthaLOG1* has critical roles in meristem maintenance, similar to the canonical developmental role for *LOGs* first reported in rice (16, 28). Therefore, duplication and diversification of the *LOG1* subclade in the *Solanum* may have facilitated *PL* functional co-option. To explore this hypothesis, we generated an expression atlas for prickled forest nightshade and compared it to matched-tissue gene expression data from tomato and Arabidopsis (**Fig. 5D**). In Arabidopsis, *AthaLOG1* possesses a broad expression pattern across tissues, while *Solanum PL* and *LOG1a* have evolved more tissue-biased expression patterns. Compared to its ortholog in tomato, forest nightshade *SprPL* has evolved enriched expression in flowers and, consistent with its co-opted function, in developing prickles. Therefore, paralog diversification in the *Solanum* likely enabled functional co-option and redeployment of ancestral *LOG1* clade function in prickle development.

### Non-pleiotropic removal of prickles with gene editing

We reasoned that the co-option of *PL* could facilitate the engineering of agriculturally desirable loss-of-function prickless mutants, even in the Australian spiny *Solanum* taxa in which we did not detect naturally occurring *PL* mutations. The duplication leading to *PL* and its subsequent expression divergence from its ancestral copy would prevent undesirable pleiotropic effects on other traits. Alternatively, cryptic background modifiers in prickless lineages may have been required to specifically suppress prickle development, and thus eliminating *PL* would leave prickles intact, or result in pleiotropy. To distinguish between these two possibilities, we devised a pan-genus CRISPR-Cas9 editing strategy to target *PL* in Forest nightshade, Desert raisin, and tomato, the latter of which harbors a *PL* ortholog, likely performing an ancestral function outside of prickle development. Adapting techniques previously established in tomato (29), we developed plant regeneration, transformation, and genome editing for Forest nightshade and Desert raisin, thereby elevating these two species into new *Solanum* genetic systems. We engineered multiple loss-of-function mutations in *PL* (*pl<sup>CR</sup>*) in all three species and compared their phenotypes (**Fig. 5E**). In both Forest nightshade and Desert raisin, *pl<sup>CR</sup>* individuals showed strong suppression of prickle development in all tissues and organs where prickles normally develop in wild type plants, though we observed small sporadic prickles (**Figs. 5F,G and S4**). Meanwhile, in tomato, *Slycpl<sup>CR</sup>* plants resembled the wild type, likely due to genetic redundancy with *SlycLOG1a* and possibly other *LOG* family members prior to the *PL* co-option event ~6 Mya (**Fig. 5H**). Fruit morphology and sweetness remained unchanged (Brix sugar content ~30% compared to ~50% in

grape raisins and ~9% in cherry tomato) and trichome density and morphology appeared unaffected in WT and *pl<sup>CR</sup>* Desert raisin lines. These results suggest that *PL* targeting is an effective strategy for first line improvement of harvestability in wild or partially domesticated prickled species bearing edible fruits, including additional locally-important cultivated indigenous *Solanum* such as vila-vila (*S. sisymbriifolium*) and naranjilla (*S. quitoense*) (**Fig. 5I, S5, and Table S8**).

## Discussion

Here we showed multiple, phylogenetically independent reuses of *LOG* family members in prickle development across 150 My of plant evolution. Studies addressing convergent trait evolution at these timescales have hinted that similar and divergent genetic programs can underpin phenotypic convergence (3, 30). For example, the convergent evolution of echolocation in bats and cetaceans is associated with positive selection on variation in shared orthologous genes (31). In plants, convergent evolution of floral asymmetry has been shown in numerous species to occur by modified expression of the transcription factor encoding gene *CYCLOIDEA* (32). On the other hand, different loci were reported to underlie convergent adaptation to marine habitats in mammals (33). The repeated use of the same genetic program seen in some traits such as prickles may in part be due to their relative simplicity. Unlike composite traits (34), where selection has the potential to act on many different loci affecting many different organismal systems, convergent traits that arise from selection on fewer potentially relevant loci may exhibit greater genetic convergence by virtue of sheer probability. However, even traits of modest complexity, such as animal eye lenses composed of homomeric crystallins (35), can have many distinct genetic origins, indicating that trait complexity alone cannot fully account for observed patterns of convergent evolution.

Genotype-phenotype convergence may also rely on developmental constraints imposed on morphological innovation, which often depends on the re-purposing of ancestral genetic mechanisms (36, 37). Gene co-option may allow key developmental regulators to take on new roles via non-pleiotropically partitioning gene function, particularly when standing paralog diversity exists. This has been suggested as an explanation for the repeated evolution of limbs, for example, by co-option of *Hox* genes. We suggest that functionally redundant *LOG* paralogs that arose through lineage-specific or shared ancestral duplication events may acquire specialized functions, as we found with prickles. The lack of an apparent *pl* mutant phenotype in tomato, coupled with the strong suppression of prickles in *pl* mutants in prickled lineages without obvious effects on other traits is consistent with *PL* functional co-option. Even after co-option in prickle development, *LOGs* may retain some functional redundancy, as engineered and natural (i.e. rice and barley) *LOG* mutants still produce sporadic small prickles (**Figs. 4A and S4**). Even partial paralog redundancy may increase the odds of phenotype-genotype convergence by allowing selection for gains and losses of prickles while avoiding developmental pleiotropy.

Perhaps most importantly, as an essential plant hormone with key developmental functions, cytokinin is well-suited to serve a recurrent role in morphological adaptation. Like the plant hormones auxin and florigen, cytokinins have cell-type and stage-specific effects. For example, beyond its role in promoting cell proliferation in shoot meristems (16), cytokinin contributes to microtubule reorientation in maturing root epidermal cells (38) and promotes growth cessation associated with cell wall stiffening in the root differentiation zone (39). The results presented here endow cytokinin activation by *LOGs* with a central and repeated role in morphological innovation. This could occur by canonical cytokinin activation of cell proliferation but could also involve cytokinin promotion of the differentiation program leading to the hard, lignified structure of the

prickle. Other plant morphological innovations are also controlled by cytokinin-related gene activity. Overexpression of a *LOG* gene is sufficient to induce the ectopic formation of shoot-borne tubers in axillary meristems in tomato (40), while a dominant mutation in a gene encoding a cytokinin receptor protein induces the ectopic formation of root nodules in the legume *Lotus japonicus* (41), both of which depend on localized cell proliferation. Unlike typical “master” regulators that often coordinate complex programs, such as floral homeotic genes (42), the repeated loss of prickles reported here relies on an enzymatic gene family involved in the activation of several types of cytokinins. Whether redeployment of such hormone activation genes in new developmental contexts is sufficient to generate morphological novelty warrants further study.

Finally, we propose that targeted gene editing of cytokinin biosynthesis and signaling components, as demonstrated here, is likely a predictable and efficient strategy for eliminating prickles in various flowering plant lineages. This approach is particularly promising for roses, where the labor intensive, manual removal of prickles is a common practice for most cut varieties. Though roses have variable ploidy (43) and genome editing in elite germplasm can be challenging (44), we demonstrated this potential using VIGS to suppress prickle development (Fig. 4D). Beyond the species presented here, the observed subclade bias in *LOG* homolog co-option will likely aid in selection of *LOG* genes for site-directed mutagenesis in other taxa. However, in principle, the general role of *LOG* proteins in cytokinin activation could allow more distantly related *LOG*s to carry out their role in prickle development, as occurs in barley (Fig. 4A,E). This necessitates consideration of both *LOG* gene expression and phylogenetic context for targeting prioritization. Overall, continued efforts to unite genetics, genomics, and genome editing across diverse plants, as illustrated in this study, will both advance our ability to track evolutionary changes over a broad range of time scales and empower the engineering of novel phenotypes to expand our use of plant diversity in agriculture.

## Methods Summary

For the mapping of *pl* in *S. melongena*, previously generated introgressions of prickled *S. insanum* into the prickleless *S. melongena* background were screened for small prickle-associated genomic intervals on chromosome 6 (45). An individual with the narrowest identified interval was then selfed and a total of 622 resulting progeny were used for fine mapping of *pl* by PCR-based marker genotyping.

For genome assembly, high molecular weight DNA was extracted from flash-frozen, dark-treated 4-week-old seedlings. A combination of long-read sequencing (Pacific Biosciences, CA, USA) and optical mapping (Bionano Genomics, CA, USA) data were used for assembly. Sequencing reads from each sample were assembled with hifiasm (46) exact parameters and software version varied between samples based on the level of estimated heterozygosity and are reported in Table S4. In addition, high-throughput chromosome conformation capture (Arima Genomics, CA, USA) was performed for one sample, *S. prinophyllum*, to finalize scaffolding. Using merqury (47), the final consensus quality (QV) of the assemblies was 51.1, on average.

For genome annotation, orthologs with coverage above 50% and 75% identity were lifted from Heinz v4.0 Heinz v4.0 (48) and Eggplant v4.1 (17) via Liftoff (49) and refined using protein and gene microsynteny support. The completeness of the gene models was determined by assessing single-copy orthologs using BUSCO5 (50).

Plant regeneration and *Agrobacterium tumefaciens*-mediated transformation of *S. prinophyllum* and tomato were performed according to (51). The same methods were also used for *S. aethiopicum* and *S. cleistogamum* with two modifications. For *S. cleistogamum*, plant regeneration, the medium was supplemented with 0.5 mg/L zeatin instead of 2 mg/L and for the

selection medium, 75 mg/L kanamycin was used instead of 200 mg/L. For *S. aethiopicum*, the protocol was the same as for *S. cleistogamum*, except the fourth transfer of transformed plantlets was performed onto media supplemented with 50 mg/L kanamycin.

## References and Notes

- 5      1. Y. F. Chan *et al.*, Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a *Pitx1* Enhancer. *Science* **327**, 302–305 (2010). doi:10.1126/science.1182213.
2. V. Konečná *et al.*, Parallel adaptation in autopolyploid *Arabidopsis arenosa* is dominated by repeated recruitment of shared alleles. *Nat. Commun.* **12**, 4979 (2021). doi:10.1038/s41467-021-25256-5.
- 10     3. M. Bohutínská, C. L. Peichel, Divergence time shapes gene reuse during repeated adaptation. *Trends Ecol. Evol.*, S0169-5347(23)00325–7 (2023). doi:10.1016/j.tree.2023.11.007.
4. J. Arendt, D. Reznick, Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32 (2008).  
15      doi:10.1016/j.tree.2007.09.011.
5. D. M. De Vienne, Lifemap: Exploring the Entire Tree of Life. *PLoS Biol.* **14**, e2001624 (2016). doi:10.1371/journal.pbio.2001624.
6. T. C. Coverdale, Defence emergence during early ontogeny reveals important differences between spines, thorns and prickles. *Ann. Bot.* **124**, iii–iv (2019). doi:10.1093/aob/mcz189.
- 20     7. F. Gallenmüller, A. Feus, K. Fiedler, T. Speck, Rose Prickles and Asparagus Spines – Different Hook Structures as Attachment Devices in Climbing Plants. *PLoS ONE* **10**, e0143850 (2015). doi:10.1371/journal.pone.0143850.
8. Water Worlds, *The Green Planet* (2022).
9. M.-C. Zhong *et al.*, Rose without prickle: genomic insights linked to moisture adaptation.  
25      *Natl. Sci. Rev.* **8**, nwab092 (2021). doi:10.1093/nsr/nwab092.
10. E. Gagnon *et al.*, Phylogenomic discordance suggests polytomies along the backbone of the large genus *Solanum*. *Am. J. Bot.* **109**, 580–601 (2022). doi:10.1002/ajb2.1827.
11. R. Hilgenhof *et al.*, Morphological trait evolution in *Solanum* (Solanaceae): Evolutionary lability of key taxonomic characters. *Taxon* **72**, 811–847 (2023). doi:10.1002/tax.12990.
- 30     12. M. D. Whalen, E. E. Caruso, Phylogeny in *Solanum* sect. *Lasiocarpa* (Solanaceae): Congruence of Morphological and Molecular Data. *Systematic Botany* **8**, 369 (1983). doi:10.2307/2418356.
- 35     13. S. Knapp, M. S. Vorontsova, A revision of the “African Non-Spiny” Clade of *Solanum* L. (*Solanum* sections *Afrosolanum* Bitter, *Benderianum* Bitter, *Lemurisolanium* Bitter, *Lyciosolanum* Bitter, *Macronesiotes* Bitter, and *Quadrangulare* Bitter: Solanaceae). *PhytoKeys*, 1–142 (2016). doi:10.3897/phytokeys.66.8457.

14. A. Frary *et al.*, QTL hotspots in eggplant (*Solanum melongena*) detected with a high resolution map and CIM analysis. *Euphytica* **197**, 211–228 (2014). doi:10.1007/s10681-013-1060-6.
- 5 15. S. Li *et al.*, Fine mapping an *AUXIN RESPONSE FACTOR*, *SmARF18*, as a candidate gene of the *PRICKLE LOCUS* that controls prickly absence/presence on various organs in eggplant (*Solanum melongena* L.). *Sci. Hortic.* **327**, 112874 (2024). doi:10.1016/j.scienta.2024.112874.
16. T. Kurakawa *et al.*, Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature* **445**, 652–655 (2007). doi:10.1038/nature05504.
- 10 17. L. Barchi *et al.*, G. Giuliano, Improved genome assembly and pan-genome provide key insights into eggplant domestication and breeding. *Plant J.* **107**, 579–596 (2021). doi:10.1111/tbj.15313.
18. J. G. Monroe *et al.*, Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* **602**, 101–105 (2022). doi:10.1038/s41586-021-04269-6.
- 15 19. L. Hua *et al.*, *LABA1*, a Domestication Gene Associated with Long, Barbed Awns in Wild Rice. *Plant Cell* **27**, 1875–1888 (2015). doi:10.1105/tpc.15.00260.
20. S. G. Milner *et al.*, Genebank genomics highlights the diversity of a global barley collection. *Nat. Genet.* **51**, 319–326 (2019). doi:10.1038/s41588-018-0266-x.
- 20 21. M.-J. Liu *et al.*, The complex jujube genome provides insights into fruit tree biology. *Nat. Commun.* **5**, 5315 (2014). doi:10.1038/ncomms6315.
22. J. Huang *et al.*, The Jujube Genome Provides Insights into Genome Evolution and the Domestication of Sweetness/Acidity Taste in Fruit Trees. *PLoS Genet.* **12**, e1006433 (2016). doi:10.1371/journal.pgen.1006433.
- 25 23. L.-Y. Shen *et al.*, Chromosome-Scale Genome Assembly for Chinese Sour Jujube and Insights Into Its Genome Evolution and Domestication Signature. *Front. Plant Sci.* **12**, 773090 (2021). doi:10.3389/fpls.2021.773090.
24. S. Cheng *et al.*, The *Tarenaya hassleriana* Genome Provides Insight into Reproductive Trait and Genome Evolution of Crucifers. *Plant Cell* **25**, 2813–2830 (2013). doi:10.1105/tpc.113.113480.
- 30 25. Y. Liu, M. Schiff, S. P. Dinesh-Kumar, Virus-induced gene silencing in tomato. *Plant J.* **31**, 777–786 (2002). doi:10.1046/j.1365-313x.2002.01394.x.
26. P. K. Latz, *Bushfires & Bushtucker: Aboriginal Plant Use in Central Australia* (IAD Press, Alice Springs, Australia, 1995).
- 35 27. J. Lee, M. Shah, S. Ballouz, M. Crow, J. Gillis, CoCoCoNet: conserved and comparative co-expression across a diverse set of species. *Nucleic Acids Res.* **48**, W566–W571 (2020). doi:10.1093/nar/gkaa348.

28. Y. Han, Y. Jiao, *APETALA1* establishes determinate floral meristem through regulating cytokinins homeostasis in *Arabidopsis*. *Plant Signal. Behav.* **10**, e989039 (2015). doi:10.4161/15592324.2014.989039.
- 5 29. C. Brooks, V. Nekrasov, Z. B. Lippman, J. Van Eck, Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. *Plant Phys.* **166**, 1292–1297 (2014). doi:10.1104/pp.114.247577.
30. A. Martin, V. Orgogozo, The Loci of Repeated Evolution: A Catalog of Genetic Hotspots of Phenotypic Variation. *Evolution* **67**, 1235–1250 (2013). doi:10.1111/evo.12081.
- 10 31. J. Parker *et al.*, Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* **502**, 228–231 (2013). doi:10.1038/nature12511.
32. L. C. Hileman, Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Phil. Trans. R. Soc. B* **369**, 20130348 (2014). doi:10.1098/rstb.2013.0348.
- 15 33. A. D. Foote *et al.*, Convergent evolution of the genomes of marine mammals. *Nat. Genet.* **47**, 272–275 (2015). doi:10.1038/ng.3198.
34. G. Chomicki *et al.*, Convergence in carnivorous pitcher plants reveals a mechanism for composite trait evolution. *Science* **383**, 108–113 (2024). doi:10.1126/science.ade0529.
- 20 35. G. Wistow, Lens crystallins: gene recruitment and evolutionary dynamism. *Trends Biochem. Sci.* **18**, 301–306 (1993). doi:10.1016/0968-0004(93)90041-K.
36. N. Shubin, C. Tabin, S. Carroll, Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818–823 (2009). doi:10.1038/nature07891.
37. D. L. Stern, The genetic causes of convergent evolution. *Nat. Rev. Genet.* **14**, 751–764 (2013). doi:10.1038/nrg3483.
- 25 38. J. C. Montesinos *et al.*, Phytohormone cytokinin guides microtubule dynamics during cell progression from proliferative to differentiated stage. *EMBO J.* **39**, e104238 (2020). doi:10.15252/embj.2019104238.
39. S. Liu *et al.*, Cytokinin promotes growth cessation in the *Arabidopsis* root. *Curr. Biol.* **32**, 1974–1985.e3 (2022). doi:10.1016/j.cub.2022.03.019.
- 30 40. T. Eviatar-Ribak *et al.*, A Cytokinin-Activating Enzyme Promotes Tuber Formation in Tomato. *Curr. Biol.* **23**, 1057–1064 (2013). doi:10.1016/j.cub.2013.04.061.
41. L. Tirichine *et al.*, A Gain-of-Function Mutation in a Cytokinin Receptor Triggers Spontaneous Root Nodule Organogenesis. *Science* **315**, 104–107 (2007). doi:10.1126/science.1132397.



42. V. Irish, The ABC model of floral development. *Curr. Biol.* **27**, R887–R890 (2017). doi:10.1016/j.cub.2017.03.045.
43. O. Raymond *et al.*, The *Rosa* genome provides new insights into the domestication of modern roses. *Nat. Genet.* **50**, 772–777 (2018). doi:10.1038/s41588-018-0110-3.
- 5 44. A. Giovannini, M. Laura, B. Nesi, M. Savona, T. Cardi, Genes and genome editing tools for breeding desirable phenotypes in ornamentals. *Plant Cell Rep.* **40**, 461–478 (2021). doi:10.1007/s00299-020-02632-x.
45. M. Plazas *et al.*, Introgression breeding from crop wild relatives in eggplant landraces for adaptation to climate change. *Crop wild relative*, 32–36 (2020).
- 10 46. H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021). doi:10.1038/s41592-020-01056-5.
47. A. Rhie, B. P. Walenz, S. Koren, A. M. Phillippy, Merquy: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245  
15 (2020). doi:10.1186/s13059-020-02134-9.
48. P. S. Hosmani *et al.*, “An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps” (preprint, Genomics, 2019); <https://doi.org/10.1101/767764>.
49. A. Shumate, S. L. Salzberg, Liftoff: accurate mapping of gene annotations. *Bioinformatics*  
20 **37**, 1639–1643 (2021). doi:10.1093/bioinformatics/btaa1016.
50. M. Manni, M. R. Berkeley, M. Seppey, E. M. Zdobnov, BUSCO: Assessing Genomic Data Quality and Beyond. *Curr. Protoc.* **1**, e323 (2021). doi:10.1002/cpz1.323.
51. J. Van Eck, P. Keen, M. Tjahjadi, “*Agrobacterium tumefaciens*-Mediated Transformation of Tomato” in *Transgenic Plants*, S. Kumar, P. Barone, M. Smith, Eds. (Springer New  
25 York, New York, NY, 2019) vol. 1864 of *Methods in Molecular Biology*, pp. 225–234.
52. R. H. G. Ranil *et al.*, Improving seed germination of the eggplant rootstock *Solanum torvum* by testing multiple factors using an orthogonal array design. *Sci. Hortic.* **193**, 174–181 (2015). doi:10.1016/j.scienta.2015.07.030.
53. A. M. McClung *et al.*, Enhancing the searchability, breeding utility, and efficient management of germplasm accessions in the USDA–ARS rice collection. *Crop Sci.* **60**,  
30 3191–3211 (2020). doi:10.1002/csc2.20256.
54. B. Kouassi *et al.*, Development of backcross generations and new interspecific hybrid combinations for introgression breeding in eggplant (*Solanum melongena*). *Sci. Hortic.* **213**, 199–207 (2016). doi:10.1016/j.scienta.2016.10.039.

55. M. Plazas *et al.*, Interspecific Hybridization between Eggplant and Wild Relatives from Different Genepools. *J. Amer. Soc. Hort. Sci.* **141**, 34–44 (2016). doi:10.21273/JASHS.141.1.34.
56. L. Barchi *et al.*, Single Primer Enrichment Technology (SPET) for High-Throughput Genotyping in Tomato and Eggplant Germplasm. *Front. Plant Sci.* **10**, 1005 (2019). doi:10.3389/fpls.2019.01005.
57. P. Gramazio *et al.*, Whole-Genome Resequencing of Seven Eggplant (*Solanum melongena*) and One Wild Relative (*S. incanum*) Accessions Provides New Insights and Breeding Tools for Eggplant Enhancement. *Front. Plant Sci.* **10**, 1220 (2019). doi:10.3389/fpls.2019.01220.
58. S. Vilanova *et al.*, SILEX: a fast and inexpensive high-quality DNA extraction method suitable for multiple sequencing platforms and recalcitrant plant species. *Plant Meth.* **16**, 110 (2020). doi:10.1186/s13007-020-00652-y.
59. L. Barchi *et al.*, A chromosome-anchored eggplant genome sequence reveals key events in Solanaceae evolution. *Sci. Rep.* **9**, 11769 (2019). doi:10.1038/s41598-019-47985-w.
60. C. T. Wittwer, G. H. Reed, C. N. Gundry, J. G. Vandersteen, R. J. Pryor, High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. *Clin. Chem.* **49**, 853–860 (2003). doi:10.1373/49.6.853.
61. M. Kanakachari *et al.*, Evaluation of Suitable Reference Genes for Normalization of qPCR Gene Expression Studies in Brinjal (*Solanum melongena* L.) During Fruit Developmental Stages. *Appl. Biochem. Biotechnol.* **178**, 433–450 (2016). doi:10.1007/s12010-015-1884-8.
62. J. J. Doyle, J. L. Doyle, A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15 (1987).
63. H. Takagi *et al.*, QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **74**, 174–183 (2013). doi:10.1111/tpj.12105.
64. M. Alonge *et al.*, Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. *Cell* **182**, 145–161.e23 (2020). doi:10.1016/j.cell.2020.05.021.
65. T. R. Ranallo-Benavidez, K. S. Jaron, M. C. Schatz, GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020). doi:10.1038/s41467-020-14998-3.
66. M. Alonge *et al.*, Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022). doi:10.1186/s13059-022-02823-7.
67. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi:10.1093/bioinformatics/bts635.

68. S. Kovaka *et al.*, Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019). doi:10.1186/s13059-019-1910-1.
69. D. Mapleson, L. Venturini, G. Kaithakottil, D. Swarbreck, Efficient and accurate detection of splice junctions from RNA-seq with Portcullis. *GigaScience* **7** (2018). doi:10.1093/gigascience/giy131.
70. T. D. Wu, C. K. Watanabe, GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005). doi:10.1093/bioinformatics/bti310.
71. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018). doi:10.1093/bioinformatics/bty191.
72. The Potato Genome Sequencing Consortium, Genome sequence and analysis of the tuber crop potato. *Nature* **475**, 189–195 (2011). doi:10.1038/nature10158.
73. L. Venturini, S. Caim, G. G. Kaithakottil, D. L. Mapleson, D. Swarbreck, Leveraging multiple transcriptome assembly methods for improved gene structure annotation. *GigaScience* **7** (2018). doi:10.1093/gigascience/giy093.
74. A. Kuno *et al.*, DAJIN enables multiplex genotyping to simultaneously validate intended and unintended target genome editing outcomes. *PLoS Biol.* **20**, e3001507 (2022). doi:10.1371/journal.pbio.3001507.
75. D. M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019). doi:10.1186/s13059-019-1832-y.
76. H. Li, Protein-to-genome alignment with miniprot. *Bioinformatics* **39**, btad014 (2023). doi:10.1093/bioinformatics/btad014.
77. A. J. Hart *et al.*, ENTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Mol. Ecol. Resour.* **20**, 591–604 (2020). doi:10.1111/1755-0998.13106.
78. M. Van Bel *et al.*, PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res.* **50**, D1468–D1474 (2022). doi:10.1093/nar/gkab1024.
79. R. Apweiler *et al.*, UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, D115–119 (2004). doi:10.1093/nar/gkh131.
80. P. Jones *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014). doi:10.1093/bioinformatics/btu031.
81. M. Van Bel *et al.*, TRAPID: an efficient online tool for the functional and comparative analysis of de novo RNA-Seq transcriptomes. *Genome Biol.* **14**, R134 (2013). doi:10.1186/gb-2013-14-12-r134.

82. R.-G. Zhang *et al.*, TESorter: An accurate and fast method to classify LTR-retrotransposons in plant genomes. *Hortic. Res.* **9**, uhac017 (2022). doi:10.1093/hr/uhac017.
83. T. Särkinen, M. Staats, J. E. Richardson, R. S. Cowan, F. T. Bakker, How to Open the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. *PLoS ONE* **7**, e43808 (2012). doi:10.1371/journal.pone.0043808.
84. L. D. Shepherd, T. G. B. McLay, Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *J. Plant Res.* **124**, 311–314 (2011). doi:10.1007/s10265-010-0379-5.
85. S. Li *et al.*, The development of a high-density genetic map significantly improves the quality of reference genome assemblies for rose. *Sci. Rep.* **9**, 5985 (2019). doi:10.1038/s41598-019-42428-y.
86. M. Stanke, B. Morgenstern, AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **33**, W465–467 (2005). doi:10.1093/nar/gki458.
87. G. S. C. Slater, E. Birney, Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* **6**, 31 (2005). doi:10.1186/1471-2105-6-31.
88. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014). doi:10.1093/bioinformatics/btu170.
89. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi:10.1093/molbev/mst010.
90. A. Stamatakis, T. Ludwig, H. Meier, RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**, 456–463 (2005). doi:10.1093/bioinformatics/bti191.
91. M. A. Miller, W. Pfeiffer, T. Schwartz, “Creating the CIPRES Science Gateway for inference of large phylogenetic trees” in *2010 Gateway Computing Environments Workshop (GCE)* (IEEE, 2010), pp. 1–8.
92. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Lam, GGTREE : an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017). doi:10.1111/2041-210X.12628.
93. L. Chen, G. B. Jameson, Y. Guo, J. Song, P. E. Jameson, The *LONELY GUY* gene family: from mosses to wheat, the key to the formation of active cytokinins in plants. *Plant Biotechnol. J.* **20**, 625–645 (2022). doi:10.1111/pbi.13783.
94. Z. H. Lemmon *et al.*, The evolution of inflorescence diversity in the nightshades and heterochrony during meristem maturation. *Genome Res.* **26**, 1676–1686 (2016). doi:10.1101/gr.207837.116.

95. T. Tamura, M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526 (1993). doi:10.1093/oxfordjournals.molbev.a040023.
96. D. Kuznetsov *et al.*, OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. *Nucleic Acids Res.* **51**, D445–D451 (2023). doi:10.1093/nar/gkac998.
97. M. Crow, H. Suresh, J. Lee, J. Gillis, Coexpression reveals conserved gene programs that co-vary with cell type across kingdoms. *Nucleic Acids Res.* **50**, 4302–4314 (2022). doi:10.1093/nar/gkac276.

**Acknowledgments:** We thank members of the Lippman laboratory and critical friend M. Bartlett for discussions and feedback on the manuscript. We acknowledge B. Seman and R. Santos from the Lippman lab for technical support, T. Mulligan, K. Schlecht, S. Qiao for assistance with plant care, S. Goodwin, S. Eskipehlivan, and D. McCombie for next-generation sequencing services, M. Frank and N. Lippman for photographs of *V. amazonica*, J. Emerson for his photograph of *S. pyrocanthos*, students in the Frary laboratory at Mt. Holyoke College for validation of genome edited plant phenotypes, and H. Golan for his work developing solpangenomics.com. We thank N. Tarnowsky, M. Pace, and the New York Botanical Garden for facilitating tissue sampling at the Steere Herbarium and reproduction of digitized collections held in the C.V. Starr Virtual Herbarium. We acknowledge the USF Institute for Systemic Botany for use of the digitized herbarium image of *T. hassleriana* and the use of public domain photos from R. Hannawacker of *A. greggii*, Zayda C. of *D. stramonium*, and P. Gibellini of *T. hassleriana*. We express our gratitude to the Peruvian government for permission to collect and sequence indigenous *Solanum* species (Ministerio de Agricultura, Dirección General Forestal y de Fauna Silvestre collection permits 084-2012-AG-DGFFSDGEFFS and 096-2017-SERFOR/DGGSPFFS, and genetic resource permit 008-2014-MINAGRI-DGFFS/DGEFFS). We are also grateful to our colleagues L.L. Giacomini and J.R. Stehmann for providing access to their *Solanum* collections and the Brazilian government for granting collecting and export permits (Cadastro A456CF7) for the RPPN Santuário do Caraça and Alto da Serra de Paranapiacaba. Lastly, we gratefully acknowledge the First Nations people of Australia for their care of the land and biodiversity where *S. cleistogamum* and *S. prinophyllum* were originally collected.

### Funding:

National Science Foundation Postdoctoral Fellowship in Biology (IOS-2305651) (JWS)

Postdoctoral fellowship (RYC2021-031999-I) funded by MCIN/AEI /10.13039/501100011033 and by "European Union NextGenerationEU/PRTR" (PG)

Predocotrual fellowship (FPU18/01742) from Spanish Ministerio de Ciencia, Innovación y Universidades (AA)

Predocotrual fellowship (PRE2019-089256) funded by MCIN/AEI/10.13039/501100011033 and by "ESF investing in your future" (GV)

Biologie et amélioration des Plantes departments of the French National Institute for Agriculture, Food, and Environment (INRAE) (MB1) (MB2)

The William Randolph Hearst Scholarship from the Cold Spring Harbor School of Biological Sciences (MJP)

5 Fonds de recherche du Québec en Nature et Technologies Postdoctoral Fellowship (EG)

The Sibbald Trust Research Fellowship (RH)

Grant of the German Research Association DFG (STE 1120/17-1) (NS)

NSF Planetary Biodiversity Initiative grant "PBI Solanum: a worldwide initiative" (DEB-0316614) (SK)

10 National Geographic Society Northern Europe Award GEFNE49-12 (TS)

National Institutes of Health grant R01MH113005 (JG)

Grant CIPROM/2021/020 funded by Conselleria d'Innovació, Universitats, Ciència i Societat Digital of the Generalitat Valenciana (JP)

15 Grant PID2021-128148OB-I00 funded by MCIN/AEI/10.13039/501100011033/ and "ERDF A way of making Europe" (JP)

Grant PDC2022-133513-I00 funded by MCIN/AEI/10.13039/501100011033/ and by "European Union NextGenerationEU/PRTR" (JP)

National Science Foundation Plant Genome Research Program grant IOS-2216612 (AF, JG, JVE, MCS, ZBL)

20 The Howard Hughes Medical Institute (ZBL)

### **Author contributions:**

Conceptualization: JWS, ZBL

Formal analysis: MJP, JG

25 Funding acquisition: JWS, PG, AA, GV, MB1, MP, EG, RH, NS, MB2, SK, TS, JG, JVE, MCS, JP, ZBL

Investigation: JWS, DA, PG, KMJ, JH, AA, GV, MP, SR, MB1, IG, AH, HS, BF, GMR, YG, KS, MJP, AP, EG, RH, SK, MB2, TS, JG, JVE, MCS, JP, SV, ZBL

30 Methodology: JWS, DA, PG, KMJ, AA, GV, MP, SR, IG, AH, BF, YG, KS, MJP, AP, SK, TS, JG, JVE, MCS, JP, SV, ZBL

Project administration: JWS, MCS, JP, ZBL

Resources: JWS, DA, PG, KMJ, JH, AA, GV, MP, SR, MB1, IG, AH, HS, BF, GMR, YG, KS, MJP, EG, RH, TDH, GCE, AG, TR, NS, SY, CB, MB2, AF, SK, TS, JG, JVE, MCS, YE, JP, SV, ZBL

Software: KMJ, SR, MJP, JG, MCS

Supervision: JP, SV, ZBL

Validation: JH, TDH, GCE, TR, NS, CB, MB2

Visualization: JWS, DA, PG, MJP, JG, TS, JP, SV, ZBL

Writing – original draft: JWS, ZBL

Writing – review & editing: JWS, AF, YE, ZBL

**Competing interests:** Authors declare that they have no competing interests.

**Data and materials availability:** Short-read sequencing data and genome assemblies reported in this paper are deposited at the NCBI Short Read Archive under BioProject ID PRJNA1073673. All other data are available in the main text or the supplementary materials. Plant materials used in this study are available upon request.

## Supplementary Materials

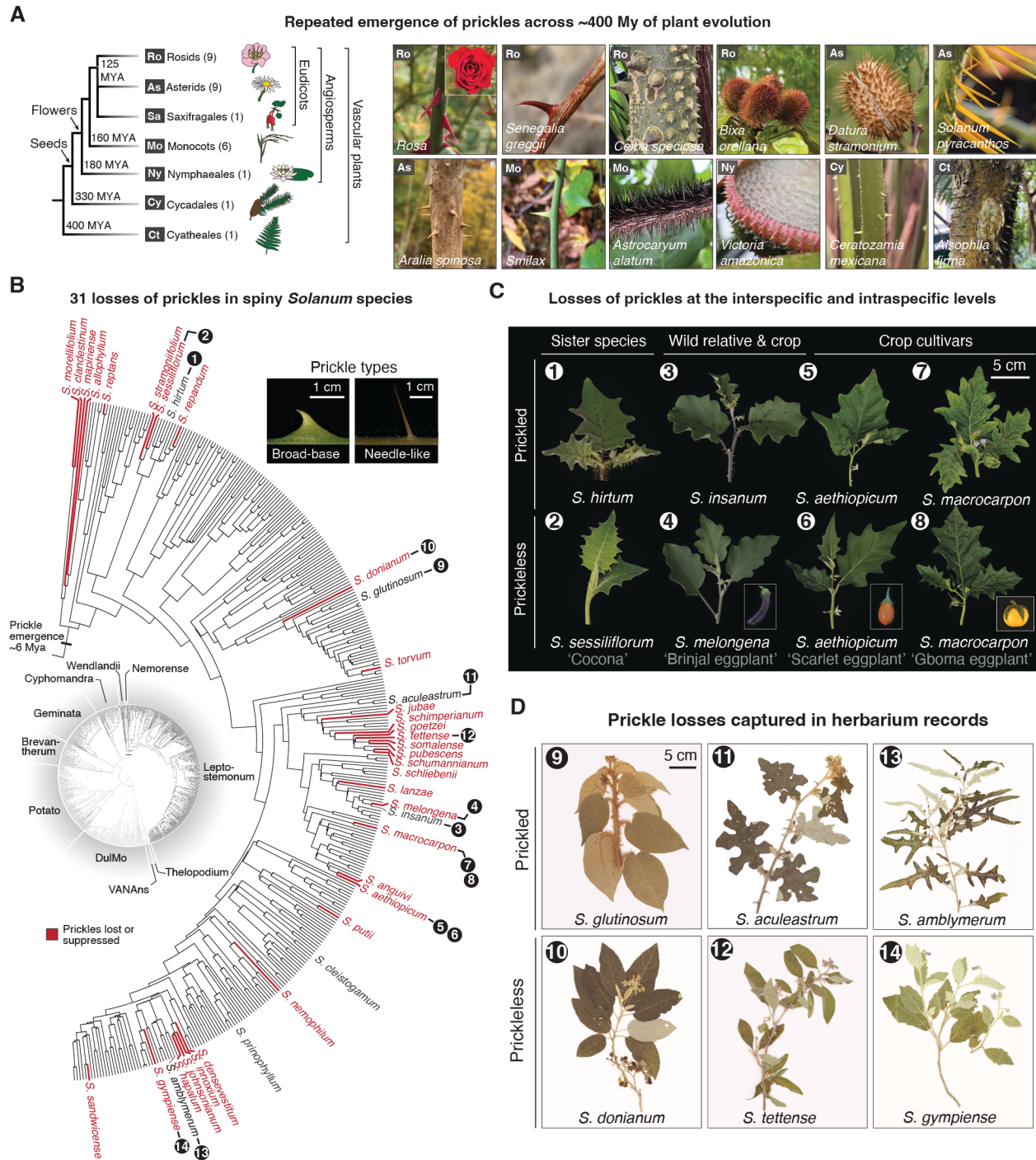
Materials and Methods

Supplementary Text

Figs. S1 to S5

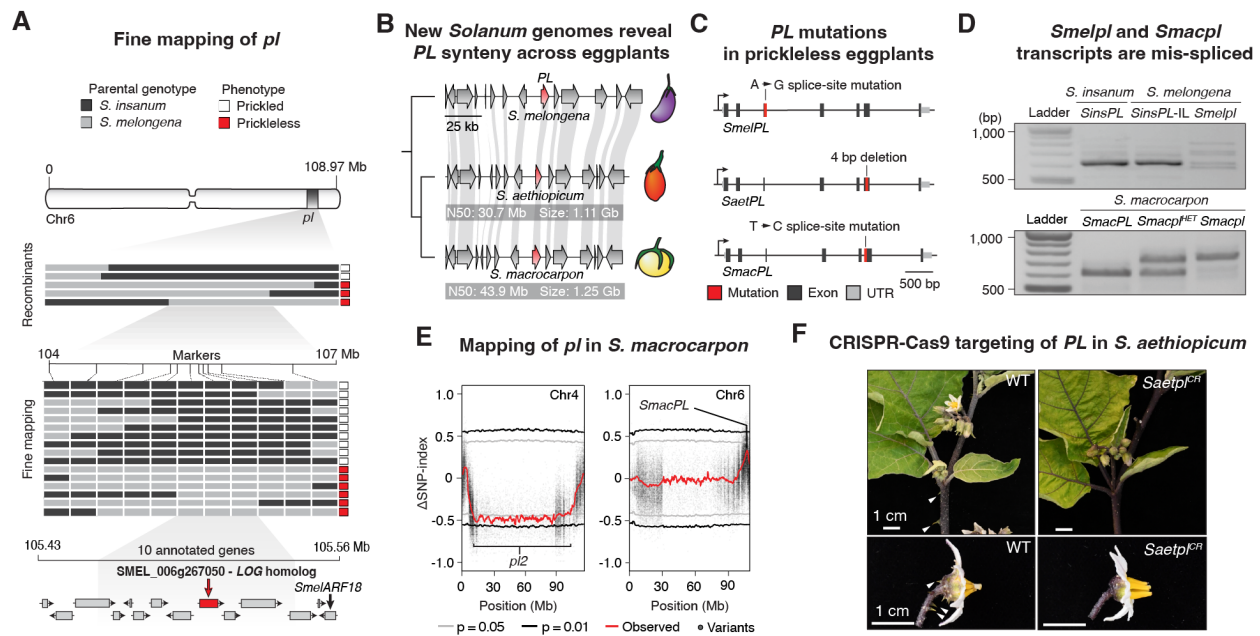
Tables S1 to S9

References (52–97)

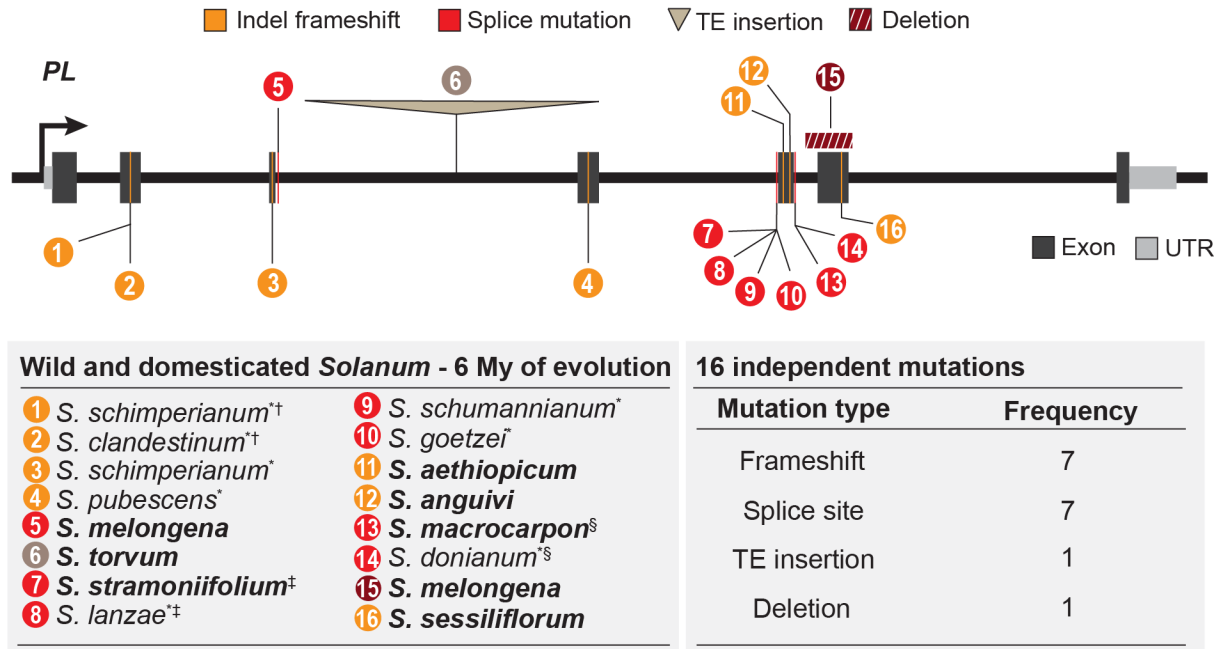


**Fig. 1. Prickles evolved convergently across vascular plants and were lost repeatedly in the spiny *Solanum* lineage.** (A) Phylogeny, from (5), and corresponding images of representative vascular plants that independently evolved prickles. Number in parentheses indicates number of identified independent evolutionary origins of prickles (B) Phylogenetic tree [adapted from (10, 11)] of the spiny *Solanum* (subclades Wendlandii, Nemorensis, and Leptostemonum) with species having lost prickles highlighted in red. Representative images of narrow and broad-based prickle morphologies are shown. (C and D) Images of *Solanum* taxa that have lost prickles captured from living (C) and herbarium (D) collections. Numbers correspond to species shown in (B).

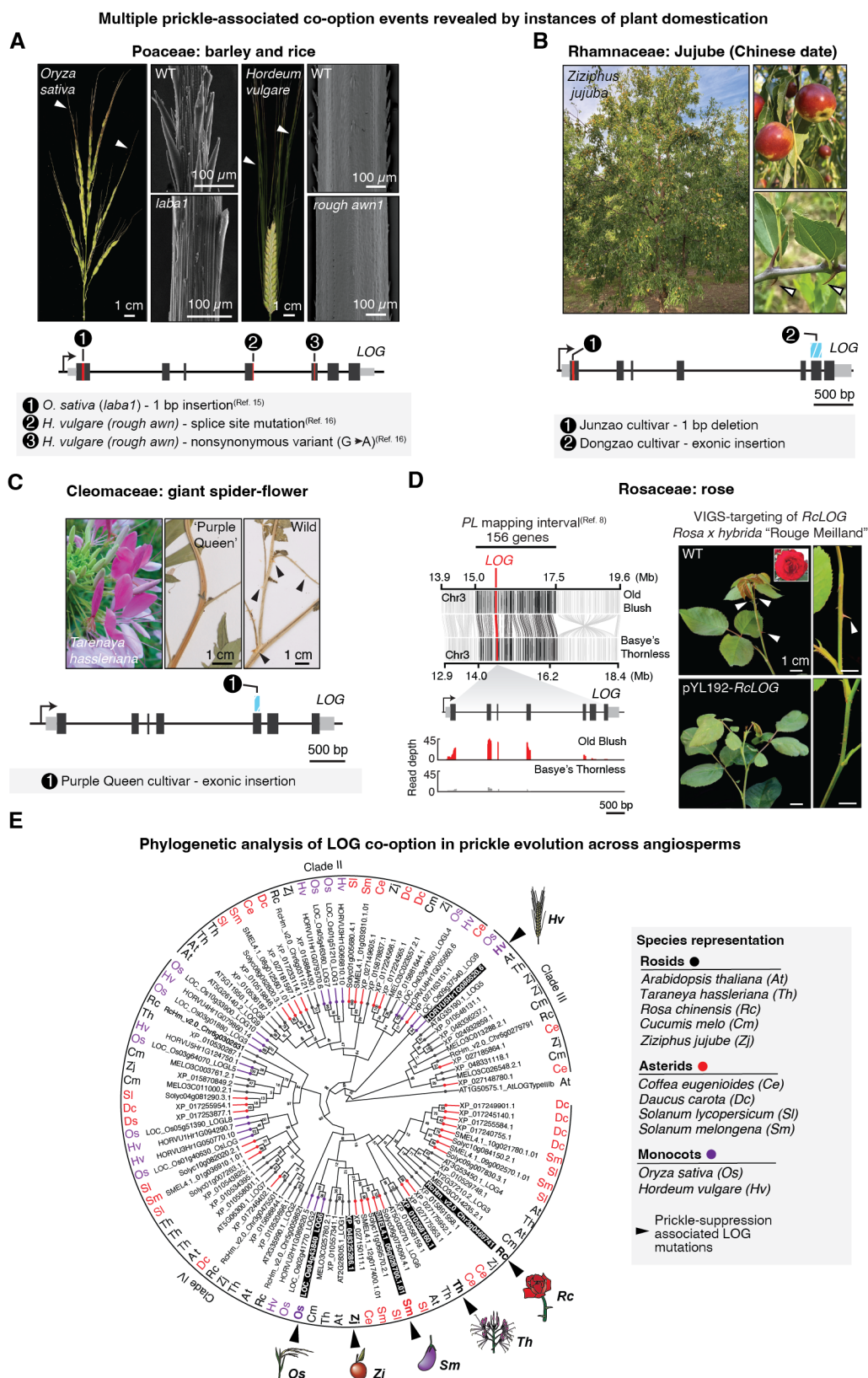




**Fig 2. Losses of prickles in three domesticated *Solanum* species are caused by independent mutations in a *LOG* cytokinin biosynthetic gene.** (A) Fine-mapping of *pl* in a Brinjal eggplant (*S. melongena*) x wild progenitor species (*S. insanum*) mapping population. (B) Genome sequencing and chromosome-scale assemblies of two African eggplants, the Scarlet eggplant (*S. aethiopicum*) and the Gboma eggplant (*S. macrocarpon*) reveals synteny of the *pl* locus. Genome summary statistics are indicated. (C) Independent mutations in a *LOG* gene in the *pl* interval in all three prickless crop species. (D) Mis-splicing of *PL* transcripts caused by the *pl* mutations in Brinjal eggplant *pl* (*Smelpl*) and Gboma eggplant *pl* (*Smacpl*) confirmed by RT-PCR. *SinsPL-IL* denotes an introgression of *S. insanum PL* into the Brinjal eggplant genomic background. (E) QTL-Seq identifies two loci that independently cause the prickless phenotype in Gboma eggplant. (F) Phenotypes resulting from CRISPR-Cas9 genome editing of *SaetPL* in a prickled *S. aethiopicum* accession. Arrowheads indicate prickles.

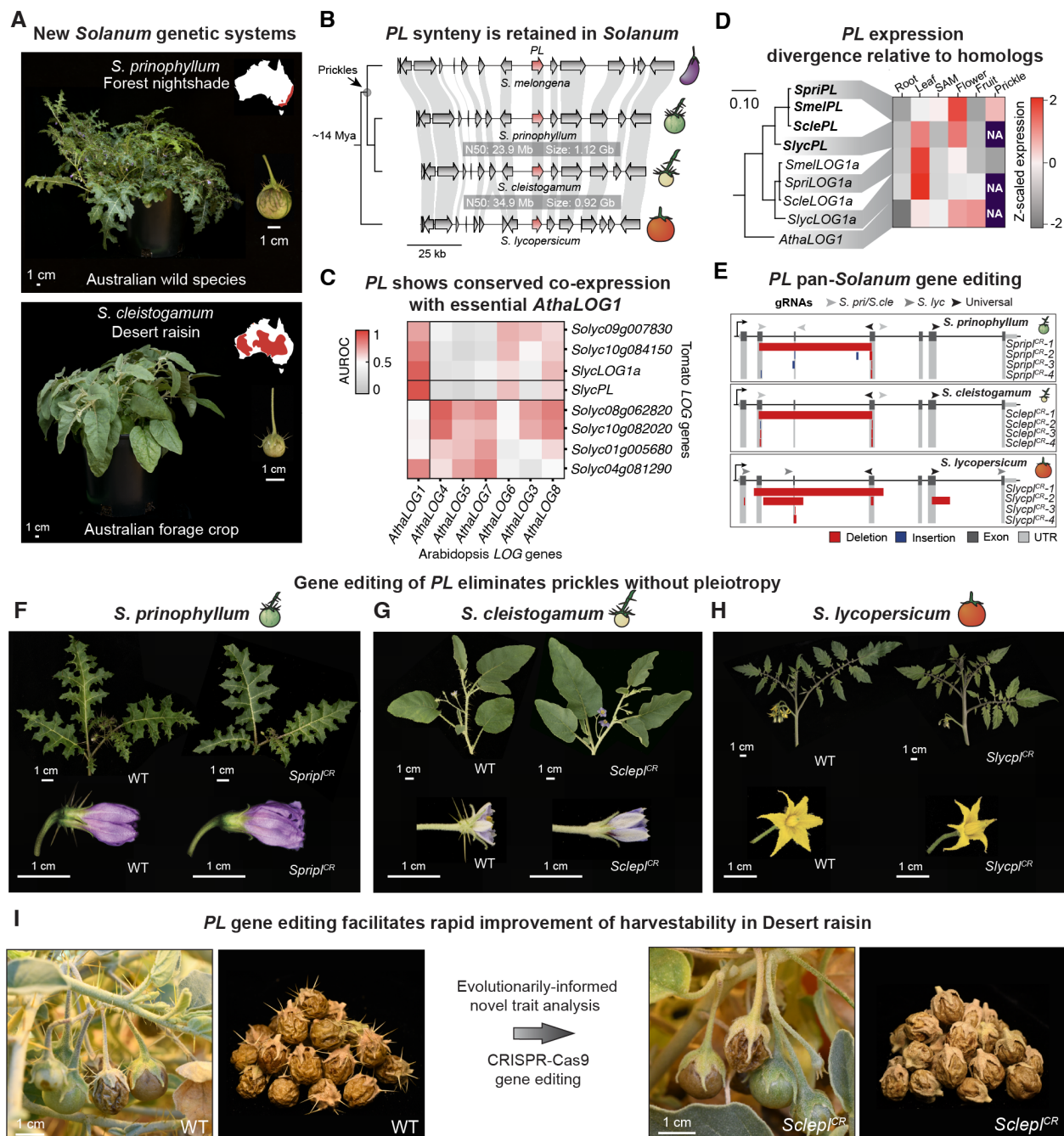


**Fig. 3. Mutations in *PL* are associated with prickles suppression across the spiny *Solanum*.** *PL* variants with strong probable deleterious effects on gene function identified in prickles-suppressed taxa but not in closely-related prickled sister taxa. Mutations are numbered and shown along with their corresponding species name and sample source in the table below. In the tables, bold text indicates cultivated species, (\*) indicates that genotyping was performed on archival herbarium samples, (†,‡,§) indicate species pairs that share identical but not necessarily ancestral mutations.



**Fig. 4.** Losses of convergently evolved prickles across angiosperms are associated with *LOG* mutations. (A to D) Instances of prickler suppression in angiosperms associated with *LOG* mutations depicted in corresponding *LOG* gene diagrams. (A) Images of rice and barley WT

inflorescences. Arrowheads indicate awns, which are shown for WT and mutant genotypes (rice, *laba1*; barley, *rough awn1*) by SEM. (B) Images of jujube trees, fruits, and stipular spines (arrowheads). Two less spiny cultivated varieties harbor two independent *LOG* mutations. (C) The ornamental giant spider flower (pictured) carries a mutated *LOG* gene in the sequenced ‘Purple Queen’ cultivar. Cultivated varieties bear fewer smaller prickles (arrowheads) than wild varieties, as reflected in herbarium samples. (D) (Left) Loss of prickles in rose maps to a ~2.5 Mb interval harboring a *LOG* gene with severely reduced expression in the prickless cultivar relative to the prickled cultivar. Syntenic genes within the mapping interval of the prickled ‘Old Blush’ and prickless ‘Basye’s Thornless’ parental lines are shown in black. Read pileups show average *LOG* expression in leaves of the parental genotypes ( $N = 3$ ). (Right) VIGS targeting of the candidate *LOG* gene leads to suppression of prickles in an ornamental rose hybrid. (E) Protein-based phylogenetic tree of the Arabidopsis LOG1 orthogroup defined by Orthofinder, from the indicated asterid (red), rosid (black), and monocot (purple) species. LOGs encoded by genes with mutations in prickle-suppressed taxa are indicated by arrowheads.



**Fig. 5. The *Solanum PL* gene was co-opted from an ancestral gene duplication event enabling non-pleiotropic editing of *PL* for crop improvement.** (A) Whole-plant and fruit images of the prickled wild species Forest nightshade (*S. prinophyllum*, top) and its close foraged berry-producing relative Desert raisin (*S. cleistogamum*, bottom). Red-shaded region in map insets indicates approximate species ranges in Australia based on reported observations (<http://www.flora.sa.gov.au/>). (B) Genome sequencing and chromosome-scale assemblies of Forest nightshade and Desert raisin reveals that *PL* interval synteny is conserved in Brinjal eggplant and tomato (*S. lycopersicum*). Genome summary statistics are indicated. (C) Heatmap depicting the predictability of identifying cross-species co-expressed genes among cross-species pairs of *LOG* homologs based on their respective co-expression relationships in tomato and Arabidopsis. A higher Area Under the Receiver Operating Characteristic (AUROC) curve score

indicates *LOG* homologs with increased conservation of their corresponding orthologous co-expressed genes. **(D)** Coding-sequence based maximum-likelihood phylogenetic tree of *Solanum PL* orthologs, their closely related paralog *LOG1a*, and *AthaLOG1* in comparable tissue types. Heatmap shows expression in matched tissues. **(E)** CRISPR-Cas9 gene editing strategy and resulting mutant alleles generated in Forest nightshade, Desert raisin, and tomato. **(F to H)** Phenotypes of WT and gene edited *pl* null mutants in Forest nightshade (F), Desert raisin (G), and tomato (H). Prickles are nearly completely suppressed (Forest nightshade) and eliminated (Desert raisin) obvious pleiotropic consequences. In tomato where *PL* was not co-opted for prickle development, *Slycpl<sup>CR</sup>* mutants resemble wild type. **(I)** Evolutionarily-informed trait analysis enables rapid and expedient removal of prickles for improved harvestability in *Solanum* crops.



Supplementary Materials for

**Convergent evolution of plant prickles by repeated gene co-option over deep time**

**Authors:** James W. Satterlee<sup>1,2</sup>, David Alonso<sup>3</sup>, Pietro Gramazio<sup>3</sup>, Katharine M. Jenike<sup>4,5</sup>, Jia He<sup>1,2</sup>, Andrea Arrones<sup>3</sup>, Gloria Villanueva<sup>3</sup>, Mariola Plazas<sup>3</sup>, Srividya Ramakrishnan<sup>4,5</sup>, Matthias Benoit<sup>6</sup>, Iacopo Gentile<sup>7</sup>, Anat Hendelman<sup>1,2</sup>, Hagai Shohat<sup>1</sup>, Blaine Fitzgerald<sup>1,2</sup>, Gina M. Robitaille<sup>1,2</sup>, Yumi Green<sup>8</sup>, Kerry Swartwood<sup>8</sup>, Michael J. Passalacqua<sup>7</sup>, Edeline Gagnon<sup>9,10</sup>, Rebecca Hilgenhof<sup>10</sup>, Trevis D. Huggins<sup>11</sup>, Georgia C. Eizenga<sup>11</sup>, Amit Gur<sup>12</sup>, Twan Rutten<sup>13</sup>, Nils Stein<sup>13,14</sup>, Shengrui Yao<sup>15,16</sup>, Adrien Poncet<sup>17</sup>, Clement Bellot<sup>17</sup>, Amy Frary<sup>18</sup>, Sandra Knapp<sup>19</sup>, Mohammed Bendahmane<sup>17</sup>, Tiina Särkinen<sup>10</sup>, Jesse Gillis<sup>1,20</sup>, Joyce Van Eck<sup>8,21</sup>, Michael C. Schatz<sup>4,5</sup>, Yuval Eshed<sup>22</sup>, Jaime Prohens<sup>3</sup>, Santiago Vilanova<sup>3</sup>, Zachary B. Lippman<sup>1,2,7\*</sup>

Corresponding author: [lippman@cshl.edu](mailto:lippman@cshl.edu)

**The PDF file includes:**

Materials and Methods

Figs. S1 to S5

Tables S1 to S9

References 52-97

## Materials and Methods

### Plant growth conditions

Brinjal eggplant and wild eggplant (*Solanum insanum*) used for the mapping of *pl* were grown in a growth chamber under 25 °C/18 °C day/night temperatures and a 16/8h light/dark photoperiod, following the protocol for wild materials described in (52). The other eggplant species used for mapping and phenotyping were sown and germinated in 96-cell flats in a greenhouse, under natural and supplemental ~250  $\mu\text{mol m}^{-2} \text{s}^{-1}$  light from high-pressure sodium bulbs (16 h light, 8 h dark) at daytime temperatures between 26-28 °C and night-time temperatures between 18-20 °C with a relative humidity of 40-60%. At ~4 weeks post-germination, seedlings were transplanted to 4 L pots containing high drainage PRO-MIX HP Mycorrhizae Growing Mix soil (Pro-Mix, PA, USA) supplemented with Osmocote fertilizer. Prickle phenotypes recorded for co-segregation analysis were performed in the field for the Scarlet eggplant (*S. aethiopicum*) and both the field (West Neck Field, Lloyd Harbor, NY, USA) and greenhouse for Gboma eggplant (*S. macrocarpon*). Field conditions included drip irrigation, regular soil fertilization, and manual weeding. All other *Solanum* species used in this study were grown in either the field or greenhouse, as specified (see **Table S6** for accession information).

Other plants imaged for this study were grown as follows. Rose plants were grown in the field or grown in an environmentally controlled greenhouse conditions at the Ecole Normale Supérieure of Lyon, France with 16 h light, 8 h dark day/night periods and 25 °C/19 °C day/night temperatures. Chinese date (*Z. jujuba*) images were taken of plants grown in an orchard at Los Lunas, NM, USA. Rice plants (*O. sativa*) were grown in field conditions or the greenhouse (*O. rufipogon*) and managed according to practices outlined in (53).

### Generation of interspecific *pl* mapping population in Brinjal eggplant (*S. melongena*)

The prickless Brinjal eggplant *S. melongena* 8104, hereafter MEL5, and the prickled eggplant wild relative *S. insanum* SLKINS-1, hereafter INS1 (54, 55), were selected as founding parents to map *pl*. Advanced backcrosses (BC3) of MEL5 with introgressions from INS1 were obtained as part of a previous project (45). The parents, their interspecific F1 hybrid, and 90 BC3 advanced backcrosses were phenotyped for the presence of prickles at the cotyledon and 3<sup>rd</sup>-4<sup>th</sup> true leaf development stages. Of these, 19 BC3 advanced backcrosses with prickles and 71 without prickles were genotyped with the high-throughput eggplant SPET platform (56, 57) to associate the presence of prickles with shared introgression intervals. After identifying the most promising genomic region, the prickled advanced backcross material with the shortest shared introgressed interval (BC3-33-3-1) was selfed to obtain a segregating BC3S1 population, which was used for fine mapping.

### Mapping of *pl* in Brinjal eggplant

A total of 622 BC3S1 seeds were germinated in growth chambers and individuals were screened for prickles two weeks post-germination. After one month, seedlings were transplanted to a greenhouse where they were phenotyped again. DNA from leaves was extracted following the SILEX protocol (58) and DNA yield and quality were measured spectrophotometrically using a NanoDrop™ ND-1000 (Thermo Scientific, Waltham, MA, USA). DNA integrity was checked by electrophoresis on an agarose gel (0.8%) (Condalab, Madrid, Spain) and diluted with ultra-pure water to a final concentration of 50 ng/ $\mu\text{L}$ .

Subsequently, primers were designed for the INS1 introgression based on the genome position of target SNPs used to perform SPET in eggplant (56, 57) and the v3.0 eggplant genome sequence (59). To further delimit the ~2.5 Mb *pl* interval and identify recombination breakpoints,



we performed PCR genotyping by HRM (High Resolution Melting) genotyping (60). Primers used for HRM genotyping are listed in **Table S9**. PCR reactions were performed on a Roche LightCycler® 480 System (Roche Diagnostics, Rotkreuz, Switzerland) using the MasterMix qPCR No-ROX PyroTaq EvaGreen 5x for HRM (CMB-Bioline, Madrid, Spain). Two microliters of the diluted DNA sample were mixed with 0.3 µL (10 µM) of each primer, 2 µL of MaterMix qPCR qPCR No-ROX PyroTaq EvaGreen 5x for HRM and ultra-pure water until a final volume of 10 µL. The PCR reaction conditions were as follows: an initial a pre-incubation step at 95°C for 15 min followed by 55 cycles of denaturation at 95°C for 10 s, annealing at 60°C for 15 s and elongation at 68°C for 15 s. A melting curve analysis (60°C to 95°C) was performed after amplification to check specificity of the reaction. The HRM generated data were analysed using the LightCycler 480 Software release 1.5.1.62 via  $T_m$  values and melting curve shapes. MEL5 and INS1 parents were used as a control to detect differences between homozygote and heterozygote genotypes. The *pl* interval was thereby narrowed to a region between SNP13 and SNP14 markers, corresponding to the 105,450,271-105,543,990 bp position of chromosome 6, with a total size of 93,719 bp.

#### RT-qPCR of PL in Brinjal eggplant (*S. melongena*)

Total RNA was isolated from the leaves of prickled and prickless Brinjal eggplant (*S. melongena*) and prickled *S. insanum* using 700 µL of Extrazol® EM300 (Blirt DNA, Gdansk, Poland), according to the manufacturer's specifications. Integrity was checked by 1% agarose gel electrophoresis, and purity and quantity were determined spectrophotometrically using a NanoDrop™ ND-1000 (Thermo Scientific, MA, USA). Total RNA (1 µg) was treated with DNase I (RNase-free) (Thermo Scientific, MA, USA) and used as template in a FastGene Scriptase Basic cDNA Synthesis kit reaction (Nippon Genetics Europe, Düren, Germany) primed with Oligo dT. RT-qPCR was performed using a Roche LightCycler® 480 System (Roche Diagnostics, Rotkreuz, Switzerland) thermocycler using qPCR MasterMix No-ROX PyroTaq EvaGreen 5x for HRM (CMB-Bioline, Madrid, Spain). The PCR program consisted of an initial step of pre-incubation denaturation at 95°C for 15 min followed by 45 cycles of denaturation at 95°C for 15 s, annealing at 60°C for 20 s and extension at 72°C for 20 s. Melt curve analysis was performed to evaluate and confirm the specificity of the PCR reactions. An endogenous control, the eggplant housekeeping CAC gene (Clathrin adaptor complex medium subunit, SMEL\_008g297560.1.01) was used as a normalization reference (61). Gene transcript expression level was analysed according to the relative quantitative accumulation by the  $2^{(-\Delta\Delta Ct)}$  method (See **Table S9** for primers). A two-tailed Student's t-test was used to determine statistical significance of inter-species differential expression.

#### RT-PCR of *PL* transcripts for splice isoform analysis in Brinjal eggplant and Gboma eggplant

Approximately 0.5 cm long leaf primordia were harvested from MEL5, INS1, and a line harboring a minimal *PL* introgression from INS1 in the MEL5 background. Similarly staged tissue was also harvested from prickless Gboma eggplant (*S. macrocarpon* PI 441914), prickled *SmacPL<sup>HET</sup>* individuals identified in a stock of Gboma eggplant (*S. macrocarpon* SOLA112), and prickled homozygous wild-type *SmacPL* individuals in an F2 population derived from a parental cross between *S. macrocarpon* PI 441914 and *S. macrocarpon* SOLA112. All tissues were collected from greenhouse grown plants at approximately 10:00-11:00 AM and flash frozen in liquid nitrogen in 1.5 mL microfuge tubes containing a 5/32 inch (~3.97 mm) 440 stainless steel ball bearing (BC Precision, TN, USA). Tubes containing tissue were placed in a -80 °C stainless steel tube rack and ground using an SPEX™ SamplePrep 2010 Geno/Grinder™ (Cole-Parmer, NJ, USA) for 2 min at 1440 RPM. Total RNA was extracted using TRIzol (Invitrogen, MA, USA)

according to the manufacturer's instructions for ground tissue. RNA was treated with DNase I and cleaned using an RNA Clean and Concentrator Kit (Zymo Research, CA, USA) according to the manufacturer's instructions. Purity and concentration of the resulting total RNA was assessed using a NanoDrop One<sup>C</sup> spectrophotometer (Fisher Scientific, MA, USA). Next, 1 µg of total RNA was used as input for Poly-T primed reverse transcription using the SuperScript IV VILO Master Mix (ThermoFisher, MA, USA) kit according to the manufacturer's instructions. PCR amplification was done with 1 µL of cDNA and a primer concentration of 10 µM in a 20 µL reaction using the KOD One<sup>TM</sup> PCR Master Mix (Toyobo, Osaka, Japan). Amplification was performed using a Mastercycler X50 thermocycler (Eppendorf, Hamburg, Germany) with an initial 2 min 98 °C denaturation step, 35 cycles consisting of 10s denaturation at 98 °C, 10 s annealing at 55 °C, and 10 s extension at 68 °C, followed by a final extension step of 2 min at 68 °C. Equal volumes (5 µL) of the resulting PCR product were loaded onto a 2% biotechnology grade Agarose I (VWR International, PA, USA) gel prepared with 1X TBE buffer and 0.0025% ethidium bromide. Electrophoresis was performed in an Owl<sup>TM</sup> D3-14 electrophoresis box (Thermo Scientific, MA, USA) containing 1X TBE buffer for 45 min at 130 V delivered from an Owl<sup>TM</sup> EC 300 XL power supply (Thermo Scientific, MA, USA). The electrophoresis results were visualized under UV light using a Bio-Rad ChemiDoc<sup>TM</sup> XRS+ (Bio-Rad, CA, USA) imaging platform and ImageLab<sup>TM</sup> (Bio-Rad, CA, USA) software. Relevant primer sequences are listed in **Table S9**.

To determine the identity and relative abundance of transcript isoforms across the different genotypes, PCR products were cloned using a StrataClone Blunt PCR Cloning Kit (Agilent, CA, USA) according to the manufacturer's instructions. PCR was conducted on single colonies swabbed into 10 µL sterile water and a 1 µL aliquot of the suspended culture was used as a PCR reaction template using 10 µM M13F/R primers with KOD One<sup>TM</sup> PCR Master Mix. PCR amplification was performed with a 2 min 98 °C denaturation step, 35 cycles consisting of 10 s denaturation at 98 °C, 10 s annealing at 55 °C, and 10 s extension at 68 °C and followed by a final extension step of 2 min at 68 °C using a Mastercycler X50 thermocycler. PCR products were purified using 1.8X volume Ampure XP beads according to the manufacturer's instructions and submitted for Sanger sequencing (Azenta Genewiz, NJ, USA).

#### Co-segregation and mapping of *pl* in Scarlet eggplant (*S. aethiopicum*) and Gboma eggplant (*S. macrocarpon*)

The *SaetPL* allele co-segregation analysis in Scarlet eggplant was performed on an F2 population derived from an intraspecific cross between the prickled accession SC102 and the prickless accession PI 424860. Field grown plants were phenotyped for the presence/absence of prickles and young leaf tissue was collected and DNA extracted using the cetrimonium bromide (CTAB) method (62). Primers (see **Table S9**) were designed to amplify the exonic region of *PL* harboring the mutant *Saetpl* deletion allele identified in the genome of the prickless parental accession. PCR reactions were prepared using 1 µL of DNA and 10 µM primers in a 10 µL reaction volume with KOD One<sup>TM</sup> PCR Master Mix (Toyobo, Osaka, Japan). Amplification was performed on a Mastercycler X50 thermocycler (Eppendorf, Hamburg, Germany) with a 2 min 98 °C denaturation step, 35 cycles consisting of 10 s denaturation at 98 °C, 10 s annealing at 55 °C, and 10 s extension at 68 °C, followed by a final extension step of 2 min at 68 °C. PCR products were purified using a 1.8X volume of AMPure XP beads (Beckman Coulter, CA, USA) and submitted for Sanger sequenced (Azenta Genewiz, NJ, USA) for *SaetPL* genotyping. Genotypes were then associated with prickle phenotypes. Chi-squared goodness of fit tests were performed with GraphPad.

For *SmacPL* in Gboma eggplant the same initial co-segregation analysis was performed as was done for Scarlet eggplant, above (see **Table S9** for primers). Multiple F2 populations derived from an intraspecific cross between prickless Gboma eggplant (*S. macrocarpon* PI 441914) and prickled Gboma eggplant (*S. macrocarpon* SOLA112) were generated. QTL-Seq was performed on a single F2 population made up of the progeny of a single prickled F1 parent. From a population of 132 individuals, the DNA from a random selection of 25 prickless and 31 prickled F2 individuals, along with 8 prickless parental accession individuals (*S. macrocarpon* PI 441914) was extracted from young leaf tissue using a DNeasy Plant Pro Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions for high-polysaccharide content plant tissue. Tissue used for extraction was ground using an SPEX<sup>TM</sup> SamplePrep 2010 Geno/Grinder<sup>TM</sup> (Cole-Parmer, NJ, USA) for 2 min at 1440 RPM. Sample DNA (1 µL assay volume) concentrations were quantified with Qubit 1X dsDNA HS buffer (ThermoFisher, MA, USA) on a Qubit 4 fluorometer (ThermoFisher, MA, USA) according to the manufacturer's instructions. Separate pools were made for the parental sample, the prickled F2 individuals, and the prickless F2 individuals, with an equivalent mass (3 µg) of DNA pooled from each individual. Next, the DNA pools were purified using a 1.8X volume of AMPure XP beads (Beckman Coulter, CA, USA) and the DNA concentration and purity assayed by Qubit and a NanoDrop One<sup>C</sup> spectrophotometer (Fisher Scientific, MA, USA), respectively.

Paired-end sequencing libraries for QTL-Seq analysis were prepared using >1 µg of pooled DNA with a KAPA HyperPrep PCR-free kit (Roche, Basel, Switzerland) according to the manufacturer's instructions. Indexed libraries were pooled for sequencing on a NextSeq 2000 P3 chip (Illumina, CA, USA). Mapping was performed using the end-to-end pipeline implemented in the QTL-Seq software package (v2.2.4, <https://github.com/YuSugihara/QTL-seq>) (63) with reads aligned against the *S. macrocarpon* (Smac3, PI 441914) genome assembly. After QTL-Seq analysis revealed a second interval, *pl2*, an additional co-segregation analysis was performed across multiple F2 populations, genotyping *SmacPL* and the *pl2* locus on chromosome 4 identified by QTL-Seq. For the *pl2* interval, an A-to-T SNP at position Chr4:20,401,670, which modifies a DraI restriction site, was used as the basis for a cleaved amplified polymorphic sequences (CAPS) genotyping assay. Following PCR amplification using the same conditions described for the *SaetPL* genotyping above (see **Table S9** for primer details), 5 µL of PCR product was added to a 10 µL reaction containing 0.2 µL DraI (New England BioLabs, MA, USA) and 1 µL rCutSmart<sup>TM</sup> Buffer (New England BioLabs, MA, USA) and incubated for 2-4 hours at 37 °C. The reactions were then loaded onto a 1% agarose gel and electrophoresed in an Owl<sup>TM</sup> D3-14 electrophoresis box (Thermo Scientific, MA, USA) containing 1X TBE buffer for 30 min at 180 V delivered from an Owl<sup>TM</sup> EC 300 XL power supply (Thermo Scientific, MA, USA). The electrophoresis results were visualized under UV light using a Bio-Rad ChemiDoc<sup>TM</sup> XRS+ (Bio-Rad, CA, USA) imaging platform and ImageLab<sup>TM</sup> (Bio-Rad, CA, USA) software. Genotypes assigned based on visualized PCR product cleavage patterns and associated with prickle phenotypes. Chi-squared goodness of fit tests were performed with GraphPad.

### Genome assembly and scaffolding

Reference quality genome assemblies for *S. aethiopicum*, *S. macrocarpon*, *S. prinophyllum*, and *S. cleistogamum* (see **Table S6** for accession information) were generated using a combination of long-read sequencing (Pacific Biosciences, CA, USA) for contigging and optical mapping (Bionano Genomics, CA, USA) for scaffolding. High-molecular weight DNA used for sequencing was extracted from greenhouse grown 4 week-old seedlings germinated in 96-cell flats and dark treated for 48 hrs prior to flash freezing. High molecular weight DNA was extracted according to the method described in (64). Two PacBio Sequel IIe flow cells (Pacific Biosciences,

CA, USA) were used for the sequencing of each sample (average read N50 = 11,221 bp, average coverage = 53X, average read QV = 83.28). Prior to assembly, we counted k-mers from raw reads with KMC3 (version 3.2.1) and estimated genome size, sequencing coverage, and heterozygosity with GenomeScope2.0 (65). Sequencing reads from each sample were assembled with hifiasm (46) exact parameters and software version varied between samples based on the level of estimated heterozygosity and are reported in **Table S4**. Post assembly, the draft contigs were screened for possible microbial contamination as described in (64).

Optical mapping (Bionano Genomics, CA, USA) was performed for each sample to facilitate scaffolding. Scaffolding with optical maps was performed using the Bionano solve Hybrid Scaffold pipeline with default parameters (<https://bionanogenomics.com/support/software-downloads/>). Hybrid scaffold N50s ranged from 75,028,200 bp to 107,486,353 bp (see **Table S4** for more detail including Bionano molecules per sample). High-throughput chromosome conformation capture (Hi-C) from Arima Genomics, CA, USA was performed for one sample, *S. prinophyllum*, to finalize scaffolding. 311,616,288 reads were integrated with the Juicer (v0.7.17-r1198-dirty) pipeline. Next, misjoins and chromosomal boundaries were manually curated in the Juicebox (v1.11.08) application. Chromosomes were named based on sequence homology, determined with RagTag (66) scaffold (v2.1.0, default parameters), with the phylogenetically-closest finished genome (see **Table S4** for details). Finally, small contigs (< 50,000 bp) with > 95% of the sequence mapping to a named chromosome were removed. Additionally, small contigs (< 100,000 bp) with > 80% of the sequence mapping to a named chromosome that contained one or more duplicated BUSCO genes, but no single BUSCO genes, were also removed using a python script. Using merquary (47) with the HiFi data, the final consensus quality of the assemblies was estimated as QV=51.1333 on average and a completeness of 99.2741% on average.

## Genome annotation

The gene annotation pipeline involved several crucial steps. Initially, the quality of raw RNA-Seq reads underwent assessment using FastQC v0.11.9. Subsequently, reference-based transcripts were generated using STAR v2.7.5c (67) and Stringtie2 v2.1.2 (68) workflows. To refine the data, invalid splice junctions from the STAR aligner were filtered out utilizing Portcullis v1.2.0 (69). Orthologs with coverage above 50% and 75% identity were lifted from Heinz v4.0 (48) and Eggplant v4.1 (17) via Liftoff v1.6.3 (49) using parameters --copies,--exclude\_partial and employing both Gmap version 2020-10-14 (70) and Minimap2 v2.17-r941 (71) aligners. In addition, protein evidence from several published Solanaceae genomes (17, 48, 72), and the UniProt/SwissProt database were utilized to support gene annotation. Structural gene annotations were generated through the Mikado v2.0rc2 (73) framework, leveraging evidence from the Daijin pipeline (74). Additionally, microsynteny and orthology to Heinz v4.0 and Eggplant v4.0 were assessed using Microsynteny and Orthofinder v2.5.2 (75). Correction of gene models with inframe stop codons utilized Miniprot2 (76) protein alignments from Heinz v4.0 and Eggplant v4.1. Furthermore, gene models lacking start or stop codons were adjusted by placing them within 300 base pairs of the nearest codon location using a custom python script.

For functional annotation, ENTAP v0.10.8 (77) integrated data from diverse databases such as PLAZA dicots (5.0) (78), Uniprot/Swissprot (79), TREMBL, RefSeq, Solanaceae proteins, and InterProScan5 (80) with Pfam, TIGRFAM, Gene Ontology, and TRAPID (81) annotations. Finally, the annotated data underwent a series of filtering steps, excluding proteins shorter than 20 amino acids, those exceeding three times the length of functional orthologs, proteins lacking assigned orthologs that have unknown function, and transposable element (TE) genes, which were removed using the TESorter (82) pipeline.

The completeness of the gene models was determined by assessing single-copy orthologs through BUSCO5 (50) in protein mode, comparing against the solanales\_odb10 database. Additionally, the presence or absence of a curated set of 180 candidate genes known to be crucial in QTL studies was examined. Genome annotation summary statistics are presented in **Table S5**.

#### DNA extraction, *PL* PCR amplification, and sequencing of *Solanum* species

For herbarium samples, tissue (~ 1-2 cm<sup>2</sup>) was excised with a scalpel from mounted herbarium specimens held in the Steere Herbarium Collection at the New York Botanical Garden, New York, USA (see **Table S6** for voucher information). Silica gel-preserved tissues were obtained from the Royal Botanic Garden Edinburgh. Tissue was obtained from species documented to be prickless and samples were inspected for the absence of prickles. Likewise, tissue was also obtained from closely-related prickled sister species for sequence comparison. In total, 8 putative loss-of-function coding or splice-site mutations were identified in herbarium samples across 7 species. Two distinct mutations were identified in separate *S. schimperianum* specimens. For technical reasons, it was not possible to survey the complete exonic space of all genes - primer binding site divergence and low DNA sample quality precluded complete survey of exonic survey in most cases. Among prickless herbarium specimens in which a complete exonic survey was possible, 7 lacked an obvious putative loss-of-function coding or splice-site mutation. Among the remaining prickless herbarium material with incomplete exonic sequence coverage, 22 specimens had no detectable putative loss-of-function coding or splice-site mutations. However, this does not preclude the possibility that *PL* mutations exist in these samples. Such mutations could lie in unsampled regions of the gene body or in non-coding regions. In addition, other loci (i.e. *pl2* in *S. macrocarpon*) may also contribute to unexplained prickle losses in *Solanum*. Meanwhile, across 30 sampled prickled sister taxa, none of the reported loss-of-function *pl* alleles were found and no other putative loss-of-function coding or splice-site mutations identified.

To extract DNA from herbarium samples, ~25 mg of tissue was added to a 2 mL microfuge tube containing a 5/32 inch (~3.97 mm) 440 stainless steel ball bearing (BC Precision, TN, USA) and ground for 2 min at 1440 RPM in an SPEX<sup>TM</sup> SamplePrep 2010 Geno/Grinder<sup>TM</sup> (Cole-Parmer, NJ, USA). Before beginning the extraction the bench area and pipettes were cleaned with a 10% bleach solution to reduce the risk of environmental DNA contamination. DNA extraction was performed using a modified version of the CTAB/STE-based extraction methods described in (83) and (84). First, 1 mL of freshly-prepared STE (0.25 M sucrose, 0.03 M Tris pH 8, 0.05 M ethylenediaminetetraacetic acid [EDTA]) was added to the ground tissue and vortexed for 5 s. The sample was then centrifuged at 2000 x g for 10 min and the wash and centrifugation in STE buffer was repeated once. Next, 1 mL 2X CTAB and 4 µL of 1 M dithiothreitol (DTT) were added to the precipitated tissue and incubated with gentle rotating agitation in a hybridization oven for 1 hr at 65 °C. Following incubation, organic extraction was performed by addition of 750 µL of 24:1 chloroform/isoamyl alcohol and the samples were inverted 10 times. Phase separation was achieved by centrifugation for 10 min at 16,200 x g and ~900 µL of the aqueous phase was transferred to a new tube. To precipitate the DNA, 3 µL of 15 mg/mL GlycoBlue<sup>TM</sup> coprecipitant (ThermoFisher Scientific, MA, USA) was added and well-mixed followed by 600 µL of chilled -20 °C isopropanol. DNA was allowed to precipitate at -20 °C for 1 hr. After precipitation, the samples were centrifuged at 20,000 x g for 30 min in a centrifuge chilled to 4 °C. The resulting DNA pellet was washed once with 70% ethanol and re-suspended in 50-100 µL 0.1X TE pH 8 buffer.

Because DNA from archived dried plant tissues is often fragmented (83), an exon screening based approach was taken to identify mutations in the CDS or immediately flanking splice

acceptor/donor sites. Primers were designed for intronic regions ~20-30 bp surrounding exons based on the sequence of *PL* from closely-related species for which whole-gene amplification of *PL* was possible due to the availability of intact DNA isolated from fresh, living tissue (see **Table S9** for primers). PCR amplification was performed with 1-2  $\mu$ L of DNA, primers at 10  $\mu$ M, in a 10  $\mu$ L reaction using the KOD One<sup>TM</sup> PCR Master Mix (Toyobo, Osaka, Japan). Amplification was performed using on a Mastercycler X50 thermocycler (Eppendorf, Hamburg, Germany) with a 2 min 98 °C denaturation step, 35-40 cycles consisting of 10 s denaturation at 98 °C, 10 s annealing at 48-62 °C, and 10 s extension at 68 °C, followed by a final extension step of 2 min at 68 °C. PCR products were purified using 1.8X volume Ampure XP beads (Beckman Coulter, MA, USA) according to the manufacturer's instructions. For high-yielding PCR reactions, the PCR products were directly Sanger sequenced (Azenta Genewiz, NJ, USA). For low-yielding PCR reactions, the PCR products were blunt cloned and sequenced as done for the *PL* transcript splicing isoform analysis described above. Sequences were then aligned (Geneious alignment) to either Brinjal eggplant *PL* gene body and screened for mutations using Geneious version 2022-07-07 software (see **Table S7**).

For the genotyping of *PL* in prickless species for which fresh tissue was available, a CTAB-based method was used to obtain DNA (62). Primers (see **Table S9**) anchored in the first and final exon of *PL* were used to PCR amplify the entire gene body. Touchdown PCR amplification was performed using on a Mastercycler X50 thermocycler (Eppendorf, Hamburg, Germany) with an initial 2 min 98 °C denaturation step, 16 cycles consisting of an initial 10 s annealing at 68 °C thereafter descending 1 °C/cycle, a 1 min extension at 68 °C, and a 10 s 98 °C denaturation. Next, the reaction proceeded through 14-19 cycles of annealing for 10 s at 60 °C, extension for 1 min at 68 °C and 10 s denaturation at 98 °C followed by a final extension of 2 min at 68 °C. PCR products were purified using 1.8X volume Ampure XP beads (Beckman Coulter, MA, USA) according to the manufacturer's instructions. PCR products were then blunt cloned, as described above, and whole-plasmid sequenced (Plasmidsaurus, OR, USA) to detect variants across the *PL* gene body. In all 4 such cases, candidate *PL* mutations were identified. See **Table S7** for identified exonic *PL* mutations in prickless species and the sequences of aligned prickled sister taxa).

#### Analysis of published mapping and RNA-Seq data from rose

The major QTL for prick formation on rose stems was previously mapped on chromosome 3 using a mapping population derived from a cross between of *R. chinensis* 'Old Blush' (OB) and *R. wichurana* Basye's Thornless (BT) (9, 85). All genes in the Prickles QTL interval were *de novo* annotated in the genome of *R. wichurana* BT using the gene prediction tool AUGUSTUS (86). Predicted annotations were refined by mapping the OB protein to the BT genomic regions with Exonerate (87). Synteny is based on the match found by Exonerate. Previously published RNA sequencing data were used to examine expression of the LOG homolog contained within the rose mapping interval. Reads were trimmed using trimmomatic v0.39 (88) and then mapped to the *Rosa chinensis* genome using STAR v2.7.5c (67). Read pileups at 1 bp resolution were visualized by converting resulting bam files to bigwig format.

#### Virus induced gene silencing (VIGS) in rose

The TRV constructs pYL192 (TRV1) and pYL156 (TRV RNA2) were described in (25). A 300 bp sequence between nucleotide 49 and 348 after start codon of the *LOG* gene was synthesized (idtDNA.com), then cloned into pYL156 using EcoRI and BamHI restriction sites (see **Table S9** for sequence of insert). The integrity of the resulting LOGpTRV2 construct was then confirmed by sequencing. pYL192 (TRV1) and LOGpTRV2 were transferred into *Agrobacterium*

*tumefaciens* EHA105 and then used for rose transformation. Agrobacteria harboring pYL192 (TRV1) or LogpTRV2 were plated on YEB medium containing 50 mg/L Kanamycin and 25 mg/L Rifampicin, and then grown for three days at 28°C. Agrobacteria were then collected in infiltration medium (YEB supplemented with 10mM MgCl<sub>2</sub>, 10mM MES, 200μM acetosyringone) at a final concentration of OD<sub>600</sub>=1.5.

After 2-3 h incubation at room temperature, *Rosa x hybrida* 'Rouge Meilland' cuttings were transfected by vacuum infiltration as follows. Three-month-old cuttings were immersed in a solution containing agrobacteria harboring pYL192 (TRV1) and LOGpTRV2 (1:1 ratio) supplemented with 300μL/L Silwet. Vacuum was maintained for 1 min at -0.04mPA twice with agitation in between. Cuttings were then planted in soil and grown in a growth chamber at a temperature of 17/24 °C (night/day) with a photoperiod of 8h/16h (night/day). Developing shoots were analyzed for prickle development three weeks after infiltration.

#### Identification of LOG mutations in Chinese date (*Ziziphus jujuba*) and giant spider-flower (*Tarenaya hassleriana*)

Published genomes of Chinese date (*Z. jujuba*) were screened for mutations in LOG homologs. Candidate exonic loss-of-function mutations were identified in the cultivated spine-suppressed *Z. jujuba* cultivar Dongzao (GenBank assembly: GCA\_000826755.1) and *Z. jujuba* cultivar Junzao (GenBank assembly: GCA\_001835785.2) in a *LOG* homolog corresponding to LOC107411352. Mutations were not detected in this gene in the spiny wild relative *Z. jujuba* var. *spinosa* (GenBank assembly: GCA\_020796205.1). For giant spider-flower a mutation was identified in a *LOG* homolog (LOC104826905) in the reference genome (GenBank assembly: GCA\_000463585.1).

#### Phylogenetic analysis of angiosperm *LOG* homolog

*LOG* genes were identified across plant genomes using Orthofinder (75). for the chosen eleven species, the protein sequences of ninety-nine proteins were aligned using MAFFT v7.402 (89). The best scoring maximum likelihood tree was inferred using default parameters, with 1000 bootstrap replicates in RAxML v8.2.12 (90) via CIPRES (91). The tree was visualized in R using the ggtree package (92). The AtLOG\_TypeIIIb (AT1G50575.1) protein sequence was used as an outgroup. Branch support was provided by 1000 replicate bootstrapping. The *LOG* sub-clades were designated after the naming convention used in (93).

#### Forest nightshade (*S. prinophyllum*) expression atlas

Washed and dried roots, leaves, flowers, fruits (with seeds removed), and developing prickles (~0.25-0.50 cm in length) were collected from 2 month-old Forest nightshade (*S. prinophyllum*) plants. All tissues were collected in 3-4 replicates, each from different plants. Total RNA was extracted using Quick-RNA MicroPrep Kit (Zymo Research) and treated with DNase I (Zymo Research) according to the manufacturer's instructions. RNA concentration and quality were analyzed using Thermo Scientific™ NanoDrop™ One<sup>C</sup> Spectrophotometer. Libraries for RNA-sequencing were prepared by KAPA mRNA HyperPrep Kit (Roche, Basel, Switzerland). Paired-end 100-base sequencing was conducted on the NextSeq 2000 P3 sequencing platform (Illumina, CA, USA). Reads were trimmed using trimmomatic v0.39 (88) and then mapped to the Spr1 genome using STAR v2.7.5c and expression computed in transcripts per million (TPM). Expression data from the Forest nightshades in vegetative meristem generated in a previous study was processed in the same fashion (94). Gene expression data from ATHENA ([https://athena.proteomics.wzw.tum.de/master\\_arabidopsisshiny/](https://athena.proteomics.wzw.tum.de/master_arabidopsisshiny/)) for the *Arabidopsis LOG1* ortholog in equivalent tissues was downloaded and data for *LOG* homologs in tomato was



downloaded from <https://www.solpangenomics.com>. Expressed genes were clustered phylogenetically based on alignment of their CDS sequences using the Clustal Omega function in Geneious (version 2022-07-07). A maximum likelihood tree using the Tamura-Nei model (95) was used for tree building.

#### Generation of CRISPR-Cas9-induced mutants and phenotyping

CRISPR Guide RNAs to target *PL* across *Solanum* species were designed using Geneious. The Golden Gate cloning approach as described in (29) was used to create multiplexed gRNA constructs. Plant regeneration and *Agrobacterium tumefaciens*-mediated transformation of *S. prinophyllum* and tomato were performed according to (51). The same methods were also used for *S. aethiopicum* and *S. cleistogamum* with two modifications. For *S. cleistogamum*, plant regeneration, the medium was supplemented with 0.5 mg/L zeatin instead of 2 mg/L and for the selection medium, 75 mg/L kanamycin was used instead of 200 mg/L. For *S. aethiopicum*, the protocol was the same as for *S. cleistogamum*, except the fourth transfer of transformed plantlets is done onto media supplemented with 50 mg/L kanamycin. Seed germination time in culture can vary between species and batches of harvested seeds. Typically, *S. prinophyllum* germination took 8 – 10 days and *S. cleistogamum* behaved similarly to tomato, germinating in 6 - 8 days.

First generation transgenic (T0) plants ( $N \geq 3$ ) were screened for prickles suppression phenotypes relative to non-transgenic regenerated control plants. DNA was extracted from T0 individuals using the CTAB method (62) and primers (see **Table S9**) were used to amplify the entire *PL* gene body, which was cloned, and whole-plasmid sequenced as was done for the pan-*Solanum* screening for mutations in the *PL* gene body. For each T0 individual, 5-8 clones were sequenced to verify the association of the pricklesless phenotype and null multiallelic or monoallelic T0 phenotype. For *S. prinophyllum* and *S. lycopersicum*, CRISPR alleles advanced to the T1 generation showed similar phenotypes to the T0.

#### Scanning Electron Microscopy

For prickles imaging in barley (*Hordeum vulgare*), freshly isolated awn samples were fixed with 2% glutaraldehyde in 50 mM phosphate buffer pH 7.0 for 16h at 8°C. After two 5 min washes with distilled water, samples were dehydrated in an ascending ethanol series of 30%, 50% 70%, 90%, 100%, and a second 100% ethanol wash with each step lasting 10 min after which they were critical point dried in a Quorum K850 critical point dryer (Quorum Technologies Ltd., <https://www.quorumtech.com>). Dehydrated samples were placed onto carbon adhesive discs, gold coated in an Edwards S150B sputter coater (Edwards High Vacuum Inc., Burgess Hill, UK) and examined in a Zeiss Gemini300 scanning electron microscope (Carl Zeiss Microscopy GmbH, Jena, Germany) at 5 kV acceleration voltage. For scanning electron microscopy of rose (*Rosa*), *Solanum*, and rice (*Oryza sativa* and *O. rufipogon*) prickles, fresh tissue was adhered to stubs using non-conducting adhesive tape and imaged using a NanoScope JCM-7000 Benchtop SEM (JEOL, Tokyo, Japan).

#### Brix assay to measure fruit sugar content

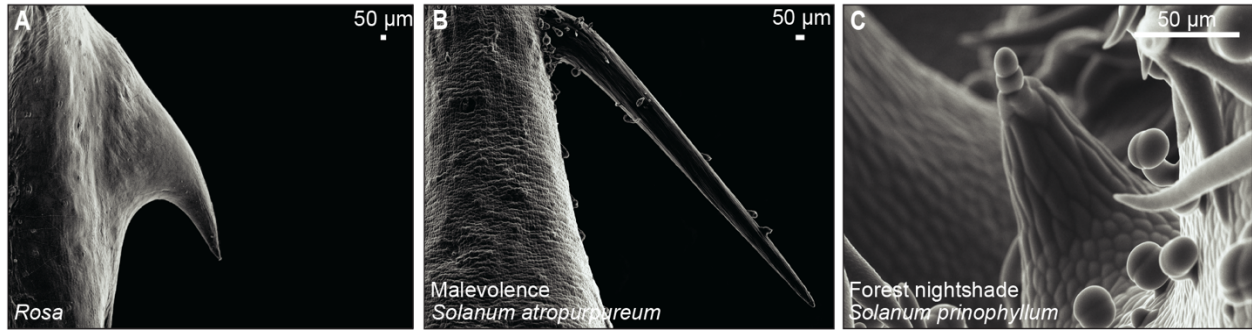
From 6 Desert raisin (*S. cleistogamum*) plants, 12 fully ripened and dried fruits were randomly selected to determine their soluble sugar content (Brix). Fruits were briefly rehydrated by soaking in DI water for 60 minutes. Following rehydration, the fruit were patted dry and placed into a 50 mL tube. Rehydrated fruit were then crushed using a spatula. The slurry was centrifuged at 15,000 rpm for 5 min. The supernatant was drawn from the samples and diluted 1:30 to obtain a working volume of 300 µL. The Brix value (%) was quantified from three technical replicates



with an ATAGO Palette digital Brix refractometer (ATAGO, Tokyo, Japan). See **Table S8** for data.

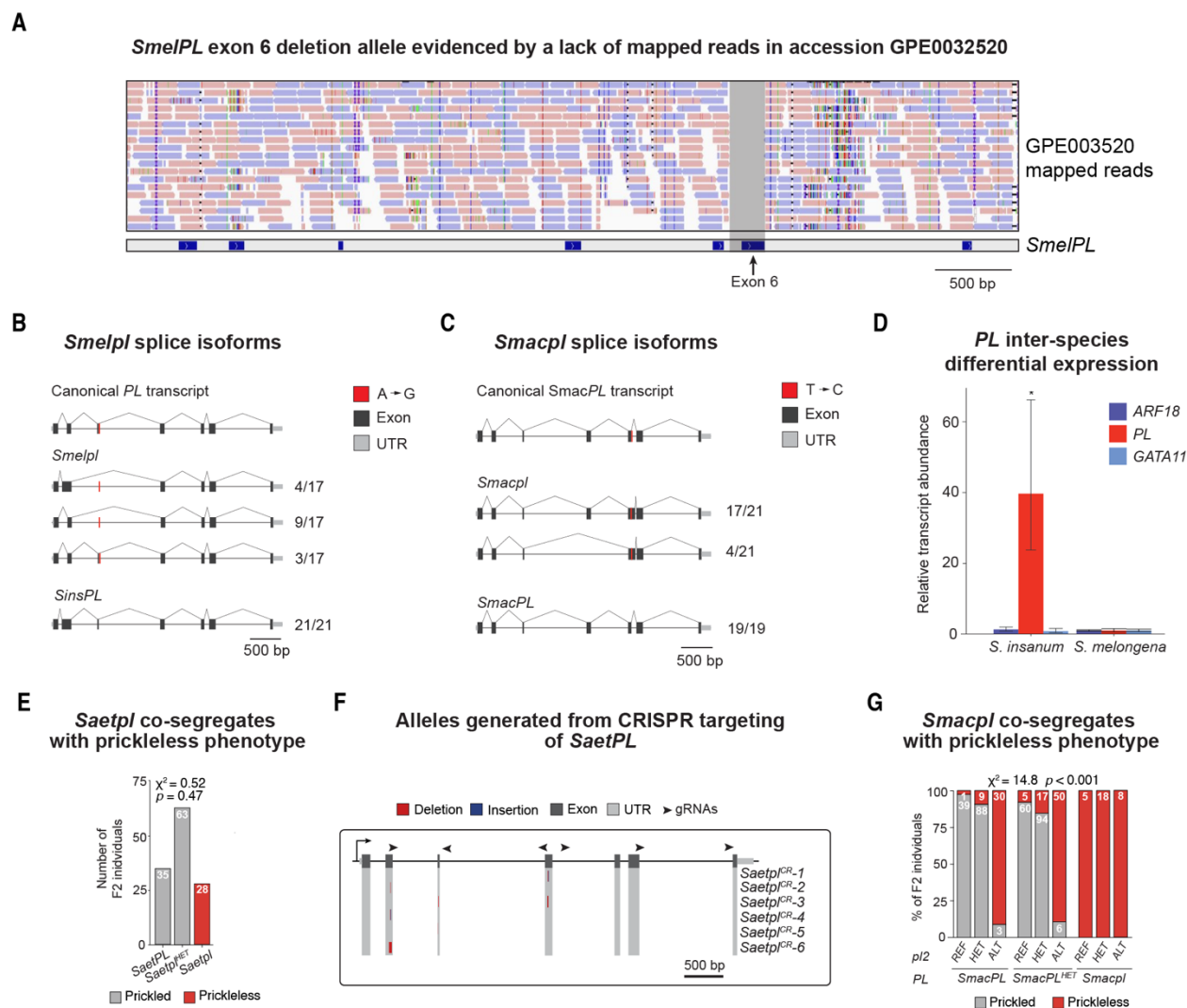
Cross-species *LOG* family co-expression conservation analysis

5 To calculate the co-expression conservation between tomato genes and *Arabidopsis*  
*thaliana* genes, gene family orthology information was downloaded from OrthoDB V11(96).  
Using one-to-one focal gene pairs between the two species, the degree to which co-expression is  
conserved between all *LOG* genes in each species was calculated (97). Briefly, a robust gold  
10 standard coexpression network for each species was sourced (27), which consists of coexpression  
networks built from hundreds of samples across dozens of experiments. For each *Arabidopsis LOG*  
gene, the top 10 genes it was coexpressed with were identified and used to predict the coexpression  
partners of all *LOG* genes in tomato. The degree to which the top coexpressed genes in *Arabidopsis*  
predicted the top coexpressed genes for a given *LOG* homolog in tomato yielded the co-expression  
conservation score, which is defined as the area under the receiver operating characteristic  
15 (AUROC) curve. This was then repeated in the other direction, taking the top 10 coexpression  
partners of each tomato *LOG* gene and predicting the coexpression partners of all *Arabidopsis*  
*LOG* genes. The two resulting AUROC statistic values are averaged, resulting in the final co-  
expression conservation scores.



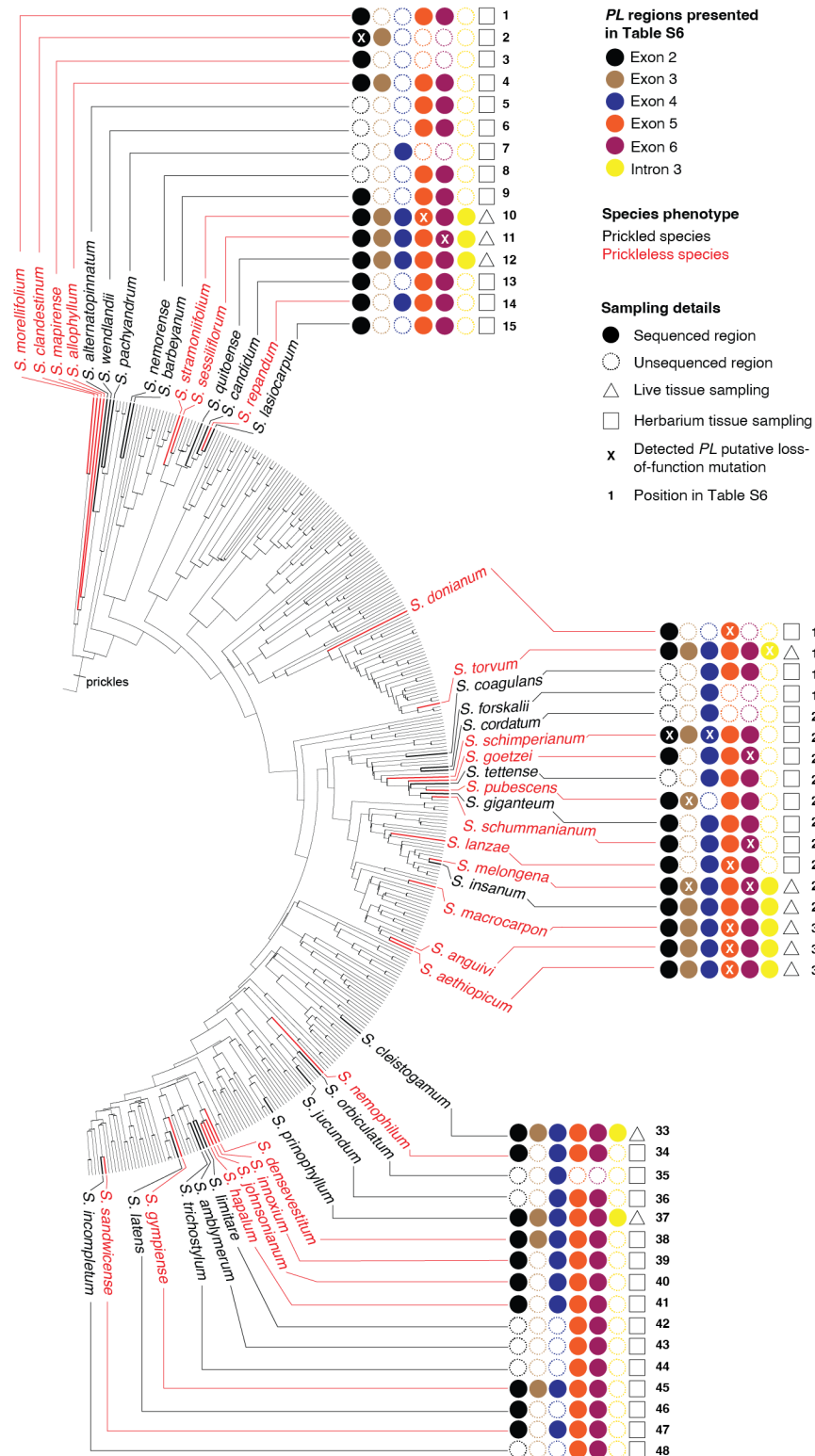
**Fig. S1.**

**Prickles are morphologically similar.** (A-C) Scanning electron micrographs (SEMs) of prickles in rose (*Rosa* spp.) (A), Malevolence (*Solanum atropurpureum*) (B), and at an early stage of development in Forest nightshade (*Solanum prinophyllum*) (C).



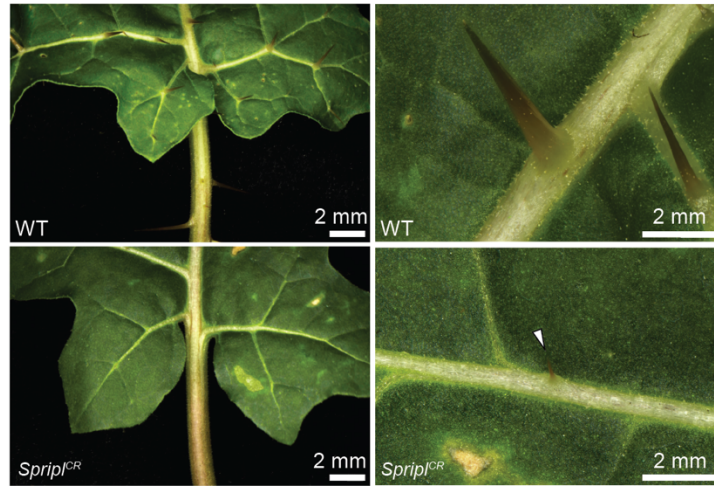
**Fig. S2.**

**Characterization of *PL* alleles in Brinjal (*S. melongena*) and Gboma eggplant (*S. macrocarpon*).** (A) A second *Smelpl* allele in the re-sequenced accession GPE0032520 removed exon 6 as evidenced by a lack of whole-genome sequencing reads mapped to this region. Reads mapping to the forward strand are in red reads mapping to the reverse strand are in blue. (B, C) Gene body diagrams illustrate the canonical and *PL* isoforms for common eggplant (A) and Gboma eggplant (B). Below are detected *PL* isoforms from the leaf tissue of plants homozygous for the indicated alleles. Fractions represent the proportion of isoforms of each type detected. (D) RT-qPCR analysis of *PL* expression in the leaves of wild eggplant (*S. insanum*) and common eggplant. The expression of two transcription factors nearby the *pl* mapping interval are also shown. Error bars reflect the standard deviation about the mean, Student's two-tailed t-test, \*  $p < 0.05$ ,  $n \geq 3$ . (E) *Saetpl* co-segregates with the prickless phenotype in a Scarlet eggplant F2 mapping population. (F) Alleles generated from CRISPR-Cas9 genome editing of *SaetPL* in *S. aethiopicum* (accession PI 666076) (G) F2 population confirms co-segregation of *Smacpl* and *pl2* with the prickless phenotype. Numbers within bars (E and G) indicate count totals of individuals within each genotype class.



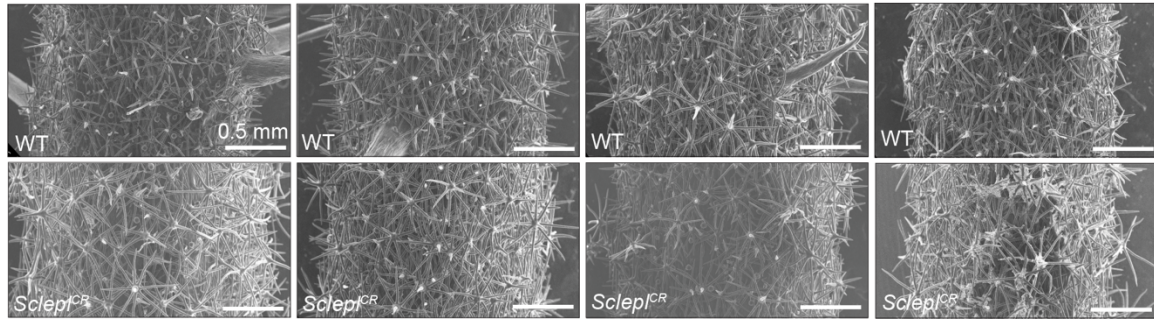
**Fig. S3.**

**Identification of PL mutations in the spiny *Solanum*.** The phylogeny from Fig. 1B is reproduced at left and species for which all or part of PL were sequenced are shown. Sequences of the indicated regions are numbered according to their position in Table S6.



**Fig. S4.**

**Small prickles are occasionally observed on  $pl^{CR}$  plants.** Bright-field images of leaves of Forest nightshade (*S. prinophyllum*) WT and *Spripl<sup>CR</sup>* plants. Arrowhead indicates the presence of a small prickles along a *Spripl<sup>CR</sup>* leaf vein.



**Fig. S5.**

**Trichome density and morphology is qualitatively unaffected in genome edited *Sclepl* mutants.** Scanning electron micrographs of the *S. cleistogamum* petiole in the fully-expanded fourth leaf just below the lamina-petiole boundary ( $N = 4$  plants for each indicated genotype).