Successive duplication-divergence mechanisms at the RCO locus contributed to leaf shape diversity in the Brassicaceae

Susanna Streubel1,∗, Michael André Fritz1, Melanie Teltow1, Christian Kappel1 and Adrien Sicard1,2,†

ABSTRACT
Gene duplication is a major driver for the increase of biological complexity. The divergence of newly duplicated paralogs may allow novel functions to evolve, while maintaining the ancestral one. Alternatively, partitioning the ancestral function among paralogs may allow parts of that role to follow independent evolutionary trajectories. We studied the REDUCED COMPLEXITY (RCO) locus, which contains three paralogs that have evolved through two independent events of gene duplication, and which underlies repeated events of leaf shape evolution within the Brassicaceae. In particular, we took advantage of the presence of three potentially functional paralogs in Capsella to investigate the extent of functional divergence among them. We demonstrate that the RCO copies control growth in different areas of the leaf. Consequently, the copies that are retained active in the different Brassicaceae lineages contribute to define the leaf dissection pattern. Our results further illustrate how successive gene duplication events and subsequent functional divergence can increase trait evolvability by providing independent evolutionary trajectories to specialized functions that have an additive effect on a given trait.

KEY WORDS: Plant development, Gene duplication, Leaf shape, Morphological evolution, Capsella, Arabidopsis

INTRODUCTION
The increase in biological complexity of organisms and their ability to adapt to new environments depend on the emergence of new functions. New genes can evolve through lateral gene transfer, de novo from noncoding sequences or through the modification of existing genes (Andersson et al., 2015). In the latter case, new functions often evolve from gene duplication and the subsequent functional divergence of one of the duplicated copies (Fisher, 1935; Haldane, 1932). Indeed, since the two incipient paralogous genes fulfill the same function, one of them may be released from selective pressure and therefore be ‘allowed’ to accumulate mutations that may or not lead to the emergence of a new function (Lynch and Conery, 2000; Lynch and Force, 2000; Olino, 1970). An increase in copy number due to gene duplication may, nevertheless, be beneficial or detrimental, depending on the circumstances (Innan and Kondrashov, 2010; Kondrashov, 2012; Kondrashov et al., 2002; Rogozin, 2014). This antagonism may lead in the first place to the rebalancing of the ancestral gene expression between the two copies and, thus, to a partitioning of the ancestral function rather than the emergence of a new function (Rogozin et al., 2014). To fully understand how gene duplication contributes to evolutionary processes it is necessary, therefore, to study the functional divergence between orthologs and paralogs along phylogenetic trees and within lineage-specific contexts (Rogozin, 2014).

Within recent years, plant leaf shape has emerged as an ideal model with which to study morphological evolution and phenotypic novelty (Mentink and Tsiantis, 2015; Rast-Somssich et al., 2015). Leaf forms vary enormously in nature both at the level of overall geometry and in the extent to which the margin is dissected (Bar and Ori, 2014). In the Brassicaceae, leaf shape varies from simple to more complex forms, such as the compound leaves of Cardamine hirsuta (Piazza et al., 2010). In simple leaves, the margin is continuous and entire, whereas in compound leaves the blade is divided into small units called leaflets. Many intermediate forms can be found and their classification mainly depends on the level of leaf margin dissection. Some species, such as Arabidopsis thaliana, display small irregularities around the leaf margin called serrations, whereas other species, such as Arabidopsis lyrata, develop more pronounced deformations called lobes. The pattern in which the lobes are positioned along the leaf margin also varies between species. For example, A. lyrata develops lyrate leaves in which the lobes are concentrated in the proximal (basal) part of the leaf with an enlarged terminal lobe, whereas other species such as Capsella rubella develop a rather regular pinnatifid lobing with deep dissection all around the margin (Sicard et al., 2014). Whether this difference in lobe positioning around the leaf margin reflects the existence of leaf type/lineage-specific functions or the divergence of conserved genetic networks is still unclear.

Differential expression of the class I KNOTTED-LIKE HOMEOBOX (KNOXI) homeodomain transcription factor has been shown to underlie the evolution of leaf complexity even beyond the Brassicaceae family (Bharathan et al., 2002; Blein et al., 2008; Furumizu et al., 2015; Hay and Tsiantis, 2010). More recently, several lines of evidence have highlighted the LATE MERISTEM IDENTITY1 (LMI1)/REDUCED COMPLEXITY (RCO) locus as a major driver of leaf shape diversification. This locus has been shown to underlie the major leaf morphs in cotton (Andres et al., 2017). In the Brassicaceae, this locus comprises a cluster of three class I HD-ZIP genes (referred to here as the RCO cluster or locus) that has arisen through two independent gene duplication events. The copy in position 1 (referred to here as LMI1) of this cluster was initially identified as a regulator of meristem identity in A. thaliana and named LATE MERISTEM IDENTITY1 (LMI1) (Saddic et al., 2006). In addition to its function in regulating inflorescence meristem identity, LMI1 also influences leaf serration and bract formation. The copy in position 2 (copy 2) underlies the formation of leaflets in the compound leaf species Cardamine hirsuta through its function as a growth repressor (Vlad et al., 2014).
Quantitative differences in the expression of the copy in position 3 (copy 3) underlie natural leaf shape variation in the genus *Capsella* (Sicard et al., 2014). The loss of copies 2 and 3 has led to the evolution of simple leaves in *A. thaliana* (Sicard et al., 2014; Vlad et al., 2014). The extent to which these two copies are functionally similar, and therefore the level of divergence among the copies at the *RCO* locus, has remained unclear.

Here, we took advantage of the presence of three expressed and potentially functional copies in *Capsella* to investigate the extent of post-duplication functional divergence within the *RCO* locus. Using interspecific gene transfer and geometric morphometrics, we demonstrate that the two events of gene duplication have been followed by cis-regulatory diversification of the incipient paralogs. This has created a cluster of three genes that regulate growth in different parts of the leaf. Copies 2 and 3 act independently and additively, and the maintenance of one or other copy in different Brassicaceae lineages contributes to specifying the pattern of leaf dissection. Our results also suggest that regulatory mechanisms repressing the function of copy 2 have evolved in both *C. rubella* and *Neslia paniculata*. Together with previous studies, our results indicate that successive duplication-divergence mechanisms and the subsequent functional divergence of the paralogs provided flexibility in the evolution of leaf morphology within the Brassicaceae.

**RESULTS**

**Evolutionary history of the *Capsella RCO* locus**

To gain further insights into the gene ancestry of the *C. rubella* (Cr) *RCO* locus, we reanalyzed the synteny relationship and sequence homology between the *RCO* paralogs of *A. thaliana* (At), *A. lyrata* (Al) and *C. hirsuta* (Ch) using *Aethionema arabicum* (Aa) as an outgroup and adding the *RCO* sequences of the closest *Capsella* relative *N. paniculata* (Np) (Fig. 1A). We first reconstituted the *RCO* phylogeny using open reading frame sequences and subsequently reconciled the resulting gene tree with the corresponding species tree (Fig. S1). Consistent with published data (Vlad et al., 2014), the maximum likelihood gene tree splits into two well-supported clades reflecting the first duplication event that occurred at this locus. The first clade includes all the *LMII* genes, while the sequences of paralogs in position 2 and 3 are grouped in a second clade. The phylogenetic signal within the latter clade is, however, less clear. The *A. lyrata*, *N. paniculata* and *C. rubella* sequences split into two clades that group the paralogs according to their position within the *RCO* cluster, suggesting a second duplication event before the split between these lineages (Fig. 1A and Fig. S1C,D). *ChRCO* and *ChLMII-LIKE3*, however, were not confidently assigned to either of these two clades. This could reflect an independent second duplication event in the *C. hirsuta* lineage or an early post-duplication allelic divergence and incomplete lineage sorting, as suggested by the reconciliation of the gene and species trees (Fig. 1A and Fig. S1C). In addition to *A. thaliana*, which has lost copies 2 and 3, structural variations in the *N. paniculata* lineage has led to the loss of copy 2. Successive events of gene duplication and independent loss of *RCO* copies have, therefore, shaped the genetic diversity at this locus.

**Cis-regulatory divergence of the three *RCO* copies**

We next investigated the functional divergence between *CrRCO-B* and *CrRCO-A*. Because leaf morphology is controlled through *RCO* expression patterns (Sicard et al., 2014; Vlad et al., 2014), we first asked whether *CrRCO-B* and *CrRCO-A* have retained identical transcriptional regulation after duplication. We started by quantifying the expression levels of the three *RCO* paralogs in different *C. rubella* organs by quantitative PCR (qPCR) (Fig. 1B and Fig. S2). Clustering of the genes based on their expression level in the different organs was consistent with their phylogenetic relationships. Indeed, *CrRCO-B* and *CrRCO-A* are more similar to each other in their expression profile than they are to *LMII*. The latter shows strong expression in the inflorescences, which is not seen for *CrRCO-B* and *CrRCO-A*. Expression of both *CrLMII* and *CrRCO-B* was detected in young leaves, early elongating leaves (4th to 6th leaves on the rosette) and late elongating leaves (12th to 15th leaves on the rosette) but not in any other tested organs. *CrRCO-A* shows a very similar expression profile to *CrRCO-B*, with the exception that its expression could not be detected in early elongating leaves and that the abundance of its mRNA in both young leaves and late elongating leaves was significantly lower than that of *CrRCO-B*.

To determine whether this difference in mRNA abundance reflects changes in expression territories, we generated reporter constructs in which the coding sequence of β-glucuronidase fused to a nuclear localization signal, NLGSU, was inserted in frame at the stop codon position of *CrLMII*, *CrRCO-B* or *CrRCO-A*. We transformed these recombinant fragments into *A. thaliana* Col-0 plants and compared the spatial distribution of the *RCO* proteins during leaf development using at least three independently transformed lines for each construct. This analysis revealed that the three *Capsella* paralogs drive the expression of the reporter in different areas of the young developing leaves (Fig. 1C). *CrLMII* induces the accumulation of NLGSU at the leaf tips and at the developing teeth (Fig. 1C). By contrast, *CrRCO-B* reporter is expressed all around the leaf margin except for the tips and the serrations. During leaf development, *CrRCO-B* expression becomes stronger in the proximal parts and weaker at the distal margin. Finally, *CrRCO-A* reporter signal was mostly restricted to a few cells at the sinuses of the developing leaf serrations. These results indicate that the expression patterns of the three paralogs have clearly diverged after duplication (Fig. 1C).

**Functional diversification of *RCO* paralogs**

In accordance with its expression at the serration tips, the loss of function of *LMII* in *A. thaliana* (such as in the *lmi1-1* mutant) leads to a decrease in the intensity of the leaf margin irregularities and, thus, to the formation of a smoother outline (Fig. 2A) (Saddiqi et al., 2006; Sicard et al., 2014). The simple leaves and the absence of functional *RCO* paralogs renders the *A. thaliana* *lmi1-1* mutant an ideal genetic background in which to determine the level of divergence between *CrRCO-B* and *CrRCO-A*.

Introducing the *CrRCO-B* gene into *lmi1-1* resulted in a global reduction in overall leaf area, in the dissection of the lower part of the blade, and occasionally in the formation of leaflet-like structures (Fig. 2 and Fig. S3). Principal component (PC) analysis (PCA) on elliptic Fourier descriptors (EFDs) of leaf outlines identified PC1 as separating *lmi1-1* from *lmi1-1; CrRCO-B* leaves (Fig. 2B and Fig. S3A). This PC mainly reflects a decrease in the overall leaf area due to a reduction of leaf margin growth and lobe formation in the basal part of the leaves. Note that we chose not to size-normalize the leaf in this analysis in order to assess the effect of *RCO* on overall leaf geometry. We measured a significant decrease in area ratio values (the ratio between the leaf area and the area of its convex hull), which was mostly due to a reduction in leaf surface area rather than to any increase in perimeter (Fig. 2C,D and Fig. S3C-E). Thus, consistent with its expression pattern, we concluded that *CrRCO-B* regulates the extent of leaf margin growth as well as the dissection in the proximal region.

The transformation of *CrRCO-A* into *lmi1-1* resulted in the formation of deep lobes in the medial leaf region (Fig. 2A). Comparing the outlines of the *lmi1-1* and *lmi1-1; CrRCO-A* leaves identified two PCs that significantly differ between the two
**genotypes:** PC1 and PC3 (Fig. 2B and Fig. S3B). PC1 reflects differences in leaf length and in the surface area of the medial part of the leaf, while PC3 reflects the regularity of the leaf outlines. Note that because of variability in lobe positioning along the proximal-distal axis in *lim1-1; CrRCO-A* plants, the reduction in width of the medial leaf region reflected by PC1 and PC3 most likely results from the formation of lobes in this area. The leaves of *lim1-1; CrRCO-A* plants have higher PC1 and lower PC3 values than *lim1-1*, indicating that *CrRCO-A* induces the formation of lobes and reduces the blade surface in the medial region (Fig. S3B). The measurement of leaf shape parameters reflecting the extent of leaf margin dissection [dissection index (DI), area ratio] supported this conclusion and indicated that it was due to both a decrease in the blade surface and an increase in the total perimeter (Fig. 2C,D and Fig. S3C-E). Therefore, *CrRCO-A* regulates the formation of lobes in the medial region of the leaf blade. Similar results were obtained by transforming *C. grandiflora CgRCO-B* and *CgRCO-A* into *lmi1-1* (Fig. 3 and Fig. S4).

Thus, *RCO-B* and *RCO-A* have distinct effects on leaf shape that are consistent with their divergent expression patterns. Although the function of *Capsella LMI1* in leaf shape determination has not been assessed directly here, the conservation of its expression pattern across the Brassicaceae strongly suggests that it fulfills the same function as its homolog in *A. thaliana* [this study compared with Saddic et al. (2006)]. The *cis*-regulatory diversification among the *RCO* paralogs has therefore resulted in the evolution of a three-gene cluster, in which each member regulates distinct aspects of leaf growth.

**PC3** and **PC1** have additive effects on leaf dissection in *A. thaliana*

We next tested the genetic interaction between the two paralogs by crossing *lmi1-1; CgRCO-B* with *lmi1-1; CgRCO-A* and analyzed...
lobes at the leaf base, whereas CgRCO-A promotes the formation of deep lobes in the medial region (Fig. 3A and Fig. S4). Leaf dissection and blade area appear to be more severely affected in lmi1-1; CgRCO-B; CgRCO-A plants. PCA on EFDs of the leaf outline identified several components that separated the different genotypes (Fig. 3B,C and Fig. S5A,B). PC1, which reflected changes in the overall leaf surface area and dissection of the blade, discriminated all genotypes, with lmi1-1; CgRCO-B; CgRCO-A plants having the lowest PC1 values, as expected from a cumulative effect of the two transgenes. PC3, which summarizes mainly the outgrowth of the medial part of the leaf blade, separated lmi1-1; CgRCO-B from both lmi1-1; CgRCO-A and lmi1-1; CgRCO-B; CgRCO-A leaves. Leaves of lmi1-1; CgRCO-B also differed from lmi1-1; CgRCO-A along PC5, which reflects the shift from ‘lyrate’ to ‘pinnatifid’ lobing (as defined above). The lmi1-1; CgRCO-B; CgRCO-A plants exhibited intermediate PC5 values.

When plotted on the PC1-PC3 morphospace, the co-expression of the two paralogs moves the leaf further away from lmi1-1 and towards PC values reflecting highly dissected leaves, suggesting an independent and additive effect of the two genes (Fig. 3C). This was further supported by the quantification of leaf parameters in the different genotype classes (Fig. 3D and Fig. S5C). Indeed, lmi1-1; CgRCO-B; CgRCO-A plants show a higher DI and lower leaf area ratio. As expected from the above experiment, the total length of the leaf perimeter was only affected by RCO-A. Linear models fitting the leaf area ratio in our population did not reveal a significant contribution of an interaction between RCO-A and RCO-B but suggested a simple additive effect of the two genes (Fig. S5D).

Thus, we concluded that RCO-B and RCO-A independently influence leaf dissection and act additively.

**Functional diversification of the RCO paralogs through cis-regulatory evolution**

The functional diversification of RCO-A and RCO-B was accompanied by the acquisition of very specific expression patterns, suggesting that cis-regulatory evolution underlies their diversification. The functional specialization of C. hirsuta RCO (in position 2 in the RCO cluster) from ChLM11 was previously reported to be the result of both cis-regulatory element and protein coding sequence evolution (Vlad et al., 2014; Vuolo et al., 2016). We confirmed that the promoter of the Capsella CrRCO-B gene also carries all the specific information required to induce the formation of lyrate lobed leaves (Fig. 4A,C,E and Fig. S6). lmi1-1 plants transformed with ProCrRCO-B:CrRCO-A (in which the CrRCO-B promoter drives CrRCO-A expression) develop similar leaves to lmi1-1; RCO-B. Indeed, it led to an increase in margin dissection in the basal part of the leaves and to a reduction in leaf area but had no significant effect on the overall leaf perimeter (Fig. 4 and Fig. S6). These parameters are similarly affected by the introduction of the full CrRCO-B construct. This was also reflected in the PCA, in which ProCrRCO-B:CrRCO-A affects only PC1 significantly, similar to CrRCO-B, whereas CrRCO-A affects both PC1 and PC3. However, ProCrRCO-B:CrRCO-A did not affect leaf growth to the same extent as CrRCO-B, suggesting that other regulatory elements, contained in the intronic or coding sequences, might have diverged between the two paralogs.

We next asked whether the functional divergence of RCO-A had also been caused by the evolution of its promoter sequences. In contrast to CrRCO-A, the ProCrRCO-A:CrRCO-B construct did not have a strong influence on leaf shape (Fig. 4B,D,F and Fig. S6). Only one transgenic line out of three showed a significant increase in DI and decrease in area ratio compared with lmi1-1 plants. This
increase was associated with a slight increase in total perimeter length due to the formation of weak lobes around the leaf margin (Fig. 4B and Fig. S6). Although not significant for all independently transformed lines, there was a trend towards reduced PC3 values for this genotype, which reflects the formation of lobes in the medial part of the leaves, usually strongly affected by CrRCO-A (Fig. 4D,F and Fig. S6E).

While screening over 20 lines independently transformed with ProCrRCO-A:CrRCO-B, we never observed leaf dissection resembling that obtained by transforming CrRCO-A in the same background. We therefore concluded that ProCrRCO-A:CrRCO-B induces the formation of weak lobes in the medial region of the leaves at the same location as the lobes induced by the full CrRCO-A genomic fragment. The fact that the dissection of the leaf margin occurred at the same position as in lmi1-1; CrRCO-A, but with a weaker intensity, suggests that ProCrRCO-A::CrRCO-B expression is correctly patterned but that the RCO level does not reach that provided by the full-length CrRCO-A genomic fragment. These results suggest that CrRCO-A function has diverged from CrRCO-B through both cis-regulatory evolution in its promoter sequence and through the evolution of its coding or intronic sequences, leading to an increase in its overall protein level.

**RCO-A is the copy with the strongest influence on Capsella leaf shape**

To determine the contribution of the RCO paralogs to Capsella leaf shape, we screened ∼1500 M2 families from an EMS-mutagenized C. rubella population for mutants with reduced leaf complexity. We then screened the selected leaf shape mutants for the presence of a point mutation within CrRCO-B or CrRCO-A by TILLING. Following this approach, we identified EMS lines mutated in CrRCO-A (Fig. S7). We identified two lines: rco-a-1, with a C-to-T conversion at the end of exon 2, and rco-a-2, harboring a G-to-A conversion in the first exon of CrRCO-A. Both mutations led to a premature stop codon (Fig. S7B). rco-a-1 and rco-a-2 display identical phenotypes marked by the absence of deep lobes (Fig. 5).
A complementation test indicated that the leaf phenotype of these two mutants is caused by mutations affecting the same genes (Fig. S7G). The \textit{rco-a-1} mutation co-segregated with the loss of deep leaf dissection in F2 progenies ($n=143$) of a cross between \textit{rco-a-1} and the \textit{C. rubella} accession \textit{Cr1504}. Furthermore, transferring a functional \textit{RCO-A} allele into \textit{rco-a-1} by \textit{Agrobacterium}-mediated transformation rescued lobe formation (Fig. S7E,F). We therefore concluded that the loss of leaf dissection in these lines was caused by inactivation of \textit{CrRCO-A}.

Both \textit{rco-a-1} and \textit{rco-a-2} lacked deeply lobed leaves throughout their development (Fig. 5A,D). The DI and area ratio values were considerably lower and higher, respectively, in all mutant leaves when compared with wild type (Fig. 5D). As expected from the above experiments, this decrease in leaf margin dissection was mainly associated with a strong reduction in perimeter ratio (Fig. 5D). PCA on EFDs of the leaf outlines identified PC2 as separating wild-type plants from both \textit{rco-a-1} and \textit{rco-a-2} (Fig. 5C). This PC mainly reflects the intensity of leaf dissection in the medial part of the leaves (Fig. 5B). PC2 values were higher in both mutants compared with wild type, reflecting the decrease in the prominence of the lobes. Since no decrease in \textit{CrRCO-B} or \textit{CrLMII} expression was detected in the young leaves of \textit{rco-a-1} and \textit{rco-a-2} (Fig. S7H), these results indicated that \textit{CrRCO-A} is the main contributor to \textit{Capsella} leaf shape. We nevertheless observed that the leaf outlines of both \textit{rco-a-1} and \textit{rco-a-2} were not completely smooth, but rather serrated (Fig. 5A). The serration is not visible in early developing leaves but intensifies along the plant lifespan. It is therefore plausible that \textit{CrLMII} or \textit{CrRCO-B} also contributes to leaf dissection, but to a weaker extent than \textit{CrRCO-A}.

### Lineage-specific inhibition of NpRCO activity contributes to the evolution of smooth leaves in \textit{Neslia paniculata}

\textit{N. paniculata} develops completely smooth leaves, without any lobes or teeth (Fig. 6A). We therefore examined whether the structural variation that has occurred within the \textit{RCO} locus has caused the loss of lobed leaves (Fig. 1A and Fig. S1D). Indeed, parts of the promoter and the majority of the coding sequence of copy 2, as well as part of the promoter of \textit{NpRCO} (copy 3), have been deleted in \textit{N. paniculata} (Fig. 6B, Fig. S1D and Fig. S9). It is therefore plausible that \textit{NpRCO} has also lost its functionality, possibly because of misexpression due to the altered promoter sequence, leaving only \textit{NpLMI1} functional (as in \textit{A. thaliana}).

To test this idea, we cloned the sequence from the 3’ end of \textit{NpLMII} to the 3’ end of \textit{NpRCO} and transformed it into \textit{lmi1-1} mutants. We compared the leaves of the transgenic plants with those of \textit{lmi1-1}, \textit{lmi1-1; CrRCO-B} and \textit{lmi1-1; CrRCO-A}. Expression of \textit{NpRCO} in \textit{lmi1-1} led to an increase in the dissection of the proximal part of the leaves, resulting in the formation of basal lobes that resembled those observed in \textit{CrRCO-B} transgenic plants (Fig. 6C). PCA of EFDs of the leaf outlines identified two PCs that differed significantly among the genotypes: PC1 and PC3 (Fig. 6E). These PCs were similar to those previously identified and reflect either the
extent of leaf margin growth or the formation of lobes in the medial region of the leaves. When plotted on the PC1×PC3 morphospace, lmi1-1; CrRCO-B leaves moved towards the geometric space occupied by lmi1-1; CrRCO-B leaves (Fig. 6D). Indeed, as for CrRCO-B, NpRCO tends to increase PC1 values without affecting PC3 values, whereas CrRCO-A affects both PCs (Fig. 6E). Again as for CrRCO-B, NpRCO affected the area ratio of the leaves without strongly affecting their perimeter ratio, DI or roundness, whereas all these parameters were strongly affected by CrRCO-A (Fig. S8A-D). This indicates that despite sharing the same ancestry with CrRCO-A, NpRCO fulfills a CrRCO-B-like function. Because N. paniculata develops smooth leaves even though NpRCO expression could be detected in its young leaves, this observation indicates that lineagespecific factors have evolved to repress NpRCO function in leaf margin dissection (Fig. S8E).

**DISCUSSION**

Although both RCO copies 2 and 3 have been shown to underlie leaf shape diversification in Cardamine and Capsella, respectively, it was unclear to what extent the second event of gene duplication that occurred at the RCO locus had contributed to diversifying leaf shape. Here, we took advantage of the fact that both RCO-B and RCO-A are expressed and potentially functional in Capsella to address this question.

The gene in position 1 within the RCO cluster, LMI1, which is believed to have retained most of the ancestral function, regulates both flowering time and the degree of leaf serration (Sadie et al., 2006; Sicard et al., 2014; Vlad et al., 2014). Re-introducing Capsella RCO-B into A. thaliana indicated that it regulates the formation of lobes in the proximal part of the leaves as well as the extent of growth in both the proximal and distal margins (Fig. 2). Capsella RCO-A expression in A. thaliana had different consequences on leaf shape, and led to the formation of deep lobes regularly distributed all around the leaf margin (Fig. 2). Based on these results, the three paralogs appear to regulate the growth of different leaf parts. Indeed, although the transcriptional regulation of RCO genes may be slightly different in A. thaliana compared with C. rubella, the fact that CrRCO-A and CrRCO-B genomic fragments had different effects on leaf shape indicates that sequence evolution after duplication within these fragments has led to the functional divergence of the two paralogs. This was further supported by the observation that mutating RCO-A in C. rubella inhibited the formation of leaf lobes qualitatively rather than quantitatively. The latter would have been expected if the two genes were fulfilling the same function. Our expression studies and promoter-swap experiments indicated that both cis-regulatory evolution and intronic or coding sequence polymorphisms underlie the functional divergence of the Capsella RCO paralogs. This is consistent with previous findings in C. hirsuta that demonstrated that CrRCO has diverged from ChLMI1 through both cis-regulatory and protein sequence evolution (Vuolo et al., 2016). The expression of ChRCO, CrRCO-B or NpRCO (which appears to be controlled by the promoter of copy 2 in Neslia owing to the deletion of the NpRCO promoter and NpRCO-B coding sequence as described below) in A. thaliana has similar effects on leaf shape (Vlad et al., 2014) (Figs 2 and 6), suggesting that the
function of the different RCO genes relative to their respective cluster position is mostly conserved.

At this stage, it is still difficult to firmly conclude whether the functional divergence of the three RCO paralogs reflects the partitioning of the ancestral function or the neo-functionalization of the incipient copies. Nevertheless, the fact that A. arabicum, in which no duplication has occurred, develops smooth leaves, and the fact that expression of the unique AalLM1 gene in this species is excluded from the proximal leaf region, suggest that the ancestral RCO function was not to regulate leaf dissection (Vlad et al., 2014). Neo-functionalization of the incipient copies, rather than partitioning of the ancestral function, is therefore likely to have occurred, at least after the first event of duplication. It is, however, noteworthy that the expression profile of the three paralogs is very similar, especially in terms of organ specificity. All three paralogs are expressed at similar stages during, and mostly restricted to, leaf development (except in inflorescence for LMI1). It is therefore tempting to speculate that the functional divergence of the RCO paralogs was built on an ancient expression pattern, upon which new duplication has contributed to the evolution of compound and lyrate leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves. N. paniculata also does not develop smooth margins. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Despite this, we could not by itself explain the evolution of smooth margins in N. paniculata. Therefore, the functional divergence of the three RCO genes underlies interspecific quantitative changes in the level of leaf dissection (Sicard et al., 2014). Within the species analyzed so far, two events of structural rearrangement at the RCO locus have been shown to underlie leaf shape evolution: one in N. paniculata leading to the loss of copy 2, and a second in A. thaliana leading to the loss of copies 2 and 3 (Fig. 1) (Sicard et al., 2014; Vlad et al., 2014). Although the deletion of both copies 2 and 3 could explain leaf simplification in A. thaliana, the deletion of copy 2 could not by itself explain the evolution of smooth margins in N. paniculata. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves. N. paniculata also does not develop smooth margins. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves.

Fig. 6. Deletion and regulatory inhibition of RCO genes underlies simple leaf evolution in N. paniculata. (A) N. paniculata rosette. Scale bar: 1 cm. (B) The structure of the N. paniculata RCO locus compared with that of the C. rubella RCO locus. (C) Silhouette of a representative 9th leaf of lmi1-1 transformed with NpRCO. (D,E) PCA on EFDs of leaf outlines. (D) Distribution of individual leaves projected on the PC1/PC3 morphospace. Each symbol of each color represents an independently transformed line. (E) Effects along PCs 1 and 3 (top) and the corresponding distribution of PC values for each genotype (bottom). Letters indicate significant differences as determined by Tukey’s HSD test. n is indicated.

This duplication-divergence mechanism might also have contributed to increasing the evolvability of leaf shape in the Brassicaceae by providing different regulatory mechanisms and evolutionary trajectories to each of the paralogs. Indeed, since each paralog has an independent influence on leaf shape and their functions rely mainly on their expression patterns and expression levels, any accumulated variant affecting expression or function has the potential to evolve a new leaf type. The first event of gene duplication has contributed to the evolution of compound and lyrate leaves (Vlad et al., 2014). Here, we demonstrated that the second event of duplication-divergence has contributed to the evolution of pinnatifid lobing. We previously demonstrated that variation in the expression level of RCO genes underlies interspecific quantitative changes in the level of leaf dissection (Sicard et al., 2014). Within the species analyzed so far, two events of structural rearrangement at the RCO locus have been shown to underlie leaf shape evolution: one in N. paniculata leading to the loss of copy 2, and a second in A. thaliana leading to the loss of copies 2 and 3 (Fig. 1) (Sicard et al., 2014; Vlad et al., 2014). Although the deletion of both copies 2 and 3 could explain leaf simplification in A. thaliana, the deletion of copy 2 could not by itself explain the evolution of smooth margins in N. paniculata. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves. N. paniculata also does not develop smooth margins. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves. N. paniculata also does not develop smooth margins. Indeed, the structural rearrangement that occurred at the N. paniculata RCO locus has left copy 3 intact, and we observed that NpRCO is able to promote the formation of lobed leaves when transformed into A. thaliana. Nevertheless, and despite the fact that NpRCO expression could be detected in young Neslia leaves (Fig. S8B), this species does not develop lobed leaves.
the main contributor to leaf shape in this species. Since rco-a leaf outlines were not totally smooth, but serrated, we concluded that CrLM1 and/or CrRCO-B also contribute to Capsella leaf shape, but to a weaker extent. Nevertheless, no lobes were observed in the basal part of rco-a leaves, in contrast to the effect of RCO-B in A. thaliana. It is therefore plausible that, as in Nelsia, lineage-specific changes in the gene regulatory network controlling leaf margin dissection have evolved in Capsella, limiting the activity of RCO-B. Thus, RCO cis- and trans-regulatory evolution, structural variation and polymorphisms have contributed to leaf shape evolution in the Brassicaceae.

Our results suggest that the two events of duplication have originated a cluster of three genes that have since diverged functionally and become specialized in regulating the growth of different leaf regions. The basic pattern of leaf dissection in the different Brassicaceae species will depend on which of the RCO-like genes has been retained in a functional state. Keeping LM1 highly active will promote the formation of a dentate/serrate margin, as is the case in A. thaliana. Maintaining RCO-B active will result in the formation of lobes in the proximal part of the leaf, whereas favoring the function of RCO-A will lead to the formation of deep lobes all along the leaf margin. Since these genes also appear to function in a quantitative manner (Sicard et al., 2014), modulating their expression also has the potential to further diversify the leaf shape among lineages. This model of leaf shape evolution within the Brassicaceae is supported by the observation that introgressing copy 2 from A. lyrata, C. hirsuta and C. rubella in A. thaliana results in the formation of proximal lobes (Fig. 2) (Vlad et al., 2014). In Capsella, which displays medially dissected leaves, the function of CrRCO-A seems to have been favored. Whether this model holds true for additional species still needs to be investigated. For instance, in A. lyrata, AlLM1-like3 (in position 3) appears to be functional when fused to the promoter of the second paralog, despite being truncated at the beginning of its third exon (Vlad et al., 2014). Our results would suggest that the basal dissection of A. lyrata leaves is due to high activity of copy 2 and low activity of the third paralog. Whether copy 3 is functional in A. lyrata is undetermined. C. hirsuta plants harboring a mutation in copy 2 still display some proximal lobing (Vlad et al., 2014). It remains unclear whether this reflects a different function of the copy in position 1 or 3, or the presence of an independent factor that influences proximal leaf dissection in this species. It would, therefore, be important to formally test the functionality and expression pattern of additional RCO copies in other Brassicaceae species in order to fully understand the extent of functional diversification at this locus. Indeed, the present model does not fully explain all leaf shape diversification in this family. For instance, both N. paniculata and C. rubella rco-a mutants lack proximal lobes despite the presence of an RCO gene, which can induce their formation. Furthermore, C. hirsuta and A. lyrata, both of which have a functional copy 2, exhibit very different leaf shapes. These observations suggest the existence of lineage-specific factors regulating the activities of the different RCO copies. Further genetic studies that aim to confirm and identify such factors will be crucial to fully understand the genetic mechanisms underlying leaf shape evolution.

MATERIALS AND METHODS
Biological materials and growth conditions
The A. thaliana accession Columbia 0 (Col-0) was used as genetic background in this study. The lm1-1 T-DNA insertion mutant in a Col-0 background (Saddiq et al., 2006) (N656213/SALK_016682) was obtained from NASC (Nottingham Arabidopsis Stock Centre). The C. rubella accession Cr22.5 (Tenoa, Italy, gift from Tanja Slotte) was used as wild-type genetic background for phenotyping analysis and mutagenesis. rco-a-1 and rco-a-2 alleles were isolated from an ethyl-methanesulphonate (EMS) mutagenized population in the Cr22.5 background. rco-a-1 and rco-a-2 were backcrossed twice to Cr22.5 before phenotypic analysis.

Plants were grown in a growth chamber under a 16 h day/8 h night photoperiod at 22°C during the day and 18°C during the night and in 70% humidity with a light intensity of 250 μmol m⁻² s⁻¹.

Molecular cloning and plant transformation
Genomic fragments and reporter constructs were introduced into pBluescript II KS vector (Stratagene, pBlueMPLUCAP) using Seamless Ligation Cloning Extract (SLICE) (Zhang et al., 2012). For plant transformation, the different recombinant fragments were transferred as Ascl fragments into the plant transformation vector pBarMAP, a derivative of pGPTVBAR (Becker et al., 1992). Further details on the molecular cloning are provided in the supplementary Materials and Methods. A. thaliana Col-0 and lm1-1, as well as C. rubella Cr22.5 (rco-a-1) plants were then transformed by floral dip (Bartholome et al., 2008; Clough and Bent, 1998).

Phylogenetic analysis
For detailed information on the sequence used in the phylogenetic analysis, see the supplementary Materials and Methods. Multiple sequence alignments were made using MUSCLE (Edgar, 2004). Maximum likelihood trees were reconstructed using RAxML (GTRCAT model) with 1000 bootstrap iterations (Stamatakis, 2014). Trees were visualized and edited using Dendroscope (Huson and Scornavacca, 2012). Gene-species tree reconciliation to estimate duplication and deletion events was performed using Notung (Stolzer et al., 2012). To study the synteny among RCO loci within the Brassicaceae, we aligned the sequences globally using the needle program from the EMBOSS suite (Rice et al., 2000). Gene structure and highly similar regions with at least 50% identity over a 100 bp window along the alignment and using a 50 bp sliding window were illustrated using R (R Core Team, 2016).

Gene expression analysis
The expression of the RCO paralogs was investigated by qPCR as described (Sicard et al., 2014). The expression of the C. rubella β-TUBULIN 6 gene (CrTUB6) was used to normalize expression values. Primer pairs used are shown in Table S1. Average values were based on at least three biological replicates.

Expression patterns of the different RCO genes were examined by detecting the activity of β-Glucuronidase as described (Lenhard et al., 2001), with the exception that higher concentrations of potassium ferricyanide (5 mM) and potassium ferricyanide (5 mM) were used.

Leaf shape phenotyping
Different parameters were used to characterize the shape of Capsella and Arabidopsis leaves. Fully elongated leaves were harvested and flattened on a white paper sheet and their silhouettes were digitalized using a Perfection V600 scanner (Epson) at 600 dpi. Images were then converted into binary pictures and leaf shape parameters were quantified using ImageJ (http://imagej.nih.gov/ij/). The dissection index was calculated as DI=(perimeter²)/(4π×area) (Bilsborough et al., 2011). The area ratio value was defined as the ratio between the area of the leaf and the area of its convex hull. The perimeter ratio corresponds to the ratio between the perimeter of the leaf and the perimeter of its 2D hull. PCA on EFDs for closed outlines (Kuhl and Giardina, 1982) were performed as described (Sicard et al., 2014). PCs linked to the genotype were identified through a Kruskal–Wallis test with the formula ‘principal component score–genotype’. Variations along the PCs were illustrated by reconstructing the mean and extreme shapes (based on minimal and maximal PC scores) using an inverse elliptical Fourier transformation as described (Sicard et al., 2014). Further information on leaf shape phenotyping can be found in the supplementary Materials and Methods.

Identification of rco-a-1 and rco-a-2 mutants
Thirty-four leaf shape mutants were isolated by screening ~1500 Cr22.5 EMS mutagenized lines and tested for mutations within RCO-B and RCO-A
by targeted induced local lesions in genomes (TILLING) using the primer pairs oAS1318/1319 and oAS633/1316 (Table S1), respectively (Till et al., 2006). The presence of mutations was then confirmed by Sanger sequencing (LGC genomics).

**Statistical analyses and descriptive statistics**

The distributions of the phenotypes investigated among the different genotypes are presented with box plots. In all box plots, middle lines represent the median, and the lower and upper hinges represent the first and third quartiles, respectively. The upper and lower whiskers extend to the largest and lowest values, respectively, but no further than 1.5 times the interquartile range (IQR; the distance between the 25th and 75th percentile) from the hinges; values beyond 1.5×IQR are considered as possible outliers and are displayed as dots.

Within each experiment, plants of the different genotypes were randomly assigned to positions in the trays, and trays were rotated once per week in the growth rooms. No blinding was performed. Statistical analyses were conducted in R. We performed a Tukey’s honest significant difference (HSD) test using the agricolae package add-ons implemented in R software for multi-comparison tests. The lm() function implemented in R was used to fit a linear model. For qPCR analysis, the data are presented as mean±s.e.m. P<0.05 was considered statistically significant.

**Acknowledgements**

We thank Doreen Märker and Christiane Schmidt for plant care, and members of the Lenhard and Bäurle group (Institut für Biochemie und Biologie, Universität Potsdam, Germany) for discussion and comments on the manuscript.

**Competing interests**

The authors declare no competing or financial interests.

**Author contributions**


**Funding**

This work was supported by a Deutsche Forschungsgemeinschaft grant to A.S. (SI1967/2).

**Supplementary information**

Supplementary information available online at [dev.biologists.org/lookup/doi/10.1242/dev.164301.supplemental](http://dev.biologists.org/lookup/doi/10.1242/dev.164301.supplemental)

**References**


Supplementary information

Supplementary Materials and methods:

Sequences used in the phylogenetic analysis.

The sequences used to reconstitute the RCO gene tree were retrieved from (https://phytozome.jgi.doe.gov/pz/portal.html, http://www.ncbi.nlm.nih.gov), Vlad et al. (2014), Sicard et al. (2014) or from whole genome sequencing data of N. paniculata (Slotte et al., 2013) or C. hirsuta (Gan et al., 2016). Sequences of 16 shared single copy genes selected based on (Duarte et al., 2010) were used to reconstitute the species tree. The mRNA sequences corresponding to these genes were retrieved from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) or whole genome sequencing data (for N. paniculata, see above) and concatenated for further analyses.

To study the synteny among RCO loci within the Brassicaceae, we extracted the corresponding genomic sequences from C. rubella, N. paniculata, C. hirsuta and A. lyrata genome assemblies (Gan et al., 2016; Hu et al., 2011; Slotte et al., 2013). The C.rubella sequence was corrected using DNA-seq reads for the Cr1504 accession ((Sicard et al., 2014); SCBI SRA study SRP044389).

Molecular cloning

Capsella genomic fragments of LMI1, RCO-B and RCO-A were amplified from Cr1504 and Cg926 DNA using the primer pairs, oAS598-599 (corresponding to ~2 kb of promoter sequence, the coding and intronic sequences as well as ~1kb of 3’UTR) oAS1164-1165 (corresponding to ~5 kb of promoter sequence, the coding and intronic sequences as well as ~1kb of 3’UTR) and oAS666-667 (corresponding to ~5.2 kb of promoter sequence, the coding and intronic sequences as well as ~0.6kb of 3’UTR) respectively. NpRCO was amplified from Neslia paniculata genomic DNA using the primers oAS1227 and oAS1224. The PCR fragments were then sub-cloned into a modified version of the pBluescript II KS vector (Stratagene, pBlueMLAPUCAP), giving rise to the plasmids pBS- LMI1r, pBS-RCO-Br, pBS-RCO-Ar, pBS-RCO-Bg, pBS-RCO-Ag and pBS-NpRCO.

For the promoter swapping experiment, RCO-B and RCO-A promoters were amplified from pBS-RCO-Br and pBS-RCO-Ar, respectively, using oAS1338-1339 and oAS1332-1333. The resulting PCR fragments were fused using Seamless Ligation Cloning Extract (SLICE) (Zhang et al., 2012) into pBS-RCO-Ar and pBS-RCO-Br respectively replacing the endogenous RCO-Ar and RCO-Br promoters.

Reporter constructs were generated by ligating a β-glucuronidase gene fused to a nuclear localization signal and a transcription terminator (GUS-NLS) at the 3’end of LMI1, RCO-B and RCO-A. To this end, the promoter and coding sequence of the RCO genes were amplified using oAS598-1188,
oAS1164-1182 and oAS666-1175. GUS-NLS fragments, amplified with oAS1189-1181, oAS1183-1181 or oAS1176-1181, were then fused to LMI1, RCO-B and RCO-A, respectively, and subcloned into a pBlueMLAPUCAP using SLICE.

**Leaf shape phenotyping**

Leaf area, perimeter and roundness \((4 \times ([\text{Area}]/(\pi \times [\text{Major axis}]^2))\) were quantified using imageJ (https://imagej.nih.gov/ij/). The Convex hull function of ImageJ followed by a visual assessment and manual corrections were applied to each object to define a hull around leaf shape. This hull, therefore, includes all the leaf lobes and reflects the two dimensional (2-D) shape of each leaf in the absence of leaf margin dissection. The area and perimeter of this convex hull was then measured in the ImageJ software. These values were used to calculate parameters below, which reflect leaf complexity:

- The dissection index was calculated as \(\text{DI} = \frac{(\text{perimeter}^2)}{(4\pi \times \text{area})}\) (Bilsborough et al., 2011). According to this parameter, a circle of the same area would have a value of 1.0, while more complex outlines have higher values.

- The area ratio value was defined as the ratio between the area of the leaf and the area of its convex hull. In this case undissected leaves have a value close to 1 whereas the formation of deep lobes results in lower values.

- When appropriate, we also measured what we call the ‘perimeter ratio’, which corresponds to the ratio between the perimeter of the leaf and the perimeter of its 2-D hull. This parameter only reflects changes in perimeter and its value increases as leaves becomes more complex.
**Supplementary Figures:**

(A) Maximum likelihood (ML) gene tree obtained from open reading frame sequences with ML bootstraps. (B) Maximum likelihood species tree, constructed based on 16 shared single-copy genes (Duarte et al., 2010). Bootstraps values and branches length are indicated in black and blue, respectively. (C) Reconciled gene-species tree bootstraps values of strong edges are indicated in green, duplication events are shown in red and gene losses in grey. Weakly-supported edges have been re-arranged to minimize the cost of gene duplication. (D) Synteny plot between *C. hirsuta*, *A. lyrata*, *C. rubella* and *N. paniculata* RCO loci.

**Fig. S1: Evolutionary history of the RCO locus.**

(A) Maximum likelihood (ML) gene tree obtained from open reading frame sequences with ML bootstraps. (B) Maximum likelihood species tree, constructed based on 16 shared single-copy genes (Duarte et al., 2010). Bootstraps values and branches length are indicated in black and blue, respectively. (C) Reconciled gene-species tree bootstraps values of strong edges are indicated in green, duplication events are shown in red and gene losses in grey. Weakly-supported edges have been re-arranged to minimize the cost of gene duplication. (D) Synteny plot between *C. hirsuta*, *A. lyrata*, *C. rubella* and *N. paniculata* RCO loci.
**Fig. S2: Expression profile of the RCO paralogs in Capsella.**

Relative expression of *LMII*, *RCO-B* and *RCO-A* in different Capsella organs determined by quantitative RT-PCR. mRNA was extracted from the apex of 4-week-old plants (YL note that this sample includes young leaves at position $\geq 10$ in the rosette and SAM), elongating (~2 cm long) leaves of 7-week-old plants at positions 15 to 17 (VOL), elongating leaves (~2cm) of 4-week-old plants at position 5-7 on the rosette (OL), developing fruit (Ft), inflorescences meristem (IN), 7-day-old seedlings (SD), roots from 7-day-old seedlings (SDD), apex of 7-day-old seedlings (SDU) and roots of 4-week-old plants (OR). Average values of at least 3 biological replicates are shown. Error bars represent s.e.m.
Fig. S3: Effects of \textit{CrRCO-B} and \textit{CrRCO-A} on \textit{A. thaliana} leaf shape.

(A-B) Principal component analysis on Elliptic Fourier Descriptors of leaf outlines. Distribution of individual leaves projected on the PC1/PC3 morphospace.

(C-E) Box plots illustrating the distribution of the 9th leaf dissection index (DI) (C), roundness (D) and perimeter (E) for each genotype (\(n\) is indicated on the graphs). Letters indicate significant differences as determined by Tukey’s HSD test.
Fig. S4: Effects of CgRCO-B and CgRCO-A on A. thaliana leaf shape.

(A-B) Principal component analysis on Elliptic Fourier Descriptors of leaf outlines. Distribution of individual leaves projected on the PC1/PC3 morphospace (A). Effects along principal components and the corresponding Box plots illustrating the distribution of PC values for each genotype are shown in (B).

(C) Box plots illustrating the distribution of the 9th leaf Area ratio, Perimeter ratio and Roundness for each genotype.

(n is indicated on the graphs). Letters indicate significant differences as determined by Tukey’s HSD test (n is indicated on the graphs).
Fig. S5: Genetic interaction between \textit{RCO-B} and \textit{RCO-A}.

(A-B) Principal component analysis on Elliptic Fourier Descriptors of leaf outlines. Effects along principal components (A). p values were determined using a Kruskal-Wallis test with the formula ‘principal component scores ~ genotype’. Box plots illustrating the distribution of PC values for each genotype are shown in (B). (C) Box plots illustrating the distribution of the 9th leaf Area, Dissection index and Perimeter for each genotype. (n is indicated on the graphs). Letters indicate significant differences as determined by Tukey’s HSD test (n is indicated on the graphs). (D) Effect of the interaction between the \textit{RCO} paralogs on leaf shape. Summary statistics of the linear model fitting leaf Area ratio in the \textit{lmi1-1;CrRCO-B x lmi1-1;CrRCO-A} population is shown on the left. An ANOVA comparing a full and additive model is display on the right.
Fig. S6: Cis-regulation and enhancer evolution underlie the functional divergence between RCO paralogs.

(A-C) Box plots illustrating the distribution of the 9th leaf area ratio (A), Perimeter (B), Area (c) dissection index (DI) (D) and roundness (E) for each genotype (n is indicated on the figure).
Fig. S7: RCO-A loss of function leads to the disappearance of deep lobes in Capsella leaves.
(A) Photograph of 7-week-old Capsella rubella wild-type, rco-a-1 and rco-a-2 plants. (B) The position of rco-a-1 and rco-a-2 mutations within CrRCO-A is shown in the upper panel. (C) An agarose gel picture showing the co-segregation of rco-a-1 with the loss of leaf lobes is shown in the lower panel. (D) Extended co-segregation analysis of the rco-a-1 mutation with the loss of leaf lobes. (E) Silhouettes of 12th leaf of wild-type and rco-a-1 and rco-a-1; CrRCO-A plants. The bar indicates 9 cm. (F) Box plots illustrating the distribution of the 13th leaf area and perimeter ratios of Capsella
wild-type, rco-a-1 and rco-a-1; CrRCO-A plants (n is indicated on the graphs). (G) Test of allelism between rco-a-1 and rco-a-2. Silhouettes of the 14th leaf of wild-type (WT) and rco-a-1 and rco-a-2, rco-a-1 x WT F1, rco-a-2 x WT F1 and rco-a-1 x rco-a-2 F1 plants are shown on the left panel and the count of phenotypic classes (plants with dissected (WT) or undissected leaves (mutant)) in the different F2s is shown on the right panel. Note that the two mutations failed to complement each other indicating that they are allelic (H) Relative expression of CrLMII and CrRCO-B in Capsella wild type, rco-a-1 and rco-a-2 plants determined by RT-PCR. mRNA was extracted from the apex of 4-week-old plants and includes young rosette leaves and SAM. Average values of at least 3 biological replicates are shown. Error bars represent s.e.m.
Fig. S8: Functional analysis of NpRCO.

(A) Box plot illustrating the distribution of Area ratio (A), Perimeter ratio (B), Roundness (C) and Dissection Index (DI) (D) (n is indicated on the graphs). Letters indicate significant differences as determined by Tukey’s HSD test. (B) Expression of NpLMI1 and NpRCO in N. paniculata young leaves.
<table>
<thead>
<tr>
<th>Development 145: doi:10.1242/dev.164301: Supplementary information</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. paniculata</td>
</tr>
<tr>
<td>------------------------------------</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>110</td>
</tr>
<tr>
<td>210</td>
</tr>
<tr>
<td>310</td>
</tr>
<tr>
<td>410</td>
</tr>
<tr>
<td>510</td>
</tr>
<tr>
<td>610</td>
</tr>
<tr>
<td>710</td>
</tr>
<tr>
<td>810</td>
</tr>
<tr>
<td>910</td>
</tr>
<tr>
<td>1010</td>
</tr>
<tr>
<td>1110</td>
</tr>
<tr>
<td>1210</td>
</tr>
<tr>
<td>1310</td>
</tr>
<tr>
<td>1410</td>
</tr>
<tr>
<td>1510</td>
</tr>
<tr>
<td>1610</td>
</tr>
</tbody>
</table>

AGAGAGAGAGTATGAGTCAAAATAGAGCTATAGACTAGGACTTTGAGTAGTG-TTCGACGAAGACAA--GCTTGTCCCTCATTGTACGATCTTACAACAT
ATGGTAGCTGCTTCCTCTGGCTGGTCGTCTTACCCGTGATAGCTGTTGCTTATCTCAAGGAGAGAAGGAAAGAGAATAGAAGAGAGAGAGAGCTAGA-
GCTTCCATCAGTGGTGGTAGCTCATCCAAGAACAGAGAATCTGAACACCGGAGGGAATCAAATCTACGGTACTGATCATGATCAATACAACAACCCAATG
GCTTCCATCGGTTGTGGTAGCTCATCCAAGAACAGAGAATCTGAACACCGGTGTCAATCAAGTCTACCGTA------TTGATCAATACAACAATCCGATG
AGGTGAAGAAGCTGAGAGCTATACTAAGAGACCAAGGTTTGATCAAGAAGCCAATCTCTGCCGGGATCATCAAAGTTTCCAGTGAGGAAGACACGGCGGA
AGGTGAAAAAGCTGAGAGCTATACTAAGAGAGCAAGGTTTGAGTAAGAAGCCACTCTCTGCCGGGACCATCAAAGCTTGCAGTGAGGAAGACACCGAGGA
CTAATTAAATTAAATCAAATCATT---GTTTT----TTTATCCCTTAGTTCTTGATTCTATCAGACTGCCCCCTTACAAAATTGGACTGGT---CAATGC
CTCGCTTAGACAAGAGTACAACGCCGTTTCTAGAGAGAAGCAGATGTTACACGAGGAGGTCTG---TGTACCATATTGCACTATTTTAAAACCATGGGGC
CTCACTTAGACAAGAGTACGACGCCGTTTCTAGGGAGAAGCAGATGTTACACGAGGAGGTCTGTACTGTACCATATTGCACTCTTTTAAAATCGCGGGGC
AGCTGTCGAGAGAGCTCGGTCTGCAACCACGTCAGATAGCGGTTTGGTTCCAGAACCGCCGTGCACGCTGGAAGGCGAAGCAGCTCGAGCAGTTGTACGA
AGCTGTCGAGAGAGCTCGGTCTGCAGCCACGTCAGATAGCGGTTTGGTTCCAGAACCGCCGTGCACGGTGGAAGGCGAAGCAGCTCGAGCAGTTGTACGA
GAGATGATGATAAAAAAGAAGAGACTAACGAGTGGACAATTAGCTTCACTTGAACGGAGTTTTCAAGAAGAGATCAAACTGGATTCAGACAGGAAGGTGA
GAGATGATG---AAGAAGAAGAGACTGACGAGTGGGCAATTAGCTTCACTTGAACGGAGTTTTCAAGAAGAGATCAAACTGGATTCAGACAGGAAGGTGA
ATACACGCCGGCAACCGATACGCAAACCGGACCGGTTATTGCTGTACCGGAACCAGAAAAGATCATGAATGCGTACCGATTTCCGAA---CAACAACAAC
ATACACGCCGCCTGCCGATACGCAAACCGGACCGGTTATTTCTGTACCGGAACCTGAAAAGATCATGAATGCGTACCGATTTCCGAACAGCAACAACAAC
GGAGGGTATAGTTCTCTCTAAGATAATGGTTGCGG----------ACATATATAATTTTATTTTTTCTTACTAACTCACGGAAGTTATTTTATTCGGTTT
CTTTTTTTTCTTCTTGGGATCCCAATATTTTTATTCATCGGGATCCAAATGCAAAAAGTTAAAACGTCACTTTTGAAAAGTTGAGACTTGAGAGATTCGATG
----------TTCTTGGGATCCCAATTTTTATTCATCGGGATCCAAATGCAAAAAGTTAAAACGTCACTTTTGAAAAG-------TTGAGAGATTCGTTG
ACCTTGGGATATGTC--ATCTAAAAACGTACTTTTCTGATGGGAAGTTAAGATTTTTTCTTATCTCATCAGCTTAGACAATT------------------
Fig. S9: Sequence alignment between the C. rubella and N. paniculata RCO locus.

Nucleotide Identities are illustrated by a color shading.
**Supplementary Table:**

**Table S1: List of primers used in this study**

<table>
<thead>
<tr>
<th>Primer</th>
<th>sequence</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>oAS1316</td>
<td>GGGATTGTCTCGCTTCTCTCTGAAGATGTTGG</td>
<td>TILLING RCO-A</td>
</tr>
<tr>
<td>oAS633</td>
<td>GTAGGGGCTGCGATATC</td>
<td>TILLING RCO-A</td>
</tr>
<tr>
<td>oAS 1318</td>
<td>TCTAGTATCGATGATGACAGTTTTTCGGTGG</td>
<td>TILLING RCO-B</td>
</tr>
<tr>
<td>oAS 1319</td>
<td>CATAACACTCTGCTCTGCAAATGCCAATGCC</td>
<td>TILLING RCO-B</td>
</tr>
<tr>
<td>oAS666</td>
<td>CTACCATGGTAAATCCAGAAGGTGCTTCTCTGATGCTGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS667</td>
<td>ATGCCCTGCAGTGACATTGATAGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1164</td>
<td>CTACCATGGTAAATCCAGAAGGTGCTTCTCTGATGCTGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1165</td>
<td>ATGCCCTGCAGTGACATTGATAGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1227</td>
<td>CTACCATGGTAAATCCAGAAGGTGCTTCTCTGATGCTGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1224</td>
<td>ATGCCCTGCAGTGACATTGATAGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1338</td>
<td>GCGATATGGGCTGCTTCTCAGGACATAGCCGAGAAGGAAGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1339</td>
<td>TTTCTTAATTATTTACAAGAAGACACG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1332</td>
<td>ATGCCCTGCAGTGACATTGATAGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1333</td>
<td>TATTTTTTCTTTAATAAGAAGCCGAGAAGGAAGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS599</td>
<td>ATGCCCTGCAGTGACATTGATAGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS598</td>
<td>CTACCATGGTAAATCCAGAAGGTGCTTCTCTGATGCTGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1188</td>
<td>GAGCCATCGGTAAGACGACCACCG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1182</td>
<td>GAGCCATAGATTCTATCTCTGCAGAAGGATCCATTACAAT</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1175</td>
<td>GAGCCATCCAGGTTAGGGCGCTTCTCTGATGCTGCTTGTTTTCGGTGG</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1189</td>
<td>CTTACCGATGGGCTTCTCCAAGAAGAAGAGA</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1183</td>
<td>TAGAATCTAGGCTTCCAAGAAGAGA</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1176</td>
<td>ACCGCTGAGGTGCTTCTCCAAGAAGAAGAGA</td>
<td>Cloning</td>
</tr>
<tr>
<td>oAS1086</td>
<td>TTCGAACCAGCTTCTGAAGCTG</td>
<td>qPCR</td>
</tr>
<tr>
<td>oAS1087</td>
<td>CCTCACAAAGGTTGCTGAC</td>
<td>qPCR</td>
</tr>
<tr>
<td>oAS628</td>
<td>TCTGAGCGAGCTGAGATAG</td>
<td>qPCR LMI1</td>
</tr>
<tr>
<td>oAS629</td>
<td>TCGTACAGTTGCTGAGCAGT</td>
<td>qPCR</td>
</tr>
<tr>
<td>oAS630</td>
<td>ATCGGGGATACAAACACGAG</td>
<td>qPCR RCO-B</td>
</tr>
<tr>
<td>oAS631</td>
<td>ATCGGGGATACAAACACGAG</td>
<td>qPCR RCO-B</td>
</tr>
<tr>
<td>oAS632</td>
<td>ATCGGGGATACAAACACGAG</td>
<td>qPCR RCO-A</td>
</tr>
<tr>
<td>oAS633</td>
<td>GTAGGGGCTGCGATATC</td>
<td>qPCR</td>
</tr>
<tr>
<td>oAS1207</td>
<td>CTGCAACCACGTCAGATAG</td>
<td>qPCR NpLMI1</td>
</tr>
<tr>
<td>oAS1208</td>
<td>CTACAGGATGAGGAGGAGGCTGAGA</td>
<td>qPCR</td>
</tr>
<tr>
<td>oAS625</td>
<td>TACACGGGATACAAACACGAG</td>
<td>qPCR NpRCO</td>
</tr>
<tr>
<td>oAS1520</td>
<td>ATCGGAGCATATGCTCCGCTTCT</td>
<td>genotyping</td>
</tr>
<tr>
<td>oAS1521</td>
<td>GTTGGTGTCGTAAGACCCCGTG</td>
<td>RCO-A</td>
</tr>
</tbody>
</table>

*Used with oAS629*
Supplementary References:


