SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
LEAF DEVELOPMENT
EVOLUTIONARY BIOLOGY

Received
20 January 2014

Accepted
31 January 2014

Published
18 February 2014

Correspondence and requests for materials should be addressed to J.L.M. (jlmicol@umh.es)

Combined haploinsufficiency and purifying selection drive retention of *RPL36a* paralogs in Arabidopsis

Rubén Casanova-Sáez, Héctor Candela & José Luis Micol

Instituto de Bioingeniería, Universidad Miguel Hernández, Campus de Elche, 03202 Elche, Alicante, Spain.

Whole-genome duplication events have driven to a large degree the evolution of angiosperm genomes. Although the majority of redundant gene copies after a genome duplication are lost, subfunctionalization or gene balance account for the retention of gene copies. The Arabidopsis 80S ribosome represents an excellent model to test the gene balance hypothesis as it consists of 80 ribosomal proteins, all of them encoded by genes belonging to small gene families. Here, we present the isolation of mutant alleles of the *APICULATA2* (*API2*) and *RPL36aA* paralogous genes, which encode identical ribosomal proteins but share a similarity of 89% in their coding sequences. *RPL36aA* was found expressed at a higher level than *API2* in the wild type. The loss-of-function *api2* and *rpl36aa* mutations are recessive and affect leaf development in a similar way. Their double mutant combinations with *asymmetric leaves2-1* (*as2-1*) caused leaf polarity defects that were stronger in *rpl36aa as2-1* than in *api2 as2-1*. Our results highlight the role of combined haploinsufficiency and purifying selection in the retention of these paralogous genes in the Arabidopsis genome.

Ancient genome duplication (polyploidization) events have played a central role in the evolution of angiosperm genomes[1]. Several models account for the fate of individual genes after a duplication occurs. The redundant gene copies might be inactivated and lost or, alternatively, they might be retained in the genome by different evolutionary processes. In the duplication-degeneration-complementation model, subfunctionalization (i.e. the distribution of the ancestral functions among the duplicated genes) accounts for the evolutionary maintenance of duplicate gene copies[2]. Duplicated genes might also be maintained, retaining the same (redundant) function and expression pattern, when the duplication leads to a new balance between the products of dosage-dependent genes (gene balance hypothesis)[3,4].

The availability of complete plant genome sequences has revealed the presence of recent and ancient whole-genome duplication events in the evolutionary history of numerous plant species. Three rounds of polyploidization, followed by extensive chromosomal rearrangements and gene loss to reach the present diploid state, have been documented in the genome of *Arabidopsis thaliana* (hereafter, Arabidopsis)[5]. The phenomenon by which individual genes are lost from one of the two homeologous chromosome pairs after a whole-genome duplication event is called fractionation[6]. Somewhat counter-intuitively, it has been demonstrated that fractionation is a biased phenomenon, with a majority of genes being preferentially lost from one of the homeologous chromosomes[7]. Interestingly, some genes appear to have escaped fractionation, so that two or more gene copies resulting from duplication events have persisted in the genome[8]. While retention of some gene pairs might be attributed to subfunctionalization, the gene balance hypothesis accounts for the retention in other cases. Indeed, the retained genes have been found to preferentially belong to certain over-represented functional categories[7], which typically comprise proteins that are dosage-dependent or form part of multisubunit complexes, often with regulatory roles, suggesting that haploinsufficiency has played a major role in their retention. Loss-of-function mutations at haploinsufficient loci are typically dominant because the level of gene function in a heterozygote is below the threshold to produce a wild-type phenotype.

The eukaryotic ribosome represents an excellent model for testing the gene balance hypothesis, as it consists of a large subunit (60S) with 46 ribosomal proteins and 3 ribosomal RNA (rRNA) molecules (25S, 5.8S and 5S), and a small subunit (40S) with 33 ribosomal proteins and a single 18S rRNA molecule[9]. The structure of the prokaryotic and eukaryotic ribosomes has been determined at an atomic-level resolution[10–15]. Prokaryotic and eukaryotic ribosomes share a general macroorganization but they differ in size and number of ribosomal proteins and rRNA molecules. In Arabidopsis, all cytosolic ribosomal proteins are encoded by small gene families, each with 2 to 7 members[16]. Although 249 genes encoding ribosomal proteins have been annotated in the Arabidopsis genome, a maximum of 80 different ribosomal proteins has been identified in isolated 80S ribosomes[17–19].

Here, we report the isolation of mutant alleles of two genes, *APICULATA2* (*API2*) and its closest paralog (i.e. homologous gene resulting from a duplication event) in the Arabidopsis genome, At3g23390. These genes encode two identical ribosomal proteins, RPL36aB and RPL36aA, despite they carry nucleotide substitutions at 34% of their codons. Loss-of-function mutations at each of these unlinked loci are recessive and affect leaf development in a similar way. We found that the *api2* and *rpl36aa* mutations represent a rare example of non-allelic non-complementation, indicating that two (out of the four) functional gene copies present in a diploid plant are not sufficient to produce a wild-type phenotype. We demonstrate that both paralogous genes are expressed in the same tissues and show that both loci behave as a single haploinsufficient unit with four alleles (combined haploinsufficiency) on which natural selection has acted to filter out deleterious non-synonymous substitutions (i.e. nucleotide substitutions that change the protein sequence), a process known as purifying selection.

Results

Cloning and molecular characterization of the *api2* mutation. In a large-scale screen for ethylmethane sulfonate (EMS)-induced mutants with abnormal leaf shape, we previously isolated 42 mutants with a pointed-leaf phenotype. These mutants were assigned to phenotypic classes dubbed Apiculata (Api), Denticulata (Den) and Angusta (Ang)[20,21]. The Api class comprises recessive alleles of seven different genes, named *API1* to *API7*[20]. Here, we focus on the *API2* gene, which was initially defined by a single mutant allele, *api2*. To identify *API2* at the molecular level, we undertook a positional approach. We first mapped the *api2* mutation to an 850-kb interval on chromosome 4, flanked by two insertion-deletion (InDel) markers, cer459267 and cer453616 (Figure 1a), encompassing 248 annotated genes. Because the *api2/api2* mutant has pointed leaves (Figure 2b, d, g, i), and is very similar in phenotype to other mutants affected in genes that encode ribosomal proteins[22–27], we focused on the only two genes within the interval that encode subunits of the cytosolic ribosome, At4g13170 and At4g14320. After sequencing the transcriptional unit of both genes, we found a single G-to-A transition mutation in the 5′ untranslated region (UTR) of the At4g14320 gene (Figure 1b and Supplementary Figure S1), also known as *RPL36aB*[17]. This mutation creates an out-of-frame ATG codon upstream of the normal translation start site of *RPL36aB* transcripts (Supplementary Figure S1). We used real-time qRT-PCR to measure the relative *RPL36aB* expression level in rosettes collected 14 days after stratification (das), and found a 0.57-fold decrease in *api2/api2* plants compared with the wild type (Figure 3). To determine whether the lesion in *RPL36aB* is responsible for the mutant phenotype, we followed a transgenic approach to complement the *api2* mutation. The phenotype of *api2/api2* plants expressing a *35S$_{pro}$:RPL36aB* transgene was indistinguishable from wild type in 4 independent lines (Figure 2l, o), indicating that loss of *RPL36aB* function causes the mutant phenotype, either by the observed down-regulation of transcript levels or because the *api2* mutation interferes with normal translation initiation.

Two paralogous genes encode the RPL36a protein in Arabidopsis. The Arabidopsis genome contains two *RPL36a* paralogous genes that encode identical proteins, *RPL36aA* (At3g23390) and the above-mentioned *RPL36aB* (hereafter, *API2*)[17]. We also searched for mutant alleles of *RPL36aA*, and identified a publicly available T-DNA insertion line (SALK_148438) that, when homozygous for the insertion, exhibited a pointed-leaf phenotype identical to that caused by the *api2* mutation in the Col-0 genetic background

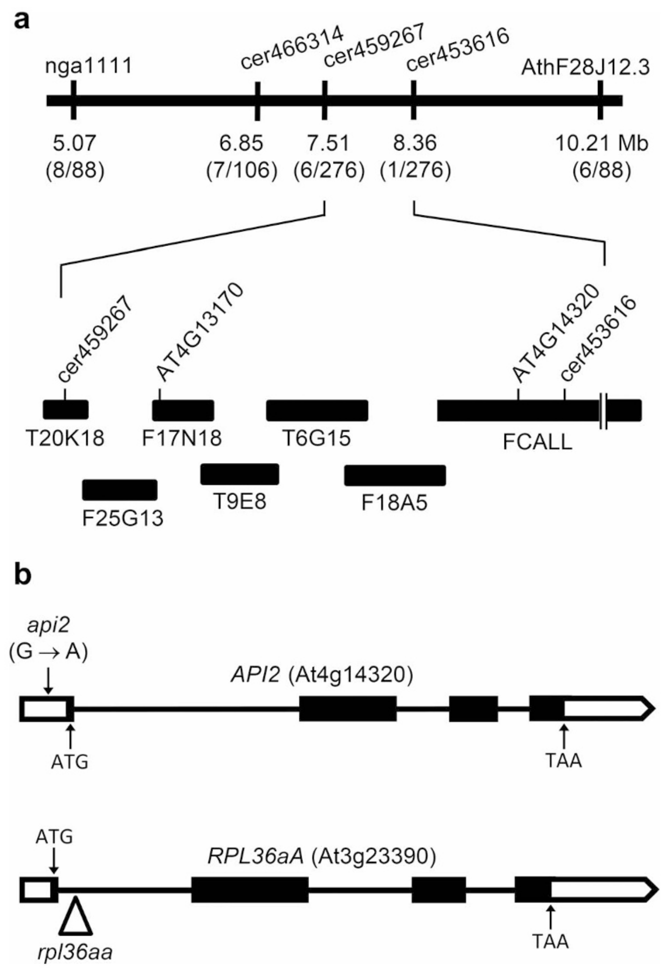

Figure 1 | Positional cloning of the *API2* gene and molecular nature and position of the *api2* and *rpl36aa* mutations. (a) A mapping population of 138 F_2 plants derived from an *api2/api2* × Col-0 outcross was used to define a candidate interval of 850 kb on chromosome 4. The names and map positions of the molecular markers used are indicated. The number of recombinant chromosomes found and the total of chromosomes analyzed are shown in parentheses. Black boxes represent BAC clones. (b) Structure of the *API2* and *RPL36aA* paralogous genes, with indication of the nature and position of the *api2* and *rpl36aa* mutations. A triangle indicates the position of a T-DNA insertion. Boxes and lines between boxes indicate exons and introns, respectively. Open boxes correspond to the 5′ and 3′ untranslated regions. The predicted translation start (ATG) and stop (TAA) codons are indicated.

(Figure 2d, e, i, j). We refer to this line as *rpl36aa/rpl36aa*. The presence of the T-DNA insertion in the first intron of *RPL36aA* was confirmed by PCR genotyping (Figure 1b). The inflorescences of *rpl36aa/rpl36aa* and *api2/api2* plants were shorter than those of wild-type plants (Figure 2k). We used real-time qRT-PCR to measure the relative *RPL36aA* expression level in *rpl36aa/rpl36aa* rosettes collected 14 das, and only found a residual expression level (Figure 3), suggesting that *rpl36aa* is a null allele. To demonstrate that the insertion causes the defects observed in the *rpl36aa* line, we also followed a transgenic approach to complement its mutant phenotype. Four independent *rpl36aa/rpl36aa;35S$_{pro}$:RPL36aA* transgenic lines exhibited a wild-type leaf phenotype (Figure 2m, p), confirming that the mutant phenotype results from the inactivation of *RPL36aA*. We concluded that identical recessive leaf mutant phenotypes occur when either *API2* or *RPL36aA* are inactivated.

***API2* and *RPL36aA* are dosage-dependent genes.** Because the API2 and RPL36aA proteins have identical amino acid sequences

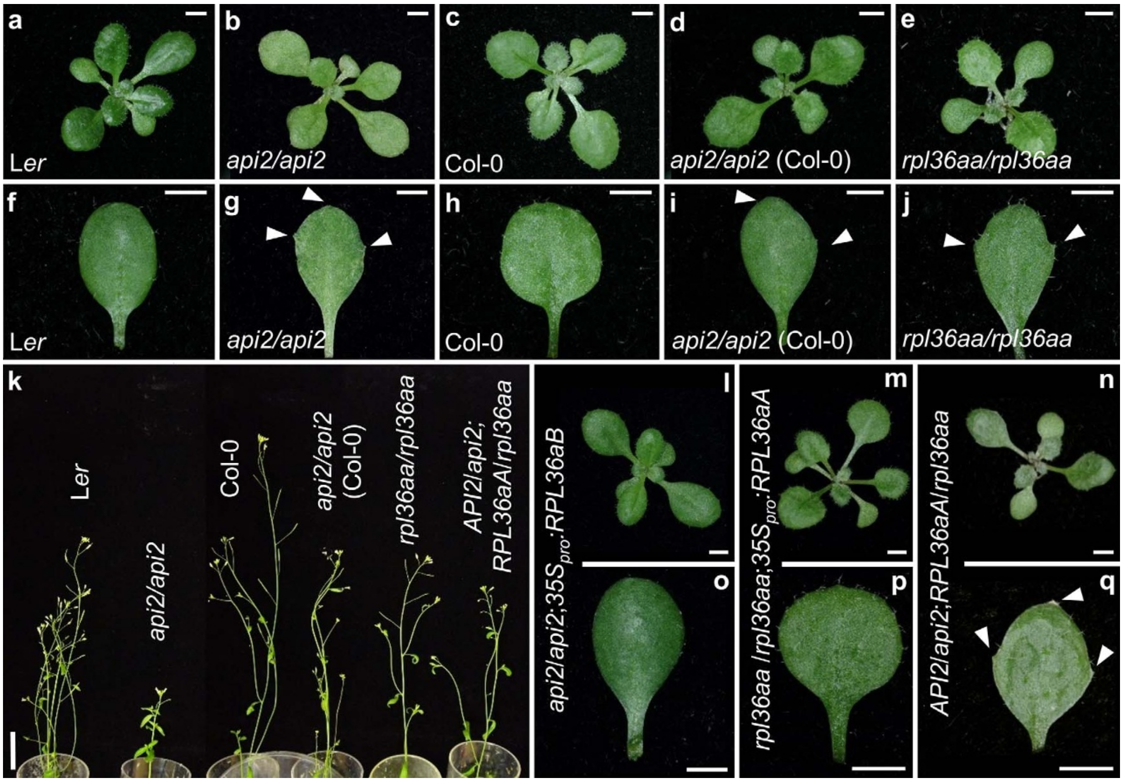

Figure 2 | Phenotypes and phenotypic rescue of the *api2* and *rpl36aa* mutants. (a–e), (l–n) Rosettes and (f–j), (o–q) first-node leaves from (a), (f) L*er*, (b), (g), (d), (i) *api2/api2* (b), (g) in a L*er* background and (d), (i) in a Col-0 background, (c), (h) Col-0, (e), (j) *rpl36aa/rpl36aa*, (l), (o) *API2/api2;RPL36aA/rpl36aa*, (m), (p) *api2/api2;35S$_{pro}$:RPL36aB* and (n), (q) *rpl36aa/rpl36aa;35S$_{pro}$:RPL36aA* plants. Arrowheads indicate marginal teeth in mutant pointed leaves. (k) Adult plants. All plants are in a Col-0 background, the only exceptions being (b), (g) the *api2* mutant and (a), (f) the L*er* wild type. Pictures were taken (a–j), (l–q) 14 and (k) 35 das. Scale bars indicate (a–j), (l–q) 2 mm and (k) 3 cm.

(Supplementary Figure S2), they are expected to perform identical functions in the ribosome, offering a unique opportunity to ask questions about the functional redundancy and haploinsufficiency of paralogous ribosomal proteins, with no interference due to the divergence of their sequences. The *api2/api2* and *rpl36aa/rpl36aa* single mutants are viable, suggesting that both genes contribute quantitatively to a global pool of RPL36a ribosome subunits and that two functional copies of any one gene can alleviate the effects of having two inactive copies of its paralog. We hypothesized that the simultaneous loss of one copy of each locus should also lead to a mutant phenotype similar to those of *api2/api2* or *rpl36aa/rpl36aa* single mutants. We crossed the single mutants and found that all the doubly heterozygous plants (*API2/api2;RPL36aA/rpl36aa*) in the F_1 progeny exhibited a pointed-leaf phenotype identical to those of the *api2/api2* and *rpl36aa/rpl36aa* homozygous parentals (Figure 2n, q). This non-allelic non-complementation involving hypomorphic[28] (partial loss-of-function) alleles of *API2* and *RPL36aA* requires that both genes play redundant functions and are co-expressed in at least some cells.

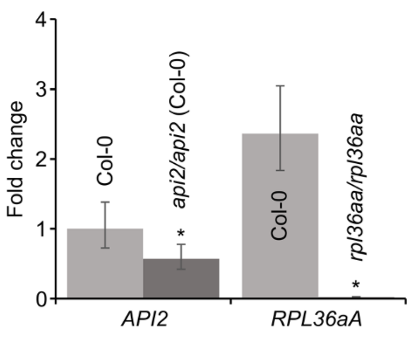

Figure 3 | Expression of the *API2* and *RPL36aA* genes in the Col-0 wild type and the *api2/api2* and *rpl36aa/rpl36aa* mutants. Real-time qRT-PCR analysis of the expression of the *API2* and *RPL36aA* genes was performed in *api2/api2* and *rpl36aa/rpl36aa* vegetative leaves. Relative expression levels are shown, determined by the comparative C_T method, and normalized with the expression of the *18S rRNA* housekeeping gene. Error bars indicate the interval delimited by $2^{-(\Delta\Delta C_T \pm SD)}$. Asterisks indicate ΔC_T values significantly different from those of Col-0 in a Mann-Whitney U test ($p < 0.01$; $n = 9$).

In order to investigate the effects of carrying less than two wild-type alleles, we genotyped 169 mutant plants from an F_2 family segregating mutant alleles of both loci. We did not find plants with a single functional copy of either gene (*API2/api2;rpl36aa/rpl36aa* or *api2/api2;RPL36aA/rpl36aa*) or lacking wild-type copies (*api2/api2;rpl36aa/rpl36aa*), despite these genotypes are collectively expected to appear in 31.25% (5 out of 16) of the F_2 plants. Interestingly, the progeny of self-pollinated *API2/api2;RPL36aA/rpl36aa* plants segregated an excess of aborted ovules (Figure 4), with a 13% increase relative to the single mutants and the wild type.

To test any male- or female-gametophytic lethality, we performed reciprocal crosses between doubly heterozygous *API2/api2;RPL36aA/rpl36aa* and Col-0 plants. When we pollinated *API2/api2;RPL36aA/rpl36aa* pistils with Col-0 wild-type pollen, we observed an average 23% excess of aborted ovules in the siliques relative to the values obtained in control crosses, indicating female gametophytic lethality. Reciprocally, we also pollinated Col-0 pistils with pollen from *API2/api2;RPL36aA/rpl36aa* plants. If the four possible types of haploid male gametophytes are equally transmitted to the progeny, then one fourth of the F_1 plants should have the same genotype (*API2/api2;RPL36aA/rpl36aa*) and exhibit the same mutant phenotype as their male parental. However, all 140 F_1 plants

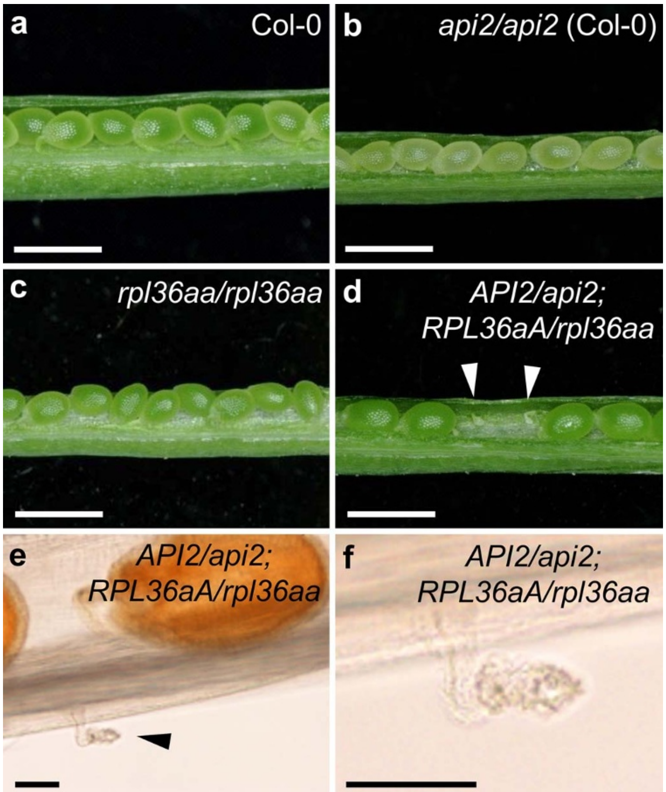

Figure 4 | Aborted ovules in *API2/api2;RPL36aA/rpl36aa* siliques. (a–d) Dissected mature siliques from (a) Col-0, (b) *api2/api2* (Col-0), (c) *rpl36aa/rpl36aa* and (d) *API2/api2;RPL36aA/rpl36aa* plants. (e), (f) Detail of an aborted ovule in an *API2/api2;RPL36aA/rpl36aa* cleared silique. Arrowheads in (d) and (e) indicate aborted ovules. Scale bars indicate (a–d) 1 mm and (e), (f) 100 μm.

studied showed a wild-type phenotype, a result that indicates that the *api2 rpl36aa* gametophytes are not viable.

***RPL36a* genes are subjected to strong purifying selection.** The identical protein sequences encoded by different *RPL36a* paralogous genes might occur as a consequence of their recent evolutionary origin; alternatively, they might be explained by concerted evolution (e.g. gene conversion) or by strong purifying selection[29]. The coding regions of both *API2* and *RPL36aA* transcripts consist of 105 codons. A pairwise alignment of the *API2* and *RPL36aA* coding sequences uncovered 36 nucleotide substitutions, each affecting a different codon (Supplementary Figure S1). All these substitutions were synonymous, affecting the third base of a codon in 35 cases and the first base only in 1 case (Supplementary Figure S2). Because the affected codons were scattered along the sequence, this level of divergence is unlikely to have arisen as a consequence of concerted evolution or a recent duplication. We determined the rates of synonymous (d_S) and non-synonymous substitutions (d_N) from a pairwise alignment of the *API2* and *RPL36aA* coding sequences. The low value of the d_N/d_S ratio (Table 1) is a signature of strong purifying selection, and is in line with the values obtained for other pairs of paralogous genes encoding identical ribosomal proteins.

***API2* and *RPL36aA* are expressed in overlapping patterns.** As mentioned above, the non-allelic non-complementation between *API2* and *RPL36aA* is best understood when there is at least some overlap of expression in the same cells. To experimentally demonstrate this, we fused the promoter regions of both genes to the GUS reporter gene. GUS staining of 5 $API2_{pro}$:*GUS* and 5 $RPL36aA_{pro}$:*GUS* independent transgenic lines in the Col-0 background uncovered an identical expression pattern for both genes. Both constructs were ubiquitously expressed, with particularly intense GUS staining in young and actively proliferating tissues, such as those of developing leaves, floral buds and root apices (Figure 5). Interestingly, the relative intensities of the GUS staining matched our qRT-PCR results (Figure 3), with the *RPL36aA* promoter conferring higher expression levels in wild-type vegetative leaves than the *API2* promoter.

***api2* and *rpl36aa* genetically interact with *asymmetric leaves* mutants.** Some ribosomal proteins are known to be involved in leaf developmental programs, as their mutations enhance the adaxial-abaxial polarity defects of the *asymmetric leaves1* (*as1*) and *as2* mutants[24,25]. Moreover, Horiguchi *et al.*[26] showed that the extent of this enhancement varies among mutants affected in different ribosomal proteins. To determine whether *api2* and *rpl36aa* can enhance the polarity defects of *as* mutants, we isolated the *api2 as1-1*, *api2 as2-1*, *rpl36aa as1-1* and *rpl36aa as2-1* double mutants. Because both paralogous proteins share an identical sequence and expression pattern, any phenotypic differences observed in these mutants are likely to reflect the contribution of individual paralogs to the pool of RPL36a subunits. Both *api2* and *rpl36aa* enhanced the phenotype of the *as2-1* mutant, although to a different extent (Figure 6e–h). Both *api2 as2-1* and *rpl36aa as2-1* double mutants exhibited trumpet-shaped (peltate) leaves (Figure 6f, g), which we did not observe in the *as2-1* single mutant (Figure 6d, e). In addition to this, the *rpl36aa as2-1* double mutant exhibited radial leaves (Figure 6h), which can be interpreted as a consequence of a more severe abaxialization. These phenotypes were not observed in the double mutant combinations involving the *as1-1* mutation (Figure 6b, c).

Table 1 | Selected ribosomal protein paralogs and their synonymous and non-synonymous substitution rates

Paralog pairs	AGI gene codes	Protein length (aa)*	Amino acid sequence identity (%)	Nucleotide sequence identity (%)	d_S	d_N	d_N/d_S
API2, RPL36aA	At4g14320, At3g23390	105	100	89	1.1338	0.0011	0.0010
RPS18A, RPS18C	At1g22780, At4g09800	152	100	88	1.1036	0.0056	0.0051
RPS28A, RPS28B	At3g10090, At5g03850	64	100	88	0.9059	0.0009	0.0010
RPL21A, RPL21D	At1g57860, At1g57660	164	100	99	0.0339	0.0000	0.0010
RPL38A, RPL38B	At2g43460, At3g59540	69	100	90	0.7566	0.0008	0.0010
RPL41A, RPL41B	At2g40205, At3g08520	25	100	95	0.3252	0.0003	0.0010
RPS13A, RPS13B	At4g00100, At3g60770	151	99	85	1.5649	0.0029	0.0018
RPS21A, RPS21B	At3g53890, At5g27700	82	94	83	3.9235	0.0297	0.0076
RPL11A, RPL11C	At2g37190, At5g60670	166	93	81	1.0114	0.0572	0.0566
RPL28A, RPL28B	At2g19730, At4g29410	143	90	85	0.7896	0.0597	0.0756

*All paralog pairs shown in this table have identical protein length.

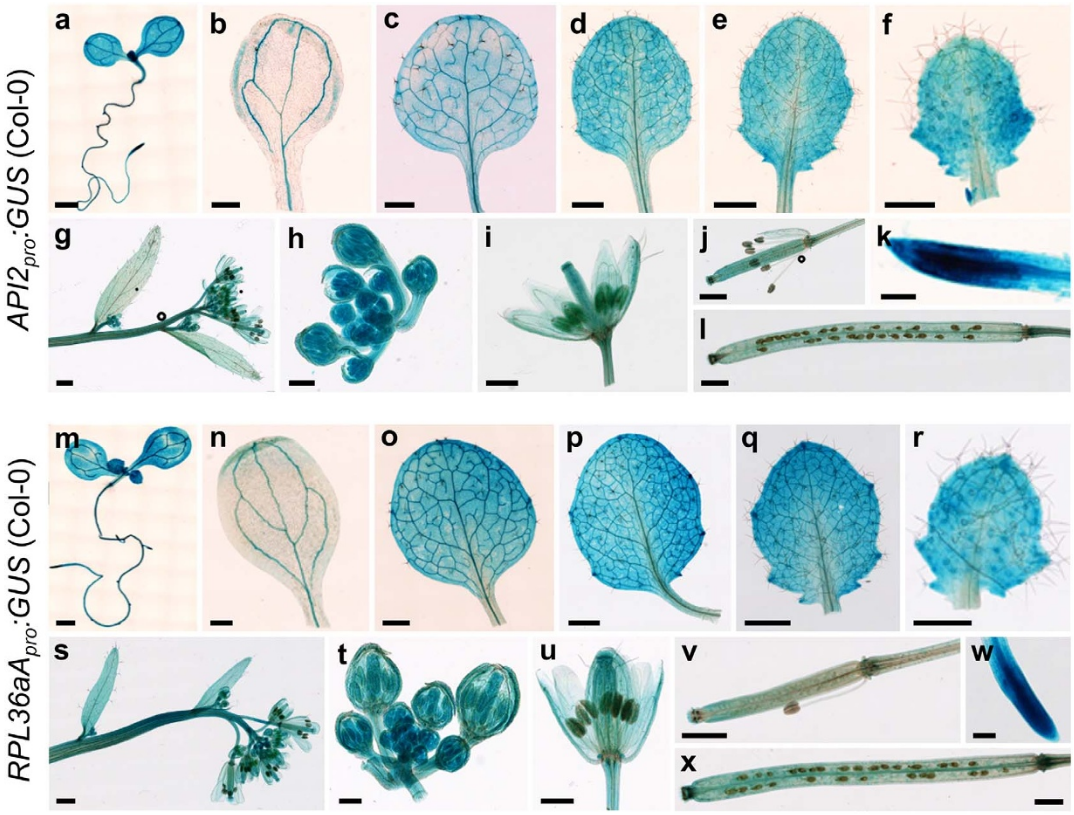

Figure 5 | Visualization of $API2_{pro}$:*GUS* and $RPL36aA_{pro}$:*GUS* transgene activity in the Col-0 background. (a), (m) Seedlings and (b), (n) cotyledons. (c–f), (o–r) Vegetative leaves from the (c), (o) first, (d), (p) third (e), (q) fifth and (f), (r) seventh nodes. (g), (s) Inflorescences. (h), (t) Immature and (i), (u) mature flowers. (j), (v) Immature siliques, (k), (w) root tips and (l), (x) mature siliques. Pictures were taken (a), (m) 6, (b–f), (k), (n–r), (w) 14 and (g–j), (l), (s–v), (x) 35 das. Scale bars indicate (a), (c–e), (g), (j), (l), (m), (o–q), (s), (v), (x) 1 mm, (b), (f), (i), (n), (r), (u) 500 μm, (h), (t) 200 μm and (k), (w) 100 μm.

Discussion

Following a positional cloning approach, we have determined that the *api2* mutant carries a lesion at a gene encoding the RPL36a ribosomal protein. This ribosomal protein is encoded by two paralogous genes, *API2* (*RPL36aB*) and At3g23390 (*RPL36aA*), which have not previously been characterized in Arabidopsis. Our results indicate that both paralogs are functionally equivalent (redundant), as shown by the non-allelic non-complementation observed in the F_1 progeny of crosses involving mutant alleles of both genes. This functional equivalence is also supported by the identical protein sequences and expression patterns conferred by the promoter regions of *API2* and its paralog. This non-allelic non-complementation indicates that a critical level of protein function is not reached in doubly heterozygous plants, a quantitative effect that is expected only when the two genes are expressed in the same cells and perform identical function. Although this type of non-complementation is a relatively rare phenomenon, numerous additional instances have been documented in Arabidopsis, often involving ribosomal proteins of the large and small subunits of the cytosolic ribosome[25,27,30,31]. The observation that loss-of-function mutations damaging ribosomal proteins are usually recessive in Arabidopsis[26], with only a few examples of semi-dominance[32,33], is in striking contrast with the dominance or haplolethality of loss-of-function (haploinsufficient) *Minute* mutations of *Drosophila melanogaster*[34]. The *Minute* mutations represent the most abundant phenotypic class in Drosophila and affect genes coding for ribosomal proteins. However, unlike in Arabidopsis, each ribosomal protein is typically encoded by a single gene in the genome of Drosophila[34]. This fundamental difference reflects how the need to keep a balance between the different ribosomal proteins after a whole-genome duplication has shaped the genome during plant evolution.

Loss of *API2* or *RPL36aA* functions led to a deleterious recessive syndrome that was very similar, not to say undistinguishable, from the phenotypes of many other mutants carrying lesions in genes encoding ribosomal proteins[22,24–26,30,31,35]. Our reciprocal crosses indicate that the simultaneous inactivation of both paralogs causes lethality and prevents transmission of the mutant alleles through the male and female gametophytes, as has also been reported for other pairs of paralogous ribosomal proteins in Arabidopsis[36]. Indeed, we were not able to isolate plants with fewer than two active (out of the four) copies of the genes encoding the RPL36a ribosomal protein. The observation that only synonymous nucleotide substitutions have accumulated in the coding region of these paralogous genes (none of which changes the protein sequence) suggests that a matching number of non-synonymous changes have also occurred but must have been selected against by strong purifying selection. A similar situation is likely to hold for other ribosomal proteins for which lethality, non-allelic non-complementation and/or dosage effects have been described. The presence of identical sequences in both paralogs (at the amino acid level) is not a peculiarity of the RPL36a paralogy group, as several instances of gene pairs encoding identical ribosomal proteins have been reported in the literature (e.g. RPS18, RPS29, RPS30, RPL11, RPL21, RPL23, RPL38, RPL40 and RPL41)[37]. We have found that these proteins have also been subjected to strong purifying selection, as inferred from the low d_N/d_S values for pairs of paralogous genes, highlighting the critical roles played by such proteins in the ribosome. Purifying selection has also been shown to account for the sequence conservation of other proteins that play essential, highly conserved functions in eukaryotes, such as members of the histone H4 protein family[29]. We propose that combined haploinsufficiency and strong purifying selection have driven the biased retention of genes coding for paralogous ribosomal proteins after the

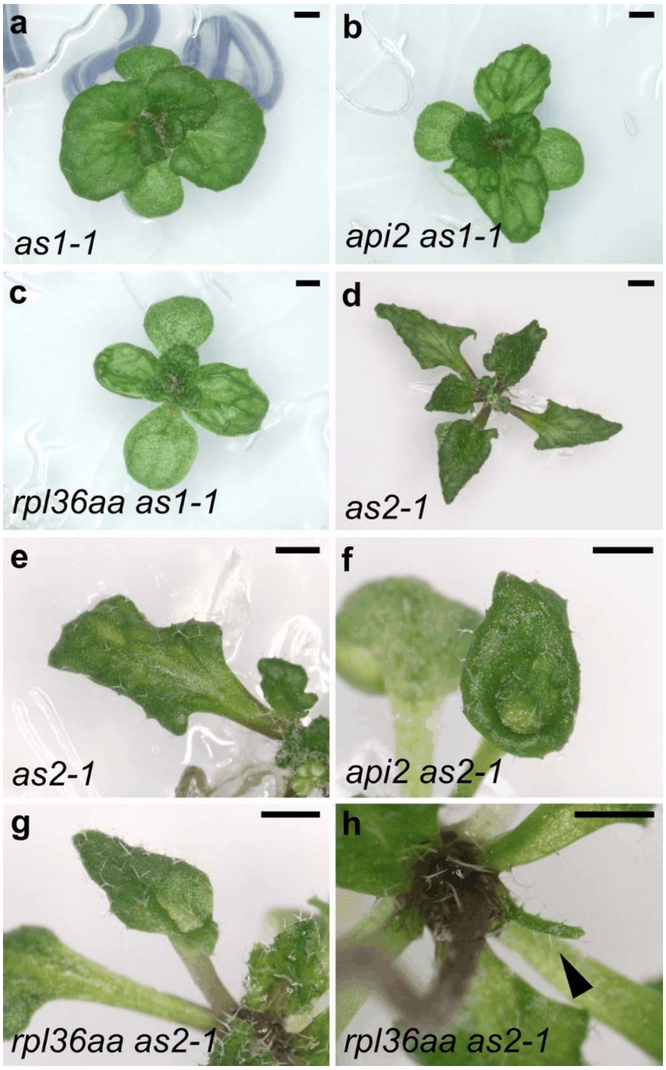

Figure 6 | Double mutant combinations of *api2* and *rpl36aa* with *as1-1* and *as2-1*. (a–d) Rosettes of the (a) *as1-1*, (b) *api2 as1-1*, (c) *rpl36aa as1-1* and (d) *as2-1* mutants. (e–f) Details of vegetative leaves from (e) *as2-1*, (f) *api2 as2-1* and (g), (h) *rpl36aa as2-1* mutant plants. All these mutants were homozygous for the mutations indicated. Pictures were taken (a–c) 14 and (d–f) 18 das. Scale bars indicate 1 mm.

duplication of plant genomes, as has been documented for the Arabidopsis genome[7,8]. The stoichiometry and structural constrains of a multisubunit complex, such as the ribosome, are likely the ultimate causes behind this retention[16].

Because API2 and RPL36aA are identical at the amino acid level, any phenotypic difference between the *api2* and *rpl36aa* loss-of-function mutant alleles should be attributed to differences in the spatio-temporal expression pattern or expression levels of the two genes. In our particular case, we observed that *rpl36aa* enhanced the polarity defects of the *as2-1* mutant (as inferred from the fully radialized lateral organs seen in the *rpl36aa as2-1* double mutant) to a greater extent than *api2* (as the *api2 as2-1* double mutant showed trumpet-shaped, but not fully radialized leaves). Interestingly, we found *RPL36aA* to be expressed at higher levels than *API2* in wild-type plants, suggesting that the stronger effect caused by the *rpl36aa* mutation is due to the fact that *RPL36aA* is the main contributor of RPL36a subunits. Our results favor the hypothesis that, at least in plants, the phenotypic differences observed between mutant alleles of paralogous genes, which have been reported in some cases[36,38], might result from differences in the expression patterns or the expression levels of individual paralogous proteins, rather than from the differential specificity for certain mRNAs of an heterogeneous population of ribosomes[17,19,39–41].

Methods

Plant material and growth conditions. *Arabidopsis thaliana* (L.) Heynh. Columbia-0 (Col-0) and Landsberg *erecta* (L*er*) wild-type accessions and the SALK_148438 line (named here *rpl36aa*; accession number N648438), were obtained from the Nottingham Arabidopsis Stock Centre (NASC). The *api2/api2* mutant was isolated after EMS mutagenesis of Landsberg *erecta* (L*er*) seeds[20]. To standardize genetic backgrounds, we outcrossed the *api2/api2* mutant three times to Col-0 [*api2/api2* (Col-0)]. Growth conditions and crosses were performed as described previously[20,42].

Map-based cloning. The *api2* mutation was low-resolution mapped as described in Ponce *et al.*[43]. For fine-mapping, we used the cer466314, cer459267 and cer453616 InDel polymorphisms (Table S1) from Monsanto (http://www.arabidopsis.org/browse/Cereon). To identify the *api2* mutation, the entire transcriptional unit of At4g14320 was amplified using the API2_F1 and API2_R1 primers (Table S1). The resulting 1660-bp PCR product was sequenced on an ABI PRISM 3130xl Genetic Analyser (Applied Biosystems). The presence of the T-DNA insertion in *rpl36aa* was confirmed by PCR using the rpl36aa_F/R primers designed with the T-DNA Primer Design tool (http://signal.salk.edu/tdnaprimers.2.html) (Table S1).

RNA isolation, cDNA synthesis and qRT-PCR. Total RNA from 20–30 mg of rosette leaves, collected 14 das, was isolated using TRI Reagent (Sigma) and first-strand cDNA synthesis and relative expression quantification were performed as described in Jover-Gil *et al.*[44], using the primers listed on Table S1.

Genetic constructs and plant transformation. To make the *$35S_{pro}$:RPL36aB* and *$35S_{pro}$:RPL36aA* transgenes, we amplified the full-length coding sequence of At4g14320 and At3g23390 genes from Col-0 cDNA using Phusion polymerase (Thermo Scientific) and the primers shown in Table S1. The resulting products were cloned into the pENTR/D-TOPO vector (Invitrogen) and transferred into the pMDC32 destination vector[45] in a reaction with LR clonase II (Invitrogen). To make the *$RPL36aB_{pro}$:GUS* and *$RPL36aA_{pro}$:GUS* transgenes, 389-bp and 716-bp fragments from the upstream region of the starting codon of At4g14320 and At3g23390 genes, respectively, were amplified from Col-0 genomic DNA using Phusion polymerase and the API2pro_F/R and RPL36aApro_F/R primers, which include the attB1 and attB2 sequences at their 5′ ends (Table S1). The amplification products were cloned into the pGEM-T Easy221 entry vector (kindly provided by B. Scheres), by recombination with BP clonase II (Invitrogen), and then transferred into pMDC163 destination vector[45] in an LR clonase II (Invitrogen) reaction. Competent *Agrobacterium tumefaciens* LBA4404 cells were transformed with these constructs. Col-0, *api2* and *rpl36aa* plants were transformed by the floral dip method[46].

GUS assay and microscopy. GUS staining was performed as described in Robles *et al.*[47], with minor modifications. In brief, plant tissue was incubated in ice-cold 90% acetone for 15 min and then in X-Gluc buffer solution (2 mM 5-bromo-4-chloro-3-indolyl-β-glucuronic acid, 50 mM sodium phosphate, pH 7.2, 5 mM potassium ferrocyanide, 50 mM potassium ferricyanide, and 0.2% Triton X-100) for 14 h at room temperature. After GUS detection, the tissue was cleared with 96% ethanol and mounted in an 8 : 2 : 1 (chloral hydrate : glycerol : water) solution. Samples were examined with a Nikon D-Eclipse C1 microscope.

Bioinformatic analysis of selection. To determine the impact of selection, we calculated synonymous (d_S) and non-synonymous (d_N) substitution rates for pairs of aligned sequences using the codeml program of the PAML (Phylogenetic Analysis by Maximum Likelihood) package[48], as implemented in the PAL2NAL program (http://www.bork.embl.de/pal2nal/)[49].

1. Soltis, D. E. *et al.* Polyploidy and angiosperm diversification. *Am. J. Bot.* **96**, 336–348 (2009).
2. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
3. Birchler, J. A. & Veitia, R. A. The gene balance hypothesis: from classical genetics to modern genomics. *Plant Cell* **19**, 395–402 (2007).
4. Birchler, J. A. & Veitia, R. A. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc. Natl. Acad. Sci. USA* **109**, 14746–14753 (2012).
5. Simillion, C., Vandepoele, K., Van Montagu, M. C., Zabeau, M. & Van de Peer, Y. The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **99**, 13627–13632 (2002).
6. Langham, R. J. *et al.* Genomic duplication, fractionation and the origin of regulatory novelty. *Genetics* **166**, 935–945 (2004).
7. Thomas, B. C., Pedersen, B. & Freeling, M. Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Proc. Natl. Acad. Sci. USA* **16**, 934–946 (2006).

8. Freeling, M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* **60**, 433–453 (2009).
9. Wilson, D. N. & Doudna Cate, J. H. The structure and function of the eukaryotic ribosome. *Cold Spring Harb. Perspect. Biol.* **4**, a011536 (2012).
10. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
11. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920 (2000).
12. Yusupov, M. M. *et al.* Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
13. Ben-Shem, A. *et al.* The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* **334**, 1524–1529 (2011).
14. Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A. & Ban, N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331**, 730–736 (2011).
15. Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. & Ban, N. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* **334**, 941–948 (2011).
16. Barakat, A. *et al.* The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant Physiol.* **127**, 398–415 (2001).
17. Carroll, A. J., Heazlewood, J. L., Ito, J. & Millar, A. H. Analysis of the Arabidopsis cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol. Cell. Proteomics* **7**, 347–369 (2008).
18. Chang, I. F., Szick-Miranda, K., Pan, S. & Bailey-Serres, J. Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol.* **137**, 848–862 (2005).
19. Giavalisco, P. *et al.* High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S ribosome. *Plant Mol. Biol.* **57**, 577–591 (2005).
20. Berná, G., Robles, P. & Micol, J. L. A mutational analysis of leaf morphogenesis in *Arabidopsis thaliana*. *Genetics* **152**, 729–742 (1999).
21. Serrano-Cartagena, J., Robles, P., Ponce, M. R. & Micol, J. L. Genetic analysis of leaf form mutants from the *Arabidopsis* Information Service collection. *Mol. Gen. Genet.* **261**, 725–739 (1999).
22. Van Lijsebettens, M. *et al.* An S18 ribosomal protein gene copy at the *Arabidopsis PFL* locus affects plant development by its specific expression in meristems. *EMBO J* **13**, 3378–3388 (1994).
23. Ito, T., Kim, G. T. & Shinozaki, K. Disruption of an *Arabidopsis* cytoplasmic ribosomal protein S13-homologous gene by transposon-mediated mutagenesis causes aberrant growth and development. *Plant J.* **22**, 257–264 (2000).
24. Pinon, V. *et al.* Three *PIGGYBACK* genes that specifically influence leaf patterning encode ribosomal proteins. *Development* **135**, 1315–1324 (2008).
25. Yao, Y., Ling, Q., Wang, H. & Huang, H. Ribosomal proteins promote leaf adaxial identity. *Development* **135**, 1325–1334 (2008).
26. Horiguchi, G. *et al.* Differential contributions of ribosomal protein genes to *Arabidopsis thaliana* leaf development. *Plant J.* **65**, 724–736 (2011).
27. Creff, A., Sormani, R. & Desnos, T. The two Arabidopsis *RPS6* genes, encoding for cytoplasmic ribosomal proteins S6, are functionally equivalent. *Plant Mol. Biol.* **73**, 533–546 (2010).
28. Muller, H. J. Further studies on the nature and causes of gene mutations. *Proceedings of the Sixth International Congress of Genetics, Ithaca, New York.* **1**, 213–255 (1932).
29. Piontkivska, H., Rooney, A. P. & Nei, M. Purifying selection and birth-and-death evolution in the histone H4 gene family. *Mol. Biol. Evol.* **19**, 689–697 (2002).
30. Rosado, A. *et al.* Auxin-mediated ribosomal biogenesis regulates vacuolar trafficking in *Arabidopsis*. *Plant Cell* **22**, 143–158 (2010).
31. Fujikura, U., Horiguchi, G., Ponce, M. R., Micol, J. L. & Tsukaya, H. Coordination of cell proliferation and cell expansion mediated by ribosome-related processes in the leaves of Arabidopsis thaliana. *Plant J.* **59**, 499–508 (2009).
32. Imai, A., Komura, M., Kawano, E., Kuwashiro, Y. & Takahashi, T. A semi-dominant mutation in the ribosomal protein L10 gene suppresses the dwarf phenotype of the acl5 mutant in Arabidopsis thaliana. *Plant J.* **56**, 881–890 (2008).
33. Weijers, D. *et al.* An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a *RIBOSOMAL PROTEIN S5* gene. *Development* **128**, 4289–4299 (2001).
34. Marygold, S. J. *et al.* The ribosomal protein genes and *Minute* loci of *Drosophila melanogaster*. *Genome Biol.* **8**, R216 (2007).
35. Szakonyi, D. & Byrne, M. E. Ribosomal protein L27a is required for growth and patterning in *Arabidopsis thaliana*. *Plant J.* **65**, 269–281 (2011).
36. Degenhardt, R. F. & Bonham-Smith, P. C. Arabidopsis ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are disparately required for normal development. *Plant Physiol.* **147**, 128–142 (2008).
37. Carroll, A. J. The Arabidopsis cytosolic tibosomal proteome: from form to function. *Front. Plant Sci.* **4**, 32 (2013).
38. Falcone Ferreyra, M. L., Casadevall, R., Luciani, M. D., Pezza, A. & Casati, P. New evidence for differential roles of l10 ribosomal proteins from Arabidopsis. *Plant Physiol.* **163**, 378–391 (2013).
39. Komili, S., Farny, N. G., Roth, F. P. & Silver, P. A. Functional specificity among ribosomal proteins regulates gene expression. *Cell* **131**, 557–571 (2007).
40. Mauro, V. P. & Edelman, G. M. The ribosome filter hypothesis. *Proc. Natl. Acad. Sci. USA* **99**, 12031–12036 (2002).
41. Horiguchi, G., Van Lijsebettens, M., Candela, H., Micol, J. L. & Tsukaya, H. Ribosomes and translation in plant developmental control. *Plant Sci.* **191–192**, 24–34 (2012).
42. Ponce, M. R., Quesada, V. & Micol, J. L. Rapid discrimination of sequences flanking and within T-DNA insertions in the *Arabidopsis* genome. *Plant J.* **14**, 497–501 (1998).
43. Ponce, M. R., Robles, P., Lozano, F. M., Brotons, M. A. & Micol, J. L. in *Arabidopsis protocols, 2nd edition* edited by Salinas, J. & Sánchez-Serrano, J. J. (Humana Press, 2006), pp. 105–113.
44. Jover-Gil, S. *et al.* The microRNA pathway genes *AGO1*, *HEN1* and *HYL1* participate in leaf proximal-distal, venation and stomatal patterning in Arabidopsis. *Plant Cell Physiol.* **53**, 1322–1333 (2012).
45. Curtis, M. D. & Grossniklaus, U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462–469 (2003).
46. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
47. Robles, P. *et al.* The *RON1/FRY1/SAL1* gene is required for leaf morphogenesis and venation patterning in Arabidopsis. *Plant Physiol.* **152**, 1357–1372 (2010).
48. Goldman, N. & Yang, Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736 (1994).
49. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–612 (2006).

Acknowledgments

We wish to thank J.M. Serrano, F.M. Lozano, T. Trujillo, L. Serna and J.M. Sánchez-Larrosa for their excellent technical assistance. Research in the laboratory of J.L.M. is supported by grants from the Ministerio de Economía y Competitividad of Spain [BFU2011-22825 and CSD2007-00057 (TRANSPLANTA)], the Generalitat Valenciana (PROMETEO/2009/112) and the European Commission [LSHG-CT-2006-037704 (AGRON-OMICS)]. H.C. is a recipient of a Marie Curie International Reintegration Grant (PIRG03-GA-2008-231073). R.C.S. holds a fellowship from the Ministerio de Economía y Competitividad of Spain (BES-2009-014106).

Author contributions

H.C. and J.L.M. conceived and designed the research. R.C.S. and H.C. performed the research. R.C.S, H.C. and J.L.M. wrote the article.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Casanova-Sáez, R., Candela, H. & Micol, J.L. Combined haploinsufficiency and purifying selection drive retention of *RPL36a* paralogs in Arabidopsis. *Sci. Rep.* **4**, 4122; DOI:10.1038/srep04122 (2014).