Supplementary Notes and Supplementary Figures

Contents

Overview

We initiated the Mendel Pea G2P (Genomics to Phenomics) project as an international collaboration between China and the UK. We dedicated to the exploration of *Pisum* population genomics and phenomics for trait and gene discovery, based on a core set of 697 representative accessions selected from the JI *Pisum* Germplasm Collection (Fig. 1). To associate the genetic diversity with the phenotypic diversity in this study, the starting point was to combine the whole-genome re-sequencing (20X) and an extensive phenotyping work across the UK and China, linking variations in genome with variations in traits. A high-quality genomic variation map (including SNPs, Indels, CNV and SV), particularly a linkage disequilibrium-based haplotype map was constructed. In parallel, we performed phenotyping survey of 83 agronomic traits (from seeds, leaves, pods, flowers, to plant architecture) coupled with large-scale haplotype-phenotype association genomics studies. We focused on 9 traits in a Mendel-centric manner, particularly on what is known about the seven contrasting pairs of traits and the underlying genetic difference that Gregor Mendel presented in his seminar paper with the segregation data (1866) (Extended Data Fig. 1). These 9 traits include Mendel's four characterized genes (*R, I, A* and *Le*) for which we focused on the identification of novel recessive alleles. The three remaining uncharacterized traits with five genetic loci involved (*Gp* for pod colour, *P* and *V* for pod shape, *Fa* and *Mfa* for flower position) were studied for the elucidation of the corresponding gene identities and causal alleles. For two additional traits, relevant to Mendel's comments, we identified the allelic variants of a *MYB* gene cluster corresponding to the *D* locus, underlying axil ring pigmentation, and defined allelic variations regulating organ size, as reflected in both pod width and grain weight.

As part of Mendel Pea G2P, we further expanded our effort to many other traits with contemporary importance for crop improvement through breeding. The key concept here was to identify candidate genes and novel allelic variation using whole genome trait association,

locating them using existing and newly developed mapping populations and using RNA-seq experiments, allelism tests, and mutagenesis coupled with genome sequence, for functional genomic validation of the identified candidates. The results of these studies provide insights for pea fundamental research, education in biology and genetics, and applied breeding practices.

Seed Shape: Round vs Wrinkled, *R* vs *r*

The first of Mendel's genes to be described at the molecular level was *R* corresponding to 'the difference in the shape of the ripe seed'. As described in the main text (Extended Data Fig. 2), we associated the haplotype map with the phenotypic changes in seed shape across the entire diversity panel, showing that there is only one *r* allele, corresponding to the *Ips-r* element, a non-autonomous type II transposon inserted in the last exon (#22) of the Starch Branching Enzyme I (*PsSBE*) gene, as described by Bhattacharyya et al. (1990)[1]. The GWAS peak corresponding to this allele is surprisingly broad (ca. 1Mb), and it is not identified by SNP allelic variation (Extended Data Fig. 2b); the SNP variants are shared with Hap5, which lacks the Ips-r insertion. The breadth of this peak is not diminished if the presence /absence of *Ips-r* is subject to GWAS analysis, suggesting that this is a structural feature of the allelic diversity around this gene differentiating *R* vs *r* genotypes. A previous study of 2792 *Pisum* accessions had determined that all *r* mutations corresponded to the same insertion allele of *SBEI* (Rayner et al. 2017)[2].

The sequence of the *Ips-r* element is very AT rich and bounded by the 8bp target site duplication, as previously described. These properties have meant that *Ips-r* has not previously been fully characterised, and here were able to present the complete sequence of the 1021 bp *Ips-r* element. This is larger than previously estimated (0.8kb)[1] presumably reflecting anomalous migration in agarose gels of DNA fragments carrying this sequence, but similar to that determined by Rayner et al. (1026 bp)[2], where the repetitive nature of the insertion element was noted to hinder amplification and determination of its sequence.

Note that there are five known genes which, when mutant, confer a wrinkled seed phenotype (*r, rb, rug3, rug4* and *rug5*)[3], of these only the *r* and *rb* alleles are known outside mutagenesis projects[4]; the *rb* mutation arose in the 1930s[5] and has limited representation in pea germplasm (12 *rb* lines identified among the 2792 accessions[2]). The variety Kebby (JI2110) is wrinkled, but the genetic basis of this is complex[2]. The phenotypic record of seed shape for each accession is summarized in Supplementary Table 12 and can be viewed in Supplementary Fig. 1.

Cotyledon Colour: Yellow vs Green, *I* vs *i*

The difference between yellow and green cotyledons (Supplementary Fig. 1) is determined by the allele of the gene *I* which corresponds to the gene encoding a stay green gene (*PsSGR*), Mg-dechelatase[6-8]. The mutants fail to degrade chlorophyll and so remains green. In the *Ii* heterozygote, the enzymatic function is present thus explaining the recessiveness of the green allele. In the present study, we found that variation in a single genomic region was responsible for most of the heritable green/yellow seed colour variation in pea; a strong genetic effect with a significant signal can explain this phenotypic difference (Extended Data Fig. 3). We did not include JI2775 in the set of accessions we studied, but did not find the i^{JI2775} allele in any of the 697 accessions in our panel, so we conclude that the 6bp insertion event of the i^{JI2775} allele[6], is rare and therefore cannot be considered as the major cause of green cotyledons. Neither did we find the group 4 *sgr* variant described previously[6].

We discovered an insertion of a 5,696 bp transposable element, a *Ty1*-copia element, together with its 5bp target site duplication; presumably this corresponds to the ca. 6.5 kb insertion of the group 3 allele described by Sato et al.[6]. The coding region of the *Ty1*-copia element is intact and the LTRs are identical, indicating an evolutionarily recent insertion.

We discovered a deletion of 408 nucleotides, 143 bp upstream of the ATG of Mg-dechelatase, together with a G to A silent substitution in exon 1. This deletion disrupts the region of the 5' UTR and therefore may interfere with the expression of this gene or the efficient production of an effective protein (Supplementary Fig. 2). Therefore, these novel *PsSGR* (*I/i*) alleles we discovered in this study are widely distributed in the *Pisum* diversity panel, and have strong explanatory power for most of the accessions with green cotyledons. The phenotypic record of green cotyledons for each accession is summarized in Supplementary Table 13 and can be viewed in Supplementary Fig. 1.

Flower Colour: Presence vs absence of anthocyanin pigmentation, *A* vs *a*

The difference in the colouration of the seed coat is the third character for which Mendel studied the pattern of inheritance; he noted that this was pleiotropic. In one of his three factor crosses, he referred to this as 'flower colour', which is the more usual description today. This corresponds to the gene *A* which encodes a *bHLH* transcription factor[9] required for the expression of chalcone synthase in epidermal tissues[10]; *a* mutants lack anthocyanin pigmentation throughout the plant. In the initial GWAS analysis based on the Caméor v1a genome reference (Extended Data Fig. 4b), we were surprised to find two GWAS peaks for white vs pigmented flowers. This turned out to correspond to a mis-assembly as revealed by mapping the Caméor v1a assembly to the JI0281xCaméor recombinant inbred population[11]. For this reason, when the ZW6 assembly became available we re-mapped our SNP matrix with respect to ZW6 (Extended Data Fig. 4a) and used this assembly in the Manhattan plots for other traits.

For the accessions that have the wild-type splice donor site (of Hap1), 359 are clearly pigmented with anthocyanins, although 5 are very pale but show axil ring pigmentation, corresponding to the *am* mutation. Of the Hap1 accessions, 9 appear white but the genetic basis for this is not clear; none of the accessions in this panel was known to carry the *a2* mutation.

A further 35 accessions carry the a^{JI1987} allele (Hap2). Four samples including JI0735 are white-flowered while having the wild-type splice donor site, but harbour a deletion of exons 1 and 2. The splice donor site allele could not be called in JI1775 (Hap4) which is white flowered and has a deletion of exons 1 to 6. The gene *A* and its *a* alleles are thus entirely consistent with this *bHLH* transcription factor being required for the production of anthocyanin pigmentation, the single exception being JI0233 as described in the main text. The JI0233 allele, has an insertion of an additional T nucleotide (which is validated by cDNA sequencing and read mapping of JI0233 in this study, Supplementary Fig. 3) between the mutant splice donor site and the splice donor site used in Caméor. This additional T nucleotide restores the exon 7 reading frame and incorporates an additional two amino acids from the wild-type intron sequence. This single nucleotide insertion is therefore an intragenic suppressor mutation. The phenotypic record of flower colours for each accession is summarized in Supplementary Table 14 and deposited to https://www.seedstor.ac.uk/ public database.

Internode length: Tall vs Dwarf, *Le* vs *le*

Several factors affect plant height in pea, of which variation in internode length under the control of gibberellic acid is a major contributor. Mendel described the segregation of plant height, but did not specifically refer to internode length. The GA-related genes, either involved in GA signalling or synthesis, are well-known to be responsible for genetic changes in plant height, such as in the semi- dwarfing wheat (*Della*)[12] and rice (*sd1*)[13], equivalent to *La* and *Cry* in pea[14]. Our GWAS analysis of internode length (stem length, or plant height) confirmed *Le* as the major determinant of the difference in stem (internode) length in pea. The gene identity of *Le* is a GA3-oxidase (GA 3β-hydroxylase); we found only one *le* allele (a G to A substitution in the coding region), consistent with previous reports. We identified one tall plant (JI0019) among the *le* genotypes, this corresponds to a 'cryptodwarf' (Reid et al. 1983)[15]. There are

several accessions carrying the *Le* alleles but showing short internode length. The haplotype clusters and the haplotype-phenotype analysis are summarized in Supplementary Table 15.

Pod Colour: Green vs Yellow, *Gp* vs *gp*

The yellow colour of *gpgp* pea pods, as Mendel noted, is one aspect of a wider syndrome affecting several parts of the plant, young leaf tips, the flowering inflorescence and the sepals as well as the pods, which is illustrated by the comparison between *GpGp* and *gpgp* near-isogenic lines in Fig. 3a and in Supplementary Fig. 4. The disturbance of the thylakoid membranes in leaflets of *gpgp* suggested the overall photosynthetic activity of these plants, and not just pod photosynthesis, might be impaired[16]. Pod dimensions, pod number, seed number and seed weight were all affected in *gp* plants, suggesting that plant productivity was lower in *gpgp* if compared with the *GpGp* plants (Supplementary Fig. 5). Not all peas with yellow pods are *gpgp* homozygotes, these can have a generally yellow or golden appearance such as chlorotic mutants or Goldkonig (JI0799), which has an overall yellowish appearance, a phenotype distinct from *gp*. Similar variant pea plants, such as cv. Doré (Vilmorin 1856, https://en.wikipedia.org/wiki/Louis_de_Vilmorin) were available at the time of Mendel's work, but as White[17] noted, crosses involving Goldkonig show that the genetic determinant of this trait is distinct from *gp* or *i*.

There are nineteen yellow-podded lines in total in the JI *Pisum* germplasm collection: JI0013, JI0015, JI0046, JI0056, JI0058, JI0073, JI0128, JI0132, JI0136, JI0816, JI1341, JI1343, JI2082, JI2273, JI2659, JI2716 and JI3541, and 12 representatives were selected in this study; crosses between these 19 lines show that all are allelic to *gp* (data not shown). Pellew and Sverdrup (1923)[18] reported a spontaneous origin of this phenotype in the offspring of cv. Duke of Albany (JI0313) and JI0013 is recorded as 'Pellew's *gp'*, suggesting that it might have been an independent *gp* allele, which if true would have been very useful.

Genetic mapping and GWAS analysis of Gp

As described in Online Method and summarized in Supplementary Fig. 6, the markers PsCam005046, PsCam056084, AX-183865165, AX-183571028 and AX-183571050 do not co-segregate with *Gp* in the cross JI0015xJI0399 (Supplementary Table 20-21). However, in the JI0816 x JI2822 F2 population (Supplementary Table 17), the marker AX-183879077 (ZW6 Chr3:324762848) does co-segregate with *Gp*. These data are consistent with *ChlG* co-segregating with *Gp* and primers designed to test *ChlG* confirmed that this is indeed the case in the JI0015 x JI0399 RIL population (JI0015xJI0399) (Supplementary Table 20-21). Thus, *Gp* is located as lying within the 4.5Mb interval defined by AX-183571050 and AX-183571028 (ZW6 Chr3:321020350-325580858). Two genes, corresponding to 3'exonucleases, have previously been shown to be associated with the *Gp* vs *gp* allelic difference[19]. These genes were located to a scaffold in the Caméor v1a assembly, which is part of the chromosome 3 assembly of ZW6 at a location close to our GWAS peak (Supplementary Fig. 6). Genomic analysis of this region in genome assemblies of JI0015 and JI2366 revealed a ca. 100kb deletion with the coordinates ZW6 Chr3:324692139-324798134 (Supplementary Fig. 7).

From our association genomics analysis, we can see that in the SNP-based GWAS, there is an unexpected shoulder that extends to most of the short arm of this chromosome, which requires further exploration. The SNP-based GWAS gives a 10Mb broad genomic interval (Chr3, 315-325Mb) with high significance, whereas our haplotype-based GWAS gives a single strong and narrow interval with only 1Mb (324-325Mb). The 10 Mb interval identified in the SNP-based GWAS analysis includes the 4.5 Mb region identified in the genetic analysis. Notably, the SV (structural variation) -based GWAS directly detected a ca. 100kb deletion within the minimum interval defined for *Gp* in the genetic analysis described above, as shown in Supplementary Fig. 7. This 100 kb deletion was found in all 19 *gp* accessions but in no *Gp* accessions (Supplementary Fig. 8). All genes within the deletion, *Psat03G0413800, Psat03G0413900, Psat03G0414000* plus a transcript within the LTRs of an *Ogre* element (at

chr3:324696949-324696718 in ZW6), which has not been annotated in the ZW6 genome, together with the truncated *TIR-NBS-LRR* (ZW6 *Psat03G0414100*, Caméor *Psat3g126360*) gene, are therefore candidates for *Gp*. The gene adjacent to the deletion is annotated as Chlorophyll Synthase (*ChlG*) and a disruption to the transcription of this gene is expected to result in a yellow phenotype.

A VIGS experiment silencing *ChlG* did indeed result in the yellowing of leaves, consistent with its annotation (Supplementary Fig. 11); however, this phenotype is not consistent with the green leaflets of *gpgp* mutants. While the VIGS experiment is not consistent with the *gp* phenotype, it would be consistent with a reduction in *ChlG* activity in leaves as an explanation for the inefficient thylakoid membrane assembly shown in Fig. 3b, **c** and also the yellowness of the pods. This suggested that a study of the transcription of *ChlG* in *Gp* and *gp* pods and leaves could be informative.

Aberrant transcripts were found in *gpgp* genotypes, including the truncated *TIR-NBS-LRR* gene (*Psat03G0414100*) to *ChlG* (*Psat03G0413700*) fused transcripts. A transcript variant with read-through of intron 1 of *ChlG* was confirmed by RT-PCR for *gp* RIL segregants, along with the expected wild-type *ChlG* transcript which is also evident in *Gp* RILs (Supplementary Fig. 9b). An premature stop codon is evident within the *ChlG* intron 1 read-through. The presence of the intron 1-containing transcript was shown to cosegregate with *gp*, indicating that intron 1 read-through is not an effect of a separate locus acting in trans, and is probably a consequence of the overall rearrangement of transcription in this region. The transcript isoforms detected are indicated in Fig. 3e-f. The relative transcript abundance from each of the exons of these two genes was assessed in leaves and pods of the *gp* accession JI2366 as shown in Fig. 3g. In addition, we detected two other alternative transcripts (T2, T3) derived from the truncated *TIR-NBS-LRR* gene in the *gp* lines (Supplementary Fig. 10); these two short transcripts are of low transcript abundance, and their biological function is unknown.

We tested whether *ChlG* is implicated in the *gp* phenotype. A TILLING mutant in *ChlG* was obtained from the pea TILLING resource at INRAe[20]. This mutant was W121* corresponding to a premature stop codon between the first and second predicted transmembrane helices of *ChlG*. The mutant was obtained as a heterozygote and no homozygous mutants were obtained upon selfing of this line**.** This is consistent with the W121* non-functional *ChlG* being an embryonic lethal. Note that there was no difference in the transmission rate of the mutant allele from the male or female gametophyte, indicating that *ChlG* is not required for gametophytic viability. A complementation/allelism test was performed by crossing the heterozygous TILLING mutant (as female) with JI0015 *gpgp* as male and the phenotype of the F1 was examined (Fig. 3h-i). The failure of $ChlG^{W121*}$ - *WT* to complement *gp* indicated that the TILLING mutant does not harbour a wild-type (Caméor) *Gp* allele; rather the TILLING allele gp^{lethal} is a lethal allele recessive to gp^{JI0015}. This establishes that *ChlG* is allelic to *Gp.*

Summary and discussion

These results suggest that a reduction in the abundance, and the presence of novel isoforms, of *ChlG* transcripts accounts for the yellow tissues in *gp* mutants.

The elevated transcript abundance in leaves vs pods correlates with the green colour of expanded leaflets. Several different aberrant transcripts were discovered from the *TIR-NBS LRR-CHLG* region, and they are of different types. One is an aberrant transcript that includes sequences from both genes, which has been validated by our PCR and sequence analysis. Another type of aberrant transcript includes intron read-through. This intron read-through aberrant transcript co-segregates with *gp*, demonstrating that it is determined by the *gp* allele. This is an important point, strongly suggesting that this type of transcript is not the consequence of another gene acting *in trans*.

Therefore, we propose that the aberrant transcripts were the major reason for lower *ChlG* transcript abundance in *gp* pods compared with leaves (Fig. 3g). We propose that aberrant transcripts carrying an in-frame early stop codon before the final exon of *ChlG* would be expected to enter into the nonsense-mediated decay pathway, resulting in an increase in the degradation rate of *bona fide ChlG* transcripts. Nonsense-mediated decay would therefore be one of the mechanistic connections between the aberrant transcripts and the lower transcript abundance. On the other hand, these unusual transcripts may have altered stability. We suspect that when pol II is reading from the *TIR-NBS-LRR* gene to *ChlG*, the polymerase will prevent access to the normal *ChlG* promoter simply because it is occupying that position during translation. In that sense, there will be interference with the normal initiation of translation at *ChlG*. It is also possible that the deletion in *gp* impacts an extended promoter or enhancer element upstream of the ChlG promoter; This model is quite hypothetical without evidence yet as the *ChlG* promoter or enhancer region has not been defined, there is ~ 3.7kb between the deletion boundary and *ChlG*. Further studies of the regulation of *ChlG* transcription from the wild type and *gp* mutant allele are needed to resolve these issues.

Here we provide evidence for aberrant transcripts of various types, and lower overall transcript abundance, suggesting that, the yellow pod phenotype of *gp* mutants is likely caused by some transcriptional interference mechanism.

Pod shape: Inflated vs Constricted, *P* and *V*

It is not clear whether Mendel studied the segregation of *P* or *V*. The gene *P* has been favoured because *V* is linked to *Le* (tall vs dwarf), but Mendel reported only one cross in which this phenotype segregated together with tall vs dwarf; this was not reported in his 1866 paper, but was part of a set of seeds, derived from a four-factor cross, which he sent to his collaborator Carl Nägeli[21]. The F2 genotypes of the list of F3 seed Mendel sent to Nägeli includes one

population segregating for tall vs dwarf and inflated vs constricted pods, and the genotypic frequencies are consistent with *P vs p* rather than *V vs v* segregation[22].

However, the similarity in the segregation ratio of tall vs dwarf and the presence vs absence of pod parchment in Mendel's crosses to determine the proportion of heterozygotes in the F2 could be consistent with these representing mostly the same plants in a multifactor cross[23] where *Le* and *V* were in repulsion phase. The issue therefore remains equivocal, and Mendel may have had some *pp* and some *vv* plants; we do not know of him having performed an allelism test.

Within our panel of accessions, both *pp* and *vv* genotypes were present as is revealed in the GWAS analysis, where significant associations were found at the expected position of *P* (chromosome 1 Linkage group VI) and *V* (chromosome 5 linkage group III) (Extended Data Fig. 6, 7).

P

The GWAS interval corresponding to *P* was initially defined as an approximately 11 Mb interval, chr1: 374889451-385871425 in the ZW6 genome[24]. This included the previously identified candidate *PsPS1* (*Psat01G0417600*, chr1:379493242-379495200)[25]. The gene *PsPS1* encoding a pectate lyase superfamily protein. We investigated this further in the F2 segregating population of the reciprocal cross JI0816 x JI2822 (N= 404, Supplementary Table 17), *P* was mapped between the markers AX-183563747 and AX-183563750 (chr1: 380049894-380967975) within the GWAS peak, narrowing the interval a little more than 10-fold: chr1 380Mb-381Mb, an only 1Mb interval (JI0816xJI2822). This interval excluded *PsPS1* as a candidate for *P*.

There are in total 8 genes in this interval, we screened each of the 8 genes in this interval for haplotype variation that could account for *p* and found that the gene *Psat01G0420500*, related to the tracheary element differentiation inhibition factor CLE41/44 or TDIF (tracheary

element differentiation inhibitory factor) had one haplotype that included an in-frame premature stop codon (235A>T, R79*). The R79* allele was designated haplotype 3, and all haplotype 3 accessions (JI0033, JI0066, JI0134, JI0340, JI0466, JI0467, JI0794, JI0806, JI1041, JI1071, JI1086, JI2100, JI2273, JI2366, JI2693 and JI3120) had parchmentless pods (Extended Data Fig. 6).

V

The second major GWAS signal for the presence or absence of pod parchment defines a broad ca. 50 Mb region at chr5:594837588-644213386 (Extended Data Fig. 7), corresponding to the expected position of *V*[26]. When the accessions carrying the R79* haplotype of *Psat01G0420500* were removed, this GWAS peak remained in the Manhattan plot and exhibited an elevated significance (Extended Data Fig. 7). All (JI1865, JI0136, JI1863, JI2683, JI1132, JI0077, JI3605, JI2262, JI0074, JI0122, JI0061, JI2177, JI2583, JI3120, JI0066, JI0134) of the remaining accessions scored as parchmentless carried haplotype 2 of the gene *Psat05G0805200*, encoding a β-fructofuranosidase. Three accessions (JI0066, JI0134 and JI3120) carrying the R79* allele of *Psat01G0420500* also had the haplotype 2 of *Psat05G0805200*, consistent with their being *ppvv* double mutants. JI0066 is a representative accession of the cultivar English Sabel WBH33, a designated *ppvv* double mutant[27], no independent information is available for JI0134 and JI3120. A few candidate genes were proposed to be involved in pod sclerenchyma development, like the fasciclins gene cluster (*Psat05G0769200*, *Psat05G0769000*, *Psat05G0768700*, *Psat05G0768600*) possibly regulated by TDIF / PXY-l representing *P*, the pea homologue of the interfascicular fibre mutant *Ifl*[28] (*Psat05G0771900*), and *WRKY*[29] (*Psat05G0839000*). The fasciclin domain is an extracellular domain involved in cell adhesion, and in Arabidopsis these have been shown to be involved in regulating cell wall elasticity and one class has been shown to be involved in the regulation of cell expansion. However, the haplotype clustering and haplotype-phenotype association

analyses for all of these genes excluded them as the candidates (Supplementary Fig. 12). There is a consistent pattern in our haplotype-phenotype association analysis for our target gene *Psat05G0805200*, a cell wall invertase (also known as β-fructofuranosidase), located at the most significant GWAS region and considered the most likely candidate for V. However, further analyses to narrow down the 2Mb target interval awaits the development of additional specific segregating populations, for fine genetic mapping a narrower interval corresponding to *V* within which the candidate gene must lie; and more functional validation work is still required.

Summary and discussion

We have identified the R79* allele of a CLE41/44 peptide which fully explains the *p* allele and therefore propose that *P* encodes this class of regulatory peptide. Removing the accessions carrying the R79* allele of a CLE41/44 from the GWAS analysis abolished the GWAS peak at this position and enhanced the significance of the peak corresponding to *V*, suggesting that *P* and *V* are two independent genetic loci controlling pod parchmentless. We propose that gene candidate *Psat05G0805200*, a cell wall invertase, is the most likely candidate for *V* locus that could explain the *v* allele in the haplotype-phenotype association study. The gene *V* was not identified unequivocally mostly because of the breadth of the corresponding GWAS peak and the absence of appropriate segregating populations or independently characterised mutations, more functional validation work is required.

Flower position: Axial vs Terminal (Fasciation), *Fa* and *Mfa*

Mendel scored ‘the position of the flowers’ on the pea stem (Supplementary Fig. 13); he used the terms axial vs terminal flowers although what he was discussing was fasciation, a condition that has been known and used in pea varieties at least since the 1500s[30]. The term ‘terminal flower’ is more appropriately used to refer to the *det* mutant where the usual apex (I_1) is replaced by the secondary (terminal) inflorescence (I_2) sometimes this terminal inflorescence

may have a terminal flower, but usually it terminates with the I_2 stub. Although several genes are known which, when mutant, confer a fasciated phenotype in pea, only *Fa* (chromosome 4 linkage group IV) and *Fas* (chromosome 5 linkage group III) are considered candidates for the gene which Mendel studied[31]. Within our panel of accessions, GWAS analysis revealed one strong peak on the Manhattan plots (Extended Data Fig. 8) corresponding to the expected peak of *Fa*; *Fas* was not obvious in these analyses, but there is a minor signal from the BSA analysis at Chr5 (Supplementary Fig. 14). These results indicate that *Fa* rather than *Fas* is the most common determinant of fasciation.

Fa

Previously, the gene *Psat4g010240* annotated as AP2 domain was considered as a candidate for Fa[32]. The authors report a mutant "which displays premature termination of the shoot meristem and differentiation of the stem cells". However, this *AP2* gene is located outside of the genetic interval and is thus excluded as a candidate. The *sym28* nutant is deficient in nodulation and also confers a fasciated phenotype, but is not the gene studied by Mendel[33], which is more likely to be *Fa* or *Fas*. Among our panel of accessions, the line JI2671 has been shown to be allelic to *fa*[34] and this is consistent with the identification of the strong GWAS signal at the end of chromosome 4 as being *Fa***.** Analysis of the segregation of *Fa/fa* in an F2 population (Caméor x JI0814) positioned *Fa* in the interval Chr4 0-20Mb of ZW6. Fasciation segregates in the F2 of the cross JI0816xJI2822 (Supplementary Tables 17-19). JI0816 is fasciated and JI2822 is wild type. The segregation of this phenotype is largely explained by a genetic locus on chromosome 4 linkage group IV which we assume to be *Fa* (Extended Data Fig. 9a). In the F2, almost all fasciated plants carry the JI0816 allele (*fafa*) from a region of chr4LGIV between the markers AX-183636277 and AX-183633456. This places *Fa* between chr4 18,180,969 and 19,506,907. Combining with another F2 population (a cross between Caméor and JI0814) and mapping of *Fa* to Chr4 18.14-19.94Mb places *Fa* between 18.18 and

19.51 Mb on chr4. There are 35 genes within the *Fa* region 18.14Mb – 19.51Mb interval on chromosome 5; one of these genes is *Psat04G0031700* at chr4: 18,548,435 – 18,551,779, annotated as *Clavata3 insensitive receptor kinase 3*, was found to have a 5bp deletion within the coding sequence of all (JI0208, JI0568, JI0780, JI0782, JI0795, JI0814, JI0817, JI0819, JI0820, JI0825, JI0853, JI0958, JI1713, JI2671) but two (JI0815/ JI1713) fasciated accessions, the latter two of which can be explained by *Mfa* as described below. The predicted truncated protein lacks the tyrosine and serine/threonine kinase domain (Extended Data Fig. 8h), and would thereby be unable to phosphorylate other proteins and the inability to transmit signals between *CLV* and *WUS*, which could result in abnormal development of the shoot apical meristem and a fasciated phenotype.

Mfa

There is no position in the genetic map of JI0816xJI2822 that co-segregates exactly with *Fa*; nevertheless a χ^2 test for its association with AX-183635452 (within the AX-183636277 and AX-183633456 interval and close to ZW6 *Psat04G0031700*) is highly significant $\chi^2 = 253$, $p \sim 0$. Initially we thought this imprecision in mapping was due to mis-scoring of the phenotype, but further analysis revealed that a location on chromosome 6 linkage group II was non-randomly associated with the discrepancy. This is not expected of mis-scores, which should be randomly distributed. We therefore proposed that this region carries a *Modifier of fasciation* (*mfa*) such that the double recessive *fa/fa mfa/mfa* is not fasciated. *Mfa* could not be located precisely because relatively few individuals are informative. The genetic map of the cross JI0816xJI2822 shows that adjacent to *Mfa* there is a region of identity by descent in JI0816 and JI02822 which positions one border of the *Mfa* interval at about chr6: 244,689,457 and the mapping positions the other at about chr6: 253,701,016. Interestingly, a minor peak at this position is also seen in the Manhattan plots of the GWAS analysis and G' plots from the other F2 population (Fig. 2b, Supplementary Fig. 14). The

Extended Data Fig. 9a shows the genotyping which corresponds to these data. The fasciated lines JI0816 (WBH 1185) and JI0817 (Wisconsin-711) are genetic stocks from Stig Blixt formerly at the Nordic Gene Bank, while JI0814 is from GA Marx in the USA. JI0815 is not fasciated, but is the type line for *undulatifolius*[27].

This model proposes that JI0816 is *fa/fa Mfa/Mfa* and that JI2822 is *Fa/Fa mfa/mfa*, but in the progeny of this cross *fa/fa mfa/mfa* genotypes are predicted to be wild type. In the Extended Data Fig. 9b,c there is clearly a deficiency of *fa/fa mfa/mfa* genotypes scored as fasciated, and most *fa/fa mfa/mfa* genotypes scored as wild type. There are, however, several F2 individuals which do not agree with this model, which can be explained by mis-scoring or by other unknown genetic interactions. Note that we do not know why *Mfa* is required for fasciation to be seen, it could be that *Mfa* is directly involved with *Fa* or it could be that it causes the plant to mature before the onset of fasciation. *PsWUS*[35] is *Psat02G0115700* at chr2:74831848-74832478, which is outside this interval, which is firmly ruled out as a candidate for *Mfa*.

Summary and discussion

The *fa* allele is strongly associated with a 5bp deletion in the gene *Psat04G0031700*, annotated as *Clavata3 insensitive receptor kinase 3*. However, the penetrance of this mutant phenotype appears to be modified by an unknown gene on chromosome 6 linkage group II which we have designated *Mfa*. This model suggest that *Mfa* is involved in meristem maintenance although the *mfamfa* mutant, in an otherwise wild-type background, is phenotypically wild type. We found no strong evidence for the gene *Fas*, a minor signal from the BSA analysis reportedly on chromosome 5 linkage group III. The line JI2771 carrying *fas* was not included in these studies.

The orthologous gene of *Psat04G0031700* in Arabidopsis is *AtCIK3*, which plays an important role in *CLV3*-mediated regulation of apical meristem homeostasis. The *atcik1 2 3 4*

quadruple mutant also has a fasciated phenotype[36]. Through transient expression of the pea orthologue in leaves of *Nicotiana benthamiana* and subcellular localization, it was found that the *PsCIK* is located on the cell membrane, as expected for receptor kinases (Supplementary Fig. 15). Based on differential gene expression patterns among *PsCIK1*, *PsWUS*, and *PsCLV3* in various tissues in pea, we can propose that PsCIK1 may interact with CLV1, CRN, PPK2 and ERL2 that are located on the membrane, thereby regulating the expression pattern of WUS. CLAVATA (*CLV*)–WUSCHEL (*WUS*) plays an important role in the proliferation and differentiation of stem cells maintained by a negative feedback loop[37-39]. Therefore, a possible *mfa-fa-CLV3-WUS* regulatory network can be proposed, the biochemical genetics of which requires further investigation.

Axil ring pigmentation: *D* vs. *d*

Mendel noted that axil ring pigmentation is absent in white flowered varieties. A dominant allele of the gene *D* is required for the presence of axil ring pigmentation[27]. In *Pisum abyssinicum* flowers are pigmented, but axil ring pigmentation is absent; this taxon is a naturally occurring *d* mutant [11,27]. There however are several alleles of *D: D^{w} D^{co} D^{tet} D^{ma}* and *d* with distinct phenotypes, as illustrated in Supplementary Fig. 16.

Our GWAS study uncovered a single major peak in Manhattan plots, at chr2:91435801-110725232 in the ZW6 assembly (Fig. 4e-f). This complements a genetic analysis of the segregation of *D* vs *d* in a recombinant inbred population of a *P. abyssinicum* x *P. sativum* cross which placed *D* in an interval corresponding to Chr2:97476978 (AX-183581090) and Chr2:113478796 (AX-1835643150), overlapping with the GWAS signal (Fig. 4). This region includes several genes encoding *MYB* transcriptionfactors[11].

In addition to the *D* allelic variation in the diversity panel (Supplementary Table 30), we obtained three Fast Neutron induced mutants (FN1091/4, FN1218/6, and FN2073/5), in the JI2822 background[40], which exhibited the *d* phenotype (Supplementary Fig. 20). Allelism tests

were performed among these and with respect to JI0073 (a multiply marked line) and JI2202 (*P. abyssinicum*). The phenotypes of the axil rings of these accessions and F1 plants are shown in Supplementary Fig. 20; the results indicate that FN1091/4 and FN1218/6 were allelic to *d*, but FN2073/5 was not.

These studies have defined *d* as corresponding to either *PsMYB104* or *PsMYB106.* The line FN2073/5 remains to be characterised, but this line has a deletion nearby that includes *Psat02G0252700*, a *MYB* gene at chr2:221575409-221576389 in the ZW6 assembly (This gene was not described by Yang et al. 2022b)[24]. Święcicki (1990)[41] obtained a genetic map position for *And* with a recombination fraction corresponding to 15.5 ± 4.5 cM with respect to *D*. The positions of *Psat02G0252700* and *PsMYB106* on the genetic map of JI0281xCaméor correspond to a distance of ca. 22 cM, so it is possible that a mutation in *Psat02G0252700* corresponds to this d-like phenotype and an allele of *And*.

A VIGS experiment was undertaken with the gene *Psat02G0138300* annotated as *MYB*-type HTH DNA-binding domain profile in the Caméor genome assembly. This is within the GWAS interval, but is excluded as a candidate for *d* by the genetic mapping and Fast Neutron mutant analysis. Nevertheless, this experiment provided an interesting further insight into ring pattern formation as shown in Supplementary Fig. 18 and 19.

Summary and discussion

The results presented here suggest that *PsMYB104* and/or *PsMYB106* are required for axil ring pigmentation as is another unknown gene defined by FN2073/5. Furthermore, the *MYB* related gene *Psat02G0138300* has a role in patterning this pigmentation. It has been noted[27] that not only are there multiple alleles of *D*, but that there are also many spontaneous conversions from one allelic form to another. These spontaneous interconversions may represent recombination or conversion events within this close association of functionally

related genes. The VIGS experiment showed us that the *MYB* gene regulating the outer ring of pigment is different from that regulating the presence or absence of this pigment.

Organ Size: Pod Width and Seed Weight, *PsOs1*

A correlation analysis within our panel shows a significant and previously unrecognised positive relationship between pod width and seed weight (r = 0.65). In addition, we analysed an F_2 population from the cross between JI0074 (wide pod and large seed, male) and JI1995 (narrow pod and small seed, female), which identified a QTL on chromosome 2 that aligned with the GWAS peak (Supplementary Figs. 22-23) and in the approximate location of a previously identified seed size QTL[42]. Fine mapping of the pod width QTL narrowed down the interval to a ca. 1Mb region containing 11 genes in the ZW6 reference genome (Fig. 4). Transcriptome data in pods and qRT-PCR analysis revealed that *Psat02G0011300* was the most significantly differentially expressed gene in a comparison between the parental lines (Supplementary Fig. 25).

Transverse sections of pods (12 days post-flowering) in JI0074 and JI1995 showed that cell size in JI0074 was notably larger than that in JI1995 (Supplementary Fig. 26). Flow cytometry revealed a relative increase in the number of cells in the G2/M phase in JI0074 compared to JI1995, indicating cell cycle suppression (Supplementary Fig. 26). Flow cytometry also revealed a decreased proportion of cells in the G2 phase in *PsOs1*-silenced lines. These studies suggest that *PsOs1* inhibits the cell cycle at G2/M and induces endoreduplication, a process potentially crucial for rapid fruit growth by increasing cell size[43,44]. The number of higher ploidy nuclei and endoreduplication indices were significantly reduced in VIGS-*PsOs1* silenced lines as compared to the control VIGS-Con lines (Supplementary Fig. 27).

We further cloned the coding sequence of *PsOs1* from the broad-pod variety JI0074 and introduced it into Arabidopsis (Col-0) under the CaMV 35S promoter, resulting in stable transgenic plants. The siliques of the overexpression lines were not only significantly wider

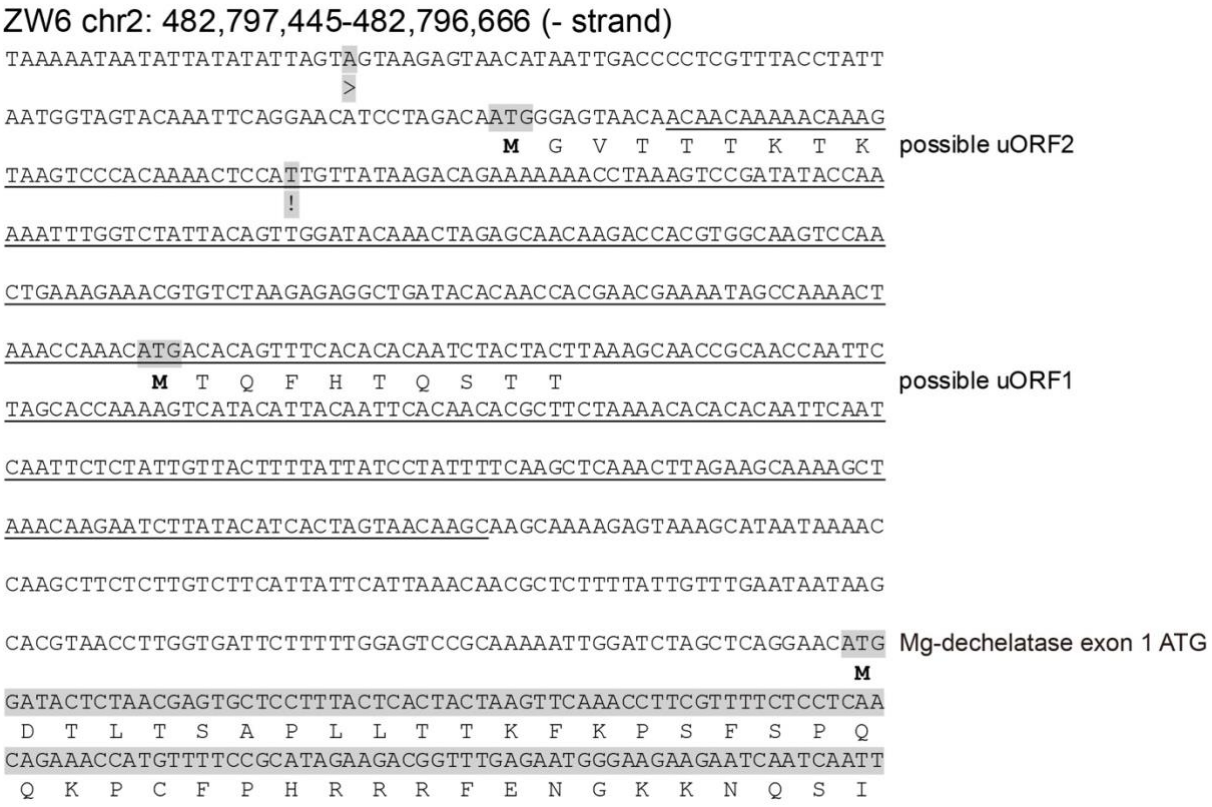
than those of the wild type but also featured mature seeds that were larger and heavier, albeit with a shorter silique length compared to the wild type (Supplementary Fig. 29). Flow cytometry data demonstrated an increased proportion of higher ploidy cells and a significant elevation in the endoreduplication index in the overexpressing lines**.** The relationship between pod and seed size is moderate for pea (r = 0.65). The transgenic Arabidopsis lines suggest that both traits are altered together under the influence of *Psat02G0011300*.

In light of significant differences in the expression of *PsOs1* between the parents, we performed Sanger sequencing of the *PsOs1* coding region from JI0074 and JI1995, and cloned a 3000 bp promoter region upstream of the coding sequence from both parents. We did not find any functional variation within the coding sequences but identified multiple differences in the promoter region (Supplementary Fig. 30), notably at position chr2:5261351 and chr2:5261587, the most significant SNPs detected in both pod width and seed weight GWAS analysis ($7.701225E^{-19}$ and $1.44E^{-19}$ respectively) (Fig. 4). Evaluation of *PsOs1* promoter activity using β-glucuronidase (GUS) contructs and expression analysis revealed that the JI0074 promoter exhibited higher activity than the corresponding JI1995 promoter. Expression analysis by qRT-PCR indicated predominant expression of *PsOs1* in pea pods (Supplementary Figs. 29-30), further confirmed by GUS staining in transgenic Arabidopsis plants expressing *PsOs1* promoter-driven GUS in roots, trichomes, sepals, and siliques, but not in seeds (Supplementary Fig. 28).

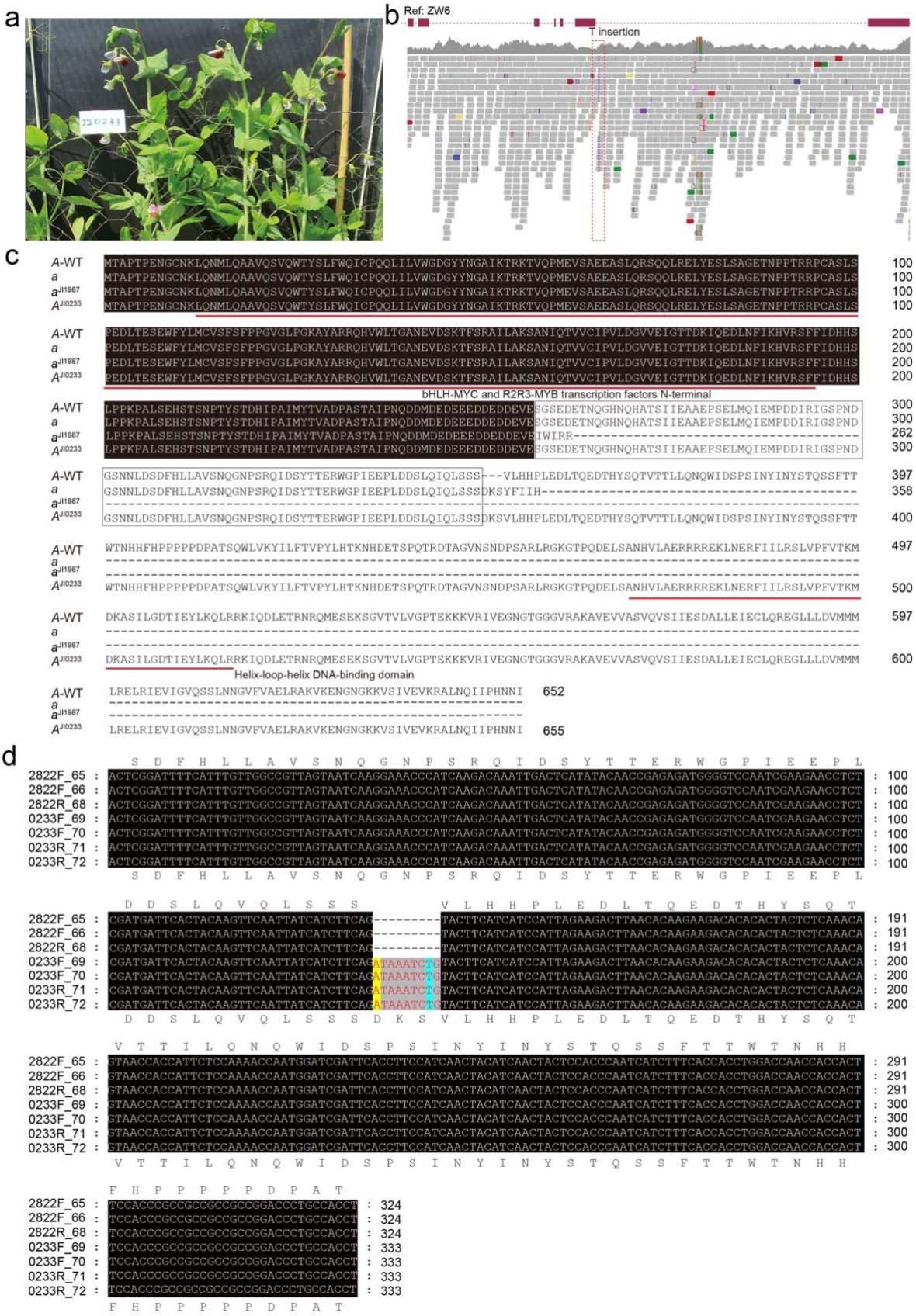
We further classified our pea diversity panel into two haplotypes based on pod width: Narrow (N) including JI1995, and Wide (W) including JI0074. Pod width and seed weight were significantly greater in W than N (Supplementary Fig. 30). RT-PCR results showed that *PsOs1* transcripts were more abundant in wide-pod samples compared to N. Variation in the promotor region of *PsOs1* was consistent with its role in driving these phenotypic changes.

Supplementary Figures

Supplementary Fig. 1 | Photo gallery to illustrate the phenotypic survey of seed characters within the diversity panel. The phenotypic records for the rest of Mendel's pea traits were summarized in Supplementary Table 6. The overall phenotypic dataset was summarized in Supplementary Table 27.

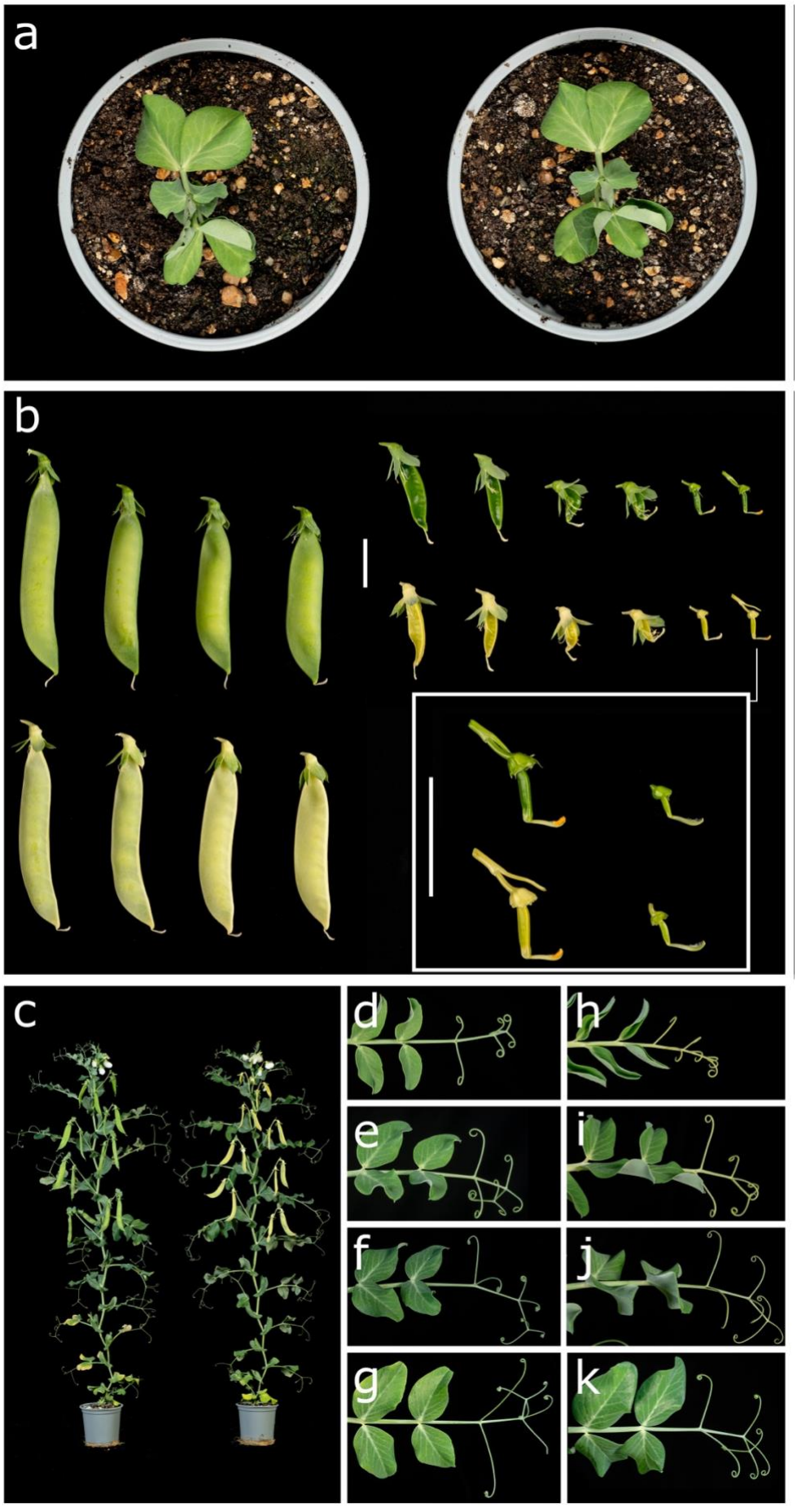


Supplementary Fig. 2 | Illustration of the 408bp deletion upstream of the Mg-dechetatase sequence, as revealed in the *i-2* allele in this study. The region immediately upstream of *SGR*, *Psat02G0529500* of ZW6 is presented. The 408 bp deleted in haplotype (Hap) 2 (*i-2* allele as presented in Extended Data Fig. 3c) is underlined. Two possible regulatory upstream open reading frames are indicated (these are in the same frame as exon 1). The proposed start of the 5' UTR in the Caméor annotation[45] is at the A marked '>' immediately after a potential TATA box. In the same assembly the most 5' nucleotide in a transcript is at the T marked '!'.



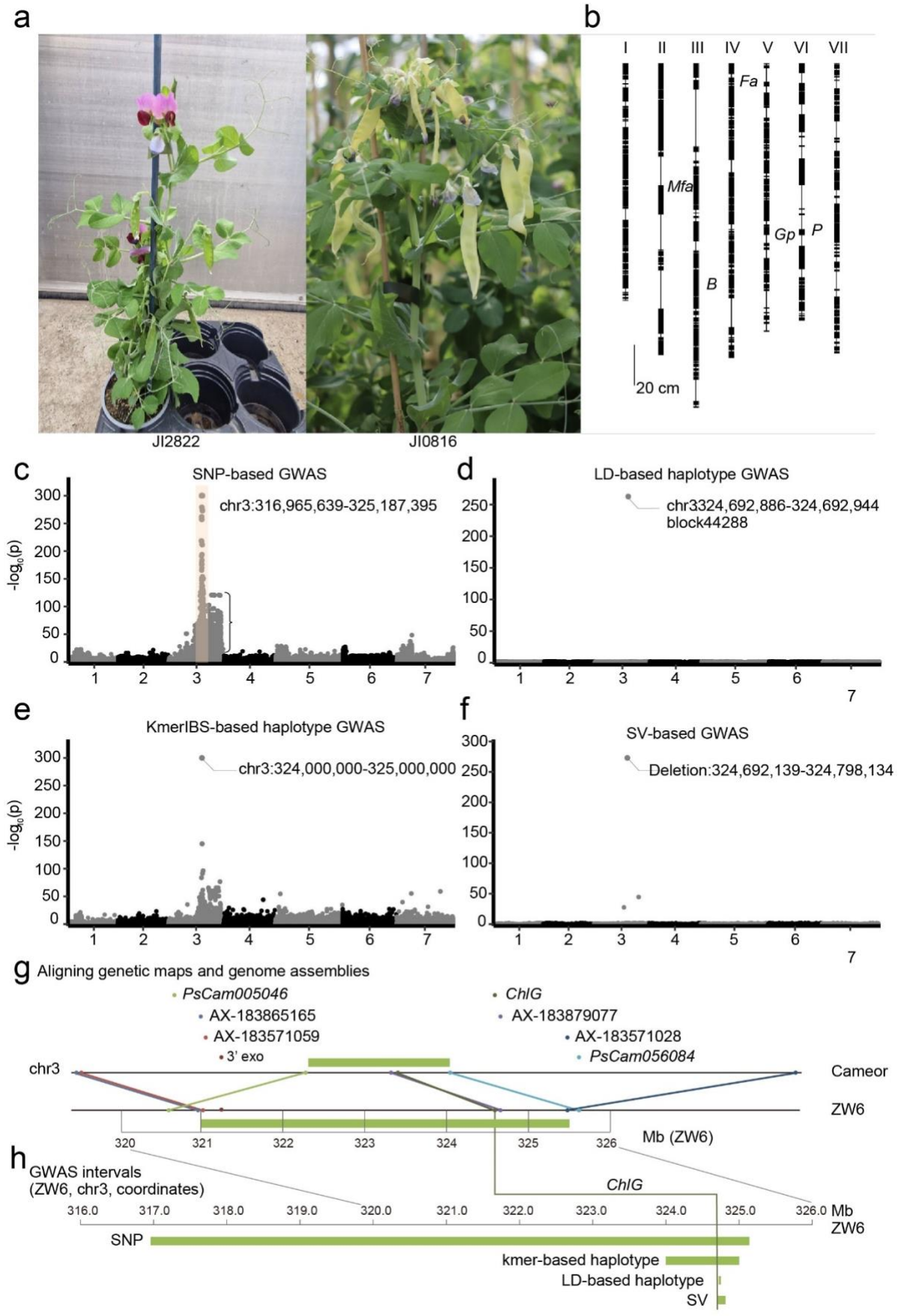
Supplementary Fig. 3 | The intragenic suppressor mutation, *A*^{JI0233}. The reading frame in the JI0233 transcript (Hap5) is restored by a single nucleotide insertion. (**a**) The phenotype of

JI0233 mature plant, clearly showing the purple flower. (**b**) Read mapping shows the insertion of a nucleotide (T in the coding strand) in JI0233 with respect to the Caméor a allele. (**c**) Alignment of the predicted amino acid sequences of the wild type, A^{JI2822} (A-WT) and the *a* alleles of Caméor (*a*) and JI1987 (a^{JI1987}) compared to the *a*-derived *A* allele of JI0233 (A^{JI2033}). (Extended Data Fig. 4e); A^{JI0233} has an insertion of an additional nucleotide (T in the coding strand) in what would be in the sixth intron of the *bHLH* transcription factor corresponding to the wild type *A* allele. With respect to the progenitor *a* allele this results in restoration of the wild type reading frame in JI0233, adding nine nucleotides to the transcript as described in the main text. (**d**). cDNA sequence of the transcript from JI2822 and JI0233 showing the most common G to A SNP in splice donor site (yellow-shaded) and the T insertion (blue-shaded) in JI0233, which restores the wild type reading frame. This confirmed that the A^{JI0233} allele uses the same splice donor site as in the *a* allele of Caméor.


a
b
c
d
h
e
i
f
j
g
k

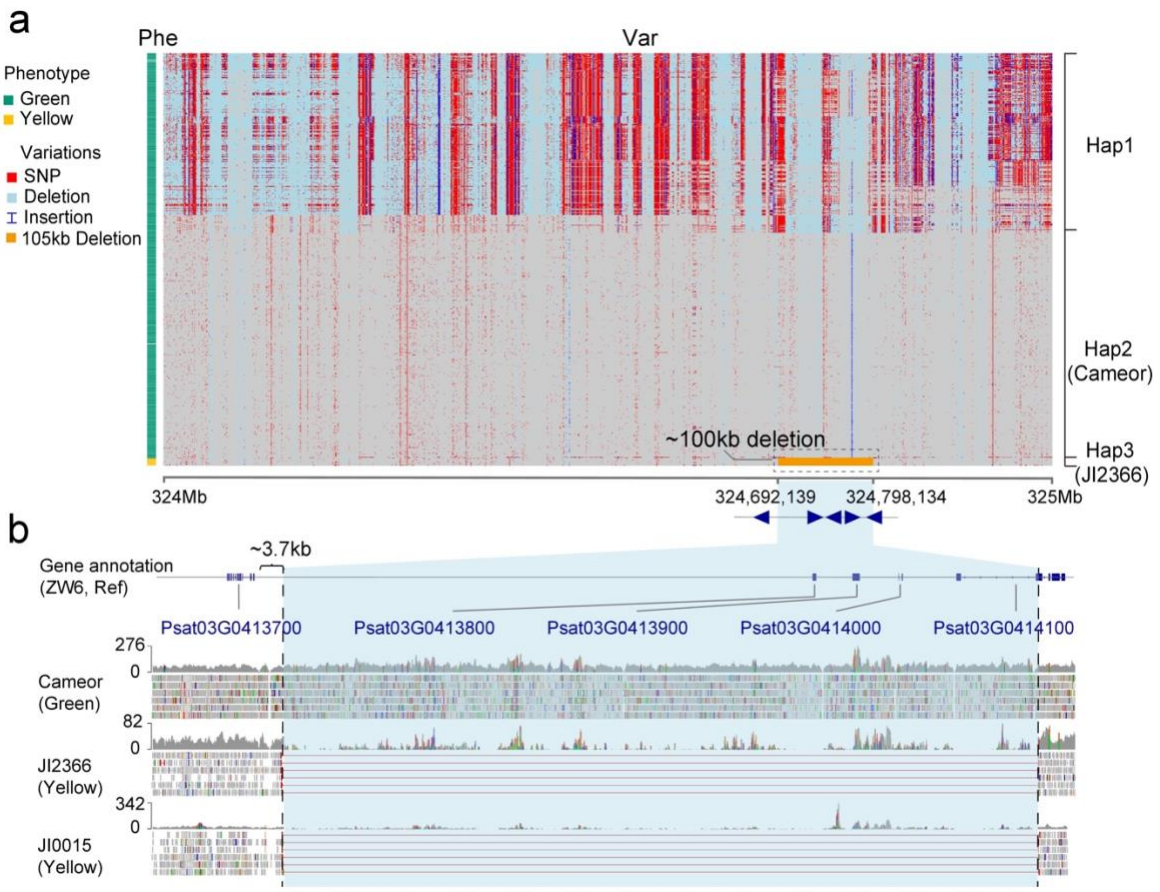
Supplementary Fig. 4 | Plant phenotype of the *gp* mutant. **(a)** Phenotype of seedlings 15 days after sowing (DAS), BC6 *Gp* (left) and BC6 *gp* (right). At the early vegetative stage of development, *gp* mutant and wild type seedlings are indistinguishable. **(b)** Phenotype of BC6 *Gp* pods (upper) and BC6 *gp* pods (lower) showing pod colour at successive stages of development. Stages are 13 days post anthesis (DPA), 11 DPA, 9 DPA, 7 DPA, 5 DPA, 4DPA, 3DPA, 2DPA, 1DPA, 0DPA from left to right. Inset shows a magnified view of *Gp* carpels (upper) and BC6 *gp* carpels (lower), post-fertilisation (left) and pre-fertilisation (right). At all stages shown, *gp* mutant carpels and pods are distinguishable from wild type by their yellow colour. **(c)** Phenotype of mature plants 72 DAS, BC6 *Gp* (left) and BC6 *gp* (right). Stage 72 DAS is the middle frame of a 7-frame time-lapse series (7frames_BC6Gp3gp21.gif), showing the development of these plants at 29 DAS, 44 DAS, 58 DAS, 72 DAS, 86 DAS, 105 DAS and 177 DAS. **(d – k)** Uppermost four leaves (each one including the rachis, two pairs of leaflets, three pairs of tendrils, the terminal tendril, and excluding stipules) on BC6 *Gp* **(d – g)**, and BC6 *gp* **(h – k)** plants shown in (C). **(d, h)** Uppermost upright leaves; **(e, i)** fully unfolded leaves at node below uppermost; **(f, j)** fully expanded leaves located two nodes below uppermost; **(g, k)** fully expanded leaves located three nodes below uppermost. Leaves at the two uppermost nodes (and leaf primordia, not shown) of the BC6 *gp* mutant have yellow tendrils, rachis and leaflet midribs and petiolules compared to more mature, expanded leaves at lower nodes, which are fully green. In contrast, leaves at all nodes of BC6 *Gp* wild type plants are green. Pots in panels (a) and (c) are 9 cm in diameter, scale bars in panel (b) are 2 cm.

Supplementary Fig. 5 | *Gp* vs *gp* phenotypes. Pod length, pod width, ratio of pod length to width, the number of pods, total seed number and seed weight of *GpGp* and *gpgp* genotypes of BC6 S2 plants from the Cameor x JI0015 cross (n=3).

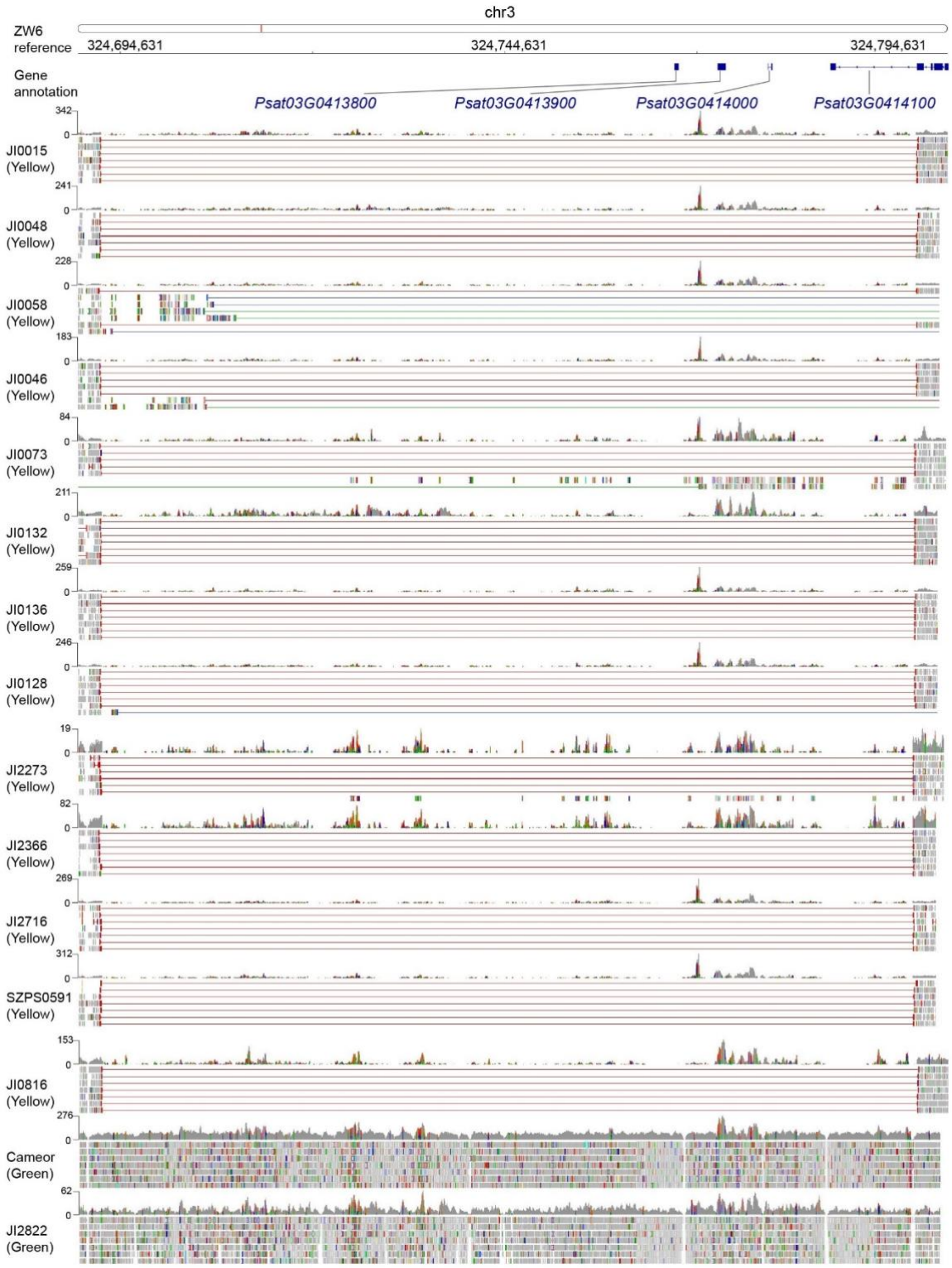


Supplementary Fig. 6 | Identification of the position of *Gp*. (**a**) The phenotype of JI2822

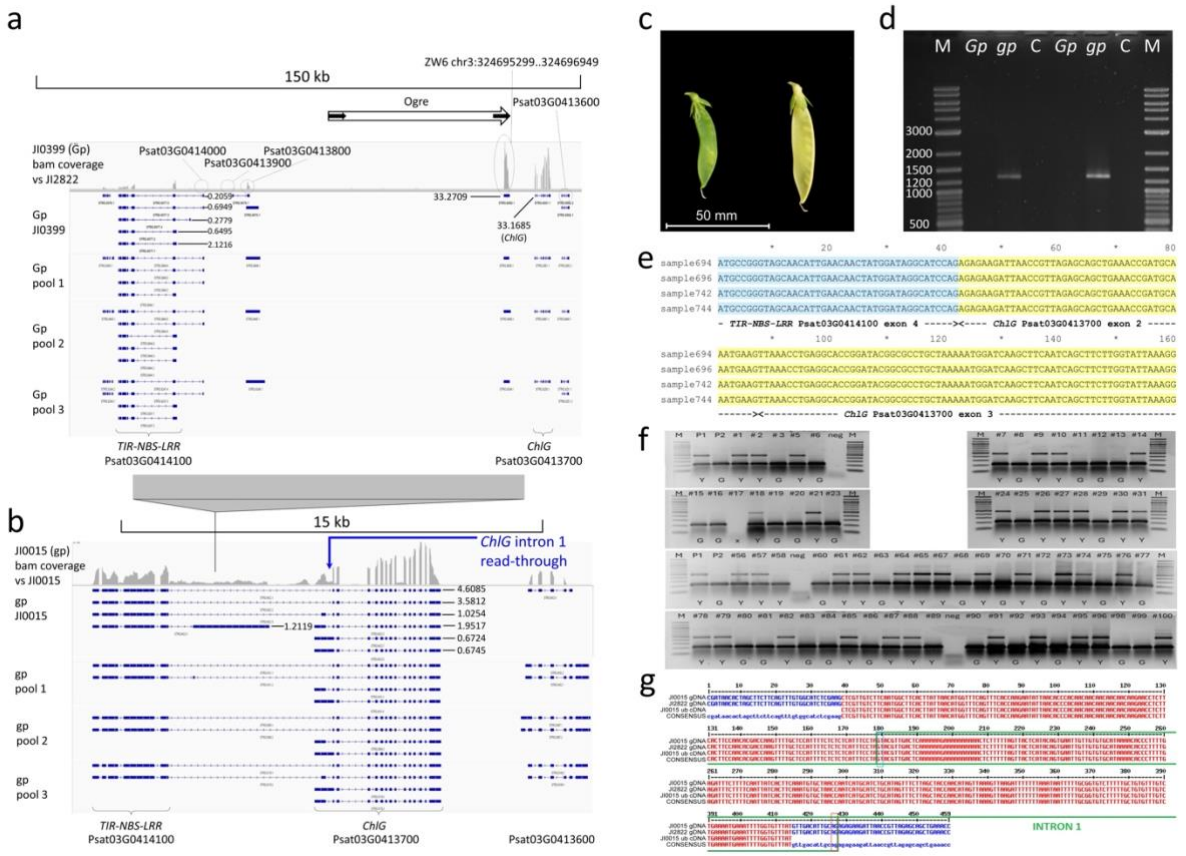
(*Gp*) and JI0816 (*gp*) showing different pod colour. **(b)** The genetic markers including those associated with *Mfa, B, Fa, Gp* and *P* locus are shown on the chromosomes. **(c)** SNP-based GWAS. **(d)** LD-based haplotype GWAS. (**e**) Kmer IBS-based haplotype GWAS. **(f)** SV-based GWAS. A significant single signal, corresponding to a deletion variant located at Chr3: 324,692,139-324,798,134 has been identified in these different approaches. (**g**) Based on JI0015 x JI0399 RILs and the two RIL cross (Supplementary Tables 20-21), genetic markers flanking *Gp* are positioned with respect to each other on a 10 Mb region of the assemblies of two pea genomes: Caméor v1a (Kreplak et al. 2019)[45], and ZW6 (Yang et al. 2022b)[24]. In addition to the genetic markers discussed above, the positions of chlorophyl synthase *ChlG* and the gene encoding a 3' exonuclease identified by Shirasawa et al. (2021)[19] that lies within this region are indicated. (**h**) the genomic intervals identified by different approaches of GWAS; the location where *CHLG* is situated is highlighted.



Supplementary Fig. 7 | (**a**) Haplotype clustering and haplotype-phenotype association analysis near the large deletion region across the entire diversity panel in this study. In total, three major haplotypes were identified (Hap1, Hap2, Hap3) with respect to the ZW6 genome reference, in which, the green podded-line Caméor (*Gp*) belongs to Hap2 and the yellow-podded mutant lines JI2366/JI0015 (*gp*) belong to Hap3. All the *gp* lines carrying the ca. 106kb deletion (marked in orange, coordinate: 324,692,139 – 324,798,134 at Chr3) are clustered into Hap3, affecting 5 protein-coding genes marked by blue triangles. The distance between the deletion boundary and the *ChlG* coding sequence is indicated (~3.7 kb). Note this map is in the opposite orientation to that shown in Fig. 3. (**b**) Top row the exon structure and position of the 5 genes within or adjacent to the deletion. Lower three rows: IGV display of short-read mapping against the genome reference (ZW6): Caméor (green pod, *Gp*), JI2366 (yellow pod, *gp*), and JI0015 (yellow pod, *gp*).



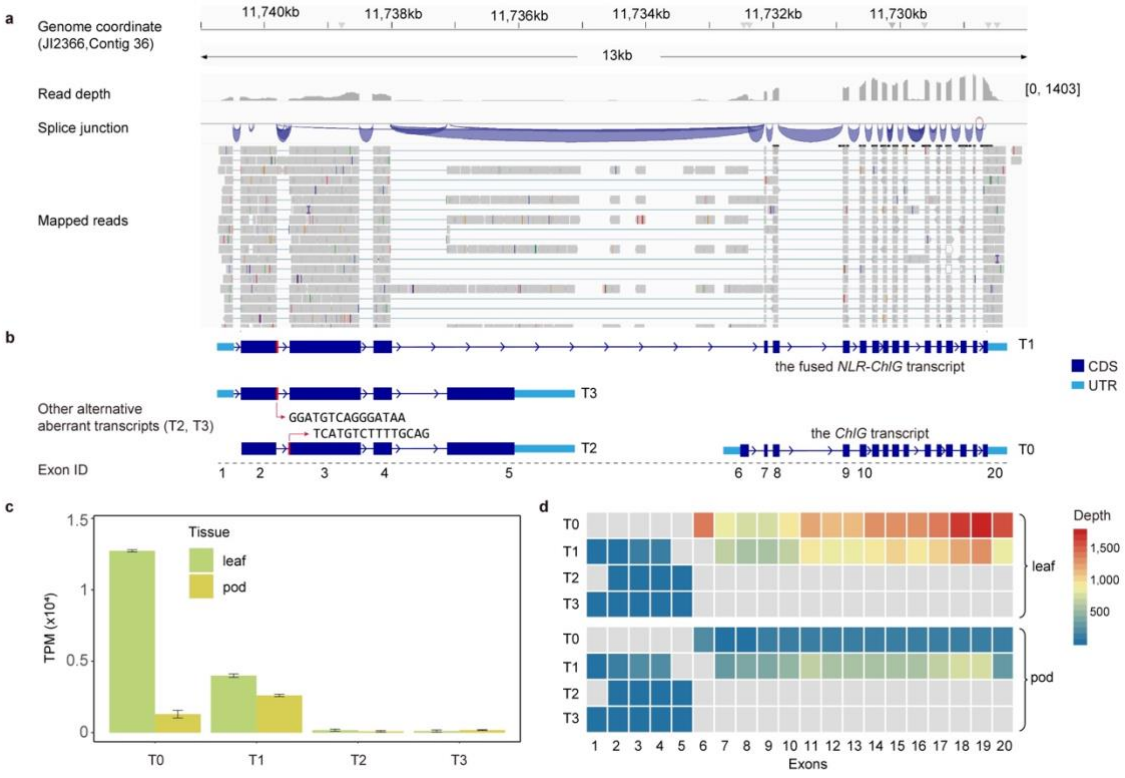
Supplementary Fig. 8 | IGV display of short-read mapping against the ZW6 reference genome, showing the ca. 105kb deletion event, found in all 12 yellow podded lines (*gpgp*), but in no *GpGp* lines within our panel. (*Gp*, Caméor and JI2822 for comparison at the bottom).



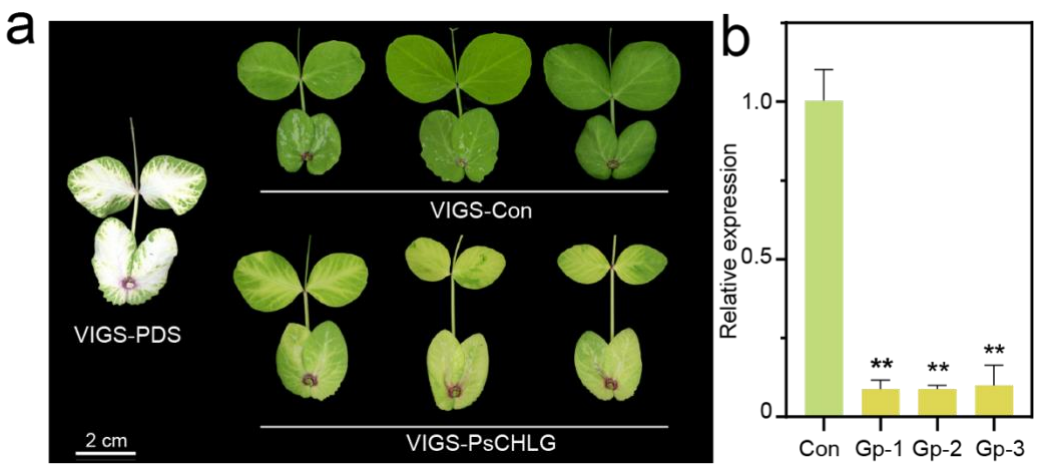
Supplementary Fig. 9 | Aberrant transcripts in *gp* pods. **(a)** Predicted transcripts from wild type (JI0399) green pods and three RIL green pod pools, mapped vs JI2822 (*Gp* allele), in a 150 kb region encompassing the *TIR-NBS-LRR* gene (Psat03G0414100) and *ChlG* (*Psat03G0413700*). An *Ogre* retrotransposon with LTRs represented by black arrows, and an expressed captured sequence in its 3' LTR (corresponding to ZW6 chr3:324695299-324696949), lies upstream of *ChlG* as indicated. Data are displayed in IGV with RNA-Seq read coverage and StringTie-predicted transcripts shown. Circled transcripts are annotated according to the ZW6 genome (Yang et al 2022). TPM values of five predicted *TIR-NBS-LRR* (*Psat03G0414100*) transcripts, one predicted *ChlG* transcript and one predicted *Ogre LTR2* transcript are marked with black numbers. **(b)** Predicted transcripts from *gp* mutant (JI0015) yellow pods and three RIL yellow pod pools, mapped vs JI0015 (*gp* allele), in a 15 kb region spanning the truncated *TIR-NBS-LRR* gene (*Psat03G0414100*) and intact *ChlG* (*Psat03G0413700*). The large deletion, which removes the 3' terminal 11 bp of

Psat03G0414100 exon 5 and ca. 100 kb downstream, is represented by a grey polygon. The *Ogre* retrotransposon lying upstream of *ChlG* has been deleted, as have three annotated genes (*Psat03G041400*, *Psat03G0413900* and *Psat03G0413800*). Data are displayed in IGV with RNA-Seq read coverage and StringTie-predicted transcripts shown. TPM values of seven predicted *gp* allele-specific transcripts are marked with black numbers. Three predicted aberrant *TIR-NBS-LRR-ChlG* transcripts, splicing *TIR-NBS-LRR* exon 4 to either *ChlG* exon 2 (2 isoforms), or *ChlG* exon 1 (1 isoform), are shown. A predicted transcriptional fusion of *TIR-NBS-LRR* exon 5 with deletion-adjacent non-coding sequence to create a novel predicted translational fusion product (1 isoform) is shown, however this is not predicted in the three yellow pod pools. Predicted aberrant intron 1 read-through *ChlG* transcripts (2 isoforms) with extended 5' UTRs are shown. One conventionally spliced *ChlG* isoform with extended 5' UTR is shown. **(c)** Examples of BC6 *Gp* (left) and *gp* (right) pods 5 days after anthesis. Scale bar is 50 mm. **(d)** RT-PCR amplification of predicted aberrant transcripts from BC6 *Gp* and *gp* pods 5 days post anthesis, in duplicate, with water negative controls (C). Primers TIRexon3F and ChlGexon4R were located in *Psat03G0414100 TIR-NBS-LRR* exon 3 and Psat03G0413700 *ChlG* exon 4 (Supplemental Table 32). Sizes of markers (M) shown on left. **(e)** Sanger sequenced PCR products amplified in (D) showing *TIR-NBS-LRR* exon 4 spliced to *ChlG* exon 2. **(f)** RT-PCR amplification of pod cDNA in the JI0015 (*gp*) x JI0399 (*Gp*) RIL population. A 210 bp PCR product amplified by primers ChlGF18 and ChlGR7 (Supplemental Table 32), corresponding to conventionally-spliced ChlG exons 1 and 2, is seen in all RILs. A 459 bp PCR product, corresponding to an aberrant intron 1 read-through transcript (upper band), occurs only in yellow podded lines of the population. RILs (#1, #2, #3 etc.) with green pods (G) and yellow pods (Y) are indicated. The last three RILs (#98, #99 and #100) were run on a separate gel. M, 100 bp ladder lanes; P1, JI0015 parent 1 of the RIL population; P2, JI0399 parent 2 of the RIL population; x, reaction failed. **(g)** Sanger sequenced isolated 459 bp upper

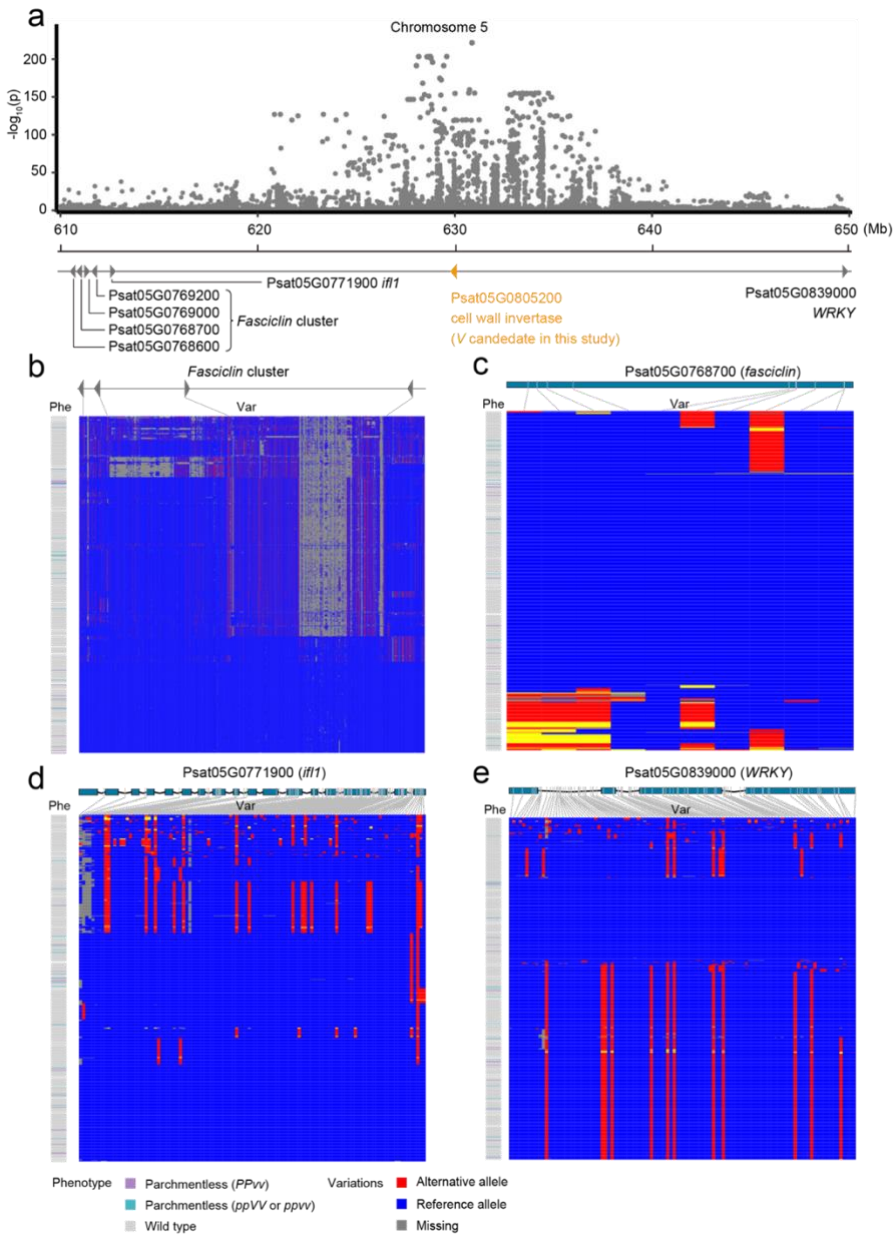
band (ub) from JI0015 yellow pod cDNA in (F) aligned with JI0015 and JI2822 genomic DNA (gDNA). *ChlG* Intron 1, boxed in green, is unspliced in the isolated JI0015 yellow pod cDNA upper band.



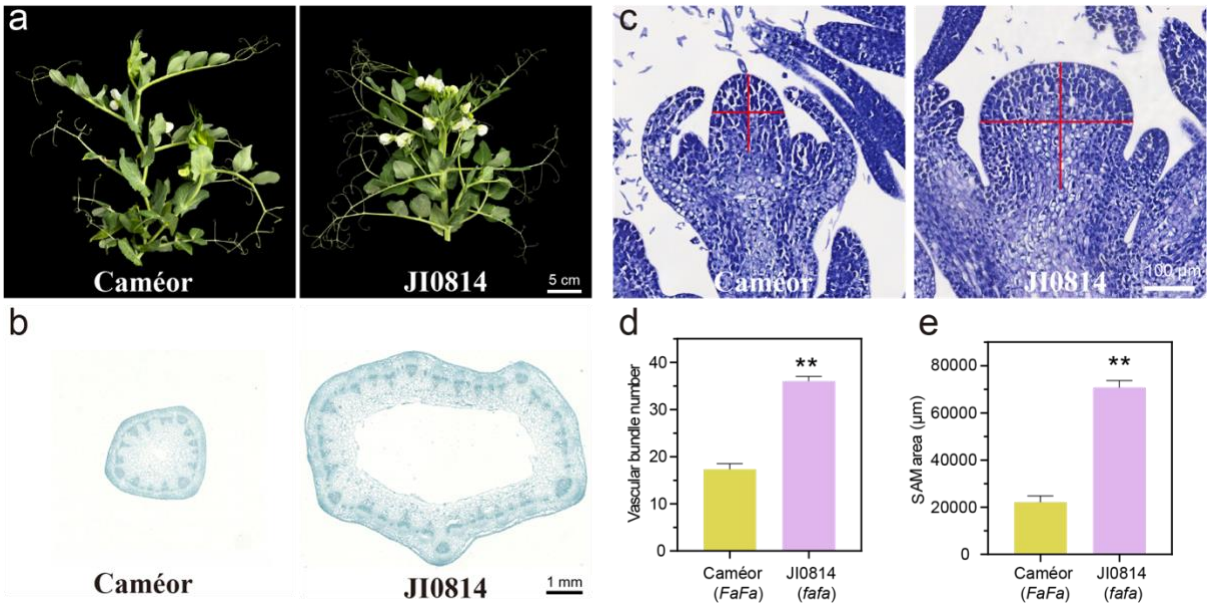
Supplementary Fig. 10 | The gene model and transcript annotation for JI2366 (*gp*) in the *NLR-ChlG* region near the ~106kb deletion, and relative gene expression of different transcripts in leaf and pod. (**a**) displays an Integrated Genomics Viewer (IGV) snapshot showing RNA reads aligned to the JI2366 (*gp*) genome, including genome coordinates, read depth, splice junctions, and mapped reads. (**b**) illustrates the structures of 4 different transcripts along with exon numbering. (**c**) quantifies the expression levels of these transcripts in *gp* leaf and pod tissues. **(d**) provides a detailed visualization of read depth by RNA-seq short-read mapping across each exon for both *gp* tissue types.



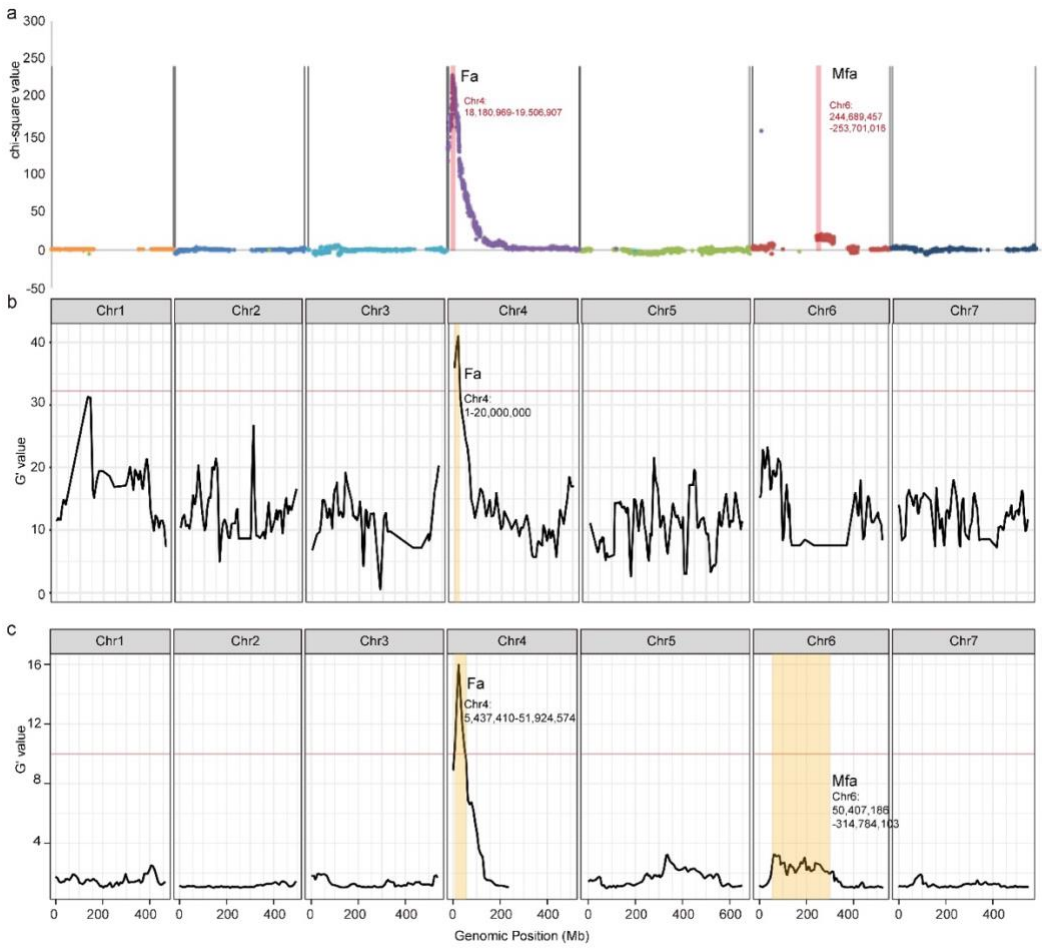
Supplementary Fig. 11 | Verification of the gene for *Gp* locus. (**a**) Silencing of *PsCHLG* by VIGS, with VIGS-con and VIGS-PDS serving as negative and positive control, respectively, after 20 days of infection of Agrobacterium with different VIGS vectors (scale bar = 2 cm). (**b**) Measurement of expression levels of *PsCHLG* in the leaves of VIGS assay after 20 days of infection of Agrobacterium with VIGS vector (** indicate a significant level at $P < 0.01$ in a Students' t-test).



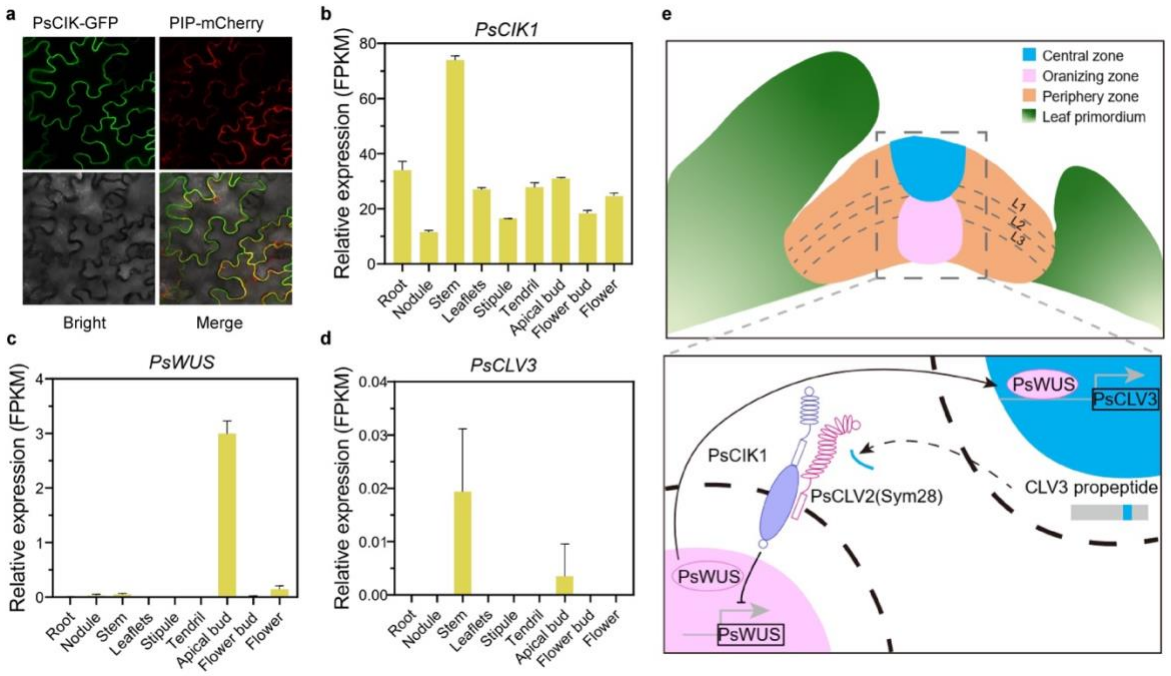
Supplementary Fig. 12 | Haplotype analysis of proposed alternative candidate genes for *V*. (**a**) GWAS local Manhattan plot showing a strong signal corresponding to *V*. The genes located within this region are shown in the bottom and *Psat05G0805200* encoding a cell wall invertase (colored in yellow) was proposed to be the candidate gene of *V*. Haplotype clustering and haplotype-phenotype association analysis for: (**b**) the entire fasciclin gene cluster. (**c**) one of the representative fasciclin gene, *Psat05G0768700*. (**d**) *Psat05G0771900* (*ifl1*) and (**e**) *Psat05G083900* (*WRKY*). The phenotypes and variations are colour coded, which are shown in the bottom of the figure.



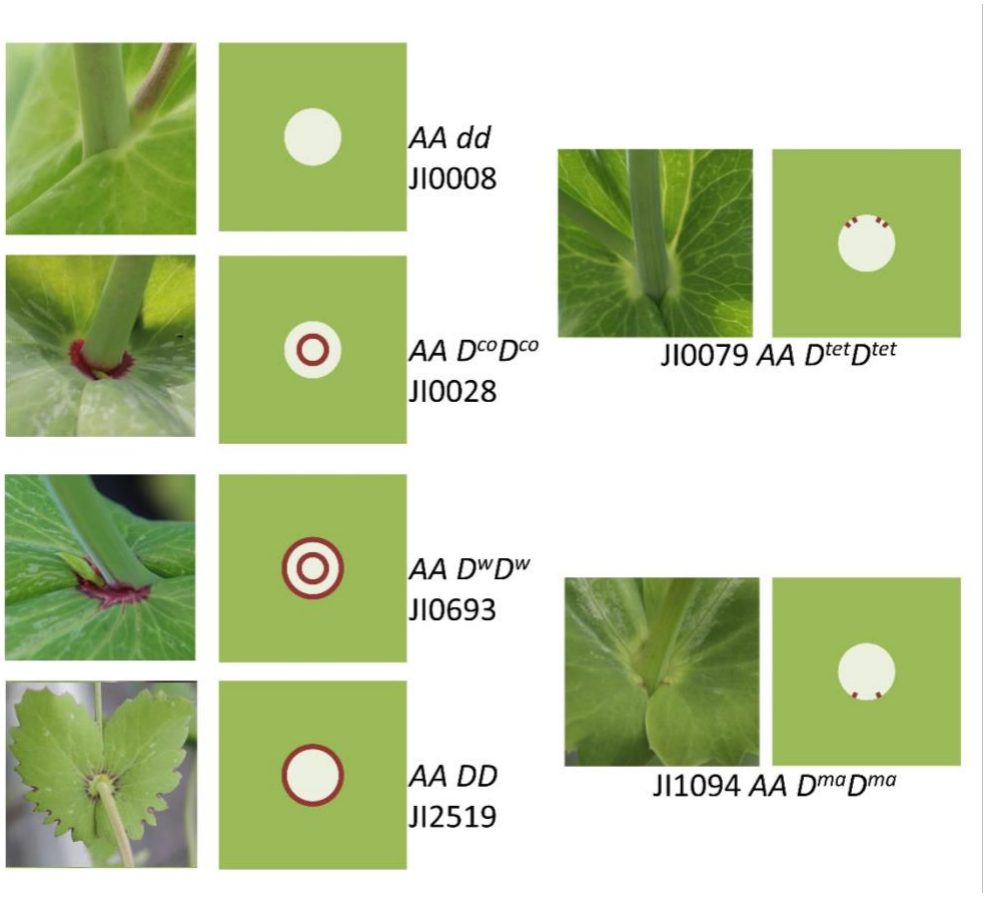
Supplementary Fig. 13 | Comparative analysis of field phenotype and microscopic observation in one of the wild types (Caméor, *Fa*) and one fasciated types (JI0814, *fa*). (a) Flowering stage phenotype of Caméor (*FaFa*, left) and JI0814 (*fafa*, right) showing axial (wt) vs terminal flowers (*fa*). Also see study[46]. **(b)** Transverse section of 40-day-old stems of Caméor (wild type - left) and JI0814 (*fa* - right), stained with safranin and fast green; **(c)** Longitudinal section of apical meristem of paraffin-embedded 14-day-old stems, stained with toluidine blue, Caméor (left) and JI0814 (right). The red crosses mark the region of shoot apical meristem (SAM) in Caméor and JI0814; **(d)** Number of vascular bundles in the transverse section of 40-day-old stems of Caméor and JI0814; **(e)** Cross sectional area of apical meristem of 14-day-old stems between Caméor and JI0814. ** represent a significant level at $P < 0.01$ using a Student's t-test in **(d)** and **(e)**.



Supplementary Fig. 14 | Chi-square test for each marker in the F2 population derived from the cross of JI2822 (*FaFa*) and JI0816 (*fafa*) and Bulk Segregant Analysis (BSA) analyses between two F2 mapping populations to locate genetic loci underlying fasciation. *Fa*, narrowing the target region into about 1Mb at the end of Chr4. (**a**)**.** Chi-square value for the 3x2 contingency test of G, H, or Y vs Dominant vs Recessive phenotype, where G is the homozygous JI2822 allele, H is the heterozygote and Y is homozygous for the JI0816 allele (Supplementary Table 17). Where there were fewer than expected of the recessive, JI0816 homozygote combination then we made the sign of the chi-square negative. The candidate intervals of *Fa* and *Mfa* are marked in red in the figure. (**b**). BSA analysis of the F2 population, a cross between Caméor (*FaFa*) x 0814 (*fafa*); (**c**). BSA analysis of the F2 population derived from the cross of JI2822 (*FaFa*) and JI0816 (*fafa*). The loci for *Fa* (Chr4) and *Mfa* (Chr6) are marked by yellow bars and the regions were shown on the plot.



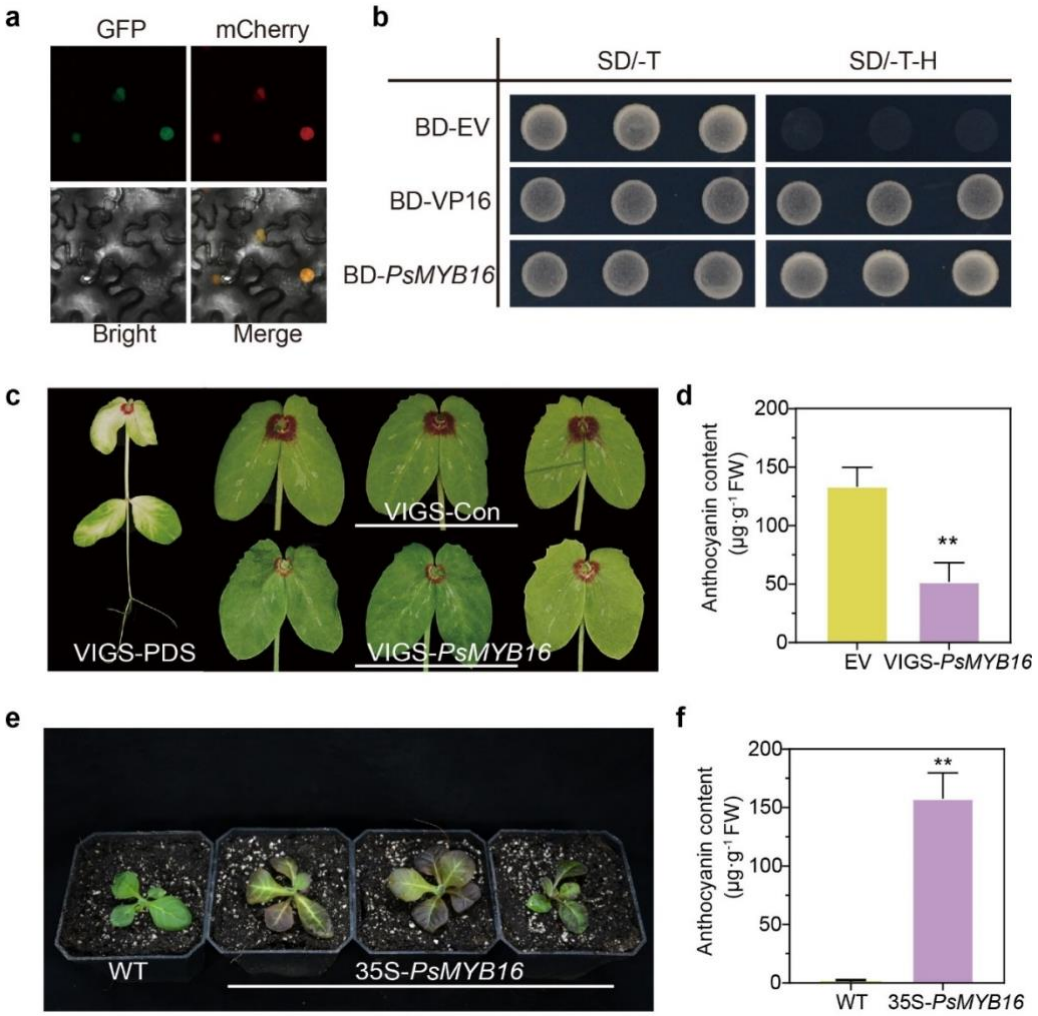
Supplementary Fig. 15 | Differential gene expression and the proposed regulatory module for *PsCIK1*. (a) subcellular location of the *PsCIK1* showing that it co-locates with the cell membrane in *Nicotiana benthamiana*. **(b, c, d)** measurement of gene expression level (RNA read counts) of *PsCIK1*, *PsWUS*, and *PsCLV3* in various tissues of the Caméor plant, showing that *PsCIK1* is highly expressed in stem development. **(e)** a proposed genetic regulatory module for the roles played by *PsCIK1*, *PsWUS*, and *PsCLV3*, also see study[47]. Organization of the shoot apical meristem (SAM) and the *CLV3*-*WUS* feedback regulation loop. The SAM is composed of the upper central zone (blue), the organizing center (pink) underneath, and the flanking peripheral zones (orange). Stem cells in the central zone occupy three clonally distinct cell layers (L1, L2 and L3). *CLV3* is expressed specifically in the central zone. Mature *CLV3* is then secreted into the apoplast and binds to the plasma membrane-localized receptor complexes *CIK*-*CLV2* to repress *WUS* transcription in the organizing center. *WUS* then moves to the *CLV3* expression domain and directly binds the *CLV3* promoter to activate its transcription.



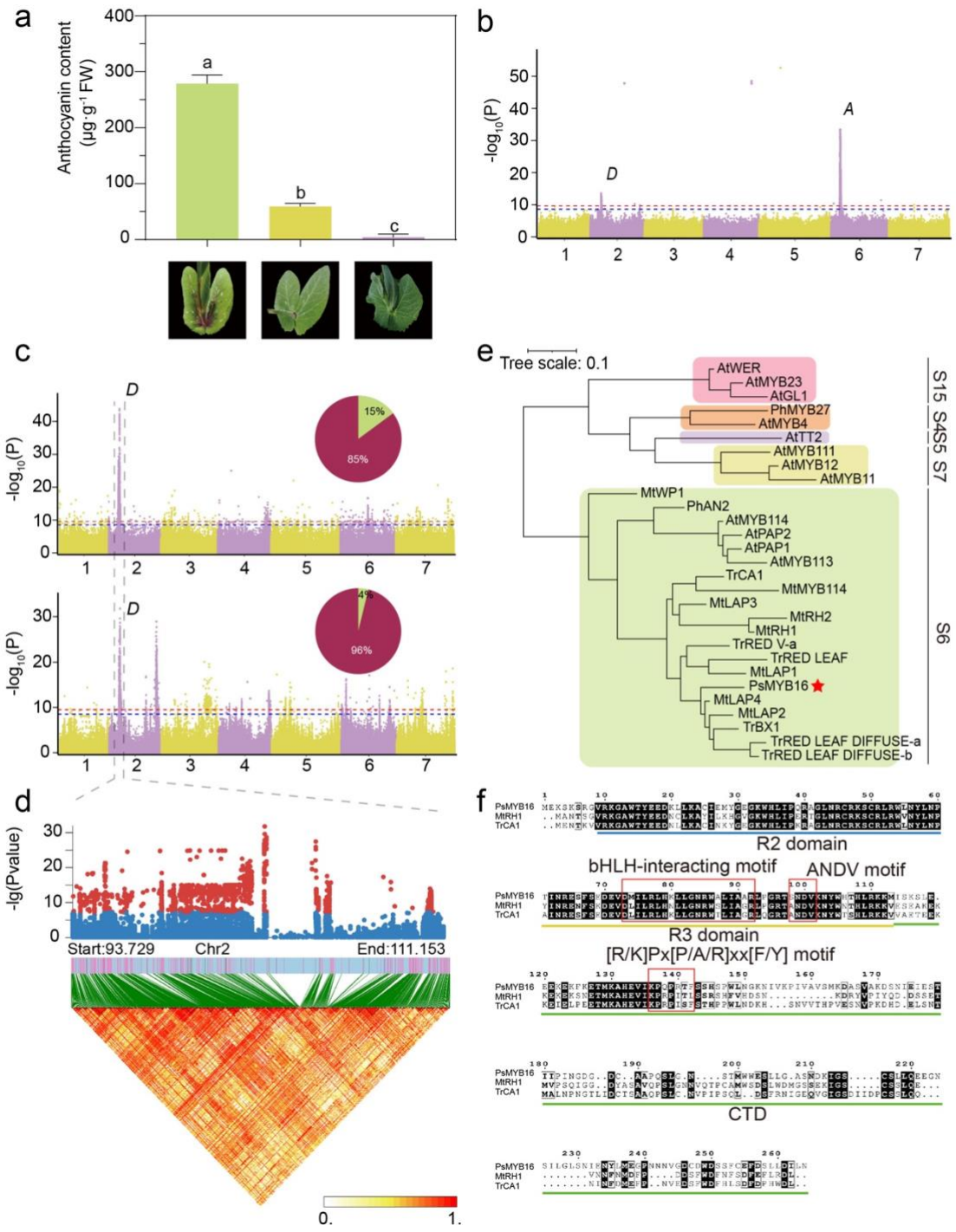
Supplementary Fig. 16 | The different *D* allele axil ring pigmentation phenotypes. The alleles D^{w} D^{co} D^{tet} D^{ma} and *d* are illustrated as displayed in the JI accessions as indicated. Note that in JI0008, the type line for *d*, the stem has anthocyanin pigmentation indicating the presence of the dominant *A* allele. JI2519 is an example of an accession where the inner axil ring is absent. These accessions often have additional anthocyanin pigmentation at the leaf margin, along the veins or as spots on the stipule lamina. These extensive anthocyanin pigmentation patterns on the lamina and on the veins have been assigned separate gene symbols *An* and *And* by Święcicki (1990)[41]. *And* is linked to, but distinct from *D*, while the genetics of *An* have yet to be described.

Gene Name			Start NT (ZW6)
Yang et al. 2022b	Cameor	ZW6	
PsMYB13	Psat2g031320	Psat02G0106500-T1	67,627,687
N/A	Psat2g040240	Psat02G0132200-T1	94,088,126
PsMYB14	Psat2g041080	Psat02G0134900-T1	95,841,702
PsMYB15	Psat2g042040	Psat02G0136200-T1	99,623,426
PsMYB104	Psat0s340g0040	Psat02G0136700-T1	101,612,599
PsMYB106	Psat0s855g0040	Psat02G0136800-T1	102,131,473
PsMYB116	Psat0ss3069g0040	Psat02G0136900-T1	102,726,853
PsMYB16	Psat2g043680	Psat02G0138300-T1	105,183,397
PsMYB17	Psat2g047760	Psat02G0151400-T1	114,012,753
PsMYB18	Psat2g058360	Psat02G0186800-T1	145,255,097

Supplementary Fig. 17 | A summary of the *MYB* gene members (cluster) within the *D* locus. All except *PsMYB13* are encoded on the + strand of the ZW6 assembly (Yang et al. 2022a)[24].

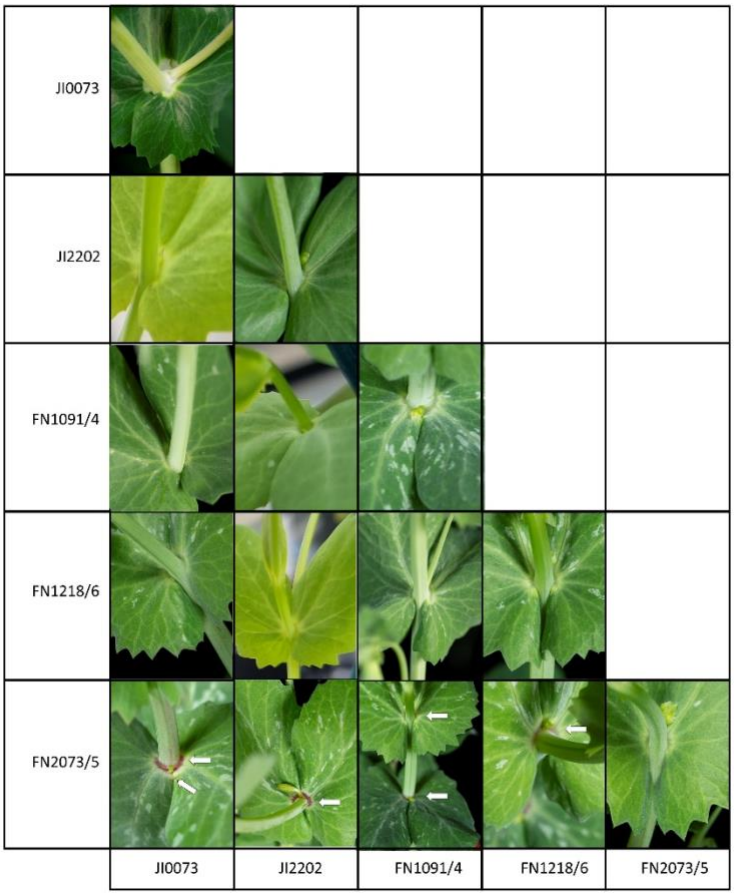


Supplementary Fig. 18 | *D* VIGS experiment. (a) Subcellular localization of *PsMYB16* (*Psat02G0138300*). **(b)** Transcriptional activation experiment of *PsMYB16*. **(c)** Phenotypes after gene silencing of *PsMYB16* using VIGS. The left panel shows a PDS control (VIGS-PDS), the central panel is an empty vector control (VIGS-Con) and the rightmost panel shows the knock-down of *Psat02G0138300*. The plants in which these experiments were carried out have the D^w allele and so have both an inner and outer ring of axil pigmentation. In the VIGS knockdown the outer ring of pigmentation is no longer present, suggesting that *Psat02G0138300* has a role in patterning axil ring pigmentation. **(d)** Anthocyanin measurement after gene silencing of *PsMYB16*. **(e)** Phenotypes in a transient overexpression assay of *PsMYB16* in *N. Benthamiana*. **(f)** Anthocyanin measurement after overexpression of *PsMYB16*. ** represent a significant level at $P < 0.01$ using a Student's t-test in (d) and (f).

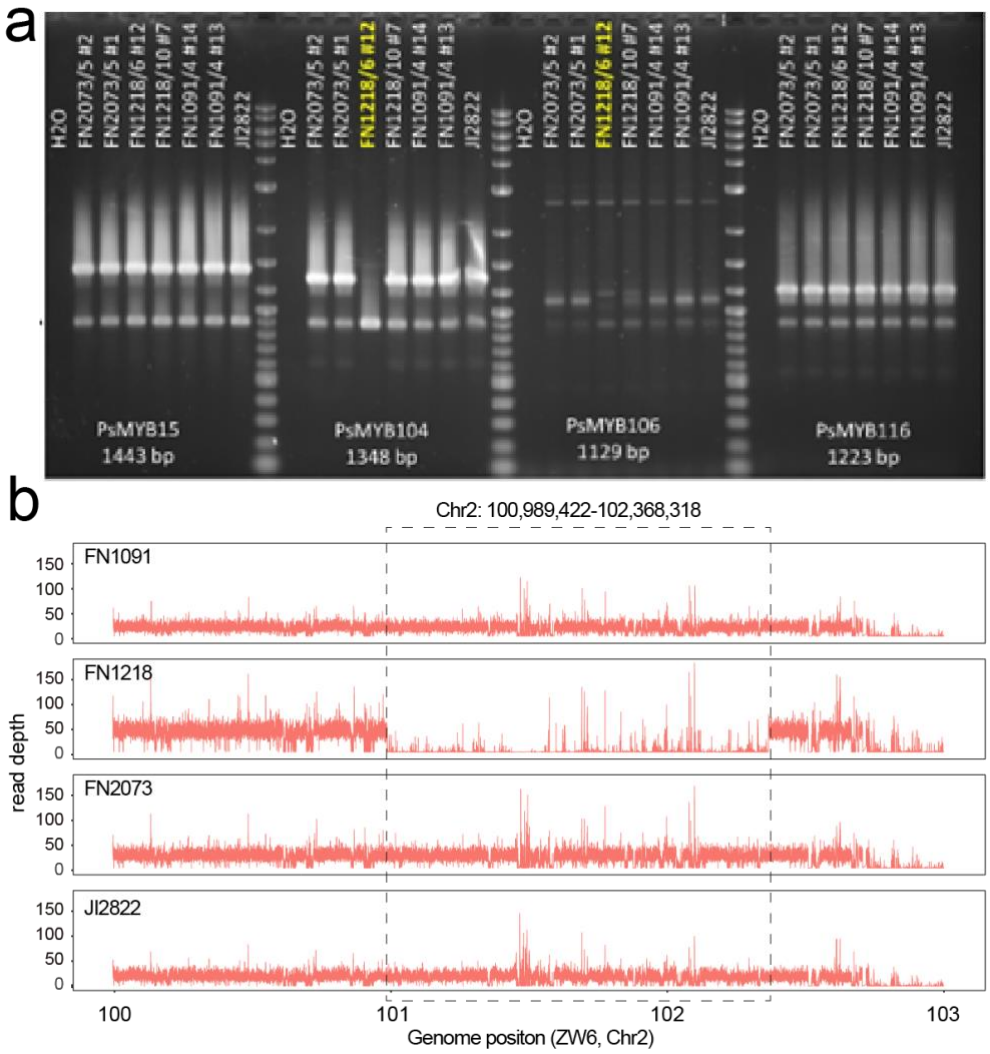


Supplementary Fig. 19 | Characterisation of one of the *MYB* genes within the gene cluster related to axillary anthocyanin. (**a**) Determination of total anthocyanin content in leaf axillary within the population. (**b**) Genome-wide association analysis of leaf axillary anthocyanins. (**c**) Genome-wide association analysis of leaf axillary anthocyanins after excluding white-flowered samples. The top row represents data from Shenzhen in 2021, and the bottom row

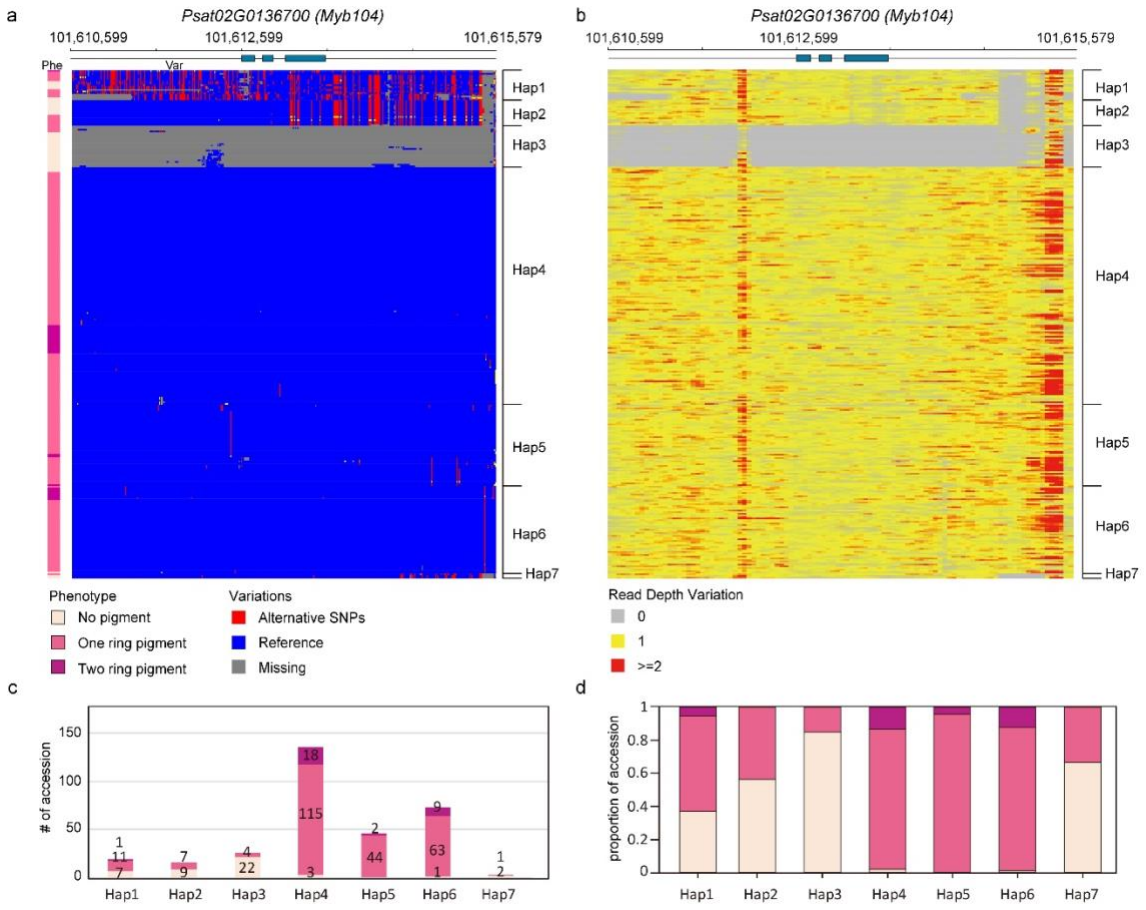
represents data from Harbin in 2022. **(d)** Phylogenetic tree construction using *PsMYB16* and the *MYB* transcription factor family. **(e)** Multiple sequence alignment of *PsMYB16* with other genes controlling leaf anthocyanin.



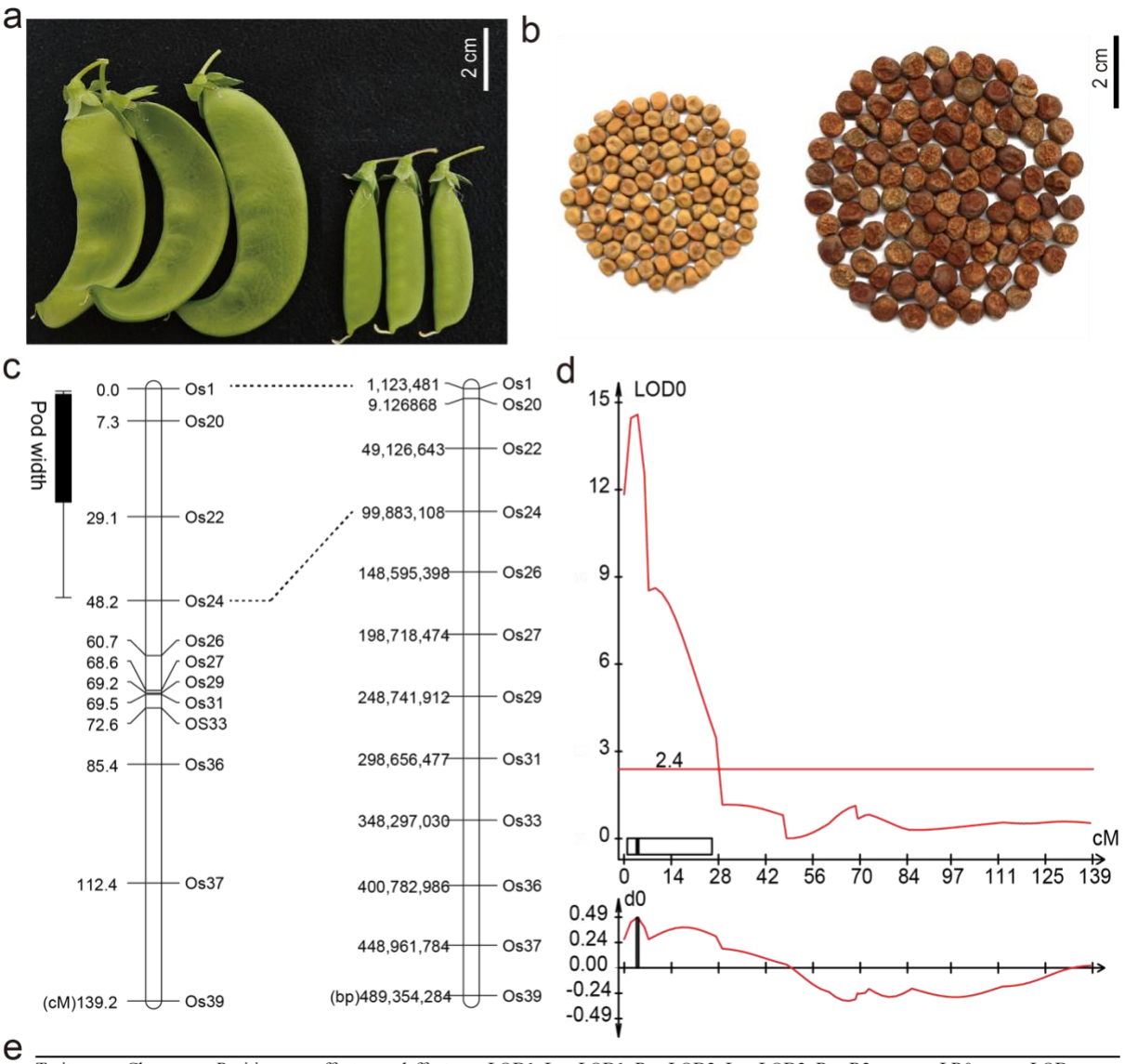
Supplementary Fig. 20 | Allelism tests of *d* mutants. The axils of self-seeded plants and F_1 hybrids of crosses between the lines indicated by the rows and columns are shown. There are two complementation groups, one of which corresponds to *d* from JI0073 and JI2202 and the other to FN2073/5.



Supplementary Fig. 21 | *MYB* gene deletions. Confirmation of the deletion of *MYB104* and *MYB106* in FN1218 mutant. (a) PCR amplification of *MYB* genes, confirming the deletion of *MYB104* and *MYB106* in some of the lines. (b) The read depth of FN lines and JI2822 aligned against the ZW6 reference genome. PCR confirmation of the deletion of *PsMYB104* and *PsMYB106*, but not *PsMYB15* or *PsMYB116* in the FN1218/6 plant #12 (plant #10 was a wild type M4 segregant). This further refined the extent of the deletion in FN1218/6. No deletion was detected in FN1091/4 (see Supplementary Figure 19) in either assay, suggesting that its deletion is small, is a base substitution rather than a deletion, or involves a sequence repeated within this region.



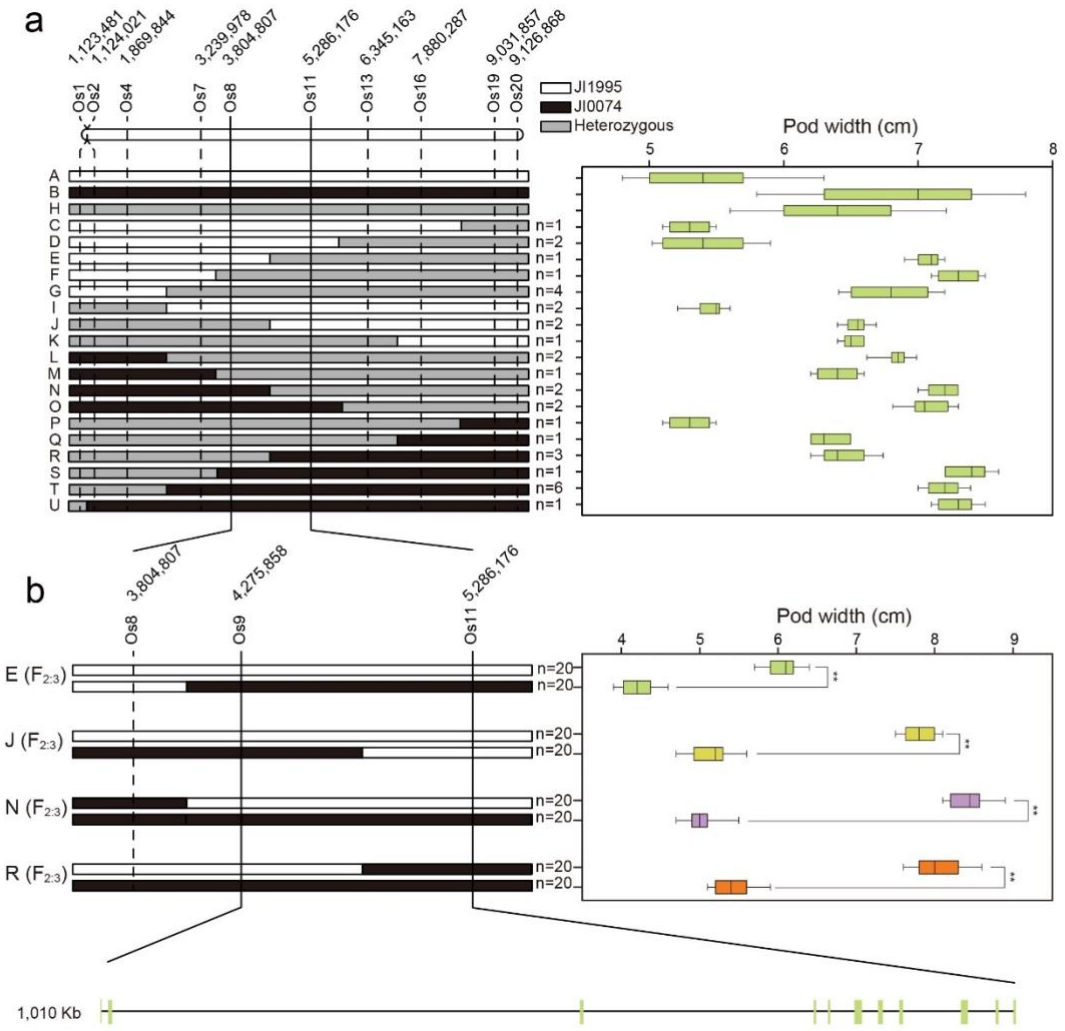
Supplementary Fig. 22 | *MYB104* is the mostly like gene candidate of *D*, explaining most of the phenotypic changes. **(a)** Haplotypes of *Psat02G0136700* (*MYB104*). The phenotype and variant are colour coded and shown in the bottom of the haplotype plot. **(b)** Variation of read depths of SNPs around *Psat02G0136700* (*MYB104*). The read depth is colour coded, with grey, yellow and red showing read depth of 0, 1 and a value greater than 2, respectively. The haplotypes are shown on the right of the plot. **(c)** The number of lines with different phenotypes categorized in each haplotype. **(d)** The proportion of lines with different phenotypes in each haplotype. The colour code of phenotype in (c) and (d) is the same as the code shown in (a).



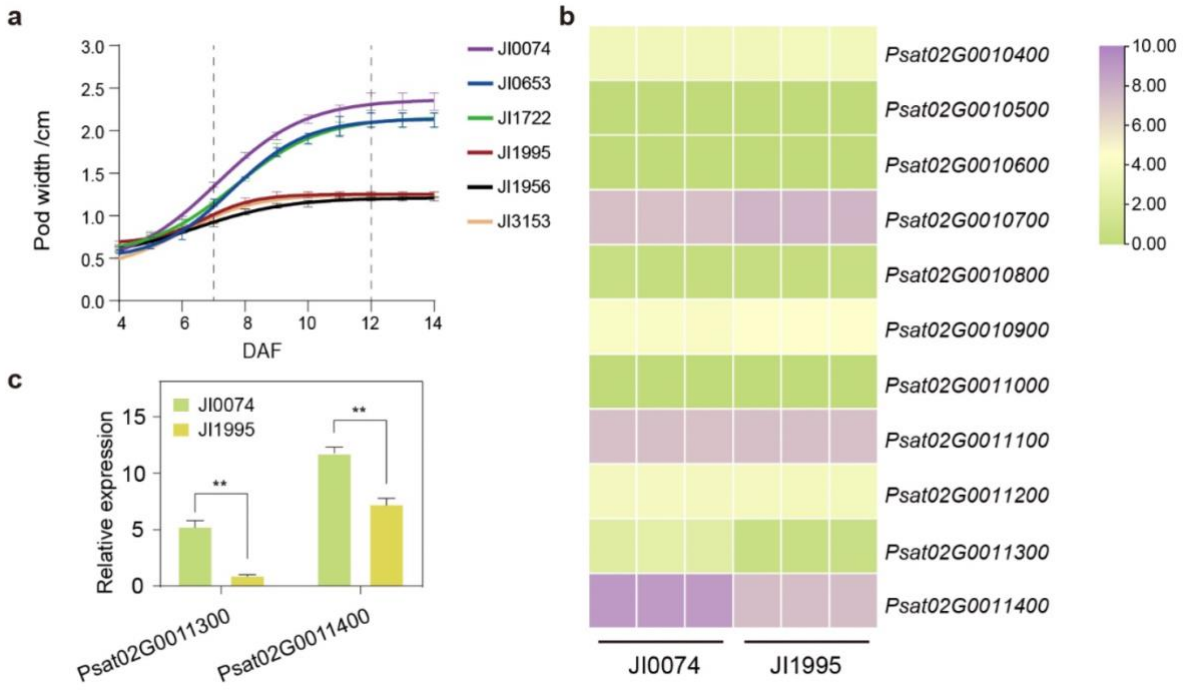
e

Trait	Chr	Position	aeffect	deffect	LOD1 L	LOD1 R	LOD2 L	LOD2 R	R2	LR0	LOD
Pod width	Chr2LG1	4	-2.4274	0.4884	1.3	25.8	0.6	47.5	0.253444	67.16635	14.60138

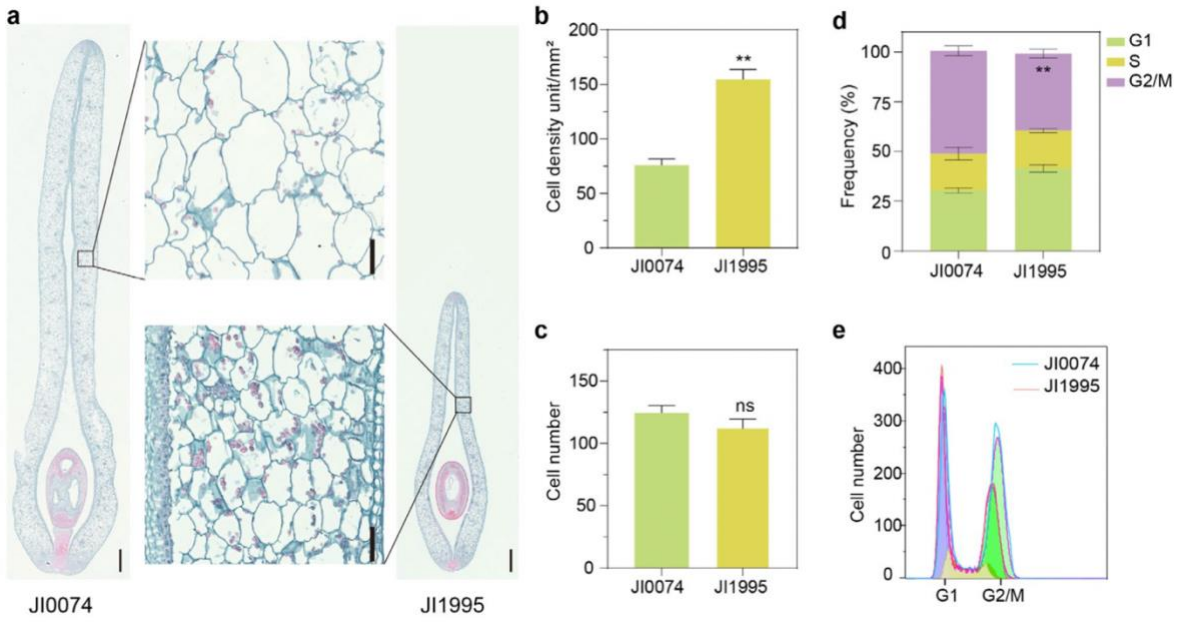
Supplementary Fig. 23 | QTL mapping of *Os1* locus. Comparison of the phenotypes of (**a**) wide and narrow pods, **(b)** large and small grains, in the two parental lines of JI0074 and JI1995, respectively, which were used to build the F2 mapping population. The scale bars in **(a)** and **(b)** represent 2 cm. **(c).** The linkage group of KASP makers (left) mapped using a F2 population derived from the cross between JI0074 and JI1995. The physical positions of these markers in this candidate region were shown in the right-hand map. **(d)**. Interval Mapping scores (logarithm of odds, LOD) of markers for pod width. Horizontal dash lines indicated the threshold of LODs. **(e)** Location and statistics of the target *Os1* locus.



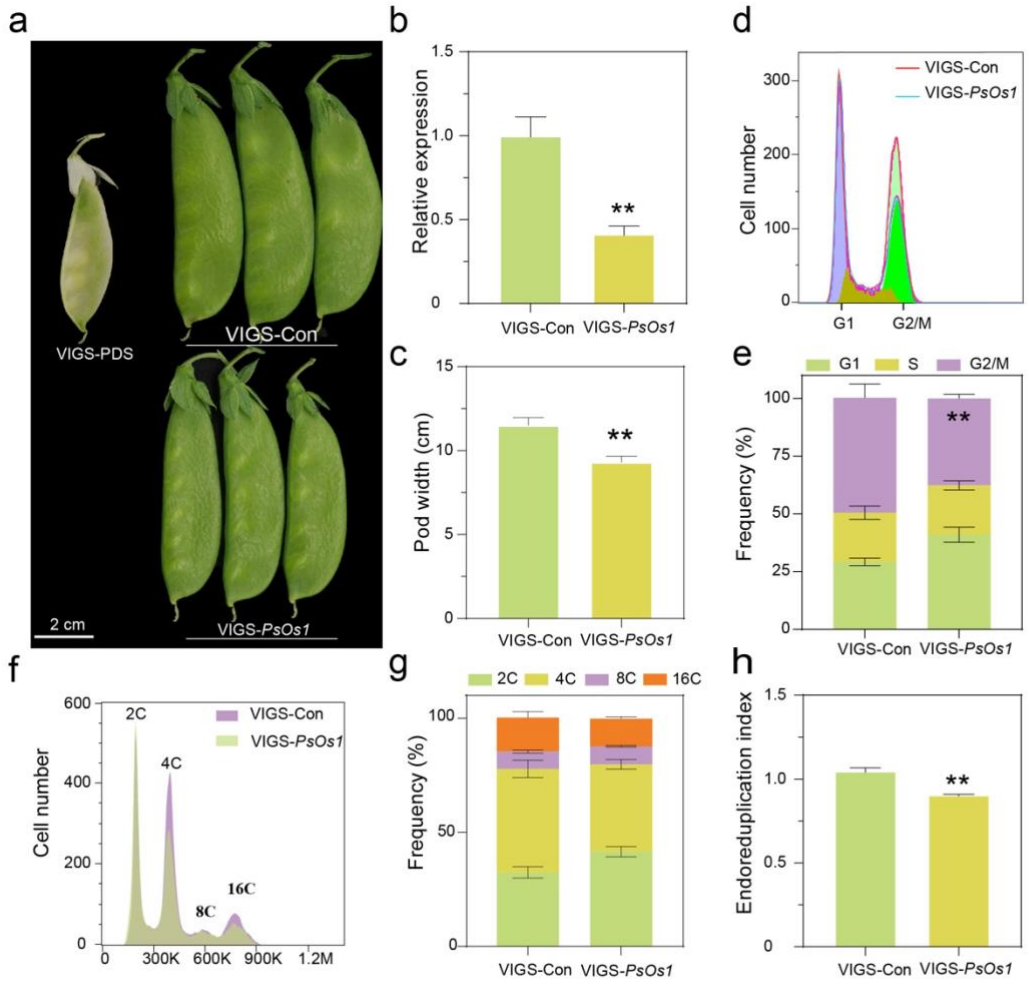
Supplementary Fig. 24 | Fine mapping of the *Os1* locus. (a) The graphical distribution of the genotypes in the candidate regions (left) and distribution of phenotypes of pod width (right) of the lines in a F2 population derived from the cross of JI0074 and JI1995. The white, black, and grey bars in the box represent genotypes of JI1995, JI0074 and heterozygotes, respectively. The number of individual plants (n) of each genotype was shown on the left of the corresponding boxes. **(b)** The graphical genotypes in the candidate regions (left) and phenotypes of pod width (right) of the lines in F2:3 populations derived from the cross of JI0074 and JI1995, which enabled the narrowing down of the candidate regions into a locus containing only 11 genes, as shown beneath. The statistics of pod width are shown on the right and the significant differences are indicated by * ($P < 0.05$), **($P < 0.01$) (Student's t-test).



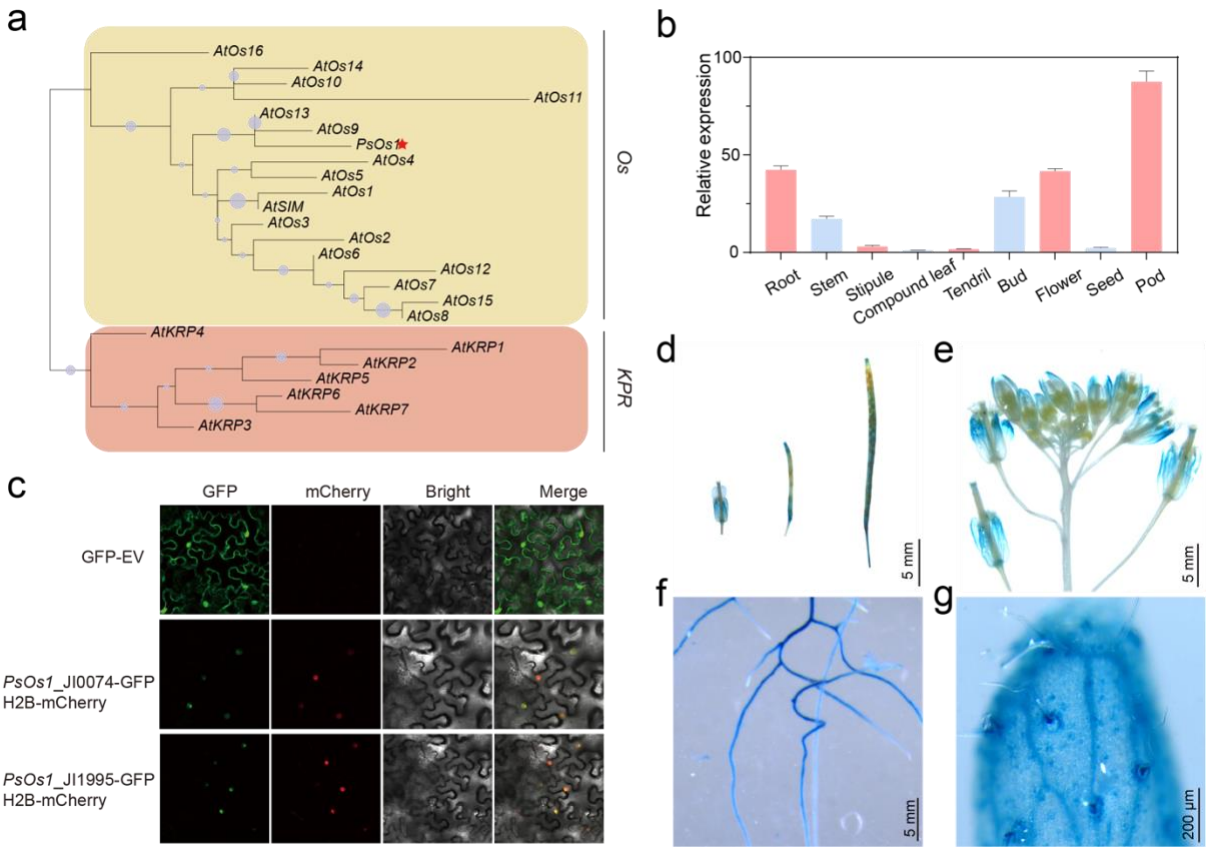
Supplementary Fig. 25 | Expression analysis of candidate genes at *PsOs1* loci. **(a)** Growth curves of wide-pod and narrow-pod varieties. **(b)** Heatmap of relative gene expression within candidate intervals. **(c)** Fluorescence quantitative detection of differentially expressed genes.



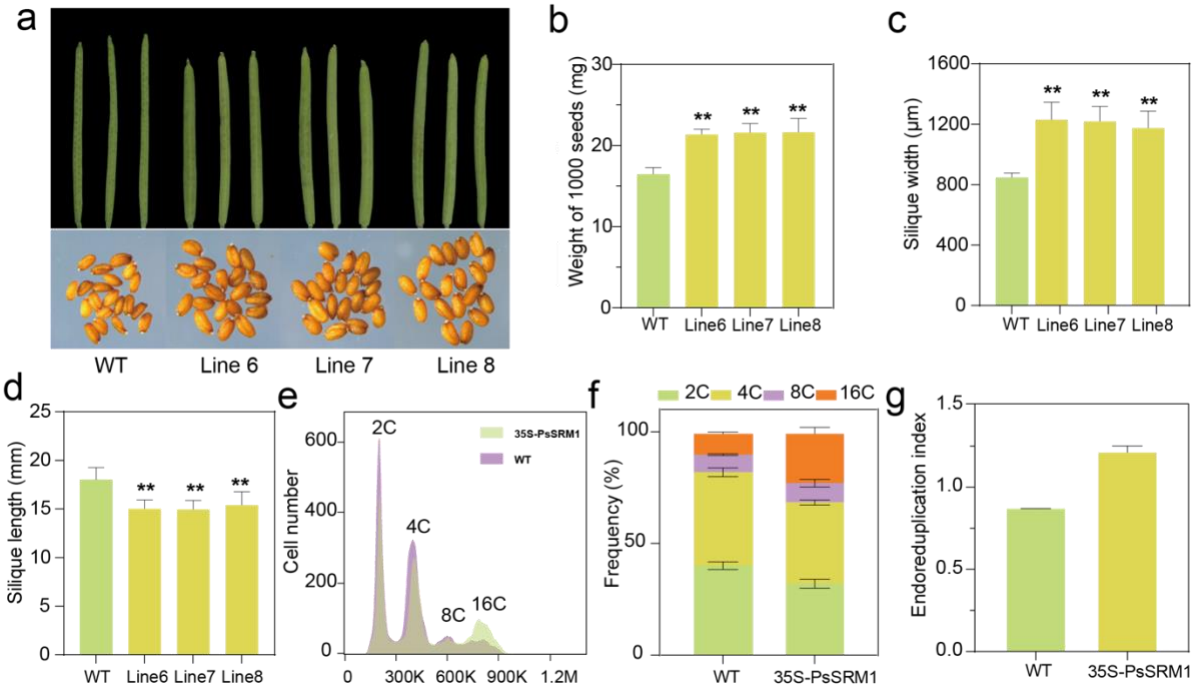
Supplementary Fig. 26 | Cellular observations and comparison in the two parental lines differing in pod width. **(a)** Cross-section of pods from JI0074 and JI1995 at 12 days after flowering, stained with Fuchsin Green dye. **(b)** Cell density per unit area within the pod cross-sections of JI0074 and JI1995. **(c)** Number of cells in transverse arrangement within the pod cross-sections of JI0074 and JI1995. **(d)** Cell cycle statistics for pods of the two lines. **(e)** Flow cytometry display of cells at different cell cycles in pods of the two lines.



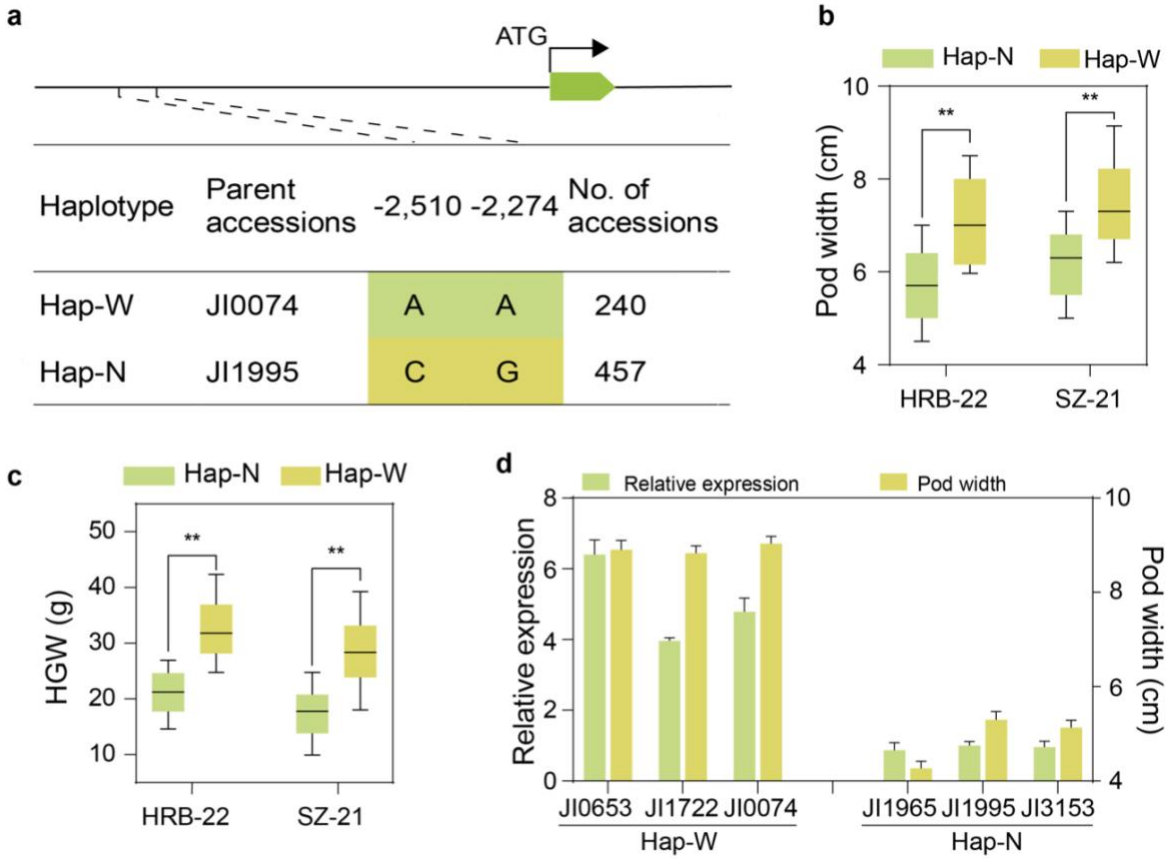
Supplementary Fig. 27 | VIGS experiment of *PsOs1*. (a) Phenotypic changes in *PsOs1*-silenced lines of pea (VIGS-*PsOs1*) in comparison with the negative control, a line inoculated with VIGS-Con. **(b)** Relative expression level of *PsOs1*gene in the lines inoculated with VIGS-Con (control) and VIGS- *PsOs1*. **(c)** Pod width of lines of VIGS-Con and VIGS- *PsOs1*. **(d)** Quantification of cells at different phases of cell cycle in lines of VIGS-Con and VIGS- *PsOs1*. **(e)** Proportion of cells at G1, S and G2/M phases in lines of VIGS-Con and VIGS- *PsOs1*. **(f)** Quantification of cells at different ploidies (2C-16C) in lines of VIGS-Con and VIGS- *PsOs1*. **(g)** Proportion of cells with different ploidies (2C, 4C, 8C and 16C) in lines of VIGS-Con and VIGS- *PsOs1*. **(h)** Endoreduplication index of lines of VIGS-Con, VIGS- *PsOs1*. '**' indicates statistical significance at $P < 0.01$ (Student's t-test).



Supplementary Fig. 28 | Expression analysis of *PsOs1*. (**a**) Clustering of *PsOs1* and homologous genes in *Arabidopsis* in the phylogenetic tree. (**b**) Relative expression levels of *PsOs1* in different tissues, including pod of pea. (**c**) Subcellular location of *PsSMR1* indicated by the expression of *PsOs1*-GFP fusion protein in tobacco. (**d**) GUS staining of the silique of transgenic Arabidopsis (PsOs1pro::GUS)expressing GUS promoted by the promoter of *PsOs1*. (**e**) GUS staining of the inflorescence of transgenic Arabidopsis (PsOs1pro::GUS) expressing GUS promoted by the promoter of *PsOs1*. (**f**) GUS staining of the root of transgenic Arabidopsis (PsOs1pro::GUS) expressing GUS promoted by the promoter of *PsOs1*. (**g**) GUS staining of the trichomes on the sepals of transgenic Arabidopsis (PsOs1pro::GUS) expressing GUS promoted by the promoter of *PsOs1*.



Supplementary Fig. 29 | Analysis of the phenotype of *PsOs1* transgenic *Arabidopsis*. (a) Overexpression of *PsOs1* in *Arabidopsis* resulted in an increasing in silique width and seed weight. **(b)** Difference in weight of 1000 seeds between the wild-type (WT) and transgenic lines of Arabidopsis expressing *PsOs1*. **(c)** Difference in silique width between the wild-type and transgenic lines of *Arabidopsis* overexpressing *PsOs1*. **(d)** Difference in silique length between the WT and transgenic lines of Arabidopsis overexpressing *PsOs1*. **(e)** Quantification of cells at different ploidies (2C-16C) in Arabidopsis overexpressing *PsOs1* and wild type. **(f)** Proportion of cells with different ploidies (2C, 4C, 8C and 16C) in *Arabidopsis* overexpressing *PsOs1* compared to WT. **(g)** Endoreduplication index of Arabidopsis lines of Arabidopsis overexpressing *PsOs1* compared to WT.



Supplementary Fig. 30 | Analysis of the haplotypes of the promoter region of *PsOs1*. (a) Haplotypes (Hap-W and Hap-N) of *PsOs1* separates the germplasms with wide pods from those with narrow pods. (b) The pod width of germplasms with Hap-W and Hap-N. (c) Hundred-grain weight (HGW) of the seeds of germplasms with Hap-W and Hap-N. (d) Relative expression levels and pod width of *PsOs1* in the pods of selected germplasms of Hap-W and Hap-N.

References:

1 Bhattacharyya, M. K., Smith, A. M., Ellis, T. H., Hedley, C. & Martin, C. The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. *Cell* **60**, 115-122 (1990). https://doi.org/10.1016/0092-8674(90)90721-p

2 Rayner, T. *et al.* Genetic Variation Controlling Wrinkled Seed Phenotypes in Pisum: How Lucky Was Mendel? *Int J Mol Sci* **18** (2017). https://doi.org/10.3390/ijms18061205

3 Wang, T. L., Bogracheva, T. Y. & Hedley, C. L. Starch: as simple as A, B, C? *Journal of Experimental Botany* **49**, 481-502 (1998). https://doi.org/10.1093/jxb/49.320.481

4 Casey, R. *et al.* The effect of modifying carbohydrate metabolism on seed protein gene expression in peas. *Journal of Plant Physiology* **152**, 636-640 (1998). https://doi.org/https://doi.org/10.1016/S0176-1617(98)80023-0

5 Kooistra, E. On the differences between smooth and three types of wrinkled peas. *Euphytica* **11**, 357-373 (1962).

6 Sato, Y., Morita, R., Nishimura, M., Yamaguchi, H. & Kusaba, M. Mendel's green cotyledon gene encodes a positive regulator of the chlorophyll-degrading pathway. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14169-14174 (2007). https://doi.org/10.1073/pnas.0705521104

7 Shimoda, Y., Ito, H. & Tanaka, A. Arabidopsis STAY-GREEN, Mendel's Green Cotyledon Gene, Encodes Magnesium-Dechelatase. *Plant Cell* **28**, 2147-2160 (2016). https://doi.org/10.1105/tpc.16.00428

8 Armstead, I. *et al.* Cross-species identification of Mendel's I locus. *Science* **315**, 73 (2007). https://doi.org/10.1126/science.1132912

9 Hellens, R. P. *et al.* Identification of Mendel's white flower character. *PLoS One* **5**, e13230 (2010). https://doi.org/10.1371/journal.pone.0013230

10 Harker, C. L., Ellis, T. H. & Coen, E. S. Identification and genetic regulation of the chalcone synthase multigene family in pea. *Plant Cell* **2**, 185-194 (1990). https://doi.org/10.1105/tpc.2.3.185

11 Ellis, N. *et al.* Recombinant inbred lines derived from wide crosses in Pisum. *Sci. Rep.* **13**, 20408 (2023). https://doi.org/10.1038/s41598-023-47329-9

12 Liu, Q., Wu, K., Harberd, N. P. & Fu, X. Green Revolution DELLAs: From translational reinitiation to future sustainable agriculture. *Mol Plant* **14**, 547-549 (2021). https://doi.org/10.1016/j.molp.2021.03.015

13 Spielmeyer, W., Ellis, M. H. & Chandler, P. M. Semidwarf (sd-1), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9043-9048 (2002). https://doi.org/10.1073/pnas.132266399

14 Weston, D. E. *et al.* The Pea DELLA proteins LA and CRY are important regulators of gibberellin synthesis and root growth. *Plant Physiol.* **147**, 199-205 (2008). https://doi.org/10.1104/pp.108.115808

15 Reid, J. B., Murfet, I. C. & Potts, W. C. Internode Length in Pisum. II. Additional Information on the Relationship and Action of Loci Le, La, Cry, Na and Lm. *J. Exp. Bot.* **34**, 349-364 (1983). https://doi.org/10.1093/jxb/34.3.349

16 Price, D. N. & Hedley, C. L. The effect of the gp gene on fruit development in Pisum sativum L. *New Phytol.* **110**, 271-277 (1988). https://doi.org/https://doi.org/10.1111/j.1469-8137.1988.tb00262.x

17 White, O. E.

18 Neff, D. I. & White, O. E. INHERITANCE STUDIES IN PISUM VI. MULTIPLE ALLELOMORPHISM AND THE INHERITANCE OF GREEN AND YELLOW FOLIAGE AND POD COLOR. *Am. J. Bot.* **14**, 379-394 (1927).

19 Shirasawa, K., Sasaki, K., Hirakawa, H. & Isobe, S. Genomic region associated with pod color variation in pea (Pisum sativum). *G3 (Bethesda)* **11** (2021). https://doi.org/10.1093/g3journal/jkab081

20 Dalmais, M. *et al.* UTILLdb, a Pisum sativum in silico forward and reverse genetics tool. *Genome Biol* **9**, R43 (2008). https://doi.org/10.1186/gb-2008-9-2-r43

21 Lesley, M. M. Mendel's Letters to Carl Nageli. *The American Naturalist* **61**, 370-378 (1927).
22 Gregor Johann Mendel 1822-1884. In centenary commemoration. *Hereditas* **100**, Ii-xiii (1984).
23 Ellis, T. H. N., Hofer, J. M. I., Swain, M. T. & van Dijk, P. J. Mendel's pea crosses: varieties, traits and statistics. *Hereditas* **156**, 33 (2019). https://doi.org/10.1186/s41065-019-0111-y
24 Yang, T. *et al.* Improved pea reference genome and pan-genome highlight genomic features and evolutionary characteristics. *Nat. Genet.* (2022). https://doi.org/10.1038/s41588-022-01172-2
25 Zhang, P. *et al.* Fine mapping PsPS1, a gene controlling pod softness that defines market type in pea (Pisum sativum). *Plant Breeding* **141**, 418-428 (2022). https://doi.org/https://doi.org/10.1111/pbr.13020
26 Lamprecht, H. & Mrkos, H. The inheritance of the bracteole in Pisum and linkage of the gene Br. (1950).
27 Blixt, S. Mutation genetics in Pisum. (1972).
28 Zhong, R. & Ye, Z. H. IFL1, a gene regulating interfascicular fiber differentiation in Arabidopsis, encodes a homeodomain-leucine zipper protein. *Plant Cell* **11**, 2139-2152 (1999). https://doi.org/10.1105/tpc.11.11.2139
29 Sussmilch, F. C., Ross, J. J. & Reid, J. B. Mendel: From genes to genome. *Plant Physiol.* **190**, 2103-2114 (2022). https://doi.org/10.1093/plphys/kiac424
30 Gerard, J., Dodoens, R., Priest, R. & Wright, S. S. *The Herball, or, Generall historie of plantes*. (By Iohn Norton, 1597).
31 Sinjushin, A. & Gostimskii, S. Relationship between different fasciated lines of pea. *Pisum Genetics* **39**, 16-18 (2007).
32 Würschum, T., Groß-Hardt, R. & Laux, T. APETALA2 Regulates the Stem Cell Niche in the Arabidopsis Shoot Meristem. *The Plant Cell* **18**, 295-307 (2005). https://doi.org/10.1105/tpc.105.038398
33 Krusell, L. *et al.* The Clavata2 genes of pea and Lotus japonicus affect autoregulation of nodulation. *Plant J.* **65**, 861-871 (2011). https://doi.org/10.1111/j.1365-313X.2010.04474.x
34 Siniushin, A. A. & Gostimskiĭ, S. A. [Genetic control of fasciation in pea (Pisum sativum L.)]. *Genetika* **44**, 807-814 (2008).
35 Osipova, M. A. *et al.* WUSCHEL-RELATED HOMEOBOX5 Gene Expression and Interaction of CLE Peptides with Components of the Systemic Control Add Two Pieces to the Puzzle of Autoregulation of Nodulation *Plant Physiol.* **158**, 1329-1341 (2012). https://doi.org/10.1104/pp.111.188078
36 Hu, C. *et al.* A group of receptor kinases are essential for CLAVATA signalling to maintain stem cell homeostasis. *Nat Plants* **4**, 205-211 (2018). https://doi.org/10.1038/s41477-018-0123-z
37 Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M. & Simon, R. Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. *Science* **289**, 617-619 (2000). https://doi.org/10.1126/science.289.5479.617
38 Fletcher, J. C., Brand, U., Running, M. P., Simon, R. & Meyerowitz, E. M. Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**, 1911-1914 (1999). https://doi.org/10.1126/science.283.5409.1911
39 Schoof, H. *et al.* The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**, 635-644 (2000). https://doi.org/10.1016/s0092-8674(00)80700-x
40 Domoney, C. *et al.* Exploiting a fast neutron mutant genetic resource in Pisum sativum (pea) for functional genomics. *Funct. Plant Biol.* **40**, 1261-1270 (2013). https://doi.org/https://doi.org/10.1071/FP13147
41 Święcicki, W. K. Two new anthocyanin genes in the D segment of chromosome 1. *Pisum Newsletter* **22**, 56-61 (1990).
42 Sawada, C. *et al.* An Integrated Linkage Map of Three Recombinant Inbred Populations of Pea (Pisum sativum L.). *Genes (Basel)* **13** (2022). https://doi.org/10.3390/genes13020196

43 Nomoto, Y. *et al.* A hierarchical transcriptional network activates specific CDK inhibitors that regulate G2 to control cell size and number in Arabidopsis. *Nat Commun* **13**, 1660 (2022). https://doi.org/10.1038/s41467-022-29316-2

44 Tourdot, E., Mauxion, J. P., Gonzalez, N. & Chevalier, C. Endoreduplication in plant organogenesis: a means to boost fruit growth. *J Exp Bot* **74**, 6269-6284 (2023). https://doi.org/10.1093/jxb/erad235

45 Kreplak, J. *et al.* A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* **51**, 1411-1422 (2019). https://doi.org/10.1038/s41588-019-0480-1

46 Sinjushin, A. A. & Gostimsky, S. A. [Fasciation in pea: basic principles of morphogenesis]. *Russian Journal of Developmental Biology* **37**, 375-381 (2006).

47 Zhuang, L. L. LATHYROIDES, Encoding a WUSCHEL-Related Homeobox1 Transcription Factor, Controls Organ Lateral Growth, and Regulates Tendril and Dorsal Petal Identities in Garden Pea (Pisum sativum L.). *Molecular Plant* **5**, 1333-1345 (2012).