

1 **Validating Genome-Wide Association candidates through quantitative variation in nodulation**

2
3 Shaun J. Curtin^{1†}, Peter Tiffin^{2†}, Joseph Guhlin², Diana Trujillo², Liana Burghart², Paul Atkins³, Nicholas J.
4 Baltes³, Roxanne Denny¹, Daniel F. Voytas³, Robert M. Stupar⁴, and Nevin D. Young^{1,2*}

5
6 ¹Department of Plant Pathology, University of Minnesota, St. Paul, MN, 55108. ²Department of Plant
7 Biology, University of Minnesota, St. Paul, MN, 55108. ³Department of Genetics, Cell Biology &
8 Development and Center for Genome Engineering, University of MN, Minneapolis, MN 55455.

9 ⁴Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN, 55108. [†]These
10 authors contributed equally to this work. *email: nevin@umn.edu

11
12 ORCID ID: 0000-0002-9528-3335 (S.J.C.); 0000-0002-8836-2924 (R.M.S.); 0000-0001-6463-4772 (N.D.Y.)

13
14
15
16 **One Sentence Summary:**

17
18 GWAS combined with multiple mutagenesis technologies, including CRISPR/Cas9, were used to
19 functionally validate novel candidate genes contributing to phenotypic variation in symbiosis between
20 legume plants and rhizobial bacteria.

21
22
23
24
25 **Author contributions:**

26 S.J.C., P.T., N.D.Y., and R.M.S. designed research;
27 S.J.C., P.T., J.G., R.D., P.A., N.J.B., J.M., A.D.F., and D.F.V. performed research;
28 S.J.C, P.T., and N.D.Y. analyzed data;
29 S.J.C, P.T., and N.D.Y wrote the paper.

30
31 The authors declare no conflict of interest.

32 This work was supported by NSF award IOS-1237993.

33 **Abstract**

34 Genome wide association (GWA) studies offer the opportunity to identify genes that contribute to
35 naturally occurring variation in quantitative traits. However, GWA relies exclusively on statistical
36 association, so functional validation is necessary to make strong claims about gene function. We used a
37 combination of gene-disruption platforms (Tnt1 retro-transposons, hairpin RNA-interference constructs,
38 and CRISPR/Cas9 nucleases) together with randomized, well-replicated experiments to evaluate the
39 function of genes that an earlier GWAS in *Medicago truncatula* had identified as candidates contributing
40 to variation in the symbiosis between legumes and rhizobia. We evaluated ten candidate genes found in
41 six clusters of strongly associated SNPs, selected on the basis of their strength of statistical association,
42 proximity to annotated gene models, and root or nodule expression. We found statistically significant
43 effects on nodule production for three candidate genes, each validated in two independent mutants.
44 Annotated functions of these three genes suggest their contributions to quantitative variation in nodule
45 production occur through processes not previously connected to nodulation, including phosphorous
46 supply and salicylic acid-related defense response. These results demonstrate the utility of GWA
47 combined with reverse mutagenesis technologies to discover and validate genes contributing to
48 naturally occurring variation in quantitative traits. The results highlight the potential for GWA to
49 complement forward genetics in identifying the genetic basis of ecologically and economically important
50 traits.

51

52

53

54 Genome wide association (GWA) studies offer the promise of identifying genes responsible for
55 quantitative trait variation in plants and animals (Mackay et al., 2009). Relative to forward genetic
56 screens, GWA has the potential to identify genes of much smaller phenotypic effect. Moreover, the
57 genes identified through GWA are expected to contribute to naturally occurring phenotypic variation. By
58 contrast, forward genetic screens identify genes that are functionally necessary but may or may not
59 contribute to natural variation. An important limitation of GWA analyses, however, is that the statistical
60 approach used to identify candidate genes is also expected to identify many false positives. Therefore,
61 functional validation of GWA candidates is essential for confirming the biological importance of the
62 identified genes (Ioannidis et al., 2009).

63 The symbiosis between legume plants and rhizobial bacteria is responsible for large inputs of
64 nitrogen into both natural and agricultural systems through symbiotic nitrogen fixation (Vance, 2001).
65 The ecological and economic importance of this symbiosis has made identifying its genetic basis an
66 important goal for plant biologists (Oldroyd et al., 2011). More than 30 plant genes that play central
67 roles in the formation and growth of plant nodules, the site of rhizobial symbiosis, and nitrogen fixation
68 have been identified (Popp and Ott, 2011; Pislariu et al., 2012). These genes were identified primarily
69 through forward genetic screens. In standard forward genetic screens, a mutant phenotype is identified
70 based on the effects seen within individual organisms, something that will bias these screens towards
71 detecting genes of large effect. Forward genetic screens are likely to miss genes of small phenotypic
72 effects, such as those that contribute to naturally occurring quantitative variation in nodulation
73 (Stanton-Geddes et al., 2013; Friesen et al., 2014). Identification of these genes of subtle effect will
74 provide a fuller understanding of the symbiosis between legumes and rhizobia, and may provide targets
75 for selection in important agronomic legumes.

76 Here we report on the application of three reverse genetics mutagenesis tools: *M. truncatula* Tnt1
77 retrotransposon mutant collection (Tadege et al., 2008; Pislariu et al., 2012), hairpin RNA-interference
78 knock-down constructs (Wesley et al., 2001), and CRISPR/Cas9 site-specific nuclease (SSN) (Baltes et al.,
79 2014) to empirically evaluate the functional importance of ten genes that a previous sequence-based
80 GWA study identified as candidates for naturally occurring phenotypic variation in the legume-rhizobia
81 symbiosis (Stanton-Geddes et al., 2013). We used 17 whole-plant, stable mutants (six Tnt1, three hairpin,
82 and eight CRISPR mutants) in well replicated phenotypic assays to evaluate the function of each of the
83 ten GWA candidates. The functional importance of three (*PHO2-like*, *PNO1-like* and *PEN3-like*) was
84 validated by two independent mutations, confirming the importance of genes found in three of the six
85 clusters investigated. These results not only identify genes not previously identified as being involved in

86 legume-rhizobia symbiosis, they also illustrate the power of combining GWA with newly-developed
87 gene-disruption technologies to identify genes that contribute to naturally occurring variation in
88 quantitative traits.

89

90

91

92 RESULTS

93

94 Association mapping SNF-associate candidate genes in *Medicago*

95 To generate a list of candidates underlying variation in nodulation, we previously carried out GWA
96 analysis involving over 6 million SNPs on a panel of 226 *M. truncatula* accessions (Stanton-Geddes et al.,
97 2013). In brief, nodulation data were collected on eight replicates of each accession and GWA was
98 conducted using the GLMM implemented in TASSEL (Bradbury et al., 2007). From this GWA we selected
99 the 100 SNPs with the strongest association to variation in the number of nodules per plants, producing
100 an initial list of nodulation candidates. This list then was filtered based on several criteria: 1) Statistical
101 support and ranking for an association from the GWA, 2) Candidate being in linkage-disequilibrium (LD)
102 with other SNPs with strong statistical support (*i.e.*, located in a peak), 3) Location within or proximate
103 to annotated coding regions, and 4) Expression in root and or nodule tissue. GWA was performed twice
104 over the course of the research as a new assembly of the *M. truncatula* genome (Mt4.0) was released
105 mid-experiment (Tang et al., 2014), replacing the earlier assembly (Mt3.5) (Young et al., 2011) used in
106 the original analysis.

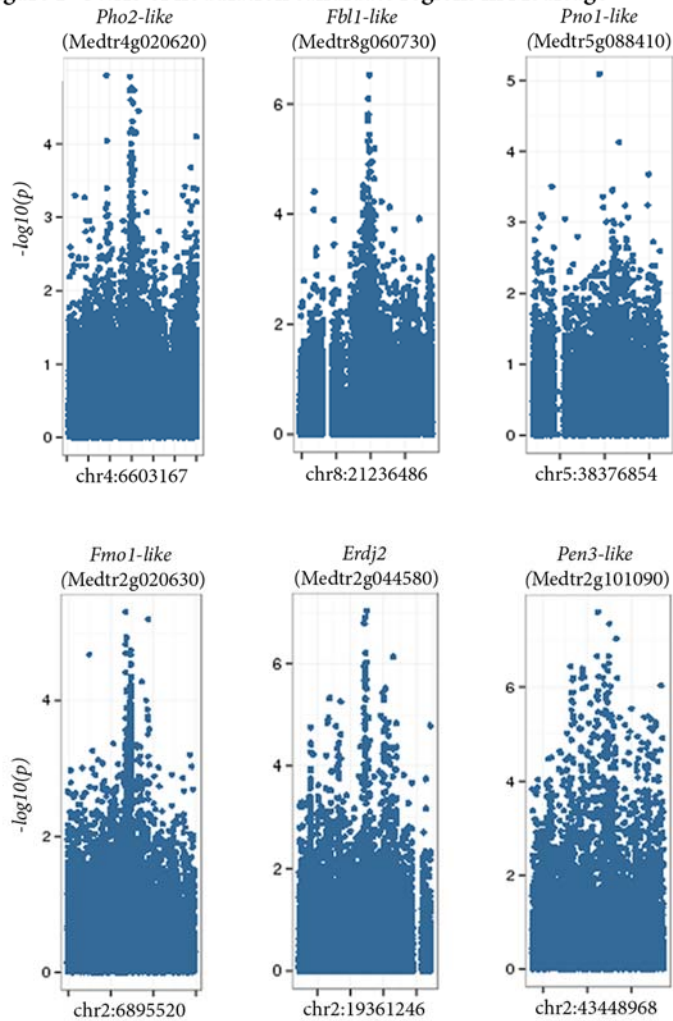
107 Based on these criteria, we selected ten candidates (noting that we did not require all candidates to
108 meet all of the filtering criteria) to target for mutagenesis and phenotypic validation (Table 1). Six of the
109 candidates were physically proximate to other candidates (two within 15 kb of each other and four
110 within 80 kb of one another). These candidates were chosen to evaluate the potential to identify the
111 causative gene when clusters of statistically associated SNPs identified by GWA spanned multiple genes
112 (Fig. 1).

113

114 Generation of whole plant mutants

115 To empirically evaluate the functional importance of GWA-identified candidate genes underlying
116 naturally occurring variation in nodulation, we generated a suite of whole-plant mutants using three
117 mutagenesis platforms: Tnt1 retrotransposons, RNA hairpin knockdown, and CRISPR/Cas9 (SSNs) (Fig.
118 2a-c). We identified Tnt1 insertions in seven candidates by either querying the Noble Foundation Tnt1
119 database (Tadege et al., 2008) or through PCR screening of pools of Tnt1 mutant lines (Table S1). For
120 each of the seven putative Tnt1 mutants, we used a PCR assay with gene and insertion-specific primers
121 to identify lines that were either homozygous for the Tnt1 insertion or wild-type at the respective
122 candidate locus; hereafter referred to as *pho2-like*_{Tnt}, *fb11-like*_{Tnt}, *pno1-like*_{Tnt}, *fmo1-like*_{Tnt}, *rfp1-like*_{Tnt},
123 *hlz1-like*_{Tnt} and *pen3-like*_{Tnt} (Fig. 3a-c, Fig. S1, Fig. S9). We then used homozygous insertion and WT lines

Figure 1 GWAS of nodulation candidate regions in Medicago

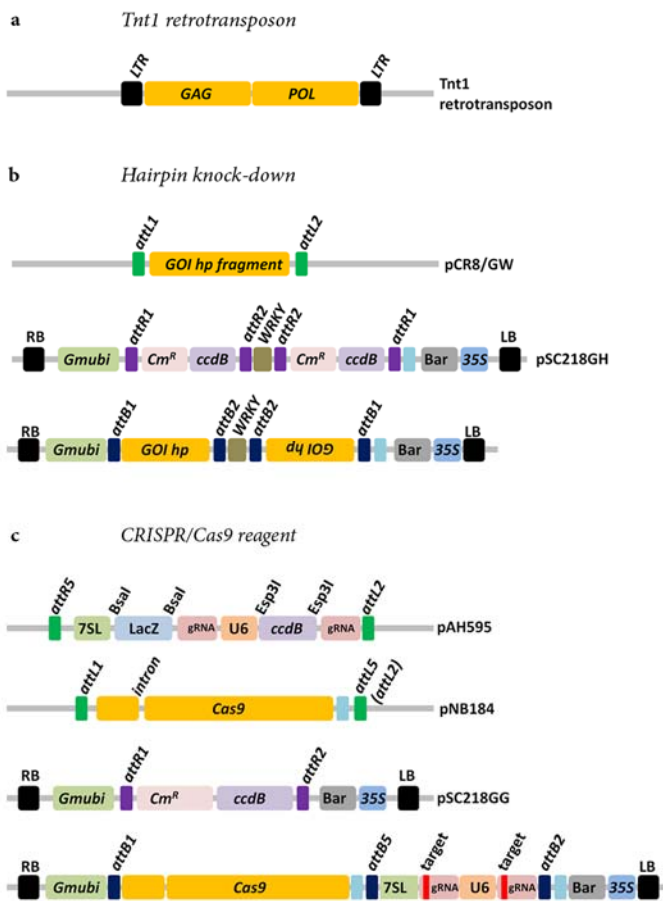


124 identified in this screen as parent plants to generate self-pollinated seeds for use in phenotype assays
125 described below. Preliminary analyses revealed that the *fb1*_{Tnt} line as well as the WT line derived from
126 the same grandparent plant were also segregating for a super-nodulating phenotype (> 100 nodules on
127 root systems < 5 cm long), and therefore these were excluded from further analyses.

128 We also attempted to generate hairpin knock-downs for all ten candidates and successfully
129 recovered multiple independently transformed whole plant mutant lines for four: *PHO2-like*, *FBL-like*,
130 *PNO1-like*, and *ERDJ2*. Using TAIL-PCR we were able to determine the genomic location of hairpin
131 transgene and identify homozygous hairpin lines for three of these genes, hereafter referred to as *fb1-*
132 *like*_{HP}, *pno1-like*_{HP}, and *erdj2*_{HP} (Fig. 3b and Fig. S2a-c). These lines were grown to produce self-pollinated
133 seed that was used for phenotype assays.

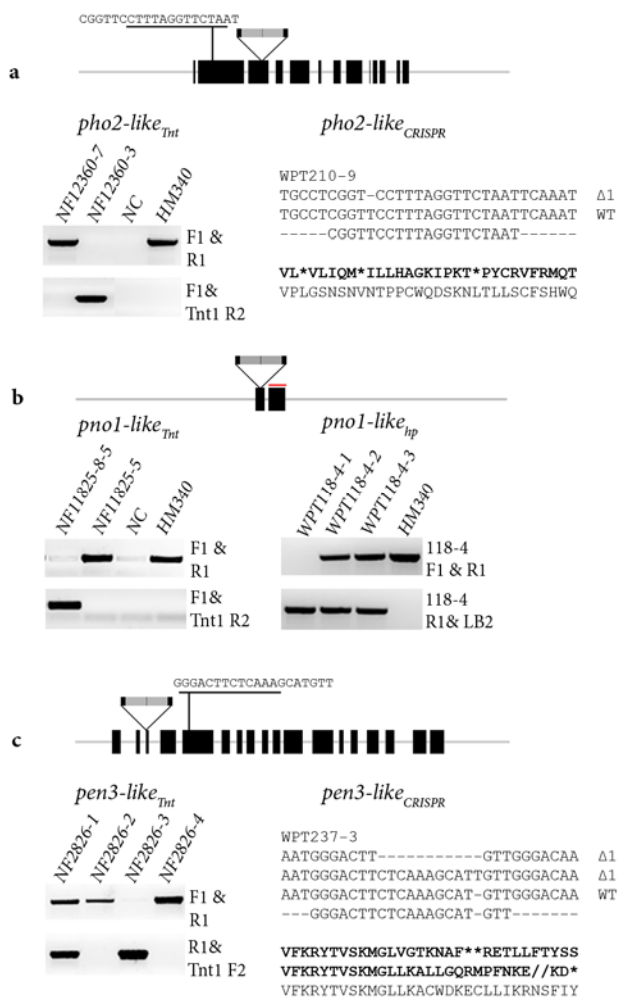
134 Because hairpin technologies are expected to reduce but not eliminate expression of target genes,
135 we also targeted all ten candidates for whole-plant knock-outs using CRISPR/Cas9 and successfully
136 generated whole-plant knock-out lines for seven candidate genes: *PHO2-like*, *FMO1-like*, *ERDJ2*, *ACRE1*,

Figure 2 Mutagenesis platforms



137 *HLZ1-like*, *MEL1* and *PEN3-like* (Fig. 3a-c and Fig. S3-S7). We recovered multiple T₀ homozygous mutant
 138 plants with mutation efficiencies ranging from 50-70% for many candidates genes including *PHO2-like*,
 139 *FMO1-like*, *ERDJ2*, and *PEN3-like* candidates, whereas germ-line mutations in *ACRE1*, *HLZ1-like*, *MEL1*
 140 mutant plants had to be confirmed in later generations (Fig. S3-S7). The CRISPR/Cas9 reagent was useful
 141 for mutating smaller sized genes (<500 bp) such as *Acre1* since the mutagenesis of these genes are
 142 typically more challenging using random mutagenesis strategies. Indeed, we could not identify a Tnt1
 143 insertion in this gene by querying the Noble database and were also unsuccessful using the Tnt1 PCR-
 144 reverse-screen (Fig. S7a). By contrast, using the CRISPR system, we generated several *Acre1* candidate
 145 mutants and could generate mutations at the *ACRE2* paralog located 16-kbp upstream of *ACRE1* (Fig.
 146 S7b-c). For the *ERDJ2* candidate we recovered two T₀ homozygous mutant plants, one with a bi-allelic 3-
 147 bp and 85-bp frame-shift deletion and a second plant with two 6-bp mutant alleles (Fig. S4) and tested
 148 both mutants since we suspected the plant might be homozygous lethal. We hereafter referred to the

Figure 3 Screening of candidate mutant lines.



149 CRISPR mutant lines as *pho2-like*_{CRISPR}, *fmo1-like*_{CRISPR}, *erdj2*_{CRISPR-3/85}, *erdj2*_{CRISPR-6}, *acre1-like*_{CRISPR}, *hlz1-*
 150 *like*_{CRISPR}, *mel1*_{CRISPR} and *pen3-like*_{CRISPR}. We were unable to confirm homozygous mutant plants for the
 151 remaining three candidates (*FBL1-like*, *PNO1-like* and *RFP1-like*). The apparent poor mutation rate of
 152 these three targets might be due to errors in the sequence used to construct the CRISPR reagent.
 153 However, we grew all of the stable transformants to produce self-pollinated seeds that were then used
 154 for phenotype assays.

155

156 Mutation and knockdown of gene candidates affect nodulation

157 We assayed nodule production in 17 mutant and paired control lines in a series of randomized and
 158 well-replicated phenotype experiments. The assays were performed with an average of 22 mutant and
 159 22 WT plants including six candidates tested with two independent mutation events to evaluate
 160 whether results were robust to the specific mutation and genomic background.

161 Phenotype assays revealed that lines carrying mutations in three of the candidates, *PHO2-like*
162 (4g020620), *PNO1-like* (5g088410), and *PEN3-like* (2g101090), produced significantly fewer nodules than
163 control plants ($P < 0.03$ for each mutant by wild-type line comparison; Fig. 4a-c and Table S2). In each of
164 these cases, the phenotypic effects were validated by two independent mutations (Fig. 3). The *PHO2-*
165 *like* mutants were generated by the insertion of a Tnt1 element (*pho2-like_{Tnt}*) in one plant and a
166 CRISPR/Cas9 generated mutant harboring a homozygous 1-bp frame-shift mutation (*pho2-like_{CRISPR}*) in
167 the other. *PNO1-like* mutant plants were generated by insertion of a Tnt1 element (*pno1-like_{Tnt}*) and a
168 plant homozygous for a hairpin transgene targeting the *PNO1-like* transcript (*pno1-like_{HP}*). Similarly, two
169 *PEN3-like* mutants were identified by Tnt1 insertion (*pen3-like_{Tnt}*) and a CRISPR/Cas9 induced 11-bp and
170 1-bp bi-allelic frame-shift mutation in each allele (*pen3-like_{CRISPR}*) (Fig. 3a-c).

171 Surprisingly, two Tnt1 insertion lines (*fmo1-like_{Tnt}* and *hlz1-like_{Tnt}*) produced significantly more
172 nodules than control lines (Table S2). However, second mutations in these candidates, both generated
173 using CRISPR/Cas9, exhibited results inconsistent with the Tnt1 mutants ($P_{mut_wt \times assay} < 0.05$). Indeed, the
174 number of nodules formed by lines carrying CRISPR mutations in these genes did not differ from
175 controls. Potentially, the greater nodule number in the Tnt1 lines may have been due to Tnt1 insertions
176 in non-target genes (two of the control Tnt1 plants when assaying the *hlz1-like_{Tnt}* had no nodules and
177 five others had ≤ 10 nodules – many fewer than expected). Consequently, we decided to focus on the
178 three candidates with mutants showing reproducibly reduced nodule numbers compared than controls
179 (*PHO2-like*, *PNO1-like*, and *PEN3-like*) for further analysis.

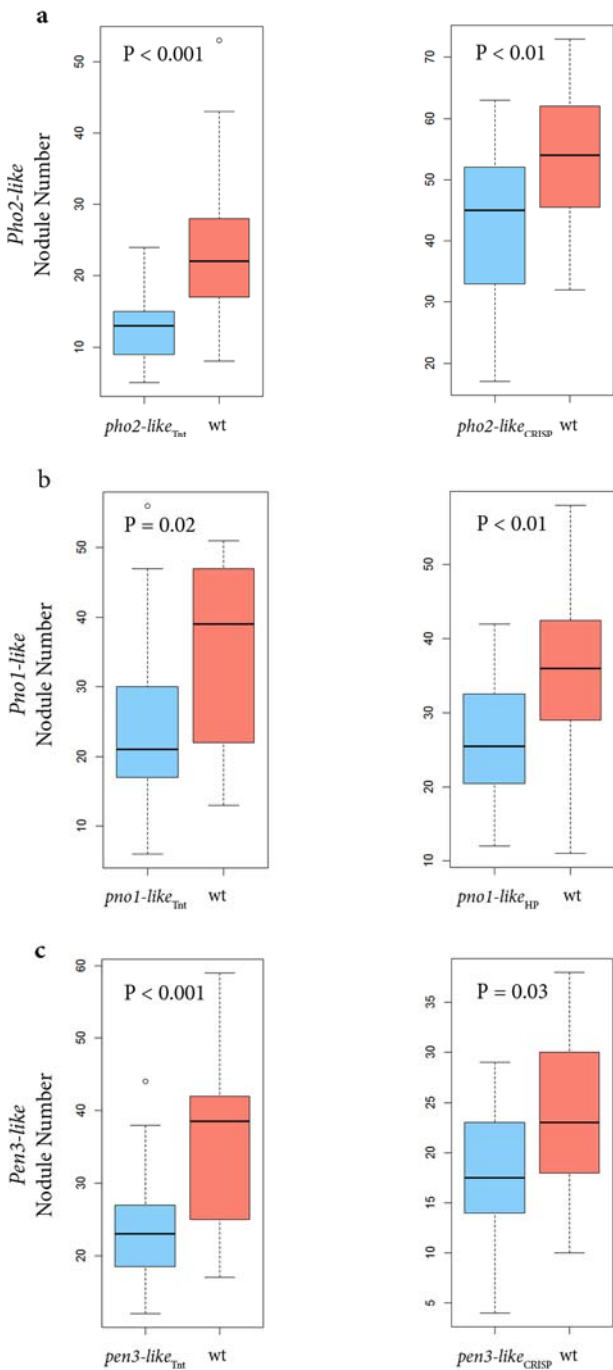
180

181 **Nitrogen content and nodule morphology of validated candidate lines**

182 Although plants with mutations in the three validated candidate genes (*PHO2-like*, *PNO1-like*,
183 *PEN3-like*) produced significantly fewer nodules, N concentration and nodule morphology were not
184 affected. Foliar nitrogen concentrations did not differ significantly between any of the mutants and their
185 control (t-tests, all $p_{df=4} > 0.1$, Supplemental Table S2). Moreover, longitudinal sections of nodules from
186 all validated mutant lines indicated they were structurally sound, with distinct meristematic, infection,
187 and pink nitrogen fixation zones (Supplemental Figs. S10, S11). Both *pho2-like_{tnt1}* mutants and their
188 corresponding WT controls had lighter pink nodules overall, however *pho2-like_{CRISPR}* nodules exhibited
189 similar color intensity compared with the HM340 control (Fig. S11). This suggests that differences in the
190 pink color of *Pho2-like_{tnt1}* is probably due to the Tnt1 background rather than disruption of the *PHO2-like*
191 gene itself.

192

Figure 4 Distribution of nodule number produced by mutant and wild



193 Effects of forward genetic mutants

194 To better understand the quantitative nature of nodulation mutants, we compared effects of the
 195 validated GWA mutants to nodulation mutants in genes originally discovered through forward genetics.
 196 Considering only “confirmed” candidates from our GWA study (*PHO2-like*, *PNO1-like*, *PEN3-like*), GWA
 197 mutants produced ~ 60% of the number of nodules on wild-type plants. These results were then

198 compared to phenotype assays on two well-known and previously characterized nodulation mutant
199 lines: the calcium and calmodulin-dependent kinase *dmi3* (Catoira et al., 2000) and *ipd3-2*, which
200 encodes for a protein that interacts with DMI3 (Messinese et al., 2007). Under exactly the same
201 experimental conditions used to test the GWA mutant lines, these forward genetics mutants exhibited
202 drastic reductions in nodule number – an average of just 0.5 nodules on *dmi3* lines and 1.5 nodules on
203 *ipd3* lines compared with 39 on WT plants ($P < 0.0005$, Fig. S8).

204

205

206

207 **DISCUSSION**

208 Genome wide association analyses provide an opportunity for identifying functionally important
209 genes that contribute to naturally occurring variation in quantitative traits, genes that are unlikely to be
210 detected by forward genetic screens. However, the approach is challenged by several statistical issues
211 including an expected high number of statistical false positives due to the large number of statistical
212 tests and the difficulty of fully eliminating confounding effects of relatedness among assayed individuals.
213 This means that GWA is best viewed as a statistical approach for identifying genomic regions that are
214 likely to contribute to variation. Moving beyond such a candidate status requires empirical validation of
215 gene function (Ioannidis et al., 2009). In this work, we applied a suite of mutagenesis platforms,
216 retrotransposon insertions, hairpin knock-down RNAi constructs, and CRISPR/Cas9, to disrupt the
217 function of ten genes that GWA had identified as candidates in contributing to variation in the legume-
218 rhizobia symbiosis through nodulation (Stanton-Geddes et al., 2013). Our evaluation experiments
219 revealed that mutations in three of the ten candidate lines produced significantly fewer nodules than
220 wild-type plants, thereby identifying new genes contributing to quantitative variation in nodule
221 production.

222 Three out of ten candidate genes were empirically validated in this study. However, there was more
223 than one candidate gene for some genomic intervals. Therefore, the validation rate is better
224 represented as three validated genes among the six intervals evaluated. These findings demonstrate
225 that GWA coupled with well-replicated validation experiments is an effective means to identify genes
226 contributing to quantitative variation. The high rate of success in validating the functional importance of
227 GWA candidates is likely due to at least two reasons. First, in creating an initial list of candidates for
228 functional evaluation, we did not limit our efforts to only those variants with the strongest statistical
229 support from the GWA. Rather, we considered the 100 SNPs with the strongest statistical associations
230 with phenotypic variation as potential candidates and then applied genetically informed criteria for
231 identifying which candidates to pursue for validation. These criteria included SNPs in or near annotated
232 genes, genes expressed in roots or nodules, and SNPs that formed clusters of statistical associations,
233 (peaks in the Manhattan plots) (Fig. 1).

234 Despite the criterion that chosen SNPs either were in or near a coding region, only four of the GWA
235 variants that we pursued were actually in annotated coding regions, with the remainder anywhere from
236 168 to 5819 bp from annotated CDS. Our data do not allow us to determine whether the SNPs near, but
237 not in, coding regions cause functionally important variation in gene expression as opposed to linkage-
238 disequilibrium (LD) with un-assayed SNPs in coding regions. Nevertheless, non-coding SNPs are often

239 associated with phenotypic variation (Hindorff et al., 2009) through effects on gene expression (Cookson
240 et al., 2009; Cubillos et al., 2012). We also prioritized candidates that showed expression in root or
241 nodule tissue, although such a criterion might have biased against lowly expressed genes (like many
242 transcription factors) or gene products expressed elsewhere in the plant affecting nodulation (Krusell et
243 al., 2002). The third criterion we applied, that candidates be physically proximate to other SNPs showing
244 strong associations, is likely to favor alleles that are maintained in contemporary populations through
245 balancing selection where the optimal trait value varies among populations or that have recently
246 increased in frequency (Barton and Turelli, 1989; Yeaman and Whitlock, 2011).

247 Also central to our evaluation experiments was the use of well-replicated experiments when
248 evaluating phenotypic effects. Each of the phenotypic assays we conducted involved an average of 22
249 mutant and 22 WT (minimum of 7 mutant and 6 wild-type) individuals. This was important because the
250 distributions of mutant and WT phenotypes often showed considerable overlap (Fig. 4a-c). The
251 importance of replication also helps to explain why the genes we identified had not been detected in
252 earlier forward-genetic screens conducted on nodulation in *M. truncatula* (Catoira et al., 2000; Oldroyd
253 and Long, 2003; Schnabel et al., 2005; Pislariu et al., 2012; Domonkos et al., 2013). If the mutations we
254 evaluated had been tested in only a single individual, the mutant phenotype would likely have fallen into
255 the range of phenotypes exhibited by WT plants generally and thus would not have been picked up in a
256 standard genetic screen. The phenotypic effects of genes identified through forward genetic screens
257 relative to those we validated here is highlighted by the comparison of *dmi3* and *ipd3* to the three GWA-
258 selected candidates: *dmi3* and *ipd3* plants produced an average of approximately 2% of the number of
259 nodules produced on WT plants whereas the validated genes first identified as candidates through GWA
260 produced roughly 60% of the number of nodules compared with WT plants.

261

262 **Genetics of nodulation**

263 Although we identified candidates without prior knowledge of their function, the molecular
264 functions of homologs for the validated candidates are known and all are interesting in the context of
265 nodulation. One of the validated candidates, *PHO-like* (Medtr4g020620), suggests a potentially
266 important role for phosphorous in quantitative variation in nodulation. *PHO2-like* is annotated as
267 encoding an E2 ubiquitin conjugating enzyme and its *Arabidopsis thaliana* ortholog (35% AA identity, E
268 value = 4e-104) is involved in phosphorous accumulation (Park et al., 2014). Not only can phosphorous
269 availability affect nodulation and N₂ fixation (Graham and Rosas, 1979; Jakobsen, 1985; Graham, 1992,
270 Sulleman 2013) but in *A. thaliana*, *AtPHO2* ortholog interacts directly with the E3 product of Nitrogen

271 Limitation Adaption (*Nla*), which mediates plant responses to N limitation (Park et al., 2014). Moreover,
272 E2 ubiquitin conjugate enzymes are presumed to interact with multiple E3 partners that could lead to
273 the proteolytic degradation of a larger range of uncharacterized substrates (Smalle and Vierstra, 2004).
274 *Medicago PHO2-like* is expressed at modest levels throughout nodule development (and also in
275 uninoculated roots), but reaches a distinct peak in expression at 20 dpi (Fig. S12). This is a timepoint
276 when nodules are fully differentiated, actively fixing nitrogen, and still prior to the initiation of
277 senescence.

278 A second interesting validated gene is the annotated ABC transporter (*PEN3-like*) (Medtr2g101090).
279 ABC transporters are involved in a wide range of plant activities, including pathogen metabolite
280 detoxification, non-host pathogen resistance, stomatal function, and Ca²⁺ sensing associated plant
281 immunity (Rea, 2007; Campe et al., 2016). The *A. thaliana* gene with greatest similarity to
282 Medtr2g101090 is *PEN3* (66% identical AA sequence across 97% of the gene length. E-value = 0), which
283 has been shown to be an essential component of salicylic acid-related defense against biotrophic fungi
284 (Stein et al., 2006; Johansson et al., 2014). The formation of nodules requires direct interaction with
285 bacteria and there is growing evidence that defense responses can negatively affect nodule formation
286 (Ramu et al., 2002; Lopez-Gomez et al., 2012; Larrainzar et al., 2015). Interestingly, ABC transporters
287 have been shown to be significantly and highly up-regulated in several nodulation mutants relative to
288 wild-type controls (Lang and Long, 2015). *PEN3-like* is strongly expressed in *Medicago* nodules
289 throughout development, though even more highly expressed in roots, with a slight peak in nodule
290 expression at 14 dpi (Fig. S12).

291 The final validated candidate is annotated as encoding an RNA-binding *PNO1-like* protein
292 (Medtr5g088410). How this protein might affect nodulation is not clear, although recent work has
293 identified roles for members of this gene family in the regulation of miRNAs (Karlsson et al., 2015),
294 suggesting that Medtr5g088410 could affect the expression of genes involved in nodulation. It is
295 expressed at a comparatively low level in both nodules and roots with no distinct peak in time of
296 expression (Fig. S12).

297

298

299 **Mutation platforms for GWA validation**

300 Efficient means to validate the functional importance of GWA candidate genes are necessary for
301 GWAS to be effective in discovering genes of subtle effects, fulfilling the promise of linking molecular
302 genetics with quantitative trait variation. Here we applied multiple genomic technologies

303 (retrotransposon mutagenesis, hairpins, and CRISPR/Cas9) to make this link. Each of these technologies
304 has advantages and disadvantages when being applied as a reverse genetics tool. The *M. truncatula*
305 Tnt1 collection has the advantage of ease: the Noble Foundation in Ardmore, OK has over 21,000
306 mutant lines that have largely saturated the genome with random Tnt1 insertion events and the
307 collection is easily searchable online. Although easy to use for validation experiments, there is the
308 problem of background mutations (e.g., Fig. S9). On average, there are 25 insertions in each Tnt1 lines
309 and as many as 97 independent Tnt1 loci have been reported (Veerappan et al., 2016). These
310 background mutations can make phenotyping data potentially unreliable, complicating their use as the
311 sole validation tool in reverse genetics phenotyping experiments. In our experiments, one Tnt1 line
312 expressed an (unexpected) super-nodulation phenotype, while two Tnt1 lines produced more nodules
313 than control plants, though neither of these phenotypes could be validated with a second (non-Tnt1)
314 mutation.

315 Both hairpin and CRISPR require more effort to create, involving construction of vectors, infecting
316 plants with *Agrobacterium*, and regeneration of whole plants from callus. Despite these challenges, we
317 had a high rate of success in creating stable whole-plant transformants using both hairpin and
318 CRISPR/Cas9 at high efficiency and frequency. For example, CRISPR/Cas9 targeting the *PHO2-like* and
319 *PEN3-like* genes induced homozygous mutations in multiple independent T₀ plants (Fig S3- S7).

320 Hairpin technologies seem promising in two situations: when genes of interest are members of
321 multi-gene families where gene redundancy might mask phenotype effects and for genes that are
322 gametophytic or homozygous lethal. This is illustrated by the *ERDJ2* candidate in our mutant screening.
323 We identified a CRISPR/Cas9 *ERDJ2* mutant with an in-frame 3-bp and a frame-shift 85-bp mutant alleles
324 and confirmed that the gene is homozygous lethal by segregation analysis. PCR analysis of
325 approximately 30 *erdj2*_{CRISPR_3/85} mutants failed to identify plants with both frame-shift 85-bp mutant
326 alleles, suggesting that the combination of frame shift mutations results in a homozygous lethal
327 phenotype (Fig. S4). Nevertheless, we successfully created a hairpin mutant that decreased transcript
328 expression without the lethality of a knock-out to use in candidate validation. For some genes, however,
329 “knocking down” rather than “knocking out” expression could also limit the phenotypic effects of the
330 mutations.

331 CRISPR technologies have advanced rapidly in recent years and hold promise for altering gene
332 sequence and expression in many systems including plants (Baltus and Voytas, 2015; Luo et al., 2016).
333 Although producing stable germ-line whole-plant mutants has proven challenging in some species,
334 including *A. thaliana* (Fauser et al., 2014; Feng et al., 2014), we were able to produce stable mutant lines

335 with high efficiency. The mechanism behind this success rate is not clear, although we suspect that the
336 transformation method might be a factor. *Medicago* transformation generates callus tissue that
337 undergoes strong herbicide selection in contrast to floral dip transformation that requires infection of
338 individual embryos. Moreover, we were often able to remove the CRISPR transgene from the mutant
339 lines by genetic segregation to obtain non-transgenic mutant plants that can be used for future
340 complementation studies. Given the rapid development of CRISPR multi-plex platforms, it should soon
341 be feasible to mutate and test large numbers of GWA identified candidates. Doing so would allow
342 researchers to move beyond validating small numbers of candidates and allow robust empirical
343 evaluation of the genes that control variation underlying complex quantitative traits.

344

345 **MATERIALS AND METHODS**

346

347 **Insertional mutant screening and genotyping**

348 To identify mutants in the 10 candidates selected for validation, we first searched for Tnt1
349 insertional mutants using the Noble Foundation's *M. truncatula* Mutant Database ([http://medicago-](http://medicago-mutant.noble.org/mutant/blast/blast.php)
350 [mutant.noble.org/mutant/blast/blast.php](http://medicago-mutant.noble.org/mutant/blast/blast.php)) (Tadege et al., 2008; Pislariu et al., 2012). Searches of this
351 database, which is comprised of more than 21,000 mutants, identified insertions in five of seven
352 candidates with the remaining two candidates identified by a PCR reverse-screen performed at the
353 Noble Foundation (Table S1). Seeds for Tnt1 mutant lines were obtained from the *M. truncatula* Tnt1
354 collection, planted, and genotyped for homozygous and wild-type (null) insertion alleles at the locus of
355 interest. Mutant plants for *Erdj2*, *Acre1* and *Mel1* candidates could not be identified either from the
356 Tnt1 database or from the PCR-reverse screen and we therefore relied upon the CRISPR/Cas9 platform
357 to generate these mutants.

358

359 **Hairpin and CRISPR/Cas9 transgenes and whole plant transformation**

360 We attempted to create hairpin constructs to reduce gene expression in all 10 candidates and
361 successfully identified homozygous hairpin transgenic plant lines for three: *FBL1-like*, *PNO1-like*, and
362 *ERDJ2*. To create hairpins, targeted coding sequences were amplified by PCR using primer sequences
363 (Table S3) from cDNA derived from *M. truncatula* accession R108. These amplicons were cloned into the
364 pCR8/GW/TOPO entry vector (Life Technologies, CA) and incorporated into the hairpin destination
365 vectors pSC218GH (see below) by a Gateway™ LR clonase reaction.

366 Finally, we generated CRISPR/Cas9 SSN to introduce frame-shift mutations in all 10 candidates and
367 recovered homozygous mutants for seven. Target sequences were identified by BLAST searches of the
368 *Medicago* R108 (HM340) genome assembly (www.medicagohapmap.org/downloads/r108) using the Mt
369 4.0 reference sequence as a query. Resulting sequences were then used to query the sgRNA designer v.1
370 for selection of target guide RNAs (www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design-v1)
371 (Doench et al., 2014). The targets were created by a primer annealing assay and cloned into the *Esp3I*
372 and *BsaI* sites of the U6/At7SL guide RNA entry vector (pAH595) (Table S3). The completed guide RNA
373 entry vector sequence was confirmed and combined with a second entry vector that harbored the
374 *Arabidopsis* optimized Cas9 cassette (pNJB184) by a multi-site Gateway LR clonase reaction into the
375 pSC218GG destination vector (see below) for *Agrobacterium* transformation (Li et al., 2013; Baltes et al.,
376 2014). A detailed protocol can be found in the Supplemental Methods.

377 Whole plant transformation of R108 plants, the same genetic background used in the *M. truncatula*
378 Tnt1 collection, with hpRNA and CRISPR/Cas9 constructs was carried out using a slightly modified
379 version of an established protocol (Cosson et al., 2006). In brief, transformation was carried out by
380 inoculating leaf tissue explants with *A. tumefaciens* strain EAH105 transformed with either a hairpin or
381 CRISPR SSN. The transformations were conducted using a modified version of the binary vector
382 backbone pNB96 (pSC218) (Curtin et al., 2011) with an introduced inverted-repeat cassette to facilitate
383 cloning of hairpin RNA fragments (pSC218GH) and a reading-frame cassette to incorporate both the
384 Cas9 and target guide RNA entry vectors (pSC218GG). The inverted-repeat and reading-frame cassettes
385 were derived from previously reported pTDT-DC-RNAi and pTDTO vectors (Valdes-Lopez et al., 2008)
386 and were driven by the constitutively expressed *Glycine max* Ubiquitin promoter (Hernandez-Garcia et
387 al., 2010).

388 After co-cultivating the leaf tissue with *Agrobacterium* for three days, explants were washed in an
389 antibiotic liquid media and transferred to a media for callus induction with bi-weekly transfers. After
390 approximately six weeks, callus tissue was transferred to a shoot induction media and incubated in 16:8
391 light:dark hour photoperiod. Developing shoots were transferred to a rooting media to encourage root
392 development before planting to soil. Both binary vectors pSC218GH and pSC218GG contain the 35S::BAR
393 cassette for phosphinothricin herbicide selection of transgenic plants. T₀ plants were rapidly screened
394 using a previously reported DNA extraction method and PCR-CAPS assay to identify target mutations
395 (Curtin et al., 2011).

396

397 **Nodulation assays**

398 To evaluate the effects of candidate gene mutations on nodulation we grew mutant and wild-type
399 plants in well-replicated (an average of 22 mutant and 22 wild type plants per assay, see Table S2 for
400 details) in completely-randomized experiments. Prior to planting, seeds were treated in concentrated
401 sulfuric acid for 5 min, washed with sterile dH₂O, and then kept in the dark on wet filter paper at 4°C for
402 7–10 days before planting into 650 ml conetainers with 4:1 parts autoclaved sand:perlite mixture. Five
403 days after planting, seedlings were inoculation with a 30mL suspension (OD₆₀₀ ~0.015) of *S. meliloti*
404 strain Sm2011 carrying the *hemA::LacZ* reporter (Ardourel et al., 1994). Plants were grown at 22-25°C,
405 75 % humidity and 200-350 μmolm⁻²s⁻¹ 16h-8 h photoperiod, fertilized 15 and 21 days post inoculation
406 with Fährus nutrient solution, and otherwise watered as needed. Thirty-five to 41 days after
407 inoculation, plants were harvested and nodule number was determined. Each mutant was compared
408 against a separate set of WT control plants. To reduce the effects of background insertions, controls for
409 the Tnt1 mutants were grown from seeds from a full-sib of the mutant plants. R108 seeds were used for
410 the control treatment for hairpin and CRISPR lines. For each assay experiment, we tested for statistically
411 significant differences between the phenotype of the mutant line and the phenotype of the wild-type
412 control line using a t-test using R (R Development Core Team, 2013). For the six candidates (*PHO2-like*,
413 *PNO1-like*, *FMO1-like*, *ERDJ2*, *HLZ1-like* and *PEN3-like*) where we evaluated two independent mutation
414 lines, we also used ANOVA to test for between-line differences in nodulation. Similarly, for the two
415 mutant lines (*pho2*_{CRISPR} *mel1*_{CRISPR}) for which the phenotypic assays were repeated, we used ANOVA to
416 test for differences between mutant and wild-type lines between assays. All data are available at
417 datadryad.org (dx.doi.org xxxxxxxx).

418

419 **Nodule morphology and nitrogen content**

420 To complement phenotypic data collected during initial validation experiments, we also grew the
421 three mutants exhibiting validated nodulation phenotypes (*PHO2-like*, *PNO1-like*, and *PEN3-like*) to test
422 for differences in foliar N concentration and to examine nodule morphology. These plants were grown
423 under the same conditions as those for initial phenotype assays. At 14 dpi, three plants per treatment
424 were harvested in order to make visual observations of nodules. Plants were washed to remove
425 surrounding substrate and 10 cm of roots with nodules were isolated to obtain close up pictures of
426 nodules (Figure S10). One nodule (second or third from the top) was then harvested from each plant,
427 sliced in half longitudinally by hand, and then observed under a stereo microscope (Figure S11). At 28
428 dpi, four plants grown in parallel were harvested for N concentration assays, which were conducted by
429 the University of Minnesota Research Analytical Lab using the Dumais method (Matejovic, 1995).

430
431
432
433

434 **Table 1.** Genome Location, GWAS statistics, and Expression of Candidate Genes

435

Mt4.0 Designation	Candidate Gene	Rank	P-value	GWAS	Expression	Confirmed
Medtr2g020630	Flavin-containing Monooxygenase-like (FMO1-LIKE)	41	8.14E-06	Mt3.5	Root & nodule specific	
Medtr2g044570	Ring finger protein1-like (RFP1-like)	8	1.62E-07	Mt3.5	All tissues	
Medtr2g044580	Endoplasmic reticulum DNAJ protein (ERDJ2)	5	1.30E-07	Mt3.5	Nodule upregulated	
Medtr2g101040	Malic enzyme-like1 (MEL1)	15	2.22E-07	Mt4.0	All tissues	
Medtr2g101090	Penetration3-like (PEN3-like)	1	2.56E-08	Mt4.0	Root & nodule upregulated	YES
Medtr2g101120	AVR9/CF-9 rapidly elicited protein 1 (ACRE1)	23	3.55E-07	Mt4.0	All tissues	
Medtr2g101190	Homeodomain leucine zipper1-like (HLZ1-like)	72	9.65E-07	Mt4.0	All tissues	
Medtr4g020620	Ubiquitin conjugate24 like (PHO2-like)	12	5.58E-06	Mt3.5	Nodule upregulated	YES
Medtr5g088410	Partner of NOB1-like (PNO1-like)	36	7.08E-06	Mt3.5	Root & nodule specific	YES
Medtr8g060730	F-box like 1 (FBL1-like)	2	4.26E-08	Mt3.5	Below detection limit	

436

437

438

439 FIGURE LEGENDS

440 **Figure 1. GWAS of nodulation candidate regions in *Medicago*.** From left to right, Manhattan
441 plots extending 300,000 bp either side of the candidate SNP for each of the six candidate gene
442 regions.

443 **Figure 2. Mutagenesis platforms. a.** Tnt1 retrotransposon insertional mutagenesis in *Medicago*
444 was developed in the R108 (HM340) accession. The original transformant harbored five Tnt1
445 copies. The Tnt1 element can transpose by a copy and paste mechanism during callus
446 regeneration leading to progeny lines with random stable insertions. **b.** Schematic
447 representations of the hairpin entry and destination vectors. Typically, a 300-500 bp amplicon
448 from the candidate gene of interest is amplified from *Medicago* whole-plant cDNA template
449 and cloned into the pCR8/GW entry vector. The entry vector is then used to clone the hairpin
450 fragment into the pSC218GH binary vector, followed by *Agrobacterium* transformation. **c.** The
451 construction of the CRISPR/Cas9 reagent requires a multi-site Gateway cloning reaction utilizing
452 two entry vectors and a destination binary vector. The CRISPR targets are cloned into pAH595,
453 and along with the Cas9 entry vector (pNB184) are cloned into pSC218GG vector, followed by
454 *Agrobacterium* transformation. A detail protocol of the CRISPR/Cas9 assembly is provided in the
455 Supplemental Methods.

456

457 **Figure 3. Screening and identification of candidate mutant alleles. a.** Two independent *PHO2-*
458 *like* mutant alleles, *pho2-like*_{Tnt} and *pho2-like*_{CRISPR}. Seeds from the Tnt1 line NF12360 were
459 screened by PCR assay to determine homozygous and null insertions at the *PHO2-like* locus,
460 identifying null insertion NF12360-8 plant and homozygous mutant NF12360-7 (Supplemental
461 Figure S9). A second mutant allele for *PHO2-like* was identified using plants transformed with a
462 CRISPR/Cas9 reagent that targeted the open-reading frame of *PHO2-like*. Several homozygous
463 T₀ mutant plants were identified and the WPT210-9 plant was selected for downstream
464 phenotype analyses. WPT210-9 had a 1-bp deletion in the coding region of both alleles,
465 effectively disrupting the reading frame as indicated by the three stop codons depicted in the
466 amino acid sequence. Heritable transmission of the 1-bp mutation was confirmed by screening
467 T1 plants by a PCR-digestion assay and sequencing. **b.** Two independent *PNO1-like* mutant
468 alleles, *pno1-like*_{Tnt} and *pno1-like*_{HP}. Tnt1 mutant alleles were identified for *PNO1-like* by
469 screening seed from the NF11825 line. *Pno1-like* mutant and null insertion lines were identified,
470 NF11825-8 and NF11825-5 respectively. Since an insufficient number of seed was recovered
471 from the NF11825-8 plant for phenotype analysis, seed from the next generation (NF11825-8-5)
472 was used as wild-type plant. For the second mutant allele, a hairpin line was identified. TAIL

473 PCR was used to identify the genomic location of the transgene and a PCR assay using both the
474 genome location and transgene specific primers were used to identify a plant homozygous for
475 the hairpin transgene (WPT118-4-1). **c**, Two independent *PEN3-like* mutant alleles, *pen3-like*_{Tnt}
476 and *pen3-like*_{CRISPR}. Seed from the NF2826 Tnt1 line were used to identify homozygous (NF2826-
477 4) and null (NF2826-2) insertion mutant plants using both *PEN3-like* and Tnt1 specific primers.
478 For the second mutant allele, we identified several CRISPR/Cas9 homozygous T₀ mutant plants
479 such as WPT237-3 which had a segregating a bi-allelic mutation consisting of a 11bp deletion
480 and a 1bp insertion. Screening of WPT237-3 T1 mutant plants confirmed heritable transmission
481 of both mutations.

482
483 **Figure 4. Nodule number produced by mutant and wild type plants.** Nodules were counted 32-
484 42 days post inoculation. Data are shown for two independent mutants for each of three
485 candidate genes. *a*, *PHO2-like*, *b*, *PNO1-like*, and *c*, *PEN3-like*. For each figure, the horizontal
486 black bar represents the median value, the box the middle quartiles, and the whiskers the range
487 of data, except for outliers shown as a single data point.

488
489
490
491
492
493
494

495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. The characterization of three Tnt1 mutant lines

Supplemental Figure S2. The characterization of two whole-plant hairpin lines

Supplemental Figure S3. The characterization of the Pho2-like CRISPR/Cas9 mutants

Supplemental Figure S4. The characterization of the Erdj2 CRISPR/Cas9 mutants

Supplemental Figure S5. The characterization of the Acre-1 CRISPR/Cas9 mutants

Supplemental Figure S6. The characterization of the Pen3-like CRISPR/Cas9 mutants

Supplemental Figure S7. The characterization of the Fmo1-like, Hlz-1 and Mel1-like CRISPR/Cas9 mutants

Supplemental Figure S8. Nodulation phenotype of dmi3 and ipd3-2

Supplemental Figure S9. The Tnt1 insertion background of the NF12360 Tnt1 mutant line

Supplemental Figure S10. WT and Pho2-like, Pen3-like, and Pno1-like mutant nodules at 14 dpi

Supplemental Figure S11. WT and Pho2-like, Pen3-like, and Pno1-like mutant dissected nodules at 14 dpi

Supplemental Figure S12. The expression analysis of validated genes, derived from the Medicago truncatula Gene Expression Atlas (<http://mtgea.noble.org/v3/index.php>)

Supplemental Table S1. List of Tnt1 mutant lines obtained from Noble Foundation Tnt1 mutant database

Supplemental Table S2. Statistical tests comparing phenotype between mutant and wild-type control plants

Supplemental Table S3. Primer sequences for PCR assays and vector construction (SEPARATE EXCEL FILE)

Supplemental Methods 1. CRISPR reagent design and construction

ACKNOWLEDGMENTS

535 We thank Catalina Pislariu, Jianqi Wen, Sarah Hoerth for assistance with mutant characterization,
536 Nathan Henning for plant care assistance, Aaron Hummel for the pAH595 vector, Ryan Morrow for
537 technical assistance with plant transformation and media preparation and Jean-Michel Ané for *dmi3* and
538 *ipd3* mutant seeds.

539

540

541

542

Parsed Citations

Ardourel M, Demont N, Debellé F, Maillet F, de Billy F, Promé JC, Dénarié J, Truchet G (1994) Rhizobium meliloti lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. Plant Cell 6: 1357-1374

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Baltes NJ, Gil-Humanes J, Cermak T, Atkins P a, Voytas DF (2014) DNA replicons for plant genome engineering. Plant Cell 26: 151-63

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Baltes NJ, Voytas DF (2015) Enabling plant synthetic biology through genome engineering. Trends Biotechnol 33: 120-131

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Barton NH, Turelli M (1989) Evolutionary Quantitative Genetics: How Little Do We Know? Annu Rev Genet 23: 337-370

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Campe R, Langenbach C, Leissing F, Popescu G V., Popescu SC, Goellner K, Beckers GJM, Conrath U (2016) ABC transporter PEN3/PDR8/ABCG36 interacts with calmodulin that, like PEN3, is required for Arabidopsis nonhost resistance. New Phytol 209: 294-306

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Catoira R, Galera C, de Billy F, Penmetsa R V, Journet EP, Maillet F, Rosenberg C, Cook D, Gough C, Dénarié J (2000) Four genes of Medicago truncatula controlling components of a nod factor transduction pathway. Plant Cell 12: 1647-66

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M (2009) Mapping complex disease traits with global gene expression. Nat Rev Genet 10: 184-194

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cosson V, Durand P, D'Erfurth I (2006) Medicago truncatula transformation using leaf explants. Methods Mol Biol. 343: 115-127

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cubillos F, Coustham V, Loudet O (2012) Lessons from eQTL mapping studies: non-coding regions and their role behind natural phenotypic variation in plants. Curr Opin Plant Biol 15: 192-198

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, Reyon D, Dahlborg EJ, Goodwin MJ, Coffman AP, et al (2011) Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol 156: 466-73

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32: 1262-7

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Domonkos A, Horvath B, Marsh JF, Halasz G, Ayaydin F, Oldroyd GED, Kalo P (2013) The identification of novel loci required for appropriate nodule development in Medicago truncatula. BMC Plant Biol 13: 157

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Fauser F, Schimi S, Puchta H (2014) Both CRISPR/Cas9-based nucleases and nickases can be used efficiently for genome

engineering in *Arabidopsis thaliana*. Plant J 79: 348-359

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang D-L, Wang Z, Zhang Z, Zheng R, Yang L, et al (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. Proc Natl Acad Sci U S A 111: 4632-4637

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Friesen ML, von Wettberg EJ, Badri M, Moriuchi KS, Barhoumi F, Chang PL, Cuellar-Ortiz S, Cordeiro MA, Vu WT, Arraouadi S, et al (2014) The ecological genomic basis of salinity adaptation in Tunisian *Medicago truncatula*. BMC Genomics 15: 1160

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Graham PH (1992) Stress tolerance in *Rhizobium* and *Bradyrhizobium*, and nodulation under adverse soil conditions. Can J Microbiol 38: 475-484

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Graham PH, Rosas JC (1979) Phosphorus fertilization and symbiotic nitrogen fixation in common bean. Agron J 71:925-926

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hernandez-Garcia CM, Bouchard R a, Rushton PJ, Jones ML, Chen X, Timko MP, Finer JJ (2010) High level transgenic expression of soybean (*Glycine max*) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. BMC Plant Biol 10: 237

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci 106 : 9362-9367

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Ioannidis JPA, Thomas G, Daly MJ (2009) Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 10: 318-329

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Jakobsen I (1985) The role of phosphorus in nitrogen fixation by young pea plants (*Pisum sativum*). Physiol Plant 64: 190-196

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Johansson ON, Fantozzi E, Fahlberg P, Nilsson AK, Buhot N, T??r M, Andersson MX (2014) Role of the penetration-resistance genes PEN1, PEN2 and PEN3 in the hypersensitive response and race-specific resistance in *Arabidopsis thaliana*. Plant J 79: 466-476

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Karlsson P, Christie MD, Seymour DK, Wang H, Wang X, Hagemann J, Kulcheski F, Manavella PA (2015) KH domain protein RCF3 is a tissue-biased regulator of the plant miRNA biogenesis cofactor HYL1. Proc Natl Acad Sci U S A 1-6

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F, et al (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420: 422-426

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Lang C, Long SR (2015) Transcriptomic analysis of *Sinorhizobium meliloti* and *Medicago truncatula* symbiosis using nitrogen fixation-deficient nodules. Mol Plant Microbe Interact 28: 856-68

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Larrainzar E, Riely B, Kim SC, Carrasquilla-Garcia N, Yu H-J, Hwang H-J, Oh M, Kim GB, Surendrarao A, Chasman D, et al (2015) Deep sequencing of the *Medicago truncatula* root transcriptome reveals a massive and early interaction between Nod factor and

ethylene signals. *Plant Physiol.* doi: 10.1104/pp.15.00350

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol* 31: 688-91

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Lopez-Gomez M, Sandal N, Stougaard J, Boller T (2012) Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. *J Exp Bot* 63 : 393-401

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Luo M, Gilbert B, Ayliffe M (2016) Applications of CRISPR/Cas9 technology for targeted mutagenesis, gene replacement and stacking of genes in higher plants. *Plant Cell Rep* 35: 1439-1450

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10: 565-577

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Matejovic I (1995) Total nitrogen in plant material determined by means of dry combustion: A possible

alternative to determination by Kjeldahl digestion. *Commun Soil Sci Plant Anal* 26: 2217-2229

Messinese E, Mun J, Yeun LH, Jayaraman D, Rougé P, Lougnon G, Schornack S, Bono J, Cook DR, Ané J (2007) A novel nuclear protein interacts with the symbiotic DMI3 Calcium- and Calmodulin- dependent protein kinase of *Medicago truncatula*. *Mol Plant Microbe Interact* 20: 912-921

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Oldroyd GED, Murray JD, Poole PS, Downie JA (2011) The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* 45: 119-44

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Oldroyd GED, Long SR (2003) Identification and characterization of Nodulation-Signaling Pathway 2, a gene of *Medicago truncatula* involved in Nod Factor Signaling. *Plant Physiol* 131: 1027-1032

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Park BS, Seo JS, Chua N-H (2014) NITROGEN LIMITATION ADAPTATION recruits PHOSPHATE2 to target the phosphate transporter PT2 for degradation during the regulation of *Arabidopsis* phosphate homeostasis. *Plant Cell* 26: 454-464

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Pislariu CI, Murray JD, Wen J, Cosson V, Muni RRD, Wang M, Benedito V a, Andriankaja A, Cheng X, Jerez IT, et al (2012) A *Medicago truncatula* tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. *Plant Physiol* 159: 1686-99

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Popp C, Ott T (2011) Regulation of signal transduction and bacterial infection during root nodule symbiosis. *Curr Opin Plant Biol* 14: 458-467

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. R Found. Stat. Comput. Vienna, Austria.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Ramu SK, Peng H-M, Cook DR (2002) Nod Factor induction of reactive oxygen species production is correlated with expression of the Early Nodulin gene rip1 in *Medicago truncatula*. *Mol Plant-Microbe Interact* 15: 522-528

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Rea PA (2007) Plant ATP-binding cassette transporters. Annu Rev Plant Biol 58: 347-375

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Schnabel E, Journet EP, De Carvalho-Niebel F, Duc G, Frugoli J (2005) The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. Plant Mol Biol 58: 809-822

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Smalle J, Vierstra RD (2004) The ubiquitin 26S proteasome proteolytic pathway. Annu Rev Plant Biol 55: 555-90

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Stanton-Geddes J, Paape T, Epstein B, Briskine R, Yoder J, Mudge J, Bharti AK, Farmer AD, Zhou P, Denny R, et al (2013) Candidate genes and genetic architecture of symbiotic and agronomic traits revealed by whole-genome, sequence-based association genetics in Medicago truncatula. PLoS One 8: e65688

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Stein M, Dittgen J, Sánchez-Rodríguez C, Hou B-H, Molina A, Schulze-Lefert P, Lipka V, Somerville S (2006) Arabidopsis PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration. Plant Cell 18 : 731-746

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Suliman S, Ha CV, Schulze J, Tran LS (2013) Growth and nodulation of symbiotic Medicago truncatula at different levels of phosphorus availability. J Exp Bot 64: 2701-2712

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M, et al (2008) Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume Medicago truncatula. Plant J 54: 335-347

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, et al (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15: 1-14

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Valdes-Lopez O, Arenas-Huertero C, Ramirez M, Girard L, Sanchez F, Vance CP, Luis Reyes J, Hernandez G (2008) Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots. Plant Cell Environ 31: 1834-1843

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Vance CP (2001) Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. Plant Physiol 127 : 390-397

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Veerappan V, Jani M, Kadel K, Troiani T, Gale R, Mayes T, Shulaev E, Wen J, Mysore KS, Azad RK, et al (2016) Rapid identification of causative insertions underlying Medicago truncatula Tnt1 mutants defective in symbiotic nitrogen fixation from a forward genetic screen by whole genome sequencing. BMC Genomics 1-11

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wesley S, Helliwell C, Smith N (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J 27: 581-590

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. Evolution 65: 1897-1911

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Young ND, Debelle F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, et al (2011)
The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520-524

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)