Transcriptome dynamics at the *Arabidopsis* graft junction reveal an inter-tissue recognition mechanism that activates vascular regeneration

Charles W Melnyk[1,2], Alexander Gabel[3], Thomas J Hardcastle[4], Sarah Robinson[5], Shunsuke Miyashima[6], Ivo Grosse[3], Elliot M Meyerowitz[1,7]

1- Sainsbury Laboratory, University of Cambridge, Bateman Street, Cambridge CB2 1LR, UK

2 – Department of Plant Biology, Swedish University of Agricultural Sciences, Almas allé 5, 756 51 Uppsala, Sweden

3 - Institute of Computer Science, Martin Luther University Halle–Wittenberg, 06120 Halle (Saale), Germany

4 – Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, United Kingdom

5 - Institute of Plant Science, University of Bern, Altenbergrain 21, 3013 Bern, Switzerland

6 – Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

7 – Howard Hughes Medical Institute and Division of Biology and Biological Engineering, California Institute of Technology, 1201 East California Boulevard, Pasadena, CA 91125, USA

Contact: charles.melnyk@slu.se

ABSTRACT

The ability for cut tissues to join together and form a chimeric organism is a remarkable property of many plants, however, grafting is poorly characterized at the molecular level. To better understand this process we monitored genome-wide temporal and spatial gene expression changes in grafted *Arabidopsis thaliana* hypocotyls. Tissues above and below the graft rapidly developed an asymmetry such that many genes were more highly expressed on one side than the other. This asymmetry correlated with sugar responsive genes and we observed an accumulation of starch above the graft that decreased along with asymmetry once the sugar-transporting vascular tissues reconnected. Despite the initial starvation response below the graft, many genes associated with vascular formation were rapidly activated in grafted tissues but not in cut and separated tissues indicating that a recognition mechanism activated that was independent of functional vascular connections. Auxin which is transported cell-to-cell, had a rapidly elevated response that was symmetric, suggesting that auxin was perceived by the root within hours of tissue attachment to activate the vascular regeneration process. A subset of genes were expressed only in grafted tissues, indicating that wound healing proceeded via different mechanisms depending on the presence or absence of adjoining tissues. Such a recognition process could have broader relevance for tissue regeneration, inter-tissue communication and tissue fusion events.

INTRODUCTION

For millennia people have cut and rejoined plants through grafting. Generating such chimeric organisms combines desirable characteristics from two plants, such as disease resistance, dwarfing and high yields, or can propagate plants and avoid the delays entailed by a juvenile state (1). Agriculturally, grafting is becoming more relevant as a greater number of plants and species are grafted to increase productivity and yield (2). However, our mechanistic understanding of grafting and the biological processes involved, including wound healing, tissue fusion and vascular formation, remain limited.

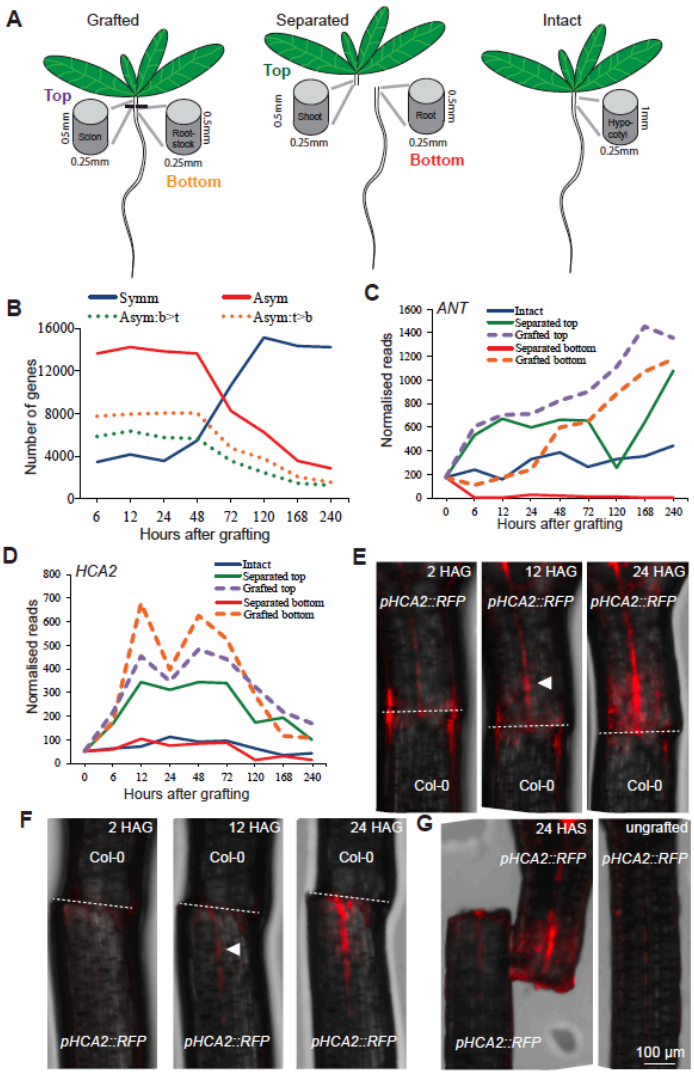
Plants have efficient mechanisms to heal wounds and cuts, in part through the production of wound-induced pluripotent cells termed callus. The callus fills the gap

69 or seals the wound, and later, differentiates to form epidermal, mesophyll and
70 vascular tissues (3). In grafted *Arabidopsis* hypocotyls, tissues adhere 1-2 days after
71 grafting and the phloem, the tissue that transports sugars and nutrients, connects after
72 three days (4, 5). The xylem, tissue that transports water and minerals, connects after
73 seven days (4). Plant hormones are important regulators of vascular formation, and at
74 the graft junction, both auxin and cytokinin responses increase in the vascular tissue
75 (4-6). Auxin is important for differentiation of vascular tissues whereas cytokinin
76 promotes vascular stem cells, termed cambium, to divide and proliferate in a process
77 known as secondary growth (7, 8). Auxin is produced in the upper parts of a plant and
78 moves towards the roots via cell-to-cell movement. Auxin exporters, including the
79 PIN proteins, transport auxin into the apoplast, whereas auxin importers, such as the
80 AUX and LAX proteins, assist with auxin uptake into adjacent cells (8). Disrupting
81 this transport, such as by mutating *PIN1*, inhibits healing of a wounded stem (9).
82 Blocking auxin transport with the auxin transport inhibitor TIBA (2,3,5-
83 triiodobenzoic acid) in the shoot inhibits vascular formation and cell proliferation at
84 the *Arabidopsis* graft junction (6). In addition to auxin, other compounds, including
85 sugars, contribute to vascular formation. The localised addition of auxin to callus
86 induces phloem and xylem but requires the presence of sugar (10, 11). In plants,
87 sugars are produced in the leaves and transported through the phloem to the roots
88 (12). The role of sugars in vascular formation and wound healing is not well
89 established, however, sugars promote cell division and cell expansion (13), processes
90 important for development including vascular formation.
91
92 The molecular and cellular mechanisms for wound healing, tissue reunion and graft
93 formation remain largely unknown. One emerging theme is that the top and bottom of
94 the cut do not behave similarly. Such tissue asymmetry occurs in other plant tissues,
95 most notably leaves. Developing leaf primordia have an inherent asymmetry that is
96 thought to derive from positional signals from the shoot apical meristem that specifies
97 differences between the top and the bottom of the leaf. The molecular mechanisms
98 that establish asymmetry are not well described, though one hypothesis is that auxin
99 contributes towards or acts on the putative signal (14). Asymmetry also appears in cut
100 *Arabidopsis* inflorescence stems where the transcription factor *RAP2.6L* expresses
101 exclusively below the cut whereas the transcription factor *ANAC071* expresses
102 exclusively above the cut (9). Both were important for stem healing and *ANAC071*

103 and a close homologue, *ANAC096*, were important for graft formation (6).
104 Asymmetry also exists in genetic requirements, since *ALF4* and *AXR1*, two genes
105 involved in auxin perception, are important below but not above the graft junction for
106 phloem connection (4). However, *ANAC071* is expressed symmetrically around the
107 hypocotyl graft junction three days after grafting (6) so the extent of asymmetry and
108 the mechanistic basis for it during wound healing remains largely uncharacterised.
109
110 Previous efforts have characterised wound healing and tissue reunion using
111 transcriptomic analyses. Mechanical wounding altered approximately 8% of the
112 *Arabidopsis* transcriptome and showed a high degree of overlap with transcriptomic
113 changes elicited by pathogen attack and abiotic stress (15). Stem wounding and
114 wound-induced callus formation altered the expression of hundreds or thousands of
115 genes (9, 16, 17), whereas grafting grape vines, lychee trees and hickory trees induced
116 hundreds or thousands of differentially expressed genes involved in hormone
117 response, wound response, metabolism, cell wall synthesis and signal transduction (5,
118 18-21). These grafting studies provide limited information, as tissues from above and
119 below the graft junction were not isolated to test whether these tissues behaved
120 differently, and controls were not performed to distinguish how grafting and tissue
121 fusion might differ from a response associated with cut tissues that remained
122 separated. Here, we perform an in-depth analysis to describe the spatial and temporal
123 transcriptional dynamics that occur during healing of cut *Arabidopsis* tissues that are
124 joined (grafted) or left unjoined (separated). We find that the majority of genes
125 differentially expressed are initially asymmetrically expressed at the graft junction
126 and that many of these genes are sugar responsive, which correlates with severing of
127 the phloem tissue and the accumulation of starch above the junction. However, genes
128 associated with cell division and vascular formation activate on both sides of the graft
129 and, similarly, auxin responsiveness activates equally on both sides. We propose that
130 the continuous transport of substances, including auxin, independent of functional
131 vascular connections, promoted division and differentiation, while the enhanced auxin
132 response and blocked of sugar transport provided a unique physiological condition to
133 activate genes specific to graft formation that promote wound healing.
134
135 **RESULTS**
136 **Genes are asymmetrically expressed around the graft**

137 Previous analyses revealed an asymmetry in several genes important for tissue
138 reunion or graft formation and identified a number of genes expressed above a cut
139 that were not expressed below (4, 9). To investigate whether asymmetry was a
140 common feature of grafting and tissue reunion, we generated RNA deep sequencing
141 libraries from *Arabidopsis thaliana* hypocotyl tissues immediately above and
142 immediately below the graft junction 0, 6, 12, 24, 48, 72, 120, 168 and 240 hours after
143 grafting (HAG)(Figure 1A). Prior to RNA extraction, we separated top and bottom
144 tissues at the graft junction. We found that the strength required to break apart the
145 graft junction increased linearly (Figure S1) similarly to previously reported breaking
146 strength dynamics of grafted *Solanum pennellii* and *Solanum lycopersicum* (22, 23).
147 When pulling apart grafts to separate top and bottom for sample preparation, grafts
148 broke cleanly with minimal tissue from one half present in the other half (Figure S1,
149 Movie S1, S2). We measured the amount of tissue from tops adherent to bottoms and
150 *vice versa* (Figure S1) and found less than 4% cross-contamination. In addition to
151 grafting, we also prepared libraries from ungrafted hypocotyls ("intact" treatment)
152 and cut plants that had not been reattached ("separated" treatment)(Figure 1A). We
153 herein refer to tissues harvested above the graft junction or from the shoot side of
154 separated tissue as "top" and that from below the graft or from the root side of
155 separated tissue as "bottom" (Figure 1A).
156
157 RNAs that were differentially expressed as a consequence of grafting (compared to
158 intact hypocotyls) equally in tops and bottoms of grafts (symmetrically expressed) or
159 were more highly expressed in one tissue than the other (asymmetrically expressed),
160 were identified by performing a pairwise analysis of the protein-coding transcriptome
161 datasets that were differentially expressed relative to the intact group. Several
162 thousand RNAs were identified that fit either pattern of expression including the
163 transcript of the cambial markers *HCA2* that was induced symmetrically, and *ANT*
164 that was induced asymmetrically (Figure 1B-G, Figure S2). 6 to 48 hours after
165 grafting, the number of graft-differentially expressed genes that were asymmetrically
166 expressed was roughly 3-fold greater than those symmetrically expressed indicating
167 that tissues above the cut changed their expression dynamics relative to below the cut.
168 However, at 72 hours the numbers were nearly equal, and by 120 hours, the number
169 of symmetrically differentially expressed genes was 3-fold greater than those
170 asymmetrically expressed (Figure 1B). As a second approach, we performed a

171 hierarchical clustering analysis that indicated that the grafted top and grafted bottom
172 became most similar after 72 hours (Figure S3), consistent with the symmetry
173 analysis (Figure 1B). Thus, graft healing and tissue reunion promoted a shift from
174 asymmetry to symmetry (Figure 1B).
175



176

177 **Figure 1. Grafting rapidly activates genes in both asymmetric and symmetric**
178 **patterns.** (A) Intact, cut and separated, or cut and grafted *Arabidopsis* tissues were
179 harvested approximately 0.5mm above (top), 0.5mm below (bottom) the cut site or for
180 intact plants 1mm segments spanning the region where cuts were made in grafted and
181 separated plants. (B) Pairwise analysis between the grafted top and grafted bottom
182 identified sets of protein-coding genes symmetrically or asymmetrically expressed at
183 the graft junction, with an FDR < 0.05 and a likelihood of symmetric/asymmetric
184 expression of greater than 50%. Asymmetrically expressed genes were further divided
185 into those whose RNAs were higher in the top (orange dotted line) or higher in the
186 bottom (purple dotted line), again with an FDR < 0.05 and a likelihood of asymmetric
187 expression greater than 50%. (C,D) Expression profiles for transcripts of cambium-
188 associated genes that initially activate symmetrically (*HCA*) or asymmetrically (*ANT*)
189 plotted for intact, separated and grafted samples. (E-G) *HCA2* transcription

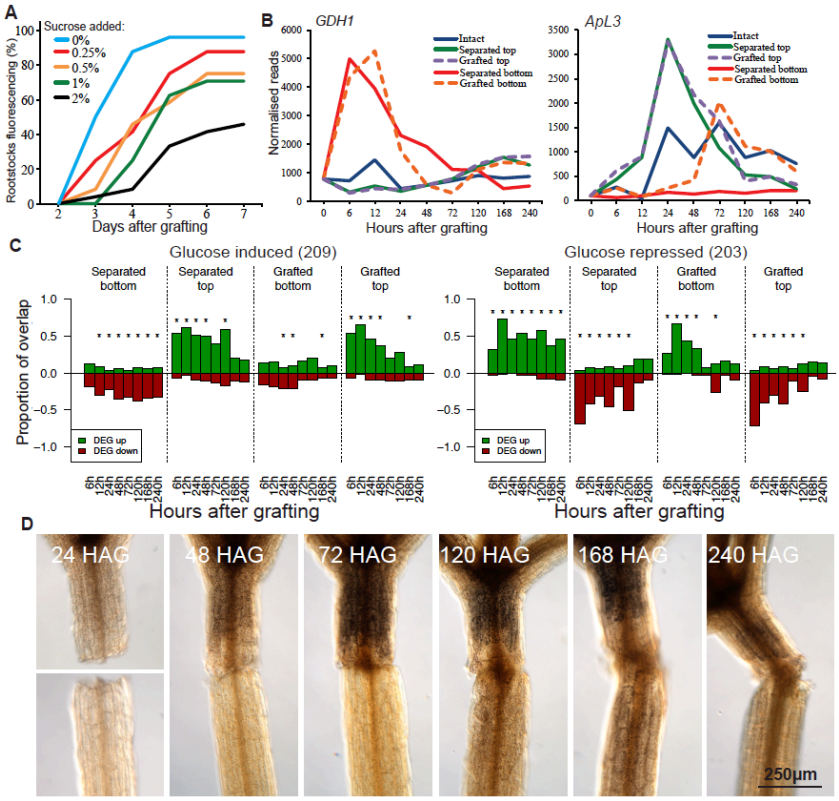
190 upregulates above the graft junction (E) and below the graft junction. (F)
191 *pHCA2::RFP* was grafted to Col-0 roots (E) or Col-0 shoots (F) to avoid ambiguity of
192 signal origin at the junction. *HCA2* was also upregulated in separated shoots but not in
193 intact samples or in high levels in separated roots (G). HAG, hours after grafting.
194 HAS, hours after separation. White triangle denotes initial fluorescent signal, dashed
195 lines denote the graft junction.

196

197 **Sugar response correlates with asymmetric gene expression**

198 The shift from asymmetry to symmetry could be due to phloem reconnection at 72
199 hours ((4), Figure 2A) and the resumption of hormone, protein and sugar transport.
200 We tested a role for sugar by grafting in the presence of exogenous sucrose which has
201 previously been reported to affect grafting success (24). Low levels of exogenous
202 sucrose lowered grafting efficiency (Figure 2A), suggesting that differential sugar
203 responses at the graft junction might be important for vascular reconnection.
204 Expression of *ApL3*, a gene whose expression is induced by sugar (25), was rapidly
205 upregulated in separated tops and grafted tops, whereas expression of *DIN6*, *GDH1*
206 and *STP1*, genes whose expression is repressed by sugar (25-27), was rapidly
207 upregulated in separated bottoms and grafted bottoms (Figure 2B, Figure S4). These
208 observations were consistent with sugar accumulation in the grafted top and sugar
209 depletion in the grafted bottom. The expression of these genes returned to levels
210 similar to intact samples by 120 hours and, with the exception of *ApL3*, the grafted
211 samples normalised expression more rapidly than the separated tissues. Genes
212 associated with photosynthesis increase expression in separated bottoms 24 hours
213 after cutting, a common response to starvation (13), but likely too late to affect sugar
214 levels before 24 hours (Figure S4). A transcriptional overlap analysis with RNAs
215 from known glucose-responsive genes (Table S1) revealed a strong overlap with
216 genes differentially expressed by grafting. RNAs from known glucose-induced genes
217 were upregulated in separated tops and grafted tops, whereas transcripts from known
218 glucose-repressed genes were upregulated in separated bottoms and grafted bottoms
219 (Figure 2C, Figure S4). This trend was not observed with genes differentially
220 expressed by mannitol treatment (Figure S4), suggesting the effect was specific to
221 metabolically active sugars. To further investigate this effect, we stained grafted,
222 separated and intact plants with Lugol solution to assay for the presence of starch.
223 Staining above the graft junction increased 48-72 hours after grafting (Figure 2D). By
224 120 hours, staining was equal on both sides of the graft whereas in separated tops
225 staining became stronger after 72 hours (Figure 2D, Figure S4). We concluded that

226 starch accumulated above the graft junction, but after 72 hours, this asymmetry
227 disappeared. To test whether the accumulation of starch and increased sugar
228 responsiveness could explain the observed transition from asymmetry to symmetry,
229 we compared our datasets to previously published genes that are induced by
230 starvation or are induced by sucrose re-addition (Table S1). In early time points, 20-
231 31% of asymmetrically expressed genes are genes known to respond to sugars
232 compared to 2-5% of symmetrically expressed genes (Table 1). However, at 72 hours,
233 the overlap between asymmetrically expressed genes and sugar responsive genes
234 reduced substantially (Table 1).



235

236 **Figure 2. Asymmetric changes in accumulation of sugar-responsive RNAs and of**
237 **starch occur at the graft junction; equalizing sucrose between the grafted**
238 **segments reduces and delays phloem reconnection.** (A) *pSUC2::GFP*-expressing
239 *Arabidopsis* shoots were grafted to Col-0 wild type roots and GFP movement to the
240 roots was monitored over 7 days for phloem connection in the presence or absence of
241 various concentrations of sucrose. (B) Expression profiles for transcripts of a sugar-
242 repressed gene (*GDH1*), and a sugar-induced gene (*ApL3*) were plotted for intact,
243 separated and grafted samples. (C) Transcriptional overlap between previously
244 published glucose-induced or glucose repressed genes and our dataset. The number in
245 brackets represents the number of glucose-responsive genes identified in the previous
246 dataset, and overlap is presented as a ratio out of 1.0 for differentiation expressed
247 genes (DEG) up- or down-regulated in our dataset relative to intact samples. Asterisks
248 represent a significant difference ($p < 0.05$) between the ratio of up- and down-
249 regulated genes in a previously published transcriptome dataset compared to the ratio

250 of all up- and down- regulated genes in our grafting dataset at a certain time point. (D)
251 Lugol staining of grafted plants at various time points revealed dark brown staining
252 associated with starch accumulation. HAG, hours after grafting.
253

254 **Table 1. Sugar response overlaps with asymmetry.** Symmetrically and
255 asymmetrically differentially expressed genes were compared to previously published
256 sugar-responsive genes (28)[1] and the percent overlap calculated (* p<.05). HAG,
257 hours after grafting.
258

HAG	Sugar Induced[1]	Graft Bottom=Top	Overlap	%	Graft Top>Bottom	Overlap	%	Graft Bottom>Top	Overlap	%
6	2243	4988	263	5	6679	1403	21*	4971	99	2
12	2243	3473	165	5	7111	1444	20*	5657	231	4
24	2243	4135	171	4	6873	1414	21*	4942	195	4
48	2243	3689	197	5	6601	1241	19*	4915	218	4
72	2243	10421	1005	10*	2459	228	9*	2019	138	7
120	2243	15012	1063	7	1510	220	15*	941	115	12*
HAG	Sugar Repressed[1]	Graft Bottom=Top	Overlap	%	Graft Top>Bottom	Overlap	%	Graft Bottom>Top	Overlap	%
6	1998	4988	107	2	6679	53	1	4971	1563	31*
12	1998	3473	68	2	7111	112	2	5657	1526	27*
24	1998	4135	88	2	6873	72	1	4942	1525	31*
48	1998	3689	113	3	6601	93	1	4915	1427	29*
72	1998	10421	530	5	2459	111	5	2019	538	27*
120	1998	15012	979	7	1510	84	6	941	210	22*

259

260

261 **Vascular formation and cell division activates on both sides of the graft**

262 Since the grafted bottom samples exhibited a starvation response up to 48 hours after
263 grafting, we reasoned that pathways associated cell division and cell differentiation
264 would be delayed or inhibited. We looked at the expression of markers associated
265 with vascular formation and cell division in the transcriptome datasets. Cambium,
266 phloem and provascular markers activated within 6 hours in grafted top samples, but
267 activation was only delayed 0-24 hours in grafted bottom samples depending on the
268 gene (Figure 1B-C, Figure 3, Figure S5, Figure S6). Expression of phloem markers
269 peaked in both grafted tops and grafted bottoms at 72 hours (Figure 3, Figure S5,
270 Figure S6), the time when phloem reconnections form in grafted *Arabidopsis* (Figure
271 1, (4, 5)). Notably, the early phloem marker *NAC020* activated before the mid phloem
272 marker *NAC086* which activated before the late phloem marker *NEN4*, consistent
273 with the dynamics of phloem transcriptional activation during primary root
274 development and leaf vascular induction (Figure S5)(29, 30). Certain markers
275 associated with xylem formation, such as *VND7* and *BFN1*, activated early in the
276 grafted top. Other xylem markers, such as *IRX3* and *CESA4*, activated late in grafted
277 samples. By 120 hours after grafting, xylem markers were activated in top and
278 bottom, consistent with when the first xylem strands differentiate at the graft junction

279 (4). Genes associated with cell division were activated by 12 hours in the grafted top
280 and by 24 hours in the grafted bottom (Figure 3, Figure S5). On the other hand,
281 control genes whose expression does not typically vary between tissues and
282 treatments (31) were not differentially expressed in grafted tops or bottoms (Figure
283 S6). The RNAseq expression data appeared to correlate well with transcriptional
284 fluorescent reporters for both activation dynamics and the localisation of expression
285 (Figure 1D-F, Figure S2).
286
287 The similar activation dynamics of vascular differentiation genes in grafted tops and
288 grafted bottoms prompted us to test whether this phenomenon occurred with other
289 known developmental processes. We obtained lists of genes whose expression is
290 associated with various biological processes from previous publications (Table S1)
291 and tested how many of the genes differentially expressed in our transcriptomes
292 overlapped with the previously published lists. Differentially expressed genes in
293 grafted samples and separated tops partially overlapped with those whose expression
294 is associated with phloem, xylem and procambium formation (Figure 4, Figure S7).
295 There was a high overlap between *Arabidopsis* inflorescence stem healing and
296 grafting, as well as between vascular induction in leaf disk cultures and grafting
297 (Figure 4). Various genes expressed in a cell type-specific manner also showed a high
298 transcriptional overlap with grafting, including phloem, endodermis and protoxylem
299 (Figure 4, Figure S7). In nearly all cases, the separated top, grafted top and grafted
300 bottom samples showed similar activation dynamics. The separated bottom samples
301 were exceptional though since gene expression associated with vascular development
302 and cell-specific processes was downregulated (Figure 4, Figure S7). We also
303 compared our datasets to RNAs expressed in longitudinal cross sections of the
304 *Arabidopsis* root (32). There was little overlap between grafted bottoms and sections
305 from the root meristemic zone, whereas overlap existed between grafted tops and the
306 root meristemic zone in early time points, and between grafting and the root
307 maturation zone (Figure S8). Our analysis also revealed that two genes expressed in
308 cambium, *WOX4* and *PXY*, were induced by grafting but the primary root marker
309 *WOX5* and lateral root marker *LBD18* were not (Figure 3, Figure S5, Figure S6).

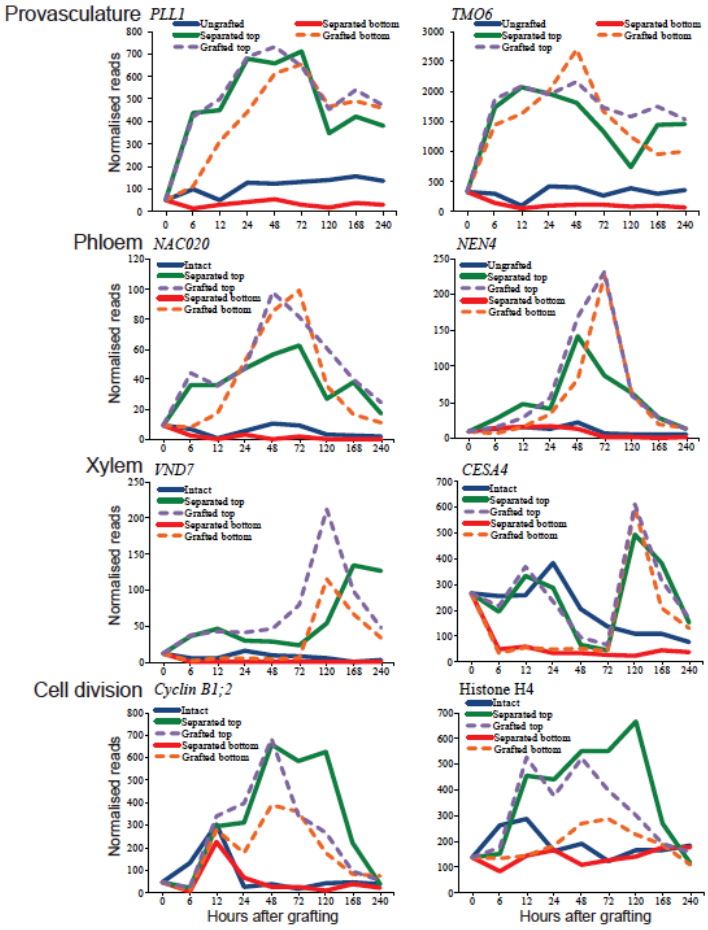


Figure 3. Transcriptional dynamics of genes associated with provasculature, phloem, xylem and cell division. Expression levels were plotted over time for intact, separated and grafted samples.

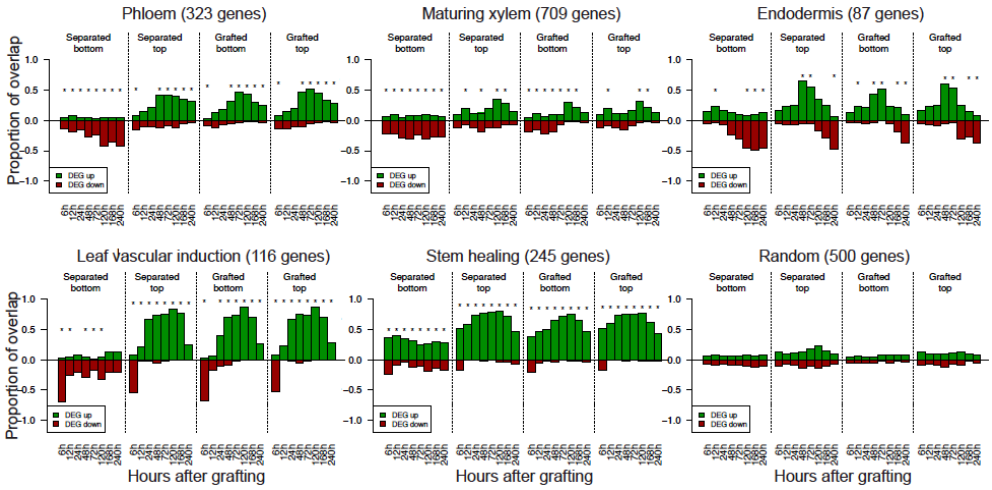


Figure 4. Transcriptional overlap between previously published vascular datasets and the grafting datasets. Genes whose transcripts are associated with various cell types or biological processes were taken from previously published datasets (see Table S1) and compared to the transcriptomic datasets generated here. The number in brackets represents the number of cell type-specific or process-specific genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0

322 for differentiation expressed genes (DEG) up- or down-regulated in our dataset
323 relative to intact samples. Asterisks represent a significant difference ($p < 0.05$)
324 between the ratio of up- and down- regulated genes in a previously published
325 transcriptome dataset compared to the ratio of all up- and down- regulated genes in
326 our grafting dataset at a certain time point.

327

328 **Auxin response is symmetric at the graft**

329 The rapid activation of vascular markers in the grafted bottoms despite the starvation
330 response promoted us to investigated whether other mobile substances such as
331 phytohormones could play a role in gene activation. We compared lists of genes
332 known to respond to cytokinin, ethylene or methyl jasmonate (33) and found no
333 substantial overlap between these lists and genes differentially expressed by grafting
334 (Figure S9)(Table S1). Abscisic acid-responsive and brassinosteroid-responsive genes
335 showed overlap with genes differentially expressed in our datasets, but this overlap
336 was of a similar magnitude in both separated and grafted datasets suggesting the
337 effect was not specific to grafting (Figure S9). Auxin responsive transcripts were
338 exceptional though, as they showed a substantial overlap with RNAs differentially
339 expressed in our data sets that varied depending on the treatment (Figure 5A-B,
340 Figure S9). Auxin-induced genes were upregulated in separated tops, grafted bottoms
341 and grafted tops whereas they were repressed in separated bottoms (Figure 8). Auxin
342 responsive genes such as *IAA1* and *IAA2* (34) were induced to similar levels in grafted
343 tops and grafted bottoms at 24 hours. To further investigate whether auxin response
344 was uniform between grafted tops and grafted bottoms, we grafted the auxin-
345 responsive fluorescent reporter *p35S:DII-Venus* whose fluorescent protein is degraded
346 in the presence of auxin (35). DII-Venus fluoresced in the separated bottoms but did
347 not fluoresce in grafted bottoms 14 hours after cutting (Figure 5C) indicating
348 separated bottoms had a low level of auxin response but grafted tops, grafted bottoms
349 and separated tops had a high level of auxin response.

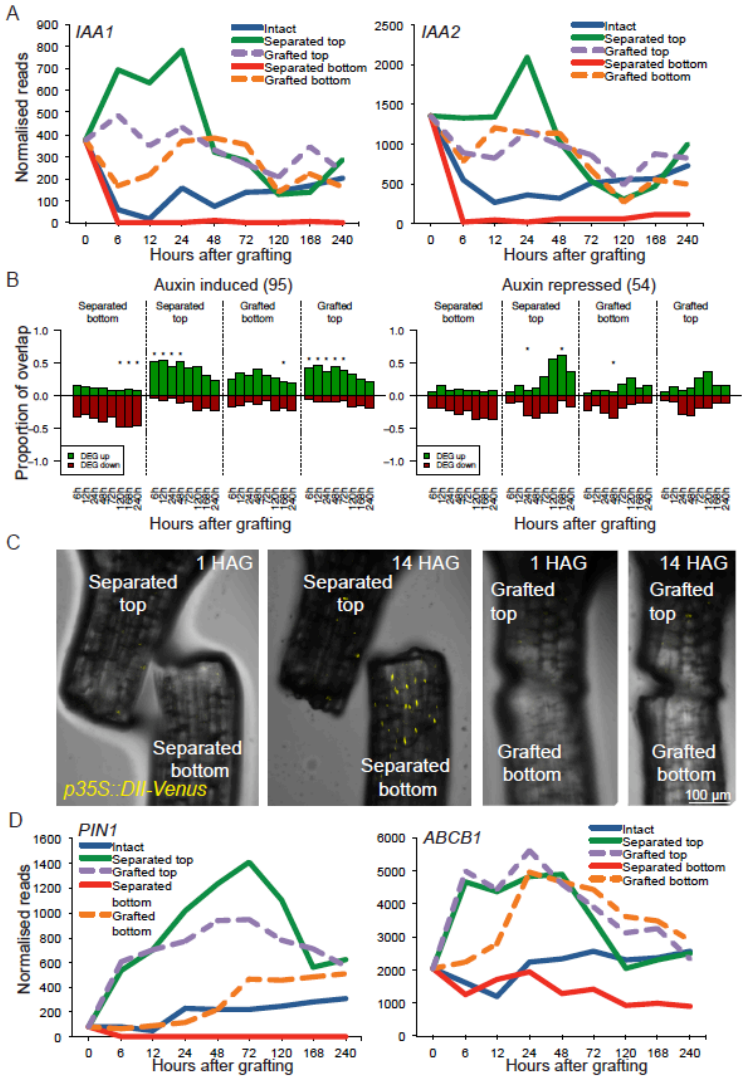


Figure 5. Auxin response is symmetric at the graft junction. (A,D) Expression profiles for various auxin-responsive genes (*IAA1*, *IAA2*) or auxin transporter genes (*PIN1* and *ABCB1* genes) were plotted for intact, separated and grafted samples. (B) Overlap between previously published auxin-induced or auxin-repressed RNAs and our dataset. The number in brackets represents the number of auxin-responsive genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0 for differentiation expressed genes (DEG) up- or down-regulated in our dataset relative to intact samples. Asterisks represent a significant difference ($p < 0.05$) between the ratio of up- and down- regulated genes in a previously published transcriptome dataset compared to the ratio of all up- and down- regulated genes in our grafting dataset at a certain time point. (C) Grafted and separated plants expressing the auxin responsive *p35S::DII-Venus* transgene that is degraded in the presence of auxin reveal a reduction of auxin response in cut bottoms, but not in grafted bottoms. HAG, hours after grafting.

368 **Tissue fusion imparts a unique physiological response that differs from tissue**
369 **separation**
370 We hypothesized that the symmetric auxin response and asymmetric sugar response at
371 the graft junction could allow a unique transcriptional response since neither
372 separated plants nor intact plants had similar response dynamics to sugars and auxin
373 as measured by genome-wide gene response, starch accumulation and *p35S::DII-*
374 *Venus* expression (Figure 2, Figure 5). To uncover protein-coding genes differentially
375 expressed specifically by grafting, we performed an empirical Bayesian analysis (36)
376 where at each time point we computed likelihoods for each gene for each possible
377 pattern of differential expression in the intact, grafted top, grafted bottom, separated
378 top and separated bottom tissues (Figure 6A). These likelihoods were used to define a
379 similarity score between pairs of genes, which was used to cluster genes with similar
380 patterns of gene expression across the five tissues (37). Clusters were then identified
381 by the predominant pattern of gene expression observed within a cluster. This analysis
382 produced 113 clusters containing at least 10 RNAs at a single time point (Table S2;
383 Dataset S1). Approximately 6000 genes were differentially expressed in at least one
384 tissue whereas between 1000 and 4000 genes were not differentially expressed
385 (Figure 6B).
386
387 To simplify the analysis, we considered clusters where gene expression was grouped
388 into patterns consisting of one comparison. At early time points, the cluster containing
389 genes differentially expressed in grafted tops, grafted bottoms, separated tops and
390 separated bottoms had high numbers that decreased with time and could represent an
391 general wound response (Figure 6C). Similarly, clusters containing separated tops and
392 grafted tops and clusters containing separated bottoms and grafted bottoms initially
393 had high numbers that decreased with time. This observation indicated the grafted top
394 was initially transcriptionally similar to the separated top, whereas the grafted bottom
395 was initially transcriptionally similar to the separated bottom. After the 48-hour time
396 point, clusters containing genes only differentially expressed in separated tops or only
397 differentially expressed in separated bottoms increased in numbers, suggesting these
398 tissues gained a unique pattern of gene expression. The cluster containing genes
399 differentially expressed in grafted tops and grafted bottoms increase in numbers
400 throughout the healing process (Figure 6C). We searched for cluster categories that
401 contained genes only differentially expressed by grafting and found very few genes

402 downregulated by grafting or upregulated only in the grafted top (Figure 7A). Instead,
403 there were clusters that contained several hundred differentially expressed genes
404 present either in the grafted bottom only, or present in both grafted bottom and grafted
405 top (Figure 7A). Genes whose expression changed only in the grafted bottom sample
406 were prevalent early during grafting and were most common at 48 hours, whereas
407 genes activated in both top and bottom became prevalent at 48 hours and were most
408 common at 120 hours (Figure 7A). We performed a gene ontology (GO) analysis on
409 the genes whose expression was grafting-specific and found that genes coding for
410 RNAs differentially expressed only in the grafted bottom sample were enriched in the
411 immune response and chitin response biological process categories (Table S3).
412 Previously published chitin-induced RNAs had a high proportion of overlap with
413 differentially expressed graft bottom-specific genes (Figure 7E, Figure S10).
414 Grafting-specific RNAs expressed symmetrically, in both the grafted top and grafted
415 bottom, were enriched in vascular-related biological processes (Table S3). Previously
416 published phloem-enriched, endodermal-enriched, vascular-induction and stem-
417 wounding associated RNAs had a high proportion of overlap with differentially
418 expressed graft-specific genes (Figure 7C-D, Figure S10). Since few genes were
419 grafting-specific and grafted tissues were initially transcriptionally quite similar to
420 separated tissues (Figure 6C, Figure 7A), we reasoned that tissues separated for short
421 periods (<48 hours) could be grafted with similar reconnection dynamics as tissues
422 that had been grafted immediately. To test this hypothesis, plants were cut at day 0,
423 and grafted at 1-5 days after separation. Separated tissues did not speed up vascular
424 reconnection, and instead, it always took three days from the point of tissue
425 attachment for vascular connections to form (Figure S10). Furthermore, the shoot lost
426 competence to graft 2-3 days after separation whereas the root remained competent to
427 graft up to 5 days after separation (Figure S10). Together, it appears the grafted shoot
428 and root have a unique physiological response that differ from separated plants and
429 that tissue attachment is required to activate graft formation and tissue fusion.

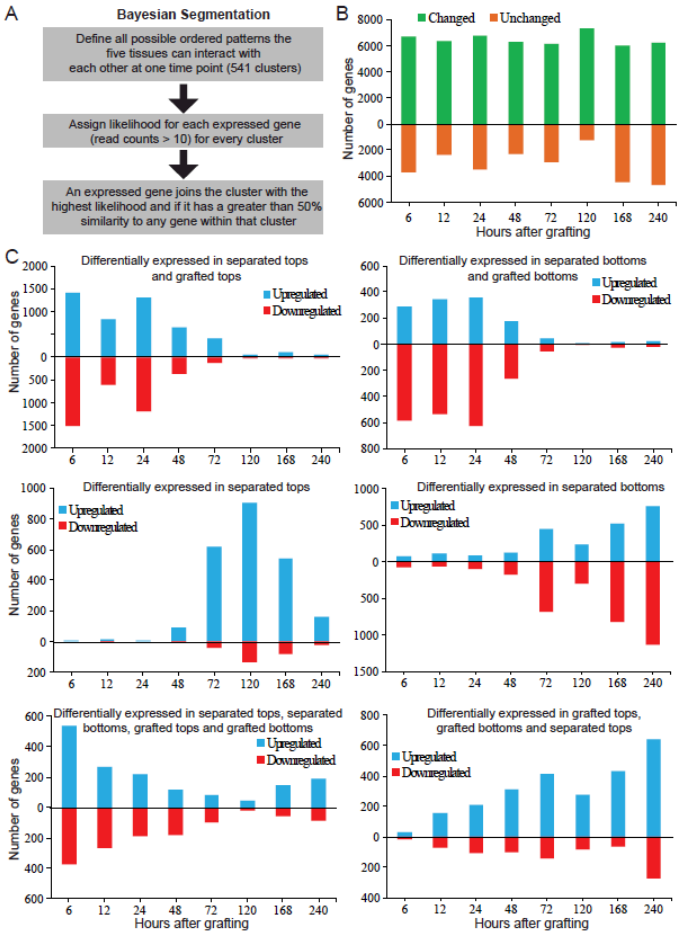


Figure 6. Clustering the transcriptome at each time point, based on likelihoods of all possible patterns of differential expression. (A) Cartoon depicting the Bayesian segmentation. (B) Analysis of differential behaviour produced 113 categories containing at least 10 genes whose expression was expressed in a specific differential pattern in at least one time point (see Table S2). One group is comprised of genes whose transcript levels are not substantially changed in the five tissues (unchanged) whereas the other group is comprised of the sum of the other 112 groups (genes whose transcript levels changed after treatment in a least one tissue) over the time points tested. (C) Major categories in the segmentation revealed RNAs whose levels changed in all the treatments listed relative to intact samples. Note that a gene can be represented only in one category for a given time point, that category in which the transcript level changes best fit the category.

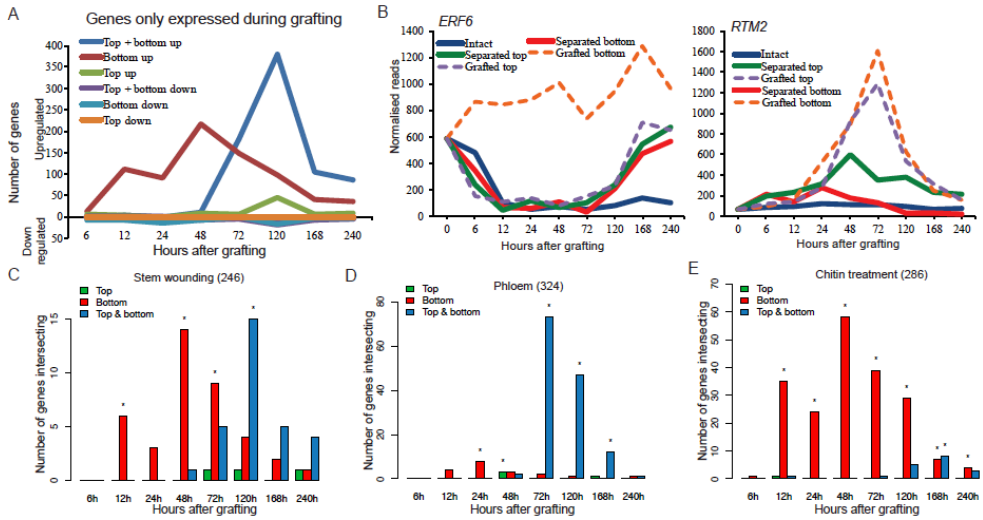


Figure 7. A subset of genes is differentially expressed only during graft formation when compared to 0 hour grafts, or intact or separated tissues. (A) Certain genes were only differentially expressed in grafted tops, grafted bottoms or both in grafted tops and grafted bottoms. (B) Expression profiles for a graft bottom-specific (*ERF6*) or a graft top and bottom differentially regulated gene (*RTM2*) were plotted for intact, separated and grafted samples. (C) Grafting-specific genes are also expressed in other processes of tissue fusion, such as stem healing. Here, 246 previously published RNAs whose differential expression is associated with stem healing were compared to our dataset to assess transcriptional overlap with the grafting-specific genes. Asterisks represent a significant high overlap ($p < 0.05$) of previously published genesets that are also differentially expressed in the grafted samples at a certain time point. (D-E) Genes differentially expressed in grafted tops and grafted bottoms show high overlap to 324 previously published genes whose transcripts are associated with phloem, whereas genes expressed in grafted bottoms show high overlap to 286 previously published genes whose encoded RNAs are associated with chitin treatment. Asterisks represent a significant high overlap ($p < 0.05$) of previously published genes that are also differentially expressed in the grafted samples at a certain time point.

DISCUSSION

To better understand how plants graft, we analysed in depth an RNA deep sequencing dataset that spatially and temporally distinguished genes activated by cutting followed by tissue attachment or continuous tissue separation. Cutting promoted a similar wound response in both grafted and separated tissues, however, by 72 hours after cutting the grafted and separated tissues became transcriptionally dissimilar (Figure 6C) indicating that tissue fusion was mechanistically different from healing an unattached cut surface. During graft formation, tissues had a very high transcriptional overlap with genes differentially expressed by inflorescence stem healing and by vascular induction in leaves (Figure 4, Figure S7) (9, 29) suggesting grafting is closely related to tissue reunion and vascular formation processes. Graft formation

474 had little transcriptional overlap with lateral root formation (Figure S6)(32) and
475 appeared to follow a pathway similar to secondary root growth (Figure S6, Figure S8)
476 since cambium markers *WOX4* and *PXY* that are specific to secondary growth were
477 expressed during grafting (38)(Figure S5). Grafted tops initially showed a short-
478 lasting and small transcriptional overlap with genes expressed during primary root
479 formation which may be related to the accumulation of substances activating
480 adventitious root formation, a common response in failed grafts or in cut shoots
481 (Figure S4C). Thus, we conclude that grafting proceeds via a pathway involving
482 secondary growth with radial meristems activating in the mature cambium to heal the
483 wound. Vascular formation genes including those specifying cambium and phloem
484 were activated early, followed by an activation of cell division genes, suggesting that
485 the start of cellular differentiation preceded activation of cell division. Xylem identity
486 genes showed an early and a late activation peak (Figure 3, Figure S5). There is no
487 visible xylem differentiation at the graft junction during the first peak of expression
488 (4) and these genes might be suppressed by phloem differentiation genes such as *APL*
489 and *CLE41* that activate early after grafting and are known to suppress protoxylem
490 formation (39, 40)(Figure S6)(Table S4). Alternatively, the first peak could represent
491 programmed-cell death that does not lead to xylem differentiation. The second
492 expression peak of xylem-expressed genes at 120 hours occurred after the
493 differentiation of functional phloem and coincided with the differentiation of xylem
494 strands at the graft junction (4). Previous studies highlighted the importance of callus
495 and pericycle cells during regeneration (16, 41), but we see little evidence that genes
496 expressed in the pericycle or during callus formation have high transcriptional overlap
497 with genes differentially expressed by grafting (Figure S4). Expression profiles for all
498 protein-coding genes can be found in table S4.
499
500 We observe only a slight delay in phloem, cambium and cell division activation
501 below the graft junction compared to above it (Figure 3, Figure S5, Figure S6)
502 whereas several genes associated with vascular formation, such as *HCA2* (42) and
503 *TMO6* (43), activated equally in both grafted top and grafted bottom at 6 hours after
504 grafting (Figure 1D-G, Figure 3). These data indicate that, at least transcriptionally,
505 the grafted root rapidly responded to the presence of the grafted shoot and this
506 response was independent of functional vascular connections. This response was not
507 present in separated roots, indicating that attachment was key for recognition. Sugars

508 are known activators of cell division and cell elongation (13) and in our datasets, a
509 large proportion of asymmetrically expressed genes are genes that are sugar-
510 responsive. However, sugars are transported in the phloem (12) that is severed upon
511 grafting whereas the grafted root exhibited a sugar starvation response and showed
512 similar sugar-response dynamics as the separated root. Instead, we infer that not sugar
513 but auxin, or some other molecular that is transported in the absence of vascular
514 connections, could be the signal in the grafted bottom that activates *HCA2*, *TMO6* as
515 well as cell division, phloem- and cambium-related genes. Alternatively, separation
516 might cause a build-up of molecules from the root that suppress vascular induction
517 genes such as *HCA2* and *TMO6*.
518
519 Given auxin's role in vascular patterning and formation ((38), it is a strong candidate
520 for the activating signal. Auxin response was largely symmetric, particularly from 12
521 hours after grafting (Figure 6) consistent with previous findings that the auxin-
522 inducible *DR5* and *ANAC071* genes are symmetrically expressed around the graft
523 junction (4-6, 44). One idea is that auxin transport was not substantially interrupted by
524 the grafting process, and instead, where opposing tissues adhered, auxin moved
525 regardless of vascular connections since auxin is transported from cell to cell through
526 the apoplast (8). The genes encoding auxin efflux proteins PIN1 and ABCB1 were
527 transcriptionally activated above the graft junction (Figure 5D), similar to putative
528 *Pisum sativum* PIN1 protein accumulating above a cut stem prior to vascular
529 reconnection (45), and could reflect a role for these proteins in exporting auxin across
530 the cut. Consistent with these observations, adding an auxin transport inhibitor to
531 grafted *Arabidopsis* shoots prevented the expression of grafting-induced genes below
532 the graft junction (6). Since auxin is fundamental for patterning vasculature (38), it is
533 plausible that auxin movement across the graft occurs prior to sugar movement since
534 the flow of auxin could assist with reconnecting the phloem. Although auxin response
535 was symmetric, our previous work demonstrated that the auxin response factors *ALF4*
536 and *AXR1* affected grafting only below the graft junction (4). Mutating *ALF4* below
537 the graft junction more strongly reduced auxin response than mutating *ALF4* above
538 the junction (4). Thus, proteins such as ALF4 or AXR1 might enhance or promote
539 rootstock-specific auxin response and vascular regeneration that could be particularly
540 important when there is incomplete attachment, cellular damage or inefficient
541 transport. All higher plants transport auxin from shoot to root, yet not all plant species

542 can be successfully grafted (3) so the response to auxin rather than the transport itself
543 may be a determining factor in the ability to graft. A role for sugars is not completely
544 ruled out though, since the magnitude of differential expression was often lower in the
545 grafted bottom and exogenous sugar reduced grafting efficiency (Figure 2A).
546 Pericycle cells require auxin to divide and the addition of sugars enhances the rate of
547 divisions (46), suggesting that the presence of sugars in the grafted top enhanced cell
548 division and differentiation. Another hypothesis that warrants testing is whether graft
549 healing activates via a mechanical signal provided by the physical presence of the
550 opposing tissues.
551
552 Our analyses identified two groups of genes whose expression changes were unique
553 to graft formation in our experiments (Figure 7). One group activated shortly after
554 grafting below the graft junction and was enriched in immune-responsive and chitin-
555 responsive genes (Table S3, Figure 7, Figure S10). The breakdown products of cell
556 walls are potent elicitors of defence responses (47), so it is possible that the grafted
557 bottom upregulates pathways specific to wound damage response. This group was not
558 upregulated in separated bottoms though, so the unique physiological state of the
559 grafted root, indicated by the presence of auxin response but the absence of sugar
560 response, could have promoted their upregulation. The second group activated both
561 above and below the graft junction and became highly expressed later during graft
562 formation (Figure 7). This group was enriched in RNAs associated with vascular
563 development (Table S3, Figure 7) and we suggest that the products of these genes are
564 involved in the vascular reconnection processes between the two tissues. Despite
565 many transcriptional similarities between separated tops and grafted tissues, tissues
566 had to be attached for at least three days for phloem connections to form, regardless
567 of when cutting occurred (Figure S10). Thus, it appears that RNAs expressed in the
568 separated top or separated bottom are insufficient to drive graft formation. Instead
569 attachment and the genes activated by this recognition process including grafting-
570 specific RNA changes (Figure 7A) or RNAs expressed in the grafted bottom, grafted
571 top and separated top (Figure 6C) are those that contribute to a response that
572 distinguishes attached from separated plant tissues. Future work should focus on these
573 RNAs to identify the pathways required for grafting that could be modified to
574 improve graft formation, wound healing and vascular regeneration. Likewise, the
575 rapid transcriptional changes below the graft indicate a recognition system that

576 promotes tissue regeneration. Identifying the exogenous cue that triggers recognition
577 and understanding how it is perceived should be priorities, as should understanding
578 whether this phenomenon applies more broadly to inter-tissue communication, tissue
579 regeneration or tissue fusion events, such as parasitic plant infections (48), epidermal
580 fusions (49, 50) or petal fusions (51).
581
582 **METHODS**
583 **Plant material and microscopy**
584 *Arabidopsis thaliana* accession Columbia was used throughout. The *p35S::GFP-ER*
585 (52), *pSUC2::GFP* (53), *pUBQ10::PM-tdTomato* (54), *pARR5::GFP* (55),
586 *pANT::H2B-YFP* (56), *pLOG4::n3GFP* (57), *pCASP1::NLS-GFP* (58), *p35S::DII-*
587 *Venus* (35) lines have been previously published. For the construction of
588 *pHCA2::RFP*, a 2.9kb 5' upstream region of the *HCA2* gene was cloned into
589 pDONRp4-p1R donor vector, and recombined with tagRFPer into a destination vector
590 by the Multisite Gateway system (59). *Arabidopsis thaliana* micrografting and
591 grafting assays were performed according to previously published protocols (60, 61).
592 Fluorescent images were taken on a Zeiss LSM-700 or LSM-780 confocal
593 microscope. Black and white fluorescent images of graft junctions were taken on a
594 Zeiss V12 dissecting microscope. FIJI software (Fiji.sc) was used to process images.
595 A micro-extensometer was used for all breaking force graft measurements according
596 to a previously published protocol (62).
597
598 **RNAseq sample and library preparation**
599 Grafted wild type *Arabidopsis thaliana* accession Col-0 were harvested at the
600 respective time points and care taken to separate grafts by gently pulling plants apart.
601 Approximately 0.5mm of tissue was taken above or below each cut site and kept
602 separate. Intact plants had 1mm of tissue taken in a similar location on the hypocotyl
603 as separated or grafted plants. Grafted, separated or intact tissues were pooled into
604 groups of approximately 24 tissues . Tissues were ground using a microcentrifuge
605 pestle frozen in liquid nitrogen. RNA was extracted using an RNeasy Kit (Qiagen,
606 UK) following the manufacturer's instructions. 90-100ng of RNA was used to prepare
607 RNAseq libraries using the TruSeq® Stranded mRNA LT kit (Illumina, UK)
608 according to the manufacturer's instructions. The final PCR was for 15 cycles and 11-
609 12 barcoded samples were randomly mixed to make a total of 7 mixes for 7 flow

610 lanes, one mix per lane. Samples were sequenced on the HiSeq 4000 platform
611 (Illumina, UK) with Paired End 75bp transcriptome sequencing (BGI Tech Solutions,
612 Shenzhen, China).
613
614 **Iodine staining**
615 *Arabidopsis* seedlings were placed in a fixation solution (3.7% formaldehyde, 50%
616 ethanol, 5% acetic acid) for 1 hour at room temperature, then transferred to 70%
617 ethanol for 10 minutes. Afterwards, plants were transferred to 96% ethanol and stored
618 at -20°C for up to a week. Samples were rehydrated in 50% ethanol for 1 hour at
619 room temperature, transferred to distilled water for 30 minutes, then stained for 10
620 minutes in Lugol solution (Sigma) at room temperature. Plants were rinsed with water
621 and mounted on microscope slides. Images were taken on a Zeiss Axioimager.M2
622 microscope.
623
624 **Bioinformatic analyses**
625 The reads acquired through high-throughput sequencing were quality trimmed with
626 sickle (63) and aligned using the eXpress tool to protein-coding gene sequences
627 acquired from TAIR10 using Bowtie2. Library scaling factors were inferred from the
628 sum of the number of reads assigned to the genes in the lowest seventy-five
629 percentiles of expressed genes for each library (64). Analyses of the data were carried
630 out using the R package baySeq (36)and clustering based on the posterior
631 probabilities acquired from this package. The gene ontology analysis (GO)
632 enrichment analysis on grafting-specific genes was done with a customized R script
633 using the package GOstats (65).
634
635 **ACKNOWLEDGEMENTS**
636 We thank Niko Geldner, Dolf Weijers, Paul Tarr, Yka Helariutta, Ruth Stadler and
637 The Arabidopsis Information Resource for providing seeds. Funding for this work
638 was provided by Gatsby Charitable Trust grants GAT3272/C and GAT3273-PR1, by
639 a Knut and Alice Wallenberg Academy Fellowship KAW2016.0274 (to C.W.M), and
640 by the Howard Hughes Medical Institute and Gordon and Betty Moore Foundation
641 grant GBMF3406 (to E.M.M.)
642
643

644 **REFERENCES**

645 1. Goldschmidt EE (2014) Plant grafting: new mechanisms, evolutionary
646 implications. *Front Plant Sci* 5:727.
647 2. Lee J*, et al.* (2010) Current status of vegetable grafting: Diffusion, grafting
648 techniques, automation. *Sci Hortic-Amsterdam* 127:93-105.
649 3. Melnyk CW (2017) Plant grafting: insights into tissue regeneration.
650 *Regeneration (Oxf)* 4(1):3-14.
651 4. Melnyk CW, Schuster C, Leyser O, & Meyerowitz EM (2015) A
652 developmental framework for graft formation and vascular reconnection in
653 Arabidopsis thaliana. *Curr Biol* 25(10):1306-1318.
654 5. Yin H*, et al.* (2012) Graft-union development: a delicate process that involves
655 cell-cell communication between scion and stock for local auxin
656 accumulation. *J Exp Bot* 63(11):4219-4232.
657 6. Matsuoka K*, et al.* (2016) Differential Cellular Control by Cotyledon-Derived
658 Phytohormones Involved in Graft Reunion of Arabidopsis Hypocotyls. *Plant*
659 *Cell Physiol* 57(12):2620-2631.
660 7. Matsumoto-Kitano M*, et al.* (2008) Cytokinins are central regulators of
661 cambial activity. *P Natl Acad Sci USA* 105(50):20027-20031.
662 8. Leyser O (2011) Auxin, self-organisation, and the colonial nature of plants.
663 *Curr Biol* 21(9):R331-337.
664 9. Asahina M*, et al.* (2011) Spatially selective hormonal control of RAP2.6L and
665 ANAC071 transcription factors involved in tissue reunion in Arabidopsis. *P*
666 *Natl Acad Sci USA* 108(38):16128-16132.
667 10. Wetmore RH & Rier JP (1963) Experimental induction of vascular tissues in
668 callus of angiosperms. *Am J Bot* 50(5):418-430.
669 11. Aloni R (1980) Role of auxin and sucrose in the differentiation of sieve and
670 tracheary elements in plant tissue cultures. *Planta* 150(3):255-263.
671 12. Lough TJ & Lucas WJ (2006) Integrative plant biology: Role of phloem long-
672 distance macromolecular trafficking. *Annu Rev Plant Biol* 57:203-232.
673 13. Wang L & Ruan YL (2013) Regulation of cell division and expansion by
674 sugar and auxin signaling. *Front Plant Sci* 4:163.
675 14. Chitwood DH, Guo M, Nogueira FT, & Timmermans MC (2007) Establishing
676 leaf polarity: the role of small RNAs and positional signals in the shoot apex.
677 *Development* 134(5):813-823.
678 15. Cheong YH*, et al.* (2002) Transcriptional profiling reveals novel interactions
679 between wounding, pathogen, abiotic stress, and hormonal responses in
680 Arabidopsis. *Plant Physiol* 129(2):661-677.
681 16. Iwase A*, et al.* (2011) The AP2/ERF transcription factor WIND1 controls cell
682 dedifferentiation in Arabidopsis. *Curr Biol* 21(6):508-514.
683 17. Ikeuchi M*, et al.* (2017) Wounding triggers callus formation via dynamic
684 hormonal and transcriptional changes. *Plant Physiol.*
685 18. Cookson SJ*, et al.* (2013) Graft union formation in grapevine induces
686 transcriptional changes related to cell wall modification, wounding, hormone
687 signalling, and secondary metabolism. *J Exp Bot* 64(10):2997-3008.
688 19. Cookson SJ*, et al.* (2014) Heterografting with nonself rootstocks induces
689 genes involved in stress responses at the graft interface when compared with
690 autografted controls. *J Exp Bot* 65(9):2473-2481.
691 20. Zheng BS*, et al.* (2010) cDNA-AFLP analysis of gene expression in hickory
692 (Carya cathayensis) during graft process. *Tree Physiol* 30(2):297-303.

21. Chen Z, *et al.* (2017) Transcriptome changes between compatible and incompatible graft combination of Litchi chinensis by digital gene expression profile. *Sci Rep* 7(1):3954.
22. Lindsay DW, Yeoman MM, & Brown R (1974) An analysis of the development of the graft union in Lycopersicon esculentum. *Ann Bot-London* 38:639-646.
23. Moore R (1984) Graft formation in Solanum pennellii (Solanaceae). *Plant Cell Rep* 3(5):172-175.
24. Marsch-Martinez N, *et al.* (2013) An efficient flat-surface collar-free grafting method for Arabidopsis thaliana seedlings. *Plant Methods* 9(1):14.
25. Villadsen D & Smith SM (2004) Identification of more than 200 glucose-responsive Arabidopsis genes none of which responds to 3-O-methylglucose or 6-deoxyglucose. *Plant Mol Biol* 55(4):467-477.
26. Thum KE, Shin MJ, Palenchar PM, Kouranov A, & Coruzzi GM (2004) Genome-wide investigation of light and carbon signaling interactions in Arabidopsis. *Genome Biol* 5(2):R10.
27. Cordoba E, Aceves-Zamudio DL, Hernandez-Bernal AF, Ramos-Vega M, & Leon P (2015) Sugar regulation of SUGAR TRANSPORTER PROTEIN 1 (STP1) expression in Arabidopsis thaliana. *J Exp Bot* 66(1):147-159.
28. Osuna D, *et al.* (2007) Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings. *Plant J* 49(3):463-491.
29. Kondo Y, *et al.* (2016) Vascular Cell Induction Culture System Using Arabidopsis Leaves (VISUAL) Reveals the Sequential Differentiation of Sieve Element-Like Cells. *Plant Cell* 28(6):1250-1262.
30. Furuta KM, *et al.* (2014) Plant development. Arabidopsis NAC45/86 direct sieve element morphogenesis culminating in enucleation. *Science* 345(6199):933-937.
31. Czechowski T, Stitt M, Altmann T, Udvardi MK, & Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* 139(1):5-17.
32. Brady SM, *et al.* (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318(5851):801-806.
33. Nemhauser JL, Hong F, & Chory J (2006) Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. *Cell* 126(3):467-475.
34. Abel S, Nguyen MD, & Theologis A (1995) The PS-IAA4/5-like family of early auxin-inducible mRNAs in Arabidopsis thaliana. *J Mol Biol* 251(4):533-549.
35. Brunoud G, *et al.* (2012) A novel sensor to map auxin response and distribution at high spatio-temporal resolution. *Nature* 482(7383):103-106.
36. Hardcastle TJ & Kelly KA (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11:422.
37. Hardcastle TJ & Papatheodorou I (2017) clusterSeq: methods for identifying co-expression in high-throughput sequencing data. *bioRxiv* 188581.
38. De Rybel B, Mahonen AP, Helariutta Y, & Weijers D (2016) Plant vascular development: from early specification to differentiation. *Nat Rev Mol Cell Biol* 17(1):30-40.

39. Bonke M, Thitamadee S, Mahonen AP, Hauser MT, & Helariutta Y (2003) APL regulates vascular tissue identity in Arabidopsis. *Nature* 426(6963):181-186.
40. Ito Y*, et al.* (2006) Dodeca-CLE peptides as suppressors of plant stem cell differentiation. *Science* 313(5788):842-845.
41. Sugimoto K, Jiao Y, & Meyerowitz EM (2010) Arabidopsis regeneration from multiple tissues occurs via a root development pathway. *Dev Cell* 18(3):463-471.
42. Guo Y, Qin G, Gu H, & Qu LJ (2009) Dof5.6/HCA2, a Dof transcription factor gene, regulates interfascicular cambium formation and vascular tissue development in Arabidopsis. *Plant Cell* 21(11):3518-3534.
43. Gardiner J, Sherr I, & Scarpella E (2010) Expression of DOF genes identifies early stages of vascular development in Arabidopsis leaves. *Int J Dev Biol* 54(8-9):1389-1396.
44. Pitaksaringkarn W, Ishiguro S, Asahina M, & Satoh S (2014) ARF6 and ARF8 contribute to tissue reunion in incised Arabidopsis inflorescence stems. *Plant Biotechnol-Nar* 31(1):49-53.
45. Sauer M*, et al.* (2006) Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Genes Dev* 20(20):2902-2911.
46. Skylar A, Sung F, Hong F, Chory J, & Wu X (2011) Metabolic sugar signal promotes Arabidopsis meristematic proliferation via G2. *Dev Biol* 351(1):82-89.
47. Souza CA*, et al.* (2017) Cellulose-Derived Oligomers Act as Damage-Associated Molecular Patterns and Trigger Defense-Like Responses. *Plant Physiol* 173(4):2383-2398.
48. Musselman LJ (1980) The biology of Striga, Orobanche, and other root-parasitic weeds. *Annual Review of Phytopathology* 18:463-489.
49. Lolle SJ, Hsu W, & Pruitt RE (1998) Genetic analysis of organ fusion in Arabidopsis thaliana. *Genetics* 149(2):607-619.
50. Becraft PW, Stinard PS, & McCarty DR (1996) CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. *Science* 273(5280):1406-1409.
51. Zhong J & Preston JC (2015) Bridging the gaps: evolution and development of perianth fusion. *New Phytol* 208(2):330-335.
52. Nelson BK, Cai X, & Nebenfuhr A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *The Plant journal* 51(6):1126-1136.
53. Imlau A, Truernit E, & Sauer N (1999) Cell-to-cell and long-distance trafficking of the green fluorescent protein in the phloem and symplastic unloading of the protein into sink tissues. *Plant Cell* 11(3):309-322.
54. Segonzac C*, et al.* (2012) The shoot apical meristem regulatory peptide CLV3 does not activate innate immunity. *Plant Cell* 24(8):3186-3192.
55. Yanai O*, et al.* (2005) Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Curr Biol* 15(17):1566-1571.
56. Randall RS*, et al.* (2015) AINTEGUMENTA and the D-type cyclin CYCD3;1 regulate root secondary growth and respond to cytokinins. *Biol Open*.
57. De Rybel B*, et al.* (2014) Plant development. Integration of growth and patterning during vascular tissue formation in Arabidopsis. *Science* 345(6197):1255215.

58. Roppolo D*, et al.* (2011) A novel protein family mediates Casparian strip formation in the endodermis. *Nature* 473(7347):380-383.
59. Siligato R*, et al.* (2016) MultiSite Gateway-Compatible Cell Type-Specific Gene-Inducible System for Plants. *Plant Physiol* 170(2):627-641.
60. Melnyk CW (2017) Grafting with Arabidopsis thaliana. *Methods Mol Biol* 1497:9-18.
61. Melnyk CW (2017) Monitoring Vascular Regeneration and Xylem Connectivity in Arabidopsis thaliana. *Methods Mol Biol* 1544:91-102.
62. Robinson S*, et al.* (2017) An automated confocal micro-extensometer enables in vivo quantification of mechanical properties with cellular resolution. *bioRxiv* 183533.
63. Joshi NA & Fass JN (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)).
64. Hardcastle TJ, Kelly KA, & Baulcombe DC (2012) Identifying small interfering RNA loci from high-throughput sequencing data. *Bioinformatics* 28(4):457-463.
65. Falcon S & Gentleman R (2007) Using GOstats to test gene lists for GO term association. *Bioinformatics* 23(2):257-258.
66. Drost HG & Paszkowski J (2017) Biomartr: genomic data retrieval with R. *Bioinformatics* 33(8):1216-1217.
67. Carlson M (2017) org.At.tair.db: Genome wide annotation for Arabidopsis), R package version 3.4.1.

832 **SUPPORTING INFORMATION**

833 **SI Materials and Methods**

834

835 **Plant material and grafting**

836 *Arabidopsis thaliana* accession Columbia was used throughout. The *p35S::GFP-ER*
837 (52), *pSUC2::GFP* (53), *pUBQ10::PM-tdTomato* (54), *pARR5::GFP* (55),
838 *pANT::H2B-YFP* (56), *pLOG4::n3GFP* (57), *pCASP1::NLS-GFP* (58), *p35S::DII-*
839 *Venus* (35) lines have been previously published. For the construction of
840 *pHCA2::RFP*, a 2.9kb 5' upstream region of the *HCA2* gene was cloned into
841 pDONRp4-p1R donor vector, and recombined with tagRFPer into a destination vector
842 by the Multisite Gateway system (59). The following primers were used for *pHCA2*
843 cloning: attB4_HCA2(-)2958
844 GGGACAACTTTGTATAGAAAAGTTGtcgatacgcgggacagatatac
845 attB1_HCA2ProEnd ggggACTGCTTTTTTGTACAAACTTGttttgtgttctgtatgtttg.
846 *Arabidopsis thaliana* micrografting was performed according to a previously
847 published protocol (60). Briefly, seven day old *Arabidopsis* seedlings were grown
848 vertically on Murashige and Skoog (MS) medium + 1% bacto agar (pH5.7; no
849 sucrose) in short day conditions (8 hours of 80-100 µmol m^{-2} s^{-1} light) at 20°C.
850 Seedlings were placed on one layer of 2.5x4cm sterile Hybond N membrane (GE
851 Healthcare) on top of two 8.5cm circles of sterile 3 Chr Whatman paper (Scientific
852 Laboratory Supplies) moisten with sterile distilled water in a 9cm petri dish. In a
853 laminar flow hood using a dissecting microscope, one cotyledon was removed and a
854 transvers cut through the hypocotyl was made with a vascular dissecting knife (Ultra
855 Fine Micro Knife; Fine Science Tools). Grafts were assembled by aligning the two
856 cut halves and joining them together, after which, the petri dishes was sealed with
857 parafilm and placed vertically under short day conditions at 20°C. For grafting on a
858 microscope coverslip to image the graft junction, a 10cm square Petri dish modified
859 by gluing a microscope coverslip in place of a section of plastic from the back. On top
860 of the microscope cover slip was placed a 2.5x4 cm rectangle of Hybond N
861 membrane. At the edges and base of the Petri dish three 3 x 8cm strips of Whatman
862 paper were placed. Sterile water moistened both Whatman paper and Hybond N.
863 After which, roots were placed on the Hybond N membrane and hypocotyls on the
864 coverslip. Grafting then proceeded as above. Graft junctions were imaged through the

865 coverslip with a Plan-Apochromat 20X/0.8 objective on a Zeiss LSM-700 or LSM-
866 780 confocal microscope.
867
868 **Fluorescent assays and microscopy**
869 To test the effect of sugars on grafting, *Arabidopsis thaliana* Col-0 plants were grown
870 on 1/2MS+ 1% bacto agar (pH5.7; no sucrose) for seven days in short day conditions.
871 Wild type roots were grafted to scions expressing *pSUC2::GFP* using the protocol
872 described above but either water or water containing 0.25%, 0.5% 1% or 2% sucrose
873 was added to the grafted plates. Roots were observed for fluorescence 2-7 days after
874 grafting with a Zeiss V12 dissecting microscope equipped with a GFP filter. Roots
875 were scored daily and the same plants were observed during the 7-day assay as
876 previously described (61).
877
878 Fluorescent images were taken on a Zeiss LSM-700 or LSM-780 confocal microscope
879 with a Zeiss Plan-Apochromat 20X/0.8 dry objective. A 488nm argon laser (Zeiss
880 780) or 488nm solid-state laser (Zeiss 700) was used for excitation of GFP and YFP.
881 A 561nm solid-state laser was used for excitation of the tdTomato fluorescent protein.
882 A T-PMT detector obtained bright-field transmitted light. Black and white fluorescent
883 images of graft junctions were taken on a Zeiss V12 dissecting microscope fitted with
884 a Hamamatsu EM-CCD camera and RFP and YFP filters. FIJI software (Fiji.sc) was
885 used to process images. Image contrast and brightness were adjusted for controls and
886 samples equally. For longitudinal images of the graft junction, z-stack projections are
887 shown and made with the average intensity function in FIJI from stacks containing the
888 hypocotyl vascular tissues, mesophyll and epidermis.
889
890 **RNAseq sample and library preparation**
891 Wild type *Arabidopsis thaliana* accession Col-0 were grafted as above taking care to
892 switch shoot and root between different plants. All grafting and cutting was
893 performed in the morning to minimize circadian effects. For the 0 hour time points,
894 plants were transferred from 1/2MS plates to the grafting plates and immediately
895 harvested. Only intact plants were harvested at the 0 hour time since at this point,
896 there would be insufficient time to reasonably expect the separated or grafted samples
897 to be transcriptionally different (the time between cutting and freezing is less than two
898 minutes). For all other time points, after cutting plants were left on grafting plates for

899 the respective amount of time. Tissues were harvested and care taken to separate
900 grafts by gently pulling plants apart. Approximately 0.5mm of tissue was taken above
901 or below each cut site and kept separate. Intact plants had 1mm of tissue taken in a
902 similar location on the hypocotyl as separated or grafted plants. Grafted, separated or
903 intact tissues were pooled into groups of approximately 24 tissues (1 plates with 24
904 plants) which were immediately placed in 96% ethanol on dry ice. After harvesting,
905 microcentrifuge tubes were briefly centrifuged and the ethanol removed before
906 storing at -80°C. Plants were grafted over two months to get sufficient material.
907
908 Tissues were ground in the microcentrifuge tube using a microcentrifuge pestle frozen
909 in liquid nitrogen. RNA was extracted using an RNeasy Kit (Qiagen, UK) following
910 the manufacturer's instructions including on column DNase digestion. RNA was
911 eluted from the column with 50ul of sterile water. Quality and quantity of RNA was
912 checked using an Agilent 2200 TapeStation and High Sensitivity (HS) RNA
913 screentapes (Agilent, UK). After RNA extraction, two to four biological replicates
914 were combined (50-100 plants) to get sufficient RNA. 90-100ng of RNA was used to
915 prepare RNAseq libraries using the TruSeq® Stranded mRNA LT kit (Illumina, UK)
916 according to the manufacturer's instructions. The final PCR was for 15 cycles and
917 samples were resuspended in 23ul of distilled water. Quantity and quality of DNA
918 libraries was checked on the Agilent 2200 TapeStation using D1000 screentapes
919 (Agilent, UK). Each sample had two libraries prepared from grafted tissues or
920 separated tissues at different times so that independent biological replicates were
921 made. Samples were diluted to 10nM and 11-12 barcoded samples randomly mixed to
922 make a total of 7 mixes for 7 flow lanes, one mix per lane. Samples were sequenced
923 on the HiSeq 4000 platform (Illumina, UK) with Paired End 75bp transcriptome
924 sequencing (BGI Tech Solutions, Shenzhen, China).
925
926 **Iodine staining**
927 *Arabidopsis* seedlings were placed in a fixation solution (3.7% formaldehyde, 50%
928 ethanol, 5% acetic acid) for 1 hour at room temperature, then transferred to 70%
929 ethanol for 10 minutes. Afterwards, plants were transferred to 96% ethanol and stored
930 at -20°C for up to a week. Samples were rehydrated in 50% ethanol for 1 hour at
931 room temperature, then transferred to distilled water for 30 minutes. Samples were
932 then transferred to a solution of Lugol solution (Sigma) and stained for 10 minutes at

933 room temperature. Plants were rinsed with water, then mounted on microscope slides.
934 Images were taken on a Zeiss Axioimager.M2 microscope with a PlanApochromat
935 20x objective and SPOT Flex camera (Imsol, UK).
936
937 **Pairwise and Bayseq analyses**
938 The reads acquired through high-throughput sequencing were quality trimmed with
939 sickle (63) to increase the read quality before mapping. Reads were aligned to
940 protein-coding gene sequences acquired from TAIR10 using Bowtie2. Read
941 assignment was performed using the eXpress tool, which also provided effective gene
942 lengths for use in normalisation. Library scaling factors were inferred from the sum of
943 the number of reads assigned to the genes in the lowest seventy-five percentiles of
944 expressed genes for each library (64).
945
946 Analyses of the data were carried out using the R package baySeq (36)and clustering
947 based on the posterior probabilities acquired from this package. For each timepoint,
948 all possible patterns of differential expression between the graft types were
949 considered, where a 'pattern' defines similarity and difference between different
950 experimental conditions. For example,
951 *'{Col_cut_bottomGenes=Col:Col_bottomGenes=ungraftedGenes},{Col_cut_topGenes=Col:Col_topGenes}'*
952 defines a pattern in which gene expression is equivalent in the separated bottoms, the
953 grafted bottoms and the intact plant, but different to the equivalently expressed
954 separated top and grafted top. The total number of possible patterns for five
955 experimental conditions (as in this analysis) is fifty-two.
956
957 For a given timepoint, posterior likelihoods on the likelihood of each pattern of
958 expression are calculated for every gene with greater than ten reads across all
959 experimental conditions. The patterns were then modified to include orderings
960 (denoted by < or >), for example, the pattern described would lead to the ordered
961 pattern
962 *'{Col_cut_bottomGenes=Col:Col_bottomGenes=ungraftedGenes}>{Col_cut_topGenes=Col:Col_topGenes}'*
963 in which gene expression is equivalent in the separated bottoms, the grafted bottoms
964 and the intact plant and greater than the equivalently expressed separated top and
965 grafted top. In total, 541 ordered patterns exist in this data set. Posterior likelihoods
966 for an ordered pattern were assigned to that of the unordered pattern for genes in

967 which the (normalised) mean expressions within the equivalently expressed groups
968 conformed to the ordering, and to zero otherwise.
969
970 Based on the posterior likelihoods for the ordered patterns, a similarity score s_{ij} was
971 established between two genes i and j as the sum over the products of their likelihoods
972 of each ordered pattern. A single-link agglomerative clustering of genes, in which a
973 gene will join a cluster if it has a greater than 50% similarity to any gene within that
974 cluster was then performed based on these similarity scores. We label each cluster
975 according to the predominant ordered pattern with high likelihood amongst the genes
976 that comprise it. The change in size of these clusters over time is shown for the major
977 clusterings in Figure 6.
978
979 We can also find likelihoods on comparisons between pairs of experimental
980 conditions by summing the likelihoods over combinations of patterns. Figure 1B
981 shows the number of genes identified at each time point in a pairwise analysis
982 between the grafted top and grafted bottom samples. The likelihood of symmetric
983 expression (i.e., expression which is equivalent across the graft junction) is calculated
984 as the sum of the likelihoods of all patterns in which the grafted top and grafted
985 bottom samples are equivalent. Conversely, asymmetric expression is calculated as
986 the sum of the likelihoods of all patterns in which the grafted top and grafted bottom
987 samples are not equivalent. Additional sets can be formed by considering the ordering
988 of the grafted top and grafted bottoms samples. Sets of genes are identified at each
989 time point with an FDR of less than 0.05 and a likelihood of symmetric/asymmetric
990 expression greater than 50%. Genes in this analysis were only included if they were
991 differentially expressed relative to intact samples.
992
993 **Gene overlap analyses of up and down regulated genes**
994 To measure if the ratio of up- and down- regulated genes from a previously published
995 dataset is significantly different to the ratio of up- and down-regulated genes in our
996 grafting dataset we only took into account genes that are differentially expressed at a
997 certain time point. A gene was called differentially expressed at a certain time point if
998 the marginal likelihood, calculated by BaySeq, was greater than 0.9 and if the
999 absolute $\log_2$-foldchange was greater than 1. Hence, we only consider genes that are
1000 significantly two-fold up- or down- regulated. This definition of differentially

1001 expressed genes was also used to filter the published datasets according to the
1002 expression values in our transcriptome dataset. Hence, some genes were filtered out
1003 from the original published datasets because they did not show a significant up- or
1004 down- regulation during a certain time point in our expression data based on our
1005 criteria. The histograms (Figure 2,4, 5, S4, S7, S8, S9) show the relative number of
1006 up- and down- regulated genes from the published datasets during a certain time point
1007 and a certain condition (separated top, separated bottom, grafted top, grafted bottom)
1008 based on the number of genes in the published dataset after filtering. To calculate the
1009 significance of the difference of the ratios between the published DEGs and all up-
1010 and down- regulated genes, we used a two-sided Fisher's exact test. To correct for
1011 multiple testing we used the Benjamini-Yekutieli (BY) correction method. Hence, the
1012 asterisks in the barplots highlight that the corrected p-value is below 0.05.
1013
1014 **Dealing with probe ids from microarray datasets**
1015 Due to the fact that some published datasets only used probe ids instead of gene ids to
1016 represent their differentially expressed genes we first had to match these probe ids to
1017 their corresponding gene ids. This step was done with the R package biomartr (66). If
1018 one probe id matched more than one gene id we used all the corresponding gene ids
1019 and tested afterwards if these genes were actually differentially expressed in our
1020 dataset. In some cases, one probe id was represented by more than one gene id.
1021 Hence, some gene sets contained slightly more gene ids than published probe ids. In
1022 contrast, some probe ids did not match to currently existing gene ids. Hence, some
1023 gene sets contained slightly fewer gene ids than published probe ids.
1024
1025 **Gene overlap analyses of gene sets involved in graft formation**
1026 For this analysis we used the previously calculated gene sets from baySeq of
1027 differentially expressed genes clustered for each time point into grafted bottom,
1028 grafted top, and grafted bottom and top and calculated their overlap to previously
1029 published data sets. The significance of the number of overlapping genes between the
1030 grafted samples and the published datasets was determined by a one-sided Fisher's
1031 exact test, to prove if the overlap is greater than expected. The resulting p-values were
1032 corrected for multiple testing by using the Benjamini-Yekutieli method.

1033 This procedure was also applied to generate Table 1 to study the overlaps of
1034 symmetrically and asymmetrically expressed genes in the grafting dataset with
1035 previously published sugar-responsive genes.
1036
1037 **GO enrichment analysis**
1038 The gene ontology analysis (GO) enrichment analysis on grafting-specific genes was
1039 done with a customized R script using the package GOstats (65). Gene ontology
1040 annotation was used from the Bioconductor package org.At.tair.db (67). The p-values
1041 calculated by a hypergeometric test were corrected for multiple testing with the
1042 Bonferroni correction. A GO category was called enriched if the corrected p-value
1043 was below 0.05.
1044
1045 A detailed description, the required data and R scripts to reproduce the clustering
1046 (hierarchical clustering and PCA), the statistical analyses regarding the overlap
1047 studies, and the GO enrichment analysis are available via the GitHub repository
1048 (https://github.com/AlexGa/GraftingScripts).
1049
1050 **Breaking force measurements**
1051 A micro-extensometer was used for all breaking force graft measurements according
1052 to a previously published protocol (62). Briefly, grafted Col-0 plants were attached to
1053 a moving plate using tough tags (0.94X0.50 inches, white, 679 catalog no. TTSW-
1054 1000, DiversifiedBiotech) and a cyanoacrylate glue. The plates were moved apart and
1055 the force was measured using a force sensor (Futek LSB200 10g load cell, Futek
1056 Inc.). Upon breaking the force dropped. The maximum force measured before
1057 breaking was recorded. Images were captured using a Leica SP5 confocal microscope.
1058 During the experiment single z-place images were captured and made into a movie in
1059 ImageJ.
1060
1061 To measure levels of tissue contamination between grafted top and bottom, grafted
1062 green fluorescent protein(*p35S::GFP*) and red fluorescent protein (*pUBQ10::PM-*
1063 *tdTomato*) expressing plants were pulled apart manually and imaged on a Zeiss LSM
1064 700 confocal microscope. Z-stacks comprising the majority of the hypocotyl were
1065 made and regions approximately 0.5mm above and 0.5mm below the graft junction
1066 were made into projections using the average intensity function on FIJI software, and

then the mean intensity quantified for red and green channels on FIJI. The mean intensity of one colour in the tissue tested for contamination was divided by the mean intensity of the same colour in the tissue expressing that transgenes to get a percentage of contamination. A region away from the cut site was also quantified to get a percentage of spectral overlap between red and green channels. The percent spectral overlap was then subtracted from the percentage contamination to get an overall percentage for how much contamination was present.

1101 **Movie S1. Breaking weight measurements at the graft junction, related to figure**
1102 **1.** A micro-extensometer pulled apart an Arabidopsis plant 9 days after grafting while
1103 force was measured using a force sensor and images captured using a Zeiss Sp5
1104 confocal microscope. Here, the transmitted light field is shown. Shoot is on top and
1105 root towards the bottom.
1106
1107 **Movie S2. Breaking weight measurements at the graft junction, related to figure**
1108 **1.** A micro-extensometer pulled apart an *Arabidopsis* plant 9 days after grafting while
1109 force was measured using a force sensor and images captured using a Zeiss Sp5
1110 confocal microscope. The same experiment is shown in Movie S1, but here, the GFP
1111 and RFP channels are shown. Shoot is on top (*p35S::GFP* expressing) and root
1112 towards the bottom (*pUBQ10::PM-tdTomato* expressing).
1113
1114 **Table S1.** Details of previously published datasets used to compare to the grafting
1115 datasets.
1116
1117 **Table S2.** Numbers and categories associated with the Bayseq analysis. Categories
1118 are defined as grafted (Col:Col), separated (Col_cut) or intact (ungrafted).
1119
1120 **Table S3.** GO analysis for biological process (BP). Shown are the top 20 BP GO
1121 terms for the grafting-specific genes. Time point selected are those when there are the
1122 most genes in grafted bottom samples (48hrs) or grafted top + bottom samples (120
1123 hrs).
1124
1125 **Table S4.** Normalised reads for all the protein-coding genes in the datasets. By
1126 entering the ATG number of interest, a plot is made which shows a differential gene
1127 expression profile for the gene of interest.
1128
1129 **Dataset S1.** Details of the Bayesian segmentation analysis providing files containing
1130 the ATGs for every cluster for every time point.
1131

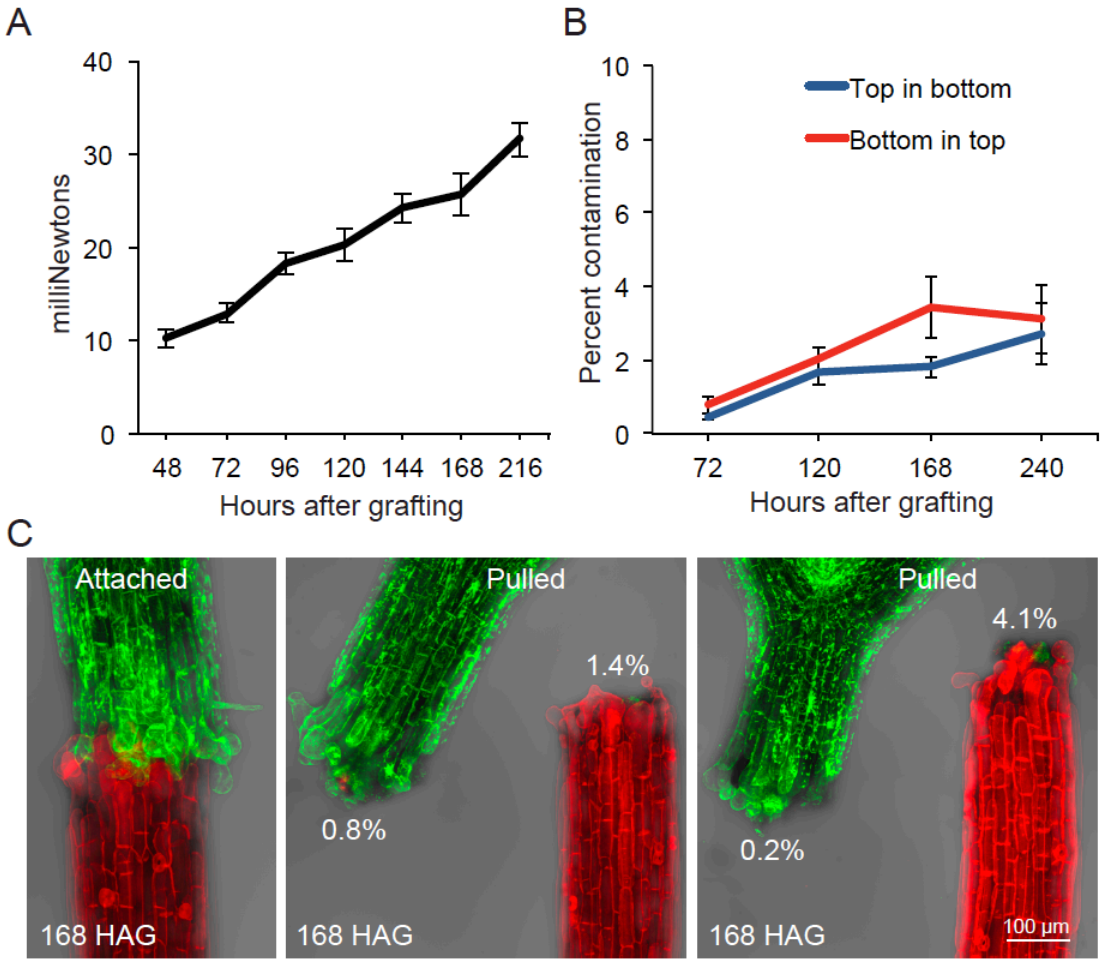


Figure S1. Experimental approach, related to figure 1. (A) The breaking weight required to pull apart grafted plants measured by a micro-extensometer. (B-C) Contamination of top in bottom or bottom in top was less than 4% as measured by grafting green fluorescent protein-expressing plants (*p35S::GFP*) to tomato fluorescent protein-expressing plants (*pUBQ10::PM-tdTomato*) and measuring amounts of fluorescence at the different wavelengths of emission in the top segments relative to the bottom. Images show different plants prior to and after pulling with the percent contamination indicated.

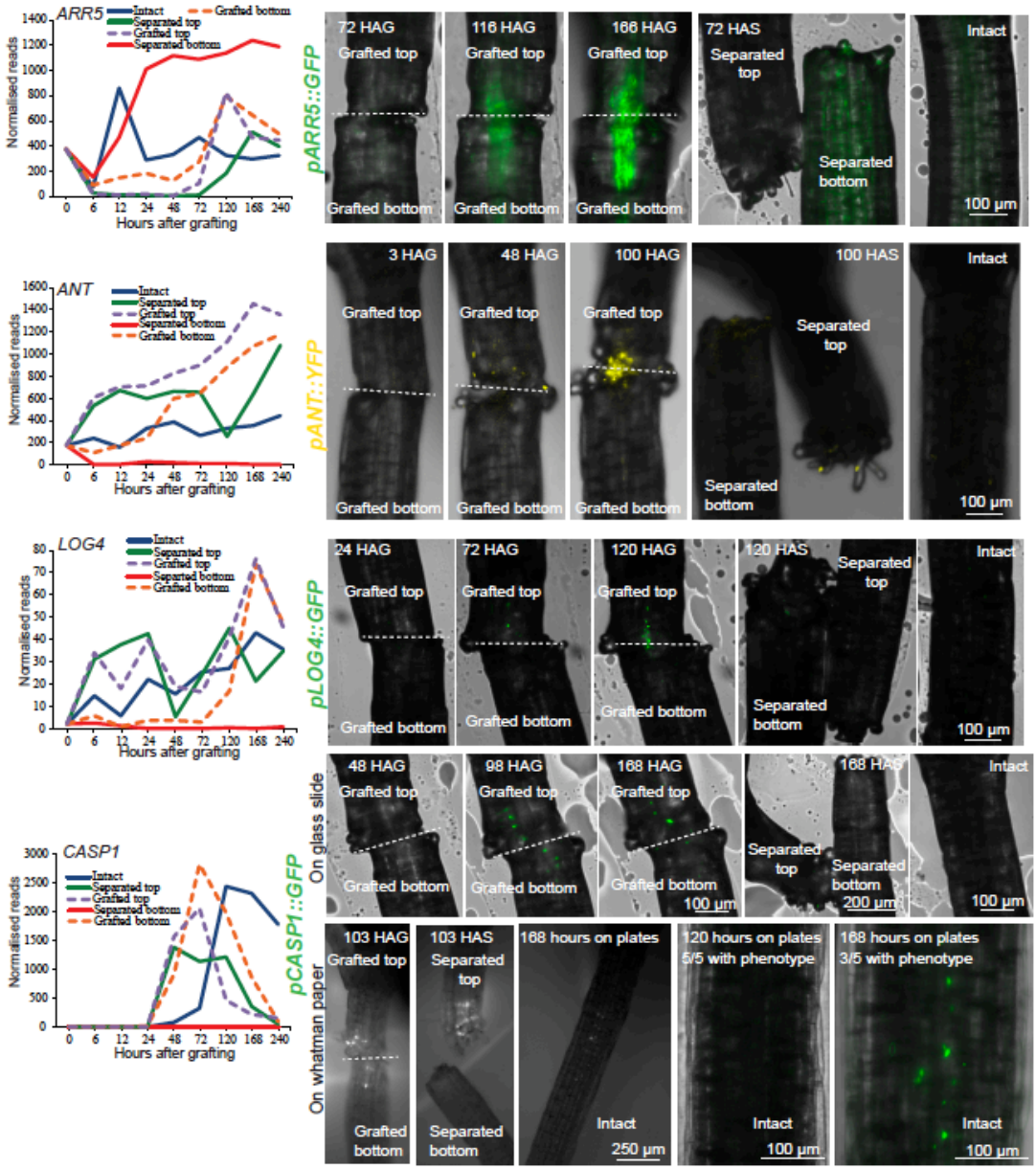


Figure S2. Comparison between RNAseq transcriptome expression profiles and transcriptional fluorescent reporters, related to figure 1. RNAseq expression profiles for various genes upregulated during graft formation were plotted for intact, separated and grafted samples (left panels). Transcriptional reporter-expressing plants were cut and separated, cut and grafted, or left intact. After cutting, plants were imaged and z-projections made at various time points (right panels). For *pCASP1::GFP*, we did not observe a signal in intact plants grafted on glass slides (see Materials and Methods), but observed a signal with 3/5 plants 7 days after grafting on Whatman-nylon membrane, the same condition used for transcriptome library preparation. HAG, hours after grafting. HAS, hours after separation. Dashed lines denote the graft junction.

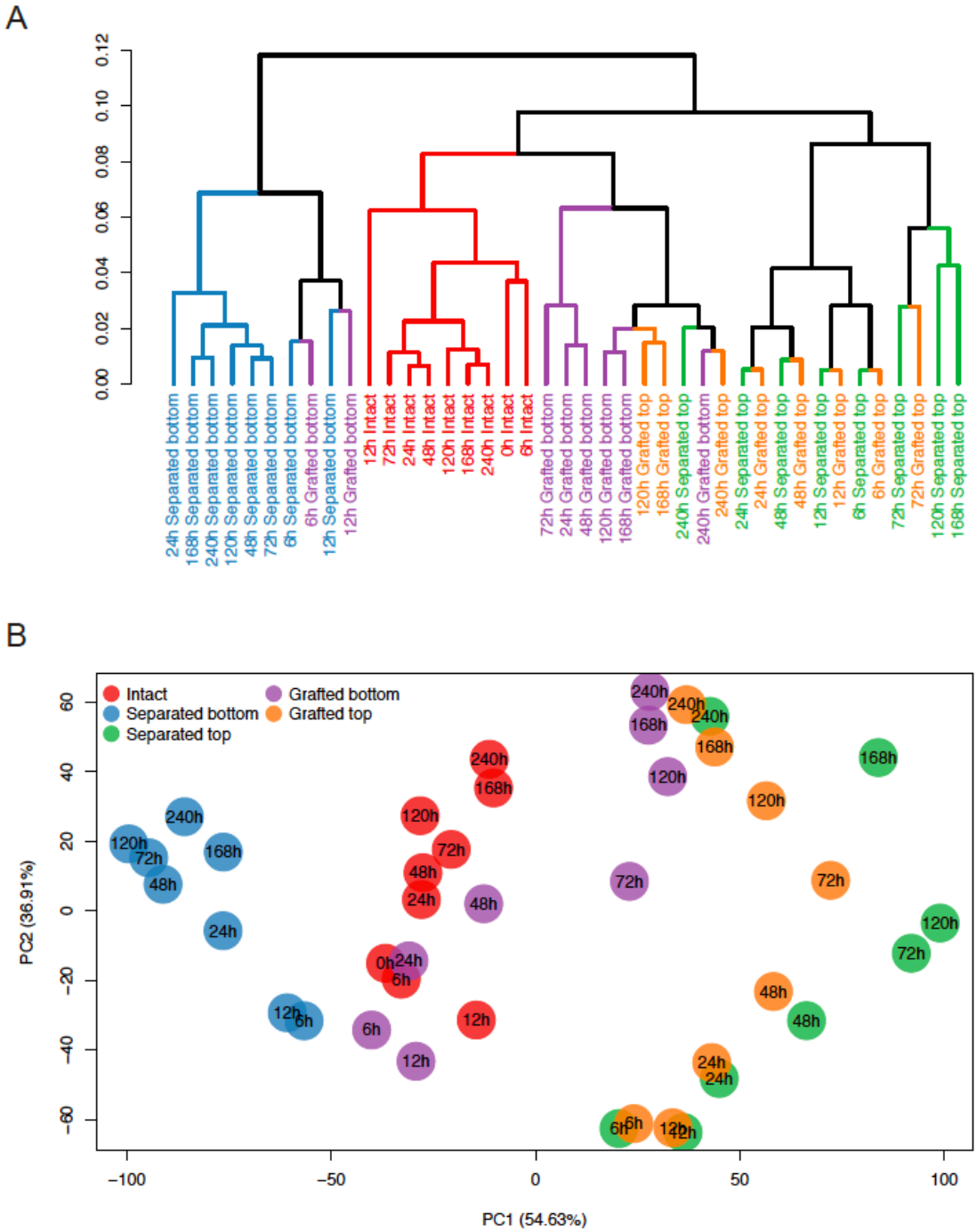


Figure S3. Hierarchical cluster analysis and principle component analysis (PCA) for the transcriptome libraries, related to figure 1. (A) Hierarchical clustering of samples based on log10 transformed TPM values. Similarity between samples was measured by 1 – spearman correlation coefficient. (B) PCA of expression data shows clustering of similar samples. Especially the grafted top and grafted bottom samples are very similar from 120 hours onwards.

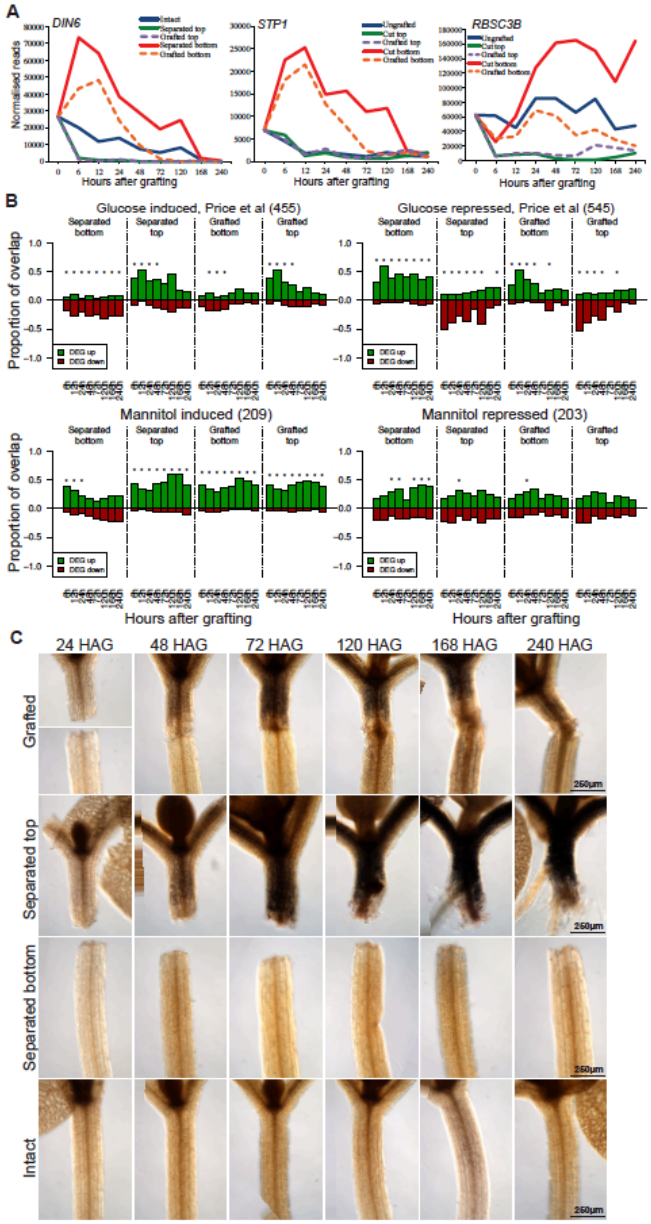


Figure S4. Asymmetry is a feature observed with sugar response and starch accumulation, related to figure 2. (A) Expression profiles for RNAs coded by sugar-repressed genes (*DIN6*, *STP1*) or photosynthetic gene (*RBSC3B*) were plotted for intact, separated and grafted samples. (B) Transcriptional overlap between previously published mannitol-induced, mannitol-repressed, sugar-induced or sugar-repressed genes and our datasets. The numbers in brackets represents the number of glucose or mannitol-responsive genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0 for differentially expressed genes (DEG) up or down regulated in our dataset relative to intact samples. Asterisks represent a significant difference ($p < 0.05$) between the ratio of up- and down- regulated genes in a previously published transcriptome dataset compared to the ratio of all up- and down-regulated genes in our grafting dataset at a certain time point. (C) Lugol staining of grafted plants at various time points reveals dark brown staining associated with starch accumulation. Upper grafted panels are the same as those present in Figure 2 and shown here to compare with controls. HAG, hours after grafting.

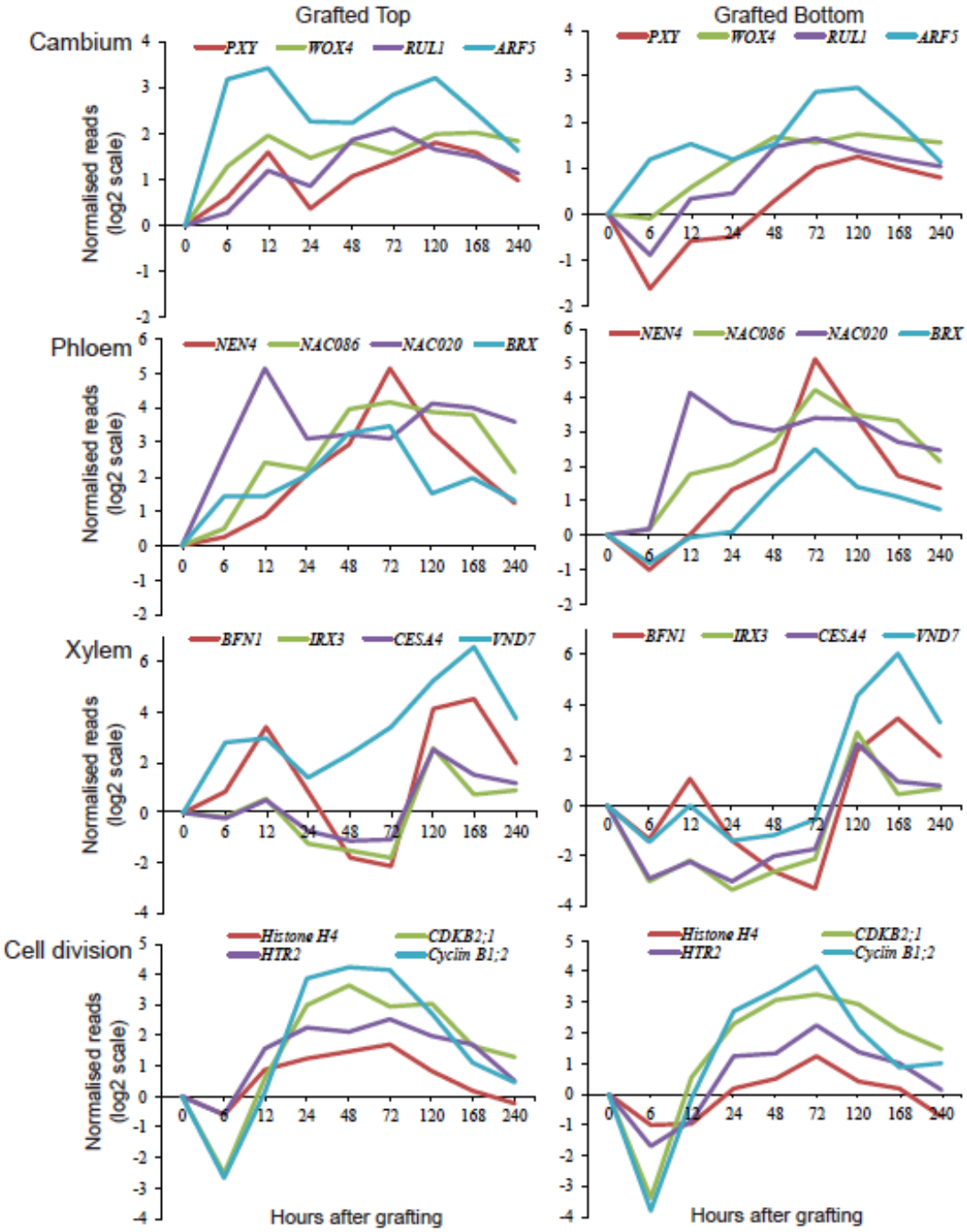


Figure S5. Transcriptional dynamics of genes associated with cambium, phloem, xylem and cell division, related to figure 3. Expression profiles in grafted tops or grafted bottoms for various genes of interest were plotted, normalised to intact samples and plotted on a $\log_2$ scale.

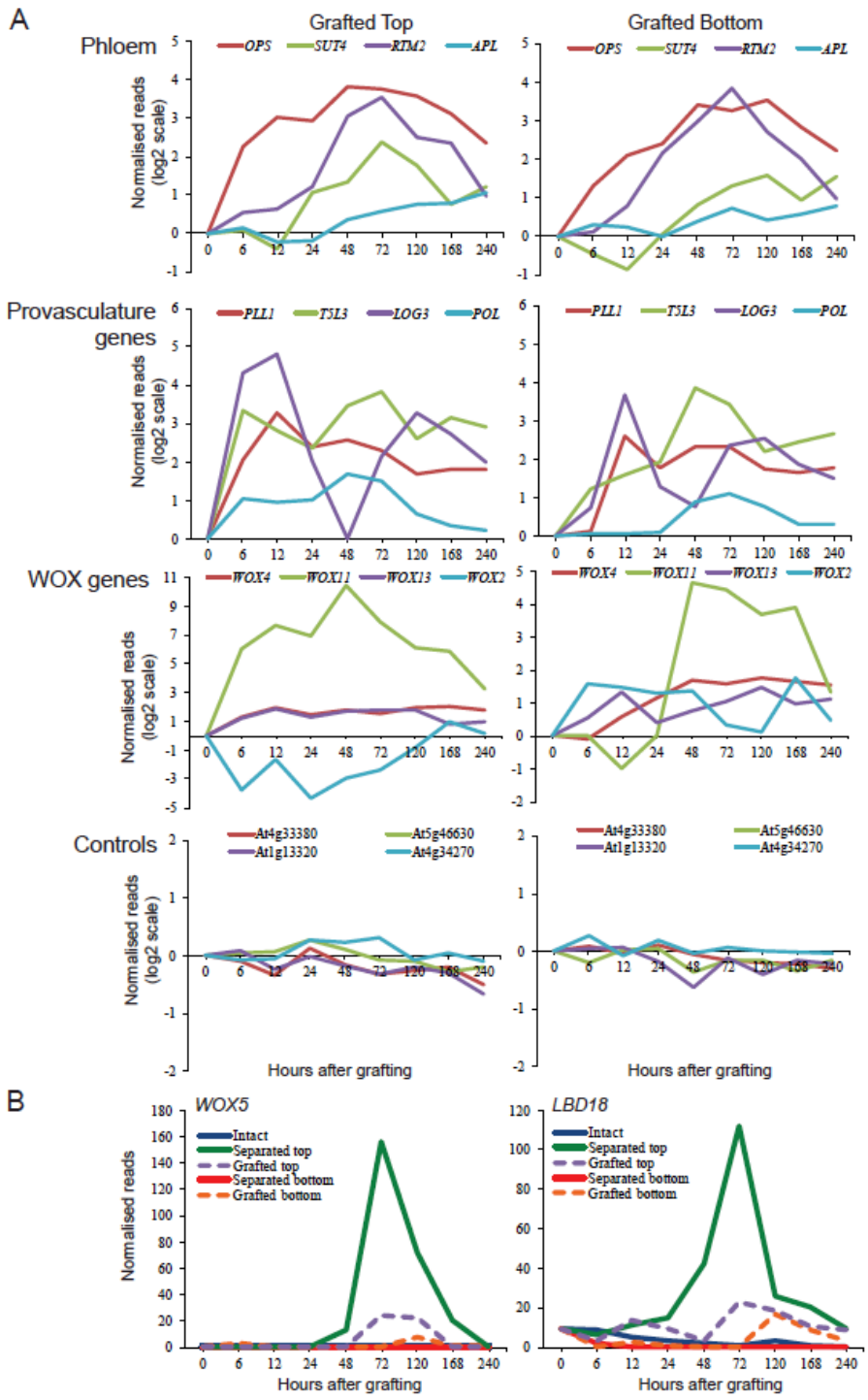


Figure S6. Accumulation of RNA from genes associated with provasculature and phloem, and of RNAs from WOX transcription factors, related to figure 3. (A) RNA accumulation profiles for various genes of interest were plotted over time as measured in grafted tops or grafted bottoms, normalised to intact samples and plotted on a $\log_2$ scale. The grafting datasets could also be used to investigate the transcriptional dynamics of related genes, such as the sequential activation of WOX transcription factors at the graft junction. (B) Expression levels of a primary root specific transcript (of the *WOX5* gene) or a lateral root specific transcript (from the *LBD18* gene) were plotted over time for intact, separated and grafted samples.

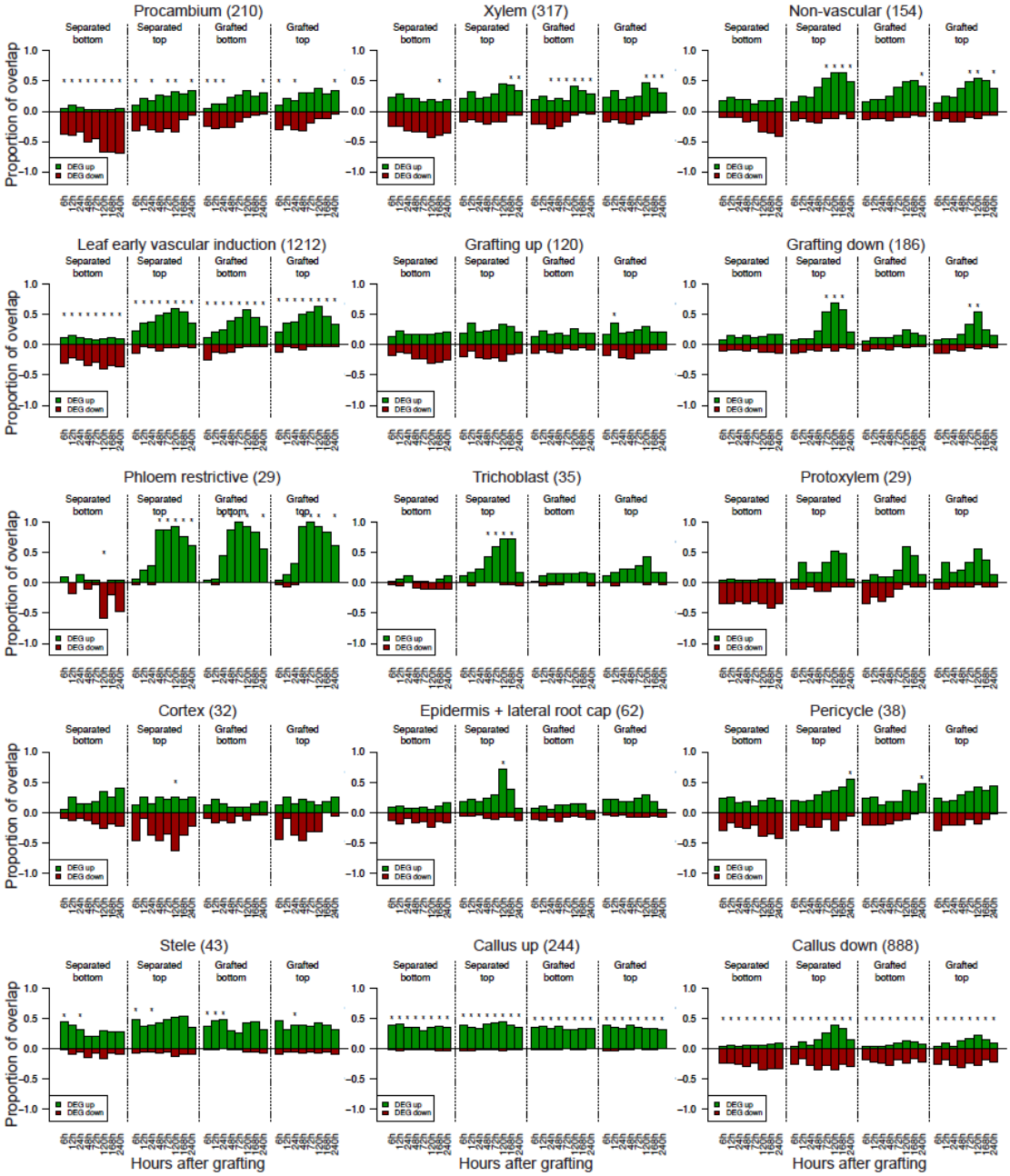


Figure S7. Transcriptional overlap between previously published datasets and the grafting datasets, related to Figure 4. Genes whose transcripts are associated with various cell types or biological processes were taken from previously published datasets (see Table S1) and compared to the datasets generated here. The number in brackets represents the number of cell type-specific or process-specific genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0 for differentiation expressed genes (DEG) up- or down-regulated in our dataset relative to intact samples. Asterisks represent a significant difference ($p < 0.05$) between the ratio of up- and down- regulated genes in a previously published transcriptome dataset compared to the ratio of all up- and down- regulated genes in our grafting dataset at a certain time point.

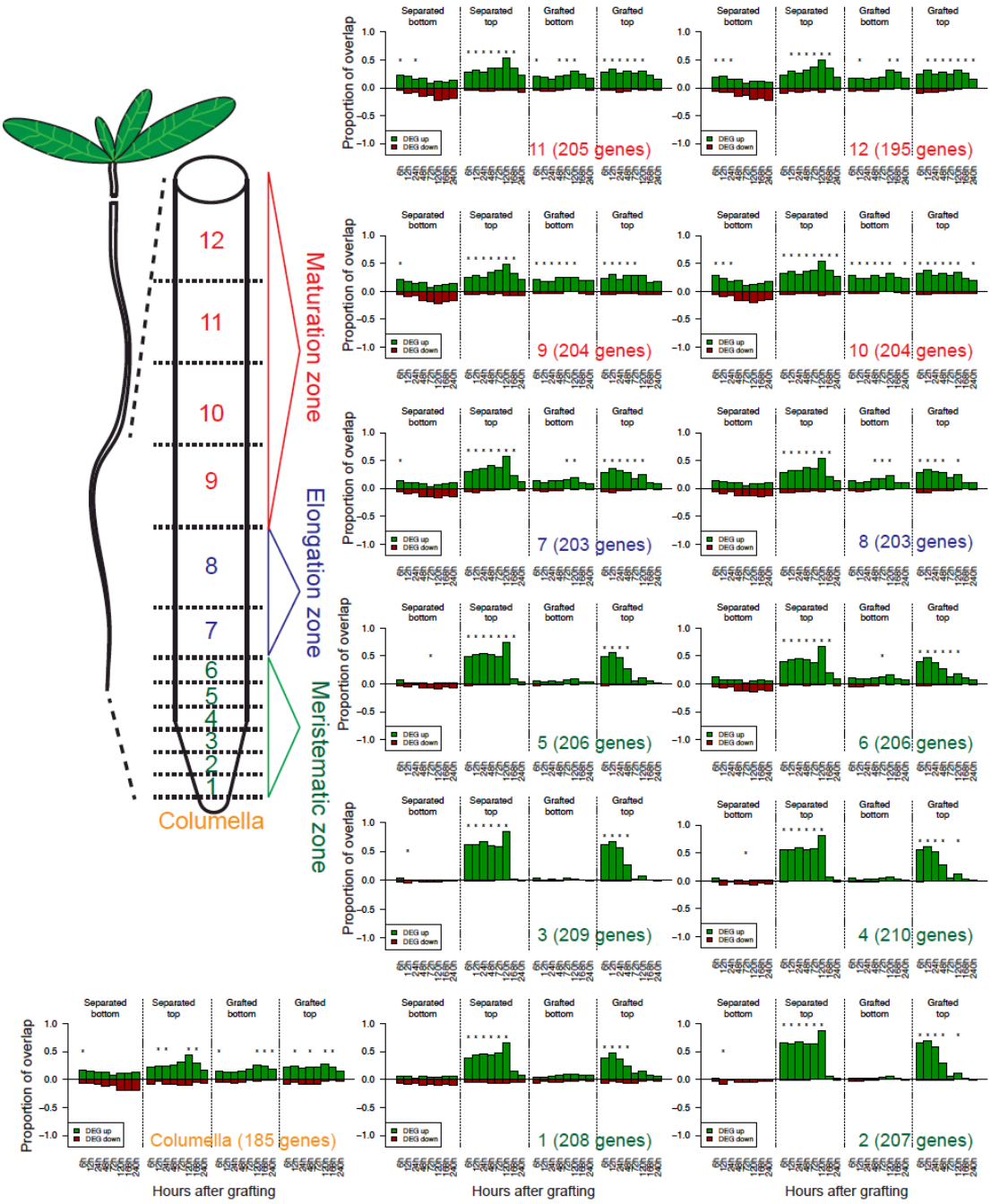


Figure S8. Transcriptional overlap between previously published root transcriptomic datasets and the grafting RNA-seq dataset, related to Figure 4. Genes whose transcripts are associated with various root regions were taken from previously published datasets (see Table S1) and compared to the datasets generated here. Root layer is indicated in the bottom right of each histogram with reference to the cartoon. The number in brackets represents the number of root layer-specific genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0 for differentiation expressed genes (DEG) up- or down-regulated in our dataset relative to intact samples. Asterisks represent a significant difference ($p < 0.05$) between the ratio of up- and down- regulated genes in a previously published transcriptome dataset compared to the ratio of all up- and down- regulated genes in our grafting dataset at a certain time point.

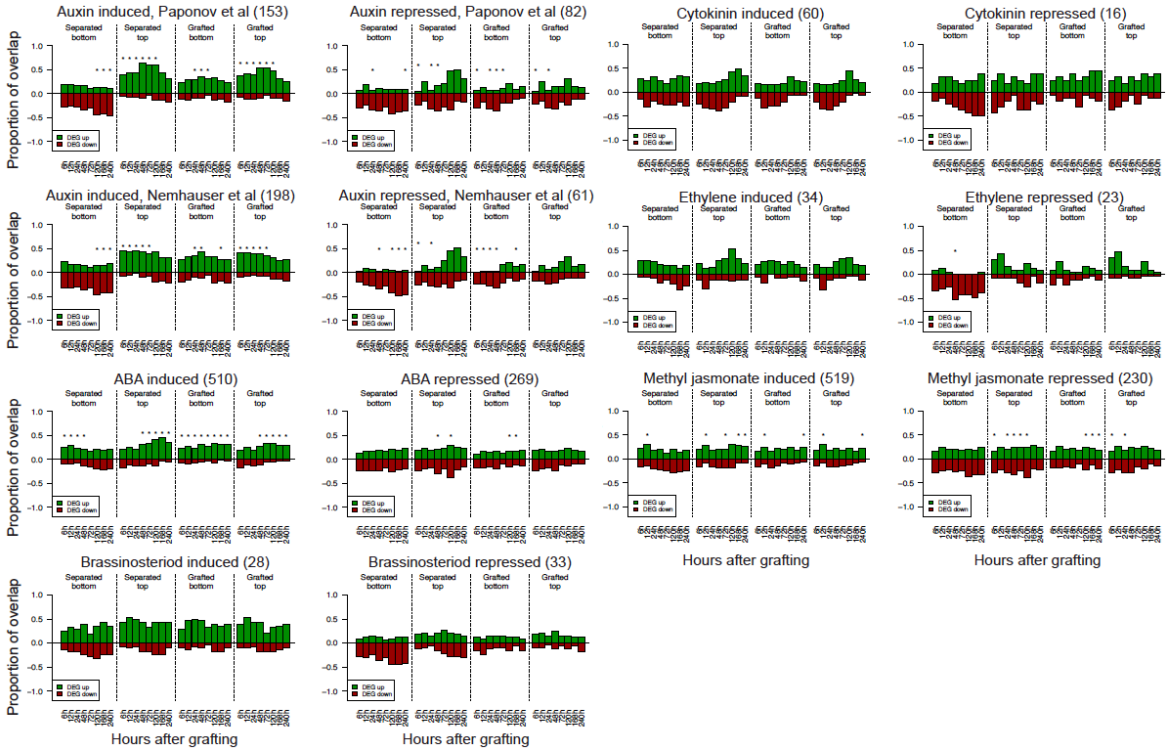


Figure S9. Transcriptional overlap between previously published hormone responsive RNA datasets and the grafting RNA datasets, related to Figure 5. Genes whose differential expression is associated with various hormone responses were taken from previously published datasets (see Table S1) and compared to the genes represented in the RNA-seq datasets generated. The number in brackets represents the number of cell type-specific or process-specific genes identified in the previous dataset, and overlap is presented as a ratio out of 1.0 for differentiation expressed genes (DEG) up- or down-regulated in our dataset relative to intact samples. Asterisks represent a significant difference ($p < 0.05$) between the ratio of up- and down- regulated genes in a previously published transcriptome dataset compared to the ratio of all up- and down- regulated genes in our grafting dataset at a certain time point.

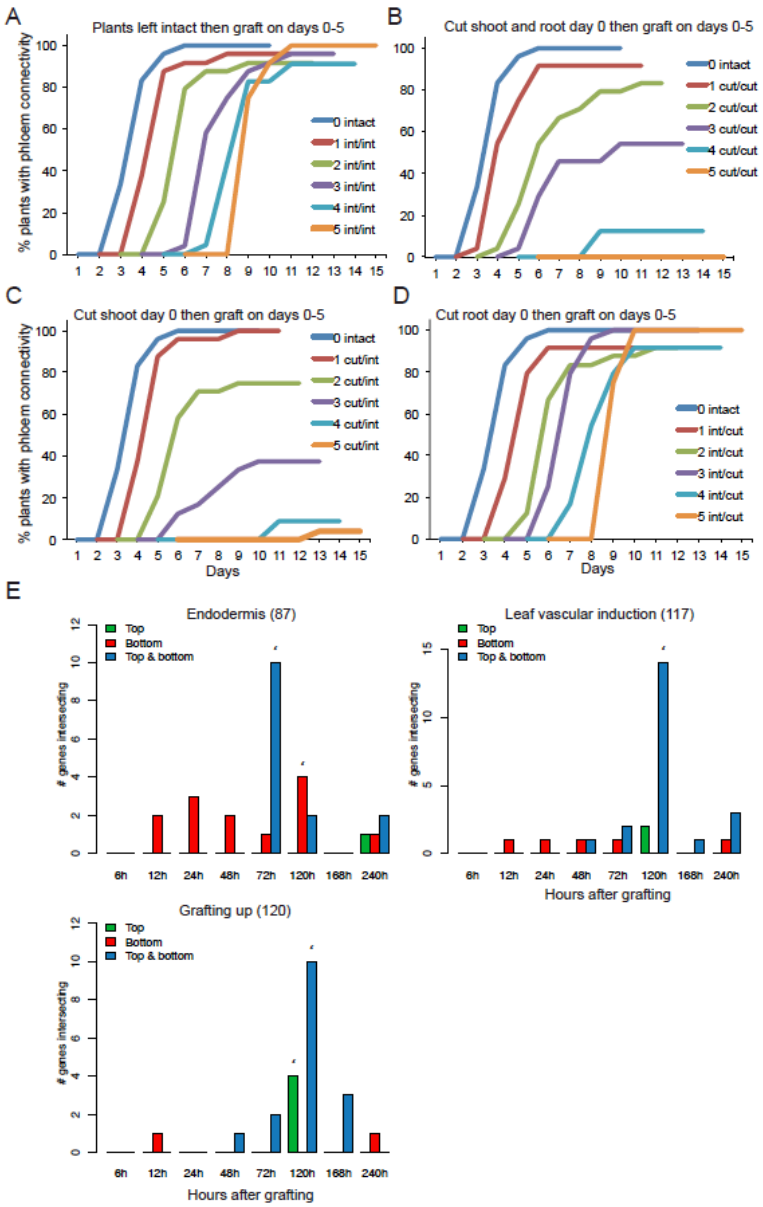


Figure S10. A subset of genes is differentially expressed only during graft formation, related to Figure 7. (A) *pSUC2::GFP* expressing *Arabidopsis* shoots were grafted to Col-0 wild type roots 0-5 days after transferring to grafting plates and kept intact until grafting (intact treatment). GFP movement to the roots for phloem connection assays was monitored over 7 days. (B) Plants were cut at day 0 and grafted at days 0-5 after cutting. (C) Shoots were cut at day 0 and grafted at days 0-5 after cutting. Roots were kept intact until immediately before grafting. (D) Roots were cut at day 0 and grafted at days 0-5 after cutting. Shoots were kept intact until immediately before grafting. (E) Genes differentially expressed in grafted tops and grafted bottoms show overlap with previously published genes whose transcripts are associated with the endodermis, vascular induction and grafting (see Table S1 for treatment information). Asterisks represent significant high overlap ($p < 0.05$) of previously published gene sets that are also differentially expressed in the grafted samples at a certain time point.