the plant journal

SEB Society for Experimental Biology

The Plant Journal (2013) **75,** 104–116 doi: 10.1111/tpj.12204

Self-rescue of an EXTENSIN mutant reveals alternative gene expression programs and candidate proteins for new cell wall assembly in *Arabidopsis*

Prasenjit Saha[1,†], Tui Ray[2], Yuhong Tang[2], Indrajit Dutta[1,†], Nicole R. Evangelous[1], Marcia J. Kieliszewski[3], Yuning Chen[3] and Maura C. Cannon[1,*]

[1]*Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, MA 01003, USA,*
[2]*Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA, and*
[3]*Department of Chemistry and Biochemistry, Ohio University, Athens, OH 45701, USA*

Received 4 December 2012; revised 21 March 2013; accepted 9 April 2013; published online 11 April 2013.
*For correspondence (e-mail mcannon@biochem.umass.edu).
†Present address: Department of Plant Sciences, University of California, Davis, CA 95616, USA.

SUMMARY

Plants encode a poorly understood superfamily of developmentally expressed cell wall hydroxyproline-rich glycoproteins (HRGPs). One, EXTENSIN3 (EXT3) of the 168 putative HRGPs, is critical in the first steps of new wall assembly, demonstrated by broken and misplaced walls in its lethal homozygous mutant. Here we report the findings of phenotypic (not genotypic) revertants of the *ext3* mutant and in-depth analysis including microarray and qRT-PCR (polymerase chain reaction). The aim was to identify EXT3 substitute(s), thus gaining a deeper understanding of new wall assembly. The data show differential expression in the *ext3* mutant that included 61% ($P \leq 0.05$) of the HRGP genes, and ability to self-rescue by reprogramming expression. Independent revertants had reproducible expression networks, largely heritable over the four generations tested, with some genes displaying transgenerational drift towards wild-type expression levels. Genes for nine candidate regulatory proteins as well as eight candidate HRGP building materials and/or facilitators of new wall assembly or maintenance, in the (near) absence of *EXT3* expression, were identified. Seven of the HRGP fit the current model of EXT function. In conclusion, the data on phenotype comparisons and on differential expression of the genes-of-focus provide strong evidence that different combinations of HRGPs regulated by alternative gene expression networks, can make functioning cell walls, resulting in (apparently) normal plant growth and development. More broadly, this has implications for interpreting the cause of any mutant phenotype, assigning gene function, and genetically modifying plants for utilitarian purposes.

Keywords: cell wall assembly, phenotype, gene expression, extensins, proline-rich proteins, hydroxyproline-rich glycoprotein, transgenerational drift, *Arabidopsis thaliana*.

INTRODUCTION

Plant cells have walls to keep each cell intact giving it shape and the turgidity needed to grow and develop. Structurally these walls are composites of interpenetrating polymers, of cellulose, hemicellulose, and pectins, while glycoprotein based polymers, the extensins (EXT) (Lamport *et al.*, 2011) and possibly the proline-rich proteins (PRP) (Fowler *et al.*, 1999; Bernhardt and Tierney, 2000), are minor albeit significant components. EXTs and PRPs as well as arabinogalactan proteins (AGP) are subfamilies of the cell wall hydroxyproline-rich glycoprotein (HRGP) superfamily (Showalter *et al.*, 2010; Kieliszewski *et al.*, 2011), and are expressed developmentally (Zimmermann *et al.*, 2004). EXTs are defined by the presence of multiple SO3-5, YXY and VYK motifs (O = Hyp, hydroxyproline; Y = tyrosine; V = valine; K = lysine). Of the 65 (putative) EXTs in Arabidopsis, only one, EXT3 (At1g21310), was shown to have a vital role (Hall and Cannon, 2002), while seven members are associated with non-lethal root hair phenotypes (Baumberger *et al.*, 2003; MacMillan *et al.*, 2010; Velasquez *et al.*, 2011). PRPs are defined by multiple variant motifs of P, O, V, Y and K. There are 18 (putative) PRPs in Arabidopsis; four of them have experimental data available showing that they localize to the cell wall (Fowler *et al.*, 1999; Bernhardt and Tierney, 2000). The 85 (putative)

© 2013 The Authors
The Plant Journal © 2013 John Wiley & Sons Ltd

Table 1 Fold change (log2) in expression of classical *EXT* genes ($P \leq 0.05$), arranged in EXT classification groups (by qRT-PCR)

EXT Group, features	Name	Gene ID	Mutant versus WT	*ANP4*-F2 versus mutant	*ANP4*-F2 versus WT	*ANP4*-F3 versus WT-2	*ANP10*-F3 versus WT-2
Group I: Idt poor	*EXT1*	At1g76930	3.33	−2.50	0.83	ns	0.40
	EXT23	At5g19810	−3.07	4.14	1.08	ns	ns
Group IIa: Idt-rich, simple	*EXT3*[a]	At1g21310	−4.05	ns	−6.20	−7.16	−7.31
	EXT18	At1g26250	ns	ns	ns	ns	−2.58
	EXT19	At1g26240	5.03	−5.13	ns	ns	ns
	EXT20	At4g08370	1.89	−0.99	ns	ns	ns
	EXT21	At2g43150	0.44	0.65	1.09	ns	0.92
	EXT22	At4g08380	ns	ns	ns	ns	ns
Group IIb: Idt-rich, SPSP motifs	*EXT6*	At2g24980	−6.68	5.39	−1.29	ns	ns
	EXT7	At4g08400	−6.76	5.19	−1.57	ns	ns
	EXT8	At4g08410	−7.12	5.38	−1.73	ns	−0.28
	EXT9	At5g06630	−6.26	5.04	−1.22	ns	ns
	EXT10	At5g06640	−6.18	4.84	−1.35	ns	ns
	EXT11	At5g49080	−0.86	0.72	ns	ns	−1.05
	EXT12	At4g13390	−7.44	5.87	−1.57	ns	ns
	EXT13	At5g35190	−8.29	6.55	−1.74	ns	−0.66
Group IIc: Idt-rich, SPSP motifs, 1 tri-C motif	*EXT2*	At3g54590	−5.82	4.35	−1.48	ns	ns
	EXT15	At1g23720	−5.18	3.73	−1.45	ns	ns
	EXT16	At3g28550	−5.70	4.23	−1.47	ns	ns
	EXT17	At3g54580	−2.48	ns	−1.45	ns	ns

[a]Mutated in the *rsh* mutant and *ANP* lines; ns, not significant.

AGP-type HRGPs are defined by being Pro-rich, by AG-glycomodules, and arabinosylated Hyp-Hyp-glycomodules (Seifert and Roberts, 2007; Showalter *et al.*, 2010; Kieliszewski *et al.*, 2011), and in most cases with glycosylphosphatidylinositol (GPI) anchors binding them to the plasma membrane (Youl *et al.*,1998; Eisenhaber *et al.*, 2003). These acidic, more highly glycosylated AGPs have been implicated in growth and development (Ellis *et al.*, 2010). In particular, a sub-group, the chimeric AGP containing fascicilin-like domain(s) (FLA) are associated with normal cell expansion (Shi *et al.*, 2003), integrity of the cell wall matrix (MacMillan *et al.*, 2010), and lateral root and shoot development from tissue culture (Johnson *et al.*, 2011). The deficit of available mutants and the recalcitrance of HRGPs to genetic and biochemical studies have been slowing factors in understanding what they do and how they do it.

There are 20 classical EXTs in Arabidopsis (Table 1, columns 1–3); they are amphiphilic with regular periodicity of repetitive motifs. The current model for EXT polymer formation and its role in cell wall assembly is: (i) these amphiphilic glycopeptides favor like-with-like recognition; (ii) strict periodicity enhances their staggered lateral alignment creating a dendritic shaped structure; (iii) Tyr covalent crosslinks stabilize the polymer, (Brady *et al.*, 1996, 1998; Held *et al.*, 2004); and (iv) an EXT polymeric network with its extreme regularity of positively charged amino acids (K and H) provides a scaffold to attract negatively charged pectic acids, thus resulting in the electrostatic deposition of a pectic polymeric network via ion pairing between pectic carboxyls and epsilon amino groups of EXTs. Evidence so far supports this model (Cannon *et al.*, 2008; Valentin *et al.*, 2010), as does the fact that EXTs and pectins are among the first cell wall material detected in new cell walls, i.e. the cell plate (Hall and Cannon, 2002). This model places a Tyr-rich EXT protein network at the start of cell wall assembly processes. While there is evidence to support the role of repetitive lone Tyr and Idt motifs as well as positively charged amino acids (Lys and His) in EXT network formation and function, there is no indication yet of a function for the SPSP nor the lone tri-Cys motifs.

Most EXTs and other HRGPs are of the non-classical type: they are chimeras with non-HRGP domains or hybrids with other HRGP subfamilies (Johnson *et al.*, 2003), or short HRGPs with non-typical and irregularly placed motifs (Showalter *et al.*, 2010). Their presence in cell walls indicates significance and it is assumed that their glycomodules bind to other cell wall carbohydrate polymers including network-forming EXTs, but their role(s) are unknown. Long-standing questions are: 'Why are there so many EXT and other HRGPs?' and 'Why has it been so hard to find strong HRGP mutant phenotypes?' The data reported here give new insights that help answer these questions.

Here we report the identification of *ext3* phenotypic revertants, and results of gene expression analyses initially using microarrays followed by qRT-PCR of the *ext3* embryo defective, seedling lethal mutant *root-*, *shoot-*, *hypocotyl-defective* (*rsh*) and its self-revertants, over three

© 2013 The Authors

generations compared with wild type. The main aim was to find substitute(s) for the Tyr-rich network-forming EXT3, and to identify putatively relevant gene expression differences. First we cast a wide net, and then focused on all (putative) 65 *EXT* and 18 *PRP* genes, as well as selections of the following: *AGP*; extensin peroxidase (*PER*), because PERs catalyze intermolecular crosslinking of EXTs containing Tyr (Schnabelrauch *et al.*, 1996; Held *et al.*, 2004); prolyl-4-hydroxylases (*P4H*), because P4Hs catalyzes post-translational conversion of Pro to Hyp (Tiainen *et al.*, 2005); and (v) expansin(-like) (*EXP/EXPL*) genes because, like classical EXTs, the EXPs are non-enzymatic (McQueen-Mason *et al.*, 1992; McQueen-Mason and Cosgrove, 1994), but unlike EXTs they are involved in wall loosening thus enabling plant cell growth. The data obtained in this study support the claim that plants can use different combinations and relative quantities of some HRGPs, regulated by alternative gene expression networks, to build apparently normal walls. Specifically, we identified three non-classical *EXTs (EXT34*, *LRX4*, *PEX2)*, four *PRPs (PRP2*, *-4*, *-11*, and *-15)*, and one *AGP (FLA13)*, as well as seven transcription factors and two kinase genes as candidates for building and/or modulating apparently normal cell walls in the (near) absence of *EXT3* expression.

RESULTS

Self-rescue of *rsh* identified

We previously showed that *rsh* was a recessive mutant, and that the mutated *ext3* gene was responsible (Hall and Cannon, 2002). In this mutant, the *ext3* gene carries a mutating insert that has the kanamycin-selectable marker (Figure 1a). In summary, when plated on kanamycin, progeny of self-fertilized *EXT3/ext3* plants segregate in a 1:2:1 ratio of white (W):green normal looking (G):*rsh* mutant (M), respectively. Because the recessive *ext3* mutant, *rsh*, provides an ideal background to test the functionality of extensin analogs, it became important to examine the progeny of heterozygous *rsh* in more detail. Here we present more detailed analyses of the progeny of heterozygous *rsh*. Starting with 30 heterozygous *rsh* plants (*EXT3/ext3*), in which the zygosity was confirmed by PCR, seed were collected from each individual plant and germinated on kanamycin plates. They segregated in a ratio of 1:2.15:0.78, which is 1:2:1 rounded to whole numbers, i.e. the expected mendelian ratios of (wild type [*EXT3/EXT3*]:heterozygote [*EXT3/ext3*]:homozygote [*ext3/ext3*], respectively) (Table 2). The segregation data were consistent in repetition experiments regardless of whether the parent plant was heterozygous F2 [seed stored since 2001, when *rsh* was first identified (Hall and Cannon, 2002)] or subsequent generations out to the F10 generation.

A genotype analysis of the phenotypically normal looking green plants (G) – presumed to be heterozygous – on

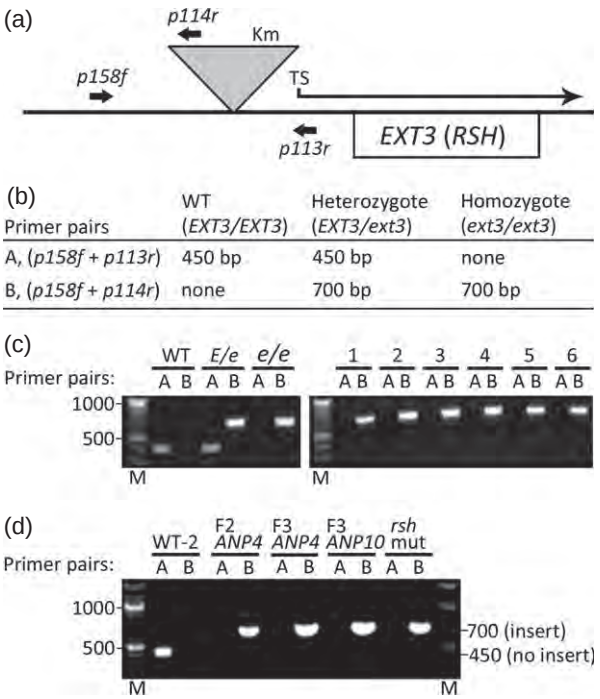

Figure 1. Genetic analysis of the *rsh* mutant and its *ANP* revertants.
(a) Map (not to scale) of the *EXT3* gene showing the enhancer-trap insert, which resulted in the *rsh* mutant phenotype, at −109 bp upstream from the transcription start (TS), and the primers (Table S10) used to determine the presence or absence of the insert.
(b) Scheme to detect the insert.
(c) Amplicons separated by 1% agarose gel electrophoresis. M, molecular size markers (base pairs). Gel on left shows DNA extracted from the controls: WT, *rsh* heterozygote (*E/e*), and *rsh* mutant (*e/e*). Gel on right shows F1 generation of six independently arising phenotypic revertants of *rsh*. Note: they are homozygous for the *ext3* insert.
(d) Amplicons generated as in (c) using DNA extracted from the next generation of wild type (WT-2), the next two generations of *ANP4* (*ANP4*-F2, and *ANP4*-F3), and from the F3 generation of the independently arising *ANP10* (*ANP10*-F3).

Table 2 Segregation of progeny of (PCR) confirmed heterozygous *rsh* plants on kanamycin plates

Number of plants tested	30
Total no. of seed plated	3814
Average no. of seed plated per plant	127.13 (±0.32)[a]
Average no. of white seedlings (W) per plant	31.57 (±0.91)
Average no. of green seedlings (G) per plant	67.90 (±0.44)
Average no. of *rsh* mutants (M) per plant	24.57 (±1.19)
Average no. of non-germinated seed per plant	3.10 (±0.70)
Segregation ratio W/G/M	1/2.15/0.78

[a]Standard deviation in parentheses.

the kanamycin-selection plates above showed that approximately 10% of these seedlings on every plate were homozygous for the mutation-causing insert (Figure 1b,c). As these seedlings arose from the self-fertilization of individual confirmed heterozygous plants, they are each an independent revertant of the *rsh* mutant. The discovery of these apparently normal phenotype (*ANP*) revertants

© 2013 The Authors

Table 3 Phenotypic characteristics of *ANP* lines compared with WT

	WT	*ANP4*-F2		*ANP4*-F3		*ANP10*-F3	
Measurement of	Mean ± SD	Mean ± SD	*P*-value	Mean ± SD	*P*-value	Mean ± SD	*P*-value
Root length (mm)[a]	24.06 ± 2.26	24.26 ± 2.32	0.66	23.84 ± 1.71	0.58	23.40 ± 1.74	0.11
Shoot length (mm)[a]	7.02 ± 1.38	6.94 ± 1.54	0.79	6.92 ± 1.37	0.72	7.12 ± 1.59	0.74
Dry weight (mg)[b]	6.25 ± 0.86	6.38 ± 6.80	0.61	5.92 ± 0.79	0.21	5.90 ± 1.17	0.29
Plant height (cm)[a]	25.52 ± 2.40	25.93 ± 2.80	0.43	24.92 ± 2.05	0.19	25.10 ± 2.30	0.38
No. of silique per plant[a]	144.66 ± 15.55	145.08 ± 16.60	0.90	146.58 ± 17.74	0.57	147.72 ± 14.39	0.31
Silique length (mm)[a]	10.59 ± 1.25	10.43 ± 1.28	0.53	10.39 ± 1.33	0.44	10.35 ± 1.11	0.32
Seed yield per plant (mg)[b]	87.40 ± 14.46	87.65 ± 14.13	0.95	90.55 ± 11.56	0.45	95.50 ± 13.41	0.06
Seed germination (%)[b]	94.50 ± 2.67	94.10 ± 1.69	0.83	95.72 ± 1.69	0.54	96.70 ± 0.71	0.24

SD, standard deviation.
[a]$n = 50$.
[b]$n = 20$.

among the progeny of PCR-confirmed heterozygous *rsh*, is consistent with the segregation ratio of 1:2.15:0.78, and the other data in Table 2. In repetition experiments in which the apparently heterozygous progeny of individual PCR-confirmed-*ext3* heterozygous plants from the F2 to F10 generations were tested, approximately 10% were *ANP* revertants.

The *ANP* lines showed no significant phenotype differences to WT, borne out by the high *P*-values (Table 3). All *ANP* lines tested were shown to have no *EXT3* transcript by RT-PCR, and were propagated for four generations. The stable *ANPs* with the *ext3/ext3* genotype persisted, showing no reversion to a mutant phenotype (Figure 1d). Based on the segregation and genotype data a genetic explanation seems unlikely. If a suppressor mutation existed, one would predict that it would become fixed (all *ext* homozygous progeny would have *ANP* phenotypes) or lost (the *ANP* phenomenon would cease to occur in subsequent generations) from the population. An alternative explanation of the data is an epigenetic model (see later). In either model, EXT3 could be compensated for, by another gene product, hypothetically another structural glycoprotein. Finding its identity would be useful in assigning gene function and towards understanding the role of HRGPs in cell wall assembly; identification of the *ANP* self-revertants provided an opportunity to do this regardless of the cause.

Global gene expression in the *rsh* mutant, *ANP*-revertant and WT

Microarray analysis of the germinated severely defective *rsh* mutant (mut), and seedlings of *ANP* line 4, F2 generation (*ANP4*-F2), showed that approximately 13.9 and 2.8% of the genes, respectively, had altered gene expression levels compared with that of WT (significance set at $P \leq 0.05$) (Figures S1 and S2). With focus on the genes that may be directly relevant to EXT function, it is notable that the top 50 *down*-regulated genes in the mutant compared with WT included (putative) *EXT*, *PRP*, *AGP* and *PER* (Table S1, '*Down in mutant versus WT*'), while the top 50 *up*-regulated genes included *EXT*, *PER* and genes for regulatory proteins (Table S2, '*Up in mutant versus WT*'). Similar comparisons of the *ANP* with WT showed that *EXT*, *AGP*, *PER* and *EXP(L)* genes were among the most differentially expressed (Tables S3 and S4). Three of the seven down-regulated *PER*, one *PRP*, and *RHS13* (ROOT HAIR SPECIFIC) have been shown to co-express with several *EXTs* (Velasquez *et al.*, 2011). It is also notable in the comparison of *ANP4*-F2 with the mutant that a reversal in gene expression levels occurred for most of the top 100 down- and up-regulated genes in the *ANP* (Tables S1 and S2, '*Change in ANP4-F2 versus mutant*'). Homing in on the microarray expression of several cell wall gene families likely associated with wall assembly, the *EXT*, *PRP*, *AGP*, *PER*, and *EXP (L)* families showed the greatest percent of genes with differential expression ($P \leq 0.05$) compared with WT. Several members of these families were either down- or up-regulated in the mutant (Tables 4 and S6 and Figure S3), thus

Table 4 Numbers of (putative) genes tested and numbers showing down-regulation (dn) or up-regulation ($P \leq 0.05$) (dn|up) in the comparisons listed, by microarray analysis (Table S6, Figure S3)

Numbers of genes and comparisons made	*EXT*	*PRP*	*AGP*	*PER*	*EXP, EXPL*	*P4H*
In Arabidopsis genome	65	18	85	106	32	20[a]
Genes represented	58	16	77	96	29	14
Discriminating primers	50	16	76	89	26	14
Differentially expressed	16	11	33	42	19	3
Mutant versus WT: dn|up	10|5	7|4	21|10	29|10	12|5	0|3
ANP4-F2 versus mutant: dn|up	5|9	4|7	8|20	9|30	4|11	3|0
ANP4-F2 versus WT: dn|up	9|3	4|4	17|8	26|8	8|4	1|0

[a]Includes all (putative) *P4H* with alpha subunit domain (InterPro: IPR006620) found in the Arabidopsis genome.

© 2013 The Authors

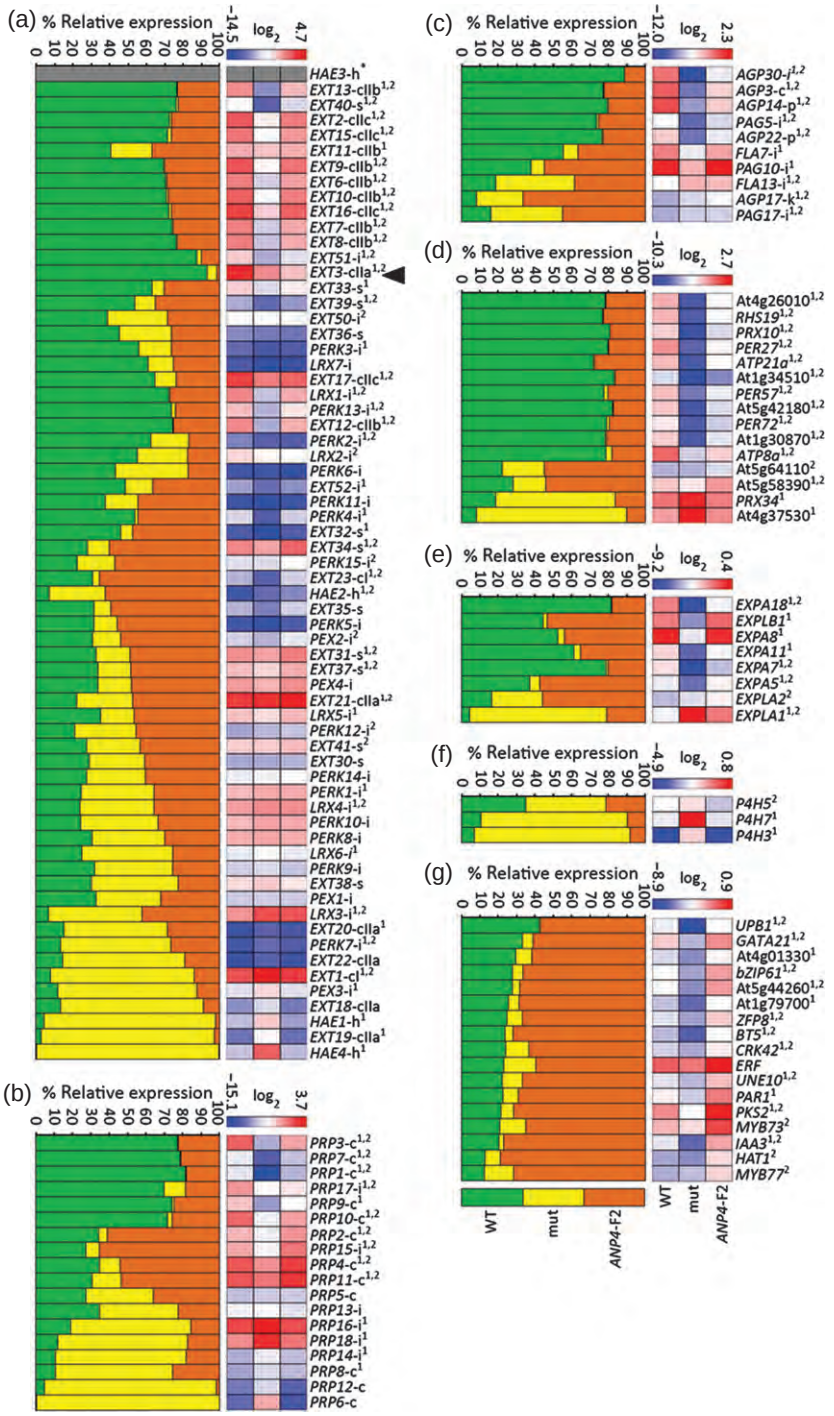

Figure 2. Comparative gene expression in seedlings of wild-type (WT), *rsh* mutant and *ANP4*-F2 (by qRT-PCR).
(a) All 65 *EXT*.
(b) All 18 *PRP*.
(c) 10 *AGP*.
(d) 15 *PER*.
(e) 8 *EXP/EXPL*.
(f) 3 *P4H*.
(g) 17 genes for regulatory proteins (14 DNA binding proteins (DBPs) and three kinases). Each group of genes was ordered according to hierarchical clustering of mean relative expression values calculated based on Pearson correlation coefficients using MultiExperiment Viewer4 (MeV4). Bar graphs represent percentages of the total mean relative expression ratio for each gene in the three samples being compared, i.e. one bar graph row compares expression of an individual gene between the three samples. Arrowhead marks the position of EXT3. Superscript numbers indicate significance ($P \leq 0.05$) of *rsh* mutant versus WT, and *ANP4*-F2 versus WT, respectively (see also Table S7). The heatmaps are $\log_2$ transformed values of these same mean relative expression ratios for each gene used in the bar graphs to show the expression levels of the transcripts. The color scale shows the highest (red) to lowest (blue) transcript expression values obtained. Each heatmap column shows the expression level of each gene within one sample, while each row allows a broad color image comparison of the expression of each individual gene between the three samples. **HAE3*-h consistently showed no expression; unnamed gene, used gene ID. Name tags: -s, short (<182 amino acid residues); -i, chimera; -c, classical; -k, lysine-rich. Name abbreviations also see Table S5.

suggesting wide-ranging consequences of *EXT3* gene expression knock-down.

Homing in on *EXT*s and related genes in the *rsh* mutant versus WT

To further explore the growing number of questions and to verify and extend the microarray data, gene expression in the *rsh* mutant was compared with that of WT by qRT-PCR analysis of all (putative) *EXTs* and *PRPs*, and for selections of genes from the other families of focus (Figure 2 and Table S7, '*mutant versus WT*'). The qRT-PCR data for the most part verified that of the microarray; combined they showed that 43–78% of the focus gene family members: *EXT* (63%), *PRP* (78%), *AGP* (43%), *PER* (47%), *EXP/EXPL* (73%), had differential expression in the mutant versus the WT, compared with approximately

© 2013 The Authors

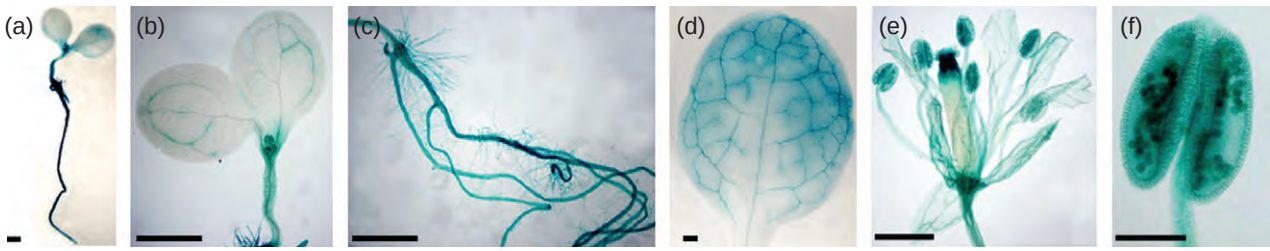

Figure 3. Expression of the GUS reporter gene linked to the *EXT3* promoter in *ANP4*-F2.
(a) Seedling at 6 days after sowing (DAS).
(b) Shoot of seedling at 12 DAS.
(c) Roots of seedling at 12 DAS.
(d) Mature rosette leaf at 30 DAS.
(e) Mature flower.
(f) Single anther. Bar = 1 mm in (a) and (d); 500 μm in (b), (c), (e) and (f).

14% for the total 21 560 genes examined on the microarray. Other randomly selected gene families examined had considerably lower percentages of their genes showing differential expression, including the non-cell wall localized P4H family. The data indicate that a switch occurred from WT to a 'mutant-*ext3* gene expression program' and it included, but was not limited to, the co-ordinated regulation of expression of the gene family members of focus.

Down-regulation and up-regulation patterns of expression in the *rsh* mutant

Apart from *EXT3*, expression of 28 and 13 *EXT* genes, were down-regulated or up-regulated ($P \leq 0.05$), respectively, in the *rsh* mutant compared with WT, by qRT-PCR analysis. The following list are some of the remarkable observations (Table 1, '*mut versus WT*'): (i) of the only two genes in the EXT Group I, one (*EXT23*) was *down*-regulated >threefold, while the other (*EXT1*) was *up*-regulated >threefold; and (ii) 3/5 members of Group IIa classical EXTs, excluding *EXT3*, (the mutated *EXT3* is in Group IIa) were *up*-regulated, while all 12 members of Groups IIb and IIc classical EXTs were *down*-regulated in the mutant. The presence of the SPSP amino acid motif (function unknown) occurs with regular periodicity in all members of Groups IIb and IIc distinguishing them from the other classical EXTs. Previous classification of the 20 classical EXTs (Cannon *et al.*, 2008) was based solely on the occurrence of amino acid motifs (Table 1, columns 1–3). The data here substantiate this, and show that EXTs *with* (contiguous Pro) and *without* SPSP motifs (non-contiguous Pro) have distinct and coordinated regulation of gene expression, thereby suggesting that they have distinct function(s) from each other as well. The chimeric, hybrid, and short EXT groups, as well as classical and chimeric PRP groups, each had members that were either down- or up-regulated in the mutant compared with WT. This 'down or up' pattern of regulation also occurred with *AGP*, *PER*, and *EXP(L)* genes.

Gene expression in the *ANP-ext3* revertant

Microarray (Figure S3 and Table S6, '*Change in ANP4-F2 versus mut*') and qRT-PCR (Figure 2 and Table S7, '*Change in ANP4-F2 versus mut*') data show that most but not all genes analyzed in the *ANP* compared with the mutant had a reversal of the gene expression that had occurred in going from the WT to the mutant. The following are some notable observations (Table 1, '*ANP4-F2 versus mut*', and '*ANP4-F2 versus WT*'): (i) the two Group I EXTs exceeded WT levels in the *ANP*; 2/3 EXT Group IIa members, which had increased gene expression in the mutant, showed a decreased expression in the *ANP*; and (ii) all 12 members of EXT Groups IIb and IIc had increased expression in the *ANP*. Many but not all the other *EXTs* and *PRPs*, as well as the *AGPs*, *EXP/EXPLs*, *PERs* and *P4Hs* genes also had reversed gene expression in the direction of WT levels.

These gene expression profiles indicate alternative coordinated expression networks in the *ANP* compared with the mutant and in both compared with the WT. This finding is further supported by the results of GUS (*uidA*) reporter gene assays, facilitated by the fact that the engineered Ds insert in the *EXT3* gene is an 'enhancer-trap' insertion (Sundaresan *et al.*,1995), in which the *uidA* gene monitors expression of the *EXT3* promoter (Hall and Cannon, 2002): unlike the *EXT3* gene expression pattern found in *rsh* heterozygotes (see Figure 6 in Hall and Cannon, 2002), the *EXT3* promoter in the *ANP* lines clearly showed a different pattern of expression. In particular note that, unlike heterozygous *ext3*, the *ANP* plants showed *uidA* gene expression in the vascular system of hypocotyl and cotyledons (Figure 3b), mature rosette leaves (Figure 3d), petals and sepals (Figure 3e), and in pollen (Figure 3e,f).

Regulatory genes associated with the alternative expression networks

For the purpose of identifying genes associated with regulatory function in the alternative gene expression programs, we examined the microarray data and verified

© 2013 The Authors

expression of candidate genes by qRT-PCR. The greatest gene expression differences between the mutant, *ANP* and WT were seen in genes for 14 DNA binding proteins (DBP) and three kinases (Figure 2g and Table S7). UPBEAT1 (*UPB1*) is of particular note as it regulates transcription of PER genes and is associated with cell differentiation (Tsukagoshi *et al.*, 2010). These data further support the presence of alternative gene expression networks resulting in a gene expression landscape that allows apparently normal plant growth in the (near) absence of EXT3; some or all of the 17 (putative) regulatory genes are likely to be involved based on their up-regulation in the *ANP* lines.

Is the '*ANP-ext3* gene expression program' stably inherited and is it reproducible?

These questions were addressed by comparing gene expression: (i) in the next generation of *ANP4*-F2 i.e. *ANP4*-F3: and (ii) in an independently arising *ANP*; i.e. *ANP10*-F3, with the next generation WT (called WT-2) (Figure 4 and Table S8).

(i) The relative expression *level* of each gene in an *ANP4* line was similar, but not identical, for most genes for the two generations tested (compare heatmaps in Figure 2, '*ANP4-F2*' and in Figure 4, '*ANP4-F3*'). The relative expression pattern showed that some genes were expressed at levels closer to that of WT-2 in the F3 generation (compare bar graphs for individual genes in Figures 2 and 4). This transgenerational drift towards wild-type levels of expression was particularly obvious in the case of the classical *EXTs* (Table 1, column 7), where all (except *EXT3*) showed no significant difference in gene expression to that of WT-2. Therefore, a classical EXT is not solely responsible for self-rescue of the *rsh* mutant.

(ii) The independently arising *ANP10*-F3 revertant compared with WT-2, showed a mostly similar, but not identical, relative expression levels (see heatmaps in Figure 4) as well as patterns of gene expression (see bar graphs in Figure 4) to that of *ANP4*-F3. With focus on the 20 classical *EXT* genes, all but one was at or very close to wild-type levels (Table 1, last column). Based on the drift towards wild-type levels of expression in going from *ANP4*-F2 to F3, it is reasonable to expect that the next generation (F4) of *ANP10* would also show no significant difference in classical *EXT* expression compared with wild type (except for *EXT3*). These data support the '*ANP-ext3* gene expression program' being reproducible, but with a transgenerational drift toward wild-type levels of gene expression for some genes, especially the classical *EXTs* (Table 5).

In comparing the differential gene expression with wild type for *ANP10* and the two generations of *ANP4*: we identified three non-classical *EXTs*, four *PRPs* and one *AGP* as well as seven transcription factors and two kinase genes, which were up-regulated ($P \leq 0.05$), in all comparisons made with wild type, by qRT-PCR. This set of genes had a $\geq$twofold increase in expression in the F3 generations compared with wild type (Table 6). This finding is important, because when other genes of focus showed little or no significant difference in expression compared with wild type, genes from this set were up-regulated in all tests in F3 generations, thus associating their products with successful wall assembly in the (near) absence of EXT3. The presence of an alternative gene expression program resulting in the assembly of apparently normal walls in *ANP* lines was further supported by the fact that GUS assays of *ANP4*-F3 and *ANP*10-F3 showed the same pattern of expression as those of *ANP4*-F2 (Figure 3).

Amino acid composition of cell walls of WT, *rsh* mutant and *ANP* lines

Given this differential expression of HRGP genes in the *rsh* mutant compared with the *ANP* and WT, in which some members of each subfamily of focus showed decreased and others increased expression, and the fact that total mean HRGP gene transcript values, as well as that of *EXT*, *PRP* and *AGP* subfamilies showed no significant difference ($P \leq 0.05$) between the three samples being compared (WT, mutant and *ANP*), we hypothesized that quantities of Pro and Hyp in cell walls of these samples would not vary significantly. Amino acid composition analysis (Table S9) showed that this was indeed the case, thus supporting the 'down–up' regulation patterns of different HRGP genes in alternative HRGP gene expression programs.

DISCUSSION

Alternative gene expression networks

An intriguing question arising from the data is 'How does a (near) knock-out of *EXT3* result in the altered gene expression seen in the *rsh* mutant?' *EXT3* is highly expressed in several plant parts including the developing embryo (Hall and Cannon, 2002). The *rsh* mutant defect was traced to a defective wall in the first division of the zygote, which produces embryos with misshapen cells with misplaced and/or incomplete walls (Cannon *et al.*, 2008). As cell shape and turgor pressure are critical for cell growth and development, a different intracellular environment in the apical and suspensor cells of that first division of the *rsh* zygote and/or early embryo cells, probably leads to the gene expression changes. Thus, the (near) absence of *EXT3* is the initial cause of gene expression changes, and the observed *rsh* mutant phenotype is the consequence of some or all of these changes. The gene expression changes that took place affect a disproportionate number of HRGP genes and genes of relevance to their function, supporting the claim that the knock-down of *EXT3* expression is responsible for the *rsh* mutant phenotype, and that there is a particular

© 2013 The Authors

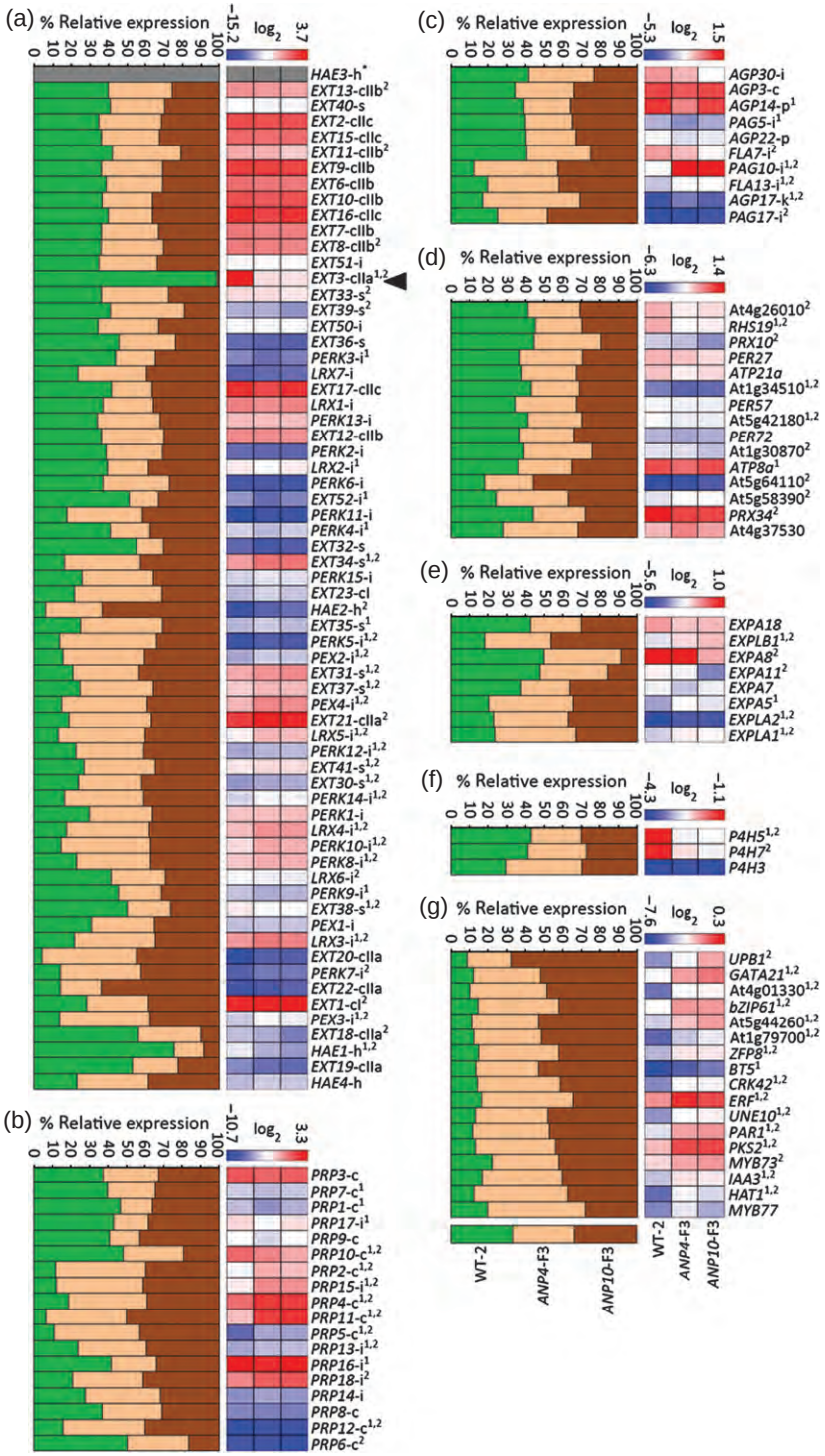

Figure 4. Comparative gene expression in seedlings of WT-2 (progeny of WT), *ANP4*-F3 and *ANP10*-F3 using qRT-PCR analysis.
(a) All 65 *EXT*.
(b) All 18 *PRP*.
(c) 10 *AGP*.
(d) 15 *PER*.
(e) 8 *EXP/EXPL*.
(f) 3 *P4H*.
(g) Seventeen genes for regulatory proteins (14 DNA binding proteins (DBPs) and three kinases). Gene order within groups is as Figure 2. Further details are described in Figure 2 legend, except here superscript numbers indicate significance ($P \leq 0.05$) of *ANP4*-F3 and *ANP10*-F3 versus WT-2, respectively (see also Table S8).

gene expression network associated with this cell wall mutant, which we call a '*rsh-ext3* gene expression program'. Further support for this result could be obtained by analyzing additional individual lines with *ext3* alleles. However, none was available.

A second intriguing question arising from the data is: 'What exactly is rescuing the *rsh* mutant?' The data show that a change in gene expression of a *classical EXT* alone is not responsible. As in the *rsh* mutant, *EXT3* expression is knocked-down in the *ANP* lines but, unlike the *rsh* mutant, the *ANP* lines show that other classical *EXT*s are expressed at or near WT levels, and that eight other HRGP genes (as well as nine putative regulatory genes) are up-regulated in all *ANP* lines. Little or no information is known specifically

© 2013 The Authors

Table 5 Number of genes tested by qRT-PCR and those showing down-regulation or up-regulation (x|y) ($P \leq 0.05$) in the comparisons listed (summary of Figures 2 and 4)

Number of genes and comparisons made	*EXT*	*PRP*	*AGP*	*PER*	*EXP, EXPL*	*P4H*	DBP genes	Kinase genes
Line 1 Number of genes tested	63[a]	18	10	15	8	3	14	3
Line 2 Mutant versus WT (Figure 2)	28|13	10|4	7|3	12|2	6|1	0|2	10|0	3|0
Line 3 ANP4-F2 versus mutant (Figure 2)	10|33	4|9	0|7	2|12	1|5	3|0	0|14	0|3
Line 4 ANP4-F2 versus WT (Figure 2)	19|14	5|4	5|3	13|0	2|3	1|0	0|11	0|2
Line 5 ANP4-F3 versus WT-2 (Figure 4)	7|17	5|7	2|3	4|0	0|4	1|0	0|11	0|3
Line 6 ANP10-F3 versus WT-2 (Figure 4)	9|20	2|8	1|4	7|2	2|3	2|0	0|12	0|3

Line 2 shows the numbers of genes down-regulated or up-regulated in the mutant compared with WT.
Line 3 compared with *Line 2* shows a reversal of this expression pattern.
Line 4 shows the number of genes down- or up-regulated in the *ANP* revertant compared with WT.
Line 5 compared with *Line 4* shows a drift toward WT levels of gene expression in the next generation of *ANP4.*
Line 6 compared with *Line 5* shows the reproducibility of the *rsh* revertant phenomenon, as well as the drift toward wild-type levels of gene expression.
[a]Excluded *EXT3 (rsh* mutant) and *HAE3*-h (consistantly not expressed) from the 65 total *EXT*s.

Table 6 Fold change in gene expression of *ANP* lines versus WT where all comparisons were different to WT ($P \leq 0.05$), and F3 generations were $\geq$twofold differentially expressed (by qRT-PCR)

Gene family	Name[a]	Gene ID	*ANP4*-F2 versus WT	*ANP4*-F3 versus WT-2	*ANP10*-F3 versus WT-2
EXT	*EXT34*-s	At3g06750	2.1	2.5	2.5
	LRX4-i	At3g24480	1.5	2.5	2.1
	PEX2-i	At1g49490	1.7	2.8	2.5
PRP	*PRP2*-c	At2g21140	1.8	4.2	3.4
	PRP4-c	At4g38770	1.6	2.3	2.1
	PRP11-c	At5g15780	1.8	6.7	7.8
	PRP15-i	At2g10940	2.4	3.9	3.4
AGP	*FLA13*-i	At5g44130	2.1	2.0	2.2
DBP	*GATA21*	At5g56860	1.8	3.0	4.4
	bZIP61	At3g58120	2.4	3.0	2.9
	At5 g44260	At5g44260	2.5	3.3	4.8
	ZFP8	At2g41940	2.7	3.0	2.9
	UNE10	At4g00050	3.0	3.0	3.7
	IAA3	At1g04240	3.8	2.6	2.4
	HAT1	At4g17460	6.3	4.1	3.1
Kinase	*CRK42*	At5g40380	2.6	3.2	3.0
	PKS2	At1g14280	3.4	3.3	3.4

[a]See Figure 2 legend, and Table S5.

about the functions of these gene products, thus the data here give an 'entree' to identifying their roles. Each one may have a necessary and unique role to play and/or some or all could be making fractional contributions to successful wall assembly in the (near) absence of *EXT3* expression. Future phenome relative to transcriptome analyses of knock-outs of each of these putative regulators in WT and in *homozygous ext3* backgrounds will decipher their roles. As the data in this manuscript show, such research approaches a big data project for accurate interpretation.

Analysis of the amino acid sequences of the putative *ext3* compensators shows that EXT34, PRP2, PRP4 and the EXT domains of chimeric LRX4 and PEX2 have key features in common with Group IIa classical EXTs, i.e. proline-rich, abundant Y, K and/or H residues and low, or no, SPSP motifs. Unlike Group IIa (to which EXT3 belongs) these five HRGPs, although they are Tyr-rich, have no Idt (YXY) sequence and the periodicity of all their EXT motifs is considerably less precise. Theoretically, they have capacity to align, albeit imprecisely, and cross-link to a 'classical' EXT network via their lone Tyr residues, thereby forming intermolecular/interpolymeric Idt (2 Y) and/or pulcherosine (3 Y), and to provide positively charged residues to template pectin deposition. Thus, they could compensate for the lack of EXT3. Their increased expression in all *ANP* comparisons with WT suggest that this outcome is the case.

PRP11 and the PRP domain of PRP15 are Pro-rich and have an abundance of positively charged amino acids; they resemble Group I classical EXTs in that they have very few

© 2013 The Authors

Tyr residues and no SPSP motifs. It is less clear how Tyr-poor Group I classical extensins (of which there are only two in Arabidopsis) fit into the EXT polymeric network model. One suggestion is that they form intermolecular/interpolymeric bridges by covalent bonding of their sparse Tyr residues with those of the EXT network proper (Kieliszewski *et al.*, 2011), thereby providing continuity and more depth to EXT polymeric network throughout the cell wall matrix. Support for this proposal comes from the fact that Group I, as well as Group II, EXTs are insoluble in the wall (Smith *et al.*, 1986) and that ionic desorption rapidly elutes them (Smith *et al.*, 1984).

The five 'EXT Group IIa-like', and two 'EXT Group I-like' HRGP are a mechanistic fit with this model of the EXT polymeric network. FLA13 has no obvious features of structural relevance to network-forming EXTs, as is the case with all AGPs. The main aim of this work was to identify *ext3* compensators; evidence strongly supports seven of these eight HRGP as candidates as they fit the most up-to-date model of EXT function. It would now be informative to examine the phenotypes of homozygous polymorphic lines of these eight HRGP in the WT as well as in the ANP (*ext3/ext3*) background. Based on the data in this report, each line would have a different gene expression landscape, thus reliable interpretation of the data arising would require extensive gene expression analysis of many plant lines over three generations (due to the drift phenomenon mentioned above).

A third intriguing question is: 'What causes the "*rsh-ext3* gene expression program" to switch to an "*ANP-ext3* gene expression program"?' The data allude to the possibility of an epigenetic mechanism regulating gene expression in response to the changed internal cell environment. The data are suggestive of a regulatory switch occurring in a proportion of homozygous *rsh* zygotes, or early stage embryo cells, setting rescue in motion. This event is likely, as the shape, size and integrity of the *rsh* walls are variable as seen in embryo sections beginning with the first division of the zygote (Hall and Cannon, 2002), and by electron microscopy of defective *rsh* embryo cells (Cannon *et al.*, 2008). The data here support involvement of the seven co-regulated DBP and two kinases, whose expression were up-regulated in all *ANP* lines tested. This change is heritable, with no mutant phenotypes arising in the subsequent generations tested. There are several examples of altered and heritable gene expression patterns, including reports of transgenerational memory in relation to disease resistance in plants (Luna *et al.*, 2012; Pieterse, 2012; Rasmann *et al.*, 2012; Slaughter *et al.*, 2012).

The alternative explanation to be considered for the *ANP* lines arising is the presence of a preexisting or a newly acquired suppressor mutation. The phenotype and genotype data presented here argue against such genetic revertants and in favor of the epigenetic model described. This epigenic model is novel because it proposes that stochastic variation in gene expression, in defective embryo cells, is sufficient to cause approximately 20% of homozygous *ext3* progeny to appear completely wild type while the other approximately 80% appear completely mutant, in every generation. The general expectation for an epigenetic explanation would be that expression reprogramming results in a continuum between the two extreme phenotypes. In the situation here, this situation would require embryo cells up to the point of rescue to continue developing. As *rsh* mutant embryo cells are variable in shape, size and wall integrity and, consequently, very likely to have different gene expression landscapes in different cells, thus not all of the defective cells contribute to the rescued *ANP*. As *ANP* embryos appear normal, this factor suggests that rescue occurs at an early stage of embryo development and that the rescued cell(s) form a normal looking embryo. We know that spontaneous epigenetic variation occurs from examining the Arabidopsis methylome (Becker *et al.*, 2011) and also there is transgenerational epigenetic instability (Schmitz *et al.*, 2011). We also know that methylation is reduced during gamete development, and re-established at embryo development in Arabidopsis (Jullien *et al.*, 2012). Therefore, theoretically, *ANP* revertants could occur by epigenetic reprogramming of a proportion of *rsh* embryo cells. It has been suggested that reprogramming during embryo development could provide plants with a means to respond to both internal and external stresses (Calarco and Martienssen, 2013).

Broader implications

The data show that a concert of players must be considered when analyzing any single gene and its product if we are to understand its role. At anytime and, for many reasons, including an altered internal cellular environment, an alternative gene expression program could be switched on, alleviating a requirement for a gene of interest. Consequently, it would be inconclusive to assign function by examining phenotype along with the expression of the gene of interest alone, even when a transgene of the gene-of-interest is expressed.

Self-rescue is a plausible explanation why cell wall mutant phenotypes in general have been hard to find. Plant cell shape is fundamental to growth and development, and it relies on cell wall structure. Cell walls, even in one tissue, are variable and dynamic structures; for this outcome, the plant has large gene families with regulation to provide varying quantities of wall building materials, where and when needed. In keeping with this situation, the data here show that there is more than one combination of materials and/or 'facilitators' suited to building functional walls. Apart from being informative regarding cell wall biology, this situation opens up an interesting line of research for utilitarian purposes.

© 2013 The Authors

EXPERIMENTAL PROCEDURES

Plant material

The wild-type genotype used was *Arabidopsis thaliana* Landsberg *erecta* (obtained from Lehle Seeds, Round Rock, TX, USA; http://www.arabidopsis.com/main/cat/!ct_seat.html), and called WT in this paper; 'WT-2' is next generation WT; the mutant used was '*root-, shoot-, hypocotyl-defective*' (*rsh*), a recessive embryo defective and seedling lethal mutant caused by an enhancer-trap insert in the *EXT3* (At1g21310) gene (Hall and Cannon, 2002); *ANP4* and *ANP10* lines were phenotypic but not genotypic revertants of the *rsh* mutant first described in this manuscript. Seed were surface-sterilized when germinated in Petri dishes on half strength Murashige and Skoog (MS) salts and vitamins (Murashige and Skoog, 1962) with 3% (w/v) sucrose and solidified with 0.8% (w/v) agar (half-MS). Kanamycin (40 mg l^{-1}) was added to the medium where mentioned after sterilization. Following stratification for 2 days in the dark at 4°C, plates were incubated under controlled environmental conditions (16 h of 140 μmol m^{-2} sec^{-1} light and 8 h of dark at 22°C) for 12 days (or more where mentioned) before harvesting. Days after sowing (DAS) excludes stratification time.

Segregation of heterozygous *rsh* (*EXT3/ext3*) progeny

Starting with 30 *EXT3/ext3* plants, where the zygosity of the mutating insert was confirmed by PCR analysis (as in Figure 1), seed were collected from each individual plant and germinated on half-MS kanamycin plates in three replicates with 10 plants per replicate. At 12 DAS, white (kanamycin sensitive), green (kanamycin resistant), and mutant seedlings, as well as non-germinated seed were scored. The mean values, standard deviation (SD) and segregation ratio were calculated using Microsoft Excel™.

Phenotyping

Plants were grown either in individual pots of soil or aseptically on half-MS in Petri dishes, positioned randomly, and grown as described in Plant material. For each phenotype analysis, five separate sets of experiments with four replicates (total, $n = 20$), or 10 replicates (total, $n = 50$), were conducted at different times and the data were combined. The mean values, SD, and *t*-test (*P*-value) were calculated using Microsoft Excel™:

(i) Root and shoot length measurement: seeds were arranged on half-MS along a line drawn on the base of the petri plate (Figure S2a). Following vertical incubation for 12 days, the distances from the line to the bottom of the root and top of the shoot were measured.
(ii) Dry weight: plants of 21 DAS were cut at soil level, placed between blotting paper sheets, and dried in an oven at 80°C for 7 days before weighing.
(iii) Plant height: this phenotype was measured from soil to top of plant at 45 DAS.
(iv) Individual silique length and number per plant: these were measured and counted at 45 DAS, respectively.
(v) Seed yield: harvested seed from fully senescent plants were separated from dust and weighted.
(vi) Germination rate: seed on half-MS plates were incubated horizontally for 7 days. Rate was calculated as percentage of seed germinated out of total number plated.

Genotyping

To test seedlings for the presence/or not of the enhancer-trap insert, genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Germantown, MD, USA; http://www.qiagen.com/), and test DNA was amplified by PCR in a thermal cycler (MJ Research PTC-200) (strategy in Figure 1). The 15 μl reaction contained 100 ng DNA, 10 μM dNTPs, 1 mM each of forward and reverse primers and one unit *Taq* polymerase. The amplification program was as follows: 5 min at 95°C; followed by 30 cycles of 95°C for 1 min, 58°C for 30 sec, 72°C for 30 sec; and a final extension of 72°C for 10 min. A 5 μl sample was electrophoresed on 1% agarose to separate and visualize amplicons.

RNA isolation

For gene expression analyses, one biological replicate consisted of approximately 100 seedlings grown on a horizontally incubated plate. Replicates and samples were grown under identical growth conditions and at the same time. Twelve-day-old seedlings were collected and immediately frozen in liquid nitrogen before storage at −80°C. Total RNA was isolated from 100 mg frozen tissue using the RNeasy Plant Mini Kit (Qiagen). RNA integrity was checked using the Agilent 2100 Bioanalyzer with a 6000 nanochip (Agilent Technologies, Palo Alto, CA, USA; http://www.home.agilent.com/agilent/home.jspx?cc=US&lc=eng) and RNA was quantified using a NanoDrop Spectrophotometer ND-8000 (NanoDrop Technologies, Wilmington, DE, USA; http://www.nanodrop.com/).

Microarray analysis

Affymetrix GeneChip® Arabidopsis ATH1 Genome Array (Affymetrix, Santa Clara, CA, USA; http://www.affymetrix.com/estore/) was used for expression analysis. Probe labelling, chip hybridization and scanning were performed according to the manufacturer's instructions for the IVT Express Labeling Kit (Affymetrix) using 500 ng of total RNA for each sample. Data normalization among chips was conducted using Robust Multichip Average (RMA) (Irizarry *et al.*, 2003). Presence and absence calls for each probe set were obtained using dCHIP (Li and Wong, 2001). Gene selections for pairwise comparison were made based on Associative Analysis (Dozmorov and Centola, 2003) in Matlab (MathWorks, Natick, MA, USA; http://www.mathworks.com/products/matlab/). In this method, the background noise presented between replicates and technical noise during microarray experiments was measured by the residual presented among a group of genes whose residuals are homoscedastic. Genes with residuals between the compared sample pairs that are significantly higher than the measured background noise level were considered as differentially expressed. A selection threshold of two for transcript ratios and a Bonferroni-corrected *P*-value threshold of 2.19202E–06 were used. The Bonferroni-corrected *P*-value threshold was derived from 0.05/N in these analyses, where N is the number of probe sets (22810) on the chip in order to correct the family wide false discovery rate (Abdi, 2007). Similarity of gene expression pattern for microarray was calculated based on Euclidean distance using MultiExperiment Viewer4 (MeV4) (Saeed *et al.*, 2006).

Real-time quantitative reverse transcription-PCR (qRT-PCR) and data analysis

First-strand cDNA was synthesized from 1 μg of total RNA using oligo-dT20 primer (Life Technologies, Grand Island, NY, USA; http://www.lifetechnologies.com/us/en/home.html) and SuperScript III reverse transcriptase (Life Technologies) after an extensive Turbo-DNase (Ambion, Austin, TX, USA; http://www.invitrogen.com/site/us/en/home/brands/ambion.html) treatment as described in manufacturer's protocol. Absence of DNA contamination was confirmed by qRT-PCR for each sample using two

© 2013 The Authors

pairs of forward (#At3g18780-f: ACTTTCATCAGCCGTTTTGA and #At5g65080-f: TTTTTTGCCCCCTTCGAATC) and reverse (#At3g-18780-r: ACGATTGGTTGAATATCATCAG and #At5g65080-r: AT-CTTCCGCCACCACATTGTAC) primers designed from the intron sequence of two control genes At3g18780 and At5g65080, respectively, prior to reverse transcription.

Primer Express® software (version 3.0) (Applied Biosystems, Foster City, CA, USA; http://www.invitrogen.com/site/us/en/home/brands/Applied-Biosystems.html) was used to design primer sets using default parameters according to the manufacturer's recommendation. For most of the genes primers were designed to anneal near the 3′ end or at the 3′ UTR (Table S10). The qRT-PCR reactions were performed as described (Czechowski *et al.*, 2004) using the ABI PRISM 7900HT sequence detection system (Life Technologies) in an optical 384 well plate. The qRT-PCR reaction was done in 10 μl final volume with 1 μM each of gene specific primers, 2 μl cDNA (5 ng μl^{-1}) and 5 μl of 2× Power SYBR Green PCR Master Mix (Life Technologies). No-template control reactions were performed for each pair of primers. Amplicon dissociation curves i.e. melting curves were recorded to confirm the gene specific amplification by a single dominant peak. Data were collected by SDS2.4 software (Life Technologies). Gene expression was normalized against the expression of eukaryotic initiation factor4A-2 (EIF4A-2: At1g54270) serving as reference gene. PCR efficiency (E) was calculated by LinRegPCR software (Ramakers *et al.*, 2003). Relative expression of each gene of interest was calculated using the ΔC_T values as previously described (Czechowski *et al.*, 2005). Values are means of three biological replicates with three technical replicates for each gene. All data were analyzed using Microsoft Excel™. For clustering of differentially expressed genes in both the microarray and qRT-PCR data, average expression values of replicates were transformed into Log_2 and hierarchical clustering was performed using Pearson correlation method using MeV4.

GUS localization

Seedlings and various plant parts were tested for GUS expression as described previously (Jefferson *et al.*, 1987; Hall and Cannon, 2002). Additionally, the tissues were cleared by repeated washing with 3:1 ethanol/acetic acid (v/v) to ensure the new GUS localization sites found in the *ANP* plants were authentic.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (NSF) IOS-0622428 and IOS-0955805 (to M.C.C.), and NSF IOS-0955569 (to M.J.K.). The Samuel Roberts Noble Foundation and the BioEnergy Science Center (BESC), Office of Biological and Environmental Research in the Office of Science, US Department of Energy supported T.R. and Y.T.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Comparison of transcriptomes of three bioreplicates each of WT, *rsh* mutant and *ANP4*-F2.

Figure S2. Phenotypes of seedlings used in this study and summary data from differential gene expression tested by microarray analysis.

Figure S3. Comparative expression of genes from selected families with ≥twofold differential expression ($P \leq 0.05$) in at least one of the three comparisons made between three samples i.e. WT, *rsh* mutant and *ANP4*-F2, using microarray analysis of seedling.

Table S1. Top 50 (putative) genes down-regulated in the *rsh* mutant versus WT, and their expression in the *ANP4*-F2 versus the *rsh* mutant, by microarray analysis of seedlings.

Table S2. Top 50 (putative) genes up-regulated in the *rsh* mutant versus WT, and their expression in the *ANP4*-F2 versus the *rsh* mutant, by microarray analysis of seedlings.

Table S3. Top 50 (putative) genes down-regulated in the *ANP4*-F2 versus WT, by microarray analysis of seedlings.

Table S4. Top 50 (putative) genes up-regulated in *ANP4*-F2 versus WT, by microarray analysis of seedlings.

Table S5. Abbreviations, full names/descriptions of (putative) gene products.

Table S6. *EXT, PRP, AGP, PER, EXP/EXPL* and *P4H* (putative) genes showing ≥twofold differential expression in at least one of the following three comparisons (and reverse comparisons) of WT, the *rsh* mutant, and *ANP4*-F2 seedlings, by microarray analysis (also see Figure S3).

Table S7. Expression comparisons of all 65 *EXT* and 18 *PRP,* and selected 10 *AGP,* 15 *PER,* 8 *EXP/EXPL*, 3 *P4H,* and 14 DBP and three kinase encoding (putative) genes in WT, *rsh* mutant and *ANP4*-F2, by qRT-PCR analysis of seedlings (also see Figure 2).

Table S8. Expression comparisons of *ANP4*-F3 and *ANP10*-F3 with WT-2 (next generation WT), for all 65 *EXT* and 18 *PRP,* and for selected 10 *AGP,* 15 *PER,* 8 *EXP/EXPL*, 3 *P4H,* and 14 DBP and three kinase encoding (putative) genes, by qRT-PCR analysis of seedlings (also see Figure 4).

Table S9. Amino acid composition as average mole % of total amino acids in cell walls of seedling roots from WT, *rsh* mutant, *ANP4*-F2 and *ANP10*-F2.

Table S10. Primers used for qRT-PCR, and PCR (last two lines).

Data S1. Experimental procedures for supporting information.

Data S2. References for supporting information.

REFERENCES

Abdi, H. (2007) Bonferroni and Sidak corrections for multiple comparisons. In *Encyclopedia of Measurements and Statistics* (Salkind, N.J., ed.). Thousand Oaks, CA: Sage, pp. 103–107.

Baumberger, N., Steiner, M., Ryser, U., Keller, B. and Ringli, C. (2003) Synergistic interactions of the two paralogous Arabidopsis genes LRX1 and LRX2 in cell wall formation during root hair development. *Plant J.* **35**, 71–78.

Becker, C., Hagmann, J., Mueller, J., Koenig, D., Stegle, O., Borgwardt, K. and Weigel, D. (2011) Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*, **480**, 245–249.

Bernhardt, C. and Tierney, M.L. (2000) Expression of AtPRP3, a proline-rich structural cell wall protein from Arabidopsis, is regulated by cell-type-specific developmental pathways involved in root hair formation. *Plant Physiol.* **122**, 705–714.

Brady, J.D., Sadler, I.H. and Fry, S.C. (1996) Di-isodityrosine, a novel tetrameric derivative of tyrosine in plant cell wall proteins: a new potential cross-link. *Biochem. J.* **315**, 323–327.

Brady, J.D., Sadler, I.H. and Fry, S.C. (1998) Pulcherosine, an oxidatively coupled trimer of tyrosine in plant cell walls: its role in cross-link formation. *Phytochemistry*, **47**, 349–353.

Calarco, J.P. and Martienssen, R.A. (2013) Imprinting: DNA methyltransferases illuminate reprogramming. *Curr. Biol.* **22**, R929–R931.

Cannon, M.C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B.L., Chen, L., Lamport, D.T.A., Chen, Y. and Kieliszewski, M.J. (2008) Self-assembly of the plant cell wall requires an extensin scaffold. *Proc. Natl Acad. Sci. USA*, **105**, 2226–2231.

Czechowski, T., Bari, R.P., Stitt, M., Scheible, W.-R. and Udvardi, M.K. (2004) Realtime RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J.* **38**, 366–379.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. and Scheible, W.-R. (2005) Genome-wide identification and testing of superior reference

© 2013 The Authors

genes for transcript normalization in Arabidopsis. *Plant Physiol.* **139**, 5–17.

Dozmorov, I. and Centola, M. (2003) An association analysis of gene expression array data. *Bioinformatics*, **19**, 204–211.

Eisenhaber, B., Wildpaner, M., Schultz, C.J., Borner, G.H.H., Dupree, P. and Eisenhaber, F. (2003) Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for Arabidopsis and rice. *Plant Physiol.* **133**, 1691–1701.

Ellis, M., Egelund, J., Schultz, C.J. and Bacic, A. (2010) Arabinogalactan-proteins: key regulators at the cell surface. *Plant Physiol.* **153**, 403–419.

Fowler, T.J., Bernhardt, C. and Tierney, M.L. (1999) Characterization and expression of four proline-rich cell wall protein genes in Arabidopsis encoding two distinct subsets of multiple domain proteins. *Plant Physiol.* **121**, 1081–1091.

Hall, Q. and Cannon, M.C. (2002) The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. *Plant Cell*, **14**, 1161–1172.

Held, M.A., Kamyab, A., Hare, M., Shpak, E. and Kieliszewski, M.J. (2004) Di-isodityrosine is the intermolecular cross-link of isodityrosine-rich extensin analogs cross-linked *in vitro*. *J. Biol. Chem.* **279**, 55474–55482.

Irizarry, R.A., Hobbs, B. and Speed, T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249–264.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: ß-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

Johnson, K.L., Jones, B.J., Schultz, C.J. and Bacic, A. (2003) Non-enzymic cell wall (glyco)proteins. In *The Plant Cell Wall* (Rose, J.K.C., ed.). Boca Raton: CRC Press LLC, pp. 111–154.

Johnson, K.L., Kibble, N.A.J., Bacic, A. and Schultz, C.J. (2011) A fasciclin-like arabinogalactan-protein (FLA) mutant of *Arabidopsis thaliana*, *fla1*, shows defects in shoot regeneration. *PLoS ONE*, **6**, e25154.

Jullien, P.E., Susaki, D., Yelagandula, R., Higashiyama, T. and Berger, F. (2012) DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* **22**, 1825–1830.

Kieliszewski, M.J., Lamport, D.T.A., Tan, L. and Cannon, M.C. (2011) Hydroxyproline-rich glycoproteins: form and function. In *Plant Polysaccharides: Biosynthesis and Bioengineering* (Ulvskov, P., ed.). Oxford: Wiley-Blackwell, pp. 321–342.

Lamport, D.T.A., Kieliszewski, M.J., Chen, Y. and Cannon, M.C. (2011) Role of the extensin superfamily in primary cell wall architecture. *Plant Physiol.* **156**, 11–19.

Li, C. and Wong, W. (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA*, **98**, 31–36.

Luna, E., Bruce, T.J.A., Robert, M.R., Flors, V. and Ton, J. (2012) Next-generation systemic acquired resistance. *Plant Physiol.* **158**, 844–853.

MacMillan, C.P., Mansfield, S.G., Stachurski, Z.H., Evans, R. and Southerton, S.G. (2010) Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus. *Plant J.* **62**, 689–703.

McQueen-Mason, S. and Cosgrove, D.J. (1994) Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proc. Natl Acad. Sci. USA*, **91**, 6574–6578.

McQueen-Mason, S., Durachko, D.M. and Cosgrove, D.J. (1992) Two endogenous proteins that induce cell wall extension in plants. *Plant Cell*, **4**, 1425–1433.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497.

Pieterse, C.M.J. (2012) Prime time for transgenerational defense. *Plant Physiol.* **158**, 545.

Ramakers, C., Ruijter, J.M., Deprez, R.H. and Moorman, A.F. (2003) Assumption free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **13**, 62–66.

Rasmann, S., De Vos, M., Casteel, C.L., Tian, D., Halitchke, R., Sun, J.Y., Agrawal, A.A., Felton, G.W. and Jander, G. (2012) Herbivory in the previous generation primes plants for enhanced insect resistance. *Plant Physiol.* **158**, 854–863.

Saeed, A.I., Bhagabati, N.K., Braisted, J.C., Liang, W., Sharov, V., Howe, E.A., Li, J., Thiagarajan, M., White, J.A. and Quackenbush, J. (2006) TM4 microarray software suite. *Methods Enzymol.* **411**, 134–193.

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A., Libiger, O., Schork, N.J. and Ecker, J.R. (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science*, **334**, 369–373.

Schnabelrauch, L.S., Kieliszewski, M.J., Upham, B.L., Alizedeh, H. and Lamport, D.T.A. (1996) Isolation of pI 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link sites. *Plant J.* **9**, 477–489.

Seifert, G.J. and Roberts, K. (2007) The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* **58**, 137–161.

Shi, H., Kim, Y., Guo, Y., Stevenson, B. and Zhu, J.-K. (2003) The Arabidopsis SOS5 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell*, **15**, 19–32.

Showalter, A.M., Keppler, B., Lichtenberg, J., Gu, D. and Welch, L.R. (2010) A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* **153**, 485–513.

Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B. and Mauch-Mani, B. (2012) Descendants of primed Arabidopsis plants exhibit resistance to biotic stress. *Plant Physiol.* **158**, 835–843.

Smith, J.J., Muldoon, E.P. and Lamport, D.T.A. (1984) Isolation of extensin precursors by direct elution of intact tomato cell suspension cultures. *Phytochemistry*, **23**, 1233–1239.

Smith, J.J., Muldoon, E.P., Willard, J.J. and Lamport, D.T.A. (1986) Tomato extension precursors P1 and P2 are highly periodic structures. *Phytochemistry*, **25**, 1021–1030.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D.G., Dean, C., Ma, H. and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **9**, 1797–1810.

Tiainen, P., Myllyharju, J. and Koivunen, P. (2005) Characterization of a second *Arabidopsis thaliana* prolyl 4-hydroxylase with distinct substrate specificity. *J. Biol. Chem.* **280**, 1142–1148.

Tsukagoshi, H., Busch, W. and Benfey, P.N. (2010) Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. *Cell*, **143**, 606–616.

Valentin, R., Cerclier, C., Geneix, N., Aguié-Béghin, V.R., Gaillard, C.D., Ralet, M.-C. and Cathala, B. (2010) Elaboration of extensin–pectin thin film model of primary plant cell wall. *Langmuir*, **26**, 9891–9898.

Velasquez, S.M., Ricardi, M.M., Dorosz, J.G. *et al.* (2011) *O*-glycosylated cell wall proteins are essential in root hair growth. *Science*, **332**, 1401–1403.

Youl, J.J., Bacic, A. and Oxley, D. (1998) Arabinogalactan-proteins from *Nicotiana alata* and *Pyrus communis* contain glycosylphosphatidylinositol membrane anchors. *Proc. Natl Acad. Sci. USA*, **95**, 7921–7926.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol.* **136**, 2621–2632.

© 2013 The Authors