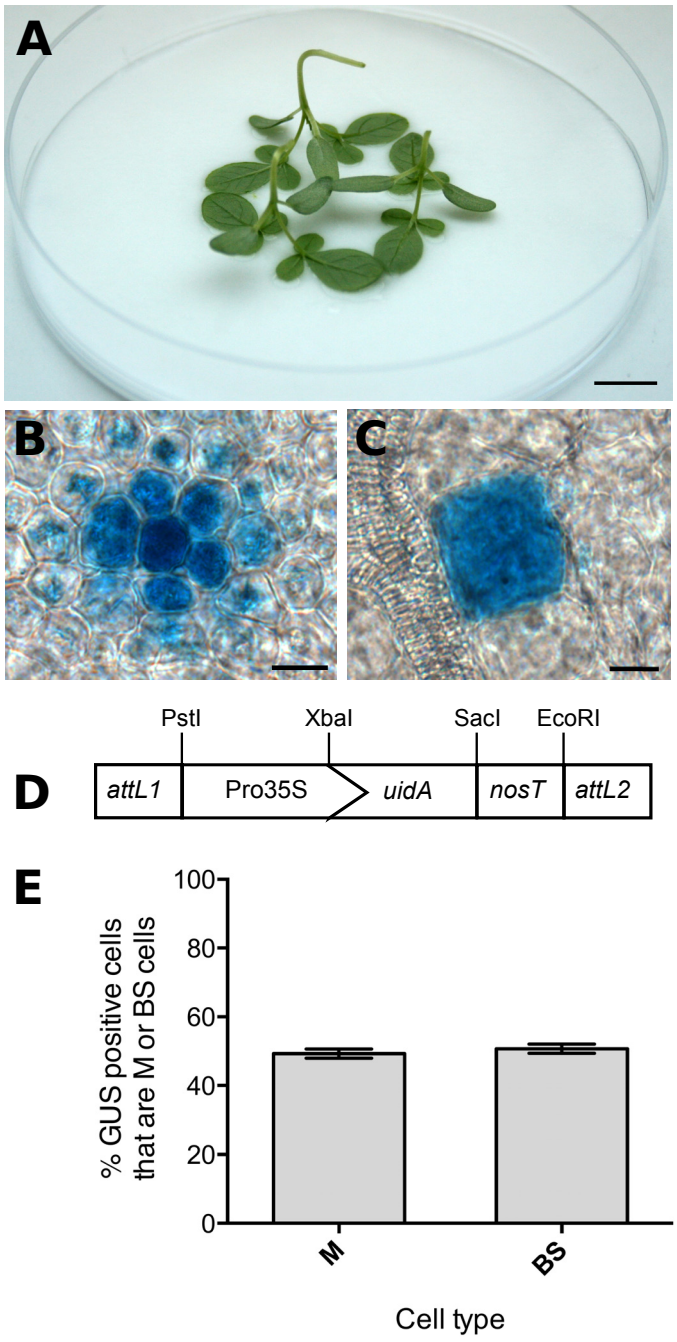
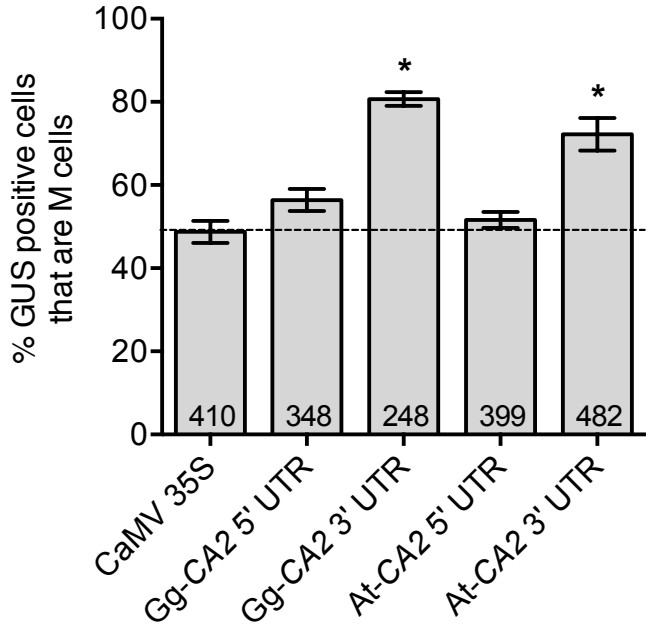
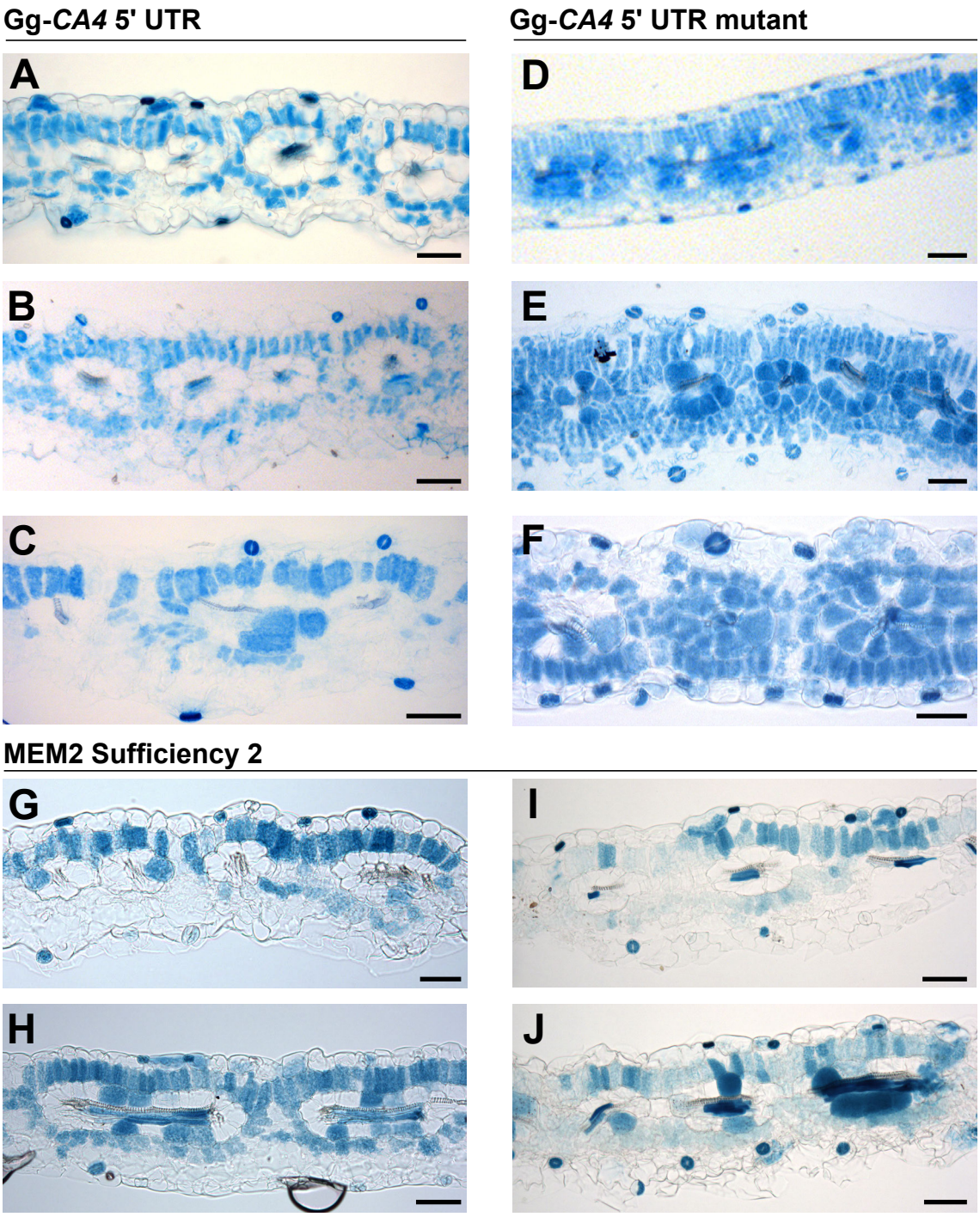


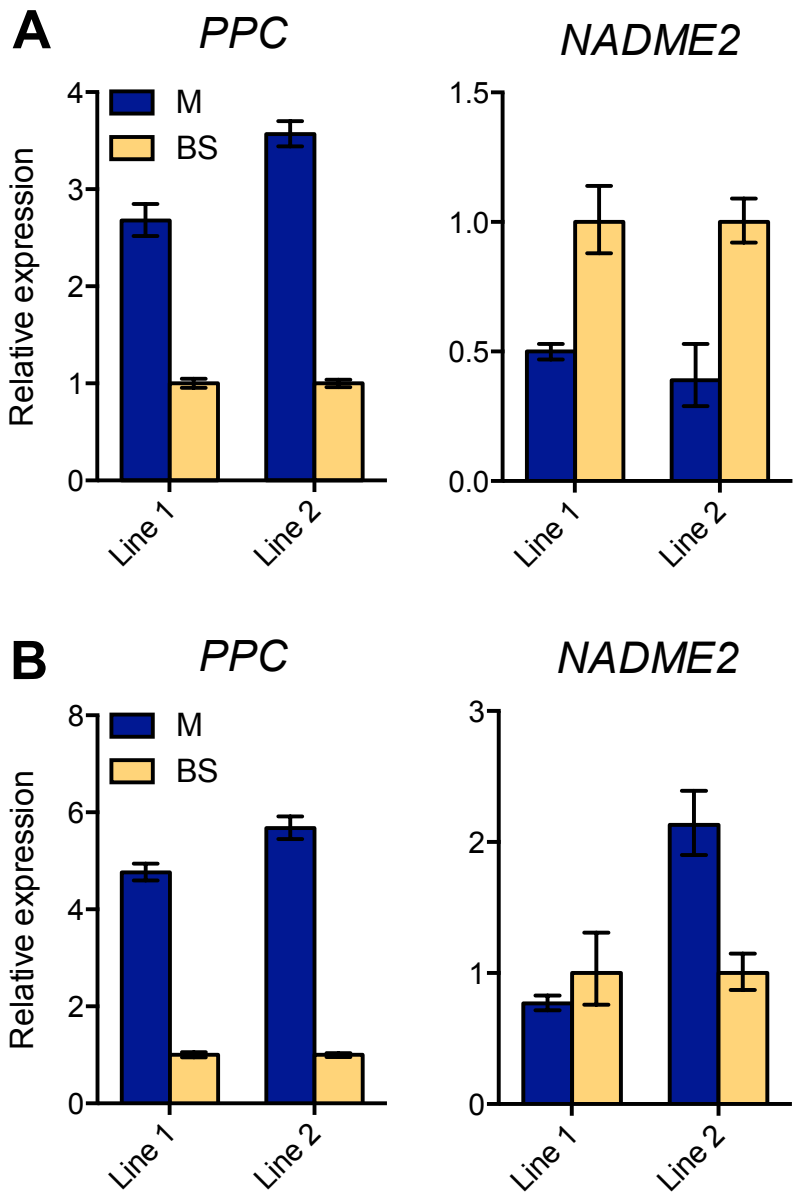
Supplemental Figure 1: Transformation of *G. gynandra* M and BS cells. Abaxial leaf surfaces of *G. gynandra* seedlings (A) were subjected to microprojectile bombardment. This transformed many cells with the reporter GUS, including M (B) and BS (C) cells. Scale bars in A = 1 cm, B-C = 50 µm. The CaMV35S promoter control construct (D) transformed an equal number of M and BS cells (number of cells counted = 394) (E). Error bars represent one standard error.



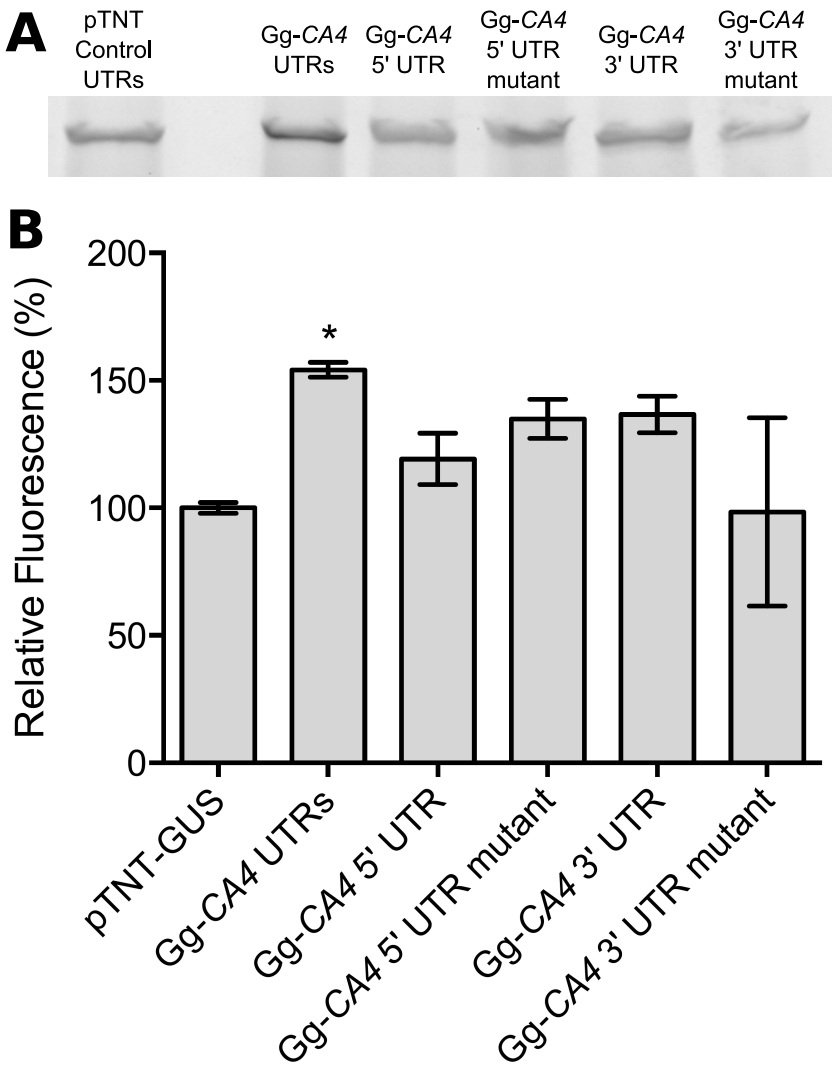
Supplemental Figure 2: 3' UTRs of Gg-*CA2* and At-*CA2* generate M-specificity. Transient transformation of *G. gynandra* leaves by microprojectile bombardment demonstrates that 3' UTRs of Gg-*CA2* and At-*CA2* genes increase the proportion of M-cells expressing GUS relative to a CaMV35S control (dashed lines), whereas 5' UTRs do not direct M-specificity. Numbers within histogram bars represent the number of independently transformed cells for each construct. Asterisks denote statistical significance compared to the control ($p<0.005$, two-tailed student's t-test), error bars denote standard error.



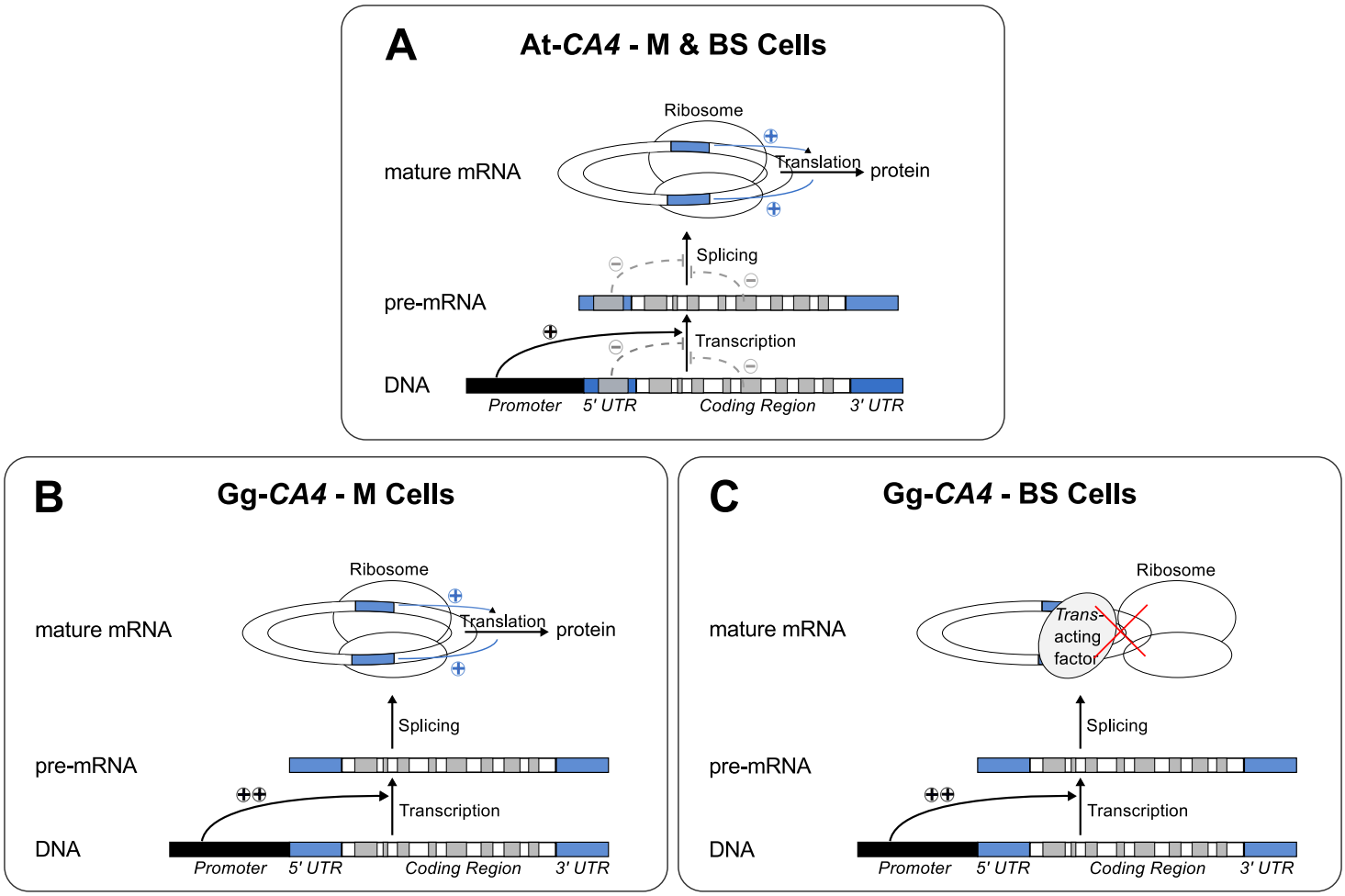
Supplemental Figure 3: Representative transverse sections from stable transgenic lines. Transverse sections of GUS stained leaves from each of three independent stable transgenic lines expressing the Gg-*CA4* 5' UTR fused to GUS (A-C), and three lines expressing the Gg-*CA4* 5' UTR with a 5 nt mutation (D-F). Two representative transverse sections are shown for each of two lines expressing a chimaeric Gg-*CA4* sequence sufficient to direct GUS accumulation primarily in M cells (termed MEM2 sufficiency 2). Line 1 (G-H), Line 2 (I-J). Scale bars = 100 μm.



Supplemental Figure 4: Quantification of marker transcripts in isolated M and BS cells. qRT-PCR quantification of the M cell marker *PPC* and the BS cell marker *NADME2* in A) lines expressing *GUS* under the control of the Gg-*CA4* 5' UTR and B) lines expressing *GUS* under the control of a mutated Gg-*CA4* 5' UTR. Error bars represent one standard error. Data are presented as mean of three replicates for each line and cell type.



Supplemental Figure 5: Mutation of MEM2 does not effect translation *in vitro*. *In vitro* translation of *uidA* encoding GUS sequence with or without Gg-*CA4* UTRs. Translation in vitro produced more protein with Gg-*CA4* UTRs compared with the pTNT vector control UTRs (A). Quantification by fluorescent gel analysis showed a 60% increase in protein synthesized from RNA with both 5' and 3' Gg-*CA4* UTRs (B). Mutation of MEM2 did not alter translation rate. Asterisks denote statistical significance compared to the control ($p = <0.05$, two-tailed student's t-test), error bars denote one standard error. Results in (B) represent the mean of two independent experiments.



Supplemental Figure 6: Hypothesis for mechanisms regulating the abundance and cell-specificity of *CA4*. In the ancestral C_3 state (A), high levels of *CA4* expression in *A. thaliana* are conferred by the promoter region (black). Introns (grey), exons (white) and UTRs (blue) have little net effect on At-*CA4* abundance, but repressive elements in the 5' UTR intron and 3' UTR reduce expression. The promoter of *CA4* from C_4 *G. gynandra* generates higher expression than At-*CA4* (B and C). The repressive 5' UTR intron is not present in Gg-*CA4*, perhaps contributing to this increased abundance. M-specificity is conferred by sequences present in both the 5' and 3' UTR of Gg-*CA4*. Gg-*CA4* transcripts are equally abundant in M and BS cells, but preferentially translated in M cells (B). We hypothesize the presence of a *trans*-acting factor that reduces translation (denoted by red cross) of Gg-*CA4* in BS cells (C). Pointed arrowheads represent mechanisms conferring increased abundance; flat arrowheads represent mechanisms conferring decreased abundance. Dashed arrow lines represent mechanisms that may operate transcriptionally or post-transcriptionally. (+) symbols denote a positive effect on overall expression, (-) symbols denote a negative effect.

UTR Construct	**TOTAL M Cells**	**Total BS Cells**	**Total Cells**
CaMV 35S Control	595	613	1208
Gg-*CA4* 5' UTR	323	114	437
Gg-*CA4* 5' 60F deletion	132	64	196
Gg-*CA4* 5' 80F deletion	441	194	635
Gg-*CA4* 5' 100F deletion	174	186	360
Gg-*CA4* 5' 20R deletion	107	104	211
Gg-*CA4* 5' 14R deletion	272	352	624
Gg-*CA4* 5' 7R deletion	320	413	733
Gg-*CA4* 5' mutant	234	187	421
Gg-*CA4* 3' UTR	283	81	364
Gg-*CA4* 3' mutant	153	165	318
Gg-*CA2* 3' UTR	199	49	248
Gg-*CA2* 3' mutant	210	188	398
Gg-*PPDK* 5' UTR	224	101	325
Gg-*PPDK* 5' mutant	223	203	426
Gg-*PPDK* 3' UTR	143	64	207
Gg-*PPDK* 3' mutant	202	149	351
Gg-*CA2* 5' UTR	198	150	348
At-*CA2* 5' UTR	204	195	399
At-*CA2* 3' UTR	310	172	482
MEM2 Sufficiency 1	547	297	844
MEM2 Sufficiency 2	609	248	857

Supplemental Table 1

RNA-Seq replicate	Transcript abundance (effective counts)	
	M	BS
Rep 1	2307	3403
Rep 2	2013	2242
Rep 3	1650	2491
Mean	2253	2337
Fold Change	1.04	
log2 Fold Change	0.05	
p-value (Fisher's exact test)	0.729	

Supplemental Table 2: Gg-*C* *A4* transcript abundance, measured by transcriptome sequencing of M and BS cells

Construct	No. of replicates	Median GUS activity (pmol MU/min/µg protein	Standard deviation	p value (vs. mutant, two-tailed t-test)
Gg-*CA4* 5' UTR	18	9.5	16.4	0.89
Gg-*CA4* 5' UTR mutant	14	11.2	23.9	
Gg-*CA4* 3' UTR	11	4.1	18.2	0.15
Gg-*CA4* 3' UTR mutant	12	4.9	15.9	

Supplemental Table 3