

Figure 1 Enhancers from photosynthesis genes show light-responsive activity. **A** and **B**, Full-length (FL) enhancers, as well as 169-bp long segments from their 5' or 3' end, of the *Pisum sativum* AB80 and *rbcS-E9* genes and the *Triticum aestivum* Cab-1 gene were cloned upstream of the 35S minimal promoter driving the expression of a barcoded GFP reporter gene (**A**). All constructs were pooled and the viral 35S enhancer was added as an internal control. The pooled enhancer library was subjected to Plant STARR-seq in tobacco leaves with plants grown for 2 days in normal light/dark cycles (+ light) or completely in the dark (- light) prior to RNA extraction (**B**). Enhancer strength was normalized to a control construct without an enhancer (\log_2 set to 0). **C**, Light-responsiveness ($\log_2[\text{enhancer strength}^{\text{light}}/\text{enhancer strength}^{\text{dark}}]$) was determined for the indicated enhancer segments. **D**, Transgenic *Arabidopsis* lines were generated with T-DNAs harboring a constitutively expressed luciferase (Luc) gene and a nanoluciferase (NanoLuc) gene under control of a 35S minimal promoter coupled to the 35S enhancer or the 3' segments of the AB80, Cab-1, or rbcS-E9 enhancers. **E**, Nanoluciferase activity was measured in 5 T2 plants from these lines and normalized to the activity of luciferase. The NanoLuc/Luc ratio was normalized to a control construct without an enhancer (none; \log_2 set to 0). **F**, The mean NanoLuc/Luc ratio was compared to the mean enhancer strength determined by STARR-seq. Pearson's R^2 , Spearman's ρ , and number (n) of enhancers are indicated. A linear regression line is shown as a dashed line. Error bars represent the 95% confidence interval. Box plots in **B**, **C**, and **E** represent the median (center line), upper and lower quartiles (box limits), 1.5 \times interquartile range (whiskers), and outliers (points) for all corresponding barcodes (**B** and **C**) or plant lines (**E**) from two (**B** and **C**) or three (**E**) independent replicates. Numbers at the bottom of each box plot indicate the number of samples in each group.

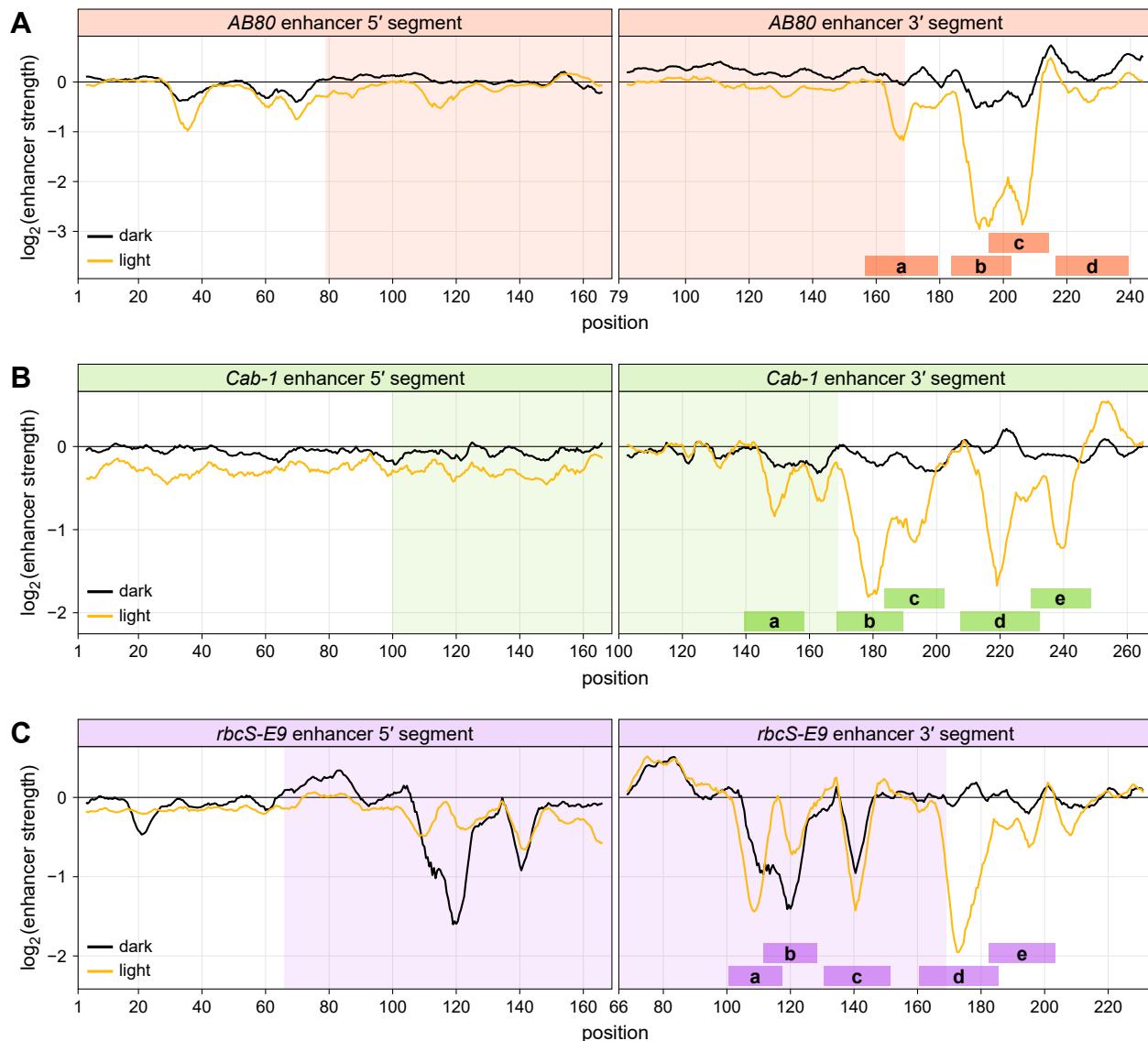


Figure 2 The *AB80*, *Cab-1*, and *rbcS-E9* enhancers contain multiple mutation-sensitive regions. **A–C**, All possible single-nucleotide substitution, deletion, and insertion variants of the 5' and 3' segments of the *AB80* (**A**), *Cab-1* (**B**), and *rbcS-E9* (**C**) enhancers were subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (light) or completely in the dark (dark) for two days prior to RNA extraction. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0). A sliding average (window size = 6 bp) of the mean enhancer strength for all variants at a given position is shown. The shaded area indicates the region where the 5' and 3' segments overlap. Mutation-sensitive regions in the 3' enhancer segments are indicated by shaded rectangles labeled a–e.

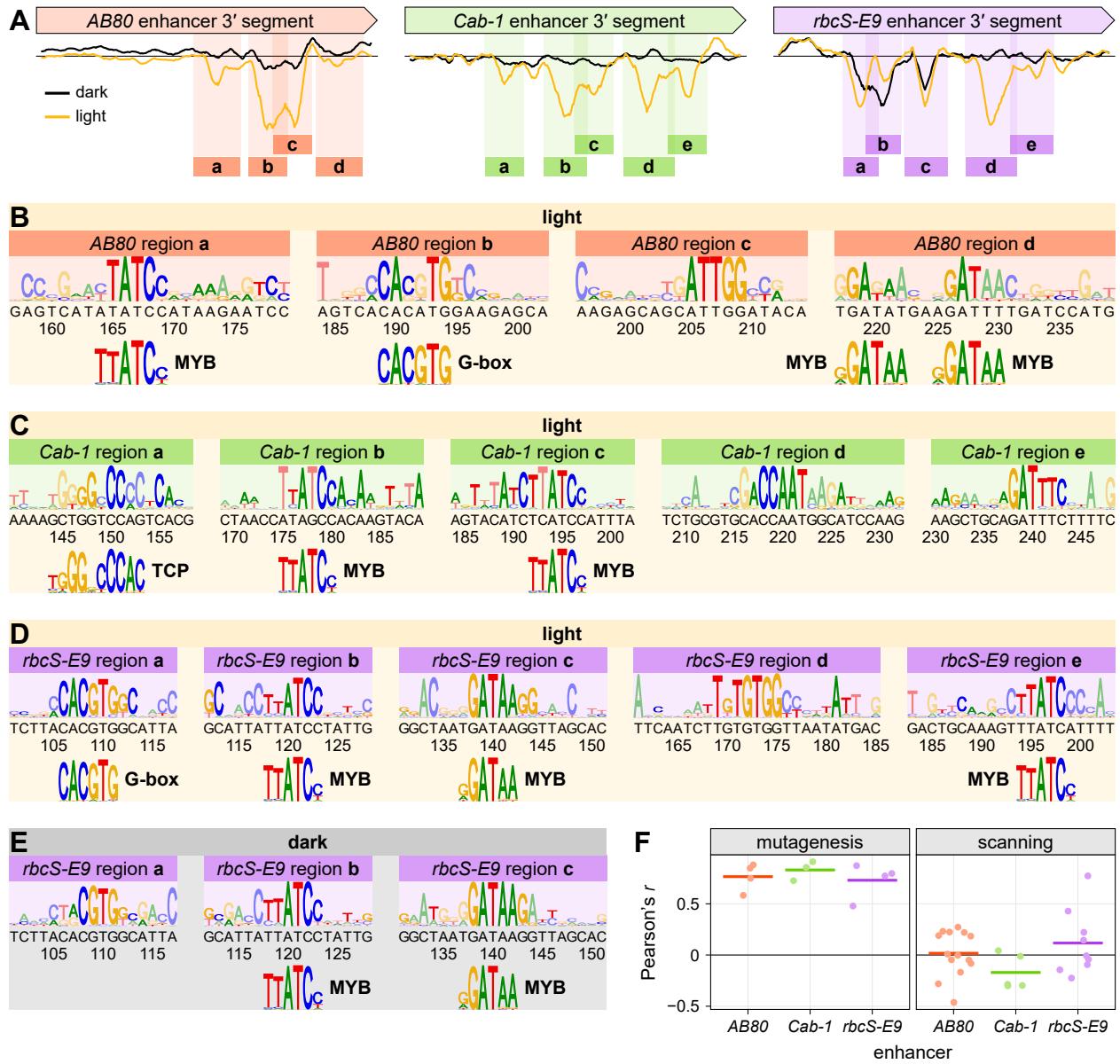


Figure 3 Mutation-sensitive regions harbor transcription factor binding sites. **A**, 4 to 5 mutation-sensitive regions (shaded rectangles; labeled a–e) were defined for the 3' segments of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers. The mutational sensitivity plots are reproduced from Figure 2. **B–E**, Sequence logo plots were generated from the enhancer strength in the light (**B–D**) or dark (**E**) of all possible single-nucleotide substitution variants within the indicated mutation-sensitive regions of the *AB80* (**B**), *Cab-1* (**C**), or *rbcS-E9* (**D** and **E**) enhancers. The sequence of the wild-type enhancer and the position along it is shown on the x axis. Letters with dark colors in the logo plot represent wild-type bases. The sequence logos for each region were compared to known transcription factor binding motifs and significant matches are shown below the plots. **F**, For each transcription factor binding motif matching a sequence logo plots derived from the saturation mutagenesis data in the light (mutagenesis; see **B–D**) or identified by the motif-scanning approach (scanning; see Supplemental Figure S7), the correlation (Pearson's r) between the strength of an enhancer variant and the score of how well the variant sequence matches this motif is plotted as points. The lines represent the average correlation for all motifs of a given enhancer.

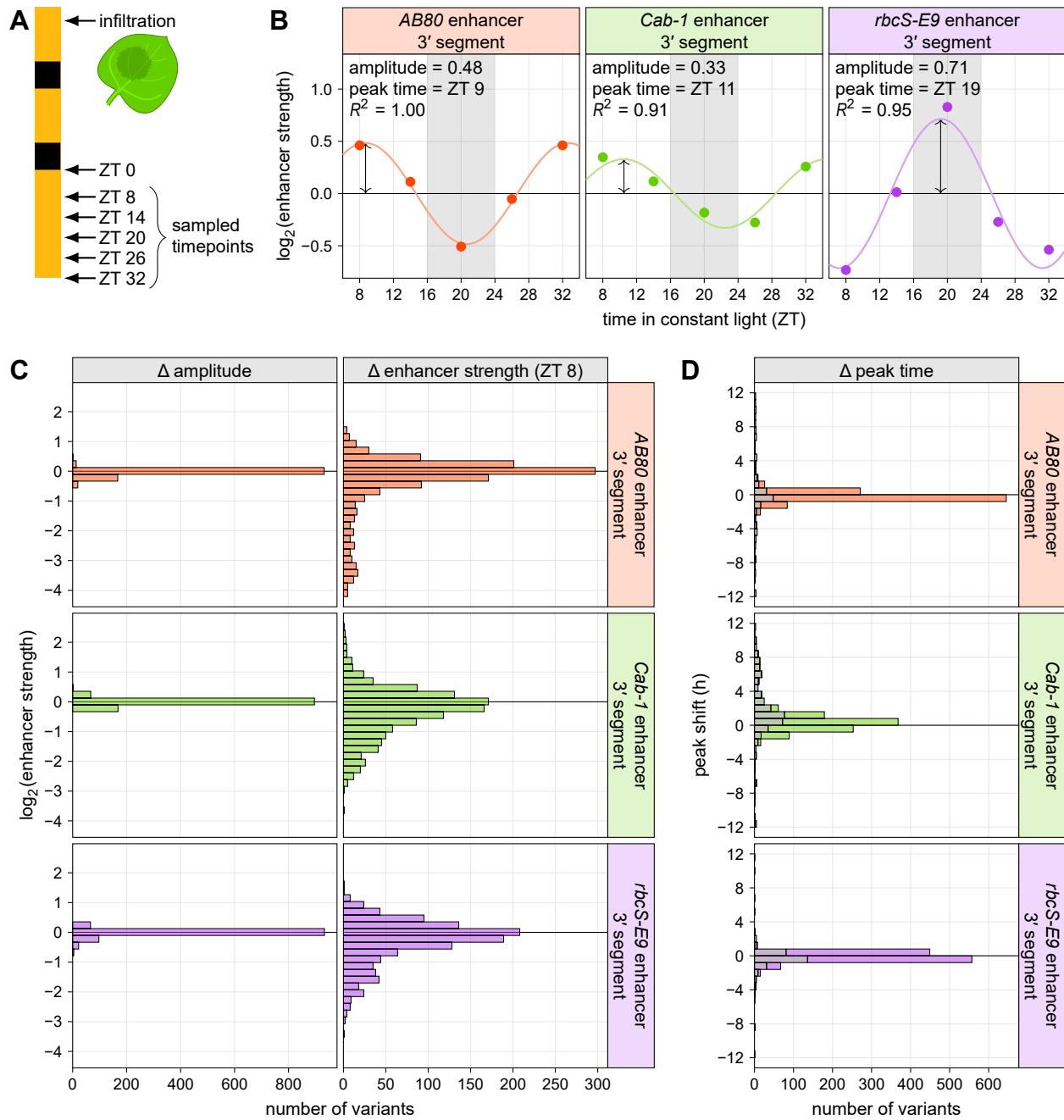


Figure 4 Circadian oscillation is robustly encoded in the *AB80*, *Cab-1*, and *rbcS-E9* enhancers. **A**, All possible single-nucleotide variants of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers were subjected to Plant STARR-seq in tobacco leaves. On the morning of the third day after transformation (ZT 0), the plants were shifted to constant light. Leaves were harvested for RNA extraction starting at mid-day (ZT 8) and in 6 hour intervals (ZT 14, 20, 26, and 32) afterwards. **B**, A sine wave with a period of 24 h was fitted to the enhancer strength of a given variant across all sampled time points. The fitted line is plotted together with the measured data points for the wild-type enhancers. The equilibrium point of the curves was set to 0. The amplitude is shown as a two-sided arrow at the time of highest enhancer strength (peak time). The goodness-of-fit (R^2) is indicated. The shaded gray area represents the timing of the dark period if the plants had not been shifted to constant light. **C** and **D**, Histograms of the difference between the amplitude (**C**) and peak time (**D**) of each single-nucleotide variant relative to the wild-type enhancer. For comparison, the difference in enhancer strength at ZT 8 is also shown in **C**. Variants with a below average goodness-of-fit are grayed out in **D**. Only data for the 3' enhancer segments is shown.

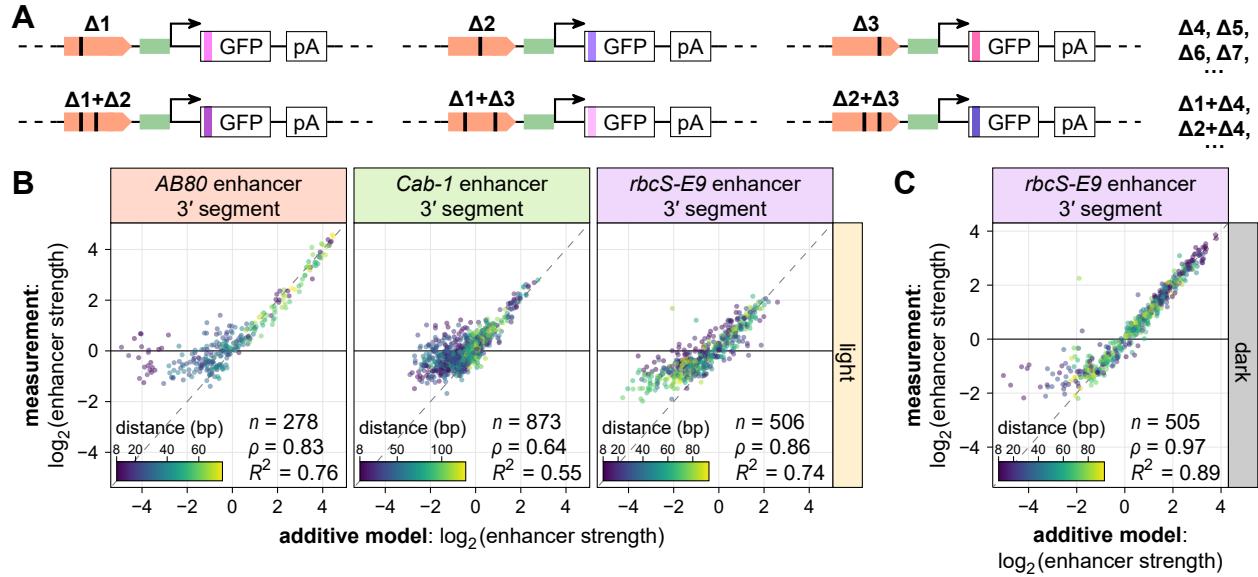


Figure 5 Epistatic interactions between single-nucleotide deletions. **A–C**, Selected single-nucleotide deletion variants (**A**; $\Delta 1, \Delta 2, \Delta 3, \dots$) of the 3' segment of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers and all possible combinations with two of these deletions (**A**; $\Delta 1+\Delta 2, \Delta 1+\Delta 3, \Delta 2+\Delta 3, \dots$) were subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (**B**) or completely in the dark (**C**) for two days prior to RNA extraction. For each pair of deletions, the expected enhancer strength based on the sum of the effects of the individual deletions (additive model) is plotted against the measured enhancer strength. The color of the points represents the distance between the two deletions in a pair.

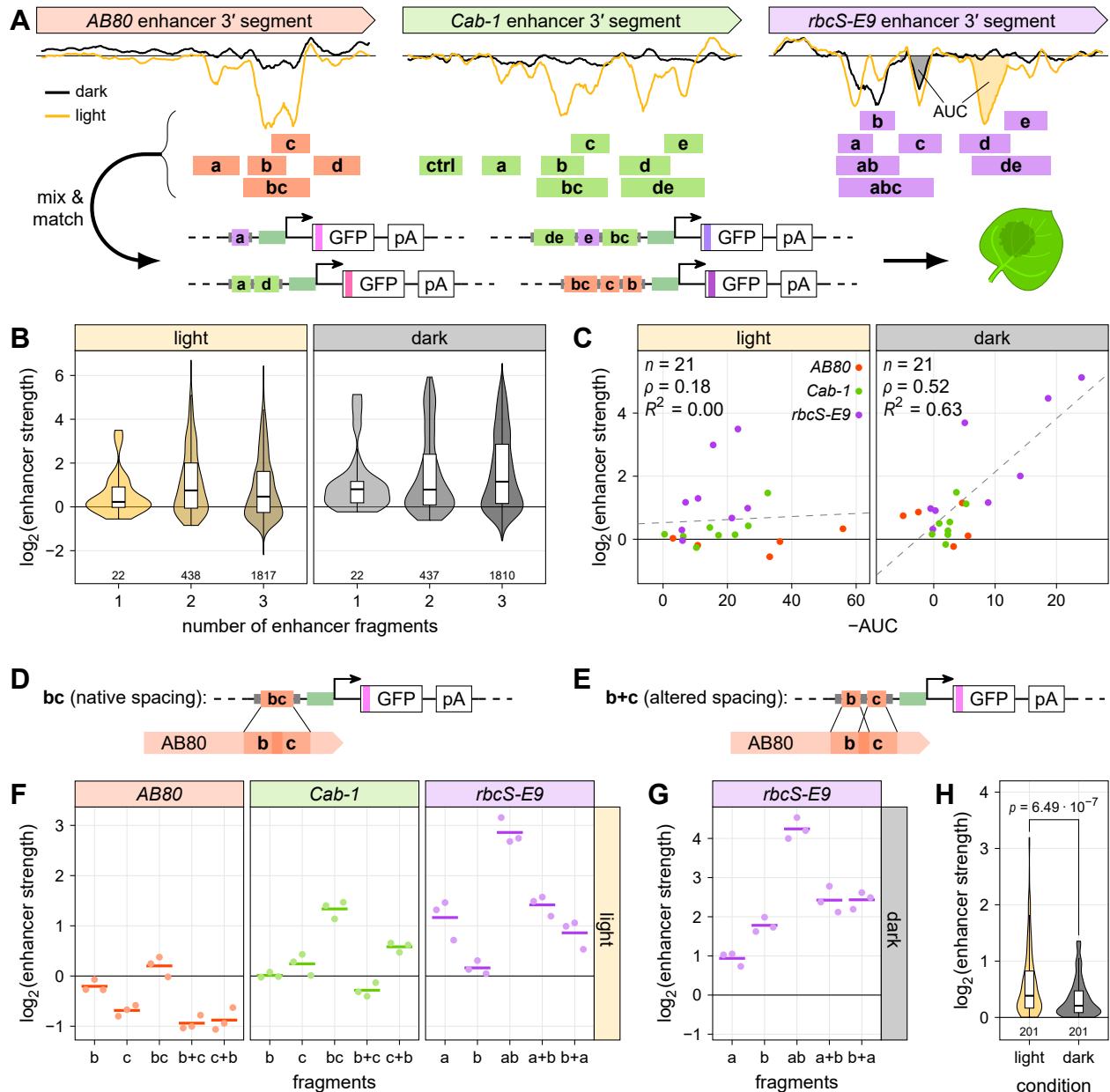


Figure 6 The number, spacing, and order of mutation-sensitive regions affects enhancer strength. **A**, Fragments of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers spanning 1–3 mutation-sensitive regions (shaded rectangles; labeled a–e, ab, abc, bc, de) as well as a control fragment (ctrl) from a mutation-insensitive region in *Cab-1* and a shuffled version of the *AB80* fragment d were ordered as oligonucleotides. These fragments were randomly combined to create synthetic enhancers with up to three fragments which were then subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (light) or completely in the dark (dark) for two days prior to RNA extraction. The mutational sensitivity plots are reproduced from Figure 2. **B**, Violin plots of the strength of the synthetic enhancers grouped by the number of contained fragments. **C**, For each enhancer fragment, the area under the curve (AUC) in the mutational sensitivity plots was calculated and plotted against the fragment's enhancer strength. AUCs in the dark or light for *rbcS-E9* fragments c and d, respectively, are shown in **A**. Pearson's R^2 , Spearman's ρ , and number (n) of enhancer fragments are indicated. A linear regression line is shown as a dashed line. **D–G**, Plots of the strength of enhancer fragments (**D**) or fragment combinations (separated by a + sign and shown in the order in which they appear in the construct; **E**) in three replicates (points) and the mean strength (lines). Enhancer strength was determined using tobacco plants grown in the light (**F**) or dark (**G**) prior to RNA extraction. **H**, Violin plots of the difference in enhancer strength between synthetic enhancers harboring the same two enhancer fragments but in different order. The p -value from a two-sided Wilcoxon rank-sum test comparing light and dark results is indicated (p). Violin plots in **B** and **H** represent the kernel density distribution and the box plots inside represent the median (center line), upper and lower quartiles and $1.5 \times$ interquartile range (whiskers) for all corresponding synthetic enhancers. Numbers at the bottom of each violin indicate the number of elements in each group. Enhancer strength in **B–G** was normalized to a control construct without an enhancer (\log_2 set to 0).

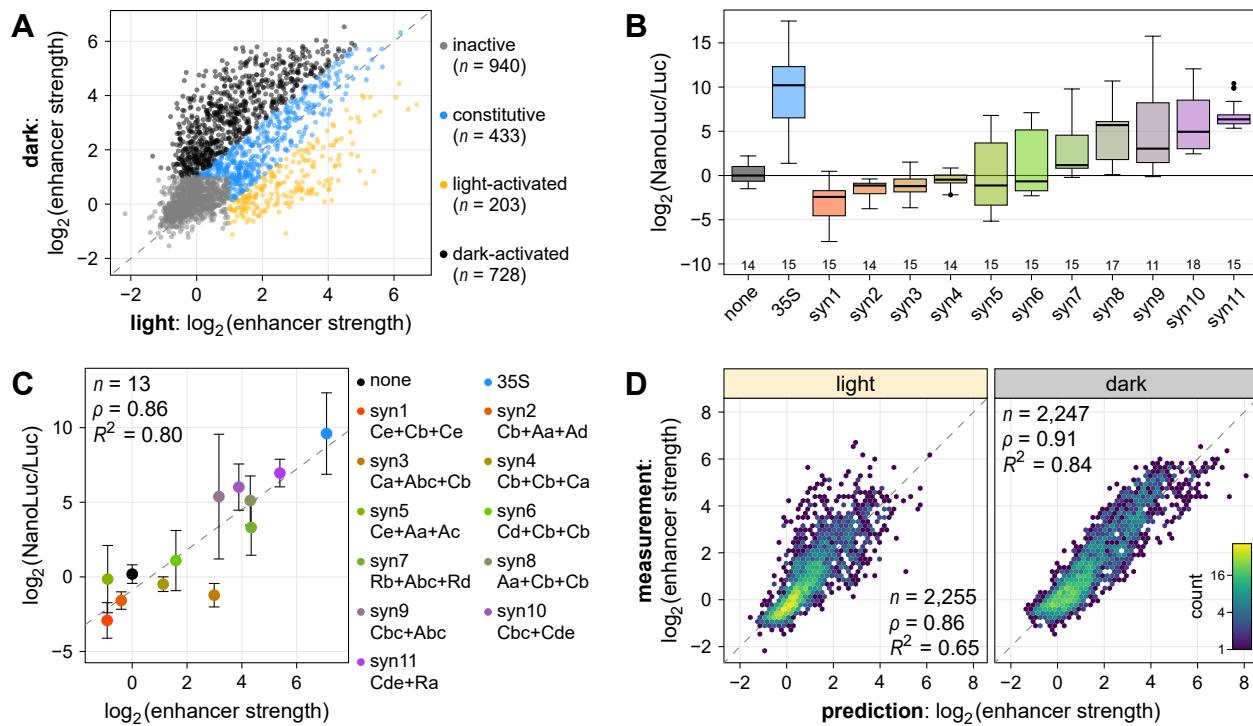
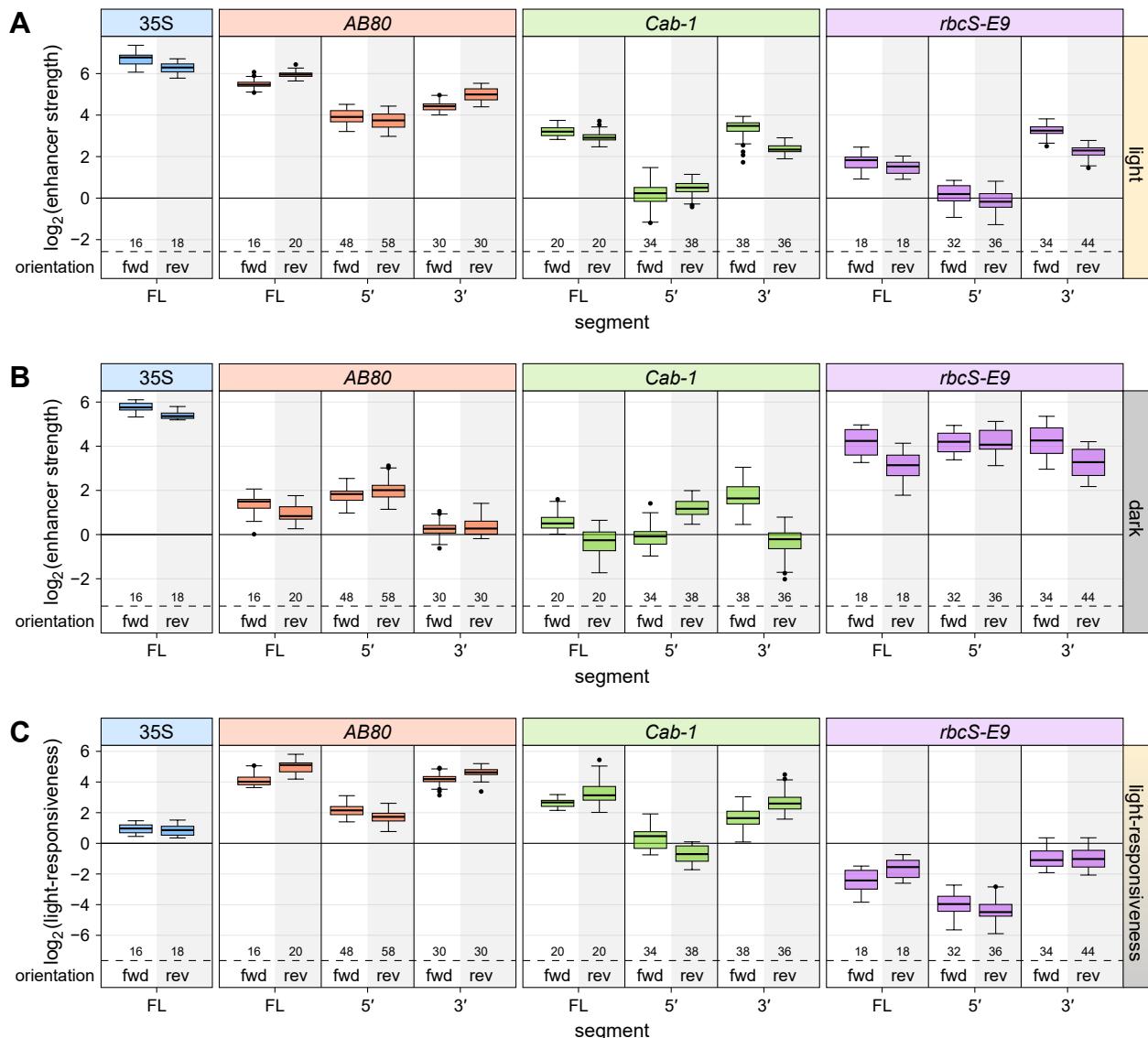
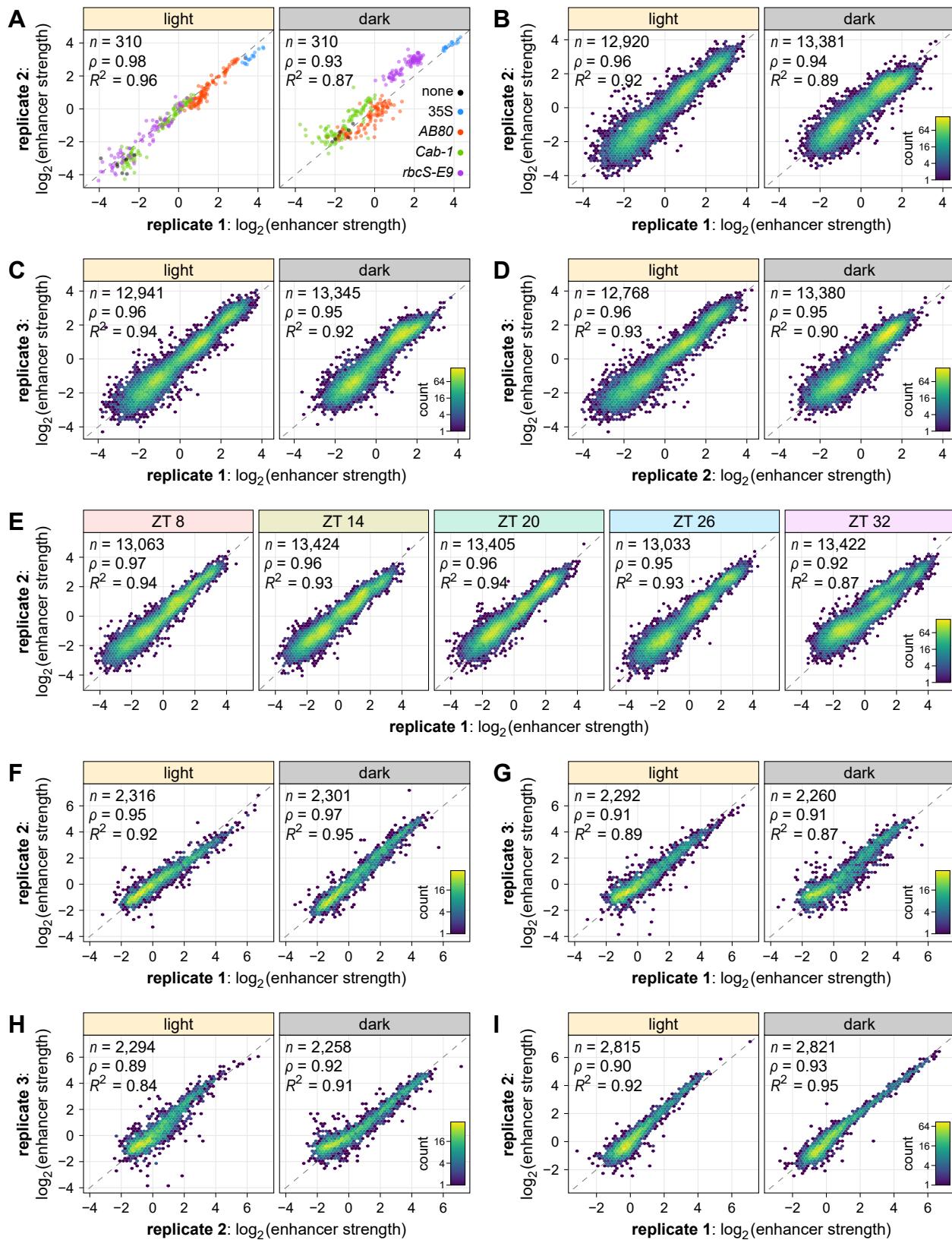


Figure 7 Enhancer fragments can be used to build condition-specific synthetic enhancers. **A**, Plot of the strength of synthetic enhancers created by randomly combining up to three fragments derived from mutation-sensitive regions of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers (see Figure 6A) as measured by Plant STARR-seq in the light or dark. The synthetic enhancers were grouped into four categories: inactive, $\log_2(\text{enhancer strength}) \leq 1$ in both conditions; constitutive, similar strength in both conditions; light-activated, at least two-fold more active in the light; dark-activated, at least two-fold more active in the dark. The number (n) of synthetic enhancers in each category is indicated. **B**, Dual-luciferase reporter constructs (see Figure 1D) were created for 11 synthetic enhancers (syn1–11). Nanoluciferase activity was measured in at least 4 T2 plants from these lines and normalized to the activity of luciferase. The NanoLuc/Luc ratio was normalized to a control construct without an enhancer (none; \log_2 set to 0). Box plots are as defined in Figure 1E. **C**, The mean NanoLuc/Luc ratio was compared to the mean enhancer strength determined by STARR-seq. A linear regression line is shown as a dashed line. Error bars represent the 95% confidence interval. The constituent fragments of the synthetic enhancers are indicated with fragments separated by a + sign. The first letter indicates the enhancer from which the fragment is derived (A, *AB80*; C, *Cab-1*; R, *rbcS-E9*) and the lowercase letters represent the fragment name. **D**, A linear model was built to predict the strength of the synthetic enhancers based on the strength of the constituent individual fragments. Hexbin plots (color represents the count of points in each hexagon) of the correlation between the model's prediction and the measured data are shown. In **C** and **D**, Pearson's R^2 , Spearman's ρ , and number (n) of synthetic enhancers are indicated.

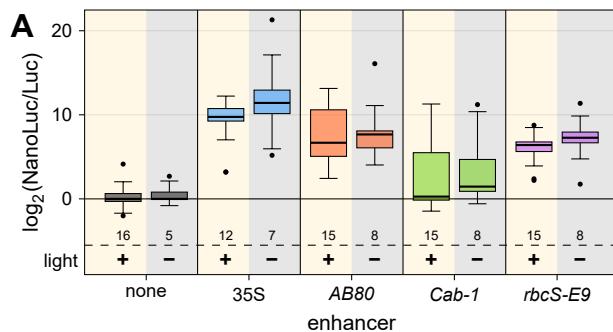


Supplemental Figure S1 Enhancer strength and light-responsiveness is orientation-independent. (Supports Figure 1)

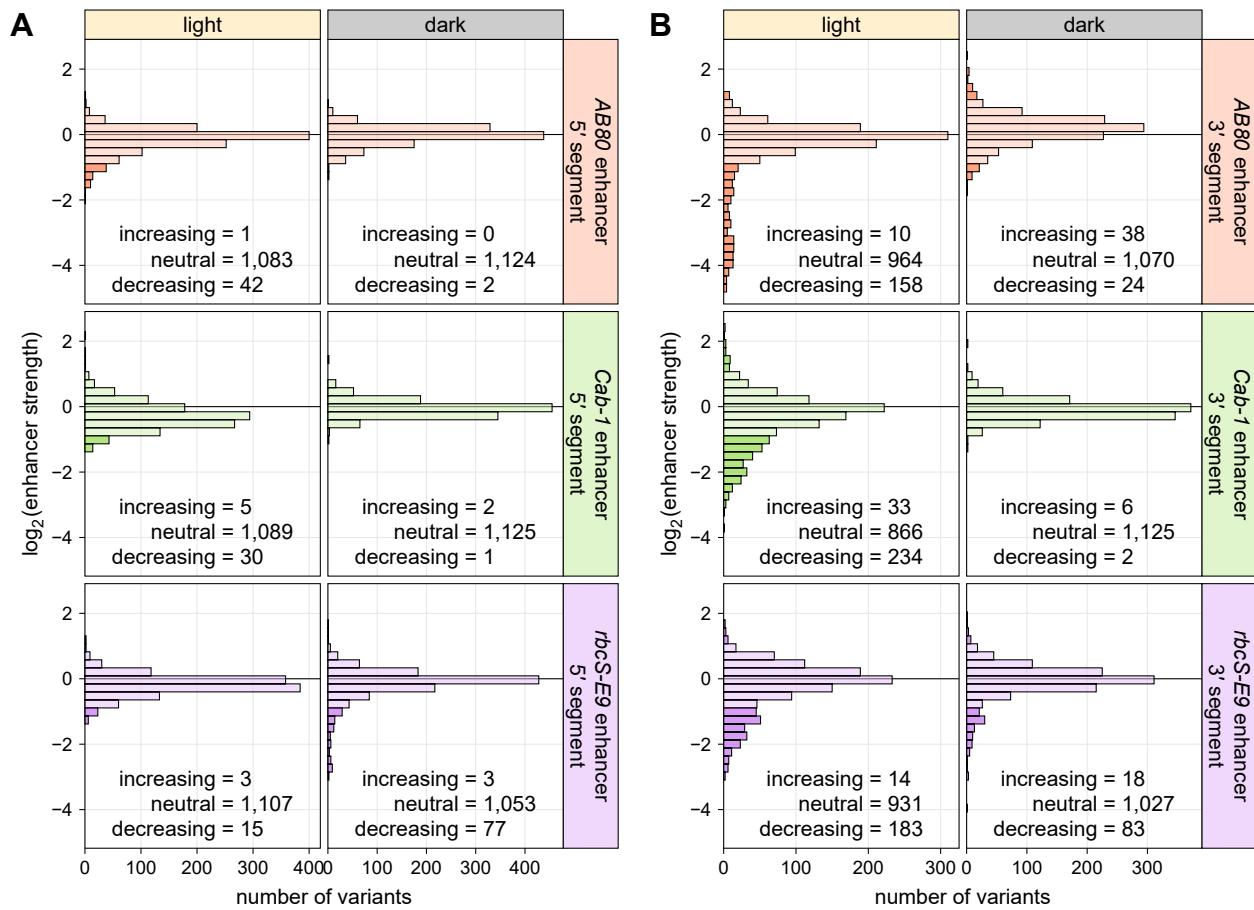
A and **B**, Full-length (FL) enhancers, as well as 169-bp long segments from their 5' or 3' end, of the *Pisum sativum* AB80 and *rbcS-E9* genes and the *Triticum aestivum* Cab-1 gene were cloned in the forward (fwd; data reproduced from Figure 1, B and C) or reverse (rev) orientation upstream of the 35S minimal promoter driving the expression of a barcoded GFP reporter gene. All constructs were pooled and the viral 35S enhancer was added as an internal control. The pooled enhancer library was subjected to Plant STARR-seq in tobacco leaves with plants grown for 2 days in normal light/dark cycles (**A**) or completely in the dark (**B**) prior to RNA extraction. Enhancer strength was normalized to a control construct without an enhancer (\log_2 set to 0). **C**, Light-responsiveness ($\log_2[\text{enhancer strength}^{\text{light}}/\text{enhancer strength}^{\text{dark}}]$) was determined for the indicated enhancer segments. Box plots represent the median (center line), upper and lower quartiles (box limits), 1.5 \times interquartile range (whiskers), and outliers (points) for all corresponding barcodes from two replicates. Numbers at the bottom of each box plot indicate the number of barcodes in each group.



Supplemental Figure S2 Plant STARR-seq yields highly reproducible results. (Supports all figures) **A–I**, Correlation between biological replicates of Plant STARR-seq for the enhancer library used in Figure 1 and Supplemental Figure S1 (**A**), the single-nucleotide enhancer variants library used in Figures 2–4 and Supplemental Figures S4–S13 (**B–E**), the synthetic enhancer library used in Figures 6 and 7 and Supplemental Figures S13–S15 (**F–H**), and the combined double-deletion and synthetic enhancer validation library used in Figure 5 and Supplemental Figure S13 (**I**) performed under the indicated condition or at the indicated time points. Pearson's R^2 , Spearman's ρ , and number (n) of enhancer variants or enhancer fragment combinations are indicated. The color in the hexbin plots in (**B–I**) represents the count of points in each hexagon.

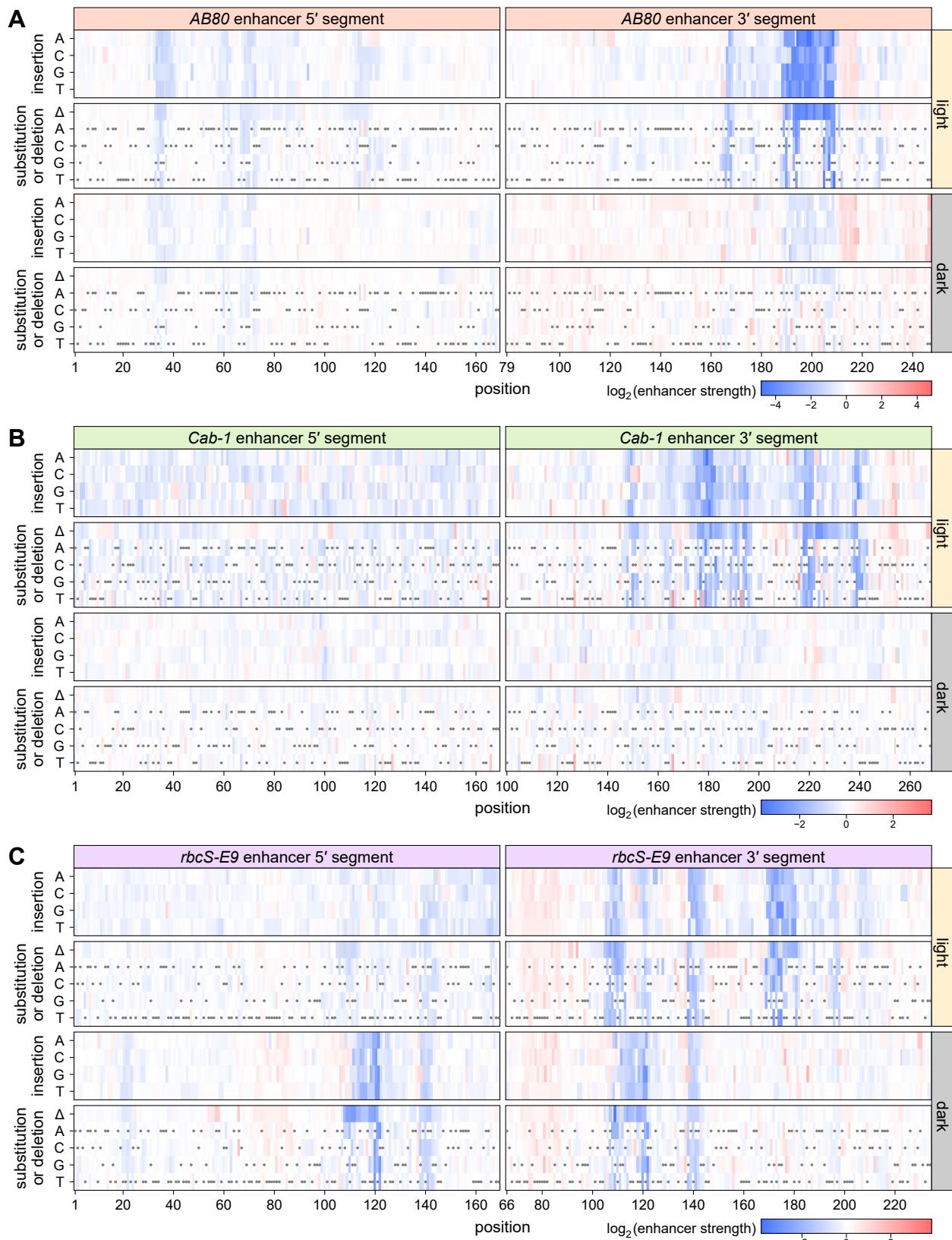


Supplemental Figure S3 The dual-luciferase assay cannot detect light-responsive enhancer activity. (Supports Figure 1) **A**, Transgenic *Arabidopsis* lines were generated with constructs harboring a constitutively expressed luciferase (Luc) gene and a nanoluciferase (NanoLuc) gene under control of a 35S minimal promoter coupled to the 35S enhancer or the 3' segments of the AB80, Cab-1, or *rbcS-E9* enhancers (see Figure 1D). Nanoluciferase activity was measured in 3–5 T2 plants from these lines and normalized to the activity of luciferase. Plants were either grown in normal light/dark cycles (+ light; data reproduced from Figure 1E) or shifted to complete darkness (− light) for 4 days prior to sample collection. The NanoLuc/Luc ratio was normalized to a control construct without an enhancer (none; \log_2 set to 0). Box plots represent the median (center line), upper and lower quartiles (box limits), 1.5× interquartile range (whiskers), and outliers (points) for all corresponding plant lines from two (− light) or three (+ light) independent replicates. Numbers at the bottom of each box indicate the number of samples in each group.

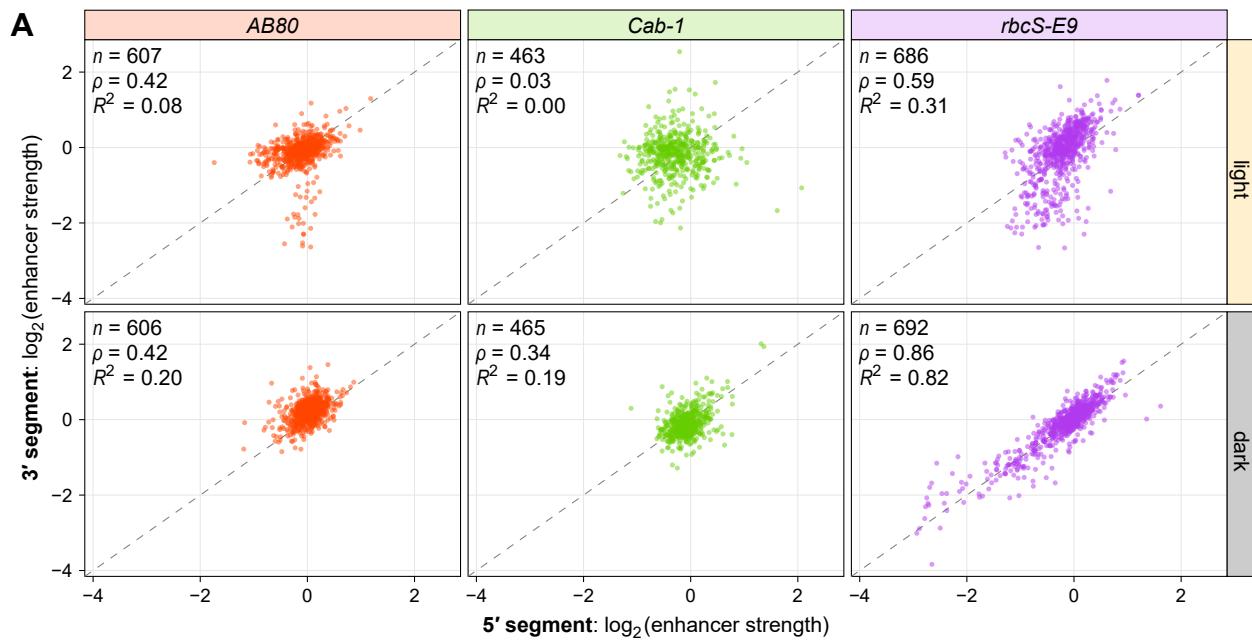


Supplemental Figure S4 Few single-nucleotide mutations have a strong effect on enhancer strength. (Supports Figures 2 and 3) **A** and **B**, All possible single-nucleotide substitution, deletion, and insertion variants of the 5' (**A**) and 3' (**B**) segments of the AB80, Cab-1, and rbcS-E9 enhancers were subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (light) or completely in the dark (dark) for two days prior to RNA extraction. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0). Variants were grouped into three categories: increasing, $\log_2(\text{enhancer strength}) > 1$; neutral, $\log_2(\text{enhancer strength})$ between -1 and 1 ; decreasing, $\log_2(\text{enhancer strength}) < -1$. The number of variants in each category is indicated. Neutral variants are shown in a lighter color than increasing and decreasing variants in the histograms.

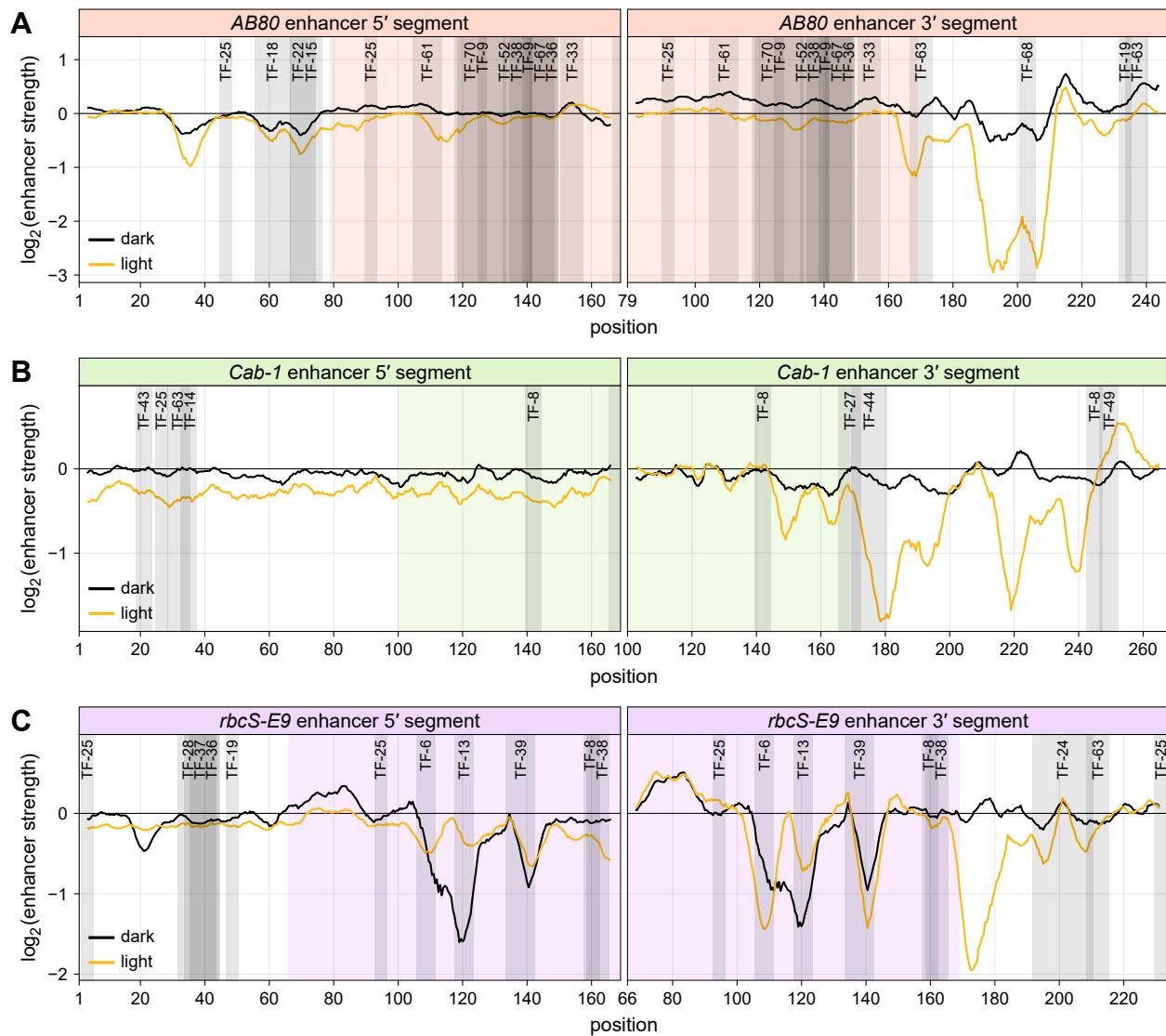
Supplemental Data. Jores et al.



Supplemental Figure S5 Saturation mutagenesis reveals mutation-sensitive patches in plant enhancers. (Supports Figures 2 and 3) **A–C**, All possible single-nucleotide substitution, deletion, and insertion variants of the 5' and 3' segments of the *AB80* (**A**), *Cab-1* (**B**), and *rbcS-E9* (**C**) enhancers were subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (light) or completely in the dark (dark) for two days prior to RNA extraction. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0) and plotted as a heatmap. Missing values are shown in light gray and wild-type variants are marked with a gray dot.

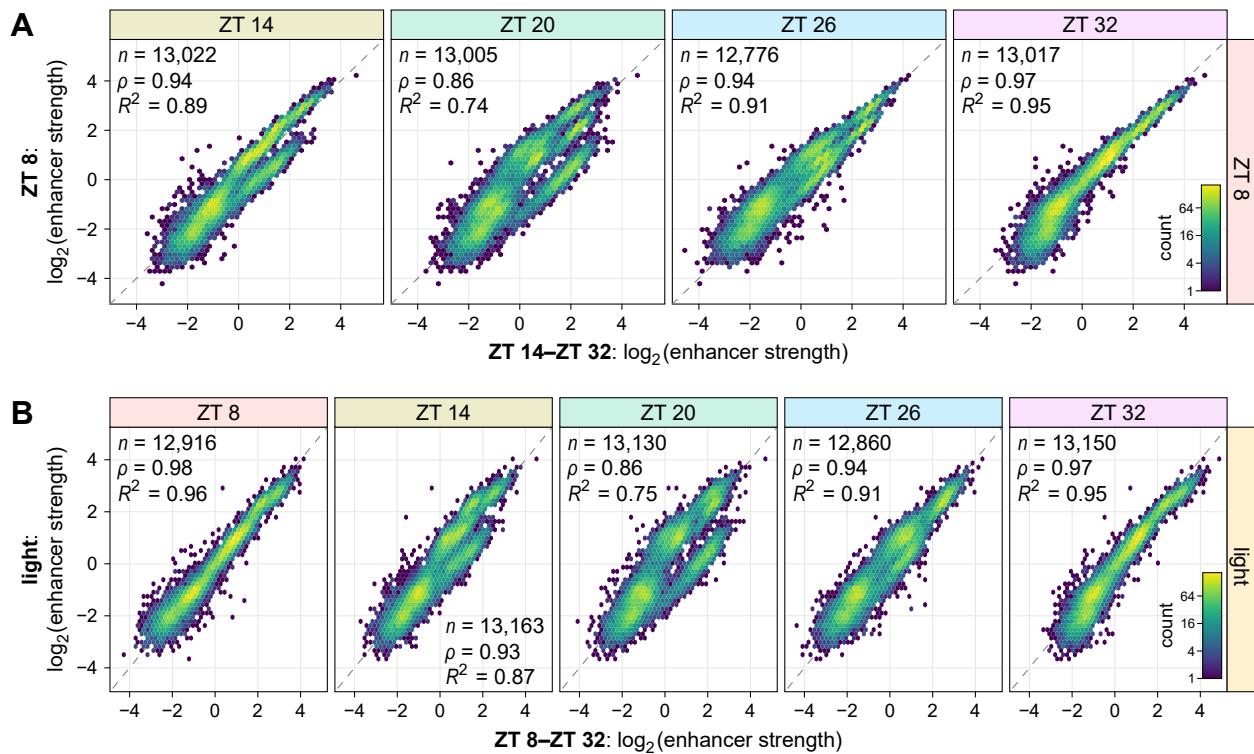


Supplemental Figure S6 Effects of mutations in the overlap region of the 5' and 3' enhancer segments are more similar in the dark. (Supports Figure 2) **A**, All possible single-nucleotide substitution, deletion, and insertion variants of the 5' and 3' segments of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers were subjected to Plant STARR-seq in tobacco plants grown in normal light/dark cycles (light) or completely in the dark (dark) for two days prior to RNA extraction. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0). For mutations located in the overlap region between the two segments (positions 79–169 in *AB80*, positions 100–169 in *Cab-1*, and positions 66–169 in *rbcS-E9*), the normalized enhancer strength measured in the context of the 5' segment (x axis) is compared against the normalized enhancer strength measured in the context of the 3' segment (y axis). Pearson's R^2 , Spearman's ρ , and number (n) of enhancer variants are indicated.

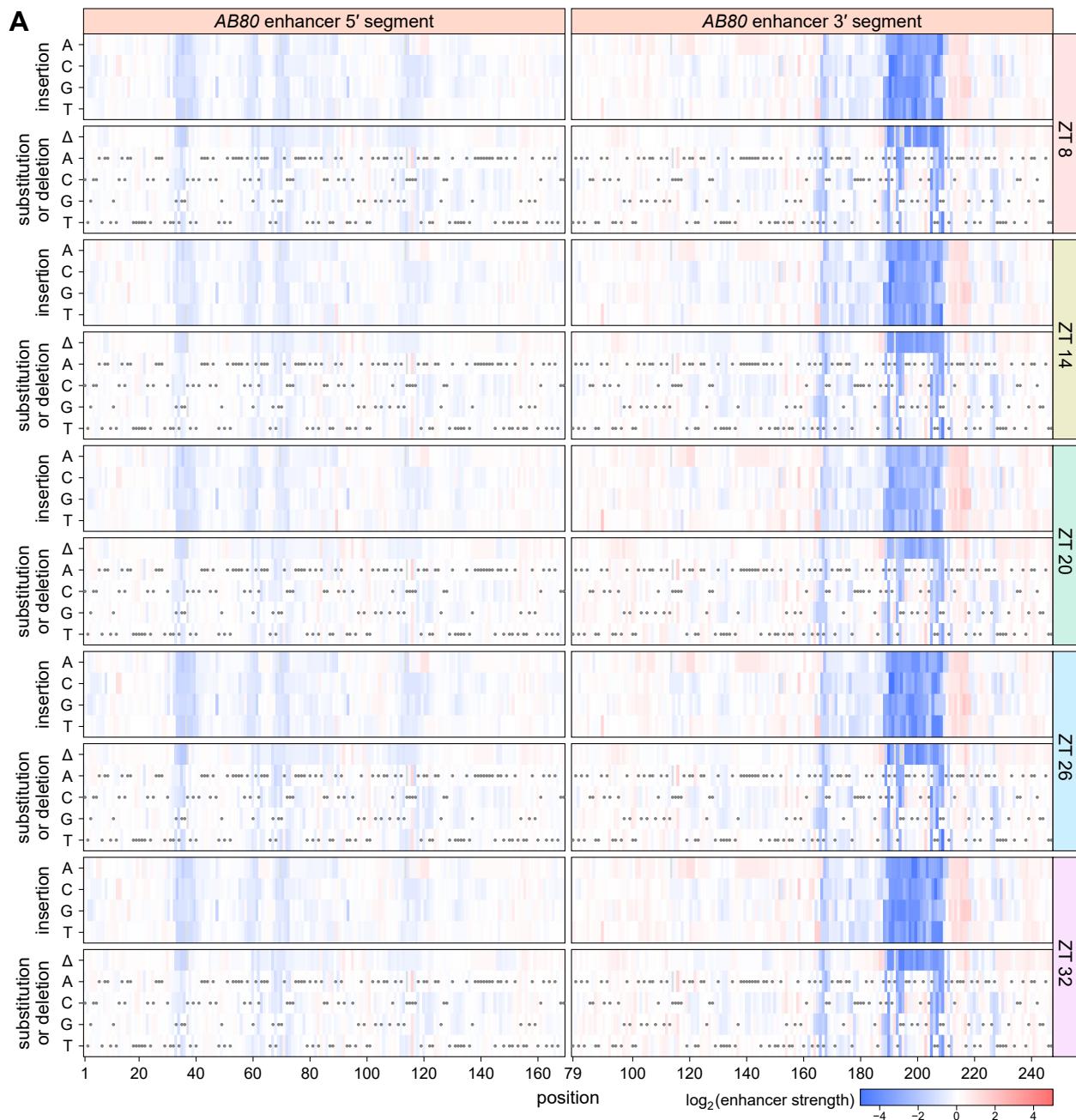


D	ID	family	consensus	ID	family	consensus	
	TF-6	bHLH/BES1/bZIP/Trihelix	CACGTG		TF-36	HD-ZIP/YABBY	ATAATAATw
	TF-8	Dof/C3H	AAAAG		TF-37	WOX/HD-ZIP	CAATCATAws
	TF-9	ZF-HD/HD-ZIP	ATATTdrnnnnnwwwT		TF-38	NF-YB/C2H2/G2-like	TTGAAAAA
	TF-13	MYB related/MYB/GeBP	GGATAA		TF-39	HD-ZIP/YABBY	TAATCATTA
	TF-14	SBP/AP2/C2H2/bHLH	CGTAC		TF-43	E2F/DP	GCGCC
	TF-15	TCP	TGGGCCACn		TF-44	Trihelix	TAACCATGTTTr
	TF-18	C2H2/GRAS	AAAGACAAAAm		TF-49	M-type MADS	TCACCA
	TF-19	GATA	GATC		TF-52	B3	AGAAAnwnnnnAAGAAAn
	TF-22	TCP	GGGACCAC		TF-61	C2H2	rAACACAGAG
	TF-24	GATA/MIKC MADS	CATCATCATCATCATCATC		TF-63	NF-YB	TCCATCA
	TF-25	C2H2	CACT		TF-67	B3	AAAAAAAAAAAA
	TF-27	MYB related/Trihelix	TTAGGGy		TF-68	Nin-like	CAGCA
	TF-28	CPP	TTTAATTTrAww		TF-70	WOX	AnTTAATTA
	TF-33	RAV	TmTGTTG				

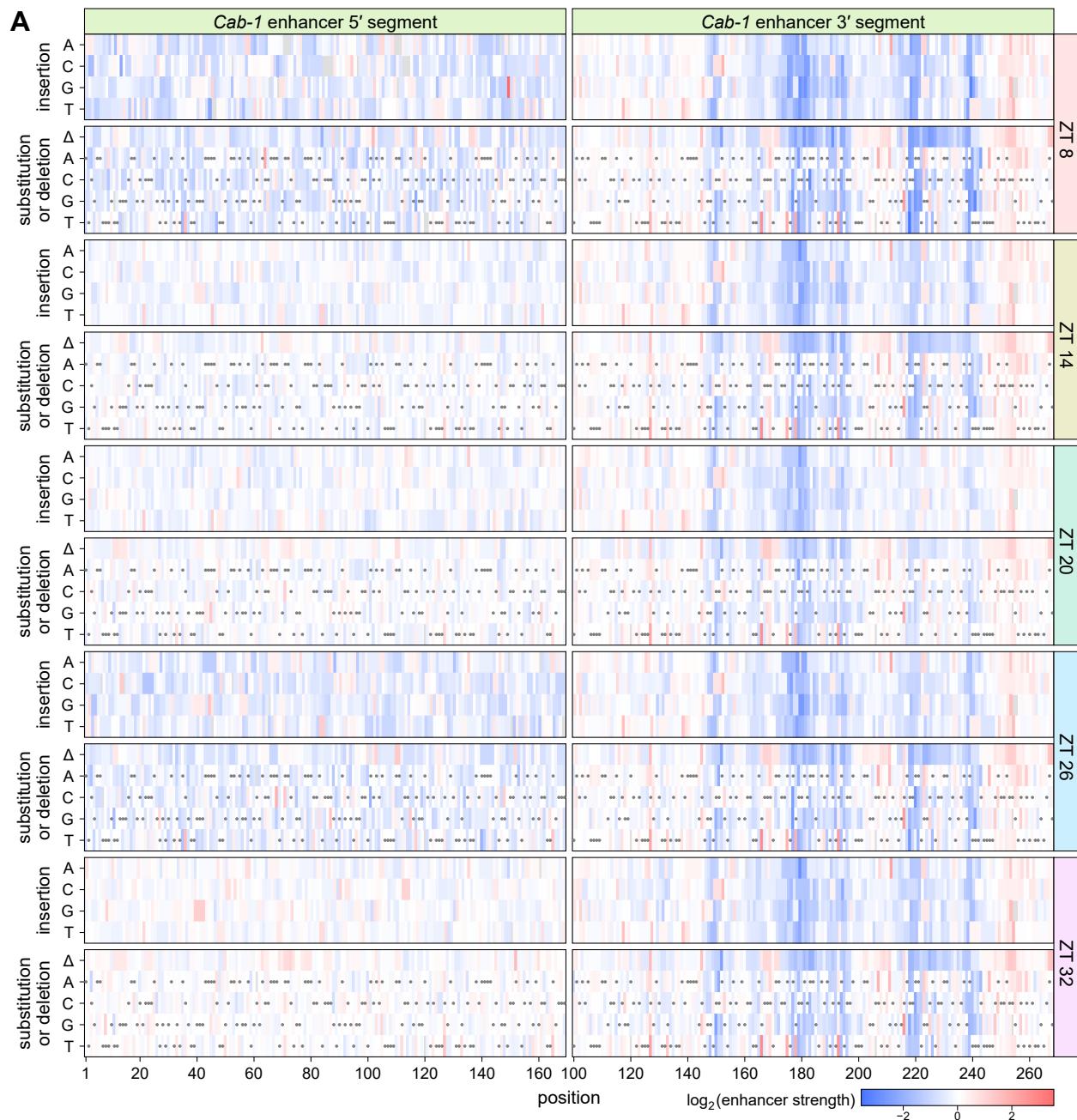
Supplemental Figure S7 Mutation-sensitive regions contain few strong matches to known transcription factor binding motifs. (Supports Figure 3) **A–C**, The wild-type sequences of the 5' and 3' segments of the *AB80* (**A**), *Cab-1* (**B**), and *rbcS-E9* (**C**) enhancers were scanned for significant matches to known transcription factor binding motifs. Hits are shown as gray areas overlaid on the mutational sensitivity plots reproduced from Figure 2. The ID of the matching transcription factor motif is indicated. **D**, For each transcription factor motif ID, the table lists the families of transcription factors that can bind to it and the consensus binding site sequence. Ambiguous nucleotides in the consensus sequence are shown as lowercase and correspond to: n = A, C, G, or T; d = A, G, or T; r = A or G; w = A or T; m = A or C.



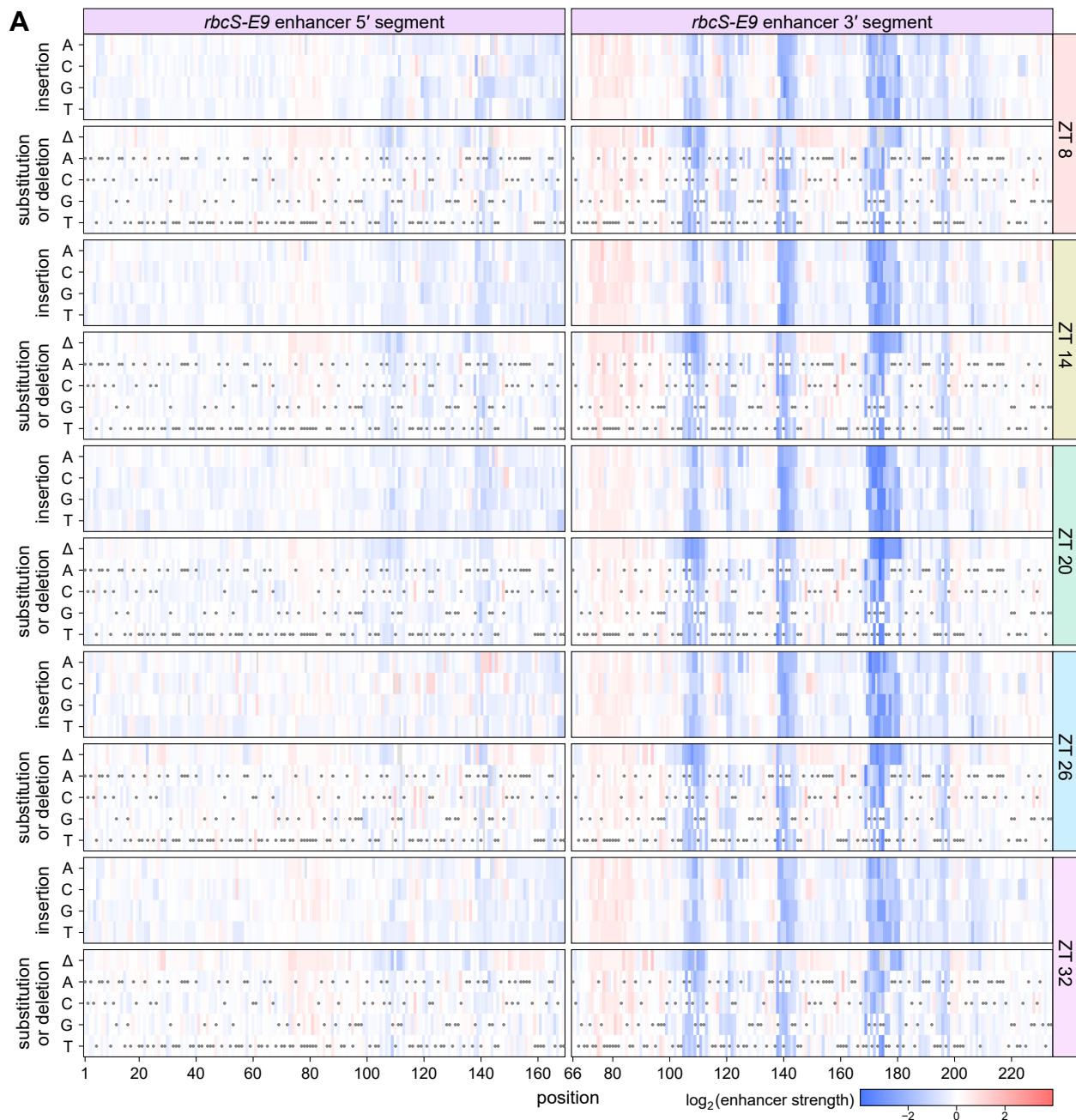
Supplemental Figure S8 Strong correlation between Plant STARR-seq samples obtained 24 hours apart from each other. (Supports Figure 4) **A**, All possible single-nucleotide variants of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers were subjected to Plant STARR-seq in tobacco leaves. On the morning of the third day after transformation (ZT 0), the plants were shifted to constant light. Leaves were harvested for RNA extraction starting at mid-day (ZT 8) and in 6 hour intervals (ZT 14, 20, 26, and 32) afterwards. Hexbin plots (color represents the count of points in each hexagon) of the correlation between samples obtained at the indicated time points are shown. Pearson's R^2 , Spearman's ρ , and number (n) of enhancer variants are indicated. **B**, Hexbin plots of the correlation between the samples from the time course experiment described above compared to the same library tested in normal light/dark cycles (light) as described in Figure 2. The "light" samples were harvested at ZT 8.



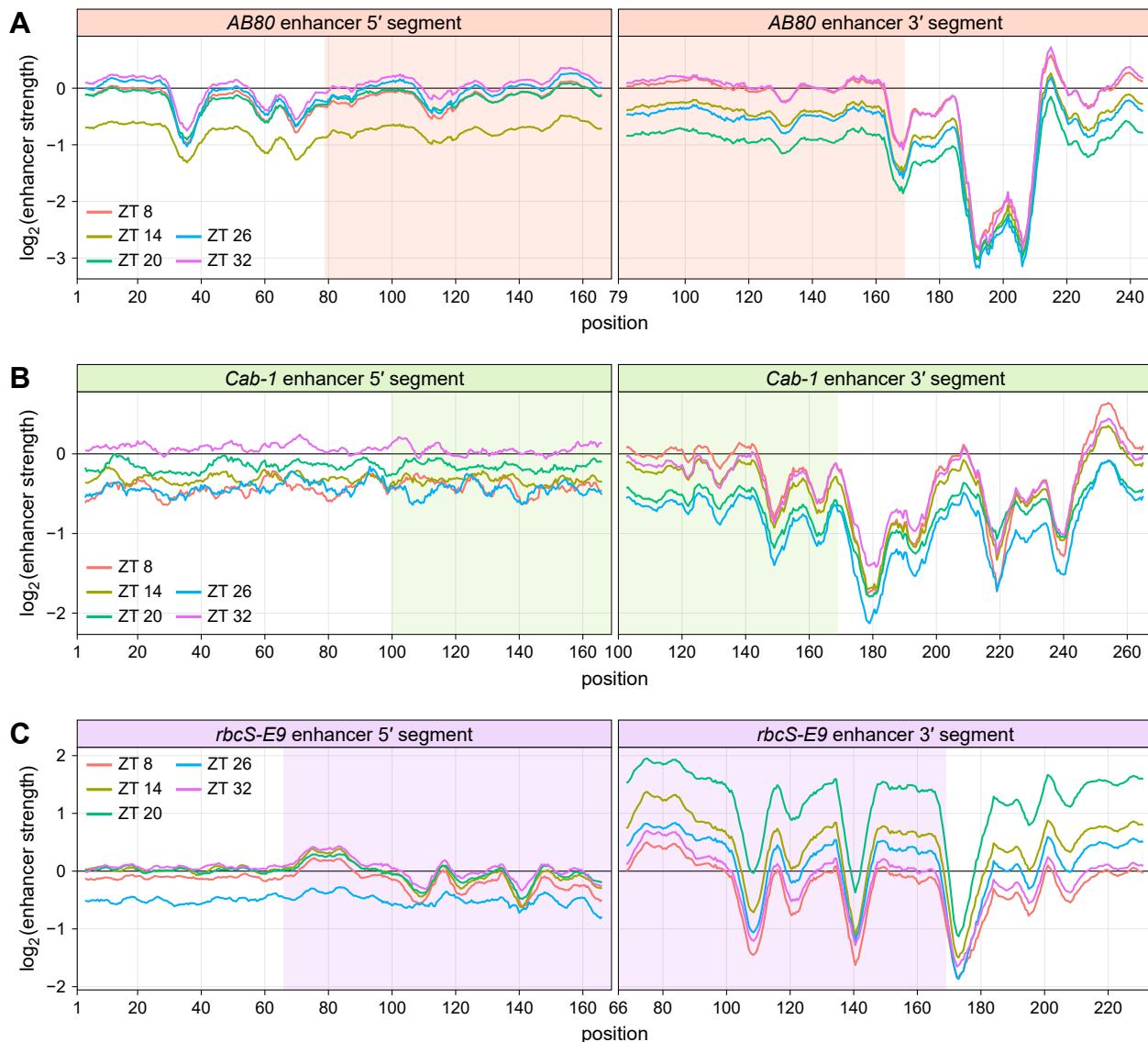
Supplemental Figure S9 Saturation mutagenesis maps of the AB80 enhancer do not change much over time in constant light. (Supports Figure 4) **A**, All possible single-nucleotide variants of the AB80 enhancer were subjected to Plant STARR-seq in tobacco leaves. On the morning of the third day after transformation (ZT 0), the plants were shifted to constant light. Leaves were harvested for RNA extraction starting at mid-day (ZT 8) and in 6 hour intervals (ZT 14, 20, 26, and 32) afterwards. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0) and plotted as a heatmap. Missing values are shown in light gray and wild-type variants are marked with a gray dot.



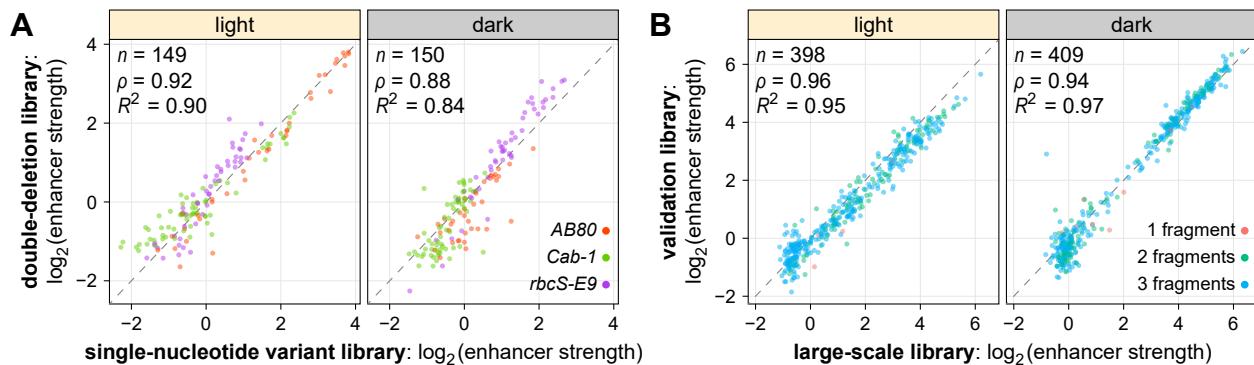
Supplemental Figure S10 Saturation mutagenesis maps of the *Cab-1* enhancer do not change much over time in constant light. (Supports Figure 4) **A**, The *Cab-1* enhancer was subjected to the same experiment as the *AB80* enhancer in Supplemental Figure S9.



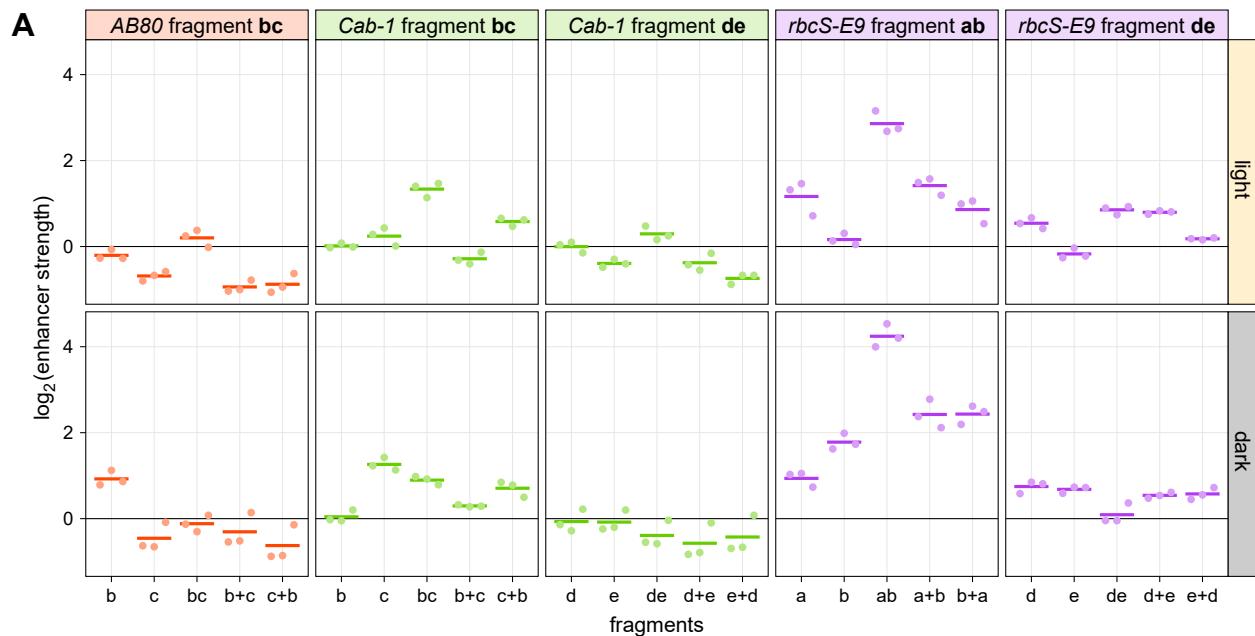
Supplemental Figure S11 Saturation mutagenesis maps of the *rbcS-E9* enhancer do not change much over time in constant light. (Supports Figure 4) **A**, The *rbcS-E9* enhancer was subjected to the same experiment as the *AB80* enhancer in Supplemental Figure S9.



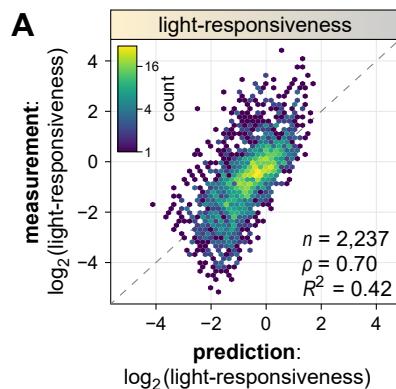
Supplemental Figure S12 Mutation-sensitive regions of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers are conserved over time in constant light. (Supports Figure 4) **A–C**, All possible single-nucleotide substitution, deletion, and insertion variants of the 5' and 3' segments of the *AB80* (**A**), *Cab-1* (**B**), and *rbcS-E9* (**C**) enhancers were subjected to Plant STARR-seq in tobacco leaves. On the morning of the third day after transformation (ZT 0), the plants were shifted to constant light. Leaves were harvested for RNA extraction starting at mid-day (ZT 8) and in 6 hour intervals (ZT 14, 20, 26, and 32) afterwards. Enhancer strength was normalized to the wild-type variant (\log_2 set to 0). A sliding average (window size = 6 bp) of the mean enhancer strength for all variants at a given position is shown. The shaded area indicates the region where the 5' and 3' segments overlap.



Supplemental Figure S13 Plant STARR-seq experiments are reproducible across libraries. (Supports Figures 2, 3 and 5–7) **A**, Correlation between the enhancer strength of single-nucleotide deletion variants of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers present in the comprehensive single-nucleotide enhancer variants library (described in Figure 2) and in a second, independent library with single- and double-deletion enhancer variants (described in Figure 5). **B**, Correlation between the strength of synthetic enhancers created by combining fragments of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers as measured in the large-scale library (described in Figure 6) and in a second, smaller validation library. Pearson's R^2 , Spearman's ρ , and number (n) of enhancer variants are indicated.



Supplemental Figure S14 Correct spacing between mutation-sensitive regions is required for full activity. (Supports Figure 6) **A**, Plots of the strength in the indicated condition of enhancer fragments (Figure 6D) or fragment combinations (separated by a + sign and shown in the order in which they appear in the construct; Figure 6E) in three replicates (points) and the mean strength (lines). Enhancer strength was normalized to a control construct without an enhancer (log₂ set to 0). Some plots are reproduced from Figure 6, F and G.



Supplemental Figure S15 A linear model can predict the light-responsiveness of synthetic enhancers. (Supports Figure 7) **A**, A linear model was built to predict the light-responsiveness of synthetic enhancers created by randomly combining up to three fragments derived from mutation-sensitive regions of the *AB80*, *Cab-1*, and *rbcS-E9* enhancers (see Figure 6A) based on the light-responsiveness of the constituent individual fragments. A hexbin plot (color represents the count of points in each hexagon) of the correlation between the model's prediction and the measured data is shown. Pearson's R^2 , Spearman's ρ , and number (n) of enhancer fragment combinations are indicated.