Supplemental material

Table S1. ChIP-seq scores for each CArG-box region. The table shows location of the first base on the forward strand sequence of the Arabidopsis genome of all perfect CArG-box (CCW_6GG) sequences, their particular 10bp sequence, the length of the A-tract element (m+n), and the ChIP-seq (or -chip) score associated to that region.

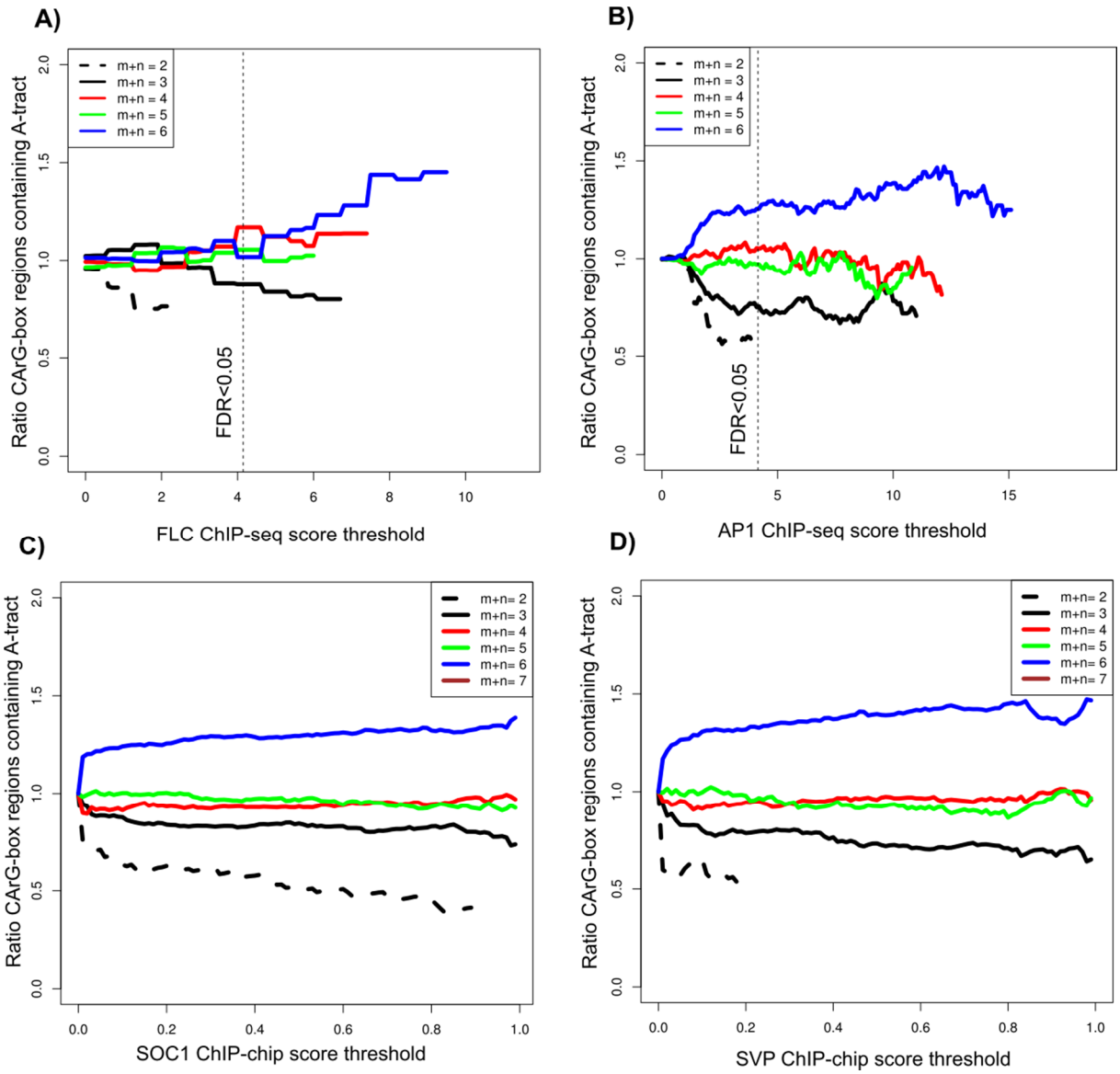

Fig. S1. Enrichment of A-tract elements in CArG-box sequences (CCW_6GG) bound by various MADS-domain proteins. The proportion of CArG-box motifs with a particular A-tract element inside (length 4-6, red, green, and blue color respectively) or no A-tract element (length<4; black color) normalized by the proportion of CArG-boxes with each particular element at genome-wide level and plotted against the ChIP-seq score threshold used. Values are only plotted until a ChIP-seq score where there are at least 15 CArG-boxes to calculate the ratio. This was done for: A) FLC ChIP-seq, B) AP1 ChIP-seq, C) SOC1 ChIP-chip, D) SVP ChIP-chip.

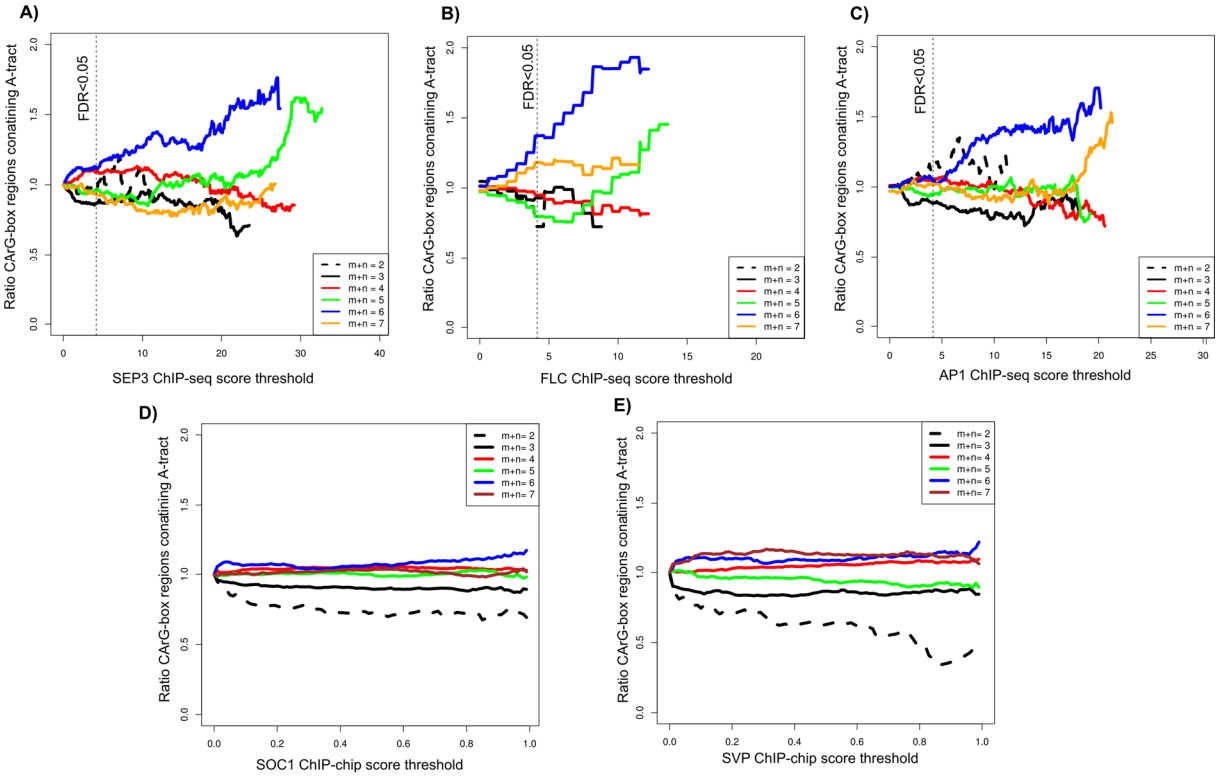

Fig. S2. Enrichment of A-tract elements in CArG-box sequences (CCW_7G) bound by various MADS-domain proteins. The proportion of CArG-box motifs with a particular A-tract element inside (length 4-6, red, green, and blue color respectively) or no A-tract element (length<4; black color) normalized by the proportion of CArG-boxes with each particular element at genome-wide level and plotted against the ChIP-seq score threshold used. Values are only plotted until a ChIP-seq score where there are at least 15 CArG-boxes to calculate the ratio. A) SEP3 ChIP-seq, B) FLC ChIP-seq, C) AP1 ChIP-seq, D) SOC1 ChIP-chip, E) SVP ChIP-chip.

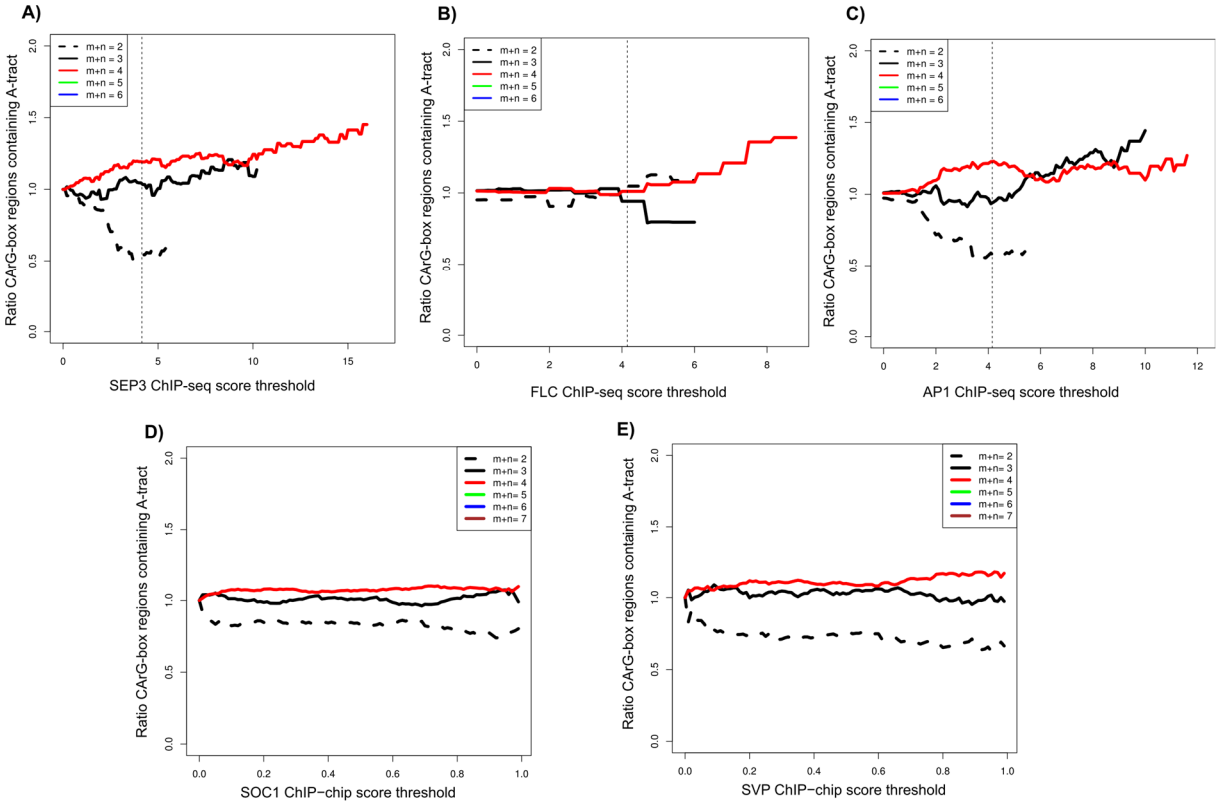

Fig. S3. Enrichment of A-tract elements in CArG-box sequences (CCW_4SSGG) bound by various MADS-domain proteins. The proportion of CArG-box motifs with a particular A-tract element inside (length 4-6, red, green, and blue color respectively) or no A-tract element (length<4; black color) normalized by the proportion of CArG-boxes with each particular element at genome-wide level and plotted against the ChIP-seq score threshold used. Values are only plotted until a ChIP-seq score where there are at least 15 CArG-boxes to calculate the ratio. This was done for: A) SEP3 ChIP-seq, B) FLC ChIP-seq, C) AP1 ChIP-seq, D) SOC1 ChIP-chip, E) SVP ChIP-chip.

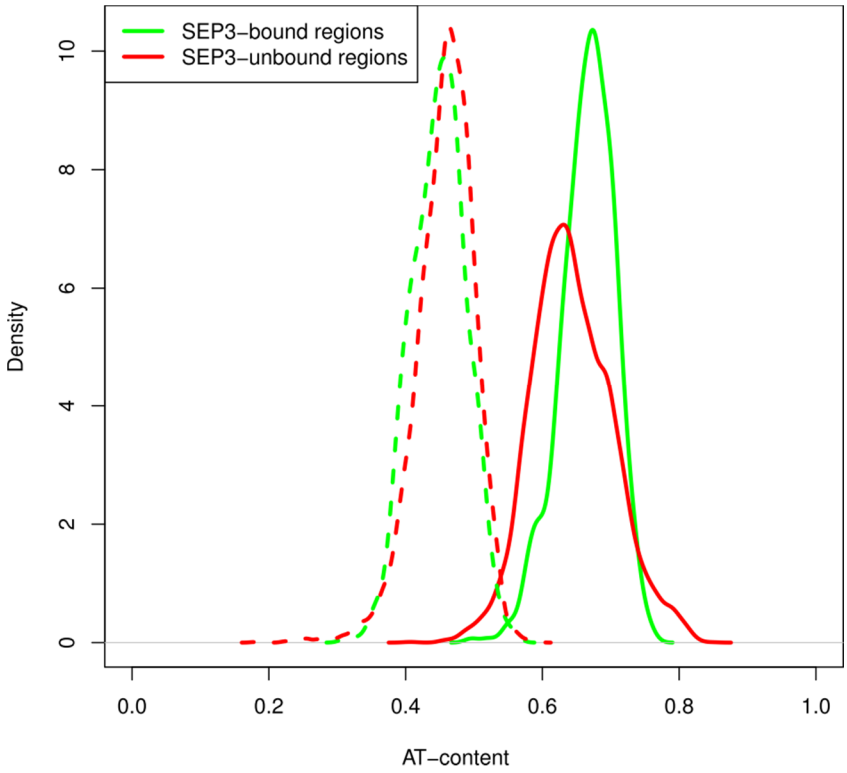

Fig. S4. Distribution of AT content in CArG-box regions bound and unbound by SEP3. The different number of A-tract elements (m+n>3) in regions bound and not bound by SEP3 is not due to a different AT content of these regions, since when the A-tracts sequences are eliminated, both sets of regions have the same AT-content (dashed line), only when the A-tracts elements are considered the set of regions have a different AT-content distribution (continuous line)

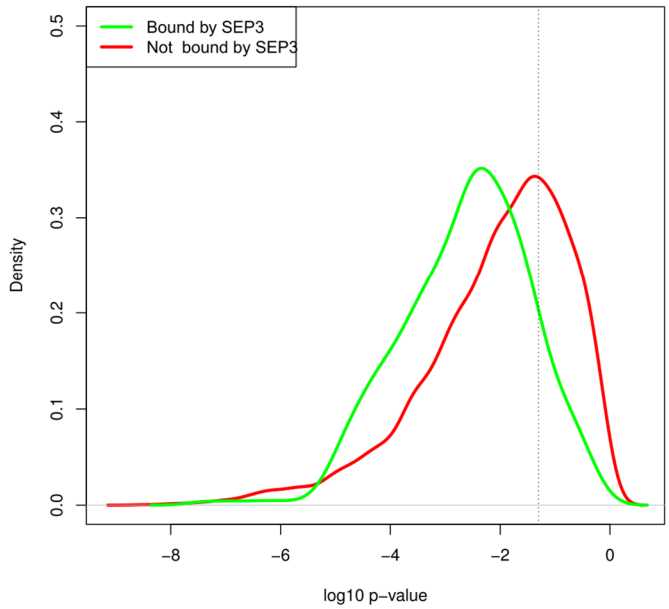

Fig. S5. Periodicity of A-tract elements around the CArG-box motif. Distribution of log p-values for testing periodicity using Fisher's g-test for each sequence (510bp) in the group of regions bound by SEP3 (green), and not bound (red). P-values lower

than 0.05 (indicated with a dash vertical line) indicate a statistically significant periodicity for the A-tact location. Regions bound by SEP3 show a distribution with more significant p-values.

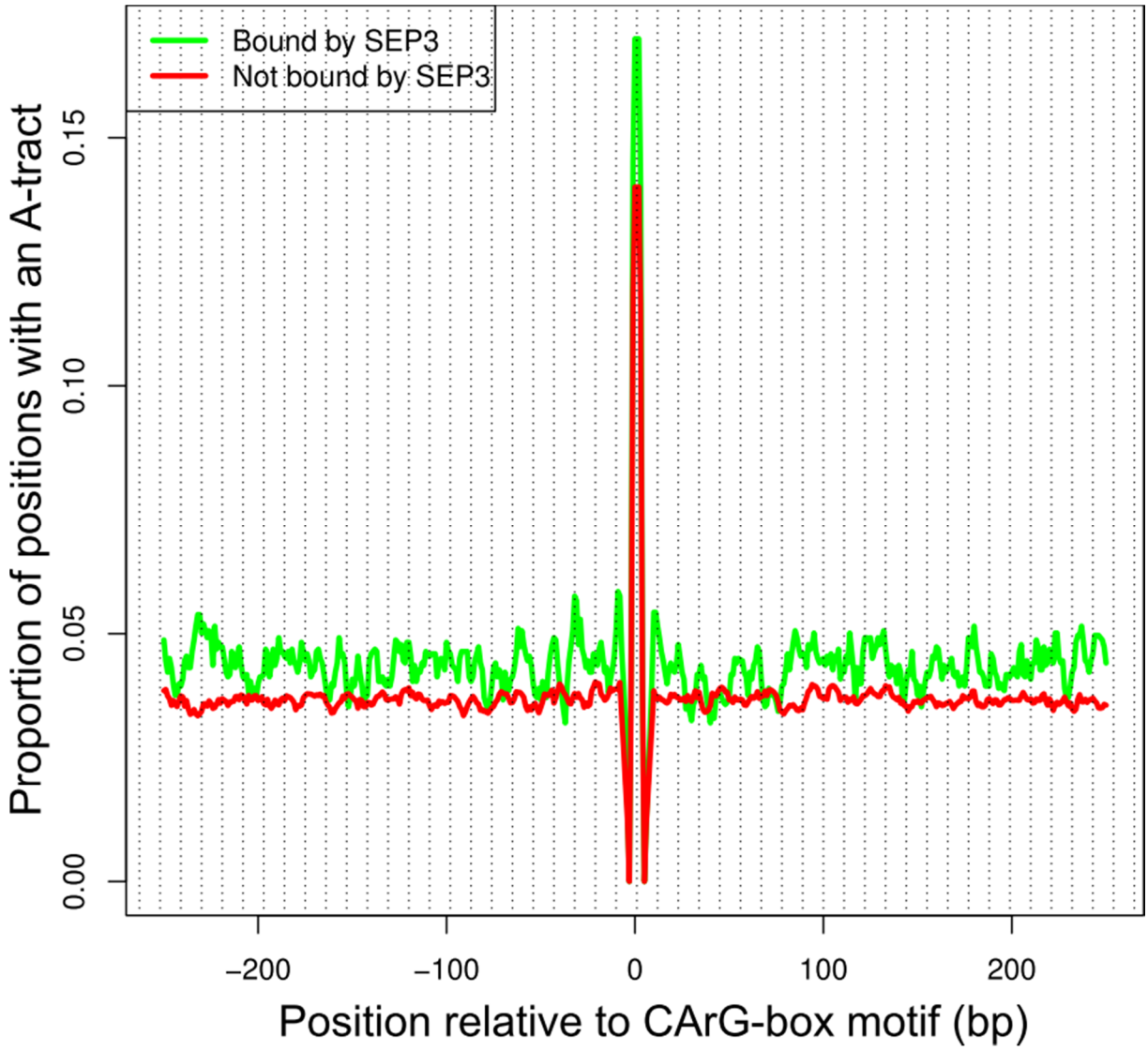

Fig. S6. Multiple A-tracts in SEP3-bound CArG-box regions. Proportion of CArG-box regions with an A-tract element in a particular position. A moving average of length 5 bp was applied to obtain a more smooth representation of the data. Regions with a SEP3 ChIP-seq binding event (FDR<0.05) are indicated in green, and regions without a binding event are indicated in red. Dashed lines are located each 11 bp from the middle of the CArG-box motif, which represents a helical turn.

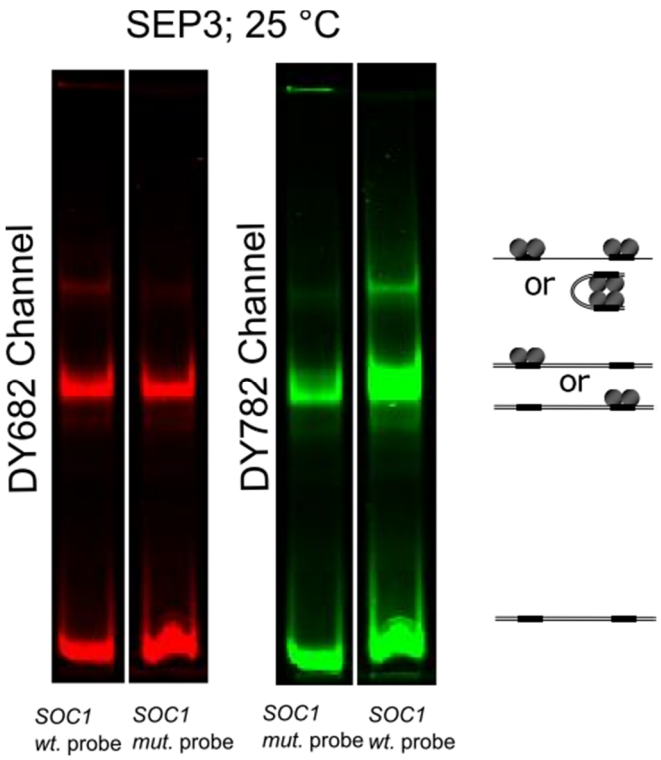

Fig S7. Gel pictures for QuMFRA interrogating a probe representing *SOC1* promoter. A gel with two lanes is represented under the DY782 and DY682 channel. *In vitro* synthesized SEP3 protein was incubated at 25° C with two different probes representing the *SOC1* promoter labelled with two dyes (see Mat. & Met.), and therefore the probes are competing with each other to bind SEP3. In the first lane the *SOC1 wt.* probe was labelled with DY682 (red) and *SOC1 mut*. with DY782 (green), in the second lane the dye and probe were swap.

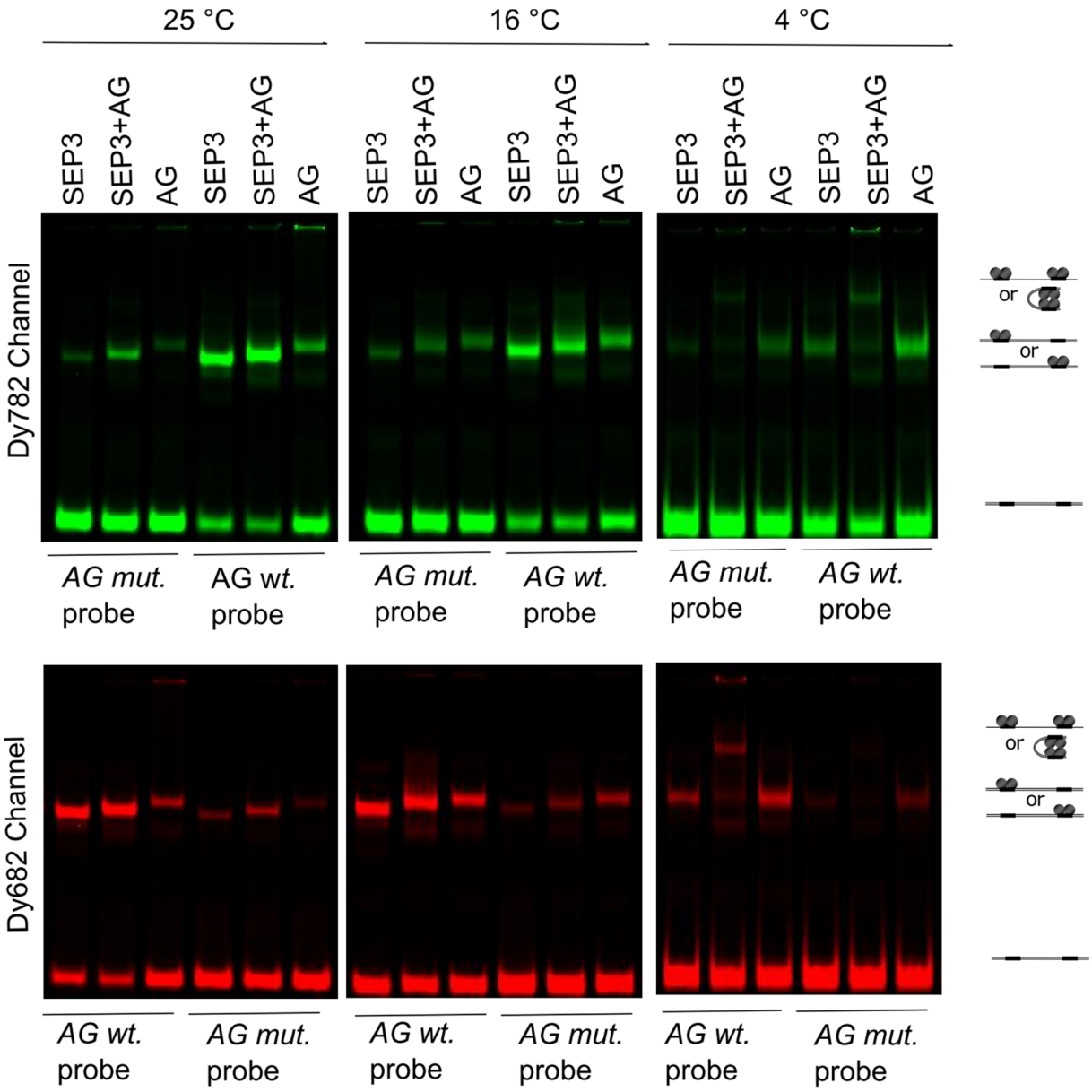

Fig S8 Gel pictures for QuMFRA interrogating a probe representing *AG* promoter at different temperatures. Three gels with six lanes each are represented under the DY782 and DY682 channel. *In vitro* synthesized SEP3, SEP3+AG, or AG protein mix was incubated with two probes representing the *AG* promoter and labelled with two different dyes (see Mat. & Met.). Therefore the probes are competing with each other to bind to the protein complexes. For each lane the top picture represents the DY782 channel and below it is represented the same lane but with the Dy682 channel. Each gel contains the two possible combinations between the two probes and the two dyes.