GABPα and CREB1 binding to double nucleotide polymorphisms of their consensus motifs and cooperative binding to the composite ETS⇔CRE motif (ACCGGAAGTGACGTCA)

Nima Assad[1], Desiree Tillo[1], Sreejana Ray[1], Alexa Dzienny[1], Peter C. FitzGerald[2], Charles Vinson[1,*]

[1]Laboratory of Metabolism, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.

[2]Genome Analysis Unit, Genetics Branch, National Cancer Institute, National Institutes of Health, Building 37, Bethesda, MD 20892.

*To whom correspondence should be addressed

Tel: 1-301-496-8783, Fax: 1-301-496-8419, E-mail: vinsonc@mail.nih.gov

Supporting Figure Captions

Category	Solvent - Variable 36-mer | Constant 24-mer - Glass	GABPa Intensity	CREB1 Intensity	Cooperativity
Shorten Constant Region	TCCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	4,300	50,000	1.7
	CCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	4,700	48,000	1.6
	.	.	.	.
	ACCGGAAGTGACGTCAGACTCAGGTG|GGACACACTTTAACACATGGAGAG	1,700	56,000	3.0
	ACCGGAAGTGACGTCAACTCAGGTG|GGACACACTTTAACACATGGAGAG	1,600	53,000	3.2
	ACCGGAAGTGACGTCACTCAGGTG|GGACACACTTTAACACATGGAGAG	1,500	63,000	3.1
	.	.	.	.
	ACCGGAAGTGACGTCA|GGACACACTTTAACACATGGAGAG	2,400	45,000	2.4
Slide Consensus Motif	**ACCGGAAGTGACGTCA**GTCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG	1,000	49,000	3.0
	G**ACCGGAAGTGACGTCA**TCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG	3,000	57,000	1.9
	.	.	.	.
	GTCCTCAAGAGACTCAGGTG**ACCGGAAGTGACGTCA**|GGACACACTTTAACACATGGAGAG	3,600	59,000	1.6
Delete Consensus Motif	GTCCTCAAGA**CCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	4,500	50,000	1.6
	GTCCTCAAGAC**GGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	700	44,000	3.6
	.	.	.	.
	GTCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG	400	1,400	0.6
TRE	GTCCTCAAGA**ACCGGAAGTGAGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	5,400	3,400	0.7
	GTCCTCAAGA**ACCGGAAGTGACTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	5,200	3,200	0.7
	GTCCTCAAGA**TGACTCACTTCCGGT**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	4,400	9,700	1.0
	GTCCTCAAGA**TGAGTCACTTCCGGT**GACTCAGGTG|GGACACACTTTAACACATGGAGAG	5,200	8,700	0.8

Table S1. ETS⇔CRE DNP Microarray Control Probes. The control microarray DNA probes varying the probe length, the composite motif position, sequentially removing nucleotides from the composite motif, and replacing the CRE with the TRE consensus sequence. The 117 control features examine several parameters: 1) probe length, 2) motif position, 3) deletion of the motif, and 4) binding to the ETS⇔TRE composite motif (**ACCGGAA**$\underline{\textbf{GTGA}}$$^{\textbf{C}}$/$_{\textbf{G}}$**TCA**) which is a one pair deletion of the ETS⇔CRE. 40 control probes sequentially shorten the 60-mer to a 40-mer, maintaining an intact ETS⇔CRE or CRE⇔ETS 16-mer and a constant 24-mer primer. 42 control probes systematically slide the ETS⇔CRE or CRE⇔ETS 16-mer along the variable 36-mer portion of the 60-mer DNA. 31 control probes sequentially remove nucleotides from the ETS⇔CRE or CRE⇔ETS motif. A final 4 control probes replace the CRE (**GTGACGTCA**) motif with the TRE motif (**GTGA**$^{\textbf{C}}$/$_{\textbf{G}}$**TCA**) to produce the composite ETS⇔TRE (**ACCGGAGTGA**$^{\textbf{C}}$/$_{\textbf{G}}$**TCA**) motif in both orientations. Binding of GABPα-GST, CREB1-GST, and cooperative binding is shown. When the ETS⇔CRE motif was deleted, binding decreased 15-fold (**Figure S4**). When the CRE is replaced with the TRE motif, GABPα binding intensity remains at 5,000. Binding of CREB1-GST is reduced 12-15 fold. The ETS⇔TRE motif diminishes cooperative binding between GABPα and CREB1.

References

[1] X. He, K.S. Syed, D. Tillo, I. Mann, M.T. Weirauch, C. Vinson, GABPαlpha Binding to Overlapping ETS and CRE DNA Motifs Is Enhanced by CREB1: Custom DNA Microarrays, G3 (Bethesda), 5 (2015) 1909-1918.
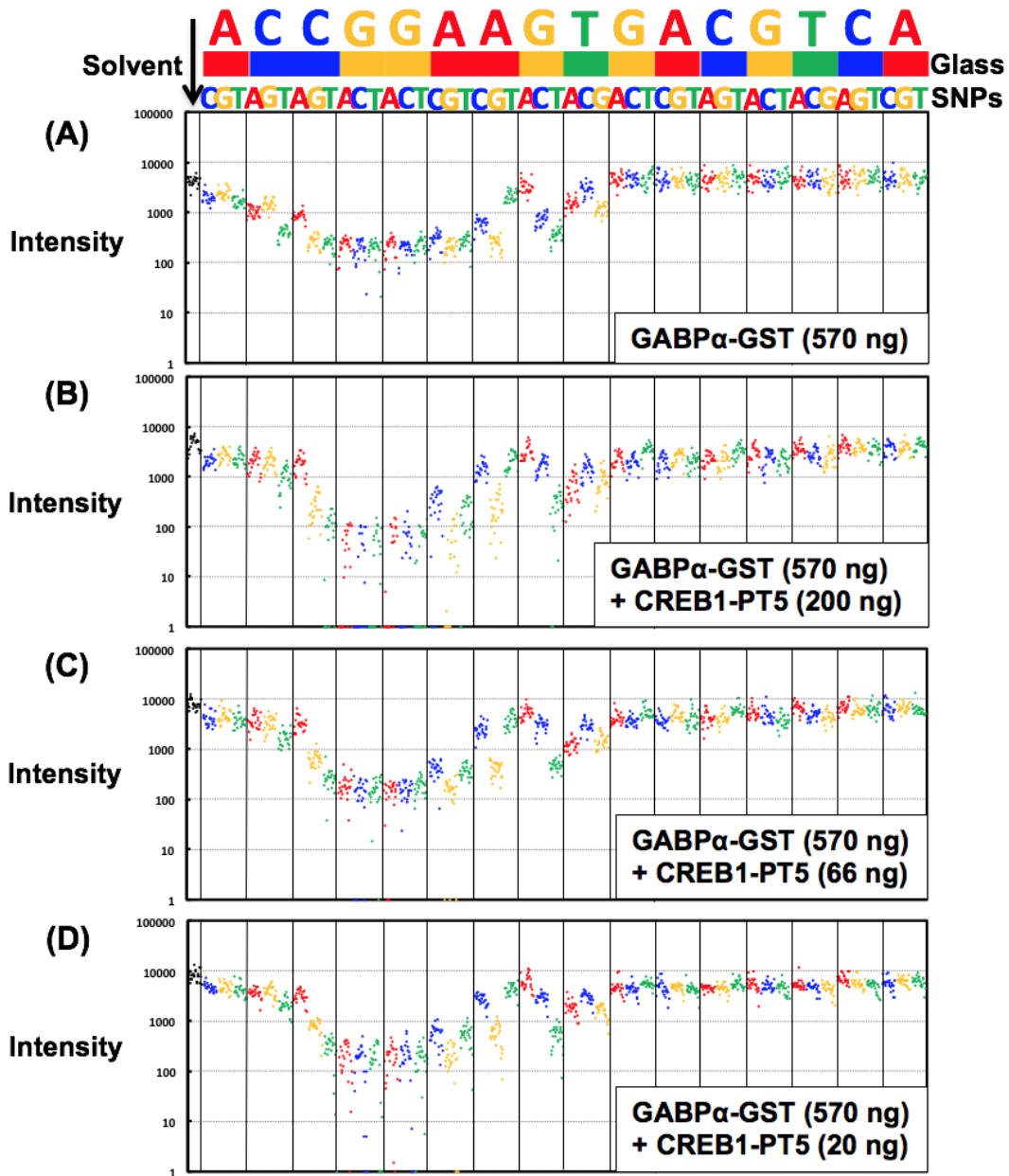
Fig S1. GABPα-GST binding to the ETS⇔CRE motif and SNPs with varying amounts of CREB1. GABPα-GST binding intensity to ETS-CRE motif and SNPs in a background of **(A)** 0 ng, **(B)** 200 ng, **(C)** 66 ng, **(D)** 20 ng CREB1-PT5
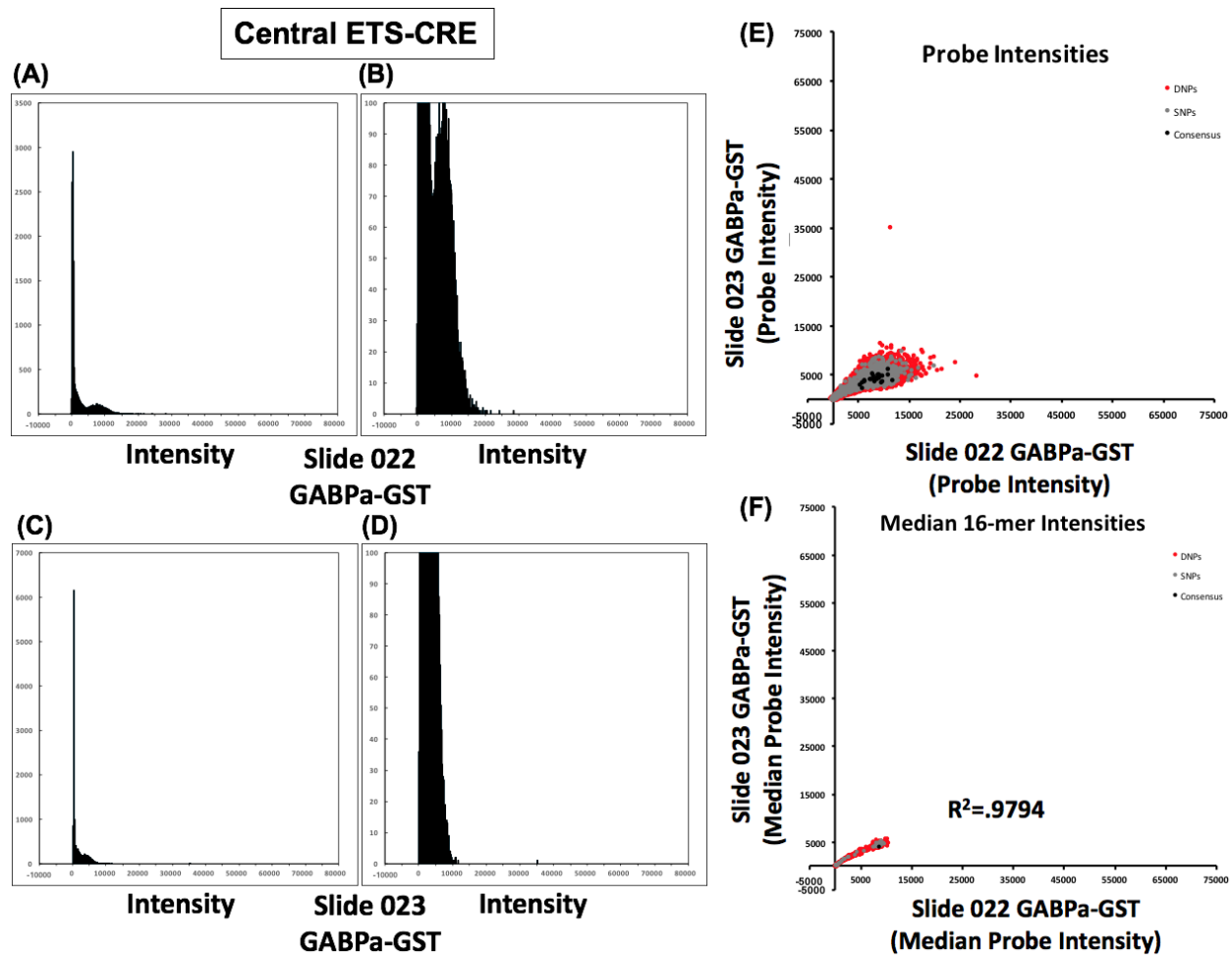
Fig S2. Duplicate GABPα-GST PBM experiments. **(A)** Histogram of raw intensity values for 570ng GABPα-GST binding to central ETS-CRE probes, including the consensus ETS-CRE, SNPs, and DNPs. **(B)** Same as in (A) with a zoomed-in y-axis to emphasize outliers. **(C-D)** Same as (A-B) for duplicate data. **(E)** Scatterplot comparing all probes from the above two duplicate datasets. **(F)** Scatterplot comparing median 16-mer intensities for each of the 24 occurrences of each probe from the above two duplicate datasets.
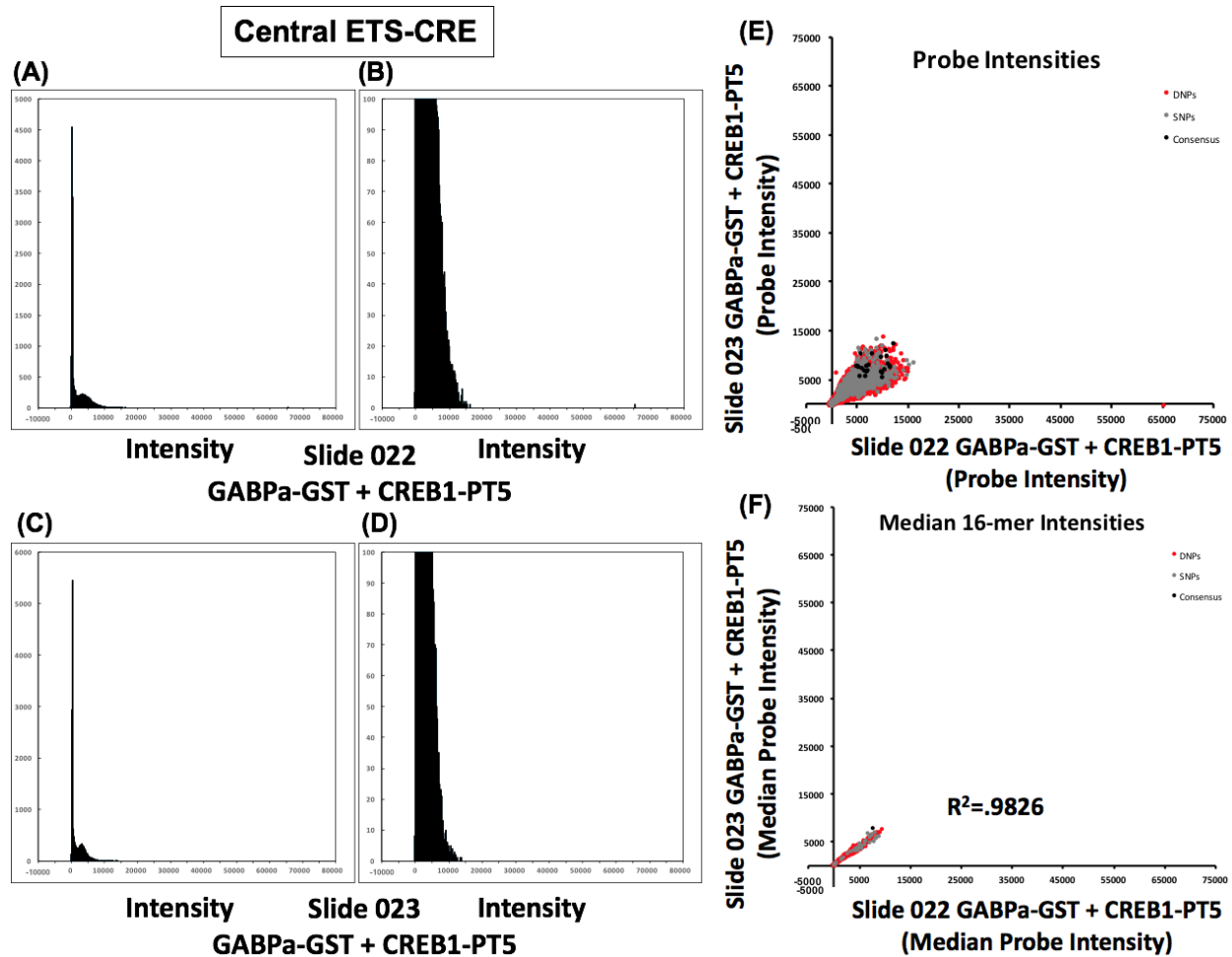
Fig S3. Duplicate GABPα-GST and CREB1-PT5 PBM experiments. **(A)** Histogram of raw intensity values for 570ng GABPα-GST plus 66ng CREB1-PT5 binding to central ETS-CRE probes, including the consensus ETS-CRE, SNPs, and DNPs. **(B)** Same as in (A) with a zoomed-in y-axis to emphasize outliers. **(C-D)** Same as (A-B) for duplicate data. **(E)** Scatterplot comparing all probes from the above two duplicate datasets. **(F)** Scatterplot comparing median 16-mer intensities for each of the 24 occurrences of each probe from the above two duplicate datasets.
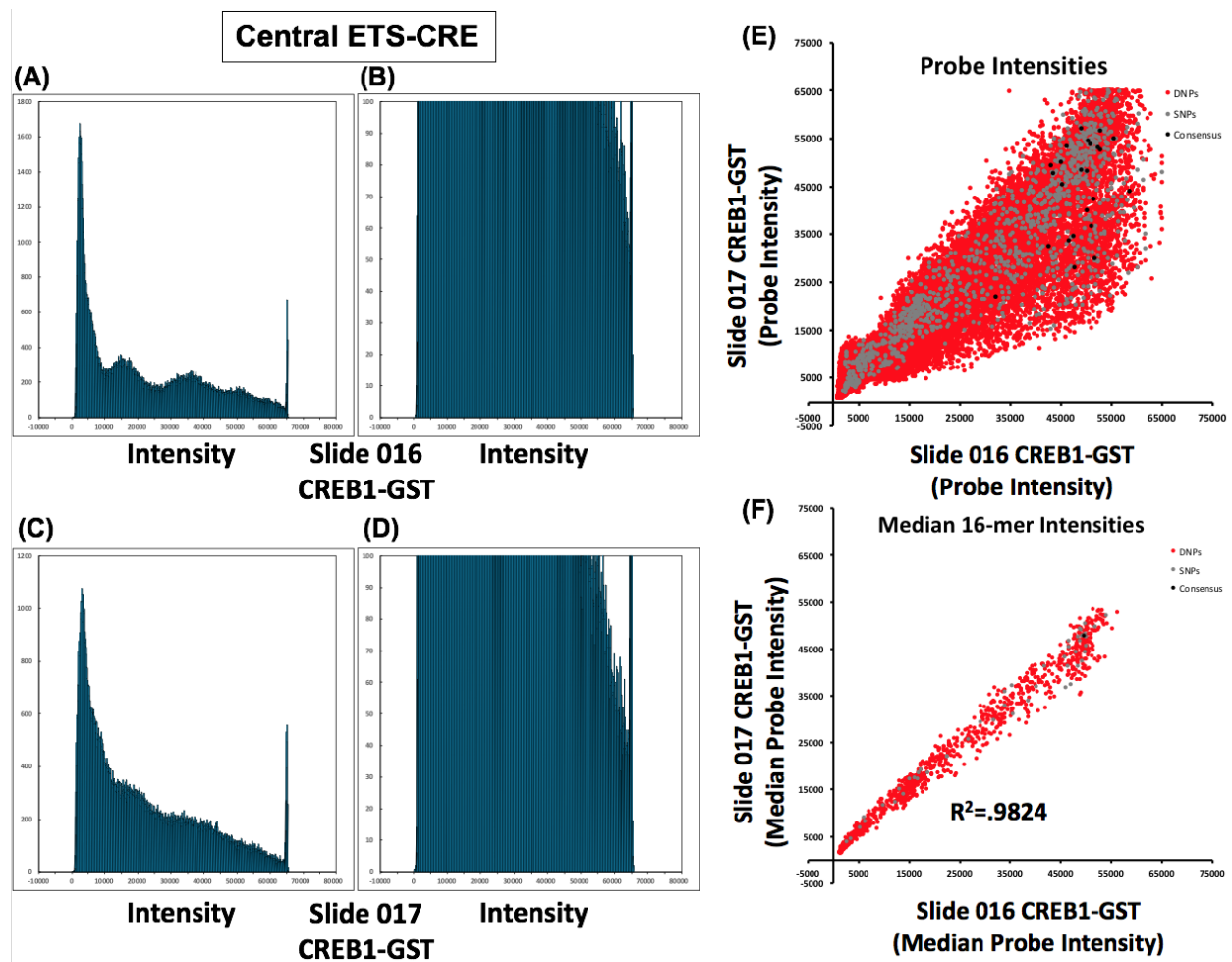
Fig S4. Duplicate CREB1-GST PBM experiments. **(A)** Histogram of raw intensity values for 570ng CREB1-GST binding to central ETS-CRE probes, including the consensus ETS-CRE, SNPs, and DNPs. **(B)** Same as in (A) with a zoomed-in y-axis to emphasize outliers. **(C-D)** Same as (A-B) for duplicate data. **(E)** Scatterplot comparing all probes from the above two duplicate datasets. **(F)** Scatterplot comparing median 16-mer intensities for each of the 24 occurrences of each probe from the above two duplicate datasets.
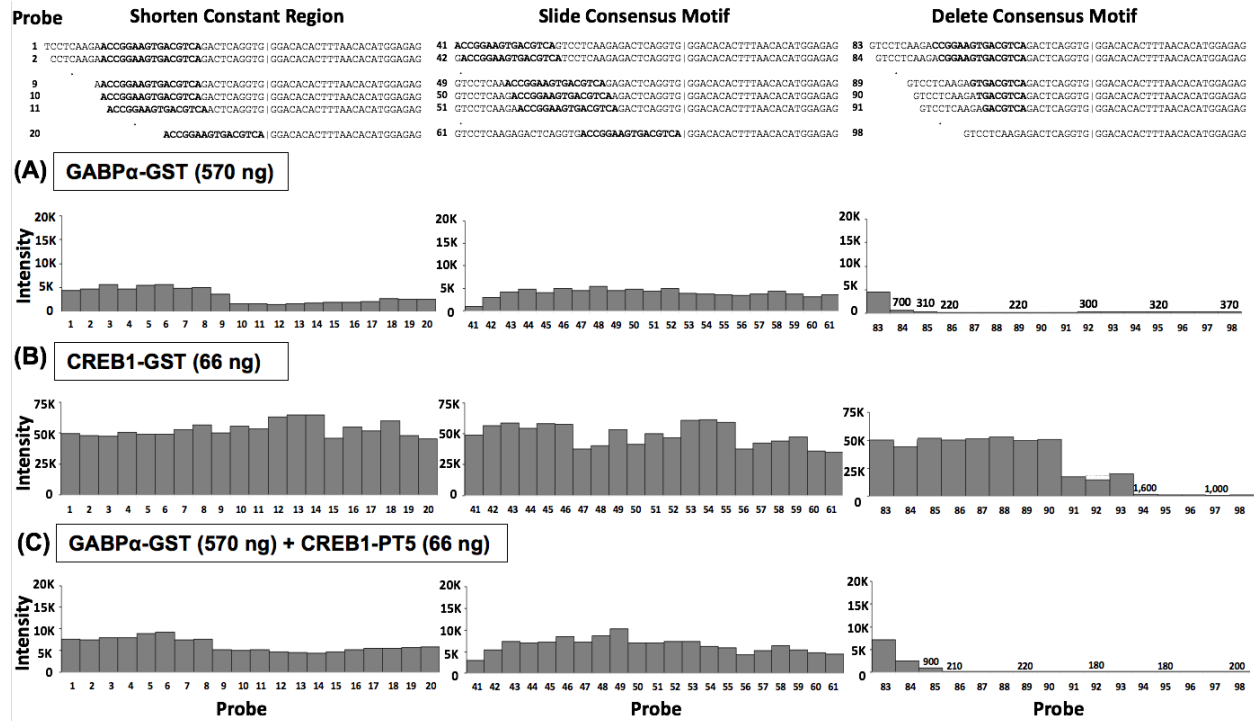
Probe | **Shorten Constant Region** | **Slide Consensus Motif** | **Delete Consensus Motif**

1 TCCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG   41 **ACCGGAAGTGACGTCA**GTCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG   83 GTCCTCAAGA**CCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG

2 CCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG   42 GA**CCGGAAGTGACGTCA**TCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG   84 GTCCTCAAGA**CGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG

9   AACCGGAAGTGACGTCAGACTCAGGTG|GGACACACTTTAACACATGGAGAG   49 GTCCTCAA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG   89 GTCCTCAAGA**GTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG

10   ACCGGAAGTGACGTCAGACTCAGGTG|GGACACACTTTAACACATGGAGAG   50 GTCCTCAAG**ACCGGAAGTGACGTCAA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG   90 GTCCTCAAGA**TGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG

11   ACCGGAAGTGACGTCAACTCAGGTG|GGACACACTTTAACACATGGAGAG   51 GTCCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG   91 GTCCTCAAGA**GACGTCA**GACTCAGGTG|GGACACACTTTAACACATGGAGAG

20   ACCGGAAGTGACGTCA|GGACACACTTTAACACATGGAGAG   61 GTCCTCAAGAGACTCAGGTG**ACCGGAAGTGACGTCA**|GGACACACTTTAACACATGGAGAG   98 GTCCTCAAGAGACTCAGGTG|GGACACACTTTAACACATGGAGAG

(A) GABPα-GST (570 ng)

(B) CREB1-GST (66 ng)

(C) GABPα-GST (570 ng) + CREB1-PT5 (66 ng)

Fig S5. Binding to control probes. Histograms of binding intensity for **(A)** GABPa-GST, **(B)** CREB1-GST, and **(C)** GABPa-GST and CREB1-PT5 binding probes varying the probe length (top), the composite ETS⇔CRE motif position (middle), and deleting the ETS⇔CRE motif (bottom).
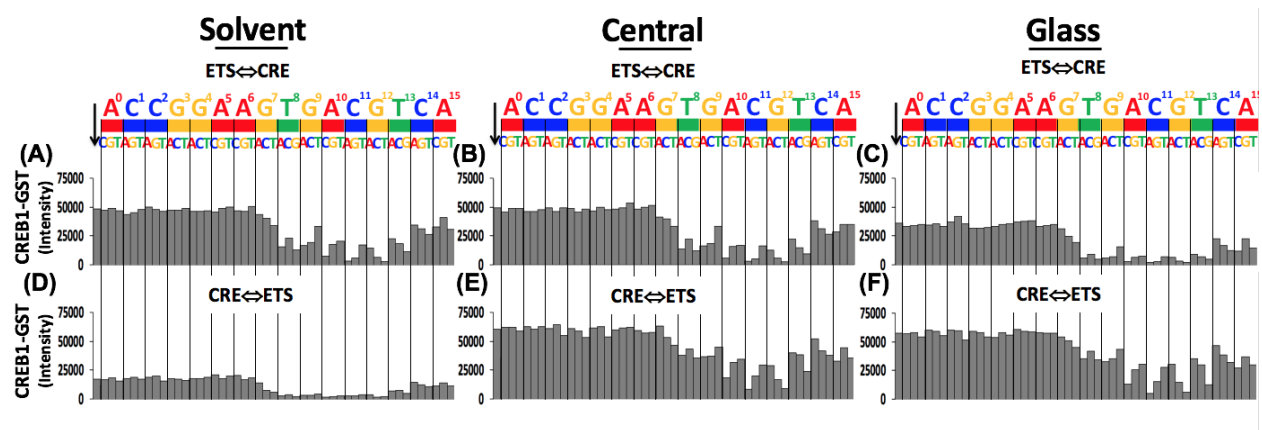
Fig S6. CREB1-GST binding ETS⇔CRE SNPs at 6 positions. Histograms of CREB1-GST binding to the ETS⇔CRE motif and SNPs at the **(A)** solvent, **(B)** central, and **(C)** glass positions on the ETS⇔CRE DNP array. Intensity values for binding to the canonical ETS⇔CRE motif are indicated with an arrow at the far left of each histogram and are shown with purple dashed lines. **(D-F)** Same as (A-C) for the CRE⇔ETS orientation of the motif.
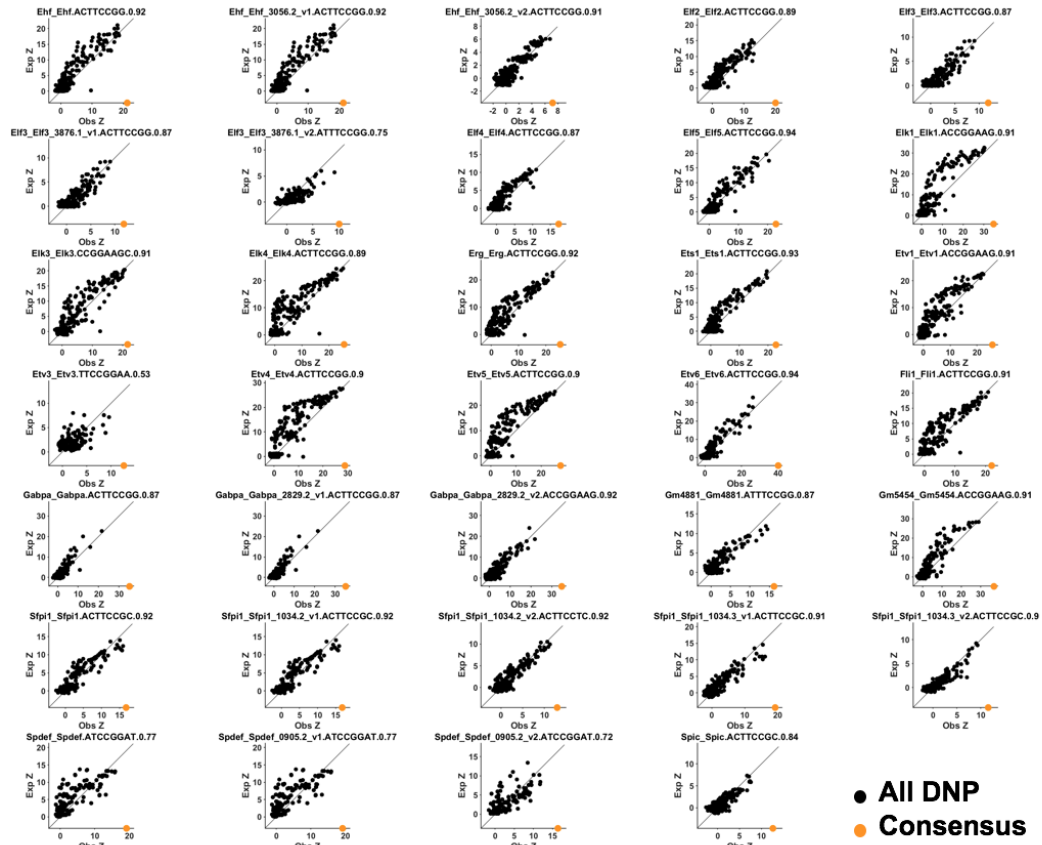
Fig S7. Observed vs. expected DNP Z-score for 34 mouse Ets domain proteins. Z-score of the consensus is shown as an orange spot.
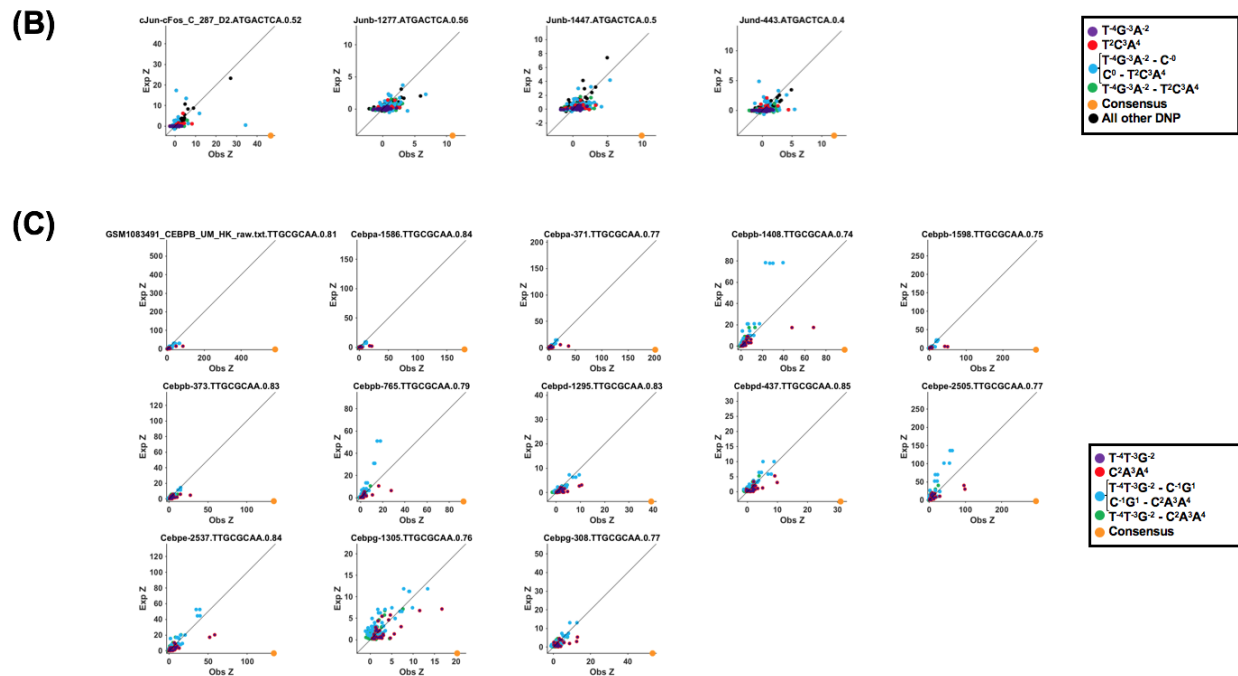
(A)
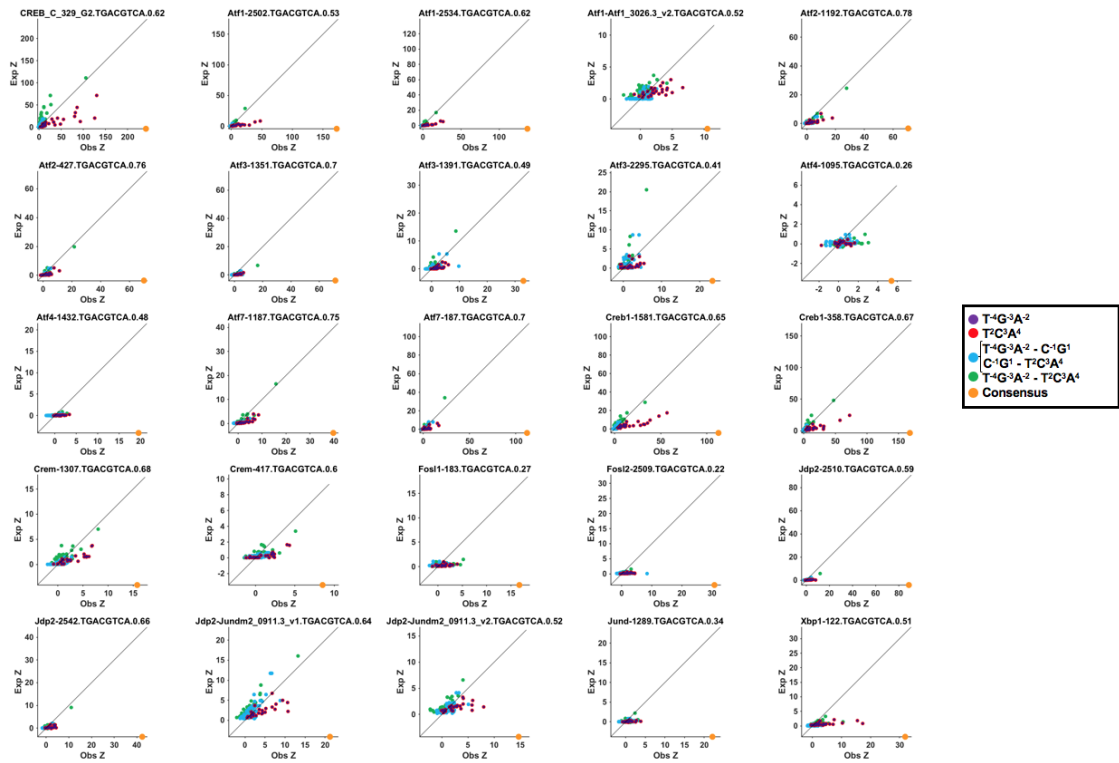
Fig S8A. Observed vs. expected DNP Z-score for mouse bZIP domain proteins. DNPs at different locations of the consensus are indicated. The Z-score of the consensus is shown as an orange spot. Plots are separated by consensus motif: **(A)** CRE ($T^{-4}G^{-3}A^{-2}$ $C^{-1}G^{1}T^{2}C^{3}A^{4}$)
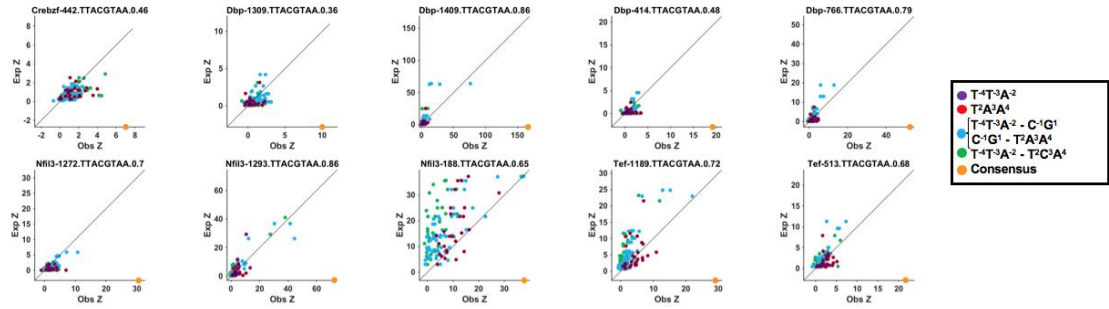
Fig S8B-C. Observed vs. expected DNP Z-score for mouse bZIP domain proteins. DNPs at different locations of the consensus are indicated. The Z-score of the consensus is shown as an orange spot. Plots are separated by consensus motif: **(B)** TRE ($T^{-4}G^{-3}A^{-2}C^0T^2C^3A^4$), **(C)** C/EBP ($T^{-4}T^{-3}G^{-2}C^{-1}G^1C^2A^3A^4$)

(D)

Fig S8D. Observed vs. expected DNP Z-score for mouse bZIP domain proteins. DNPs at different locations of the consensus are indicated. The Z-score of the consensus is shown as an orange spot. Plots are separated by consensus motif: **(D)** PAR ($T^{-4}T^{-3}A^{-2}C^{-1}G^1T^2A^3A^4$)
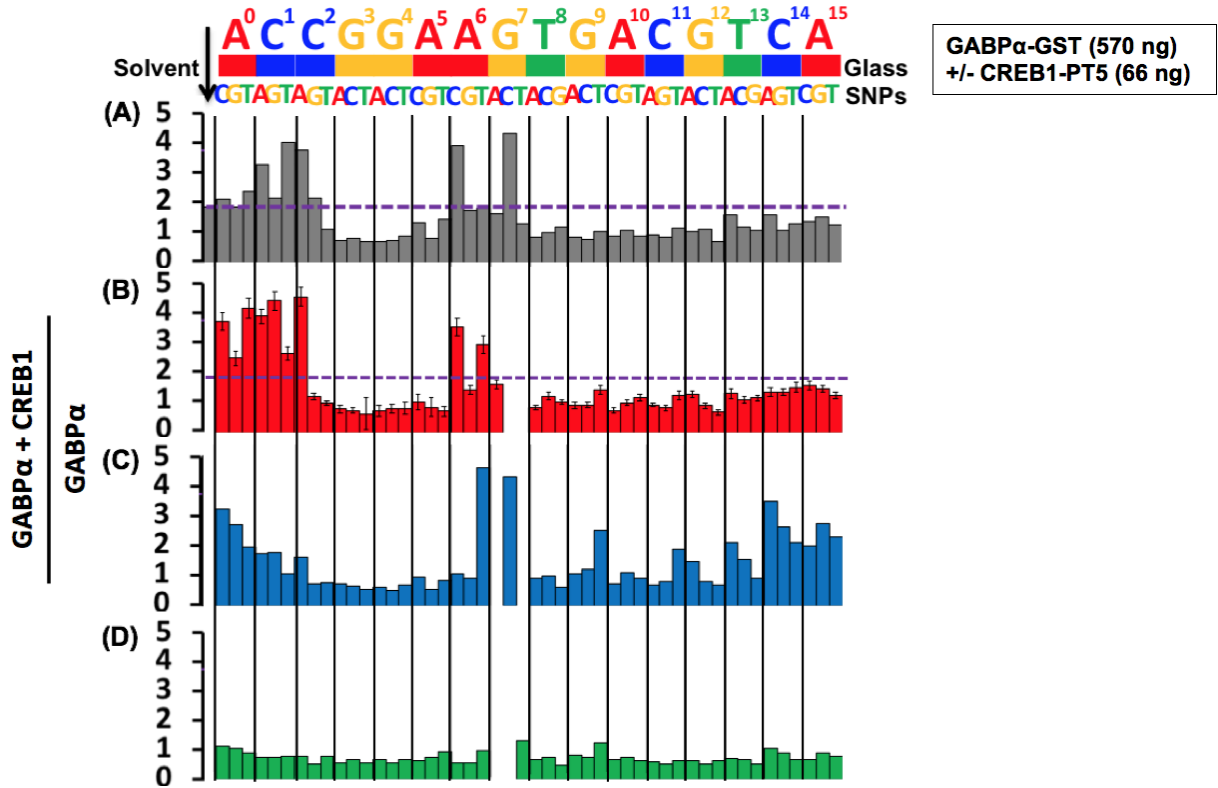
Fig S9. Cooperative binding of GABPα and CREB1 to Central ETS⇔CRE, SNPs, and position 7 DNPs. **(A)** Histogram of GABPα and CREB1 binding cooperativity, defined as the ratio of GABPα binding in the presence of CREB1 to binding in the absence of CREB1, to central ETS-CRE (indicated with arrow) and SNPs. A purple dashed line across the histogram represents the binding cooperativity to the canonical ETS-CRE motif. Histograms of GABPα and CREB1 binding cooperativity to central ETS-CRE SNPs and DNPs at position G^7 for **(B)** A^7 (red), **(C)** C^7 (blue), and **(D)** T^7 (green). A purple dashed line across the histogram represents the binding cooperativity to the canonical ETS⇔CRE motif. Error bars represent standard error from mean for the 24 occurrences of each sequence on the microarray.
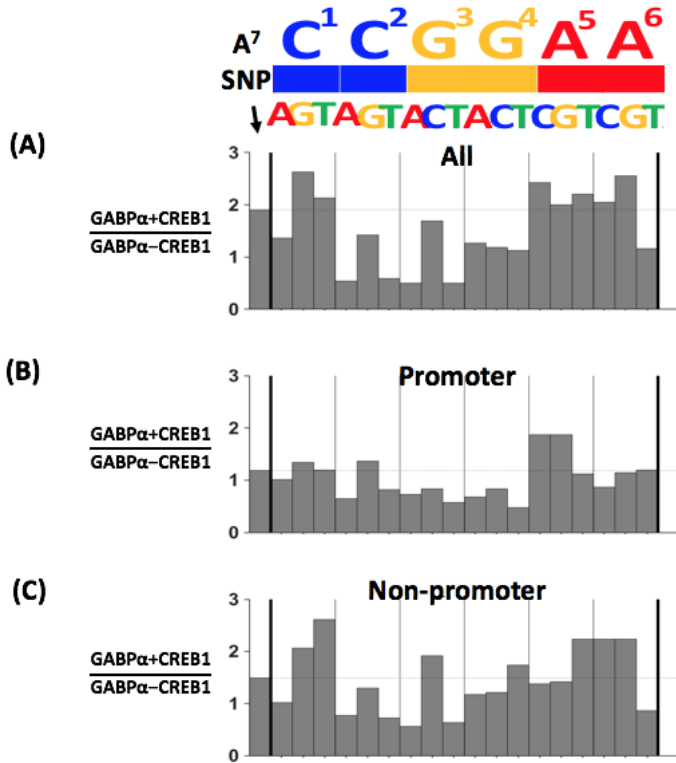
Figure S10. Enrichment of A⁷ DNPs at co-bound GABPα+CREB1 peaks *in vivo*. Histograms

showing the ratio of enrichment scores (see Methods) for the ETS motif SNP $A^0C^1C^2G^3G^4A^5A^6\underline{A}^7$ and

DNPs in **(A)** all, **(B)** promoter, and **(C)** non-promoter GABPα ChIP-seq peaks ± CREB1 measured in

A549 cells.