Supplementary Materials

Msn2 Coordinates a Stoichiometric Gene Expression Program

Supplemental Experimental Procedures

Microscopy and image analysis

Cells expressing Msn2-YFP or related constructs were plated in SD complete media onto ConcanavalinA coated 96 well glass bottom plates, allowed to settle and then washed twice with fresh media. Samples were imaged on a Nikon Ti inverted scope with arc-lamp illumination using RFP(560/40nm excitation, 630/75nm emission, Chroma) and YFP (510/10nm excitation, 542/27nm emission, Semrock) filters. Images were processed and analyzed with ImageJ and custom built Matlab scripts. Nuclear enrichment was computed by dividing the average intensity of the brightest 10pixels in the cell by the median intensity of the cell.

Yeast Strain Construction

Promoter constructs were integrated at the TRP1 locus of a HIS3+ MATa strain. Overexpression constructs were integrated into the TRP1 locus of a LEU2+ MATalpha strain which contained the estradiol inducible construct. These strains were then mated and diploids selected in SD-leu/his media. All strains were constructed using standard yeast protocols and LioAc/PEG transformation. For a complete list of strains and plasmids see supplementary tables 1 and 2.

Construction of Promoter Library

One kilobase upstream of a gene of interest was PCRed from genomic DNA with primers containing pspOMI (or Not1) and Xho1 (or Sal1) sites, digested, and cloned into a TRP1 marked single integration vector directly upstream of a yeast optimized Venus YFP. Clones were sequence verified and integrated into yeast by digestion with Nae1, transformation using a standard PEG/LioAc protocol and selection on SD-TRP plates.

Deletion of STRE elements was accomplished by quick change mutagenesis. The binding sites AGGGG were mutated to AGaGG, a mutation which appeared to eliminate activity *in vivo* and dramatically reduce the affinity for MSN2 *in vitro*.

Insertion of polyT(12x) elements was accomplished by quickchange mutagenesis, insertion sites were chosen based on location in the promoter (250-400bp upstream of the ATG) and availability of suitable regions for design of efficient quickchange primers (16GC basepairs within a 40bp region).

Construction of synthetic promoters

We constructed a vector with the prCYC1(1-243) sequence cloned directly upstream of Venus in a TRP1 marked vector. Additionally, we included the prCYC1(1000-684) sequence to reduce background expression. In between these sequences (which lack UASs), we inserted our STRE sequences or STRE like sequences.

4xSTRE: GGGCCCCTNCATTACCCCTNCTTTACCCCTNCAAACCCCTN

Where the N nucleotide was varied

3x6C: GGGCCCATTTACCCCCCATTTACCCCCCATTTACCCCCCA

These constructs were then integrated into yeast and assayed as above.

Construction of over expression constructs

Genes of interest were PCRed from yeast genomic DNA and cloned downstream of a prGAL1 promoter in a TRP1 marked single integration plasmid. Site directed mutagenesis was performed as necessary to construct the constitutively active alleles of MSN2 (S288A, S582A, S620A, S625A, S633A), Msn4 (S263A, S316A, S531/2A, S558A), or dominant negative alleles of RAS2 (S24N or G22A) using a standard QuickChange protocol and the pfuTurbo enzyme mix (Stratagene). Constructs were sequence verified, cut with Pme1 and integrated into yeast with selection on SD-TRP plates.

Measurement of Msn2 abundance

To convert arbitrary intensity units to absolute number of molecules, we constructed strains where we tagged one of NUF2, ASC1, and SPC42 with YFP. These proteins localize to the kinetocore and spindle pole body structures where they have a well described abundance. AS a result, they have been used previously as molecular standards (Karpova et al., 2008). We compared fluorescence values expressed from prGal1-Msn2(5A)-YFP and those from Nuf2, Asc1, or Spc42-YFP, and untagged or uninduced strains from the same background, all measured by Flow Cytometry. We assumed that the nuclear volume is 10% of the cell volume and after background subtraction, we used the nuclear enrichment measured by microscopy (Figure S2A) to compute the amount of fluorescence in arbitrary units of Msn2 in the nucleus. To convert this value to absolute numbers, we divided this by the background subtracted fluorescence of the Spc42-YFP strains. These calculations resulted in values ranging from 579-2300 nuclear Msn2 molecules depending on the estimate of the concentration of each standard we used and the particular Msn allele.

Zinc Finger Engineering H/T/HT alleles

To construct MSN2 alleles with altered affinity but identical binding preferences, we compared the sequence of the MSN2 DNA binding domain to that of other *S. cerevisiae* 2xC2H2 zinc finger proteins MSN4, COM2, GIS1, RPH1, MIG1, MIG2, ADR1 (Figure S3E). The consensus sequence (TGEKP) of the linker region is highly conserved across zinc fingers and mutations to this sequence are strongly associated with reduced affinity. Msn2 has an unusual spacing between the two histidines in the first finger (highlighted in green in Figure S3E. It makes up four residues, not the more common three we observed in Mig1 or Adr1 for example). To see if the increased spacing of histidines could explain the non-cannonical spacing we queried a protein structure database HMMER (http://hmmer.janelia.org/). Using the MSN2 DBD(residues 648-705) as a query, we downloaded all matching 2xC2H2 proteins and compared the linker structure for fingers with a HX_3H (101864) and those with a HX_4H spacing (4626) by constructing consensus seqlogs with standard matlab scripts (Figure S3F). We observe substantially less conservation of the core linker sequence in the HX_4H fingers, perhaps due to the disrupted helix which has been observed in the NMR structure of the second finger of ADR1 (Bernstein et al., 1994).

Zinc Finger Engineering 6G alleles

To alter the Msn2 binding preference towards 'GGGGGG', we examined the second finger (as the first already recognizes GGG). The recognition helix in this finger is the somewhat unusual RSDNLSQ, the residues highlighted in red are largely responsible for DNA recognition in a canonical zinc finger. A scan of the literature suggests that this helix should recognize AAG (with the Q and N both recognizing A and the R, G), consistent with its in vivo binding properties. Previously a RSDHLTR helix has been shown to recognize GGG, as has a RSDKLTR with less affinity (Pomerantz et al., 1995; Segal et al., 1999). We therefore constructed a Q693R, N690H/K MSN2 which did indeed have strong affinity for a 6xG site.

MITOMI devices and experiments

MITOMI devices were made as described in Maerkl and Quake, 2007 and Fordyce et al., 2010. Devices were based on the designs from Maerkl and Quake, 2007. The two layers of the device were made from RTV615 PDMS casts from the silicon molds. The two-layer device was aligned and bonded to a glass substrate with a contact-printed array of the DNA library (ssDNA template strands were ordered from IDT Coralville, Iowa, and Alexa-647 end-labeled second strands were synthesized with Klenow exo- enzyme). Finished devices were run as described previously (Maerkl and Quake 2007, Fordyce et al Nat Biotech 2010 and PNAS 2012). His-tagged Msn2 protein was synthesized in wheat germ extract (Promega), and fluorescently labeled by incorporation of Bodipy-lysine. After running the devices the fluorescence intensities were

scanned using an arrayWoRx scanner. Fluorescence data for bound DNA and protein and free DNA in the DNA chamber were extracted from the scanned images with Genepix 6.1 . A dilution series of the labeled primer flowed onto the DNA chambers was used as a standard curve to calibrate the relationship of Alexa-fluor signal to free DNA concentration in the DNA storage chamber on the devices. Binding curves were fit to a hyperbolic saturation curve with global nonlinear regression in Graphpad Prizm 4.00, to derive K_d values.

Model of MSN2 promoter regulation (figure 2)

Consider a model of MSN2 gene regulation whereby MSN2 is produced in the cytoplasm (at rate α), degraded (at rate γ_1) and translocated into the nucleus (at rate k_{in}). Nuclear MSN2 then is degraded (at the higher rate γ_n) and can re-enter the cytoplasm (at rate k_{out}). We assume that RFP is produced at a rate proportional to the amount of MSN2n and degraded (at a rate γ_3). This simple model is described by the differential equations:

$$\frac{dMsn2_c}{dt} = \alpha + k_{out}Msn2_n - k_{in}Msn2_c - \gamma_1\, Msn2_c$$

$$\frac{dMsn2_n}{dt} = -k_{out}Msn2_n + k_{in}Msn2_c - \gamma_n\, Msn2_n$$

$$\frac{dRFP}{dt} = k.Msn2_n + \gamma_3\, RFP$$

Solving this system of equations for at steady state, we get the following for RFP

$$RFP = \frac{k}{\gamma_3}\alpha\frac{k_{in}}{(k_{out}+\gamma_n)(kin+\gamma_1)} * \frac{1}{1-\frac{k_{in} * k_{out}}{(k_{out}+\gamma_n)(k_{in}+\gamma_1)}}$$

Showing a linear relationship between the production rate of MSN2 and RFP concentration with a slope governed by the relative rates of import/export and degradation.

For the simulations shown in Figure 2 we use the following parameters

k=1;	Arbitrary, scaling term
k_{in}=0.1-1	Arbitrary, import rate
k_{out}=0.4	Arbitrary
α =0.1-1	Arbitrary, scaling term
γ_1= γ_3=0.0077	Degradation rate = division rate of yeast, 90 minute half life
γ_2=5*0.0077	Degradation rate = nuclear degradation of Msn2, 15 minute half life

Model to generate heat map in Figure 4A

If X is a transcription factor and Y is its cognate binding cite, then gene expression is a function of the bound complex $X{:}Y$. The binding model is:

$$\frac{dx}{dt} = -\alpha * x * y + \gamma * xy$$
$$\frac{dxy}{dt} = \alpha * x * y - \gamma * xy$$
$$\frac{dy}{dt} = -\alpha * x * y + \gamma * xy$$

For our simulations we set α=10^3 (on rate, diffusion limited), γ=10^8 (off rate). X_{total}=1, and Y_{total} (number of binding sites) was varied over 9 log2s centered on 10^-7, the off rate (γ) was similarly varied over 9log2s centered on 10^8. The linearity score (reported in Figure 4A) at each point was computed as the R^2 coefficient between the concentration of X_{total} and XY.

Supplemental Figure Legends

S1: Many different methods for modulating Msn2 activity are equivalent. (A-D) Expression of 90 promoters enriched for their susceptibility to stress conditions was measured in exponential growth and in response to overexpression of Msn2(5A), Msn4(4A), Ras2(S24N), Ras2(G22A), or PDE2. The figures show the Log2 fold change in expression normalized to exponentially growing cells for each perturbation. (E) Expression of the same 90 promoters in a WT or in an *msn2/4Δ* strain with overexperssion of Ras2(S24N). The size of each circle represents the number of Msn2 binding sites in that promoter.

S2: Sequential removal of repressive PKA phosphorylation sites results in graded increase in Msn2 nuclear localization and transcriptional activity but does not change the essentially linear character of the transcriptional regulation of Msn2 of its target genes. We constructed six Msn2 alleles with serial substation of Alanine in each of the five PKA phosphorylation sites (WT, 1-5A). Each of these alleles tagged with YFP was expressed in a strain also harboring prTPS1-mKate2. Regulated expression of the Msn2 alleles was achieved using the estradiol titratable system (A) Quantification of the nuclear localization and abundance of each Msn2 allele. Increased nuclear localization results in lower protein levels likely due to high nuclear degradation. Errorbars are standard deviation across single cells (N=20-50) (B) A plot of prTPS1-mKate2 expression as a function of Msn2 for the different alleles.

S3: Dissection of Msn2 and its consensus binding sequence. (A) To explore the role of flanking nucleotides *in vivo* we scanned the genome for all potential Msn2 binding sequences ('NAGGGG') and compared the frequency of nucleotides in the 1st position for promoters with varying numbers of binding sites. Promoters with many Msn2 sites tend to have a mild increase in the number of strong 'AAGGGG' sites and a decline in the number of weak 'TAGGGG' sites. Errorbars represent standard errors computed from bootstrapping of the genomic dataset. (B) K_d values of wild type and mutant alleles of Msn2 from MITOMI 2.0 measurements for five DNA oligos lacking any consensus Msn2 binding sites show no correlation. (C) K_d measurements for the three Msn2 alleles binding to 34 oligos containing at least one 'AGGGG' sequence. Measurements are also carried out for five negative controls oligos lacking this sequence. (D) The location of Msn2 binding sites measured by MITOMI 2.0 are indicated on each promoter by red dashes where the height represents the value of the K_d. (E) Alignments of the zinc finger domains of different transcription factors in *S. cerevisiae* show similar structural arrangements of the DNA binding domain. The critical C2H2 residues which coordinate the zinc ion are marked in red, the linker between the two fingers in purple. (F) Seqlogos of the linker region of proteins with two adjacent zinc fingers with HX_3H (N=101864) or HX_4H (N=4626) spacing.

S4: Affinity, not cooperativity drive saturating transcription from Msn2 promoters. To compare Msn2 binding to other transcription factors we replaced the Msn2 DNA binding domain with that of the highly cooperative and dimeric transcription factor Gal4. (a) The Gal4-Msn2-YFP fusion was expressed from a tetracyclin regulated promoter and expression of mCherry driven by the galactose responsive Gal1 promoter was measured in response to a titration of the inducer anhydrotetracyclin (aTC). We observe a rapidly saturating interaction that resembles the high affinity 6G mutants of Msn2. This result suggests that cooperative interactions could lead to the saturating expression from the target promoter. (b) To test if cooperative interactions play a role in Msn2 we compared the transcription from a promoter with one, two, or three binding sites for

Msn2-6G mutant. We observe a graded increase in expression from the promoter as the number of binding sites is increased suggesting a lack of cooperativity between Msn2 molecules on promoter binding.

Figure S1

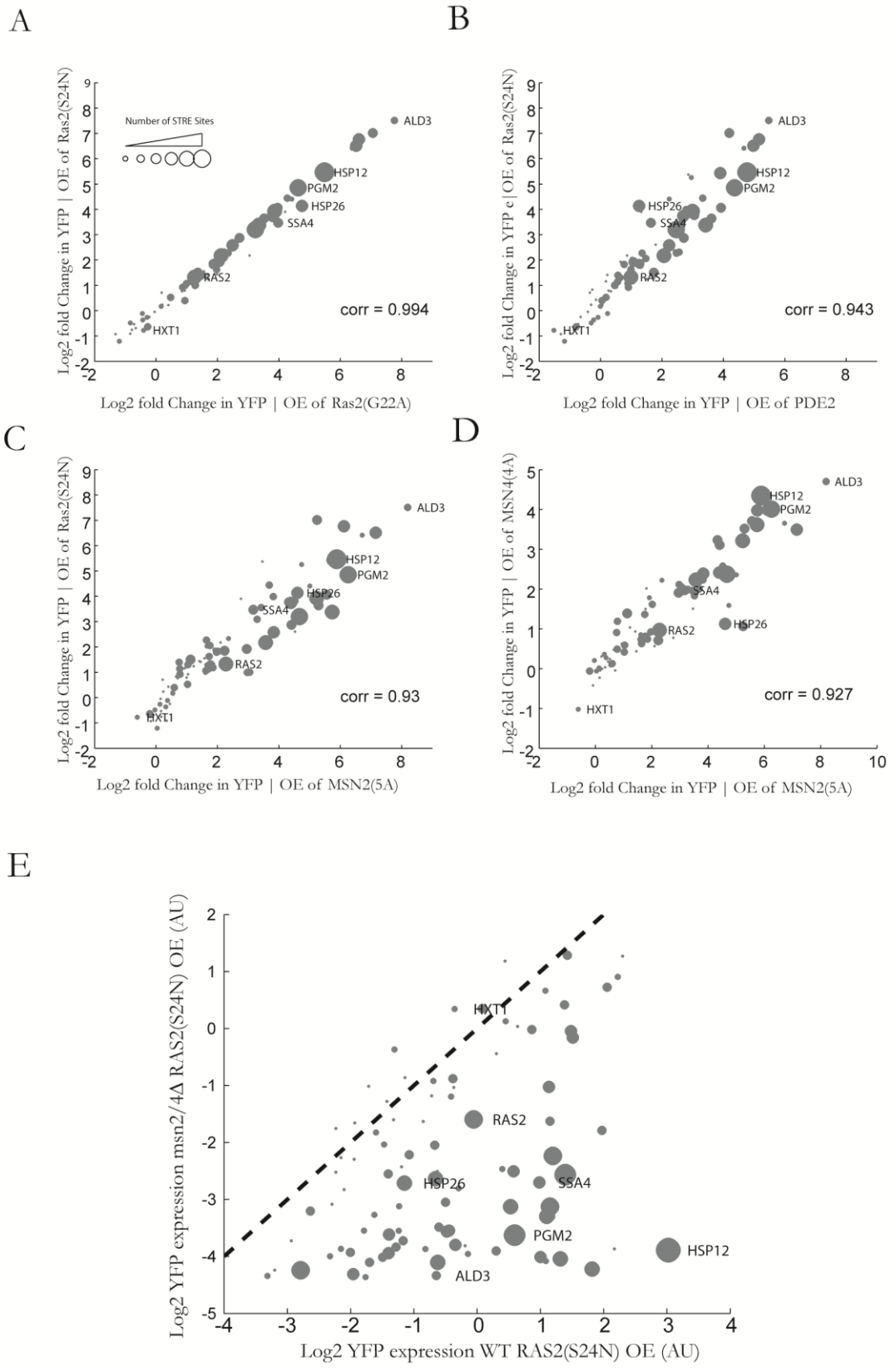

A
B
C
D
E
Number of STRE Sites
ALD3
HSP12
PGM2
HSP26
SSA4
RAS2
HXT1
corr = 0.994
corr = 0.943
corr = 0.93
corr = 0.927
Log2 fold Change in YFP | OE of Ras2(S24N)
Log2 fold Change in YFP | OE of Ras2(G22A)
Log2 fold Change in YFP | OE of PDE2
Log2 fold Change in YFP | OE of MSN2(5A)
Log2 fold Change in YFP | OE of MSN4(4A)
Log2 YFP expression msn2/4Δ RAS2(S24N) OE (AU)
Log2 YFP expression WT RAS2(S24N) OE (AU)

Figure S2

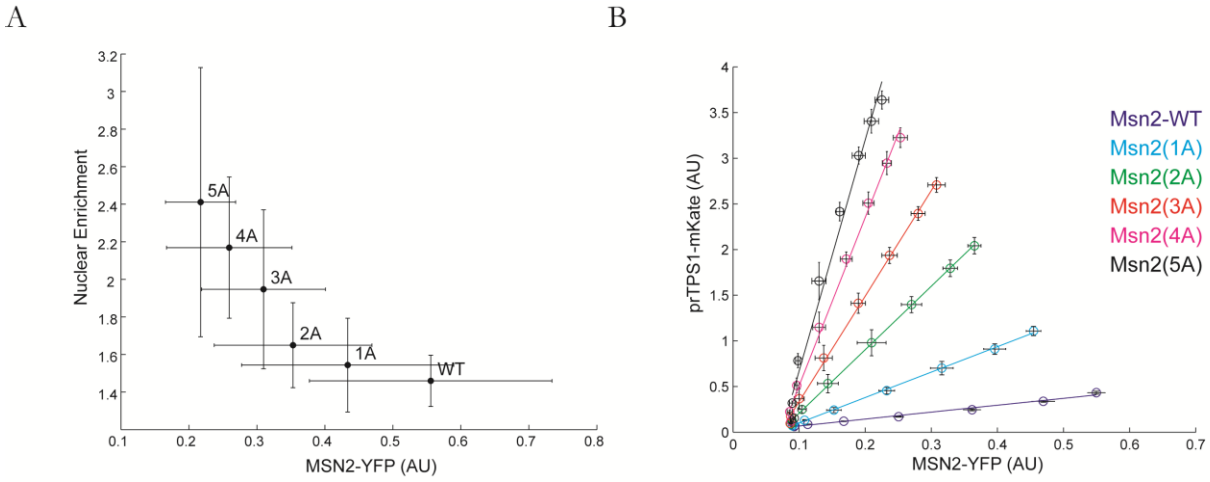

A
B
Nuclear Enrichment
MSN2-YFP (AU)
5A
4A
3A
2A
1A
WT
prTPS1-mKate (AU)
MSN2-YFP (AU)
Msn2-WT
Msn2(1A)
Msn2(2A)
Msn2(3A)
Msn2(4A)
Msn2(5A)

Figure S3

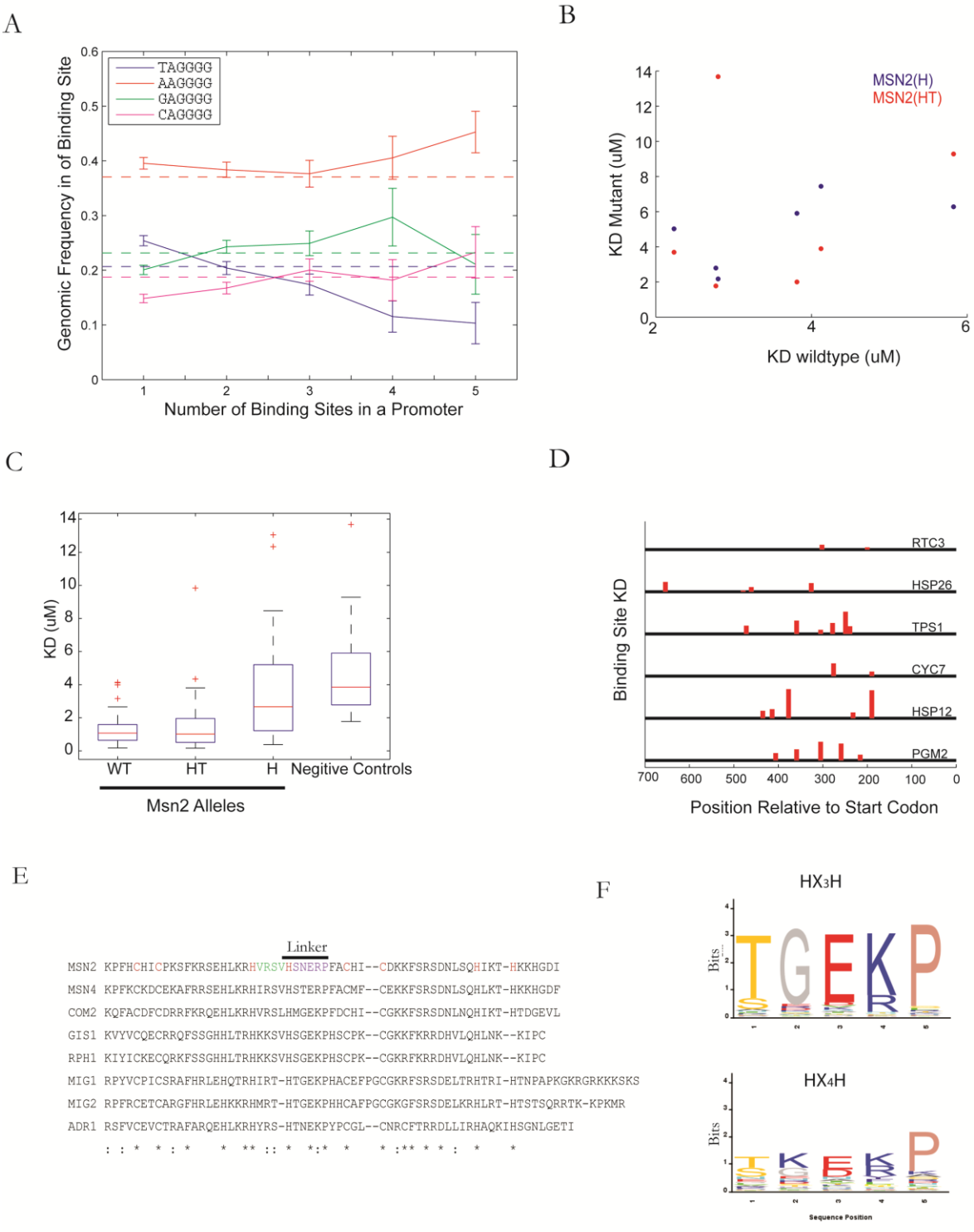
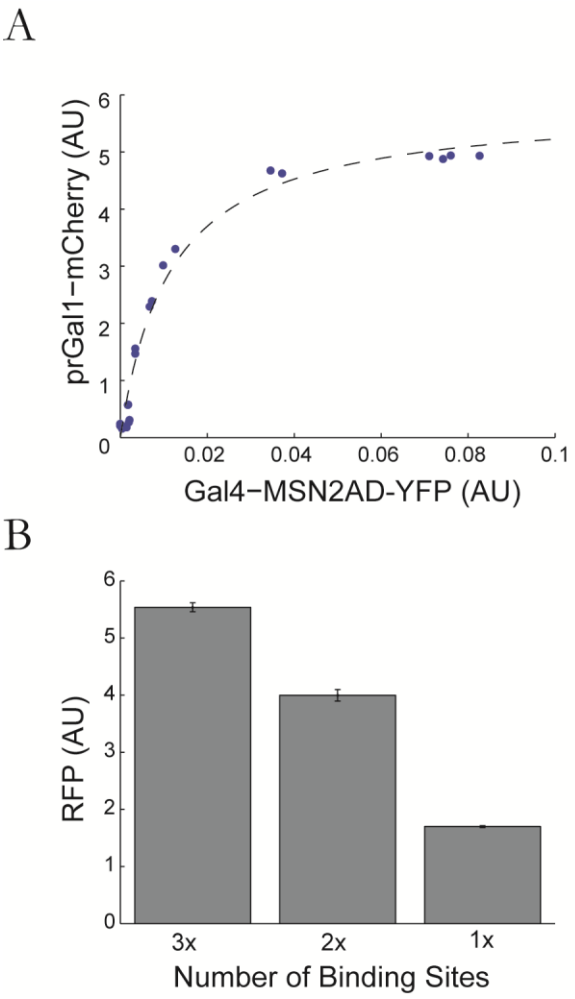

A
TAGGGG
AAGGGG
GAGGGG
CAGGGG
Genomic Frequency in of Binding Site
Number of Binding Sites in a Promoter
B
MSN2(H)
MSN2(HT)
KD Mutant (uM)
KD wildtype (uM)
C
KD (uM)
WT
HT
H
Negitive Controls
Msn2 Alleles
D
Binding Site KD
RTC3
HSP26
TPS1
CYC7
HSP12
PGM2
Position Relative to Start Codon
E
Linker
MSN2 KPFHCHICPKSFKRSEHLKRHVRSVHSNERPFACHI--CDKKFSRSDNLSQHIKT-HKKHGDI
MSN4 KPFKCKDCEKAFRRSEHLKRHIRSVHSTERPFACMF--CEKKFSRSDNLSQHLKT-HKKHGDF
COM2 KQFACDFCDRRFKRQEHLKRHVRSLHMGEKPFDCHI--CGKKFSRSDNLNQHIKT-HTDGEVL
GIS1 KVYVCQECRRQFSSGHHLTRHKKSVHSGEKPHSCPK--CGKKFKRRDHVLQHLNK--KIPC
RPH1 KIYICKECQRKFSSGHHLTRHKKSVHSGEKPHSCPK--CGKRFKRRDHVLQHLNK--KIPC
MIG1 RPYVCPICSRAFHRLEHQTRHIRT-HTGEKPHACEFPGCGKRFSRSDELTRHTRI-HTNPAPKGKRGRKKKSKS
MIG2 RPFRCETCARGFHRLEHKKRHMRT-HTGEKPHHCAFPGCGKGFSRSDELKRHLRT-HTSTSQRRTK-KPKMR
ADR1 RSFVCEVCTRAFARQEHLKRHYRS-HTNEKPYPCGL--CNRCFTRRDLLIRHAQKIHSGNLGETI
F
HX3H
Bits
HX4H
Bits
Sequence Position

Figure S4


A
prGal1−mCherry (AU)
Gal4−MSN2AD-YFP (AU)
0
1
2
3
4
5
6
0.02
0.04
0.06
0.08
0.1
B
RFP (AU)
3x
2x
1x
Number of Binding Sites