Supplemental Information, Takahashi et al.

Supplemental Figure Legends

Figure S1. **Identification of residues critical for interaction of Med26 NTD with EAF1 or TFIID. Related to Figure 4**. (A) Activation of a GAL4-dependent luciferase reporter by GAL4-NTD. Relative luciferase activities were measured in extracts of cells transiently transfected with expression plasmids encoding empty vector (vector), the GAL4 DNA binding domain (GAL4) or the indicated GAL4-Med26 NTD proteins. (B) Mutation of Med26 NTD residues R61,K62 or K74,K75 interferes with recruitment of EAF1 or TFIID to immobilized GAL4x5-MLT template. Binding reactions included EAF1 or TFIID, and either GAL4 DNA binding domain or the indicated GAL4-Med26 NTD fusion proteins.

Figure S2. Effect of Med26 depletion on ES cell proliferation and gene expression. Related to Figure 5. (A) Knockdown of endogenous Med26 and other Mediator subunits in ES cells. 3 days after transfection with the indicated siRNAs, cells were lysed and subjected to immunoblotting with anti-Med26, CDK8, or Med6 antibodies. (B) Med26 knockdown inhibits ES cell proliferation. 4 days after transfection with the indicated siRNA, KH2 ES cells were stained with a cell-permeable green fluorescent dye to identify live cells (Live-Dead Cell staining kit, BioVision, Cat.# K501-100). (C) Effect of Med26 depletion on *Hsp70*, *BMP2* and *JUN* expression in HEK293T cells. Gene expression was measured by real time qPCR and normalized to *GAPDH*. Data are averages from three independent experiments; error bars show standard deviations.

Figure S3. Localization SEC and Mediator subunits on the c-*MYC* and *HSP70* genes. Related to Figure 6. (A) Occupancy of Med26 and endogenous or exogenously expressed SEC compornents on the c-*MYC* gene. (B) Recruitment of Pol II, Mediator subunits and SEC components to the *HSP70* gene by heat shock. Bars indicate average ChIP / input obtained from two biological replicates, error bars show the data range.

Extended Experimental Procedures

Mass spectrometry. Identification of proteins was accomplished using a modification of the multidimensional protein identification technology (MudPIT) procedure (Florens and Washburn, 2006; Washburn et al., 2001). TCA-precipitated proteins were urea-denatured, reduced, alkylated and digested with endoproteinase Lys-C (Roche) followed by modified trypsin (Roche) as described (Washburn et al., 2001). Peptide mixtures were loaded onto 100 µm fused silica microcapillary columns packed with 5-µm C_{18} reverse phase (Aqua, Phenomenex), strong cation exchange particles (Partisphere SCX, Whatman), and reverse phase (McDonald et al., 2004). Loaded microcapillary columns were placed in-line with an LCQ or LTQ ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Fully automated MudPIT runs were carried out on the electrosprayed peptides, as described by (Florens and Washburn, 2006). Tandem mass (MS/MS) spectra were interpreted using SEQUEST (Eng et al., 1994) against a database of a database of 30709 human proteins (downloaded from NCBI on 2009-10-27) , and complemented with 177 sequences from usual contaminants (human keratins, IgGs, proteolytic enzymes). In addition, to estimate false positive discovery rates, each sequence was randomized (keeping AA composition and length the same) and the resulting "shuffled" sequences were added to the "normal" DB (doubling its size) and searched at the same time.

Peptide/spectrum matches were sorted and selected using DTASelect (Tabb et al., 2002) with the following criteria set: spectra/peptide matches were only retained if they had a DeltaCn of at least 0.08, and minimum XCorr of 1.8 for singly-, 2.0 for doubly-, and 3.0 for triply-charged spectra. In addition, peptides had to be fully-tryptic and at least 7 amino acids long. Combining all runs, proteins had to be detected by at least 2 such peptides, or 1 peptide with 2 independent spectra. Under these criteria, the overall false discovery rate was 0.18%, suggesting a >99.82% confidence level for positive protein identifications. Peptide hits from multiple runs were compared using CONTRAST (Tabb et al., 2002). To estimate relative protein levels, Normalized Spectral Abundance Factors (NSAFs) were calculated for each detected protein (Florens et al., 2006; Paoletti et al., 2006; Zhang et al., 2010; Zybailov et al., 2006). The NSAF for a protein *k* is proportional to the amount of the protein present in the sample and is calculated using the formula:

$$(\mathrm{NSAF})_k = \frac{(SpC/L)_k}{\sum_{i=1}^{N} (SpC/L)_i},$$

where *SpC* = spectal count, *L* = protein length in amino acids, and *i* = all proteins detected in the MudPIT runs.

Plasmid construction and generation of stable cell lines. cDNAs encoding full-length and various mutant versions of human Med26 (NM_004831.3) or human EAF1 (NM_033083.6), ELL2 (NM_012081.5), AFF4 (NM_014423.3), CDK9 (NM_001261.3) or TAF7 (BC032737) were introdruced by retroviral transduction into a Hela S3 cell line stably expressing the mouse ecotropic retrovirus receptor (mCAT-1) (Albritton et al., 1989). Full-length human NARG2 (NM_024611.4), KIAA0947 (NM_015325.1), and AF9 (IMAGE clone ID: 5298142) were amplified from HeLa cell total RNA or obtained from the American Type Culture Collection, subcloned into pcDNA5-FRT with epitope tags, introduced into Flp-In 293 cells using the Invitrogen Flp-in system, and clonal stable cell lines were established following the manufacturers' instructions.

Clonal HEK293T cell lines stably expressing TRIPZ lentiviral Dox-inducible shRNAmirs (Open Biosystems, Huntsville, AL), either nontargeting (RHS4743) or targeting the *MED26* 3'-UTR (#1, RHS1764-9695596; #2, RHS1764-9394914; #3, RHS1764-9209087), were generated according to the manufacturer's instructions. To generate cell lines stably expressing Dox-inducible shRNA with Med26 rescue contructs, HEK293T cells expressing Med26 shRNA #3 were stably transformed with pcDNA3.1/Hygro carrying wild type or mutant Med26 with an N-terminal epitope tag.

Antibodies for western blotting and immunoprecipitation. Anti-Flag (M2) and anti-HA (HA-7) antibodies were from Sigma; monoclonal anti-TBP (ab818), anti-TAF6 (ab51026), anti-AFF4 (ab57077) anti-AF9 (ab60083) antibodies were from Abcam: monoclonal anti-TAF1 (sc-735), anti-TAF4 (sc-736), anti-CDK9 (sc-13130) antibodies and rabbit polyclonal anti-GAL4 (sc-577), anti-Cyclin T1 (sc-8127) antibodies were from Santa Cruz; rabbit polyclonal anti-Cyclin T2 (A301-678A) antibody was from Bethyl. Monoclonal anti-ELL antibody has been described (Lin et al., 2010). Mouse monoclonal anti-EAF1 antibody [cite] was a gift from Michael Thirman (Department of Medicine, University of Chicago).

ES cell culture. ES cells (KH2) were cultured in 6 well tissue culture plates in Glasgow Modified Eagle Medium containing 10 % fetal bovine serum, 1/100 (v/v) L-glutamine (Gibco 25030-024), 1/100 (v/v) non-essential amino acids (Gibco 11140-035), 0.1 mM 2-mercaptoethanol (Gibco 31350-010) and 1000 U/ml LIF (Chemicon, No. ESG1107).

siRNA transfections. HEK293T cells in 6 well tissue culture plates (~$1x10^5$ cells/well) or 10 cm dishes (~$2\ x10^6$ cells/dish) were transfected with 25 nM siRNAs targeting human Med26 (Ambion/Applied Biosystems, #4, s18074; #5, s18075; #6, s18076) or 25 nM siGENOME NON-TARGETING siRNA Pool #2 (Dharmacon D-001206-14) using X-tremeGENE siRNA Transfection Reagent (Roche) and grown for 48 hours. ES cells were transiently transfected with siRNA targeting mouse Med26 (Ambion/Applied Biosystems, #7: cat. number s89018; #8; cat. number s89019, #9, cat. number s89020), siRNA designed against mouse Med6 (ON-TARGET plus SMART pool, L-055846-01), siRNA targeting mouse CDK8 (ON-TARGET plus SMART pool, L-053848-00) or siCONTROL NON-TARGETING siRNA (Dharmacon) using Lipofectamine™ RNAiMAX Transfection Reagent (Invitrogen).

Luciferase reporter assays. HEK293T cells were cotransfected with 1 μg of the GAL4-responsive luciferase reporter plasmid pG5-Luc (Promega), 100 ng of the renilla luciferase control plasmid pRL-tk (Promega), and 250 ng effector plasmid expressing wild type or mutant GAL4-Med26 NTD using FuGene 6 reagent (Roche). Firefly and renilla luciferase activities were measured using the Dual-Luciferase Reporter assay kit (Promega). Each experiment was performed in triplicate; error bars shown in the figure represent standard deviations.

Cell proliferation assays. $5x10^5$ 293T cells stably expressing Dox-inducible shRNAmir were seeded on 10 cm dishes and cultured in 10% Tet-free FBS medium with or without 2 μg/ml doxycycline. The number of live cells was determined by counting using a haemocytometer after trypan blue staining at the indicated times. Data are average from three independent experiments.

Gene Expression Analysis. Total RNA was isolated using the RNeasy Protect Mini Kit (Qiagen). For RT-qPCR, total mRNA was reverse transcribed using the iScript Select cDNA Synthesis Kit (Biorad), and real-time PCR reactions were performed using an iCycler iQ™ Real-Time PCR Detection System and iQ SYBR green supermix (Bio-Rad). Ct values were normalized to GAPDH. Primer sequences are listed in PCR primers. For genome-wide expression analyses, total RNA was amplified and labeled using the MessageAmp™ III RNA Amplification Kit (Ambion), and aRNA was hybridized to Affymetrix U133A plus 2.0 expression arrays. Hybridized arrays were washed, stained, scanned according to standard procedures (Affymetrix). Affymetrix CEL files were processed in the R statistical environment and normalized using RMA (Irizarry et al., 2003). The linear modeling package Limma (Smyth, 2004) was used to derive gene expression coefficients and calculate p-values. P-values were adjusted for multiple hypothesis testing using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995). Lists of the most highly affected genes (Table S2) were generated by applying a log2-fold change cutoff of ≤ -0.5 or ≥ 0.5 to data prefiltered for p value; these included 97 well-characterized genes that were down-regulated and 196 that were up-regulated after cells were treated independently with all three of the siRNAs.

Chromatin immunoprecipitations. Cells from one 10 cm dish (~ 1 x 10^7) of HEK293T cells or HeLa cells grown to 70-80% of confluence were used for each immunoprecipitation. Cells that were non-heat shocked or heat shocked at 42°C for the indicated times were cross-linked with 1% formaldehyde in PBS for 20 min at room temperature. Cells were resuspended and lysed in lysis buffer (0.2% or 0.5% SDS, 10mM EDTA, 150 mM NaCl, 50 mM Tris-HCl pH 8.0), and were sonicated with a Bioruptor® Sonicator (Diagenode) for 30 x 30 seconds at the maximum power setting to generate DNA fragments of ~150-400 bps. Sonicated chromatin was incubated at 4°C overnight with 5-15 μg of normal IgG or specific antibodies. Antibodies used were as follows: CDK9 (sc-8338, Santa cruz), AFF4 (A302-539A, Bethyl); Pol II total Rpb1 (N-20, sc-899, Santa Cruz); Pol II Rpb1 CTD phospho Ser5 (ab5131, Abcam); Rpb1 CTD phospho Ser2 (H5; ab24758, Abcam with Upstate IgG-IgM linker antibody 12-488, Millipore), FLAG M2 (F3165, Sigma), Med26 (sc-48776, Santa cruz), Med1 (sc-5334, Santa cruz) or TBP (ab51841, Abcam). Then, salmon sperm DNA-protein A agarose (16-157, Millipore) was added and incubated for 2 hours at 4°C. Beads were washed 2 times with IP buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100), 2 times with high salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100), 1 time with LiCl buffer (250 mM LiCl, 20 mM Tris-HCl pH8.0, 1 mM EDTA, 1% Triton X-100, 0.1% NP40 and 0.5% NaDOC) and 2 times with TE buffer. Bound complexes were eluted from the beads with 100 mM $NaHCO_3$ and 1% SDS by incubating at 50°C for 30 mins with occasional vortexing. Crosslinking was reversed by overnight incubation at 65°C. Immunoprecipitated DNA and input DNA were treated with RNase A and Proteinase K by incubation at 45°C. DNA was purified using the QIAquick PCR purification kit (28106, Qiagen) or MinElute PCR purification kit (28006, Qiagen). Immunoprecipitated and input material was analyzed by quantitative PCR. ChIP signal was normalized to total input.

Production of recombinant proteins. N-terminally 6xHis-tagged Med26 NTD (residues 1-113), GAL4-VP16, HNF4α, GAL4 DNA binding domain (1-97) or GAL4 DNA binding domain fused to Med26 NTD (GAL4-Med26 NTD) and derivatives were expressed in BL21(DE3) CodonPlus *Escherichia coli* (Stratagene) and purified as described (Yao et al., 2006). For production of recombinant proteins in Sf9 or Sf21 cells, epitope-tagged ELL, EAF1 and Med26 were subcloned into pBacPAK8 expressed using the BacPAK system (Clontech). Full length or deletion mutants of NARG2 and KIAA0947 were subcloned into pFastBac HTb with epitope tags and expressed singly or together with the BAC-to-BAC system (Clontech). Baculovirus infections, Sf9 or Sf21 culture, and affinity purifications were performed as described (Yao et al., 2006).

PCR Primers

Primers used for ChIP-qPCR

HSPA1A (-50)
5’- TCTGATTGGTCCAAGGAAGGCTG-3’ (forward)
5’- TTTCCCTTCTGAGCCAATCACCGA-3’ (reverse)

HSPA1A (+492)
5’-AGGTGATCAACGACGGAGAC-3’ (forward)
5’-ATCTCCTCGGGGTAGAATGC-3’ (reverse)

HSPA1A (+898)
5’-GGGTGGGGAGGACTTTGACAACAGG-3’ (forward)
5’-TGGCTGATGTCCTTCTTGTG-3’ (reverse)

HSPA1A (+1486)
5’-GACGCAGATCTTCACCACCT-3’ (forward)
5’-GCCCCAACAGATTGTTGTCT -3’ (reverse)

HSPA1A (+1895)
5’- AAGATCAGCGAGGCGGACAAGAA-3’ (forward)
5’- TCCTCTTGTGCTCAAACTCGTCCT-3’ (reverse)

MYC (-3kb)
5’-AACCTCCACTGCCAGAAGTCCTTA-3’ (forward)
5’-GAAATTTACCTGGCACGTGTCCCT-3’ (reverse)

MYC (+83)
5’-TTCTCAGAGGCTTGGCGGGAAA-3’ (forward)
5’-CTGCCTCTCGCTGGAATTACTACA-3’ (reverse)

MYC (+548)
5’-AAACCAGGTAAGCACCGAAGTCCA-3’ (forward)
5’-TGTCAATAGCGCAGGAATGGGAGA-3’ (reverse)

MYC (+1170)
5’-AACCTGGGTCTCTAGAGGTGTTAGGA-3’ (forward)
5’-TCAACGATTCCAGGAGAATCGGAC-3’ (reverse)

MYC (+2243)
5’-ACTCGGTGCAGCCGTATTTCTACT-3’ (forward)
5’-GCAGCAGCTCGAATTTCTTCCAGA-3’ (reverse)

Primers used for expression analysis

Human GAPDH
5’-TCGACAGTCAGCCGCATCTTCTTT-3’ (forward)
5’-GCCCAATACGACCAAATCCGTTGA-3’ (reverse)

Human MYC
5’-ACAGCTACGGAACTCTTGTGCGTA-3’ (forward)
5’-CAGCCAAGGTTGTGAGGTTGCATT-3’ (reverse)

Human HSP70
5’-TGCTGGACAAGTGTCAAGAGGTCA-3’ (forward)
5’-TCCTCTTGTGCTCAAACTCGTCCT-3’ (reverse)

Human BMP2
5’-CAACCATGGATTCGTGGTGGAAGT-3’ (forward)
5’-CCAGCTGTGTTCATCTTGGTGCAA-3’ (reverse)

Human SLC7A11
5’-TCTTCATGGTTGCCCTTTCCCTCT-3’ (forward)
5’-ACCACCTGGGTTTCTTGTCCCATA-3’ (reverse)

Human SNAIL2
5’-TTTCTGGGCTGGCCAAACATAAGC-3’ (forward)
5’-ACACAAGGTAATGTGTGGGTCCGA-3’ (reverse)

Human JUN
5’-AGATGAACTCTTTCTGGCCTGCCT-3’ (forward)
5’-ACACTGGGCAGGATACCCAAACAA-3’ (reverse)

Supplemental References

Albritton,L.M., Tseng,L., Scadden,D., and Cunningham,J.M. (1989). A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. Cell *57*, 659-666.

Benjamini,Y. and Hochberg,Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Statist. Soc. Ser. B *57*, 289-300.

Eng,J.K., McCormack,A.L., and Yates III,J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass. Spec. *5*, 976-989.

Florens,L., Carozza,M.J., Swanson,S.K., Fournier,M., Coleman,M.K., Workman,J.L., and Washburn,M.P. (2006). Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods *40*, 303-311.

Florens,L. and Washburn,M.P. (2006). Proteomic analysis by multidimensional protein identification technology. Methods Mol. Biol. *328*, 159-175.

Irizarry,R.A., Hobbs,B., Collin,F., Beazer-Barclay,Y.D., Antonellis,K.J., Scherf,U., and Speed,T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. *4*, 249-264.

Lin,C., Smith,E.R., Takahashi,H., Lai,K.C., Martin-Brown,S., Florens,L., Washburn,M.P., Conaway,J.W., Conaway,R.C., and Shilatifard,A. (2010). AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol. Cell *37*, 429-437.

McDonald,W.H., Tabb,D.L., Sadygov,R.G., MacCoss,M.J., Venable,J., Graumann,J., Johnson,J.R., Cociorva,D., and Yates,J.R., III (2004). MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. Rapid Commun. Mass Spectrom. *18*, 2162-2168.

Paoletti,A.C., Parmely,T.J., Tomomori-Sato,C., Sato,S., Zhu,D., Conaway,R.C., Conaway,J.W., Florens,L., and Washburn,M.P. (2006). Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. Proc. Natl. Acad. Sci. U. S. A. *103*, 18928-18933.

Smyth,G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. *3*, Article3.

Tabb,D.L., McDonald,W.H., and Yates III,J.R. (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. *1*, 21-26.

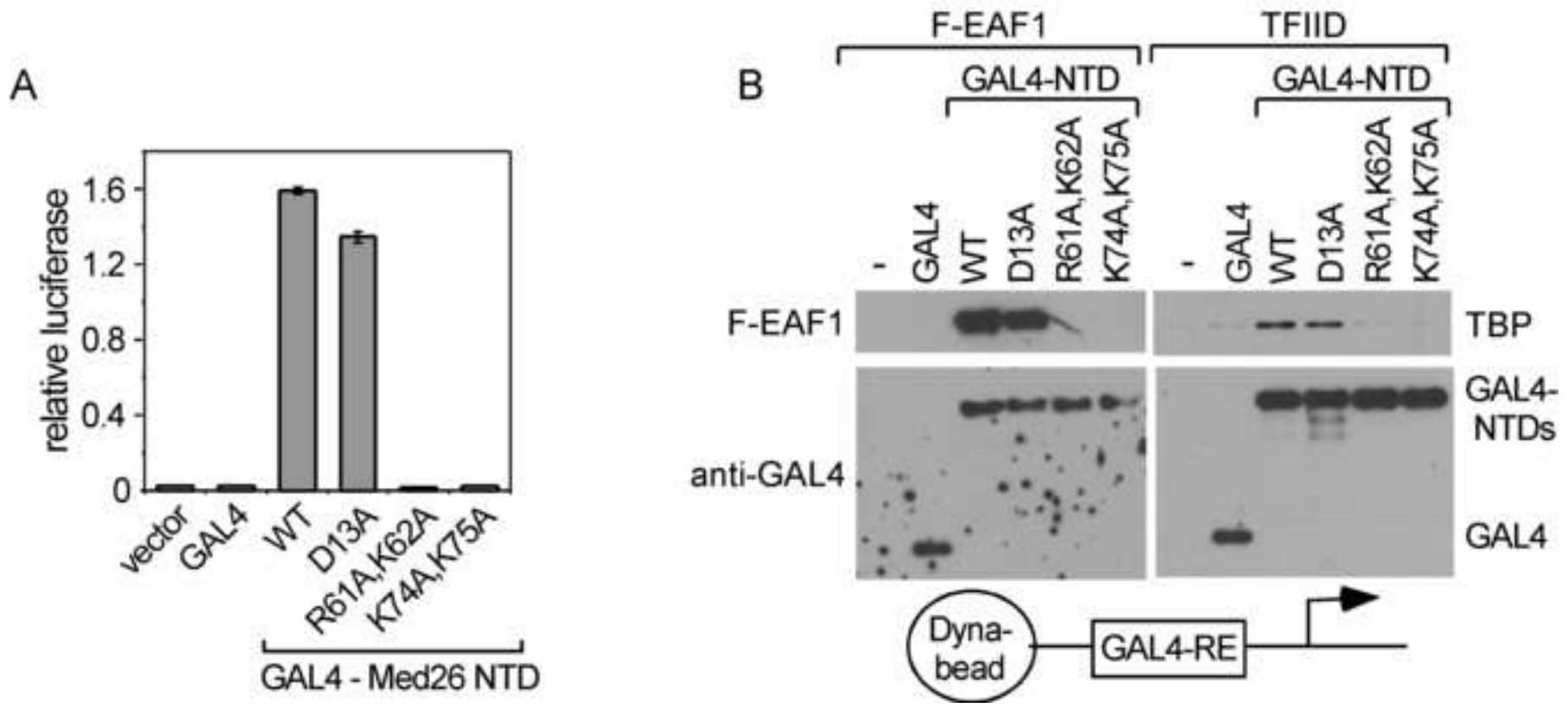
Washburn,M.P., Wolters,D., and Yates,J.R., III (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. *19*, 242-247.

Yao,T., Song,L., Xu,W., DeMartino,G.N., Florens,L., Swanson,S.K., Washburn,M.P., Conaway,R.C., Conaway,J.W., and Cohen,R.E. (2006). Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. Nat. Cell Biol. *8*, 994-1002.

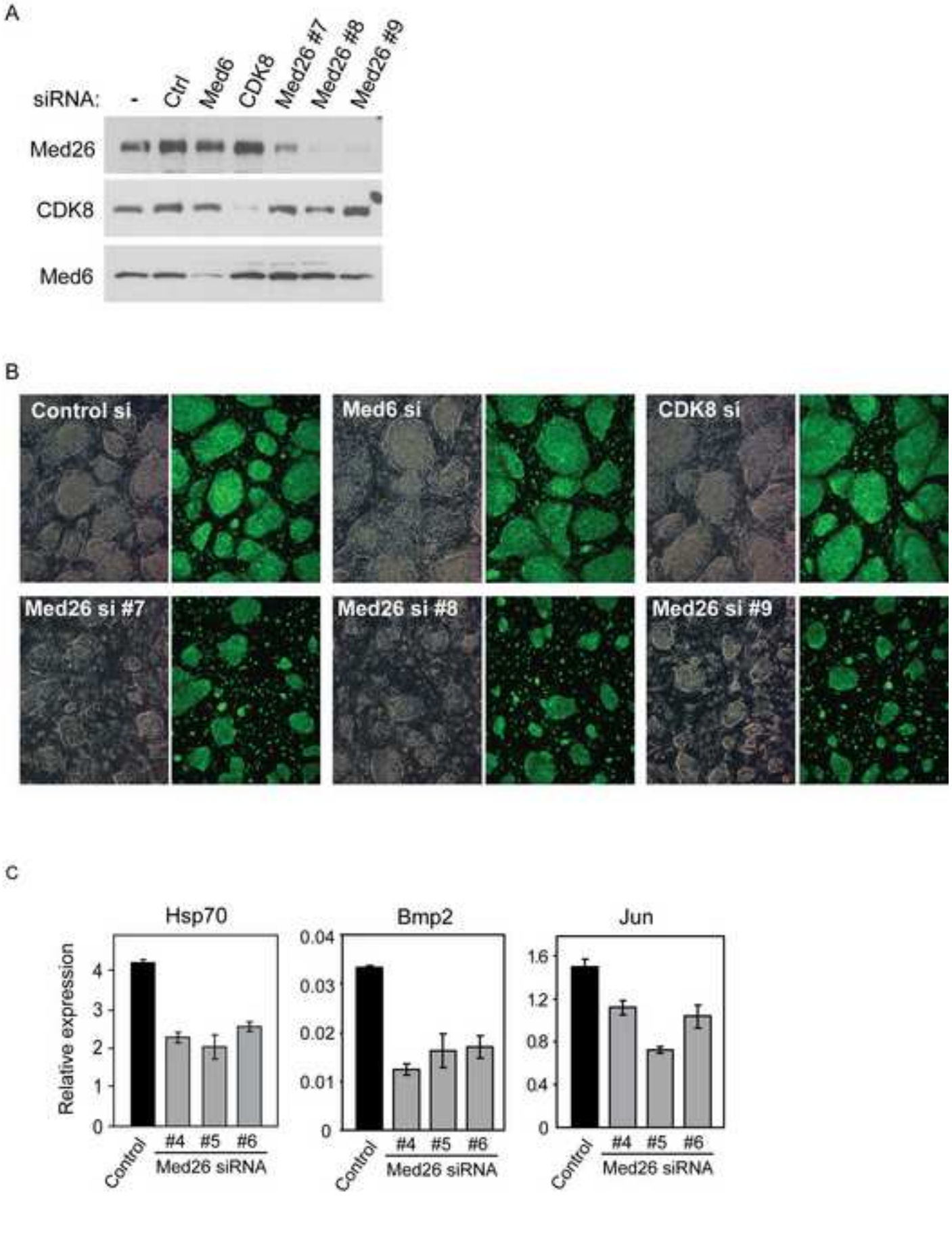
Zhang,Y., Wen,Z., Washburn,M.P., and Florens,L. (2010). Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal. Chem. *82*, 2272-2281.

Zybailov,B., Mosley,A.L., Sardiu,M.E., Coleman,M.K., Florens,L., and Washburn,M.P. (2006). Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces cerevisiae*. J. Proteome Res. *5*, 2339-2347.

Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3

Fig. S3