Landes Bioscience www.landesbioscience.com

Supplemental Material to:

Andrey Poleshko, Andrew V Kossenkov, Natalia Shalginskikh, Anna Pecherskaya, Margret B Einarson, Anna Marie Skalka, Richard A Katz

Human factors and pathways essential for mediating epigenetic gene silencing

Epigenetics 2014; 9(9)
http://dx.doi.org/10.4161/epi.32088
http://www.landesbioscience.com/journals/epigenetics/article/32088/

Additional Files:

http://www.landesbioscience.com/journals/epigenetics/ 2014EPI0224R_ST1.xls article/32088/

http://www.landesbioscience.com/journals/epigenetics/ 2014EPI0224R_ST2.xls article/32088/

http://www.landesbioscience.com/journals/epigenetics/ 2014EPI0224R_ST3.xls article/32088/

http://www.landesbioscience.com/journals/epigenetics/ 2014EPI0224R_ST4.xls article/32088/

SUPPLEMENTARY MATERIALS

Supplementary Table Legends

Table S1. **Summary of the complete genome-wide siRNA screen results.** The table contains raw data from all 21,121 genes/factors interrogated in the primary screen, and the 470 genes analyzed in the validation screen.

Table S2. **A comprehensive description of the 128 validated factor hits.** The table includes screening data, gene descriptions, intracellular localization, domain structures, Gene Ontology functional categories and other details.

Table S3. **Significantly enriched Pathways and Gene Ontology categories among the 128 factor hits.** A comprehensive matrix showing protein factor hits versus significantly enriched annotation terms. Highlighted are the functional categories used to create Figure 2.

Table S4. **A list of 18 previously uncharacterized factors that are candidate mediators of epigenetic gene silencing**.

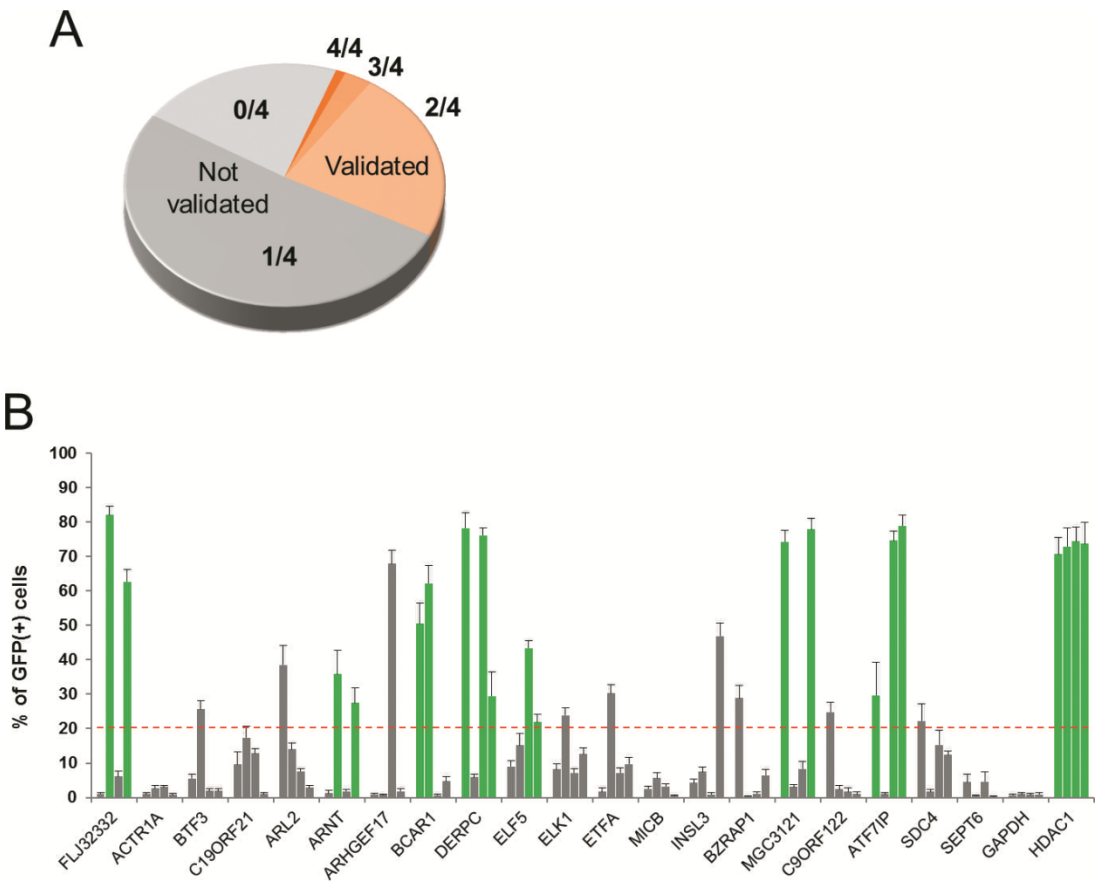

Figure S1. **Deconvolution of 4 siRNAs from SMARTpools for validation step. A.** Distribution based on the number of siRNAs scoring positive in validation. Validation analysis was performed on 470 primary factor hits. The number of individual siRNA scoring for each gene was 108 for 0/4, 234 for 1/4, 104 for 2/4, 19 for 3/4 and 5 for 4/4. Colored slices indicate validated genes. **B.** GFP reactivation, expressed as percentage of GFP positive cells, is shown for selected validated and non-validated factors. The criteria for an individual siRNA scoring positive was reactivation of GFP in 20 percent of the cells (red dashed line). A factor was considered validated if at least 2 of the 4 independent siRNAs from the pool scored positive. Green bars indicate positive siRNAs that are members of validated siRNA pools. Grey bars indicate siRNAs that are either members of non-validated pools, or failed to achieve cut-off levels of reactivation.

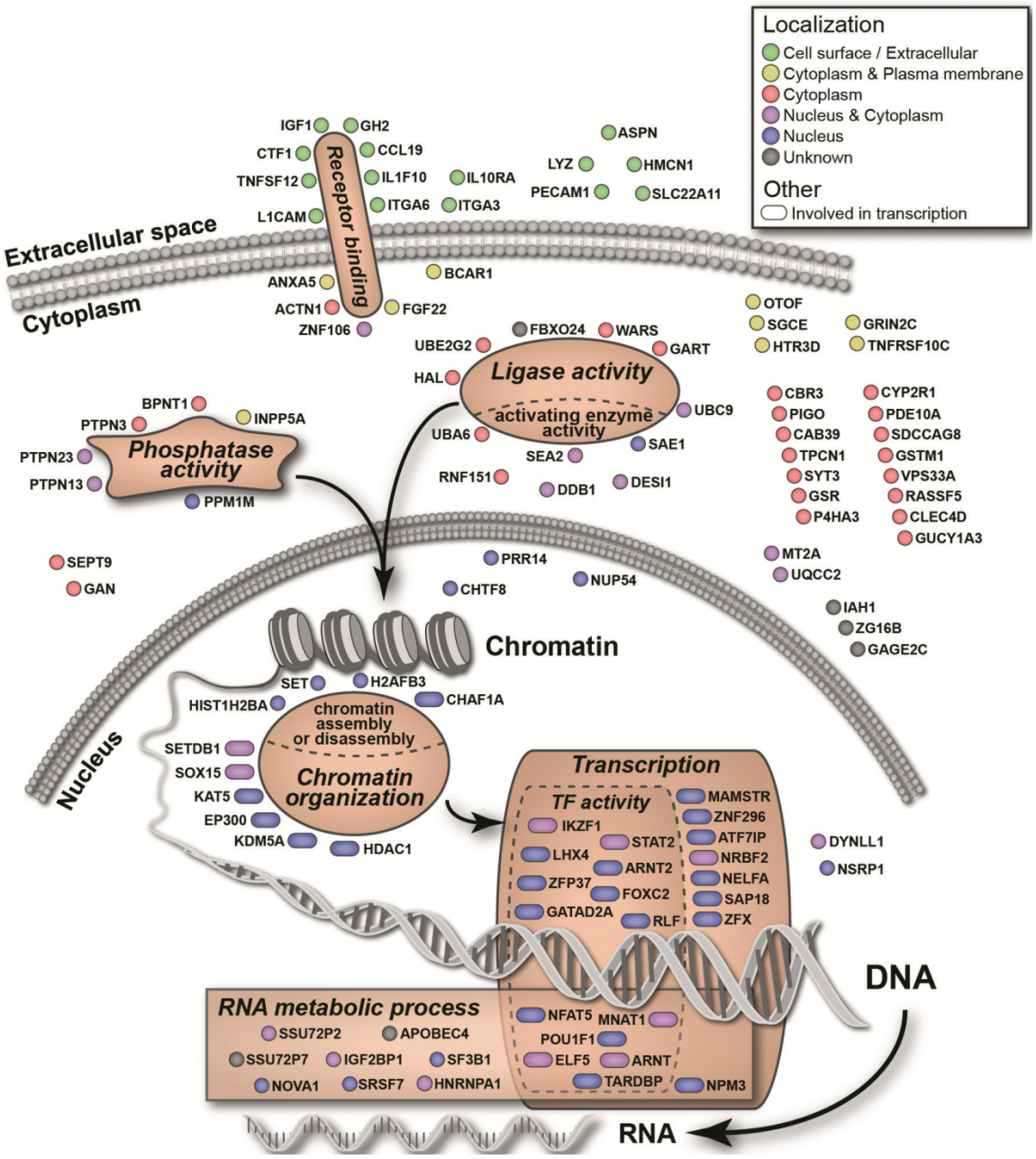

Figure S2. **A summary of factors detected in the screen, along with their membership in GO categories.** Shown is an expanded version of Figure 2 that includes factors from non-enriched GO categories. The figure shows 110 factors. The remaining 18 factors are designated as uncharacterized and are listed in Table S4.

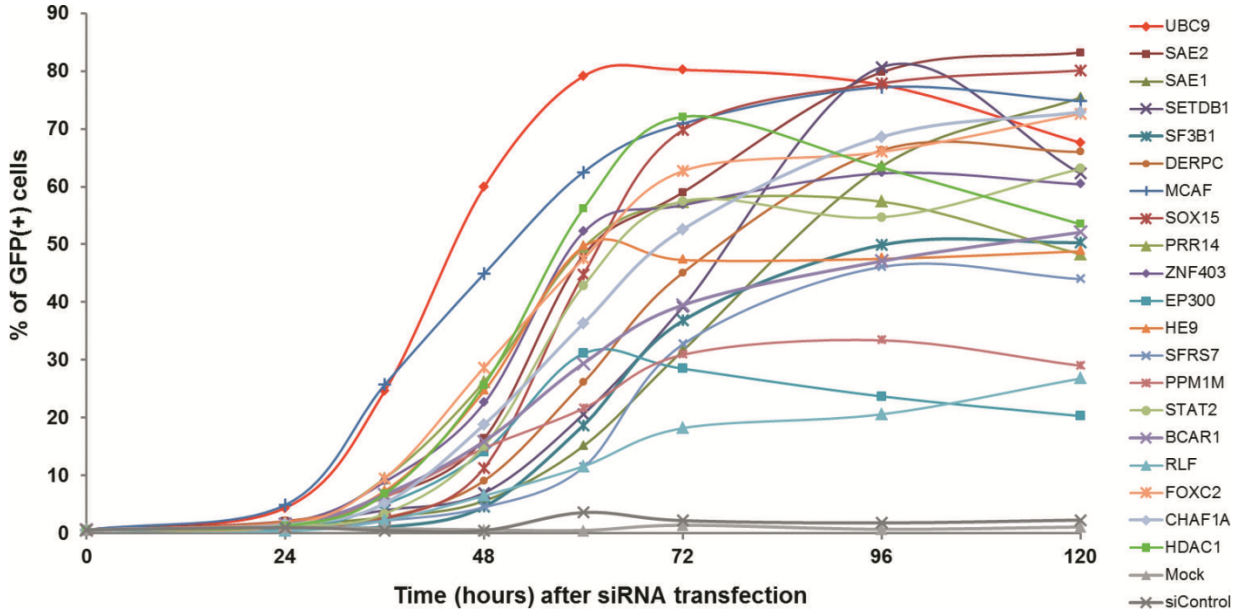

Figure S3. **A time course of GFP reporter gene reactivation.** The GFP-silent reporter cells were transfected with siRNAs and the percentage of GFP-positive cells were analyzed by flow cytometry at the indicated time points. siControl, negative control siRNA.

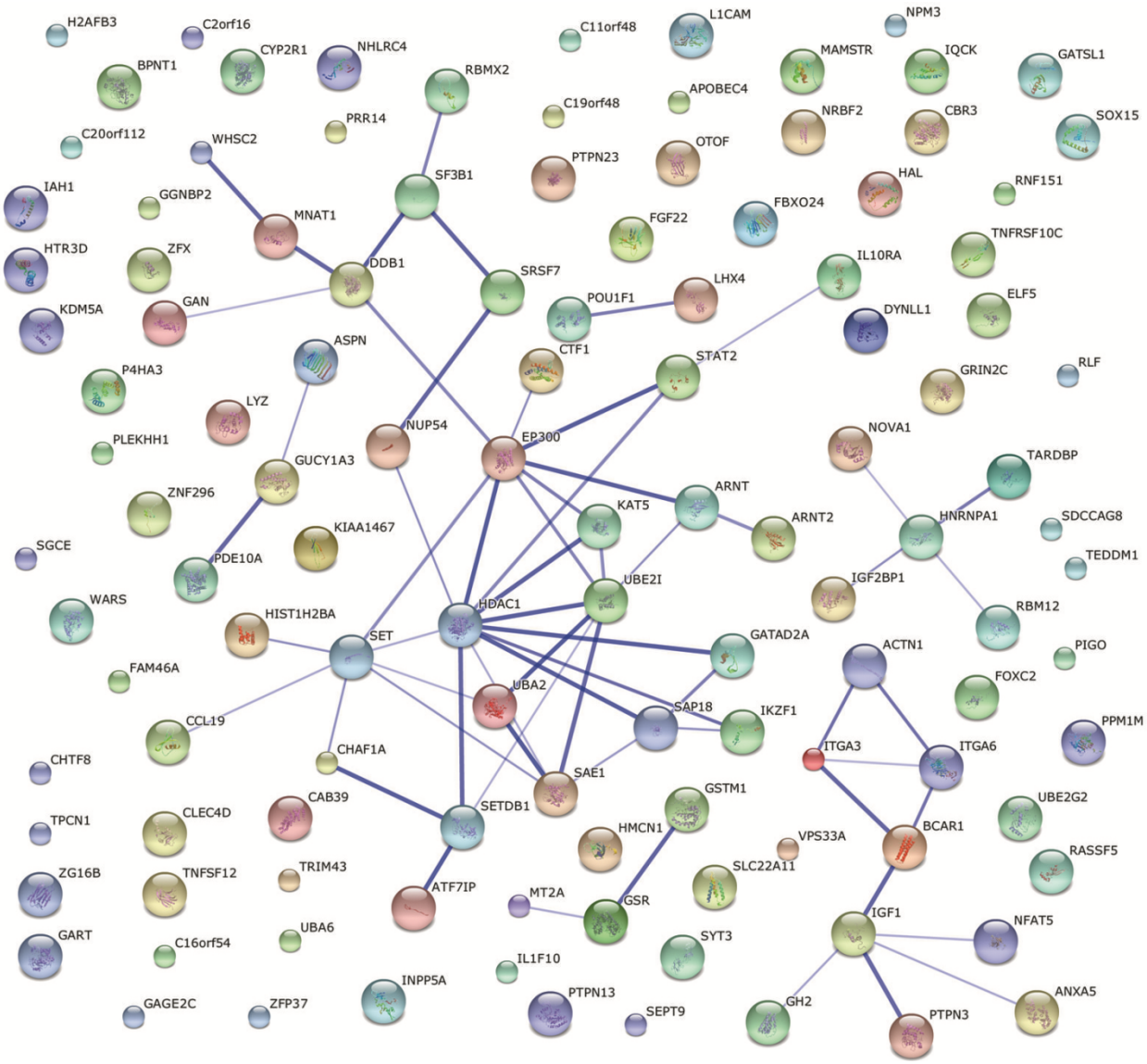

Figure S4. **EMBL String 9.1 analysis.** An unbiased analysis of the 128 factor set was carried out using EMBL String, a database of known and predicted protein interactions. Shown is a network display map. The output includes 119 factors. "Confidence View" is shown with stronger associations represented by thicker lines. All Active Prediction Methods were selected: Neighborhood, Gene Fusion, Co-occurrence, Co-expression, Experiments, Databases, and Textmining. The network display map detected a core set of chromatin factors with HDAC1 as a major node, the SUMO pathway, and the integrin-adhesome.

SUPPLEMENTARY MATERIALS AND METHODS

Design of the primary and validation siRNA screening assays

The primary screen was optimized to quantitate total GFP fluorescence intensity of individual wells in a 96-well plate format. The 96-well optical-bottom plates (165305, Nunc) were used. Each plate was washed twice with HBSS solution to eliminate serum and medium residues, and then analyzed. The GFP fluorescence intensity signal was measured using 2102 EnVision Plate Reader (Perkin Elmer). The primary screening was performed using a Dharmacon Genome-Wide siRNA SMARTpool library targeting 21,121 human genes. A 96-well sample plate contained 80 single siRNA SMARTpools along with eight positive and eight negative controls (alternating in rows 1 and 2). SiRNAs targeting GAPDH and HDAC1 were used as negative and positive control, respectively. Each plate was prepared and analyzed in duplicate. Results were calculated as a ratio of the sample GFP signal to average negative and average positive controls on each plate. The validation screen was performed using 96-well culture plates. A sample plate contained 20 deconvoluted siRNA pools, a total of 80 siRNAs (4 single siRNA from each SMARTpool), along with siRNA controls as described above. The percentage of GFP-positive cells and mean fluorescent intensity for each well was measured using a 96-well Guava EasyCyte Plus flow cytometer (Millipore-Guava). The data were scored as percent GFP-positive cells. Three independent experimental screening assays were carried out, each in duplicate (n=6).

Evaluation of assay sensitivity and reproducibility

To evaluate the efficiency and reproducibility of the screening assays, the Z'-factors were calculated [1]. The Z'-factor was measured in three independent experiments using the GFP reporter cell population. For this measurement, 96-well plates were used containing 48 positive control and 48 negative control siRNA treated wells, targeting HDAC1 and GAPDH, respectively. The Z'-factor was calculated $(1 - 3x(s_p + s_n)/(\mu_p - \mu_n))$, with the s_p and s_n being the standard deviation of positive and negative values and μ_p and μ_n the mean of

positive and negative values. The average Z'-factor was 0.58 and 0.79, for the primary and the validation screening assays, respectively.

qRT-PCR

HeLa and HeLa GFP reporter cells were transfected with siRNA as described above. AT 72 hours post transfection total RNA was isolated using RNAqueous-4PCR kit (AM1914, Ambion) according to manufacturer's protocol. cDNA were synthesized using SuperScript III Reverse Transcriptase kit (18080-044, Invitrogen). DNA detection was performed using quantitative Real-Time PCR with the KAPA SYBR FAST qPCR Kit (KK4600, KAPA Biosystems) according to manufacturer's recommendations. Quantifications were performed using ΔC_T method. Experimental data were normalized against GAPDH level and compared with siControl siRNA transfection. The heat map illustrating silent cellular gene reactivation created using color coding for demonstrating fold-difference relative to control siControl siRNA transfection with p-value < 0.05.

Primer sets:

RUNX3 – Forward: 5'-TCCTAGCAGAAGGCATTGAGGTGA

Reverse: 5'-AGGAGTCGCAAGATTTGGCTGGAT

PARK2 – Forward: 5'-TGCCTTCTGCCGGGAATGTAAAGA

Reverse: 5'-CCTGAGGCTTCAAATACGGCACT

HIC1 – Forward: 5'-TCGTGCGACAAGAGCTACAAGGA

Reverse: 5'-AACTTCTTCCCGCAGATGGTGCAT

TP73 – Forward: 5'-ATCCTCTTTGCTGATGGACTGCCA

Reverse: 5'-ACAGTCACTTGGCTGCTCACTACT

ID4 – Forward: 5'-TGTGCCTGCAGTGCGATATGAA

Reverse: 5'-TGCAGGATCTCCACTTTGCTGACT

LGI4 – Forward: 5'-TGATTGAGGACGATGCATTTGCGG

Reverse: 5'-TGGTTATTGGCCAGGCTTAGGTGT

FOXD3 – Forward: 5'-TGAAGCCGCCTTACTCGTACAT

Reverse: 5'-TGACGAAGCAGTCGTTGAGTGAGA

GLI1 – Forward: 5'-AAAGCCTTCAGCAATGCCAGTGAC

Reverse: 5'-TGCAGCCAGGGAGCTTACATACAT

GAPDH – Forward: 5'- TACTAGCGGTTTTACGGGCG

Reverse: 5'- TCGAACAGGAGGAGCAGAGAGCGA

EGFP – Forward: 5'- TCTTCTTCAAGGACGACGGCAACT

Reverse: 5'- TTGATGCCGTTCTTCTGCTTGTCG

Supplementary Information References

1. Zhang JH, Chung TD, Oldenburg KR (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen 4: 67-73.