Lawrence Berkeley National Laboratory

Biological Systems & Engineering

Title

Identification of Genes Directly Regulated by the Oncogene ZNF217 Using Chromatin Immunoprecipitation (ChIP)-Chip Assays*

Permalink

https://escholarship.org/uc/item/3r07f40q

Journal

Journal of Biological Chemistry, 282(13)

ISSN

0021-9258

Authors

Krig, Sheryl R
Jin, Victor X
Bieda, Mark C
et al.

Publication Date

2007-03-01

DOI

10.1074/jbc.m611752200

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

JBC Papers in Press. Published on January 26, 2007 as Manuscript M611752200
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M611752200

Identification of genes directly regulated by the oncogene ZNF217 using ChIP-chip assays

Sheryl R. Krig[1#], Victor X. Jin[1#], Mark C.Bieda[1], Henriette O'Geen[1], Paul Yaswen[3], Roland Green[2], and Peggy J. Farnham[1]*

1) Department of Pharmacology and the Genome Center, University of California-Davis, Davis, CA 95616; 2) NimbleGen Systems Inc, Madison, WI; 3) Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; #) co-first authors

running title: Identification of ZNF217 target genes

*Address correspondance to: Peggy J. Farnham, Department of Pharmacology and the Genome Center, University of California-Davis, One Shields Avenue, Davis, CA 95616; Tel (530) 754-4988; Fax (530) 752-0436; pjfarnham@ucdavis.edu

It has been proposed that ZNF217, which is amplified at 20q13 in various tumors, plays a key role during neoplastic transformation. ZNF217 has been purified in complexes that contain repressor proteins such as CtBP2, suggesting that it acts as a transcriptional repressor. However, the function of ZNF217 has not been well characterized due to a lack of known target genes. Using a global ChIP-chip approach, we have identified thousands of ZNF217 binding sites in three tumor cell lines (MCF7, SW480, and Ntera2). Further analysis of ZNF217 in Ntera2 cells has shown that many promoters are bound by ZNF217 and CtBP2, and that a subset of these promoters are activated upon removal of ZNF217. Thus, our in vivo studies corroborate the *in vitro* biochemical analyses of ZNF217-containing complexes and support the hypothesis that ZNF217 functions as a transcriptional repressor. Gene ontology analysis shows that ZNF217 targets in Ntera2 cells are involved in organ development, suggesting that one function of ZNF217 may be to repress differentiation. Accordingly, we show that differentiation of Ntera2 cells with retinoic acid leads to down-regulation of ZNF217. Our identification of thousands of ZNF217 target genes will enable further studies of the consequences of aberrant expression of ZNF217 during neoplastic transformation.

Amplification at 20q13 occurs in a variety of tumor types, such as breast (1), gastric (2), ovarian (3), lung (4), prostate (5), and colon (6), and is associated with aggressive tumor behavior (7). The mapping of the amplified region at 20q13.2 led to the positional cloning and characterization of ZNF217 (1), which is considered to be one of the driver genes at 20q13.2, promoting selection during the early stages of tumor development. Initial comparative genomic hybridization (CGH) studies showed that ZNF217 is amplified and over-expressed in ~40% of breast cancer cell lines and 18% of primary breast tumors (8). Further CGH studies of various tumor specimens report that ZNF217 amplification and overexpression at the 20q13 locus can display tumor type-specific profiles. For example, an analysis of 22 sporadic colorectal carcinomas detected DNA copy number changes for ZNF217 in 45% of the CIN-type (chromosomal sCRC) but not the MIN-type (microsatelite sCRC) colon tumors (6). Evidence in support of a causal role for ZNF217 in tumor formation comes from studies using normal human mammary epithelial cells (HMECs). Nonet et al. showed that introduction of ZNF217 into early passage HMECs can lead to a rare event of immortalization (9). It has been proposed that overexpression of ZNF217 may give a selective advantage to tumor cells by interfering with pathways associated with normal regulation of cell growth, cell death, differentiation, or DNA repair.

DNA sequence analysis suggests that ZNF217 encodes a transcription factor having eight C2H2 Kruppel-like zinc finger DNA-binding motifs and a proline-rich transactivation domain at the C-terminus (1). Biochemical studies support a role for ZNF217 in transcriptional regulation. For example, ZNF217 has been identified in complexes that contain repressor proteins such as CtBP and coREST (10,11), histone deacetylases, the histone methyltransferase G9a, and the histone demethylase LSD1 (11-14). The direct

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

Copyright 2007 by The American Society for Biochemistry and Molecular Biology, Inc.

interaction of ZNF217 with CtBP (15) suggests that ZNF217 could be recruited to a variety of transcription complexes through the interaction of CtBP with numerous site-specific DNA binding proteins (16).

Although both biochemical and structural studies have linked ZNF217 to transcriptional regulation, a detailed analysis of its role in transcription has been limited due to a lack of known ZNF217 target genes. Therefore, we have used a ChIP-chip method to identify thousands of ZNF217 target genes in 3 cancer cell lines; the breast cancer line MCF7, the colon cancer line SW480, and Ntera2, a teratocarcinoma line that can differentiate into neurons. To investigate the role of ZNF217 in transcriptional regulation, we have examined the expression level of ZNF217 target genes in Ntera2 cells before and after reduction of the levels of ZNF217 using siRNAs and have examined colocalization of ZNF217 with CtBP family members using ChIP-chip assays. Gene ontology analysis indicates that some ZNF217 target genes in Ntera2 cells are transcription factors that are involved in cell differentiation and organ development. We show that ZNF217 is down regulated upon treatment of Ntera2 cells with retinoic acid, suggesting that the inappropriate expression of ZNF217 in differentiated adult cells may suppress differentiation, leading to tumorigenesis.

Experimental Procedures:

Cell Culture. SW480 cells were grown in McCoy's 5A modified medium (Invitrogen), supplemented with 10% FBS (NovaTech) and 1% Penicillin/Streptomycin (Invitrogen). MCF7 and Ntera2 cells were grown in Dulbecco's Modified Eagle Medium supplemented with 2mM glutamine, 1% Penicillin/Streptomycin and 10% FBS. All cells were incubated at 37°C in a humidified 5% CO_2 incubator. For ZNF217 knockdown ChIP assays, ZNF217 siRNA (SMARTpool; Dharmacon, cat# M-004987-00) or si-GLO RISC-Free (Dharmacon, cat# D-001600-01) as a non-specific control, was transiently transfected into Ntera2 cells (100nM) plated on 100mm dishes. Transfections were carried out using Invitrogen Lipofectamine2000 according to manufacture recommendations. After 72 hours, cells were replated at 30-50% density for re-transfection and harvested after another 72 hours. For ZNF217 knockdown RNA analysis, Ntera2 cells were transfected with the siRNAs (100nM) in 6-well dishes and RNA harvested at 72h (experiments A and B) or re-transfected and harvested 48 hour later (experiment C); RNA was prepared using a Qiagen RNA easy kit (cat#74104). Ntera2 cells were differentiated using 10^{-5} M retinoic acid (Sigma). Antibodies used on western blots were anti-ZNF217 polyclonal (17) and anti-Oct4 goat polyclonal (Santa Cruz #8628 N-19).

ChIP Assays and Amplicon Preparation. ChIP assays were performed as previously described (18) with minor modifications. A complete protocol can be found on our website at http://genomics.ucdavis.edu/farnham/ and in Oberley et al. (19). Antibodies used in this study include a ZNF217 rabbit polyclonal that was generated using a GST-peptide sequence (17). Immunosera was purified on Pierce Aminolink peptide column constructed with the fusion peptide sequence. CtBP1 (cat# 612043) and CtBP2 (#612044) antibodies were purchased from BD Transductions. The secondary rabbit anti-mouse IgG (cat# 55436) was purchased from MP Biomedicals. Standard PCR reactions using 2 uls of the immunoprecipitated DNA were performed. PCR products were separated by electrophoresis through 1.5% agarose gels and visualized by ethidium bromide intercalation. Amplicons, prepared using 50-80% of a ChIP sample, were generated using Sigma's Whole Genome Amplification Kit; see our published ChIP protocol (20) and http://genomics.ucdavis.edu/farnham/ for details).

ChIP-chip Assays. ENCODE and promoter arrays were produced by NimbleGen Systems, Inc (Madison, WI). The 5kb human promoter array design is a two-array set, containing 5.0 kb of each promoter region (from build HG17) that extends 4.2 kb upstream and 800 bp downstream of the TSS. Where individual 5.0 kb regions overlap, they are merged into a single larger region, preventing redundancy of coverage. The promoter regions thus range in size from 5.0 kb to 50 kb. These regions are tiled at a 110 bp interval, using variable length probes with a target Tm of 76°. NimbleGen ENCODE oligonucleotide arrays contained ~380,000 50mer probes per array, tiled every 38 bp. The regions included on the arrays encompassed the 30 MB of the repeat masked ENCODE sequences, representing approximately 1% of the human genome. All NimbleGen arrays were hybridized and the data were extracted according to standard operating procedures by NimbleGen Systems Inc. Signal Map software provide by NimblGen was used to visualize the array peaks.

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

ASBMB The Journal of Biological Chemistry jbc ONLINE

PCR analysis of target genes. Primers were designed to flank ZNF217 binding sites identified on the arrays. PCR was performed using low cycle numbers to ensure linear amplification of the input. BioRad's Quantity One software was used for quantification of gene enrichment over input DNA.

RNA Illumina expression arrays. RNA samples were harvested using Qiagen's RNAeasy Kit (Qiagen) and then assayed using the Agilent Systems Bioanalyzer to ensure that high quality RNA was used for the array experiments. The Illumina TotalPrep RNA amplification kit from Ambion (AMIL1791) was used to generate biotinylated, amplified RNA for hybridization with the Illumina Sentrix Expression Beadchips, Human 6-v1. The Sentrix gene expression beadchips used for this study consisted of a 6-array, 2 stripe format comprising approximately 48k probes/array. In this collection 24,000 probes were from refseq sequences and 24,000 from other genbank sequences (see (http://www.illumina.com/products/arraysreagents/gwe_human6.ilmn for more details). Arrays were processed as per manufacturer's instructions, scanned at medium PMT settings as recommended by the manufacturer, and analyzed using Bead Studio Software v. 2.3.41. Data was normalized using the "average" method which simply adjusts the intensities of two populations of gene expression values such that the means of the populations become equal. Differential expression was calculated for the control versus siZNF217 data sets using an algorithm provided by Bead Studio. Fold-enrichment values were used to obtain the list of candidates with greater than 1.5-fold change.

DAVID analysis. Functional annotations were performed using the program DAVID 2.1 (21) see also http://apps1.niaid.nih.gov/david/). The same parameters were used for all analyses presented in this study. These parameters were Gene Ontology (GO) Molecular Function term, level 2; Interpro name in the Protein Domains section; and SP_PIR_Keywords in the Functional Categories section. Before performing the analyses, hypothetical genes and genes with no known function were removed from the list. After performing the analyses, all categories that represented less than 4% of the total number of genes were eliminated. In addition, redundant terms (e.g. transcriptional regulation and transcription factor activity) and non-informative terms (e.g. multi-gene family) were also eliminated.

Peak finding: For identification of the ZNF217 binding sites on the ENCODE arrays, we used the Tamalpais program described in Bieda et al., 2006 (22) and chose the L1 set of high confidence peaks for further analysis. Briefly, these binding sites are identified as peaks that have a minimum of 6 consecutive probes in the top 2% of all probes on the array. Only those peaks that were identified in at least two out of three biological replicates were considered to be binding sites. The human 5 kb promoter array consists of a set of two arrays encompassing a total of ~24,000 human promoters. We refer to the two arrays in the set as "promoter 1" and "promoter 2" arrays. Most promoters encompass ~5 Kb of genomic sequence tiled with a 50 mer probe every ~100 nts. Given that amplicons were ~300-500 nts in length, a binding site should be the center of a hill waveform of ~600-1000 nts in total length, for an average of 6-10 probes per binding sites. Given that the probes on the "edges" of the hill will have little amplification, we reasoned that the center four probes (~400 nts) should be well-enriched above background. Hence, we assigned a value to each promoter based on the highest mean of 4 consecutive probes, a procedure we termed "Maxfour." To calculate these "Maxfour values", custom software was written in Perl and bash shell (Bieda et al, manuscript in preparation). A small portion of promoters were represented by <4 probes, and hence no Maxfour value was possible; we considered analysis of these promoters to be unreliable and these promoters are assigned "-100" as a value. Further statistical exploration of this procedure will be presented elsewhere (Bieda et al., in preparation).

RESULTS

Location analysis of ZNF217 binding in the human genome. Although ZNF217 is predicted to contain eight zinc fingers and thus is thought to be a DNA binding protein, its role in transcriptional regulation has not been well-characterized due to a lack of known target genes. Using the ChIP-chip approach, previous studies have identified target genes of transcription factors in a global and unbiased manner (23,24). Commonly used platforms for ChIP-chip are arrays that contain CpG islands (25), core promoters, or 5-10 kb of upstream promoter sequences (26). However, we did not know if ZNF217 binds to CpG islands or near to transcription start sites. Therefore, to initiate our studies we first needed to perform an unbiased location analysis of ZNF217

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry ASBMB jbc

binding. For these ChIP-chip analyses, we used ENCODE oligonucleotide arrays that represent 1% of the human genome and include ~400 genes and intergenic regions (see **Methods** for details). Using a rabbit polyclonal antibody to ZNF217, we performed ChIP assays in three different human cancer cell lines; the MCF7 breast cancer line, the SW480 colon cancer line, and the teratocarcinoma cell line Ntera2. Our classification of a peak as a binding site on ENCODE arrays requires that the region be bound in at least two of three independent experiments (see (22)). Therefore, three biological replicate ChIP samples from each of the 3 cell lines were hybridized to ENCODE arrays (see Supplemental Table S1). After array normalization, the ZNF217 hybridization signals were divided by the total input signals to provide a fold-enrichment value for each 50 bp oligomer on the array. Sites that were bound by ZNF217 were identified with the Tamalpais Peaks program (22) at the L1 level (P< 0.0001); this requires that at least 6 oligos in a row be in the top 2% of all probes on the array. We identified a total of 61 ZNF217 binding sites in SW480 cells, 175 binding sites in MCF7 cells, and 178 binding sites in Ntera2 cells. To determine the location of the binding sites relative to the nearest gene, we used the Gencode Database (27). This analysis identified 44 genes in SW480 cells, 103 genes in MCF7 cells, and 101 genes in Ntera2 cells. We found that a significant percentage of the binding sites (39% in SW480, 41% in MCF7 and 49% in Ntera2) fell within 2 Kb upstream or downstream of the transcription start site of the nearest gene, although binding was also observed in the regions greater than 2 Kb upstream from the start site as well as within genes and in intergenic regions **(Figure 1A)**. Although most of the binding sites covered about 500 bp to 1 kb, there was one region on the ENCODE array of more than 150 kb that was bound by ZNF217 in Ntera2 cells (**Figure 1B**). Because these experiments are the first to identify genomic binding sites of ZNF217 using ChIP-chip assays, we thought it was important to demonstrate the specificity of the ZNF217 antibody. Therefore, we treated Ntera2 cells with siRNAs to ZNF217 (see **Figure 7** for a western blot indicating the degree to which ZNF217 protein can be reduced by the siRNA treatment) and then performed ChIP experiments using the ZNF217 antibody in the control vs the knock-down cells. Amplicons prepared from these ChIP samples were hybridized to ENCODE arrays. As shown in **Figure 1B**, ZNF217 binding throughout the HOXA gene cluster on chromosome 7 is greatly reduced in the cells treated with siRNAs to ZNF217. Thus, because reduction of ZNF217 RNA reduced the signals obtained in the ChIP assay, we are confident that we are identifying bona fide ZNF217 target genes.

***De Novo* Identification of a Putative ZNF217 Binding Site Motif.** The peaks that we identified above represent the first collection of *in vivo* ZNF217 binding sites. It is not yet known if ZNF217 is directly bound to the DNA at each of these sites or if it is recruited to the sites indirectly via interaction with partners such as CtBP. Because CtBP has been shown to bind to numerous DNA binding proteins, an indirect recruitment mechanism might result in the identification of multiple motifs in the collection of binding sites. To search for common motifs, we used a *de novo* motifs discovery approach termed ChIPMotifs (28). Briefly, the ChIPMotifs approach incorporates a statistical bootstrap re-sampling method to identify the top motifs detected from a set of ChIP-chip training data using *ab initio* motif-finding programs such as Weeder (29) and MEME (30). To obtain a training dataset, we first selected the common ZNF217 binding sites identified using the ENCODE arrays from all three cell lines. This provided a set of 53 very high confidence ZNF217 binding regions (each region was identified as a binding site in at least 2 of 3 biologically independent experiments in all 3 of the different cell types). A set of 506 sequences from the ENCODE regions of 500 bp in length that did not bind to ZNF217 was selected as a negative control dataset. After applying the ChIPMotifs approach to these training sets, we identified the motif shown in **Figure 2**. An eight-base consensus is defined as **ATTCCNAC** (reverse-complement counterpart is **GTNGGAAT, Figure 2A**) with a five-base core consensus as **ATTCC** (reverse-complement counterpart is **GGAAT)**. The scores cutoff for this motif positional weight matrix (ZNF217_PWM, **Figure 2B)** built from the ChIPMotifs is determined as 1.00 for core score and 0.86 for PWM score with a significant fisher *p*-value of 0.001 determined by our ChIPMotifs approach. Using these cutoff scores, 47% (25 out of 53) of the ZNF217 binding regions included this motif whereas 81% of the regions in the negative control data set lacked the motif. After finding a motif common to the ZNF217 binding sites of all three cell lines, we wanted to know if we would identify the same or different motifs if we analyzed the three cell lines individually. An AP1-like motif **ANGAGTCA**

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry ASBMB jbc ONLINE

was identified in MCF-7 cells with a significant *p*-value of 0.009 for the cutoff of 1.0 for core score and 0.86 for PWM score, and a core-binding motif **CATTCC** was identified in SW480 cells with a *p*-value of 1.1 x 10^{-5} for the cutoff of 1.0 for core score and 0.85 for PWM score. The SW480 core sequence is much like the motif identified using the combined datasets. This may be due to the fact that the majority of the 61 binding sites identified in SW480 were included in the 53 common binding sites in the training dataset from all three cell lines. Unfortunately we are unable to identify any significant motif using only the Ntera2 binding sites; an E2F core consensus was initially identified by the *ab initio* programs but failed to pass the significance test using a bootstrap re-sampling statistical approach. It is likely that the "common motif" was not identified using the Ntera2 set because the common 53 targets were a small portion of the total number of ZNF217 binding sites in Ntera2 cells. A recent study used in vitro casting experiments to identify a 15 base consensus sequence bound by zinc fingers 6 and 7 of ZNF217 (31). Although we could not identify the full consensus site in any of the in vivo ZNF217 binding sites, the core sequence of CAGAAY was found in 64% of the MCF7, 50% of the Ntera2, and 54% of the SW480 sites identified by ChIP-chip.

Identification of ZNF217 target genes using promoter arrays. Having shown that ~40-50% of the ZNF217 binding sites identified using ENCODE arrays localize within 2 kb upstream or downstream of the transcription start site, we felt that the 5kb NimbleGen promoter arrays would be appropriate for the identification of a large set of ZNF217 target genes. Using this 2 array set, the region spanning 4.2kb upstream to 0.8kb downstream of the transcription start sites of 24,000 human promoters can be analyzed, with each promoter being represented by 50 probes spaced approximately 100 bp apart. We next performed ChIP-chip assays using the ZNF217 antibody and two independent cultures of Ntera2 cells (see Supplemental **Table S1**). After array normalization, the ZNF217 hybridization signals were divided by the total input signals to provide a fold-enrichment value for each 50-bp oligomer on the array. To identify binding sites on these promoter arrays, we developed a analysis program (termed "Maxfour") which ranks each of the promoters based on the highest average intensity value for 4 consecutive probes (which corresponds to ~ 400 nts). This ranking system was used for the analysis of all promoter arrays in this study. The Maxfour values from the two independent ZNF217 ChIP-chip experiments for the 14,000 promoters on promoter 1 array (the promoters are present on two different arrays termed promoter 1 and promoter 2 arrays) were aligned to examine the reproducibility of the ChIP assays. The correlation graph shown in **Figure 3** demonstrates a high overlap between the two Ntera2 ChIP samples (r = 0.865) (see supplemental **Figure S1** for the promoter 2 correlation graph) with only a few points deviating from a simple linear relationship. To confirm that the top-ranked promoters we identified using the ZNF217 antibody and the promoter arrays are in fact ZNF217 targets, we performed ZNF217 ChIP experiments before and after treatment of cells with siRNAs to ZNF217. We identified 2000 promoters that show an average enrichment of greater than 1.6 fold in the control cells (note that the values shown in Table 1 are on the log2 scale). Importantly, knockdown of ZNF217 mRNA by siRNA treatment caused a reduction of the average Maxfour values for the ZNF217 top-ranked targets to the level of ZNF217 binding at non target promoters (**Table 1**), confirming that the targets identified on the promoter arrays are indeed bound by ZNF217. The ZNF217 Maxfour enrichment values for all 24,000 promoters, before and after treatment of cells with the siRNAs to ZNF217, can be found in Supplemental **Table S2A** (for the promoter 1 arrays) and Supplemental **Table S2B** (for the promoter 2 arrays).

To determine if ZNF217 binding shows any cell-type specificity, we broadened our study to include the two "adult" cancer cell lines MCF7 and SW480. **Figure 4A** shows a representative comparison analysis graph between the MCF7 Maxfour values and Ntera2 replicate 1 for the promoter1 array. (See **Figure S1** in Supplemental data for additional graphs comparing MCF7, SW480, and Ntera2 ChIP-chip data). Using stringent parameters, we analyzed the Maxfour values to produce three categories of ZNF217 target promoters; promoters only bound in MCF7 cells, promoters only bound in Ntera2 cells, and promoters bound in both MCF7 and Ntera2 cells. For example, the MCF7-exclusive ZNF217 targets on promoter1 were selected to have greater than 1.0 (log2) Maxfour values in MCF7 cells and less than 0.4 Maxfour values in Ntera2 cells. Using the parameters described in the legend to **Figure 4A**, we identified 754 promoters that were bound in both cells types, 114 ZNF217 targets that were potentially MCF7-specific, and 452 ZNF217

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry ASBMB jbc

targets that were potentially only bound in Ntera2 cells (see **Table S3** in supplemental data for complete lists of the 3 categories of ZNF217 target genes). Although fewer targets were identified in SW480 cells, the top targets showed a high degree of overlap with the MCF7 bound promoters (see panel 7 in Supplemental **Figure S1**).

It was possible that some of the promoters that fell into the cell type-specific classes were in fact false negatives on one array (i.e. should have been classified as being bound in both cell types) or false positives on one array (i.e. do not show reproducible binding in subsequent experiments). To confirm the existence of cell-type specific ZNF217 gene targets we performed additional ChIP assays using biologically independent sets of cross-linked Ntera2 and MCF7 cells. Amplified chromatin from ZNF217 ChIP assays and input chromatin was analyzed by PCR (with low cycle numbers to ensure that the assay was within linear range) using primers specific to promoters in the "cell type-specific" classes of ZNF217 targets. As shown in **Figure 4B**, all 5 of the class of Ntera2-exclusive promoters show higher ZNF217 enrichment values in Ntera2 vs MCF7 cells. However, several of the tested MCF7-specific promoters do not confirm as ZNF217 targets.

To begin to understand the function of ZNF217, we analyzed the 3 categories of ZNF217 gene targets using the DAVID analysis program (21). This analysis program uses gene ontology to classify a list of genes and provides a measure of significance for the identified categories by including a P-value that indicates the probability that the identified category is more highly enriched in the target set than would be expected by random chance. Although the MCF7-exclusive list sample size is small and, based on the PCR confirmation results, may include false positive targets, we did see a slight enrichment for genes involved in organ development and morphogenesis **(Table 2)**. The DAVID analysis for the large set of ZNF217 targets bound in both MCF7 and Ntera2 cells showed enrichment for genes with transcription factor activity and ion binding. The Ntera2-exclusive gene list is highly enriched for transcription factor activity and organ development and moderately enriched for cell differentiation and receptor binding. We found that a very large percentage of the Ntera2-exclusive ZNF217 target genes were homeodomain transcription factors such as LHX2, SIX6, and SOX2. The binding of ZNF217 in Ntera2 cells, which are derived from germ cell tumors, to the promoters of developmental regulatory genes such as homeodomain proteins suggests that ZNF217 may play a role in regulating embryonic development. An alphabetized list of all the genes in the three different categories can be found in **Table S3**.

ZNF217 and CtBP co-localize at promoters. Biochemical purification studies have suggested that ZNF217 can be found in a complex with CtBP1 and CtBP2 together with histone-modifying enzymes such as histone deacetylases and histone methyltransferases. (13,14). As ZNF217 is thought to be a putative DNA binding protein, it has been proposed that ZNF217 functions in gene repression by recruiting CtBP and an associated repressor complex to DNA (15,31). To test this hypothesis, we investigated whether ZNF217 and CtBP proteins are bound to the same target promoters. We first performed ChIP assays in Ntera2 cells using antibodies to CtBP1 and CtBP2. As expected, the sets of promoters bound by these two proteins are very similar (**Table 3**). Interestingly, we found that the majority of the ZNF217 targets in Ntera2 cells are also bound by both CtBP1 and CtBP2 (**Table 3** and supplementary **Table S4**). Similarly, the majority of ZNF217 targets in MCF7 cells are also bound by CtBP2.

It is possible that ZNF217 binds to promoters via some or all of its zinc fingers and recruits CtBP1/2 to the promoter regions through the direct protein-protein interactions that have been previously characterized (15). Alternatively, other DNA binding factors may be the primary contact point between the promoters and the ZNF217/CtBP repressor complex (**Figure 5A)**. For example, previous studies have shown that CtBP proteins can interact with other site-specific DNA binding proteins such as the homeodomain zinc finger proteins ZEB1 and 2 (32-34), with the ETS family member NET (35), and with other DNA binding factors (16,36,37). It is also possible that ZNF217 and CtBP2 are recruited to the same promoters, but to different regions. However, the identical patterns of binding seen in the genomic tiling arrays showing binding of ZNF217 and CtBP2 to the same location in the promoter region do not support this 3rd model (**Figure 5B**). We have attempted to address the cause and effect relationship between recruitment of ZNF217 and the CtBP proteins to promoters by performing ChIP-chip assays using an antibody to CtBP2 in cells treated with siRNAs specific for ZNF217. We find that, in

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

general, binding of CtBP2 is not greatly reduced by removal of ZNF217 from the set of commonly bound promoters (several examples are shown in **Figure 5B)**. Only 8 of the top 497 CtBP2 targets of promoter 1 array and 43 of the top 481 targets of promoter 2 array showed a convincing reduction in binding of CtBP2 in the cells treated with siRNA to ZNF217 (the list of the 51 genes and the CtBP2 ChIP enrichment values before and after reduction of ZNF217 are shown in Supplemental **Table S5**). Thus, the interaction between ZNF217 and CtBP2 is not required for recruitment of CtBP2 to most promoter regions. Our studies are suggestive that the 2nd model shown in **Figure 5A** may be correct for many promoters.

ZNF217 can function as a transcriptional repressor. As indicated above, ZNF217 has been purified in repressor complexes and we have shown that ZNF217 and the co-repressor CtBP bind many of the same promoters. Taken together, this suggests that ZNF217 can repress target genes. To provide support for the hypothesis that ZNF217 is a transcriptional repressor, we have determined the RNA expression levels of the target genes. For this analysis, we compared two independent ZNF217 ChIP-chip assays using Ntera2 cells and chose a set of promoters that were ranked in the top 1000 targets on promoter 1 in both of the experiments and a set of promoters that were ranked in the top 1000 targets on promoter 2 array in both of the experiments. This produced a list of 1276 promoters (631 genes from promoter 1 and 645 genes from promoter 2). We then examined the RNA expression data for this set of ZNF217 target genes using Illumina expression arrays. We isolated RNA from two independent cultures of Ntera2 cells and performed expression arrays; the RNA expression values obtained from these two experiments were very reproducible and were averaged for the analysis of ZNF217 target genes. Of the 1276 ZNF217 targets, 1077 were represented on the Illumina arrays. We divided the genes into three groups; those having very low expression values (less than 300), those have very high expression values (greater than 1000), and moderately transcribed genes (300-1000). We found that 64% of the top-ranked Ntera2 targets have low RNA values (**Table 4**), with a cut-off of less that 300 (corresponding to a P-value less than 0.00065), supporting the hypothesis that binding of ZNF217 to a promoter region can often result in transcriptional repression. A DAVID gene ontology analysis of the low expressed ZNF217 targets (**Table 5**) again shows a great enrichment for transcription factor activity, organ development, cell differentiation and embryonic development. In contrast, the small set of ZNF217 target genes in the RNA>1000 category are only modestly enriched in morphogenesis. To test the possibility that this small set of genes may be false positives for ZNF217 binding, we conducted PCR analysis of 4 potential targets: N-MYC, HAND1, TDGF3, and ZIC2 (**Figure 6**) using an independent ChIP samples from Ntera2 cells; for comparison, four "repressed" ZNF217 targets (USP32, CITED, ZHX2 and c-MYC) are also shown. Although the ZNF217 targets that were in the low expression category confirmed to be bound by ZNF217, only one of the ZNF217 targets in the high expression category (N-MYC) was confirmed to be bound by ZNF217. Therefore, the addition of RNA expression information may aid in identifying a robust list of ZNF217 target genes.

The low expression levels of the majority of the ZNF217-bound promoters suggested that binding of ZNF217 can result in transcriptional repression. To provide further support for this hypothesis, we analyzed RNA expression levels before and after reduction of ZNF217 by siRNA treatment. We performed triplicate siRNA treatments, prepared RNA samples from the control and treated cells, and hybridized the samples to Illumina expression arrays. As expected, ZNF217 RNA levels were shown to be substantially diminished in all three experiments, demonstrating the effectiveness of the siRNA treatment and the reliability of the expression arrays. In each of the three biological replicate experiments, several hundred genes were either up or down-regulated after reduction of ZNF217 levels (**Table 6**). For example, 288 genes were up-regulated (suggesting that they are normally repressed by ZNF217) in at least two of the three experiments, with 106 genes being up-regulated by reduction of ZNF217 in three independent siRNA experiments. Fewer genes were down-regulated by reduction of ZNF217, with only 29 showing consistent down-regulation in all three experiments. When we compared the list of promoters directly bound by ZNF217 and the list of genes that showed a change in expression in at least 2 of the 3 knockdown experiments, we found a total of 56 genes; the 54 genes having a known function are shown in **Table 7**.

The genes identified as being repressed by ZNF are highly enriched for neural-specific genes

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

(**Table S6**). Ntera2 cells possess characteristics of pluripotent embryonal cells and can differentiate into post-mitotic CNS neurons and a variety of other cell types upon treatment with retinoic acid (38), suggesting that one role of ZNF217 may be to assist in repressing differentiation pathways**.** If so, then one would predict that ZNF217 would be down-regulated upon retinoic acid-induced differentiation of Ntera2 cells. To test this hypothesis, we treated Ntera2 cells with retinoic acid and harvested samples at 4, 10, and 14 days. As a control for differentiation, we analyzed levels of OCT4, which is known to be down-regulated upon differentiation of Ntera2 cells (39). To ensure that we were detecting ZNF217 protein on the Western blot, we also treated Ntera2 cells (in the absence of retinoic acid) with siRNAs specific for ZNF217. As shown in **Figure 7**, ZNF217 is down-regulated by Day 4 of retinoic acid treatment to the same reduced protein level as obtained by treatment with siRNAs. The reduction in ZNF217 levels after differentiation of Ntera2 cells along the neural cell lineage supports the hypothesis that ZNF217 may be critical in repressing neural-specific genes. Future studies will be focused on determining if forced over-expression of ZNF217 can prevent differentiation in Ntera2 cells.

Discussion:

Using the global and unbiased approach of ChIP-chip assays that interrogate ~24,000 human promoters, we have gained new insight into the function of the oncogene ZNF217. We began by determining that ~half of the *in vivo* binding sites of ZNF217 are located within proximal promoter regions, and then identified thousands of promoters bound by ZNF217 in three different cancer cell lines. We have established that ZNF217 can function as a transcriptional repressor by demonstrating a) that the majority of ZNF217 target genes are bound by CtBP1/2, b) that most genes bound by ZNF217 show very low expression levels, and c) that reduction in the amount of ZNF217 bound to a promoter can, in some cases, result in increased gene expression. Interestingly, we find that many of the genes bound by ZNF217 in Ntera2 cells function in the neural cell lineage, suggesting that one role of ZNF217 may be to repress specific differentiation pathways.

Although inappropriate expression of ZNF217 has been linked to tumorigenesis (due to amplification at the 20q13 locus in multiple tumor types), the exact mechanisms by which ZNF217 might promote or enhance neoplastic transformation have not been elucidated. It has been proposed that ZNF217 functions as a transcriptional repressor due to its purification in complexes that contain co-repressors such as CtBP2 and co-REST (10,11,13,14) and histone modifying enzymes such as G9a and LSD1 (12,13). However, an understanding of how ZNF217-mediated repression might influence tumorigenesis has not been developed, in large part due to a lack of known ZNF217 target genes. The thousands of ZNF217 target genes that we have now identified provide an excellent data set for testing models of ZNF217-mediated gene regulation. For example, our preliminary analyses of histone modifications of ZNF217 target genes indicates that very low levels of H3me3K9 are found on the target promoters but that about half of the promoters repressed by ZNF217 have high levels of H3me3K27 (see **Table S6**). These results suggest that, in some cases, ZNF217 may recruit the PRC2 complex (40). Further studies examining other histone modifications are in progress.

Our studies support the previous biochemical purification experiments in that we demonstrate that there is a very large overlap between the promoters bound by ZNF217 and the promoters bound by CtBP1/2 in both Ntera2 and MCF7 cancer cells. It is possible that ZNF217 binds to promoters via some or all of its zinc fingers and recruits CtBP1/2 to the promoter regions through the direct protein-protein interactions that have been previously characterized (15). Alternatively, other DNA binding factors may be the primary contact point between the promoters and the ZNF217/CtBP repressor complex (16,36,37). Our data showing very similar binding patterns of ZNF217 and CtBP2 and our finding that, in general, binding of CtBP2 is not greatly reduced by removal of ZNF217 from a promoter, suggest that ZNF217 and CtBP2 might both be recruited to the chromatin via a different DNA binding protein. This hypothesis is supported by a motif present in many of the ZNF217 binding sites that is distinct from the motif recently identified using *in vitro* casting experiments and zinc fingers 6 and 7 of ZNF217 (31). Although the 8-base motif that we identified does not match any known binding motifs in the TRANSFAC or JASPAR databases, 3 positions of the motif (**TCC** or reverse-complement **GGA**) do match the conserved GGA that constitutes the core of the DNA binding motif of all ETS family members (41). Thus, it is possible that an ETS family member may help to recruit the ZNF217/CtBP complex to the DNA. Interestingly,

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

previous studies have shown that CtBP can interact with ELK-3 (also called Net, Sap-2, and Erp), an Ets family member (35).

Our finding that thousands of promoters are bound by ZNF217 is consistent with recent ChIP-chip studies of other human transcription factors, such as Men1 (42), Myc (43), E2F1 (22), NFkB (44), and the estrogen receptor (45). Further characterization of these target promoters have revealed that changes in binding of the factor does not necessarily lead to changes in gene expression (42,44). For example, ~2000 promoters are bound by the transcriptional activator Men1, but only ~5% of these genes showed a change in expression in mice that were nullizygous for the Men1 gene (42). Similarly, we found that ~5% of the genes identified as ZNF217 targets using ChIP-chip assays showed a change in expression when ZNF217 levels were reduced by siRNA treatment of Ntera2 cells. It is becoming increasingly clear that promoters are regulated by many factors and that loss of a single factor is usually not sufficient to alter the regulation of a promoter. Once the landscape of transcriptional regulation in the human genome has been mapped in more detail, sets of promoters that are commonly regulated by two or more different complexes may be identified. At that point, it may be possible to alter the expression of a greater percentage of genes by removing multiple complexes from the promoter regions.

A major goal of our studies of ZNF217 is to obtain insight into how inappropriate expression of this factor in human cancers can contribute to neoplasia. Interestingly, many of the genes that are bound by ZNF217 encode proteins that are critical in mediating differentiation. We also have shown that ZNF217 levels are down-regulated upon forced differentiation of Ntera2 cells, allowing for the expression of genes involved in differentiation and organogenesis. Thus, we propose that inappropriate expression of ZNF217 may lead to a down-regulation of genes that confer a differentiated phenotype, causing a de-differentiation of the cells and driving them towards a more proliferative and pluripotent phenotype. Future studies will be focused on comparing differentiation potential of cells expressing different levels of ZNF217.

Acknowledgements: RNA was labeled and the Illumina beadchip was processed and scanned at the UC Davis Genome Center Expression Analysis Core Facility. This work was supported in part by Public Health Service grant CA45250, HG003129, and DK067889 and support for SRK was from a California Breast Cancer Research Program 11FB-0029 postdoctoral fellowship. We also thank the members of the Farnham lab for helpful discussion and Sharon Squazzo and Sushma Iyengar for sharing unpublished data.

References

1. Collins, C., Rommens, J. M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S. I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., Jay, K. E., Froula, J., Cloutier, T., Kuo, W. L., Yaswen, P., Dairkee, S., Giovanola, J., Hutchinson, G. B., Isola, J., Kallioniemi, O. P., Palazzolo, M., Martin, C., Ericsson, C., Pinkel, D., Albertson, D., Li, W. B., and Gray, J. W. (1998) *Proc Natl Acad Sci U S A* **95**, 8703-8708
2. Yang, S. H., Seo, M. Y., Jeong, H. J., Jeung, H. C., Shin, J., Kim, S. C., Noh, S. H., Chung, H. C., and Rha, S. Y. (2005) *Clin Cancer Res* **11**, 612-620
3. Iwabuchi, H., Sakamoto, M., Sakunaga, H., Ma, Y. Y., Carcangiu, M. L., Pinkel, D., Yang-Feng, T. L., and Gray, J. W. (1995) *Cancer Res* **55**, 6172-6180
4. Zhu, H., Lam, D. C., Han, K. C., Tin, V. P., Suen, W. S., Wang, E., Lam, W. K., Cai, W. W., Chung, L. P., and Wong, M. P. (2006) *Cancer Lett* **in press**
5. Bar-Shira, A., Pinthus, J. H., Rozovsky, U., Goldstein, M., Sellers, W. R., Yaron, Y., Eshhar, Z., and Orr-Urtreger, A. (2002) *Cancer Res* **62**, 6803-6807
6. Lassmann, S., Weis, R., Makowiec, F., Roth, J., Danciu, M., Hopt, U., and Werner, M. (2006) *J Mol Med* **in press**
7. Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Holli, K., Collins, C., Kowbel, D., Gray, J. W., Kallioniemi, O. P., and Isola, J. (1995) *Clin Cancer Res* **1**, 1455-1461

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry ASBMB jbc

8. Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W., and Waldman, F. M. (1994) *Proc Natl Acad Sci U S A* **91**, 2156-2160
9. Nonet, G. H., Stampfer, M. R., Chin, K., Gray, J. W., Collins, C. C., and Yaswen, P. (2001) *Cancer Res* **61**, 1250-1254
10. You, A., Tong, J. K., Grozinger, C. M., and Schreiber, S. L. (2001) *Proc Natl Acad Sci U S A* **98**, 1454-1458
11. Lee, M. G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005) *Nature* **437**, 432-435
12. Hakimi, M. A., Dong, Y., Lane, W. S., Speicher, D. W., and Shiekhattar, R. (2003) *J Biol Chem* **278**, 7234-7239
13. Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005) *Mol Cell* **19**, 857-864
14. Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P., Nakatani, Y., and Shi, Y. (2003) *Nature* **422**, 735-738
15. Quinlan, K. G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M., and Crossley, M. (2006) *Mol Cell Biol* **26**, 8159-8172
16. Chinnadurai, G. (2002) *Mol. Cell.* **9**, 213-224
17. Huang, G., Krig, S., Kowbel, D., Xu, H., Hyun, B., Volik, S., Feuerstein, B., Mills, G. B., Stokoe, D., Yaswen, P., and Collins, C. (2005) *Human Mol. Genetics* **14**, 3219-3225
18. Weinmann, A. S., Bartley, S. M., Zhang, M. Q., Zhang, T., and Farnham, P. J. (2001) *Mol. Cell. Biol.* **21**, 6820-6832
19. Oberley, M. J., Tsao, J., Yau, P., and Farnham, P. J. (2004) *Methods in Enzymol.* **376**, 316-335
20. O'Geen, H., Nicolet, C. M., Blahnik, K., Green, R., and Farnham, P. J. (2006) *Biotechniques* **41**, 577-580
21. Dennis, G. J., Sherman, B. T., Hosack, D. a., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) *Genome Biol* **4**, P3
22. Bieda, M., Xu, X., Singer, M., Green, R., and Farnham, P. J. (2006) *Genome Research* **16**, 595-605
23. Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T. L., Wilson, C. J., Bell, S. P., and Young, R. A. (2000) *Science* **290**, 2306-2309
24. Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H.-M., and Farnham, P. J. (2002) *Genes & Dev.* **16**, 235-244
25. Heisler, L. E., Torti, D., Boutros, P. C., Watson, J., Chan, C., Winegarden, N., Takahashi, M., Yau, P., Huang, T. H., Farnham, P. J., Jurisica, I., Woodgett, J. R., Bremner, R., Penn, L. Z., and Der, S. D. (2005) *Nucleic Acids Res* **33**, 2952-2961
26. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) *Cell* **122**, 947-956
27. Harrow, J., Denoeud, F., Frankish, A., Reymond, A., Chen, C. K., Chrast, J., Lagarde, J., Gilbert, J. G., Storey, R., Swarbreck, D., Rossier, C., Ucla, C., Hubbard, T., Antonarakis, S. E., and Guigo, R. (2006) *Genome Biol* **7 Suppl 1**, S4 1-9
28. Jin, V. X., O'Geen, H., Iyengar, S., Green, R., and Farnham, P. J. (2007) *Genome Res* **in press**
29. Pavesi, G., Mereghetti, P., Mauri, G., and Pesole, G. (2004) *Nucleic Acids Res* **32**, W199-203
30. Bailey, T. L., and Gribskov, M. (1997) *J Comput Biol* **4**, 45-59
31. Cowger, J. J., Zhao, Q., Isovic, M., and Torchia, J. (2006) *Oncogene* **in press**
32. Furusawa, T., Moribe, H., Kondoh, H., and Higashi, Y. (1999) *Mol Cell Biol* **19**, 8581-8590
33. Postigo, A. A., and Dean, D. C. (1999) *Proc Natl Acad Sci U S A* **96**, 6683-6688
34. Turner, J., and Crossley, M. (1998) *EMBO J.* **17**, 5129-5140
35. Criqui-Filipe, P., Ducret, C., Maira, S. M., and Wasylyk, B. (1999) *Embo J* **18**, 3392-3403
36. Turner, J., and Crossley, M. (2001) *Bioessays* **23**, 683-690
37. Bergman, L. M., and Blaydes, J. P. (2006) *Apoptosis* **11**, 879-888
38. Andrews, P. W. (1984) *Dev Biol* **103**, 285-293

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry ASBMB jbc ONLINE

39. Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M., and Walker, P. R. (2005) *J Biol Chem* **280**, 6257-6260
40. Squazzo, S. L., Komashko, V. M., O'Geen, H., Krig, S., Jin, V. X., Jang, S.-W., Green, R., Margueron, R., Reinberg, D., and Farnham, P. J. (2006) *Genome Research* **16**, 890-900
41. Graves, B. J., Gillespie, M. E., and McIntosh, L. P. (1996) *Nature* **384**, 322
42. Scacheri, P. C., Davis, S., Odom, D. T., Crawford, G. E., Perkins, S., Halawi, M. J., Agarwal, S. K., Marx, S. J., Spiegel, A. M., Meltzer, P. S., and Collins, F. S. (2006) *PLoS Genet* **2**, e51
43. Cawley, S., Bekiranov, S., Ng, H. H., Kapranov, P., Sekinger, E. A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A. J., Wheeler, R., Wong, B., Drenkow, J., Yamanaka, M., Patel, S., Brubaker, S., Tammana, H., Helt, G., Struhl, K., and Gingeras, T. R. (2004) *Cell* **116**, 499-509
44. Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T. E., Luscombe, N. M., Rinn, J. L., Nelson, F. K., Miller, P., Gerstein, M., Weissman, S., and Snyder, M. (2003) *Proc Natl Acad Sci U S A* **100**, 12247-12252
45. Carroll, J. S., Liu, X. S., Brodsky, A. S., Li, W., Meyer, C. A., Szary, A. J., Eeckhoute, J., Shao, W., Hestermann, E. V., Geistlinger, T. R., Fox, E. A., Silver, P. A., and Brown, M. (2005) *Cell* **122**, 33-43

FIGURE LEGENDS

Figure 1. **Localization of ZNF217-binding sites in the ENCODE regions. (A)** Triplicate ENCODE array data sets were analyzed for all 3 cell lines and peaks were called using the Tamalpais program (22). Regions containing ZNF217 binding sites were categorized relative to the transcription start site; the divisions are within 2 kb upstream or downstream of the start site, 2-10 kb or 10-100 kb upstream of the start site, 2-10 kb or 10-100 kb downstream of the start site, and sites farther than 100 kb upstream or downstream from a start site. The total number of ZNF217 binding sites in Ntera2 cells (**white bars**) was 176, in MCF7 cells (**grey bars**) was 155, and in SW480 cells (**black bars**) was 61. Also shown (**panel B**) is the binding pattern of ZNF217 to the HOXA cluster before and after treatment of Ntera2 with siRNAs specific to ZNF217.

Figure 2. De novo motif analysis using ZNF217 ChIP-chip data. (A) Shown is a sequence log of a consensus site that was derived from regions that were identified in the ENCODE ChIP-chip assays as ZNF217 binding sites from all three cell lines (Ntera2, MCF7, and SW480). **(B)** Shown is the positional weight matrix (ZNF217_PWM) built by the ChIPMotifs approach (28), with a core score computed from position 2 to 6 (5 bases) and a PWM score from 1 to 8 (8 bases).

Figure 3. Identification of ZNF217 target promoters in Ntera2 cells. The Maxfour values (see METHODS) are plotted for the ZNF217 ChIP-chip data from promoter 1 arrays of two biological replicates using Ntera2 cells; the correlation value is r=0.865. The comparison of the replicate ZNF217 Ntera2 data from promoter 2 arrays can be found in Supplemental **Figure S1**.

Figure 4 Cell type-specificity of ZNF217 binding. (A) The Maxfour values (see METHODS) are plotted for the ZNF217 ChIP-chip data for MCF7 and Ntera2 promoter1 arrays. The boxes indicate regions used for determination of "Ntera2-exclusive" promoters [values less than 0.3 (log2) in MCF7 and greater than 1.5 (log2) in Ntera2] and "MCF7-exclusive" promoters [values less than 0.4 (log2) in Ntera2 and greater than 1.0 (log2) in MCF7]. For the promoter 2 arrays, the values used to identify "Ntera2-exclusive" promoters were less than 0.3 (log2) in MCF7 and greater than 0.8 (log2) in Ntera2 and the values for "MCF7-exclusive" promoters were less than 0.4 (log2) in Ntera2 and greater than 0.8 (log2) in MCF7; see Supplemental **Figure S1** for the promoter 2 comparison graph. **(B)** Five targets identified to be Ntera2-specific (EVX1, OLIG2, TDGF1, PAX6, and SIX6) and five targets identified to be MCF7-specific (KCNK2, FRK, AGR2, MYO1B, and SHC4) from the ChIP-chip array data were analyzed by PCR using a new set of Ntera2 (black bars) and MCF7 (white bars) ChIP samples. The fold enrichment of each promoter was determined by comparing the signal obtained using 10 ng of the

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

ZNF217 amplicon to the signal obtained using 10 ng of input chromatin; a value of 1.0 is expected for non-target promoters. The enrichment values for the target promoters were normalized to that of a negative control (the RAP22A promoter), which was not enriched for ZNF217 binding in either cell type.

Figure 5. (**A) Schematic illustrating three models for ZNF217 and CtBP2 interaction. a**. ZNF217 is required for recruitment of CtBP to DNA, **b.** another factor recruits the complex to DNA, **c**. ZNF217 and CtBP bind to adjacent sites in the promoter. **(B)** Examples of three ZNF217 and CTBP2 target promoters (**a.** EOMES, **b**. NEFL, and **c**. LMO3) in Ntera2 are shown using Signal Map software; the top panel for each promoter is from the ZNF217 ChIP array, the middle panel is from the CtBP2 ChIP array, and the third panel shows CtBP2 binding from siZNF217-treated cells.

Figure 6. Characterization of ZNF217 targets expressed at different levels.
PCR analysis of 4 ZNF217 target genes that are in the "repressed" category (left side of panel) compared to 4 from the "activated" category (right side of panel). RAP22A was used as a negative control for normalizing sample to input with a fold-enrichment equal to one.

Figure 7. **ZNF217 is down-regulated by retinoic acid in Ntera2cells.** Western blot of retinoic acid timecourse in Ntera2 shows down-regulation of ZNF217 by Day 4 of treatment to the same levels as siZNF217 treatment. Oct4 down-regulation indicates Ntera2 cells are differentiating. A section of the ponceau stain indicates levels of protein loaded in each lane.

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

Krig_Table 1. ZNF217 enrichment in siRNA-treated cells

Ranked genes*	Control Array	Knockdown Array
Top 500	1.04	0.49
500-1000	0.81	0.42
1000-2000	0.67	0.39
5000-5500	0.41	0.34

*ZNF217 ChIP-chip assays were performed using control cells treated with si-GLO or siRNAs specific to ZNF217, the ~14,000 promoters on promoter 1 array were ranked by their ZNF217 Maxfour enrichment values in the control cells, and the top half of the ranked promoters were binned into the indicated categories. Then, the average fold-enrichment for each bin was calculated for ZNF217 enrichment in the control cells and in cells treated with siRNAs to ZNF217; shown are the log2 values.

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

Krig_Table 2. Gene ontology of ZNF217 target genes.

MCF7 and Ntera2 (616)	%	P-Value
ion binding	23.5	1.6E-04
transcription factor activity	8.0	5.1E-04
nucleic acid binding	21.7	9.4E-04
helicase activity	2.0	5.6E-03
Ntera2 exclusive (379)		
transcription factor activity	17.9	6.0E-19
organ development	11.9	9.0E-15
receptor binding	7.7	2.6E-04
nucleic acid binding	25.9	3.0E-04
cell differentiation	6.3	5.5E-04
embryonic development	2.1	1.5E-03
cell adhesion	7.1	2.8E-03
ion transporter activity	7.4	8.6E-03
MCF7 exclusive (114)		
organ development	6.9	5.36E-03
morphogenesis	6.0	3.26E-02
negative regulation of biological process	6.9	3.32E-02
tissue development	3.5	4.05E-02
membrane	25.0	9.85E-02

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

Krig_Table 3. Comparison of ZNF217 and CtBP target promoters

Ntera2 Top 1000	promoter1	promoter2
CtBP1/CtBP2	78%	63%
ZNF217/ZNF217	69%	64%
CtBP1/ZNF217	71%	62%
CtBP2/ZNF217	72%	75%
IgG/ZNF217	7%	2%
MCF7 Top 1000	promoter1	promoter2
CtBP2/ZNF217	64%	56%

Krig_Table 4. ZNF217 targets are expressed at low levels

RNA Level	No. genes matching	% of Total
Low <300	689	64
300-1000	255	24
High >1000	132	12

The Journal of Biological Chemistry
Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

Krig_Table 5. Gene ontology of ZNF217 targets expressed at low vs. high levels.

RNA<300 (613)	%	P-Value
transcription factor activity	15.6	2.64E-22
system development	8.4	4.79E-13
organ development	8.4	2.85E-11
nucleic acid binding	26.9	4.01E-08
cell differentiation	6.1	1.45E-05
embryonic development	2.0	5.75E-05
transporter activity	4.9	1.65E-04
tumor suppressor	0.8	8.86E-04
sex differentiation	1.2	1.80E-03
postsynaptic membrane	1.5	8.34E-03
RNA>1000 (122)		
morphogenesis	9.9	2.1E-03
regulation of growth	4.1	1.1E-02
system development	7.4	1.4E-02
cell differentiation	6.6	5.2E-02
organ development	6.6	6.5E-02

Krig_Table 6. Identification of genes responsive to loss of ZNF217

RNAs responsive to loss of ZNF217*	A	B	C	2 of 3	3 of 3
RNAs upregulated by loss of ZNF217	462	402	540	288	106
RNAs downregulated by loss of ZNF217	452	116	318	133	29
ZNF217 targets upregulated by loss of ZNF217	43	61	77	42	18
ZNF217 targets downregulated by loss of ZNF217	27	18	23	14	3

* Up-regulated RNAs are identified as those having expression levels greater than 200 in the cells treated with siRNA to ZNF217 and a greater than 1.5 fold increase in expression after reduction of ZNF217; down-regulated RNAs are identified as those having RNA expression levels greater than 200 in the untreated samples and showing a greater than 1.5 fold decrease in expression after reduction of ZNF217. The intersection of the ZNF217 bound promoters and the upregulated RNAs are identified as ZNF217 targets upregulated by the loss of ZNF217 and the intersection of the ZNF217 bound promoters and the downregulated RNAs are identified as ZNF217 targets downregulated by the loss of ZNF217.

Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

The Journal of Biological Chemistry

Krig_Table 7. Genes directly bound by and regulated by ZNF217*

Repressed by ZNF217		Activated by ZNF217
ABHD7	NRXN3	**ARMCX5**
ADM	PAK3	**CCNE2**
ANK3	PIPOX	EOMES
ATP10D	PKP2	GAD1
CCL2	PLAT	GATA4
COL8A1	RGS20	**KRT18**
CREB5	SEC14L2	PUNC
CXXC4	SEMA3A	SOCS2
DPP6	**SH3RF2**	ST3GAL6
DSCR1	SLC6A15	ST6GAL1
EVA1	**SPAG9**	STRA6
GPRC5A	SPG3A	WNT5B
HAPLN1	**TP53AP1**	ZNF616
HOXC6	VSNL1	
IFI16		
IGFBP3		
ITM2A		
KLHL4		
LMO3		
LYPD1		
MAP2K5		
MYCBP2		
NEFL		
NLGN1		
NMNAT2		
NMU		
NRK		

*Shaded genes indicate ZNF217 targets in both MCF7 and Ntera2 cells

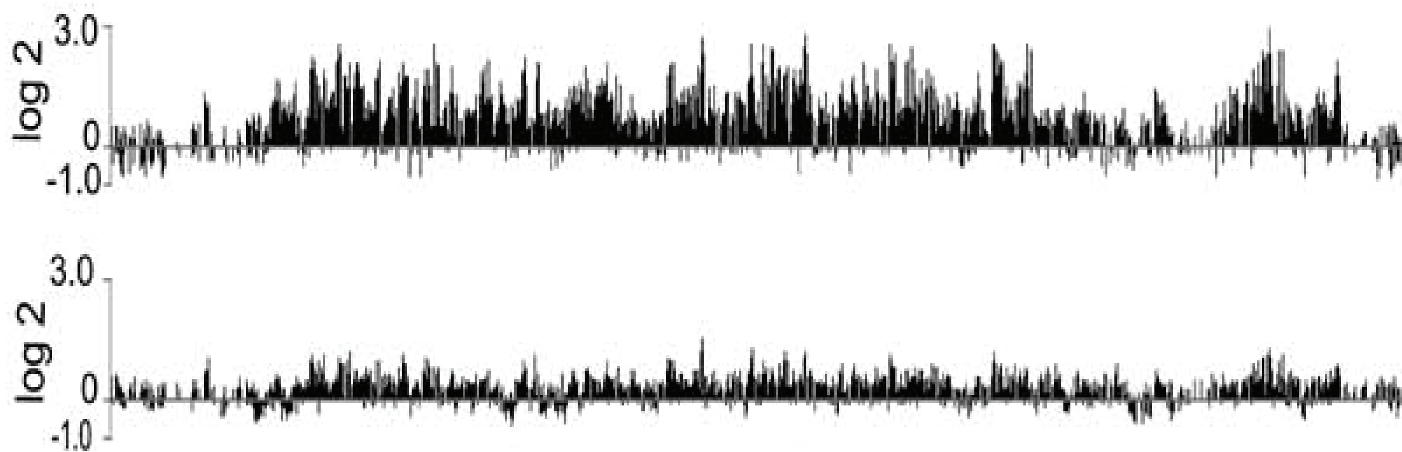
Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry
jbc ONLINE

100
90
80
70
60
50
No. of ZNF217
Binding Sites
40
30
20
10
0

Gene desert
-100 kb - -10 kb
-10 kb - -2 kb
-2 kb - 2 kb
2 - 10 kb
10 kb - 100 kb

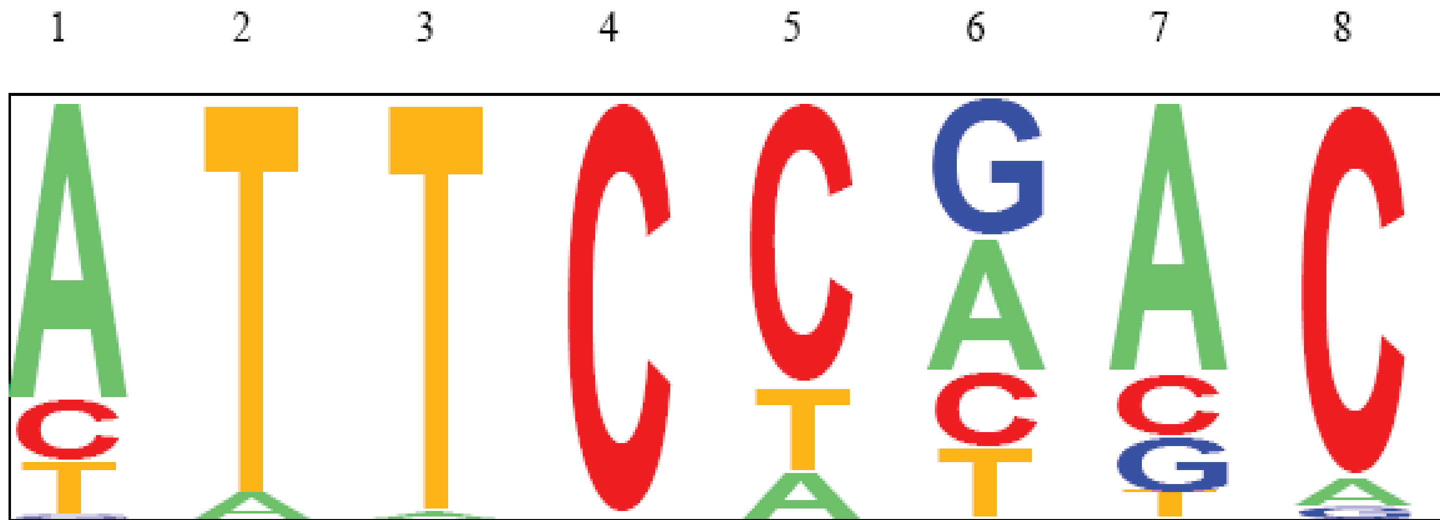
Distance relative to TSS



The Journal of Biological Chemistry

jbc ONLINE

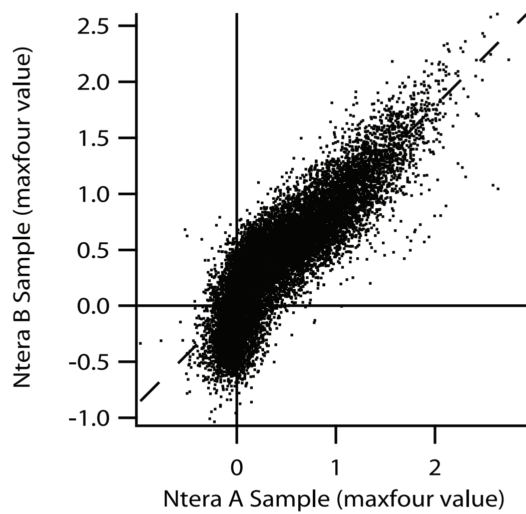
ded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007



B

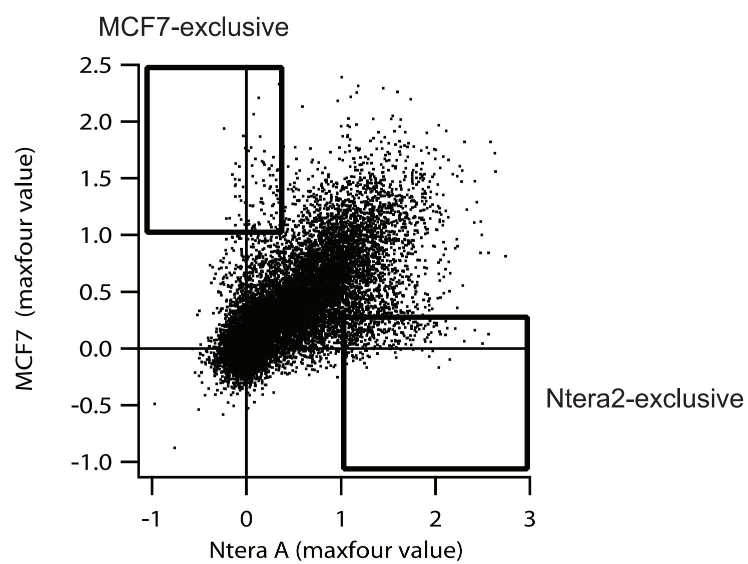
	1	2	3	4	5	6	7	8
A	69	6	2	0	11	31	63	6
C	15	0	0	97	67	18	15	88
G	1	0	0	0	0	32	13	3
T	12	91	95	0	19	16	6	0

Krig_Figure 3



The Journal of Biological Chemistry
jbc ONLINE
ASBMB
Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

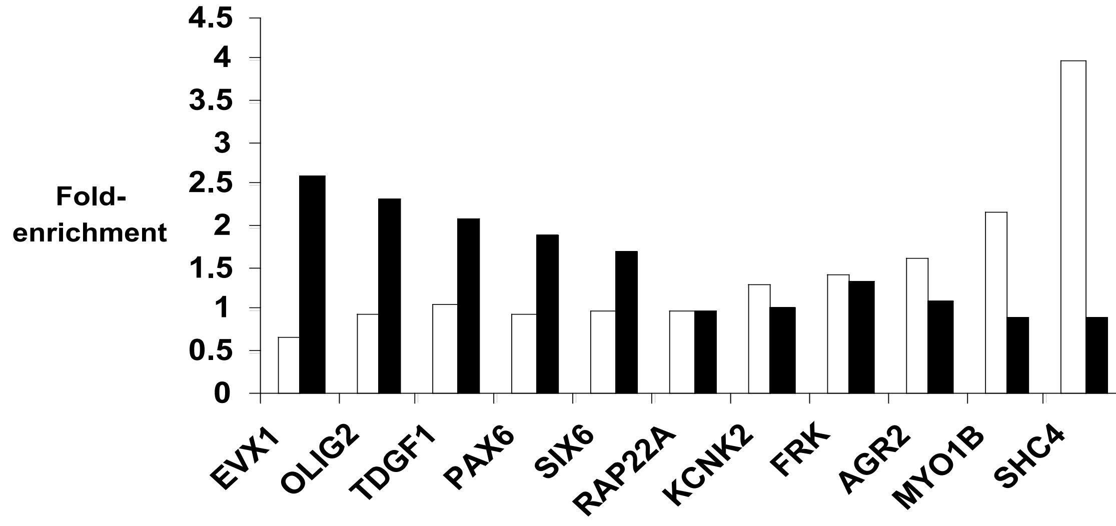
Krig_Figure 4A



Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

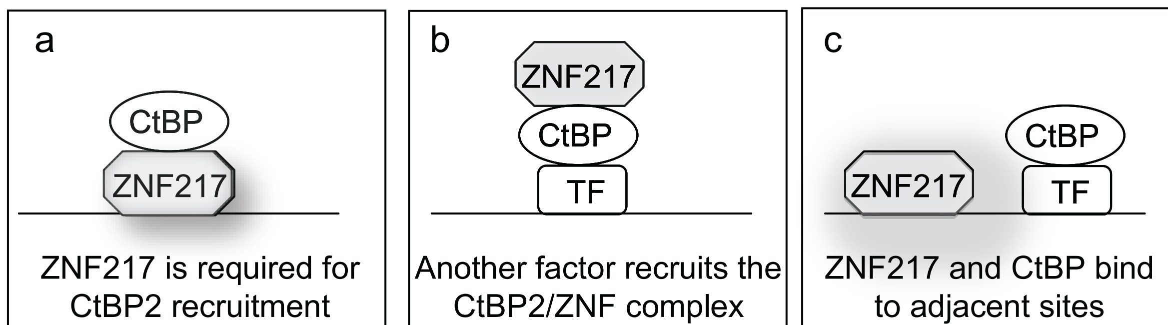
Krig_Figure 4B



Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry
ASBMB
jbc ONLINE

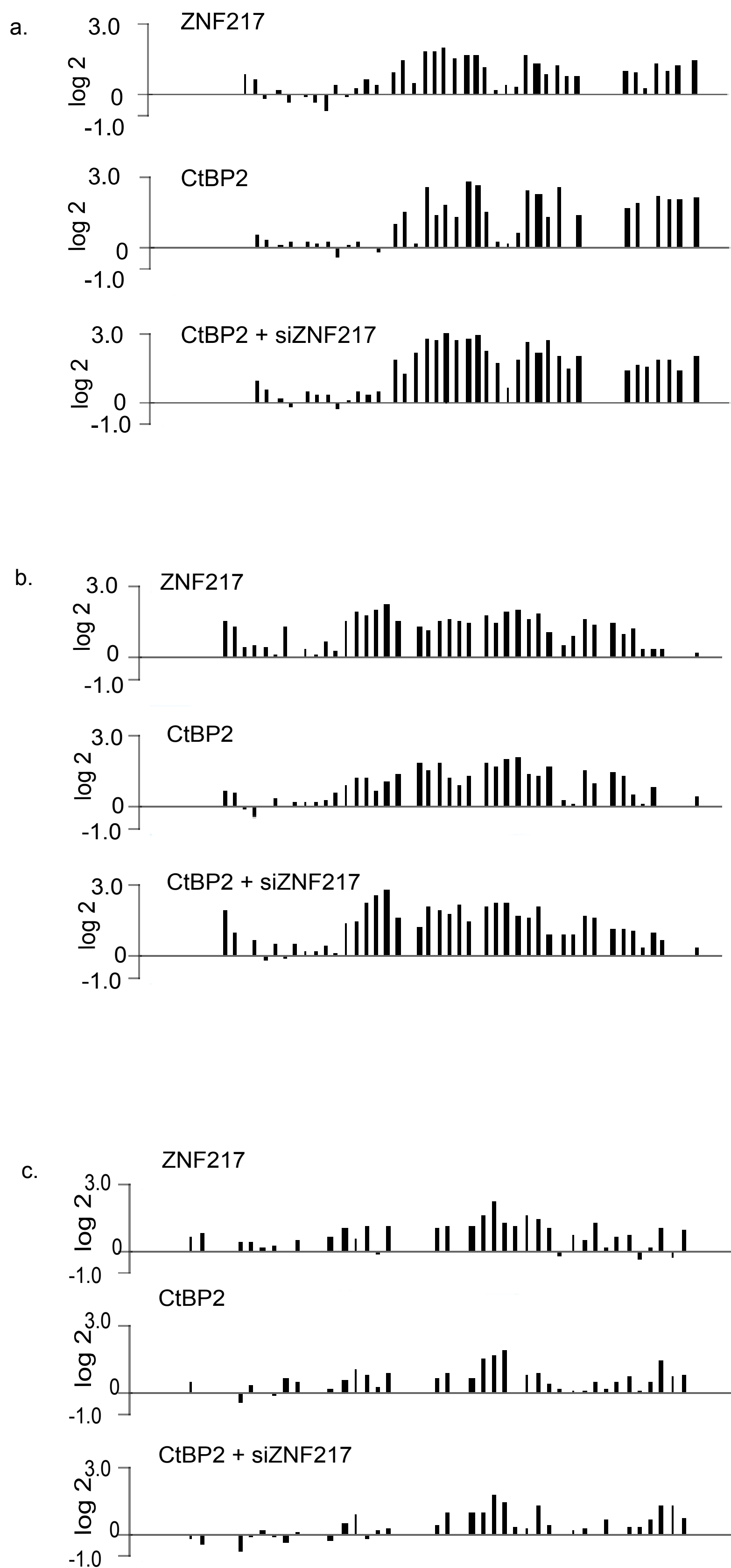
Krig_Figure 5A



Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry
ASBMB
jbc ONLINE

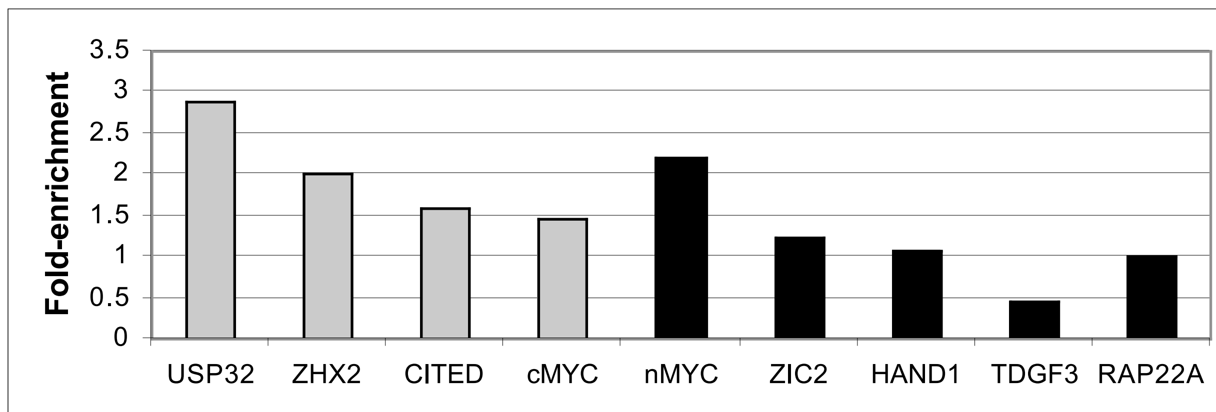
Krig_Figure 5B



Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry

Krig_Figure 6

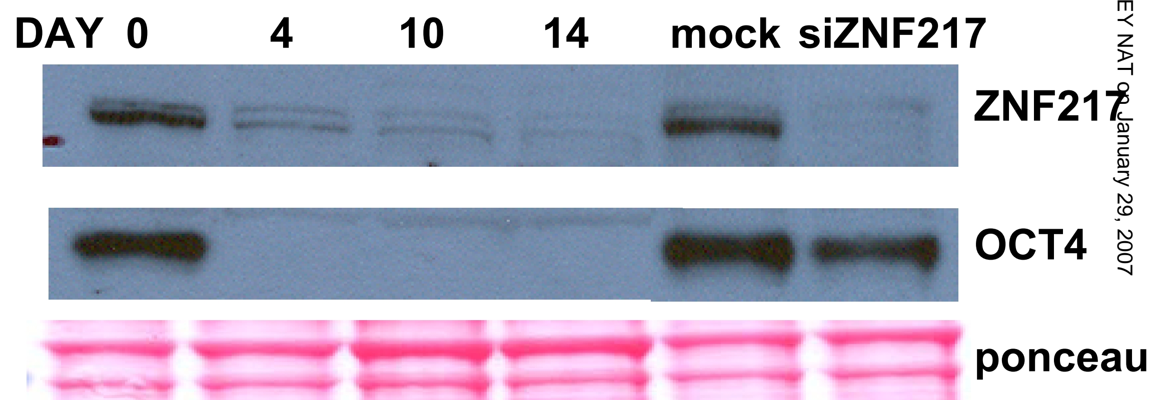


Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007

The Journal of Biological Chemistry
ASBMB
jbc ONLINE

The Journal of Biological Chemistry

Krig_Figure 7



Downloaded from www.jbc.org at E O LAWRENCE BERKELEY NAT on January 29, 2007