THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 45, pp. 31327–31335, November 6, 2009

Zfp206, Oct4, and Sox2 Are Integrated Components of a Transcriptional Regulatory Network in Embryonic Stem Cells*[S]

Received for publication, May 4, 2009, and in revised form, September 4, 2009 Published, JBC Papers in Press, September 9, 2009, DOI 10.1074/jbc.M109.016162

Hong-bing Yu[‡§], Galih Kunarso[¶], Felicia Huimei Hong[‡§], and Lawrence W. Stanton[‡§1]

From the [‡]Department of Biological Sciences, National University of Singapore, Singapore 117542 and the Departments of [§]Cell and Developmental Biology and [¶]Computational and Mathematical Biology, Genome Institute of Singapore, Singapore 138672

Zfp206 **(recently renamed *Zscan10*) encodes a zinc finger transcription factor specifically expressed in human and mouse embryonic stem cells (ESC). It has been shown that Zfp206 is required to maintain ESC in an undifferentiated, pluripotent state. Presented here are data showing that Zfp206 works together with two other transcription factors, Oct4 and Sox2, which are also essential regulators of ESC pluripotency. We show that Zfp206 binds to the *Oct4* promoter and directly regulates Oct4 expression. Genome-wide mapping of Zfp206-binding sites in ESC identifies more than 3000 target genes, many of which encode transcription factors that are also targeted for regulation by Oct4 and Sox2. In addition, we show that Zfp206 physically interacts with both Oct4 and Sox2. These data demonstrate that Zfp206 is a key component of the core transcriptional regulatory network and together with Oct4 and Sox2 regulates differentiation of ESC.**

Pluripotency, the potential to give rise to all lineages of the developing embryo, is a unique and defining characteristic of mammalian embryonic stem cells (ESC).[2] Pluripotent ESC, like the inner cell mass of the embryo from which they were derived, exist in a developmental state that is poised to respond to extracellular signals that specify unique patterns of cellular differentiation. ESC responding to extrinsic cues must undergo transitions from a self-renewing and pluripotent state to one of many alternative states of differentiation. Early genomics approaches have revealed transcriptional regulatory networks that are responsible for maintaining ESC pluripotency (1, 2). Two essential regulators of pluripotency are the transcription factors (TF) Oct4 and Sox2. Knockdown of these transcription factors results in loss of ESC pluripotency and induction of nonspecific differentiation (3). The importance of Oct4 and Sox2 in pluripotency is underscored by their ability to reprogram differentiated fibroblasts into induced pluripotent stem-like cells (4). More recently, thousands of direct target genes regulated by Oct4 and Sox2 have been identified through comprehensive, genome-wide chromatin immunoprecipitation studies (5, 6). Many of the Oct4/Sox2 targets are genes encoding other transcriptional regulators, including several that also have been found to play a role in regulating pluripotency, such as *Nanog*, *Esrrb*, *Tcf3*, *Tcl1*, *Zfp281*, *Zic3*, and *Sall4* (3, 7, 8–12).

Zfp206 is another transcription factor that is specifically expressed in ESC (13, 14) and is directly regulated by Oct4 and Sox2 (15). Zfp206 was implicated as a pluripotency factor because it was found highly expressed in undifferentiated ESC and the inner cell mass of the preimplantation embryo, but not in differentiated ESC or trophectoderm. Knockdown of *Zfp206* expression induces ESC differentiation, whereas its sustained overexpression impedes retinoic acid induced differentiation of ESC, thus establishing that Zfp206 is a regulator of pluripotency (13). Zfp206 encodes a protein that contains 14 zinc fingers, although alternative splice forms contain fewer finger domains. The presence of a SCAN domain in the amino-terminal half of Zfp206 suggests that this protein dimerizes, perhaps with itself or other binding partners. Recently, the HUGO gene nomenclature committee renamed this gene *Zscan10*, but we will continue to use the term Zfp206 throughout this paper.

It remains unknown what are the target genes that are regulated by Zfp206 in pluripotent ESC. Identification of such target genes would broaden our understanding of the transcriptional regulatory network that operates in ESC to maintain pluripotency. We present here the results of a genome-wide study to identify Zfp206-binding sites and their associated genes. Zfp206 target genes were compared with the regulatory networks that have been previously determined for Oct4, Sox2, Zfp281, and Nanog. Our results show that hundreds of genes are commonly targeted by Zfp206 and the other pluripotency transcription factors. Of note, we demonstrate that Zfp206 binds to the enhancer region of *Oct4*. We also present evidence that Zfp206 physically interacts with both Oct4 and Sox2. Thus, it is clearly demonstrated here that Zfp206 is an integral component of the transcriptional network and functions in concert with Oct4 and Sox2 to maintain ESC pluripotency.

MATERIALS AND METHODS

ESC Culture—E14 cells, a mouse ESC cell line (American Type Culture Collection), were cultured under feeder-free conditions in Dulbecco's modified Eagle's medium supplemented with 15% fetal bovine serum (ESC-qualified; Invitrogen), 0.055 mM β-mercaptoethanol (Invitrogen), 2 mM L-glutamine, 0.1 mM nonessential amino acid, 5000 units/ml penicillin/streptomycin, and 1000 units/ml of leukemia inhibitory factor (Chemicon) and maintained at 37 °C with 5% CO_2.

Chromatin Immunoprecipitation (ChIP)—ChIP assays with E14 cells were carried out as described previously (2, 16).

* This work was supported by Singapore's Agency for Science Technology and Research.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and Tables S1 and S2.

[1] To whom correspondence should be addressed: Genome Institute of Singapore, 60 Biopolis St., Singapore 138672. E-mail: stantonl@gis.a-star.edu.sg.

[2] The abbreviations used are: ESC, embryonic stem cell(s); TF, transcription factor(s); ChIP, chromatin immunoprecipitation; shRNA, small hairpin RNA.

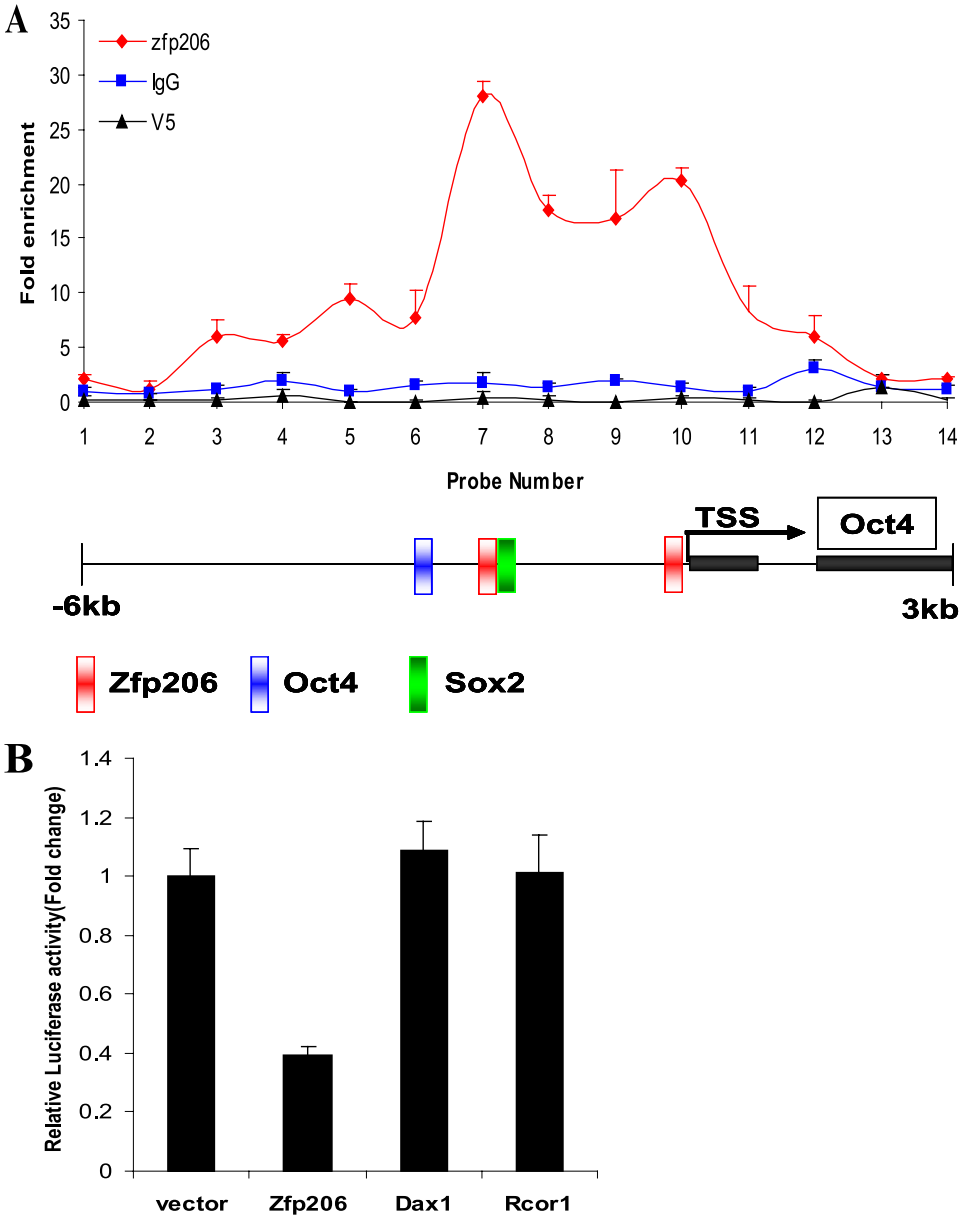

FIGURE 1. **Mapping of Zfp206-binding sites at the *Oct4* promoter.** *A*, two binding sites for Zfp206 were identified within the enhancer/promoter region of *Oct4*. Chromatin immunoprecipitations were performed in ESC using IgG, a V5 antibody, or an antibody specific to Zfp206. ChIP enrichments were measured by Q-RT-PCR and expressed as fold differences, relative to non-IP (input) chromatin. Positions of the 14 regions (*probe numbers 1–14*) amplified are indicated as distances (−6 to 3 kb) from the transcriptional start site (*TSS*) of the *Oct4* gene. *B*, the transcriptional activity of the *Oct4* promoter is dependent on Zfp206-binding sites. A luciferase reporter assay was used to measure activity of the *Oct4* promoter in ESC that expressed shRNA against *Zfp206*, *Dax1*, and *Rcor1*. The relative luciferase activity is compared with a vector that expressed no shRNA. Knockdown of Zfp206 resulted in a significant decrease in Oct4 promoter activity (*, $p < 0.01$ by Student's *t* test). Knockdown of *Dax1* or *Rcor1* had no significant affect on *Oct4* promoter activity.

were eluted, and the ChIP DNA was extracted. For all of the ChIP experiments, quantitative PCR analyses were performed in real time using the ABI PRISM 7900 sequence detection system and SYBR Green master mix (Applied Biosystems) as described previously (2, 16). Relative occupancy values were calculated by determining the apparent immunoprecipitation efficiency (ratios of the amount of immunoprecipitated DNA over that of the input sample) and then normalized to the level observed at the control region.

Commercial ChIP-on-chip promoter microarrays (Agilent Technologies, Palo Alto, CA), which tiled the genomic regions from 5.5 kb upstream to 2.5 kb downstream of the transcription start site of 17,000 annotated mouse genes, were used to interrogate genome-wide occupancy of Zfp206 from two independent biological replicates. ChIP samples were blunt-ended, ligated to linkers, and PCR-amplified. DNA was fluorophore-labeled using an Invitrogen CGH labeling kit (ChIP samples with Cy5; whole cell extract with Cy3). The labeled DNA was hybridized to Agilent mouse promoter ChIP-on-chip arrays (Agilent). The chips were washed and scanned as per the manufacturer's protocol, and the data were processed by Agilent ChIP Analytics software. A probe was considered bound if p value < 0.001.

For the Gene Ontology analysis, the PANTHER classification system was utilized. For *de novo* motif identification, the top 500 high quality, enriched probes from the Zfp206-bound regions were selected, and the 300 bp, repeat masked sequences centered on these probes were used as input for the Weeder algorithm (17).

Briefly, 1×10^8 cells were treated with 1% formaldehyde for 10 min at room temperature to cross-link transcription factors to chromatin for each ChIP experiment. Formaldehyde was inactivated by the addition of 125 mM glycine. Sonicated chromatin extracts containing DNA fragments with an average size of 500 bp were immunoprecipitated using 10 μg of Zfp206 antibody (13), rabbit IgG (Santa Cruz Biotechnology), or V5 control monoclonal antibodies (Invitrogen) with preblocked protein G-Sepharose beads with overnight incubation at 4 °C. The following day, the chromatin-protein-antibody-bead complexes

Luciferase Reporter Assays—The promoter regions, including the Zfp206-binding sites, for *Oct4*, *Klf5*, *Jarid1c*, *Pitx2*, *Meis1*, *Meis2*, *mir124a1*, *mir124a2*, *Jarid1c-truncated*, and *mir-124-a2-truncated* were cloned into pGL4-basic vector (Promega), and the *Klf4* intron region was cloned into the pGL4.23-minimal promoter vector (Promega). A dual luciferase system (Promega, Madison, WI) was used. For the luciferase assay in E14 cells, 2×10^5 cells were seeded into one well of 24-well plate. After 18 h, 275 ng of luciferase reporter plasmid,

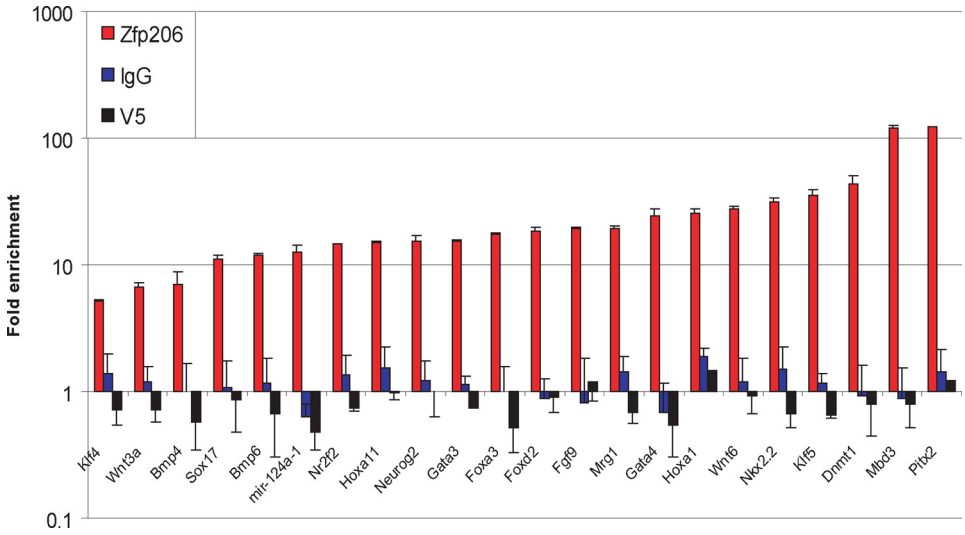

FIGURE 2. **Validation of selected Zfp206 target genes.** ChIP-PCR was performed on 22 binding sites identified by ChIP-chip. The ChIP-PCR experiments confirmed enrichments of Zfp206 binding at all 22 sites. Shown are the ChIP enrichments (fold differences relative to input DNA) using IgG, a V5 antibody, and a Zfp206-specific antibody. The primers used in this study are listed in the supplemental data.

4 ng of plasmid pRL-SV40, and Zfp206 shRNA vector or empty vector were co-transfected into the cells by Lipofectamine 2000 (Invitrogen). The pRL-SV40 plasmid served as an internal control for normalizing the transfection efficiency. After 40 h of transfection, the E14 cells were lysed, and the luciferase activity was determined with the dual luciferase system (Promega) using a Centro LB960 96-well luminometer (Berthold Technologies). Transfections were performed in quadruplicate. For luciferase assay in HEK293T cells, 4×10^5 cells were seeded into one well of 24-well plate. After 18 h, 275 ng of luciferase reporter plasmid, 2 ng of plasmid pRL-SV40, and Zfp206 overexpression or empty vector were co-transfected into the cells by Lipofectamine 2000 (Invitrogen). After 40 h of transfection, the luciferase activity was determined as in E14 cells.

Co-immunoprecipitation—Expression plasmids, which contain full-length mouse *Oct4* and *Sox2* fused to a V5 epitope tag, were transfected into E14 cells. The IgG and Rcor1 antibodies served as negative controls in this experiment. Whole cell lysates were harvested 48 h after transfection. Zfp206, V5, IgG, and Rcor1 antibodies were adsorbed to protein G-Sepharose beads (GE Healthcare) and incubated with the cell lysates for 24 h at 4 °C. The beads were washed six times with lysis buffer, added to 150 μl of lysis buffer, boiled for 10 min, and analyzed by Western blot using Zfp206 or V5 antibodies.

RESULTS

Direct Activation of Oct4 Expression by Zfp206—The expression of the transcription factor Zfp206 is tightly correlated with expression of Oct4; both are expressed predominantly in the inner cell mass of the preimplantation embryo and in ESC but down-regulated upon differentiation. We and others have demonstrated that Zfp206 is involved in maintaining ESC pluripotency (13, 14). Previously, we proposed that Zfp206 regulates pluripotency by controlling *Oct4* expression because knockdown of Zfp206 expression by RNA interference resulted in repression of *Oct4*, and conversely, overexpression of Zfp206 enhanced expression of *Oct4* (15). Those results led us to ask whether *Oct4* expression is regulated by direct binding of Zfp206 to the *Oct4* promoter. To test this hypothesis, ChIP experiments were carried out using a Zfp206 antibody. The locations of Zfp206 enrichments by ChIP were ascertained by PCR using primers designed across the *Oct4* promoter region, from 6 kb upstream and 3 kb downstream of the transcriptional start site, which includes the first exon. We identified two regions where Zfp206 binds; one was ~1000 bp upstream, and the other was very near the transcriptional start site of *Oct4* (Fig. 1*A*).

We then tested the activity of Zfp206 on the *Oct4* promoter using a luciferase reporter assay for transcription. The 1600-bp *Oct4* promoter sequence, which includes the two newly identified Zfp206-binding sites, was linked to the luciferase reporter gene. This Oct4-luciferase reporter was co-transfected with a shRNA vector that targets *Zfp206* for knockdown. The shRNA resulted in an 80% decrease in Zfp206 mRNA levels (supplemental Fig. S1). In response to *Zfp206* knockdown, a greater than 2-fold reduction in transcription activity was observed from the *Oct4* promoter (Fig. 1*B*). In control assays, no influence on the Oct4 promoter was observed upon knockdown of two other transcriptional regulators, Dax1 and Rcor1, which are also expressed in ESC. Thus, Oct4 expression is regulated through direct recruitment of Zfp206 to the *Oct4* promoter. It is interesting that, correspondingly, Oct4 binds at the *Zfp206* promoter and directly regulates *Zfp206* expression (15), indicating that these two regulators of pluripotency form reciprocal regulatory loops in ESC and control each other's expression.

Genome-wide Mapping of Zfp206 Targets in ESC—The above results demonstrate that Zfp206 and Oct4 are linked in a common transcriptional network. To define more comprehensively the regulatory network associated with Zfp206 in ESC, we performed a genome-wide ChIP analysis using a microarray-based (ChIP-chip) strategy. A Zfp206-specific antibody was used to immunoprecipitate chromatin bound by Zfp206. The ChIP DNA was then applied to a commercial DNA array that contained probes tiled across promoter regions of ~17,000 genes from −5.5 to +2.5 kb, relative to the transcriptional start site. Statistically significant ($p < 0.001$) binding sites were identified from two independent biological samples, as described under "Materials and Methods." We identified 3552 Zfp206-binding sites that were associated with 3558 genes (supplemental Table S1). Consistent with our ChIP-PCR results (Fig. 1*A*), the two Zfp206-binding sites in the *Oct4* promoter were revealed as two peaks in the ChIP-chip result. To confirm the validity of the ChIP-chip results, we performed quantitative ChIP-PCR on 22 selected targets. All of these showed substantial enrichment for Zfp206 binding (Fig. 2), thus confirming the reliability of the ChIP-chip data.

Using the ChIP-chip results we asked whether Zfp206 preferentially targeted genes of any particular biological functions in ESC. To this end we analyzed the gene ontology terms ascribed to each of the target genes to determine whether there was over-representation of genes related to certain biological processes (Table 1). It is interesting that the top three significantly over-represented ontologies were related to transcription and nucleic acid metabolism functions, suggesting that Zfp206, like Oct4, regulates pluripotency in large part by targeting other components of a transcriptional network. This is further supported by the fact that in addition to *Oct4*, several other genes that encode pluripotency-regulating transcription factors, including *Klf2*, *Klf4*, *Klf5*, *Zfp281*, and *Sall4*, were identified as direct targets of Zfp206.

TABLE 1
Gene ontology of Zfp206 target genes

The enriched gene ontology terms of Zfp206 target genes are shown. The number of observed and expected genes for each of the most significantly over-represented gene ontology terms is shown together with associated *p* value for each.

Biological process	Number observed	Number expected	*p* value
Nucleoside, nucleotide, and nucleic acid metabolism	789	533.39	1.01E-28
mRNA transcription	489	294.19	7.2E-26
mRNA transcription regulation	399	226.51	5.44E-25
Neurogenesis	185	110.4	5.1E-09
Ectoderm development	198	125.84	1.16E-07
Cell cycle	235	166.66	0.0000053
Developmental processes	490	394.23	0.0000121
Segment specification	43	19.25	0.000287
Cell cycle control	119	79.1	0.00198
Intracellular protein traffic	223	174.27	0.0047
Cell proliferation and differentiation	211	163.91	0.00514
Embryogenesis	48	27.92	0.0471

A DNA-binding motif has not previously been described for Zfp206. Our ChIP-chip data provided the opportunity to describe such a binding sequence. The ChIP-chip data were analyzed with Weeder (17), a motif-finding algorithm, to identify common DNA sequences within the enriched binding regions. The program identified a 10-bp palindromic sequence as the binding motif for Zfp206 (Fig. 3). It was observed that 2305 of the 3552 binding regions (64.9%) contain such a DNA motif at a statistically significant cut-off *E* value of <0.0001.

FIGURE 3. **Consensus DNA motif for Zfp206-binding sites.** A consensus binding site motif for Zfp206 was computationally determined by analysis of ChIP-chip data. A logo of the newly defined 10-base DNA-binding motif is presented. The motif was determined by the Weeder algorithm (17) using the top 500 binding regions identified by ChIP-chip.

There is mounting evidence that Oct4, Sox2, and Nanog are components of an interacting network that regulates ESC pluripotency by targeting a large set of common genes (5, 6, 18). Because Zfp206 regulates ESC pluripotency, we asked whether there was a preference for Zfp206 target genes to be targeted by this extended network, which also includes Zfp281 (6, 10). We compared the target genes, as identified by ChIP, for all of these factors. In each case a statistically significant overlap of target genes was found (Fig. 4). For example, 846 of 3558 Zfp206 target genes overlapped with the 2147 genes targeted by Zfp281 ($p = 1.50E-89$) (Fig. 4). Likewise, we found extremely statistically significant co-occupancies of Zfp206 with both Oct4 and Sox2, which shared 487 ($p = 3.8E-25$) and 534 ($p = 2.53E-47$) sites with Zfp206, respectively. Nanog and Zfp206 co-targeted 697 genes, which was at a lower although still significant value ($p = 3.52E-10$) than was seen for the other transcription factors. Some common targets of Oct4, Sox2, and Zfp206 have been shown to play an essential role in ES cell pluripotency such as *Oct4*, *Jarid1*, *Klf2*, as well as mouse development such as *Hoxb13*, *Meis1*, and *Pax6* (Table 2). These

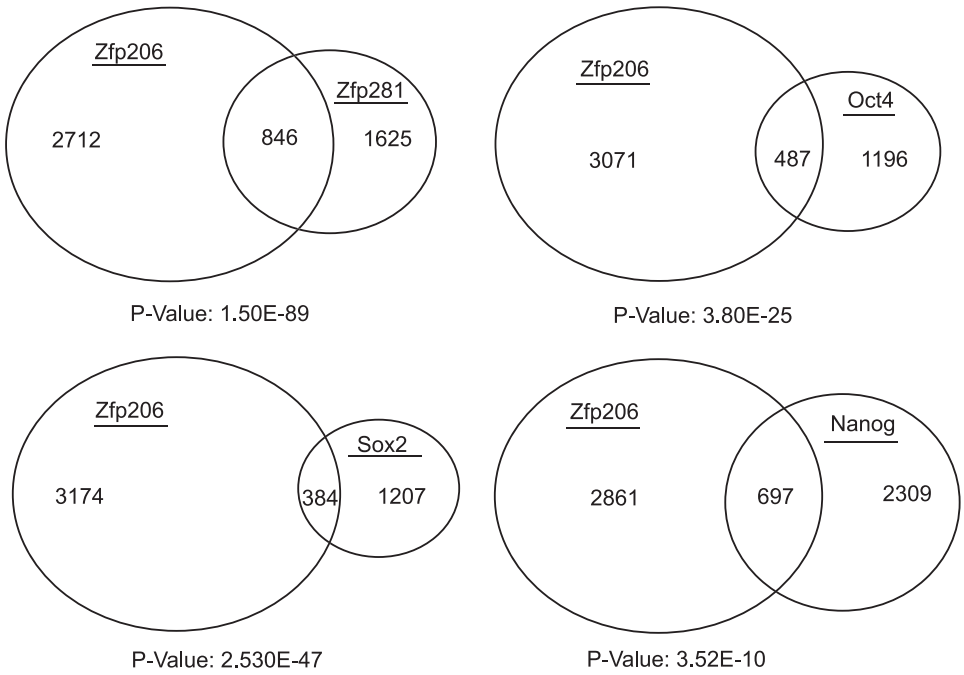

FIGURE 4. **Frequent co-targeting of genes by Zfp206 and other pluripotency transcription factors.** The Venn diagrams show the number of Zfp206 target genes that are also targets of Zfp281, Nanog, Oct4, and Sox2. The results indicate statistically significant overlap of Zfp206 target genes and genes targeted also by pluripotency transcription factors, Zfp281, Sox2, Nanog, and Oct4.

TABLE 2
Transcription factor genes commonly targeted by Oct4, Sox2, and Zfp206

Gene symbol	Accession number
Oct4	NM_013633
Sall4	NM_201396
Dll1	NM_007865
Dll4	NM_019454
Foxn3	NM_183186
Foxo3	NM_019740
Foxp4	NM_028767
Hivep2	NM_010437
Hoxb13	NM_008267
Jarid2	NM_021878
Junb	NM_008416
Klf2	NM_008452
Klf9	NM_010638
Lbxcor1	NM_172446
Lmx1b	NM_010725
Meis1	NM_010789
Meis2	NM_010825
Mtf2	NM_013827
Mybl2	NM_008652
Mycl1	NM_008506
Neurog1	NM_010896
Nfat5	NM_133957
Nfib	NM_008687
Otx2	NM_144841
Pax3	NM_008781
Pax5	NM_008782
Pbx3	NM_016768
Cited2	NM_010828
Rax	NM_013833
Sall1	NM_021390
Arid5b	NM_023598
Satb2	NM_139146
Smyd3	NM_027188
Ssrp1	NM_182990
Tle1	NM_011599
Zbtb40	NM_198248
Zbtb45	NM_001024699
Zfp521	NM_145492
Zfp536	NM_172385
Zfp623	NM_030199
Zfp64	NM_009564

results further demonstrate that Zfp206 is an integral component of pluripotency circuitry.

To investigate whether Zfp206 and other TF physically occupy overlapping regions, we sought to determine the distances between the binding sites. A distribution analysis of the binding sites of Zfp281, Oct4, Sox2, and Nanog relative to Zfp206-binding sites was performed (Fig. 5*A*). Interestingly, ~35% of the common targets for Zfp206 and Zfp281 were found to have their binding sites within 200 bp of each other. Oct4 displayed a similar co-localization with Zfp206. A few notable examples of genes co-targeted by close binding of the pluripotency regulatory factors Zfp206, Oct4, and Sox2 include *Ssrp1*, *Stmn1*, and *Cited2* (Fig. 5*B*). In contrast, co-localization of Zfp206 binding was much less frequent for Nanog. Previous studies have shown a significant co-localization of Oct4, Sox2, and Nanog (5, 6). Our data suggest that Zfp206 may be cooperating more selectively with Oct4 and Sox2 but less so with Nanog.

Zfp206 Selectively Activates or Represses Target Genes—The ChIP-chip experiment identified 3552 Zfp206-binding sites, but it remained to be established whether Zfp206 regulates gene expression as a consequence of binding to these sites. To test whether Zfp206 functionally activates or represses gene expression from promoters associated with these binding sites, eight promoter regions containing Zfp206-binding sites were linked to a luciferase reporter. In addition, one binding site (from the *Klf4* intron) was linked to a minimal promoter driving a luciferase reporter to test for enhancer activity of Zfp206. The luciferase reporters were co-transfected into ESC together with a shRNA construct for specific knockdown of endogenous Zfp206. Knockdown of Zfp206 (~80% reduction) induced a drop of luciferase reporter activity from *Oct4*, *Klf5*, and *Jarid1c* promoters and the *Klf4* enhancer (Fig. 6). In contrast, *Meis2* promoter activity was elevated by Zfp206 knockdown, which indicates that Zfp206 represses its activity. Similarly, two microRNAs, mir-124-a1 and mir-124-a2, were also elevated in response to knockdown of Zfp206. A second, independent shRNA that targets Zfp206 gave essentially identical results, confirming that this response was specific to Zfp206 knockdown (Fig. 6). Knockdown of Zfp206 in ESC had comparable effects on endogenous expression of target genes (supplemental Fig. S1).

To confirm that the modulation of promoter activities was dependent on the Zfp206-binding sites, we deleted the binding sites from two of the luciferase reporter. For the Jarid1c and mir-124-a2 reporter constructs, the region containing the Zfp206 was deleted (Fig. 7). The truncated constructs were then assayed in parallel with the reporters that contained the binding sites. Upon shRNA knockdown of Zfp206 in ESC, we observed no reduction in activity of the truncated Jarid1c promoter or activation of the truncated mir-124-a2 promoter (Fig. 7). As a control for the specificity of the luciferase reporter assay, knockdown of the transcription factor Dax1 was tested and had no impact on expression of either reporter. These results confirm that the Zfp206-binding sites are required for transcriptional regulation of these promoters by Zfp206.

We also tested the activity of Zfp206 on transcriptional regulation of target genes by enforced expression in a somatic cell line, HEK293T, which normally do not express Zfp206. The nine luciferase reporter constructs used above were co-transfected with a Zfp206 expression vector. Zfp206 expression activated expression from *Oct4*, *Klf5*, and *Jarid1c* promoters and the *Klf4* enhancer (Fig. 8). In contrast, Zfp206 activated reporter expression from the *Meis2*, *mir-124a1*, and *mir-124-a2* promoters. No activation or repression was observed on the *Meis2* and *Pitx2* promoters. These results completely agree with the knockdown results presented in Fig. 6 and thus provide strong evidence that Zfp206 is a transcriptional regulator that binds to specific genomic locations and thereby activates or represses its target genes. Our results also indicate that Zfp206 functions in ES cells as a pluripotency regulator by modulating a specific transcriptional network, which is composed of many other transcription factors.

Zfp206 Physically Interacts with Oct4 and Sox2—It has been shown that Nanog is a key component of the pluripotency network in part by directly interacting with Oct4, Sox2, Zfp281, and other transcription factors (18). Comprehensive mapping of the binding sites of these networks showed they also co-localize to a large number of genomic locations (6). This suggests that co-localization of these TF correlates with protein-protein interaction. Because Zfp206 is a pluripotency factor (13) and co-localizes substantially with other pluripotency factors, we reasoned that Zfp206 might also physically interact

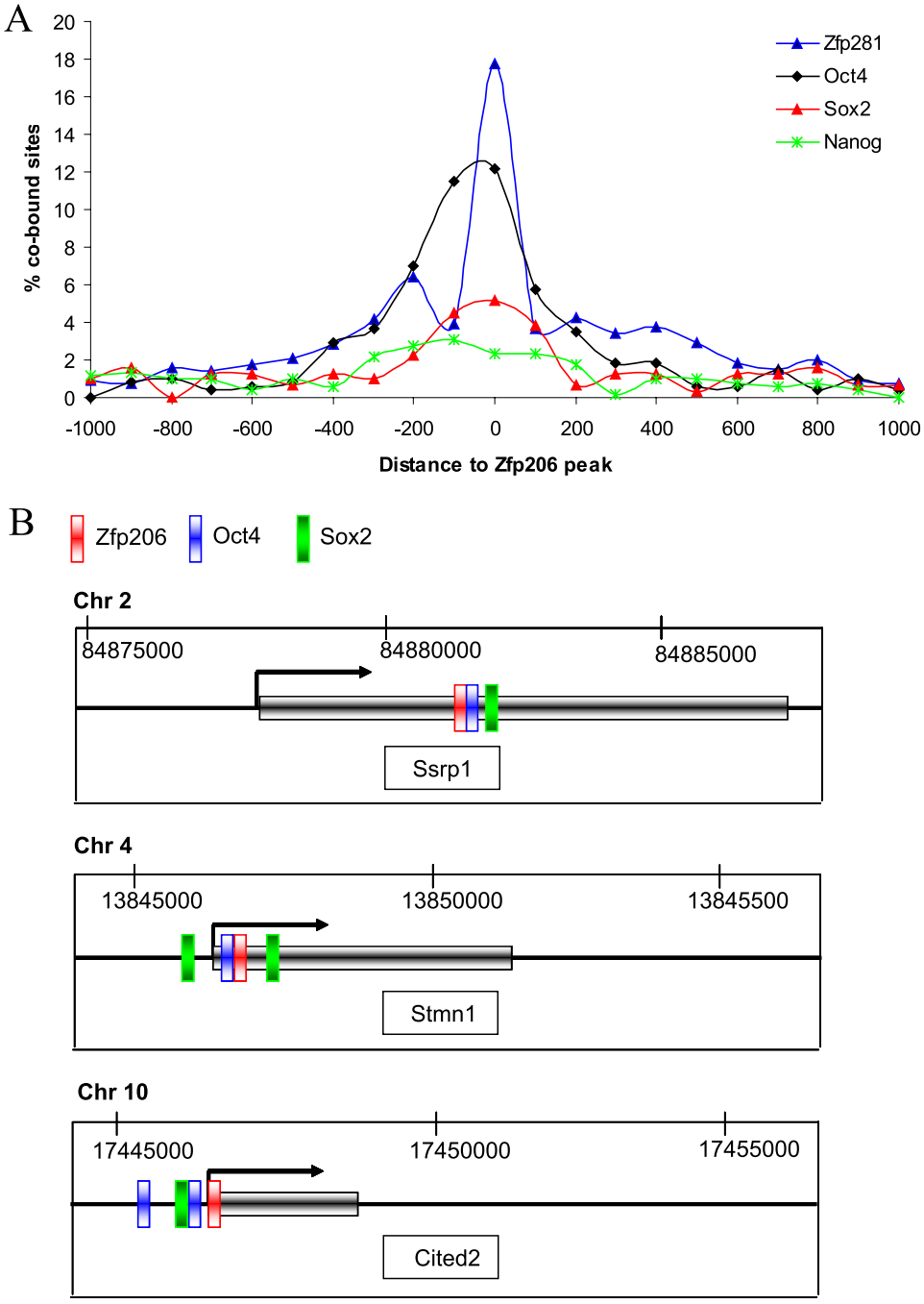

FIGURE 5. **Zfp206 binding co-localizes with Zfp281, Oct4, Sox2, and Nanog.** *A*, using Zfp206 sites as reference, the nearest binding sites for each of the other TF were calculated. The results indicate substantial overlap of Zfp206 sites with sites for Zfp281 and Oct4 but less so with Sox2 and Nanog. The data are expressed as frequencies of co-occupancies for Zfp206 with the indicated TF at interval distances indicated from the Zfp206-binding site. *B*, examples of close proximities of Oct4- (*blue bars*), Sox2- (*green bar*), and Zfp206 (*red bar*)-binding sites at three target genes, Ssrp1, Stmn1, and Cited2.

with Oct4 and Sox2. To test this hypothesis Oct4 and Sox2 were tagged with V5 epitope, cloned into an expression vector, and transfected into ESC. The V5 antibody was used to immunoprecipitate the tagged Oct4 and Sox2. The pull-down products were then probed on Western blots using a Zfp206 antibody to test whether Oct4 and Sox2 interacted with Zfp206 (Fig. 9). In a converse experiment, Zfp206 antibody was used to pull down endogenous Zfp206 protein in the transfected cells expressing Oct4-V5 or Sox2-V5. These pull down products were then tested by Western blot using V5 antibody to determine whether Zfp206 interacted with Oct4 and Sox2 (Fig. 9). In both sets of experiments, the results showed that Oct4 and Sox2 were co-immunoprecipitated with Zfp206. A control experiment with an antibody specific for the transcription factor Rcor1, which is also expressed in ESC, showed no co-IP with Zfp206, Oct4, or Sox2 (Fig. 9). These results indicate that in ESC Zfp206 associates with Oct4 and Sox2 in a macromolecular complex.

DISCUSSION

The maintenance of a pluripotent state in ESC is governed, in part, by a complex transcriptional regulatory network. Oct4 and Sox2 are often considered the core transcription factors in this network, but there are other important components including Nanog, Esrrb, Tcf3, Tcl1, Zfp281, Zic3, Sall4 (3, 7–12), and the topic of this study, Zfp206 (13). We have performed ChIP studies and identified more than 3500 binding sites for Zfp206 in ESC. A comparison of the binding landscapes of Zfp206 and those generated for Oct4 and Sox2 shows substantial overlaps of genes targeted by this trio of transcription factors (Fig. 4). There were 183 genes identified as targets of all three TF, which is highly statistically significant (p = 2.43E-150) (see supplemental Table S2 for complete list). This significant overlap of targets genes reveals that Zfp206 works in concert with Oct4 and Sox2 to regulate ESC pluripotency. Our demonstration of biochemical associations of Oct4 and Sox2 with Zfp206 (Fig. 9) further supports that these TF are components of a regulatory complex that controls expression of genes required for pluripotency. It remains to be determined whether the interactions of Zfp206 with Oct4 and Sox2 are direct protein-protein interactions and, if so, what are the domains within each protein that physically interact. It remains possible that the associations we have found are indirect, perhaps mediated through an adapter protein or, alternatively, through co-localization on the chromatin. Additional biochemical experiments are needed to clarify the biophysical nature of the regulatory complex. Such infor-

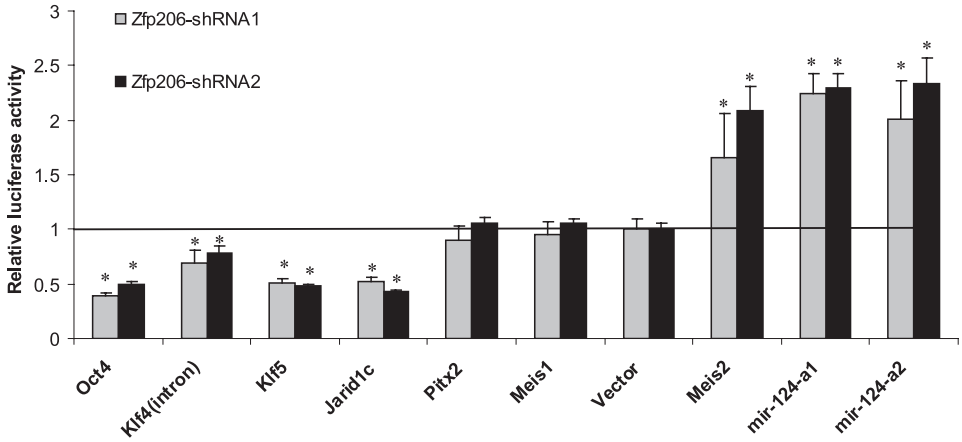

FIGURE 6. **Zfp206 activates or represses its target genes.** Luciferase reporter constructs containing the promoter and flanking Zfp206-binding site of the indicated Zfp206 target genes were transfected into Zfp206 knockdown or control ESC. For Klf4, the intronic Zfp206-binding site from intron-1 was linked to a minimal promoter to drive luciferase. Knockdown of Zfp206 expression by two shRNA constructs resulted in decreased expression of four genes, increased expression of three genes, and no change in three others. The luciferase activities are expressed as fold differences relative to an empty vector control. The mean values are shown with *error bars* (S.D.) from four independent assays. The statistical significance was assessed by comparing data to empty vector control, using Student's *t* test (*, $p < 0.01$).

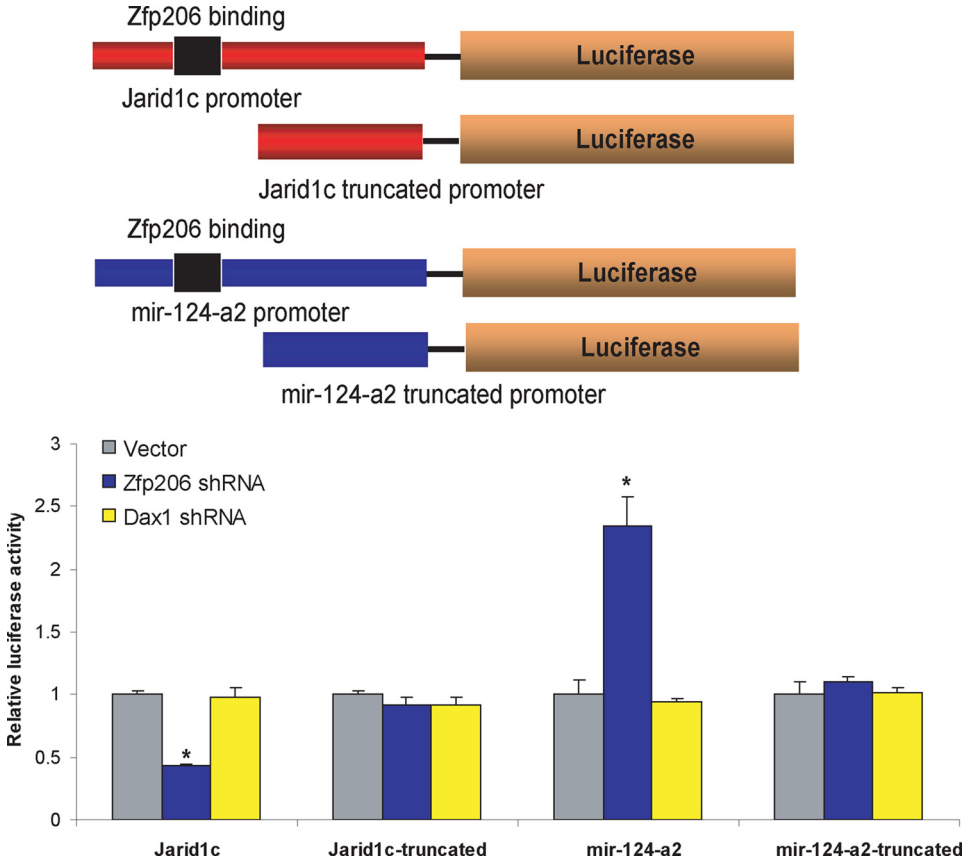

FIGURE 7. **Zfp206 binding is necessary for activation or repression of target gene expression.** The luciferase reporter constructs for Jarid1c and mir-124-a2 promoters were deleted of their Zfp206-binding sites. The activities of the truncated and full versions were compared. Removal of the Zfp206-binding sites resulted in a loss of regulation in response to Zfp206 knockdown by shRNA. An empty shRNA vector and shRNA direct against Dax-1 had no effect on Zfp206 target genes. The luciferase activities are expressed as fold differences relative to an empty vector control. The mean values are shown with *error bars* (S.D.) from four independent assays. Statistical significance was assessed by comparing data with empty vector control, using Student's *t* test (*, $p < 0.01$).

mation would potentially shed light on how Zfp206 can activate certain promoters while repressing the activity of others.

Zfp206 targets a significant ($p < 5.4E-25$) number of genes that are involved in nucleic acid metabolism and RNA transcription (Table 1). Thus, Zfp206 may maintain pluripotency by controlling the expression of other transcription factors. This is consistent with other pluripotency factors such as Oct4, Sox2, Nanog, and Klf4, which also target numerous DNA-binding proteins in ESC (5, 6). It is interesting that Zfp206 targets Oct4 and Sox2, and conversely Oct4 and Sox2 target Zfp206, thus establishing an interconnected, regulatory loop among these pluripotency-regulating transcription factors. A mutually reinforcing control of transcription factor expression in ESC was similarly observed for Nanog, Oct4, and Sox2 (16). In addition, each of these transcription factors autoregulates their own expression levels. Thus, it appears that transcriptional regulation of ESC pluripotency is tightly controlled by interconnected and autoregulatory loops involving a key set interacting transcription factors, a situation that has repeatedly been observed in the exquisite control of cell fates in development (19).

The mapping of more than 3500 Zfp206-binding sites enabled the elucidation of a consensus sequence motif for the binding of Zfp206 to DNA. The motif, GCGCATGCGC, is a perfect palindrome, suggesting that Zfp206 may bind to DNA as a homodimer. Consistent with this idea is the presence of a SCAN domain in Zfp206. SCAN domains are specific to vertebrates and highly conserved and have been shown to mediate protein-protein interactions of zinc finger transcription factors (20). Preliminary data indicate that Zfp206 does form homodimers, likely through the SCAN domains.[3] It is possible that Zfp206 also forms heterodimers with other SCAN-containing transcriptional regulators expressed in undifferentiated ESC, such as Scnd1 and Zfp110, which are also down-

[3] H.-B. Yu, G. Kunarso, F. H. Hong, and L. W. Stanton, unpublished data.

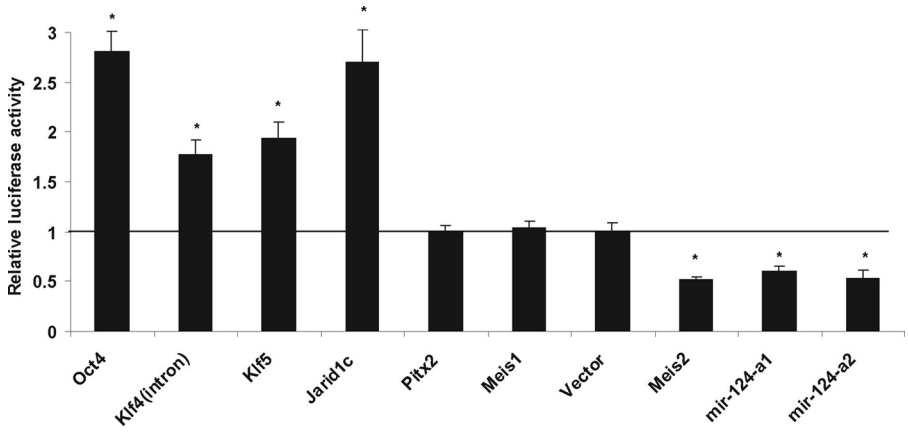

FIGURE 8. **Zfp206 regulates expression of target genes in somatic cells.** The luciferase reporter constructs described in Fig. 6 were co-transfected into HEK293T somatic cells together with a Zfp206 expression construct. The transcriptional responses to Zfp206 overexpression were perfectly opposed to the results seen by Zfp206 knockdown by shRNA treatment shown in Fig. 6. The luciferase activities are expressed as fold differences relative to HEK293 cells co-transfected with an empty expression vector. The mean values are shown with *error bars* (S.D.) from four independent assays. Statistical significance was assessed by comparing data to empty vector control, using Student's *t* test (*, $p < 0.01$).

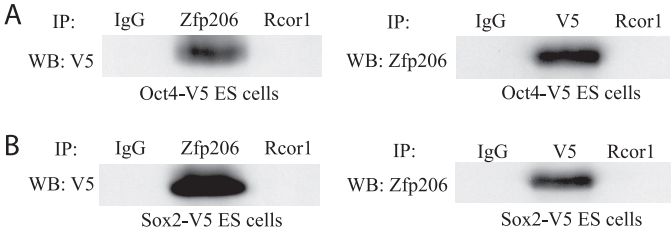

FIGURE 9. **Zfp206 physically interacts with Oct4 and Sox2 proteins.** ESC were transfected to express V5 epitope tagged Oct4 (Oct4-V5) (*A*) or Sox2 (Sox2-V5) (*B*). Zfp206 antibody was used to immunoprecipitate (*IP*) Zfp206 and the co-immunoprecipitation products were analyzed by Western blot (*WB*) with a V5 antibody to test whether Zfp206 interact with Oct4 and Sox2. V5 antibody was used to immunoprecipitate Oct4-V5 (*A*) and Sox2-V5 (*B*), and co-immunoprecipitation products were analyzed by Western blot with a Zfp206 antibody. The results show co-IP of Zfp206 with Oct and co-IP of Zfp206 with Sox2. No co-IP of Oct4 or Sox2 was observed in control experiments using an antibody directed against the ESC-expressed transcription factor Rcor1.

regulated upon differentiation of ESC,[3] or Zscan4, which is a target of Zfp206 in ESC (14).

It is interesting that a strong consensus motif was identified for Zfp206 and that 65% of the binding sites identified conformed well to this motif. It was not certain that a common sequence motif would be identifiable, given that Zfp206 has 14 zinc finger domains and several other isoforms, generated by alternative splicing, that have only a subset of these domains (5–14 fingers). It has been shown that transcription factors with C_2H_2-type zinc fingers make specific contacts to DNA through 3–5 fingers (21). Thus, Zfp206, which contains up to 14 zinc fingers, might be expected to display a diversity of binding motifs. This remains a possibility because 35% of the Zfp206 sites did not conform to the consensus binding motif. Indeed, the two regions in the *Oct4* promoter region that bind Zfp206 lack sequences that match closely to the motif. A search for a common sequence among this minority of sites did not reveal any other consensus motifs. It is also possible that the multiple zinc fingers of Zfp206 are involved in protein-protein interactions as has been demonstrated for other C_2H_2 type zinc finger proteins (22).

The results presented here have defined additional connections within the circuitry of the complex regulatory network that governs differentiation of ES cells. Our understanding of the transcriptional regulatory networks in ESC has expanded rapidly in recent years with advances in capabilities to comprehensively define the transcriptome using gene expression arrays and to map target genes genome-wide by chromatin immunoprecipitation coupled with deep sequencing. The knowledge gained about the architecture of this complex biological system continues to advance our understanding of embryonic development and will lead to improvements in engineering stem cells for production of clinically useful cells and tissues.

Acknowledgments—We are thankful for the expert technical advice of Linda Lim and the generosity of Zheng-Xu Wang for sharing of reagents. Thanks also go out to Ralf Jauch and Rory Johnson for advice on the experiments and preparation of the manuscript.

REFERENCES

1. 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
2. Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C. L., Ruan, Y., Lim, B., and Ng, H. H. (2006) *Nat. Genet.* **38,** 431–440
3. Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I. R. (2006) *Nature* **442,** 533–538
4. Takahashi, K., and Yamanaka, S. (2006) *Cell* **126,** 663–676
5. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y. H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W. K., Clarke, N. D., Wei, C. L., and Ng, H. H. (2008) *Cell* **133,** 1106–1117
6. Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S. H. (2008) *Cell* **132,** 1049–1061
7. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003) *Cell* **113,** 631–642
8. Zhang, X., Zhang, J., Wang, T., Esteban, M. A., and Pei, D. (2008) *J. Biol. Chem.* **283,** 35825–35833
9. Tam, W. L., Lim, C. Y., Han, J., Zhang, J., Ang, Y. S., Ng, H. H., Yang, H., and Lim, B. (2008) *Stem Cells* **26,** 2019–2031
10. Wang, Z. X., Teh, C. H., Chan, C. M., Chu, C., Rossbach, M., Kunarso, G., Allapitchay, T. B., Wong, K. Y., and Stanton, L. W. (2008) *Stem Cells* **11,** 2791–2799
11. Lim, L. S., Loh, Y. H., Zhang, W., Li, Y., Chen, X., Wang, Y., Bakre, M., Ng, H. H., and Stanton, L. W. (2007) *Mol. Biol. Cell* **18,** 1348–1358
12. Zhang, J., Tam, W. L., Tong, G. Q., Wu, Q., Chan, H. Y., Soh, B. S., Lou, Y., Yang, J., Ma, Y., Chai, L., Ng, H. H., Lufkin, T., Robson, P., and Lim, B. (2006) *Nat. Cell Biol.* **8,** 1114–1123

13. Wang, Z. X., Kueh, J. L., Teh, C. H., Rossbach, M., Lim, L., Li, P., Wong, K. Y., Lufkin, T., Robson, P., and Stanton, L. W. (2007) *Stem Cells* **25,** 2173–2182
14. Zhang, W., Walker, E., Tamplin, O. J., Rossant, J., Stanford, W. L., and Hughes, T. R. (2006) *Nucleic Acids Res.* **34,** 4780–4790
15. Wang, Z. X., Teh, C. H., Kueh, J. L., Lufkin, T., Robson, P., and Stanton, L. W. (2007) *J. Biol. Chem.* **282,** 12822–12830
16. Rodda, D. J., Chew, J. L., Lim, L. H., Loh, Y. H., Wang, B., Ng, H. H., and Robson, P. (2005) *J. Biol. Chem.* **280,** 24731–24737
17. Pavesi, G., Mauri, G., and Pesole, G. (2001) *Bioinformatics* **17,** (Suppl. 1) S207–S214
18. Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., and Orkin, S. H. (2006) *Nature* **444,** 364–368
19. Davidson, E. H., and Levine, M. S. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 20063–20066
20. Edelstein, L. C., and Collins, T. (2005) *Gene* **359,** 1–17
21. Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000) *Annu. Rev. Biophys. Biomol. Struct.* **29,** 183–212
22. Brayer, K. J., and Segal, D. J. (2008) *Cell Biochem. Biophys.* **50,** 111–131