nature genetics

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells

Yuin-Han Loh[1,2,7], Qiang Wu[1,7], Joon-Lin Chew[1,2,7], Vinsensius B Vega[3], Weiwei Zhang[1,2], Xi Chen[1,2], Guillaume Bourque[3], Joshy George[3], Bernard Leong[3], Jun Liu[4], Kee-Yew Wong[5], Ken W Sung[3], Charlie W H Lee[3], Xiao-Dong Zhao[4], Kuo-Ping Chiu[3], Leonard Lipovich[3], Vladimir A Kuznetsov[3], Paul Robson[2,5], Lawrence W Stanton[5], Chia-Lin Wei[4], Yijun Ruan[4], Bing Lim[5,6] & Huck-Hui Ng[1,2]

Oct4 and Nanog are transcription factors required to maintain the pluripotency and self-renewal of embryonic stem (ES) cells. Using the chromatin immunoprecipitation paired-end ditags method, we mapped the binding sites of these factors in the mouse ES cell genome. We identified 1,083 and 3,006 high-confidence binding sites for Oct4 and Nanog, respectively. Comparative location analyses indicated that Oct4 and Nanog overlap substantially in their targets, and they are bound to genes in different configurations. Using *de novo* motif discovery algorithms, we defined the *cis*-acting elements mediating their respective binding to genomic sites. By integrating RNA interference–mediated depletion of Oct4 and Nanog with microarray expression profiling, we demonstrated that these factors can activate or suppress transcription. We further showed that common core downstream targets are important to keep ES cells from differentiating. The emerging picture is one in which Oct4 and Nanog control a cascade of pathways that are intricately connected to govern pluripotency, self-renewal, genome surveillance and cell fate determination.

ES cells are pluripotent cells derived from the inner cell mass (ICM) of the mammalian blastocyst. They are capable of indefinite self-renewing expansion in culture. Depending on culture conditions, these cells can differentiate into a variety of cell types[1]. The ability to steer ES cell differentiation into specific cell types holds great promise for regenerative medicine[2–4].

Oct4, Sox2 and Nanog are key regulators essential for the formation and/or maintenance of the ICM during mouse preimplantation development and for self-renewal of pluripotent ES cells[5–10]. Oct4 is a POU domain–containing transcription factor encoded by *Pou5f1*. In the absence of Oct4, pluripotent cells *in vivo* (epiblast) and *in vitro* (ES cell) both revert to the trophoblast lineage. This implicates Oct4 as an important regulatory molecule in the initial cell fate decisions during mammalian development. Additionally, increasing the expression of Oct4 above the endogenous levels in ES cells leads to differentiation toward the extraembryonic endoderm lineage[7]. These divergent effects of Oct4 suggest that Oct4 transcriptionally regulates genes involved in coordinating multiple cellular functions. Oct4 is known to bind to a classical octamer sequence, ATGCAAAT, and in ES cells, it often binds in partnership with Sox2, which binds to a neighboring sox element[11,12]. Nanog, a homeodomain–containing protein, was identified as a factor that can sustain pluripotency in ES cells even in the absence of leukemia inhibitory factor (LIF)[9,10]. Nanog-null embryos seem to be able to initially give rise to the pluripotent cells, but these cells then immediately differentiate into the extraembryonic endoderm lineage. During development, Nanog function is required at a later point than the initial requirement for Oct4, but both are required for the maintenance of pluripotency.

To understand how Oct4 and Nanog maintain pluripotency, we sought to identify the physiological targets of these transcription factors in mouse ES cells. We made use of the recently developed paired-end ditag (PET) technology to characterize chromatin immunoprecipitation (ChIP)-enriched DNA fragments and achieved unbiased, genome-wide mapping of transcription factor binding sites. This method extracts a pair of signature tags from the 5′ and 3′ ends of each DNA fragment, concatenates these PETs for efficient sequencing and maps them to the genome[13,14]. Here we combine this ChIP-PET identification of Oct4 and Nanog binding sites with RNA interference (RNAi) analyses to demonstrate the regulation of target gene expression. Overexpression of Nanog in ES cells further identified upregulated or downregulated genes. This comprehensive analysis

[1]Gene Regulation Laboratory, Genome Institute of Singapore, Singapore 138672. [2]Department of Biological Sciences, National University of Singapore, Singapore 117543. [3]Information & Mathematical Sciences Group and [4]Cloning and Sequencing Group, Genome Institute of Singapore, Singapore 138672. [5]Stem Cell & Developmental Biology, Genome Institute of Singapore, Singapore 138672. [6]Harvard Institutes of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to Y.R. (ruanyj@gis.a-star.edu.sg), or H.-H.N. (nghh@gis.a-star.edu.sg).

Received 30 November 2005; accepted 6 February 2006; published online 5 March 2006; doi:10.1038/ng1760

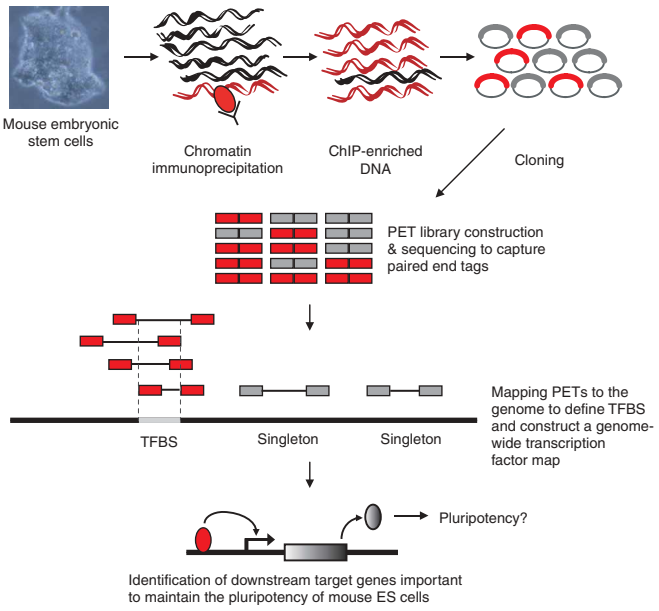

Figure 1 Schematic diagram of genome-wide mapping of Oct4 and Nanog binding sites using ChIP-PET. Mouse embryonic stem cells cultured under feeder-free conditions were treated with formaldehyde to mediate covalent cross-links between DNA and proteins. The chromatin was fragmented by sonication. Immunoprecipitation using a specific antibody was used to capture the transcription factor bound to target sites (shown in red). The ChIP-enriched DNA was first cloned into a plasmid-based library, and we then used restriction enzymes to transform this original library into one that contained concatenated paired-end ditag (PET) sequences[13]. Each tag is 18 bp in length, and each ditag represents the 5′-most and 3′-most ends of the ChIP-enriched DNA fragments cloned into the original library. This second library increases the throughput of analysis, as each sequencing read identifies 10 to 15 PETs representative of 10 to 15 ChIP-enriched genomic fragments. We refer to this as the ChIP-PET methodology[14]. The concatenated PETs were sequenced and their locations were mapped to the mouse genome to demarcate the boundaries of transcription factor ChIP-enriched DNA. PET overlaps of four or more members were empirically determined to be high-confidence transcription factor binding sites. Random recovery of genomic DNA was observed in the form of PET singletons. To further establish the importance of the selective downstream targets of Oct4 and Nanog, we depleted the transcripts encoding these factors by RNAi and demonstrated their roles in maintaining ES cells in a nondifferentiated state. TFBS, transcription factor binding site.

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

uncovers a complex network connecting the regulators important in maintaining ES cell pluripotency.

RESULTS

Global mapping of Oct4 and Nanog binding sites by ChIP-PET

To better understand the roles of Oct4 and Nanog in self-renewal and pluripotency, we set out to determine the downstream targets of these transcription factors in undifferentiated mouse ES cells by the ChIP-PET method (**Fig. 1** and **Supplementary Note** online)[14]. Although the majority of the PETs were located in the genome discretely (classified as PET singletons), about 25% of PETs from both ChIP-PET libraries were found overlapping with other PETs, thus representing clusters. These PET cluster–defined genomic loci represent potential interaction sites in the genome. We hereafter refer to PET clusters with two overlapping members as PET2, for clusters with three overlapping members as PET3, and so forth.

Next, we empirically determined the minimum required size of a PET cluster to identify an authentic binding site with high confidence (that is, not a result of background noise). For the Oct4 dataset, we selected 115 PET clusters for validation (**Supplementary Figure 1** and **Supplementary Table 1** online). All of the clusters with five or more overlapping members ('PET5+') and 38 of the 40 PET4 clusters showed enrichment above background. Among the PET3 clusters, Oct4 bound three out of the 34 loci tested. As 91% of the PET3 clusters were not enriched, a cluster size of at least four PETs was selected as a cutoff for maximum identification of high-confidence Oct4-binding sites; 1,083 clusters with four or more overlapping members (PET4+) were identified (**Supplementary Table 2** online). In further validation of the Oct4 ChIP-PET data, we found that the PET profile precisely paralleled that detected by real-time PCR on two previously characterized targets of Oct4, *Pou5f1* and *Nanog*[15] (**Supplementary Figure 2** online). This attests to the reliability of this approach for high-resolution mapping of transcription factor binding sites in living ES cells.

For the Nanog data set, we selected 100 PET clusters for validation (**Supplementary Figure 3** online). All PET5+ clusters and 20 out of the 21 PET4 clusters showed enrichment above background. Among the PET3 clusters, Nanog bound 12 out of the 16 loci tested. As 25% of the PET3 clusters were not enriched, we chose clusters of PET4+ as high-confidence Nanog binding loci; 3,006 of these were identified (**Supplementary Table 2**). To exclude the possibility that the polyclonal antibody we used cross-reacted with other proteins, we further validated these 100 loci by repeating the ChIP-PCR assay using an ES cell line expressing hemagglutinin (HA)-tagged Nanog (**Supplementary Figure 4** online). Notably, we observed PET clusters over the regulatory regions for *Pou5f1*, *Sox2* and *Nanog* (**Supplementary Figure 5** online), and the binding profiles were validated by real-time PCR quantification of Nanog ChIP DNA.

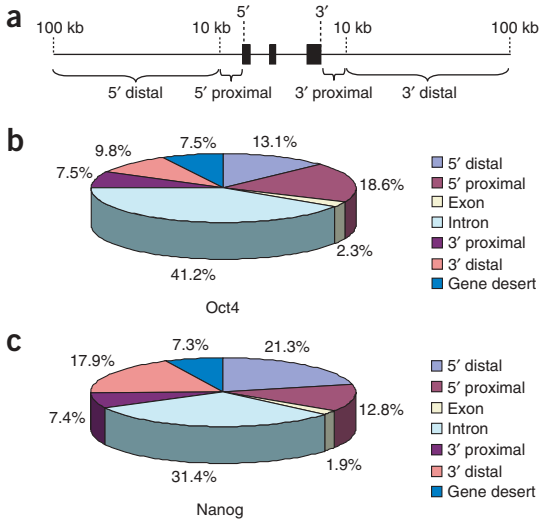

Figure 2 Distribution of Oct4 and Nanog binding sites. (**a**) Schematic diagram illustrating the definition of the location of a binding site in relation to a transcription unit. 5′ distal, 5′ proximal, 3′ proximal and 3′ distal regions are depicted in the 100 kb upstream and 100 kb downstream of the transcriptional unit. (**b**) Locations of Oct4 binding sites relative to the nearest transcription units. The percentages of binding sites at the respective locations are shown. (**c**) Locations of Nanog binding sites relative to the nearest transcription units. The percentage of binding sites at each location is shown.

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

Binding site distribution relative to gene structure

As a first step to identify genes that are potentially regulated by Oct4 or Nanog, we annotated all the binding site loci with positional information relative to the nearest gene. For loci within 100 kb of a gene, their relative positions were annotated as 5′ distal (10–100 kb upstream), 5′ proximal (0–10 kb upstream), intragenic (contained within the respective genes), 3′ proximal (0–10 kb downstream) or 3′ distal (10–100 kb downstream; **Fig. 2a** and **Supplementary Table 2**). Loci mapping >100 kb away from the nearest gene were annotated as residing in gene deserts. All the distinct genes associated with the binding sites were further annotated with the Panther classification system[16].

About 44% of the Oct4 binding sites mapped within a gene, with 437 mapping to introns and 25 to exons (**Fig. 2b**). The 5′ proximal region contained 196 Oct4 loci (19%), whereas 140 Oct4 loci (13%) mapped in the 5′ distal regions of genes. The number of Oct4 binding sites mapped downstream of genes was 79 in the 3′ proximal (including *Sox2*) and 104 in the 3′ distal regions. Of the Nanog clusters, 2,786 clusters were located within 100 kb of transcription units (**Fig. 2c**). Nine hundred forty-four (31%) of the Nanog binding sites were found within introns. Six hundred forty-one loci (21.3%) and 386 loci (12.8%) were bound by Nanog at 5′ distal and 5 proximal regions, respectively. Seven hundred fifty-eight Nanog loci (25.3%) were found at 3′ downstream regions of the genes.

Targeting of Oct4 and Nanog to the genome

As Oct4 and Nanog are among the key regulators in ES cells, we examined whether there is cross-talk between the two factors and how they extend their circuitries to the different genes. Notably, Nanog was found to bind to an extended region of the *Pou5f1* promoter covering conserved regions 2 to 4, whereas Oct4 was found only at conserved region 4 (**Fig. 3a**)[15,17]. To further investigate the relationship of Nanog and Oct4 occupancies on a global scale, we generated a list of genes containing Nanog and Oct4 binding sites anywhere within the vicinity of 50 kb of a transcription unit (**Fig. 3b**). A substantial proportion of the genes (345, representing 44.5% of Oct4-bound genes) were occupied by both Nanog and Oct4 (**Supplementary Table 3** online). The result also showed Nanog-Oct4 colocalization as well as independent binding of Nanog and Oct4 to the targeted genes (**Fig. 3c**).

Besides protein-coding genes, both Oct4 and Nanog localized to genes encoding microRNAs (**Fig. 3d**). Nanog binds to sites within 6 kb of four microRNA genes: *mir296*, *mir302*, *mir124a* and *mir9-2*. For *mir296*, *mir124a* and *mir9-2*, there were no other known genes in close proximity to the Nanog loci. For *mir135*, the Nanog cluster was found to bind 30 kb away. Oct4 bound in juxtaposition with Nanog at sites near the *mir296* and *mir302* genes.

Defining the *cis* elements mediating Oct4 and Nanog binding

The ChIP-PET method provides high-resolution mapping of binding sites, and the average length of the PET cluster overlaps for binding loci was around 100 bp. This high resolution increases the likelihood of finding motifs using *de novo* motif discovery algorithms such as Weeder and NMICA[18,19]. Notably, the predominant motif found in our computational search of the Oct4 data set (**Fig. 4a**)

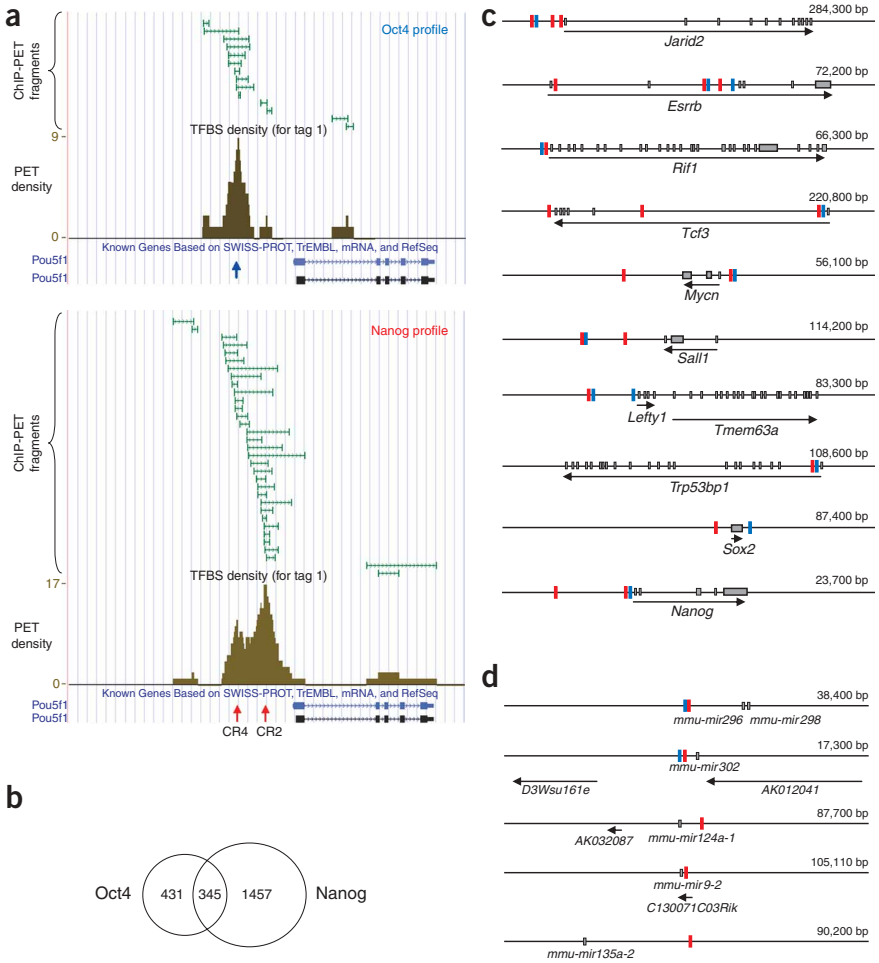

Figure 3 Oct4 and Nanog binding site configurations at genomic locations. (**a**) A screen shot of the T2G browser showing Oct4 (upper panel) and Nanog (lower panel) PET clusters at *Pou5f1*. Each horizontal green line represents a DNA fragment mapped to the genome. PET density (in brown) shows the profile of the transcription factor binding and is based on the number of overlapping DNA fragments. The peaks of Nanog binding are highlighted by red arrows, and the peak of Oct4 binding is highlighted by a blue arrow. CR2 refers to conserved region 2. CR4 contains a Sox2-Oct4 motif[15]. (**b**) Common targets (overlap) between Nanog- and Oct4-bound genes (analyzed 50 kb upstream and 50 kb downstream of each gene) (**c**) Different configurations of Oct4 (blue block) and Nanog (red block) binding to genes. Exons are depicted as gray boxes. The arrow indicates the direction and body of a gene, extending from first exon to last exon based on University of California, San Diego mouse genome coordinates. The numbers on the right indicate the window span represented by each plot. (**d**) Plots showing the presence of Oct4 binding sites (blue block), Nanog (red block) binding sites or both at genomic regions containing microRNA genes. The microRNAs are depicted as gray blocks. Each arrow represents a gene. The numbers on the right indicate the window span represented by each plot. All known genes within the respective windows are shown.

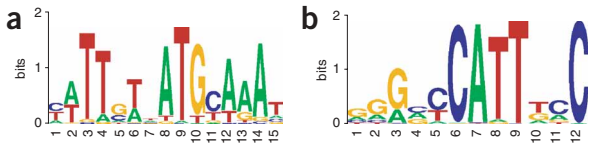

Figure 4 *De novo* prediction of motifs that mediate specific transcription factor–DNA interaction. (**a**) A Sox2-Oct4 joint motif identified from the Oct4 ChIP-PET dataset. (**b**) A motif identified from the Nanog ChIP-PET dataset.

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

was a perfect match to the sox-oct composite element consensus derived from six previously characterized Oct4-Sox2 target genes[15]. This motif, discovered by both algorithms, was present in a high percentage of the Oct4 binding loci (**Supplementary Note**), suggesting a Sox2-Oct4 binary complex binding to these target genes. Sequential ChIP of Oct4 and Sox2 at six loci (**Supplementary Figure 6** online), three of which had not previously been described (*Tcf3*, *Trp53*, *Mycn*), further demonstrates that both Sox2 and Oct4 bind to these sites. We therefore suggest that one of the main mechanisms for targeting Oct4 to its genomic sites is through the sox-oct motif via a cooperative interaction with Sox2.

We also predicted a CATT-containing motif enriched over genomic background in the Nanog ChIP-PET dataset using the NMICA algorithm (**Fig. 4b**)[19]. Notably, this CATT-containing motif has some overlap with an ATTA motif previously defined biochemically[10]. The interaction between Nanog and this CATT-containing motif was confirmed by EMSA using probes to a number of the Nanog binding loci (**Supplementary Figure 7** online). Notably, this motif was not found by Weeder, which we suspect is due to the algorithm (**Supplementary Note**) and may be related to the strength or length of the specific signal.

Genome-wide analyses of gene regulation by Oct4 and Nanog

To determine the functional relevance of the Oct4 and Nanog binding sites on the transcriptional regulation of their associated genes, we perturbed Oct4 and Nanog expression in mouse ES cells by two methods. First, we induced ES cells to differentiate. As our goal is to examine the change in expression profiles associated with differentiation, we used three chemical treatments (retinoic acid (RA), dimethyl sulfoxide (DMSO) and hexymethyl-bis-acetamide (HMBA)) to avoid chemical-specific modulation of gene expression. Microarrays with over 16,000 gene probes were used to interrogate gene expression changes. We first clustered the gene expression to separate differentiation-induced and differentiation-repressed genes (**Supplementary Note**). Both Oct4 and Nanog were substantially repressed in all three treatments. We subsequently scanned all of these genes (50 kb upstream to 50 kb downstream) for the presence of the Oct4 and Nanog binding sites that we had identified by ChIP-PET (**Supplementary Table 4** online). The data showed enrichment of Oct4- or Nanog-bound genes that were induced and repressed upon differentiation (**Fig. 5**). This suggests that Oct4 and Nanog can activate or repress transcription. The genome-wide analysis also showed that there are more Oct4- or Nanog-bound genes downregulated than induced upon differentiation, suggesting that Oct4 and Nanog have a dominant role in activating the transcription of ES cell–specific genes. In addition, a third plot interrogates the presence of both factors and showed that binding of two factors was more strongly correlated with genes that were downregulated upon differentiation than with genes that were upregulated.

The second method to determine functional relevance of binding sites was to deplete ES cells of Oct4 or Nanog by RNAi and examine differential gene expression again by microarray analysis. Our Oct4 and Nanog siRNAs were specific, as the effects of knockdown could be rescued by coexpression of the respective RNAi-immune ORFs (**Supplementary Figures 8** and **9** online). For each differentially expressed gene, we determined if a Oct4 or Nanog binding site was present (**Fig. 6a**,**b**). Of the 4,711 statistically selected genes (median false discovery rate < 0.001) from the *Pou5f1* knockdown experiment, 394 contained Oct4 binding sites (**Supplementary Table 5** online). After *Nanog* knockdown, 475 of the 2,264 differentially expressed genes were

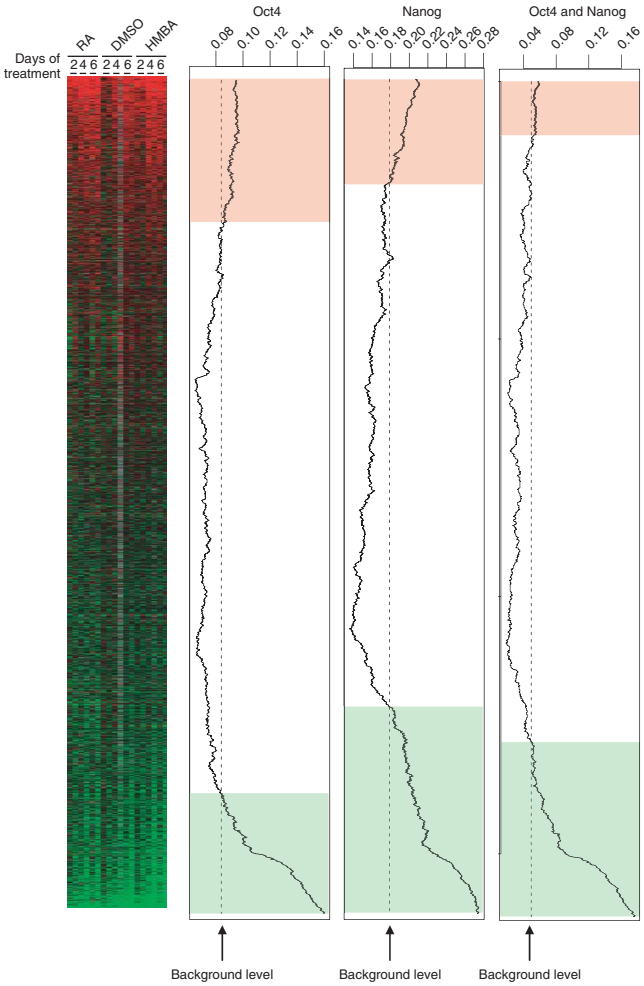

Figure 5 Genome-wide association of Oct4 and Nanog binding sites with differentiation profiles of mouse ES cells. Hierarchical clustering was performed on gene expression data for 16,223 probes obtained from differentiated mouse ES cells. Genes are rank ordered by degree of induction (red) and repression (green) by RA, DMSO and HMBA, relative to undifferentiated control cells at days 2, 4 and 6 (leftmost panel). The three plots at right show the corresponding numbers (moving averages) of gene probes that have associated Oct4 and/or Nanog binding sites. The pink and light green shaded areas indicate those genes that have Oct4 and/or Nanog binding sites at frequencies significantly greater than background ($P < 10^{-6}$). A value of 0.16 on the *y*-axis means that 320 genes were bound in a sliding window of 2,000. The dashed line (background level) indicates the expected average (that is, the ratio of number of gene probes with associated binding regions over total number of interrogating gene probes).

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

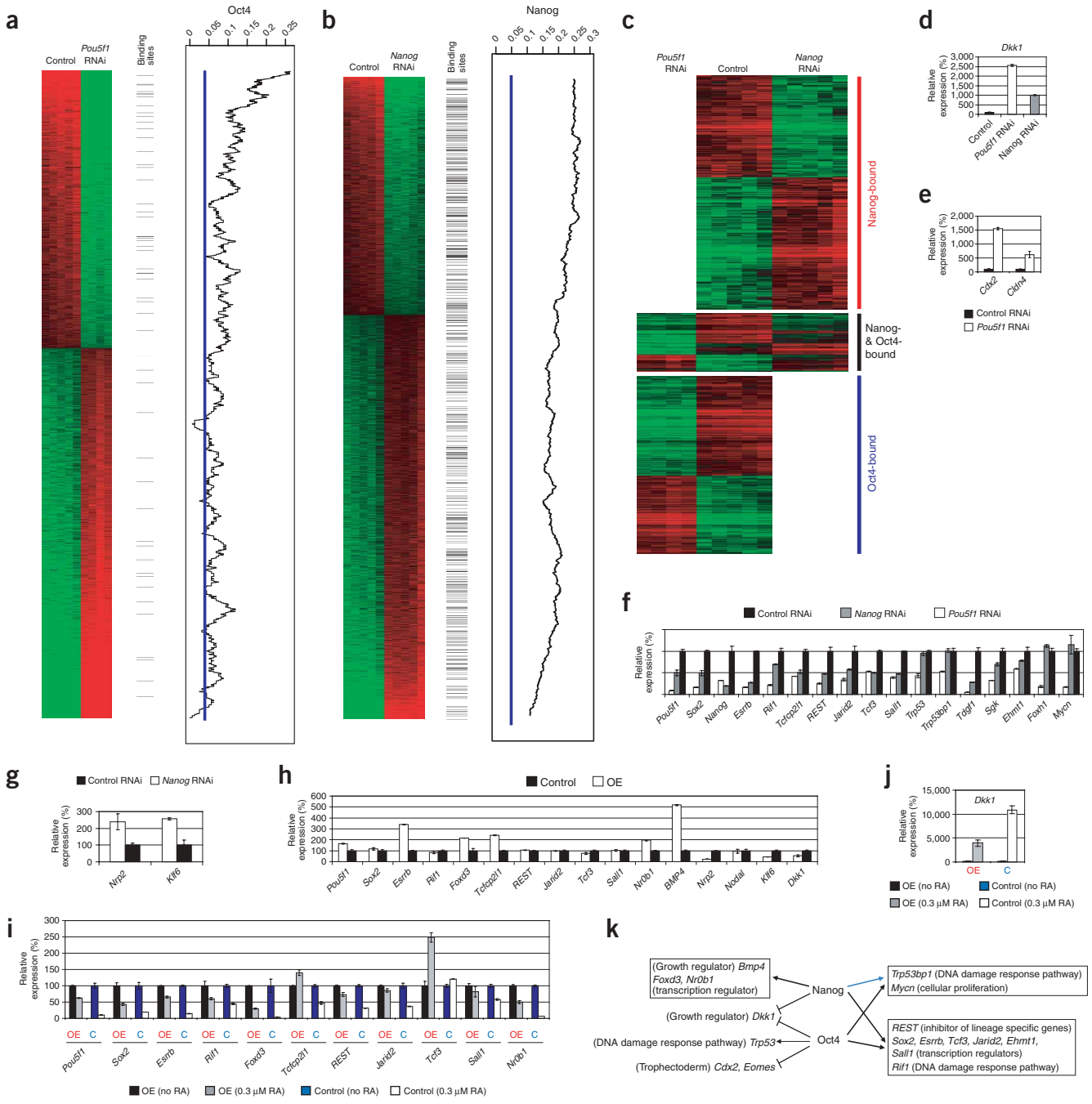

Figure 6 Genome-wide association of Oct4 and Nanog binding sites with expression profiles of mouse ES cells depleted of Oct4 or Nanog. (**a**) Expression profile of genes differentially expressed after *Pou5f1* knockdown that were selected using Significance Analysis of Microarrays (SAM) analysis. The genes were sorted by the average expression ratio and mean centered. The horizontal black lines denote the presence of Oct4 binding sites. At right is the moving window average of number of probes associated with Oct4 binding sites. The expected number of genes with binding site association is defined as the ratio of the number of genes with binding sites over the total number of genes interrogated (vertical blue line). (**b**) The same analysis described in **a** was also done for Nanog. Blue line represents background level. (**c**) Expression profiles of the genes that were differentially expressed in both *Pou5f1* and *Nanog* RNAi experiments. The upper panel shows the expression profile of the genes differentially expressed from the *Nanog* RNAi experiment, the lower panel shows genes that are differentially expressed from the *Pou5f1* RNAi experiment and the center panel shows the expression profiles of the 77 genes bound and differentially regulated by both factors. (**d**) Validation of change in expression of *Dkk1* after *Pou5f1* or *Nanog* RNAi. (**e**) Upregulation of Oct4-bound genes *Cdx2* and *Cldn4* after *Pou5f1* RNAi. (**f**) Changes in gene expression after Oct4 or Nanog depletion. The levels of the transcripts were normalized against values derived from control RNAi-transfected ES cells. *Trp53* and *Foxh1* were bound by Oct4 but not Nanog. (**g**) Depletion of Nanog induced *Nrp2* and *Klf6* expression. (**h**) Changes in gene expression after Nanog overexpression (OE). (**i**) Changes in gene expression after Nanog overexpression and Nanog overexpression with RA-induced differentiation. (**j**) Induction of *Dkk1* in treated cells from **i**. β-actin was used as an internal control for all real-time PCR measurements. (**k**) Model of how Oct4 and Nanog regulate genes involved in different pathways. Oct4 and Nanog occupy *Trp53bp1* and *Mycn*, but Nanog does not regulate their expression (link shown as blue arrow). Black arrows signify regulation by the transcription factors, as shown by RNAi depletion.

bound by Nanog (**Supplementary Table 5**). These genes thus represent direct targets regulated by the respective factors. As those genes whose expression was affected in the knockdown experiments did not preferentially contain binding sites within the 5′ proximal region (**Supplementary Figure 10** online), functional transcription factor binding seems not to be limited to the proximal promoter region. Our analysis also identified 77 genes that were bound and regulated by both Oct4 and Nanog (**Fig. 6c** and **Supplementary Table 6** online). *Rcor2*, *Esrrb* and *Phc1* are examples of transcriptional regulators positively regulated by both factors. The *Dkk1* gene, encoding for a Wnt antagonist, is negatively regulated by both Oct4 and Nanog (**Fig. 6d**). One interesting demonstration of Oct4-repressed genes is that of the trophectoderm marker genes *Cdx2* and *Cldn4*: both were markedly upregulated upon *Pou5f1* reduction (**Fig. 6e**).

Among the Nanog-bound genes, notable ones are *Pou5f1*, *Sox2*, *Rif1* and *REST*. Depletion of Nanog resulted in downregulation of their expression (**Fig. 6f**), indicating that Nanog activates transcription of these genes. Our previous work has shown that the Oct4/Sox2 binary complex has a role in regulating *Pou5f1*, *Sox2* and *Nanog*. The data presented here showed the reverse links from Nanog to *Pou5f1*, *Sox2* and *Nanog*. Notably, we found that Nanog can also have a repressive role in transcription. For example, neuropilin 2 (*Nrp2*) and core promoter element binding protein (*Klf6*, also known as Kruppel-like factor 6) were induced after Nanog depletion (**Fig. 6g**).

In addition to the knockdown experiment, we performed the reciprocal experiment, that of Nanog overexpression. This was to determine if the expression of any of the genes associated with Nanog binding sites was altered. Gene expression was compared between two ES cell lines, both stably transfected, one with a Nanog expression construct and the other with a parental vector control. Quantitative real-time PCR indicated that mRNA levels of *Pou5f1*, *Esrrb*, *Foxd3*, *Tcfcp2l1*, *Nr0b1* and *BMP4* were all increased to at least 150% of that of the control cells (**Fig. 6h**). The expression of other genes with associated Nanog binding sites remained unchanged (*Sox2*, *Rif1*, *Sall1*, *REST*, *Tcf3* and *Jarid2*). A third group of genes was downregulated upon Nanog overexpression (*Nrp2*, *Klf6* and *Dkk1*). These data suggest that a higher cellular concentration of Nanog within ES cells can modulate the transcription of a subset of target genes, though not all target genes.

It is known that ES cells overexpressing Nanog are resistant to differentiation induced by RA[9]. We asked whether Nanog can sustain the expression of several key genes identified in our study in the presence of RA. The cells were treated with 0.3 μM of RA for 2 d to induce differentiation. Control cells underwent rapid changes in

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

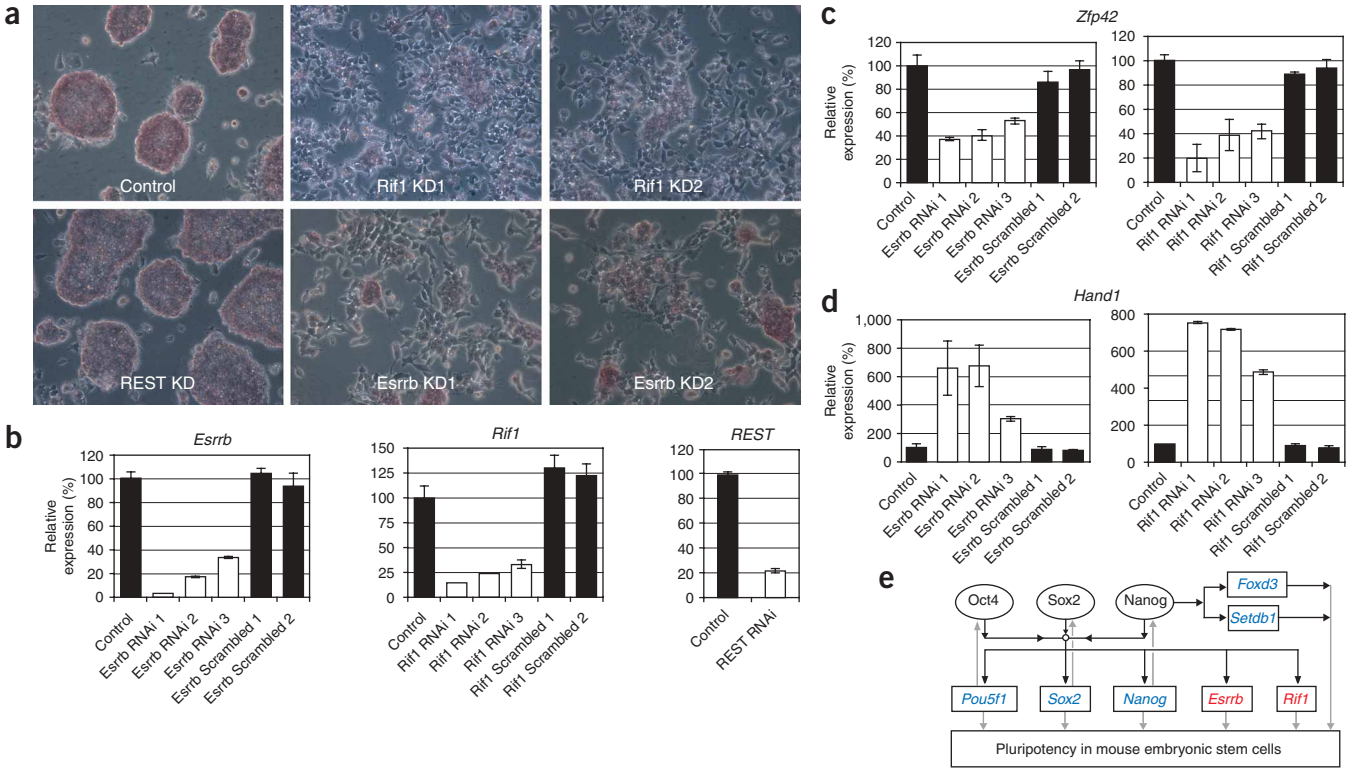

Figure 7 Regulation of pluripotency by downstream targets of Oct4 and Nanog. (**a**) Knockdown (KD) of *Esrrb* or *Rif1* led to differentiation of ES cells. Note the presence of flattened epithelial-like cells in the knockdown cells not seen in the vector control and *REST* knockdown ES cells. Cells were stained for alkaline phosphatase (pink), which is characteristic of nondifferentiated cells. (**b**) The levels of *Esrrb* or *Rif1* after knockdown using three constructs that target different regions of the respective genes were determined by real-time PCR quantification of reverse-transcribed RNAs. The third graph shows the level of *REST* after *REST* knockdown (**a**). (**c**) Reduction of ES cell marker *Rex1* after *Esrrb* or *Rif1* knockdown by RNAi. (**d**) Induction of trophectoderm marker *Hand1* after *Esrrb* or *Rif1* knockdown by RNAi. (**e**) Oct4 and Nanog regulatory network controlling pluripotency in ES cells. Transcription factors are represented by ovals, and the genes (printed in italics) are represented by rectangles. A black arrow indicates a transcription factor binding to a gene and positively regulating that gene. These links are largely based on evidence derived from ChIP and RNAi experiments. *Esrrb* and *Rif1* were also bound by Sox2 (data not shown). Gray arrows denote the synthesis of gene products from their respective genes. The genes printed in red (*Esrrb* and *Rif1*) are novel functional nodes in this network. All the factors shown in this model are required to maintain ES cell pluripotency. Foxd3 and ESET have been shown to be important in maintaining pluripotency of mouse ES cells[39,41].

morphology and became fibroblast-like in appearance. However, most of the Nanog-overexpressing cells retained ES cell morphology (**Supplementary Figure 11** online).

The expression of *Pou5f1*, *Sox2*, *Esrrb*, *Rif1 FoxD3*, *Tcfcp2l1*, *Sall1*, *REST*, *Jarid2*, *Tcf3* and *Nr0b1* was reduced by a smaller amount compared with the reduction of expression of these genes in the control cells treated with RA (**Fig. 6i**). This indicates that Nanog was able to sustain the expression of these genes. Consistent with the repression of *Dkk1* transcription by Nanog, the induction of *Dkk1* upon RA induction was lower for Nanog-overexpressing cells (**Fig. 6j**). However, genetic manipulations such as RNAi-mediated knockdown or overexpression may have had indirect effects.

In summary, we show that Oct4 and Nanog bind to and regulate diverse classes of genes. Of particular interest are genes encoding transcriptional regulators, growth factors, signaling molecules, DNA damage response sensors and suppressors of lineage-specific genes (**Fig. 6k**). It is noteworthy that there are genes such as *Trp53bp1* and *Mycn* that are bound by Nanog but are not regulated by it, as observed through genetic manipulation. Hence, independent validations such as these knockdown experiments are critical in distinguishing functional from nonfunctional circuitries.

Functional importance of downstream targets

Oct4 and Nanog are two important regulators in the maintenance of pluripotency in ES cells, targeting a core set of 345 genes (**Fig. 3b**). Among these genes, 30 of them encode known or putative DNA-binding regulators, including key genes *Pou5f1*, *Sox2* and *Nanog*. To determine if the regulatory network identified in our study has additional functional nodes, we asked if other common targets of Oct4 and Nanog are required to maintain mouse ES cells in a nondifferentiated state (**Fig. 7**). *Esrrb*, *Rif1* and *REST* are genes shown to be regulated by both Oct4 and Nanog (**Fig. 6f**). Notably, the *Esrrb* and *Rif1* knockdown cells became flattened and fibroblast-like, with a loss of alkaline phosphatase staining of nondifferentiated ES cells (**Fig. 7a,b**). *REST* knockdown changed neither the morphology of ES cells nor the level of alkaline phosphatase. The expression of the ESC-specific gene *Zfp42* was reduced in *Esrrb* and *Rif1* knockdown cells, whereas the trophectoderm marker *Hand1* was induced (**Fig. 7c,d**). The effect of RNAi was specific, as we observed the same phenotypic change with three siRNA targeting different regions of the *Esrrb* or *Rif1* genes. Scrambled siRNA sequences had no effect on the ES cells (**Fig. 7b–d**; **Supplementary Figure 12** online). In summary, we identified two new nodes in the Oct4 and Nanog circuitries that are important for maintaining the nondifferentiated state of mouse ES cells.

Oct4 and Nanog circuitries in mouse and human ES cells

Recently, the binding sites of OCT4 and NANOG at promoter regions in human ES cells have been reported[20]. Although the two studies used different approaches to identify binding sites, it is useful to compare the Oct4 and Nanog circuitries in mouse and human ES cells (**Fig. 8** and **Supplementary Table 7** online). First, we compared the bound genes identified in that study[20] with ours. Notably, we found that only 9.1% of Oct4-bound genes and 13% of Nanog-bound genes overlapped between the two studies (**Fig. 8a,c**). From our Oct4 ChIP-PET data set, we found 233 Oct4 sites in the 10-kb upstream regions of known genes (we termed these 'promoters'), and of these, only 33 of the corresponding human promoters were bound by OCT4 (**Fig. 8b**). Among the 434 Nanog sites within mouse promoters, NANOG bound to 92 of the corresponding human promoters (**Fig. 8d**). The limited overlap between the mouse and human datasets suggests that there

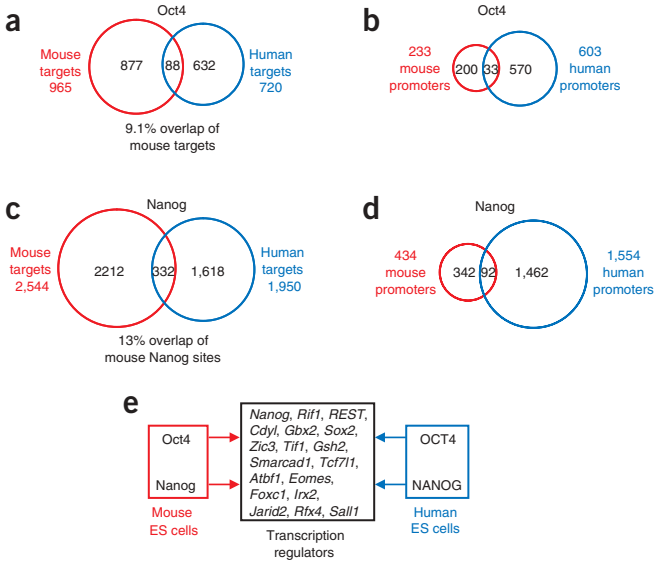

Figure 8 Conserved and diverged Oct4 and Nanog circuitries of mouse and human ES cells. (**a**) Venn diagram showing the overlap between Oct4 putative gene targets in mouse (red) and OCT4 putative gene targets in human ES cells (blue). (**b**) Venn diagram showing the overlap between Oct4 bound mouse promoters (red) and the promoters bound by OCT4 (blue) in human ES cells. Of the 1,083 Oct4 binding sites in mouse, 233 (22%) fall in the promoter region of known genes (defined as being less than 8 kb upstream and less than 2 kb downstream of transcription start site). Out of these, only 33 can be associated to a human promoter-bound region. (**c**) Venn diagram showing the overlap between Nanog putative gene targets in mouse (red) and NANOG putative gene targets in human ES cells (blue). (**d**) Venn diagram showing the overlap between Nanog-bound mouse promoters (red) and the promoters bound by NANOG (blue) in human ES cells. Of the 3,006 Nanog binding sites in mouse, 434 (14%) fall in the promoter region of known genes. Out of these, only 92 can be associated to a human promoter-bound region. (**e**) Common genes that encode for transcription regulators bound by Oct4 and Nanog in both mammalian ES cells.

may exist differences in the networks controlled by Oct4 and Nanog between species. For instance, here we have found Oct4 and Nanog binding to the proximal promoter of *Mycn* in mouse ES cells[21], but these interactions have not been detected in human ES cells[20].

Nevertheless, the human promoter datasets provides us with a unique opportunity to investigate the Oct4 and Nanog binding circuitries conserved in pluripotent cells from two mammalian species. There are 32 genes that were bound by Oct4 and Nanog in both mouse and human ES cells. Among this list, 18 of them encode for transcription regulators (**Fig. 8e**), including *Nanog*, *Sox2* and *Rif1*, further highlighting the importance of these genes in mammalian ES cells.

DISCUSSION

Unbiased mapping of binding sites in ES cells by ChIP-PET

An unbiased genome-wide location mapping approach is very powerful in elucidating the physiological targets of transcription regulators[22–27]. In the context of mammalian systems, this is particularly important because regulatory elements do not always fall within the 5′ proximal region of the first exon[22]. Our method is unique in that the technique allows for the detection of overlapping ChIP fragments that can then be used to precisely define the binding sites in living cells. Based on the empirically determined criteria of taking

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

only PET clusters with at least four overlaps of the PET fragments, we obtained about 1,000 and 3,000 high-confidence binding sites for Oct4 and Nanog, respectively.

We find that Sox2 sites are present to a great extent at Oct4-bound genomic loci. It has been shown that Sox2 and Oct4 occupy key regulatory regions of *Pou5f1*, *Sox2*, *Nanog*, *Fgf4*, *Fbxo15* and *Utf1* at adjacent *cis* elements[15,28–33]. The predominant motif uncovered by a *de novo* motif prediction algorithm is a sox-oct composite element present in approximately 70% of the Oct4 ChIP-PET clusters containing six or more PET overlaps (**Supplementary Note**). We also show empirical evidence for the *in silico* prediction that Oct4 and Sox2 occupy the same binding sites. Indeed, using Sox2 ChIP, we have detected Sox2 binding at the majority of Oct4-bound loci (J.-L.C. & H.-H.N., unpublished data). Sequential ChIP analysis for a number of genes further demonstrated that Oct4 and Sox2 are bound to the same target DNA molecules (**Supplementary Note**). The evidence we presented suggests that Oct4 and Sox2 work in tandem to regulate gene expression for a majority of their target genes.

Similarly, we have predicted *de novo* a Nanog motif from the Nanog ChIP-PET data. Nanog belongs to the Q50 homeoprotein family with the amino acid glutamine at position 50 of the homeodomain making direct contact with the nucleotides just 5′ of the ATTA sequence[34,35]. The ATTA tetramer has been reported to be the preferred sequence for Nanog[10]. Using a combination of mutagenesis and EMSA experiments, we determined that the CATT residues within the *Pou5f1* Nanog binding region are important for interaction between Nanog and DNA. Sequences containing a related CATT motif are also bound by Nanog *in vitro*.

Regulation of gene expression by Oct4 and Nanog

The global survey approach in this study demonstrated the targeting of two structurally unrelated transcription factors to genes on an extensive scale, indicating a high degree of cooperation between the two factors. Our data shows, for the first time, the different configurations of Oct4 and Nanog binding sites (**Fig. 3c**). This study represents a starting point of how to decipher the combinatorial binding site architectures of mammalian genes.

In order to understand transcription regulation by these factors, we must understand whether the bound genes are indeed regulated, as binding alone does not imply regulation. Using genome-wide microarray analysis, we find a notable association of Oct4 or Nanog binding sites with genes that are repressed and induced during differentiation. As an additional level of validation that the bound genes are *bona fide* targets, we examined the transcripts in cells with and without RNAi depletion of the respective factors. The data indicate that only a subset of the bound genes is regulated by Oct4 or Nanog. The nonresponsive genes could reflect nonfunctional sites or functional redundancy of transcription regulators.

Oct4 and Nanog circuitries in mouse and human ES cells

There are several plausible explanations for the limited conservation of Oct4 or Nanog-bound sites and genes between species. First, on the basis of transcriptome analyses that include microarrays, serial analysis of gene expression (SAGE) and massively parallel signature sequencing (MPSS), it is known that mouse and human ES cells show key differences[36,37]. Second, the disparity may arise from the scope of the transcription factor binding sites being mapped. A previous study[20] has surveyed 10-kb upstream regions of approximately 18,000 annotated genes, roughly 6% of the human genome. Previous work on mapping transcription factor binding sites using unbiased approaches shows that certain mammalian transcription factors can target sites outside proximal promoter elements[14,22,38]. Here we have performed unbiased surveys of transcription factor binding sites and find that Oct4 and Nanog binding sites are not restricted to upstream regions of genes.

Third, different technology platforms and reagents may contribute to the discrepancy. We chose a cutoff of four or more overlapping PET clusters to ensure >95% true positive binding sites. Clearly, there are true positives in PET clusters with three or fewer overlaps.

How do Oct4 and Nanog maintain pluripotency?

Both binding and genetic evidence presented in this study showed that Nanog regulates the expression of *Pou5f1* and *Sox2*. One likely mechanism for how Nanog sustains self-renewal and the undifferentiated state is through the modulation of Oct4 and Sox2 levels. These two transcription factors in turn control the downstream genes important for maintaining pluripotency or inhibiting differentiation (**Fig. 7e**). In addition, Nanog also controls important molecular effectors of ES cell fate, as exemplified by *Foxd3* and *Setdb1*. *Foxd3* encodes for a transcriptional repressor important for the maintenance of the inner cell mass or epiblast and the *in vitro* establishment of ES cell lines[39,40]. The *Setdb1* gene encodes for a histone H3 Lys9 methyltransferase that is required for survival of mouse ES cells[41]. Oct4 and Nanog both bind to *Mycn*, which has recently been reported to be among the key mediators in the self-renewal and proliferation of ES cells[21]. Further illustrating the central role of Oct4 and Nanog as key regulators, we have identified two downstream targets, *Esrrb* and *Rif1*, that are important for maintaining pluripotency of mouse ES cells. Essrb belongs to the superfamily of nuclear hormone receptors, and homozygous mutant embryos show abnormal trophoblast proliferation, precocious differentiation toward the giant cell lineage and reduction in primordial germ cells[42,43]. Rif1 is an ortholog of a yeast telomeric protein and is upregulated in mouse ES and germ cells[44]. In human cells, Rif1 associates with dysfunctional telomeres and has a role in DNA damage response[45,46]. Notably, Rif1 is also a target of OCT4 and NANOG in human ES cells, further implicating its functional importance in ES cell biology. The exact nature of how Esrrb and Rif1 regulate pluripotency of mouse ES cells remains to be studied. The location maps generated in this study should serve as useful guides in identifying additional components in the regulatory network important for self-renewal, pluripotency and differentiation of ES cells.

METHODS

Cell culture. E14 mouse ES cells, either cocultured with mouse primary embryonic fibroblast feeders or cultured under feeder-free conditions, were maintained in Dulbecco's modified Eagle medium (DMEM; GIBCO), supplemented with 15% heat-inactivated fetal bovine serum (FBS; GIBCO), 0.055 mM β-mercaptoethanol (GIBCO), 2 mM L-glutamine, 0.1 mM MEM nonessential amino acid, 5,000 units/ml penicillin/streptomycin and 1,000 units/ml of LIF (Chemicon). HEK293T cells were cultured in DMEM supplemented with 10% FBS and maintained at 37 °C with 5% CO_2. Detection of alkaline phosphatase, which is indicative of the nondifferentiated state of ES cells, was carried out using a commercial ES Cell Characterization Kit from Chemicon.

ChIP-PET analysis. Affinity-purified polyclonal Nanog antibody was purchased from Cosmo Bio and characterized as shown in the **Supplementary Note**. Antibodies against Oct4 and Sox2 have been characterized previously[15]. ChIP was performed as described previously[15]. The ChIP-PET analysis was performed as previously described[14]. The locations of the ChIP-enriched DNA present in the library were visualized using our in-house genome browser (T2G browser) which was implemented in the context of the University of California, Santa Cruz (UCSC) genome browser.

Electrophoretic mobility shift assays (EMSA). Full-length mouse Nanog cDNA and mutants were amplified with appropriate primers, and the resulting DNA fragments were cloned into the expression vector pET42b (Novagen). The recombinant Nanog proteins were expressed in BL21 after induction with 0.2 mM IPTG at 20 °C and purified with GST beads followed by Ni-NTA beads. The purified proteins were dialyzed against dialysis buffer (20 mM HEPES, pH 7.9, 20% glycerol, 100 mM KCl, 0.83 mM EDTA, 1.66 mM DTT, Protease Inhibitor Cocktail (Roche)) at 4 °C for 4 h. The concentrations of the proteins were measured with a Bradford assay kit (Bio-Rad). Oligonucleotides labeled with biotin at the 5′ termini of sense strands were annealed with reverse strands in annealing buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA) and purified with an agarose gel DNA extraction kit (Qiagen). EMSA was performed in 10-μl mixtures containing 10 mM HEPES, pH 7.5, 10 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM EDTA, 10% glycerol, 0.5 ng of biotin-labelled oligonucleotide, 100 ng recombinant proteins and 1 μg of poly(dI-dC). If indicated, antibodies or unlabeled competitor DNA were added after the initial incubation for additional 20 min. After incubation for 10 min at RT, the binding mixtures were subjected to electrophoresis on pre-run 5% native PAGE gels in 0.5× TBE buffer. The gels were transferred to Biodyne B nylon membranes (Pierce Biotechnologies) and the binding signal was detected with LightShift Chemiluminescent EMSA kit (Pierce Biotechnologies).

RNA extraction, reverse transcription and quantitative real-time PCR. Total RNA was extracted using TRIzol Reagent (Invitrogen) and purified with the RNAeasy Mini Kit (Qiagen). Reverse transcription was performed using SuperScript II Kit (Invitrogen). DNA contamination was removed by DNase (Ambion) treatment, and the RNA was further purified by an RNAeasy column (Qiagen). Quantitative PCR analyses were performed in real time using an ABI PRISM 7900 Sequence Detection System and SYBR Green Master Mix as described[15]. Two pairs of primers were used to quantify the amount of cDNA, and both primer pairs showed identical results. For all the primers used, each gave a single product of the right size. In all our controls lacking reverse transcriptase, no signal was detected (Threshold cycle (Ct) >35). Each RNAi experiment was repeated at least twice with different batches of ES cells. For ChIP experiments, relative occupancy values were calculated by determining the apparent IP efficiency (ratios of the amount of ChIP enriched DNA over that of the input sample) and normalized to the level observed at a control region, which was defined as 1.0. The error bars shown are 1 s.d. and were calculated from technical replicates based on triplicate real-time PCR measurements of DNA. The validation for ChIP-PET data was performed at least twice from independent ChIP. The sequences of the primers are available upon request.

Accession codes. GEO: GSE4189.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to the Biomedical Research Council (BMRC) and Agency for Science, Technology and Research (A*STAR) for funding. Y.-H.L is supported by the A*STAR graduate scholarship. J.-L.C is supported by the Singapore Millennium Foundation graduate scholarship. W.Z. and X.C. are supported by the National University of Singapore graduate scholarship. B.L. is partially supported by grants from the US National Institutes of Health (DK47636 and AI54973). We thank E. Cheung, T. Lufkin, N. Clarke, C.-A. Lim, P. Melamed and J. Buhlman for critical comments on the manuscript. We are grateful to E. Ng, A. Ang and Y.-C. Chong for assistance in annotating the binding sites.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Smith, A.G. Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* **17**, 435–462 (2001).
2. Pera, M.F., Reubinoff, B. & Trounson, A. Human embryonic stem cells. *J. Cell Sci.* **113**, 5–10 (2000).
3. Donovan, P.J. & Gearhart, J. The end of the beginning for pluripotent stem cells. *Nature* **414**, 92–97 (2001).
4. Loebel, D.A., Watson, C.M., De Young, R.A. & Tam, P.P. Lineage choice and differentiation in mouse embryos and embryonic stem cells. *Dev. Biol.* **264**, 1–14 (2003).
5. Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K. & Gruss, P. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**, 435–439 (1990).
6. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).
7. Niwa, H., Miyazaki, J. & Smith, A.G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).
8. Avilion, A.A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003).
9. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).
10. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
11. Pesce, M. & Scholer, H.R. *Oct-4*: gatekeeper in the beginnings of mammalian development. *Stem Cells* **19**, 271–278 (2001).
12. Chambers, I. & Smith, A. Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23**, 7150–7160 (2004).
13. Ng, P. *et al.* Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nat. Methods* **2**, 105–111 (2005).
14. Wei, C.L. *et al.* A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**, 207–219 (2006).
15. Chew, J.L. *et al.* Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell. Biol.* **25**, 6031–6046 (2005).
16. Mi, H. *et al.* The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res.* **33**, D284–D288 (2005).
17. Yeom, Y.I. *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894 (1996).
18. Pavesi, G., Mauri, G. & Pesole, G. An algorithm for finding signals of unknown length in unaligned DNA sequences. *Bioinformatics* **17** (Suppl.), 207–214 (2001).
19. Down, T.A. & Hubbard, T.J. NestedMICA: sensitive inference of over-represented motifs in nucleic acid sequence. *Nucleic Acids Res.* **33**, 1445–1453 (2005).
20. Boyer, L.A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
21. Cartwright, P. *et al.* LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**, 885–896 (2005).
22. Cawley, S. *et al.* Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499–509 (2004).
23. Kim, T.H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436**, 876–880 (2005).
24. Pollack, J.R. & Iyer, V.R. Characterizing the physical genome. *Nat. Genet.* **32**, 515–521 (2002).
25. Buck, M.J. & Lieb, J.D. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83**, 349–360 (2004).
26. Impey, S. *et al.* Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* **119**, 1041–1054 (2004).
27. Kim, J., Bhinge, A.A., Morgan, X.C. & Iyer, V.R. Mapping DNA-protein interactions in large genomes by sequence tag analysis of genomic enrichment. *Nat. Methods* **2**, 47–53 (2005).
28. Rodda, D.J. *et al.* Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.* **280**, 24731–24737 (2005).
29. Kuroda, T. *et al.* Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol. Cell. Biol.* **25**, 2475–2485 (2005).
30. Okumura-Nakanishi, S., Saito, M., Niwa, H. & Ishikawa, F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J. Biol. Chem.* **280**, 5307–5317 (2005).
31. Ambrosetti, D.C., Scholer, H.R., Dailey, L. & Basilico, C. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J. Biol. Chem.* **275**, 23387–23397 (2000).
32. Tokuzawa, Y. *et al.* Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. *Mol. Cell. Biol.* **23**, 2699–2708 (2003).
33. Nishimoto, M., Fukushima, A., Okuda, A. & Muramatsu, M. The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell. Biol.* **19**, 5453–5465 (1999).
34. Kornberg, T.B. Understanding the homeodomain. *J. Biol. Chem.* **268**, 26813–26816 (1993).
35. Affolter, M., Schier, A. & Gehring, W.J. Homeodomain proteins and the regulation of gene expression. *Curr. Opin. Cell Biol.* **2**, 485–495 (1990).
36. Brandenberger, R. *et al.* MPSS profiling of human embryonic stem cells. *BMC Dev. Biol.* **4**, 10 (2004).
37. Wei, C.L. *et al.* Transcriptome profiling of human and murine ESCs identifies divergent paths required to maintain the stem cell state. *Stem Cells* **23**, 166–185 (2005).
38. Martone, R. *et al.* Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc. Natl. Acad. Sci. USA* **100**, 12247–12252 (2003).

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

39. Hanna, L.A., Foreman, R.K., Tarasenko, I.A., Kessler, D.S. & Labosky, P.A. Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev.* **16**, 2650–2661 (2002).
40. Guo, Y. *et al.* The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endodermal-specific promoter expression. *Proc. Natl. Acad. Sci. USA* **99**, 3663–3667 (2002).
41. Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H. & Li, E. Histone H3–K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* **24**, 2478–2486 (2004).
42. Luo, J. *et al.* Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. *Nature* **388**, 778–782 (1997).
43. Mitsunaga, K. *et al.* Loss of PGC-specific expression of the orphan nuclear receptor ERR-beta results in reduction of germ cell number in mouse embryos. *Mech. Dev.* **121**, 237–246 (2004).
44. Adams, I.R. & McLaren, A. Identification and characterisation of mRif1: a mouse telomere-associated protein highly expressed in germ cells and embryo-derived pluripotent stem cells. *Dev. Dyn.* **229**, 733–744 (2004).
45. Xu, L. & Blackburn, E.H. Human Rif1 protein binds aberrant telomeres and aligns along anaphase midzone microtubules. *J. Cell Biol.* **167**, 819–830 (2004).
46. Silverman, J., Takai, H., Buonomo, S.B., Eisenhaber, F. & de Lange, T. Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. *Genes Dev.* **18**, 2108–2119 (2004).

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics