Cell

A DNA Repair Complex Functions as an Oct4/Sox2 Coactivator in Embryonic Stem Cells

Yick W. Fong,[1] Carla Inouye,[1] Teppei Yamaguchi,[1] Claudia Cattoglio,[1] Ivan Grubisic,[2] and Robert Tjian[1,3,*]
[1]Howard Hughes Medical Institute, Department of Molecular and Cell Biology
[2]UC Berkeley-UCSF Graduate Program in Bioengineering
[3]Li Ka Shing Center for Biomedical and Health Sciences
University of California, Berkeley, Berkeley, CA 94720, USA
*Correspondence: jmlim@berkeley.edu
DOI 10.1016/j.cell.2011.08.038

SUMMARY

The transcriptional activators Oct4, Sox2, and Nanog cooperate with a wide array of cofactors to orchestrate an embryonic stem (ES) cell-specific gene expression program that forms the molecular basis of pluripotency. Here, we report using an unbiased in vitro transcription-biochemical complementation assay to discover a multisubunit stem cell coactivator complex (SCC) that is selectively required for the synergistic activation of the *Nanog* gene by Oct4 and Sox2. Purification, identification, and reconstitution of SCC revealed this coactivator to be the trimeric XPC-nucleotide excision repair complex. SCC interacts directly with Oct4 and Sox2 and is recruited to the *Nanog* and *Oct4* promoters as well as a majority of genomic regions that are occupied by Oct4 and Sox2. Depletion of SCC/XPC compromised both pluripotency in ES cells and somatic cell reprogramming of fibroblasts to induced pluripotent stem (iPS) cells. This study identifies a transcriptional coactivator with diversified functions in maintaining ES cell pluripotency and safeguarding genome integrity.

INTRODUCTION

The molecular events leading to the maintenance of pluripotency in embryonic stem (ES) cells and reacquisition of a stem-like state in induced pluripotent stem (iPS) cells during somatic reprogramming represent mechanistically distinct processes that converge on a set of remarkably similar transcriptional events that underpin the pluripotent state. Both ES and iPS cells depend on fundamental transcription frameworks that are governed by a common set of "core" stem cell-specific transcription factors, namely Oct4, Sox2, and Nanog (Jaenisch and Young, 2008). These activators, in turn, collaborate with both ubiquitous and cell type-specific transcription factors to orchestrate complex gene expression programs that confer upon stem cells the unique ability to safeguard stemness while remaining poised to execute a broad range of developmental programs that drive lineage specification (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Marson et al., 2008).

Proper execution of these highly regulated processes by sequence-specific transcription factors often requires the coordinated recruitment of coactivator proteins to their cognate promoters. For example, transcriptional activators direct histone modifiers (e.g., CBP/p300) and chromatin remodelers (e.g., PBAF/BAF) to gene promoters to alter chromatin structure toward a state that is more permissive to transcriptional activation (Näär et al., 2001). Independent of chromatin, a variety of activators recruit other classes of coactivators, such as the multisubunit Mediator, various TBP/TAF complexes, SRC, etc., via direct protein-protein interactions to execute specific transcriptional programs. This class of coactivators often serves as molecular "adaptors" by bridging activators to the general transcription machinery, thereby mediating the synergistic response by these activators (Näär et al., 1999). Interestingly, subunits of Mediator have also been shown to interact with cohesin possibly to promote DNA looping and thereby facilitate long-distance interactions between enhancers and core promoters in vivo (Kagey et al., 2010). Indeed, such coactivators are often multifunctional and can activate transcription through chromatin-dependent as well as independent mechanisms. Further expanding the transcriptional repertoire of coactivator complexes, their protein levels and subunit compositions are frequently modulated in a developmental stage and cell type-specific manner (Roeder, 2005; Taatjes et al., 2004). Additionally, these protein-protein-driven coactivator-activator transactions are often critical nodes in various signal transduction pathways and can serve as molecular "sensors" by integrating cell-intrinsic and -extrinsic cues, thereby coupling gene networks with specific cellular responses to produce complex biological programs of gene expression (Rosenfeld et al., 2006).

Totipotent ES cells employ these same sets of coactivators in conjunction with special activators such as Oct4 and Sox2 to regulate transcription of a large number of genes, including *Nanog*, that form the molecular basis of pluripotency (Gao et al., 2008; Kagey et al., 2010; Kidder et al., 2009; Tutter et al., 2009). The transcription of *Nanog* is exquisitely dependent on

Oct4 and Sox2 (Kuroda et al., 2005; Rodda et al., 2005). However, coexpression of Oct4 and Sox2 failed to robustly activate a *Nanog* promoter reporter construct in differentiated cells like 293 or NIH 3T3 cells, even though Mediator, p300/CBP, and PBAF/BAF complexes remain abundantly expressed and active (Rodda et al., 2005). This led us to speculate that one or more as yet unidentified stem cell-specific cofactors may be required to activate the transcription of *Nanog* and other Oct4/Sox2 target genes in ES cells. Indeed, recent studies of germ cells and differentiated somatic cells revealed that even parts of the general transcriptional machinery may be radically altered in a tissue- or cell-specific context (Goodrich and Tjian, 2010; Müller et al., 2010). Diversification of the transcriptional apparatus may therefore represent a fundamental strategy, particularly in ES cells, to cope with the multidimensional nature of transcription programs that must be precisely tuned to both maintain pluripotency and, at the same time, allow for lineage-specific programs of differentiation (Liu et al., 2011).

The human *Nanog* promoter contains a prototypic composite oct-sox *cis*-acting regulatory element located immediately upstream of the transcription start site that is conserved across several mammalian species (Kuroda et al., 2005; Rodda et al., 2005). A *Nanog* promoter-GFP reporter construct containing a DNA fragment encompassing this promoter-proximal oct-sox element is sufficient to recapitulate the robust expression pattern of endogenous Nanog in ES cells in an Oct4-, Sox2-dependent manner (Kuroda et al., 2005; Rodda et al., 2005). Unbiased genome-wide motif searching analyses of Oct4 in both mouse and human ES cells identified an oct-sox composite consensus sequence element, confirming that Oct4 likely orchestrates an ES-specific gene expression program primarily through cooperation with Sox2 (Chen et al., 2008; Loh et al., 2006). Because the oct-sox *cis*-control element in the *Nanog* promoter represents a common configuration that is present in the promoters of many other Oct4- and Sox2-activated genes in ES cells, the well-characterized *Nanog* proximal promoter provided us with a useful model template for identifying uncharacterized transcriptional cofactors required for Oct4- and Sox2-directed activation. Therefore, we took advantage of a fully reconstituted in vitro transcription system in which one can unambiguously and systematically test and identify transcriptional cofactors that may be directly required to potentiate Oct4- and Sox2-dependent gene activation of *Nanog*. Here, we report the biochemical purification and identification of a multisubunit stem cell coactivator (SCC) that is required for the synergistic activation of *Nanog* by Oct4 and Sox2 in vitro. After extensive biochemical characterization, we surprisingly found that SCC is none other than the XPC-RAD23B-CETN2 (XPC) nucleotide excision repair (NER) complex. SCC/XPC interacts directly with Oct4 and Sox2 and co-occupies a majority of Oct4 and Sox2 targets genome-wide in mouse ES cells. Importantly, SCC/XPC is required for stem cell self-renewal and efficient somatic cell reprogramming. Thus, our findings unmask an unanticipated selective coactivator role of an NER complex in transcription in the context of ES cells and may provide a previously unknown molecular link that couples stem cell-specific transcription to DNA damage response with potential implications for enhanced ES cell genome stability.

RESULTS

Detection of an Oct4- and Sox2-Dependent Coactivator Activity in EC and ES Cells

Having chosen the *Nanog* promoter as our model template, we next set out to develop an in vitro reconstituted transcription assay that could recapitulate the Oct4- and Sox2-dependent transactivation at the *Nanog* promoter observed in vivo. To enhance the sensitivity of the assay, we inserted four copies of the *Nanog* oct-sox-binding sites immediately upstream of the native oct-sox element found in the human *Nanog* promoter. Our basal in vitro transcription assay consisted of purified recombinant TFIIA, -B, -E and -F together with immunoaffinity-purified native RNA polymerase II, TFIID, and TFIIH (Figure S1A available online). When purified Oct4 and Sox2 were added to this reconstituted transcription system, only a very weak activation of the *Nanog* promoter was detected (Figure 1A, lanes 1 and 2). As a control, we could show that the same complement of general transcription factors (GTFs) was able to support strong Sp1-dependent activation from a GC box-containing "generic" transcription template (G3BCAT) (Figure 1A, lanes 5 and 6). This initial result suggested that efficient activation of *Nanog* by Oct4 and Sox2 may require additional cofactors to potentiate a full activator-dependent response.

We reasoned that such a putative coactivator ought to be selectively active in pluripotent cell types that express Nanog under the control of Oct4 and Sox2. For example, NTERA-2 (NT2) is a pluripotent human embryonal carcinoma (EC) cell line that expresses Oct4, Sox2, and Nanog and shares with ES cells core molecular mechanisms that govern self-renewal (Pal and Ravindran, 2006). Detailed expression profiling of NT2 and bona fide human ES cell lines revealed many similarities, including robust expression of Nanog (Schwartz et al., 2005; Sperger et al., 2003). However, unlike human ES cells, NT2 cell culture can be more readily scaled up, a prerequisite to generating sufficient quantities of starting materials for the biochemical purification of putative Oct4/Sox2 coactivators. We therefore chose extracts derived from NT2 cells as our starting material in our efforts to develop a "biochemical complementation" assay to hunt for pluripotent stem cell-selective cofactors.

We first fractionated NT2 nuclear extracts by conventional phosphocellulose ion exchange chromatography. Next, we supplemented our "basal" reconstituted transcription reactions with various salt-eluted fractions from the phosphocellulose column to see whether there was any activity that could restore Oct4/Sox2-dependent activation of our *Nanog* promoter. This strategy allowed us to unmask an activity in the high salt phosphocellulose fraction (P1M) prepared from NT2 nuclear extracts (but not HeLa extracts) (Figure S1B) that strongly potentiated transcription of the *Nanog* promoter in an Oct4- and Sox2-dependent manner using either a naked (Figure 1A, lanes 3 and 4) or a *Nanog* chromatin template assembled with a crude *Drosophila* cytosolic extract (data not shown). This new cofactor activity is selectively required for transcription of *Nanog*, as it had no effect on either basal- or Sp1-activated transcription from a control G3BCAT template (Figure 1A, lanes 5–8). Importantly, this P1M fraction also stimulated the Oct4/Sox2-dependent transcription from a native *Nanog* promoter template (Figure 1B),

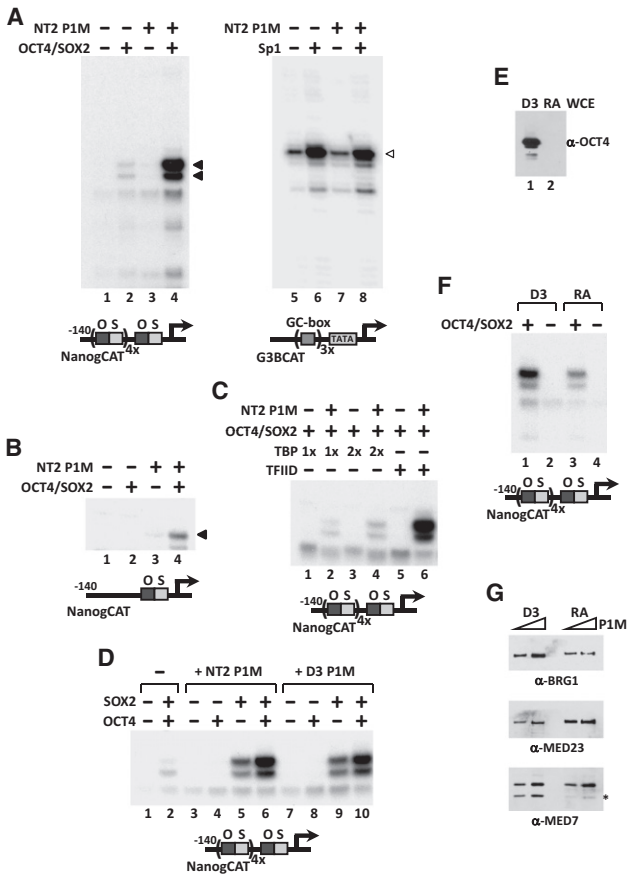


Figure 1. Transcriptional Activation of *Nanog* by Oct4 and Sox2 Requires a Stem Cell-Specific Cofactor

(A) Reconstituted in vitro transcription reactions supplemented with Oct4 and Sox2 (lanes 2 and 4) or Sp1 (lanes 6 and 8) plus a phosphocellulose 1 M KCl fraction derived from NT2 nuclear extracts (NT2 P1M, lanes 3, 4, 7, and 8) and programmed with either a *Nanog* template engineered with four extra copies of the oct-sox composite element (NanogCAT, lanes 1–4), or a GC box-containing template (G3BCAT, lanes 5–8). Oct4/Sox2, NT2 P1M-dependent transcripts are indicated by filled arrowheads and Sp1-dependent transcriptions by open arrowheads.

(B) Transcription of the native *Nanog* promoter requires Oct4, Sox2, and NT2 P1M fraction (lane 4).

(C) TFIID and NT2 P1M fraction are needed to potentiate Oct4/Sox2-dependent activation. Transcription reactions contain Oct4 and Sox2 (lanes 1–6), NT2 P1M fraction (lanes 2, 4, and 6) with increasing amounts of recombinant TBP (1× or 2×, lanes 1–4), or TFIID (lanes 5 and 6).

(D) Synergistic activation of *Nanog* by Oct4 and Sox2 requires P1M fractions prepared from NT2 or mouse ES cell line D3 nuclear extracts. In vitro transcription reactions contain equal amounts (~0.7 μg) of NT2 (lanes 3–6) or D3 P1M fractions (lanes 7–10), with Oct4 alone (lanes 4 and 8), Sox2 alone (lanes 5 and 9), or both activators (lanes 2, 6, and 10).

(E) Immunoblotting analysis of Oct4 levels in whole-cell extracts (WCE) prepared from pluripotent D3 cells (D3, lane 1) and cells treated with retinoic acid for 6 days (RA, lane 2).

(F) P1M fractions prepared from pluripotent (D3, lanes 1 and 2) and differentiated (RA, lanes 3 and 4) D3 nuclear extracts were added to transcription reactions with or without Oct4 and Sox2.

(G) Western blots (2-fold titration) of P1M fractions prepared from pluripotent (D3) and differentiated (RA) D3 nuclear extracts using anti-BRG-1, anti-MED23, and anti-MED7 antibodies. Asterisk indicates a nonspecific band or a breakdown product recognized by anti-MED7 antibody.

See also Figure S1.

as well as two other Oct4/Sox2-dependent templates derived from the mouse *Fbxo15* promoter (Tokuzawa et al., 2003) (mFbxo15CAT) (Figure S1C, lanes 1–4) and the human *HESX1* promoter (Chakravarthy et al., 2008) (HESX1CAT) (Figure S1C, lanes 5–8). Thus, our in vitro complementation assay programmed with naked DNA templates revealed at least one potential coactivator activity that directs Oct4/Sox2-dependent activation of *Nanog*. We decided to pursue characterization of this cofactor that does not appear to require chromatin-based functions. To the best of our knowledge, this finding also demonstrates for the first time a fully reconstituted, in vitro transcription system that can faithfully recapitulate stem cell-specific gene activation.

We next investigated the relative requirements for other cofactors in our assay system. Consistent with previous studies demonstrating that TAFs in the TFIID complex are often required for transcriptional activation by a variety of activators, including nuclear receptors (Lemon et al., 2001), Sp1 (Ryu et al., 1999), and SREBP-1 (Näär et al., 1998), substituting holo-TFIID with recombinant human TBP resulted in a near complete loss of activation by Oct4 and Sox2 (Figure 1C). The very weak residual activation that we see using TBP (Figure 1C, lanes 2 and 4) is most likely due to trace amounts of TFIID present in the NT2 P1M fraction (data not shown). These findings suggest that TAFs/holo-TFIID and the putative cofactor detected in the NT2 P1M fraction are both required for optimal transcription of *Nanog* elicited by Oct4 and Sox2. Interestingly, in this reconstituted system, the addition of CRSP/Mediator complex was not required to obtain robust Oct4/Sox2 activation at the *Nanog* promoter. However, it is likely that some CRSP/Mediator is present in the P1M fraction, and it remains possible that some other component of the reconstituted system (i.e., Pol II) may have some residual amount of CRSP/Mediator contamination (Näär et al., 2002). We found, however, that adding purified CRSP/Mediator instead of the NT2 P1M factor to these reactions completely failed to enhance Oct4/Sox2-dependent activation of *Nanog* transcription (Figure S1D). This finding indicates that the NT2 cofactor must be distinct from Mediator. Furthermore, addition of other transcriptional activators implicated in *Nanog* expression (i.e., Nanog, Sall4 [Zhang et al., 2006], Klf4 [Jiang et al., 2008] and Esrrb [van den Berg et al., 2008; Zhang et al., 2008]) also did not replace or enhance Oct4/Sox2-dependent transcription of *Nanog* in vitro (Figure S1E).

To confirm that this newly detected cofactor activity in NT2 cells is also present in bona fide ES cells, P1M fractions were prepared from the pluripotent D3 mouse ES cell line and assayed for transcription. We found that the D3 P1M fraction was as active as the NT2 P1M fraction in potentiating Oct4/Sox2-activated transcription of *Nanog* (Figure 1D, compare lane 2 to 6 and 10). Interestingly, the highest levels of transactivation by the NT2 or D3 P1M fractions were observed only when both activators were added to the transcription reaction, whereas no activation was detected with Oct4 alone and a moderate level of activation was seen with Sox2 alone (Figure 1D, lanes 3–10). Apparently, this cofactor mediates the synergistic activation of *Nanog* by Oct4 and Sox2. If, as we postulated, this new coactivator functions selectively in pluripotent cells, one might expect that its presence or activity would need to be downregulated

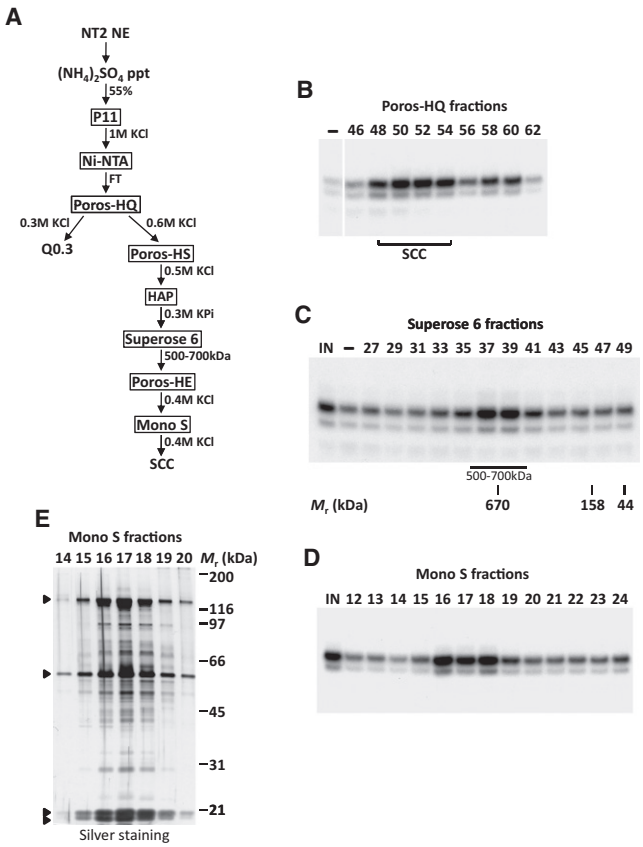


Figure 2. Purification of Stem Cell Coactivator

(A) Chromatography scheme for partial purification of Q0.3 and purification of SCC from NT2 nuclear extracts (NT2 NE). NT2 NE is first subjected to ammonium sulfate precipitation (55% saturation) followed by a series of chromatographic columns as indicated.

(B) Buffer (–) and fractions containing SCC eluted from a Poros-HQ anion exchanger (top) assayed in the presence of Oct4 and Sox2 in in vitro transcription assays.

(C) Coactivator SCC migrates as a large complex. Input (IN), buffer (–), and Superose 6 fractions (top) assayed as in (B) except that all reactions are supplemented with Q0.3 (A). Mobilities of peak activity (500–700 kDa) and gel filtration protein standards are shown (bottom).

(D) Transcription profile of stem cell coactivator (SCC) activity after the final Mono S chromatography step. Reactions contain input (IN) and Mono S fractions (top) and are assayed as in (C).

(E) Silver-stained SDS-PAGE gel of the active Mono S fractions. Filled arrowheads indicate polypeptides that comigrate with SCC activity.

See also Figure S2.

upon differentiation, as is the case for Oct4. To investigate whether the cofactor activity is restricted to the pluripotent state of ES cells, D3 cells were induced to differentiate by removal of LIF and treatment with retinoic acid (RA). The extent of differentiation was monitored by the loss of Oct4 expression that was complete after 6 days (Figure 1E). Nuclear extracts and P1M fractions were then prepared from D3 cells before and after differentiation. When compared to pluripotent D3 P1M fractions, an equivalent amount of P1M fraction prepared from differentiated D3 nuclear extracts showed significantly decreased cofactor activity in our in vitro transcription assay (Figure 1F, compare lanes 1 and 3). This decrease is not due to a wholesale loss of transcription factors and other cofactors during stem cell differentiation because the levels of PBAF/BAF (BRG-1) and the Mediator complex (MED23 and MED7) were largely unchanged in the two extracts (Figure 1G).

Purification and Identification of a Stem Cell Coactivator

Starting with 200–400 L of NT2 cells, we were able to separate the cofactor activity into two distinct chromatographic fractions. One cofactor activity eluted from an anion exchanger (Poros-HQ) at ~0.3 M KCl (Q0.3; data not shown), whereas a second distinct activity eluted at ~0.6 M KCl (stem cell coactivator [SCC] (Figures 2A and 2B). Full synergistic Oct4/Sox2-dependant activation of *Nanog* required both fractions in our in vitro reconstituted transcription reactions (Figure S2). Using this biochemical complementation system, we sequentially purified the more robust activity, SCC, over eight chromatographic columns, resulting in > 50,000-fold increase in specific activity (Figure 2A). Because SCC activity migrated with an apparent native molecular mass (M_r) of ~600 kDa during size-exclusion chromatography (Figure 2C), it seemed likely that this coactivator was a multiprotein complex. Accordingly, SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the most purified Mono S fractions revealed a distinct pattern of four major polypeptides (along with multiple breakdown products) that consistently copurified with the SCC activity (Figures 2D and 2E). For the remainder of this report, we focus on the identification and functional characterization of SCC in vitro and in vivo.

To identify polypeptides comprising the SCC complex, peak Mono S-purified fractions were pooled and separated by SDS-PAGE. Surprisingly, tryptic digests of excised gel bands followed by high-sensitivity mass spectrometry revealed all detectable constituents of SCC to be none other than the Xeroderma pigmentosum group C (XPC)-RAD23B-Centrin 2 (CETN2) nucleotide excision repair (NER) complex (Araki et al., 2001) (Figure 3A). We next carried out western blot analysis with antibodies specific to XPC, RAD23B, and CETN2 to confirm the identities of the purified SCC subunits (Figure 3B). As expected, these three polypeptides were highly enriched in the purified SCC Mono S peak fractions when compared to the crude NT2 P1M fraction (Figure 3B). Because identification of SCC as being identical to the XPC-NER complex was so unexpected, particularly as this repair complex has not been associated with any cell type-specific function nor linked to stem cell transcription, we next wanted to compare the relative amounts of this factor in different cell types. Consistent with the notion that SCC may be functioning in an unusual way in pluripotent stem cells, we found that these three proteins are highly enriched in ES and EC cells. For example, the levels of XPC, RAD23B, and CETN2 in the NT2 P1M fraction are much higher than in an equivalent amount of P1M fraction prepared from HeLa nuclear extracts (Figure 3B). Accordingly, in in vitro transcription reactions, Oct4/Sox2-dependent activation of *Nanog* by HeLa P1M fraction is much lower than that of NT2 P1M fraction (Figure S1B). XPC and RAD23B were rapidly downregulated upon RA-induced differentiation of mouse D3 ES cells, whereas CETN2, components of the basal transcription machinery (TBP and TFIIE), and other NER factors (XPA and XPB) decreased only slightly while

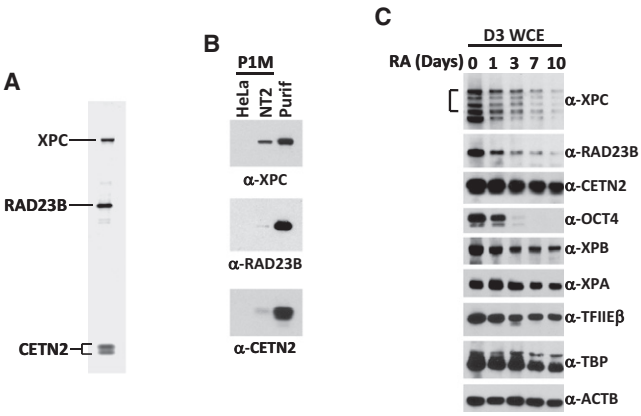


Figure 3. SCC Is the XPC-RAD23B-CETN2 Nucleotide Excision Repair Complex

(A) Mass spectrometry analysis of Mono S peak activity fractions (16–18) in Figure 2E with protein identities indicated.

(B) SCC is highly enriched in NT2 P1M fraction. Comparative western blot analysis of HeLa and NT2 P1M fractions (1.5 μg each) and purified Mono S SCC fraction (Purif, ~30 ng) using anti-XPC, anti-RAD23B, and anti-CETN2 antibodies.

(C) Downregulation of XPC and RAD23B upon RA-induced differentiation of mouse D3 ES cells. Western blot analysis of whole-cell extracts prepared from D3 cells (D3 WCE) collected at indicated days post-RA treatment using antibodies against XPC, RAD23B, CETN2, OCT4, XPB, XPA, TFIIEβ, TBP, and loading control β-actin (ACTB).

the loading control β-actin remained unchanged (Figure 3C). This finding is consistent with our previous observation that the D3 P1M fraction from differentiated cells is significantly less active than the pluripotent D3 P1M fraction in potentiating *Nanog* transcription (Figure 1F).

Reconstitution and Mechanism of Coactivation by SCC

While we were in the process of further characterizing the role of the XPC-RAD23B-CETN2 complex in transcription, Le May et al. reported that XPC and other components of the NER apparatus can be recruited to a gene promoter (e.g., *RARβ2*) upon nuclear hormone induction (Le May et al., 2010). Although the mechanism by which XPC and other NER factors mediate gene activation remains unclear, these recent studies and our new findings have unmasked a hitherto unknown and potentially important role for XPC that is directly linked to transcription. In our case, the most striking finding was the direct requirement for the SCC/XPC complex in selectively potentiating the transcriptional activation of *Nanog* by Oct4 and Sox2 in ES cell extracts. However, to more firmly establish this exciting new connection, we first needed to eliminate the possibility that trace amounts of contaminants present in our purified SCC fraction were responsible for the coactivator activity detected in our in vitro transcription assays. Therefore, we set about to reconstitute the heterotrimeric XPC-RAD23B-CETN2 complex from recombinant gene products expressed in insect (Sf9) cells following co-infection with baculoviruses expressing His-tagged XPC, FLAG-tagged RAD23B, and untagged CETN2. Using an efficient two-step affinity purification procedure, we were able to purify the recombinant heterotrimeric complex to near homogeneity (Figure 4A). Our ability to generate pure polypeptide subunits, as well as various combinations of dimeric and trimeric complexes, allowed us to address a number of important questions, such as whether known functional domains of XPC required for NER are also necessary for the cofactor activity. It is well established that XPC's ability to interact nonspecifically with DNA is essential for its NER function. Indeed, a single point mutation in the DNA-binding domain (W690S) of XPC, identified in an XP patient (XP13PV), abolishes binding to damaged (and undamaged) DNA and is defective in repair in vivo and in vitro (Maillard et al., 2007; Yasuda et al., 2007). To address whether XPC's nonspecific DNA-binding activity is also important for its coactivator function, a mutant DNA-binding-defective XPC (W690S) complex (that had been independently confirmed to be compromised for DNA binding in vitro; Figures S3A and S3B) was reconstituted in Sf9 cells and tested along with the wild-type complex for their ability to support Oct4/Sox2-dependent transcriptional activation of *Nanog* in vitro. Surprisingly, both the recombinant wild-type and mutant complexes exhibited specific activities for coactivation comparable to that observed for purified native endogenous SCC from NT2 cells (Figure 4B). Taken together, these results confirm that the XPC-RAD23B-CETN2 complex is indeed SCC and suggest that its DNA binding (and repair) activity is dispensable and functionally separable from its transcriptional cofactor activity at least in vitro. It has also been reported that XPC can interact directly with TFIIH (Uchida et al., 2002) and thus might provide a DNA-independent mechanism by which SCC can be recruited to gene promoters. To test this possibility, a C-terminal truncation of XPC that abolishes TFIIH (and CETN2) but retains RAD23B binding (amino acids 1–813, C814St) (Bernardes de Jesus et al., 2008) was used in our in vitro assay and was found to have no adverse affect on the ability of a XPC (C814St)-RAD23B heterodimer to mediate Oct4/Sox2-activated transcription of *Nanog* (Figures S3C and S3D). We therefore speculate that SCC/XPC is most likely targeted to its cognate promoters via potential interactions with specific activators such as Oct4 and Sox2.

To probe for a potential direct interaction between SCC and Oct4 and/or Sox2, mouse SCC subunits were overexpressed with Oct4, Sox2, Klf4, and c-Myc (STEMCCA) (Sommer et al., 2009) in 293T cells. SCC coimmunoprecipitated with Oct4, but not with control IgG (Figure 4C). To examine whether the DNA-binding property of SCC is required for its interaction with Oct4 and other activators, both the wild-type (WT) and DNA-binding-defective (W683S in mouse) XPC/SCC complexes were coexpressed with STEMCCA. Immunoprecipitation of WT and mutant SCC complexes using an anti-RAD23B antibody pulled down both Oct4 and Sox2, but not Klf4 or XPA (Figure 4D). These data indicate a direct and specific protein-protein binding between SCC and select activators, thus providing a mechanism by which SCC may serve as a transcriptional coactivator for Oct4 and Sox2 (but not Klf4; see Figure S1E) in potentiating *Nanog* transcription. These findings may also explain why the DNA-binding activity of the XPC subunit of SCC is dispensable for transcription in vitro. However, we were unable to reproducibly detect a stable interaction between SCC and Oct4/Sox2 in D3 ES cell extracts. It is worth noting, though, that other coactivators implicated in Oct4/Sox2-directed transcriptional activation

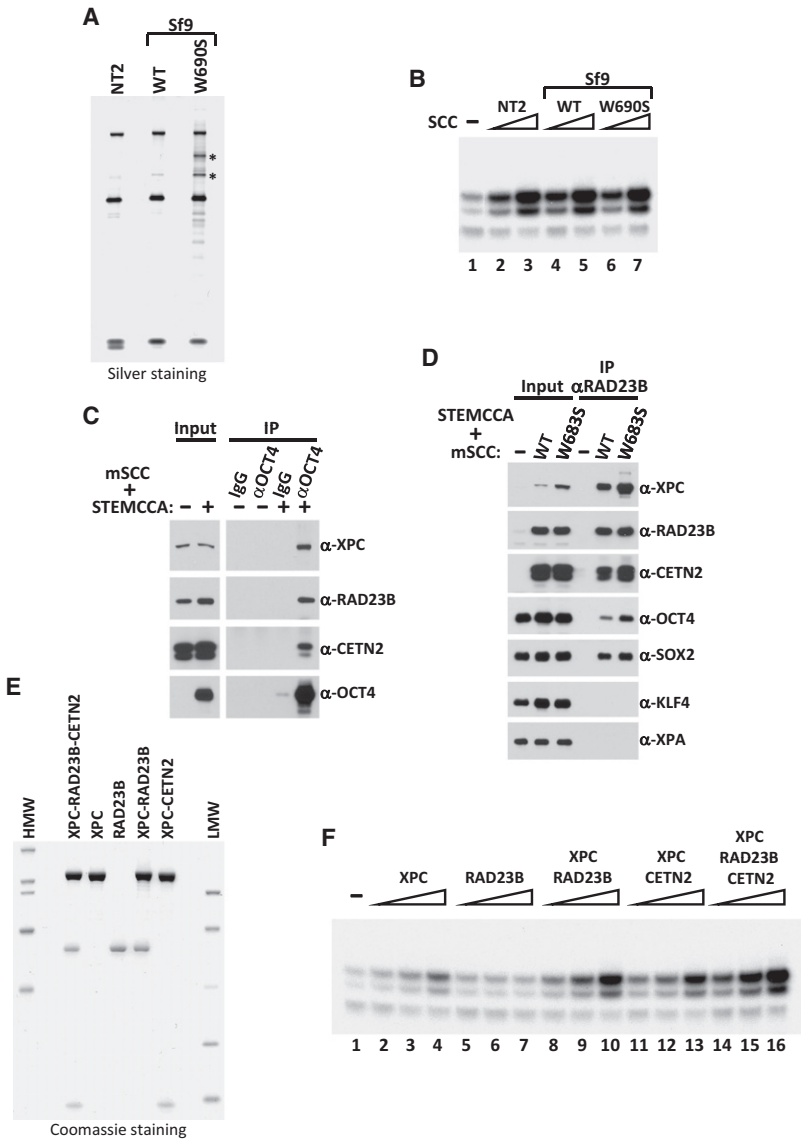


Figure 4. Reconstitution of Recombinant SCC Complexes

(A) Silver-stained SDS-PAGE gel of purified NT2 SCC (NT2), recombinant wild-type (WT), and DNA-binding-defective mutant (W690S) XPC-containing SCC complexes reconstituted in insect Sf9 cells by coinfection with baculoviruses expressing His-tagged XPC, FLAG-tagged RAD23B, and untagged-CETN2. Major proteolytic fragments of mutant XPC are indicated by asterisks.

(B) Recombinant SCC complex enhances Oct4/Sox2-activated transcription of *Nanog* independent of DNA binding. Buffer (–), NT2 (Mono S peak activity fractions; lanes 2 and 3), recombinant WT (lanes 4 and 5), and W690S mutant (lanes 6 and 7) SCC complexes are assayed (over a 3-fold concentration range). All transcription reactions contain Oct4, Sox2, and Q0.3 (lanes 1–7).

(C) Oct4 interacts with SCC. Western blot analysis of input lysates (2%) and coimmunoprecipitated proteins from extracts of 293T cells transfected with a polycistronic expression plasmid encoding all three subunits of mouse SCC (mSCC) with or without a polycistronic plasmid expressing mouse Oct4, Sox2, Klf4, and c-Myc (STEMCCA) using normal IgG or anti-Oct4 antibody. See also Figure S3.

(D) SCC-B interacts directly with Oct4 and Sox2 independent of DNA binding. Control vector (–), plasmids expressing wild-type (WT), or mutant (W683S) XPC-containing mSCC complexes were cotransfected with STEMCCA into 293T cells and immunoprecipitated with anti-RAD23B antibody. Input lysates (2%) and RAD23B-bound proteins were detected by immunoblotting.

(E) Coomassie-stained SDS-PAGE gel of purified recombinant XPC, RAD23B, dimeric (XPC-RAD23B and XPC-CETN2), and holo-SCC (XPC-RAD23B-CETN2) complexes.

(F) Titrations (over a 4-fold concentration range) of XPC (lanes 2–4), RAD23B (lanes 5–7), XPC-RAD23B (lanes 8–10), XPC-CETN2 (lanes 11–13), and XPC-RAD23B-CETN2 (lanes 14–16) in in vitro transcription reactions supplemented with Q0.3 (lanes 1–16) and assayed as in (B). See also Figure S3.

(e.g., Mediator and p300/CBP) have not been identified in recent "interactome" studies on Oct4-, Sox2-, or Nanog-associating factors (Engelen et al., 2011; van den Berg et al., 2010; Wang et al., 2006), supporting the idea that functional coactivator-activator interactions can often be weak and transient.

The ability to reconstitute active SCC from purified recombinant subunits also provided us with a unique opportunity to examine the contribution of individual subunits, as well as different dimeric combinations, in supporting Oct4/Sox2 transcriptional activation. Purified individual subunits (XPC or RAD23B), partial dimeric complexes (XPC-RAD23B or XPC-CETN2), and holo-SCC complexes (Figure 4E) were assayed over a 4-fold dose-response range in our fully reconstituted in vitro transcription reactions containing Oct4, Sox2, and a partially purified Q0.3 fraction (Figure 4F). The large XPC subunit alone only slightly activated transcription above background at the highest concentrations tested (Figure 4F, compare lanes 1 and 4), whereas RAD23B alone was essentially inactive. The XPC-CETN2 dimer was slightly more active than XPC alone. By contrast, a marked gain in specific activity was observed with the XPC-RAD23B dimeric complex that was nearly as active as the holo-complex (Figure 4F). These results suggest that the minimal active complex likely consists of XPC and RAD23B, whereas CETN2 may enhance the activity of the complex by providing structural support or stability.

SCC Coactivator Function in ES Cell Self-Renewal and Somatic Cell Reprogramming

We next set out to determine the role of the SCC/XPC complex on gene expression and *Nanog* transcription by loss-of-function studies in ES cells. Lentiviruses containing two independent short hairpin RNAs (shRNAs) specifically targeting XPC, RAD23B, and CETN2 were used to infect mouse D3 ES cells to selectively deplete SCC (Figures 5A, S4A, and S4B). Knockdown of SCC subunits resulted in pronounced cellular morphological abnormalities and decreased alkaline phosphatase (AP) activity

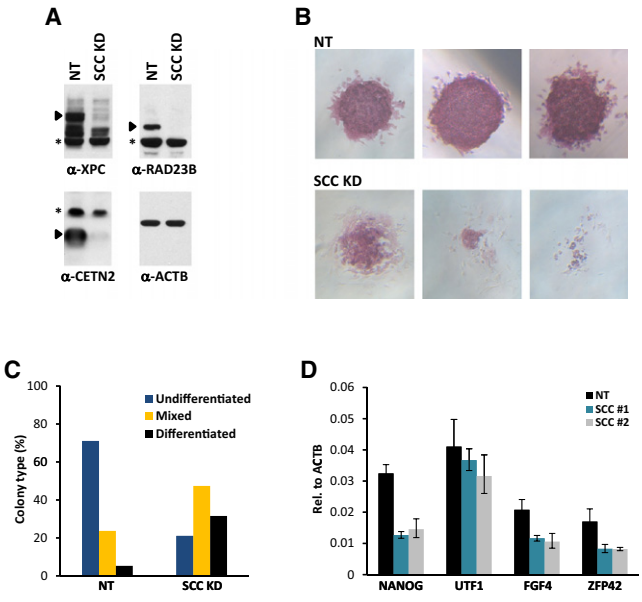


Figure 5. SCC Is Required for ES Cell Maintenance

(A) Efficiency of shRNA-mediated depletion of SCC in mouse ES cell line D3. Whole-cell extracts of mouse D3 cells infected with nontarget (NT) lentiviruses (MOI of 300) or with an equal mixture of three lentiviruses (MOI of 100 each) targeting XPC, RAD23B, and CETN2 (SCC KD) are analyzed by western blotting. Specific bands recognized by their respective antibodies are indicated by filled arrowheads. Asterisks denote nonspecific signals.

(B) ES cell colony morphology and alkaline phosphatase (AP) activity (red) are maintained in control D3 cells (NT, top) but are compromised in SCC-depleted D3 cells (SCC KD, bottom). See also Figure S4C.

(C) Clonal assays on SCC-depleted D3 ES cells. Stable nontarget (NT) and SCC-depleted (SCC KD) D3 cell pools were plated at 300 cells per well in 6-well plates, and emerging colonies were stained for AP activity. Differentiation status was scored based on AP staining intensity, ES cell morphology, and colony integrity after 6 days.

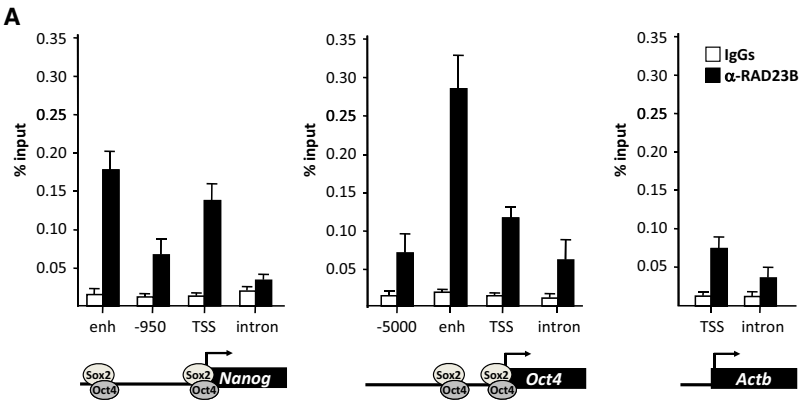
(D) Two nonoverlapping sets of shRNAs targeting SCC (SCC #1 and SCC #2) are used to deplete SCC. Quantification of *Nanog*, *Utf1*, *Fgf4*, and *Zfp42* mRNA levels are analyzed by real-time quantitative PCR (qPCR) and normalized to *Actb*. Data from representative experiments are shown; error bars represent standard deviations. n = 3.

See also Figure S4.

(Figures 5B and S4C). These knockdown cells also showed reduced proliferation rates when compared to control ES cells infected with nontarget viruses, indicating that the self-renewal capacity of ES cells depleted of SCC may also be compromised (data not shown). Indeed, prolonged depletion of SCC resulted in the apoptosis of flattened, fibroblastic AP-negative cells surrounding the collapsing ES cell colonies (Figure 5B and data not shown). Therefore, knockdown of SCC in ES cells likely promotes differentiation followed by rapid apoptosis, two processes that are often coupled. Quantification of colony assays revealed that ES cells depleted of SCC formed fewer undifferentiated colonies, with a corresponding increase in partially and fully differentiated colonies (Figure 5C). Consistent with the observed morphological changes associated with compromised stem cell identity, double and triple knockdown of XPC, RAD23B, and CETN2 resulted in a 2- to 3-fold reduction in the mRNA level of *Nanog* (Figures 5D and S4D) as well as several other stem cell markers (*Fgf4*, *Zfp42*, and *Utf1*) (Figure 5D). Knockdown of individual subunits of SCC resulted in only mild effects on *Nanog* expression (Figure S4D). Accordingly, we did not observe overt defects in self-renewal in these single-subunit knockdown ES cells (data not shown).

To further probe the molecular mechanism underpinning the function of SCC as a transcriptional coactivator for Oct4 and Sox2 in vivo, we investigated whether regulatory regions of *Nanog* and *Oct4* might serve as direct SCC targets by performing chromatin immunoprecipitation (ChIP) assays in D3 cells using a RAD23B antibody. ChIP-qPCR analysis revealed that RAD23B (and presumably XPC/SCC) occupancy sites coincide with those of Oct4 (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008) and Sox2 (Figures 6A and S5A). By contrast, we failed to detect any significant enrichment of RAD23B at housekeeping genes β-actin (*Actb*) (Figure 6A) and dihydrofolate reductase (*Dhfr*) (Figure S5B) or an intergenic region on chromosome 1 (Figure S5B).

To evaluate the extent to which Oct4 and Sox2 target sites overlap those of RAD23B on a genome-wide scale, we performed RAD23B ChIP assays followed by high-throughput sequencing (ChIP-seq) to identify an entire range of RAD23B/SCC-bound genomic regions in D3 cells. RAD23B ChIP-seq results were then compared with published Oct4 and Sox2 ChIP-seq data, along with those of Nanog and Tcf3 (Marson et al., 2008), to assess any potential bias in RAD23B occupancy in relation to these transcription factors. This analysis revealed a striking binding preference of RAD23B/SCC to genomic sites that are also co-occupied by Oct4 and Sox2, but not Nanog or Tcf3 only (~70% versus ~28%, $p < 10^{-15}$, ANOVA). This strong bias is maintained whether the ChIP-seq data sets are analyzed by the degree of peak overlap (defined by any two peaks with at least one nucleotide of overlap) (Figure 6B) or base pair coverage (Figure 6C), indicating that the majority of RAD23B/SCC-binding sites align with those of Oct4 and Sox2. Importantly, the same analyses performed on ChIP-seq samples obtained from control IgG immunoprecipitations yielded only background correlation (between 4% and 8%), confirming the specificity of the RAD23B/SCC association. We further validated the colocalization among RAD23B/SCC, Oct4, and Sox2 by measuring the distance between overlapping RAD23B/SCC and Oct4/Sox2 peaks (see Extended Experimental Procedures). The majority of them (76%) lie within close proximity (≤ 200 base pairs) of each other (Figure 6D). Even though most of RAD23B/SCC-bound regions overlap poorly with those bound by Nanog/Tcf3 (~28%), those that do are still largely (64%) positioned within 200 base pairs from each other but with a noticeably different distribution pattern than that of Oct4/Sox2 ($p < 10^{-15}$, ANOVA, Figure 6D). However, upon a closer look at the Nanog/Tcf3 "only" genomic coordinates that overlap with RAD23B-bound sites, we found that many of them (~40%) could, in fact, contain Oct4 and/or Sox2 when an alternative peak calling strategy (MACS) was used. Taken together, these data strongly suggest a classical coactivator function rather than a purely NER function of SCC both in vitro with naked DNA and in the context of chromatin in ES cells, as XPC/RAD23B-mediated DNA damage repair generally involves transient interactions with DNA (Camenisch et al., 2009) that would not show either sequence or promoter specificity.



B

Peak Overlap	RAD23B (n=33,259)	IgGs (n=15,677)
OCT4/SOX2 (n=10,138)	70.9% (7,196)	8.5% (865)
NANOG/TCF3 (n=6,424)	27.9% (1,794)	4.1% (265)

C

Base-pair Coverage	RAD23B	IgGs
OCT4/SOX2 (5,030,849 bp)	75.1% (3,776,285 bp)	8.3% (420,442 bp)
NANOG/TCF3 (1,151,570 bp)	28.5% (327,992 bp)	3.7% (42,672 bp)

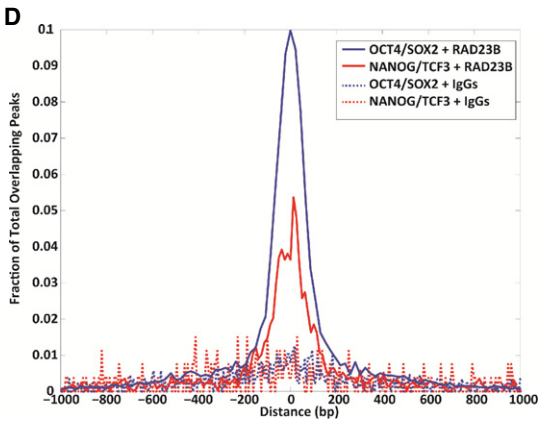


Figure 6. SCC Is Recruited to the *Nanog* and *Oct4* Promoters and Genomic Regions Occupied by Oct4 and Sox2

(A) Co-occupancy of SCC, Oct4, and Sox2 on the promoters of *Nanog* and *Oct4*. ChIP analysis of RAD23B occupancy on distal enhancers (enh), proximal promoter (transcription start site, TSS), and upstream (positions indicated by numbers) and downstream intronic regions of the *Nanog* (left), *Oct4* (middle), and *Actb* (right) gene loci. Representative data (n > 5) showing the enrichment of RAD23B (black bars) compared to normal IgGs (white bars) are analyzed by qPCR and expressed as percentage of input chromatin. Schematic diagrams of Oct4- and Sox2-binding sites on the *Nanog* and *Oct4* regulatory regions (TSS and enhancers; see also Figure S5A) are indicated at the bottom. Error bars represent standard deviations. n = 3.

(B) Percent peak overlap between RAD23B and control IgG ChIP-seq data relative to published Oct4/Sox2 and Nanog/Tcf3 peak data.

(C) Percent base pair overlap between RAD23B and control IgG ChIP-seq data relative to Oct4/Sox2 and Nanog/Tcf3 ChIP-seq data sets.

(D) Distribution of distance (in base pair) of RAD23B and control IgG peaks from Oct4/Sox2 and Nanog/Tcf3 peaks.

See also Figure S5.

Given the importance of SCC in stem cell maintenance, we next asked whether it might also play a role in the reacquisition of pluripotency during somatic cell reprogramming. Downregulation of either XPC or RAD23B in Oct4-GFP mouse embryonic fibroblasts (MEFs)—which express some SCC, albeit at significantly lower levels than ES cells—led to a dramatic reduction in the reprogramming efficiency. We observed a significant decrease in the number of AP-positive colonies, as well as a marked reduction in the percentage of partially (SSEA-1$^+$, GFP$^-$) and fully (SSEA-1$^+$, GFP$^+$) reprogrammed cells, as determined by FACS sorting (Figures 7A, 7B, and S6A). Consistent with our in vitro reconstitution result showing that the CETN2 subunit may not be essential for the transcriptional activity of SCC (Figure 4F), knockdown of CETN2 had minor effects on iPS cell derivation efficiency. As expected, reprogramming efficiency using MEFs derived from XPC and RAD23B knockout (KO) mice (Ng et al., 2003) was also highly compromised. Surprisingly, RAD23A KO MEFs were nearly as efficient as wild-type or RAD23A and B double-heterozygous MEFs in generating AP-positive colonies upon iPS cell induction (Figures 7C and S6B). This result may point to a nonredundant function of RAD23B in somatic reprogramming independent of its role in DNA repair, as RAD23B KO (and RAD23A KO) MEFs are NER proficient (Ng et al., 2003). Importantly, depletion of XPC (knockdown and knockout) and CETN2 in MEFs did not affect proliferation rates when compared to nontarget or Oct4 knockdown MEFs. However, RAD23B-depleted MEFs displayed noticeable changes in growth rates, which may partially account for the marked reduction in reprogramming efficiency (data not shown). These data suggest that efficient reprogramming may require SCC/XPC in conjunction with Oct4 and Sox2 to re-establish ES-specific gene expression programs.

DISCUSSION

Establishment of ground state pluripotency in embryonic stem cells represents one of the most remarkable events in development. Stem cells have evolved a subset of cell type-specific activators among a constellation of previously identified transcription factors and cofactors to resolve the dichotomy between self-renewal versus differentiation. Our de novo purification of the SCC/XPC complex as a potent coactivator for Oct4 and Sox2 was unanticipated but may, in part, reflect the need for stem cells to robustly expand and diversify their transcriptional repertoire while also maintaining genome integrity. Indeed, other NER factors have been shown to participate in transcriptional regulation both at the basal and activated levels.

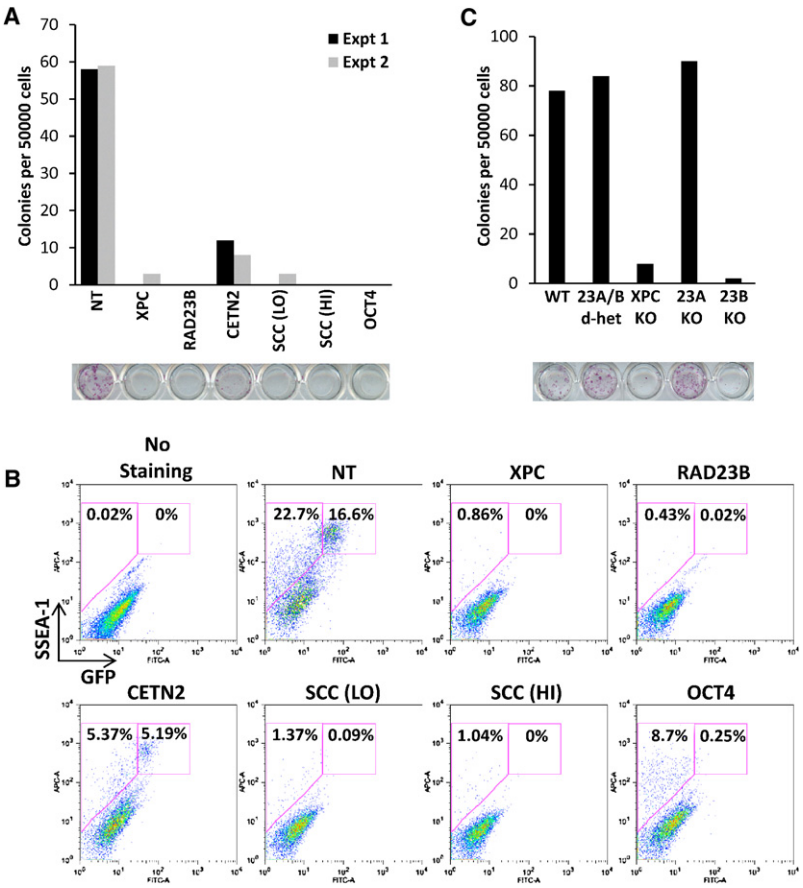


Figure 7. SCC Is Required for Efficient Somatic Cell Reprogramming

(A) Depletion of SCC blocks somatic cell reprogramming. Oct4-GFP mouse embryonic fibroblasts infected with lentiviruses expressing STEMCCA and rtTA together with nontarget shRNA (NT), shRNAs against Oct4, individual subunits of SCC, or all three subunits simultaneously at low or high multiplicity of infection (SCC LO or HI) are plated in 6-well plates for colony counting and FACS or in 24-well plates for AP staining. AP-positive (red) cells are stained and counted 17 days (14 days + dox, 3 days − dox) postinduction (dpi). Results from two separate experiments are shown.

(B) Single cell suspensions of 17 dpi Oct4-GFP MEFs as described in (A) are stained with anti-mouse SSEA-1 antibodies and analyzed by FACS.

(C) Wild-type (WT), RAD23A, and RAD23B double-heterozygous (23A/B d-Het) MEFs, together with XPC, RAD23A, and RAD23B knockout (KO) MEFs, are induced with STEMCCA. AP-positive colonies are stained and counted as in (A).

See also Figure S6.

For instance, the general transcription factor TFIIH is a classic example with established roles in both transcription initiation and NER (Schaeffer et al., 1993). Interestingly, it has recently been reported that, in HeLa cells, the entire NER complex can be assembled onto promoters of activated genes in an XPC-dependent manner. However, XPC alone is not sufficient, as other NER components appear to be responsible for RA-activated transcription (Le May et al., 2010). This finding in HeLa cells is distinct from our observation that the XPC-NER (SCC) complex plays a direct and critical role in *Nanog* transcription in vitro and in ES cells. In our studies, optimal activation of *Nanog* by Oct4/Sox2 potentiated by SCC requires a second activity present in the Q0.3 fraction. However, preliminary mass spectrometry analyses of the partially purified Q0.3 fraction failed to detect any other XP or NER factors or factors previously identified to copurify with Nanog or Oct4 in ES cells (van den Berg et al., 2010; Wang et al., 2006) (data not shown). Therefore, the SCC/XPC complex can potentiate *Nanog* transcription and likely other Oct4/Sox2-directed promoters in the absence of additional XP and NER factors in vitro. Taken together, these results suggest that the mechanism by which the SCC/XPC complex coactivates transcription in ES cells may be distinct from its function in HeLa cells.

Although XPC plays a critical role in DNA lesion recognition, XPC is not universally required for NER, as certain types of bulky DNA lesions (e.g., cholesterol-DNA adducts) can be repaired without XPC (Mu et al., 1996). Intriguingly, even though XPC is recruited to gene promoters irrespective of DNA damage signals (Le May et al., 2010), the XPC-NER complex is the only factor in the XP family that is dispensable for transcription-coupled repair (TCR) (Venema et al., 1990). Indeed, our findings suggest that the coactivator and NER duties carried out by SCC are mechanistically distinct processes, as SCC can function as part of the transcriptional cofactor apparatus via a direct interaction with Oct4 and Sox2 without requiring either DNA or TFIIH binding mediated by XPC.

It is worth noting that the effect of single knockdown of XPC or RAD23B was much more pronounced in the reprogramming of MEFs than in the maintenance of ES cells. We surmise that perhaps other redundant regulatory mechanisms in established ES cells can partially compensate for the loss of SCC. Such robust regulatory circuitries are likely to be less developed during the early phase of reprogramming in MEFs and are thus more susceptible to perturbation by SCC depletion. It is conceivable that SCC/XPC may also contribute to the process of chromatin reorganization and facilitate changes in the epigenetic landscape that are conducive to iPS conversion (Le May et al., 2010).

Also in agreement with our in vitro and cell-based studies, a mouse double KO of RAD23B and its homolog RAD23A was found to be early embryonic lethal (Ng et al., 2003). This previously puzzling phenotype can now be more readily rationalized in light of the functional role of XPC in transcriptional coactivation revealed here. Taken together, these results strongly suggest that loss of the SCC/XPC complex may indeed compromise the transcriptional integrity of pluripotent stem cells, as well as the ability of somatic cells to re-establish pluripotency. However, XPC KO mice are UV sensitive but otherwise normal, with no obvious developmental defects (Sands et al., 1995). It has been shown that RAD23B is in vast excess relative to XPC (Sugasawa

et al., 1996), suggesting that RAD23B may exist in other complexes independent of XPC that functionally replace SCC.

Embryonic stem cells are thought to be under strong selective pressure to maintain genome fidelity because accumulation and propagation of DNA errors to progenitor cells would be lethal during development; therefore, DNA damage response factors and pathways are often upregulated in ES cells (e.g., XPC, RAD23B, ERCC5, etc.) (Cervantes et al., 2002; Ramalho-Santos et al., 2002). Should DNA repair fail, UV-damaged ES cells can be eliminated first by repressing *Nanog* expression through p53 upregulation, which in turn promotes spontaneous differentiation and efficient apoptosis (Lin et al., 2005). It is interesting to note that, upon UV-induced DNA damage in HeLa cells, recruitment of XPC to non-UV-inducible genes, as well as their expression, are dramatically delayed (Le May et al., 2010). This suggests that some sort of redistribution mechanism may redirect XPC from transcription duty at promoter targets to the NER pathway in response to DNA damage. In light of these observations, it is tempting to speculate that redistribution of XPC-RAD23B-CETN2 from *Nanog* and presumably other Oct4/Sox2-regulated promoters to DNA damage sites may provide an efficient sensing mechanism to perturb stem cell-specific gene expression programs and thus provide a window of opportunity for ES cells to either repair the lesions or commit to differentiation and apoptosis. The SCC/XPC complex may therefore act as a molecular link to couple stem cell-specific gene expression programs and genome surveillance in ES cells.

EXPERIMENTAL PROCEDURES

DNA Constructs, Cell Lines, and Cell Culture

Construction of in vitro transcription templates and protein expression plasmids are described in Extended Experimental Procedures. HeLa, 293T, NTERA-2 (NT2), and mouse ES cell line D3 were maintained in standard conditions. Differentiation of D3 ES cells was carried out by LIF removal followed by retinoic acid treatment (5–10 μM, Sigma).

Purification and Identification of SCC

Nuclear extracts from ~400 l of NT2 cells were purified over eight chromatographic steps to homogeneity. Methods for purification and mass spectrometry analyses of SCC are detailed in Extended Experimental Procedures.

Western Blotting, Immunoprecipitation, and Affinity Purification

Antibodies used are described in Extended Experimental Procedures. Transcriptional activators were purified from transiently transfected HeLa cells followed by affinity purification using anti-FLAG (M2) agarose (Sigma) as described in Extended Experimental Procedures. Recombinant SCC complexes were purified from Sf9 cells infected with baculoviruses (BAC-to-BAC system, Invitrogen) expressing N-terminal His_6-tagged or FLAG-tagged XPC, N-FLAG-tagged RAD23B, and untagged CETN2. Sf9 cells were harvested 48 hr after infection, and protein complexes were purified by incubating cell lysates with Ni-NTA resin (QIAGEN), anti-FLAG (M2) agarose (Sigma), and elution by the FLAG peptides.

shRNA-Mediated Knockdown of SCC by Lentiviral Infection

Control nontarget and pLKO shRNA plasmids targeting XPC, RAD23B, and CETN2 (Sigma) were transfected with packaging vectors into 293T cells using FuGENE 6 (Roche). Supernatants were concentrated by ultracentrifugation and resuspended in PBS. Viral titer was determined by a QuickTiter Lentivirus Titer Kit (Cell Biolabs). SCC knockdown was performed by incubating lentiviral concentrates with D3 cells in the presence of 8 μg/ml polybrene followed by puromycin selection (1.5 μg/ml).

Gene Expression Analysis and ChIP

Total RNA from shRNA-mediated knockdown D3 ES cells was isolated using RNeasy Plus Kit (QIAGEN) and analyzed by qRT-PCR. Chromatin immunoprecipitation (ChIP) assays were performed in D3 cells as described in Extended Experimental Procedures. Precipitated DNA was measured by qPCR or sequenced using an Illumina HiSeq 2000 sequencing platform. Methods for gene expression and ChIP analyses are detailed in Extended Experimental Procedures.

Somatic Cell Reprogramming

Oct4-GFP MEFs (The Jackson Laboratory) were infected with lentiviruses containing STEMCCA and rtTA, followed by infection with pLKO shRNA lentiviral supernatants targeting SCC. Oct4, Sox2, Klf4, and c-Myc expressions were induced by doxycycline, and SCC knockdown MEFs were selected with puromycin. Reprogrammed cells were either detected by alkaline phosphatase activity or stained with anti-SSEA-1 antibodies conjugated to Alexa Fluor 647 (BioLegends) and analyzed by FACS. XPC, RAD23A, and RAD23B knockout MEFs were generous gifts from Dr. Hoeijmakers (Rotterdam, The Netherlands).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and six figures and can be found with this article online at doi:10.1016/j.cell.2011.08.038.

ACKNOWLEDGMENTS

The authors wish to thank A. Fischer and M. Richner at the Tissue Culture Facility (University of California, Berkeley); S. Zheng, G. Dailey, M. Haggart, and E. Bourbon for technical assistance; S. Chen at National Institute of Biological Sciences (Beijing, China) for mass spectrometry analysis; S. Zhou for initial mass spectrometry analysis; J. Hoeijmakers for knockout MEFs; G. Mostoslavsky for STEMCCA; O. Puig for pGL3-CAT plasmid; S. Ryu for purified rTBP and G3BCAT plasmid; K. Hochedlinger, M. Stadtfeld, J. de Wit, and M. Holmes for technical advice; and M. Botchan, M. Holmes, M. Levine, S. Martin, M. Rape, D. Rio, and all members of our laboratory for critical reading of the manuscript. Y.W.F. was a California Institute for Regenerative Medicine Scholar (CIRM training grant T1-00007). T.Y. was supported by the Swiss National Science Foundation and the Siebel Stem Cell Institute and is a Jane Coffin Childs fellow.

Received: July 26, 2011
Revised: August 23, 2011
Accepted: August 25, 2011
Published: September 29, 2011

REFERENCES

Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y., and Hanaoka, F. (2001). Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. J. Biol. Chem. *276*, 18665–18672.

Bernardes de Jesus, B.M., Bjørås, M., Coin, F., and Egly, J.M. (2008). Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. Mol. Cell. Biol. *28*, 7225–7235.

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., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Camenisch, U., Träutlein, D., Clement, F.C., Fei, J., Leitenstorfer, A., Ferrando-May, E., and Naegeli, H. (2009). Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. EMBO J. *28*, 2387–2399.

Cervantes, R.B., Stringer, J.R., Shao, C., Tischfield, J.A., and Stambrook, P.J. (2002). Embryonic stem cells and somatic cells differ in mutation frequency and type. Proc. Natl. Acad. Sci. USA *99*, 3586–3590.

Chakravarthy, H., Boer, B., Desler, M., Mallanna, S.K., McKeithan, T.W., and Rizzino, A. (2008). Identification of DPPA4 and other genes as putative Sox2: Oct-3/4 target genes using a combination of in silico analysis and transcription-based assays. J. Cell. Physiol. *216*, 651–662.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Engelen, E., Akinci, U., Bryne, J.C., Hou, J., Gontan, C., Moen, M., Szumska, D., Kockx, C., van Ijcken, W., Dekkers, D.H., et al. (2011). Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. Nat. Genet. *43*, 607–611.

Gao, X., Tate, P., Hu, P., Tjian, R., Skarnes, W.C., and Wang, Z. (2008). ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. Proc. Natl. Acad. Sci. USA *105*, 6656–6661.

Goodrich, J.A., and Tjian, R. (2010). Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat. Rev. Genet. *11*, 549–558.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567–582.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat. Cell Biol. *10*, 353–360.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kidder, B.L., Palmer, S., and Knott, J.G. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. Stem Cells *27*, 317–328.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049–1061.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.Y., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol. Cell. Biol. *25*, 2475–2485.

Le May, N., Mota-Fernandes, D., Vélez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010). NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. Mol. Cell *38*, 54–66.

Lemon, B., Inouye, C., King, D.S., and Tjian, R. (2001). Selectivity of chromatin-remodelling cofactors for ligand-activated transcription. Nature *414*, 924–928.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat. Cell Biol. *7*, 165–171.

Liu, Z., Scannell, D.R., Eisen, M.B., and Tijan, R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. Cell *146*, 720–731.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. *38*, 431–440.

Maillard, O., Solyom, S., and Naegeli, H. (2007). An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. PLoS Biol. *5*, e79.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521–533.

Mu, D., Hsu, D.S., and Sancar, A. (1996). Reaction mechanism of human DNA repair excision nuclease. J. Biol. Chem. *271*, 8285–8294.

Müller, F., Zaucker, A., and Tora, L. (2010). Developmental regulation of transcription initiation: more than just changing the actors. Curr. Opin. Genet. Dev. *20*, 533–540.

Näär, A.M., Beaurang, P.A., Robinson, K.M., Oliner, J.D., Avizonis, D., Scheek, S., Zwicker, J., Kadonaga, J.T., and Tjian, R. (1998). Chromatin, TAFs, and a novel multiprotein coactivator are required for synergistic activation by Sp1 and SREBP-1a in vitro. Genes Dev. *12*, 3020–3031.

Näär, A.M., Beaurang, P.A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. Nature *398*, 828–832.

Näär, A.M., Lemon, B.D., and Tjian, R. (2001). Transcriptional coactivator complexes. Annu. Rev. Biochem. *70*, 475–501.

Näär, A.M., Taatjes, D.J., Zhai, W., Nogales, E., and Tjian, R. (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. Genes Dev. *16*, 1339–1344.

Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J.H. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev. *17*, 1630–1645.

Pal, R., and Ravindran, G. (2006). Assessment of pluripotency and multilineage differentiation potential of NTERA-2 cells as a model for studying human embryonic stem cells. Cell Prolif. *39*, 585–598.

Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C., and Melton, D.A. (2002). "Stemness": transcriptional profiling of embryonic and adult stem cells. Science *298*, 597–600.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J. Biol. Chem. *280*, 24731–24737.

Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett. *579*, 909–915.

Rosenfeld, M.G., Lunyak, V.V., and Glass, C.K. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev. *20*, 1405–1428.

Ryu, S., Zhou, S., Ladurner, A.G., and Tjian, R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. Nature *397*, 446–450.

Sands, A.T., Abuin, A., Sanchez, A., Conti, C.J., and Bradley, A. (1995). High susceptibility to ultraviolet-induced carcinogenesis in mice lacking XPC. Nature *377*, 162–165.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., and Egly, J.M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science *260*, 58–63.

Schwartz, C.M., Spivak, C.E., Baker, S.C., McDaniel, T.K., Loring, J.F., Nguyen, C., Chrest, F.J., Wersto, R., Arenas, E., Zeng, X., et al. (2005). NTera2: a model system to study dopaminergic differentiation of human embryonic stem cells. Stem Cells Dev. *14*, 517–534.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543–549.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O., and Thomson, J.A. (2003). Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc. Natl. Acad. Sci. USA *100*, 13350–13355.

Sugasawa, K., Masutani, C., Uchida, A., Maekawa, T., van der Spek, P.J., Bootsma, D., Hoeijmakers, J.H., and Hanaoka, F. (1996). HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. Mol. Cell. Biol. *16*, 4852–4861.

Taatjes, D.J., Marr, M.T., and Tjian, R. (2004). Regulatory diversity among metazoan co-activator complexes. Nat. Rev. Mol. Cell Biol. *5*, 403–410.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). 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.

Tutter, A.V., Kowalski, M.P., Baltus, G.A., Iourgenko, V., Labow, M., Li, E., and Kadam, S. (2009). Role for Med12 in regulation of Nanog and Nanog target genes. J. Biol. Chem. *284*, 3709–3718.

Uchida, A., Sugasawa, K., Masutani, C., Dohmae, N., Araki, M., Yokoi, M., Ohkuma, Y., and Hanaoka, F. (2002). The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. DNA Repair (Amst.) *1*, 449–461.

van den Berg, D.L., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369–381.

van den Berg, D.L., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol. Cell. Biol. *28*, 5986–5995.

Venema, J., van Hoffen, A., Natarajan, A.T., van Zeeland, A.A., and Mullenders, L.H. (1990). The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. Nucleic Acids Res. *18*, 443–448.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

Yasuda, G., Nishi, R., Watanabe, E., Mori, T., Iwai, S., Orioli, D., Stefanini, M., Hanaoka, F., and Sugasawa, K. (2007). In vivo destabilization and functional defects of the xeroderma pigmentosum C protein caused by a pathogenic missense mutation. Mol. Cell. Biol. *27*, 6606–6614.

Zhang, J., Tam, W.L., Tong, G.Q., Wu, Q., Chan, H.Y., Soh, B.S., Lou, Y., Yang, J., Ma, Y., Chai, L., et al. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat. Cell Biol. *8*, 1114–1123.

Zhang, X., Zhang, J., Wang, T., Esteban, M.A., and Pei, D. (2008). Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. J. Biol. Chem. *283*, 35825–35833.