bioRχiv

ARTICLE PREPRINT

ATP-binding cassette protein ABCF1 couples transcription and genome surveillance in embryonic stem cells through low complexity domain

ABCF1 links transcription and genome surveillance

Eun-Bee Choi[1,2,5], Munender Vodnala[1,2,5], Madeleine Zerbato[1,2,5], Jianing Wang[1,2], Jaclyn J. Ho[3,4], Carla Inouye[3,4], Yick W. Fong[1,2,6*]

[1]Brigham Regenerative Medicine Center, Brigham and Women's Hospital, Boston, MA, USA
[2]Department of Medicine, Cardiovascular Medicine Division, Harvard Medical School, Boston, MA, USA
[3]Department of Molecular and Cell Biology, Li Ka Shing Center for Biomedical and Health Sciences, California Institute for Regenerative Medicine Center of Excellence, University of California at Berkeley, Berkeley, CA, USA
[4]Howard Hughes Medical Institute, Berkeley, CA, USA
[5]These authors contributed equally
[6]Lead Contact
[*]Correspondence: yfong@bwh.harvard.edu

OCT4 and SOX2 confer pluripotency by recruiting coactivators to activate stem cell-specific transcription. However, the composition of coactivator complexes and their roles in maintaining stem cell fidelity remain unclear. Here we report the ATP-binding cassette subfamily F member 1 (ABCF1) as a coactivator for OCT4/SOX2 critical for stem cell self-renewal. The intrinsically disordered low complexity domain (LCD) of ABCF1 forms phase-separated condensates and stimulates pluripotency gene transcription by mediating multivalent interactions with SOX2 and co-dependent coactivators XPC and DKC1. These LCD-driven transcription factor-coactivator interactions at pluripotency genes are disrupted by aberrant intracellular DNAs arising from genome instability, due to ABCF1's strong propensity to form condensates with DNA. These results support the idea that the LCD of ABCF1 directly couples pluripotency gene transcription with sensing aberrant DNAs and acts as a checkpoint for self-renewal to safeguard stem cell genome integrity.

INTRODUCTION

Stem cell pluripotency is largely driven by core transcription factors including OCT4 and SOX2 (*1*, *2*). Genome-wide studies demonstrated extensive co-binding of OCT4 and SOX2 at key pluripotency genes and across the ES cell genome (*3–6*). It has also been shown that proper reactivation of this pluripotency gene network by OCT4 and SOX2 in somatic cells is a critical barrier to cellular reprogramming (*7–9*). Therefore, OCT4 and SOX2 are essential for both maintenance and reacquisition of stem cell pluripotency.

Transcriptional activation of pluripotency genes by OCT4 and SOX2 requires transcriptional coactivators (*10–12*). Recent studies showed that multivalent interactions between intrinsically disordered low-complexity sequence domains (LCDs) found in coactivators such as Mediator and transcriptional activators OCT4 and SOX2 form protein-rich hubs at genomic loci, thereby increasing the local concentration of these factors to promote transactivation (*13–15*). Mediator is required for transcription of most RNA polymerase II (Pol II) genes in somatic and pluripotent cells, by virtue of its ability to interact with Pol II and diverse transcription factors (*16*, *17*). However, we and others previously have shown that OCT4 and SOX2 require nuclear factors enriched in pluripotent stem cells to robustly activate transcription (*18*, *19*). Therefore, LCD-driven interactions by Mediator and other cell-ubiquitous coactivators are unlikely sufficient to account for the cell-type and gene-specific transactivation by OCT4 and SOX2 in ES cells. We surmised that additional coactivators may be involved to nucleate the assembly of stem cell-specific transcriptional hubs at pluripotency gene loci.

We took an unbiased in vitro approach to identify coactivators that can reconstitute transcriptional activation by OCT4 and SOX2 (*18*). By biochemical fractionation of human pluripotent cell nuclear extracts, we uncovered three coactivators that are enriched in ES cells and specifically

required by OCT4 and SOX2 to activate transcription of the *NANOG* promoter in vitro (*18*). We reported previously that the first two stem cell coactivators (SCCs) are the XPC nucleotide excision repair complex and the dyskerin (DKC1) ribonucleoprotein complex (RNP) (*18*, *20–23*). However, we found that robust activation of the *NANOG* gene by XPC and DKC1 requires an additional coactivator activity (SCC-B) (*22*). Therefore, revealing the identity and the transcriptional mechanisms by which SCC-B coordinates with XPC and DKC1 to promote the stem cell fate is fundamental to understanding the molecular basis of pluripotency.

Here, we identify SCC-B as ATP-binding cassette subfamily F member 1 (ABCF1). We demonstrate that ABCF1 readily undergoes liquid-liquid phase separation (LLPS) mediated by its N-terminal LCD to form droplets in vitro. Using biochemical approaches and chromatin immunoprecipitation (ChIP) assays, we show that the LCD of ABCF1 potently stimulates OCT4/SOX2-dependent transcriptional activation by mediating selective multivalent interactions with XPC, DKC1, SOX2, and Pol II, thereby forming a stem cell-specific transcriptional ensemble at pluripotency gene promoters. Consistent with its role as a critical stem cell coactivator, disruption of ABCF1 compromises stem cell maintenance, pluripotency gene expression, and somatic cell reprogramming.

In somatic cells, ABCF1 has been implicated in the detection of intracellular DNA and ubiquitin conjugation in the innate immune pathway (Arora et al., 2019; Lee et al., 2013). We provide evidence that the LCD is also required for DNA binding. More importantly, formation of ABCF1 LLPS droplets is dramatically enhanced by short double-stranded DNAs (dsDNAs), suggesting several mechanisms for regulating ABCF1 assembly in cells that may compete with transcription, by preferentially forming ABCF1-dsDNA condensates. Indeed, we show that when ES cells are challenged with DNA damage or pathogen-derived DNAs, ABCF1 binds these aberrant DNAs accumulated in cells, resulting in loss of interaction with SOX2 and dissociation of ABCF1 from gene promoters targeted by SOX2 and OCT4. This leads to disruption of pluripotency gene expression and elimination of compromised ES cells through spontaneous differentiation. Our findings reveal that ABCF1 directly couples sensing of genome instability and infections to the pluripotency gene network through its LCD. Given that the canonical innate immune pathway is severely suppressed in ES cells (*28*, *29*), leveraging an innate immune DNA sensor such as ABCF1 as transcriptional coactivator could provide a means for ES cells to respond to these cellular insults and maintain a pristine pool of progenitors as they self-renew.

RESULTS

A Stem Cell Coactivator Essential for SCC-dependent Transcriptional Activation by OCT4 and SOX2

We have previously shown that robust activation of the human *NANOG* promoter by OCT4 and SOX2 in vitro requires three distinct stem cell coactivators (SCCs) present in a human pluripotent cell nuclear extract (*18*, *22*). The first SCC, the XPC complex, separated from the other two at the Poros-HQ anion exchange chromatographic step, while the second SCC, the DKC1 complex (SCC-A), segregated from the remaining unknown coactivator (SCC-B) at the Poros-Heparin (Poros-HE) step (Fig. 1A). Starting with nuclear extracts prepared from 400 L of a pluripotent human embryonal carcinoma cell line N-TERA2 (NT2), we serially fractionated the nuclear extracts over seven chromatographic columns. We tracked SCC-B activity in various protein fractions by assessing their ability to restore SCC-dependent transcriptional activation by OCT4 and SOX2 (Schematic diagram of the in vitro transcription and biochemical complementation assay is shown in Fig. S1A). In the final Mono S chromatographic step, salt-eluted fractions were assayed in in vitro transcription reactions. We found that adding fractions 14 and 15 to reactions containing purified XPC and DKC1 complexes potently stimulated transcription of the *NANOG* promoter template (Fig. 1B). These results demonstrated an important role of SCC-B in mediating cooperative coactivation by SCCs. Furthermore, our results suggested that the bulk of SCC-B likely resided in these fractions. Accordingly, SDS-PAGE of these Mono S fractions revealed that fractions 14 and 15 were highly enriched with a polypeptide at ~110 kDa along with multiple apparent breakdown products (Fig. 1C). To identify these polypeptides that co-migrate with SCC-B activity, peak Mono S fractions were pooled and separated by SDS-PAGE. Tryptic digestion of the gel slice containing the 110 kDa protein band followed by mass spectrometry identified SCC-B to be ATP-binding cassette subfamily F member 1 (ABCF1) (Fig. 1D). Identification of ABCF1 as the active constituent of SCC-B activity was unexpected because it has not been previously associated with transcriptional regulation or any cellular processes in the nucleus. To corroborate the mass spectrometry results, we showed by western blotting that ABCF1 resides in nuclear and cytoplasmic fractions of human and mouse ES cells (Fig. S1B), and is enriched exclusively in the phosphocellulose 1M KCl (P1M) nuclear fraction that contains SCC activities, but not in the transcriptionally inactive P0.3 and P0.5 fractions (Fig. 1A and Fig. S1C) (*18*). Furthermore, we found that the mRNA and protein level of ABCF1 in both human and mouse ES cells decreased sharply when these cells were induced to undergo differentiation (Fig. 1, E and F). Here, we identified ABCF1 as a stem cell-enriched coactivator and a key player in mediating the synergistic transcriptional activation by OCT4 and SOX2 (*18*).

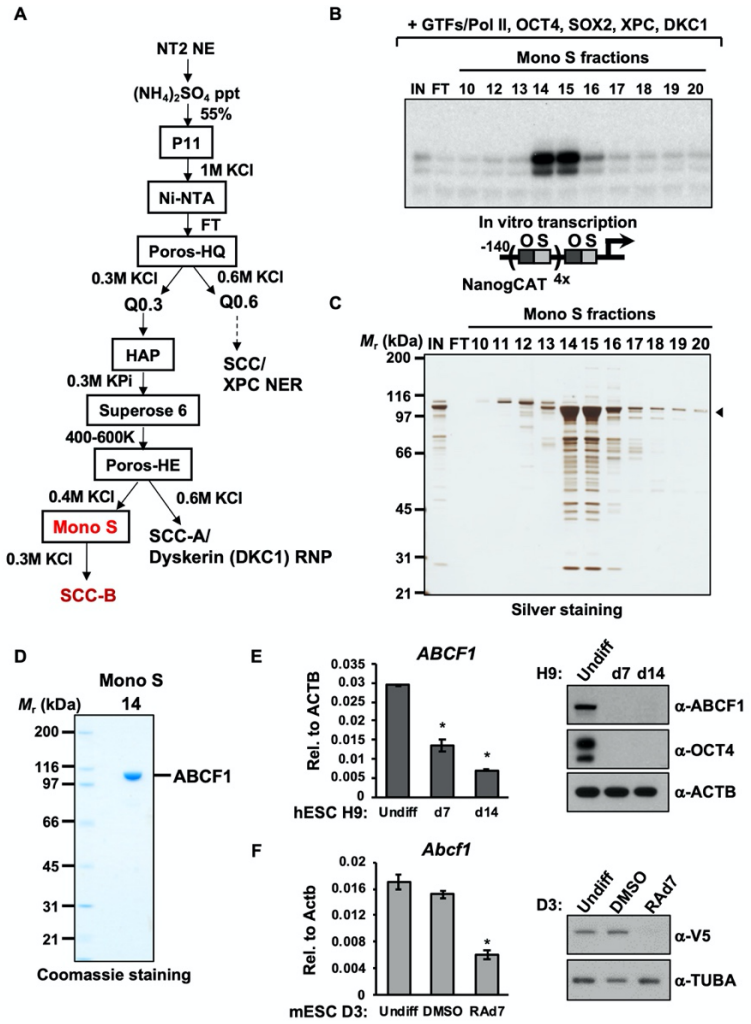

Fig. 1. Purification and identification of Stem Cell Coactivator-B (SCC-B) as ATP-binding cassette subfamily F member 1 (ABCF1). **a,** Chromatography scheme for purification of SCC-B from NT2 nuclear extracts (NT2 NE). NT2 NE is first subjected to ammonium sulfate precipitation (55% saturation) followed by a series of chromatographic columns including nickel affinity agarose (Ni-NTA), cation exchangers phosphocellulose (P11), heparin (Poros-HE), Mono S, anion exchanger Poros-HQ, hydroxyapatite (HAP), and gel filtration medium Superose 6. The XPC nucleotide excision repair (NER) complex separates from SCC-A and SCC-B at the Poros-HQ step. SCC-B activity segregates from SCC-A (Dyskerin (DKC1) RNP) at the Poros-HE step. **b,** Input fraction containing SCC-B activity from the Poros-HE step (IN), flow-through (FT) and various salt-eluted Mono S fractions were assayed for their ability to stimulate OCT4/SOX2-dependent transcription from the human *NANOG* promoter template engineered with four extra copies of the oct-sox composite binding element (bottom). All reactions contain purified general transcription factors (GTFs), Pol II, OCT4, SOX2, and recombinant XPC and DKC1 complexes. **c,** Fractions assayed in vitro transcription shown in b are separated on a 4-12% gradient polyacrylamide gel and stained with silver. Filled arrowhead indicates the predominant polypeptide at ~110 kDa that co-migrates with SCC-B transcriptional activity. **d,** Coomassie staining of Mono S fraction 14 demonstrates purification to homogeneity. The breakdown products observed in silver staining in c are minor compared to the full-length product. Mass spectrometry analysis of Mono S peak activity fractions (#14-16) in b identifies the ~110 kDa polypeptide as ABCF1. **e,** ABCF1 is enriched in human ES cells. Downregulation of ABCF1 in human ES cell line H9 upon exit of pluripotency. H9 human ES cells are induced to differentiate in differentiation medium (DMEM/F12 with FBS). Cells are collected at day 7 and 14 post induction. mRNA levels of *ABCF1* are analyzed by quantitative PCR (qPCR) and normalized to *β-actin* (*ACTB*). Western blot analyses of whole cell extracts (WCEs) prepared from H9 cells collected at indicated days are performed using antibodies against ABCF1, pluripotency marker OCT4, and ACTB as loading control. **f,** ABCF1 is enriched in mouse ES cells. mRNA and protein levels of ABCF1 are analyzed in pluripotent D3 mouse ES cells carrying V5-epitope tagged *Abcf1* alleles (V5-ABCF1 knock-in), and in the same ES cell line induced to differentiate by retinoic acid (RA) treatment for 7 days (RA d7) by qPCR and western blotting, respectively. Error bars present SEM. $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.
The following figure supplement is available for figure 1: Supplementary Figure S1.

ABCF1 is Required for Pluripotency Gene Expression and Stem Cell Pluripotency

We next set out to determine if ABCF1 is required for stem cell maintenance and pluripotency by performing loss-of-function studies using lentiviruses expressing shRNAs that target ABCF1 (Fig. 2A and Fig. S2A). Compared to control D3 mouse ES cells, ABCF1 knockdown resulted in rapid collapse of the tightly packed ES cell colonies and the appearance of flattened and elongated cells with reduced alkaline phosphatase (AP) activity, a marker for pluripotent cells (Fig. 2B) (*30*, *31*). We also found that depletion of ABCF1 in both mouse and human ES cells reduced

proliferation and viability (Fig. S2B). These observations indicated that loss of ABCF1 in ES cells compromises self-renewal capacity and promotes spontaneous differentiation. Consistent with the morphological changes indicative of compromised stem cell identity, depletion of ABCF1 in mouse ES cells resulted in significant decrease in mRNA levels of key pluripotency-associated genes, some of which are known direct targets of OCT4 and SOX2 (Fig. 2C) (*4*, *6*, *32*). Concomitant with the downregulation of these pluripotency genes, we observed increased expression of lineage-specific genes related to the ectoderm, mesoderm, and trophectoderm, at the expense of endoderm marker induction (Fig. 2D). These data suggest that loss of ABCF1 destabilizes the pluripotency gene network and restricts the ability of ES cells to efficiently differentiate into all three embryonic germ layers. Our finding is consistent with a previously unexplained observation that *Abcf1* knockout mouse embryos die at 3.5 days post coital, a developmental stage that coincides with the emergence of pluripotent cells in the inner cell mass of the blastocyst (*33*) and from which ES cells are derived (*34–36*).

Given the importance of ABCF1 in stem cell maintenance and expression of genes that are also known to promote somatic cell reprogramming (e.g. *Nanog*, *Klf4*, *Esrrb*, *Prdm14, Nr5a2*) (Fig. 2C) (*37–40*), we next assessed whether ABCF1 is required for the generation of induced pluripotent stem (iPS) cells from mouse embryonic fibroblasts (MEFs). To this end, we transduced MEFs with lentiviruses expressing non-targeting control shRNA or two independent shRNAs specific for ABCF1 and initiated reprogramming by doxycycline (dox)-induced expression of OCT4, KLF4, SOX2, and c-MYC (*41*). ABCF1 knockdown resulted in a marked decrease in the number of AP-positive iPS cell colonies formed (Fig. 2E and Fig. S2C). Flow cytometry analysis further demonstrated that loss of ABCF1 led to a decrease in cells expressing SSEA-1+, a cell surface marker that indicates early stage of reprogramming (Fig. 2F) (*30*, *42*). Our results are consistent with a role of ABCF1 in overcoming early barriers to reacquisition of pluripotency during cellular reprogramming.

We reasoned that if indeed ABCF1 functions as a transcriptional coactivator for OCT4 and SOX2 in ES cells, ABCF1 ought to occupy the genome at cis-regulatory regions that are bound by OCT4 and SOX2. Therefore, we performed chromatin immunoprecipitation (ChIP) assays to investigate the binding of ABCF1 at known pluripotency gene loci targeted by OCT4 and SOX2. Because we could not find a commercially available antibody against ABCF1 suitable for ChIP, we employed CRISPR/Cas9 genome editing to knock-in a V5 epitope tag into the N-terminus of the *Abcf1* gene to enable ChIP with an anti-V5 antibody. We confirmed D3 mouse ES cell clones with biallelic knock-in of the V5-tag by genomic sequencing and western blotting (Fig. S2D). Importantly, these knock-in lines express pluripotency and differentiation-associated genes at a level comparable to wild-type cells, indicating that epitope tagging did not compromise ABCF1 function (Fig. S2E). We performed micrococcal nuclease-ChIP (MNase-ChIP) using formaldehyde as well as the protein-protein crosslinker ethylene glycol bis[succinimidylsuccinate] (EGS) (*43*) and found that ABCF1 was enriched at OCT4/SOX2 co-bound regions in pluripotency genes such as *Oct4 (Pou5f1)*, *Sox2*, *Nanog*, *Lefty1*, and *Trim28* (Fig. 2G and Fig. S2, F and G). By contrast, we did not observe any significant enrichment of ABCF1 at housekeeping genes *β-actin* (*Actb*) and *Dhfr* (Fig. 2H and Fig. S2G), consistent with ABCF1 acting as a selective coactivator for OCT4 and SOX2. Taken together, the data presented thus far suggest a classical coactivator function of ABCF1 both in vitro with naked DNA and in the context of chromatin in ES cells.

Low Complexity Domain in ABCF1 Mediates LLPS

ABCF1 belongs to a large class of transporters that couples ATP hydrolysis to the active transport of substrates across cell membranes (*44*). While most ABC proteins contain transmembrane domains (TMDs) as expected, ABCF1 lacks TMD and is therefore not predicted to function as a transporter. In somatic cells, ABCF1 has been implicated as a regulator of mRNA translation (*25*, *45*), an E2 ubiquitin-conjugating enzyme (*27*), and a sensor for intracellular double-stranded DNAs involved in innate immune responses (*26*). It is worth noting that these reported activities of ABCF1 all reside in the cytoplasm. Our findings here thus point to a hitherto unknown nuclear function of ABCF1 in gene regulation in ES cells.

ABCF1 contains two conserved nucleotide-binding domains (NBDs) and a ~ 300 amino acid N-terminal domain that is predicted to be highly disordered (*46*) (Fig. 3A, Top). This is in agreement with a recent structural study on ABCF1 which indicated that the N-terminal domain cannot be crystallized (*47*). The N-terminal domain of ABCF1 is a low-complexity region that contains a polyglutamine (polyQ) tract, and is unusually rich in charged amino acids, of which ~40% are divided between lysine (K) and glutamic acid (E) residues (Fig. 3A, Bottom) (*24*, *48*, *49*). This unique amino acid composition is conserved among vertebrates (Fig. 3B) and consistent with one of the largest classes of intrinsically disordered region (IDR) called polyampholytes (*50*, *51*). These clusters of positively and negatively charged amino acid, interspersed with hydrophobic residues such as phenylalanine, are known to promote LLPS (*52*). Given the recent appreciation that the ability of LCDs in transcription factors and coactivators to form LLPS droplets in vitro often correlates with their transactivation activities (*13–15*, *53*), we next investigated whether ABCF1 proteins are also capable of forming liquid droplets in vitro and in cells. Recombinant GFP-fusion proteins of full-length (FL) ABCF1 were purified and added to reaction buffers containing a crowding agent (10% PEG-8000) to mimic the intracellular crowding environment in vitro (Fig. 3C and Fig. S3A). Fluorescence microscopy of the protein mixture revealed that ABCF1 readily formed micrometer-sized liquid droplets at or above physiological salt concentration (200 mM NaCl) (Fig. 3D).

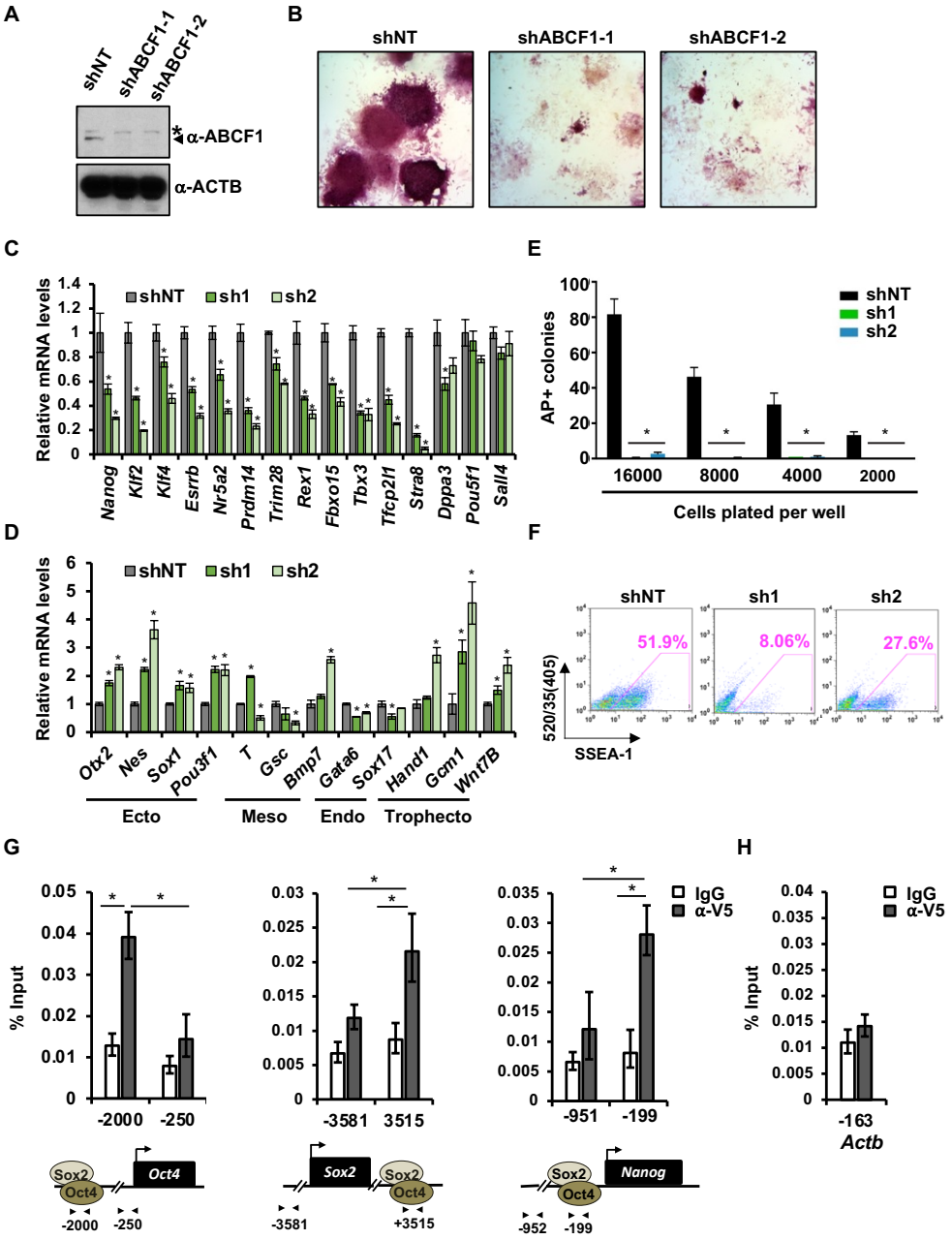

Fig. 2. ABCF1 is required for stem cell pluripotency and recruited to regulatory regions of key pluripotency genes targeted by OCT4 and SOX2 in mouse ES cells. a, shRNA-mediated knockdown of ABCF1 in mouse ES cells. WCEs of mouse ES cells transduced with lentiviruses expressing control non-targeting shRNA (shNT) or two independent shRNAs against ABCF1 (shABCF1-1 or shABCF1-2) are analyzed by western blotting to confirm knockdown efficiency. ACTB is used as loading control. Asterisk denotes a non-specific band. **b,** shRNA-mediated depletion of ABCF1 in mouse ES cells leads to colony collapse with flattened cell morphology and reduced alkaline phosphatase (AP) staining, indicating spontaneous differentiation. **c,** Loss of ABCF1 in mouse ES cells compromises pluripotency gene expression. Quantification of mRNA levels of pluripotency genes are analyzed by qPCR and normalized to *Actb*. mRNA level in each sample is expressed relative to its respective control set as 1. **d,** Depletion of ABCF1 induces expression of differentiation-associated genes. Lineage-specific genes representing the three embryonic germ layers and the trophectoderm are analyzed by qPCR as in c. **e,** Depletion of ABCF1 blocks somatic cell reprogramming. CF-1 MEFs are transduced with lentiviruses expressing either a control shRNA (NT) or shRNAs targeting ABCF1 (sh1 and sh2), together with lentiviruses expressing OCT4, KLF4, SOX2, and c-MYC. Cells are plated at the indicated number in 24-well plates; cellular reprogramming is initiated by the addition of doxycycline (dox). Cells are stained for alkaline phosphatase (AP) activity and AP + colonies are counted after 14 days (11 days with dox followed by 3 days of dox withdrawal) post induction (dpi). **f,** Single cells suspensions of 14 dpi reprogrammed CF-1 MEFs as described in e are stained with anti-mouse SSEA-1 and analyzed by flow cytometry. **g,** Micrococcal nuclease (MNase)-chromatin immunoprecipitation (ChIP) analysis of ABCF1 occupancy on control and enhancer regions of *Oct4*, *Sox2,* and *Nanog* gene loci in V5-ABCF1 knock-in D3 mouse ES cells. Representative data showing the enrichment of V5-ABCF1 (grey bars) compared to control IgGs (white bars) are analyzed by qPCR and expressed as percentage of input chromatin. **h,** ABCF1 is not recruited to the promoter of housekeeping gene *Actb*. Enrichment of ABCF1 is analyzed as in g. Error bars present SEM. $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.

The following figure supplement is available for figure 2: Supplementary Figure S2.

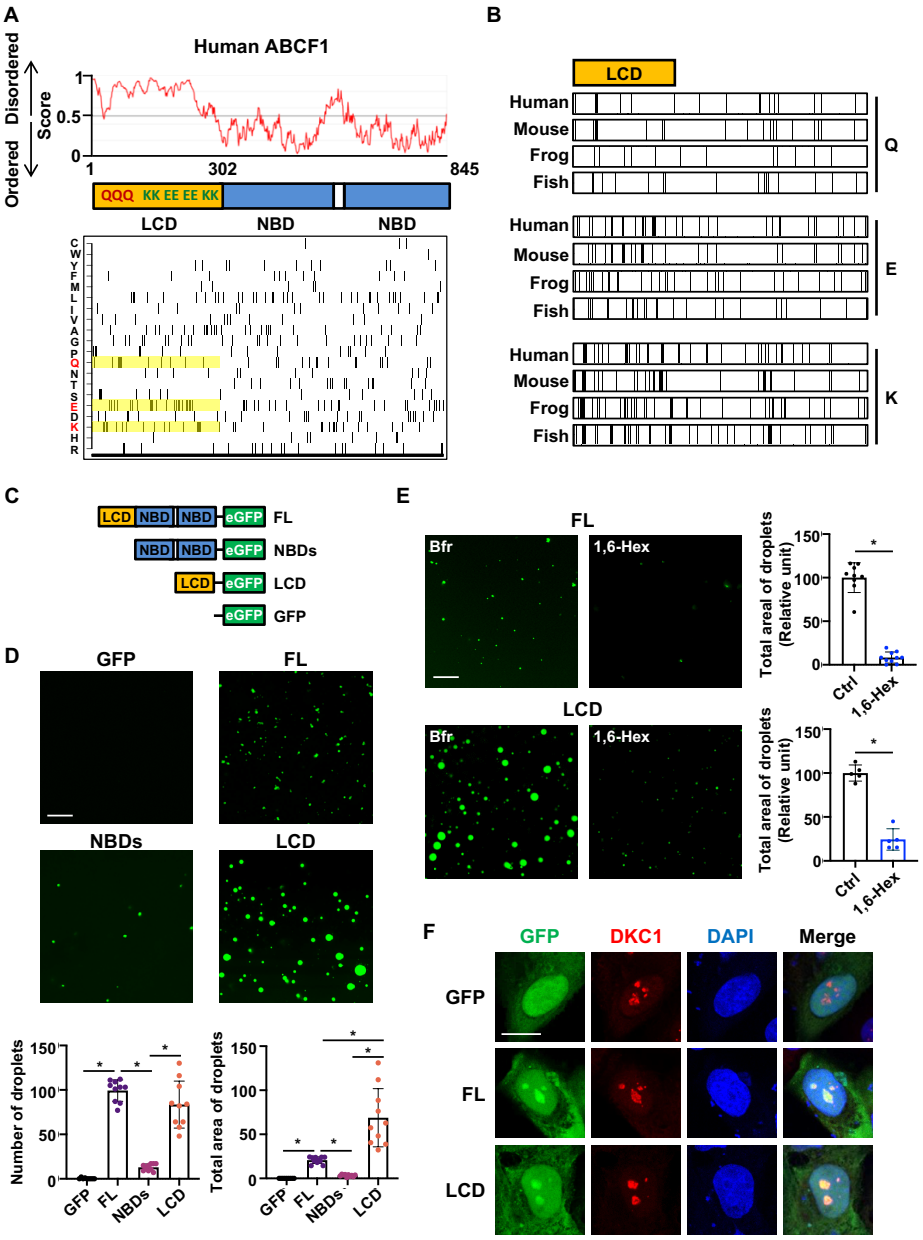

Fig. 3. The low complexity domain (LCD) of ABCF1 phase separates in vitro. a, Top: Intrinsically disordered protein prediction tool, IUPred2 is used to identify unstructured regions in human ABCF1. X-axis indicates position of the amino acids and Y-axis shows the probability of disordered sequences. Regions that are above the value of 0.5 are predicted to be unstructured and those below 0.5 are more likely structured. The schematic diagram denotes protein domains of ABCF1: intrinsically disordered low complexity domain (LCD, yellow, amino acids 1-302); two nucleotide-binding domains (NBDs, blue). Similar results are also obtained using a different prediction program DISOPRED3 (data not shown) (Jones and Cozzetto, 2015). Bottom: Amino acid composition of human ABCF1. Each of the 20 amino acids is counted and marked as a black bar at that position in ABCF1. One-letter abbreviations for amino acids are indicated (Left): C, Cys; W, Trp; Y, Tyr; F, Phe; M, Met; L, Leu; I, Ile; V, Val; A, Ala; G, Gly; P, Pro; Q, Gln; N, Asn; T, Thr; S, Ser; E, Glu; D, Asp; K, Lys; H, His; R, Arg. Q, E, and K residues in the LCD are highlighted. **b,** Comparison of distribution of glutamine (Q), glutamic acid (E), and lysine (K) residues in ABCF1 from indicated organisms, analyzed as described in a. LCD in human ABCF1 is indicated on top. **c,** Schematic diagrams of recombinant GFP fusion ABCF1 proteins used in in vitro droplet assays. **d,** Representative images of droplet formation with full-length (FL) ABCF1, NBDs, LCD, or GFP. Proteins are added to droplet formation buffer to a final concentration of 13 mM and NaCl concentration of 200 mM. ImageJ is used to quantify the average number and size of droplets in each image (n=10). **e,** Representative images of droplet formation with FL (n=10) and LCD (n=5) in the presence of buffer (Bfr) or 10% 1,6-hexanediol (1,6-Hex). Images are analyzed as described in d. **f,** Fluorescence images of GFP, FL ABCF1, and LCD transiently expressed in Hela cells. FL and LCD ABCF1 show colocalization with DKC1 in the nucleoli in Hela cells. DNA is stained with DAPI. Scale bars, 20 mm. Error bars present SEM. (*) $P < 0.05$, calculated by two-sided Student's t-test.
The following figure supplement is available for figure 3: Supplementary Figure S3.

Deletion of the LCD (i.e. NBDs-GFP) dramatically reduced the number of droplets, while LCD alone (amino acids 1-302) forms even bigger droplets that are sensitive to salt concentration (Fig. S3B). Furthermore, addition of the alcohol 1,6-hexanediol (1,6-Hex), which is known to disrupt LLPS, significantly suppressed droplet formation (Fig. 3E) (*53*). These results indicate that the LCD is essential for ABCF1 to phase separate into liquid-like bodies, but the structured NBDs may function to regulate its tendency.

In cells, ectopically expressed FL-GFP and LCD-GFP ABCF1 were found to localize both in the cytoplasm and nucleus, consistent with our data showing that endogenous ABCF1 proteins are present in both compartments (Fig. S1B). In the nucleus, we found that both FL-GFP and NBDs-GFP were enriched in the nucleolus, the site of ribosome biogenesis and a well-characterized phase-separated non-membrane-bound subnuclear compartment, in which the nucleolar protein DKC1 also primarily resides (Fig. 3F) (*54*). Taken together, these results demonstrate that ABCF1 can form homotypic LLPS droplets in vitro and is efficiently incorporated into phase-separated bodies in cells.

Transcriptional Coactivation Mediated by the LCD in ABCF1

We next asked whether the LCD important for LLPS is also required for the transcriptional activity of ABCF1. We generated recombinant full-length, wild-type (WT) ABCF1, the LCD fragment (1-302) as well as a series of N-terminal truncations of the LCD in full-length ABCF1. To address whether its catalytic activity is required for transcriptional coactivation, we also purified an ATPase-defective mutant ABCF1 protein wherein the two conserved ATP-binding lysine residues (K324 and K664) in the Walker A motif of both NBDs were substituted with methionine (2KM) (Fig. 4A and Fig. S4A) (*24*). When added to in vitro transcription assay, the full-length WT and 2KM mutant proteins were as active as endogenous ABCF1 purified from NT2 cells (Fig. 4B). However, deletion of polyQ alone (Δ115) or together with the K/E-rich domain (Δ248) led to a progressive loss of transcriptional activity, and the two NBDs by themselves (Δ302) were completely inactive. Surprisingly, the LCD fragment (1-302) also lacked transcriptional activity. These results provide several insights. First, ABCF1 is unambiguously shown to be the sole constituent of SCC-B. Second, ATP binding and hydrolysis by ABCF1 are dispensable for transcription, unlike for ABCF1's role in translation (*24*, *47*). Third, while the entire LCD is essential for full coactivator activity, the LCD by itself is not sufficient to activate transcription. These observations suggest that the NBDs also contribute to the full transcriptional activity of ABCF1.

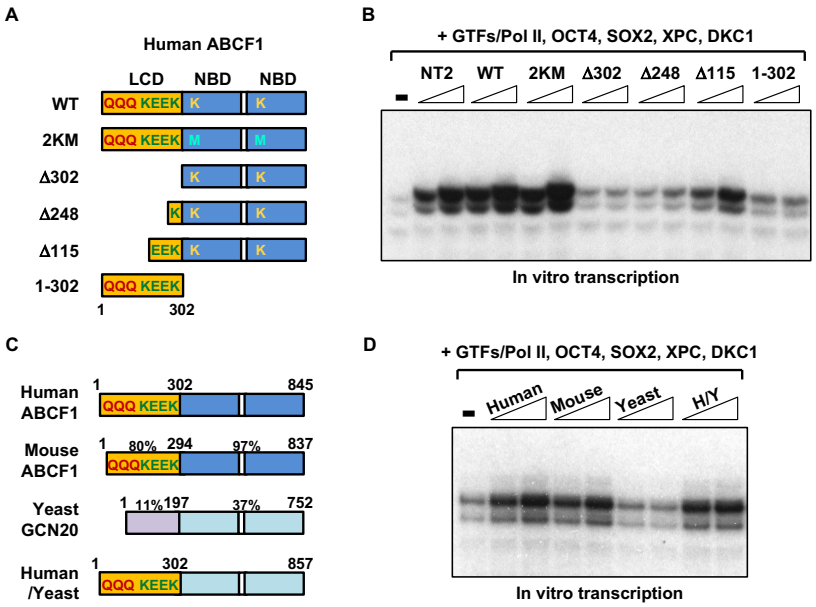

Fig. 4. Transcription coactivation by ABCF1 requires its LCD. a, Schematic diagram of full-length, wild-type (WT) ABCF1 protein depicting the N-terminal LCD (yellow) containing a polyglutamine (polyQ) tract and lysine (K)/glutamic acid (E)-rich regions. The two conserved lysine residues (K324, K664) in the Walker A motif of each of the two NBDs (blue) in ABCF1 are highlighted. Full-length WT ABCF1, ATP-binding defective lysine-to-methionine mutant (2KM), various truncated ABCF1 proteins lacking part (D248, D115) or all of the LCD (D302) as well as the LCD by itself (1-302) are purified from *E. coli*. **b,** Transcriptional activities of the various recombinant ABCF1 proteins shown in a are assayed (over a two-fold concentration range) together with recombinant XPC and DKC1 complexes in in vitro transcription as described in Fig. 1. **c,** Schematic representation of the human and mouse ABCF1, and yeast homolog GCN20. The percentage of sequence similarity among human, mouse, and yeast homolog is indicated. Domain-swapped hybrid protein between the human LCD and yeast NBDs (H/Y) is generated and purified from *E. coli*. **d,** Titration over a two-fold concentration range of human and mouse ABCF1, yeast GCN20, and human-yeast hybrid (H/Y) proteins are assayed in in vitro transcription reactions.
The following figure supplement is available for figure 4: Supplementary Figure S4.

The yeast homolog of ABCF1, GCN20, shares with its human and mouse counterpart a conserved NBD (*55*). However, the residues outside of NBD (amino acids 1-197) are highly divergent from mammalian ABCF1. This domain lacks the polyQ tract and is not K/E-rich (Fig. 4C and Fig. S4B). More importantly, this region is predicted to be structured, unlike the LCDs found in human and mouse ABCF1 proteins (Fig. S4C). As predicted based on a requirement for the LCD, we found that purified GCN20 exhibited no coactivator activity in the in vitro transcription assay (Fig. 4D and Fig. S4D). However, replacing the GCN20 N-terminal domain with the human ABCF1 LCD fully conferred transcriptional stimulatory capacity to the hybrid protein in vitro. These observations provide strong biochemical evidence that the mammalian-specific LCD confers coactivator activity.

Specificity of LCD-mediated interactions by ABCF1 Having identified the LCD as a "transactivation domain" in ABCF1, we next investigated the mechanisms by which the LCD potentiates stem cell-specific transcription. We have previously shown that transcriptional activation by OCT4 and SOX2 selectively requires SCCs but not the prototypical coactivator Mediator at least in vitro (*18*). To gain mechanistic insights into the specificity of the LCD in ABCF1, we tested whether the LCD can associate with its two co-dependent coactivators, XPC and DKC1, as well as Pol II by performing glutathione S-transferase (GST) pull-down assay. We incubated NT2 cell nuclear extract with immobilized LCD from human ABCF1 (1-302) and the transcriptionally inactive N-terminal domain from yeast GCN20 (1-197) as GST-fusion proteins. Western blot analysis revealed that only LCD/ABCF1 was able to bind XPC, DKC1, and Pol II (Fig. 5A and Fig. S5A). To further demonstrate the specificity of LCD/ABCF1 for the XPC and DKC1 complexes, we also tested the transactivation domain from human transcription factor SREBP1a, which harbors an LCD and is known to bind Mediator (*17*). No significant binding of XPC and DKC1 was observed, but the fusion protein was able to pull down Pol II, likely indirectly through a Mediator-Pol II interaction (*16*, *56*). Using transient transfection and co-immunoprecipitation experiments in 293T cells, we showed that ABCF1 interacted preferentially with SOX2 but not OCT4 (Fig. 5B). We further confirmed the interaction between ABCF1 and SOX2 in mouse ES cell extracts (Fig. 5C). Importantly, deletion of the LCD from ABCF1 (NBDs), which abrogates its transcriptional activity in vitro, completely abolished its ability to bind SOX2 (Fig. 5D). We next showed that the high-mobility group (HMG) box DNA-binding domain of SOX2 but not its activation domain (AD) is required for binding ABCF1 (Fig. 5E and Fig. S5B), consistent with studies indicating that interactions of SOX family transcription factors with cofactors are often mediated by their conserved HMG domain (*57*). There are 20 SOX genes in human and mice that are divided into different groups (A-H) (*58*, *59*). Even though these SOX proteins share an HMG box with more than 80% sequence identity, we found that ABCF1/LCD showed a strong preference for SOX2 and SOX1 from group B1, but not SOX11 (group C) or SOX10 (group E) (Fig. 5F). These results highlight the ability of ABCF1/LCD to discriminate small differences in protein sequences of the HMG boxes, despite high similarity in their overall structure. Taken together, we conclude that the transcriptional defect observed with LCD-truncated ABCF1 is likely due to its failure to establish multivalent interactions with SOX2, SCCs, and the Pol II machinery.

To provide in vivo evidence that the ABCF1 LCD is critical for OCT4/SOX2-dependent transcription, we performed knockdown-rescue experiments in mouse ES cells. Simultaneous knockdown of mouse ABCF1 and ectopic expression of full-length human ABCF1, but not LCD-truncated ABCF1 (NBDs) or control RFP, restored expression of several ABCF1-dependent pluripotency genes such as *Nanog*, *Fgf4*, and *Klf4* in mouse ES cells (Fig. 5G and Fig. S5C). These results demonstrate that the LCD in ABCF1 is required for pluripotency gene expression in vivo.

We next showed that ectopic expression of NANOG alone was sufficient to restore pluripotency gene expressions (*Fgf4* and *Klf4*) in ABCF1-deficient ES cells (Fig. 5G). This result indicates that *Nanog* is likely a critical downstream target of ABCF1, consistent with our in vitro transcription results. More importantly, unlike ABCF1, transcription factor NANOG has no other known roles beyond transcription in ES cells. The fact that NANOG can bypass the requirement of ABCF1 strongly suggests that the transcriptional defect in ABCF1 knockdown ES cells is unlikely due to disruption of other cellular processes that may be carried out by ABCF1. Rather, our data thus far suggest that ABCF1 directly controls pluripotency gene expression in vitro and in vivo and, by extension, stem cell maintenance. Therefore, we conclude that ABCF1 potentiates stem cell-specific transcriptional activation through LCD-dependent assembly of select activator and coactivators with the Pol II transcription machinery.

ABCF1 Links Gene Expression and DNA sensing in ES Cells DNA sensing by the innate immune system underpins many physiological responses to DNA (*60*), including immunity to DNA viruses and bacteria (*61–64*) and inflammatory responses to intracellular self-DNAs that arise from genome instability (*65–67*). However, ES cells lack the ability to mount a robust pro-inflammatory response to these intracellular DNAs (*68–72*). Whether there are alternative outputs in ES cells in response to aberrant intracellular DNAs remains poorly understood. Because ABCF1, an essential stem cell coactivator shown here, has also been reported to act as an important sensor for intracellular DNAs in the innate immune system in somatic cells (*26*), we next asked whether ABCF1 can also recognize these DNAs in ES cells to couple DNA sensing to pluripotency gene regulation.

We took two independent approaches to test whether ABCF1 can bind intracellular DNAs in ES cells.

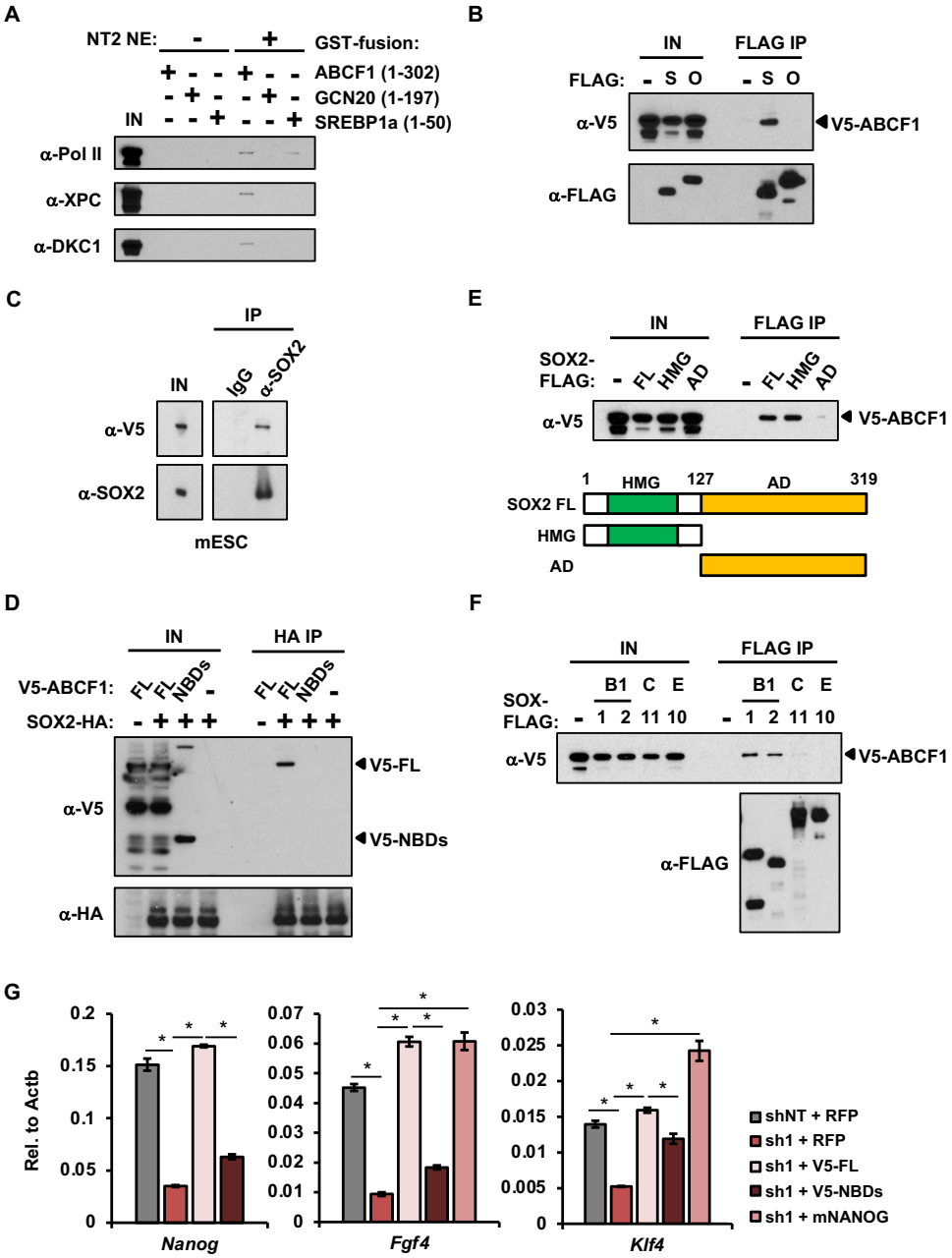

Fig. 5. Transcriptional coactivation by ABCF1 is mediated by LCD-dependent selective interactions. **a,** GST-fusion proteins containing the LCD of human ABCF1 (1-302), the N-terminal domain of yeast GCN20 (1-197), and the transactivation domain of human transcription factor SREBP1a (1-50) are incubated with buffer only (-) or NT2 nuclear extracts (NE, +). Input NE (IN) and bound proteins are analyzed by western blotting using antibodies against the largest subunit of RNA polymerase II (Pol II, RBP1), XPC, and DKC1. **b,** WCEs from 293T cells co-transfected with plasmid expressing V5-tagged ABCF1 together with either empty plasmid (-) or plasmids expressing FLAG-tagged SOX2 (S) or OCT4 (O) are immunoprecipitated with anti-FLAG antibody. Input WCEs (IN) and bound proteins are detected by western blotting. **c,** Input V5-ABCF1 KI mouse ES cell WCEs (IN) and IPs by IgG and anti-SOX2 antibodies are analyzed by western blotting using anti-V5 and anti-SOX2 antibodies. **d,** HA IPs from 293T cells overexpressing HA-tagged SOX2 (SOX2-HA) with V5-tagged full-length (FL) or LCD-truncated human ABCF1 (NBDs). SOX2-bound V5-ABCF1 proteins are detected by western blotting. **e,** WCEs from 293T cells transfected with plasmids expressing V5-ABCF1 alongside FLAG-tagged full-length (FL), high-mobility group (HMG) DNA-binding domain, or activation domain (AD) of SOX2 are immunoprecipitated with anti-FLAG antibody. Input WCEs (IN) and bound V5-ABCF1 proteins are detected by western blotting. Schematics of full-length (FL), HMG (green), and AD (yellow) of SOX2 are shown (bottom). **f,** FLAG IP from 293T cells overexpressing V5-ABCF1 with FLAG-tagged SOX1 and 2 (group B1), 11 (group C), or 10 (group E). Empty vector (-) is used as control. Input (IN) and bound proteins are detected by western blotting. **g,** ABCF1 knockdown-rescue assay. mRNAs from control mouse ES cells (shNT) overexpressing RFP, and ABCF1 knockdown ES cells (sh1) overexpressing RFP, V5-tagged full-length (FL), LCD-truncated human ABCF1 (NBDs), or mouse NANOG are analyzed for *Nanog*, *Fgf4*, and *Klf4* mRNA levels by qPCR, normalized to *Actb*. Error bars present SEM. $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.

The following figure supplement is available for figure 5: Supplementary Figure S5.

As a first approach, we incubated mouse ES cell extracts containing ectopically expressed full-length (FL) and LCD-truncated (NBDs) human ABCF1 proteins with a 5' biotinylated single-stranded (ss) or two different double-stranded (ds) DNA oligonucleotides containing sequences derived from *Listeria monocytogenes*. ds-matched (ds-M) contains a consensus sox-motif while ds-unmatched (ds-UM) does not. In neutrophils, SOX2 has been shown to act as a sequence-specific innate immune sensor for sox-motif-containing DNAs such as ds-M (*73*). Given that ABCF1 binds SOX2 in ES cells, we asked whether ABCF1 bound DNA, and if so, whether binding was direct or indirect through SOX2. Western blot analyses of protein-DNA complexes captured by streptavidin beads revealed that both ds-M and ds-UM efficiently pulled down FL ABCF1 but not NBDs, while neither binds ssDNA (Fig. 6A). Similar results were also obtained with endogenous ABCF1 (Fig. S6A). These results indicate that ABCF1 binds to short dsDNAs in an LCD-dependent, but sequence-independent, and therefore, SOX2-independent manner. Because the LCD important for LLPS is also required for DNA binding, we next investigated how DNA may influence ABCF1 droplet formation. Consistent with our DNA pull-down results, we found that fluorescently-labeled (Cy5) ds-UM but not ssDNAs were efficiently incorporated into full-length ABCF1 droplets. Strikingly, addition of dsDNAs dramatically increased both the number and size of the ABCF1 droplets (Fig. 6B). The increased propensity for LLPS in the presence of DNA suggested the formation of larger hetero-oligomers.

Indeed, we observed that many larger ABCF1 droplets colocalized with dsDNAs. These DNA-containing droplets were nonetheless sensitive to disruption by 1,6-Hex treatment, indicating that these bodies remained liquid-like. Interestingly, while droplets containing only the LCD are also capable of incorporating dsDNA, we did not observe any stimulatory effect on droplet formation compared to ssDNA (Fig. S6B). Taken together, these results demonstrate a strong preference for ABCF1 to form phase-separated droplets in the presence of dsDNAs and underscore a critical regulatory function of the NBDs in LLPS in response to dsDNA.

To evaluate the effect of binding of these dsDNAs by ABCF1 on pluripotency gene expression, we modeled pathogen infection in ES cells by transfecting 5' 6-carboxyfluorescein (6-FAM)-labeled ss, ds-M, or ds-UM oligonucleotides into mouse ES cells. Transfected, 6-FAM-positive cells were enriched by flow cytometry (Fig. S6C). Gene expression analyses by qPCR revealed that transfection of ds-M or ds-UM downregulated pluripotency genes and upregulated differentiation genes compared to ssDNAs (Fig. 6C and Fig. S6D). These data are consistent with our observation that ABCF1 selectively binds dsDNA but not ssDNA. Importantly, the fact that both ds-M and ds-UM were equally potent in eliciting changes in gene expressions in ES cells is distinct from what was observed in neutrophils, thus strongly suggesting a SOX2-independent mechanism.

For the second approach, we investigated whether ABCF1 can bind short DNAs accumulated in cells following DNA damage-induced stress. We treated V5-ABCF1 knock-in mouse ES cells with DNA damaging agent etoposide (ETO) (*66*, *67*). Intracellular self-DNA fragments were readily detected in the cell nucleus 6 hr post-treatment while ABCF1 and SOX2 protein levels remained unchanged (Fig. S6, E and F). We showed that immunoprecipitation of ABCF1 from ETO-treated ES cell extracts co-precipitated small DNA fragments that correspond to the size of a mononucleosome (Fig. 6D). Micronuclei, extra-nuclear bodies that contain mis-segregated and damaged chromosomal DNAs, were also evident in these ETO-treated cells (Fig. 6E). We found striking accumulation of ABCF1 with these micronuclei. Together with our in vitro data showing a direct binding of ABCF1 to naked dsDNAs, these data provide evidence that ABCF1 can also recognize short endogenous and extra-nuclear DNAs in ES cells upon DNA damage. In an effort to understand the consequence of endogenous DNA binding on the coactivator function of ABCF1, we showed that, under the same DNA damaging condition, interaction of ABCF1 with SOX2 was disrupted in ES cells (Fig. 6F). These results suggest that interaction of ABCF1 LCD with self-DNA fragments accumulated in cells after DNA damage may compete with SOX2. Accordingly, we observed a loss of enrichment of ABCF1 at its target genes such as *Oct4*, *Sox2*, and *Nanog* in ETO-treated ES cells by ChIP-qPCR, while histone H3 levels at these genomic loci remained unchanged (Fig. 6G and Fig. S6, G to I). Decrease in ABCF1 occupancy at promoters was accompanied by downregulation of pluripotency-associated genes and upregulation of lineage-specific genes (Fig. S6, J and K), consistent with loss of pluripotency and spontaneous differentiation in ETO-treated ES cells.

To provide further evidence that the transcription factor-coactivator hubs assembled at the SOX2/OCT4 enhancers are mediated by LCD/ABCF1-driven multivalent interactions reminiscent of droplet behavior observed in vitro, we showed that treatment of ES cells with 1,6-Hex significantly reduced the enrichment of ABCF1 and Pol II at the *Oct4* and *Nanog* enhancers, while H3K4me3 level remained unchanged (Fig. 6H and Fig. S6L). Our results suggest that the strong preference of ABCF1 to form condensates with intracellular dsDNAs may come at the cost of transcription, indicating that ABCF1 may be limiting in ES cells. If this were the case, increasing cellular concentration of ABCF1 would enhance DNA damage tolerance in ES cells, by increasing the pool of ABCF1 proteins available for transcriptional activation. Indeed, we found that, compared to RFP, ectopic expression of full-length (FL) ABCF1 significantly enhanced the self-renewal capacity of ETO-treated ES cells as shown by an increase in number of AP-positive colonies formed in limiting dilution assays (Fig. 6I and Fig. S6, M and N).

By contrast, deletion of the LCD (i.e., NBDs) compromised

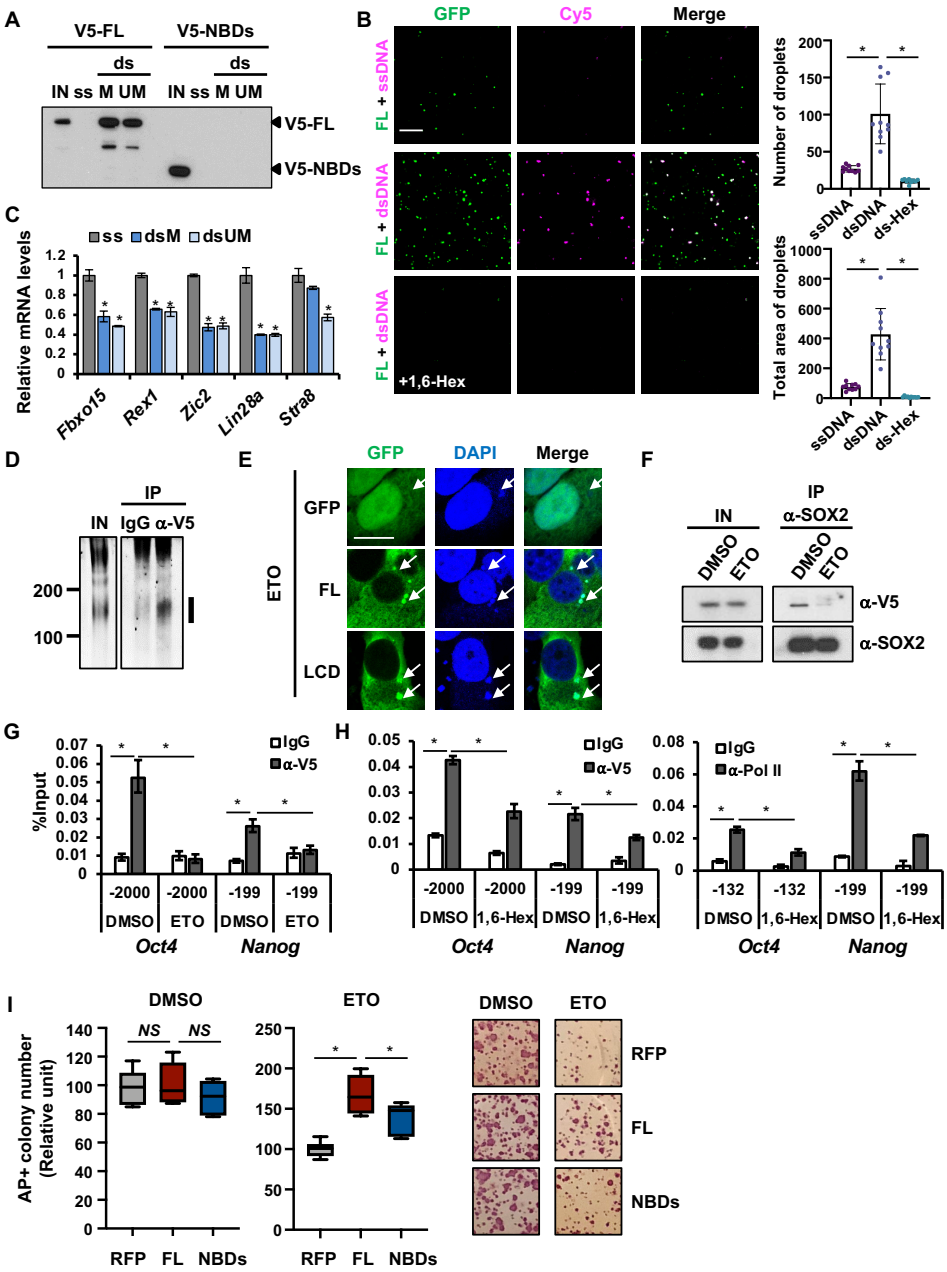

Fig. 6. Intracellular DNAs modulate phase separation of ABCF1 and pluripotency gene expressions. **a,** WCEs from D3 mouse ES cells stably expressing V5-tagged full-length (FL) ABCF1 or NBDs are incubated with three different 5' biotinylated 98mer oligonucleotides: single-stranded (ss), double-stranded (ds) with SOX2-binding motif (Matched, ds-M), or ds lacking the sox-motif (Unmatched, ds-UM). These DNA sequences are derived from *Listeria monocytogenes* genome. Input WCEs (IN) and streptavidin-beads captured, DNA-bound ABCF1 proteins are analyzed by western blotting. **b,** Representative images of droplet formation with FL ABCF1 (9 mM) in the presence of 12.5 nM of Cy5-labeled (magenta) ss or ds-UM DNAs. Fluorescence channels are indicated on top. Images are merged to show degree of colocalization. Number and size of droplets are analyzed as described in Fig. 3d. **c,** Mouse ES cells transfected with 5' 6-carboxyfluorescein (6-FAM) labeled ss, ds-M, or ds-UM on the sense strand are enriched by FACS. Expression levels of pluripotency genes are analyzed by qPCR, normalized to *Actb*. **d,** WCEs prepared from ETO-treated (20 mM) V5-ABCF1 KI mouse ES cells are incubated with IgGs or anti-V5 antibodies. Co-purified nucleic acids are treated with RNase A, separated on urea-PAGE, and stained with SYBR Gold. Vertical bar denotes DNAs specifically bound by ABCF1. **e,** Fluorescence images demonstrate colocalization of GFP-fusion FL ABCF1 and LCD proteins with micronuclei (white arrows) in ETO-treated (10 mM) Hela cells. **f,** DNA damage disrupts ABCF1-SOX2 interaction. Input (IN) and SOX2 IPs from WCEs of DMSO or ETO-treated (20 mM) V5-ABCF1 KI mouse ES cells are analyzed by western blotting. **g,** MNase-ChIP of ABCF1 in DMSO and ETO-treated (80 mM) V5-ABCF1 KI mouse ES cells. Enrichment of ABCF1 on OCT4/SOX2-targeted regions of *Oct4* and *Nanog* gene promoters is analyzed by qPCR as in Fig. 2. **h,** MNase-ChIP of ABCF1 in DMSO and 1,6-Hex (1.5%)-treated V5-ABCF1 KI mouse ES cells. Enrichment of ABCF1 (left) and Pol II (right) on OCT4/SOX2-targeted regions of *Oct4* and *Nanog* gene promoters is analyzed by qPCR as in g. **i,** Colony formation assays in control and ABCF1 gain-of-function D3 mouse ES cells. 1000 mouse ES cells stably expressing RFP, V5-FL ABCF1, or V5-NBDs are plated on 24-well plates, treated with DMSO (left) or ETO (1 mM, right) for 1 hr, and let recover for 6 days before staining for AP activity. AP-positive colonies are counted and indicated as a relative unit. Error bars represent SEM of three independent experiments. Representative images of each AP-stained cells are indicated (right). $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.

The following figure supplement is available for figure 6: Supplementary Figure S6.

the ability of ABCF1 to overcome ETO-induced stresses. Importantly, we did not observe significant differences between RFP and ABCF1 gain-of-function ES cells in the absence of DNA damage, thus suggesting a specific protective effect of ABCF1 overexpression on self-renewal capacity when ES cells encounter significant genome instability. Our studies reveal a new and important safeguard mechanism, whereby the pluripotency gene network can be coupled with intracellular DNA sensing through the LCD of ABCF1, likely by modulating the availability of ABCF1 for transcriptional coactivation.

DISCUSSION

Role of LCD in Mediating Stem Cell-specific Transcriptional Activation

In this study, we identified ABCF1 as a critical transcriptional coactivator for stem cell maintenance and showed that the LCD of ABCF1 is indispensable for both physical interaction with SOX2, XPC, DKC1, and Pol II, and functional reconstitution of ABCF1 coactivator activity both in vitro and in vivo. The flexible nature of LCDs is thought to facilitate the dynamic interaction with multiple protein partners, by virtue of their ability to rapidly adopt an ensemble of conformations (*74*, *75*). It is worth emphasizing that LCDs are not simply unstructured sequences that bind promiscuously to any proteins; instead, they can be selective for binding partners (*14*, *76*). In this regard, the conformationally flexible XPC protein also contains several highly disordered regions that we found, however, to be dispensable for transcriptional activation (*18*, *20*, *21*). These observations reveal the unique ability of ABCF1 LCD to integrate multiple lines of information encoded by SOX2, SCCs, and the Pol II machinery, likely by forming a hub of these factors at target gene promoters through selective multivalent interactions (Fig. 7) (*14*). Consistent with this hypothesis, we find that these ABCF1-mediated interactions at pluripotency gene loci are sensitive to 1,6-Hex, a chemical that interferes with LLPS of many LCDs (*53*). Our findings here thus provide a clear example of an LCD that can impart activator preference and target gene specificity to a transcriptional coactivator, and activate cell-specific transcription by a protein:protein driven local high concentration mechanism.

An Alternative Mechanism for ES Cells to Sense Infection and Genome Instability

In most somatic cell types, the innate immune pathway plays a critical role in managing the DNA damage response (*65*, *77–79*). DNA sensors recognize damage-induced accumulation of endogenous DNA fragments and activate a pro-inflammatory response to promote the apoptosis and clearance of the affected somatic cells (*80*). Yet, this canonical innate immune pathway is severely suppressed in ES cells, in part by pluripotency-associated transcription factors including OCT4 and SOX2 (*28*, *29*). A suppressed innate immune system also suggests that ES cells may lack an effective defense system against pathogen infections. While it is generally thought that pluripotent stem cells are less susceptible to infection because they are physically protected within a blastocyst, vertical transmission of pathogens to these cells can still occur, for example, through infection of the surrounding trophoblast cells, often with devastating consequences to the embryo (*81*).

Studies have shown that ES cells may be able to combat infections to some extent by constitutively expressing a subset of interferon-stimulated genes (ISGs) that function to suppress certain viral infections. However, this handful of ISGs are likely effective against only a subset of pathogens (*82*). High proliferation rate of mouse ES cells and a short gap (G1) phase of the cell cycle impose replication stress that, in turn, can lead to genome instability and DNA damage (*83–85*). Even though ES cells lack a robust innate immune system, they are nonetheless hypersensitive to DNA

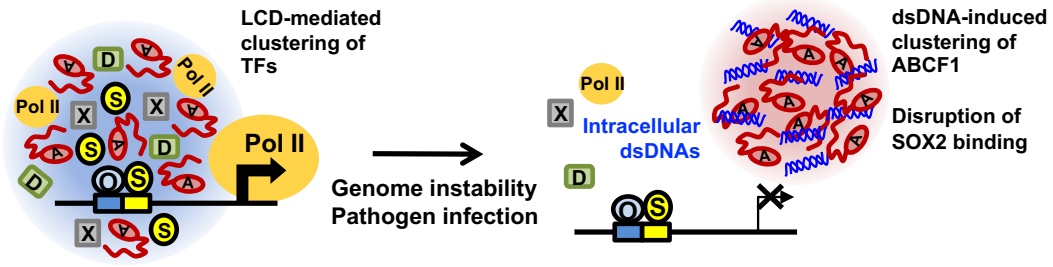

Fig. 7. Illustration of a proposed model whereby the LCD of ABCF1 couples pluripotency gene transcription with intracellular DNA sensing in ES cells. (Left) ABCF1 LCD (red wavy line) promotes specific clustering and formation of a hub comprising of SOX2 (S), XPC (X), DKC1 (D), and Pol II molecules at target gene promoter to stimulate transcription, presumably by increasing local concentration of these factors. (Right) ABCF1 proteins available for transcription are diverted to bind intracellular dsDNAs generated from genome instability or pathogen infection, due to increase propensity of ABCF1 to form condensates with dsDNAs. Decrease in ABCF1 at gene promoters destabilizes the multivalent interactions between SOX2, XPC, DKC1, and Pol II. This leads to disruption of the protein hub and decrease in gene transcription by Pol II. Downregulation of pluripotency-associated genes promotes differentiation of compromised ES cells and their elimination from the self-renewing population.

damage and readily undergo differentiation and apoptosis (*85–87*). This is not due to a diminished DNA repair capacity because many repair-related factors are in fact elevated in ES cells (*88–90*). Taken together, these observations suggest that alternative mechanisms likely exist in order to sensitize ES cells to these cellular insults, thereby protecting the pristine state of ES cells from infections and genome instability.

Biological Implications of ABCF1/LCD-mediated Coupling of DNA Sensing and Pluripotency Gene Expression

Because LCD-mediated interactions are highly dynamic (*14*, *74*, *91–93*), targeting these transient and multivalent interactions between transcription factors and SCCs could provide an effective means of modulating the pluripotency gene network in a rapid manner in response to changing cellular environment. Indeed, it has been shown that molecular crowding by LCD is particularly sensitive to changes in concentration, where a less than two-fold decrease can be sufficient to cause an LCD-mediated phase-separated body to be disrupted (*94*, *95*). Therefore, we suggest that a decrease in ABCF1 protein concentration, such as during stem cell differentiation (Fig. 1, E and F), could lead to rapid dissolution of the transcriptional apparatus, pluripotency gene expression, and the pluripotent state as a result.

The coupling between transcription and DNA sensing by ABCF1 appears to be critically dependent on its LCD. Paradoxically, we find that the LCD by itself is inactive in transcription (Fig. 4B). Likewise, unlike full-length ABCF1, binding of dsDNA to the LCD fragment fails to promote LLPS (Fig. S6B). These observations point to a potential critical regulatory function of the NBDs in both processes. We show that in the absence of the NBDs, the LCD becomes "hyperactive" in forming significantly bigger LLPS droplets (up to 10 micrometers in diameter) in vitro (Fig. 3D). We speculate that the higher propensity of the LCD to form homotypic droplets could interfere with its ability to interact with transcription factors, coactivators, and the transcription machinery. It remains possible that the NBDs provide additional contact points critical for transactivation. With regard to DNA sensing, we show that the NBDs not only curtail LLPS of the LCD, but also endow the LCD the ability to efficiently form phase-separated droplets in the presence of dsDNAs (Fig. 6, A and B, and Fig. S6B). We posit that this dynamic response to dsDNA is likely advantageous if not essential for ABCF1's ability to rapidly respond to accumulation of aberrant DNAs in cells upon infection or DNA damage. By favoring the formation of an ABCF1-dsDNA complex, likely at the expense of SOX2 binding, the transcription factor hub anchored by ABCF1 LCD becomes destabilized. This leads to disruption of pluripotency gene expression and rapid exit of pluripotency of the compromised ES cells (Fig. 7).

Under optimal growth conditions, there are likely multiple mechanisms to return ES cells experiencing low level of DNA damage or infections to homeostasis, such as DNA repair or constitutive ISG expressions. We surmise that when these cellular insults turn catastrophic, ABCF1 serves as a critical checkpoint for self-renewal to ensure that compromised ES cells do not propagate, by enforcing their exit of pluripotency and into differentiation. This safeguard mechanism allows ES cells to robustly stabilize the pluripotency gene network while remaining responsive to pathogen infections and genome instability. Based on our findings, we propose that ABCF1 represents an important regulatory nexus, wherein the constant tug-of-war between transcriptional activation and intracellular DNA sensing by ABCF1 could drive an ES cell to self-renew or commit to differentiation. This switching of cell fates may depend on whether intracellular DNA rises above a certain threshold level that irreversibly tilts the balance toward rapid exit of pluripotency.

MATERIALS AND METHODS

DNA constructs and antibodies

cDNAs for human and mouse ABCF1 were obtained from cDNA libraries generated from total RNAs isolated from human embryonic carcinoma cell line NTERA-2 (NT2) and mouse embryonic stem (ES) cell line D3. Intronless human SOX1 and SOX11 cDNAs were cloned from human genomic DNA. Human SOX10 was subcloned from a full-length cDNA clone purchased from NIH Mammalian Gene Collection (MGC). Yeast GCN20 cDNA was amplified from purified S. cerevisiae genomic DNA. For expression of GCN20, full-length human and mouse ABCF1, as well as various truncations of human ABCF1 in *E. coli*, corresponding cDNAs were cloned into a modified pMtac-His6 vector containing a His6 tag at the N-terminus and a FLAG tag at the C-terminus. Mutations of the two Walker-A motifs in human ABCF1 (K324M, K664M) were created by using Quikchange II Site Directed Mutagenesis Kit (Agilent). Human ABCF1 and yeast GCN20 domain fusion cDNAs were created by PCR-mediated ligation and cloned into modified pMtac-His6 vector. For in vitro droplet assays, C-terminal eGFP fusion proteins containing full-length, LCD (1-302), and LCD-truncated human ABCF1 (NBDs) were inserted into pMtac vector, in frame with a flexible linker (GGSGGGSG) followed by eGFP. For in vivo expression of full-length ABCF1-eGFP and LCD-eGFP fusion proteins, corresponding human ABCF1 cDNAs were inserted into pFLAG-CMV5a, in frame with a flexible linker (GGSGGGSG) followed by eGFP. GST fusion proteins containing the human ABCF1 LCD (1-302), yeast GCN20 (aa 1-197), and transcription factor SREBP1a (aa 1-50) were generated by inserting the corresponding cDNA fragments into pGEX4T-3 vector (Sigma). V5-tagged full-length and NBDs of human ABCF1, and untagged mouse *Nanog* were cloned into lentiviral overexpression vector pHAGE-IRES-Neo (*18*). cDNAs for human OCT4 and SOX2 were cloned into the pFLAG-CMV5a mammalian expression vector (Sigma).

Polyclonal antibodies against ABCF1 (13950-1-AP) were purchased from ProteinTech, XPC (A301-122A) and mouse Nanog (A300-397A) from Bethyl Laboratories, DKC1 (H-300), OCT4 (N-19), and Pol II (N-20) from Santa Cruz Biotechnology, SOX2 (AB5603) from EMD Millipore, V5 ChIP grade (ab15828) from Abcam, Histone H3 (ChIP formulated, #2650) from Cell Signaling Technology (CST). Monoclonal antibodies against ACTB (66009-1) from ProteinTech, HA-tag (C29F4) and Rbp1 (Pol II) NTD (D8L4Y) from CST, FLAG-tag (M2), α-Tubulin (T5168) and trimethyl-Histone H3 (Lys4) (H3K4me3) (05-745R-S) from Sigma, RFP (600-401-379) from Rockland, and V5-tag (R96025) from Life Technologies.

Cell culture

NT2 cell line was obtained from ATCC. NT2, 293T, and HeLa cells were cultured in DMEM high glucose with GlutaMAX (Gibco) supplemented with 10% fetal bovine serum (FBS). Large scale culture of NT2 cells were described (*18*). Feeder-independent human ES cell line H9 was purchased from WiCell Research Institute. H9 ES cells were cultured in mTeSR1 (STEMCELL Technologies) with normocin (50 μg/ml; Invivogen) on Matrigel-coated tissue culture plates (Corning). D3 mouse ES cell line was purchased from ATCC and adapted to feeder-free condition as described (*18*). Medium was changed daily. Cell cultures were passaged by StemPro accutase (Gibco) for human ES cells and trypsin for mouse ES cells.

For chromatin immunoprecipitation (ChIP) experiments, mouse ES cells were adapted to 2i/LIF culture condition. Mouse ES cells were passaged in 2i/LIF, serum-free medium composed of 1:1 mix of Advanced-DMEM /F12 and Neurobasal medium (Gibco) supplemented with N2 (Gibco), B27 (Gibco), L-glutamine (GlutaMAX, Gibco), beta-mercaptoethanol (0.1 mM; Sigma), BSA (50 μg/ml; Sigma), PD0325901 (1 μM; Sigma), CHIR99021 (3 μM; EMD Millipore), and LIF (10^2 U/ml) for at least four passages before the cells were used for ChIP and RT-qPCR analyses.

Differentiation of H9 ES cells was induced by exchanging human ES cell growth medium with DMEM/F12 (Gibco) containing 2 mM L-glutamine, 12.5% FBS, and normocin (50 μg/ml) for up to 14 days. V5-ABCF1 knock-in D3 mouse ES cells were induced to differentiate by maintaining cells in regular medium containing 5 μM all-trans retinoic acid (Sigma) for 7 days.

In vitro transcription assay

In vitro transcription reactions, the human *NANOG* transcription template, purification of activators OCT4 and SOX2, general transcription factors, RNA polymerase II, and recombinant XPC complex were described (*18*). Recombinant XPC complex purified from Sf9 and DKC1 complex reconstituted and purified from Sf9 cells or *E. coli* (*22*) were supplemented in the in vitro transcription reactions to assay for SCC-B activity.

Purification of SCC-B/ABCF1

All steps were performed at 4°C. Nuclear extracts were prepared from 400 L of NT2 cells. Partially purified P11-phosphocellulose 1 M KCl and Ni-NTA flowthrough (Ni-FT) fractions were prepared as described (*18*). The Ni-FT fraction was dialyzed against buffer D (20 mM HEPES, pH 7.9, 2 mM $MgCl_2$) at 0.2 M KCl with 0.0025% NP-40 and 10% glycerol (all buffers from then on contained 0.0025% NP-40 and 10% glycerol unless otherwise stated). This Ni-FT fraction was applied to a Poros 20 HQ column (Applied Biosystems), subjected to a 4 column volume (CV) linear gradient from 0.2 M to 0.4 M KCl (Q0.3), washed at 0.52 M KCl, and developed with a 13 CV linear gradient from 0.52 M to 1.0 M KCl. Transcriptionally active Q0.3 fraction (0.32–0.4 M) were pooled and applied directly to hydroxyapatite (HAP) type II ceramic resin (Bio-Rad), washed first at 0.38 M, then lowered to 0.1 M KCl in 3 CV. HAP column buffer was then exchanged and washed extensively with buffer D at 0.03 M KPi, pH 6.8 without KCl and NP-40. The HAP column was subjected to a 20 CV linear gradient from 0.03 M to 0.6 M KPi. Active HAP fractions eluting from 0.2–0.3 M KPi were pooled and separated on a Superose 6 XK 16/70 gel filtration column (130 ml, GE Healthcare) equilibrated with buffer D + 0.1 mM EDTA at 0.15 M KCl. Active Superose 6 fractions with an apparent molecular mass of 400 - 600 kDa were pooled and supplemented with 0.25 mg/ml insulin (Roche). Pooled fractions were applied to a Poros 20 HE column (Applied Biosystems) equilibrated in buffer D + 0.1 mM EDTA at 0.15 M KCl, subjected to a 34 CV linear gradient from 0.15 M to 1 M KCl. SCC-B containing HE fractions eluted from 0.35 - 0.43 M KCl. Active HE fractions eluting from 0.35-0.43 M KCl were supplemented with 0.3 mg/ml insulin and dialyzed to 0.15 M KCl in buffer D + 0.1 mM EDTA + 0.01% NP-40. The dialyzed HE fraction was applied to a Mono S PC 1.6/5 column (GE Healthcare), washed and developed with a 20 column volume linear gradient from 0.15 M to 0.65 M KCl. Transcriptionally active SCC-B/ABCF1 fractions eluted from 0.29 – 0.31 M KCl.

Mass spectrometry analysis

Peak Mono S fractions were pooled, concentrated using a Spin-X centrifugal concentrator, separated by SDS-PAGE, stained with PageBlue (Thermo Fisher), protein bands excised, digested with trypsin, and extracted. Peptide pools from each gel slice were analyzed by matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF MS; Bruker Reflex III). Selected mass values were used to search protein databases linked to PROWL (Rockefeller University) using ProFound and protein databases linked to ExPASy (Swiss Institute of Bioinformatics, Geneva) using PeptIdent.

Generation of endogenously V5-tagged ABCF1 knock-in mouse ES cell line

sgRNA targeting genomic region immediately downstream of the ATG translation start codon of ABCF1 was cloned into LentiCRISPRv2 vector (Addgene). A single-stranded (ss) donor oligonucleotide containing a V5-tag followed by a

flexible linker GSSG sequence in frame with the second amino acid of ABCF1, which is flanked by left and right homology arms of about 70 bp, was synthesized (Integrated DNA Technologies (IDT)). Both the LentiCRISPRv2-sgRNA plasmid and the ss donor oligonucleotides were transfected into D3 mouse ES cell line using lipofectamine 3000 (Invitrogen). Transfected cells were selected with puromycin (1.5 μg/ml) for 3 days. Cells were then expanded in the absence of puromycin to avoid integration of the LentiCRISPRv2-sgRNA plasmid into the genome. Single clones were plated into 96-well plates. Positive clones were identified by PCR and confirmed by sequencing and western blotting. Clones selected for further analysis were confirmed to be puromycin-sensitive and express similar levels of key pluripotency genes such as *Nanog*, *Oct4*, and *Sox2*. See Table S1 for sgRNA and ss donor oligo sequences.

Purification of recombinant proteins pMtac expression plasmids were transformed into BL21-Codon Plus RIPL competent cells (Agilent). Expression of His6-tagged proteins were induced at 30°C with 0.5 mM IPTG for 4 hr. Cell pellets were lysed in high salt lysis buffer HSLB (50 mM Tris–HCl, pH 7.9, 0.5 M NaCl, 0.6% Triton X-100, 0.05% NP-40, 10% glycerol) with imidazole (10 mM) and lysozyme (0.5 mg/ml). Sonicated lysates were cleared by ultracentrifugation and incubated with Ni-NTA resin for 1.5 hr at 4°C. Bound proteins were washed extensively with HSLB with 20 mM imidazole, equilibrated with 0.3 M NaCl HGN (25 mM HEPES, pH 7.9, 10% glycerol, 0.05% NP-40) with 20 mM imidazole, and eluted with 0.25 M imidazole in 0.3 M NaCl HGN. Peak fractions were pooled and incubated with anti-FLAG M2 agarose (Sigma) for 3 hr at 4°C. Immunoprecipitated proteins were washed extensively with 0.7 M NaCl HEMG (25 mM HEPES, pH 7.9, 0.1 mM EDTA, 12.5 mM $MgCl_2$,10% glycerol) with 0.1% NP-40 and equilibrated with 0.25 M NaCl HEMG with 0.1% NP-40. For recombinant proteins used in in vitro transcription assays, bound proteins were eluted in the same buffer containing 0.4 mg/ml FLAG peptides. For in vitro LLPS droplet assays, bound proteins were extensively washed with 0.25 M NaCl HEMG without NP-40, before the proteins were eluted with FLAG peptides in the same buffer without detergent. Eluted proteins were filtered through a 0.22 μm filter. Proteins used in droplet assays were concentrated by using a Spin-X UF concentrator (10K MWCO) (Corning). Purified proteins are snap frozen in liquid nitrogen, and stored at -80°C.

pGEX4T-3 expression plasmids were transformed into BL21-Codon Plus RIPL competent cells (Agilent). Protein expression was induced at 20°C for 3.5 hr. Bacterial pellets were resuspended in 1X PBS containing 3 M DTT, 0.5 mM PMSF, and complete protease inhibitor cocktail (Sigma) and snap frozen in liquid nitrogen. Sonicated lysates were cleared by ultracentrifugation. Cleared lysates were then incubated with Glutathione Sepharose 4 Fast Flow (GE Healthcare) to immobilize GST-fusion proteins.

In vitro droplet assay
Recombinant eGFP-fusion proteins were diluted into indicated protein and NaCl concentration with 10% PEG-8000 (Sigma) with or without 10% 1,6-hexanediol and incubated for 30 min at room temperature (RT). The solutions were loaded into an in-house chamber and imaged with a confocal microscope with a 63X objective.

Immunofluorescence microscopy
Cells were plated on cover glass and transfected with plasmids expressing indicated eGFP-fusion proteins. After 48 hr, cells were fixed in 4% paraformaldehyde for 10 min at RT and permeabilized with 0.1% triton X-100 in PBS for 10 min at RT. After three washes in PBS for 5 min each, cells were blocked with 5% bovine serum albumin (BSA) and 5% goat serum for 1hr at RT and incubated with primary antibodies (anti-DKC1, Santa Cruz, 1:100) overnight at 4°C. Following three washes, cells were incubated with secondary antibodies (Goat anti-Rabbit IgG Alexa Fluor-594, Life Technologies, 1:500) for 1 hr at RT. Cover glasses were mounted on slides with mountant with DAPI (Life Technologies) and imaged with a confocal microscope.

GST pull-down assay
Nuclear extracts from NT2 cells were prepared as described (*96*). Proteins were precipitated with ammonium sulfate (55% saturation) and resuspended in buffer D (20 mM HEPES, pH 7.9, 2 mM $MgCl_2$, 20% glycerol) containing 20 mM KCl and 0.01% NP-40 to about a fourth of the starting volume of nuclear extracts. Soluble extracts were dialyzed against buffer D containing 0.2 M KCl, 0.2 mM EDTA, 10% glycerol, and 0.01% NP-40 (buffer HEGN). Dialyzed nuclear extracts were cleared by centrifugation.

Bacterial lysates containing GST-fusion proteins were immobilized onto Glutathione Sepharose 4 Fast Flow in 1X PBS and 30 μg/ml BSA. Bound proteins were washed extensively with STE buffer (20 mM Tris pH 8.0, 1 mM EDTA, 1% TX-100, 0.5% NP-40, 10% glycerol) containing 1M NaCl. Approximately 20 μg of GST-ABCF1 (1-302), equimolar of GST-GCN20 (1-197), and GST-SREBP1a (1-50) were immobilized onto the sepharose beads. Bound GST-fusion proteins were equilibrated to 0.2 M NaCl HEGN.

Approximately 3 mg of NT2 nuclear extracts were incubated with the immobilized GST-fusion proteins overnight at 4°C. Bound proteins were washed extensively with 0.3 M NaCl HEGN, then with 0.1M NaCl HEGN. Proteins bound to GST-fusion proteins were eluted by incubating the sepharose slurry twice with 0.1 M NaCl HEGN containing 0.2% sarkosyl.

Coimmunoprecipitation assay
pHAGE-V5-ABCF1 (full-length and NBDs), pFLAG-CMV5a-OCT4, and pFLAG-CMV5a-SOX1, 2 (full-length, HMG, and activation domain (AD)), 10, and 11 expression plasmids were co-transfected in various combinations into 293T cells

using Lipofectamine 2000 (Invitrogen). Transfected cells on 10 cm dishes were lysed with 1 ml of lysis buffer (0.25 M NaCl, 20 mM HEPES, pH 7.9, 2 mM $MgCl_2$, 0.5% NP-40, 10% glycerol) 40 hr post transfection. Cell lysates were collected and homogenized by passing through a 25-gauge needle 5 times. Lysates were cleared by centrifugation at 15k rpm for 25 min at 4°C. Cleared lysates were then incubated with anti-FLAG M2 agarose pre-blocked with 5 mg/ml BSA for 3-4 hr at 4°C. Bound proteins were washed extensively with lysis buffer followed by FLAG peptide elution (0.4 mg/ml) in buffer HEMG containing 0.3 M NaCl and 0.1% NP-40.

To detect interaction between endogenous ABCF1 and SOX2 under normal and DNA damage conditions, V5-ABCF1 knock-in D3 mouse ES cells were treated with DMSO or etoposide (ETO, 20 μM; Sigma) for 6 hr. Whole cell extracts were prepared using lysis buffer (20mM HEPES, pH 7.9, 0.12 M NaCl, 1 mM EDTA, 1% NP-40, 1 mM DTT, 1 mM Benzamidine, 0.5 mM PMSF, complete protease inhibitors (Sigma)). Cell lysates were incubated on ice for 10 min and lysates were cleared by centrifugation. Cleared lysates were incubated with anti-SOX2 antibodies at 4°C overnight. Lysates were incubated with Protein A Sepharose (GE Healthcare) for 1-2 hr at 4°C and bound proteins were washed extensively with lysis buffer and eluted with SDS sample buffer with boiling.

Micrococcal Nuclease (MNase)-chromatin immunoprecipitation (ChIP)

V5-ABCF1 knock-in D3 mouse ES cells were first adapted to 2i/LIF conditions. Cells were treated with DMSO/ETO (80 μM, 8.5 hr) or DMSO/1,6-hexanediol (1.5%, 30 min) before proceeding with ChIP. Cells were crosslinked with EGS for 20 min and then with formaldehyde for 5 min. Nuclei were prepared as described (*22*). Nuclei were equilibrated with MNase buffer (10 mM Tris-HCl, pH 7.5, 15 mM NaCl, 60 mM KCl, 1mM $CaCl_2$, 1 mM PMSF, complete protease inhibitors (Sigma)) before nuclei were digested with MNase in the same buffer as described (*97*). Briefly, nuclei from 4 x 10^7 of cells were resuspended in 0.5 ml of MNase digestion buffer and digested with 2400 U of MNase (NEB) in the presence of 3.3 mM $CaCl_2$ and 0.2% Triton X-100 at 30°C for 12 min with vigorous shaking (1200 rpm). Digestion was terminated by adding 12.5 μl of MNase stop buffer (250 mM EDTA, 250 mM EGTA, 0.5% SDS, 125 mM NaCl). MNase-digested nuclei slurry was sonicated using Bioruptor 300 (Diagenode) to disrupt the nuclear membrane. Released chromatin was clarified by centrifugation. Input chromatin and immunoprecipitated DNA were reversed crosslinked at 65°C overnight. DNA was purified by phenol/chloroform extraction followed by chloroform extraction. DNA was precipitated by adding 1/10 volume of 3M sodium acetate, pH 5.5 and 3 volumes of ethanol. Precipitated DNA was washed once with 70% ethanol and resuspended in 1X TE buffer (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA). Purified DNA was quantified by real time PCR with KAPA SYBR FAST qPCR Master Mix (KAPA Biosystems) and gene specific primers using a CFX Touch Real-Time PCR Detection System (Bio-Rad). Primer sequences are shown in Table S2.

shRNA-mediated knockdown and rescue by lentiviral infection

For lentivirus production, non-targeting control and pLKO plasmids targeting mouse ABCF1 (Sigma), or pHAGE plasmids for overexpression were co-transfected with packaging vectors (psPAX2 and pMD2.G, Addgene) into 293T cells using lipofectamine 2000 (Invitrogen). Supernatants were collected at 48 hr and again at 72 hr. Virus preparation was performed as described (*18*). Functional viral titer was determined by transduction of limited dilution of concentrated viruses into HeLa or NIH 3T3 cells followed by counting antibiotic-resistant colonies to obtain colony formation unit (CFU)/ml. Cells were infected with viruses in the presence of 8 μg/ml polybrene. For knockdown experiments in mouse ES cells, lentiviruses expressing control non-targeting shRNA (NT), or two independent shRNAs targeting ABCF1 were used to infect cells. Infected cells were selected with puromycin (1.5 μg/ml). Pluripotency status of control and ABCF1-knockdown cells were analyzed using an alkaline phosphatase (AP) detection kit (EMD Millipore), or by RT-qPCR analysis of mRNAs purified using TRIzol reagent (Life Technologies).

For rescue experiments in ABCF1-knockdown mouse ES cells, mouse ES cells were co-infected with viruses expressing shABCF1 (pLKO-shABCF1-1) at multiplicity of infection (MOI) of 15 and viruses expressing either control RFP at MOI of 6, mouse Nanog at MOI of 6, full-length hABCF1 (FL), or N-terminally truncated ABCF1 (NBDs) at MOI of 6 and 1, respectively. We used a lower MOI for NBDs because it expresses at a substantially higher level than full-length ABCF1 in mouse ES cells even at MOI of 1 (see Fig. S5C). Co-Infected cells were selected with puromycin (1 μg/ml) and neomycin (750 μg/ml).

siRNA-mediated knockdown in human ES cells

For knockdown of ABCF1 in human ES cells, H9 cells were transfected with 3.3 μM of si-non-targeting (siNT) and siABCF1 (siGENOME SMARTpool, Dharmacon) using Lipofectamine 3000 (Invitrogen) according to manufacturer's instructions. 48 hr after transfection, cell viability and morphology were documented before RNAs were collected by TRIzol reagent (Life Technologies) and processed for RT-qPCR analysis.

Somatic cell reprogramming and flow cytometry

CF-1 MEFs (Charles River) were transduced with inducible STEMCCA and rtTA lentivirus-containing supernatants overnight in 8 μg/ml polybrene (Sigma). Doxycycline (2 μg/ml; Sigma) was supplemented to complete mouse ES cell media to induce expression of OCT4, KLF4, SOX2, and c-

MYC. Reprogramming was assayed by AP staining (EMD Millipore), or by flow cytometry analysis using anti-SSEA1 (Biolegends) on a BD LSRFortessa, performed according to the manufacturers' protocols.

RNA isolation, reverse transcription and real time PCR analysis

Cells were rinsed once with 1X PBS. Total RNA was extracted and purified using TRIzol reagent (Life Technologies) followed by DNase I treatment (Invitrogen). cDNA synthesis was performed with 1 μg of total RNA using iScript cDNA Synthesis Kit (Bio-Rad) and diluted 10-fold. Real time PCR analysis was carried out with iTaq UniverSYBR green (Bio-Rad) and gene specific primers using the CFX Touch Real-Time PCR Detection System (Bio-Rad). Results were normalized to *β-actin*. Primer sequences are shown in Table S3.

Nucleofection of oligonucleotides and flow cytometry

D3 mouse ES cells were detached and washed once with 1X PBS. For each nucleofection 3 x 10^6 cells were resuspended in 80 μl of mouse ES cell nucleofection solution (Lonza). Equimolar (0.32 nmol) of 5' 6-carboxyfluorescein (6-FAM) labeled ss, ds-M, or ds-UM DNA oligos (IDT) (Supplementary Table 4) were nucleofected into mouse ES cells using an Amaxa nucleofector (Lonza) according to manufacturer's instructions. Cells mock-nucleofected with water were used as a negative control. Nucleofected cells were recovered in 1 ml of pre-warmed ES cell medium at 37°C for 10 min before cells were transferred to 10 cm culture plates. 24 hr after plating, cells were detached, rinsed with 1X PBS, resuspended in 0.2 ml of ice-cold PBS, and filtered into BD filter cap tubes for fluorescent activated cell sorting (FACS). FITC-positive cells were sorted on a BD FACS Aria II instrument using the following settings: 70 μM nozzle, 70 psi pressure, frequency 85, amplitude 5, and 70 sheath pressure. Mock-nucleofected negative control cells were used for setting up the gate for FITC+ cells. Once the gate is adjusted, FITC+ cells were collected directly into TRIzol (Life Technologies) solution and immediately frozen in dry ice.

Cytoplasmic and nuclear fractionation

DMSO or ETO treated mouse ES cells were washed with PBS twice and lysed with Buffer A (10 mM HEPES, pH7.9, 10 mM KCl, 0.1 mM EDTA, 0.4% NP-40, 1 mM DTT, complete protease inhibitors (Sigma)) on ice for 10min. Lysates were cleared by spinning down at 15k g for 3 min at 4°C. Supernatants were kept as cytoplasmic fractions. Nuclei were washed with Buffer A three times and lysed with Buffer B (20mM HEPES, pH7.9, 0.4 M NaCl, 1 mM EDTA, 10% Glycerol, 1 mM DTT, complete protease inhibitors (Sigma)) with vigorous shaking for 2 hr or overnight at 4°C. Lysates were cleared by centrifugation at 15k g for 5min at 4°C. Supernatants were kept as nuclear fractions.

ABCF1-DNA interaction analysis

For DNA pull-down assay, approximately 5 x 10^6 V5-ABCF1 knock-in D3 mouse ES cells were lysed in 0.2 ml of lysis buffer (0.14 M NaCl, 50 mM HEPES, pH 7.9, 1 mM EDTA, 2 mM $MgCl_2$, 0.5% NP-40, 10% glycerol). Cell lysates were homogenized by passing through a 25-gauge needle 5 times. Cell lysates were cleared by centrifugation. 5' biotinylated single stranded (ss) or doubled stranded (ds) oligonucleotides labeled with biotin at the 5' end of the sense strand were synthesized. All oligos contain modified, exonuclease-resistant nucleotides (phosphorothioate bonds) in the last 5 nucleotides at the 5' and 3' ends of both sense and antisense strands (IDT). These oligonucleotides were incubated with cell lysates for 1.5 hr at 4°C. For each pull down reaction, 100 μl of Dynabeads MyOne Streptavidin T1 bead slurry was first pre-blocked in lysis buffer containing 50 mg/ml BSA for 1 hr at 4°C. The pre-blocked beads were then incubated with cell lysate containing oligonucleotides for 1 hr at 4°C. Biotinylated DNA oligonucleotides and associated protein factors captured onto the beads were washed extensively with lysis buffer. Bound proteins were eluted by boiling the beads in sample buffer. Sequences of the 5' biotinylated oligonucleotides are available in Table S4.

To analyze binding of ABCF1 to endogenous dsDNA fragments, V5-ABCF1 knock-in D3 mouse ES cells were treated with DMSO or ETO (20 μM) for 6 hr to induce DNA fragmentation. Whole cell extracts were prepared by using the same lysis buffer as for DNA pull-down assay. Cell lysates were homogenized and cleared. Cleared lysates were incubated with IgG or anti-V5 overnight at 4°C. Protein A Sepharose were added to lysates and incubated for 1-2 hr at 4°C. Bound proteins were washed extensively with lysis buffer. Bound DNAs were eluted by RNase and Proteinase K treatment. Eluted DNAs were extracted by phenol-chloroform extraction and purified by ethanol precipitation. Purified DNAs were run on a 6% Urea-denaturing gel, stained with SYBR gold (Invitrogen), and analyzed.

Colony formation assay

To examine colony forming ability of ABCF1 overexpressing cells under DNA damaging condition, D3 mouse ES cells were transduced with lentiviruses expressing untagged RFP, V5-tagged full-length human ABCF1, or V5-tagged LCD-truncated (NBDs) ABCF1. Mouse ES cells were transduced with lentiviruses at MOI 3 and selected with neomycin (500 μg/ml). 200 or 1000 RFP or ABCF1 overexpressing mouse ES cells were plated on 24-well plates. Cells were allowed to recover for 24 hr before treatment with DMSO or ETO (1 μM) for 1, 2, or 4 hr, after which fresh medium without ETO was replaced. After 6 days, cells were fixed and stained for AP activity (EMD Millipore). AP-positive cells were counted and analyzed.

Amino acid composition analysis

The amino acid composition of a protein was analyzed in R with custom scripts. The occurrence of each amino acid is counted by using the package 'stringr' and plotted with the package 'plot. matrix.'

Statistical analysis

To determine statistical significance, *P*-values were calculated by using unpaired two-sided Student's t-test. *P*-values less than 0.05 (<0.05) were considered as statistically significant and they were indicated with * (*P<0.05). All data represent the mean ± SEM (error bars) except for Fig. 6I. For Fig. 6I, box-and-whisker plot was used. The central mark represents median and edges represent 25th and 75th percentiles. The whiskers indicate 5th to 95th percentiles.

REFERENCES

1. S. J. Dunn, G. Martello, B. Yordanov, S. Emmott, A. G. Smith, Defining an essential transcription factor program for naïve pluripotency. *Science (80-.).* **344**, 1156–1160 (2014).
2. R. A. Young, Control of the embryonic stem cell state. *Cell.* **144** (2011), pp. 940–954.
3. S. J. Hainer, A. Bošković, K. N. McCannell, O. J. Rando, T. G. Fazzio, Profiling of Pluripotency Factors in Single Cells and Early Embryos. *Cell.* **177**, 1319-1329.e11 (2019).
4. A. Marson, S. S. Levine, M. F. Cole, G. M. Frampton, T. Brambrink, S. Johnstone, M. G. Guenther, W. K. Johnston, M. Wernig, J. Newman, J. M. Calabrese, L. M. Dennis, T. L. Volkert, S. Gupta, J. Love, N. Hannett, P. A. Sharp, D. P. Bartel, R. Jaenisch, R. A. Young, Connecting microRNA Genes to the Core Transcriptional Regulatory Circuitry of Embryonic Stem Cells. *Cell.* **134**, 521–533 (2008).
5. L. A. Boyer, I. L. Tong, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, R. A. Young, Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* **122**, 947–956 (2005).
6. J. Kim, J. Chu, X. Shen, J. Wang, S. H. Orkin, An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell.* **132**, 1049–1061 (2008).
7. K. Takahashi, S. Yamanaka, A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* **17** (2016), pp. 183–193.
8. J. Brumbaugh, B. Di Stefano, K. Hochedlinger, Reprogramming: Identifying the mechanisms that safeguard cell identity. *Dev.* **146** (2019), , doi:10.1242/dev.182170.
9. R. Jaenisch, R. Young, Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell.* **132** (2008), pp. 567–582.
10. Y. W. Fong, C. Cattoglio, T. Yamaguchi, R. Tjian, Transcriptional regulation by coactivators in embryonic stem cells. *Trends Cell Biol.* **22** (2012), pp. 292–298.
11. A. M. Näär, B. D. Lemon, R. Tjian, Transcriptional Coactivator Complexes. *Annu. Rev. Biochem.* **70**, 475–501 (2001).
12. R. G. Roeder, in *FEBS Letters* (Elsevier, 2005; http://www.ncbi.nlm.nih.gov/pubmed/15680973), vol. 579, pp. 909–915.
13. A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, R. A. Young, Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell.* **175**, 1842-1855.e16 (2018).
14. S. Chong, C. Dugast-Darzacq, Z. Liu, P. Dong, G. M. Dailey, C. Cattoglio, A. Heckert, S. Banala, L. Lavis, X. Darzacq, R. Tjian, Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science (80-.).* **361** (2018), doi:10.1126/science.aar2555.
15. I. Kwon, M. Kato, S. Xiang, L. Wu, P. Theodoropoulos, H. Mirzaei, T. Han, S. Xie, J. L. Corden, S. L. McKnight, XPhosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell.* **155**, 1049 (2013).
16. T. M. Harper, D. J. Taatjes, The complex structure and function of Mediator. *J. Biol. Chem.* **293** (2018), pp. 13778–13785.
17. A. M. Näär, P. A. Beaurang, S. Zhou, S. Abraham, W. Solomon, R. Tjian, Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature.* **398**, 828–832 (1999).
18. Y. W. Fong, C. Inouye, T. Yamaguchi, C. Cattoglio, I. Grubisic, R. Tjian, A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. *Cell.* **147**, 120–131 (2011).
19. D. J. Rodda, J. L. Chew, L. H. Lim, Y. H. Loh, B. Wang, H. H. Ng, P. Robson, Transcriptional regulation of Nanog by OCT4 and SOX2. *J. Biol. Chem.* **280**, 24731–24737 (2005).
20. C. Cattoglio, E. T. Zhang, I. Grubisic, K. Chiba, Y. W. Fong, R. Tjian, Functional and mechanistic studies of XPC DNA-repair complex as transcriptional coactivator in embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E2317–E2326 (2015).
21. E. T. Zhang, Y. He, P. Grob, Y. W. Fong, E. Nogales, R. Tjian, Architecture of the human XPC DNA repair and stem cell coactivator complex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14817–14822 (2015).
22. Y. W. Fong, J. J. Ho, C. Inouye, R. Tjian, The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. *Elife.* **3** (2014), doi:10.7554/elife.03573.
23. Y. W. Fong, C. Cattoglio, R. Tjian, The Intertwined Roles of Transcription and Repair Proteins. *Mol. Cell.* **52** (2013), pp. 291–302.
24. S. Paytubi, X. Wang, Y. W. Lam, L. Izquierdo, M. J. Hunter, E. Jan, H. S. Hundal, C. G. Proud, ABC50 promotes translation initiation in mammalian cells. *J. Biol. Chem.* **284**, 24061–24073 (2009).
25. S. Paytubi, N. A. Morrice, J. Boudeau, C. G. Proud, The N-terminal region of ABC50 interacts with eukaryotic initiation factor eIF2 and is a target for regulatory phosphorylation by CK2. *Biochem. J.* **409**, 223–31 (2008).
26. M. N. Lee, M. Roy, S. E. Ong, P. Mertins, A. C. Villani, W. Li, F. Dotiwala, J. Sen, J. G. Doench, M. H. Orzalli, I. Kramnik, D. M. Knipe, J. Lieberman, S. A. Carr, N. Hacohen, Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach. *Nat. Immunol.* **14**, 179–185 (2013).
27. H. Arora, S. M. Wilcox, L. A. Johnson, L. Munro, B. A. Eyford, C. G. Pfeifer, I. Welch, W. A. Jefferies, The ATP-Binding Cassette Gene ABCF1 Functions as an E2 Ubiquitin-Conjugating Enzyme Controlling Macrophage Polarization to Dampen Lethal Septic Shock. *Immunity.* **50**, 418-431.e6 (2019).
28. Y. L. Guo, The underdeveloped innate immunity in embryonic stem cells: The molecular basis and biological perspectives from early embryogenesis. *Am. J. Reprod. Immunol.* **81**, e13089 (2019).
29. J. Eggenberger, D. Blanco-Melo, M. Panis, K. J. Brennand, B. R. Ten Oever, Type I interferon response impairs differentiation potential of pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1384–1393 (2019).
30. T. Brambrink, R. Foreman, G. G. Welstead, C. J. Lengner, M. Wernig, H. Suh, R. Jaenisch, Sequential Expression

of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells. *Cell Stem Cell*. **2**, 151–159 (2008).
31. I. Ginis, Y. Luo, T. Miura, S. Thies, R. Brandenberger, S. Gerecht-Nir, M. Amit, A. Hoke, M. K. Carpenter, J. Itskovitz-Eldor, M. S. Rao, Differences between human and mouse embryonic stem cells. *Dev. Biol.* **269**, 360–380 (2004).
32. X. Chen, H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y. H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W. K. Sung, N. D. Clarke, C. L. Wei, H. H. Ng, Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. *Cell*. **133**, 1106–1117 (2008).
33. S. M. Wilcox, H. Arora, L. Munro, J. Xin, F. Fenninger, L. A. Johnson, C. G. Pfeifer, K. B. Choi, J. Hou, P. A. Hoodless, W. A. Jefferies, The role of the innate immune response regulatory gene ABCF1 in mammalian embryogenesis and development. *PLoS One*. **12** (2017), doi:10.1371/journal.pone.0175918.
34. G. R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 7634–7638 (1981).
35. M. J. Evans, M. H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos. *Nature*. **292**, 154–156 (1981).
36. F. A. Brook, R. L. Gardner, The origin and efficient derivation of embryonic stem cells in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5709–5712 (1997).
37. J. C. D. Heng, B. Feng, J. Han, J. Jiang, P. Kraus, J. H. Ng, Y. L. Orlov, M. Huss, L. Yang, T. Lufkin, B. Lim, H. H. Ng, The Nuclear Receptor Nr5a2 Can Replace Oct4 in the Reprogramming of Murine Somatic Cells to Pluripotent Cells. *Cell Stem Cell*. **6**, 167–174 (2010).
38. H. Iseki, Y. Nakachi, T. Hishida, Y. Yamashita-Sugahara, M. Hirasaki, A. Ueda, Y. Tanimoto, S. Iijima, F. Sugiyama, K. I. Yagami, S. Takahashi, A. Okuda, Y. Okazaki, Combined Overexpression of JARID2, PRDM14, ESRRB, and SALL4A Dramatically Improves Efficiency and Kinetics of Reprogramming to Induced Pluripotent Stem Cells. *Stem Cells*. **34**, 322–333 (2016).
39. J. Yu, M. A. Vodyanik, K. Smuga-Otto, J. Antosiewicz-Bourget, J. L. Frane, S. Tian, J. Nie, G. A. Jonsdottir, V. Ruotti, R. Stewart, I. I. Slukvin, J. A. Thomson, Induced pluripotent stem cell lines derived from human somatic cells. *Science (80-.).* **318**, 1917–1920 (2007).
40. K. Takahashi, S. Yamanaka, Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. **126**, 663–676 (2006).
41. C. A. Sommer, M. Stadtfeld, G. J. Murphy, K. Hochedlinger, D. N. Kotton, G. Mostoslavsky, Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. *Stem Cells*. **27**, 543–549 (2009).
42. J. M. Polo, E. Anderssen, R. M. Walsh, B. A. Schwarz, C. M. Nefzger, S. M. Lim, M. Borkent, E. Apostolou, S. Alaei, J. Cloutier, O. Bar-Nur, S. Cheloufi, M. Stadtfeld, M. E. Figueroa, D. Robinton, S. Natesan, A. Melnick, J. Zhu, S. Ramaswamy, K. Hochedlinger, A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*. **151**, 1617–1632 (2012).
43. P. Y. Zeng, C. R. Vakoc, Z. C. Chen, G. A. Blobel, S. L. Berger, In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. *Biotechniques*. **41** (2006), pp. 694–698.
44. V. Vasiliou, K. Vasiliou, D. W. Nebert, Human ATP-binding cassette (ABC) transporter family. *Hum. Genomics*. **3**, 281–290 (2009).
45. R. A. Coots, X. M. Liu, Y. Mao, L. Dong, J. Zhou, J. Wan, X. Zhang, S. B. Qian, m6A Facilitates eIF4F-Independent mRNA Translation. *Mol. Cell*. **68**, 504-514.e7 (2017).
46. G. ´ Abor, E. ˝ Os, Z. Dosztányi, D. Dosztányi, IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding B ´ alint M ´ eszárosesz´eszáros. *Nucleic Acids Res.* **46**, 329–337 (2018).
47. L. Qu, Y. Jiang, C. Cheng, D. Wu, B. Meng, Z. Chen, Y. Zhu, N. Shaw, S. Ouyang, Z. J. Liu, Crystal Structure of ATP-Bound Human ABCF1 Demonstrates a Unique Conformation of ABC Proteins. *Structure*. **26**, 1259-1265.e3 (2018).
48. J. C. Hansen, X. Lu, E. D. Ross, R. W. Woody, Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* **281** (2006), pp. 1853–1856.
49. H. J. Dyson, Making Sense of Intrinsically Disordered Proteins. *Biophys. J.* **110** (2016), pp. 1013–1016.
50. R. K. Das, R. V Pappu, Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13392–13397 (2013).
51. R. Van Der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright, M. M. Babu, Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).
52. C. W. Pak, M. Kosno, A. S. Holehouse, S. B. Padrick, A. Mittal, R. Ali, A. A. Yunus, D. R. Liu, R. V. Pappu, M. K. Rosen, Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell*. **63**, 72–85 (2016).
53. B. R. Sabari, A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, D. S. Day, J. Schuijers, E. Vasile, S. Malik, D. Hnisz, T. I. Lee, I. I. Cisse, R. G. Roeder, P. A. Sharp, A. K. Chakraborty, R. A. Young, Coactivator condensation at super-enhancers links phase separation and gene control. *Science (80-.).* **361**, eaar3958 (2018).
54. F. Frottin, F. Schueder, S. Tiwary, R. Gupta, R. Körner, T. Schlichthaerle, J. Cox, R. Jungmann, F. U. Hartl, M. S. Hipp, The nucleolus functions as a phase-separated protein quality control compartment. *Science (80-.).* **365**, 342–347 (2019).
55. C. R. Vazquez de Aldana, M. J. Marton, A. G. Hinnebusch, GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2 alpha kinase GCN2 in amino acid-starved cells. *EMBO J.* **14**, 3184–3199 (1995).
56. A. M. Näär, D. J. Taatjes, W. Zhai, E. Nogales, R. Tjian, Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev.* **16**, 1339–1344 (2002).
57. M. Wilson, P. Koopman, Matching SOX: Partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr. Opin. Genet. Dev.* **12** (2002), pp. 441–446.
58. G. E. Schepers, R. D. Teasdale, P. Koopman, Twenty pairs of Sox: Extent, homology, and nomenclature of the mouse and human Sox transcription factor gene families. *Dev. Cell*. **3** (2002), pp. 167–170.
59. A. Sarkar, K. Hochedlinger, The Sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell Stem Cell*. **12** (2013), pp. 15–30.
60. S. R. Paludan, A. G. Bowie, Immune Sensing of DNA. *Immunity*. **38**, 870–880 (2013).

61. D. B. Stetson, R. Medzhitov, Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity.* **24**, 93–103 (2006).
62. H. Ishikawa, Z. Ma, G. N. Barber, STING regulates intracellular DNA-mediated, type i interferon-dependent innate immunity. *Nature.* **461**, 788–792 (2009).
63. K. J. Ishii, C. Coban, H. Kato, K. Takahashi, Y. Torii, F. Takeshita, H. Ludwig, G. Sutter, K. Suzuki, H. Hemmi, S. Sato, M. Yamamoto, S. Uematsu, T. Kawai, O. Takeuchi, S. Akira, A toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nat. Immunol.* **7**, 40–48 (2006).
64. Y. H. Chiu, J. B. MacMillan, Z. J. Chen, RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway. *Cell.* **138**, 576–591 (2009).
65. K. J. MacKenzie, P. Carroll, C. A. Martin, O. Murina, A. Fluteau, D. J. Simpson, N. Olova, H. Sutcliffe, J. K. Rainger, A. Leitch, R. T. Osborn, A. P. Wheeler, M. Nowotny, N. Gilbert, T. Chandra, M. A. M. Reijns, A. P. Jackson, CGAS surveillance of micronuclei links genome instability to innate immunity. *Nature.* **548**, 461–465 (2017).
66. A. Härtlova, S. F. Erttmann, F. A. M. Raffi, A. M. Schmalz, U. Resch, S. Anugula, S. Lienenklaus, L. M. Nilsson, A. Kröger, J. A. Nilsson, T. Ek, S. Weiss, N. O. Gekara, DNA Damage Primes the Type I Interferon System via the Cytosolic DNA Sensor STING to Promote Anti-Microbial Innate Immunity. *Immunity.* **42**, 332–343 (2015).
67. G. Dunphy, S. M. Flannery, J. F. Almine, D. J. Connolly, C. Paulus, K. L. Jønsson, M. R. Jakobsen, M. M. Nevels, A. G. Bowie, L. Unterholzner, Non-canonical Activation of the DNA Sensing Adaptor STING by ATM and IFI16 Mediates NF-κB Signaling after Nuclear DNA Damage. *Mol. Cell.* **71**, 745-760.e5 (2018).
68. R. Wang, J. Wang, A. M. Paul, D. Acharya, F. Bai, F. Huang, Y. L. Guo, Mouse embryonic stem cells are deficient in type i interferon expression in response to viral infections and double-stranded RNA. *J. Biol. Chem.* **288**, 15926–15936 (2013).
69. A. Zampetaki, Q. Xiao, L. Zeng, Y. Hu, Q. Xu, TLR4 expression in mouse embryonic stem cells and in stem cell-derived vascular cells is regulated by epigenetic modifications. *Biochem. Biophys. Res. Commun.* **347**, 89–99 (2006).
70. J. Yu, R. Rossi, C. Hale, D. Goulding, G. Dougan, Interaction of enteric bacterial pathogens with murine embryonic stem cells. *Infect. Immun.* **77**, 585–597 (2009).
71. L. Guo, L. Lin, X. Wang, M. Gao, S. Cao, Y. Mai, F. Wu, J. Kuang, H. Liu, J. Yang, S. Chu, H. Song, D. Li, Y. Liu, K. Wu, J. Liu, J. Wang, G. Pan, A. P. Hutchins, J. Liu, D. Pei, J. Chen, Resolving Cell Fate Decisions during Somatic Cell Reprogramming by Single-Cell RNA-Seq. *Mol. Cell.* **73**, 815-829.e7 (2019).
72. W. D'Angelo, C. Gurung, D. Acharya, B. Chen, N. Ortolano, V. Gama, F. Bai, Y.-L. Guo, The Molecular Basis for the Lack of Inflammatory Responses in Mouse Embryonic Stem Cells and Their Differentiated Cells. *J. Immunol.* **198**, 2147–2155 (2017).
73. P. Xia, S. Wang, B. Ye, Y. Du, G. Huang, P. Zhu, Z. Fan, Sox2 functions as a sequence-specific DNA sensor in neutrophils to initiate innate immunity against microbial infection. *Nat. Immunol.* **16**, 366–378 (2015).
74. Z. Zhang, R. Tjian, Measuring dynamics of eukaryotic transcription initiation: Challenges, insights and opportunities (2017), doi:10.1080/21541264.2017.1363017.
75. U. B. Choi, H. Sanabria, T. Smirnova, M. E. Bowen, K. R. Weninger, Spontaneous switching among conformational ensembles in intrinsically disordered proteins. *Biomolecules.* **9** (2019), doi:10.3390/biom9030114.
76. Y. E. Guo, J. C. Manteiga, J. E. Henninger, B. R. Sabari, A. Dall'Agnese, N. M. Hannett, J. H. Spille, L. K. Afeyan, A. V. Zamudio, K. Shrinivas, B. J. Abraham, A. Boija, T. M. Decker, J. K. Rimel, C. B. Fant, T. I. Lee, I. I. Cisse, P. A. Sharp, D. J. Taatjes, R. A. Young, Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature.* **572** (2019), pp. 543–548.
77. A. Roers, B. Hiller, V. Hornung, Recognition of Endogenous Nucleic Acids by the Innate Immune System. *Immunity.* **44** (2016), pp. 739–754.
78. G. Chatzinikolaou, I. Karakasilioti, G. A. Garinis, DNA damage and innate immunity: Links and trade-offs. *Trends Immunol.* **35** (2014), pp. 429–435.
79. S. M. Harding, J. L. Benci, J. Irianto, D. E. Discher, A. J. Minn, R. A. Greenberg, Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature.* **548**, 466–470 (2017).
80. I. Jorgensen, M. Rayamajhi, E. A. Miao, Programmed cell death as a defence against infection. *Nat. Rev. Immunol.* **17** (2017), pp. 151–164.
81. C. B. Coyne, H. M. Lazear, Zika virus-reigniting the TORCH. *Nat. Rev. Microbiol.* **14** (2016), pp. 707–715.
82. X. Wu, V. L. Dao Thi, Y. Huang, E. Billerbeck, D. Saha, H. H. Hoffmann, Y. Wang, L. A. V. Silva, S. Sarbanes, T. Sun, L. Andrus, Y. Yu, C. Quirk, M. Li, M. R. MacDonald, W. M. Schneider, X. An, B. R. Rosenberg, C. M. Rice, Intrinsic Immunity Shapes Viral Resistance of Stem Cells. *Cell.* **172**, 423-438.e25 (2018).
83. A. K. Ahuja, K. Jodkowska, F. Teloni, A. H. Bizard, R. Zellweger, R. Herrador, S. Ortega, I. D. Hickson, M. Altmeyer, J. Mendez, M. Lopes, A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nat. Commun.* **7**, 10660 (2016).
84. I. I. Suvorova, B. B. Grigorash, I. A. Chuykin, T. V. Pospelova, V. A. Pospelov, G1 checkpoint is compromised in mouse ESCs due to functional uncoupling of p53-p21waf1 signaling. *Cell Cycle.* **15**, 52–63 (2016).
85. M. I. Aladjem, B. T. Spike, L. W. Rodewald, T. J. Hope, M. Klemm, R. Jaenisch, G. M. Wahl, ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. *Curr. Biol.* **8**, 145–155 (1998).
86. B. S. Heyer, A. Macauley, O. Behrendtsen, Z. Werb, Hypersensitivity to DNA damage leads to increased apoptosis during early mouse development. *Genes Dev.* **14**, 2072–2084 (2000).
87. H. Qin, T. Yu, T. Qing, Y. Liu, Y. Zhao, J. Cai, J. Li, Z. Song, X. Qu, P. Zhou, J. Wu, M. Ding, H. Deng, Regulation of apoptosis and differentiation by p53 in human embryonic stem cells. *J. Biol. Chem.* **282**, 5842–5852 (2007).
88. S. Maynard, A. M. Swistowska, J. W. Lee, Y. Liu, S.-T. Liu, A. Bettencourt, D. A. Cruz, M. Rao, N. C. De Souza-Pinto, X. Zeng, V. A. Bohr, Human Embryonic Stem Cells Have Enhanced Repair of Multiple Forms of DNA Damage. *Stem Cells.* **26**, 2266–2274 (2008).
89. I. Vitale, G. Manic, R. De Maria, G. Kroemer, L. Galluzzi, DNA Damage in Stem Cells. *Mol. Cell.* **66** (2017), pp. 306–319.
90. E. H. Choi, S. Yoon, K. S. Park, K. P. Kim, The Homologous Recombination Machinery Orchestrates Post-replication DNA Repair during Self-renewal of Mouse Embryonic Stem Cells. *Sci. Rep.* **7** (2017), doi:10.1038/s41598-017-11951-1.
91. J. Weng, W. Wang, Dynamic multivalent interactions of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **62**

(2020), pp. 9–13.
92. L. Staby, C. O'Shea, M. Willemoës, F. Theisen, B. B. Kragelund, K. Skriver, Eukaryotic transcription factors: Paradigms of protein intrinsic disorder. *Biochem. J.* **474** (2017), pp. 2509–2532.
93. D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, P. A. Sharp, A Phase Separation Model for Transcriptional Control. *Cell.* **169** (2017), pp. 13–23.
94. J. Wang, J. M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, A. A. Hyman, A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell.* **174**, 688-699.e16 (2018).
95. M. Kato, T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, N. V. Grishin, D. E. Frantz, J. W. Schneider, S. Chen, L. Li, M. R. Sawaya, D. Eisenberg, R. Tycko, S. L. McKnight, Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell.* **149**, 753–767 (2012).
96. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
97. J. Mieczkowski, A. Cook, S. K. Bowman, B. Mueller, B. H. Alver, S. Kundu, A. M. Deaton, J. A. Urban, E. Larschan, P. J. Park, R. E. Kingston, M. Y. Tolstorukov, MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat. Commun.* **7** (2016), doi:10.1038/ncomms11485.

ACKNOWLEDGEMENTS

We thank G. Dailey for help with expression constructs; S. Zheng for tissue culture assistance; and S. Zhou and D. King for mass spectrometry analysis, and E. Guo for advice on in vitro droplet assay. We thank R. Tjian, Y-C. Hsu, S. Agarwal, Y. Isogai, Z. Zhang, B. Zhang, E. Guo, S. Chong, and C. Cattoglio for valuable discussion. **Funding:** We acknowledge the following funding sources: NIH grant R01HL125527, Harvard Stem Cell Institute, Boston Biomedical Innovation Center, Charles H. Hood Foundation, Brigham Research Institute, and Brigham and Women's Hospital HVC Junior Faculty Research Awards to Y.W.F.

AUTHOR CONTRIBUTIONS

Y.W.F. conceived and supervised this study. Y.W.F., E.C., M.V., and M.Z. designed the experiments. E.C. and Y.W.F. conducted biochemical purification and in vitro assays. C.I. purified general transcription factors for in vitro transcription assays. J.W. assisted in purification of recombinant proteins. E.C., M.V., and M.Z. performed cell-based assays and gene expression analyses. M.V. and M.Z. performed ChIP and FACS analyses. J.J.H. performed somatic cell reprogramming experiments and FACS analysis. Y.W.F., E.C., and M.V. wrote the manuscript. All authors edited the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

SUPPLEMENTARY FIGURES AND TABLES

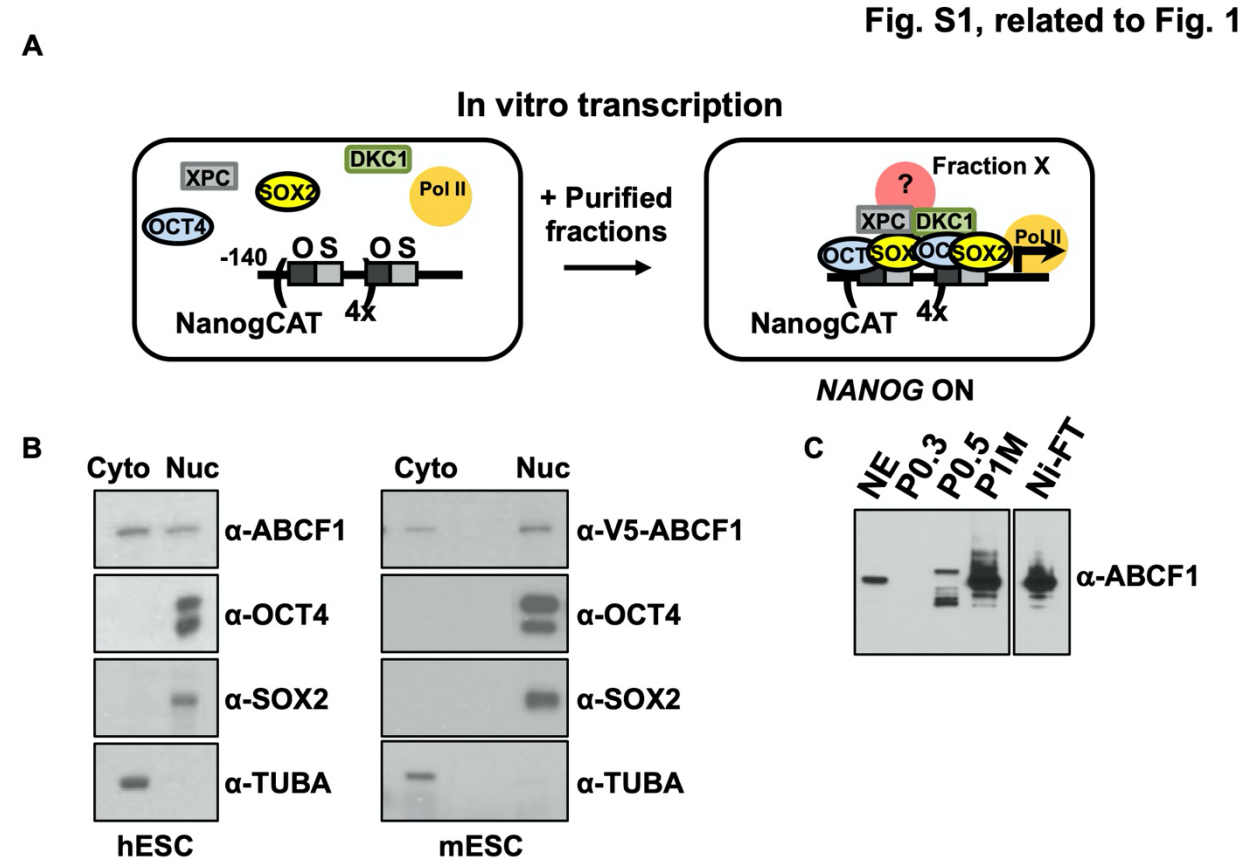

Fig. S1. Schematic diagram of the in vitro transcription assay and subcellular localization of ABCF1. (A) Schematic diagram of in vitro transcription assay to detect SCC-B activity. In vitro transcription reactions contain purified general transcription factors (TFII-A, B, D, E, F, and H), RNA polymerase II (Pol II), recombinant XPC and DKC1 complexes, and affinity-purified OCT4 and SOX2. These factors are supplemented with partially purified fractions from NT2 nuclear extracts and assayed for the ability of protein fractions to stimulate transcription of a modified transcription template derived from the human *NANOG* promoter (-140). The modified *NANOG* template contains four extra copies of the OCT4/SOX2 composite binding sites upstream of the native enhancer. We have previously shown that SCC-dependence transcriptional activation is also observed using an unadulterated, native *NANOG* promoter template (Fong et al., 2011). Transcription is measured by primer extension assay. **(B)** Western blot analyses of ABCF1 protein levels in cytoplasmic and nuclear fractions prepared from human and mouse ES cells. Effectiveness of fractionation is indicated by enrichment of OCT4 and SOX2 in nuclear fractions, and α-tubulin (TUBA) in cytoplasmic fractions. **(C)** ABCF1 is highly enriched in the transcriptionally active, partially purified nuclear extract fractions. Comparative western blot analysis of NT2 nuclear extract (NE), phosphocellulose 0.3 M, 0.5 M, 1 M KCl fractions (P0.3, P0.5, P1M, respectively), and Ni-NTA flowthrough (Ni-FT) using an antibody against ABCF1. P0.3 and P0.5 fractions do not contain SCCs (Fong et al., 2011). Each lane contains 5 mg of protein as determined by Bradford protein assay.

Fig. S2. related to Fig. 2

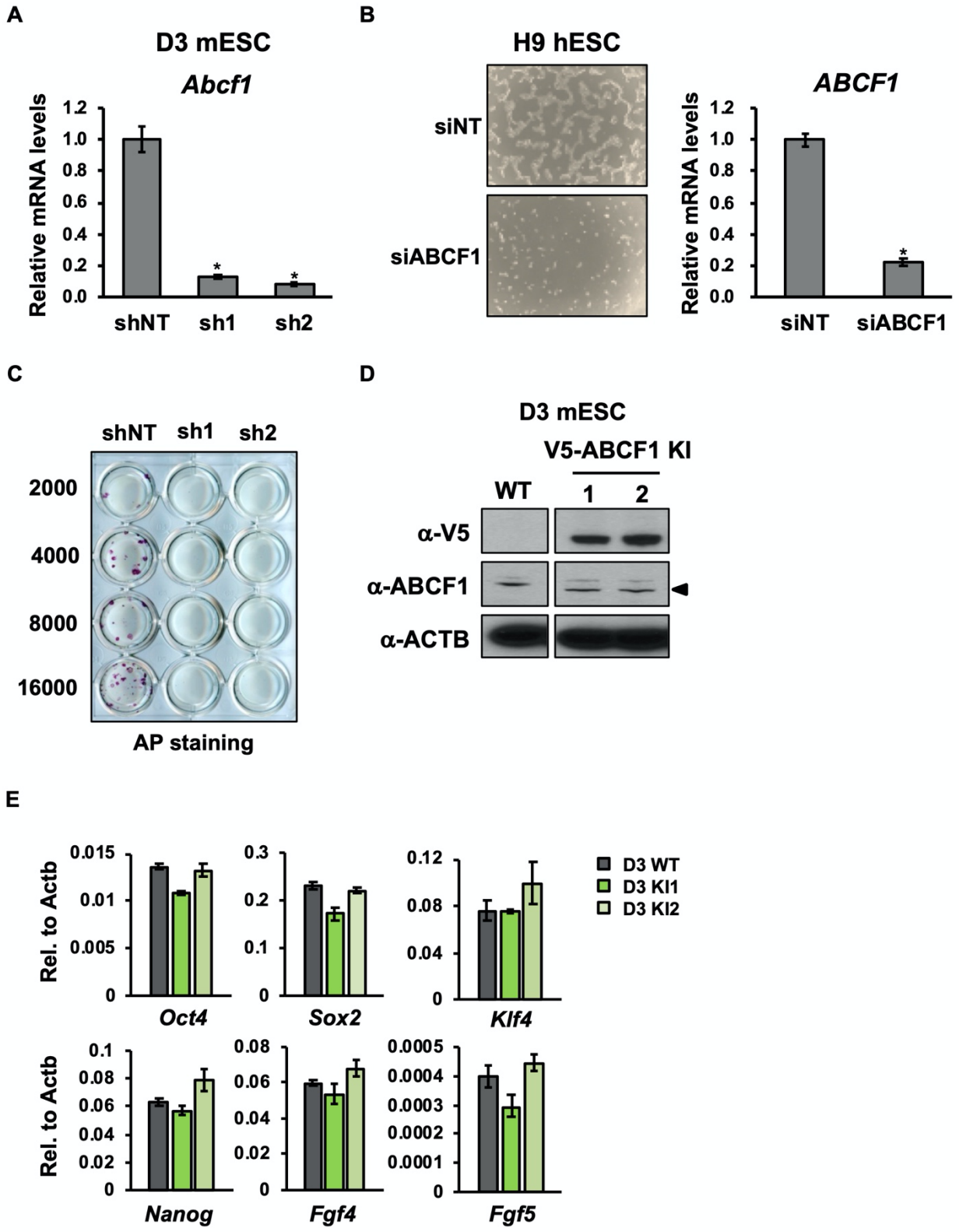

Fig. S2 cont'd, related to Fig. 2

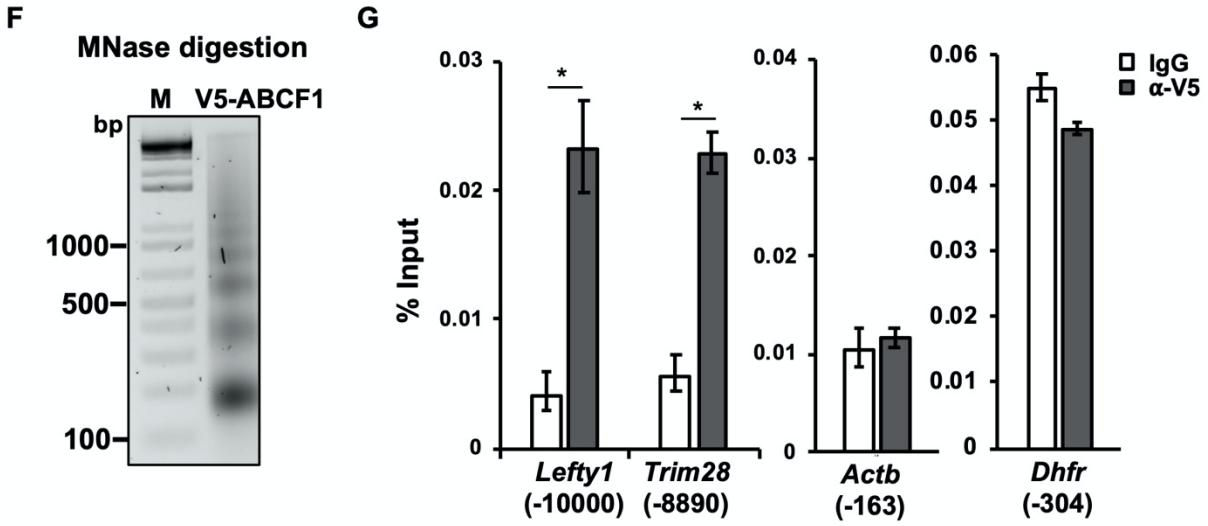

Fig. S2. ABCF1 is required for stem cell pluripotency and reprogramming and recruited to OCT4/SOX2 binding sites. (A) Efficiency of shRNA-mediated knockdown of ABCF1 in D3 mouse ES cells is determined by qPCR. *Abcf1* mRNA levels are normalized to *Actb* and expressed as fraction of control (=1). **(B)** H9 human ES cells are transfected with non-targeting siRNA (siNT) or siRNAs against ABCF1 (siABCF1). Knockdown efficiency is analyzed by qPCR as in A. Downregulation of ABCF1 in human ES cells compromises stem cell self-renewal. Phase contrast images of H9 human ES cells transfected with siNT or siABCF1. **(C)** Depletion of ABCF1 blocks somatic cell reprogramming. CF-1 MEFs are reprogrammed into iPS cells as described in Fig. 2E. Cells are plated at the indicated number in 24-well plates. Reprogrammed cells are stained for alkaline phosphatase (AP) activity. **(D)** Expression of V5-tagged ABCF1 (V5-ABCF1) in two independent knock-in (KI) D3 mouse ES cell clones is verified by western blotting using antibodies against V5 and ABCF1. ACTB is used as loading control. **(E)** V5 tagging of ABCF1 in mouse ES cells (D3 KI1, 2) does not compromise ABCF1 function. mRNA levels of *Oct4*, *Sox2*, *Nanog*, *Klf4*, and *Fgf4* are quantified by qPCR and normalized to *Actb*. V5-ABCF1 knock-in clones display low expression level of differentiation-associated gene, *Fgf5*, comparable to WT cells. **(F)** MNase-digested chromatin from EGS/formaldehyde crosslinked V5-ABCF1 KI mouse ES cell are purified and separated on an agarose gel. **(G)** ChIP-qPCR analysis of ABCF1 enrichment at previously identified strong OCT/SOX2 binding sites on *Lefty1* and *Trim28*, and at housekeeping gene *Actb* and *Dhfr* is analyzed as described in Fig. 2. Error bars present SEM. $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.

Fig. S3. related to Fig. 3

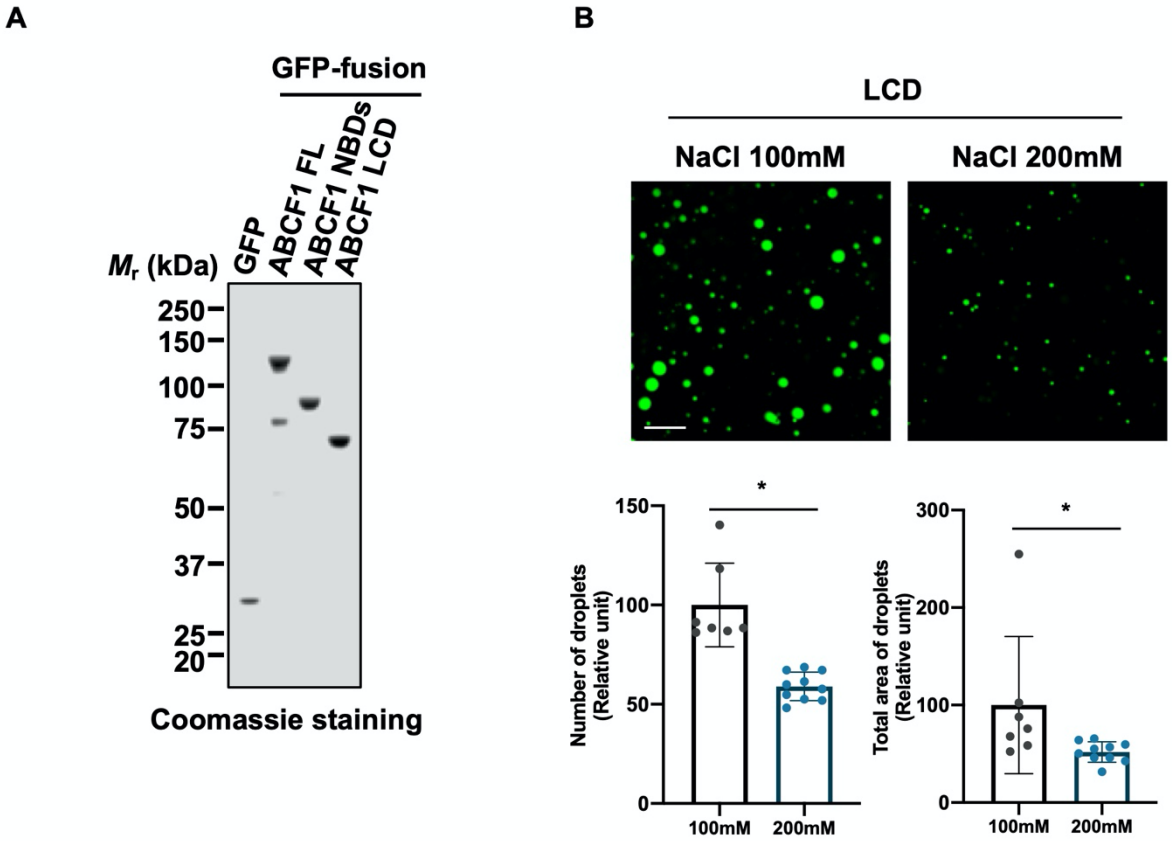

Fig. S3. Purified recombinant GFP-fusion ABCF1 proteins and droplet formation of ABCF1 LCD in different salt concentrations. (A) SDS-PAGE and Coomassie staining of recombinant GFP and various GFP-ABCF1 fusion proteins purified from *E. coli*. GFP, full-length (FL) ABCF1, NBDs, and LCD of ABCF1 are used in in vitro droplet reactions. **(B)** Fluorescence microscopy images (left) showing droplet formation of LCD at two different NaCl concentrations (100 and 200 mM). Number and size of droplets are quantified by ImageJ and indicated as relative units (right). Error bars present SEM. n = 8-10. (*) $P < 0.05$, calculated by two-sided Student's t-test.

Fig. S4, related to Fig. 4

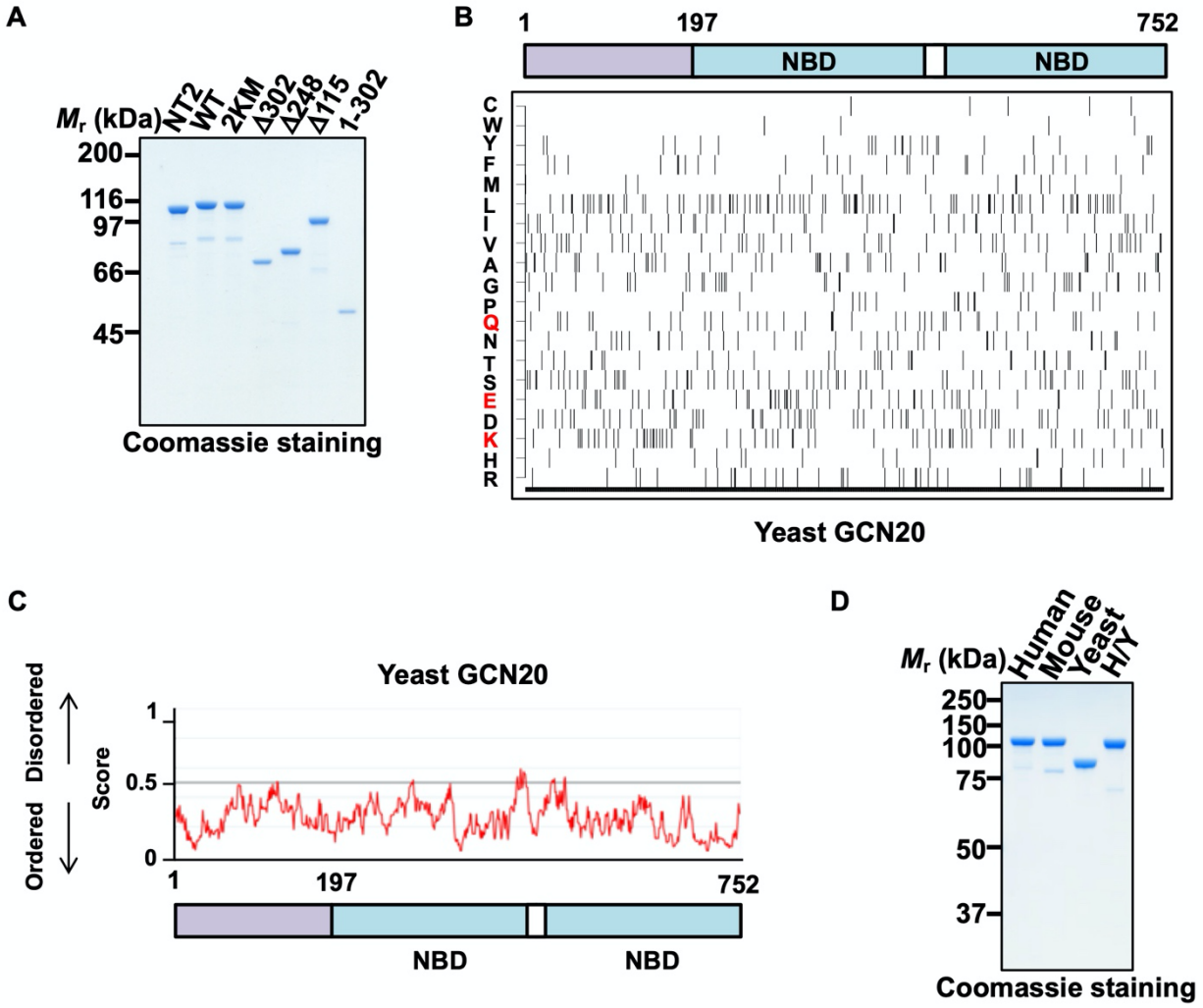

Fig. S4. Purified recombinant ABCF1 proteins used in vitro transcription assays and structural analyses of yeast GCN20. (A) SDS-PAGE and Coomassie staining of endogenous ABCF1 purified from NT2 nuclear extracts (NT2), purified recombinant full-length and truncated ABCF1 proteins described in Fig. 4A. **(B)** Amino acid composition of yeast GCN20, analyzed as described in Fig. 3A. The relative positions of NBDs and the non-conserved N-terminal region of GCN20 are indicated. One-letter abbreviations for amino acids are indicated (Left): C, Cys; W, Trp; Y, Tyr; F, Phe; M, Met; L, Leu; I, Ile; V, Val; A, Ala; G, Gly; P, Pro; Q, Gln; N, Asn; T, Thr; S, Ser; E, Glu; D, Asp; K, Lys; H, His; R, Arg. **(C)** Intrinsically disordered protein prediction tool, IUPred2 (Abor et al., 2018), is used to identify unstructured regions in yeast GCN20. X-axis indicates position of the amino acids and Y-axis shows the probability of disordered sequences. Regions that are above the value of 0.5 are predicted to be unstructured and those below 0.5 are more likely structured. Relative positions of the protein domains are depicted below. Similar results are also obtained using a different prediction program DISOPRED3 (data not shown) (Jones and Cozzetto, 2015). **(D)** SDS-PAGE and Coomassie staining of recombinant human, mouse, and yeast homolog of ABCF1 (GCN20) as well as the human LCD-yeast NBDs hybrid protein (H/Y).

Fig. S5, related to Fig. 5

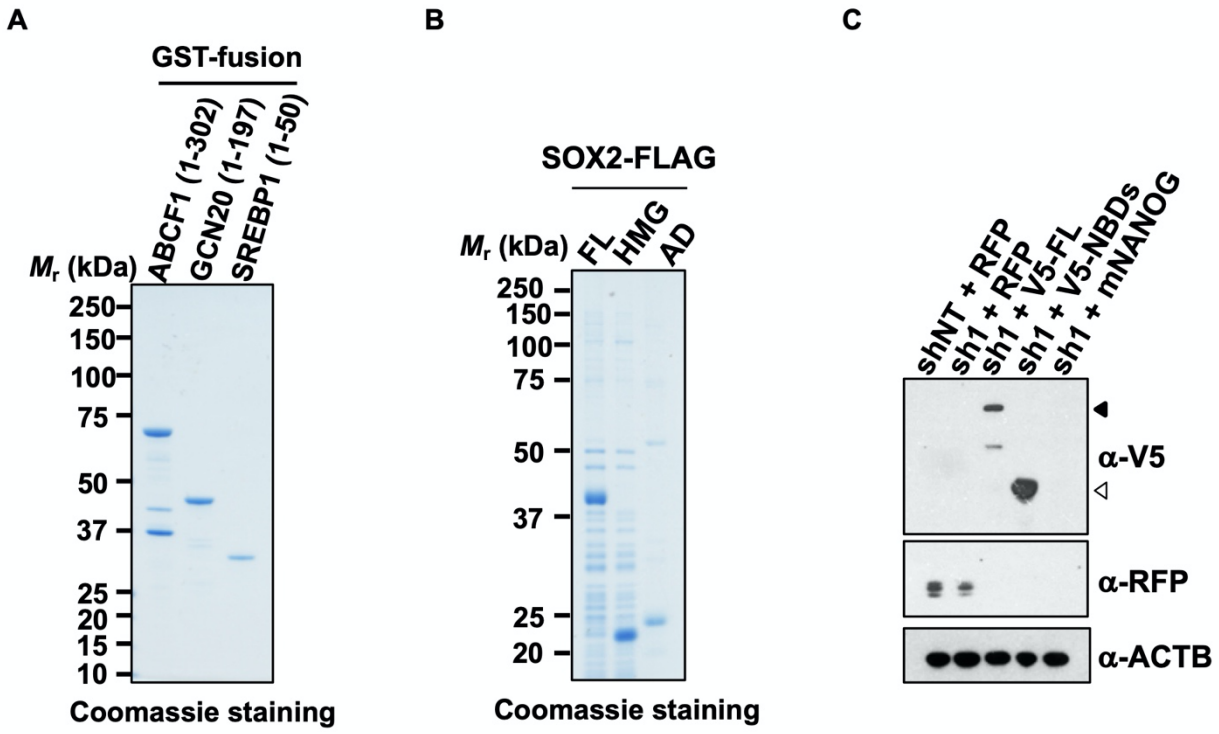

Fig. S5. Validation of purified recombinant GST-fusion proteins and SOX2 truncation proteins, and expression of ABCF1 proteins in ABCF1 knockdown ES cells. (A) SDS-PAGE and Coomassie staining of purified GST-fusion proteins containing N-terminal domain of ABCF1 (1-302), GCN20 (1-197), and SREBP1a (1-50). **(B)** SDS-PAGE and Coomassie staining of purified FLAG-tagged SOX2 full-length (FL), high-mobility group (HMG) DNA-binding domain, and activation domain (AD) of SOX2. **(C)** Expression of control RFP and V5-tagged human ABCF1 proteins described in Fig. 5G is confirmed by western blotting using antibodies against RFP and V5. Filled and white arrowheads denote V5-tagged full-length (FL) and LCD-truncated (NBDs) human ABCF1, respectively. ACTB is used as loading control.

Fig. S6, related to Fig. 6

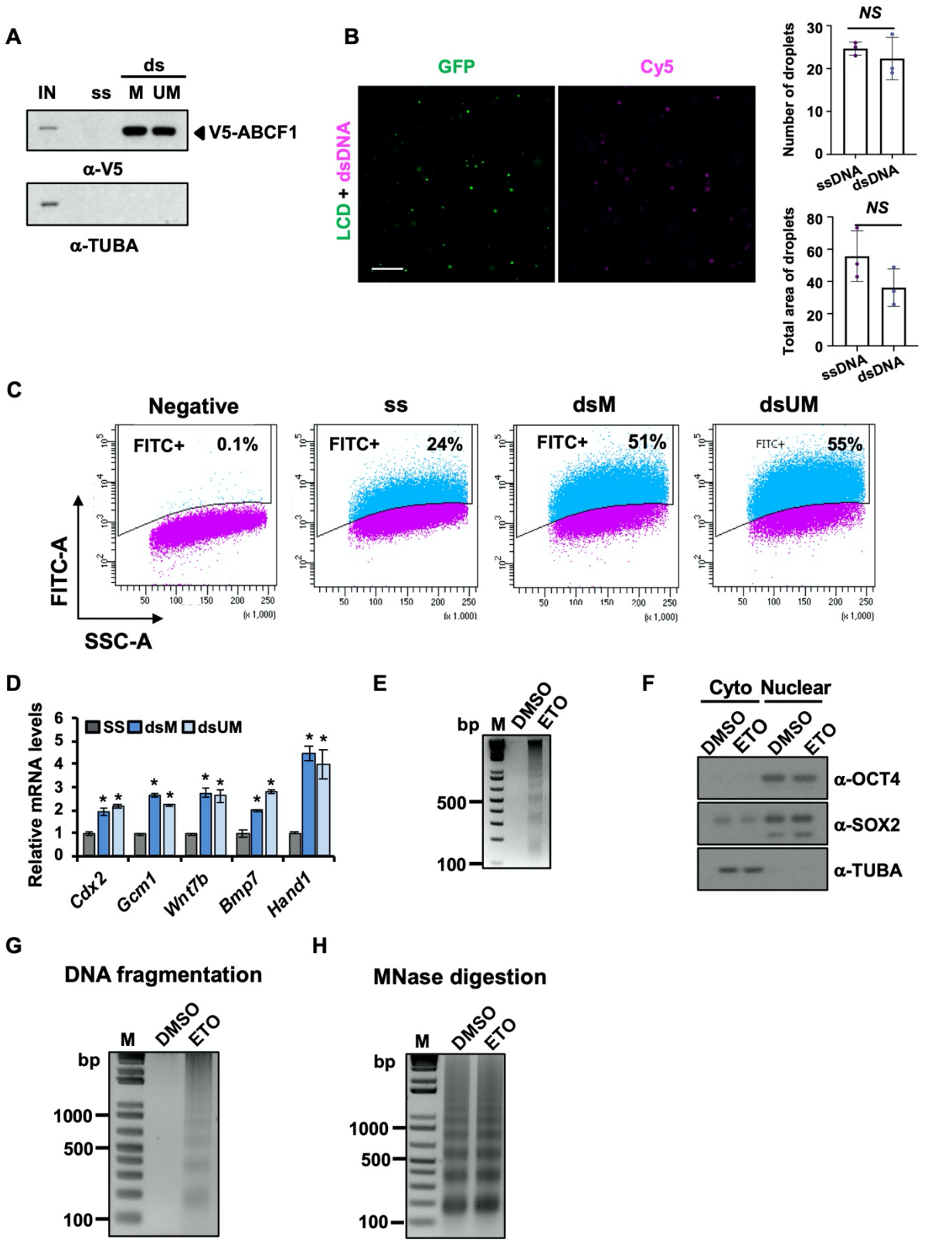

Fig. S6 cont'd, related to Fig. 6

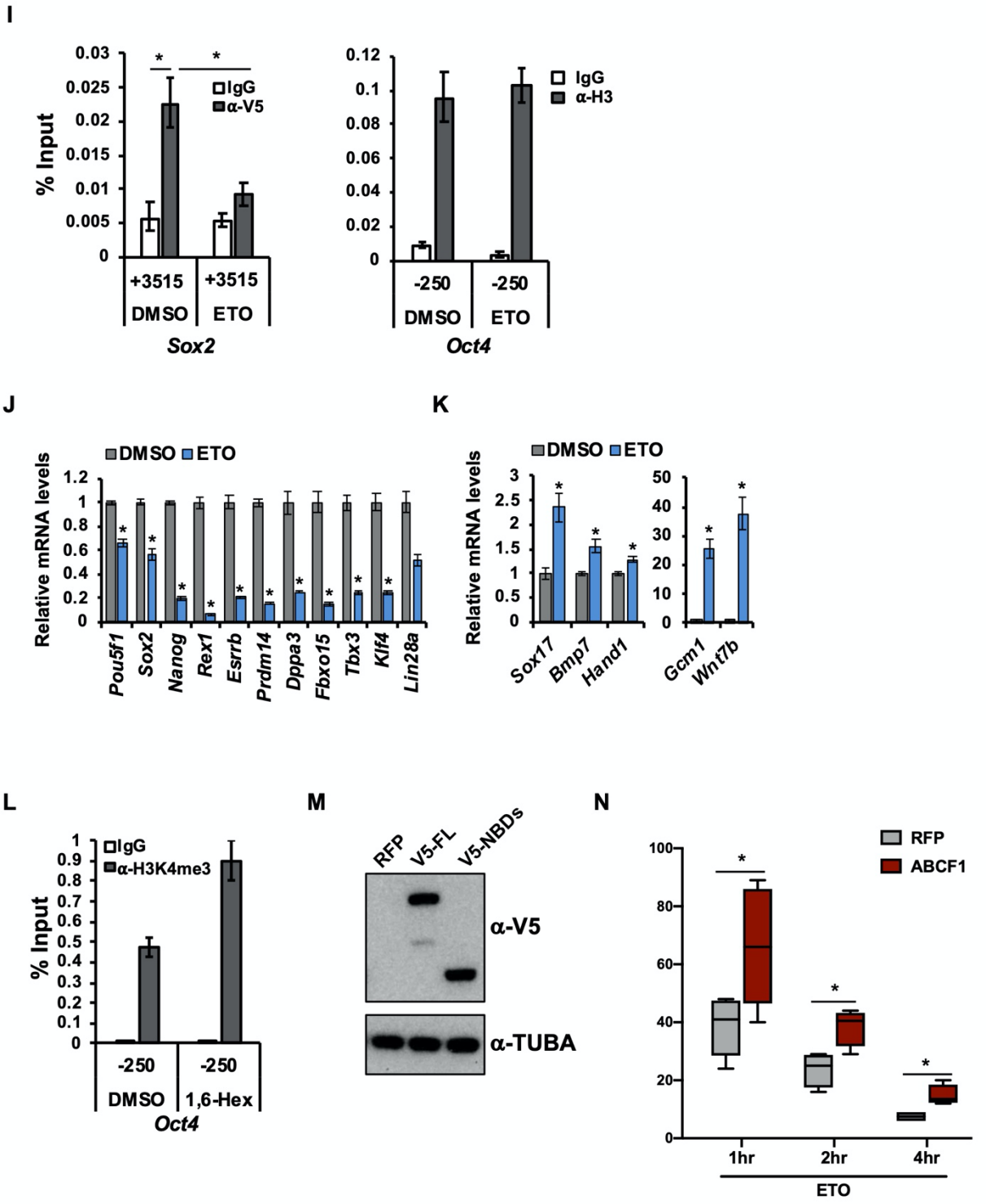

Fig. S6. Intracellular DNAs modulate phase separation of ABCF1 and pluripotency gene expressions. (A) V5-ABCF1 knock-in (KI) D3 mouse ES cell WCEs are incubated with ss, ds-M, or ds-UM as described in Fig. 6A. Input WCEs (IN) and streptavidin-beads captured, DNA-bound ABCF1 proteins are analyzed by western blotting. α-tubulin (TUBA) is used as control for binding specificity. **(B)** Fluorescence microscopy showing that ABCF1 LCD droplets colocalize with Cy5-dsDNA (12.5 nM). Number and size of droplets in the presence of ss and dsDNA are quantified by ImageJ (right). **(C)** Single cell suspensions of D3 mouse ES cells transfected with 5' 6-carboxyfluorescein (6-FAM) labeled ss, ds-M, or ds-UM on the sense strand are analyzed by flow cytometry. 6-FAM-positive cells (blue) are gated by comparing to untransfected cells (Negative). 6-FAM-positive cells are purified for further analyses. **(D)** Mouse ES cells transfected with 5' 6-carboxyfluorescein (6-FAM) labeled ss, ds-M, or ds-UM on the sense strand are enriched by FACS. Expression levels of differentiation genes are analyzed by qPCR, normalized to *Actb*, and expressed as relative units. **(E)** Genomic DNA purified from nuclear extracts prepared from DMSO and etoposide-treated (ETO, 20 μM) V5-ABCF1 KI mouse ES cells were analyzed on agarose gel and stained with ethidium bromide. **(F)** Western blot analysis of nuclear and cytoplasmic fractions prepared from V5-ABCF1 KI mouse ES cells using antibodies against OCT4, SOX2, and TUBA. Note that OCT4 and SOX2 levels remain unchanged after ETO treatment. Effective nuclear-cytoplasmic fractionation is demonstrated by enrichment of OCT4 and SOX2 in nuclear extracts and TUBA in cytoplasmic fraction. **(G)** DNAs purified from DMSO or ETO-treated (80 μM) V5-ABCF1 KI mouse ES cells grown in 2i/LIF medium are separated on an agarose gel and visualized by ethidium bromide staining. **(H)** Crosslinked nuclear chromatin from DMSO or ETO-treated (80 μM) mouse ES cells are fragmented by MNase digestion. Digested DNAs are purified, separated on an agarose gel, and stained with ethidium bromide. Both DMSO and ETO-treated chromatins are digested to a similar degree. **(I)** MNase-ChIP of ABCF1 and H3 in DMSO and ETO-treated (80 μM) V5-ABCF1 KI mouse ES cells. Enrichment of ABCF1 and H3 at *Sox2* and *Oct4* gene promoters, respectively, is analyzed by qPCR as in Fig. 2G. **(J)** Expression level of pluripotency genes in DMSO or ETO-treated (80 μM) V5-ABCF1 KI mouse ES cells are analyzed by qPCR, normalized to *Actb*. mRNA level in each sample is expressed relative to its respective control set as 1. **(K)** Expression level of lineage specific genes in DMSO or ETO-treated (80 μM) V5-ABCF1 KI mouse ES cells are analyzed as in J. **(L)** MNase-ChIP of H3K4me3 in DMSO and 1,6-hexanediol (1,6-Hex)-treated (1.5%) V5-ABCF1 KI mouse ES cells. Enrichment of H3K4me3 at *Oct4* gene promoter is analyzed by qPCR as in I. **(M)** Western blot analysis of WCEs prepared from wild-type D3 mouse ES cells stably expressing V5-tagged full-length (V5-FL) and LCD-truncated (V5-NBDs) ABCF1. D3 cells stably expressing untagged RFP is used as control. TUBA is used as loading control. **(N)** Colony formation assays in control and ABCF1 gain-of-function D3 mouse ES cells. 200 mouse ES cells stably expressing RFP or V5-ABCF1 are plated on 24-well plates, treated with ETO (1 μM) for indicated period of time (hr), and let recover for 6 days before staining for AP activity. AP-positive colonies are counted. Error bars present SEM. $n = 3$. (*) $P < 0.05$, calculated by two-sided Student's t-test.

SUPPLEMENTARY TABLES

Table S1. sgRNA and ss donor oligo sequences for generation endogenously V5-tagged ABCF1 knock-in mouse ES cell line

Name	Sequence
sgRNA-F	CACCGGGCTCGGGCGGCTGTTGCTT
sgRNA-R	AAACAAGCAACAGCCGCCCGAGCCC
ss donor oligo	CCCACATCGCCGGAAGCGGAAACGGCGCCGGGCGCCGCTGGAGGACTCTT AACTGCCGCCGCGATGGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGAT TCTACGGGATCATCTGGTCCGAAGGGTCCCAAGCAACAGCCGCCCGAGCCC GAGTGGATCGGGGACGGCGAGGGCACGAGCCCCGCGGGTGAGTGGGG

Table S2. Primers used for ChIP or MNase ChIP-qPCR analysis

Name	Forward	Reverse
Oct4-250	AGCAACTGGTTTGTGAGGTGTCCGGTG AC	TCCCCAATCCCACCCTCTAGCCTTGAC
Oct4-2000	GGAACTGGGTGTGGGGAGGTTGTA	AGCAGATTAAGGAAGGGCTAGGACGAG AG
Sox2-3581	CCCTGTTCCAAGTCTCTTTCTG	GATTTCAATCCAACACCATCATAG
Sox2+3513	TTTTCGTTTTTAGGGTAAGGTACTGG	CGTGAATAATCCTATATGCATCACAAT
Nanog-952	GGCAAACTTTGAACTTGGGATGTGGAA ATA	CTCAGCCGTCTAAGCAATGGAAGAAGA AAT
Nanog-199	GGGTCACCTTACAGCTTCTTTTGCATTA	GGCTCAAGGCGATAGATTTAAAGGGTA G
Lefty1-10kb	TAGCCAGCAGACAGGACAAA	GTTCCACCTACCCTCATCCC
Trim28-8890	GCTCAGGTCCTACTCTCTCC	CCCAGGGCTACAGGAATGTT
Actb-163	CATGGTGTCCGTTCTGAGTGATC	ACAGCTTCTTTGCAGCTCCTTCG
Dhfr-304	CTATTACCGCTGGTCGGGTA	GGAGGCCTTATGGAGAGCTT
Oct4-132	ATCCGAGCAACTGGTTTGTG	ATCCGAGCAACTGGTTTGTG

Table S3. Primers used for RT-qPCR analysis

Gene	Forward	Reverse
Abcf1	AAGTCCTGACTCGAAAACAGC	GTAGCCAAACGTCACACCG
Nanog	CCTCAGCCTCCAGCAGATGC	CCGCTTGCACTTCATCCTTTG
Klf2	ACCAAGAGCTCGCACCTAAA	GTGGCACTGAAAGGGTCTGT
Klf4	CAGTGGTAAGGTTTCTCGCC	GCCACCCACACTTGTGACTA
Prdm14	GCATCCTGGTTCCCACAGAG	GCTGCAGAACACGCCAAAGTG
Esrrb	TACCTGAACCTGCCGATTTC	CCCAGTTGATGAGGAACACA
Rex1	TGTCCTCAGGCTGGGTAGTC	TGATTTTCCGACGTATGCAAA
Trim28	CTTAGCAGGTCCTGTGGAGC	TGTGTGGAGTGCTCTGAACC
Stra8	GACTGACTCGTCAGGGCTTC	CCATCATCACTGGGTTGGTT
Tfcp2l1	AGGTGCTGACCTCCTGAAGA	GTTTTGCTCCAGCTCCTGAC
Tbx3	GAAACAGAATTCATCGCCGT	TGCAGTGTGAGCTGCTTTCT
Fbxo15	TAGATTCTTGGACTTCCGTTCA	ACCAAGGTCACCGCATCCAA
Nr5a2	GTTGAGTGGGCCAGGAGTAG	GATTGTCCCTTCCTTCCCAT
Dppa3	GGACCCTGAAACTCCTCAGA	CGGGGTTTAGGGTTAGCTTT
Pou5f1(Oct4)	GAAGCAGAAGAGGATCACCTTG	TTCTTAAGGCTGAGCTGCAAG
Sall4	GAAGCCCCAGCACATCAAC	CTGAGGCTTCATCGCAGTT
Sox2	CCGGCACGGCCATTAAC	CCCTCCCAATTCCCTTGTAT
Zic2	CAAGGTCCGGGTGCTTACC	ATTAAAGGGAGGCCCCGAATA
Lin28a	GGCATCTGTAAGTGGTTCAACG	CCCTCCTTGAGGCTTCGGA
Fgf4	GGGAGGCTACAGACAGCAAG	CTGTGAGCCACCAGACAGAA
Fgf5	TTGCGACCCAGGAGCTTAAT	CTACGCCTCTTTATTGCAGCAT
Otx2	CTTCATGAGGGAAGAGGTGG	CTGACCTCCATTCTGCTGCT
Nes	TCCCTTAGTCTGGAAGTGGCTA	GGTGTCTGCAAGCGAGAGTT
Sox1	GTGACATCTGCCCCCATC	GAGGCCAGTCTGGTGTCAG
Pou3f1(Oct6)	GCCACTTTCTCAAGTGTCCC	TCATGCGCTTCTCCTTCTG
T	CTCTAATGTCCTCCCTTGTTGCC	TGCAGATTGTCTTTGGCTACTTTG

Gsc	ATGCTGCCCTACATGAACGT	CAGTCCTGGGCCTGTACATT
Bmp7	GTGGTATCGAGGGTGGAAGA	ACAAGGCCGTCTTCAGTACC
Gata6	TTGCTCCGGTAACAGCAGTG	GTGGTCGCTTGTGTAGAAGGA
Sox17	CAGTATCTGCCCTTTGTGTA	GCAATAGTAGACCGCTGAG
Cdx2	AGGCTGAGCCATGAGGAGTA	CGAGGTCCATAATTCCACTCA
Hand1	CACCAAGCTCTCCAAGATCA	GCGCCCTTTAATCCTCTTCT
Gcm1	TGTCGTCCGAGCTGTAGATG	AACGATGTGAAACTGCCTCA
Wnt7B	CCAGGCCAGGAATCTTGTT	ACGTGTTTCTCTGCTTTGGC
Actb	AGAGGGAAATCGTGCGTGAC	CAATAGTGATGACCTGGCCGT
Tubb	GATCGGTGCTAAGTTCTGGGA	AGGGACATACTTGCCACCTGT
hABCF1	TCAGTGCCAACCAGTGATGAG	CGGCCTTATTGATCCGATTCTT
hACTB	AGAGCTACGAGCTGCCTGAC	GGATGCCACAGGACTCCA

Table S4. DNA oligos used for nucleofections and DNA pull-down assays

Name	Forward	Reverse
ss	ACTTGTCGCTAGTTTAATGCCGATTATTATGCAAATTGTTACTGCGGTTATGCAAATTGGTGCCGCGTTAATGCCGATGGTTATGCAGTTTATTAGCTTT	
ds-M	ACTTGTCGCTAGTTTAATGCCGATTATTATGCAAATTGTTACTGCGGTTATGCAAATTGGTGCCGCGTTAATGCCGATGGTTATGCAGTTTATTAGCTTT	AAAGCTAATAAACTGCATAACCATCGGCATTAACGCGGCACCAATTTGCATAACCGCAGTAACAATTTGCATAATAATCGGCATTAAACTAGCGACAAGT
ds-UM	AAATAACTGGTCGCTTATTTGATGTCCGCTTTATTGATGGCGAGCAGGAAGAAAACTTTGAATACACTGTGATTAAACCAAATCCAGCGTTAGATGAAGA	TCTTCATCTAACGCTGGATTTGGTTTAATCACAGTGTATTCAAAGTTTTCTTCCTGCTCGCCATCAATAAAGCGGACATCAAATAAGCGACCAGTTATTT

5’ end of ss and 5’ end of the sense strand of ds oligos are labeled with 6-FAM, biotin, or Cy-5 for nucleofections, DNA pull-down, or in vitro droplet assay, respectively.

Phosphorothioate bonds are introduced into the last 5 nucleotides of both 5’ and 3’ ends of the oligos to inhibit exonuclease digestion.