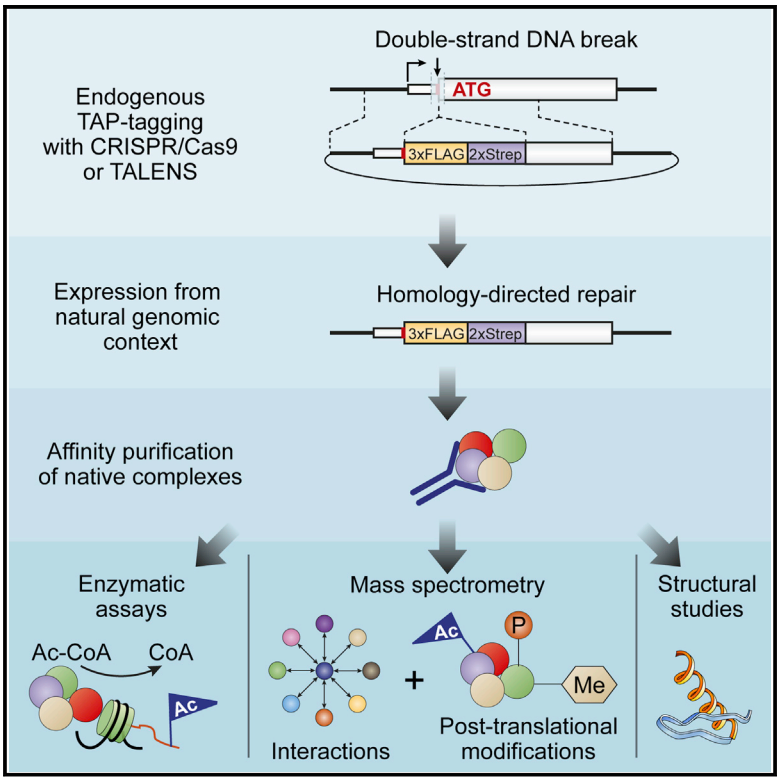
Resource

Cell Reports

A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells

Graphical Abstract



Authors

Mathieu Dalvai, Jeremy Loehr, Karine Jacquet, ..., Pauline Herst, Jacques Côté, Yannick Doyon

Correspondence

yannick.doyon@crchudequebec.ulaval.ca

In Brief

Characterization of multiprotein complexes via affinity purification is optimal when expression of the bait protein is specified by its natural genomic context. Dalvai et al. implement a robust genome-editing-based approach to tag endogenous genes and purify native complexes from human cells to near homogeneity.

Highlights

- Streamlined affinity purification of protein complexes using CRISPR/Cas9 and TALENs
- Purification of native holoenzymes regulated under physiological conditions
- Blueprint for the systematic and unbiased mapping of the human protein interactome

Dalvai et al., 2015, Cell Reports *13*, 621–633
October 20, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2015.09.009

Cell Reports
Resource

OPEN ACCESS
CellPress

A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells

Mathieu Dalvai,[1,2,3] Jeremy Loehr,[1,3] Karine Jacquet,[1,2] Caroline C. Huard,[1] Céline Roques,[1,2] Pauline Herst,[1,2] Jacques Côté,[1,2] and Yannick Doyon[1,*]
[1]Centre Hospitalier Universitaire de Québec Research Center and Faculty of Medicine, Laval University, Quebec City, QC G1V 4G2, Canada
[2]St-Patrick Research Group in Basic Oncology and Laval University Cancer Research Center, Quebec City, QC G1R 3S3, Canada
[3]Co-first author
*Correspondence: yannick.doyon@crchudequebec.ulaval.ca
http://dx.doi.org/10.1016/j.celrep.2015.09.009
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

Conventional affinity purification followed by mass spectrometry (AP-MS) analysis is a broadly applicable method used to decipher molecular interaction networks and infer protein function. However, it is sensitive to perturbations induced by ectopically overexpressed target proteins and does not reflect multilevel physiological regulation in response to diverse stimuli. Here, we developed an interface between genome editing and proteomics to isolate native protein complexes produced from their natural genomic contexts. We used CRISPR/Cas9 and TAL effector nucleases (TALENs) to tag endogenous genes and purified several DNA repair and chromatin-modifying holoenzymes to near homogeneity. We uncovered subunits and interactions among well-characterized complexes and report the isolation of MCM8/9, highlighting the efficiency and robustness of the approach. These methods improve and simplify both small- and large-scale explorations of protein interactions as well as the study of biochemical activities and structure-function relationships.

INTRODUCTION

The cell is composed of a collection of protein machines responsible for the coordinated execution of cellular functions (Alberts, 1998). Deciphering the components and activities of these molecular assemblies is crucial to understand the cellular networks that are perturbed under disease states (Gavin et al., 2011; Rolland et al., 2014). For example, the identification of cancer-driving mutations among a background of passenger mutations can be improved by tying combinations of rare mutations to specific protein complexes (Krogan et al., 2015; Leiserson et al., 2015; Rolland et al., 2014). Besides, the growing list of cancer-related genes often comprises uncharacterized proteins for which focused proteomic studies could help uncover function (Lawrence et al., 2014; Vogelstein et al., 2013).

Affinity purification followed by mass spectrometry (AP-MS) analysis is a powerful approach used to characterize protein-protein interactions and multiprotein complexes that are defined as sets of stably associated proteins isolated under standardized biochemical conditions (Gavin et al., 2011). Landmark studies describing genome-wide identification of complexes in budding yeast generated high-quality datasets by relying on two major technical innovations (Gavin et al., 2006; Krogan et al., 2006). First, the development of standardized tandem affinity purification (TAP) protocols allowed the isolation of complexes by sequential capture and elution using pairs of affinity tags (Rigaut et al., 1999). Second, systematic tagging of open reading frames on the chromosomes via homologous recombination permitted retention of the physiological regulation of gene expression for the bait proteins.

While pioneering studies in *Arabidopsis*, *Drosophila*, and human cells using ectopic expression of tagged proteins have yielded important insights into biological pathways in higher eukaryotes (Behrends et al., 2010; Goudreault et al., 2009; Guruharsha et al., 2011; Hegemann et al., 2011; Hutchins et al., 2010; Huttlin et al., 2015; Marcon et al., 2014; Sardiu et al., 2008; Sowa et al., 2009; Van Leene et al., 2010), there is a need to keep improving these methods in order to reduce perturbations in the stoichiometry of protein interactions and prevent aberrant localization, protein aggregation, dominant-negative effects, and toxicity (Doyon et al., 2011; Ho et al., 2002). For example, chromatin-modifying complexes require meticulous biochemical characterization, since they often exist in different forms with paralogous subunits and only their native assembly can recapitulate their specificity for a given histone residue within chromatin (reviewed in Lalonde et al., 2014). Hence, we developed straightforward methods to streamline the mapping of protein-protein interactions in human cells under settings minimizing deviations from their natural context. We used engineered nucleases such as zinc-finger nucleases (ZFNs), TAL effector nucleases (TALENs), and CRISPR/Cas9 to simplify the generation of cell lines with tailored modifications and enable the expression of tagged proteins from their endogenous loci (Hsu et al., 2014; Joung and Sander, 2013; Sternberg and Doudna, 2015; Urnov et al., 2010). First, we exploited a genomic safe harbor locus to rapidly and reliably generate cell lines expressing bait proteins at near-physiological levels as a surrogate to classical plasmid- or virus-based methods for stable cell line generation. Second, we introduced an affinity tag at the N and C terminus of proteins encoded by endogenous genes in order

to retain the dynamic control of gene expression specified by the native chromosomal context and the natural post-transcriptional regulation mechanisms. Under these conditions, we obtained near-homogenous preparations of native proteins complexes in sufficient amounts to perform biochemical assays and identify their subunits via mass spectrometry analysis. The tagged cell lines were also used for immunoprecipitation of cross-linked chromatin fragments (chromatin immunoprecipitation [ChIP]) to study protein-DNA interactions in vivo under normalized conditions. These tools were portable to numerous proteins and represent a general solution for protein isolation, complex identification, and genome location analysis under physiological conditions. Importantly, the scalability and adaptability of this system will open avenues for the systematic and unbiased mapping of protein-protein networks in a variety of organisms.

RESULTS

Characterization of a Potent Tandem Affinity Tag

An ideal TAP tag leads to high recovery of a fusion protein present at low concentration with minimal background contaminants. It should be functional as N- and C-terminal fusions and its size and amino acid (aa) sequence should have no impact on protein function. In preliminary experiments, we evaluated several tags, including the AC-TAP, SBP, and FLAG-hemagglutinin (FLAG-HA) tags, and settled on a combination of 3xFLAG and 2xSTREP tags, as both tags can be eluted under gentle conditions and yield to the isolation of highly purified material (see below) (Doyon et al., 2006; data not shown). Our version contains 59 aa for a predicted molecular weight of 6 kDa and does not necessitate proteolytic cleavage (Figure 1A).

To fine-tune our purification scheme, we selected the KAT5/TIP60 tumor suppressor protein, the catalytic subunit of the NuA4 (nucleosome acetyltransferase of histone H4) complex that regulates gene expression and promotes DNA repair via homologous recombination and is required for embryonic stem cell self-renewal and differentiation (Steunou et al., 2014). NuA4 is composed of more than 15 subunits ranging from 20 to 400 kDa in size (Cai et al., 2005; Doyon et al., 2004, 2006; Ikura et al., 2000). We first established two independent pools of cells expressing KAT5 fused to a C-terminal TAP tag and performed purifications from whole-cell extracts to determine the proper order of steps required to achieve high yields and purity. We observed that the high binding capacity of the anti-FLAG M2 affinity resin is best used as a first step to recover the maximum number of complexes while the Strep-Tactin resin dramatically increases the purity of the final samples, yielding near-homogenous preparations from low amounts of starting material (Figure S1).

TAP following Nuclease-Driven Gene Addition to the *AAVS1* Genomic Safe Harbor Locus

In order to rapidly generate isogenic cell lines expressing TAP-tagged cDNAs, we used nuclease-driven targeted integration into the human *PPP1R12C* gene, a safe harbor genomic locus known as *AAVS1* that allows stable transgenesis and neutral marking of the cell (DeKelver et al., 2010; Hockemeyer et al., 2009; Lombardo et al., 2011). This system is composed of a nuclease that cleaves the first intron of *PPP1R12C* and a gene-trap vector allowing puromycin selection of targeted cells. We adapted this system to integrate tagged cDNAs and used the moderately active human *PGK1* promoter to achieve slight overexpression conditions (Figure 1A and see below). We targeted two subunits of NuA4 to the *AAVS1* locus and purified the associated complexes. We chose the enhancer of polycomb homolog 1 (EPC1) and the E1A binding protein p400 (EP400) as model proteins. Due to their size (100 and 400 kDa, respectively), we reasoned that it would represent a substantial test for biochemical stability during purification. Likewise, targeting required error-free ZFN-driven addition of their relatively large expression cassettes of 4.3 kb and 11.3 kb.

K562 cells were selected as our model cell line because they can be cultivated, expanded, and transfected with high efficiency as suspension cultures. They are also permissive to genome editing events and tolerate cloning via limiting dilution or in semi-solid media. Moreover, as a designated tier 1 cell line used by all investigators of the Encyclopedia of DNA Elements (ENCODE) project, a myriad of genomic and epigenomic data is available for this cell line (ENCODE Project Consortium, 2012).

We first transfected K562 cells with *AAVS1*-targeting ZFNs and the *EPC1* donor and selected clones in methylcellulose-containing media supplemented with puromycin (Figure 1A). Single-cell-derived colonies were picked after 10 days and expanded, and transgene expression was monitored by western blot. In a typical experiment, more than 90% of the clones expressed the transgene with little variability (see Figure S1 for an example). Nuclear extracts were prepared from an EPC1-tag cell line, as well as from K562 cells expressing only the tag (mock), and subjected to TAP. EPC1 complex subunits separated by SDS-PAGE could be unambiguously identified on silver-stained gels due to the very low protein background observed in the mock purification after double-affinity purification (Figures 1B). Mass spectrometry analysis identified all known components of NuA4, in addition to MBTD1, which was not previously ascertained as a core complex subunit (Cai et al., 2005; Doyon et al., 2004, 2006; Ikura et al., 2000) (Table S1). Next, we performed a reciprocal purification of the complex by repeating the process with *EP400*, coding for the large SWI/SNF2-family ATPase that incorporates histone variant H2A.Z into chromatin (Weber and Henikoff, 2014). We obtained a complex that was indiscernible from the EPC1 assembly (Figures 1C and 1D; Table S1). In contrast to previously published data, the EP400 complex contains the KAT5 catalytic subunit (Fuchs et al., 2001). It appears that overexpression of EP400 that was achieved using retroviral-based gene delivery in the previous study resulted in the purification of a partially assembled complex.

An important characteristic of our system is that the purified fractions can be assayed biochemically in vitro. To determine whether the purified complexes are recovered in sufficient amounts and concentration, we performed histone acetyltransferase (HAT) assays. We observed robust enzymatic activity for both complexes with substrate specificity consistent with published reports for KAT5/NuA4 (Doyon et al., 2004; Ikura et al., 2000) (Figure 1E). Again, these observations contrast with the minimal traces of HAT activity observed for the EP400 complex

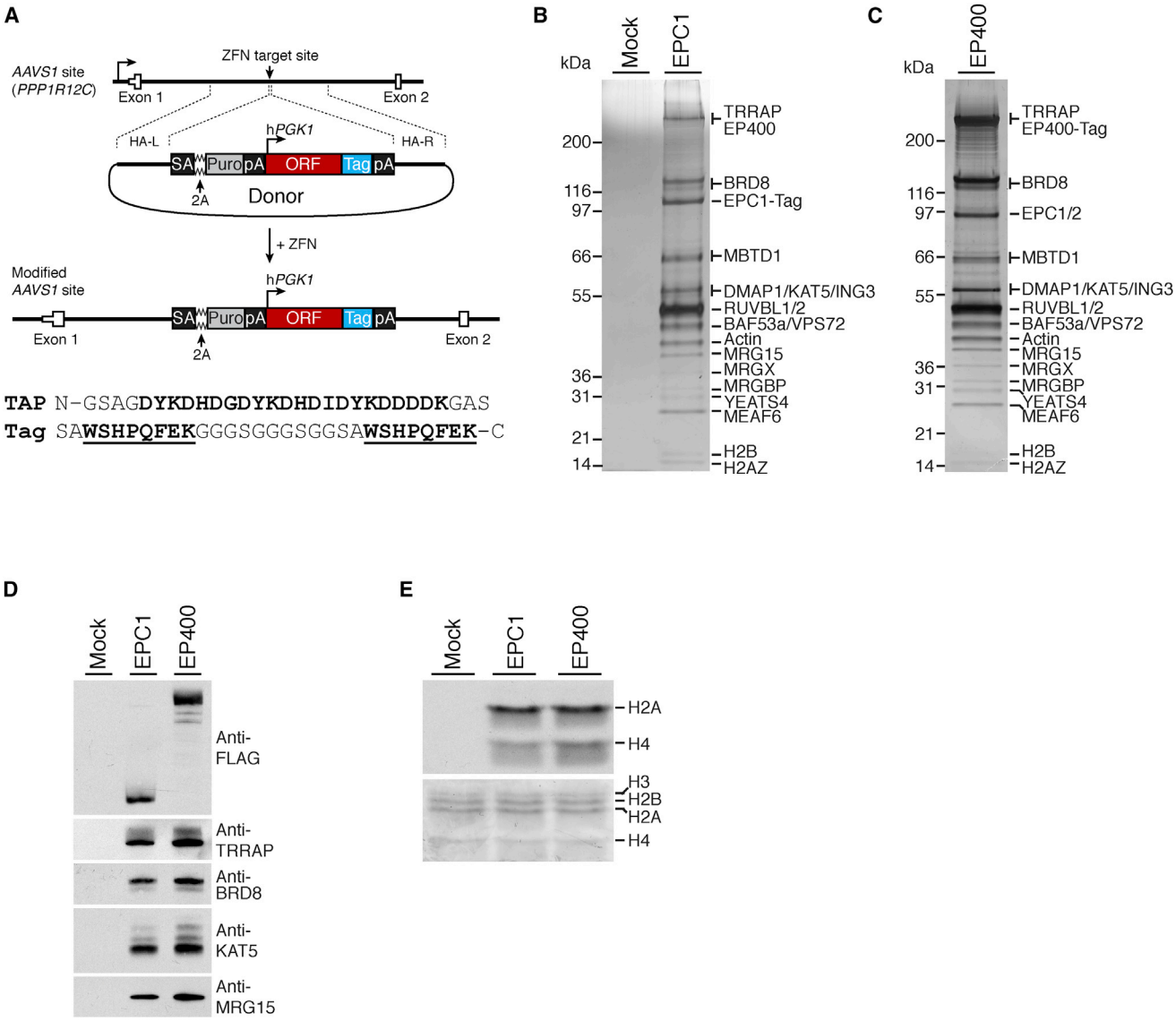


Figure 1. ZFN-Driven Gene Addition to the *AAVS1* Locus Simplifies Tandem Affinity Purification of Multisubunit Protein Complexes

(A) Schematic of the donor construct and of the *AAVS1* locus following cDNA addition. The first two exons of the *PPP1R12C* gene are shown as open boxes. Also annotated are the locations of the splice acceptor site (SA), 2A self-cleaving peptide sequence (2A), puromycin resistance gene (Puro), polyadenylation sequence (pA), human phosphoglycerate kinase 1 promoter (h*PGK1*), and 3xFLAG-2xSTREP tandem affinity tag (Tag); homology arms left and right (HA-L, HA-R) are respectively 800 and 840 bp. Sequence of the TAP tag. The 3xFLAG sequence is in bold, and 2xSTREP is in bold and underlined.

(B) Silver-stained SDS-PAGE showing the purified EPC1 complex. K562 cells expressing the tag (Mock) and a clonal cell line expressing EPC1-tag (EPC1).

(C) Silver-stained SDS-PAGE showing the purified EP400 complex from a clonal cell line expressing EP400 tag (EP400). Proteins were identified from unfractionated protein samples and assigned to specific gel bands based on extensive western blotting analysis.

(D) Western blots of selected NuA4 subunits on purified fractions.

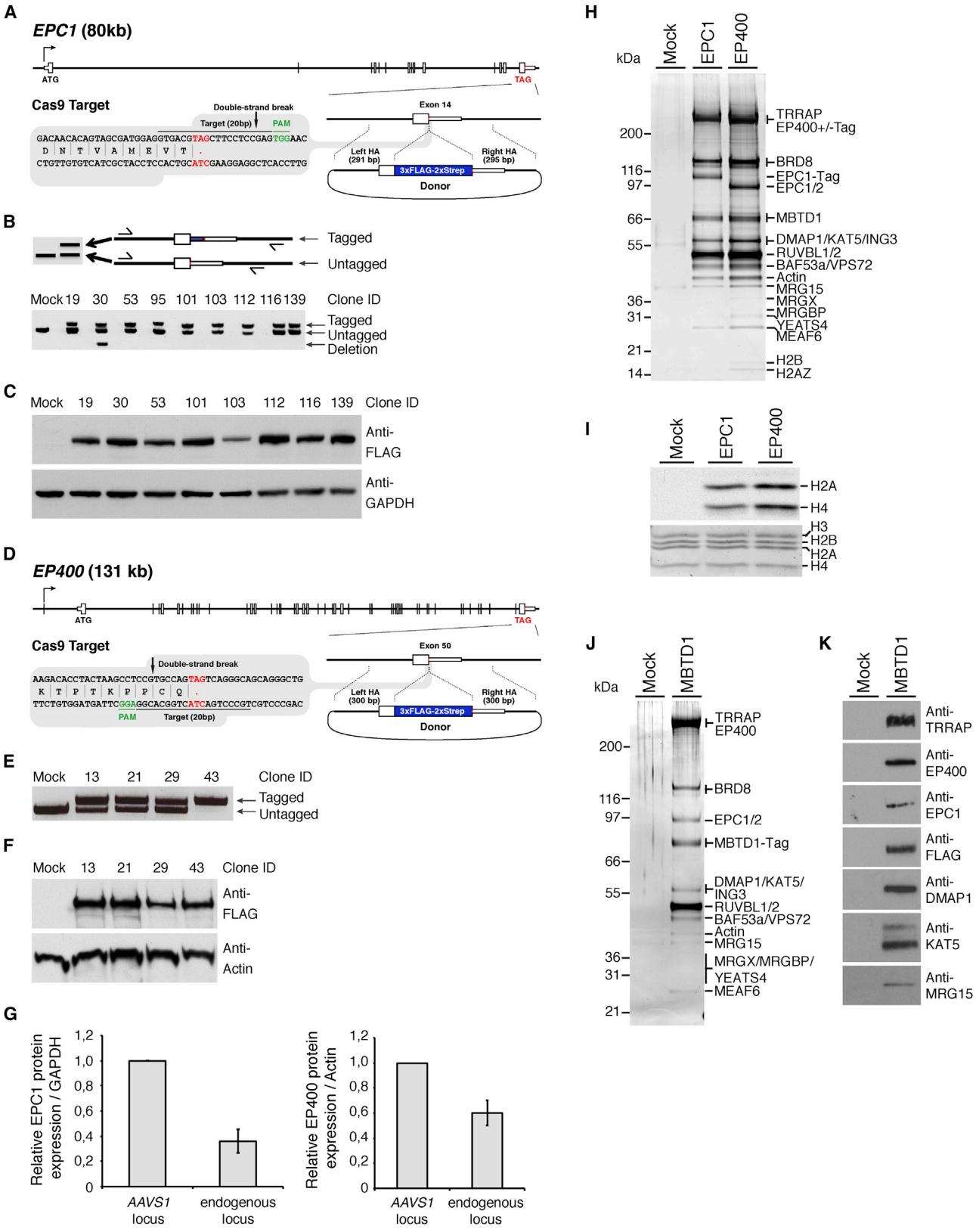
(E) Autoradiogram showing the results of HAT assays to determine the specificity of the EPC1 and EP400 complexes. Coomassie staining was used as a loading control for histones.

See also Figure S1.

and the reported disconnection between these two enzymatic activities (Fuchs et al., 2001; Park et al., 2010; Tyteca et al., 2006). Taken together, these data demonstrate that coupling our TAP approach with gene targeting at *AAVS1* creates a highly efficient surrogate method for the isolation of native protein complexes under near-physiological conditions.

TAP of Native Multisubunit Protein Complexes following CRISPR/Cas9-Driven Tagging of Endogenous Genes

Having demonstrated that our purification strategy worked efficiently and eliminated most background contaminants at the low expression levels obtained by targeting the *AAVS1* locus, we aimed to determine if proteins expressed from their natural



(legend on next page)

chromosomal contexts could be isolated with high specificity as physiologically regulated endogenous complexes. We designed CRISPR/Cas9-based nucleases to cleave DNA at the vicinity of the stop codon of *EPC1* in order to stimulate homology-directed integration of the TAP tag using a donor molecule containing short homology arms (Figures 2A and S2). Following transfection of K562 cells with the nuclease and recombination donor, we readily detected the incorporation of the tag in the pool of cells by PCR (Figure S3). Targeted clones, isolated by limiting dilution in the absence of any selection, were observed at a frequency of 6% (10/165), and accurate gene modification was confirmed by western blot analysis and sequencing (Figures 2B, 2C, and S3). We did not obtain homozygous clones for this gene, but results of our PCR analysis suggest that three copies of the locus are present in K562 cells (Figure S3). Comparison of the EPC1-tag protein levels expressed from the *AAVS1* locus versus the endogenous gene revealed an ~2.5-fold overexpression in the former case (Figures 2G and S3). We performed side-by-side TAP from an endogenously tagged EPC1 cell line and from a clone targeted at *AAVS1* and observed matching banding patterns on silver-stained gels (Figure S3). Mass spectrometry analysis confirmed the complete coverage of the subunits when the complex is purified from the endogenous locus (Table S1). To confirm these observations, we used a similar approach to perform a reciprocal purification of the complex by tagging the endogenous *EP400* gene at its C terminus (Figures 2D and S2). The efficiency of targeting in this case was 21% (13/61, including 2 homozygous clones). For the purification, a homozygous clone expressing tagged EP400 was selected (Figures 2E, 2F, and S3). This clone expressed ~2-fold less protein than the one used in the *AAVS1* purification (Figures 2G and S3). Both complexes could be purified to near homogeneity and displayed strong H4/H2A-specific HAT activity toward nucleosomes (Figures 2H and 2I; Table S1). Given that MBTD1 was reproducibly detected in these preparations, we performed reciprocal tagging to confirm its stable association with NuA4. CRISPR/Cas9-mediated integration of the TAP tag at its C terminus was efficient reaching 26% (20/78, including 6 homozygous clones) of positive clones (Figures S2; data not shown). Purification of MBTD1 using a homozygote clone confirmed its association with NuA4 (Figures 2J and 2K). Thus, our approach led to the most complete characterization of native NuA4 components to date. Not only we were able to conclusively demonstrate that EPC1 and EP400 exclusively associate with the complex, but we also identified MBTD1 as novel subunit. These data demonstrate that highly efficient purification of native protein complexes is achievable with minimal chromosomal sequence disruption and preservation of endogenous physiological regulation of the bait protein.

TALEN-Enabled Purification of the Endogenous Polycomb Repressive Complex 2

To test the generality of the TAP procedure, we undertook the purification of the enhancer of zeste homolog 2 (EZH2), the catalytic subunit of the polycomb repressive complex 2 (PRC2) responsible for the di- and tri-methylation of histone H3 at lysine 27 (H3K27me2/3), a histone mark that correlates with silent or poorly transcribed genomic regions (Margueron and Reinberg, 2011). EZH2 is a critical regulator of development and controls stem cells pluripotency and differentiation, and its deregulation is at the center of novel therapeutic strategies for a variety of cancers (Plass et al., 2013). First, EZH2 with an N-terminal TAP tag was targeted to the *AAVS1* locus (Figure S1). Next, we replaced the natural ATG of EZH2 with the TAP tag via homology-directed repair at the chromosomal locus using a pair of TAL effector nucleases (TALENs) (Figures 3A and S4) (Reyon et al., 2012). 21% (20/96, including 6 homozygous clones) of screened clones had a tagged allele as assessed by PCR, western blotting, and sequencing (Figures 3B, 3C, and S4). A homozygote clone expressing exclusively the tagged EZH2 protein was selected for TAP. Interestingly, in this context, the endogenous locus drives higher protein expression levels as compared to the ectopically expressed protein from the *AAVS1* locus (Figure S4). The complexes were isolated from nuclear extracts prepared from both cell lines and analyzed by SDS-PAGE and silver staining. Both preparations appeared similar, as a specific protein band pattern could be clearly identified from the gels when compared to the mock fractions (Figures 3D and 3E). Co-purified proteins were

Figure 2. CRISPR/Cas9-Driven Tagging of NuA4 Subunits Enables Reciprocal Tandem Affinity Purification of the Endogenous Native Complexes

(A) Schematic of the *EPC1* locus, Cas9 target site, and donor construct used to insert the TAP tag to the C terminus of the EPC1 protein. Annotated are the positions of the stop codon (TAG), the protospacer adjacent motif (PAM) that specifies the cleavage site, and homology arms left and right (HA-L, HA-R).

(B) Schematic and results of a PCR-based assay (out-out PCR) to detect targeted integration (TI) of the tag sequence in single-cell-derived K562 clones obtained by limiting dilution. Primers are located outside of the homology arms and are designed to yield a longer PCR product if the tag is inserted.

(C) Western blots showing EPC1-tag protein expression in K562 clones. Mock indicates cells treated with donor and Cas9 nuclease in the absence of gRNA. The FLAG M2 antibody was used to detect EPC1, and the GAPDH antibody was used as a loading control.

(D) Targeting scheme for *EP400*, depicted as in (A).

(E) Same as (B), but for *EP400*.

(F) EP400 expression monitored as in (C).

(G) 2-fold serial dilutions of whole-cell extracts prepared from *AAVS1*-EPC1/EP400 and CRISPR-EPC1/EP400 cell lines were analyzed by western blot to determine the relative expression of EPC1-tag proteins. Error bars indicate the SD from two independently performed experiments.

(H) Silver-stained SDS-PAGE showing the purified EPC1 and EP400 complexes. Wild-type K562 cells (Mock) and clonal cell lines expressing EPC1-tag (#112) and EP400-tag (#43) from their endogenous loci.

(I) Autoradiogram showing the results of HAT assays to determine the specificity of the complexes. Coomassie staining was used as a loading control for histones.

(J) Silver-stained SDS-PAGE showing the purified MBTD1 complex.

(K) Western blots of selected NuA4 subunits on purified fractions. Wild-type K562 cells (Mock). Proteins were identified from unfractionated protein samples and assigned to specific gel bands based on extensive western blotting analysis.

See also Figures S2 and S3.

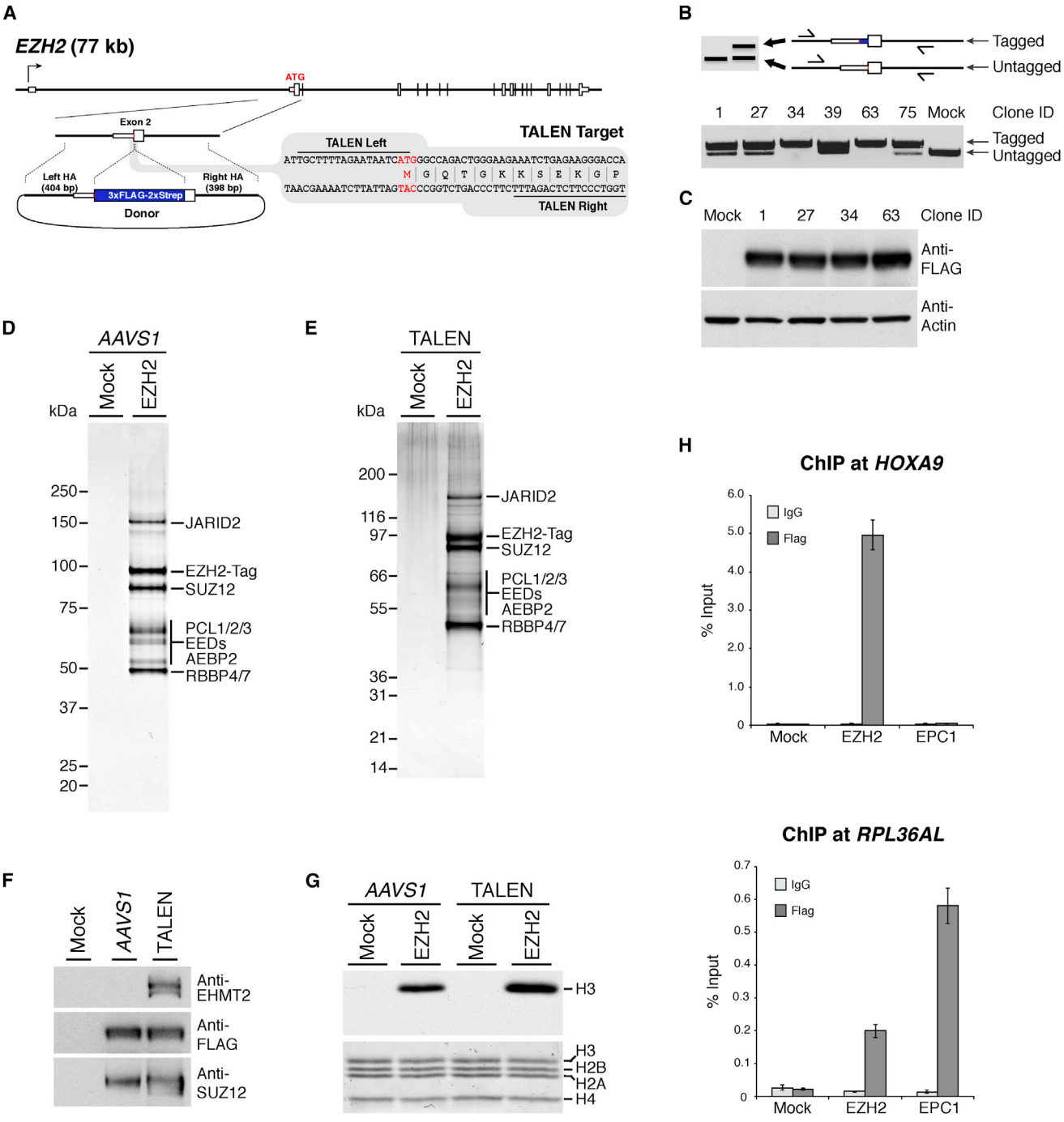


Figure 3. Tandem Affinity Purification of the Native PRC2 Protein Complex

(A) Schematic of the *EZH2* locus, TALEN target site, and donor construct used to insert an affinity tag to the N terminus of the EZH2 protein. Annotated are the positions of the start codon (ATG), the left and right TALEN target sites, and homology arms left and right (HA-L, HA-R).

(B) Schematic and results of a PCR-based assay (out-out PCR) to detect targeted integration (TI) of the tag sequence in single-cell-derived K562 clones obtained by limiting dilution following TALEN-driven gene targeting. Primers are located outside of the homology arms and are designed to yield a longer PCR product if the tag is inserted.

(C) Western blots showing tag-EZH2 protein expression in K562 clones. Mock indicates cells treated with donor only. The FLAG M2 antibody was used to detect EZH2 and the actin antibody was used as a loading control.

(D) Silver-stained SDS-PAGE showing the purified EZH2 complex expressed from the *AAVS1* locus. K562 cells expressing the tag (Mock) and a clonal cell line ectopically expressing tag-EZH2.

(legend continued on next page)

analyzed by mass spectrometry, leading to the unambiguous identification of all known PRC2 components (Table S2) (Margueron and Reinberg, 2011). Substoichiometric interactions between PRC2 and the EHMT1/2 complex, a H3K9 mono- and di-methyltransferase, were also detected in the TALEN-derived clone (Figure 3F; Table S2). The interaction is not mediated by co-localization/co-purification on DNA, since it is not sensitive to treatment with benzonase (Figure S4). This association shows a direct physical link that supports the concept of an H3K27me/H3K9me switch for the long-term repression of differentiation genes (Mozzetta et al., 2014). It also indicates that biallelic targeting of endogenous genes, leading to exclusive expression of the tagged bait protein in the cell, can reveal labile and important interactions. Lastly, the purified fractions contained robust enzymatic activity as determined by histone methyltransferase (HMT) assays (Figure 3G).

Interactions between Endogenously Tagged Chromatin-Modifying Complexes and Genomic DNA Regions

ChIP is a powerful technique used to investigate the interaction between proteins and DNA in the cell, permitting the identification of genomic regions bound by a protein of interest. We tested whether the endogenously tagged cell lines could be used for ChIP using the well-characterized anti-FLAG M2 antibody and analyzed the occupancy of tag-EZH2 and EPC1-tag to the *HOXA9* locus and *RPL36AL* ribosomal protein gene. We observed strong binding of EZH2 at *HOXA9,* a bona fide PRC2 target, and marginal enrichment at *RPL36AL* (Figure 3H). In contrast, EPC1 was specifically recruited to the highly transcribed *RPL36AL* gene, as expected for NuA4 (Figure 3H). The effectiveness of ChIP assays and genome-wide location analysis currently relies on validated antibodies to target proteins, which vary greatly. Thus, the use of a common tag added on endogenous proteins will greatly enhance reproducibility and, importantly, enable accurate comparison between different sets of factors in distinct growth conditions (Bradbury and Plückthun, 2015; Venters et al., 2011).

Purification of Endogenous DNA Repair Complexes

To further exemplify the robustness of our strategy, we extended our studies to proteins involved in DNA repair. The Fanconi anemia (FA) core complex is a multisubunit ubiquitin ligase that initiates DNA repair of interstrand crosslinks (Walden and Deans, 2014). Based on the work of several groups, its subunit composition has been well established and the core was defined as containing FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FAAP20, and FAAP100 (Walden and Deans, 2014). However, some variations in its architecture are observed, suggesting that the complex is composed of submodules (Ali et al., 2012; Rajendra et al., 2014; Walden and Deans, 2014). We attempted to purify the FA core complex through the FANCF protein, as it is suggested to act as a scaffolding subunit. We tagged the N terminus of FANCF using the CRISPR/Cas9 double-nicking strategy and selected a homozygote clone expressing exclusively the tagged FANCF protein for TAP (Ran et al., 2013) (Figures 4A–4C and S5). In this experiment, 5% (5/96, including 2 homozygous clones) of screened clones had a tagged allele. Our approach successfully resulted in the purification of a core complex that is highly similar to previously reported assemblies (Figure 4D; Table S3) (Ali et al., 2012; Rajendra et al., 2014). In addition, subunits of the anchor complex (FANCM, FAAP24, and FAAP16) were detected with the core complex in cells under normal cycling conditions (Walden and Deans, 2014) (Table S3). Next, we attempted to purify minichromosome maintenance complex component 8 (MCM8), a protein evolutionarily related to members of the MCM2–7 replicative helicase family (Maiorano et al., 2006). MCM8 has previously been shown to promote homologous recombination via its interaction with MCM9, but it is unknown whether this association is exclusive (Lutzmann et al., 2012; Nishimura et al., 2012; Park et al., 2013). C-terminal tagging of MCM8 clearly revealed the strict heterodimeric nature of the complex (Figures 5 and S5; Table S4). Taken together, these findings establish the value of using nuclease-mediated endogenous gene tagging to refine the composition and enzymatic activities of protein complexes and highlight the robustness of our TAP strategy.

Toward High-Throughput Genome-Scale Purification of Native Endogenous Protein Complexes

The combined robustness of the endogenous targeting and purification methods lead us to test whether it was possible to purify the native protein complexes from the pool of cells shortly after transfection. If such a scheme were successful, then it would offer the unique opportunity to easily and rapidly tag all human genes at their genomic location and isolate complexes. We transfected cells with the *EP400* CRISPR/Cas9 nuclease and donor vectors and expanded cells for 7 days before harvesting them for TAP purification (Figure 6A). Over the course of the experiment, we monitored the presence of targeted alleles in the cell population via PCR (Figures 6B and 6C). The EP400 complex could be efficiently purified from whole-cell extracts with all its associated subunits (Figure 6D). Then, we confirmed this observation with the EZH2 reagents and obtained an almost pure complex 10 days post-transfection (Figures 6E–6G**)**. It is critical to mention that no selection or cell sorting was used to enrich for tagged cells in the population before attempting the purifications. Thus, the requirement to generate genome-edited cell clones can be bypassed such that protein complexes

(E) Silver-stained SDS-PAGE showing the purified EZH2 complex expressed from the endogenous locus. Wild-type K562 cells (Mock) and a clonal cell line expressing tag-EZH2 (#63). Proteins were identified from unfractionated protein samples and assigned to specific gel bands based on western blotting analysis and predicted molecular weights.

(F) Western blots of selected PRC2 subunits on purified fractions shown in (D) and (E).

(G) Autoradiogram showing the results of an HMT assay used to determine the specificity of the EZH2 complex. Coomassie staining was used as a loading control for histones.

(H) Binding of endogenously tagged EZH2 and EPC1 to target gene promoters as determined by ChIP-qPCR analysis.

Values are expressed as % of input chromatin. Error bars indicate the SD from two independently performed experiments. See also Figure S4.

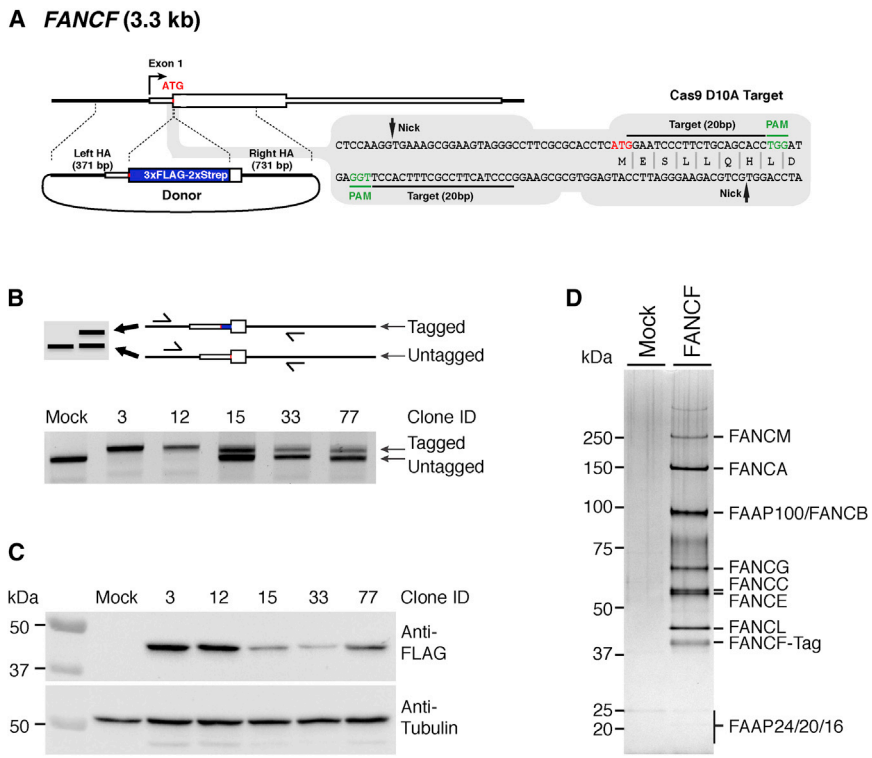


Figure 4. Tandem Affinity Purification of Endogenously Tagged Fanconi Anemia Core Complex

(A) Strategy for CRISPR/Cas9-driven insertion of the TAP tag to the N terminus of the FANCF protein. Schematic of the *FANCF* locus, Cas9 double nickase target sites, and donor construct. The positions of the start codon (ATG) and protospacer adjacent motifs (PAMs) that specify nicking sites are shown.

(B) Schematic and results of a PCR-based assay (out-out PCR) to detect targeted integration (TI) of the tag sequence in single-cell-derived clones obtained by limiting dilution following CRISPR/Cas9-driven gene targeting. Primers are located outside of the homology arms and are designed to yield a longer PCR product if the tag is inserted.

(C) Western blots showing tag-FANCF protein expression in K562 clones. Mock indicates cells treated with GFP expression vector. The FLAG M2 antibody was used to detect FANCF, and the tubulin antibody was used as a loading control.

(D) Silver-stained SDS-PAGE showing the purified FANCF complex. Wild-type K562 cells (Mock) and clonal cell line expressing tag-FANCF (#12). Proteins were identified from unfractionated protein samples and assigned to specific gel bands based predicted molecular weights.

See also Figure S5.

can be directly isolated from gene-modified cell pools. These data constitute a blueprint for high-throughput genome-scale charting of physiological protein-protein interactions in human cells.

DISCUSSION

Uses and Limitations of the *AAVS1* Safe Harbor Ectopic Expression System

The methods described here offer unique advantages over traditional approaches used to study protein complexes in metazoan cells. The purification strategy allows the identification of associated proteins by mass spectrometry, and the purity and yields of the preparations are sufficient to perform enzymatic and mechanistic studies in vitro. It is worth noting that gene targeting at *AAVS1* can be performed in any human cell line since the nuclease target site is naturally occurring, as compared to the Flp-In system, which requires prior integration of a Flp recombination target (FRT) site into the genome. Biosafety issues associated with the use of lentiviral vectors (Biosafety level 2 containment) are also avoided. Thus, it offers greater flexibility of use than traditional ectopic expression systems used for proteomic analysis. Note that ZFNs, TALENs, and CRISPR/Cas9 nucleases targeting *AAVS1* can be used interchangeably (Figure S1) (Hockemeyer et al., 2009; Hockemeyer et al., 2011; Mali et al., 2013). The generation of cell lines expressing near-physiological levels of bait proteins using the *AAVS1* targeting system is straightforward and can be used to rapidly isolate a protein of interest and perform reciprocal purifications in order to confirm the stable association of a subunit with a protein complex. It can also be used to generate panels of variants under isogenic settings and for the analysis of specific splicing isoforms. As an example, the long and short isoforms of JADE1 regulate the presence of ING4/5 tumor suppressor proteins in the KAT7/HBO1 histone acetyltransferase complex (Figure S6) (Saksouk et al., 2009). One can also contemplate using this system to test the function of several mutants in the background of a cellular gene knockout. This could be especially useful for the functional study of essential genes. In addition, it offers a surrogate to study protein fusions resulting from complex genomic rearrangements that cannot currently be modeled via genome editing. As these protein fusions are often "toxic" when expressed in cells, we established a single-vector autoregulated Tet-On expression system permitting tightly controlled inducible expression of target proteins (Figure S6). Of course, this system requires obtaining and subcloning the cDNA of interest, which could be troublesome for very large proteins. However, this does not seem to be a major limitation, as we successfully purified the EP400 complex (400 kDa) and obtained highly purified 350-kDa ataxia telangiectasia mutated (ATM) kinase, a critical activator of the DNA damage response that is notoriously challenging to purify (Figure S6) (Shiloh and Ziv, 2013). Importantly, the use of standard and characterized nucleases can minimize the risk of confounding results due to off-target mutagenesis.

Advantages and Limitations of Endogenous Tagging

Seamless tagging of genes at their natural chromosomal locations preserves the physiological regulation of bait protein

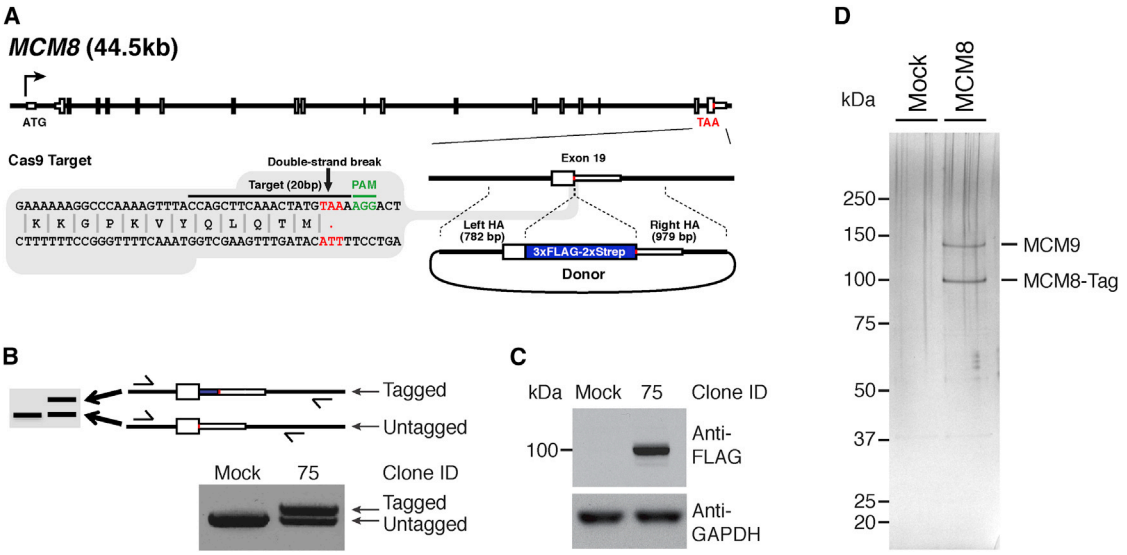


Figure 5. Tandem Affinity Purification of Endogenously Tagged Minichromosome Maintenance Complex Component 8

(A) Strategy for CRISPR/Cas9-driven insertion of the TAP tag to the C terminus of the MCM8 protein. Schematic of the *MCM8* locus, wild-type Cas9 target site, and donor construct. Annotated are the positions of the stop codon (TAA), target site, and protospacer adjacent motif (PAM) that specifies the cleavage site.

(B) Schematic and results of a PCR-based assay (in-out PCR) to detect targeted integration (TI) of the tag sequence in a single-cell-derived clone obtained by limiting dilution following CRISPR/Cas9-driven gene targeting. In this particular case, one primer is located upstream of the left homology arm and one binds the right homology arm to yield a longer PCR product if the tag is inserted.

(C) Western blots showing MCM8-tag protein expression in K562 clone #75. Mock indicates cells treated with donor and Cas9 nuclease in the absence of gRNA. The FLAG M2 antibody was used to detect MCM8, and the GAPDH antibody was used as a loading control.

(D) Silver-stained SDS-PAGE showing the purified MCM8 complex. Wild-type K562 cells (Mock) and a clonal cell line expressing MCM8 tag (#75). Proteins were identified from unfractionated protein samples and assigned to specific gel bands based on predicted molecular weights.

See also Figure S5.

expression and its splicing variants. Since the epitope tag inserted is very small and not linked to a drug resistance gene or other extraneous elements, natural 5′ and 3′ UTRs are minimally perturbed and no sequence is lost. *cis*-acting regulatory elements are maintained within (introns) and outside (promoters/enhancers) transcribed regions. Thus, native regulatory mechanisms of protein expression are retained, and the impact of various stimuli that modulate isoform ratios and interactors can be more precisely studied. Current limitations linked to constitutive overexpression of bait proteins leading to higher rates of false-positive and negative interactions are therefore avoided. As examples, our approach settled the score on the KAT5 and EP400 enzymatic activities, clearly establishing them as part of a single stable macromolecular assembly in vivo. Moreover, this approach led to the identification of MBTD1 as a novel subunit of NuA4. It also demonstrates a direct physical link, in solution, between H3K27 and H3K9 histone methyltransferases, an interaction previously suggested to occur only through colocalization on the genome during development (Alekseyenko et al., 2014). Histone and residue specificity of chromatin-modifying enzymes has been marred on multiple occasions over the years by several conflicting and debated results in the literature. Still today, the specificity of many enzymes is up to interpretation. We feel that our approach using endogenous activities along with native substrates will lead to a more coherent picture helping us to better understand the dynamic nature of the epigenome during development and disease (Lalonde et al., 2014).

The wide availability of research reagents for CRISPR/Cas9 and TALENs greatly facilitates the design and construction of custom reagents. All nuclease-based reagents described here were obtained from the plasmid repository Addgene. We provide detailed examples of donor design (Figures S2, S4, and S5) in order to facilitate the implementation and adaptation of the strategy to user-specific contexts. The use of short homology arms facilitates the construction of donor vectors, as they can be synthesized as DNA fragments smaller than 1 kb in length. For sequences that are AT- or GC-rich, PCR-based amplification of the homology arms might be required. Potential problems can be encountered if highly repetitive elements are found in proximity of the ATG or STOP codons, in which cases the length of the arms should be truncated to avoid the presence of repeats in the donor. However, this is not essential, as the MCM8 donor described in this study contains repetitive elements and successful targeting was achieved. Apart from strict biochemical considerations, one should take into account the structure of the gene, the various protein isoforms produced, and adjacent genetic elements when choosing to tag either the 5′ or 3′ end of genes.

A sensible preoccupation with genome editing techniques is the possibility of off-target mutagenesis. In our experiments, we used TALENs, wild-type Cas9, and Cas9 D10A (dual nickase) interchangeably and did not observe overt toxicity or loss of targeted cells. Continuing progress in the field aiming to increase the precision of genome editing should progressively decrease

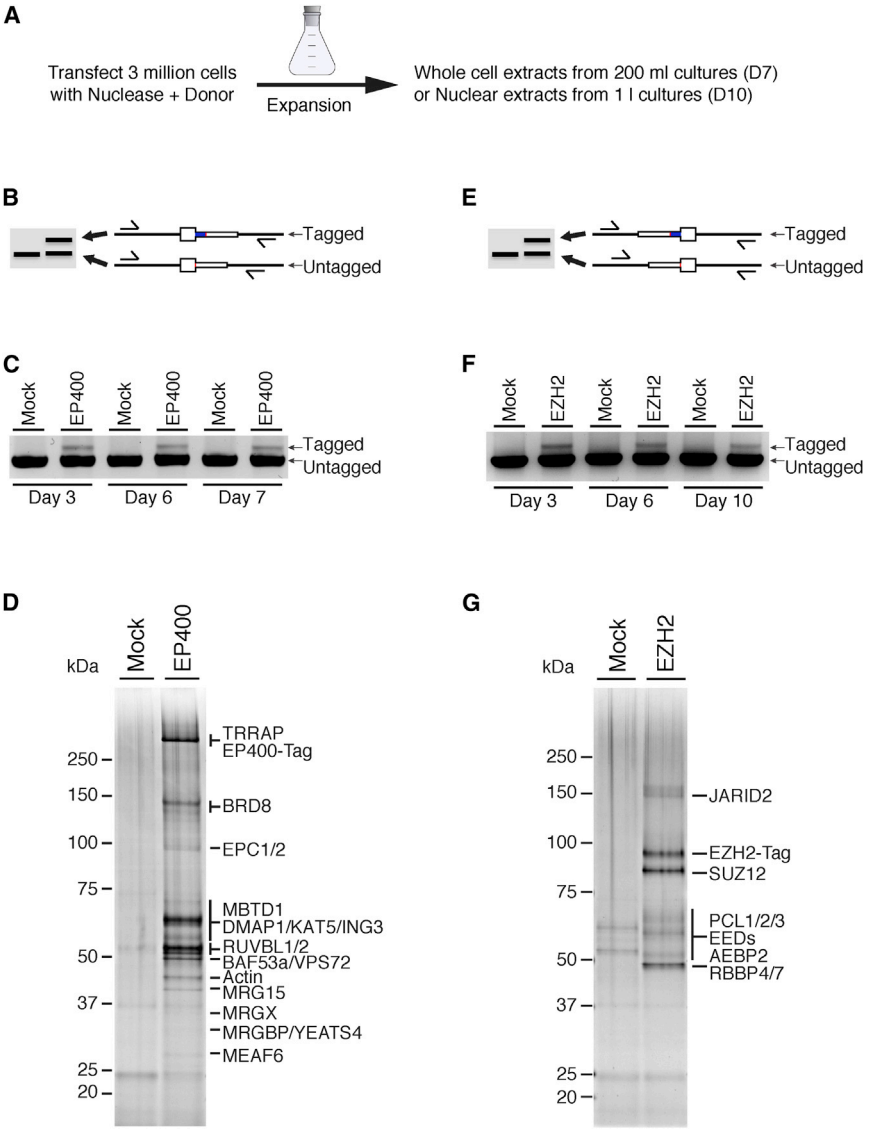


Figure 6. Efficient Complex Purification from Unselected Gene-Modified Cell Pools

(A) Timeline of the experiment.

(B) Schematic of a PCR-based assay (out-out PCR) to detect targeted integration (TI) of the tag sequence at the C terminus of EP400. Primers are located outside of the homology arms and are designed to yield a longer PCR product if the tag is inserted.

(C) Results of an out-out PCR assay conducted on genomic DNA from K562 cells transfected with (1) EGFP expression vector and EP400 donor (Mock) and (2) wild-type Cas9 expression vector, gRNA #3, and EP400 donor (EP400). Cells were collected at various time points post-transfection.

(D) Silver-stained SDS-PAGE showing the purified EP400 complex isolated from the cell pool 7 days post-transfection.

(E) Same as in (B) but for the N-terminal tagging of EZH2.

(F) Same as in (C) but using the EZH2 TALENs and donor.

(G) Silver-stained SDS-PAGE showing the purified EZH2 complex isolated from the cell pool 10 days post-transfection. Proteins were identified from unfractionated protein samples and assigned to specific gel bands based on western blotting analysis and predicted molecular weights.

the risk of obtaining confounding results. However, we note that the targeting specificity for the TAP tag using circular DNA donors benefits from the fact that homology-directed repair (HDR) mechanisms are required for integration of donor sequences. Thus, random integration of the donor is minimal, because there are no homologous sequences at these potential off-targets. Nevertheless, on-target mutagenesis via non-homologous end joining (NHEJ) can result in small insertions and deletions (indels) at the non-targeted allele, and care should be taken to carefully genotype both tagged and untagged alleles in clonally derived cells lines (see Figures S3 and S5 for examples). We note that it is possible to design the targeting strategy to minimize the potential impact of such mutagenic events (see Figures S2, S4, and S5). In this study, we used clones with biallelic integrations of the TAP tag when possible. This has an additional advantage, as all molecules of the bait proteins are tagged in the cell.

Lastly, our data suggest that high-throughput characterization of protein-protein interactions under physiological conditions is achievable, since protein complexes can be efficiently purified from mixed cell populations containing only a fraction of tagged alleles (Figure 6). As genome-scale CRISPR-Cas9 knockout screening in human cells is now a reality (Shalem et al., 2014; Wang et al., 2014), it will be possible to adapt these methods to generate nucleases targeting either the N or C terminus of every human protein. Thus, a genome-scale proteomic approach of endogenous human proteins using this strategy seems imminently feasible.

EXPERIMENTAL PROCEDURES

Cell Culture and Transfection

K562 cells were obtained from the ATCC and maintained at 37°C under 5% CO_2 in RPMI medium supplemented with 10% fetal bovine serum, penicillin-streptomycin, and GlutaMAX. When cultivated in Erlenmeyer or spinner flasks, 25 mM HEPES-NaOH (pH 7.4) was added. Cells were transfected using the Amaxa 4D-Nucleofector (Lonza) per the manufacturer's recommendations.

ZFN, TALEN, and CRISPR/Cas9 Reagents

The *AAVS1*-targeting ZFNs and *EZH2* TALENs (Addgene #36775 and #36776) have been described previously (Hockemeyer et al., 2009; Reyon et al., 2012). The cytomegalovirus (CMV)-driven human-codon optimized Cas9 nuclease and nickase (Cas9 D10A) vectors (Addgene #41815 and #41816) and the U6-driven guide RNA (gRNA) vector to target human *AAVS1* (T2 target sequence; Addgene #41818) have been described elsewhere (Mali et al.,

2013). The two U6-driven guide RNA (gRNA) vectors to target human *FANCF* have been described previously (Tsai et al., 2014). All other gRNA expression vectors were built in the MLM3636 (Addgene #43860) backbone. Target sequences for *EPC1* (5′-GTGACGTAGCTTCCTCCGAG-3′), *EP400* (5′-TGCCCTGACTACTGGCACGG-3′), *MBTD1* (5′-ATCAAACAAGAGCCATGAGG-3′), and *MCM8* (5′-CCAGCTTCAAACTATGTAAA-3′) were chosen according to a web-based CRISPR design tool (Hsu et al., 2013). The DNA sequence for the gRNA for *EP400*, *MBTD1*, and *MCM8* were modified at position 1 to encode a "G" due to the transcription initiation requirement of the human U6 promoter. When present on a nuclease construct, FLAG epitopes were removed by subcloning.

Construction of Plasmid Donors for Recombination

The *AAVS1*-TAP tagging plasmid (Addgene #68375) was assembled in the *AAVS1* SA-2A-puro-pA vector (Hockemeyer et al., 2009) by inserting a synthetic DNA fragment containing a SV40 late polyadenylation sequence, h*PGK1* promoter, and TAP tag sequence in between the puromycin resistance gene and the BGH polyadenylation site. The donor plasmids for tagging *EPC1*, *EP400*, *MBTD1*, *EZH2*, and *FANCF* were synthesized as gBlocks gene fragments (Integrated DNA Technologies) and assembled using the Zero Blunt TOPO cloning kit (Life Technologies) or cloned by restriction into pUC19. The homology arms for the MCM8 donor plasmid were amplified by PCR from K562 genomic DNA.

Targeted Integration to the *AAVS1* Locus

One million cells were transfected with 1 μg ZFN expression vector and 4 μg donor constructs. Simultaneous selection and cloning was performed for 10 days in methylcellulose-based semi-solid RPMI medium supplemented with 0.25 μg/ml puromycin starting 3 days post-transfection. Clones were picked and expanded in 96 wells for 3 days and transferred to 12-well plates for another 3 days before cells were harvested for western blot.

CRISPR/Cas9 and TALEN-Driven Targeted Integration

For targeting with the CRISPR/Cas9 system, one million cells were transfected with 2 μg gRNA plasmid, 2 μg Cas9 vector, and 4 μg donor. For TALEN-driven integration, 2.5 μg of each vector and 4 μg donor were transfected. Limiting dilution cloning was performed 3 days post-transfection, and targeted clones were identified via out-out PCR. For the experiments shown in Figure 6, 3 million cells were transfected at the same DNA ratios and the cells were cultivated in 10 ml RPMI media in a T-75 flask for 3 days. The cells were then diluted to 2E5/ml and grown in Erlenmeyer flasks with agitation until they reached a density of 1E6/ml. Under these conditions, we typically obtained 2E8 cells (200 ml culture at saturation) 7 days post-transfection and 1E9 cells (1 l culture at saturation) 10 days post-transfection.

TAP

Typically, nuclear extracts (Abmayr et al., 2006) were prepared from 1E9 to 3E9 cells (1 l to 3 l cultures at saturation), adjusted to 0.1% Tween-20, and ultracentrifuged at 100,000 × *g* for 45 min. Extracts were precleared with 300 μl Sepharose CL-6B (Sigma), then 250 μl anti-FLAG M2 affinity resin (Sigma) was added for 2 hr at 4°C. The beads were then washed in Poly-Prep columns (Bio-Rad) with 40 column volumes (CV) of buffer #1 (20 mM HEPES-KOH [pH 7.9], 10% glycerol, 300 mM KCl, 0.1% Tween 20, 1 mM DTT, 1× Halt protease and phosphatase inhibitor cocktail without EDTA [Pierce]) followed by 40 CV of buffer #2 (20 mM HEPES-KOH [pH 7.9], 10% glycerol, 150 mM KCl, 0.1% Tween 20, 1 mM DTT, 1× Halt protease and phosphatase inhibitor cocktail without EDTA [Pierce]). Complexes were eluted with 5 CV of buffer #2 supplemented with 150 μg/ml 3xFLAG peptide (Sigma) for 1 hr at 4°C. Next, fractions were mixed with 125 μl Strep-Tactin Sepharose (IBA) affinity matrix for 1 hr at 4°C, and the beads were washed with 40 CV of buffer #2 in Poly-Prep columns (Bio-Rad). Complexes were eluted in two fractions with 4 CV of buffer #2 supplemented with 2.5 mM D-biotin, flash frozen in liquid nitrogen, and stored at −80°C. Typically, 15 μl of the first elution (3% of total) was loaded on Bolt or NuPAGE 4%–12% Bis-Tris gels (Life Technologies) and analyzed by silver staining. For purifications from whole-cell extracts, cells were washed twice with PBS and lysed in buffer A (20 mM HEPES-KOH [pH 7.9], 10% glycerol, 300 mM KCl, 0.1% IGEPAL CA-630, 1 mM DTT, 1× Halt protease and phosphatase inhibitor cocktail without EDTA [Pierce]) (ratio of 100 μl of lysis buffer per 1E6 cells) for 30 min at 4°C. Extracts were centrifuged for 30 min at 17,000 × *g*, and the purifications were performed as described above.

HAT and HMT Assays

TAP-purified fractions were assayed for enzymatic activity on short oligonucleosomes isolated from HeLa S3 cells as described previously (Doyon et al., 2004; Musselman et al., 2012).

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, six figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.09.009.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.D. and J.C.; Methodology, Y.D., M.D., J.L.; Investigation, M.D., J.L., K.J., C.C.H., C.R., P.H., Y.D.; Writing, Y.D. and J.C.; Funding Acquisition, Y.D. and J.C.

ACKNOWLEDGMENTS

This study was supported by grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2014-059680) to Y.D. and the Canadian Institutes of Health Research (MOP-64289) to J.C. Salary support for Y.D. was provided by the Fonds de la Recherche du Québec-Santé (FRQS). J.C. holds a Canada Research Chair in Chromatin Biology and Molecular Epigenetics. We thank J. Laganière (Héma-Québec) and F. Urnov (Sangamo BioSciences) for helpful discussions and M. Wilson (University of Toronto) and N. Lacroix-Pepin for assistance with cloning. Protein identification by mass spectrometry was performed by the Proteomics Center at CHU de Québec Research Center.

Received: June 13, 2015
Revised: July 27, 2015
Accepted: September 2, 2015
Published: October 8, 2015

REFERENCES

Abmayr, S.M., Yao, T., Parmely, T., and Workman, J.L. (2006). Preparation of nuclear and cytoplasmic extracts from mammalian cells. Curr. Protoc. Mol. Biol. *12*, 11.

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291–294.

Alekseyenko, A.A., Gorchakov, A.A., Kharchenko, P.V., and Kuroda, M.I. (2014). Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. Proc. Natl. Acad. Sci. USA *111*, 2488–2493.

Ali, A.M., Pradhan, A., Singh, T.R., Du, C., Li, J., Wahengbam, K., Grassman, E., Auerbach, A.D., Pang, Q., and Meetei, A.R. (2012). FAAP20: a novel ubiquitin-binding FA nuclear core-complex protein required for functional integrity of the FA-BRCA DNA repair pathway. Blood *119*, 3285–3294.

Behrends, C., Sowa, M.E., Gygi, S.P., and Harper, J.W. (2010). Network organization of the human autophagy system. Nature *466*, 68–76.

Bradbury, A., and Plückthun, A. (2015). Reproducibility: Standardize antibodies used in research. Nature *518*, 27–29.

Cai, Y., Jin, J., Florens, L., Swanson, S.K., Kusch, T., Li, B., Workman, J.L., Washburn, M.P., Conaway, R.C., and Conaway, J.W. (2005). The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. J. Biol. Chem. *280*, 13665–13670.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

DeKelver, R.C., Choi, V.M., Moehle, E.A., Paschon, D.E., Hockemeyer, D., Meijsing, S.H., Sancak, Y., Cui, X., Steine, E.J., Miller, J.C., et al. (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res. *20*, 1133–1142.

Doyon, Y., Selleck, W., Lane, W.S., Tan, S., and Côté, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. Mol. Cell. Biol. *24*, 1884–1896.

Doyon, Y., Cayrou, C., Ullah, M., Landry, A.J., Côté, V., Selleck, W., Lane, W.S., Tan, S., Yang, X.J., and Côté, J. (2006). ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. Mol. Cell *21*, 51–64.

Doyon, J.B., Zeitler, B., Cheng, J., Cheng, A.T., Cherone, J.M., Santiago, Y., Lee, A.H., Vo, T.D., Doyon, Y., Miller, J.C., et al. (2011). Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. Nat. Cell Biol. *13*, 331–337.

Fuchs, M., Gerber, J., Drapkin, R., Sif, S., Ikura, T., Ogryzko, V., Lane, W.S., Nakatani, Y., and Livingston, D.M. (2001). The p400 complex is an essential E1A transformation target. Cell *106*, 297–307.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631–636.

Gavin, A.C., Maeda, K., and Kühner, S. (2011). Recent advances in charting protein-protein interaction: mass spectrometry-based approaches. Curr. Opin. Biotechnol. *22*, 42–49.

Goudreault, M., D'Ambrosio, L.M., Kean, M.J., Mullin, M.J., Larsen, B.G., Sanchez, A., Chaudhry, S., Chen, G.I., Sicheri, F., Nesvizhskii, A.I., et al. (2009). A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. Mol. Cell. Proteomics *8*, 157–171.

Guruharsha, K.G., Rual, J.F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., et al. (2011). A protein complex network of Drosophila melanogaster. Cell *147*, 690–703.

Hegemann, B., Hutchins, J.R., Hudecz, O., Novatchkova, M., Rameseder, J., Sykora, M.M., Liu, S., Mazanek, M., Lénárt, P., Hériché, J.K., et al. (2011). Systematic phosphorylation analysis of human mitotic protein complexes. Sci. Signal. *4*, rs12.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., et al. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180–183.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat. Biotechnol. *27*, 851–857.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat. Biotechnol. *29*, 731–734.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

Hutchins, J.R., Toyoda, Y., Hegemann, B., Poser, I., Hériché, J.K., Sykora, M.M., Augsburg, M., Hudecz, O., Buschhorn, B.A., Bulkescher, J., et al. (2010). Systematic analysis of human protein complexes identifies chromosome segregation proteins. Science *328*, 593–599.

Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., et al. (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell *162*, 425–440.

Ikura, T., Ogryzko, V.V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000). Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. Cell *102*, 463–473.

Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. Nat. Rev. Mol. Cell Biol. *14*, 49–55.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643.

Krogan, N.J., Lippman, S., Agard, D.A., Ashworth, A., and Ideker, T. (2015). The cancer cell map initiative: defining the hallmark networks of cancer. Mol. Cell *58*, 690–698.

Lalonde, M.E., Cheng, X., and Côté, J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. Genes Dev. *28*, 1029–1041.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

Leiserson, M.D., Vandin, F., Wu, H.T., Dobson, J.R., Eldridge, J.V., Thomas, J.L., Papoutsaki, A., Kim, Y., Niu, B., McLellan, M., et al. (2015). Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat. Genet. *47*, 106–114.

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat. Methods *8*, 861–869.

Lutzmann, M., Grey, C., Traver, S., Ganier, O., Maya-Mendoza, A., Ranisavljevic, N., Bernex, F., Nishiyama, A., Montel, N., Gavois, E., et al. (2012). MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. Mol. Cell *47*, 523–534.

Maiorano, D., Lutzmann, M., and Méchali, M. (2006). MCM proteins and DNA replication. Curr. Opin. Cell Biol. *18*, 130–136.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Marcon, E., Ni, Z., Pu, S., Turinsky, A.L., Trimble, S.S., Olsen, J.B., Silverman-Gavrila, R., Silverman-Gavrila, L., Phanse, S., Guo, H., et al. (2014). Human-chromatin-related protein interactions identify a demethylase complex required for chromosome segregation. Cell Rep. *8*, 297–310.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Mozzetta, C., Pontis, J., Fritsch, L., Robin, P., Portoso, M., Proux, C., Margueron, R., and Ait-Si-Ali, S. (2014). The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol. Cell *53*, 277–289.

Musselman, C.A., Avvakumov, N., Watanabe, R., Abraham, C.G., Lalonde, M.E., Hong, Z., Allen, C., Roy, S., Nuñez, J.K., Nickoloff, J., et al. (2012). Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nat. Struct. Mol. Biol. *19*, 1266–1272.

Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., Takisawa, H., and Kanemaki, M.T. (2012). Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. Mol. Cell *47*, 511–522.

Park, J.H., Sun, X.J., and Roeder, R.G. (2010). The SANT domain of p400 ATPase represses acetyltransferase activity and coactivator function of TIP60 in basal p21 gene expression. Mol. Cell. Biol. *30*, 2750–2761.

Park, J., Long, D.T., Lee, K.Y., Abbas, T., Shibata, E., Negishi, M., Luo, Y., Schimenti, J.C., Gambus, A., Walter, J.C., and Dutta, A. (2013). The MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination. Mol. Cell. Biol. *33*, 1632–1644.

Plass, C., Pfister, S.M., Lindroth, A.M., Bogatyrova, O., Claus, R., and Lichter, P. (2013). Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. Nat. Rev. Genet. *14*, 765–780.

Rajendra, E., Oestergaard, V.H., Langevin, F., Wang, M., Dornan, G.L., Patel, K.J., and Passmore, L.A. (2014). The genetic and biochemical basis of FANCD2 monoubiquitination. Mol. Cell *54*, 858–869.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat. Biotechnol. *30*, 460–465.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. *17*, 1030–1032.

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., et al. (2014). A proteome-scale map of the human interactome network. Cell *159*, 1212–1226.

Saksouk, N., Avvakumov, N., Champagne, K.S., Hung, T., Doyon, Y., Cayrou, C., Paquet, E., Ullah, M., Landry, A.J., Côté, V., et al. (2009). HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail. Mol. Cell *33*, 257–265.

Sardiu, M.E., Cai, Y., Jin, J., Swanson, S.K., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2008). Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc. Natl. Acad. Sci. USA *105*, 1454–1459.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87.

Shiloh, Y., and Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. Nat. Rev. Mol. Cell Biol. *14*, 197–210.

Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the human deubiquitinating enzyme interaction landscape. Cell *138*, 389–403.

Sternberg, S.H., and Doudna, J.A. (2015). Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol. Cell *58*, 568–574.

Steunou, A.L., Rossetto, D., and Côté, J. (2014). Regulating chromatin by histone acetylation. In Fundamentals of Chromatin, J.L. Workman and S.M. Abmayr, eds. (Springer-Verlag), pp. 147–212.

Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat. Biotechnol. *32*, 569–576.

Tyteca, S., Vandromme, M., Legube, G., Chevillard-Briet, M., and Trouche, D. (2006). Tip60 and p400 are both required for UV-induced apoptosis but play antagonistic roles in cell cycle progression. EMBO J. *25*, 1680–1689.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. *11*, 636–646.

Van Leene, J., Hollunder, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Stals, H., Van Isterdael, G., Verkest, A., Neirynck, S., Buffel, Y., et al. (2010). Targeted interactomics reveals a complex core cell cycle machinery in Arabidopsis thaliana. Mol. Syst. Biol. *6*, 397.

Venters, B.J., Wachi, S., Mavrich, T.N., Andersen, B.E., Jena, P., Sinnamon, A.J., Jain, P., Rolleri, N.S., Jiang, C., Hemeryck-Walsh, C., and Pugh, B.F. (2011). A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Mol. Cell *41*, 480–492.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546–1558.

Walden, H., and Deans, A.J. (2014). The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu. Rev. Biophys. *43*, 257–278.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80–84.

Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. Genes Dev. *28*, 672–682.