nature cell biology

ERCC1–XPF cooperates with CTCF and cohesin to facilitate the developmental silencing of imprinted genes

Georgia Chatzinikolaou[1,4], Zivkos Apostolou[1,2,4], Tamara Aid-Pavlidis[1], Anna Ioannidou[1,2], Ismene Karakasilioti[1,5], Giorgio L. Papadopoulos[2,3], Michalis Aivaliotis[1], Maria Tsekrekou[1,2], John Strouboulis[1,3], Theodore Kosteas[1] and George A. Garinis[1,2,6]

Inborn defects in DNA repair are associated with complex developmental disorders whose causal mechanisms are poorly understood. Using an *in vivo* biotinylation tagging approach in mice, we show that the nucleotide excision repair (NER) structure-specific endonuclease ERCC1–XPF complex interacts with the insulator binding protein CTCF, the cohesin subunits SMC1A and SMC3 and with MBD2; the factors co-localize with ATRX at the promoters and control regions (ICRs) of imprinted genes during postnatal hepatic development. Loss of *Ercc1* or exposure to MMC triggers the localization of CTCF to heterochromatin, the dissociation of the CTCF–cohesin complex and ATRX from promoters and ICRs, altered histone marks and the aberrant developmental expression of imprinted genes without altering DNA methylation. We propose that ERCC1–XPF cooperates with CTCF and cohesin to facilitate the developmental silencing of imprinted genes and that persistent DNA damage triggers chromatin changes that affect gene expression programs associated with NER disorders.

To counteract DNA damage, mammalian cells have evolved partially overlapping DNA repair systems to remove DNA lesions and restore their DNA back to its native form[1,2]. For bulky helix-distorting ultraviolet (UV)-induced DNA lesions, the principal repair mechanism is the evolutionarily conserved nucleotide excision repair (NER) pathway[3]. NER recognizes and removes helical distortions throughout the genome, that is, global genome NER, or selectively from the transcribed strand of active genes, that is, transcription-coupled repair. The NER structure-specific endonucleases XPG and ERCC1–XPF cleave on the 3′ and 5′ side of the DNA lesion, respectively, to release a 24–32-nucleotide fragment containing the damaged DNA[4–6]. Besides NER, the ERCC1–XPF complex also participates in the repair of DNA interstrand crosslinks (DNA ICLs)[7,8] and for the completion of homologous recombination at DNA replication forks stalled by DNA ICLs[9,10].

In humans, mutations in NER genes lead to the skin cancer-prone xeroderma pigmentosum or to a heterogeneous group of premature ageing-like (progeroid) disorders, including Cockayne syndrome (associated genes: *Csa*, *Csb*, *Xpd*, *Xpb*) and trichothiodystrophy (TTD; associated genes: *Xpd*, *Xpb*)[11–15]. Patients with subtle mutations in *Xpf* show mild xeroderma pigmentosum features and develop cancer during adulthood[16]. Instead, a patient with an *Ercc1* mutation showed severe clinical abnormalities coupled with a relatively mild DNA repair defect[17]. This and the fact that a complete defect in NER is compatible with life[18] argues for XPF–ERCC1 having functions outside the canonical NER[11].

Indeed, NER factors are now known to play a role, in addition to DNA repair, in the regulation of gene expression[19,20], chromatin looping[21], the transcriptional reprogramming of pluripotent stem cells[22] and the fine-tuning of growth-promoting genes during postnatal development[23]. At present, however, no solid evidence exists as to how NER is functionally involved in these processes, what are the NER-bound protein factors involved and their *in vivo* relevance to NER disorders. To tackle this, we used an *in vivo* biotinylation tagging approach in mice and mutant animals to dissect the functional contribution of ERCC1–XPF during liver development.

[1]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Nikolaou Plastira 100, 70013 Heraklion, Crete, Greece. [2]Department of Biology, University of Crete, Vassilika Vouton, GR71409 Heraklion, Crete, Greece. [3]Division of Molecular Oncology, Biomedical Sciences Research Center 'Alexander Fleming', GR 16672 Vari, Greece. [4]These authors contributed equally to this work. [5]Present address: Max Planck Institute for Metabolism Research, Gleueler str. 50, DE-50931 Cologne, Germany.
[6]Correspondence should be addressed to G.A.G. (e-mail: garinis@imbb.forth.gr)

Received 28 March 2016; accepted 24 February 2017; published online 3 April 2017; DOI: 10.1038/ncb3499

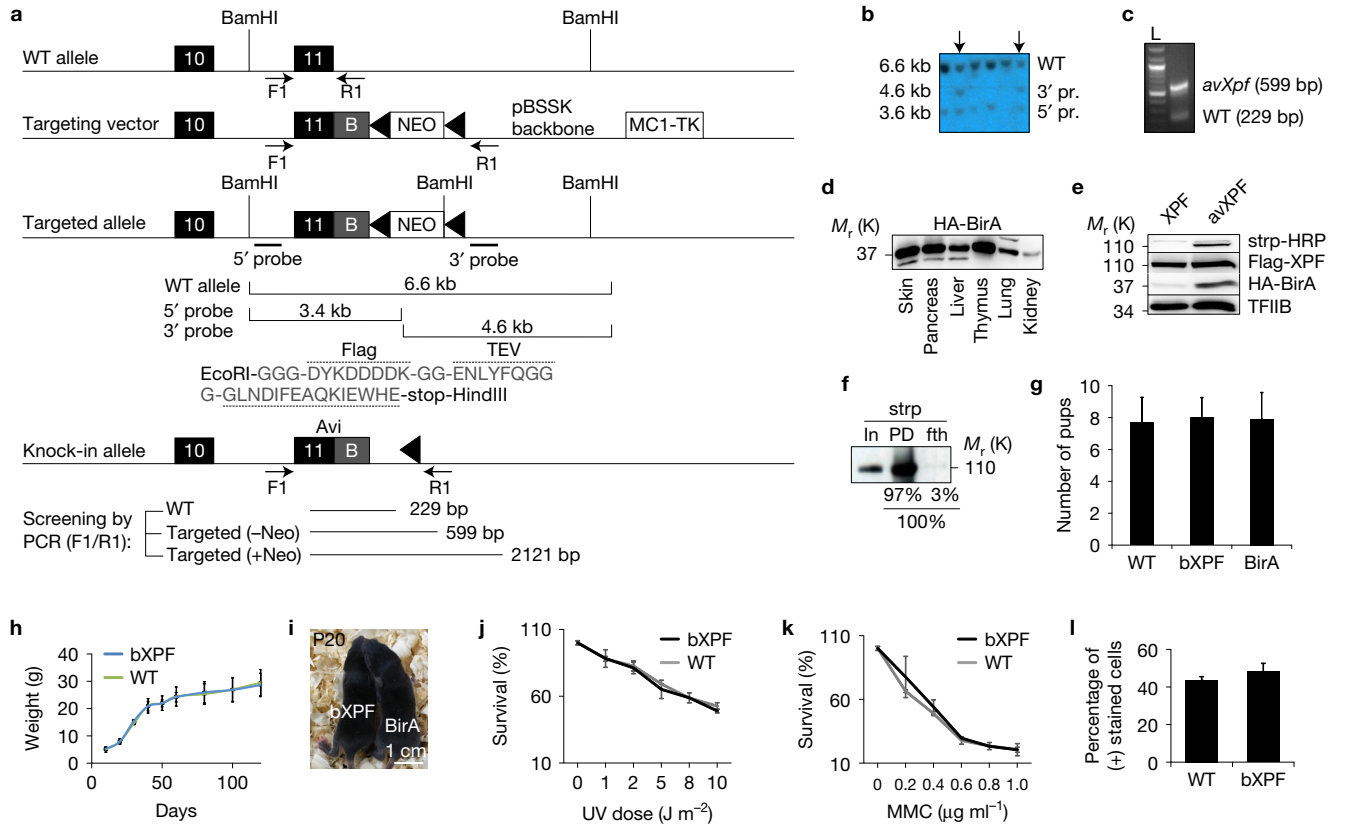


Figure 1 Generation of NER-proficient, biotin-tagged XPF animals. (**a**) Schematic representation of knock-in mice expressing the NER structure-specific endonuclease XPF fused with the 15-amino-acid Avi tag sequence, a 3×Flag tag and a tobacco etch virus (TEV) site separating the two tags. EcoRI and HindIII sites are synthetic; B, Flag-Tev-Avi fragment. (**b**,**c**) Two positive embryonic stem cell clones (marked with arrow) were used to generate germ-line-transmitting chimaeras that were backcrossed with C57Bl/6 mice to generate *avXpf*$^{+/-}$ pups, as confirmed by Southern blot (**b**) and PCR (**c**). (**d**) Anti-HA immunoblot showing expression of the BirA biotin ligase protein (relative molecular mass, 37,000) in different tissues of 2-month-old BirA transgenic animals. (**e**) *In vivo* biotinylation of the short 15-amino-acid Avi tag in bXPF animals. Nuclear extracts from P15 avXPF livers expressing either only avXPF (XPF; relative molecular mass, 110,000) or avXPF and BirA biotin ligase (bXPF) were tested by western blot. The blot was probed with streptavidin-HRP (strp-HRP), which confirms biotinylation of XPF *in vivo*, with anti-Flag and anti-HA confirming the presence of the knock-in allele and of the BirA transgene, respectively. (**f**) Biotinylation efficiency in bXPF livers. The percentage of biotinylated XPF (97%) and flow-through (3%; fth) was calculated by performing pull-down with 1 mg of nuclear extract derived from 15-day-old bXPF livers and M-280 paramagnetic beads in excess. (**g**) Number of WT, bXPF and BirA pups born. (**h**) Weight (grams; g) of WT and bXPF animals at the indicated time points ($n=3$ mice per time point and genotype). (**i**) A photograph of 20-day bXPF and BirA animals. (**j**,**k**) Survival of primary bXPF and WT MEFs to UV (**j**) or to MMC (**k**) at the indicated doses ($n=3$ dishes per dose per genotype). (**l**) UV-induced unscheduled DNA synthesis in primary bXPF and WT MEFs (number of positive stained cells; $n=3$ dishes per genotype per treatment). Error bars, standard deviation (s.d.). Statistical source data are provided in Supplementary Table 6.

RESULTS

Generation of biotin-tagged XPF mice

We generated knock-in mice expressing the NER structure-specific endonuclease XPF fused carboxy terminally before the stop codon of the last exon 11 with a 15-amino-acid tandem affinity purification (TAP) tag biotinylatable sequence[24] and a 3× Flag tag separated by a tobacco etch virus protease cleavage site for easy tag removal (Fig. 1a). After transfection in 129/SV embryonic stem cells expressing the Protamine 1-*Cre* recombinase transgene to efficiently excise the neomycin cassette in the male germ line[25] and selection of properly targeted clones (Fig. 1a and Methods), we used two independent transfected clones to generate germ-line-transmitting chimaeras (Fig. 1b,c). TAP-tag-fused heterozygous males (av*Xpf*$^{+/-}$) were backcrossed and maintained in a C57BL/6J background. Homozygous av*Xpf*$^{+/+}$ knock-in mice were then crossed with mice broadly expressing the HA-tagged bacterial BirA biotin ligase transgene under the control of the human hnRNPA2B1/CBX3 methylation-free island[26] (Fig. 1d and Supplementary Fig. 1A). BirA is a bacterial ligase that specifically recognizes and efficiently biotinylates the 15-amino-acid avidin within the short 15-amino-acid tag (Fig. 1e,f), thus creating a high-affinity 'handle' for the *in vivo* isolation of XPF-bound protein complexes from protein extracts isolated from biotin-tagged XPF (bXPF) mice by binding to streptavidin. Unlike *Ercc1*$^{-/-}$ or *Xpf*$^{-/-}$ mice[27,28], bXPF animals are born at the expected Mendelian frequency (Fig. 1g), grow normally (Fig. 1h) and show no developmental defects or other pathological features (Fig. 1i). Importantly, primary bXPF mouse embryonic fibroblasts (MEFs) show no hypersensitivity to UV (Fig. 1j) or to the DNA crosslinker mitomycin C (MMC; Fig. 1k) and no detectable differences in UV-induced unscheduled DNA synthesis when compared to wild-type (WT) control MEFs (Fig. 1l). Thus, bXPF animals develop normally to adulthood and are NER- and DNA ICL-repair-proficient.

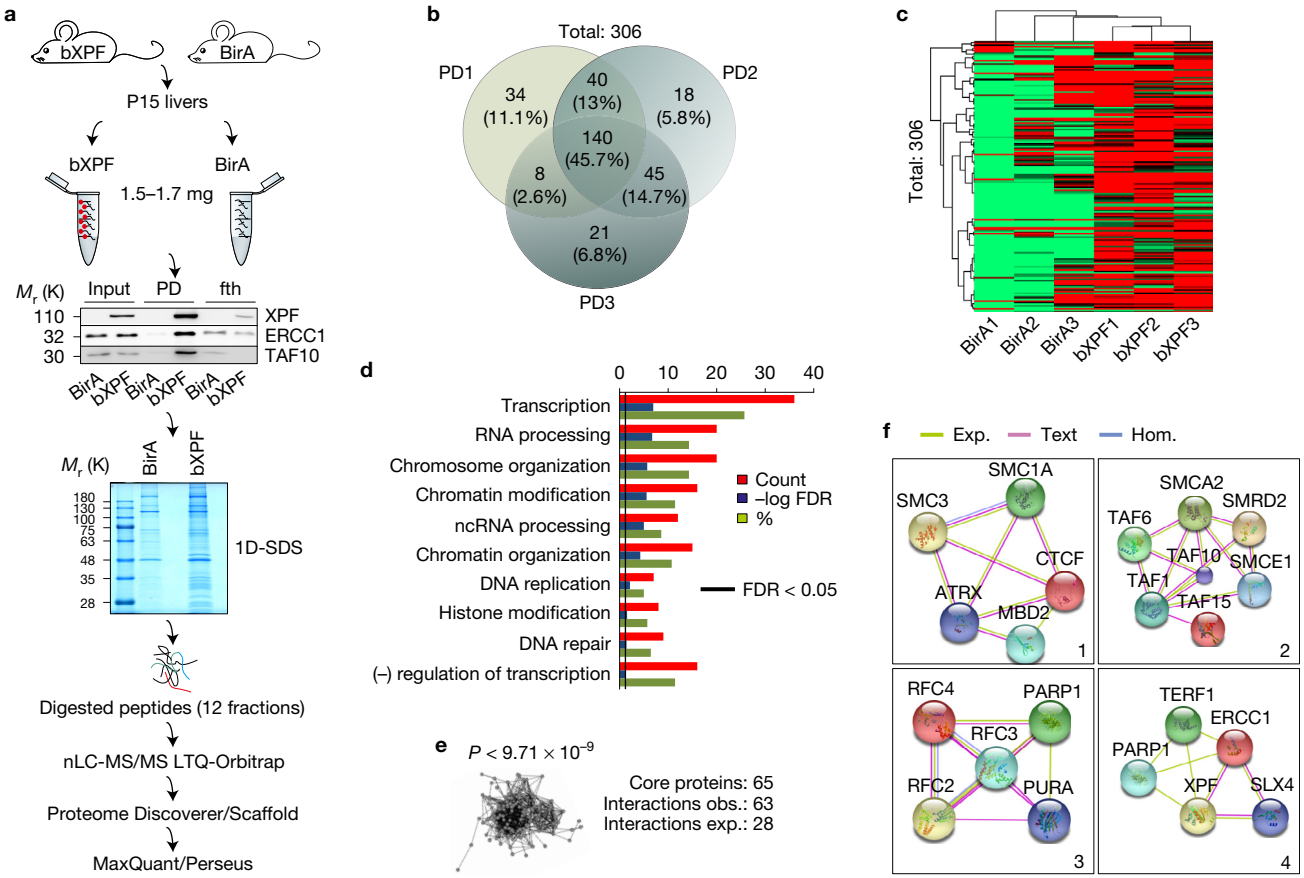


Figure 2 ERCC1–XPF interacts with protein complexes involved in genome organization and chromosome architecture. (**a**) Schematic representation of the high-throughput MS analysis in the livers derived from P15 bXPF animals expressing the BirA transgene ($n=15$) and BirA transgenic mice ($n=15$). (**b**) Venn diagram of bXPF-bound protein factors from three independent pulldowns (PD) and subsequent MS analyses. (**c**) Hierarchical clustering of the signal intensities of 306 bXPF-bound proteins in P15 bXPF and BirA livers. (**d**) List of significantly over-represented biological processes of 140 shared XPF-bound proteins. The biological processes are sorted by the −log of the *P* value, which is calculated by Fisher's exact test right-tailed. The black line marks the threshold of significance at $P \leq 0.05$. Count: the number of identified XPF-bound protein factors involved in the indicated biological process; % depicts the percentage of XPF-bound proteins involved in the indicated biological process over the total number of XPF-bound proteins (**e**). Number of observed (obs.) and expected (exp.) known protein interactions within the core 65 XPF-bound protein set. (**f**) Schematic representation of the four major XPF-bound protein complexes involved based on experimental (exp.) evidence and/or text mining (text) evidence, hom: homologous proteins. fth: flow-through. Unprocessed original scans of blots are shown in Supplementary Fig. 8.

A proteomics strategy reveals ERCC1–XPF protein conjugates involved in genome organization and chromosome architecture

To isolate and characterize NER-associated protein complexes during postnatal murine development, we prepared nuclear extracts from P15 avXPF; BirA livers (designated as bXPF) and livers expressing only the BirA transgene using high-salt extraction. Next, we confirmed that bXPF can still interact with known protein partners involved in NER, that is, ERCC1 (ref. 11), or transcription, that is, TAF10 (ref. 23) (Fig. 2a and Supplementary Fig. 8A). We then separated the bound protein interactome by one-dimensional SDS–PAGE followed by in-gel digestion (~12 slices) and peptides were further separated and analysed with high-resolution liquid chromatography–tandem mass spectrometry (nLC-ESI-MS/MS) on a hybrid linear ion trap Orbitrap mass spectrometer (Fig. 2a). From three biological replicates, we identified a total of 306 proteins (Supplementary Table 1) with 140 proteins (45.7%) shared in all three measurements under stringent selection criteria (Fig. 2b and Supplementary Table 2; see Methods). Using a hierarchical clustering approach, we confirmed that the 306 bXPF-bound proteins are capable of classifying the bXPF knock-in and BirA transgenic livers into the expected groups (Fig. 2c). At the confidence interval used, that is, false discovery rate (FDR) < 0.05, the significantly over-represented GO terms found (Fig. 2d) involved 65 out of the initial 140 XPF-bound core proteins; this set of proteins showed a significantly higher number of known protein interactions (that is, 63 interactions) than expected by chance (that is, 28 interactions; Fig. 2e) indicating a functionally relevant and interconnected protein network. Using this data set, we were able to discern four major XPF-bound protein complexes involved in: transcription silencing; transcription initiation; DNA replication; and DNA repair (Fig. 2f; shown in this order). These findings confirm previously documented interactions of ERCC1–XPF with components of the TFIID complex[23] and NER[11], whilst they reveal interactions of ERCC1–XPF with factors associated with transcription repression and DNA replication; the latter probably reflects the functional role of the ERCC1–XPF complex in homologous recombination repair.

Ablation of *Ercc1* gene triggers aberrant expression of imprinted genes during hepatic development

ERCC1–XPF is a highly conserved heterodimeric complex. *Ercc1*$^{-/-}$ mice are growth-defective, show progeroid features in several organs and die of liver failure within a month after birth[29,30]. Streptavidin pulldown identified CTCF co-purifying with bXPF under native (micrococcal nuclease digested) chromatin conditions (Supplementary Fig. 1B). This and our findings that XPF interacts with proteins known to be involved in chromatin organization and gene silencing, that is, CTCF, SMC1A, SMC3 and MBD2 (refs 31–33), prompted us to examine their relevance to the developmental defects seen in *Ercc1*$^{-/-}$ animals[34]. Co-immunoprecipitation experiments confirmed that a portion of endogenous ERCC1 interacts with CTCF, SMC1A, SMC3 and MBD2 in P15 livers; the protein–protein interactions were not affected by treating the extracts with benzonase (Fig. 3ai and Supplementary Fig. 8B). An antibody raised against CTCF confirmed the reciprocity of the interactions (Fig. 3aii and Supplementary Fig. 8B). We also challenged the specificity of ERCC1 interactions in *Ercc1*$^{-/-}$ livers (Fig. 3bi and Supplementary Fig. 8B); notably, CTCF, SMC1A, SMC3 and MBD2 interact in *Ercc1*$^{-/-}$ livers (Fig. 3bii and Supplementary Fig. 8B).

The functional role of CTCF, the cohesin and MBD2 in genomic imprinting or the postnatal silencing of distinct genes is well documented[32,33,35,36]. Imprinted genes are expressed in a parent-of-origin-specific manner already established in the gametes[37]. During embryogenesis only one parental allele is expressed, while the other allele is silenced[35]. Soon after birth, the remaining active allele of several imprinted genes is also silenced[38]. Using liver gene expression data sets[39], we evaluated the gene expression profiles of 68 imprinted genes in P15 livers derived from NER mutant animals displaying severe (*Csb*$^{m/m}$/*Xpa*$^{-/-}$, *Ercc1*$^{-/-}$), mild (*Xpd*^{TTD}, *Csb*$^{m/m}$) or no significant (*Xpa*$^{-/-}$) developmental/progeroid defects[23,30,40,41] (Supplementary Table 3). Unlike *Csb*$^{m/m}$, *Xpa*$^{-/-}$ or *Xpd*^{TTD} livers, we find that 22 out of the 68 imprinted genes show significantly aberrant gene expression profiles in P15 *Ercc1*$^{-/-}$ livers ($P < 0.05$; fold change $> \pm 1.2$) with no preference for genes expressed from the maternal or the paternal allele (Supplementary Table 4); the great majority of imprinted genes showed increased messenger RNA levels (17 out of 22; Fig. 3c) and associate with GO terms related to developmental and endocrine disorders (Supplementary Fig. 1C). A deregulation in the expression of eight imprinted genes was also observed in *Csb*$^{m/m}$/*Xpa*$^{-/-}$ livers (Fig. 3c). For further studies, we focused on insulin growth factor-2 *(Igf2)*, the paternally expressed gene 3 (*Peg3*), the protein delta homologue 1 (*Dlk1*) and the growth factor receptor-bound protein 10 (*Grb10*) genes in the *Ercc1*$^{-/-}$ and WT animals during hepatic development[42–45].

The *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes showed a progressive postnatal decline in mRNA levels compared with E18.5 that in *Csb*$^{m/m}$, *Xpa*$^{-/-}$ and WT livers led to barely detectable levels at day P60 (Fig. 3d). Instead, in *Ercc1*$^{-/-}$ livers, we find a gradual but steady postnatal increase in the mRNA levels of these genes when compared with age-matched WT livers (Fig. 3e). To test whether the increased mRNA levels in P15 *Ercc1*$^{-/-}$ livers reflect an embryonic defect in genome imprinting or a defect in the postnatal silencing of imprinted genes, we extended our studies in E18.5 *Ercc1*$^{-/-}$ fetal livers. Despite the marginally smaller size of the ERCC1-null embryos (Fig. 3f), we find no significant differences in the mRNA levels of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in E13.5 *Ercc1*$^{-/-}$ livers compared with age-matched WT controls (Fig. 3g). To determine the allelic source of imprinted gene transcripts in P15 *Ercc1*$^{-/-}$ livers, we identified a single polymorphic site within exon 9 of the *Peg3* gene in P15 *Ercc1*$^{-/-}$ livers that are generated when FVB *Ercc1*$^{-/+}$ females are crossed with C57BL/6J *Ercc1*$^{-/+}$ males; *Peg3* is maternally methylated and is expressed solely from the paternal allele in the liver[44]. Our analysis revealed that the increased expression of the *Peg3* gene was not due to reactivation of the maternal allele, because transcripts were still derived solely from the paternal allele (Fig. 3h). Further analysis revealed increased *Igf2*, *Grb10*, *Peg3*, *Meg3*, *Atp10a*, *H13* and *Airn* mRNA levels in the kidney, white adipose tissue, pancreas, cerebellum and spleen of the P15 ERCC1-defective animals, compared with age-matched WT animals (Fig. 3i and Supplementary Fig. 1D). Interestingly, *Ercc1* inactivation in aP2-*Ercc1*$^{F/-}$ animals occurs at ~P15 in mature adipocytes[46] when imprinted gene expression has already been established. Thus, unlike other NER-defective animals, a defect in ERCC1–XPF triggers the aberrant postnatal expression of a subset of imprinted genes.

Aberrant histone modifications and recruitment of factors associated with transcription initiation at the promoters of imprinted genes in *Ercc1*$^{-/-}$ livers

A series of chromatin immunoprecipitation (ChIP) studies in P15 *Ercc1*$^{-/-}$ livers revealed that RNAPII assembles together with the basal transcription factors TFIIB, XPD (TFIIH subunit) and TBP (TFIID) on promoters in P15 *Ercc1*$^{-/-}$ livers (Fig. 4a). Unlike WT livers, we also find the loss of repressive histone H3K9 trimethylation and H3K27 trimethylation marks and a concomitant increase of activating acetylated histone H3K9Ac and H3K4 trimethylation on *Igf2*, *Peg3*, *Dlk1* and *Grb10* gene promoters in P15 *Ercc1*$^{-/-}$ livers (Fig. 4b).

Genomic imprinting is established already in gametes by parent-of-origin-specific epigenetic marks, such as DNA methylation[47]. We, therefore, examined the DNA methylation status of the *Igf2*, *Peg3*, *Dlk1* and *Grb10* promoters, the *Peg3*, *H19*/*Igf2* (from now on designated as *H19*) and *Meg3*/*Dlk1* intergenic differentially methylated region (DMR) (from now on designated as *Meg3*) ICRs and the *Grb10* DMR in P15 *Ercc1*$^{-/-}$ and WT livers. Using a bisulfite conversion and sequencing assay on well-defined CTCF- and RNAPII-bound loci in *Peg3*, *H19* and *Meg3* ICRs and *Grb10 (CGI2)*/*Meg3* DMRs (Fig. 4c and Supplementary Fig. 2A,G), we detected no difference on methylation in *Ercc1*$^{-/-}$ compared to WT livers. The *H19*/*Igf2* genomic domain is regulated by long-range chromatin interactions in the liver, a process mediated in part by the DMRs and their DNA methylation state[48,49]. However, similar to ICRs, the DNA methylation state was also preserved in the *Igf2* DMR1 regulatory region in both WT and *Ercc1*$^{-/-}$ livers (Fig. 4c). Loss of ERCC1 in the liver resulted in reduced methylation of the *Peg3* promoter (Supplementary Fig. 2B; as shown), but no change in DNA methylation was observed between *Ercc1*$^{-/-}$ and WT livers at the CTCF peak within the *Grb10* region, the *Igf2* promoters 2 and 3, and the *Grb10* promoter (Supplementary Fig. 2C–F; as shown). Thus, in P15 *Ercc1*$^{-/-}$ livers, the increased mRNA levels of imprinted genes occurs without affecting DNA methylation in ICRs or the DMRs.

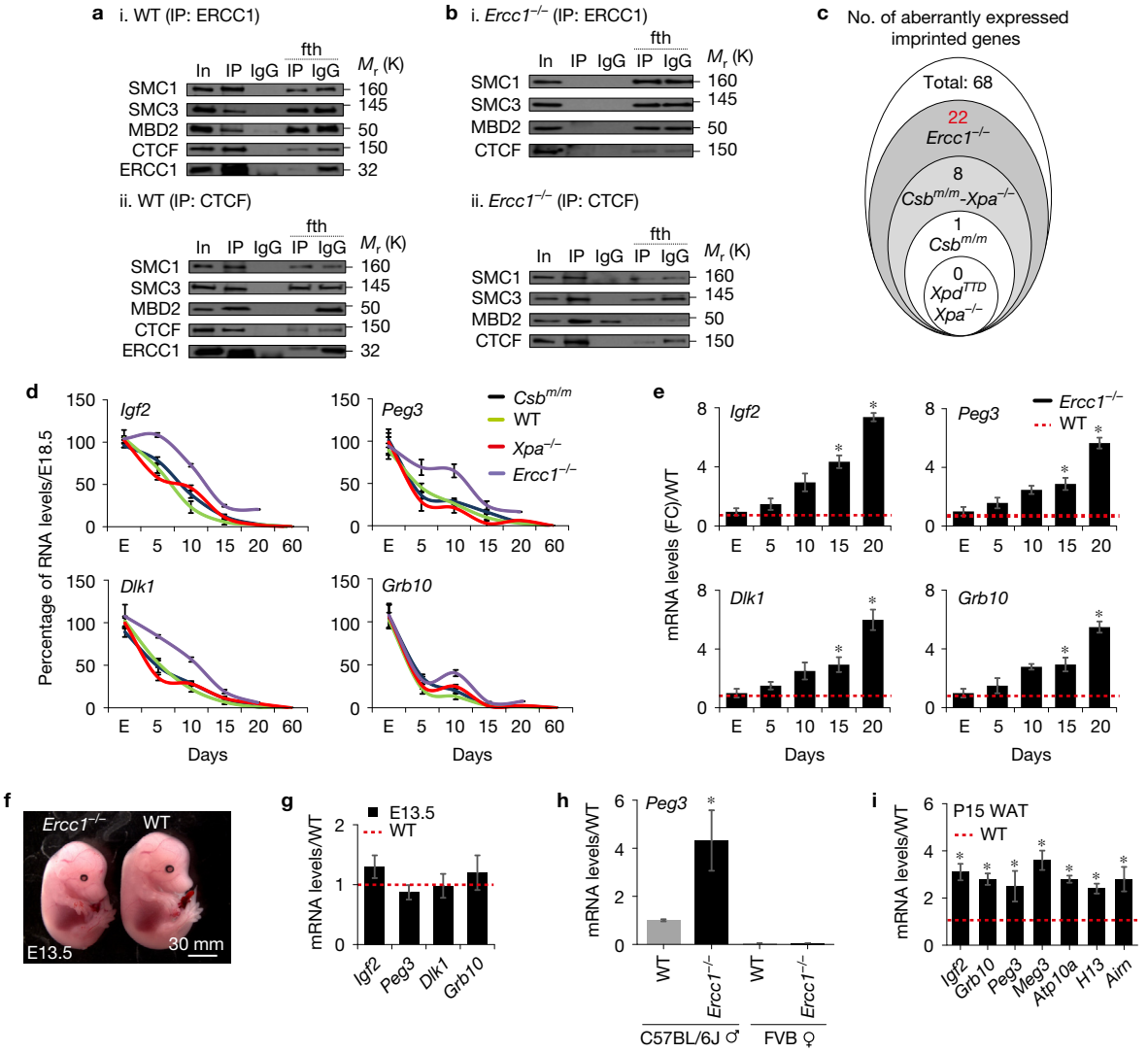


Figure 3 Ablation of the *Ercc1* gene triggers the aberrant expression of imprinted genes during postnatal hepatic development. (**a**) Co-immunoprecipitation experiments using anti-ERCC1 (i) or anti-CTCF (ii) in nuclear extracts from P15 livers analysed by western blotting for SMC1A, SMC3, MBD2, CTCF and ERCC1 as indicated. (**b**) Co-immunoprecipitation experiments using anti-ERCC1 (i) or anti-CTCF (ii) in nuclear extracts from P15 *Ercc1*$^{-/-}$ livers analysed by western blotting for the indicated protein factors. The input and flow-through are 1/20 of the extract used. Unprocessed original scans of blots are shown in Supplementary Fig. 8B. (**c**) Number of significantly aberrantly expressed imprinted genes in P15 *Ercc1*$^{-/-}$, *Csb*$^{m/m}$*/Xpa*$^{-/-}$, *Csb*$^{m/m}$, *Xpa*$^{-/-}$ and *Xpd*^{TTD} livers. (**d**) Quantitative PCR (qPCR) mRNA levels of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in *Csb*$^{m/m}$, *Xpa*$^{-/-}$, *Ercc1*$^{-/-}$ and WT livers (as shown) at the indicated time points compared with corresponding E18.5 (E), ($n=3$ biological replicates each representing a pool of 4–5 livers per time point and genotype; error bars, standard error of the mean (s.e.m.)). (**e**) qPCR mRNA levels of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in *Ercc1*$^{-/-}$ livers at the indicated time points as compared with age-matched littermate WT control livers ($n=3$ livers per time point and genotype; error bars, s.e.m.; two-tailed *t*-test). (**f**) A photograph of E13.5 *Ercc1*$^{-/-}$ and WT littermate embryos. (**g**) qPCR mRNA levels of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in E13.5 *Ercc1*$^{-/-}$ and WT livers (as shown; $n=3$ biological replicates each representing a pool of 4–5 livers per genotype). Red dotted line, WT mRNA levels. Error bars, s.d. (**h**) Allele-specific mRNA analysis revealing that increased levels of *Peg3* transcripts are derived only from the paternal (C57BL/6J) allele; no expression is detected from the maternally imprinted allele (FVB) ($n=3$ livers per genotype). Error bars, s.d.; two-tailed *t*-test. (**i**) qPCR mRNA levels of imprinted genes in P15 *Ercc1*$^{-/-}$ white adipose tissue (WAT, $n=3$ biological replicates each representing a pool of 4–5 tissues per genotype). Error bars, s.d.; two-tailed *t*-test. Red dotted line, WT mRNA levels; E, E18.5; $^{*}P \leq 0.05$. Statistical source data are provided in Supplementary Table 6.

Disruption of *Ercc1* leads to the dissociation of the CTCF–cohesin–MBD2 interacting factors from the promoters and ICRs of imprinted genes

ChIP followed by qPCR on the well-characterized CTCF- and RNAPII-bound *H19*, *Igf2*, *Dlk1/Meg3*, *Peg3* and *Grb10* target sequences (Supplementary Fig. 2A) showed that bXPF assembles with CTCF, SMC1A, SMC3 and MBD2 on promoters (Supplementary Fig. 3A) and the ICRs/DMR (Fig. 5a) in P15 WT livers but not in CTCF-negative regions (Supplementary Fig. 3B). We also find ATP-dependent helicase ATRX to recruit to the respective promoters and the ICRs (Supplementary Fig. 3A and Fig. 5a); ATRX was found to interact with bXPF in our proteomics studies and is known to cooperate with cohesin in silencing a subset of imprinted genes during postnatal murine brain development[32]. ChIP for CTCF or

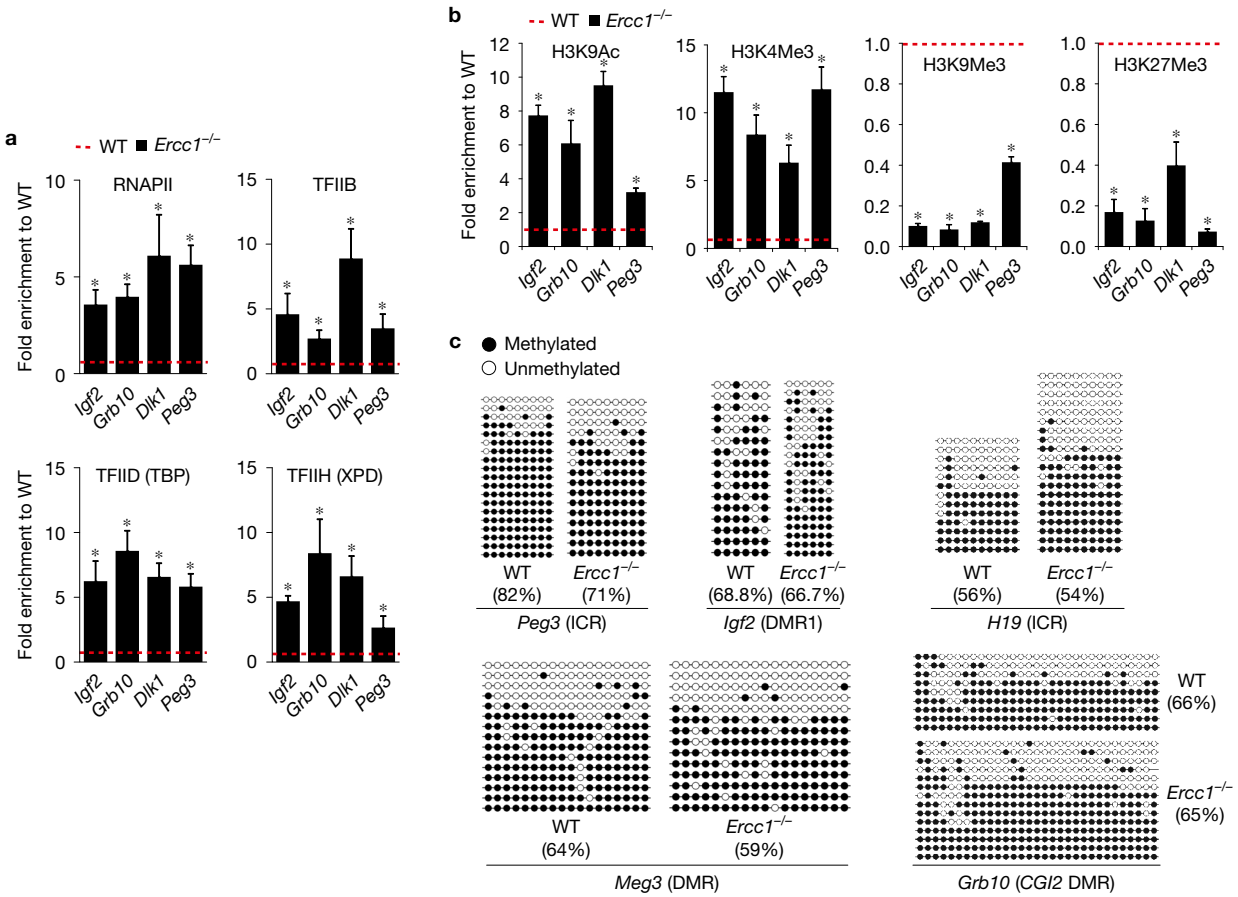


Figure 4 Aberrant histone post-translational modifications and recruitment of factors associated with transcription initiation at the promoters of imprinted genes in *Ercc1*$^{-/-}$ livers. (**a**) ChIP signals (shown as fold enrichment) of protein factors associated with transcription initiation (as indicated) at the *Igf2*, *Peg3*, *Dlk1* and *Grb10* proximal promoter regions in P15 *Ercc1*$^{-/-}$ livers as compared with age-matched WT controls; $n=3$ biological replicates each representing a pool of 4–5 livers per genotype. ChIP signals were expressed as fold enrichment over those obtained with control antibody (IgG), which were set as 1 (red dotted line). Error bars, s.d.; two-tailed *t*-test. (**b**) ChIP signals (shown as fold enrichment) of activating H3K9Ac, H3K4me3 and repressive H3K9me3, H3K27me3 histone marks at the *Igf2*, *Peg3, Dlk1* and *Grb10* proximal promoter regions in P15 *Ercc1*$^{-/-}$ livers as compared with age-matched WT controls; $n=3$ biological replicates each representing a pool of 4–5 livers per genotype. ChIP signals are expressed as in **a**. Error bars indicate s.e.m. among replicates; two-tailed *t*-test. (**c**) Schematic representation of the percentage of DNA methylation at the *Igf2*, *H19*, *Peg3*, *Meg3* and *Grb10* ICRs/DMRs in P15 *Ercc1*$^{-/-}$ and WT livers. Black circles, methylated cytosine; open circles, unmethylated cytosine, $^{*}P \leq 0.05$. Statistical source data are provided in Supplementary Table 6.

SMC1A and re-ChIP for SMC3, ERCC1, Flag-tagged XPF, MBD2 and ATRX showed that these factors co-occupy the *H19* and *Peg3* ICRs (Fig. 5b). Instead, ChIP signals for all factors tested were significantly reduced in promoters and the ICRs in P15 *Ercc1*$^{-/-}$ livers compared with WT controls (Fig. 5c and Supplementary Fig. 3C). Thus, the ERCC1–XPF complex is required for the optimal promoter and ICR assembly of protein complexes that associate with the postnatal silencing of imprinted genes in the developing liver.

CTCF, SMC1A, SMC3 and ATRX are known to display allele-specific binding at imprinted loci[32,33,35,36]. We, therefore, envisioned a similar scenario for ERCC1–XPF in the developing liver. ChIP followed by allele-specific restriction digest analysis of amplified DNA in P15 bXPF/SPRET/EiJ F1 livers that are polymorphic within the *Peg3* and *H19* ICRs showed that CTCF, SMC1A and SMC3 are preferentially enriched on the maternal and the paternal allele of the *H19* and *Peg3* ICRs, respectively (Fig. 5d); MBD2 and ATRX are preferentially recruited at the paternal and the maternal alleles of the *H19* and *Peg3* ICRs, respectively. The observation that CTCF and ATRX are localized to opposite alleles in *H19* and *Peg3* ICRs was unexpected as ERCC1–XPF assembles with these factors on ICRs (Fig. 5a) and interacts with CTCF and ATRX in P15 livers (Fig. 2 and Fig. 3a,b). We, therefore, reasoned that the ERCC1–XPF complex resides in both the paternal and the maternal alleles of *H19* and *Peg3* ICRs. In line, allele-specific restriction digest analysis of amplified immunoprecipitated or pulled-down DNA in P15 bXPF/SPRET/EiJ F1 livers showed no allele preference for either ERCC1 or bXPF on the ICRs (Fig. 5d).

Persistent DNA damage signalling triggers aberrant CTCF localization and the dissociation of the CTCF, the cohesin complex and MBD2 from promoters and ICRs

Confocal imaging in WT, *Csb*$^{m/m}$, *Xpa*$^{-/-}$ and *Xpc*$^{-/-}$ primary MEFs revealed that CTCF is distributed throughout the nucleoplasm.

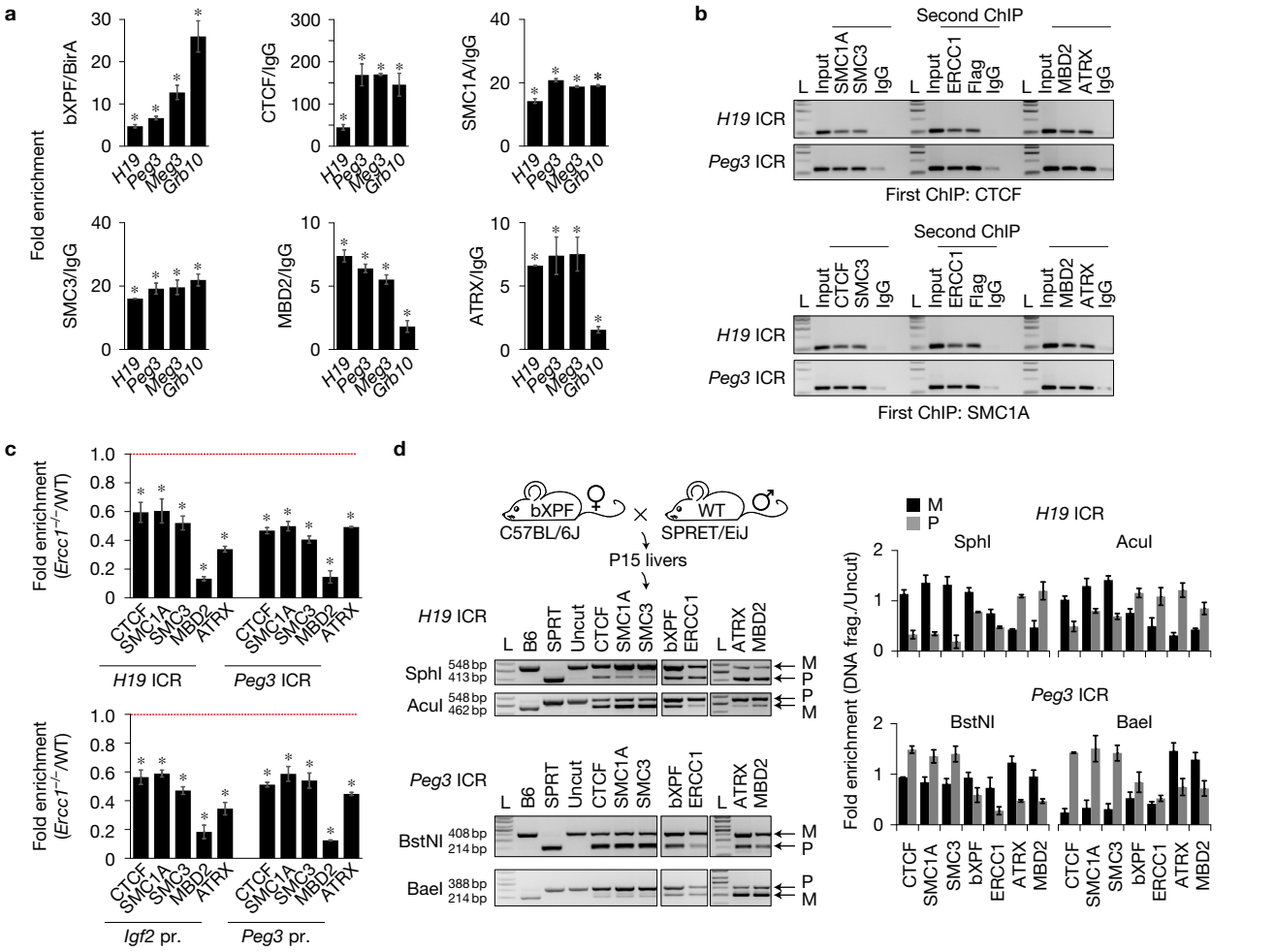


Figure 5 Dissociation of the CTCF–cohesin complex and MBD2 from the promoters and ICRs of imprinted genes in *Ercc1*$^{-/-}$ livers. (**a**) bXPF, CTCF, SMC1A, SMC3, MBD2 and ATRX ChIP signals are shown as fold enrichment of the percentage of input antibody over the percentage of input BirA (for bXPF) or control antibody (IgG) at the *H19*, *Peg3*, *Meg3* and *Grb10* ICRs/DMR in P15 WT mouse livers ($n=3$ biological replicates each representing a pool of 4–5 livers). Error bars, s.d.; two-tailed *t*-test. (**b**) ChIP with antibodies raised against CTCF (top panel) or SMC1A (bottom panel) and re-ChIP with antibodies raised against SMC1A (top panel) or CTCF (bottom panel) and SMC3, ERCC1, Flag-tagged XPF, MBD2 and ATRX (both panels). (**c**) CTCF, SMC1A, SMC3, MBD2 and ATRX ChIP signals are shown as fold enrichment of the percentage of input antibody over the percentage of input control antibody (IgG) in *Ercc1*$^{-/-}$ livers over the corresponding ChIP signals for the WT livers at the *H19/Igf2* and *Peg3* promoters and ICRs ($n=3$ biological replicates each representing a pool of 4–5 livers per genotype). Error bars indicate s.e.m. between biological replicates; two-tailed *t*-test. The red dotted line indicates the WT. (**d**) For *H19* ICR, ChIP-isolated P15 bXPF/SPRET/EiJ F1 liver DNA was amplified and digested with SphI (SPRET/EiJ paternal-specific site) or AcuI (bXPF maternal-specific site) restriction enzymes. For *Peg3* ICR, ChIP-isolated P15 bXPF/SPRET/EiJ F1 liver DNA was amplified and digested with BstNI (SPRET/EiJ paternal-specific site) or BaeI (bXPF maternal-specific site) restriction enzymes. Data are quantified and shown in corresponding graphs (right panel). Error bars indicate s.d.; $n=3$ biological replicates. L, ladder; SPRT, SPRET/EiJ; B6, C57BL/6J; M, maternal; P, paternal, $^{*}P \leq 0.05$. Statistical source data are provided in Supplementary Table 6.

Instead, in *Ercc1*$^{-/-}$ MEFs, we find that CTCF translocates to clear subnuclear landmarks identified as heterochromatin by 4′,-6-diamidine-2-phenylindole (DAPI) and HP1a staining, occasionally surrounded by the dense fibrillar component of the nucleolus (Fig. 6a–c and Supplementary Fig. 4A and Supplementary Fig. 4C). These dense clusters are known to be formed by the centromeres of different chromosomes that coalesce during interphase to form chromocentres[50]. ATRX accumulates to HP1a-stained heterochromatic regions in *Ercc1*$^{-/-}$ and to a lesser extent also in *Csb*$^{m/m}$ but not in *Xpa*$^{-/-}$ or *Xpc*$^{-/-}$ MEFs (Fig. 6d and Supplementary Fig. 4D). SMC1A accumulates only in the nucleoplasm of *Ercc1*$^{-/-}$ MEFs with minimal, sporadic localization to heterochromatin (Fig. 6e and Supplementary Fig. 4E). SMC3 showed comparable nuclear and cytoplasmic immunolocalization patterns across all NER-defective MEFs (Supplementary Fig. 4F). As for MEFs, we find that CTCF and ATRX accumulate in *Ercc1*$^{-/-}$ but not in WT primary hepatocytes (Supplementary Fig. 4G–H). For SMC1A and CTCF, we find no differences in protein levels between *Ercc1*$^{-/-}$ and WT MEFs or in MEFs exposed to MMC (Supplementary Fig. 4B). To test whether DNA damage directly contributes to the aberrant localization of CTCF and ATRX in *Ercc1*$^{-/-}$ MEFs, naive primary WT MEFs were exposed to UVC, H_2O_2 and MMC, a potent DNA crosslinker that, similar to the ERCC1 defect, triggers DNA ICLs that are efficiently processed by ERCC1–XPF[51,52] in a mechanism

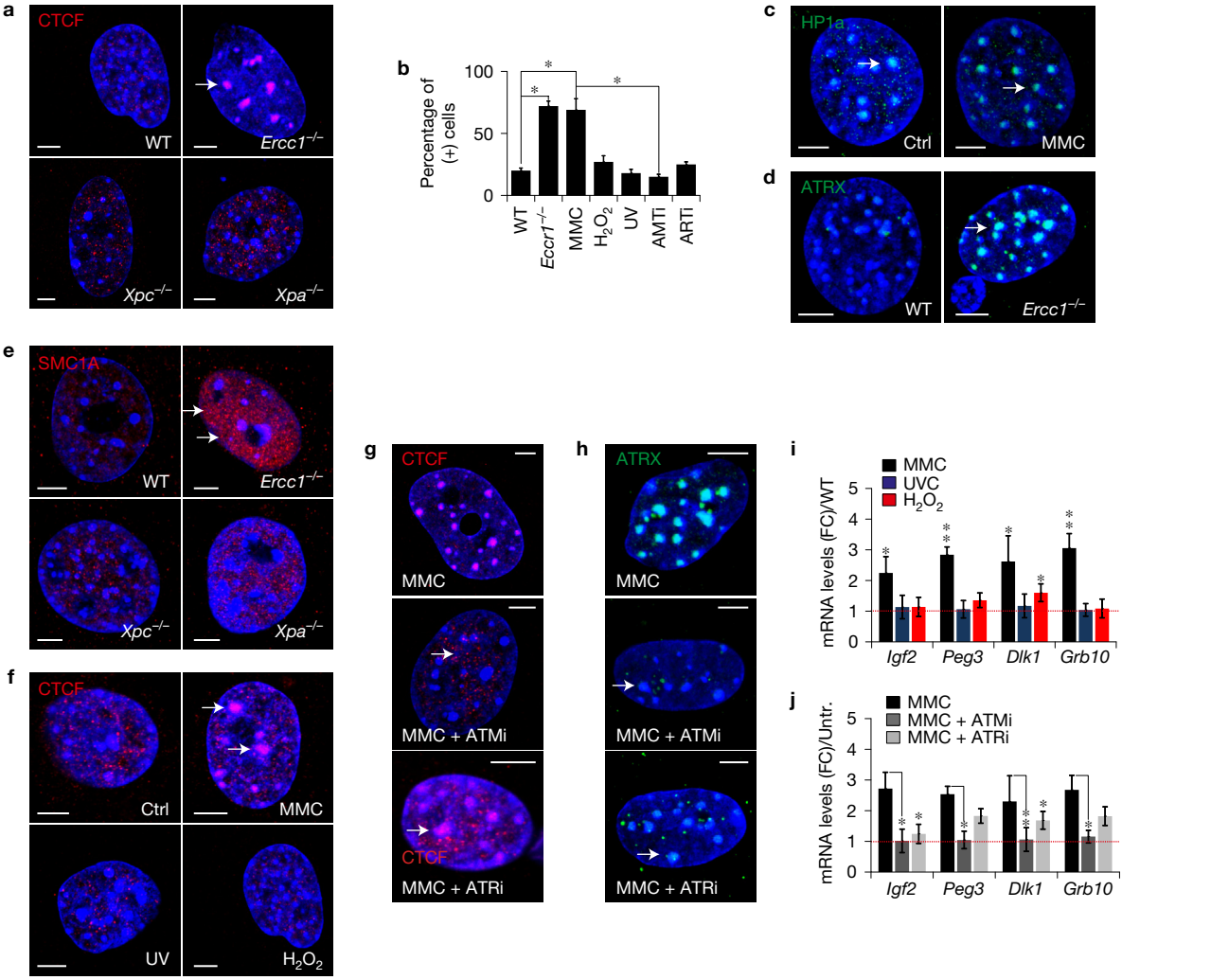


Figure 6 Persistent DNA damage signalling triggers aberrant CTCF localization and the dissociation of the CTCF–cohesin complex and MBD2 from promoters and ICRs. (**a**) Immunofluorescence detection of CTCF in WT, *Ercc1*$^{-/-}$, *Xpc*$^{-/-}$ and *Xpa*$^{-/-}$ primary MEFs. Note the distinctive translocation of CTCF to heterochromatin in *Ercc1*$^{-/-}$ MEFs. (**b**) Average number of CTCF-positive stained cells showing CTCF translocation to heterochromatin following treatment (as indicated by an arrow) from 20 fields analysed; $n=3$ biological replicates. Error bars indicate s.e.m.; two sided *t*-test. (**c**) Immunofluorescence detection of HP1a in MMC-treated and untreated control MEFs (as indicated by arrows). (**d**) Immunofluorescence detection of ATRX in *Ercc1*$^{-/-}$ and WT primary MEFs. Note the distinctive accumulation of ATRX to heterochromatin in *Ercc1*$^{-/-}$ MEFs (indicated with an arrow). (**e**) Immunofluorescence detection of SMC1A in WT, *Ercc1*$^{-/-}$, *Xpc*$^{-/-}$ and *Xpa*$^{-/-}$ primary MEFs. Note the nucleoplasmic accumulation of SMC1A in *Ercc1*$^{-/-}$ MEFs (indicated with arrows). (**f**) Immunofluorescence detection of CTCF in primary MEFs exposed to MMC, UV and H_2O_2. Note the distinctive translocation of CTCF to heterochromatin in MMC-treated MEFs (indicated by arrows). (**g**) Immunofluorescence detection of CTCF in primary MMC-treated MEFs exposed to ATM (ATMi) or ATR (ATRi) inhibitors. Note the absence of CTCF translocation to heterochromatin in MMC-treated MEFs exposed to ATMi (indicated with arrows). (**h**) Immunofluorescence detection of ATRX in primary MMC-treated MEFs exposed to ATM (ATMi) inhibitor (see also Supplementary Fig. 3). Note the absence of ATRX accumulation to heterochromatin in MMC-treated MEFs exposed to ATMi (indicated with arrows). (**i**) qPCR mRNA levels (expressed as fold change to untreated control MEFs) of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in primary MEFs exposed to MMC, UV or H_2O_2 ($n=3$ biological replicates). Error bars, s.d.; two-tailed *t*-test. Red dotted line, mRNA levels of untreated WT MEFs. (**j**) qPCR mRNA levels (expressed as fold change (FC) to untreated (Untr.) control MEFs) of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes in primary MMC-exposed MEFs treated with ATMi or ATRi ($n=3$ biological replicates). Error bars, s.d.; two-tailed *t*-test. Red dotted line, mRNA levels of untreated control MEFs. $^{*}P \leq 0.05$, $^{**}P \leq 0.01$. Scale bars, 5 μm. Statistical source data are provided in Supplementary Table 6.

distinct from NER[9] and the accumulation of DNA damage-associated γH2AX foci in the nucleus (Supplementary Fig. 4I). As with *Ercc1*$^{-/-}$ MEFs, we find that CTCF and ATRX are predominantly localized in heterochromatic regions in MMC-treated MEFs but, importantly, not following exposure of MEFs to UVC irradiation or to H_2O_2-induced oxidative DNA damage (Fig. 6f). The γH2AX foci accumulate only in the nucleoplasm and unlike with CTCF, they do not appear in the chromocentres of *Ercc1*$^{-/-}$ and MMC-treated MEFs (Supplementary Fig. 4I,J). Translocation of CTCF was also visible in G0 synchronized, serum-starved, MMC-treated MEFs (Supplementary Fig. 5A). Likewise, ATRX accumulates with CTCF in heterochromatin of MMC- but not of UVC- or H_2O_2-treated cells (Supplementary Fig. 5B,C). In line, we find increased mRNA levels for all genes tested in MMC- but not in UVC (4 J m^{-2})- or

H_2O_2-treated cells (Fig. 6i); increased mRNA levels were also detected at the much higher dose of 10 J m^{-2} of UVC probably reflecting other types of damage[53]. Similar findings were seen for a larger subset of imprinted genes in primary MEFs that also show increased mRNA levels for genes known to be induced following exposure to 4 J m^{-2} of UVC (Supplementary Fig. 6A) or following treatment with H_2O_2 (Supplementary Fig. 6B). Inhibition of ATM (ATMi) with KU-55933 inhibitor[54] in MMC-treated MEFs significantly abrogated the accumulation of CTCF and ATRX in heterochromatin (Fig. 6g,h). In ATMi-treated cells, we also evidenced the normative expression levels for *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes compared with untreated control cells (Fig. 6j). Inactivation of ATR with the ATR/CDK inhibitor NU6027 in MMC-treated MEFs led to similar results to those seen following ATM inactivation, albeit to a smaller magnitude (Fig. 6g,h,j). We find that CTCF does not interact with components of the Fanconi anaemia, that is, FANCA or FANCD2 and, unlike in *Ercc1*$^{-/-}$ MEFs (Supplementary Fig. 5D), CTCF does not translocate to heterochromatin in *Fanca*$^{-/-}$ or *Fancd2*$^{-/-}$ MEFs (Supplementary Fig. 5E). Importantly, exposure of MEFs to MMC led to substantially reduced ChIP signals on the ICRs/DMR and the promoters for all genes tested (Fig. 7a and Supplementary Fig. 6C). With the exception of the *Dlk1* promoter, we find normative ChIP signals for all gene promoters and ICRs/DMR tested in ATMi- and, in part, also in ATRi-treated MEFs exposed to MMC (Fig. 7a and Supplementary Fig. 6C), suggesting that such chromatin changes depend on functional DNA damage response signalling. In line with *Ercc1*$^{-/-}$ livers (Fig. 3h), exposure of C57Bl/6/SPRET/EiJ F1 MEFs to MMC that are polymorphic within the *H19* and *Peg3* ICRs led to substantially reduced ChIP signals for CTCF, SMC1A and SMC3 from the paternal allele (SPRET/EiJ) and the maternal (C57Bl/6) alleles of *Peg3* and the *H19* genes, respectively (Fig. 7b and Supplementary Fig. 7A) and to increased *Peg3* mRNA levels (Fig. 7c).

As for P15 *Ercc1*$^{-/-}$ livers, we find the loss of repressive histone H3K9 trimethylation and H3K27 trimethylation marks on the ICRs/DMR (Fig. 7d) and promoters (Supplementary Fig. 7B) in MMC- but not in UVC- or H_2O_2-treated MEFs. We also find the concomitant increase of activating acetylated histone H3K9 on the *H19* and *Peg3* ICRs and on all gene promoters tested (Supplementary Fig. 7C–D). Abrogation of ATM—and to a lesser extent also of ATR—led again to normative ChIP signals for these histone marks compared with untreated MEFs (Fig. 7d and Supplementary Fig. 7B–D). We find no difference for activating histone H3K4 trimethylation in *H19/Igf2*, *Peg3*, *Meg3/Dlk1* and *Grb10* ICRs/DMR or promoters in these cells (Supplementary Fig. 7C–D).

DISCUSSION

Using NER-defective animals and an *in vivo* biotinylation tagging approach in mice coupled to high-throughput proteomics, our data link ERCC1–XPF to important regulators of chromatin structure, to the proper occupancy of CTCF, the cohesin complex, MBD2 and ATRX at four imprinted loci and to the timely silencing of a subset of imprinted genes in the developing liver.

Importantly, ERCC1–XPF is not required for the assembly of the CTCF–cohesin complex and MBD2 in developing *Ercc1*$^{-/-}$ livers and is not involved in genome imprinting itself, a process already established in gametes. However, the heterodimeric complex is required for the recruitment of the CTCF–cohesin complex, MBD2 and ATRX to promoters and ICRs during hepatic development. Unlike CTCF or ATRX, the ERCC1–XPF complex is found on both alleles of *H19* and *Peg3* ICRs. This would allow ERCC1–XPF to interact with CTCF and/or ATRX for optimal gene silencing during embryogenesis or postnatal development.

A central aspect of these findings is their possible relevance to associated developmental disorders. A recently reported patient with ERCC1 deficiency presented with microcephaly, growth retardation and neurological abnormalities[17]. These pathological features also manifest in patients with CTCF haploinsufficiency[55], with Cornelia de Lange syndrome associated with mutations affecting the cohesin complex or, intriguingly, the RAD21 protein also involved in double-strand break repair[56] and in male children carrying a mutation in *ATRX*[57].

The involvement of CTCF and the cohesin in specific chromatin loop formation[49,58] and the influence of these structures in transcriptional regulation are well documented[36,59]. XPG and XPF were recently shown to be required for establishing CTCF-dependent chromatin looping between the promoter and terminator of the activated RARβ2 gene in HeLa cells[19,21]. This and the data presented herein suggest a similar role for ERCC1–XPF in facilitating long-range looping, thereby altering gene expression in the developing liver. The lack of any comparable defects in the single NER mutant *Csb*$^{m/m}$, *Xpa*$^{-/-}$ or *Xpc*$^{-/-}$ livers suggests that this mechanism does not require functional NER.

As for *Ercc1*$^{-/-}$ mice, exposure of cells to DNA ICLs—but not to other types of DNA lesion also repaired by NER—triggers chromatin changes and aberrant histone post-translational modifications associated with active transcription and the localization of CTCF and ATRX to heterochromatin. In line, animals carrying inborn defects in *Ercc1*, *Xpf* or *Xpg* (required for DNA damage incision) are the only single gene mutations in NER known to manifest with severe developmental defects[60–62]. All other single NER mutations result in phenotypically healthy animals with minor to moderate progeroid features[63]. The heterochromatic localization of CTCF and ATRX was accompanied by the dissociation of the CTCF–cohesin complex, MBD2 and ATRX from promoters and ICRs in MMC-treated or *Ercc1*$^{-/-}$ MEFs and in postnatal *Ercc1*$^{-/-}$ livers. Inactivation of ATM, and to a lesser extent of ATR, abolished the DNA damage-driven accumulation of CTCF and ATRX to heterochromatin, the release of repressor complexes from promoters and ICRs and the optimal gene silencing of the active allele. Thus, the DNA damage-driven reorganization of chromatin structure is reversible and requires functional DNA damage response. Indeed, the high affinity of the ERCC1–XPF complex for persistent DNA ICLs[64] in MMC-treated MEFs could trigger the displacement of the heterodimeric complex from high-order chromatin structures relevant for gene regulation to DNA damage sites requiring DNA repair (Fig. 7d).

It has been challenging to delineate how DNA damage drives the onset of tissue-specific, developmental defects. Here, we provide evidence for a functional link between ERCC1–XPF, DNA ICLs, chromatin architecture and gene silencing during hepatic development. Further studies are necessary to reveal how chromatin organizers respond to DNA damage during development or with disease onset. □

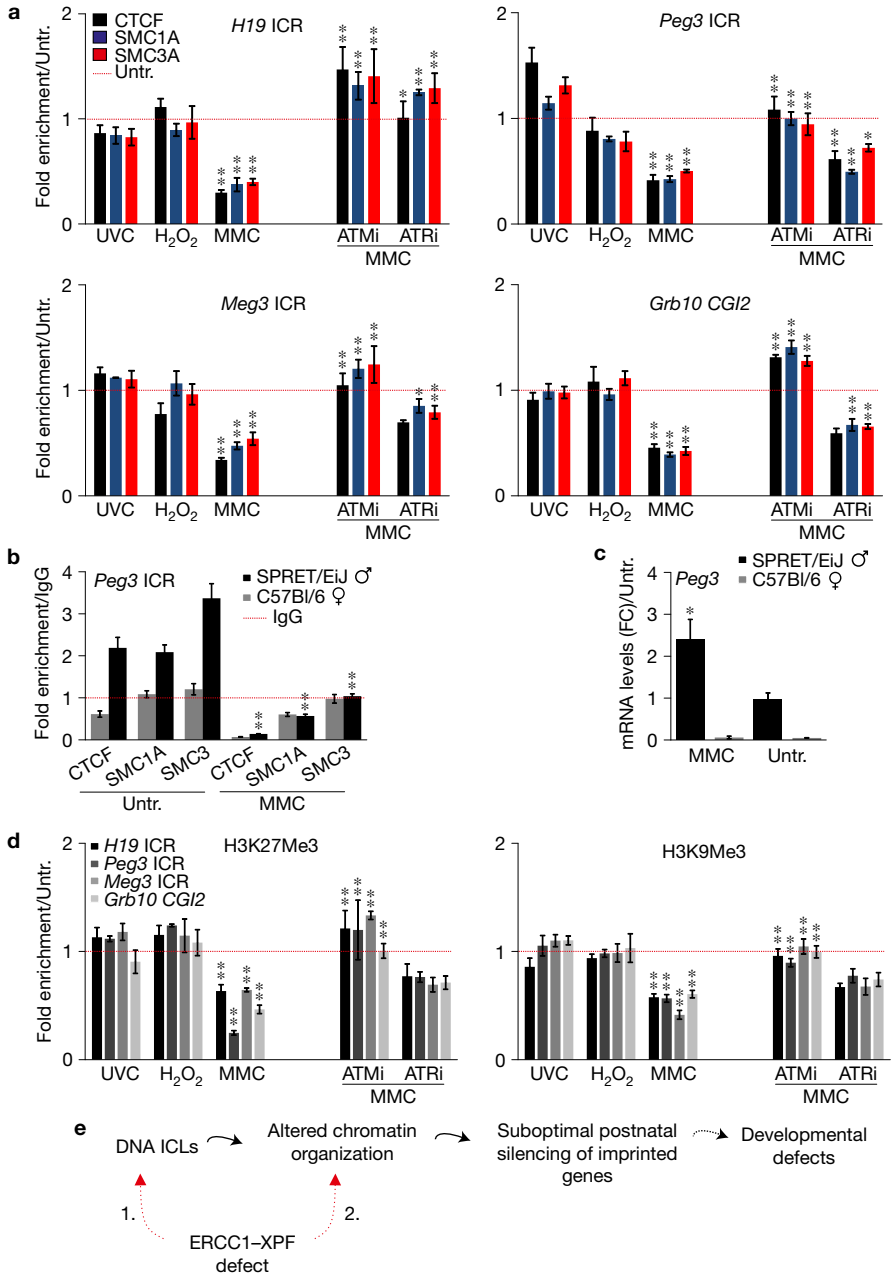


Figure 7 A working hypothesis for ERCC1–XPF, DNA damage and aberrant postnatal expression of imprinted genes during hepatic development. (**a**) ChIP signals of CTCF, SMC1A and SMC3 at the *H19*, *Peg3*, *Meg3* and *Grb10* ICRs/DMR in primary MEFs exposed to MMC, UV, H_2O_2, or to MMC and ATM or ATR inhibitor (ATMi, ATRi; $n=3$ biological replicates). ChIP signals are shown as in Fig. 5c. For ATMi- and ATRi-treated MEFs, ChIP signals are compared with MMC-treated MEFs for significance. Error bars indicate s.e.m. between 3 biological replicates; two-tailed *t*-test. (**b**). Allele-specific ChIP signals of CTCF, SMC1A and SMC3 at the *Peg3* ICR in MMC-treated C57BL/6/SPRET/Eij MEFs ($n=3$ biological replicates per genotype). ChIP signals are shown as in Fig. 5a. Error bars indicate s.e.m. among replicates ($n\geq 3$); two-tailed *t*-test. (**c**) Allele-specific mRNA levels for the *Peg3* gene in SPRET/EiJ-C57BL/6 MEFs exposed to MMC; $n=3$ biological replicates per genotype. Error bars, s.d.; two-tailed *t*-test. ChIP signals are shown as in Fig. 5c. (**d**) ChIP signals of repressive H3K27me and H3K9me3 histone marks at the *H19*, *Peg3*, *Meg3* and *Grb10* ICRs/DMR in primary MEFs exposed to MMC, UV or H_2O_2 and in MMC-treated MEFs treated with ATMi or ATRi ($n=3$ biological replicates). To test for significance, ChIP signals of MMC-treated MEFs are compared against untreated cells; ChIP signals of ATMi- and ATRi-treated MEFs are compared against those of MMC-treated MEFs. Error bars, s.d.; two-tailed *t*-test. The red dotted line indicates the untreated control. (**e**) A working model for ERCC1–XPF, irreparable DNA ICLs and aberrant postnatal expression of imprinted genes that are normally silenced in mammalian development. An inborn defect in ERCC1–XPF leads to the accumulation of persistent DNA damage (red arrow; '1') triggering chromatin changes that affect gene expression programs associated with mammalian development. ERCC1–XPF also maintains a functional role in chromatin architecture; the heterodimeric complex cooperates with chromatin organizers to facilitate the postnatal silencing (red arrow; '2') of a subset of imprinted genes during mammalian development. These findings suggest an attractive scenario whereby DNA damage in NER-proficient cells triggers the displacement of the ERCC1–XPF complex from chromatin structures involved in gene regulation to DNA damage sites requiring DNA repair. $^{*}P\leq 0.05$, $^{**}P\leq 0.01$. Statistical source data are provided in Supplementary Table 6.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of this paper.

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

The THALIS ESPA 2007–2013 'GenAge' (MIS380228) and 'miREG' (MIS380247), the ARISTEIA 'TagNER' (45) and 'Epilogeas' (3446), the FP7 Marie Curie ITN 'aDDRess (GA316390), 'CodeAge' (GA316354), 'Marriage' (GA316964) and 'Chromatin3D (GA622934), and the Horizon 2020 ERC Consolidator grant 'DeFiNER' (GA64663) supported this work. G.A.G. was supported by the EMBO Young Investigator programme. I.K. was supported by the Maria-Michail Manassakis fellowship. We thank M. Fousteri for providing the reagents for the unscheduled DNA synthesis assay and the Fanconi Anemia Research Fund (FA Cell Repository and the FA Antibody Project) for anti-FANCA and anti-FANCD2 antibodies and corresponding mutant MEFs.

AUTHOR CONTRIBUTIONS

G.C., Z.A., T.A.-P., A.I., M.T., M.A., I.K. and T.K. performed the experiments and analysed data. T.K., G.L.P. and J.S. generated new reagents. G.A.G. interpreted data and wrote the manuscript. All relevant data are available from the authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://dx.doi.org/10.1038/ncb3499

1. Harper, J. W. & Elledge, S. J. The DNA damage response: ten years after. *Mol. Cell* **28,** 739–745 (2007).
2. Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411,** 366–374 (2001).
3. Gregg, S. Q., Robinson, A. R. & Niedernhofer, L. J. Physiological consequences of defects in ERCC1–XPF DNA repair endonuclease. *DNA Rep.* **10,** 781–791 (2011).
4. Marteijn, J. A., Lans, H., Vermeulen, W. & Hoeijmakers, J. H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* **15,** 465–481 (2014).
5. van Duin, M. *et al.* Molecular characterization of the human excision repair gene ERCC-1: cDNA cloning and amino acid homology with the yeast DNA repair gene RAD10. *Cell* **44,** 913–923 (1986).
6. Sijbers, A. M. *et al.* Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86,** 811–822 (1996).
7. Hoy, C. A., Thompson, L. H., Mooney, C. L. & Salazar, E. P. Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res.* **45,** 1737–1743 (1985).
8. Klein Douwel, D. *et al.* XPF-ERCC1 acts in unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol. Cell* **54,** 460–471 (2014).
9. Niedernhofer, L. J. *et al.* The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Mol. Cell. Biol.* **24,** 5776–5787 (2004).
10. Bergstralh, D. T. & Sekelsky, J. Interstrand crosslink repair: can XPF-ERCC1 be let off the hook? *Trends Genet.* **24,** 70–76 (2008).
11. Kamileri, I., Karakasilioti, I. & Garinis, G. A. Nucleotide excision repair: new tricks with old bricks. *Trends Genet.* **28,** 566–573 (2012).
12. Bootsma, D., Kraemer, K. H., Cleaver, J. E. & Hoeijmakers, J. H. J. in *The Genetic Basis of Human Cancer* (eds Vogelstein, B. & Kinzler, K. W.) 245–274 (McGraw-Hill, 1998).
13. Bootsma, D., Kraemer, K. H., Cleaver, J. E. & Hoeijmakers, J. H. J. *The Metabolic and Molecular Basis of Inherited Disease* (McGraw-Hill, 2001).
14. Itin, P. H., Sarasin, A. & Pittelkow, M. R. Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes. *J. Am. Acad. Dermatol.* **44,** 891–920 (2001).
15. Garinis, G. A., van der Horst, G. T., Vijg, J. & Hoeijmakers, J. H. DNA damage and ageing: new-age ideas for an age-old problem. *Nat. Cell Biol.* **10,** 1241–1247 (2008).
16. Cleaver, J. E., Thompson, L. H., Richardson, A. S. & States, J. C. A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. *Hum. Mutat.* **14,** 9–22 (1999).
17. Jaspers, N. G. *et al.* First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am. J. Hum. Genet.* **80,** 457–466 (2007).
18. Bootsma, D., Kraemer, K. H., Cleaver, J. E. & Hoeijmakers, J. H. J. in *The Metabolic and Molecular Basis of Inherited Disease* Vol. 1 (eds Scriver, C. R. *et al.*) 677–703 (McGraw-Hill, 2001).
19. Le May, N. *et al.* NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol. Cell* **38,** 54–66 (2010).
20. Le May, N., Egly, J. M. & Coin, F. True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J. Nucleic Acids* **2010,** 616342 (2010).
21. Le May, N., Fradin, D., Iltis, I., Bougneres, P. & Egly, J. M. XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol. Cell* **47,** 622–632 (2012).
22. Fong, Y. W. *et al.* A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. *Cell* **147,** 120–131 (2011).
23. Kamileri, I. *et al.* Defective transcription initiation causes postnatal growth failure in a mouse model of nucleotide excision repair (NER) progeria. *Proc. Natl Acad. Sci. USA* **109,** 2995–3000 (2012).
24. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **8,** 921–929 (1999).
25. O'Gorman, S., Dagenais, N. A., Qian, M. & Marchuk, Y. Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells. *Proc. Natl Acad. Sci. USA* **94,** 14602–14607 (1997).
26. Katsantoni, E. Z. *et al.* Ubiquitous expression of the rtTA2S-M2 inducible system in transgenic mice driven by the human hnRNPA2B1/CBX3 CpG island. *BMC Dev. Biol.* **7,** 108 (2007).
27. Tian, M., Shinkura, R., Shinkura, N. & Alt, F. W. Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol. Cell. Biol.* **24,** 1200–1205 (2004).
28. McWhir, J., Selfridge, J., Harrison, D. J., Squires, S. & Melton, D. W. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Genet.* **5,** 217–224 (1993).
29. Selfridge, J., Hsia, K. T., Redhead, N. J. & Melton, D. W. Correction of liver dysfunction in DNA repair-deficient mice with an ERCC1 transgene. *Nucleic Acids Res.* **29,** 4541–4550 (2001).
30. Niedernhofer, L. J. *et al.* A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444,** 1038–1043 (2006).
31. Bonora, G., Plath, K. & Denholtz, M. A mechanistic link between gene regulation and genome architecture in mammalian development. *Curr. Opin. Genet. Dev.* **27,** 92–101 (2014).
32. Kernohan, K. D. *et al.* ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Dev. Cell* **18,** 191–202 (2010).
33. Phillips, J. E. & Corces, V. G. CTCF: master weaver of the genome. *Cell* **137,** 1194–1211 (2009).
34. Niedernhofer, L. J. *et al.* A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444,** 1038–1043 (2006).
35. Lewis, A. & Murrell, A. Genomic imprinting: CTCF protects the boundaries. *Curr. Biol.* **14,** R284–R286 (2004).
36. Nativio, R. *et al.* Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* **5,** e1000739 (2009).
37. Barlow, D. P. & Bartolomei, M. S. Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* **6,** 6:a018382 (2014).
38. Lui, J. C., Finkielstain, G. P., Barnes, K. M. & Baron, J. An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **295,** R189–R196 (2008).
39. Schumacher, B. *et al.* Delayed and accelerated aging share common longevity assurance mechanisms. *PLoS Genet.* **4,** e1000161 (2008).
40. van der Pluijm, I. *et al.* Impaired genome maintenance suppresses the growth hormone–insulin-like growth factor 1 axis in mice with Cockayne syndrome. *PLoS Biol.* **5,** e2 (2006).
41. de Boer, J. *et al.* Premature aging in mice deficient in DNA repair and transcription. *Science* **296,** 1276–1279 (2002).
42. Charalambous, M. *et al.* Disruption of the imprinted Grb10 gene leads to disproportionate overgrowth by an Igf2-independent mechanism. *Proc. Natl Acad. Sci. USA* **100,** 8292–8297 (2003).
43. Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. & Tilghman, S. M. Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* **375,** 34–39 (1995).
44. Li, L. *et al.* Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science* **284,** 330–333 (1999).
45. Moon, Y. S. *et al.* Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. *Mol. Cell. Biol.* **22,** 5585–5592 (2002).
46. Karakasilioti, I. *et al.* DNA damage triggers a chronic autoinflammatory response, leading to fat depletion in NER progeria. *Cell Metab.* **18,** 403–415 (2013).
47. Plasschaert, R. N. & Bartolomei, M. S. Genomic imprinting in development, growth, behavior and stem cells. *Development* **141,** 1805–1813 (2014).
48. Kurukuti, S. *et al.* CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl Acad. Sci. USA* **103,** 10684–10689 (2006).
49. Murrell, A., Heeson, S. & Reik, W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat. Genet.* **36,** 889–893 (2004).

50. Snapp, R. R. *et al.* Spatial organization of fibroblast nuclear chromocenters: component tree analysis. *J. Anat.* **223,** 255–261 (2013).
51. Westerveld, A. *et al.* Molecular cloning of a human DNA repair gene. *Nature* **310,** 425–429 (1984).
52. Hoy, C. A., Thompson, L. H., Mopney, C. L. & Salazar, E. P. Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res.* **45,** 1737–1743 (1985).
53. Garinis, G. A. *et al.* Transcriptome analysis reveals cyclobutane pyrimidine dimers as a major source of UV-induced DNA breaks. *EMBO J.* **24,** 3952–3962 (2005).
54. Ding, J., Miao, Z. H., Meng, L. H. & Geng, M. Y. Emerging cancer therapeutic opportunities target DNA-repair systems. *Trends Pharmacol. Sci.* **27,** 338–344 (2006).
55. Gregor, A. *et al. De novo* mutations in the genome organizer CTCF cause intellectual disability. *Am. J. Hum. Genet.* **93,** 124–131 (2013).
56. Boyle, M. I., Jespersgaard, C., Brondum-Nielsen, K., Bisgaard, A. M. & Tumer, Z. Cornelia de Lange syndrome. *Clin. Genet.* **88,** 1–12 (2015).
57. Gibbons, R. Alpha thalassaemia-mental retardation, X linked. *Orphan. J. Rare Dis.* **1,** 15 (2006).
58. Dowen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159,** 374–387 (2014).
59. Ling, J. Q. *et al.* CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312,** 269–272 (2006).
60. Niedernhofer, L. J. Nucleotide excision repair deficient mouse models and neurological disease. *DNA Repair (Amst.)* **7,** 1180–1189 (2008).
61. Shiomi, N. *et al.* Identification of the XPG region that causes the onset of Cockayne syndrome by using Xpg mutant mice generated by the cDNA-mediated knock-in method. *Mol. Cell Biol.* **24,** 3712–3719 (2004).
62. Tian, M., Shinkura, R., Shinkura, N. & Alt, F. W. Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol. Cell Biol.* **24,** 1200–1205 (2004).
63. van der Pluijm, I. *et al.* Impaired genome maintenance suppresses the growth hormone–insulin-like growth factor 1 axis in mice with Cockayne syndrome. *PLoS Biol.* **5,** e2 (2007).
64. McNeil, E. M. & Melton, D. W. DNA repair endonuclease ERCC1–XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic Acids Res.* **40,** 9990–10004 (2012).

METHODS

Biotin-tagged XPF and NER mutant animals. To generate the targeting vector for the insertion/knock-in of the Avi tag cassette before the stop codon of the last exon of the XPF gene, for the generation of the av*Xpf* knock-in mice, the following approach was used: PCR products were amplified using Phusion High-Fidelity DNA Polymerase (NEB). A quadruple ligation reaction was set up using the fragments: 5′ homology (BamHI/EcoRI 2.1-kb); Avi tag (EcoRI/Hind III 0.18-kb); lox–neomycin–lox cassette (HindIII/SalI 1.5-kb); pBSSK (BamHI/SalI 2.9-kb). The 3′ homology region (Apa I fragment 2.7-kb) was then subcloned into the vector followed by cloning of the MC1-TK gene (SacII 1.8-kb) for negative selection. The final targeting vector was linearized using NotI and used for embryonic stem cell electroporation. 129/SV embryonic stem cells carrying the Protamine 1-Cre transgene were maintained in their undifferentiated state (LIF-ESGRO 10^7 units) and grown on a feeder layer of gamma-irradiated (3,500 rads) G418^r primary mouse embryonic fibroblasts. Embryonic stem cells (0.8×10^7) were electroporated (400 V, 25 μF) with 50 μg of Not I linearized targeting vector (2 mg ml^{-1}) and homologous recombined clones were selected with G418 (300 μg ml^{-1}) and ganciclovir (2 μM). G418-resistant embryonic stem cell clones were subjected to Southern blot analysis and hybridized with 5′ and 3′ probes from their homology region. Clones with the correct homologous recombination were expanded to confirm their integrity and karyotyped to verify their euploid karyotype. Positive clones tested negative for mycoplasma (Venor GeM) were used for C57/BL6 blastocyst injection to generate chimaeric mice. Genomic DNA from embryonic stem clones was digested overnight with BamHI (MINOTECH Biotechnology) and resolved on 1% agarose gels. Samples were immobilized on Hybond-N+ nylon membranes (Amersham Bioscience) and hybridized with probes with [^{32}P]dCTP (Izotop). 5′ (1.2-kb SacI) and 3′ (1.2-kb BglII/Hind III) specific probes flanking the last exon of the XPF gene were used to identify the targeted (3.4-kb or 4.6-kb) and wild-type allele (6.6-kb). Chimaeric males were bred to C57BL/6 wild-type females for germ-line transmission. Offspring were screened by PCR for neo-deletion using primers F1: 5′-AAGAACTGTCGCTCCCTGATGAAC-3′ and R1 5′-CCTGGGGGGGAA AGAATGAATTGCT-3′ (Fig. 1a). Expression of Protamine-1 Cre transgene in the male germ line resulted in the deletion of the floxed neomycin gene in the first two out of five pups born, leaving behind a single loxP site after the Avi tag cassette. The cre recombinase transgene, derived from the PC3 embryonic stem cell background, was bred out in the process of backcrossing to C57BL/6 mice. The BirA transgenic mice were generated by PCR cloning a 3×HA NLS BirA cDNA into the HindIII/SalI restriction sites of the A2UCOE-EV-1 construct (a kind gift from M. Antoniou, Kings College, UK), which comprises the CpG methylation-free island of the human hnRNPA2B1–CBX3 locus (Supplementary Fig. 1A), previously shown to drive near-ubiquitous expression of transgenes by protecting from adverse chromosomal position effects at the sites of transgene integration[26,65]. Three independent transgenic lines were established and on the basis of the expression profile of the BirA transgene (Supplementary Fig. 1A), one transgenic line was selected for further crosses with the biotin tag XPF knock-in mice. For allele-specific expression studies, F1 pups were obtained by mating SPRET/EiJ male mice (The Jackson Laboratory) with *Mus musculus* female mice (bXPF). The generation and characterization of NER-deficient *Csb*$^{m/m}$, *Xpa*$^{-/-}$, *Csb*$^{m/m}$*; Xpa*$^{-/-}$, *Ercc1*$^{-/-}$ and *Xpd*^{TTD} mice has been previously described[11]. With the exception of *Ercc1*$^{-/-}$ mice, which were generated in an FVB:C57BL/6J (50:50) genetic background, all mice were in a C57BL/6J genetic background. Animals were kept on a regular diet and housed at the IMBB animal house, which operates in compliance with the 'Animal Welfare Act' of the Greek government, using the 'Guide for the Care and Use of Laboratory Animals' as its standard. As required by Greek law, formal permission to generate and use genetically modified animals was obtained from the responsible local and national authorities. All animal studies were approved by independent Animal Ethical Committees at FORTH and BSRC Al. Fleming.

Mass spectrometry studies. Proteins eluted from the beads were separated by SDS–PAGE electrophoresis on a 10% polyacrylamide gel and stained with colloidal blue silver (ThermoFisher Scientific). The entire lane was cut out and divided into at least 20 gel plugs, which were each further reduced to 1 mm^3 gel pieces and placed in low-bind tubes (Eppendorf UK). Proteins were in-gel-digested by using modified trypsin (Roche Diagnostics) in 50 mM ammonium bicarbonate. Peptide mixtures were analysed by nLC-ESI-MS/MS on a LTQ-Orbitrap XL coupled to an Easy nLC (Thermo Scientific). The sample preparation and the nLC-ESI-MS/MS analysis were performed as previously described[66] with minor modifications. Briefly, the dried peptides were dissolved in 0.5% formic acid aqueous solution, and the tryptic peptide mixtures were separated on a reversed-phase column (Reprosil Pur C18 AQ, Dr. Maisch GmbH), fused silica emitters 100 mm long with a 75 μm internal diameter (ThermoFisher Scientific) packed in-house using a packing bomb (Loader kit SP035, Proxeon). Tryptic peptides were separated and eluted in a linear water–acetonitrile gradient and injected into the MS.

Cells, colony formation and unscheduled DNA synthesis assays. Primary MEFs (P4) were cultured in standard medium containing Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 50 μg ml^{-1} streptomycin, 50 U ml^{-1} penicillin (Sigma) and 2 mM L-glutamine (Gibco). MEFs were rinsed with PBS, exposed to UVC irradiation (10 J m^{-2}), H_2O_2 (10 μM) or MMC (10 μg ml^{-1}) (Applichem) and cultured at 37 °C for 4 h prior to ChIP-qPCR, mRNA expression or immunofluorescence experiments. For ATM or ATR kinase inhibitor assays, cells were treated for 1 h with 10 μM inhibitor (Millipore) followed by the addition of MMC. For cell survival experiments, a total of 100–200 primary MEFs were seeded in 10 cm Petri dishes. The next day, the cells were exposed to MMC treatment for 4 h or to UVC irradiation and incubated for 10 days. Colonies were stained with Coomassie blue (0.2% Coomassie blue, 50% methanol, 7% acetic acid), and the number of colonies was counted and expressed as a percentage of the treated cells relative to that of the untreated control. At least three independent survival experiments were performed using three dishes per dose. DNA repair synthesis was determined by 5-ethynyl-2′-deoxyuridine (EdU) incorporation. Primary MEFs grown on coverslips were globally UVC irradiated and incubated for 2 h in medium supplemented with 10 μM EdU. After EdU incorporation, cells were washed with PBS followed by fixation with 2% formaldehyde in PBS. The coverslips were blocked for 30 min with 10% FBS in PBS, followed by 1 h incubation with 10 mM sodium ascorbate and 4 mM $CuSO_4$ containing Alexa Fluor594 azide (ThermoFischer Scientific A10270) and DAPI staining. The number of EdU-positive cells among 200 cells was counted, and the percentage of EdU-positive cells relative to the total number of cells was calculated.

Immunofluorescence, antibodies, westerns blots, co-immunoprecipitation, ChIP, ChIP/re-ChIP and chromatin pulldown assays. Immunofluorescence experiments were performed as previously described[46,53]. Briefly, cells (primary MEFs) were plated in 60 mm plates and cultured in standard medium containing DMEM supplemented with 10% FBS or for serum starvation with 1% FBS, 50 μg ml^{-1} streptomycin, 50 U ml^{-1} penicillin (Sigma) and 2 mM L-glutamine (Gibco). Primary hepatocytes were obtained from 15-day-old mice. Livers were minced, digested with collagenase type IV at 37 °C for 30 min and cells were plated to attach on coverslips at 37 °C, followed by immunostaining. Primary MEFs (passage 4; P4) were exposed to DNA-damaging agents for ChIP-qPCR, mRNA expression or immunofluorescence experiments. Cells were rinsed with PBS, exposed to UVC light (10 J m^{-2}), H_2O_2 (10 μM) or mitomycin C (MMC; 10 μg ml^{-1}) (Applichem) and cultured at 37 °C for 4 h. For ATM or ATR kinase inhibitor assays, cells were treated for 1 h with 10 μM inhibitor (Millipore) followed by the addition of MMC. Nuclear protein extracts from 15-day-old livers or cells were prepared as previously described[10] using the high-salt extraction method (10 mM HEPES–KOH pH 7.9, 380 mM KCl, 3 mM $MgCl_2$, 0.2 mM EDTA, 20% glycerol and protease inhibitors). For immunoprecipitation (IP) assays, nuclear lysates were diluted threefold by adding ice-cold HENG buffer (10 mM HEPES–KOH pH 7.9, 1.5 mM $MgCl_2$, 0.25 mM EDTA, 20% glycerol) and precipitated with antibodies overnight at 4°C followed by incubation for 3 h with protein G–Sepharose beads (Millipore). Normal mouse or rabbit IgG (Santa Cruz) was used as a negative control. Immunoprecipitates were washed five times (10 mM HEPES–KOH pH 7.9, 300 mM KCl, 0.3% NP40, 1.5 mM $MgCl_2$, 0.25 mM EDTA, 20% glycerol and protease inhibitors), eluted and resolved on 8% or 10% SDS–PAGE. Pulldowns were performed with 1.5–1.7 mg of nuclear extracts using M-280 paramagnetic streptavidin beads (Invitrogen) as previously described[23]. For ChIP assays, tissues or cells were crosslinked at room temperature for 5 or 2.5 min respectively with 1% formaldehyde. Chromatin was prepared and sonicated on ice for 15 min using a Covaris S220 Focused-ultrasonicator. Samples were immunoprecipitated with antibodies (5–6 μg) overnight at 4 °C followed by incubation for 3 h with protein G–Sepharose beads (Millipore) and washed sequentially. The complexes were eluted and the crosslinking was heat reversed. Purified DNA fragments were analysed by qPCR using sets of primers (primer sequences are listed in Supplementary Table 5) targeting different regions of *Igf2*, *Peg3*, *Dlk1* and *Grb10* genes. ChIP re-ChIP experiments were performed as described above with the following modifications after the first immunoprecipitation and washing: complexes were eluted with 10 mM dithiothreitol, 1% SDS in TE buffer for 30 min. Eluted samples were diluted 1:20 with re-ChIP buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 150 mM NaCl, 0.01% SDS and 1% Triton X-100) and immunoprecipitated overnight with the secondary antibody. Antibodies against HA (Y-11, western blotting (wb): 1:500), TFIIB (C-18, wb: 1:500), ERCC1 (D-10, wb: 1:500), ATRX (H-300, immunofluorescence (IF): 1:50) were from SantaCruz Biotechnology. γH2AX (05-636, IF: 1:12,000) and CTCF (07-729, wb: 1:1,000), H3K9Ac (07-352) and H3K27me3(07-449) were from Millipore. CTCF (clone 48, wb: 1:500, IF: 1:100) was from BD Transduction Laboratories. SMC1A (A300-055A, wb: 1:4,000, IF: 1:200), SMC3 (A300-060A, wb: 1:4,000, IF: 1:100) and MBD2 (A301-633A, wb: 1:500) were from Bethyl Laboratories. Streptavidin-HRP (wb: 1:12,000) was from Upstate Biotechnology. Nucleolin (ab22758, IF: 1:2,000), H3K4me3 (ab8580) and H3K9me3 (ab6001) were

DOI: 10.1038/ncb3499

from Abcam. Flag M2 (F3165, wb 1:2,000) was from Sigma-Aldrich. HP1a (2HP-1H5, IF: 1:200) was from Chemicon.

Gene expression studies. Microarray gene expression data on P15 NER-deficient (that is, $Csb^{m/m}$;$Xpa^{-/-}$, Xpd^{TTD}, $Ercc1^{-/-}$) and age-matched littermate control livers ($Csb^{m/m}$, $Xpa^{-/-}$ and WT) were collected from cited sources[23,30,40]. Total RNA was isolated from liver, heart, kidney, spleen and lung of P15 $Ercc1^{-/-}$ mice using a Total RNA isolation kit (Qiagen) as described by the manufacturer. Quantitative PCR (qPCR) was performed with a DNA Engine Opticon device according to the instructions of the manufacturer (MJ Research). The generation of specific PCR products was confirmed by melting curve analysis and gel electrophoresis. Each primer pair was tested with a logarithmic dilution of a cDNA mix to generate a linear standard curve (crossing point (CP) plotted versus log of template concentration), which was used to calculate the primer pair efficiency ($E = 10^{(-1/\text{slope})}$). Hypoxanthine guanine phosphoribosyltransferase1 (*Hprt-1*) mRNA was used as an external standard. For data analysis, the second derivative maximum method was applied: $(E_{\text{1gene of interest}}^{\Delta\text{CP(cDNA of WT mice-cDNA of } Ercc1^{-/-},\ Xpa^{-/-} \text{ or } Csb^{m/m} \text{ or } Csb^{m/m}/Xpa^{-/-}) \text{ gene of interest}})/(E_{\text{hprt-1}}^{\Delta\text{CP (cDNA WT mice-cDNA of } Xpa^{-/-} \text{ or } Csb^{m/m} \text{ or } Csb^{m/m}/Xpa^{-/-}) \text{ hprt-1}})$. Primer sequences are listed in Supplementary Table 5.

Allele-specific ChIP analysis. Allele-specific ChIP analysis was performed on P15 bXPF(C57BL/6)-SPRET/EiJ F1 livers. PCR amplification using unbiased primers was performed for the *H19* and *Peg3* ICR regions. The amplified product was purified and subsequently digested with specific restriction enzymes for 2 h. The digested DNA was resolved on a 2% agarose gel and visualized using an ultraviolet transilluminator (BioRad). Primer sequences are available in Supplementary Table 5. Allele-specific sequences were retrieved from the Mouse Genomes Project (Sanger Institute) and primers were designed with Batch Primer 3.0. All sets of primers were validated for allele specificity by qPCR amplification of strain/species pure or mixed genomic DNA from C57BL/6NJ, FVB/NJ and SPRET/Eij.

Bisulfite mutagenesis. Genomic DNA isolated from P15 livers of $Ercc1^{-/-}$ and littermate control mice was mutagenized with sodium bisulfite by using an EpiTect Bisulfite Conversion Kit (QIAGEN) according to the manufacturer's instructions. PCR amplification was carried out with primers specific for bisulfite-treated DNA. Primer sequences are available in Supplementary Table 5. All promoter and DMR regions were amplified by PCR. The resulting PCR products were ligated into the pCRII vector using a TOPO-TA cloning kit (Invitrogen), according to the manufacturer's instructions. Positive clones were sequenced, and clones were only accepted at >94% cytosine conversion. To ensure that the selected clones were derived from a unique DNA template we used non-converted cytosines and mismatched base pairs as reference points for each clone.

Data analysis. A two-tailed *t*-test was used to extract the statistically significant data by means of the IBM SPSS Statistics 19 (IBM), Spotfire (Tibco), Partek (Partek INCoR1porated) and R-statistical package (www.r-project.org). Significant over-representation of pathways and gene networks was determined by DAVID (http://david.abcc.ncifcrf.gov/summary.jsp; through BBID, BIOCARTA and KEGG annotations) as well as by means of the Ingenuity Pathway Analysis software (www.ingenuity.com). For mass spectrometry (MS), the MS/MS raw data were loaded in Proteome Discoverer 1.3.0.339 (ThermoFischer Scientific) and run using the Mascot 2.3.02 (Matrix Science) search algorithm against the *Mus musculus* theoretical proteome (last modified 6 July 2015) containing 46,470 entries in Uniprot. A list of common contaminants was included in the database[67]. For protein identification, the following search parameters were used: precursor error tolerance 10 ppm, fragment ion tolerance 0.8 Da, trypsin full specificity, maximum number of missed cleavages 3 and cysteine alkylation as a fixed modification. The resulting .dat and .msf files were subsequently loaded and merged in Scaffold (version 3.04.05, Proteome Software) for further processing and validation of the assigned MS/MS spectra. Thresholds for protein and peptide identification were set to 99% and 95% accordingly, for proteins with minimum 1 different peptides identified, resulting in a protein false discovery rate (FDR) of <0.1%. For single peptide identifications, we applied the same criteria in addition to manual validation of MS/MS spectra. Protein lists were constructed from the respective peptide lists through extensive manual curation based on previous knowledge. For label-free relative quantitation of proteins, we applied a label-free relative quantitation method between the different samples (control versus bait) to determine unspecific binders during the affinity purification. All .dat and .msf files created by Proteome Discoverer were merged in Scaffold where label-free relative quantification was performed using the total ion current (TIC) from each identified MS/MS spectra. The TIC is the sum of the areas under all the peaks contained in a MS/MS spectrum and total TIC value results by summing the intensity of the peaks contained in the peak list associated to a MS/MS sample. Protein lists containing the Scaffold-calculated total TIC quantitative value for each protein were exported to Microsoft Excel for further manual processing including categorization and additional curation based on previous knowledge. The fold change of protein levels was calculated by dividing the mean total TIC quantitative value in bait samples with the mean value of the control samples for each of the proteins. Proteins having ≥60% protein coverage, ≥1 peptide in each sample and a fold change ≥1.2 in all three measurements were selected as being significantly enriched in bXPF compared with BirA liver samples. Proteins that were significantly enriched in bait samples were considered these with P value ≤0.05 and a fold change ≥2. Significant over-representation of pathways, protein–protein interactions and protein complexes were derived by STRING[68] (http://string-db.org).

Statistics and reproducibility. Experiments were repeated at least 3 times. The data exhibited normal distribution (where applicable). There was no estimation of group variation before experiments. Error bars indicate standard deviation unless stated otherwise, that is, s.e.m. For animal studies, each biological replicate consists of 3–5 mouse tissues or cell cultures per genotype per time point or treatment. No statistical method was used to predetermine sample size. None of the samples or animals was excluded from the experiment. The animals or the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. The representative images shown in Figs 1d–f,i, 2a, 3a,b,f, 4c, 5b,d and 6a,c–h and Supplementary Figs 1A,B, 2B–G, 4 and 5 have been repeated more than 3 times.

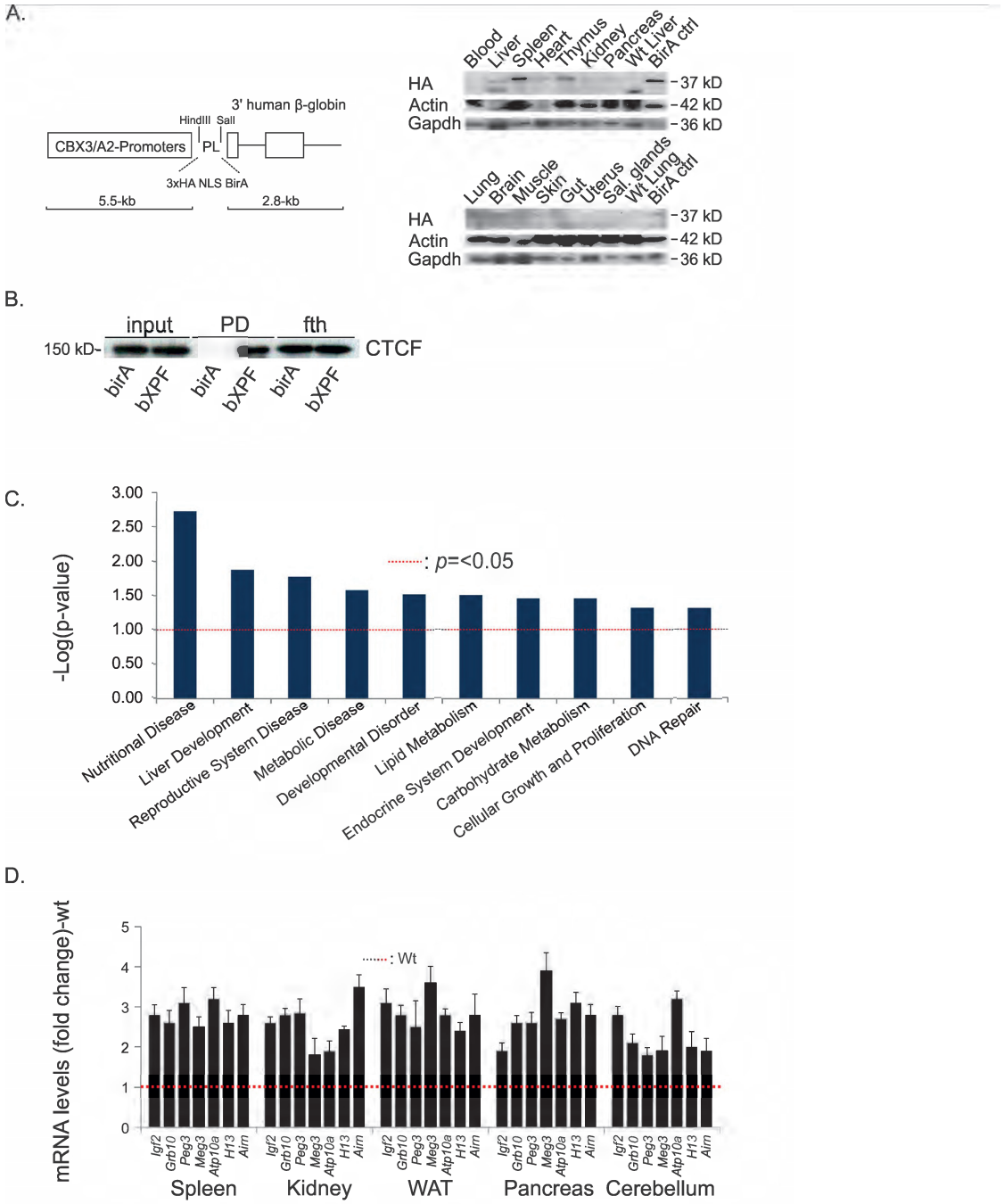
Data availability. Previously published microarray data that were reanalysed here are available under accession codes E-MEXP-3930 and E-MEXP-835 (ref. 23, 30,40). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[69] with the data set identifier PXD005897. Source data for Figs 1 and 3–7 and Supplementary Figs 1, 3, 4, 6 and 7 have been provided as Supplementary Table 6 (Statistics source data). All other data supporting the findings of this study are available from the corresponding author on reasonable request.

65. Antoniou, M. *et al.* Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. *Genomics* **82,** 269–279 (2003).
66. Aivaliotis, M. *et al.* Large-scale identification of N-terminal peptides in the halophilic archaea *Halobacterium salinarum* and *Natronomonas pharaonis*. *J. Proteome Res.* **6,** 2195–2204 (2007).
67. Rappsilber, J., Ryder, U., Lamond, A. I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12,** 1231–1245 (2002).
68. Szklarczyk, D. *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **39,** D561–D568 (2011).
69. Vizcaino, J. A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41,** D1063–D1069 (2013).

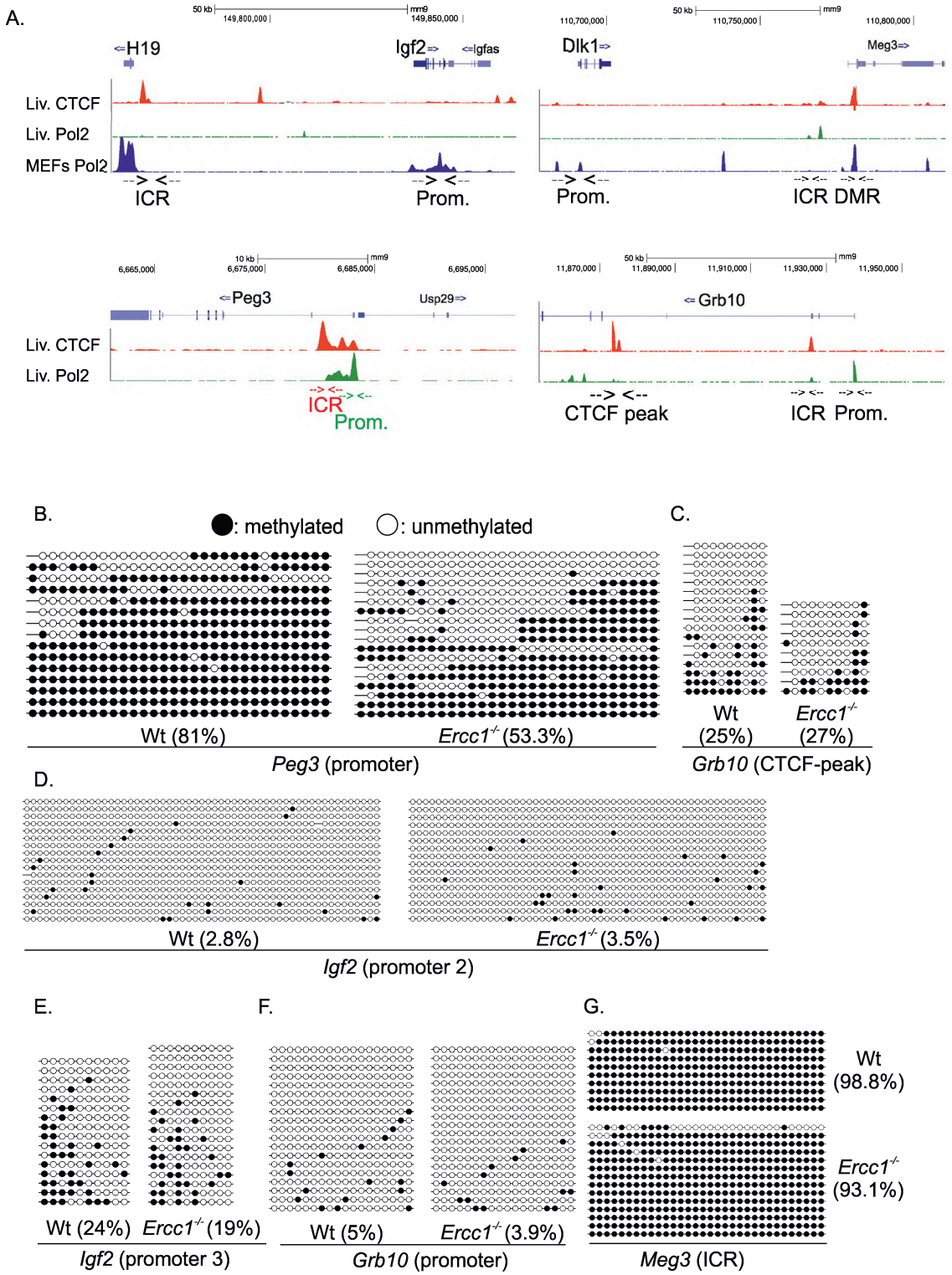
nature cell biology

DOI: 10.1038/ncb3499

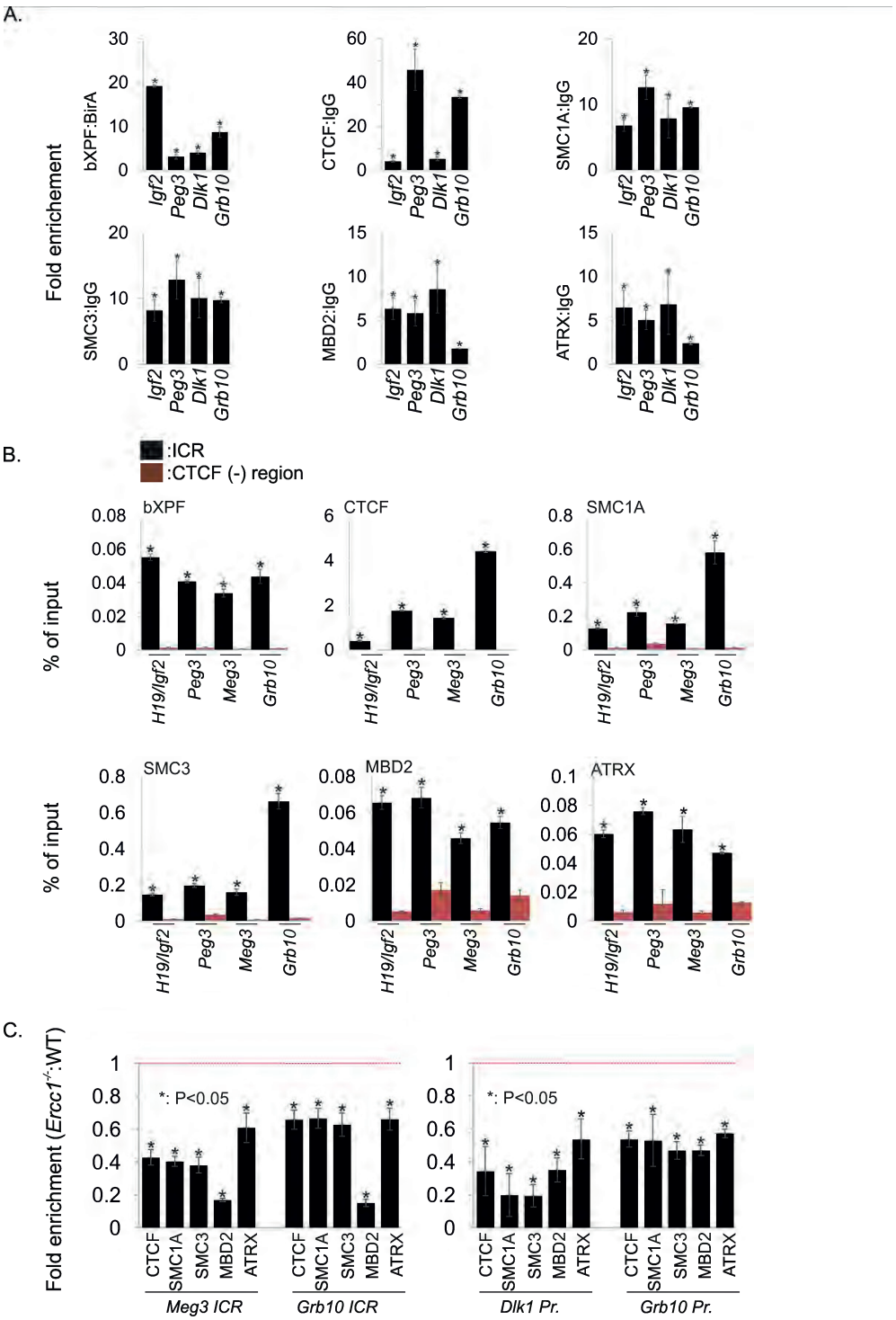
In the format provided by the authors and unedited.



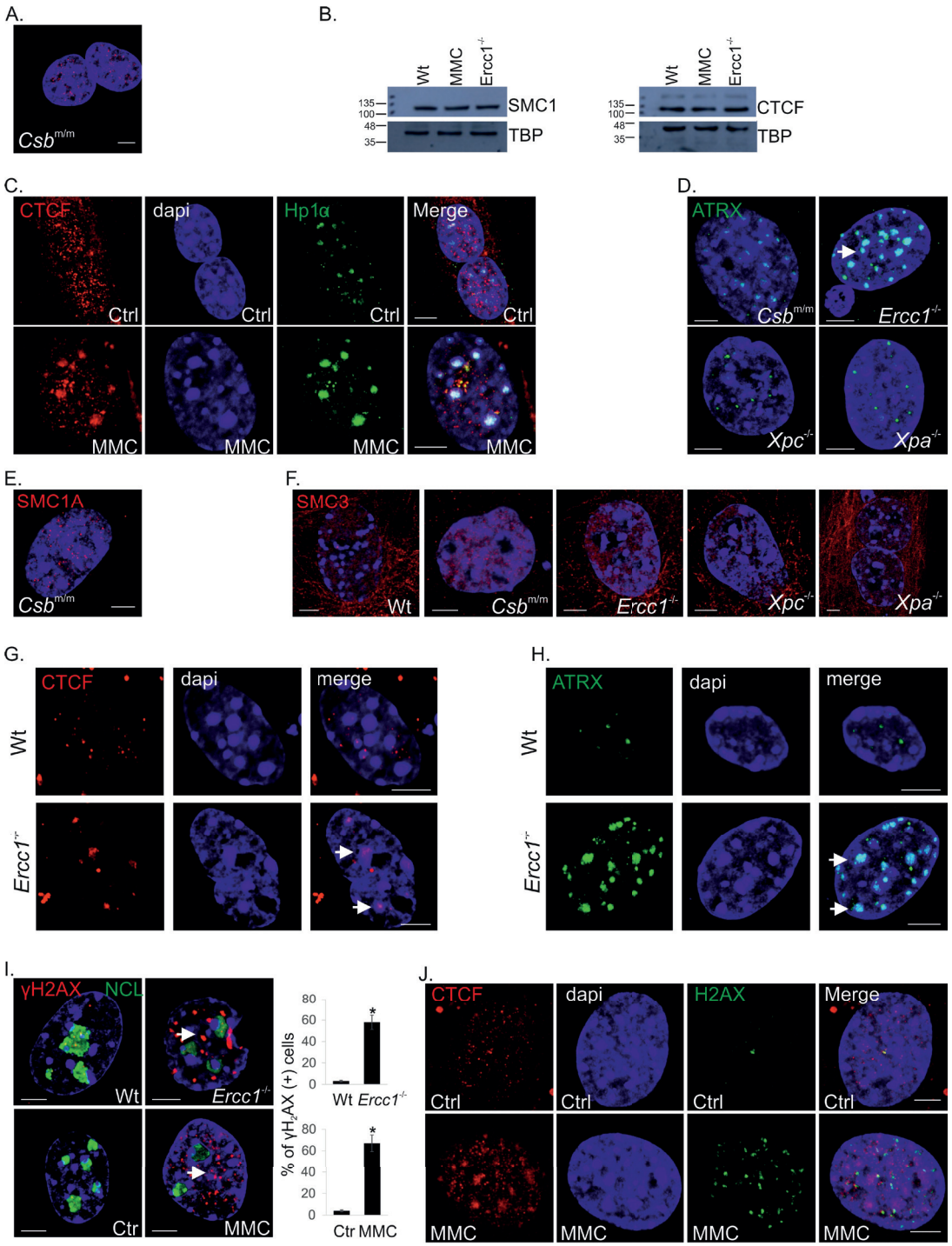
Supplementary Figure 1 Ablation of *Ercc1* gene triggers the aberrant silencing of imprinted genes during postnatal hepatic development. (A). Schematic representation of transgenic mice expressing the BirA biotin ligase transgene and anti-HA immunoblot showing expression of the BirA biotin ligase protein in different tissues of 2-month old BirA transgenic animals (as indicated). (B). Streptavidin pull-downs in nuclear or chromatin extracts under native (micrococcal nuclease digested) conditions derived from primary bXPF MEFs expressing the BirA transgene or the BirA transgenic animals (as indicated) and analysed by Western blotting for CTCF. (C). Over-represented biological processes derived from the significantly aberrantly expressed imprinted genes in P15 *Ercc1*$^{-/-}$ compared to age-matched wt livers; *p:* -log of p-value which is calculated by Fisher's exact test right-tailed. Red dotted line marks the threshold of significance at 0.05. (D). qPCR mRNA levels of imprinted genes in P15 *Ercc1*$^{-/-}$ spleen, kidney, white adipose tissue (WAT), pancreas and cerebellum (as indicated; n=3 biological replicates each representing a pool of 4-5 tissues/genotype). Red dotted line: wt mRNA levels. Error bars indicate S.D; two tailed t-test. Statistical source data are provided in Table S6.



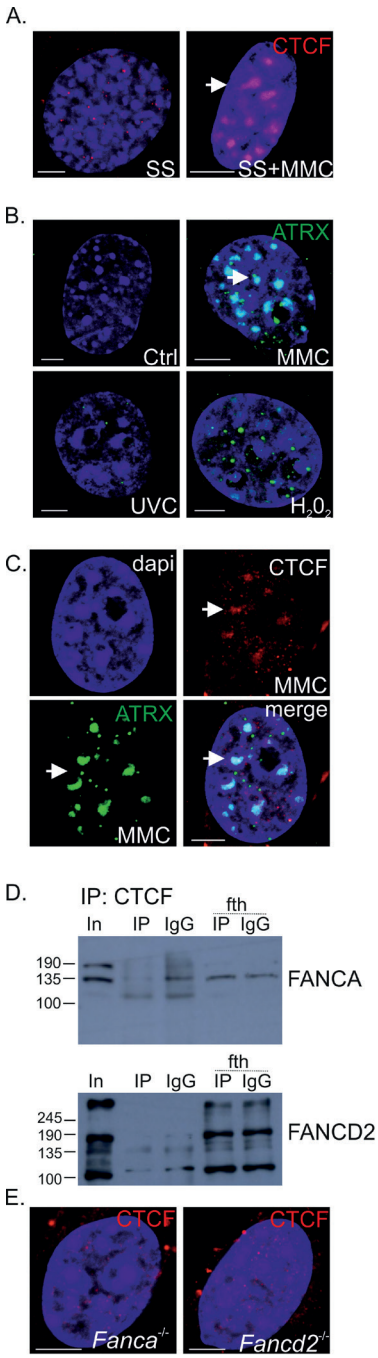
Supplementary Figure 2 DNA methylation at the I*gf2*, *Peg3*, *Dlk1* and *Peg3* proximal promoter regions in *Ercc1*$^{-/-}$ livers. (A). ChIP-seq profiles marking the recruitment of CTCF and RNAPII (Pol2) at the *H19*, *Dlk1*, *Peg3 and Grb10* promoter and ICRs/DMR regions in livers (liv) or mouse embryonic fibroblasts (MEFs) as indicated. Arrow heads mark the PCR amplified ICR or promoter (prom) regions (as indicated). (B-G). Schematic representation of the % of DNA methylation at the *Peg3* (promoter), *Grb10* (CTCF-peak), *Igf2* (promoter 2 and 3), *Grb10* (promoter) and *Meg3* (ICR) regions (as indicated) in P15 *Ercc1*$^{-/-}$ and wt livers. Black circles: methylated cytosine; open circles: unmethylated cytosine.



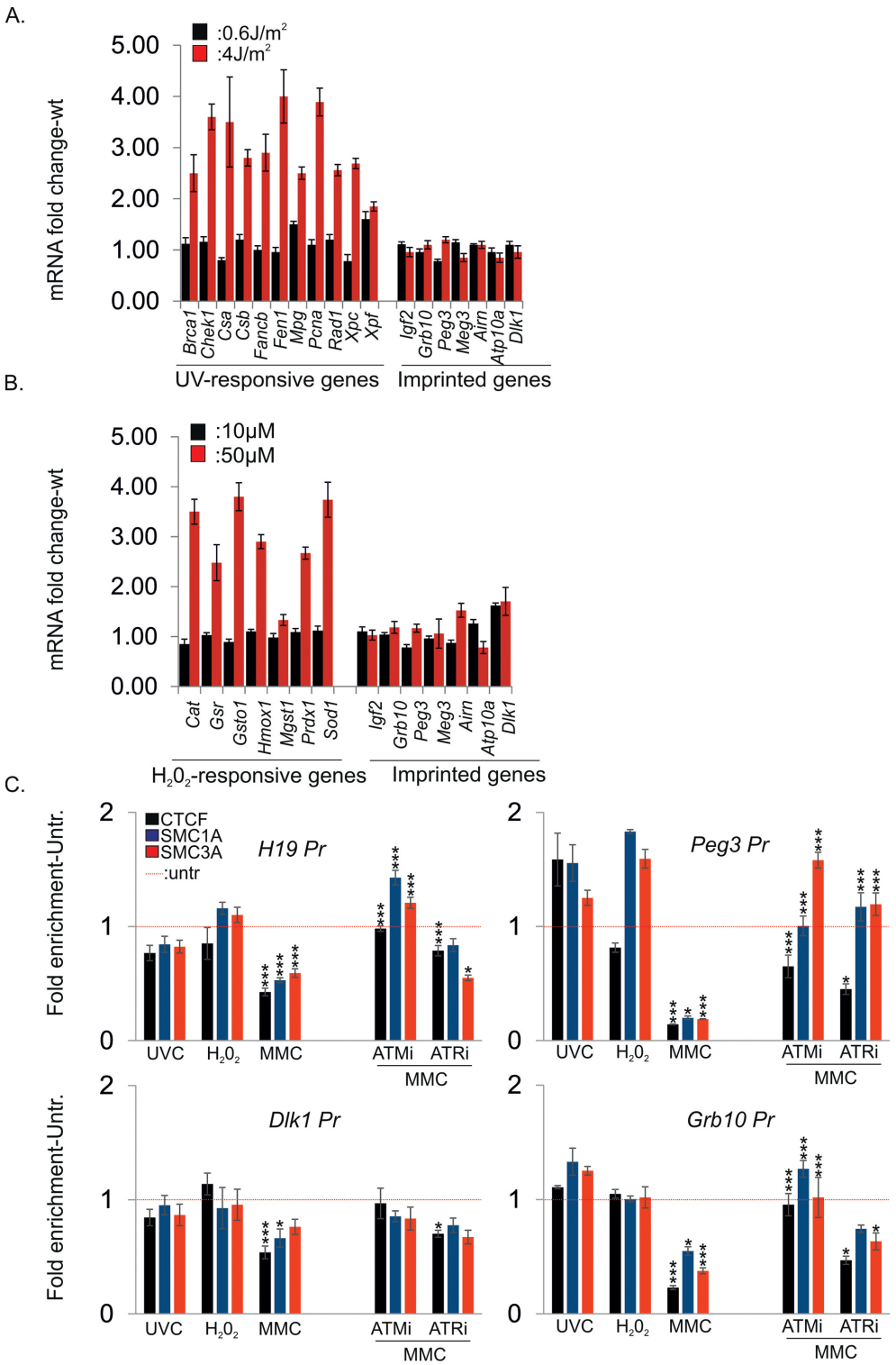
Supplementary Figure 3 Dissociation of the CTCF-cohesin complex and MBD2from the promoters and ICRs of imprinted genes in *Ercc1*$^{-/-}$ livers. (A). bXPF, CTCF, SMC1A, SMC3, MBD2 and ATRX ChIP signals expressed as fold enrichment over those obtained with BirA (for bXPF) or control antibody (IgG) at the *Igf2, Peg3, Dlk1* and *Grb10* promoters in P15 wt mouse livers (as indicated; n=3 biological replicates each representing a pool of 4-5 livers). Error bars indicate S.E.M. among replicates (n ≥ 3). (B). ChIP signals shown as % of input of CTCF, SMC1A, SMC3, MBD2 and ATRX at the *Igf2, Peg3, Meg3* and *Grb10* ICR regions and CTCF negative (-) regions (as indicated; n=3 biological replicates each representing a pool of 4-5 livers). Error bars indicate S.E.M. (C). CTCF, SMC1A, SMC3, MBD2 and ATRX ChIP signals normalized against their respective control antibody (IgG) and expressed as fold enrichment over the corresponding ChIP signals obtained for wt mouse livers at the *Dlk1/Meg3* and *Grb10* promoters and ICRs (as indicated; n=3 biological replicates each representing a pool of 4-5 livers). Error bars indicate S.E.M., *: p≤0.05; two-tailed t-test. Statistical source data are provided in Table S6.



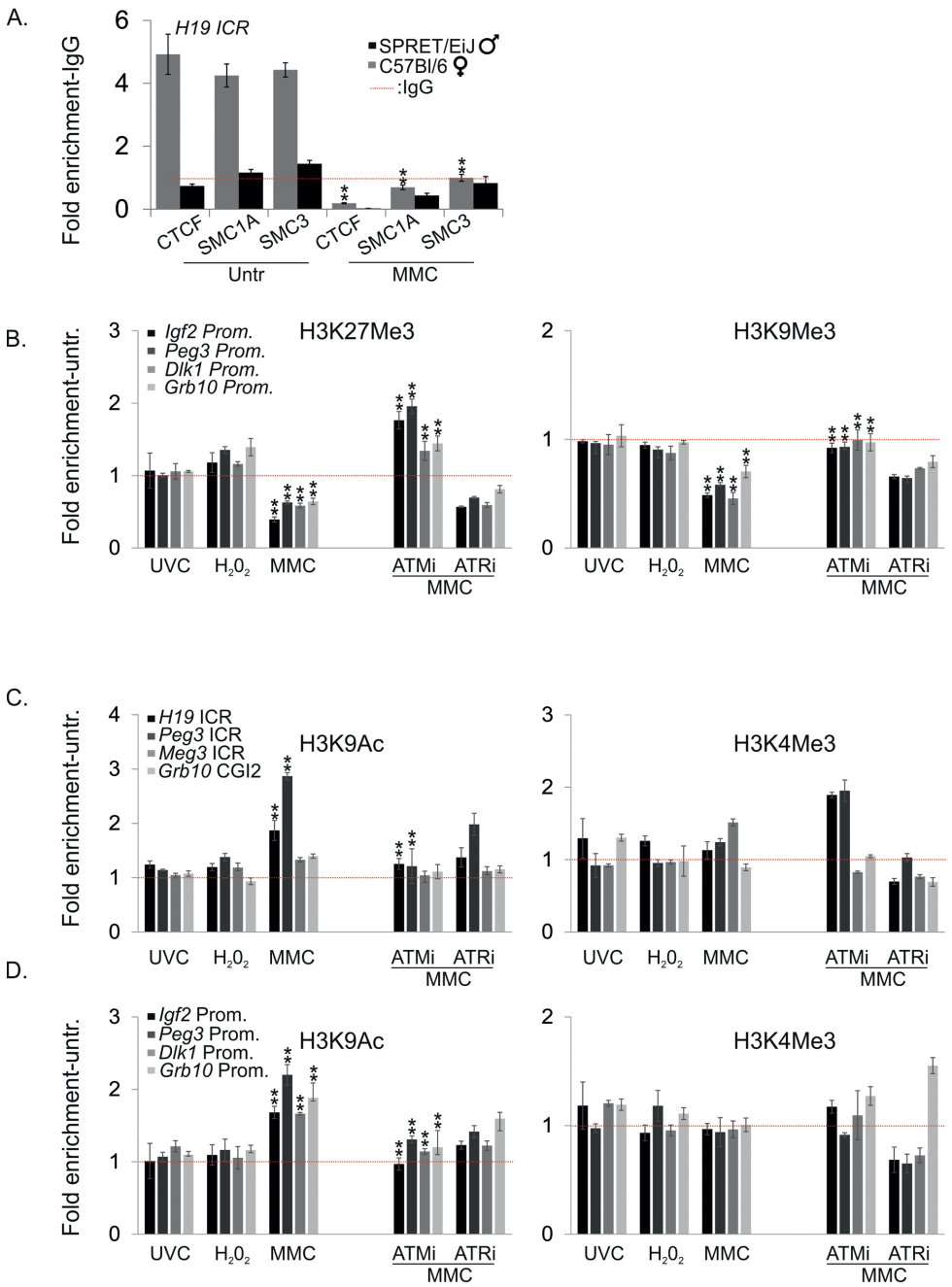
Supplementary Figure 4 Persistent DNA ICLs trigger aberrant CTCF and ATRX localization in *Ercc1*$^{-/-}$ and MMC-treated MEFs. (A). Immunofluorescence detection of CTCF in *Csb*$^{m/m}$ primary mouse embryonic fibroblasts (MEFs; 20 fields analysed from 3 biological replicates). (B). Equal amount of nuclear extracts from wt, MMC-treated and *Ercc1*$^{-/-}$ MEFs analysed by Western blotting for CTCF, SMC1A and TBP. (C). Immunocolocalization of CTCF and HP1a in MMC-treated and untreated P4 MEFs. (D). Immunofluorescence detection of ATRX in *Ercc1*$^{-/-}$, *Csb*$^{m/m}$, *Xpc*$^{-/-}$ and *Xpa*$^{-/-}$ primary mouse embryonic fibroblasts (MEFs; 20 fields analysed from 3 biological replicates). Note the distinctive accumulation of ATRX to heterochromatin in *Ercc1*$^{-/-}$ MEFs. (E). Immunofluorescence detection of SMC1A in primary *Csb*$^{m/m}$ MEFs. (F). Immunofluorescence detection of SMC3 in wt, *Ercc1*$^{-/-}$, *Csb*$^{m/m}$, *Xpc*$^{-/-}$ and *Xpa*$^{-/-}$ primary MEFs (as indicated; 20 fields analysed from 3 biological replicates). (G-H). Immunofluorescence detection of CTCF and ATRX in primary wt and *Ercc1*$^{-/-}$ hepatocytes; note the distinctive translocation of CTCF and accumulation of ATRX to heterochromatin *Ercc1*$^{-/-}$ hepatocytes. (**I**). Immunofluorescence detection of γH2aX and nucleolin in wt and *Ercc1*$^{-/-}$ MEFs (upper panel) and in MMC-treated and untreated control (ctrl) MEFs (lower panel); the graph depicts the average number of γH2aX-positive stained cells in wt and *Ercc1*$^{-/-}$ MEFs or wt MEFs exposed to MMC from 20 fields analysed representing n=3 biological replicates*: P-value≤0.05. (**J**). Immunocolocalization of CTCF and γH2aX in MMC-treated and untreated P4 MEFs. Scale bars, 5 μm. Statistical source data are provided in Table S6.



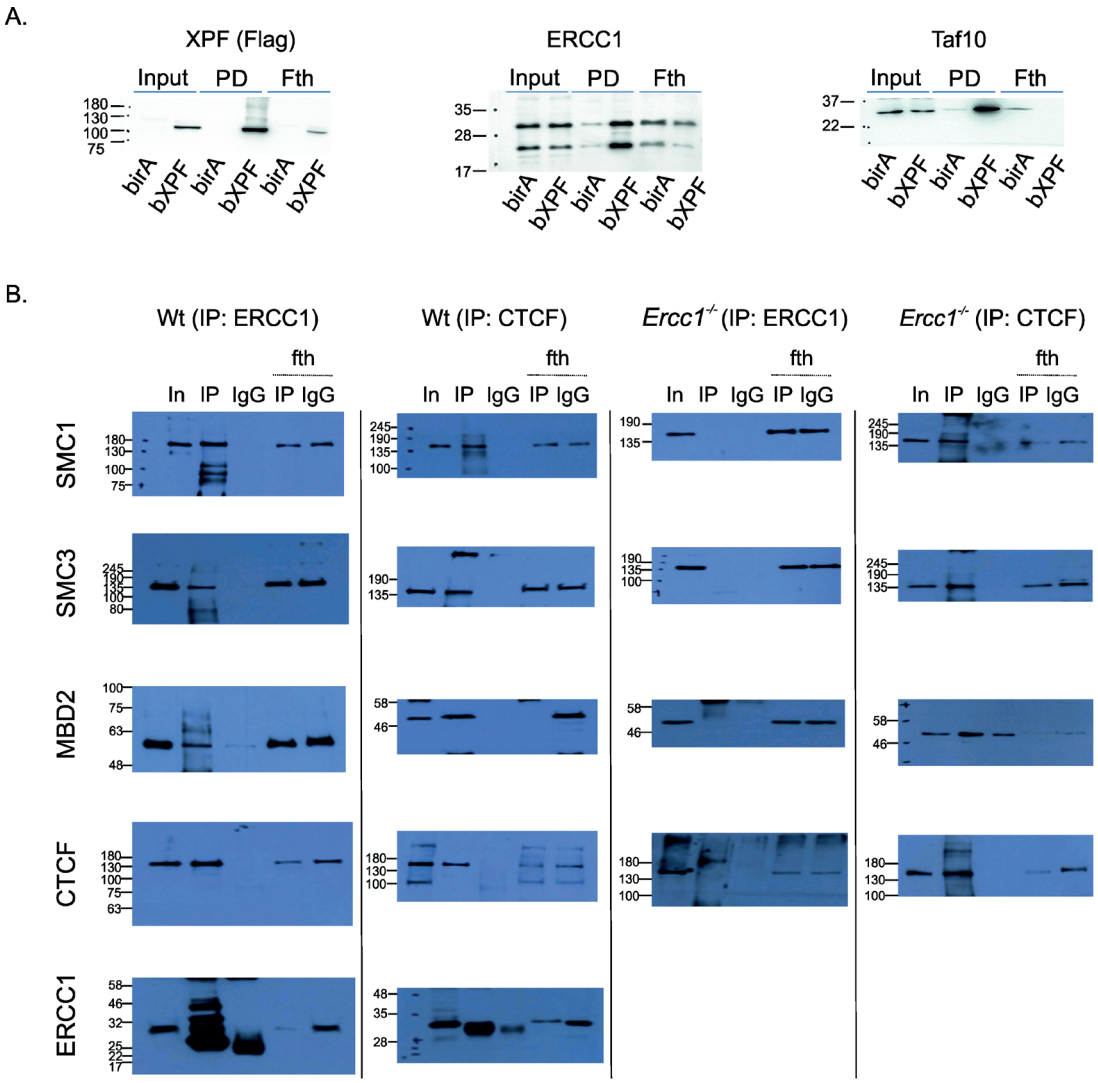
Supplementary Figure 5 Persistent DNA ICLs trigger aberrant CTCF and ATRX localization in *Ercc1*$^{-/-}$ and MMC-treated MEFs. (A). Immunofluorescence detection of CTCF in serum starved (SS) MEFs exposed to MMC (as indicated; 20 fields analysed from 3 biological replicates). (B). Immunofluorescence detection of ATRX in primary MEFs exposed to MMC, UV and H_2O_2; Ctr: untreated MEFs. Note the distinctive accumulation of ATRX to heterochromatin in MMC-treated MEFs. (C). Immunocolocalization of CTCF and ATRX in MMC-treated primary P4 MEFs. (D). Co-immunoprecipitation experiments using **α**CTCF in nuclear extracts from P15 livers analysed by Western blotting for FANCA or FANCD2. (E). Immunofluorescence detection of CTCF in *Fanca*$^{-/-}$ and *Fancd2*$^{-/-}$ MEFs (as indicated; 20 fields analysed from 3 biological replicates). Scale bars, 5 μm.



Supplementary Figure 6 Persistent DNA ICLs trigger the dissociation of the CTCF-cohesin complex and MBD2 from promoters and ICRs. (A). Quantitative (q) PCR mRNA levels of UV-responsive and imprinted genes in primary mouse dermal fibroblasts exposed to 0.6 and 4 J/m^2 of UVC irradiation (as indicated), n= 3 biological replicates/dose. Error bars indicate S.D. (B). qPCR mRNA levels of H_2O_2-responsive and imprinted genes in primary mouse dermal fibroblasts treated with 10μM or 50μM of H_2O_2 (as indicated), n= 3 biological replicates/dose. Error bars indicate S.D. (C). ChIP signals of CTCF, SMC1A and SMC3 (as indicated) at the *H19, Peg3, Dlk1* and *Grb10* promoters in primary MEFs exposed to UV, H_2O_2 and MMC or to MMC and ATM or ATR inhibitors (ATMi, ATRi as indicated). ChIP signals from treated MEFs were normalized to respective control antibody (IgG) which were set as 1 (red dotted line) and expressed as fold enrichment over those obtained from untreated MEFs, n= 3 biological replicates. Error bars indicate S.D.; two-sided t-test. Statistical source data are provided in Table S6.



Supplementary Figure 7 Persistent DNA ICLs trigger changes in histone marks associated with aberrant postnatal silencing. (A). Allele-specific ChIP signals of CTCF, SMC1A and SMC3 (as indicated) at the *H19* ICR in MMC-treated C57BL/6 /SPRET/Eij MEFs (as indicated). ChIP signals expressed as fold enrichment over those obtained with control antibody (IgG). Error bars indicate S.D. among biological replicates; (n = 3 biological replicates). (B). ChIP signals of repressive H3K27me and H3K9me3 histone marks at the *H19, Peg3, Dlk1* and *Grb10* promoters in primary MEFs exposed to MMC, UV or H_2O_2 (as indicated) and in MMC-treated MEFs treated with ATMi or ATRi (as indicated). ChIP signals are shown as in Figure 5C. To test for significance, ChIP signals of ATMi and ATRi-treated MEFs are compared against those of MMC-treated MEFs; two sided t-test. (C-D). ChIP signals of activating H3K-9Ac and H3K4me3 histone marks at the *H19/Igf2, Peg3, Meg3/Dlk1* and *Grb10* ICRs and promoters in primary MEFs exposed to MMC, UV or H_2O_2 (as indicated) and in MMC-treated MEFs treated with ATMi or ATRi (as indicated; n=3 biological replicates). ChIP signals are shown as in Figure 5C. To test for significance, ChIP signals of MMC-treated MEFs are compared against untreated cells; ChIP signals of ATMi and ATRi-treated MEFs are compared against those of MMC-treated MEFs. **: P≤ 0.01, Error bars indicate S.D.; two-sided t-test. Statistical source data are provided in Table S6.



Supplementary Figure 8 Unprocessed scanned Western blots. (A). Unprocessed images of the western blots shown in Figure 2A. (B). Unprocessed images of the western blots shown in Figure 3Ai, ii, iii and iv (as indicated).

Supplementary Table legends

Supplementary Table 1 Overview of the 306 XPF-bound proteins in P15 bXPF (bX) and birA (br) control livers.

Supplementary Table 2 Overview of the 140 shared XPF-bound proteins from the two measurements of P15 bXPF (bX) and birA (br) control livers.

Supplementary Table 3 Overview of the expression profiles of 68 imprinted genes in P15 NER-deficient livers vs. age-matched wt controls. A: allele; FC: Fold Change; P: p-value.

Supplementary Table 4 Overview of imprinted genes with significantly aberrant gene expression profiles in P15 NER-deficient livers as compared to age-matched wt controls. AI: allele; FC: Fold Change; P: p-value.

Supplementary Table 5 Primer sequences. Primer sequences for mRNA and allele specific gene expression, bisulfite mutagenesis, ChIP and allele-specific ChIP qPCR assays and targeting vector.

Supplementary Table 6 Statistics source data.