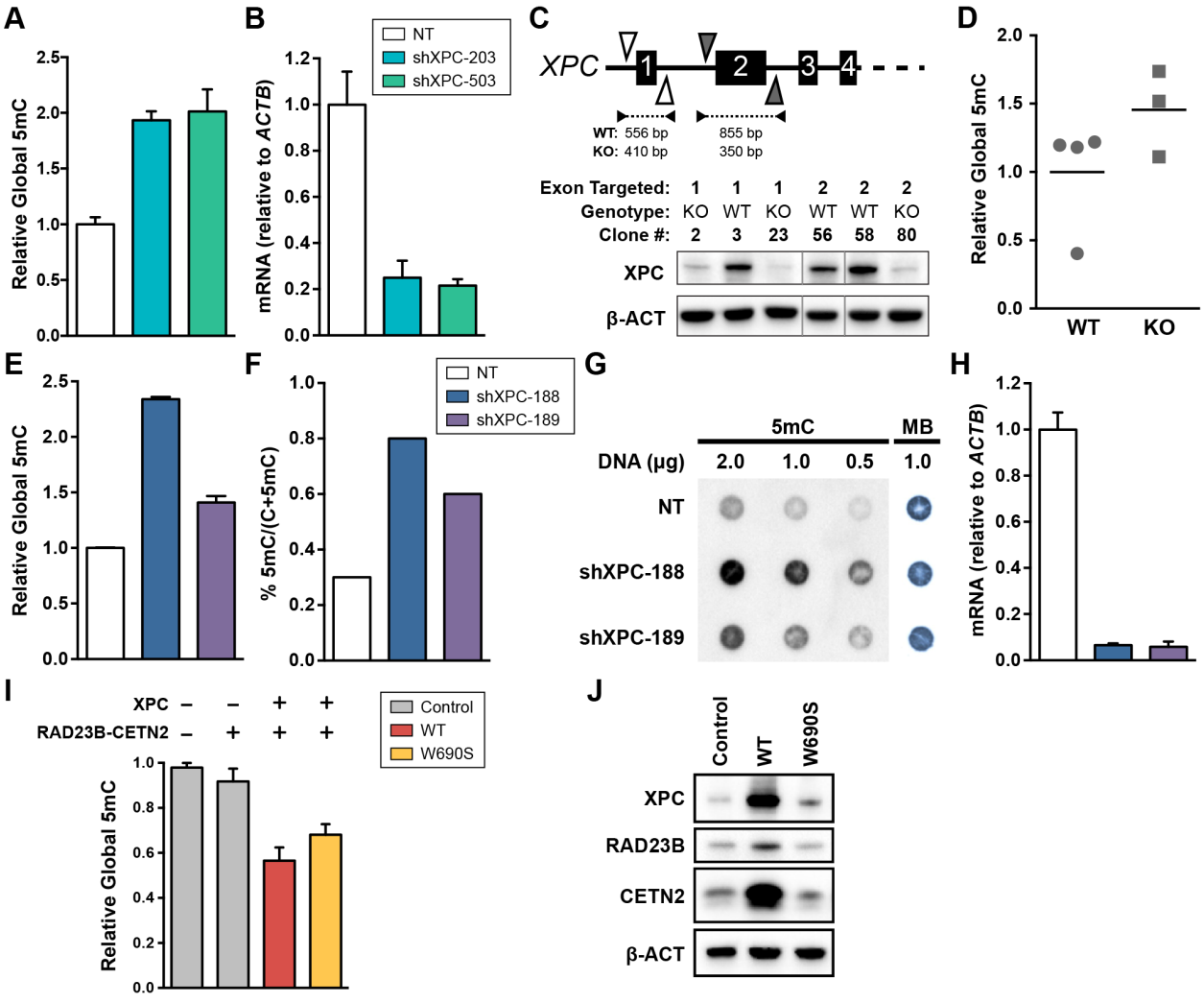
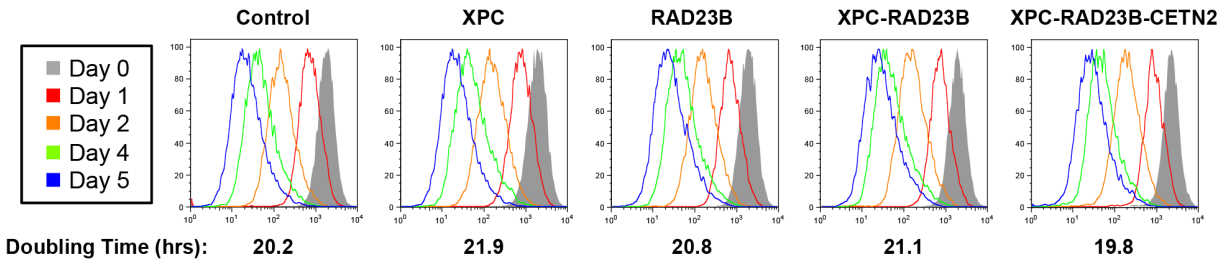
Supplemental Figure S1, related to Figure 1. Global DNA methylation is inversely correlated with XPC expression, independent of DNA repair activity. (**A**) Relative global 5mC levels in human dermal fibroblasts (HDFs) transduced with lentiviruses expressing control, non-targeting (NT) shRNA or two independent shRNAs targeting XPC (shXPC-203 and -503), assayed by 5mC-specific ELISA. (**B**) Quantification of *XPC* mRNA levels in XPC-depleted HDFs as evaluated by quantitative RT-PCR, normalized to *ACTB*. (**C**) Top, schematic representation of CRISPR/Cas9-mediated deletion of *XPC* exon 1 and 2 in H9 ESCs; bottom, summary of clones used, including genotyping validated by Sanger sequencing and immunoblot analysis for XPC. Gray lines in immunoblot indicate where the image was cropped to remove irrelevant lanes. Residual signal detected in the XPC-knockout cell extracts is due to contamination from feeder layers. (**D**) Global 5mC in XPC-knockout human ESCs, assayed by 5mC-specific ELISA. (**E-G**) Global 5mC of XPC-depleted mouse ESCs using two independent shRNAs against mouse XPC (shXPC-188 and -189), assayed by (**D**) 5mC-specific ELISA, (**F**) thin layer chromatography, and (**G**) dot blot. (**H**) Quantification of *Xpc* mRNA levels in XPC-depleted mouse ESCs as evaluated by quantitative RT-PCR, normalized to *Actb*. (**I**) Global 5mC of HDFs overexpressing the RAD23B and CETN2 subunits, with or without ectopic expression of the XPC subunit, and assayed by 5mC-specific ELISA. The first bar is duplicated from Fig. 1B for comparison purposes. (**J**) Immunoblot analysis for HDF samples depicted in Fig. 1D. Ectopic expression of XPC subunit alone increases endogenous RAD23B and CETN2 protein levels. Error bars represent the standard deviation, n = 3. MB, methylene blue.


A
Relative Global 5mC
B
mRNA (relative to ACTB)
NT
shXPC-203
shXPC-503
C
XPC
WT: 556 bp
KO: 410 bp
855 bp
350 bp
Exon Targeted: 1 1 1 2 2 2
Genotype: KO WT KO WT WT KO
Clone #: 2 3 23 56 58 80
XPC
β-ACT
D
Relative Global 5mC
WT
KO
E
Relative Global 5mC
F
% 5mC/(C+5mC)
NT
shXPC-188
shXPC-189
G
5mC
MB
DNA (μg) 2.0 1.0 0.5 1.0
NT
shXPC-188
shXPC-189
H
mRNA (relative to ACTB)
I
XPC
RAD23B-CETN2
Control
WT
W690S
Relative Global 5mC
J
Control
WT
W690S
XPC
RAD23B
CETN2
β-ACT

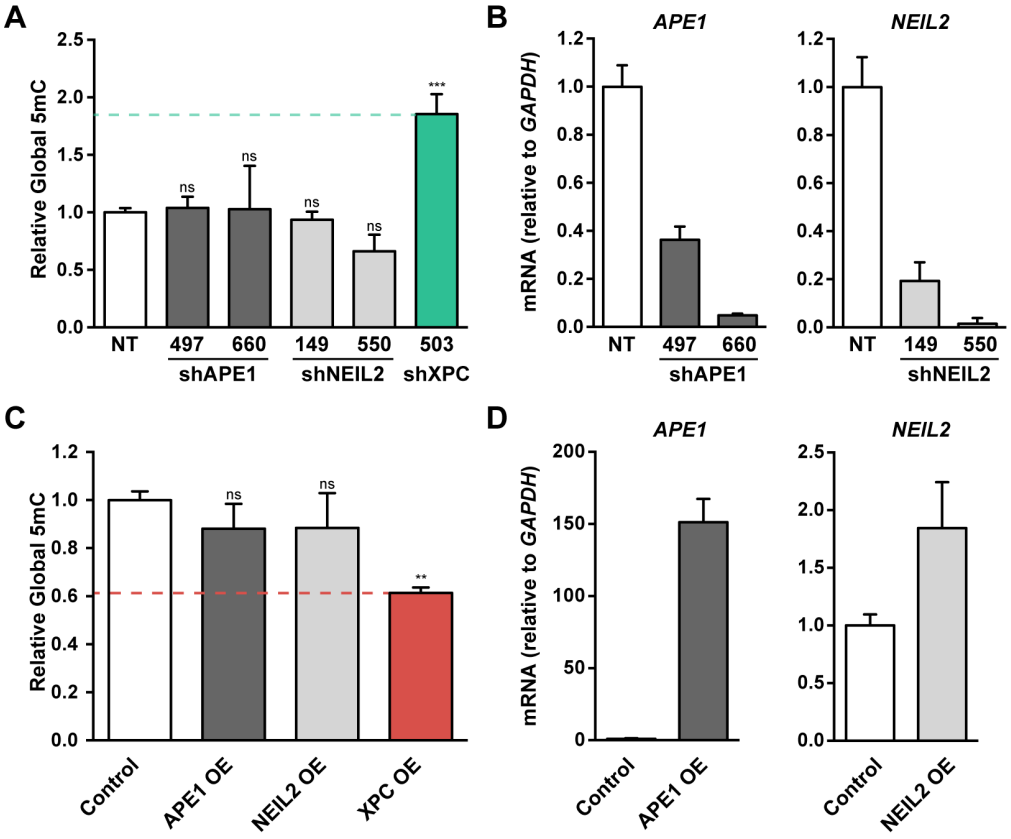
Supplemental Figure S2, related to Figure 1. Overexpression of the XPC complex does not significantly change the doubling time of HDFs. HDFs overexpressing XPC, RAD23B, or various combinations of the XPC complex subunits have similar doublings times, compared to the control. The doubling time for each population was calculated using the CellTrace CFSE Proliferation Assay (Life Technologies), where the mean fluorescence intensity was calculated over a 5-day period.

Ho_FigS2



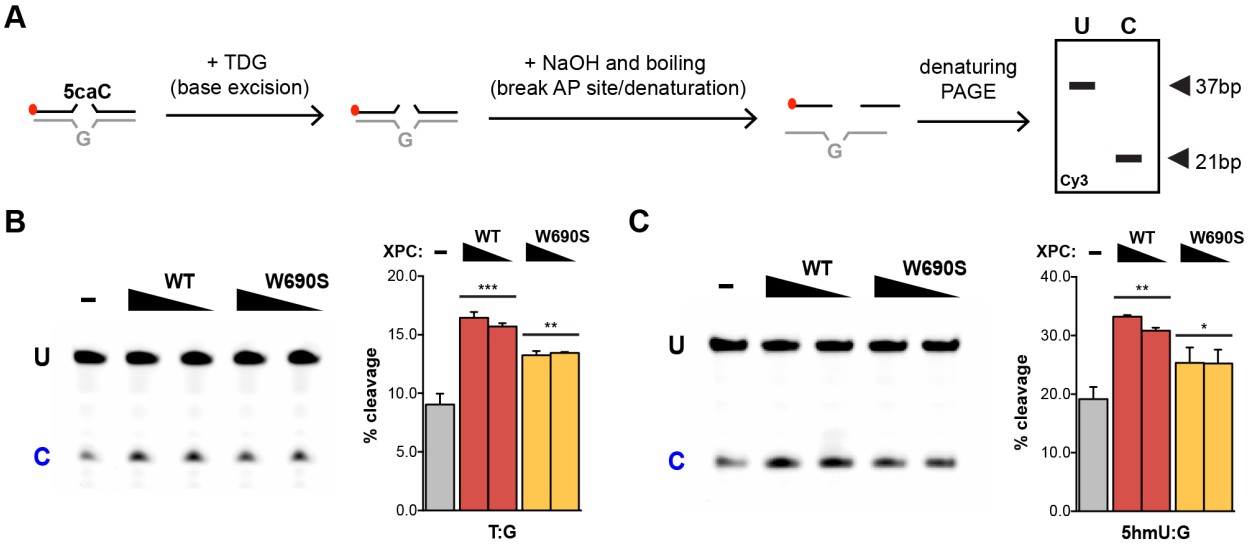
Doubling Time (hrs):	20.2	21.9	20.8	21.1	19.8

Supplemental Figure S3, related to Figure 1. APE1 and NEIL2 gain- and loss-of-function in HDFs do not affect global methylation. (**A**) Relative global 5mC levels in HDFs transduced with lentiviruses expressing control, non-targeting (NT) shRNA or independent shRNAs targeting APE1 (shAPE1-497 and -660), NEIL2 (shNEIL2-149 and -550), and XPC (shXPC-503), assayed by 5mC-specific ELISA. Dotted line represents the average 5mC level in shXPC-503 cells. (**B**) Quantification of *APE1* and *NEIL2* mRNA levels in APE1-depleted (shAPE1-497 and -660) and NEIL2-depleted (shNEIL2-149 and -550) HDFs as evaluated by RT-qPCR, normalized to *GAPDH***.** (**C**) Relative global 5mC in HDFs overexpressing APE1 (APE1 OE), NEIL2 (NEIL2 OE), and XPC (XPC OE). Dotted line represents the average 5mC level in XPC OE cells. (**D**) Quantification of *APE1* and *NEIL2* mRNA levels in APE1 OE and NEIL2 OE cells as evaluated by RT-qPCR, normalized to *GAPDH*. Error bars represent the standard deviation, n = 3. *** $p < 0.001$, ** $p < 0.01$, ns = non-significant, calculated by 1-way ANOVA.


A
Relative Global 5mC
2.5
2.0
1.5
1.0
0.5
0.0
ns
ns
ns
ns
***
NT
497
660
149
550
503
shAPE1
shNEIL2
shXPC
B
APE1
NEIL2
mRNA (relative to GAPDH)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NT
497
660
shAPE1
NT
149
550
shNEIL2
C
Relative Global 5mC
1.2
1.0
0.8
0.6
0.4
0.2
0.0
ns
ns
**
Control
APE1 OE
NEIL2 OE
XPC OE
D
APE1
NEIL2
mRNA (relative to GAPDH)
200
150
100
50
0
2.5
2.0
1.5
1.0
0.5
0.0
Control
APE1 OE
Control
NEIL2 OE

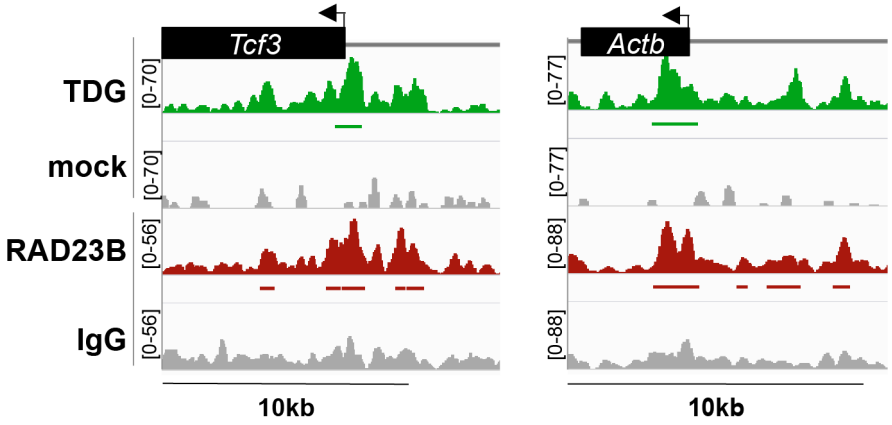
Supplemental Figure S4, related to Figure 1. The XPC complex stimulates the repair activity of TDG. (**A**) Schematic diagram for in vitro TDG glycosylase assay. Red circle, Cy3 labeling; U, uncleaved; C, cleaved. (**B** and **C**) TDG cleavage activity of 5'-labeled oligonucleotides in the presence of absence of decreasing amounts of wildtype (WT) or DNA repair-deficient (W690S) XPC complex. The 37mer dsDNA contained either a (**B**) T:G or (**C**) 5hmU:G mismatch-like base pair as TDG substrate. Representative gels are shown. The intensity of the cleaved product indicates the efficiency of base excision by TDG and is calculated in the graphs as the percentage of total labeled substrate (=C/(U+C). Error bars represent the standard deviation, n = 3. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, calculated by 2-way ANOVA.

Ho_FigS4

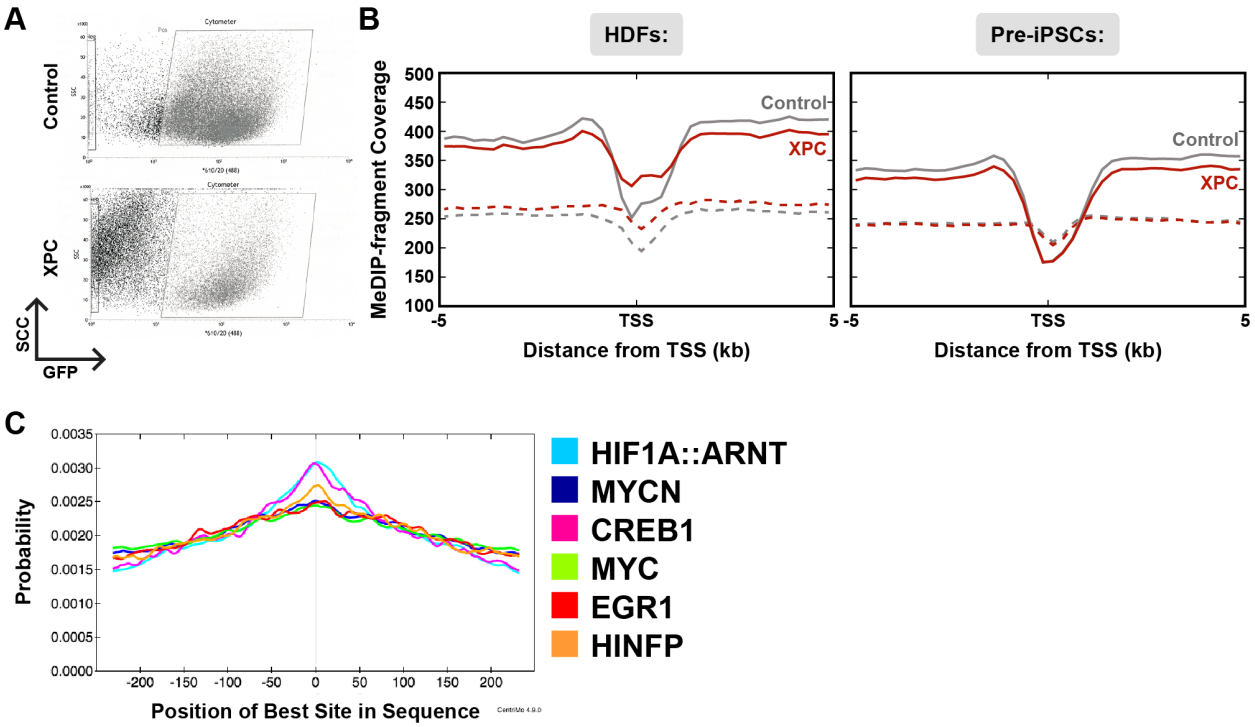

A
5caC
G
+ TDG
(base excision)
G
+ NaOH and boiling
(break AP site/denaturation)
G
denaturing
PAGE
U C
37bp
21bp
Cy3
B
WT
W690S
U
C
XPC: – WT W690S
% cleavage
T:G
C
WT
W690S
U
C
XPC: – WT W690S
% cleavage
5hmU:G

Supplemental Figure S5, related to Figure 2. TDG is enriched at enhancers and promoters genome-wide and colocalizes with the XPC complex subunit, RAD23B. IGV-computed ChIP-seq tracks are plotted as (number of reads) × [1,000,000/(total read count)] for pluripotency-related gene *Tcf3* and housekeeping gene *Actb*. Mock and normal IgG were used as specificity controls for the TDG and RAD23B ChIP, respectively.

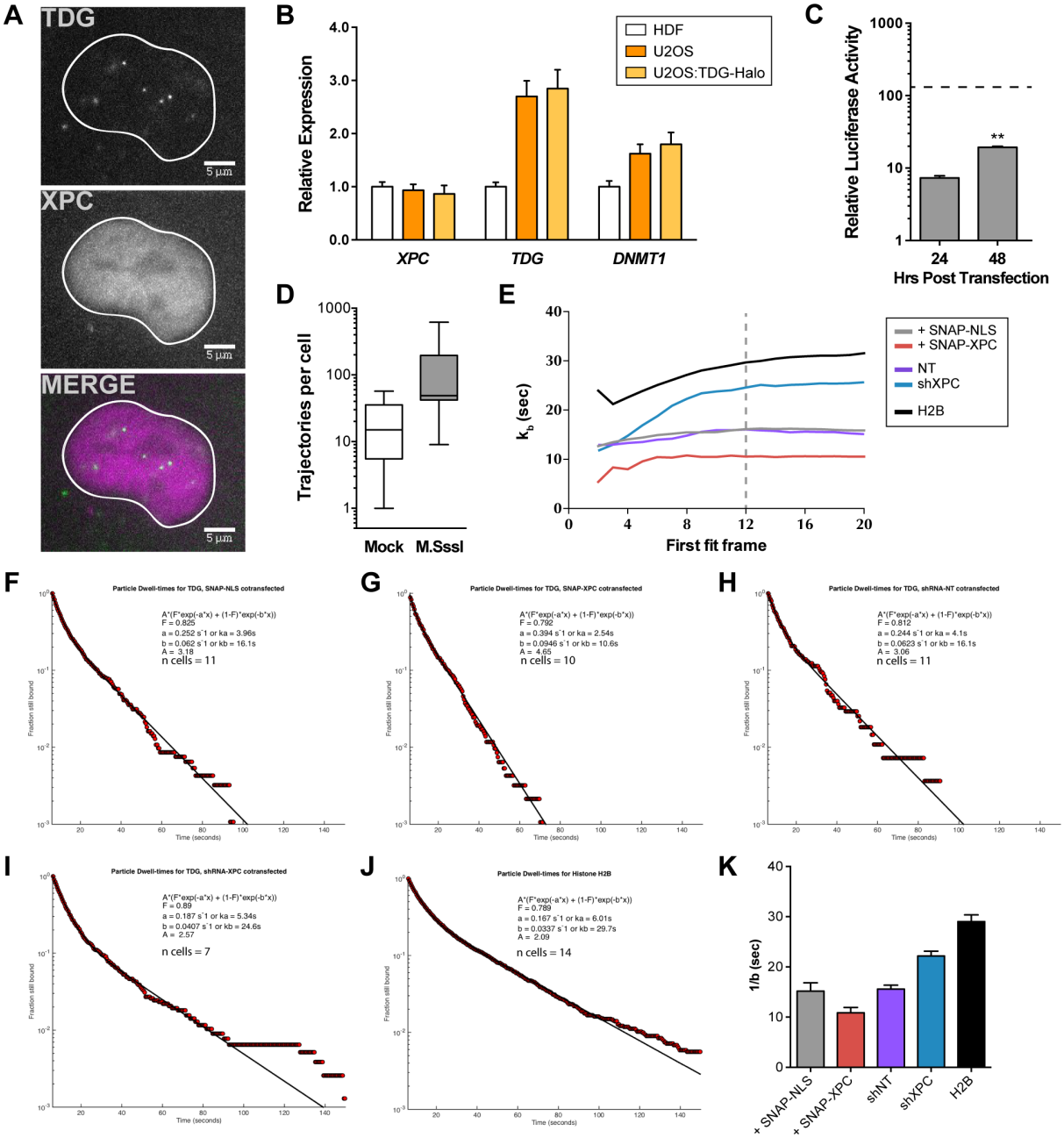
Ho_FigS5


Tcf3
Actb
TDG
mock
RAD23B
IgG
[0-70]
[0-70]
[0-56]
[0-56]
[0-77]
[0-77]
[0-88]
[0-88]
10kb
10kb

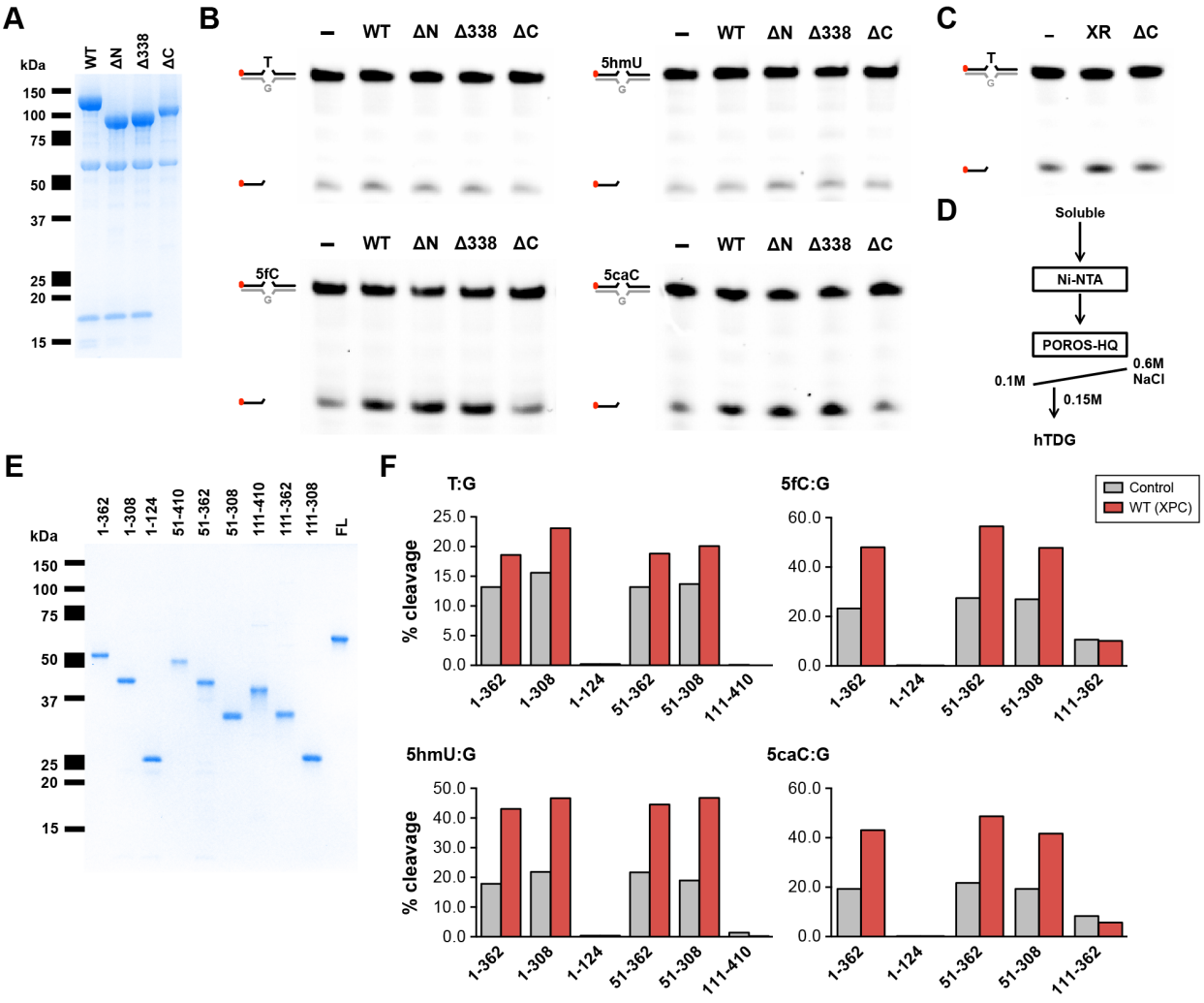
Supplemental Figure S6, related to Figure 3. Ectopic expression of XPC results in loss of DNA methylation genome-wide in both HDFs and pre-iPSCs. (**A**) FACS-sorted, GFP-positive pre-iPSC populations for MeDIP-seq experiment depicted in Fig. 3. (**B**) Distribution of MeDIP-seq reads by their distance ± 5kb from the TSS of RefSeq genes, as in Fig. 3C, prior to input (dashed lines) subtraction. (**C**) CentriMo analysis of transcription factor binding sites enriched at regions which preferentially lose 5mC upon XPC overexpression, as in Fig. 3D. Motif probability graph shows the probability of the best match to a given motif and its positional distribution from the input MeDIP-seq sequences.


A
Control
XPC
SCC
GFP
B
HDFs:
Pre-iPSCs:
MeDIP-fragment Coverage
Control
XPC
TSS
Distance from TSS (kb)
C
Probability
Position of Best Site in Sequence
HIF1A::ARNT
MYCN
CREB1
MYC
EGR1
HINFP

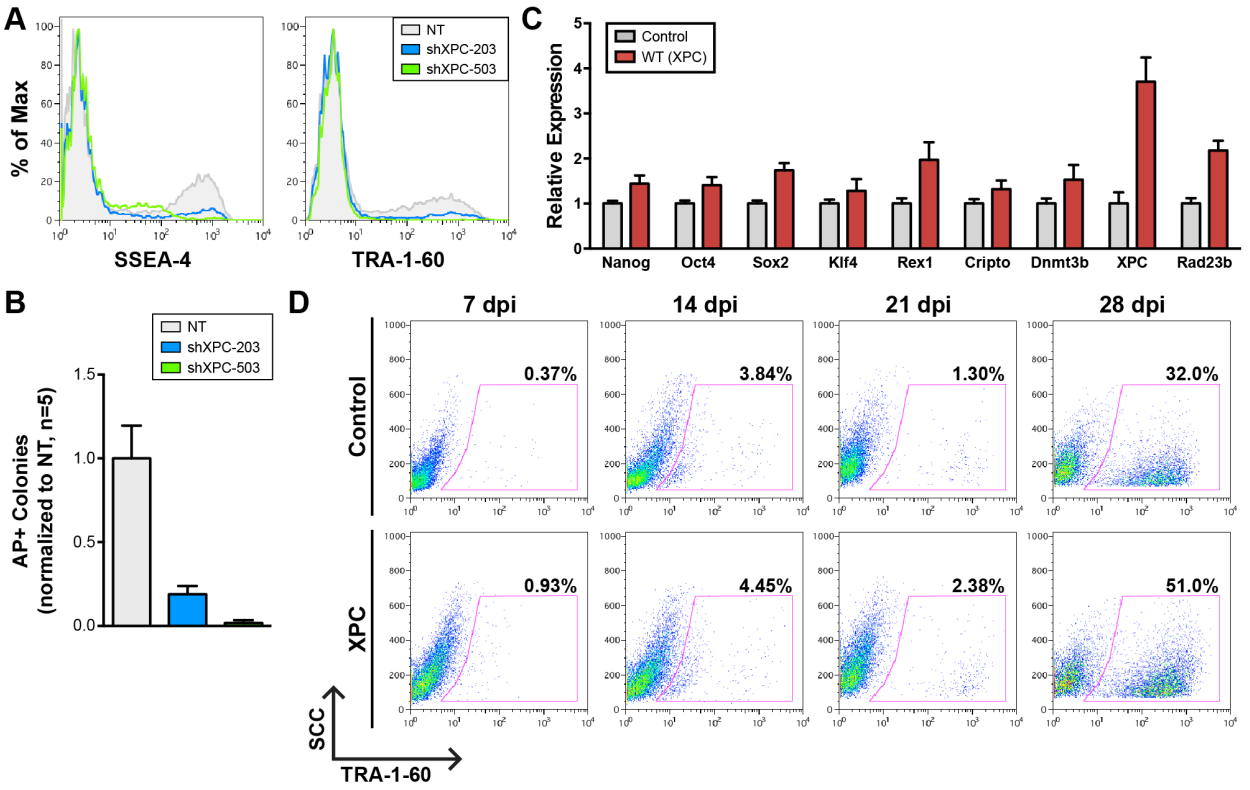
Supplemental Figure S7, related to Figure 4. XPC modulates TDG binding dynamics to DNA in vivo. (**A**) Example frame from tracking movie of TDG-Halo co-transfected with SNAP-XPC and M.SssI-treated plasmid DNA. (**B**) RT-qPCR analyses of *XPC*, *TDG*, and *DNMT1* expression in HDFs, U2OS cells, and the TDG-Halo inducible U2OS stable cell line (U2OS:TDG-Halo). (**C**) Relative luciferase activity (firefly/renilla) of TDG-Halo cells transfected with M.SssI-treated methylated pGL4.13[Luc2/SV40], 24 and 48 hours post transfection. Dotted line denotes the luciferase activity in TDG-Halo cells transfected with mock-treated, unmethylated pGL4.13 at 24 hours. (**D**) A comparison of the number of SPT trajectories for TDG-Halo after co-transfection of SNAP-NLS with mock-treated (mock) or M.SssI-treated methylated (M.SssI) plasmid DNA. The increased number of trajectories is likely due to an increase in available TDG substrates provided by the methylated plasmid DNA. (**E**) Because two-exponential fitting is dependent on the frame at which fitting begins (Mazza et al., 2012), we measured the second-exponential rate k_b as a function of first fit frame to determine where to begin the fit. At 12 frames, k_b has plateaued for all of the TDG-Halo conditions and thus was chosen for future analyses. (**F-J**) Two-exponential fitting for each experimental condition, before bootstrap resampling. Survival probability and fit for individual TDG-Halo molecules after co-transfection with (**F**) SNAP-NLS, (**G**) SNAP-XPC or after transduction with (**H**) control non-targeting shRNAs or (**I**) XPC-targeting shRNAs. (**J**) Survival probability and fit for histone H2B-Halo. All fits are constrained to molecules that were observed for at least 12 frames. (**K**) Second order rate constant for each experimental condition, prior to bootstrapping or photobleaching correction. Error bars represent the standard deviation from 1000 bootstrap iterations.


A
TDG
XPC
MERGE
5 μm
B
Relative Expression
HDF
U2OS
U2OS:TDG-Halo
XPC
TDG
DNMT1
C
Relative Luciferase Activity
**
24
48
Hrs Post Transfection
D
Trajectories per cell
Mock
M.SssI
E
k_b (sec)
First fit frame
+ SNAP-NLS
+ SNAP-XPC
NT
shXPC
H2B
F
Particle Dwell-times for TDG, SNAP-NLS cotransfected
n cells = 11
G
Particle Dwell-times for TDG, SNAP-XPC cotransfected
n cells = 10
H
Particle Dwell-times for TDG, shRNA-NT cotransfected
n cells = 11
I
Particle Dwell-times for TDG, shRNA-XPC cotransfected
n cells = 7
J
Particle Dwell-times for Histone H2B
n cells = 14
K
1/b (sec)
+ SNAP-NLS
+ SNAP-XPC
shNT
shXPC
H2B

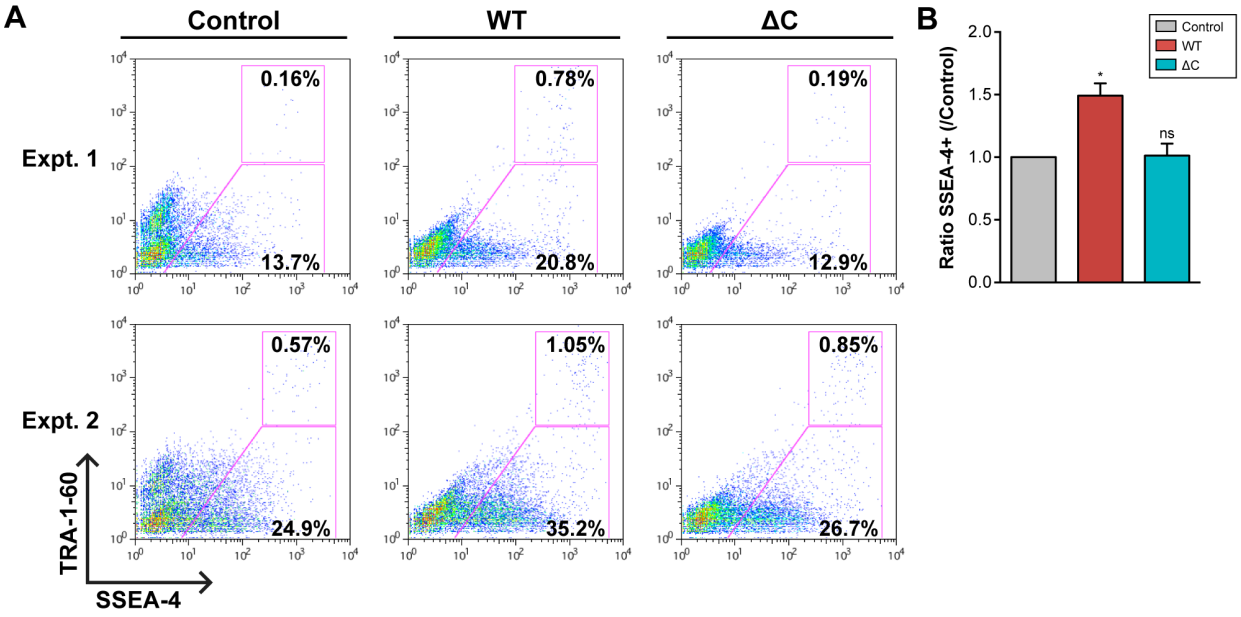
Supplemental Figure S8, related to Figure 5. XPC-mediated stimulation of TDG activity requires the C-terminus of XPC and N-terminus of TDG. (**A**) Recombinant human XPC complexes were expressed in Sf9 cells, purified to near homogeneity, and visualized by Coomassie staining. The full-length (WT) XPC subunit runs at ~120 kDa, RAD23B at ~60kDa, and CETN2 at ~17 kDa. (**B**) Representative gel for the TDG glycosylase assay shown in Fig. 5B and for similar assays using additional TDG substrates (5hmU, 5fC, and 5caC). (**C**) Representative gel corresponding to the TDG glycosylase assay shown in Fig. 5C. (**D**) Schematic representation of the purification method for recombinant TDG. (**E**) SDS-PAGE and Coomassie staining of all purified TDG truncations. Full-length (FL) TDG runs at ~65 kDa. (**F**) Glycosylase activity of additional TDG truncations to those shown in Fig. 5E in the presence or absence of WT XPC complexes, assayed on all relevant DNA demethylation intermediates.


A
kDa
WT
ΔN
Δ338
ΔC
150
100
75
50
37
25
20
15
B
–
WT
ΔN
Δ338
ΔC
T
5hmU
5fC
5caC
C
–
XR
ΔC
D
Soluble
Ni-NTA
POROS-HQ
0.1M
0.6M NaCl
0.15M
hTDG
E
kDa
1-362
1-308
1-124
51-410
51-362
51-308
111-410
111-362
111-308
FL
F
T:G
5fC:G
5hmU:G
5caC:G
% cleavage
Control
WT (XPC)
1-362
1-308
1-124
51-362
51-308
111-410
111-362

Supplemental Figure S9, related to Figure 6. XPC increases somatic reprogramming fidelity and iPSC self-renewal capacity. (**A**) Flow cytometry analysis of human iPSCs expressing a control, non-targeting shRNA or one of two specific shRNAs against XPC (shXPC-203 and -503), 28 days post induction (dpi). Cells are labeled with antibodies against human iPSC markers, SSEA-4 and TRA-1-60. (**B**) Average number of colonies obtained 28 dpi, scored by alkaline phosphastase (AP) staining, n = 5. (**C**) RT-qPCR analyses of pluripotency gene expression in bulk 30 dpi iPSCs from the experiment shown in Fig. 6B-C. (**D**) Flow cytometry analysis of the late stage human iPSC marker, TRA-1-60, on induced HDFs overexpressing the XPC subunit alone at 7, 14, 21, and 28 dpi. Error bars represent the standard deviation.


A
% of Max
SSEA-4
TRA-1-60
NT
shXPC-203
shXPC-503
B
AP+ Colonies
(normalized to NT, n=5)
C
Relative Expression
Control
WT (XPC)
Nanog
Oct4
Sox2
Klf4
Rex1
Cripto
Dnmt3b
XPC
Rad23b
D
7 dpi
14 dpi
21 dpi
28 dpi
Control
XPC
0.37%
3.84%
1.30%
32.0%
0.93%
4.45%
2.38%
51.0%
SCC
TRA-1-60

Supplemental Figure S10, related to Figure 6. The ability of XPC to enhance reprogramming depends on the same C-terminal domain required for stimulating TDG-dependent DNA demethylation. (**A**) Flow cytometry analysis of the early and late stage human iPSC markers, SSEA-4 and TRA-1-60, on induced HDFs overexpressing the wild type (WT) or C-terminally truncated (ΔC) XPC subunit at 14 days post induction. (**B**) Relative number of SSEA-4+ cells (sum of both SSEA-4 single and SSEA-4/TRA-1-60 double positive fractions) in cells overexpressing WT and ΔC XPC as shown in A, normalized to control. Error bars depict the standard deviation, n = 2. * $p < 0.05$, ns = non-significant, calculated by two-tailed Student's t-test.


A
Control
WT
ΔC
Expt. 1
Expt. 2
0.16%
13.7%
0.78%
20.8%
0.19%
12.9%
0.57%
24.9%
1.05%
35.2%
0.85%
26.7%
TRA-1-60
SSEA-4
B
Ratio SSEA-4+ (/Control)
Control
WT
ΔC
*
ns
2.0
1.5
1.0
0.5
0.0