Supplementary materials

Lead Exposure Induces Dysregulation of Constitutive Heterochromatin Hallmarks in Live Cells

Oscar F. Sánchez [1], Li F. Lin[1], Junkai Xie[1], Jennifer L. Freeman[2,3], Chongli Yuan[1,3*]

1. Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907
2. School of Health Sciences, Purdue University, West Lafayette, IN, 47907, USA
3. Purdue Center of Cancer Research, Purdue University, West Lafayette, IN, 47907

* To whom correspondence should be addressed. Tel: + 1 765 494 5824; Fax: + 1 765 494 0805; Email: cyuan@purdue.edu

Table of Contents

obtained from the comparison of the median fluorescence intensity of the total cell population analyzed between Pb-treated groups and the untreated control group (**B and D**). *, $p < 0.05$, $n \geq 3$, one-way ANOVA followed by Tukey's HSD post-hoc test. S-13

Figure S5. Quantitative determination of ^{me}CpG levels in HEK293T cells exposed to Pb via immunostaining. A. Representative 2D projected images of HEK293T cells after 24 h exposure for 24 h to Pb, 100 or 500 ppb and immunostained with anti-^{me}CpG antibody. Scale bar = 2 µm. **B.** Relative ^{me}CpG levels to the control, or 0 ppb of Pb, using the integrated intensity per nuclei (IIN) obtained from image analysis. Data is presented as the mean ± S.D., N=3. **C.** Probability density distribution of the IIN obtained from single cell image analysis, n≥150. Presented data corresponds to at least 50 cells per biological repeat, N = 3. ** represents $p < 0.01$ and *** represents $p < 0.001$ from a One-way ANOVA followed by Tukey's post hoc test. S-14

Figure S6. Relative H3K9me3 (A) and meCpG (B) levels after 24 h of Pb cessation determined by immunostaining. HEK293T cells were immunostained at the end of a Pb exposure of 24 h followed by a Pb cessation of 24 h. Relative H3K9me3 and meCpG levels were calculated against a control group, or 0 ppb of Pb, after 24 h of culture. Presented data corresponds to at least 1500 cells per biological repeat, N = 3. NS: stands for non-significant difference, *** represents $p < 0.001$ from a One-way ANOVA followed by Tukey's post hoc test. S-15

Table S1. Amino acid sequence of ^{me}CG, H3K9me3, H3K27me3 and H3K14ac protein sensors conjugated to either N- or C-mEGFP terminus, and negative control N – and C–mEGFP terminus lacking "reading" domain. SV40 NLS (italic), flexible linker (underscored), protein sensor (bold), and mEGFP N – or C–terminus (double underscored).

Name	Amino acid sequence
^{me}CpG-N-mEGFP	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGS**E FAEDWLDSPA LGPGWKRREV FRKSGATAGR** 60 70 80 90 100 **SDTYYQSPTG DRIRSKVELT RYLGPAGDLT LFDFKQGIL**G GG**QFAEDWLD** 110 120 130 140 150 **SPALGPGWKR REVFRKSGAT AGRSDTYYQS PTGDRIRSKV ELTRYLGPAG** 160 170 180 190 200 **DLTLFDFKQG IL**GGGQFVPV ATMVSKGEEL FTGVVPILVE LDGDVNGHKF 210 220 230 240 250 SVSGEGEGDA TYGKLTLKFI CTTGKLPVPW PTLVTTLTYG VQCFSRYPDH 260 270 280 290 300 MKQHDFFKSA MPEGYVQERT IFFKDDGNYK TRAEVKFEGD TLVNRIELKG 310 320 330 340 IDFKEDGNIL GHKLEYNYNS HNVYIMADKQ KNGIKVNFKI RHNIE
^{me}CG-C-mEGFP	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGS**E FAEDWLDSPA LGPGWKRREV FRKSGATAGR** 60 70 80 90 100 **SDTYYQSPTG DRIRSKVELT RYLGPAGDLT LFDFKQGIL**G GG**QFAEDWLD** 110 120 130 140 150 **SPALGPGWKR REVFRKSGAT AGRSDTYYQS PTGDRIRSKV ELTRYLGPAG** 160 170 180 190 200 **DLTLFDFKQG IL**GGGQFVPV ATDKQKNGIK VNFKIRHNIE DGSVQLADHY 210 220 230 240 250 QQNTPIGDGP VLLPDNHYLS TQSKLSKDPN EKRDHMVLLE FVTAAGITLG MDELYK
H3K9me3-N-mEGFP	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGS**E FASQEFEVEA IVDKRQDKNG NTQYLVRWKG** 60 70 80 90 100 **YDKQDDTWEP EQHLMNCEKC VHDFNRRQTE KQKKLTWTTT SRIFSNNA**GS 110 120 130 140 150 GGGGGSQLGS GGVDSCLPVA TMVSKGEELF TGVVPILVEL DGDVNGHKFS 160 170 180 190 200 VSGEGEGDAT YGKLTLKFIC TTGKLPVPWP TLVTTLTYGV QCFSRYPDHM 210 220 230 240 250 KQHDFFKSAM PEGYVQERTI FFKDDGNYKT RAEVKFEGDT LVNRIELKGI 260 270 280 290 DFKEDGNILG HKLEYNYNSH NVYIMADKQK NGIKVNFKIR HNIE
H3K9me3-C-mEGFP	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGS**E FASQEFEVEA IVDKRQDKNG NTQYLVRWKG** 60 70 80 90 100 **YDKQDDTWEP EQHLMNCEKC VHDFNRRQTE KQKKLTWTTT SRIFSNNA**GS 110 120 130 140 150

	GGGGGSQLGS GGVDSCLPVA TDKQKNGIKV NFKIRHNIED GSVQLADHYQ 160 170 180 190 200 QNTPIGDGPV LLPDNHYLST QSKLSKDPNE KRDHMVLLEF VTAAGITLGM DELYK
mEGFP N-terminus (Control)	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGSE LGSGGVDSCL PVATMVSKGE ELFTGVVPIL 60 70 80 90 100 VELDGDVNGH KFSVSGEGEG DATYGKLTLK FICTTGKLPV PWPTLVTTLT 110 120 130 140 150 YGVQCFSRYP DHMKQHDFFK SAMPEGYVQE RTIFFKDDGN YKTRAEVKFE 160 170 180 190 200 GDTLVNRIEL KGIDFKEDGN ILGHKLEYNY NSHNVYIMAD KQKNGIKVNF KIRHNIE
mEGFP C-terminus (Control)	10 20 30 40 50 M*PKKKRKV*GS GSGSGSGGSE LSVGVDSCLP VATDKQKNGI KVNFKIRHNI 60 70 80 90 100 EDGSVQLADH YQQNTPIGDG PVLLPDNHYL STQSKLSKDP NEKRDHMVLL 110 EFVTAAGITL GMDELYK

Table S2. A list of qPCR primers used to quantify mRNA level of "writer" and "eraser" proteins for DNA methylation and H3K9me3.

Gene Name	Primer Sequence 5' – 3'	Reference
DNMT1	Forward: GGTTTTCCTTCCTCAGCTACTGCGA Reverse: CACTGATAGCCCATGCGGACCA	(Liao et al., 2015)
TET1	Forward: CAGAACCTAAACCACCCGTG Reverse: TGCTTCGTAGCGCCATTGTAA	(Yang et al., 2015)
KMT1A	Forward: GCACAAGTTTGCCTACAA Reverse: CCAGGTCAAAGAGGTAGGTG	(Lee et al., 2011)
KDM4A	Forward: GAAGCCA CGAGCATCCTATGA Reverse: GCGGAACTCTCGAACAGTCA	(Wang et al., 2017)
β - ACTIN	Forward: GGAGTCCTGTGGCATCCACG Reverse: CTAGAAGCATTTGCGGTGGA	(Missiaglia et al., 2010)

Table S3. ^{me}CpG and H3K9me3 data were adjusted for unimodal and bimodal distributions using Gauss and a mixture of Gauss models, respectively.

Gauss model: $y = y_o + \left(\frac{A}{w\sqrt{\frac{\pi}{2}}}\right) e^{\left(-2\left(\frac{x-x_c}{w}\right)^2\right)}$

A. H3K9me3

[Pb] ppb	Parameters	Peak 1	Peak 2
0	y_o	-0.0161±0.099	-0.0161±0.099
	x_c	21.321±2.511	43.762±9.150
	w	12.243±1.109	46.537±5.767
	A	1.222±0.092	5.018±1.222
	R^2	0.86	
100	y_o	-0.0022±0.0026	-0.0022±0.0026
	x_c	23.198±0.617	45.729±0.360
	w	20.695±1.423	10.663±0.705
	A	2.619±0.183	1.528±0.125
	R^2	0.98	
500	y_o	0.0018±0.0056	0.0018±0.0056
	x_c	20.062±1.740	40.243±1.750
	w	17.052±3.081	13.562±2.820
	A	2.534±0.498	1.677±0.470
	R^2	0.94	

B. ^{me}CpG

[Pb] ppb	Parameters	Peak 1	Peak 2
0	y_o	0.0093±0.0073	0.0093±0.0073
	x_c	60.188±7.008	110.670±8.916
	w	35.443±9.63	43.365±15.678
	A	3.070±1.378	4.340±1.758
	R^2	0.89	
100	y_o	0.0270±0.0042	0.0270±0.0042
	x_c	43.201±1.111	74.558±1.271
	w	13.433±1.962	15.914±2.846
	A	1.996±0.300	2.133±0.346
	R^2	0.90	
500	y_o	-0.0019±1.8E-4	-0.0019±1.8E-4
	x_c	59.941±0.032	113.927±0.084
	w	33.209±0.076	28.556±0.187
	A	8.023±0.018	2.538±0.017
	R^2	0.97	

C. ^{me}CpG-N-H3K9me3-C

[Pb] ppb	Parameters	Peak 1	Peak 2
0	y_o	-0.0103±0.0134	-0.0103±0.0134
	x_c	10.620±0.479	31.113±2.269
	w	9.871±0.952	25.109±7.944
	A	1.579±0.244	1.865±0.906
	R^2	0.94	
100	y_o	-6.03E-4±0.0055	-6.03E-4±0.0055
	x_c	6.438±0.508	17.993±2.005
	w	5.344±1.912	19.874±3.763
	A	0.514±0.304	2.614±0.584
	R^2	0.95	
500	y_o	-0.0032±0.0040	N.A.
	x_c	11.757±0.544	
	w	19.877±1.403	
	A	3.616±0.273	
	R^2	0.96	

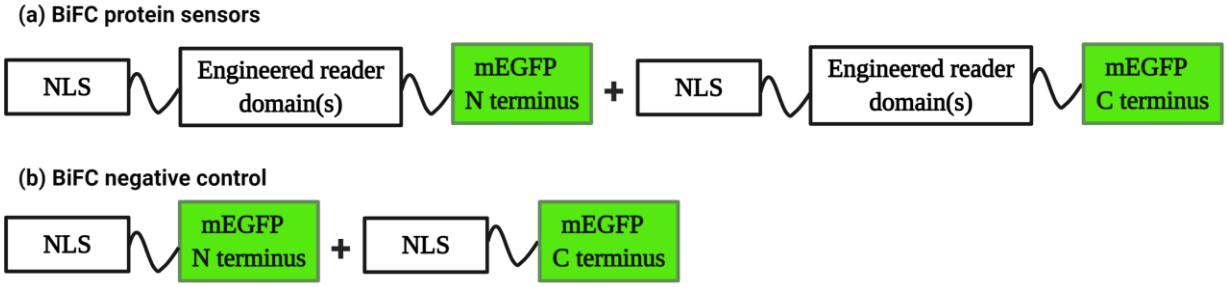
D. H3K9me3-N-^{me}CpG-C

[Pb] ppb	Parameters	Peak 1	Peak 2
0	y_o	-0.0145±0.0213	-0.0145±0.0213
	x_c	12.029±0.750	31.196±3.733
	w	9.533±1.598	24.161±11.454
	A	1.598±0.488	2.189±1.496
	R^2	0.90	
100	y_o	-0.0027±0.0060	-0.0027±0.0060
	x_c	9.414±0.479	22.316±4.931
	w	9.216±2.044	20.884±7.395
	A	1.265±0.704	1.955±0.976
	R^2	0.97	
500	y_o	0.0014±0.0055	N.A.
	x_c	9.869±1.237	
	w	24.100±3.000	
	A	3.667±0.502	
	R^2	0.94	

Table S4. The $\Delta\Delta C_t$, defined as the ΔC_t treated with Pb compared to ΔC_t of untreated samples, with 95% confidence interval (CI) and expression fold change of selected genes in HEK293T cells after Pb exposure, 100 or 500 ppb, for 48 h. Statistical difference, *P* value (*P*), in the expression fold change against the control, HEK293T cells Pb untreated, was calculated using a one-way ANOVA followed by a Tukey's HSD post-hoc test.

Epigenetic Function	Target Gene	[ATZ] ppb	$\Delta\Delta C_t \pm$ SE	95% CI	Fold Change	*P*
5-mC writer	**DNMT1**	100	-0.16 ± 0.09	-0.26 – -0.06	1.12 ± 0.06	0.636
		500	-0.15 ± 0.14	-0.31 – 0.01	1.12 ± 0.11	0.642
5-mC eraser	**TET1**	100	-0.42 ± 0.06	-0.49 – -0.35	1.34 ± 0.05	**0.034**
		500	-0.64 ± 0.11	-0.76 – -0.51	1.57 ± 0.11	**0.003**
H3K9me3 writer	**KMT1A**	100	-1.04 ± 0.09	-1.14 – -0.94	2.07 ± 0.14	**0.005**
		500	-0.90 ± 0.13	-1.05 – -0.75	1.88 ± 0.17	**0.013**
H3K9me3 eraser	**KDM4A**	100	-0.28 ± 0.18	-0.48 – -0.08	1.23 ± 0.14	0.558
		500	-0.25 ± 0.23	-0.51 – 0.01	1.22 ± 0.21	0.584

A.



B.

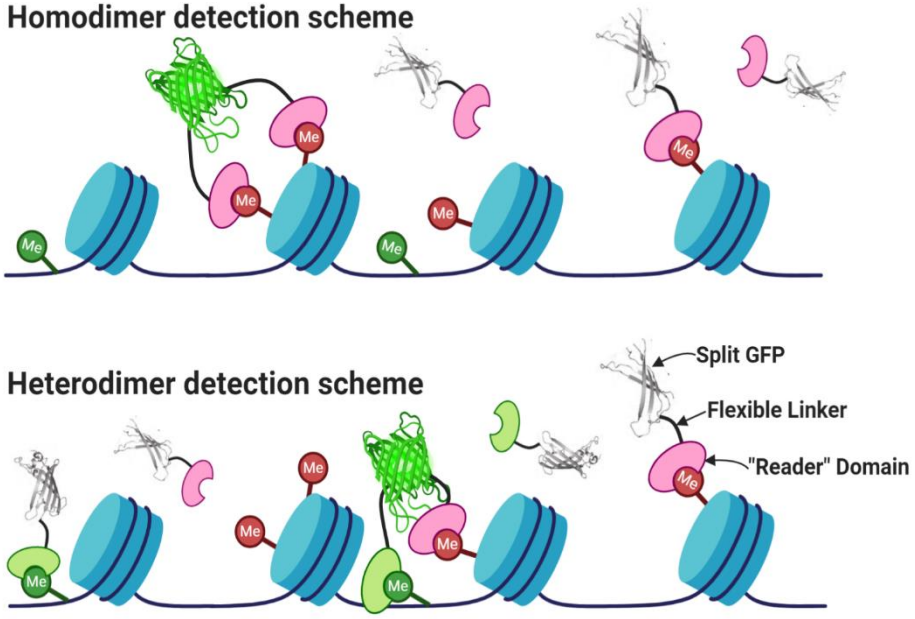


Figure S1. A. Schematic representation of designed protein sensor. B. Schematic representation of BiFC strategy for recognizing one or two different epigenetic marks closely localized. Probes fused to the split fluorescence protein (*i.e.*, N-terminal or C-terminal of mEGFP) that are close to each other will produce a fluorescence signal distinctive from background and non-interacting molecules remain dark. Homodimeric constructs refer to the same epigenetic probe fused to split mEGFP for sensing one specific epigenetic mark. Heterodimeric constructs refer to two different epigenetic probes fused to split mEGFP for sensing two different epigenetic marks.

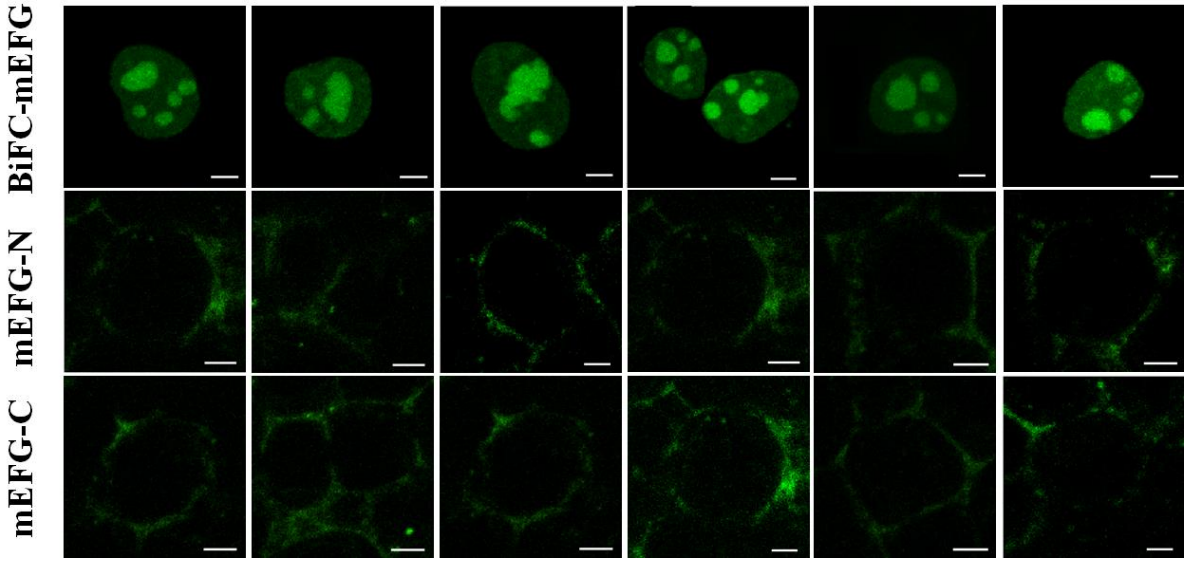


Figure S2. No particular features such as foci or fluorescent islands are observed in transfected cells with the BiFC negative control. Transfected cells with BiFC-mEGFP shows diffusive fluorescence while transfected cells with either mEGFP-N or mEGFP-C exhibited auto-fluorescence at high laser power. Scale bar = 5 μm.

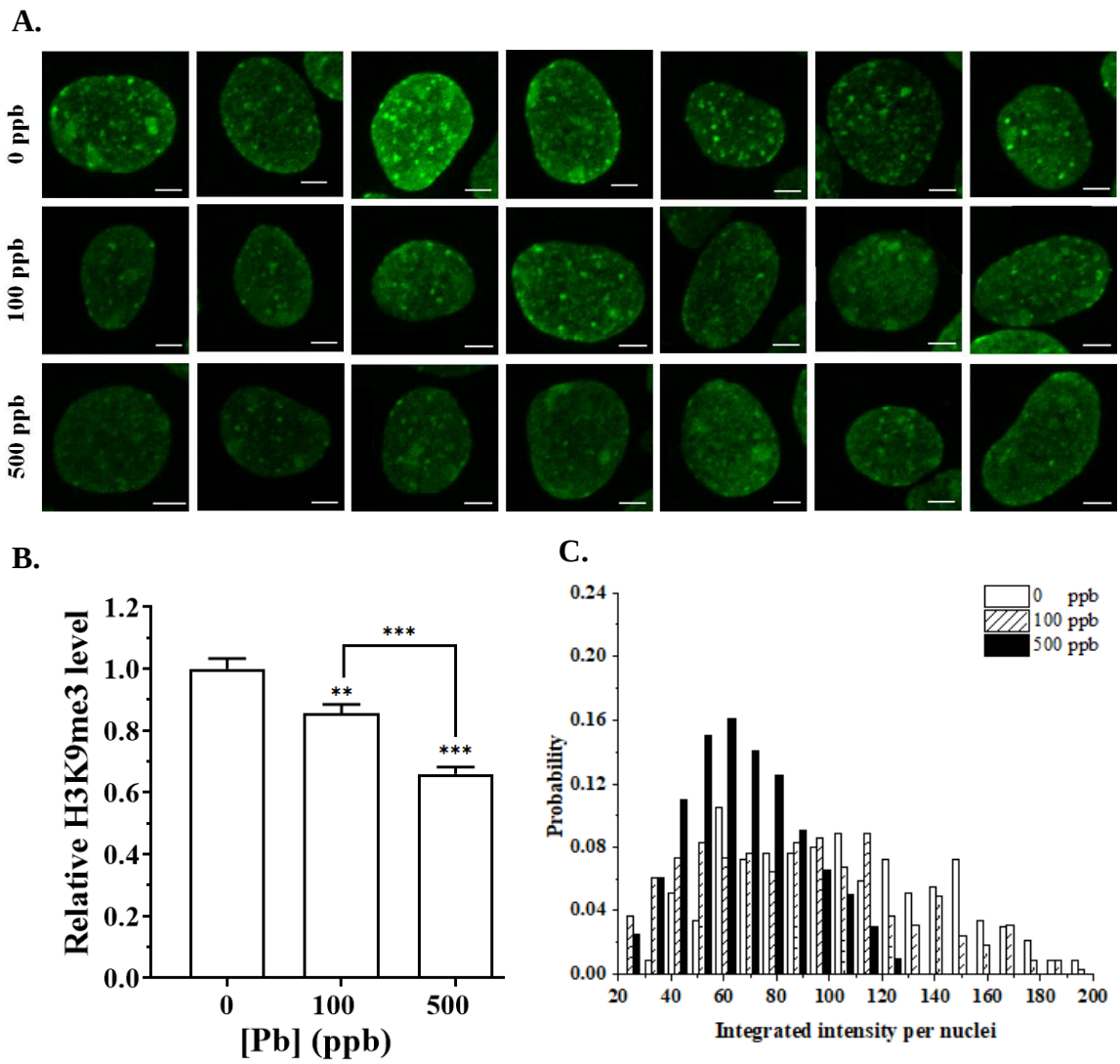


Figure S3. Quantitative determination of H3K9me3 levels in HEK293T cells exposed to Pb via immunostaining. A. Representative 2D projected images of HEK293T cells after 24 h exposure to Pb of 0, 100 or 500 ppb and immunostained with anti-H3K9me3 antibody. Scale bar = 2 µm. **B.** Relative H3K9me3 levels to the control, or 0 ppb of Pb, using the integrated intensity per nuclei (IIN) obtained from image analysis. Data = mean ± S.E., N=3. **C.** Probability density distribution of the IIN obtained from single cell image analysis, n ≥ 150. The data presented corresponds to at least 50 cells per biological repeat, N = 3. **: $p < 0.01$ and ***: $p < 0.001$ from a One-way ANOVA followed by Tukey's post-hoc test.

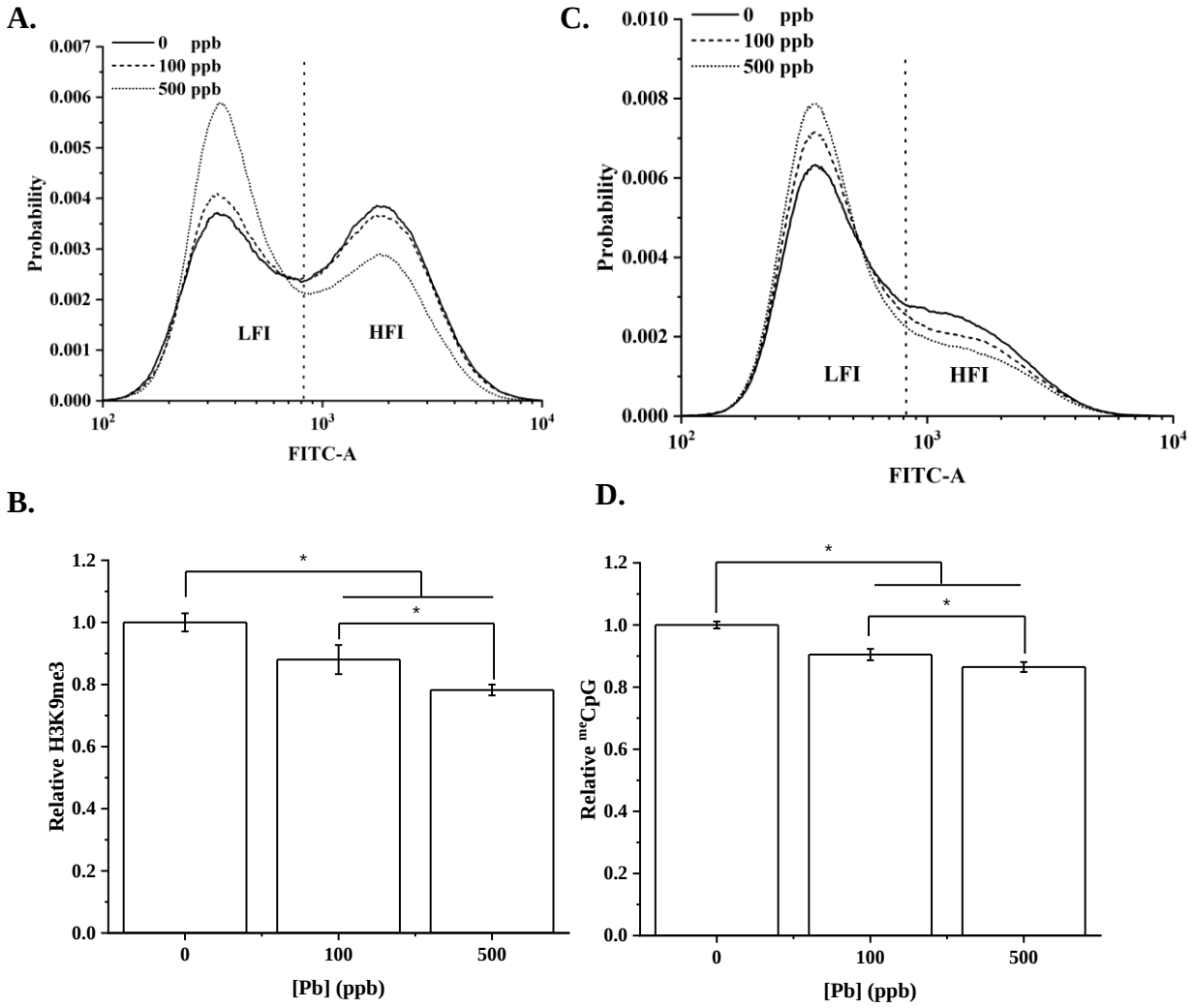


Figure S4. Change of H3K9me3 (A and B) and ^{me}CpG (C and D) levels in HEK293T cells exposed to Pb captured by homodimeric H3K9me3 and ^{me}CpG BiFC probes, respectively. For both homodimeric probes, the probability distribution from FACS histograms (**A and C**) reveals two subpopulations. The cell population consist of a low fluorescence intensity (LFI) and high fluorescence intensity (HFI) sub-population. Like the image analysis a drop in the relative H3K9me and ^{me}CpG level was obtained from the comparison of the median fluorescence intensity of the total cell population analyzed between Pb-treated groups and the untreated control group (**B and D**). *, $p < 0.05$, $n \geq 3$, one-way ANOVA followed by Tukey's HSD post-hoc test.

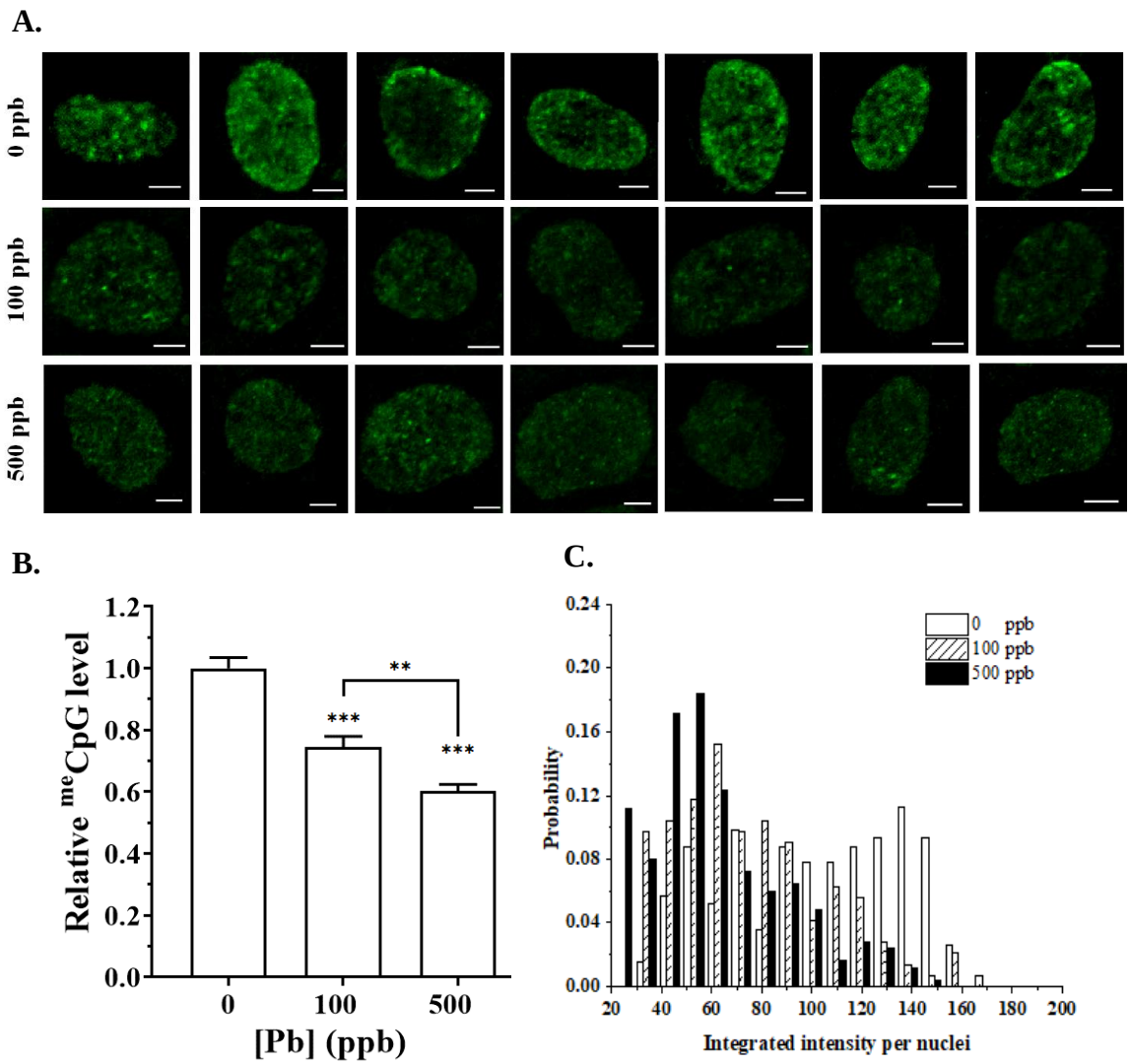


Figure S5. Quantitative determination of ^{me}CpG levels in HEK293T cells exposed to Pb via immunostaining. A. Representative 2D projected images of HEK293T cells after 24 h exposure to Pb, 100 or 500 ppb and immunostained with anti- ^{me}CpG antibody. Scale bar = 2 µm. **B.** Relative ^{me}CpG levels to the control, or 0 ppb of Pb, using the integrated intensity per nuclei (IIN) obtained from image analysis. Data is presented as the mean ± SEM, N=3. **C.** Probability density distribution of the IIN obtained from single cell image analysis, n ≥ 150. The data presented corresponds to at least 50 cells per biological repeat, N = 3. **: $p < 0.01$ and ***: $p < 0.001$ from a one-way ANOVA followed by Tukey's post-hoc test.

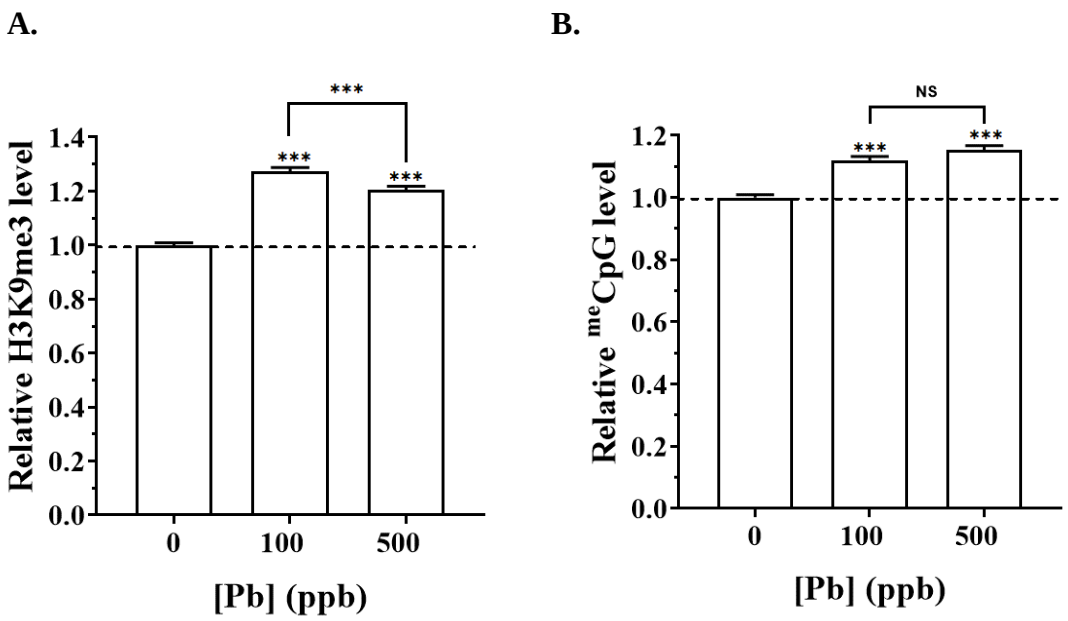


Figure S6. Relative H3K9me3 (A) and ^{me}CpG (B) levels after 24 h of Pb cessation determined by immunostaining. HEK293T cells were immunostained at the end of a Pb exposure of 24 h followed by a Pb cessation of 24 h. Relative H3K9me3 and ^{me}CpG levels were calculated against a control group, or 0 ppb of Pb, after 24 h of culture. Presented data corresponds to at least 1500 cells per biological repeat, N = 3. NS: stands for non-significant difference, *** represents $p < 0.001$ from a one-way ANOVA followed by Tukey's post hoc test.