Figure S1

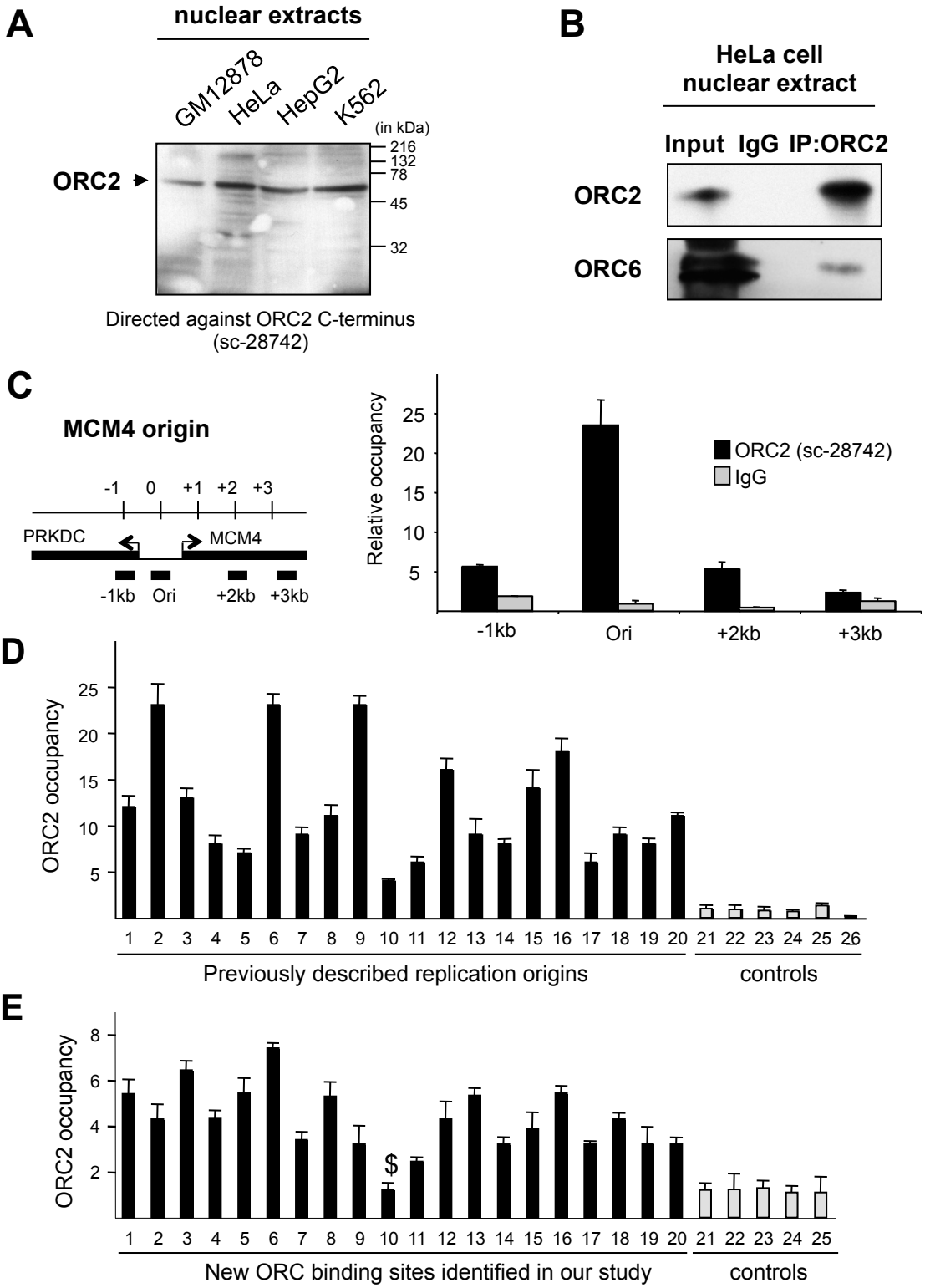

A
nuclear extracts
GM12878
HeLa
HepG2
K562
(in kDa)
216
132
78
45
32
ORC2
Directed against ORC2 C-terminus
(sc-28742)
B
HeLa cell
nuclear extract
Input IgG IP:ORC2
ORC2
ORC6
C
MCM4 origin
-1 0 +1 +2 +3
PRKDC
MCM4
-1kb Ori +2kb +3kb
Relative occupancy
ORC2 (sc-28742)
IgG
-1kb
Ori
+2kb
+3kb
D
ORC2 occupancy
Previously described replication origins
controls
E
ORC2 occupancy
$
New ORC binding sites identified in our study
controls

Figure S2

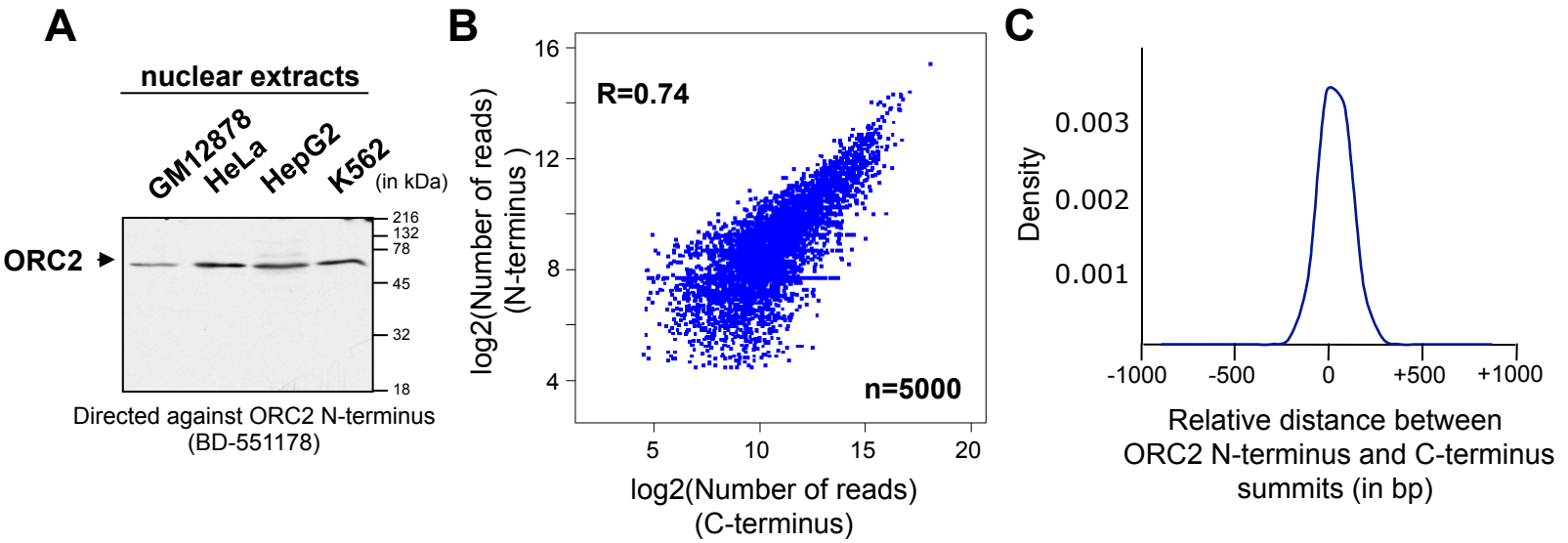

A
nuclear extracts
GM12878
HeLa
HepG2
K562
(in kDa)
ORC2
216
132
78
45
32
18
Directed against ORC2 N-terminus
(BD-551178)
B
R=0.74
n=5000
log2(Number of reads)
(N-terminus)
log2(Number of reads)
(C-terminus)
16
12
8
4
5
10
15
20
C
Density
0.003
0.002
0.001
-1000
-500
0
+500
+1000
Relative distance between
ORC2 N-terminus and C-terminus
summits (in bp)

Figure S3

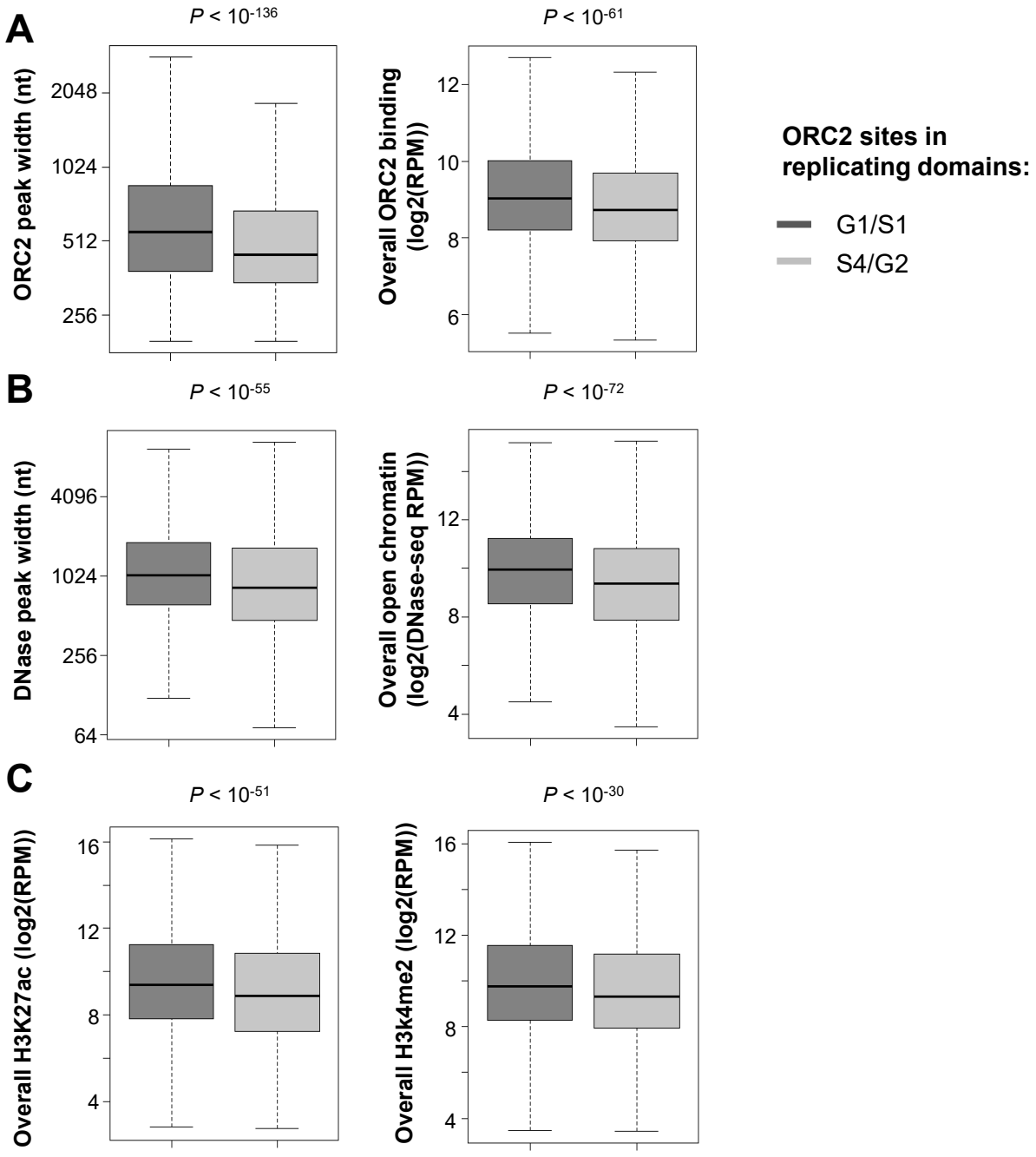
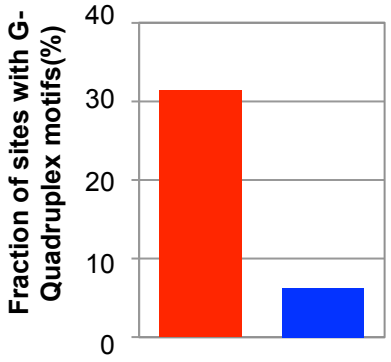
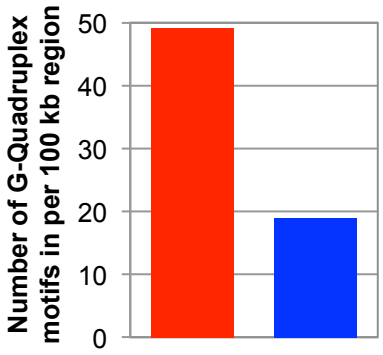

A
P < 10-136
ORC2 peak width (nt)
2048
1024
512
256
P < 10-61
Overall ORC2 binding (log2(RPM))
12
10
8
6
B
P < 10-55
DNase peak width (nt)
4096
1024
256
64
P < 10-72
Overall open chromatin (log2(DNase-seq RPM))
12
8
4
C
P < 10-51
Overall H3K27ac (log2(RPM))
16
12
8
4
P < 10-30
Overall H3k4me2 (log2(RPM))
16
12
8
4
ORC2 sites in replicating domains:
G1/S1
S4/G2

Figure S4

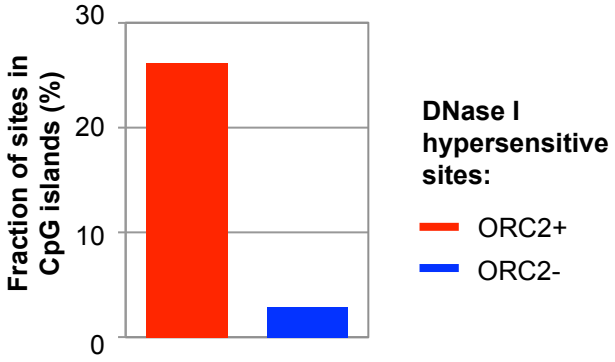
A



B



C



D

Factor	Motif	Fraction of ORC2 binding region with motif	Fold enrichment
E2F	UP00001	17.1%	9.4
HIF1A::ARNT	MA0259.1	12.1%	4.3
IRF	UP00011	8.1%	3.8
MYC	MA0147.1	22.0%	3.8
NF-KB	MA0105.1	37.5%	3.6
CREB	MA0018.2	16.4%	3.2
NFYA	MA0060.1	20.1%	2.8
TP53	MA0106.1	26.2%	2.7
STAT	MA0144.1	41.1%	2.5
TEAD	MA0090.1	26.3%	2.4
GATA	MA0140.1	33.0%	2.2
AP1	MA0099.2	29.0%	2.0

Figure S5

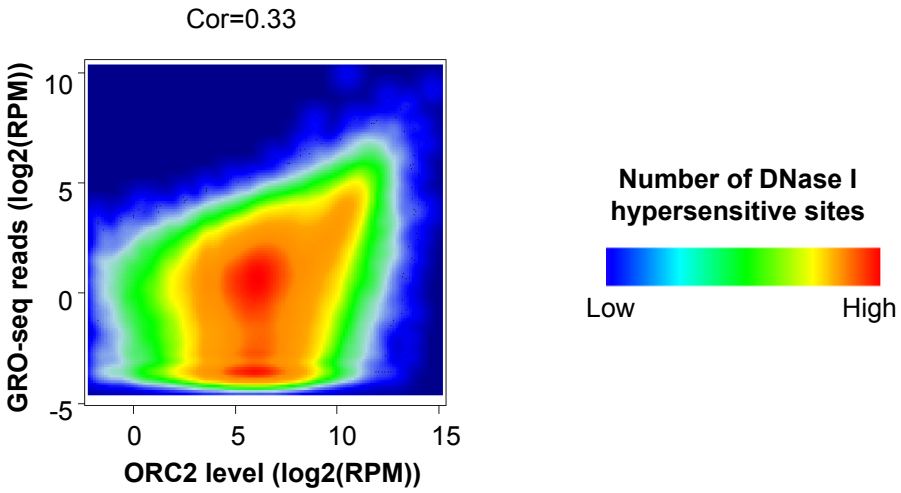

Cor=0.33
GRO-seq reads (log2(RPM))
10
5
0
-5
0
5
10
15
ORC2 level (log2(RPM))
Number of DNase I hypersensitive sites
Low
High

Figure S6

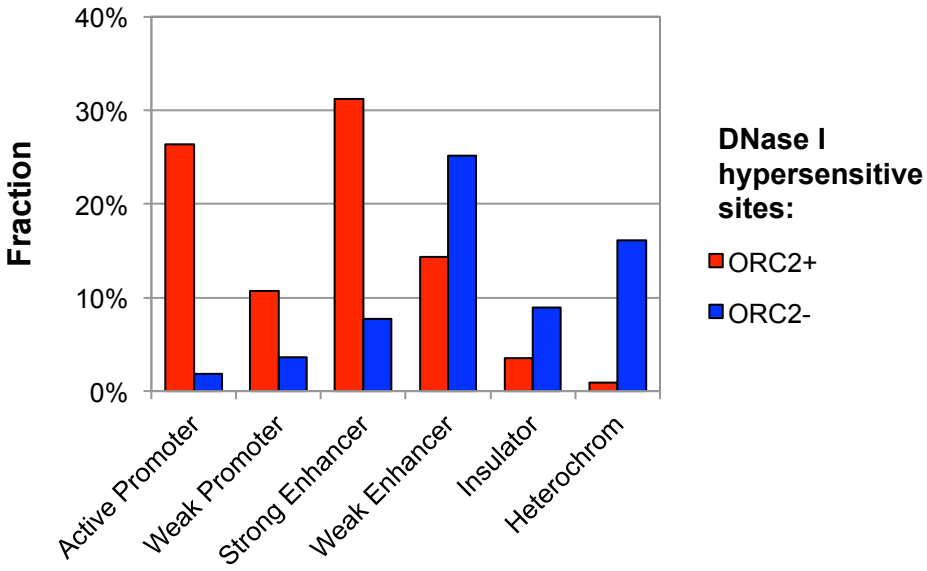
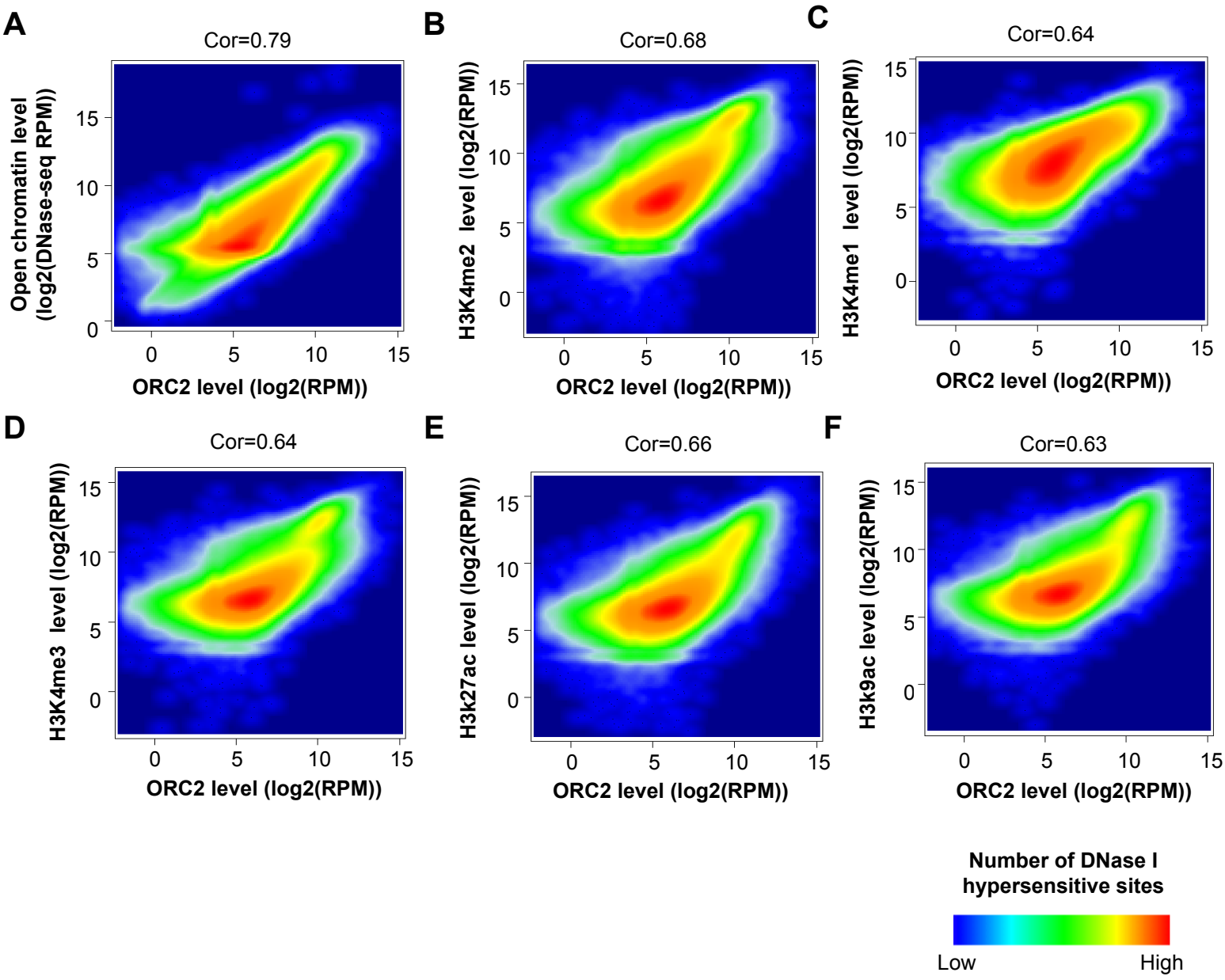

40%
30%
20%
10%
0%
Fraction
Active Promoter
Weak Promoter
Strong Enhancer
Weak Enhancer
Insulator
Heterochrom
DNase I hypersensitive sites:
ORC2+
ORC2-

Figure S7


A
Cor=0.79
Open chromatin level (log2(DNase-seq RPM))
ORC2 level (log2(RPM))
B
Cor=0.68
H3K4me2 level (log2(RPM))
ORC2 level (log2(RPM))
C
Cor=0.64
H3K4me1 level (log2(RPM))
ORC2 level (log2(RPM))
D
Cor=0.64
H3K4me3 level (log2(RPM))
ORC2 level (log2(RPM))
E
Cor=0.66
H3k27ac level (log2(RPM))
ORC2 level (log2(RPM))
F
Cor=0.63
H3k9ac level (log2(RPM))
ORC2 level (log2(RPM))
0
5
10
15

Number of DNase I hypersensitive sites
Low
High

Figure S8

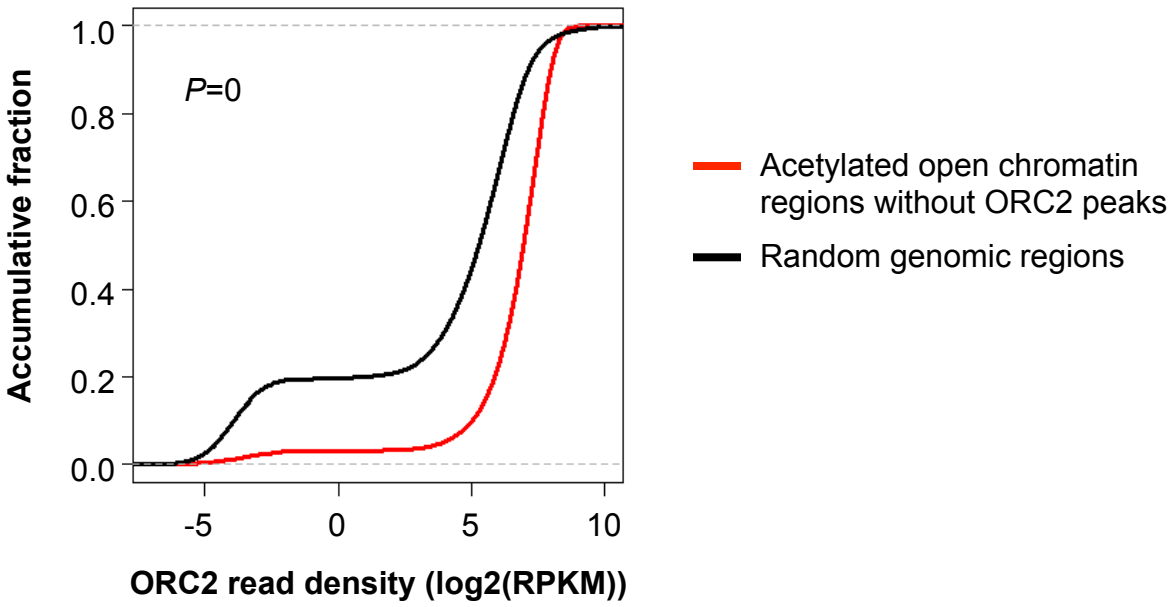

Accumulative fraction
1.0
0.8
0.6
0.4
0.2
0.0
P=0
-5
0
5
10
ORC2 read density (log2(RPKM))
Acetylated open chromatin regions without ORC2 peaks
Random genomic regions

Figure S9

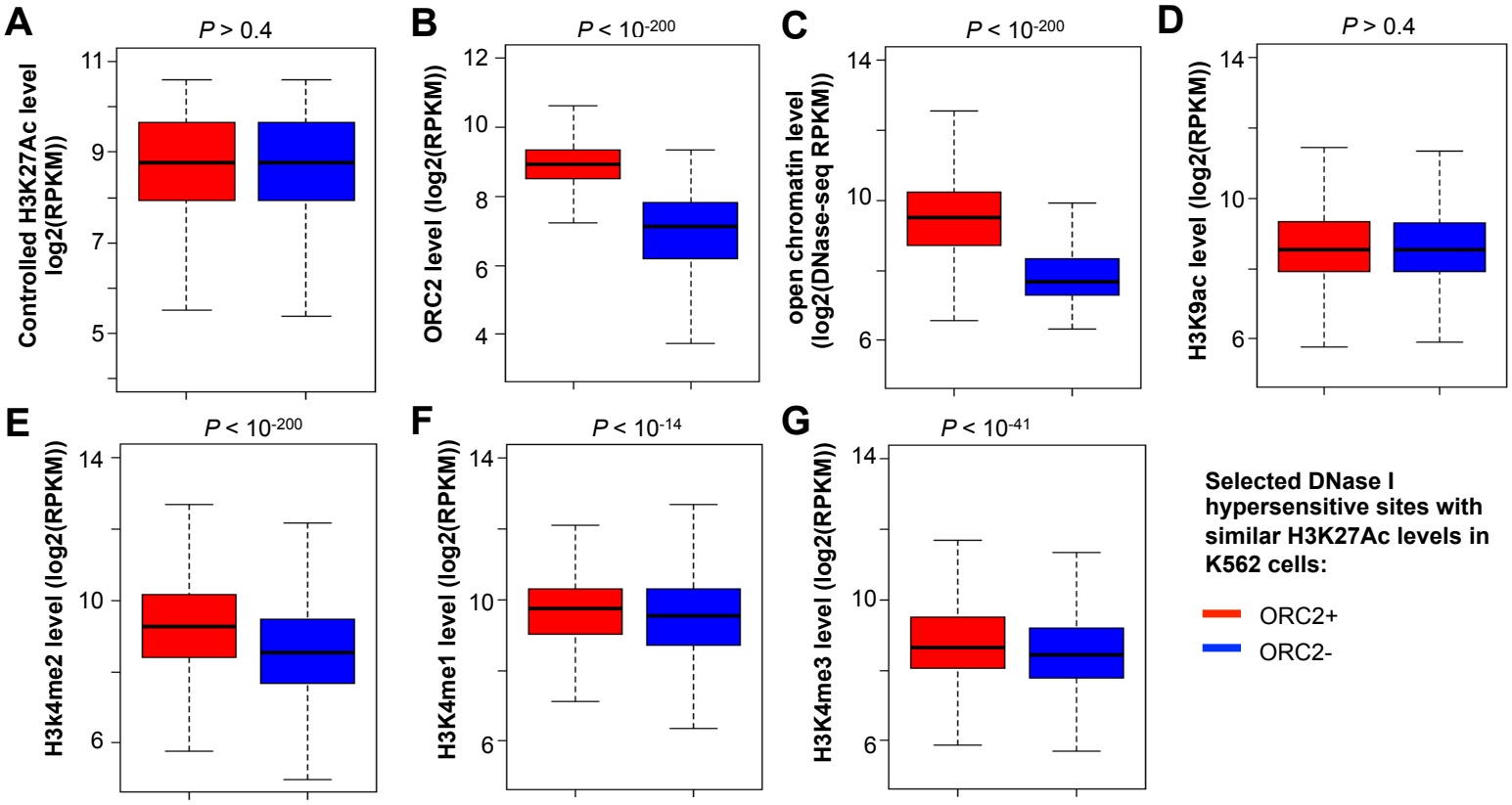

A
P > 0.4
Controlled H3K27Ac level log2(RPKM))
11
9
7
5
B
P < 10-200
ORC2 level (log2(RPKM))
12
10
8
6
4
C
P < 10-200
open chromatin level (log2(DNase-seq RPKM))
14
10
6
D
P > 0.4
H3K9ac level (log2(RPKM))
14
10
6
E
P < 10-200
H3k4me2 level (log2(RPKM))
14
10
6
F
P < 10-14
H3K4me1 level (log2(RPKM))
14
10
6
G
P < 10-41
H3K4me3 level (log2(RPKM))
14
10
6
Selected DNase I hypersensitive sites with similar H3K27Ac levels in K562 cells:
ORC2+
ORC2-

Figure S10

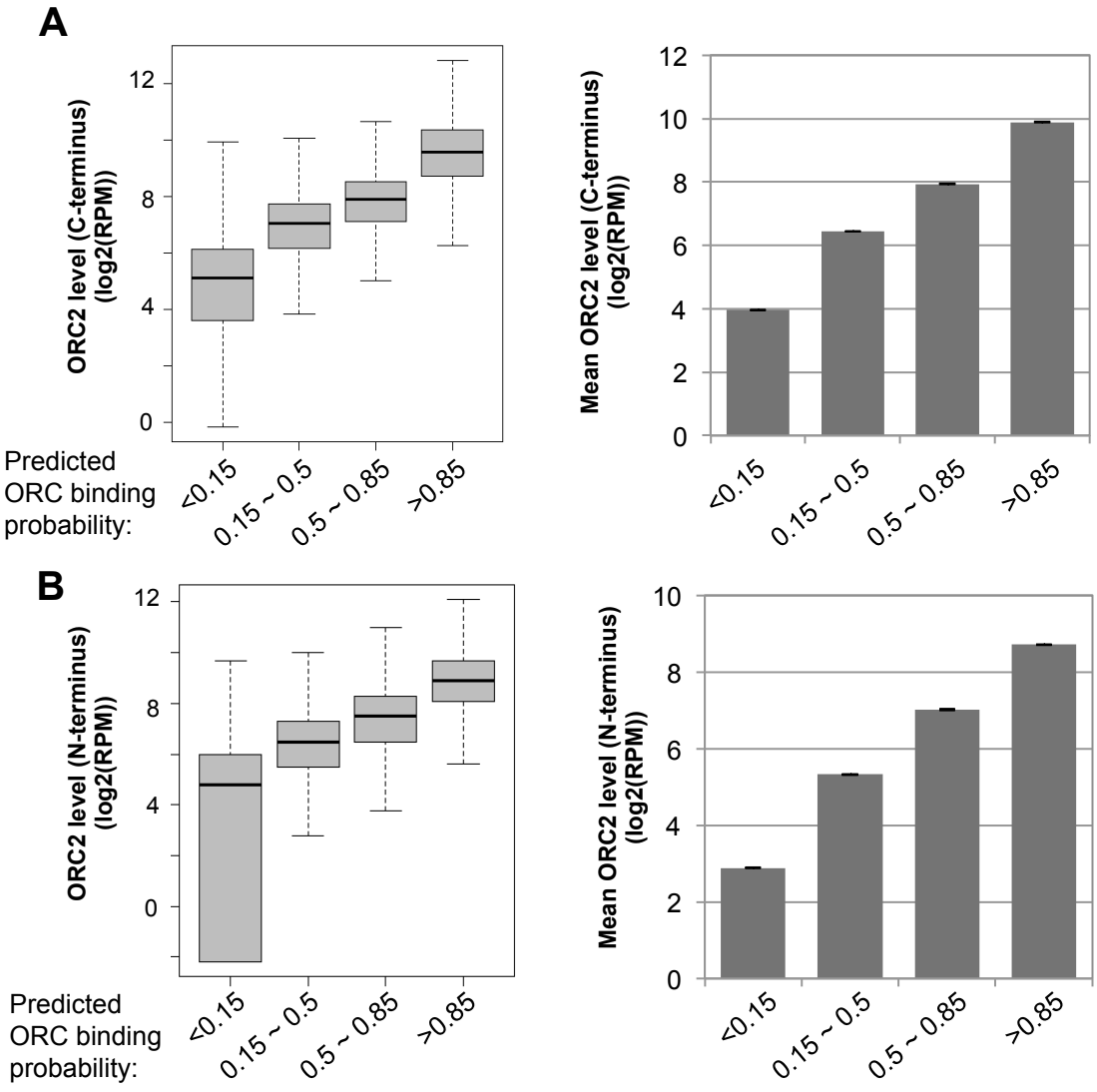

A
ORC2 level (C-terminus) (log2(RPM))
Mean ORC2 level (C-terminus) (log2(RPM))
Predicted ORC binding probability:
<0.15
0.15 ~ 0.5
0.5 ~ 0.85
>0.85
B
ORC2 level (N-terminus) (log2(RPM))
Mean ORC2 level (N-terminus) (log2(RPM))
Predicted ORC binding probability:
C
H3K4me2 only
True positive rate
False positive rate
AUC=0.87
D
H3K27ac only
AUC=0.86

Figure S11

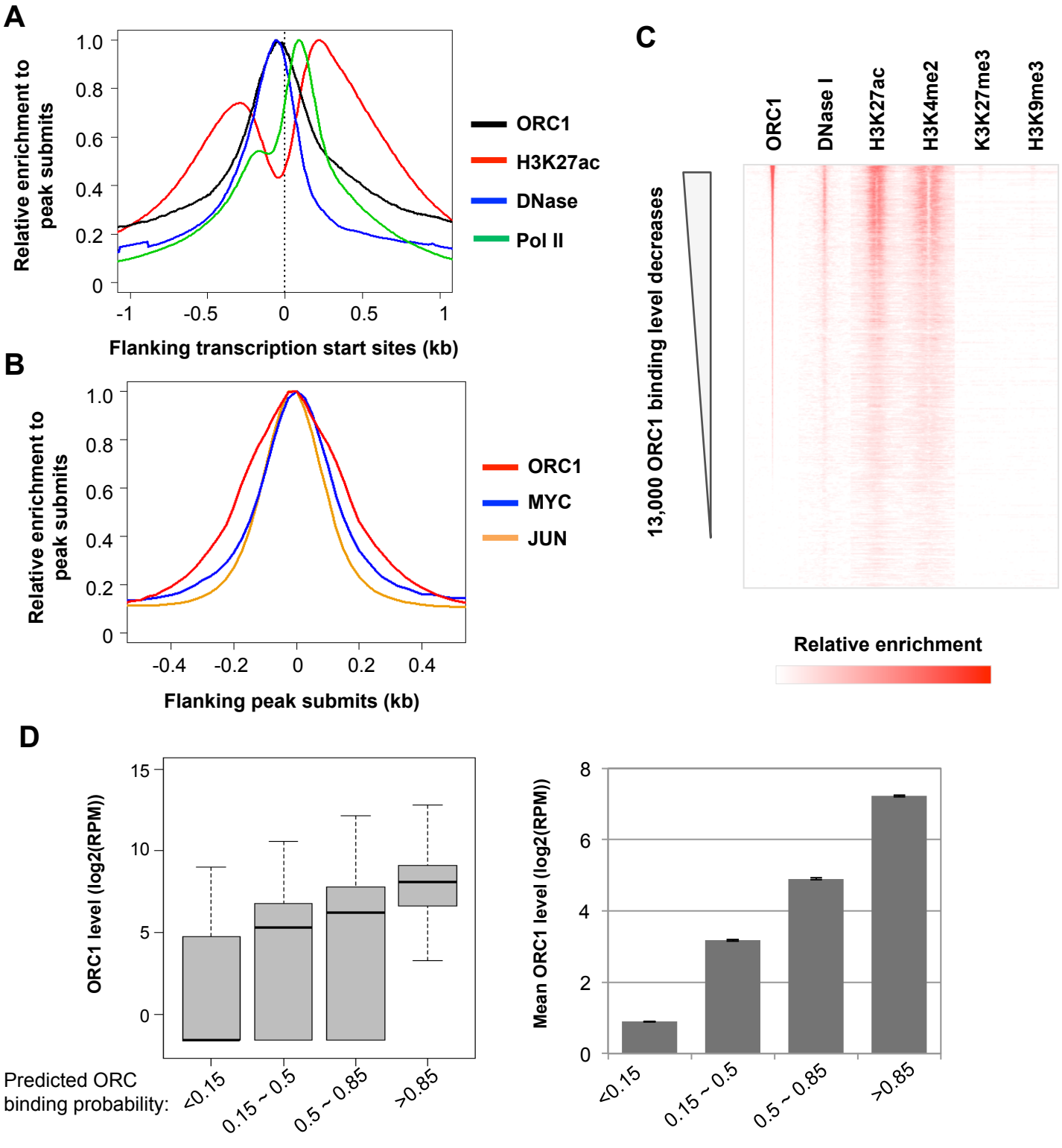
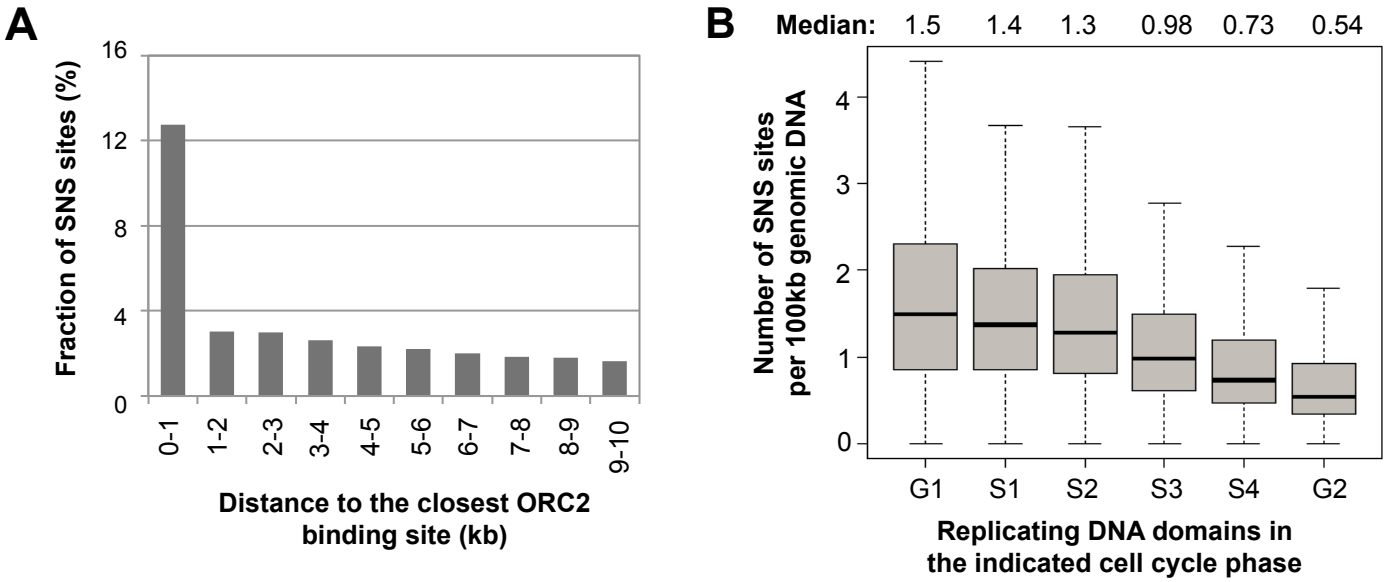

A
Relative enrichment to peak submits
1.0
0.8
0.6
0.4
0.2
0
-1
-0.5
0
0.5
1
Flanking transcription start sites (kb)
ORC1
H3K27ac
DNase
Pol II
B
Relative enrichment to peak submits
1.0
0.8
0.6
0.4
0.2
0
-0.4
-0.2
0
0.2
0.4
Flanking peak submits (kb)
ORC1
MYC
JUN
C
ORC1
DNase I
H3K27ac
H3K4me2
K3K27me3
H3K9me3
13,000 ORC1 binding level decreases
Relative enrichment
D
ORC1 level (log2(RPM))
15
10
5
0
Predicted ORC binding probability:
<0.15
0.15 ~ 0.5
0.5 ~ 0.85
>0.85
Mean ORC1 level (log2(RPM))
8
6
4
2
0
<0.15
0.15 ~ 0.5
0.5 ~ 0.85
>0.85

Figure S12


A
Fraction of SNS sites (%)
16
12
8
4
0
0-1
1-2
2-3
3-4
4-5
5-6
6-7
7-8
8-9
9-10
Distance to the closest ORC2 binding site (kb)
B
Median:
1.5
1.4
1.3
0.98
0.73
0.54
Number of SNS sites per 100kb genomic DNA
4
3
2
1
0
G1
S1
S2
S3
S4
G2
Replicating DNA domains in the indicated cell cycle phase

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Characterization and validation of the antibody directed against the C-terminus of ORC2 used in our study. (*A*) Detection of ORC2 by western blot. ORC2 antibody specificity was tested using equivalent amounts of nuclear extracts from human GM12878, HeLa, HepG2 and K562 cells. (*B*) Detection of ORC2 by western blot following immunoprecipitation and washes. HeLa cell nuclear extract were immunoprecipitated with anti-ORC2 (Santa Cruz antibody) over-night. Immunocomplexes were resolved on a SDS-Page prior to ORC2 (BD Bioscience antibody) and ORC6 (Millipore) detection. (*C*) ChIP analysis of ORC2 binding at the MCM4 replication origin in HeLa cell. Cross-linked sheared chromatin from HeLa cell was immunoprecipitated over-night with anti-ORC2 (Santa Cruz antibody). DNA present in the immunocomplex was then analyzed by real-time quantitative PCR using primers pairs amplifying the MCM4 replication origin and flanking control regions (n=3). IgG was used as a control and data presented as relative occupancy over a control background region. (*D*) ChIP-qPCR validation of ORC2 sites overlapping with HeLa replication origins and ORC ChIP binding sites previously described (n=3). Tested sites include, ORC-ChIP sites 1: TOP1, 2: MCM4, 3: LaminB2, 4: HPRT1, 5: FMR1; HeLa SNS-origins 6: ori4, 7: T1, 8: ori29, 9: T6, 10: ori25, 11: T9, 12: ori22, 13: T18, 14: ori21, 15: T19; lymphocyte SNS origins (Lucas et al., 2007) 16: TUG1, 17: TBC10D1A, 18: OSBP2 (3'), 19: OSBP2 (intron), 20: Sec14L2; "ARS" elements 21: OrBeta, 22: Or9, 23: Ors8, 24: ORS8 and control sites 25: HoxA8, 26: cJun. (*E*) Validation of ORC2 binding sites by ChIP-qPCR. 19 out of 20 were significantly enriched ($, not enriched compare to control sites).

Figure S2. ORC2 ChIP-Seq analysis with a second ORC2 antibody. (*A*) Detection of ORC2 by western blot using the second sera recognizing the N-terminus of ORC2 (BD Biosciences). ORC2 antibody specificity was tested using equivalent amounts of nuclear extracts from human GM12878, HeLa, HepG2 and K562 cells. (*B*) Correlation between the two ORC2 ChIP-Seq experiments was evaluated by plotting the number of reads at segment summits for 5,000 randomly chosen ORC2

sites. The graph represents the plot of common segments and the correlation coefficient (r) is indicated. (*C*) Analysis of the overlap between 10,000 ORC2 binding site summits between the two ChIP-Seq experiments.

Figure S3. Compare ORC2 sites in early (G1/S1) and late (S4/S2) replicating DNA domains. (*A*) ORC peak with and overall ORC2 binding level. (*B*) DNase peak width and overall open chromatin level. (*C*) Overall H3K27ac level and H3K4me2 level. The Wilcoxon Rank Sum Test *P*-values are shown.

Figure S4. Analyses of ORC2 binding sites. (*A*) Fraction of DNase I hypersensitive sites with G-Quadruplex motifs. (*B*) Number of G-Quadruplex motifs in per 100 kb DNase I hypersensitive regions. (*C*) Fraction of DNase I hypersensitive sites located in CpG islands. (*D*) Transcription factor binding sites enriched in DNase I hypersensitive sites with ORC2 binding, compared to those without ORC2 binding (Fisher's Exact Test *P*-value < 10^{-200}).

Figure S5. Correlation between ORC2 levels with nascent RNA expression level measured by GRO-seq. The Pearson correlation coefficiency value was indicated.

Figure S6. Distribution of DNase I hypersensitive sites with or without
ORC2 binding, grouped based on chromatin states. Some DNase I hypersensitive sites do not fit into the indicated categories, which is why the sum of the ORC bound (red) or ORC-unbound (blue) sites is not 100%.

Figure S7. Correlation between ORC2 levels with chromatin accessibility (*A*), and histone modification levels, including H3K4me2 (*B*), H3K4me1 (*C*), H3K4me3 (*D*), H3K27ac (*E*) and H3K9ac (*F*) at 250,000 DNase hypersensitive sites. The Pearson correlation coefficiency values were indicated.

Figure S8. Compare ORC2 level in open acetylated regions without ORC2 peaks versus in random genomic regions with matching sizes. The Wilcoxon Rank Sum Test *P*-value was shown.

Figure S9. ORC2 binding is correlated with chromatin accessibility and H3K27Ac/H3K4me2 levels. We randomly selected 6,000 open chromatin regions with similar H3K27Ac (*A*), but with or without ORC2 binding sites (*B*). Then we examined chromatin accessibility (*C*), H3K9ac (*D*), H3K4me2 (*E*), H3K4me1 (*F*) and H3K3me3 (*G*) levels around. The Wilcoxon Rank Sum Test *P*-values comparing two groups of open chromatin regions were shown.

Figure S10: A predictive model for ORC binding based on chromatin status in K562 cells. (*A-B*) Open chromatin regions were grouped based on the predicted ORC binding probabilities, and the ORC2 binding levels estimated by ChIP-seq using the antibody recognizing C-terminus (*A*) and N-terminus (*B*) of ORC2 were plotted (boxplot on the left and mean and standard error values on the right). (*C-D*) The Area Under ROC Curve (AUC) values measuring performances of logistic regression classifiers predicting ORC2 binding status, based on indicated training parameters.

Figure S11: Analyses of ORC1 binding sites in HeLa cells. (*A*) Distribution of ORC1, H3K27Ac, DNase-seq and Pol II reads around gene promoter regions. (*B*) Read distribution around submits of ORC1, MYC and cJUN peak. (*C*) The DNase-seq and ChIP-seq read distribution for active histone modifications around ORC2 binding sites, ranked by ORC2 binding levels. (*D*) Open chromatin regions were grouped based on the predicted ORC binding probabilities, and the ORC1 binding levels estimated by ChIP-seq were plotted (boxplot on the left and mean and standard error values on the right).

Figure S12. Analyses of SNS sites in K562 cells. (*A*) Distribution of distances between SNS sites and closest ORC2 binding sites. (*B*) Density of SNS sites in per 100kb replicating DNA regions in G1, S1, S2, S3, S4 and G2 phases.

SUPPLEMENTARY TABLE LEGENDS

Table S1. ORC2 binding sites in K562 cells (Genome build: hg19)

Table S2. ChIP-qPCR primers for the validation of ORC2 binding sites

Table S3. Common fragile sites are often located in large ORC2 "poor" regions of the genome

Table S4. Recurrent deletions identified in human cancer often overlap with ORC2 "poor" regions

Table S5. Primer pair sequences used for quantitative PCR validation