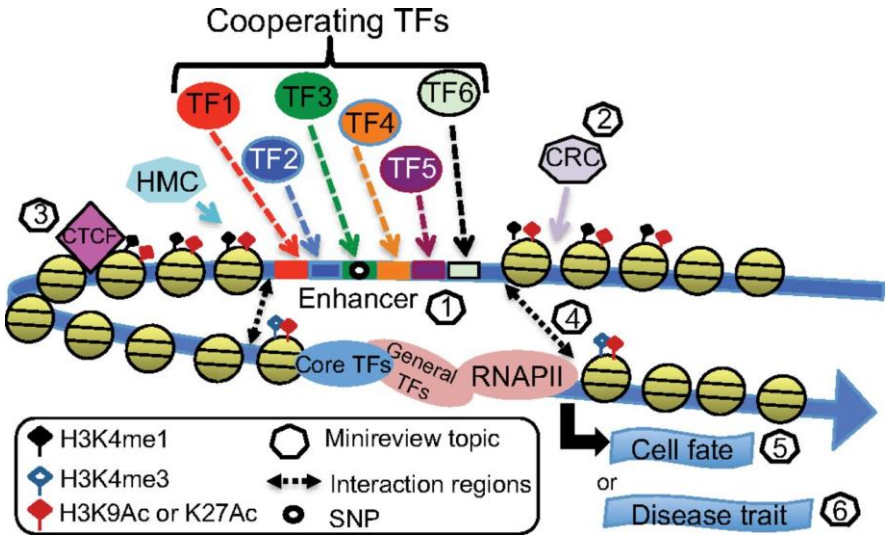
SNPs in the genomic regulatory regions may affect:

- **Enhancer Activation: loss of TFs interaction or TFs recruitment.**
- **Enhancer Selection: loss or association of LTDF**
- **Alteration of timing or specific tissues activation**
- **Long range interaction between genomic regulatory regions**



How are SNPs studying in genome-wide manner?

Super-Enhancers in the Control of Cell Identity and Disease

Denes Hnisz,[1,3] Brian J. Abraham,[1,3] Tong Ihn Lee,[1,3] Ashley Lau,[1,2] Violaine Saint-André,[1] Alla A. Sigova,[1] Heather A. Hoke,[1,2] and Richard A. Young[1,2,*]

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]These authors contributed equally to this work

*Correspondence: young@wi.mit.edu

http://dx.doi.org/10.1016/j.cell.2013.09.053

SUMMARY

Super-enhancers are large clusters of transcriptional enhancers that drive expression of genes that define cell identity. Improved understanding of the roles that super-enhancers play in biology would be afforded by knowing the constellation of factors that constitute these domains and by identifying super-enhancers across the spectrum of human cell types. We describe here the population of transcription factors, cofactors, chromatin regulators, and transcription apparatus occupying super-enhancers in embryonic stem cells and evidence that super-enhancers are highly transcribed. We produce a catalog of super-enhancers in a broad range of human cell types and find that super-enhancers associate with genes that control and define the biology of these cells. Interestingly, disease-associated variation is especially enriched in the super-enhancers of disease-relevant cell types. Furthermore, we find that cancer cells generate super-enhancers at oncogenes and other genes important in tumor pathogenesis. Thus, super-enhancers play key roles in human cell identity in health and in disease.

SUMMARY

Super-enhancers are large clusters of transcriptional enhancers that drive expression of genes that define cell identity. Improved understanding of the roles that super-enhancers play in biology would be afforded by knowing the constellation of factors that constitute these domains and by identifying super-enhancers across the spectrum of human cell types. We describe here the population of transcription factors, cofactors, chromatin regulators, and transcription apparatus occupying super-enhancers in embryonic stem cells and evidence that super-enhancers are highly transcribed. We produce a catalog of super-enhancers in a broad range of human cell types and find that super-enhancers associate with genes that control and define the biology of these cells. Interestingly, disease-associated variation is especially enriched in the super-enhancers of disease-relevant cell types. Furthermore, we find that cancer cells generate super-enhancers at oncogenes and other genes important in tumor pathogenesis. Thus, super-enhancers play key roles in human cell identity in health and in disease.

DEFINITION

AIM

1) Protein complexes

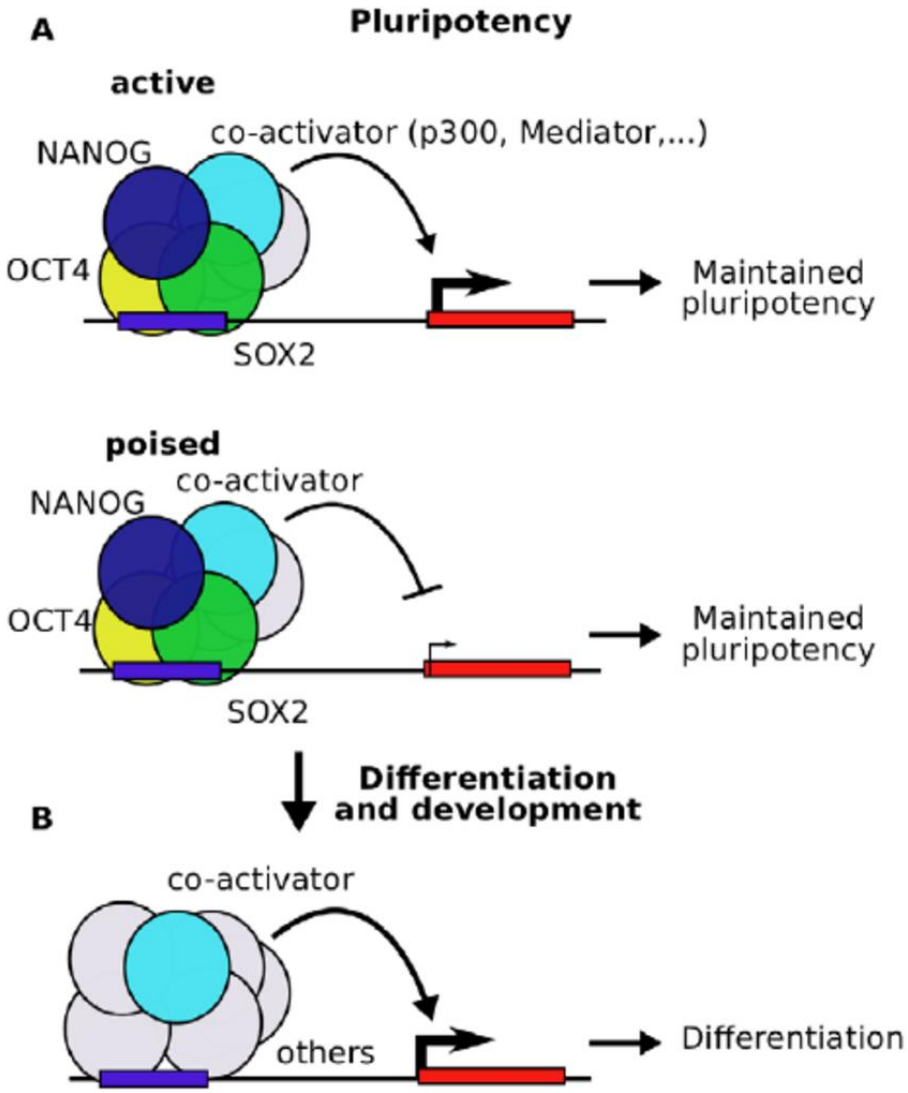
2) SE cell type-specific

3) SNPs linked to disease in SE

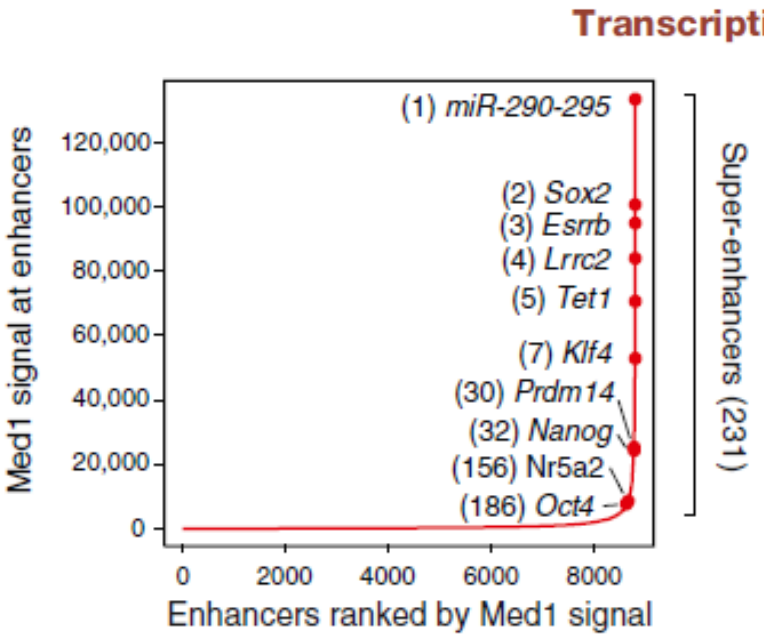
CONCLUSION

Murine Embrionic stem cells (ESC)

SOX2, Nanog, OCT4: transcription factors that bind SE controlling genes for mantained pluripotency



Mediator Coactivator Complexes and Master TFs are bound at Super-enhancers



Super-enhancers are clusters of enhancers—formed by binding of high levels of master transcription factors and Mediator coactivator—that drive high-level expression of genes encoding key regulators of cell identity (Figure 1A) (Whyte et al., 2013). Five ESC transcription factors were previously shown to occupy super-enhancers (Oct4, Sox2, Nanog, Klf4, and Esrrb) (Whyte et al., 2013), but there are many additional transcription factors that contribute to the control of ESCs (Ng and Surani, 2011; Orkin and Hochedlinger, 2011; Young, 2011). We compiled ChIP-seq data for 15 additional transcription factors in ESCs, for which high-quality ChIP-seq data were available, and investigated whether they occupy enhancers defined by Oct4, Sox2, and Nanog (OSN) co-occupancy (Whyte et al., 2013) (Table S1 avail-

(A) Distribution of Med1 ChIP-seq signal at enhancers reveals two classes of enhancers in ESCs. Enhancer regions are plotted in an increasing order based on their input-normalized Med1 ChIP-seq signal. Super-enhancers are defined as the population of enhancers above the inflection point of the curve. Example super-enhancers are highlighted along with their respective ranks and their associated genes.

Bioinformatic analysis for the definition of SE:

- Signal in proximity of the gene

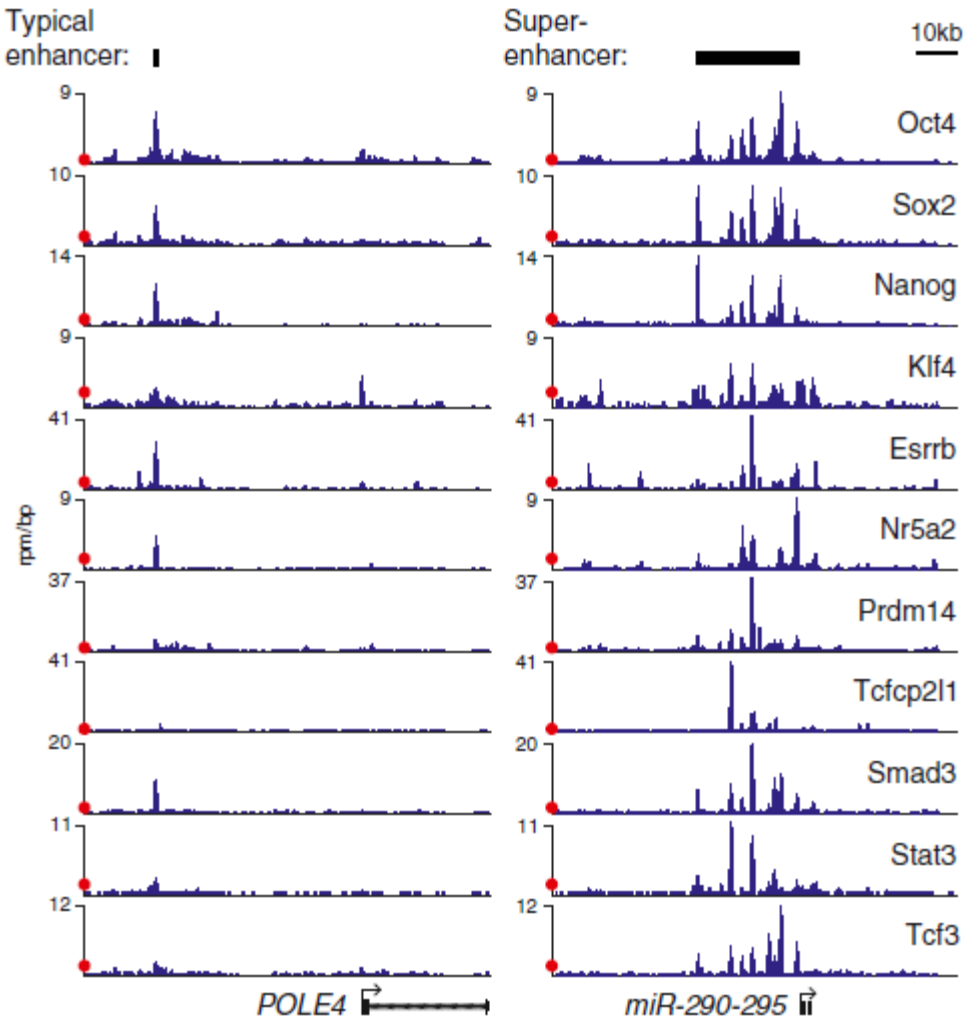
- signal extended in the genomic regions that identify SE

- increased numbers of reads into SE respect to costituent, single enhancer in the SE

- increased signal into SE respect to typical enhancer

Mediator Coactivator Complexes and Master TFs are bound at Super-enhancers

Chromatin Immunoprepitation Binding Profiles at target genes

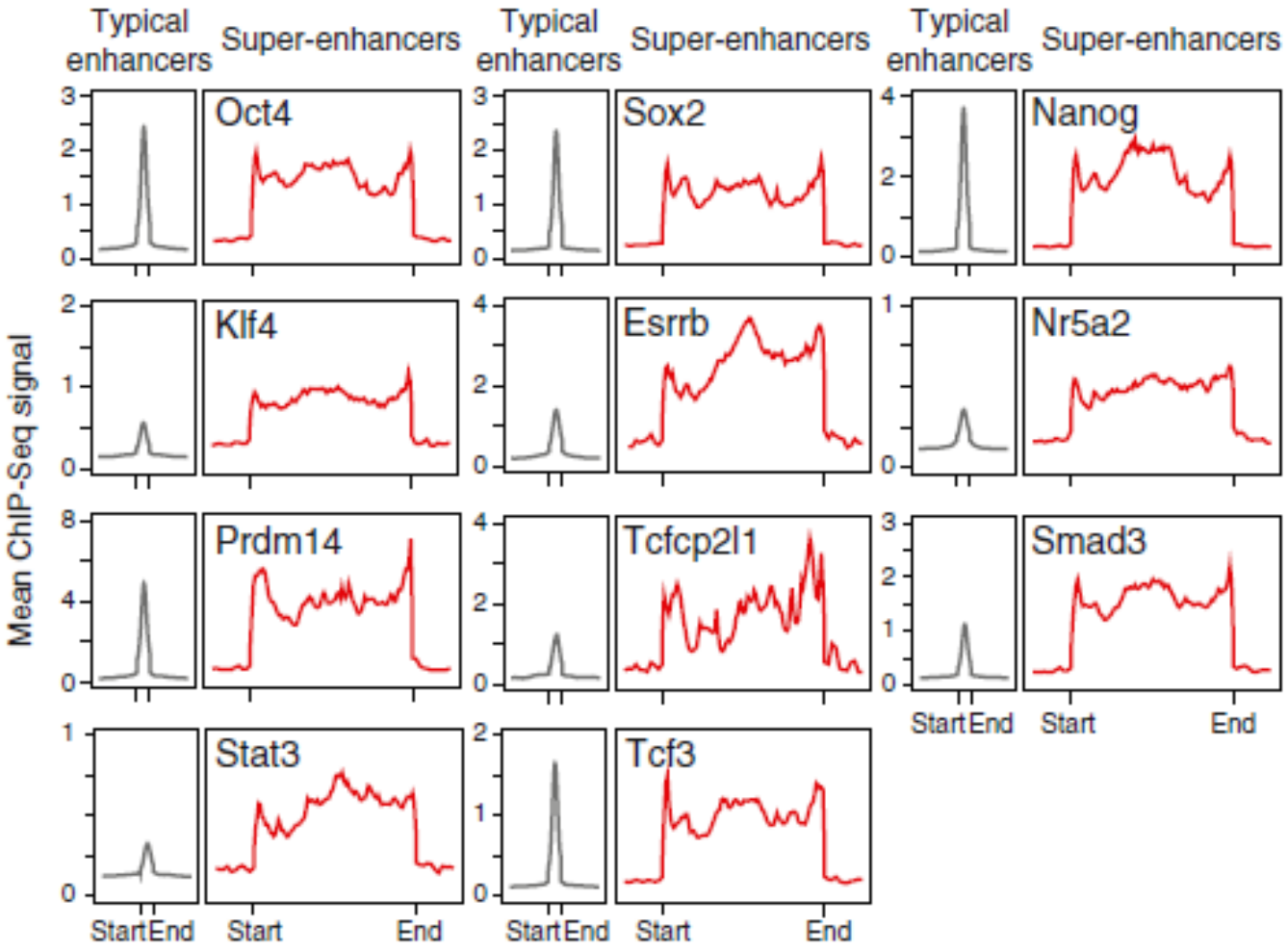


Transcription Factors

Specific Loci

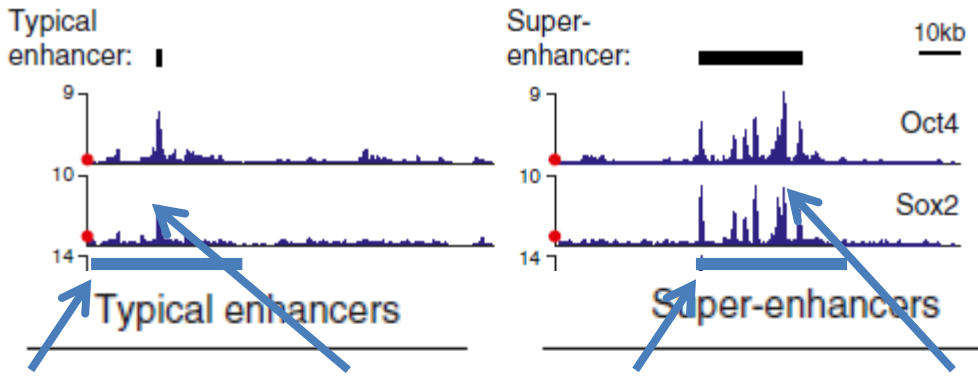
(B) ChIP-seq binding profiles for the indicated transcription factors at the *POLE4* and *miR-290–295* loci in ESCs. Red dots indicate the median enrichment of all bound regions in the respective ChIP-seq data sets and are positioned at maximum 20% of the axis height. rpm/bp, reads per million per base pair.

ChIP-seq signal across SE domains



(C) Metagene representations of the mean ChIP-seq signal for the indicated transcription factors across typical enhancers and super-enhancer domains. Metagenes are centered on the enhancer region, and the length of the enhancer reflects the difference in median lengths (703 bp for typical enhancers, 8,667 bp for super-enhancers). Additional 3 kb surrounding each enhancer region is also shown.

Reads distribution in regions and constituents (single enhancers into SE) (rpm/bp)

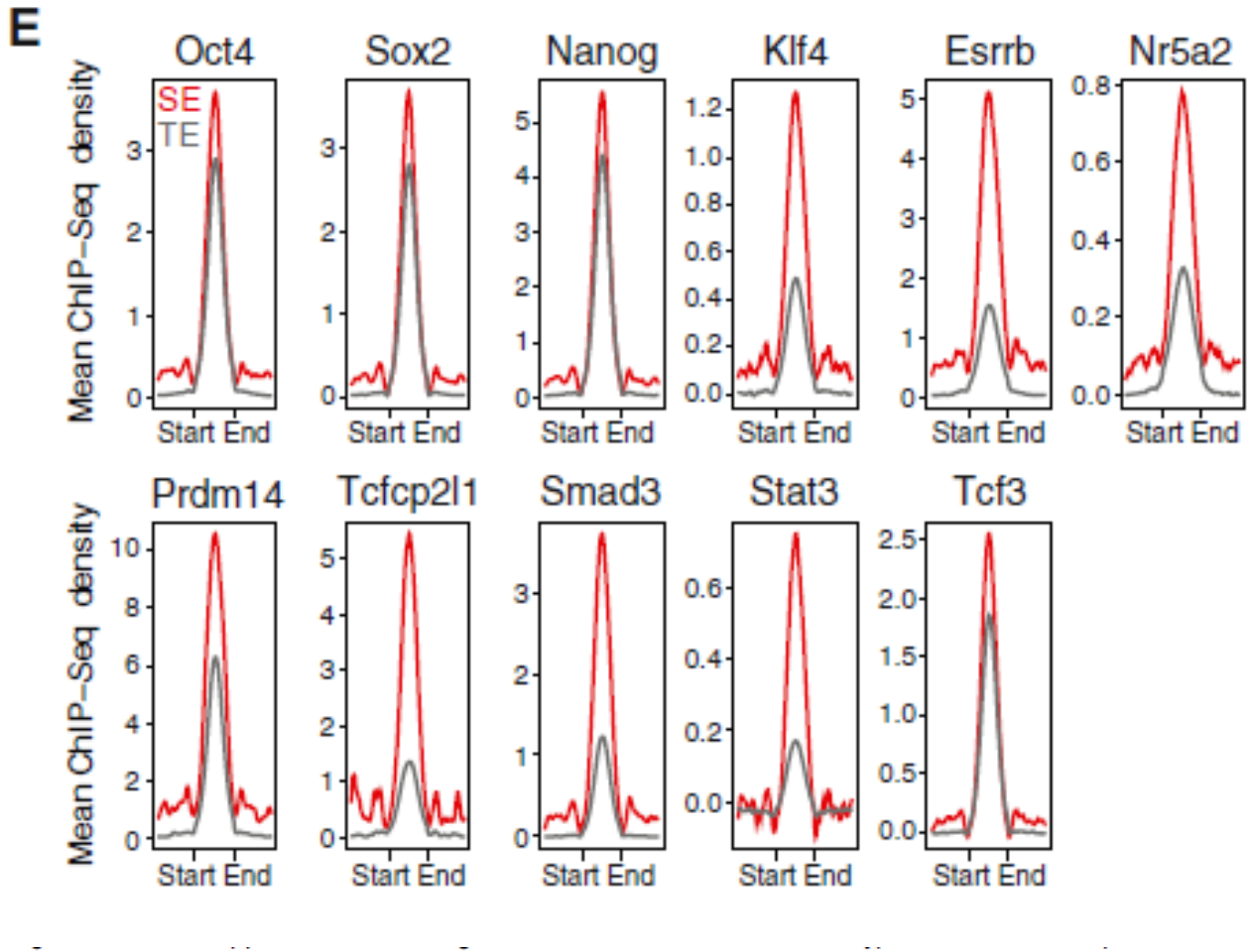


D

	Typical enhancers					Super-enhancers				
	Regions		Constituents			Regions		Constituents		
	Total signal	Density	Total signal	Density	Enhancer reads %	Total signal	Density	Total signal	Density	Enhancer reads %
Oct4:	1x	1x	1x	1x	90 %	4.5x	0.9x	1.6x	1.3x	10 %
Sox2:	1x	1x	1x	1x	90 %	4.2x	1.8x	1.7x	1.2x	10 %
Nanog:	1x	1x	1x	1x	91 %	3.9x	0.9x	1.5x	1.2x	9 %
Klf4:	1x	1x	1x	1x	85 %	8.4x	1.9x	2.9x	2.2x	15 %
Esrrb:	1x	1x	1x	1x	81 %	10.8x	2.6x	3.8x	3.0x	19 %
Nr5a2:	1x	1x	1x	1x	85 %	7.5x	1.7x	2.7x	2.1x	15 %
Prdm14:	1x	1x	1x	1x	87 %	6.2x	5.4x	2.3x	1.9x	13 %
Tcfcp2l1:	1x	1x	1x	1x	80 %	10.7x	0.9x	4.1x	3.3x	20 %
Smad3:	1x	1x	1x	1x	82 %	9.1x	2.8x	3.6x	2.8x	18 %
Stat3:	1x	1x	1x	1x	83 %	7.4x	5.8x	3.5x	2.5x	17 %
Tcf3:	1x	1x	1x	1x	90 %	4.0x	1.1x	1.7x	1.3x	10 %

(D) Fold difference values of ChIP-seq signal between typical enhancers and super-enhancers for the indicated transcription factors. Total signal indicates the mean ChIP-seq signal (total reads) at typical enhancers and super-enhancers normalized to the mean value at typical enhancers. Density indicates the mean ChIP-seq density at constituent enhancers (rpm/bp) of typical enhancers and super-enhancers normalized to the mean value at typical enhancers. Enhancer read % indicates the percentage of all reads mapped to enhancer regions that fall in the constituents of typical enhancer or super-enhancer regions.

ChIP-Seq density on TFs binding sites



(E) Metagene representations of the mean ChIP-seq density for the indicated transcription factors across the constituent enhancers within typical enhancers and super-enhancers. Each metagene is centered on enhancer constituents. Additional 2.5 kb surrounding the constituent enhancer regions is also shown.

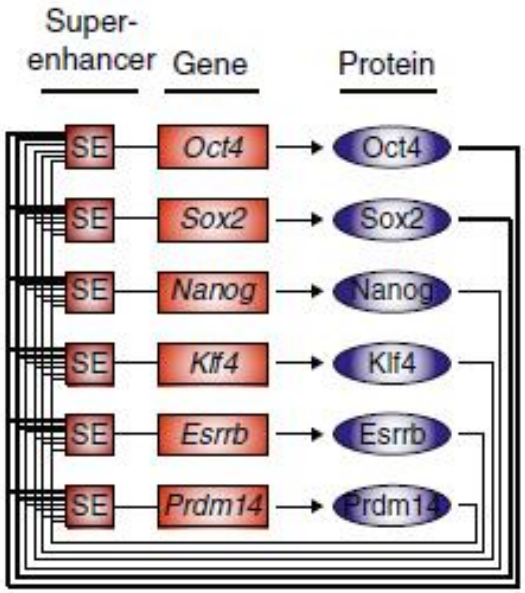
TFs motif enrichment are used to associate gene target

F

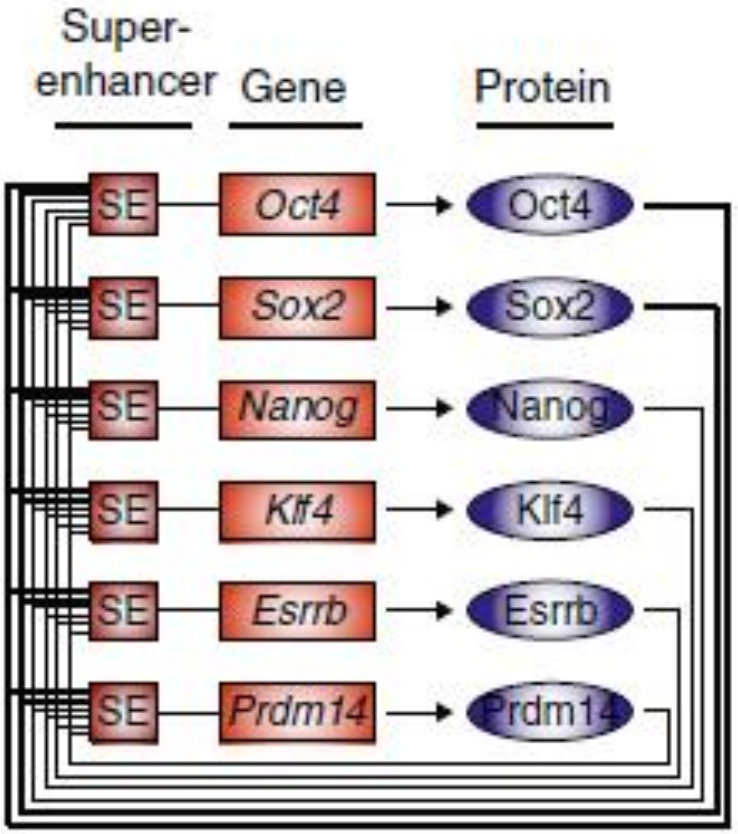
Transcription factor	Motif	P-value
Oct4		$9.19*10^{-64}$
Sox2		$3.01*10^{-67}$
Nanog		$9.46*10^{-17}$
Klf4		$4.33*10^{-6}$
Esrrb		$2.55*10^{-84}$
Nr5a2	n.a.	n.a.

Transcription factor	Motif	P-value
Prdm14	n.a.	n.a.
Tcfcp2l1		$6.83*10^{-11}$
Smad3		$9.31*10^{-11}$
Stat3		$2.90*10^{-10}$
Tcf3		$5.46*10^{-27}$

G

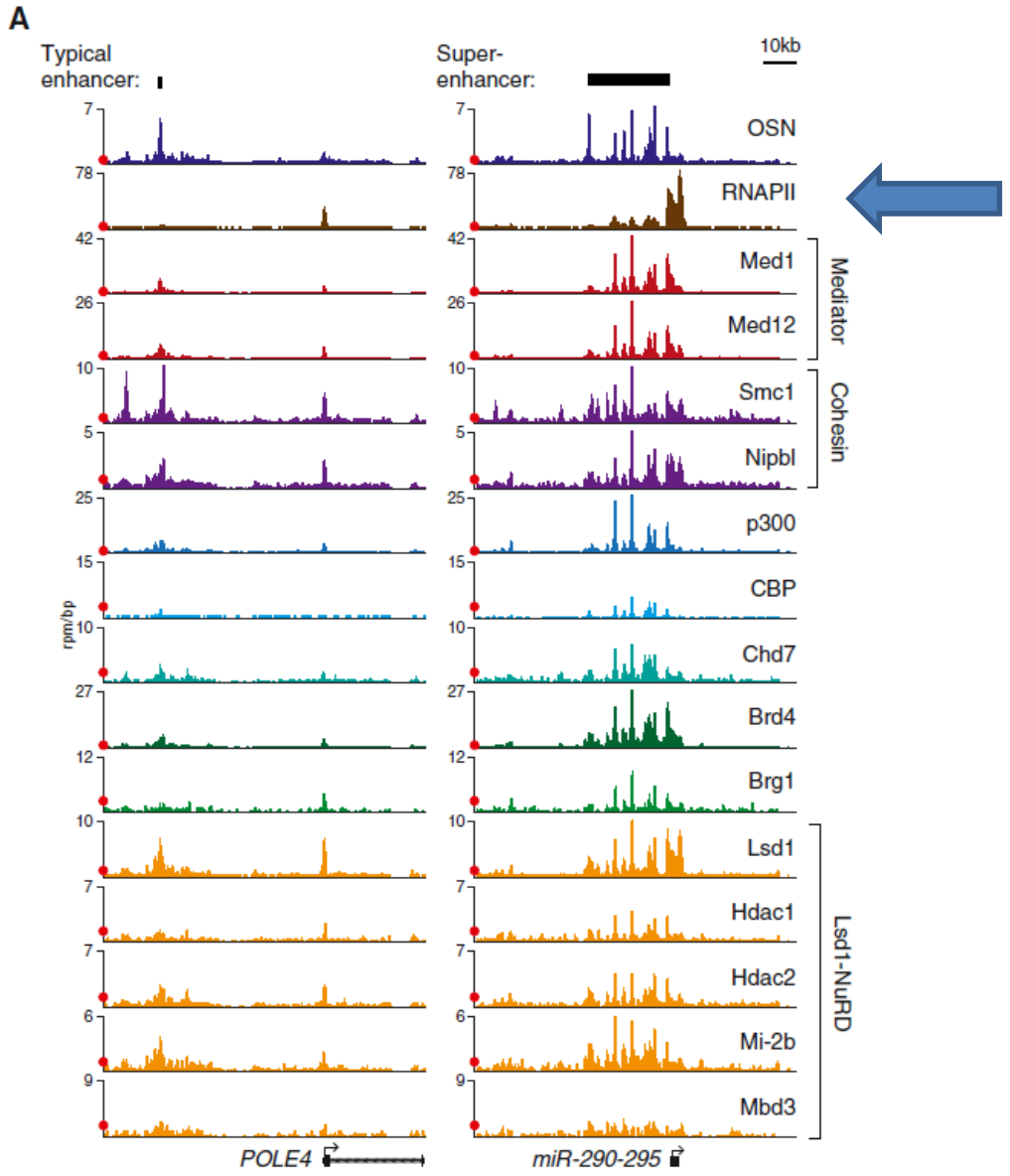


Core Transcriptional Regulatory Circuit of ESCs



(G) Revised model of the core transcriptional regulatory circuitry of ESCs. The model contains an interconnected autoregulatory loop consisting of transcription factors that meet three criteria: (1) their genes are driven by super-enhancers, (2) they co-occupy their own super-enhancers as well as those of the other transcription factor genes in the circuit, and (3) they play essential roles in regulation of ESC state and iPSC reprogramming. The layout of the circuit model was adapted from Whyte et al. (2013).

Super-enhancers are occupied by a large portion of the enhancer-associated RNA polymerase II

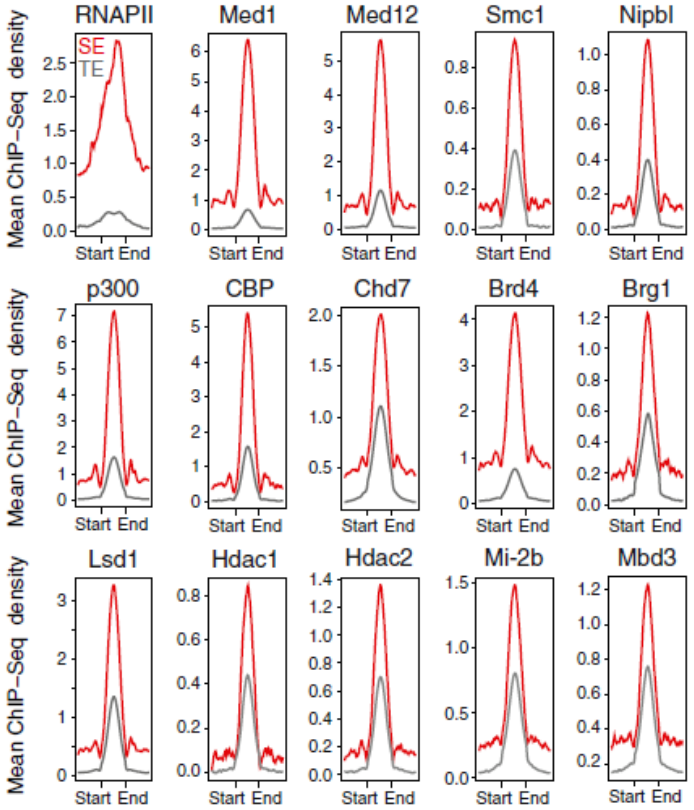


A large fraction of these enhancer cofactors are associated with super-enhancers

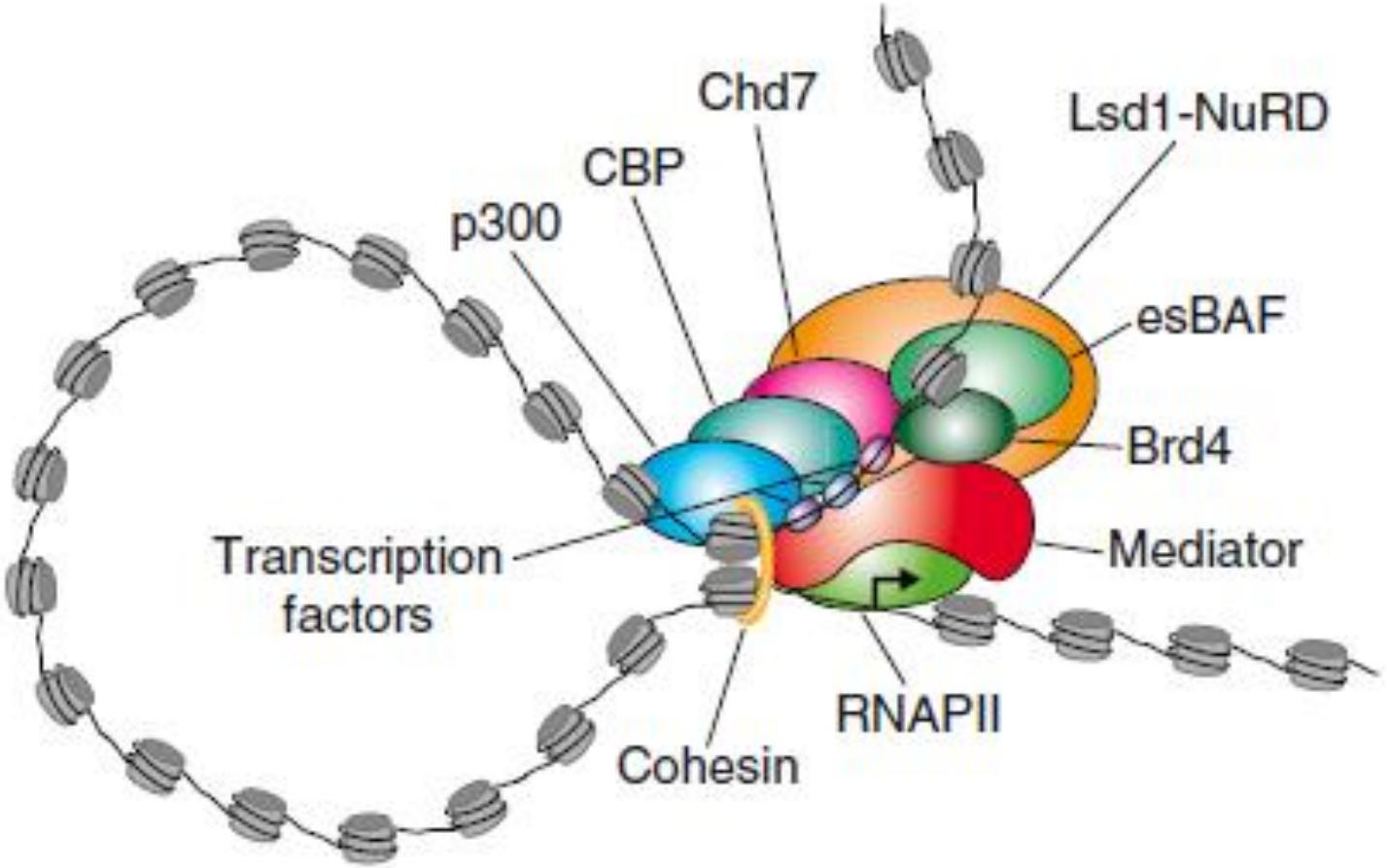
C

	Typical enhancers					Super-enhancers				
	Regions		Constituents			Regions		Constituents		
	Total signal	Density	Total signal	Density	Enhancer reads %	Total signal	Density	Total signal	Density	Enhancer reads %
RNAPII:	1x	1x	1x	1x	68 %	28.8x	8.1x	8.2x	7.1x	32 %
Med1:	1x	1x	1x	1x	64 %	28.2x	6.8x	10.3x	8.1x	36 %
Med12:	1x	1x	1x	1x	75 %	16.0x	3.5x	5.8x	4.5x	25 %
Smc1:	1x	1x	1x	1x	86 %	8.4x	2.2x	2.9x	2.3x	14 %
Nipbl:	1x	1x	1x	1x	80 %	9.4x	2.1x	3.2x	2.5x	15 %
p300:	1x	1x	1x	1x	86 %	13.3x	0.8x	2.0x	4.2x	14 %
CBP:	1x	1x	1x	1x	80 %	10.7x	2.4x	4.0x	3.1x	20 %
Chd7:	1x	1x	1x	1x	87 %	7.6x	1.5x	2.3x	1.8x	13 %
Brd4:	1x	1x	1x	1x	74 %	19.7x	4.2x	6.2x	5.0x	26 %
Brg1:	1x	1x	1x	1x	85 %	8.6x	1.5x	2.7x	2.0x	15 %
Lsd1:	1x	1x	1x	1x	85 %	9.0x	1.9x	2.9x	2.3x	15 %
Hdac1:	1x	1x	1x	1x	88 %	6.5x	1.3x	2.1x	1.6x	12 %
Hdac2:	1x	1x	1x	1x	87 %	6.9x	1.4x	2.2x	1.7x	13 %
Mi-2b:	1x	1x	1x	1x	88 %	7.4x	1.5x	2.2x	1.7x	12 %
Mbd3:	1x	1x	1x	1x	88 %	8.3x	1.5x	2.1x	1.7x	12 %
RNA:	1x	1x	1x	1x	74 %	24.3x	6.0x	5.4x	4.6x	26 %

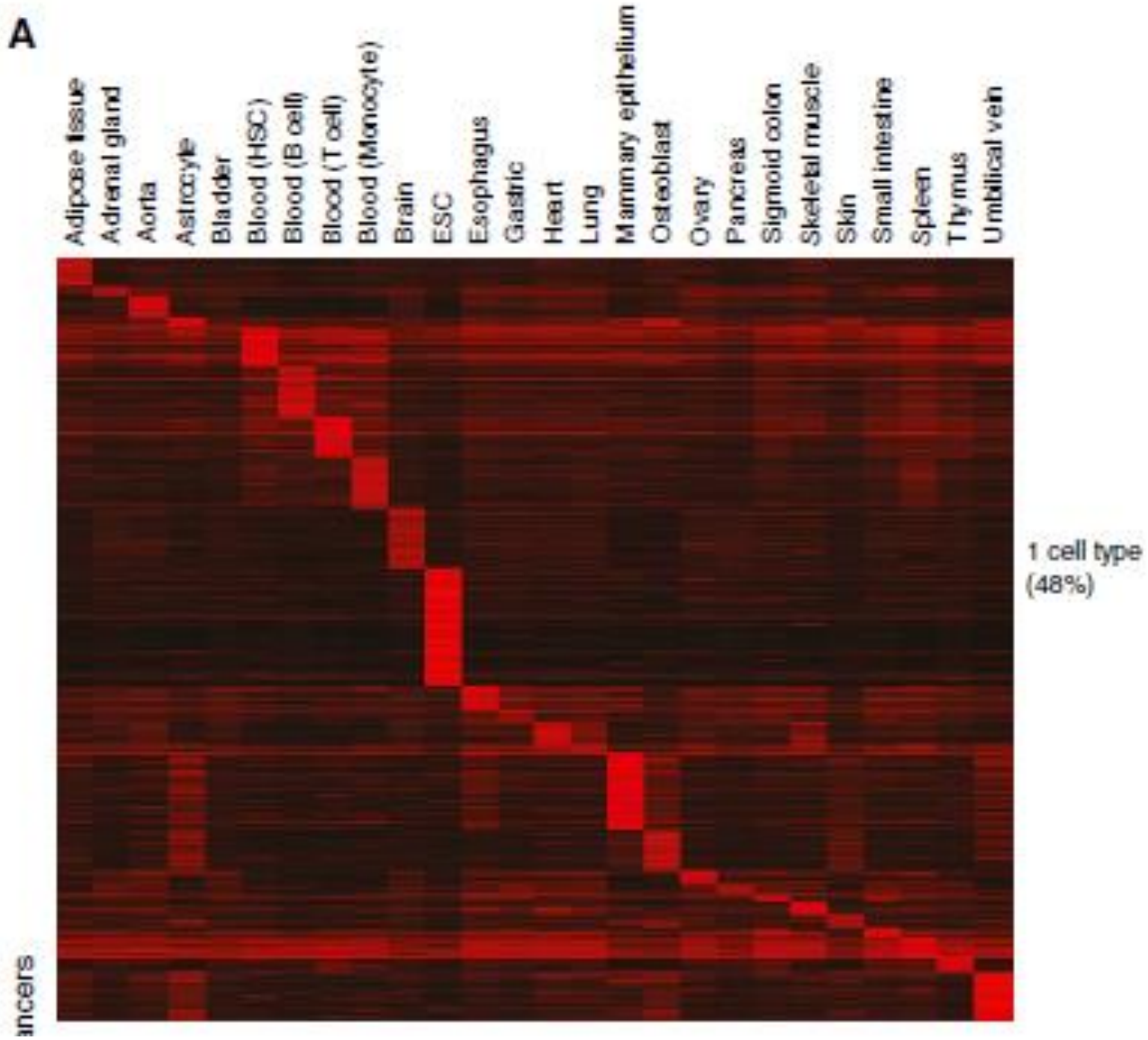
D



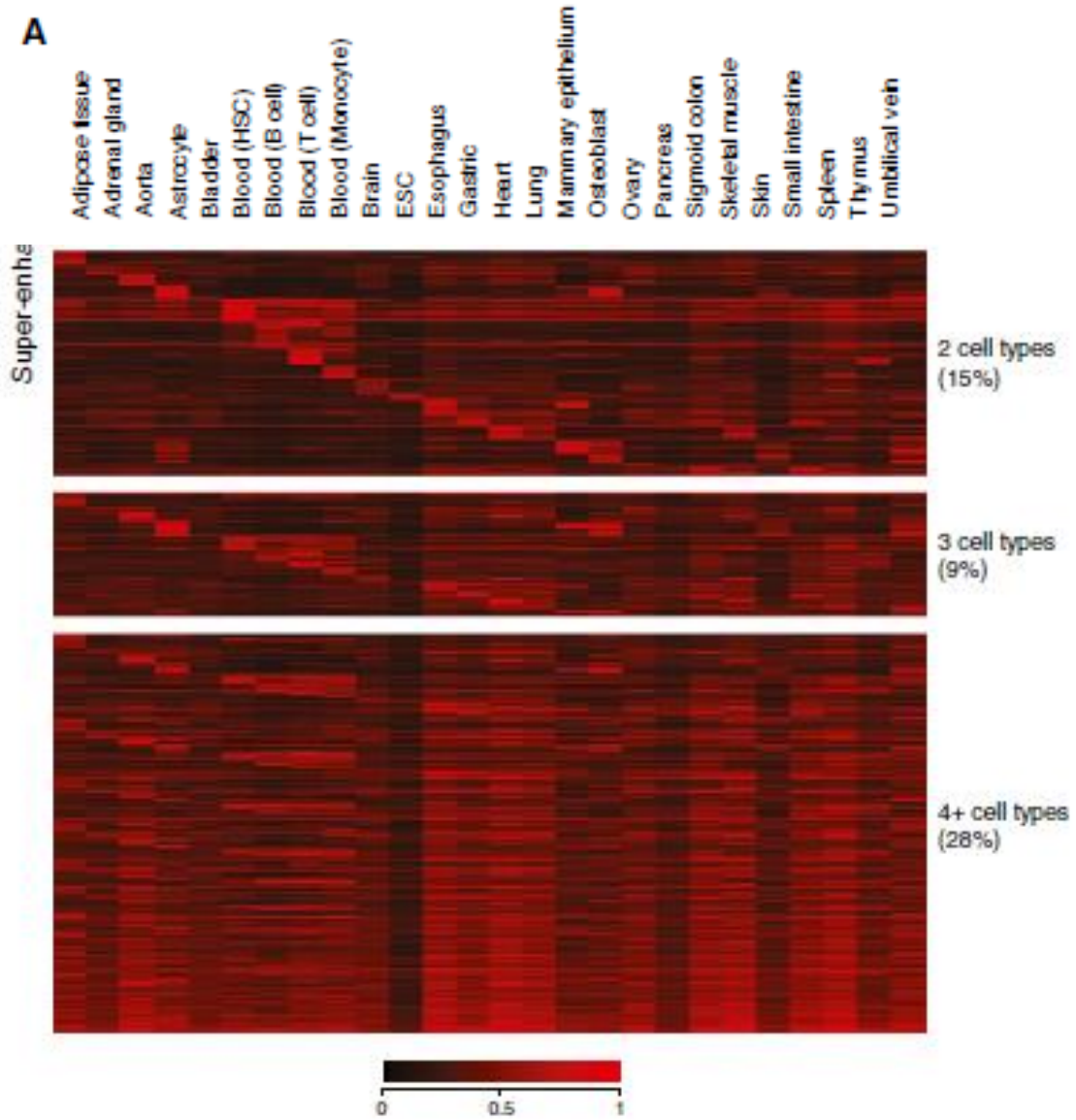
Model showing **RNAPII, transcriptional cofactors, and chromatin regulators that are found in ESC super-enhancers**. The indicated proteins are responsible for diverse enhancer-related functions, such as enhancer looping, gene activation, nucleosome remodeling, and histone modification.



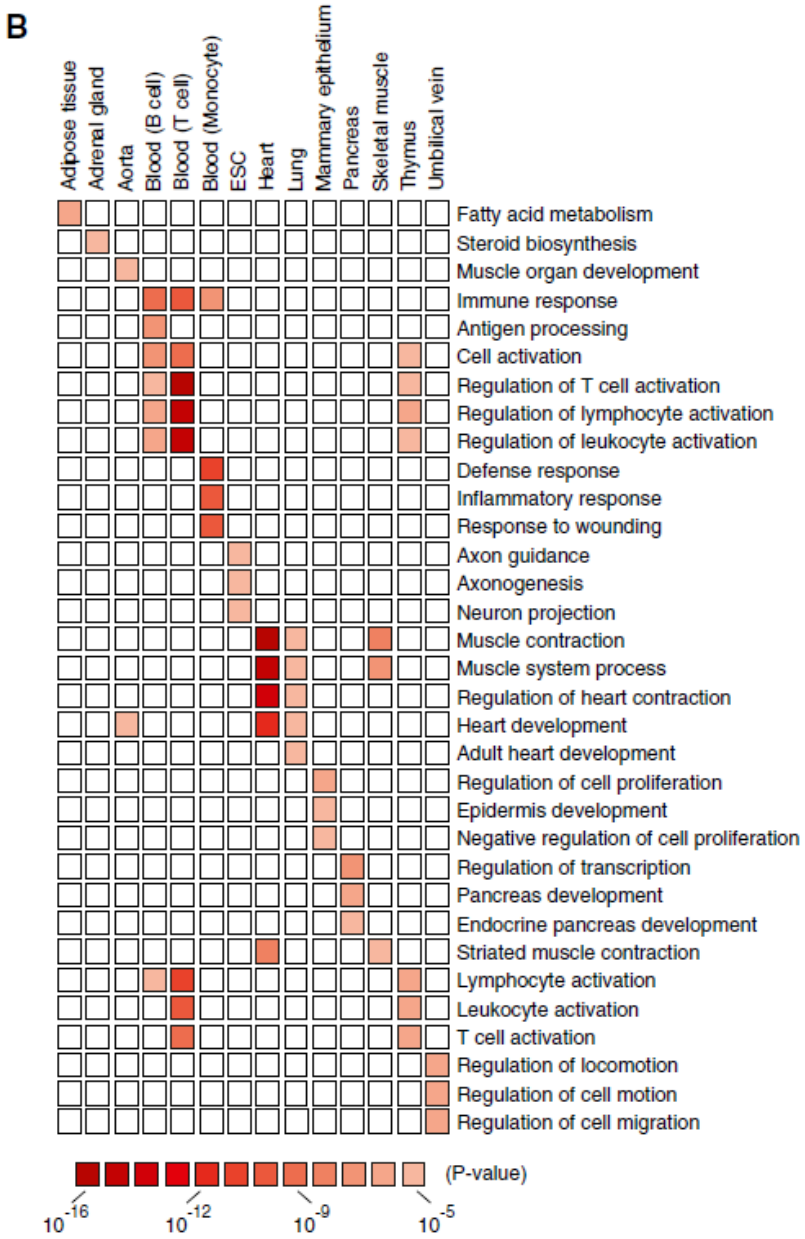
SUPERENHANCER SHARED BETWEEN SEVERAL CELL TYPES



SUPERENHANCER SHARED BETWEEN SEVERAL CELL TYPES



GENE ASSOCIATED TO SUPERENHANCER IN SEVERAL CELL TYPES: GENE ONTOLOGY



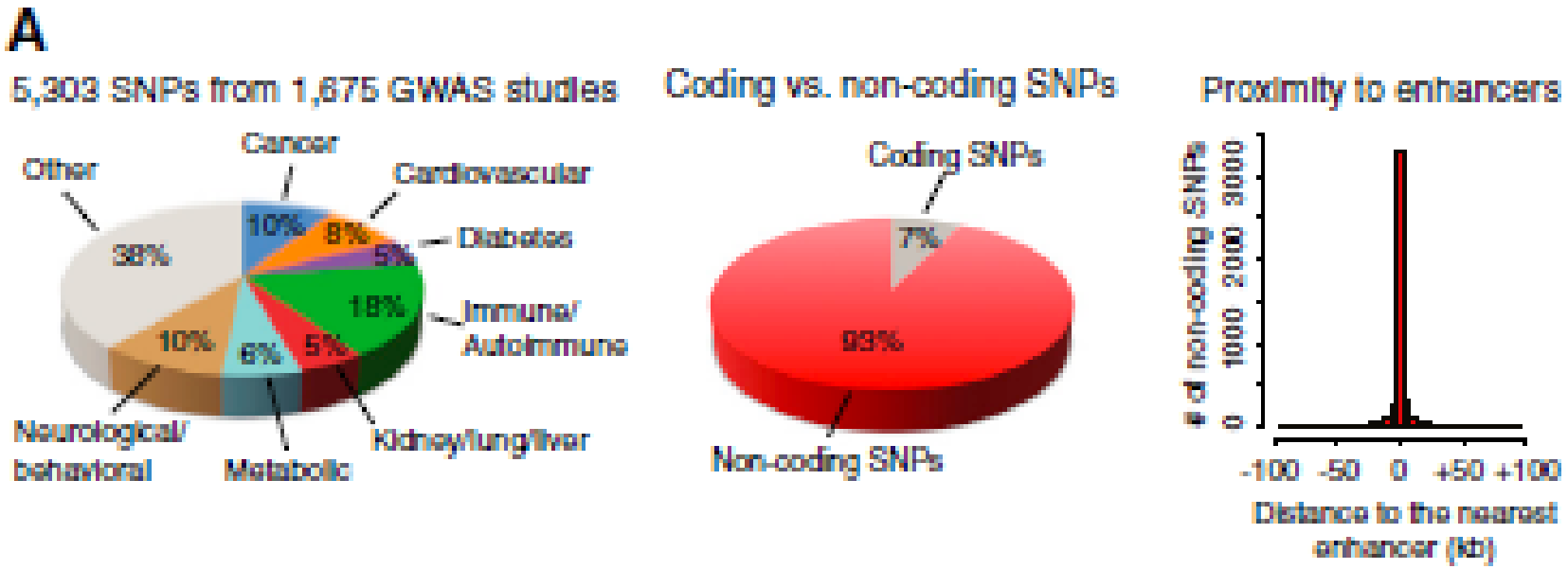
MASTER TRANSCRIPTION FACTORS IN SIX CELL TYPES

C

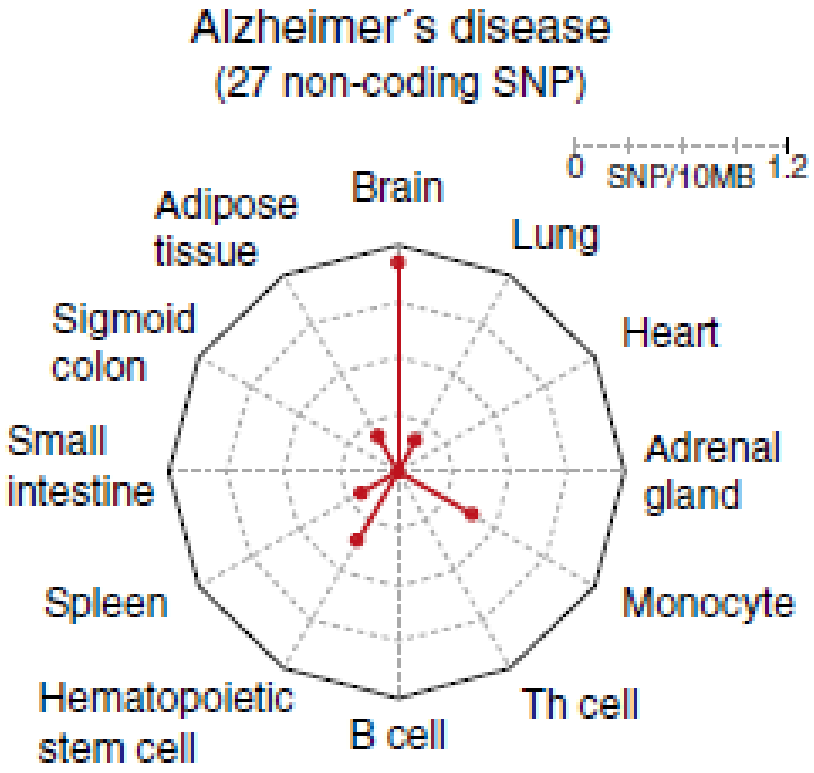
Brain	Heart	Skeletal muscle	Lung	Adipose tissue	B cell
NKX2-2	TBX20	MYOD1	NFIB	PPARG	IKZF3
OLIG1	TBX5	PITX2	TBX5	CEBPB	PAX5
BRN2	MEF2A	SIX1	CEBPA	CEBPD	BACH2
SOX10	NKX2-5	TEAD4	TBX2	CREB1	OCT2
SOX2	GATA4		TBX3		IKZF1
					IRF8

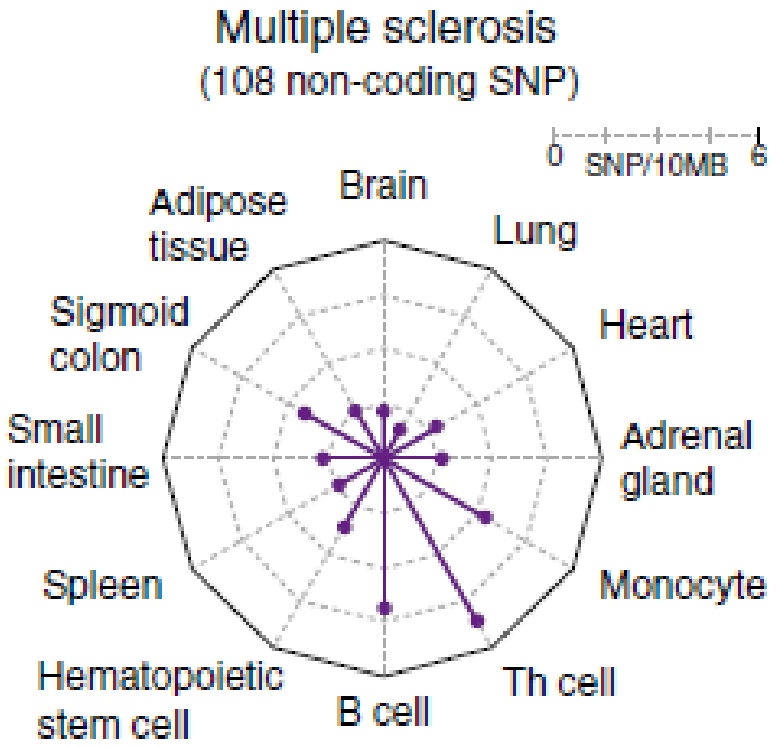
(C) Candidate master transcription factors identified in six cell types. All of these transcription factors were previously demonstrated to play key roles in the biology of the respective cell type or facilitate reprogramming to the respective cell type.

SINGLE NUCLEOTIDE MUTATIONS LINKED TO DISEASE (GWAS) ASSOCIATED TO SUPERENHANCERS



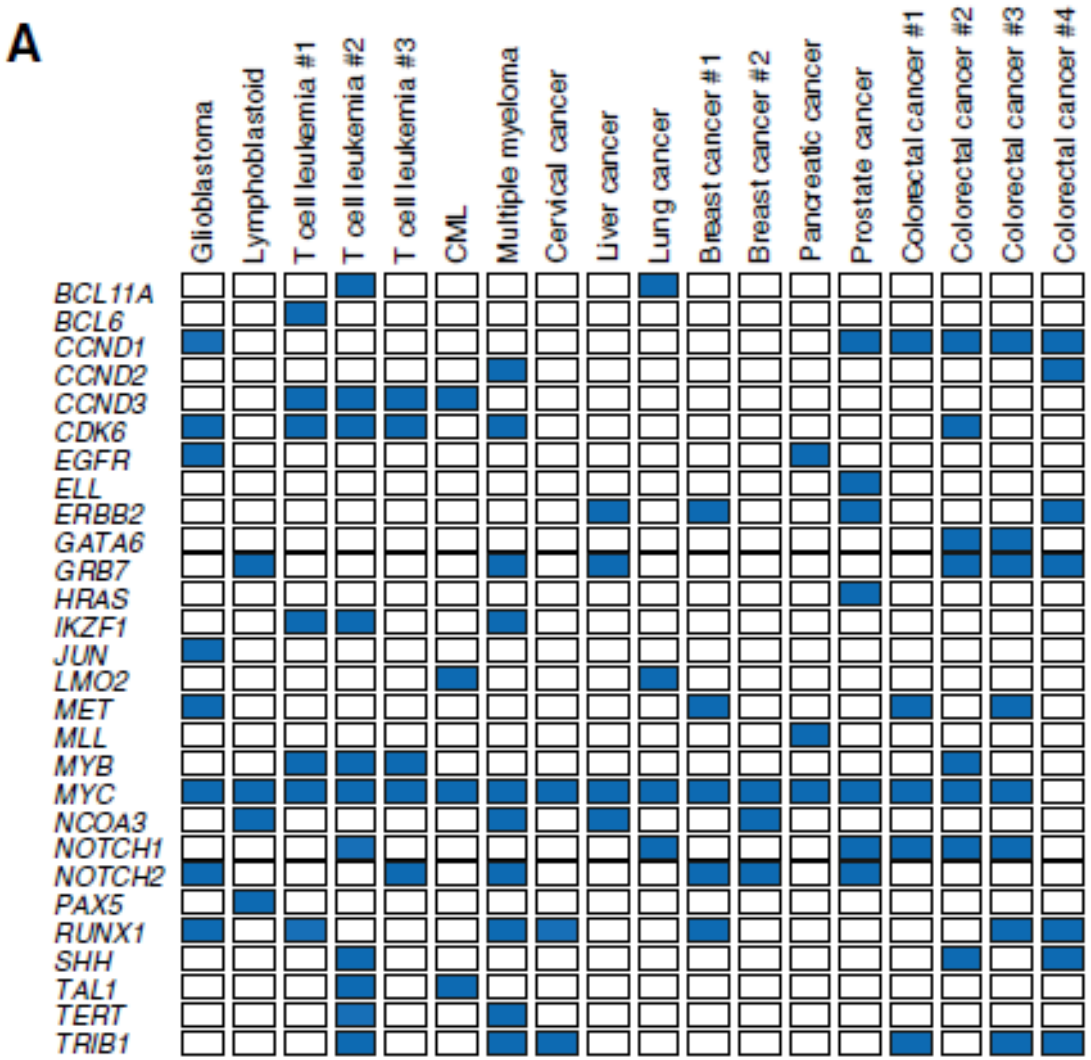
SINGLE NUCLEOTIDE MUTATIONS LINKED TO DISEASE (GWAS) ASSOCIATED TO SUPERENHANCERS



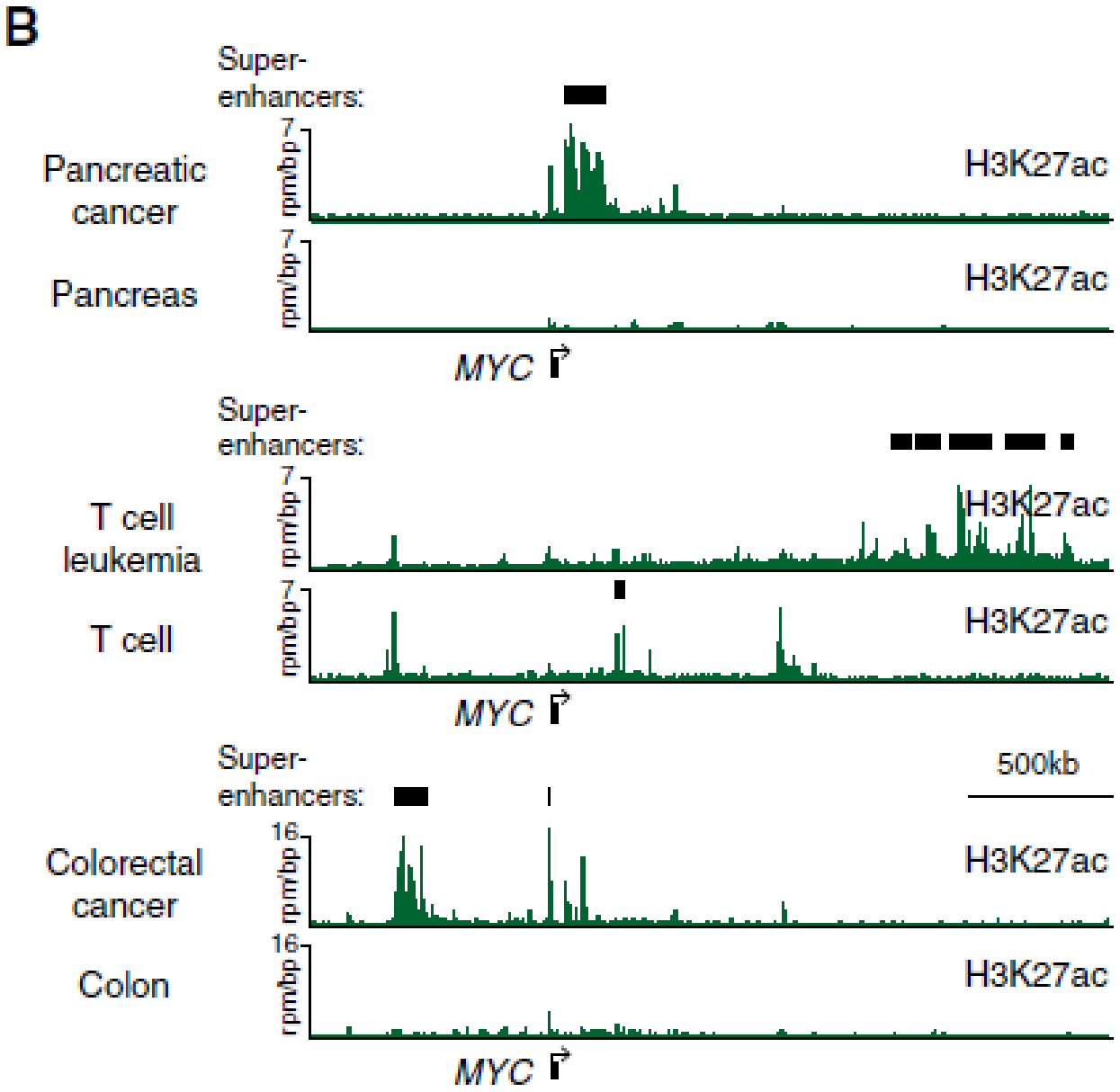


Super-enhancers in Cancer

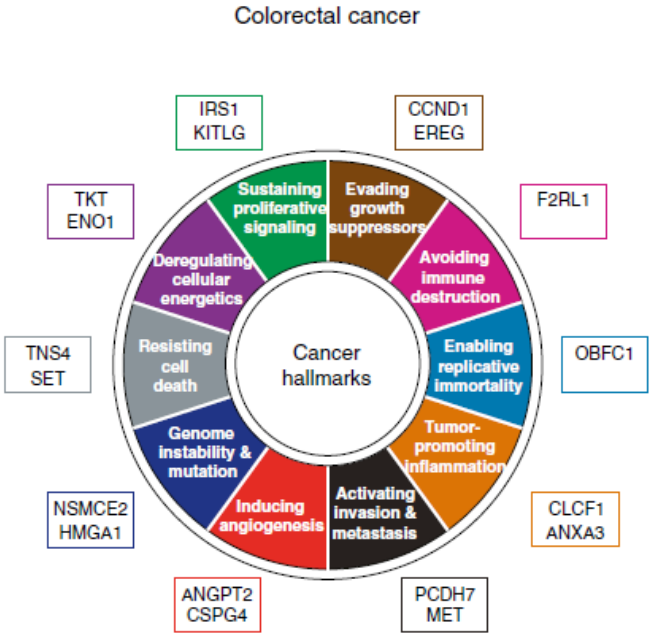
Genes associated with SE and involved in cancer progression

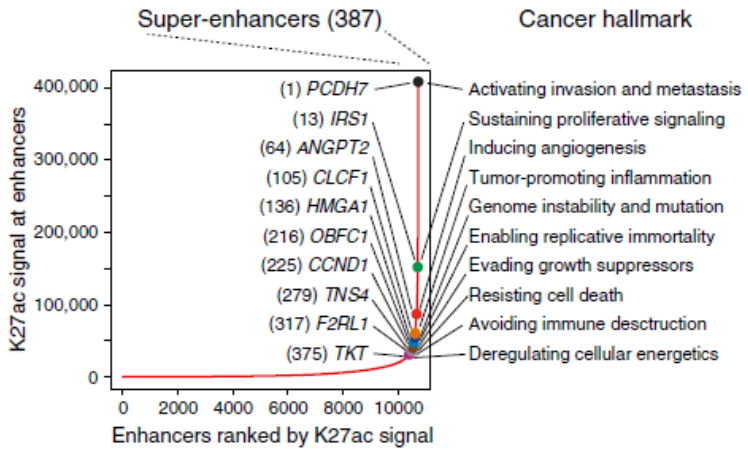


The super-enhancers formed in the MYC locus were tumor type specific



Super-enhancers are associated with genes that act as hallmarks in colonrectal cancer





Methods to study Superenhancers:
From prediction of SE by ChIP-Seq to experimental validation

Convergence of Developmental and Oncogenic Signaling Pathways at Transcriptional Super-Enhancers

Denes Hnisz,[1,4] Jurian Schuijers,[1,4] Charles Y. Lin,[2] Abraham S. Weintraub,[1,3] Brian J. Abraham,[1] Tong Ihn Lee,[1] James E. Bradner,[2] and Richard A. Young[1,3,*]
[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA
[2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

SUMMARY

Super-enhancers and stretch enhancers (SEs) drive expression of genes that play prominent roles in normal and disease cells, but the functional importance of these clustered enhancer elements is poorly understood, so it is not clear why genes key to cell identity have evolved regulation by such elements. Here, we show that SEs consist of functional constituent units that concentrate multiple developmental signaling pathways at key pluripotency genes in embryonic stem cells and confer enhanced responsiveness to signaling of their associated genes. Cancer cells frequently acquire SEs at genes that promote tumorigenesis, and we show that these genes are especially sensitive to perturbation of oncogenic signaling pathways. Super-enhancers thus provide a platform for signaling pathways to regulate genes that control cell identity during development and tumorigenesis.

BACKGROUND

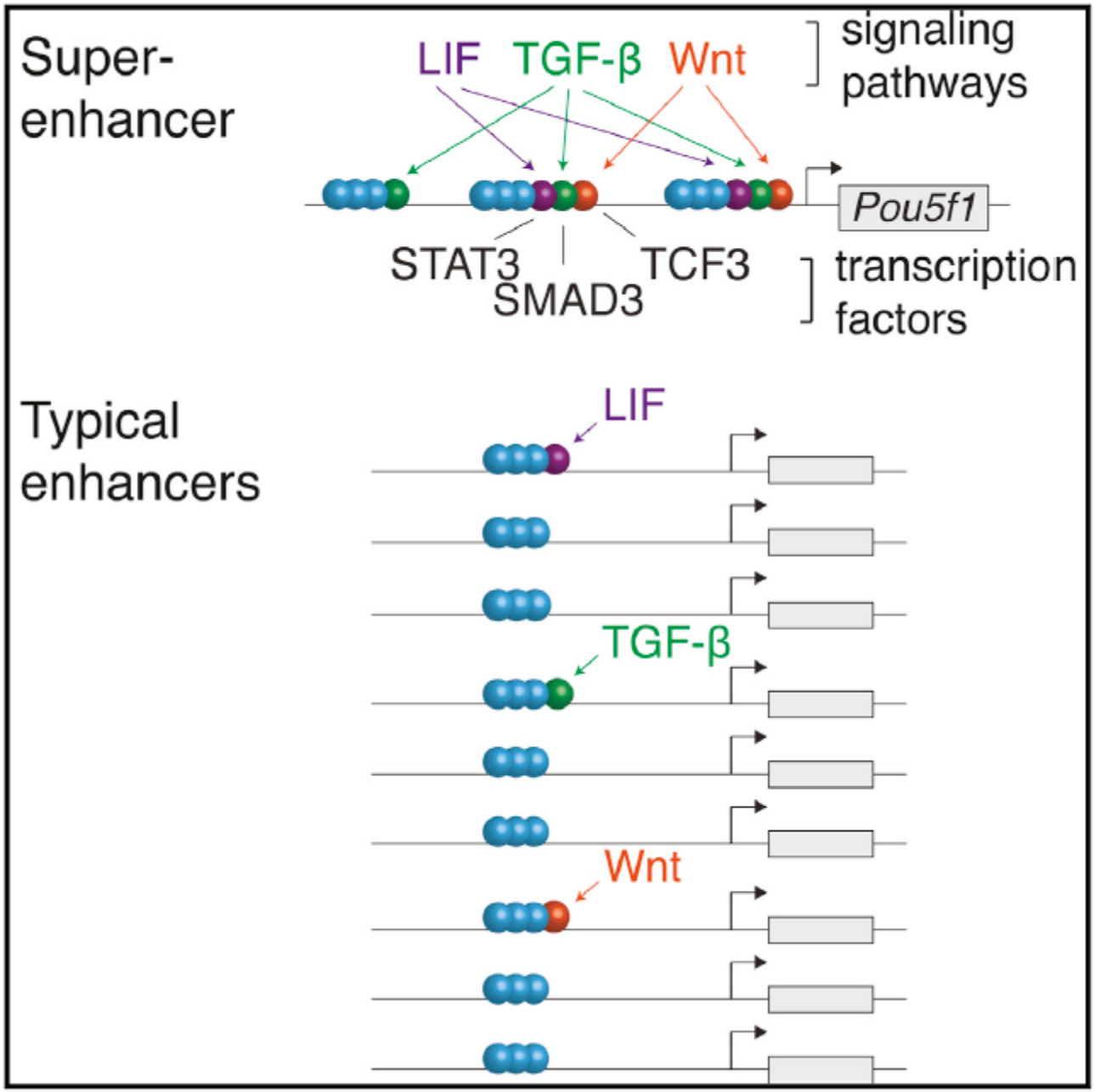
AIM

Super-enhancers (SE)
Functional costituent units

Cancer cells SE target
for oncogenic signalling

CONCLUSION

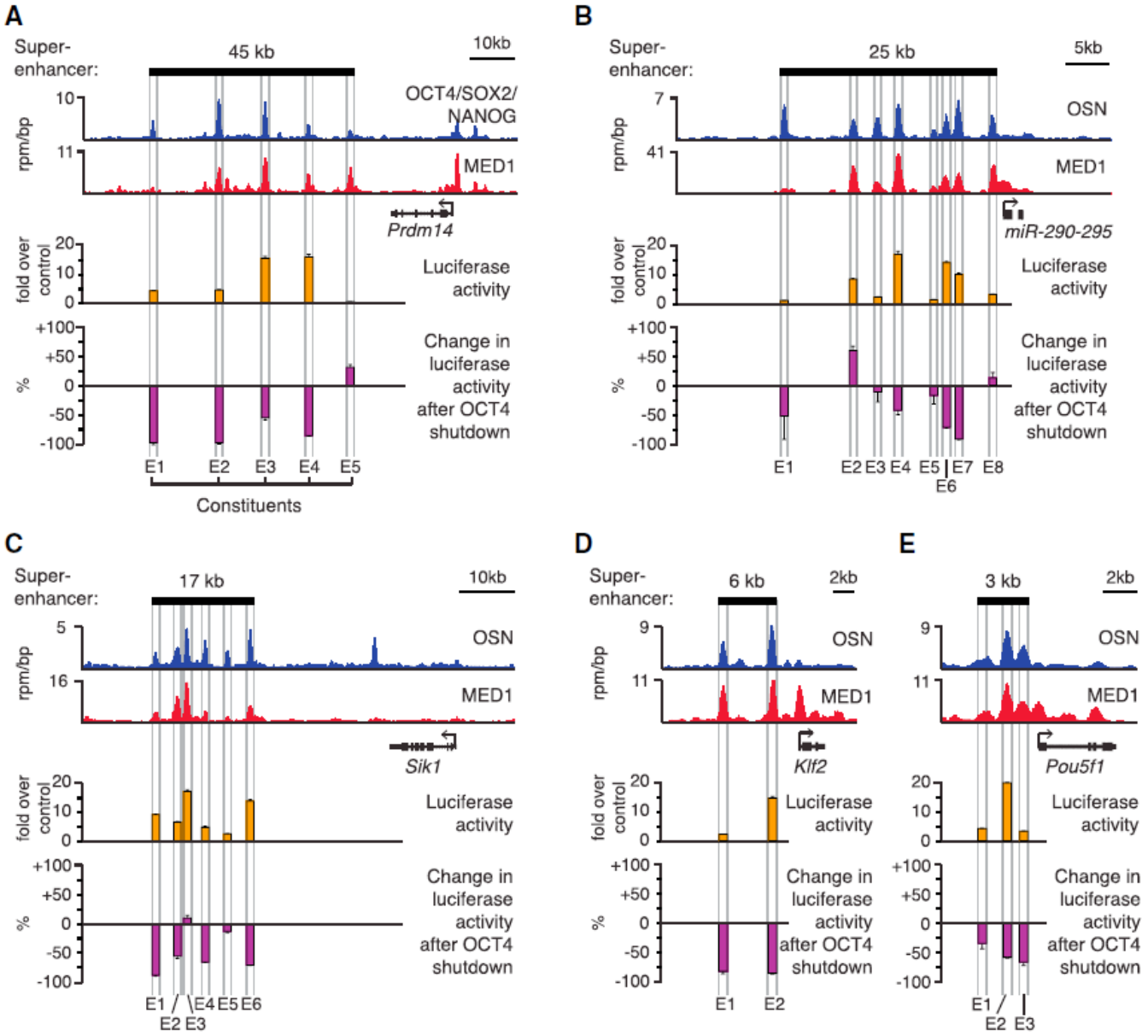
Graphical Abstract



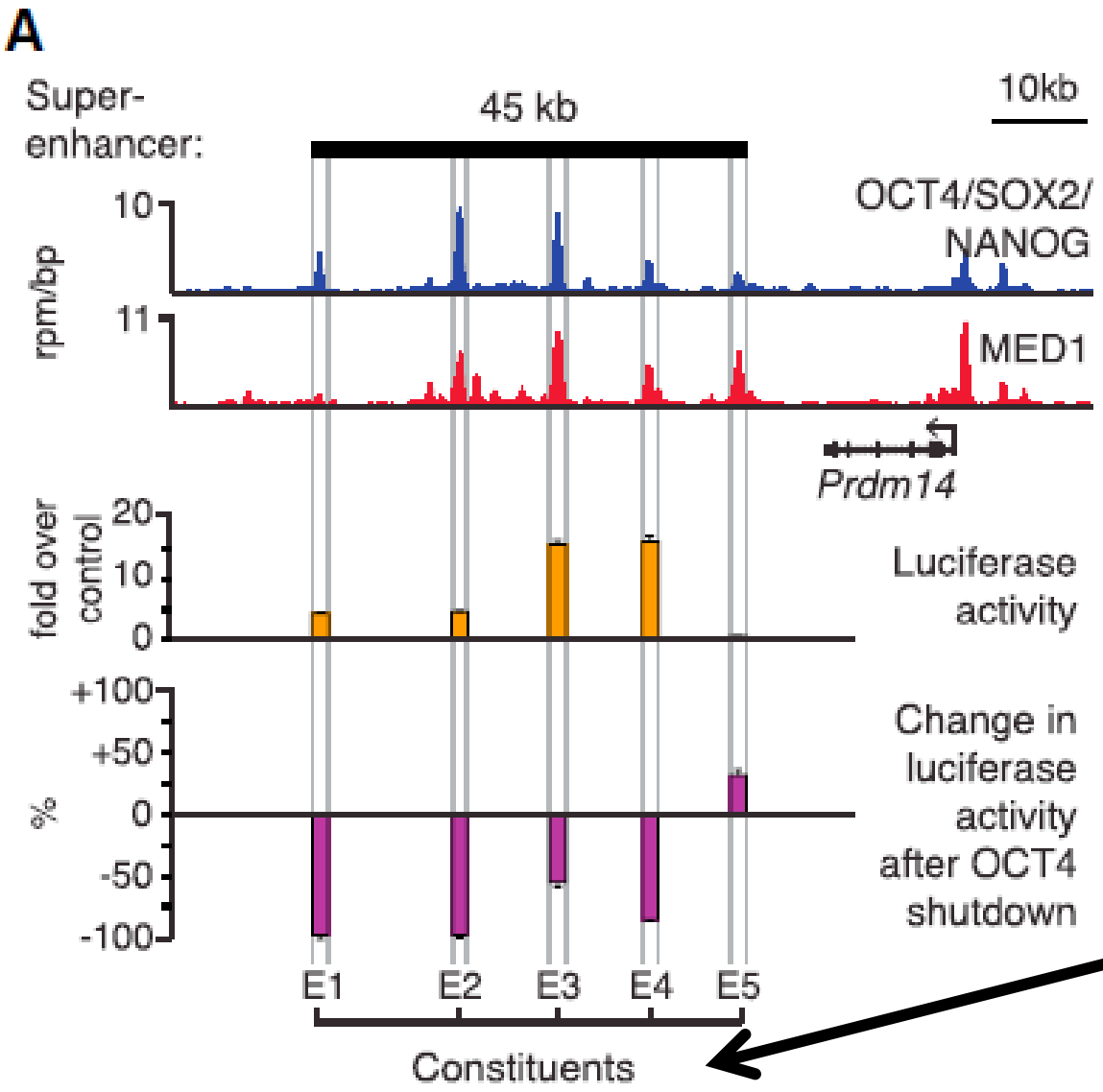
Highlights

- Super-enhancers (SEs) consist of clusters of active enhancers
- SEs are frequently bound by terminal transcription factors of signaling pathways
- SE-driven genes are especially responsive to signaling input
- SEs acquired in cancer cells are responsive to oncogenic signaling

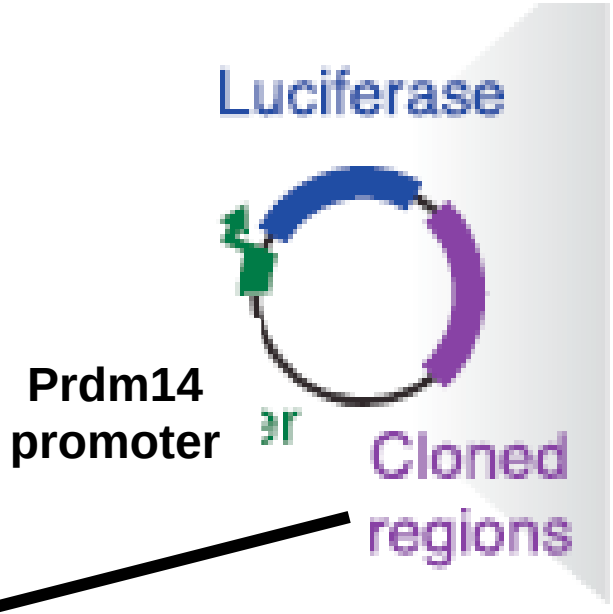
ACTIVITY OF SUPER-ENHANCER CONSTITUENTS IN SEVERAL GENOMIC LOCI



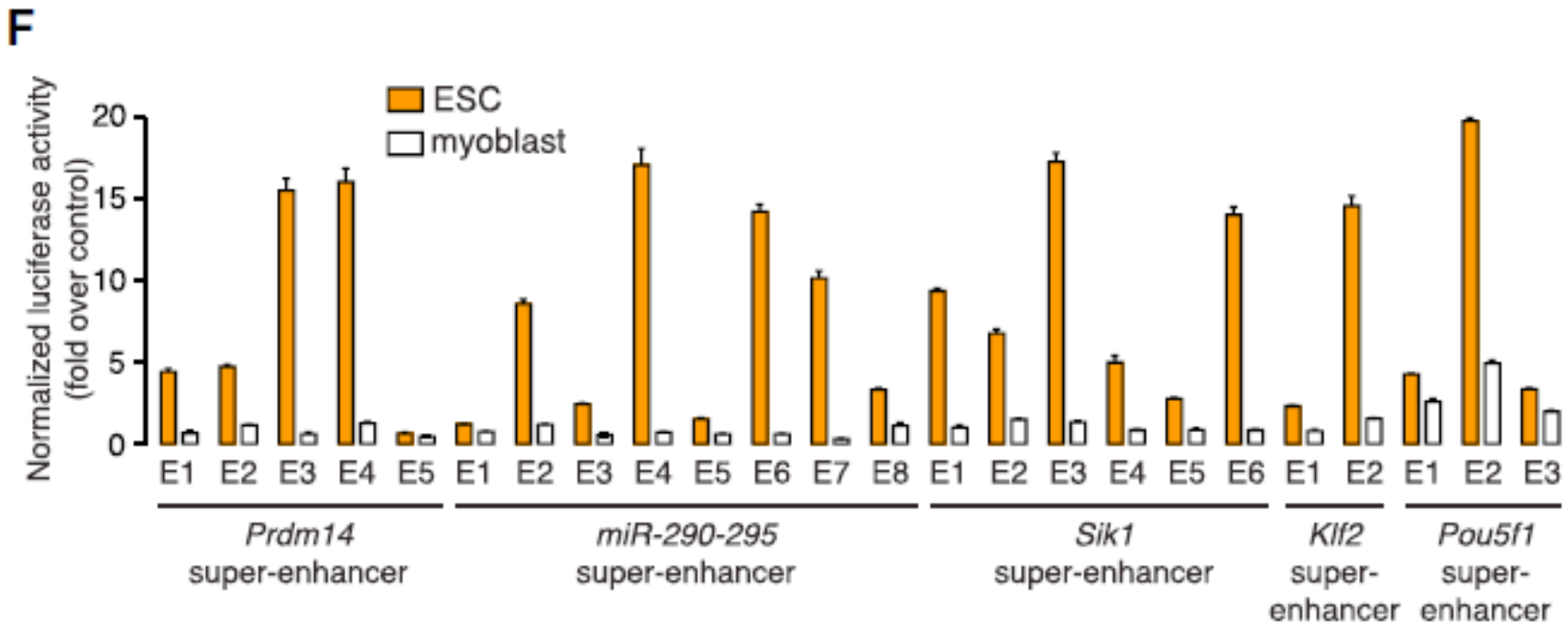
ACTIVITY OF SUPER-ENHANCER CONSTITUENTS



ChIP-Seq profile

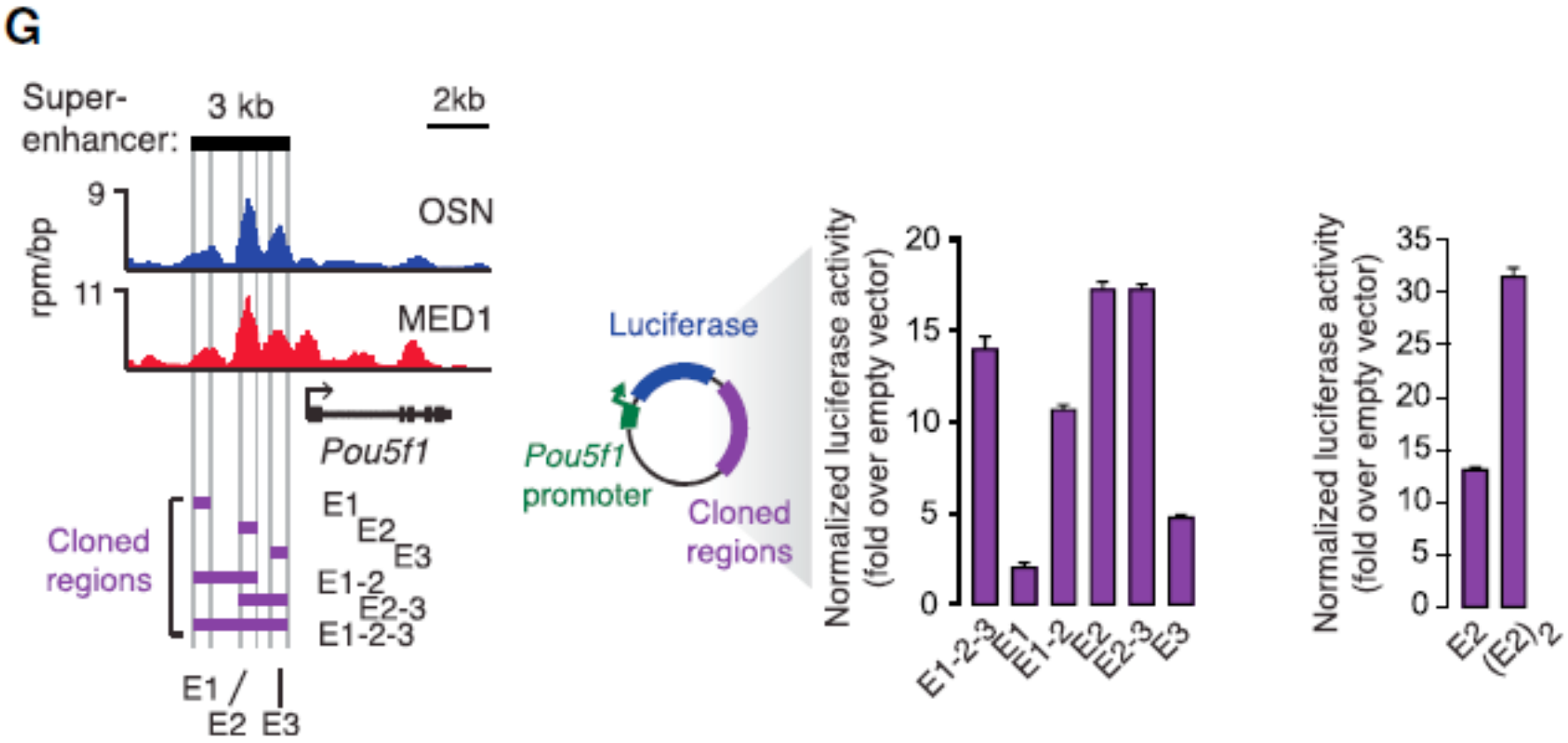


ENHANCER ACTIVITY OF SE CONSTITUENTS IS SPECIFIC FOR ESC, COMPARED TO MYOBLAST

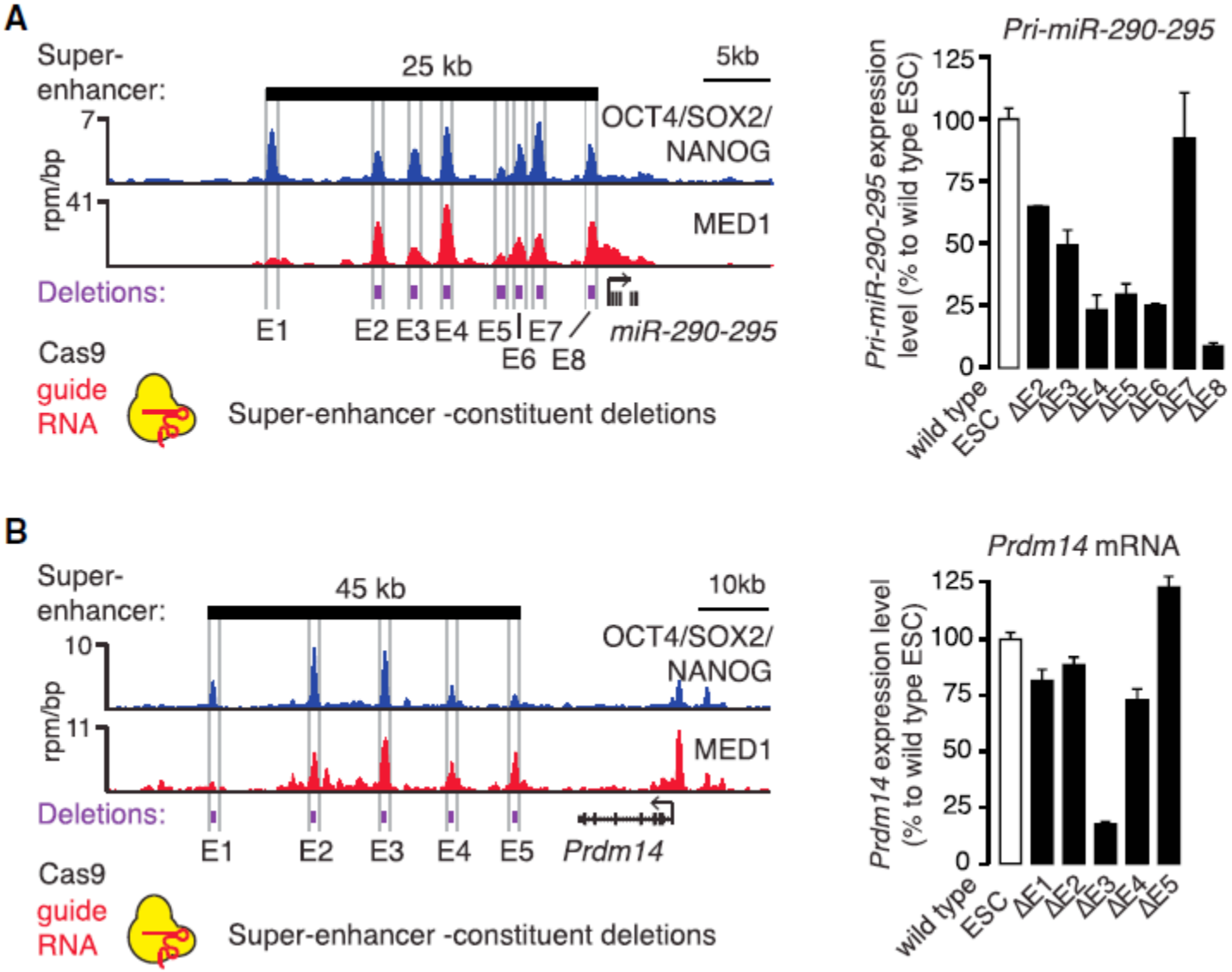


QUESTION: DO "SE CONSTITUENTS" ACT ADDITIVELY, SINERGISTICALLY OR EXERT A COMPLEX INFLUENCE?

E2 has high activity, E1 and E3 influence E2 activity



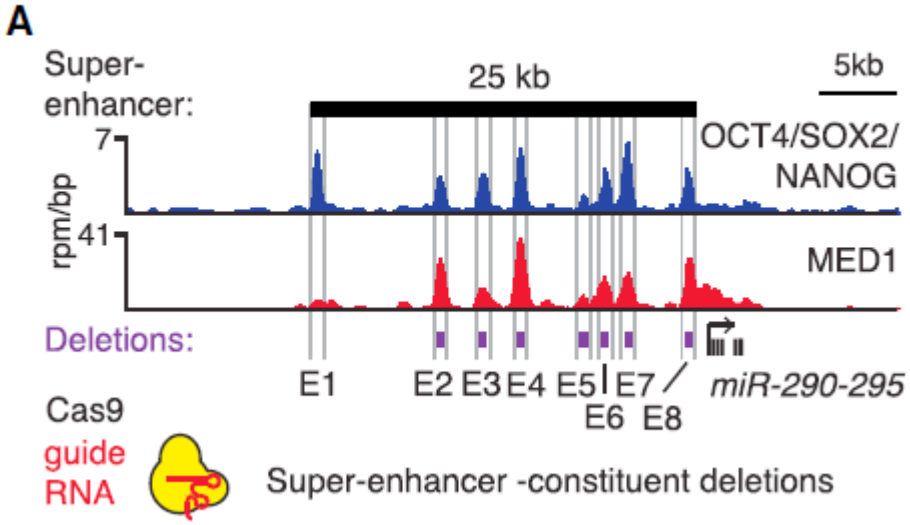
CONTRIBUTIONS OF SUPER-ENHANCER COSTITUENTS TO GENE EXPRESSION IN VIVO



QUESTION: HOW DOES SUPER-ENHANCER COSTITUENTS REGULATE GENE EXPRESSION IN VIVO?

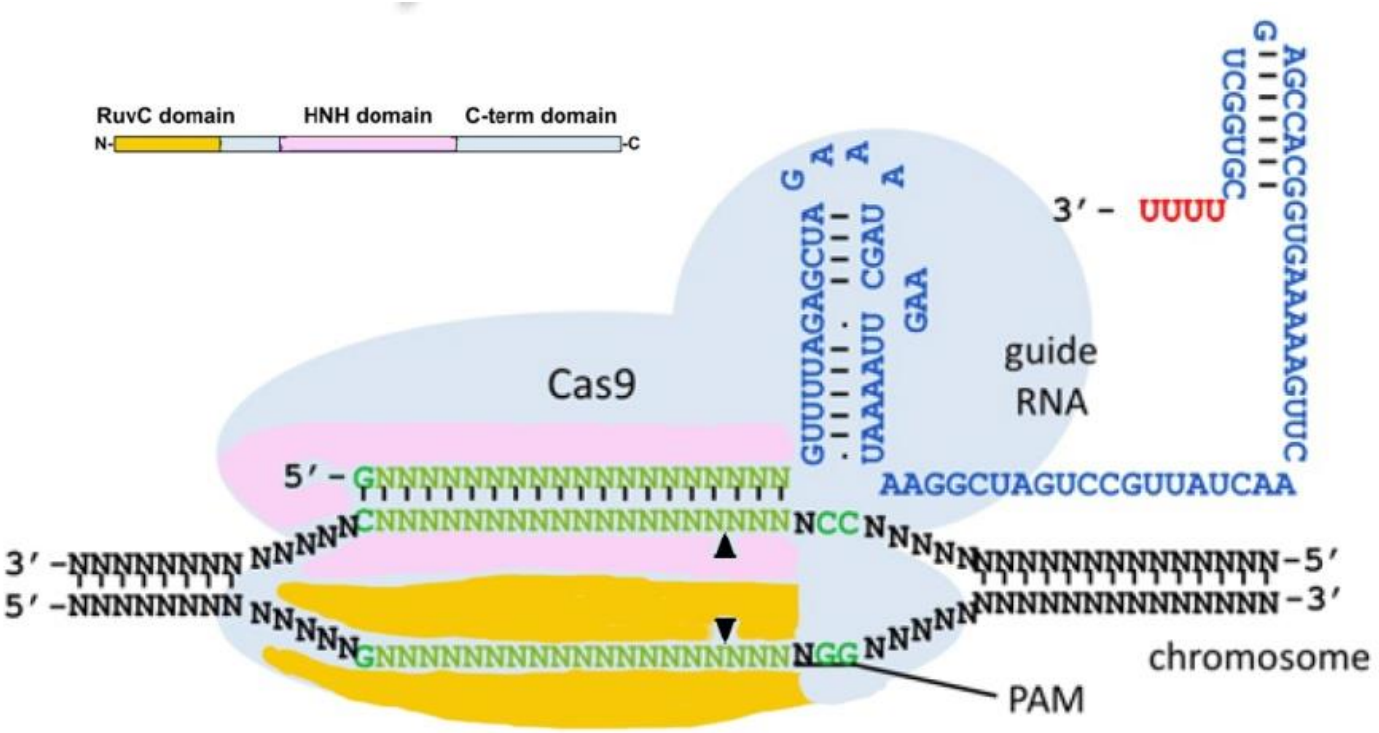
METHOD: DELETION OF SPECIFIC GENOMIC REGIONS

TECHNIQUE: CRISPR/CAS9



ChIP-Seq DATA used to design STUDY ON COSTITUENTS ENHANCERS FUNCTION

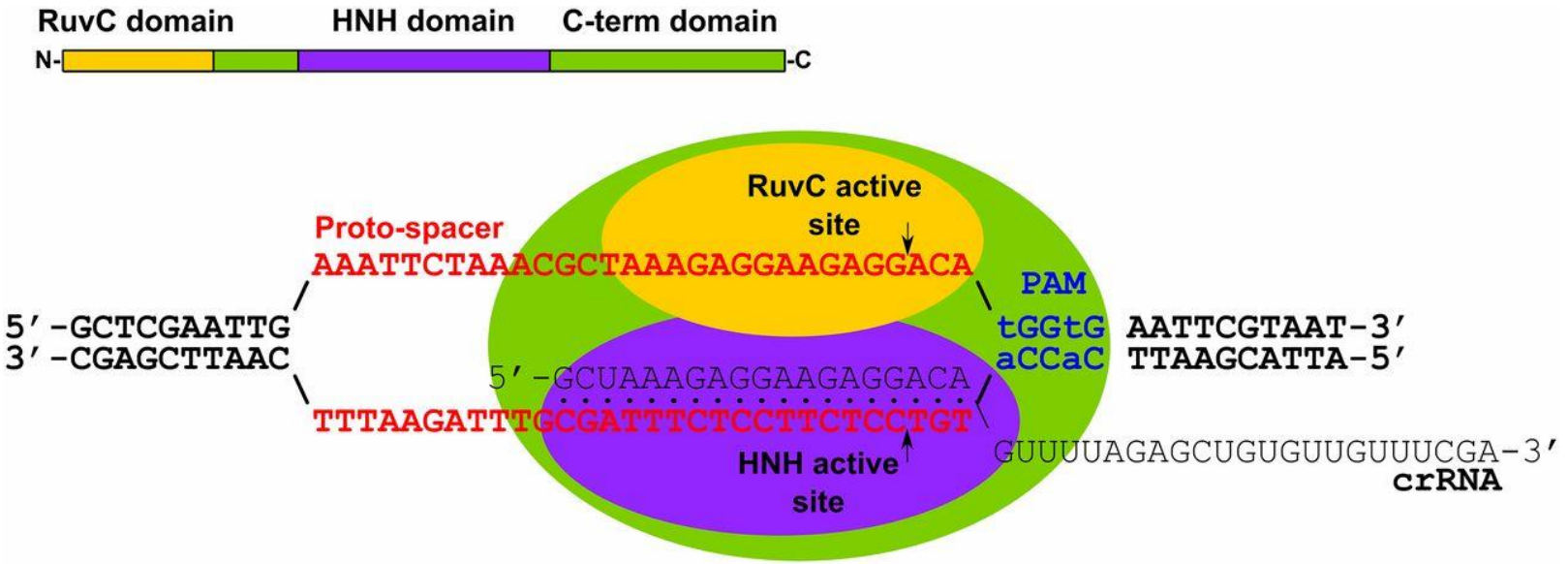
In type II **CRISPR** (**clustered regularly interspaced short palindromic repeats**)–Cas (CRISPR-associated) systems, the **endonuclease Cas9** associates with a dual-RNAguide structure consisting of a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA) to cleave double-stranded DNA (dsDNA) using its HNH and RuvC nuclease Domains.



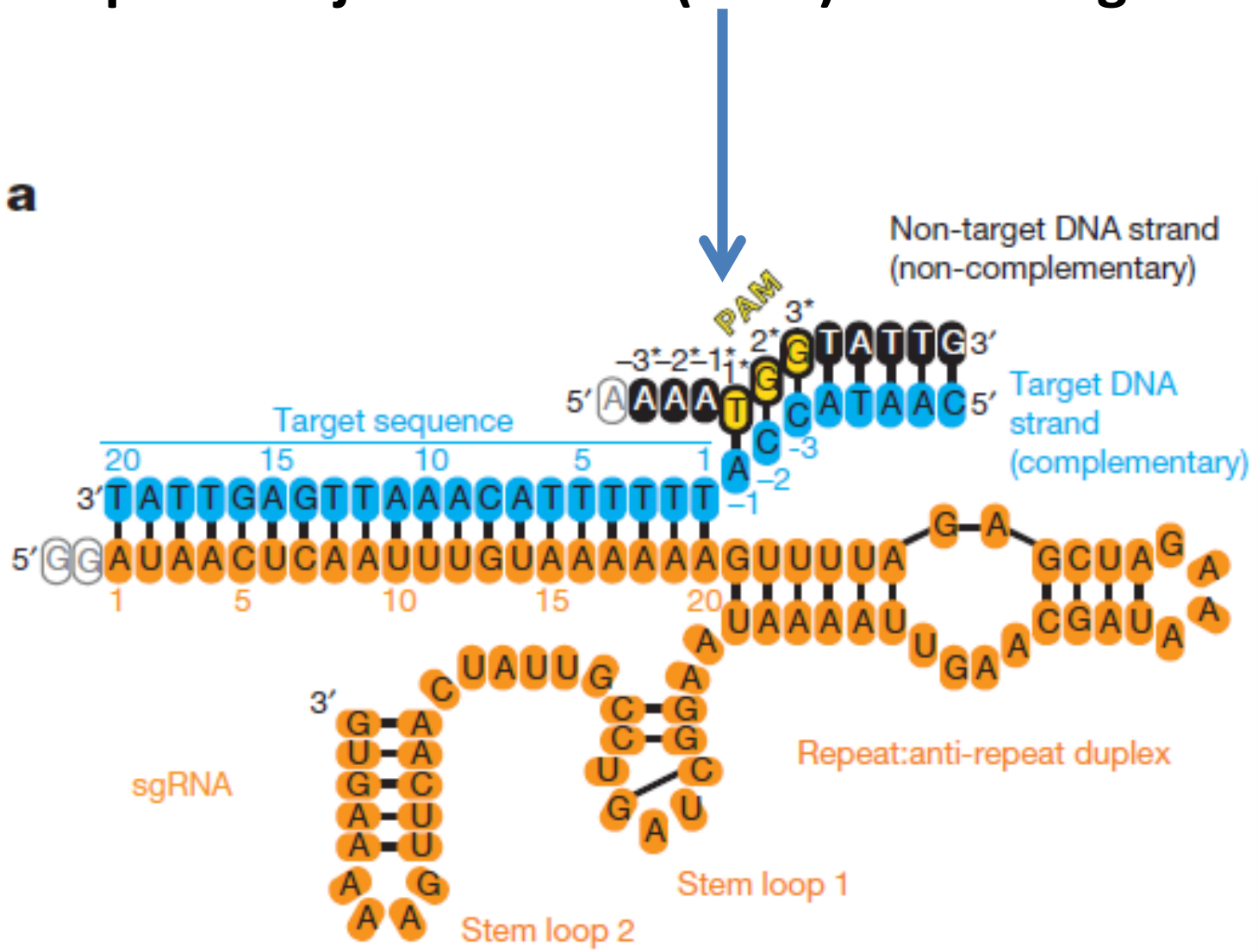
CRISP-Cas9

https://youtu.be/4YKFw2KZA5o

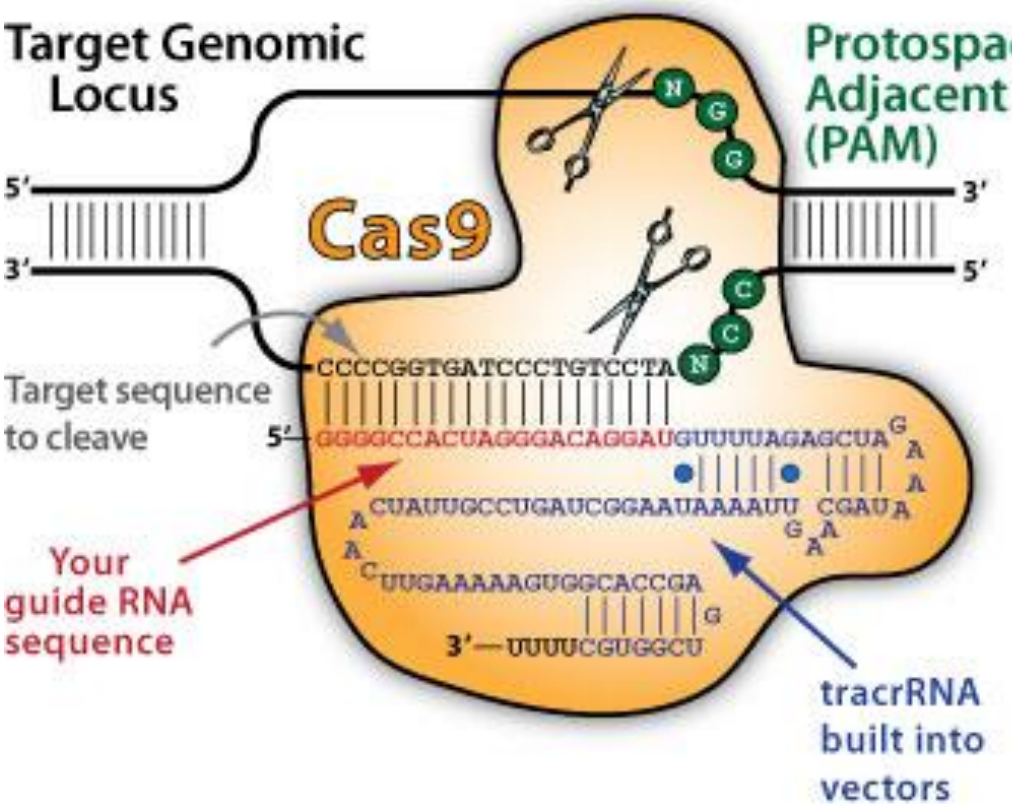
The **Cas9 nuclease** encloses the guide RNA and binds the structure with specific interactions in a number of domains. Specifically, the **REC1 and REC2 domains bind the complementary region of the guide RNA**, and eventually the **guide RNA target DNA heteroduplex upon DNA binding**. Mutations to the REC2 domain causes a small decrease in Cas9 activity, while mutations in the REC1 domain eliminate activity completely. The Rec1 domain is likely essential for Cas9 activity because it binds the repeat/anti-repeat duplex. The **Protospacer Adjacent Motif (PAM) Interacting (PI) domain** and **RuvC nuclease domain bind the stem loops on the guide RNA**. Overall, the exact process of guide RNA binding by the nuclease is unknown. However, a dramatic conformational change upon binding has been shown to produce an activated Cas9 nuclease capable of binding and cleaving specific sequences of double-stranded DNA

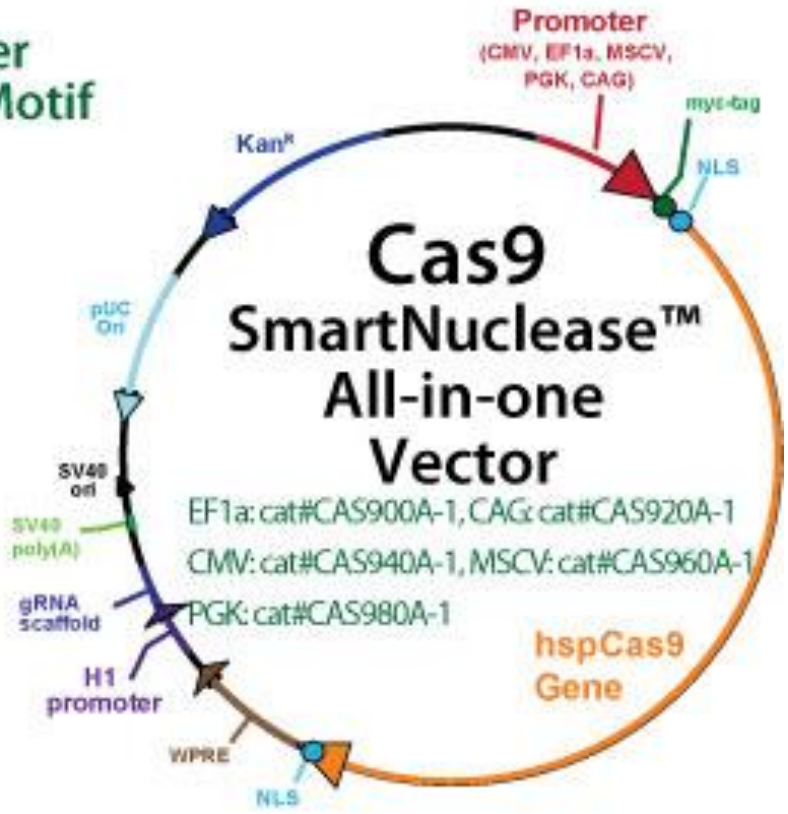


RNA-guided DNA recognition and cleavage strictly require the presence of a **protospacer adjacent motif (PAM) in the target DNA**.



CONTRIBUTIONS OF SUPER-ENHANCER COSTITUENTS TO GENE EXPRESSION IN VIVO



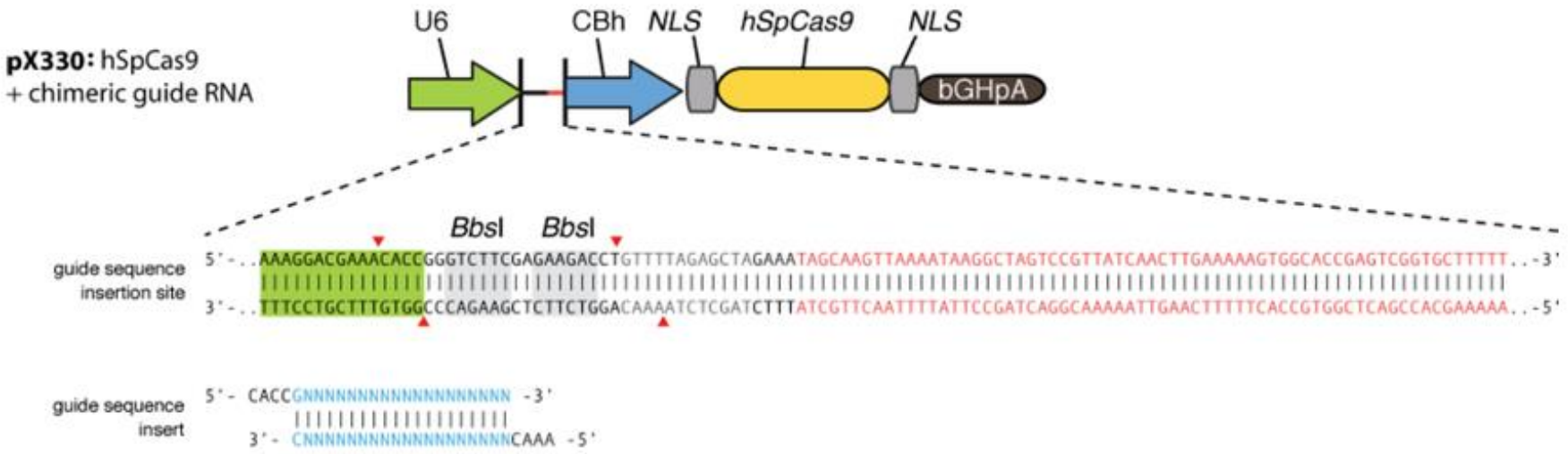


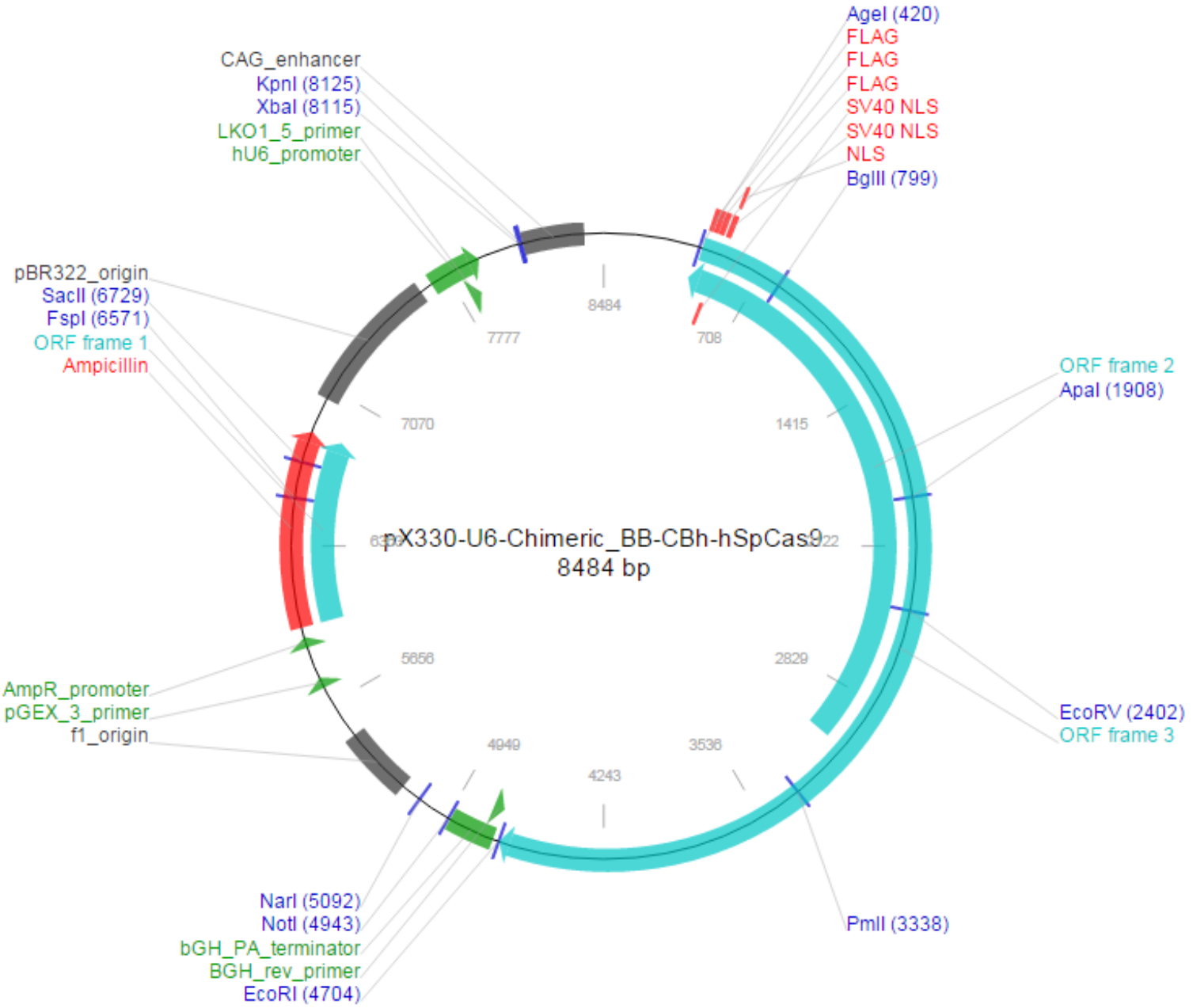
Genome Editing

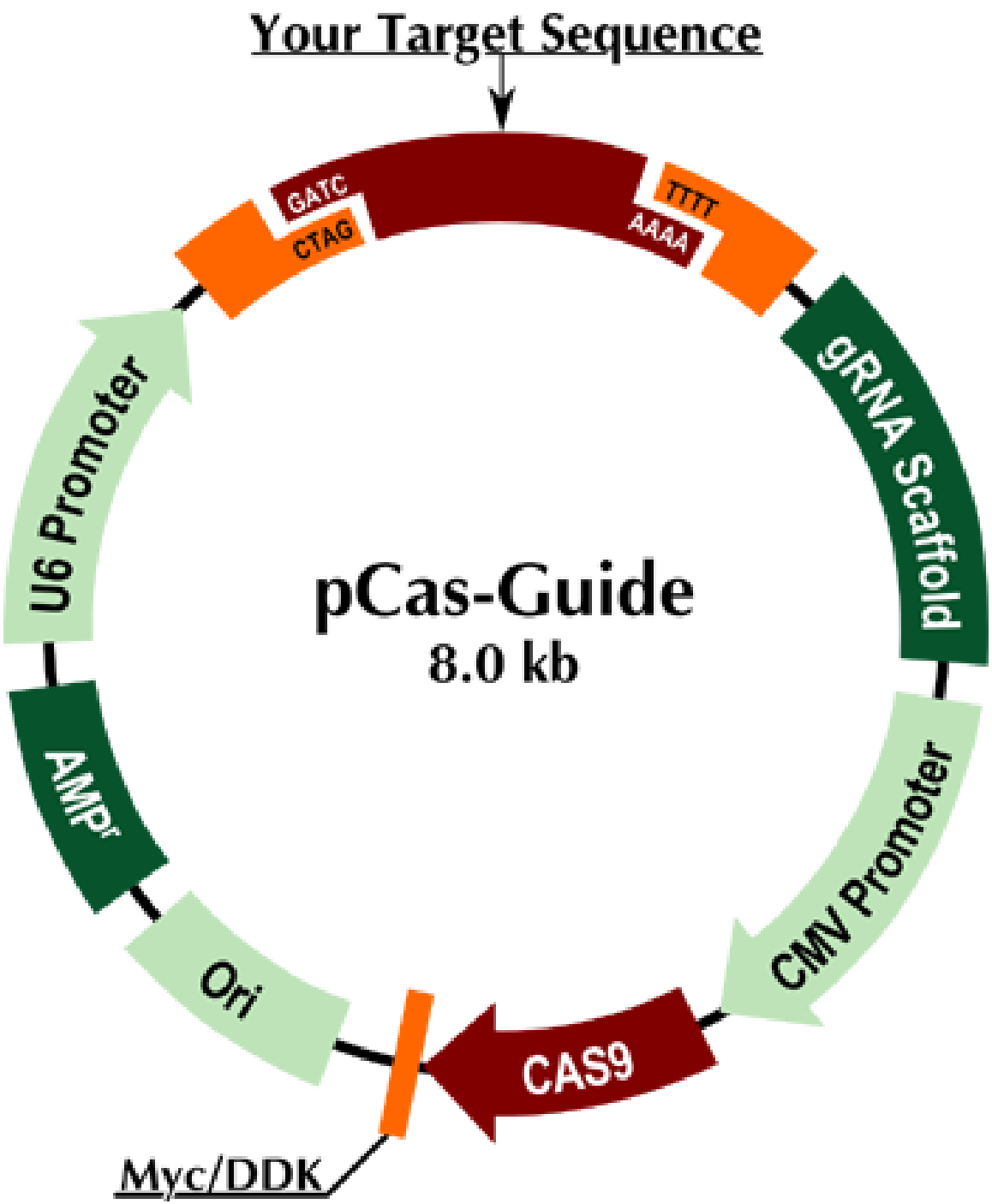
Super-enhancer constituent enhancers (~400 bp) were deleted in V6.5 murine ESCs using the CRISPR/Cas9 system. sgRNAs were cloned into the pX330 vector (Addgene: 42230) containing Cas9. Cells were transfected with two plasmids expressing Cas9 and an sgRNA complementary to each end of the targeted super-enhancer constituent using X-fect reagent (Clontech). A plasmid expressing PGK-puroR was co-transfected for selection. 1 day after transfection, cells were re-plated on DR4 MEF feeder layers. 1 day after re-plating, puromycin (2 μg/ml) was added for 3 days. Subsequently, puromycin was withdrawn for 3–4 days. Individual colonies were picked and genotyped by PCR. Deletion alleles were verified by sequencing.

Map Image for pX330-U6-Chimeric_BB-CBh-hSpCas9

Map uploaded by the depositor.

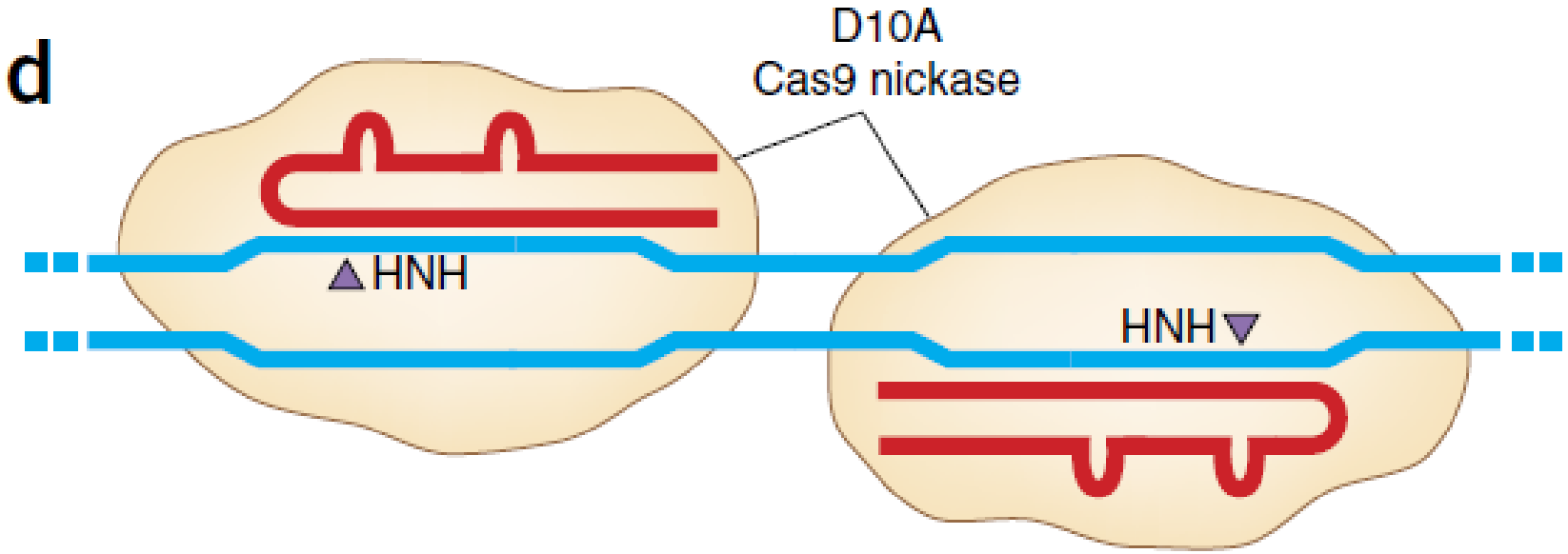


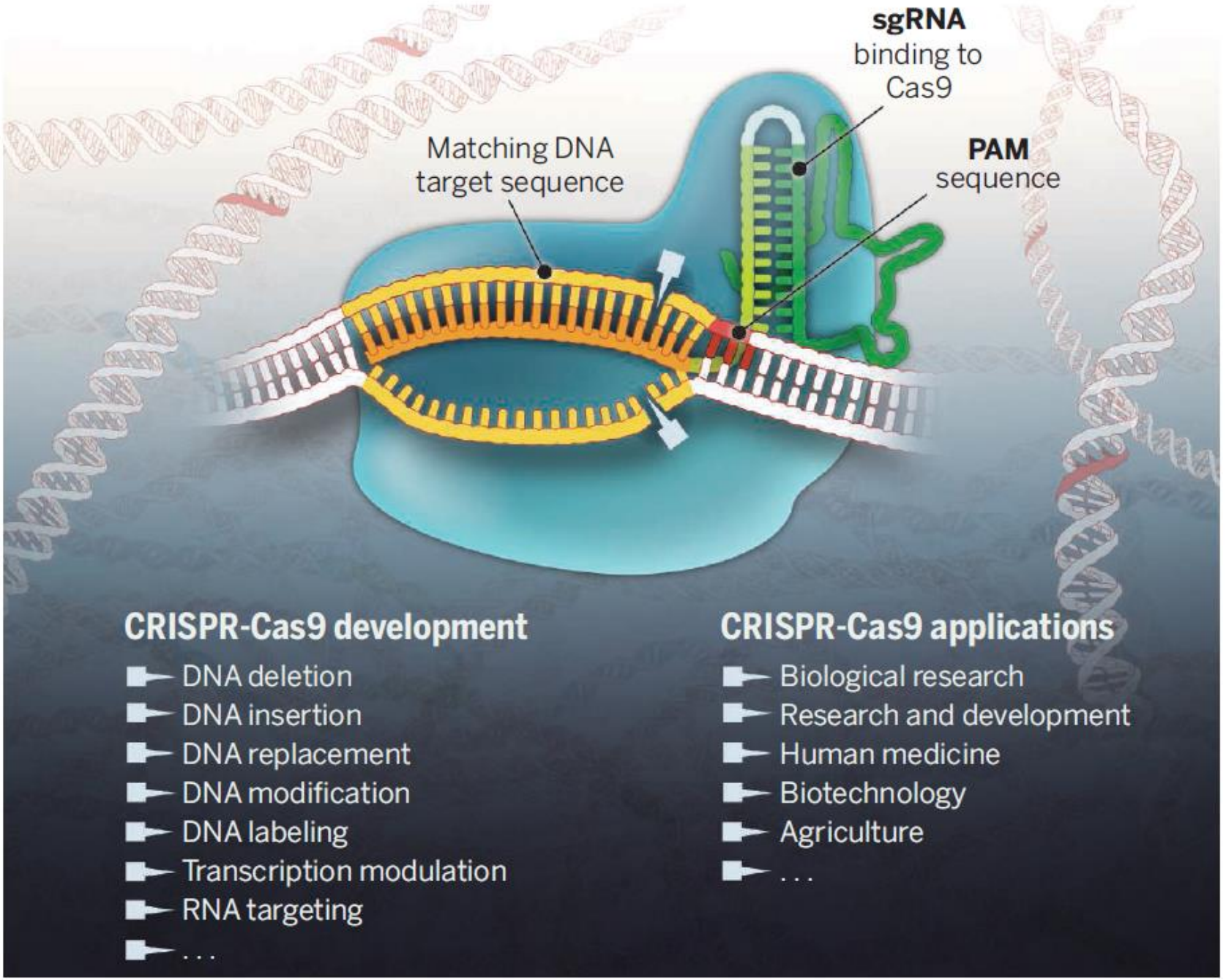

AgeI (420)
FLAG
FLAG
FLAG
SV40 NLS
SV40 NLS
NLS
BglII (799)
CAG_enhancer
KpnI (8125)
XbaI (8115)
LKO1_5_primer
hU6_promoter
pBR322_origin
SacII (6729)
FspI (6571)
ORF frame 1
Ampicillin
ORF frame 2
ApaI (1908)
AmpR_promoter
pGEX_3_primer
f1_origin
EcoRV (2402)
ORF frame 3
NarI (5092)
NotI (4943)
bGH_PA_terminator
BGH_rev_primer
EcoRI (4704)
PmlI (3338)
pX330-U6-Chimeric_BB-CBh-hSpCas9
8484 bp
8484
708
1415
2829
3536
4243
4949
5656
7070
7777


Your Target Sequence
GATC
CTAG
TTTT
AAAA
gRNA Scaffold
CMV Promoter
CAS9
Myc/DDK
Ori
AMPr
U6 Promoter
pCas-Guide
8.0 kb

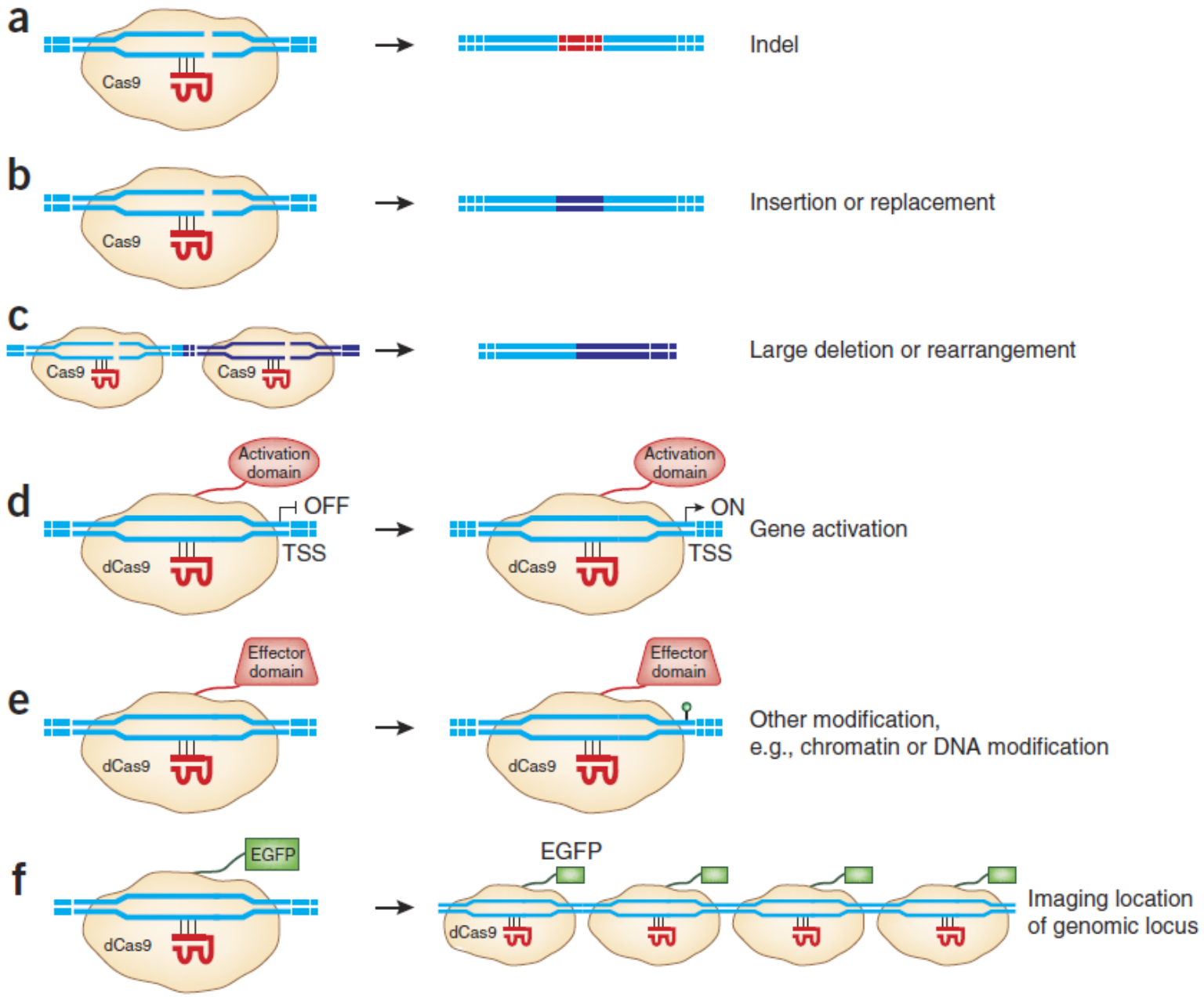
Genome editing

Genome editing was performed using CRISPR/Cas9 essentially as described (Wang et al., 2013). Briefly, target-specific oligonucleotides were cloned into a plasmid carrying a codon-optimized version of Cas9 (pX330, Addgene: 42230). The genomic sequences complementary to guide RNAs are listed in Supplemental Table 2. V6.5 murine ESCs were transfected with two plasmids expressing Cas9 and sgRNA targeting regions around 200 basepairs up- and down-stream of the center of the targeted SE-constituent (as defined by OCT4/SOX2/NANOG co-binding; see below), respectively. A plasmid expressing PGK-puroR was also co-transfected. Transfection was carried out with the X-fect reagent (Clontech) according to the manufacturer's instructions. One day after transfection, cells were re-plated on DR4 MEF feeder layers. One day after re-plating puromycin (2µg/ml) was added for three days. Subsequently, puromycin was withdrawn for three to four days. Individual colonies were picked, and genotyped by PCR, and the edited alleles were verified by Sanger sequencing. All cell lines used in subsequent experiments were homozygous deletion lines. Reference and deletion allele sequences are listed in Supplemental Table 2.


d
D10A
Cas9 nickase
HNH
HNH


sgRNA
binding to
Cas9
Matching DNA
target sequence
PAM
sequence
CRISPR-Cas9 development
DNA deletion
DNA insertion
DNA replacement
DNA modification
DNA labeling
Transcription modulation
RNA targeting
. . .
CRISPR-Cas9 applications
Biological research
Research and development
Human medicine
Biotechnology
Agriculture
. . .

OVERVIEW CRISPR-CAS SYSTEM APPLICATION

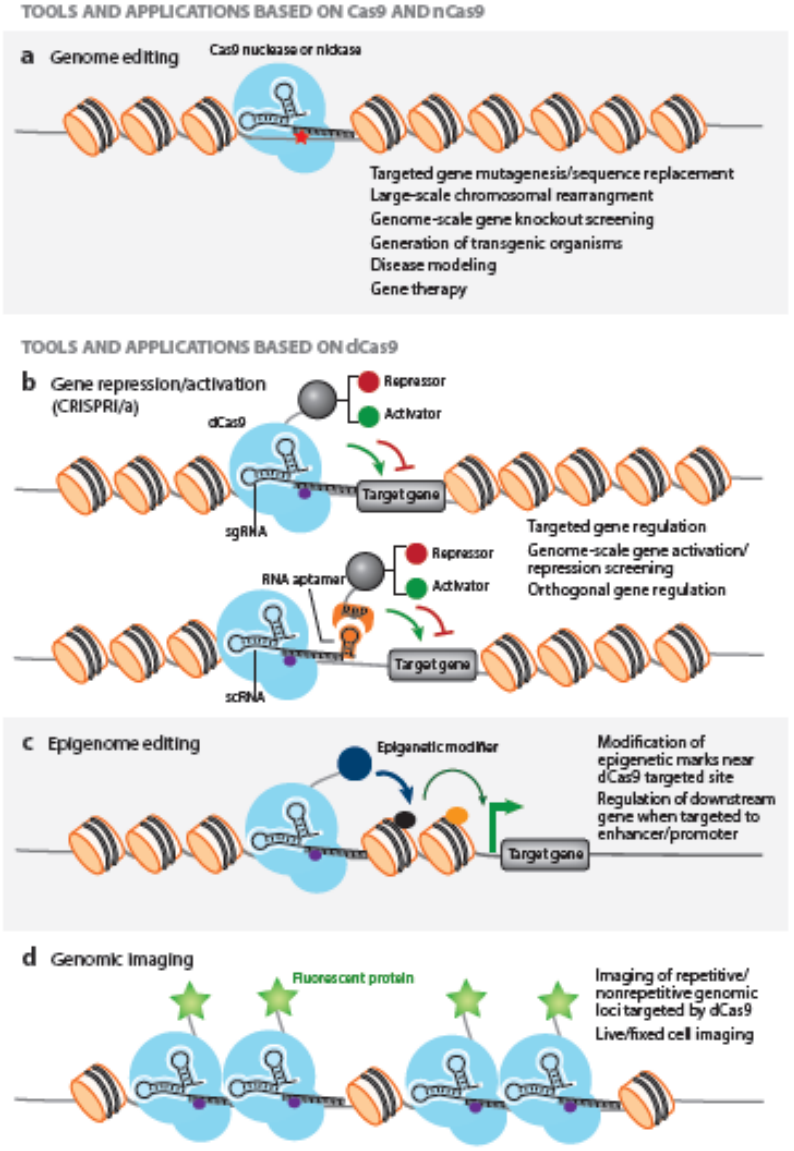


What is CRISPR, its scientific and ethical impact?

https://youtu.be/TdBAHexVYzc

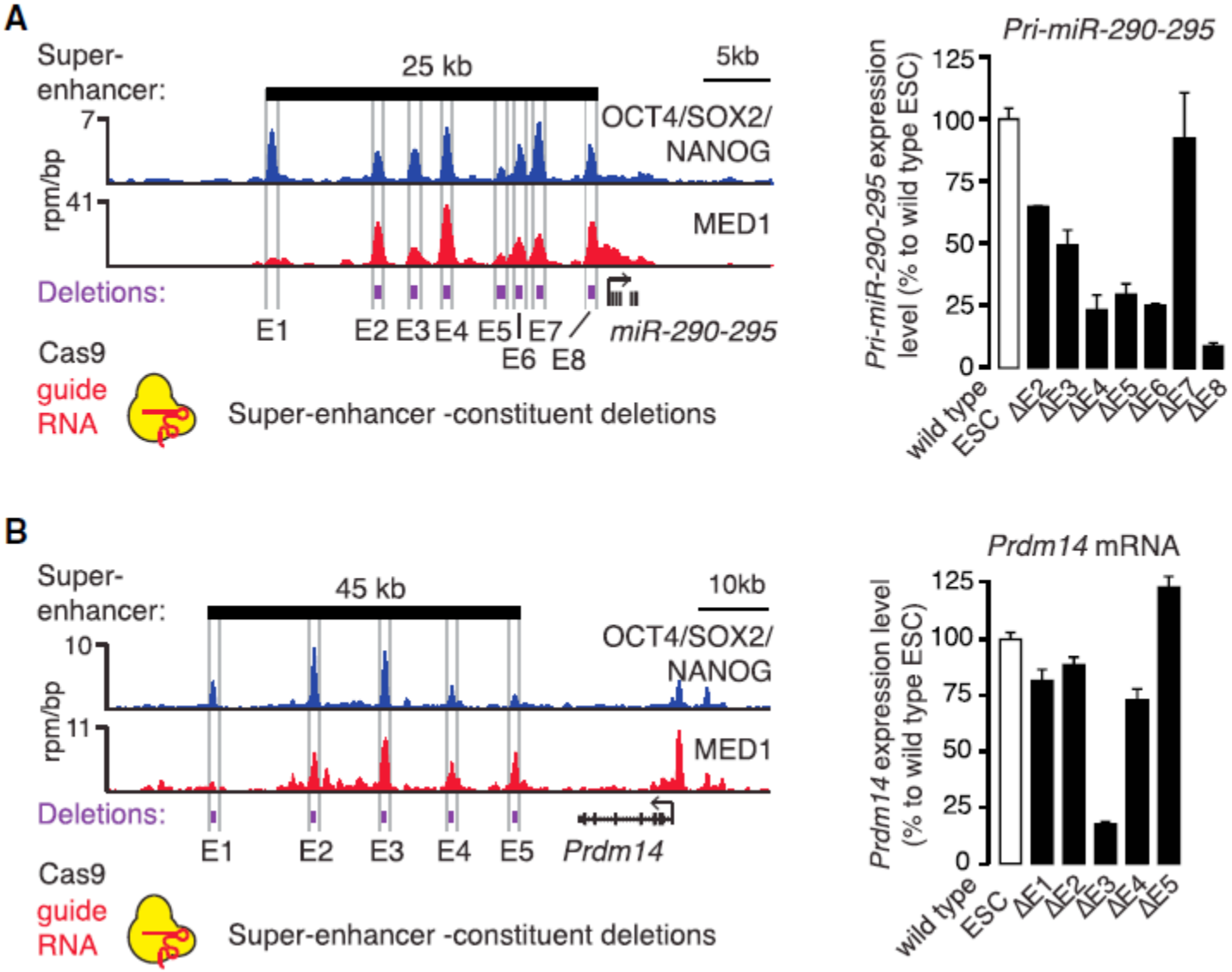
CRISPR/Cas9 in Genome Editing and Beyond

Haifeng Wang,[1] Marie La Russa,[1,2] and Lei S. Qi[1,3,4]

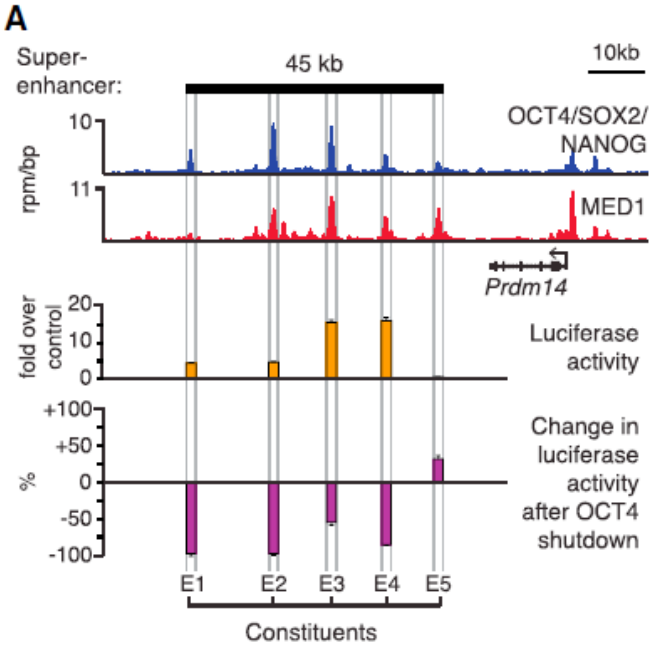


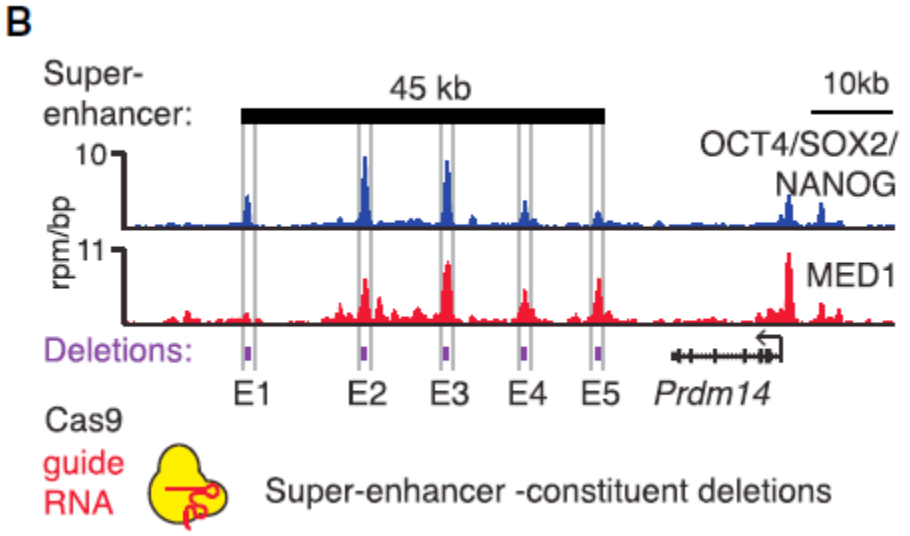
Annu. Rev. Biochem. 2016. 85:22.1–22.38

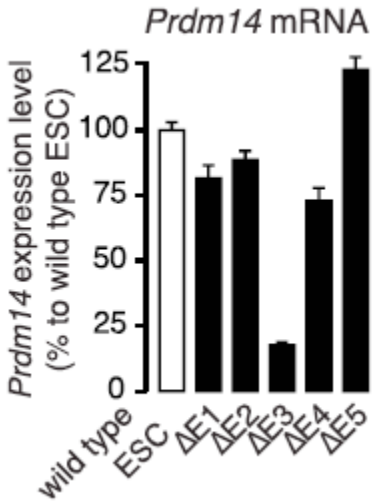
CONTRIBUTIONS OF SUPER-ENHANCER COSTITUENTS TO GENE EXPRESSION IN VIVO



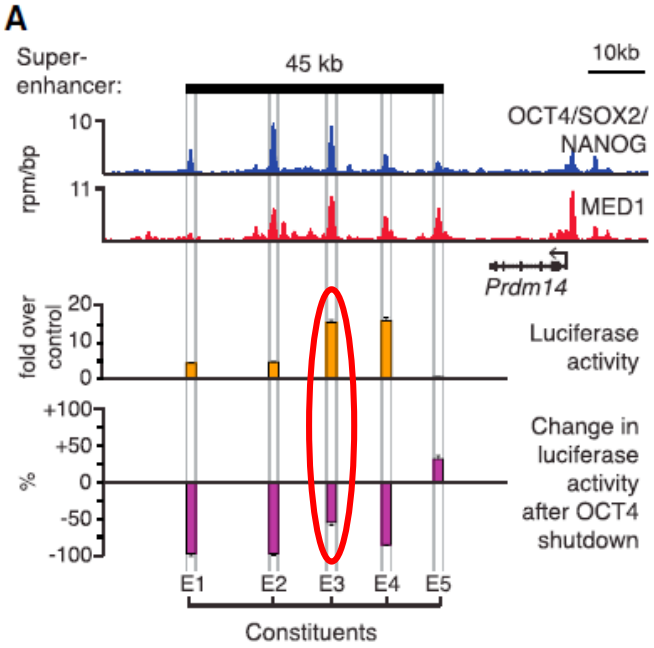
LUCIFERASE ACTIVITY AND DELETION OF SPECIFIC ENHANCER

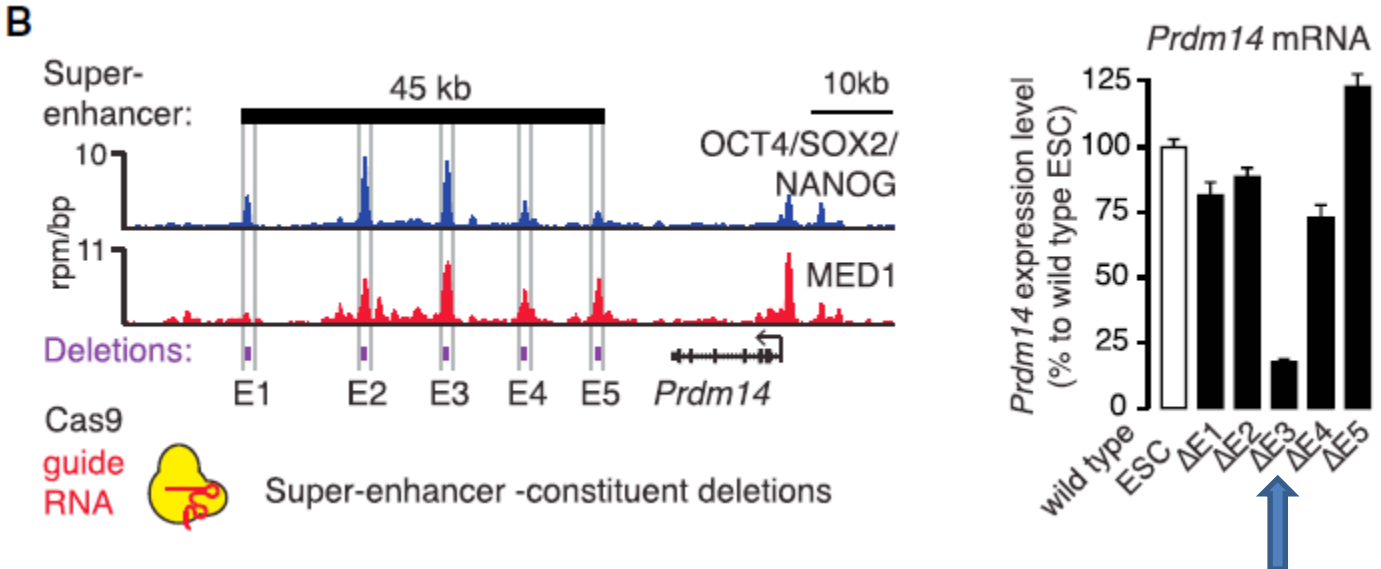




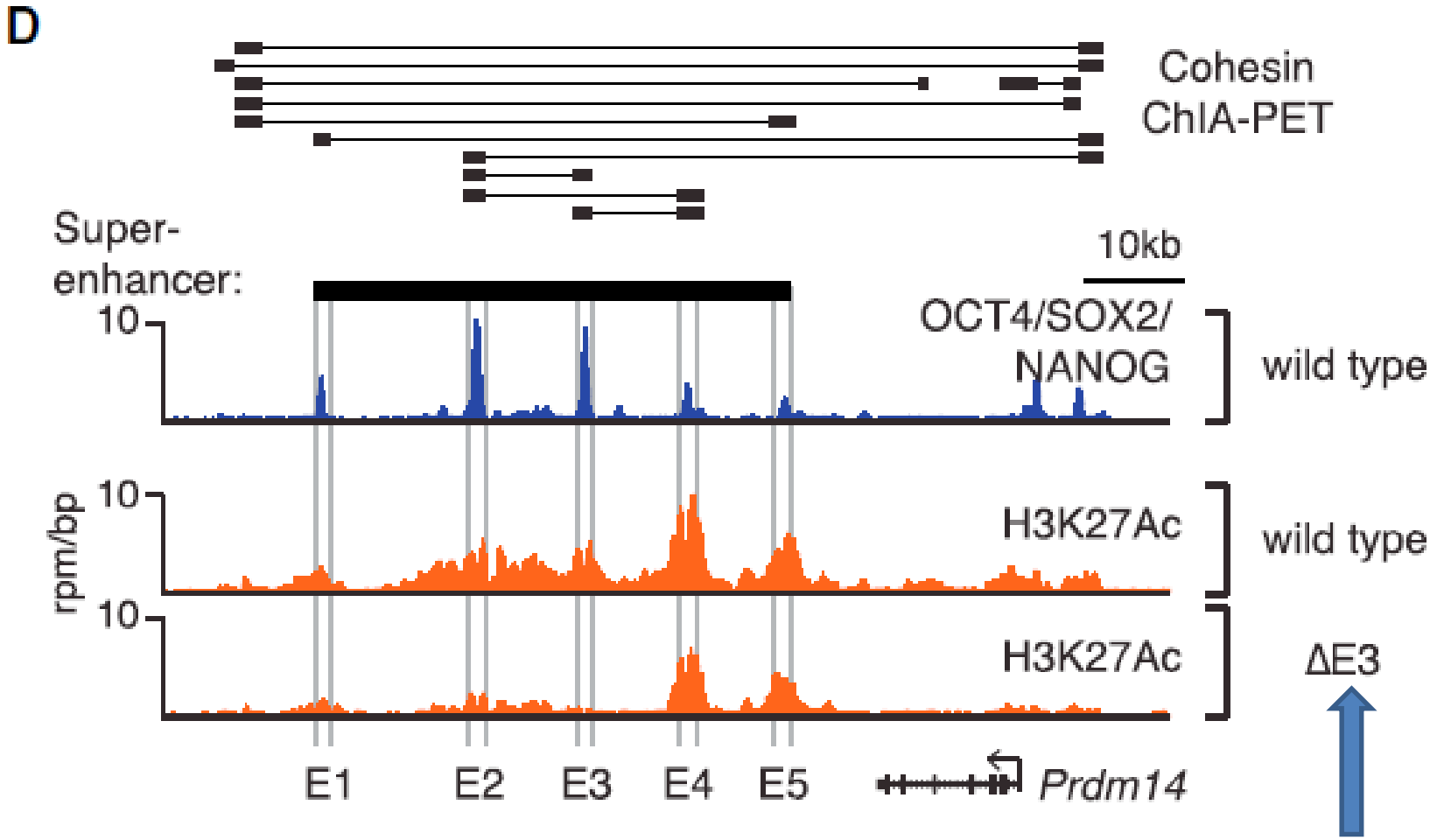


LUCIFERASE ACTIVITY AND DELETION OF SPECIFIC ENHANCER





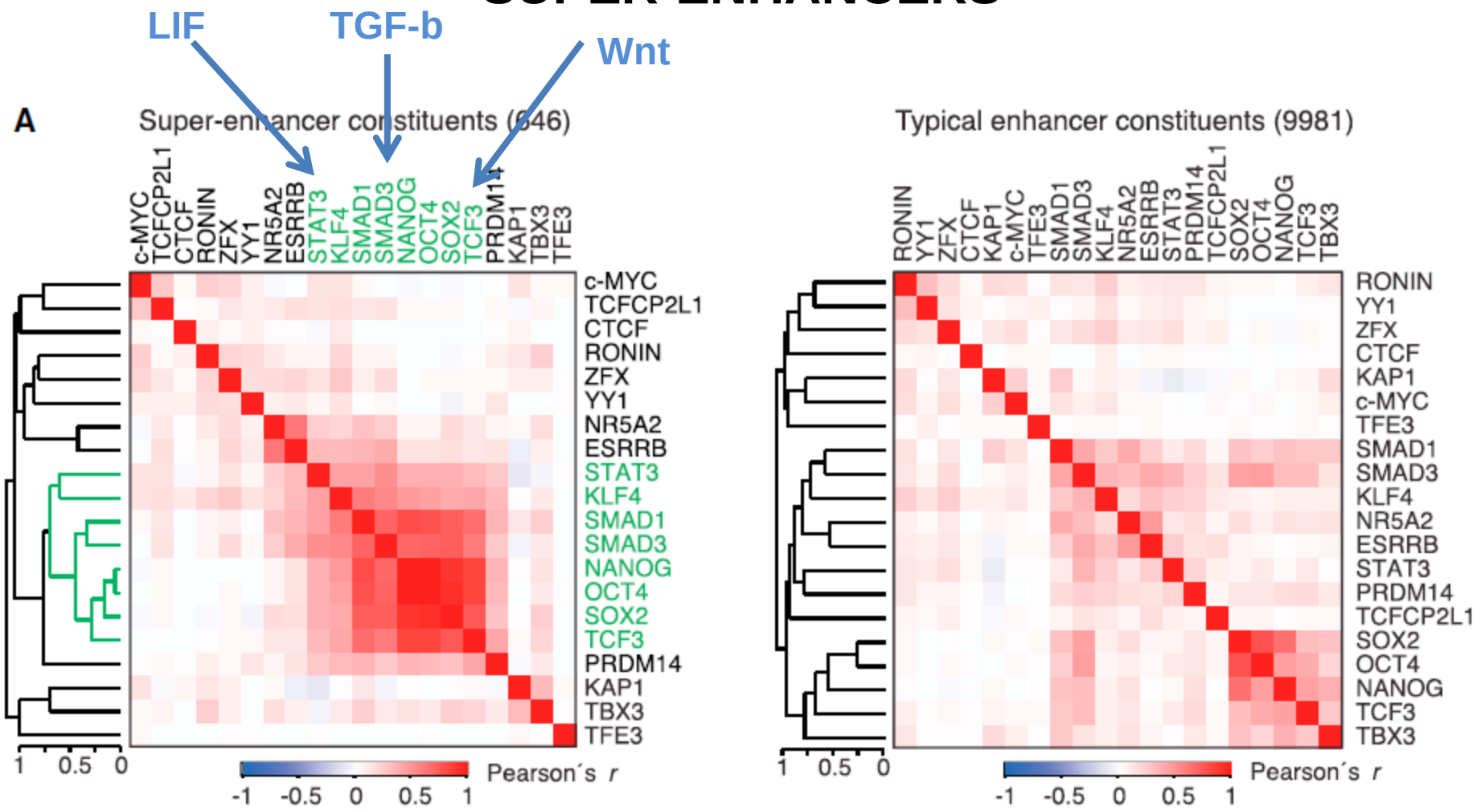
IDENTIFICATION OF INTERACTION BETWEEN SPECIFIC COSTITUENTS ENHANCERS



DELETION OF E3 INDUCES EPIGENETIC MODIFICATIONS: H3K27Ac reduction

SUPER-ENHANCERS ENRICHMENT OF TRANSCRIPTION FACTORS PATTERN

SPECIFIC MULTIPLE TRANSCRIPTION FACTORS BIND SUPER-ENHANCERS

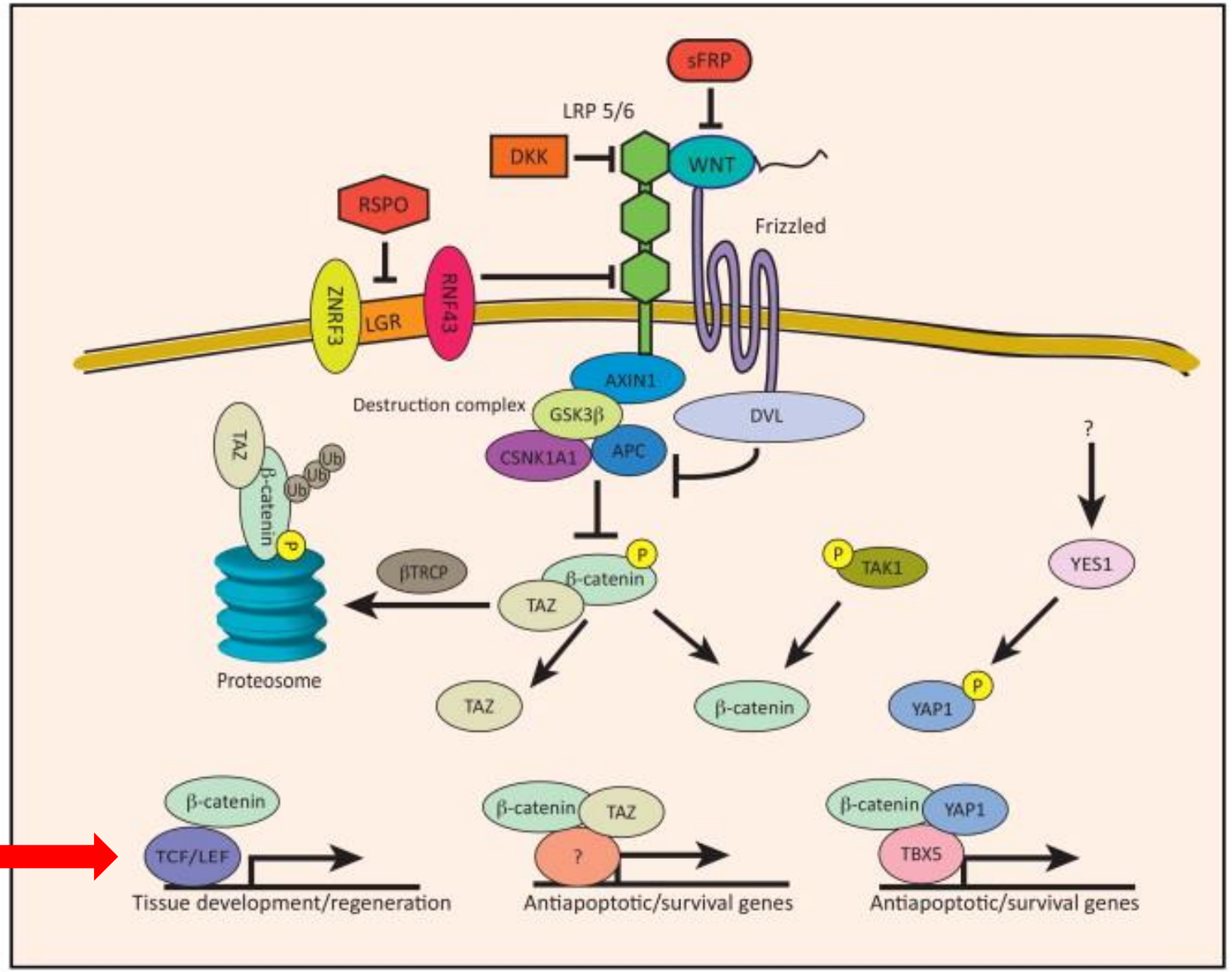


Hierarchical clustering of 20 transcription factor ChIP-seq binding profiles at super-enhancer and typical enhancer constituents. A set of factors with binding profiles similar to OCT4, SOX2, and NANOG is highlighted in green.

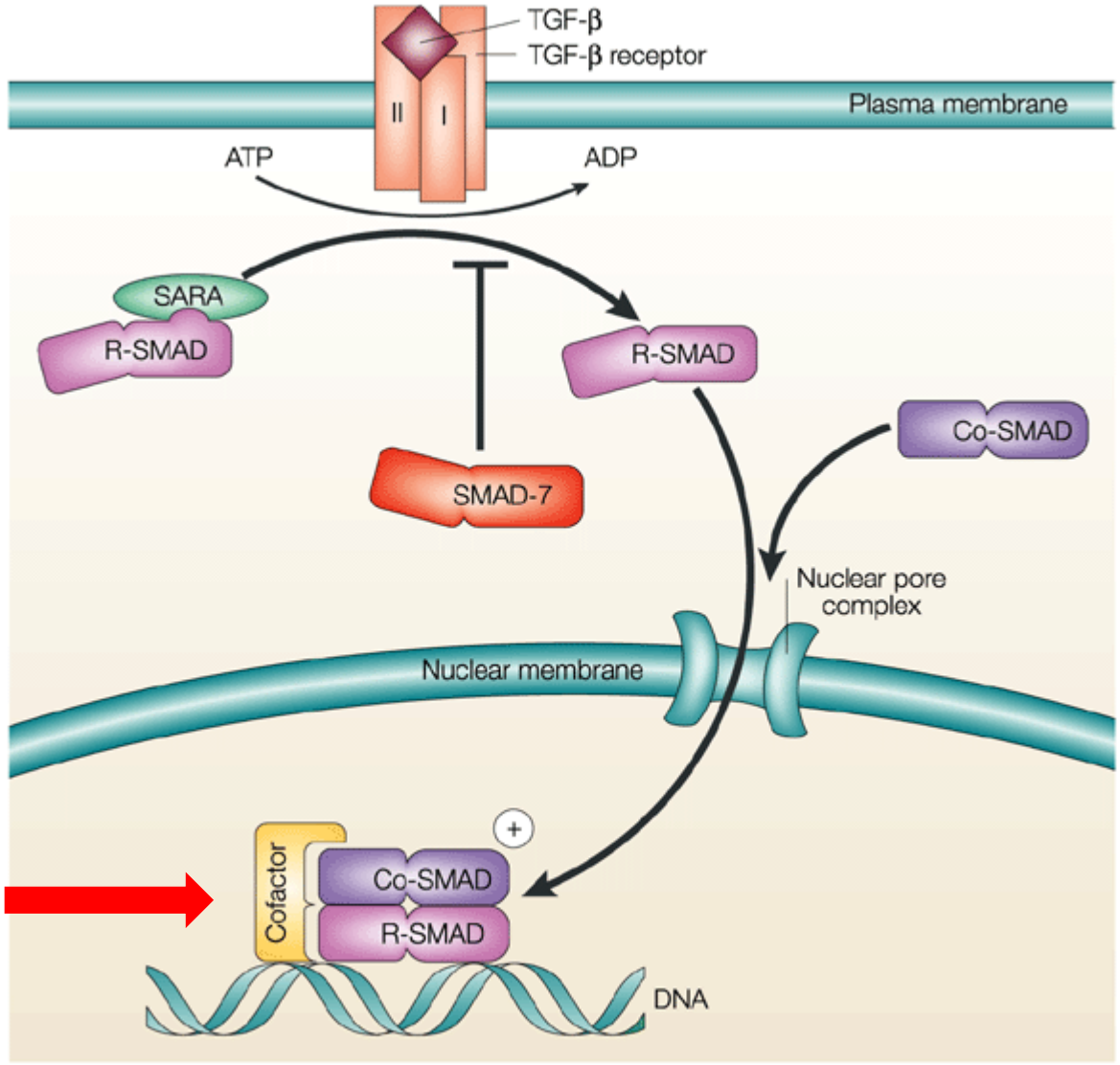
enhancers. An examination of the pattern of transcription factor binding to super-enhancer constituents provided a hypothesis to resolve this conundrum (Figure 3A, Table S3). The terminal TFs of the Wnt (TCF3), TGF-β (SMAD3), and LIF (STAT3) signaling pathways, which play essential roles in transcriptional control of the stem cell state (Ng and Surani, 2011; Young, 2011), were among the TFs whose binding pattern to SE constituents was most similar to that of OCT4, SOX2, and NANOG at SE constituents (Figure 3A). Most SE constituents (75%) were occupied by at least one of these three TFs, whereas only 43% of typical enhancer constituents were bound by one of the three (Figure S3A). More importantly, 98% of super-enhancers were bound by at least one, 86% were bound 46% were bound by all three signaling

1TF
2TFs
3TFs

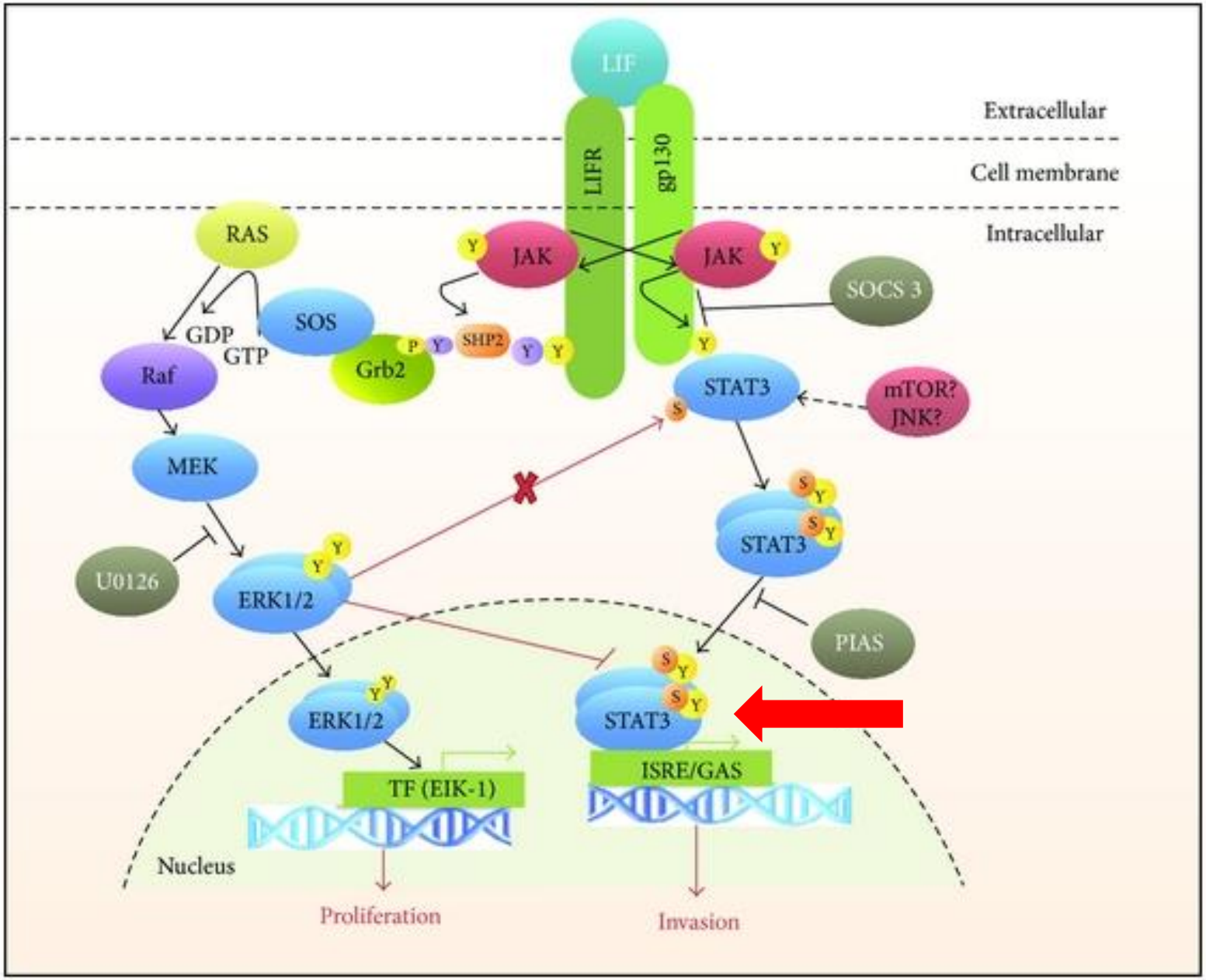
WNT PATHWAY


sFRP
LRP 5/6
DKK
WNT
RSPO
Frizzled
ZNRF3
LGR
RNF43
AXIN1
Destruction complex
GSK3β
DVL
CSNK1A1
APC
?
TAZ
β-catenin
Ub
P
βTRCP
TAK1
YES1
Proteosome
YAP1
TCF/LEF
TBX5
Tissue development/regeneration
Antiapoptotic/survival genes
Antiapoptotic/survival genes
TRENDS in Pharmacological Sciences

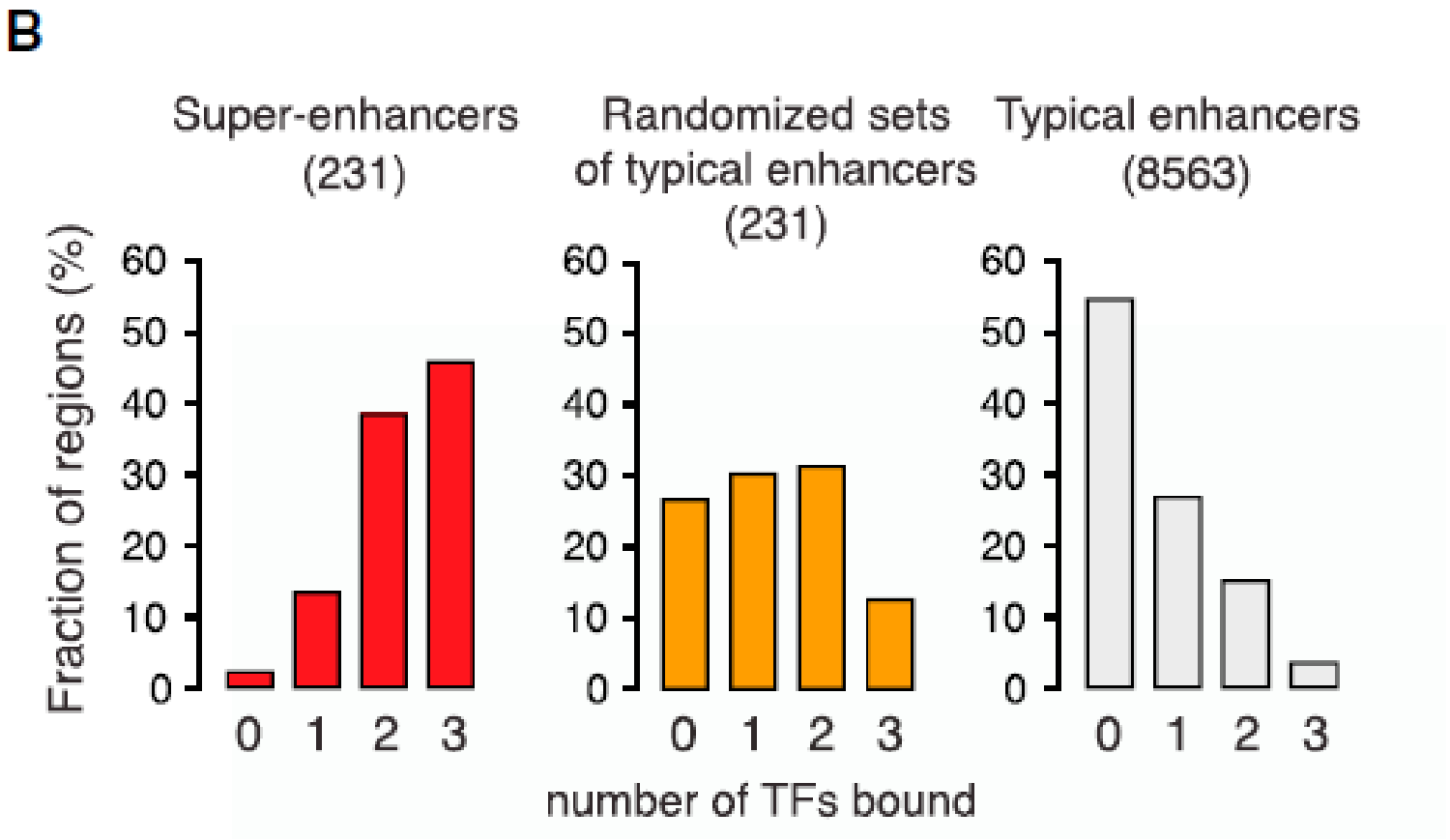
TGF beta PATHWAY



Nature Reviews | Immunology


LIF
Extracellular
Cell membrane
Intracellular
LIFR
gp130
RAS
JAK
JAK
SOCS 3
SOS
GDP
GTP
Grb2
SHP2
Raf
STAT3
mTOR?
JNK?
MEK
STAT3
U0126
ERK1/2
PIAS
ERK1/2
STAT3
TF (EIK-1)
ISRE/GAS
Nucleus
Proliferation
Invasion

TCF3, SMAD3, STAT3, regulated by oncogenic pathways, bind constituent enhancers in SE.
No same pattern in randomized set of typical enhancers.



Binding motifs for TCF3, SMAD3, and STAT3 and the p values for their enrichment in super-enhancer constituent enhancers in murine and human ESCs. The motif of CTCF is not found enriched and serves as a negative control.

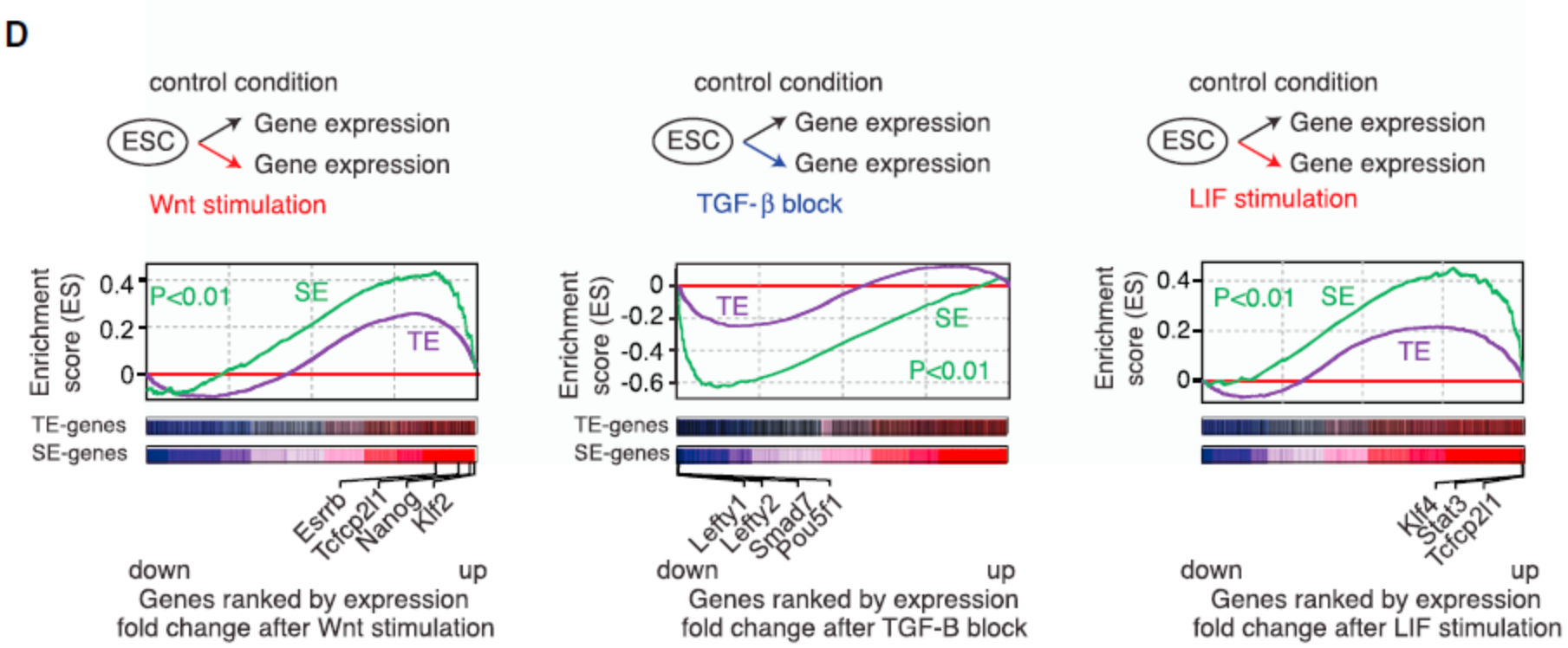
c

TF	Motif	murine ESC P-value	human ESC P-value
TCF3		$5.46*10^{-27}$	$2.23*10^{-28}$
SMAD3		$9.31*10^{-11}$	$3.34*10^{-4}$
STAT3		$2.90*10^{-10}$	$6.26*10^{-2}$
CTCF		0.45	1

Gene expression analyses

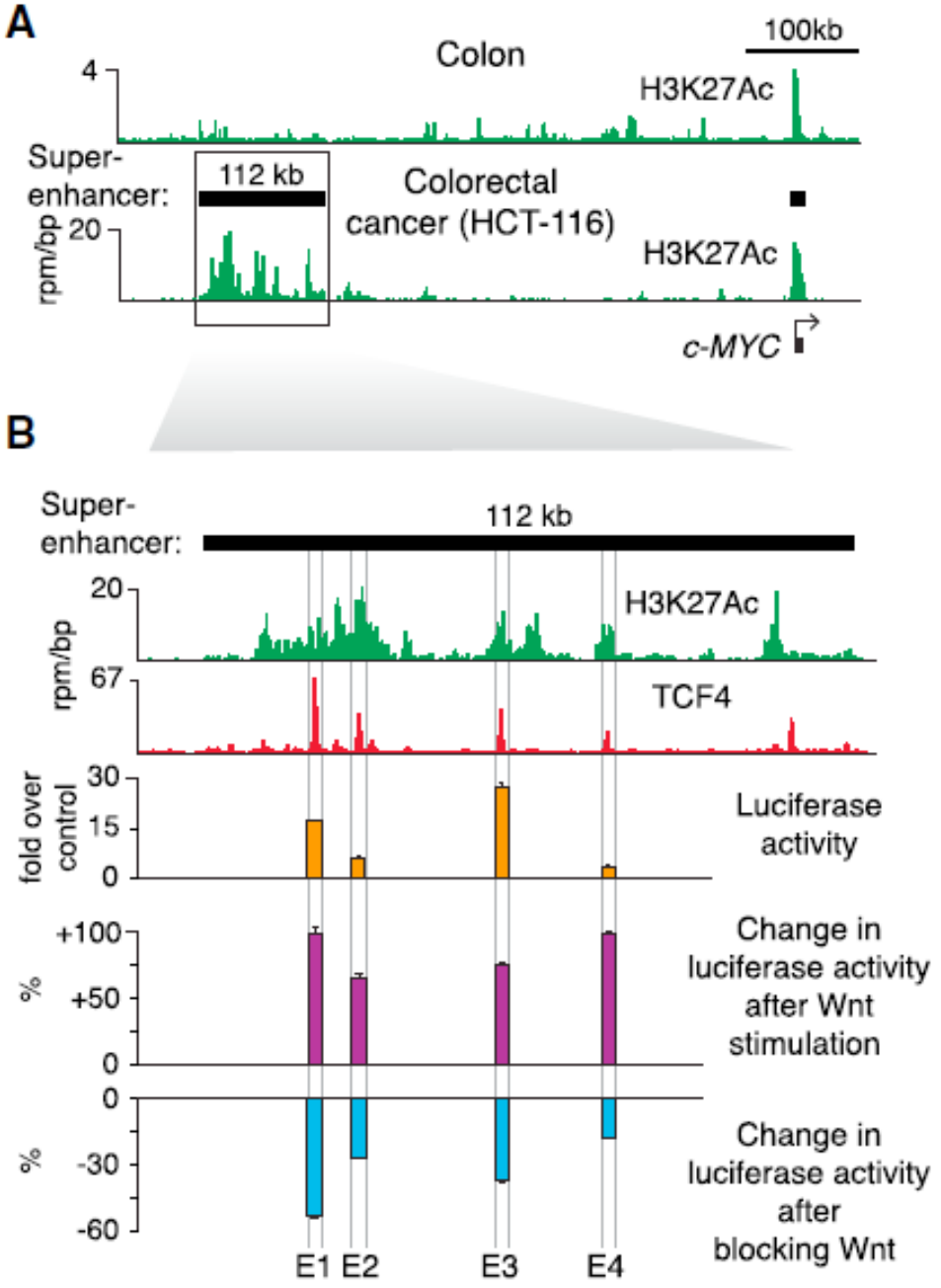
RNA-Seq RPKMs were calculated for two replicates each of murine ESCs treated with LIF for 1h (E-MTAB-1796 Arrayexpress dataset (Martello et al., 2013)) . Reads for each replicate were aligned to the mm9 reference genome using Tophat2 (Trapnell et al., 2009) version 2.0.11, using Bowtie version 2.2.1.0 and Samtools version 0.1.19.0. RPKMs per Refseq transcript were calculated from aligned reads using RPKM_count.py from RSeQC (Wang et al., 2012). Fold-changes for +/-LIF conditions were calculated by averaging RPKMs for each condition for all transcripts with the same gene name, dividing the –LIF by the +LIF average RPKM (adding one pseudocount each), and transforming by log2. Gene expression changes after blocking TGF-β signaling by the inhibitor SB431542 were downloaded from a previous study (Mullen et al., 2011). Gene expression changes after stimulation of the Wnt pathway by Wnt 3a conditioned medium were downloaded from a previous study (Cole et al., 2008).

Gene set enrichment analysis (GSEA) of gene expression changes after manipulation of the Wnt, TGF-b, and LIF pathways. “SE-genes” and “TE-genes” indicate genes associated with SEs and typical enhancers, respectively.



If super-enhancers confer responsiveness to the Wnt, TGF-β, and LIF pathways more frequently than typical enhancers, then stimulation or perturbation of these pathways should have a more profound effect on super-enhancer-associated genes than typical enhancer-associated genes. The results of transcriptional profiling and gene set enrichment analysis in ESCs confirm this prediction (Figure 3D); super-enhancer associated genes were found enriched among the genes whose expression exhibited the most profound changes after pathway stimulation or perturbation (Wnt: $p < 0.01$; TGF- β: $p < 0.01$; LIF: $p < 0.01$). In contrast, the enrichment for genes associated with typical enhancers was more moderate (Figure 3D). The super-enhancer-associated genes that showed a profound response to signaling included previously reported targets of these pathways that play key roles in ESC self-renewal and differentiation (Figure 3D, Figure S3G). A subset of the *Prdm14* SE constituents that are bound by signaling TFs were found to be responsive to perturbation of these signaling pathways in reporter assays (Figure S3H). These results lead us to propose that key cell identity genes have evolved a clustered enhancer structure to provide a means to respond directly to these developmentally important signaling pathways.

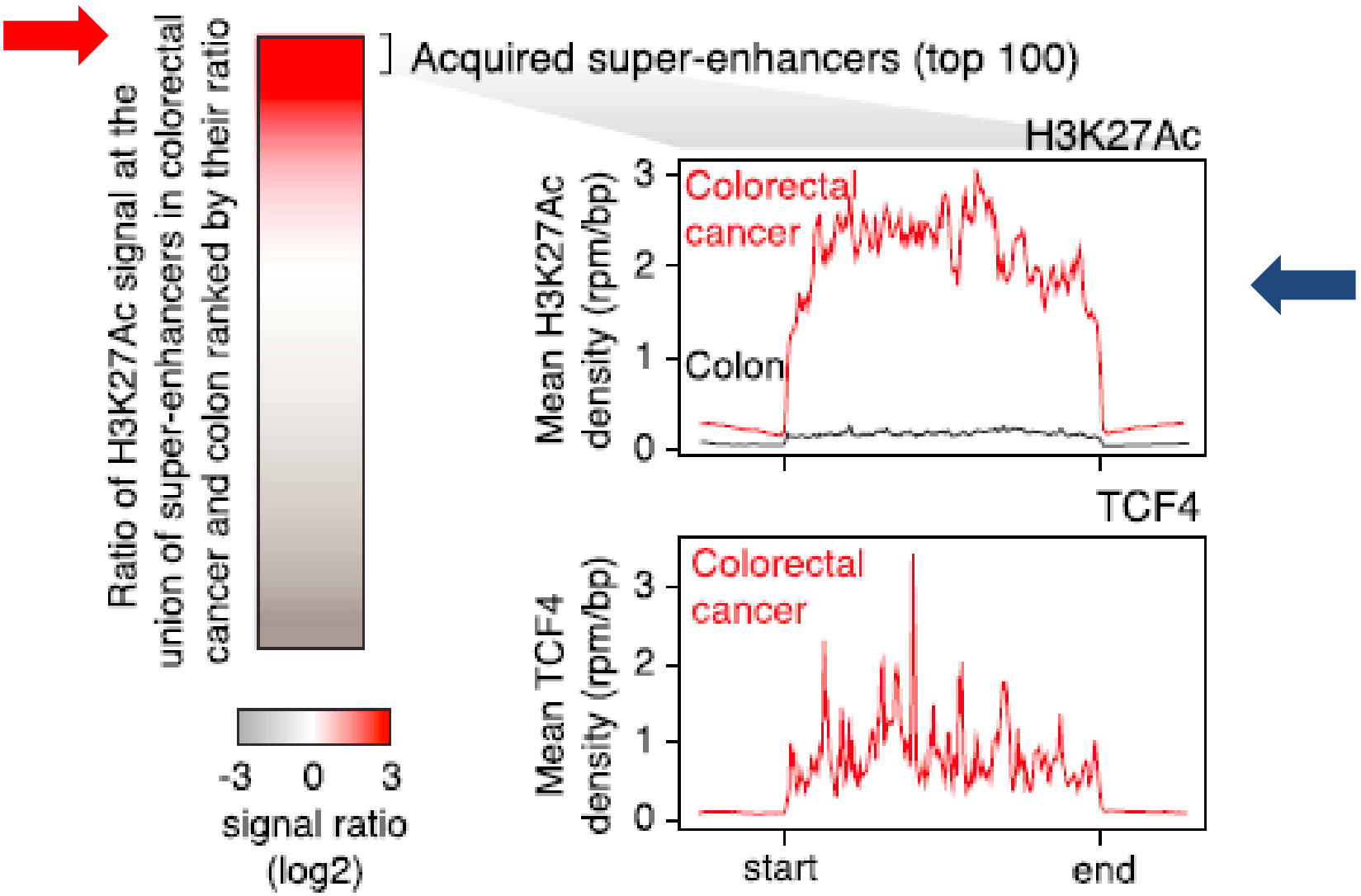
SE function on c-Myc locus



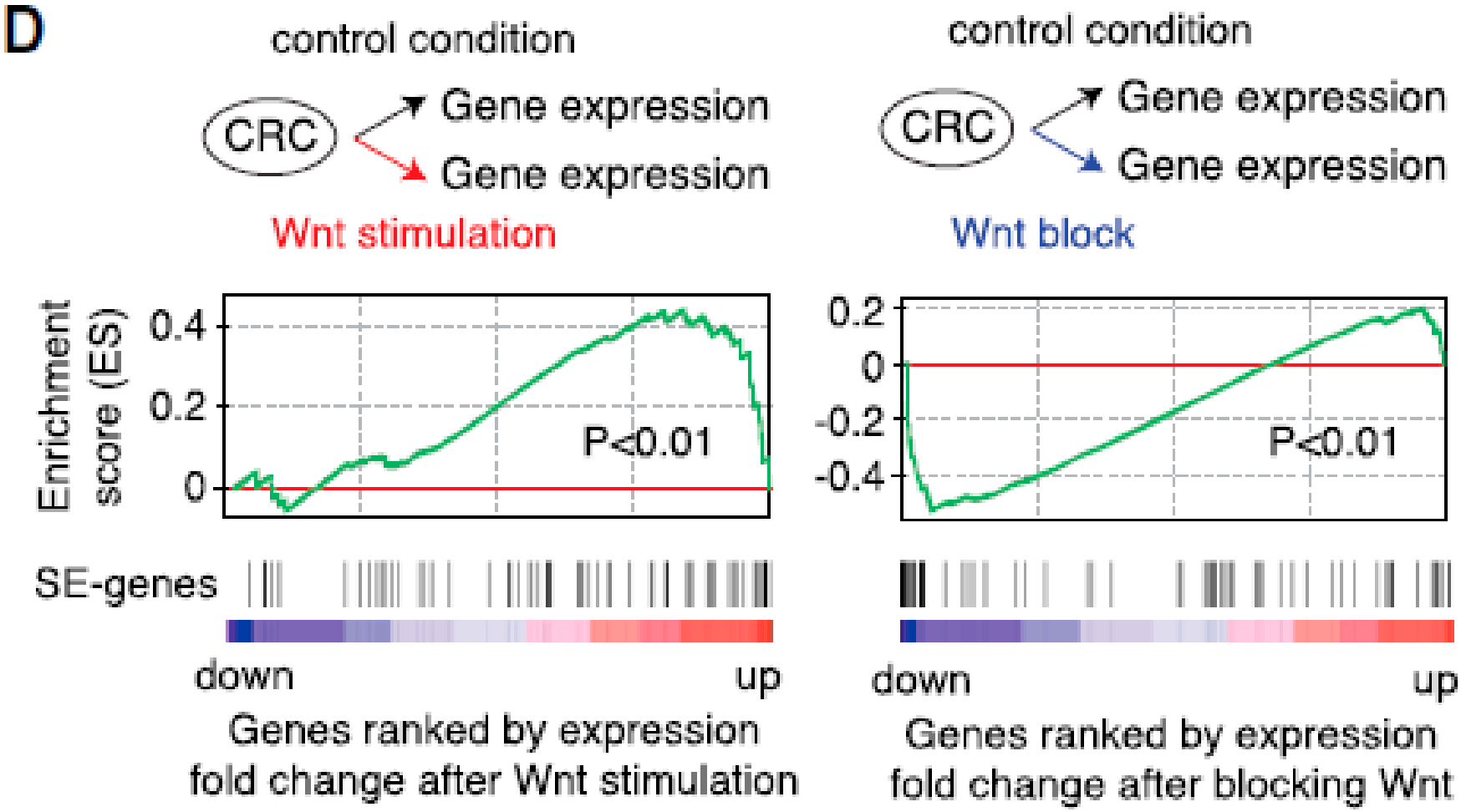
ChIP-seq binding profiles for H3K27Ac at the c-MYC locus in colon and colorectal cancer cells (HCT-116).

Wnt: V6.5 mESCs were cultured in media containing 3μM **IWP-2** *(STEMGENT) for 24* hours prior to transfection **to suppress Wnt signaling**. Cells were then transfected either in media containing 3μM IWP-2 or in media containing 50ng/μl recombinant **Wnt3a** (R&D). Transfected cells were incubated for 24 hours, and luciferase measurements were performed as described above.

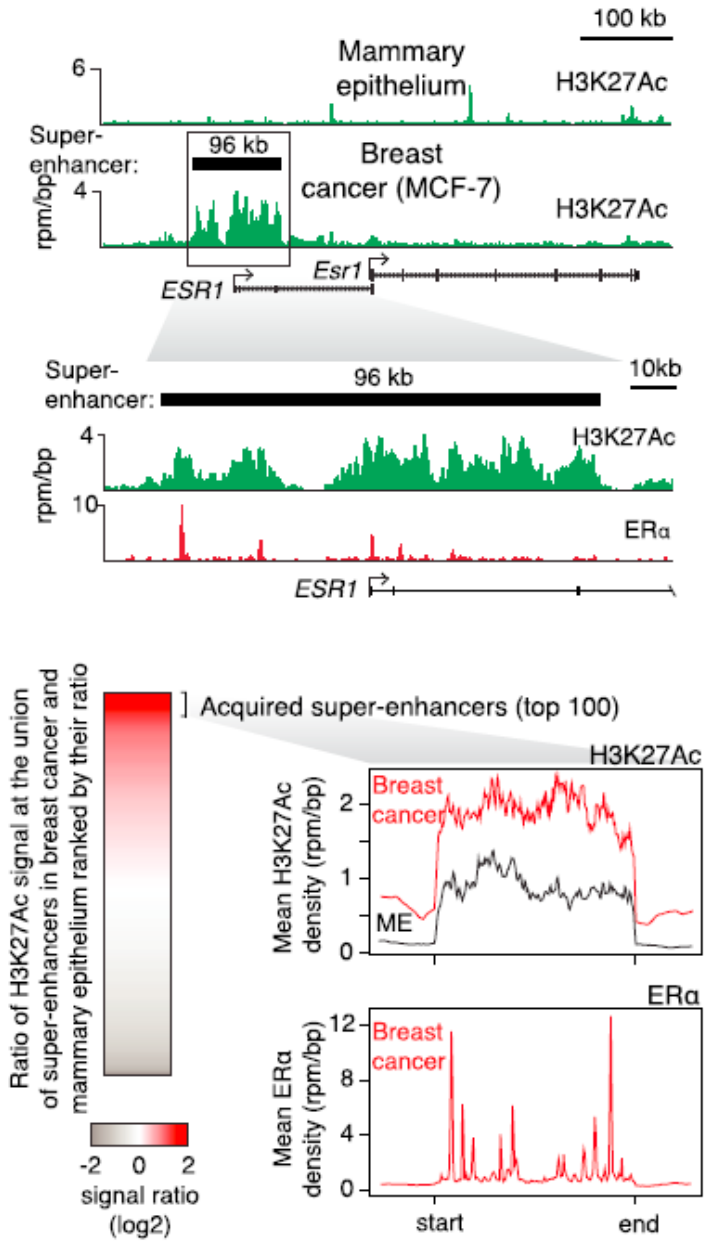
Left: ratio of H3K27Ac in CRC (HCT-116) versus normal colon tissue used densities at the union of SEs identified in the two samples. **Right:** metagene representation of H3K27Ac and TCF4 ChIP-seq densities at the regions corresponding to the top 100 acquired super-enhancers.



ure 4C). Genes associated with these acquired super-enhancers were enriched for expression changes after stimulation or blockage of the Wnt pathway (stimulation: $p < 0.01$; blockage: $p < 0.01$), although not all super-enhancer genes showed this response (Figure 4D). These results indicate that acquired su-



SEs in breast cancer cell lines



DISCUSSION

Super-enhancers control genes that play especially prominent roles in cellular physiology and disease (Brown et al., 2014; Chapuy et al., 2013; Gröschel et al., 2014; Herranz et al., 2014; Hnisz et al., 2013; Lovén et al., 2013; Mansour et al., 2014; Northcott et al., 2014; Parker et al., 2013; Siersbæk et al., 2014; Whyte et al., 2013), but there is a limited understanding of the functions of these clustered elements and, thus, why they have evolved to drive genes that play key roles in cell-type-specific biology. Our results reveal that SEs can provide a platform for signaling pathways to regulate genes that control cell identity during development and tumorigenesis.

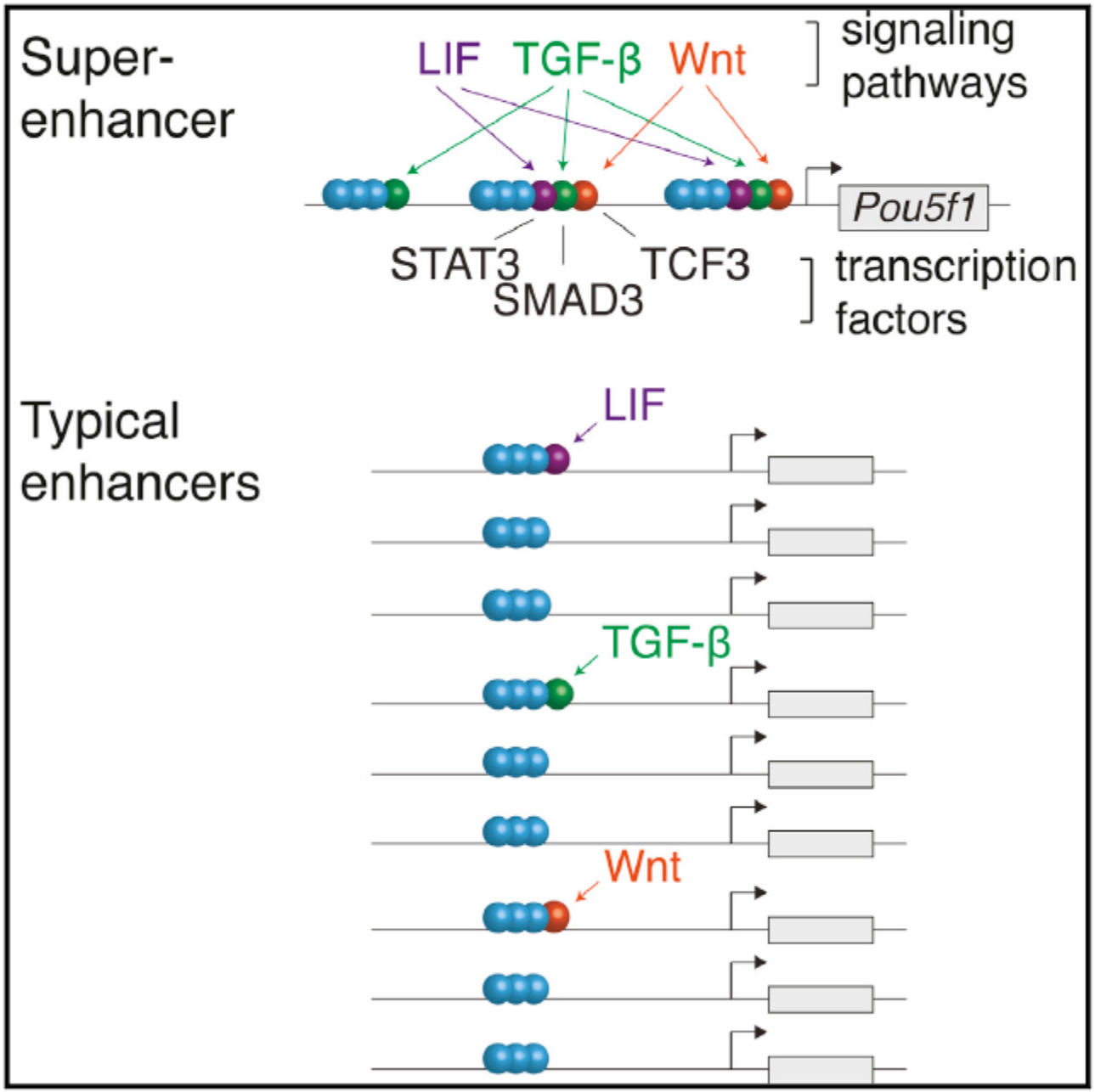
Several lines of evidence argue that the constituent enhancers of at least some super-enhancers can act as an interdependent structural and functional unit to control their associated genes. Our results show that ESC SEs generally consist of clusters of active enhancers that have OCT4-dependent and ESC-specific functions (Figure 1) and demonstrate that optimal transcriptional activity of target genes is dependent on the presence of most of the constituent enhancers (Figure 2). Chromatin interaction data indicate that constituent enhancers physically interact within the SEs; indeed, the interactions among SE constituents in ESCs appear to be more frequent than interactions between the SE constituents and their associated gene promoters, and interactions between typical enhancers (Dowen et al., 2014). We previously noted that enhancer clusters can be gained or lost as a unit during development or oncogenesis (Hnisz et al., 2013) and have shown that large tumor SEs can collapse as a unit when depleted of the enhancer cofactor BRD4 (Lovén et al., 2013) or when a constituent is deleted (Mansour et al., 2014). In some T cell acute lymphoblastic leukemia (T-ALL) cells, a small mono-allelic insertion that creates a binding site for a master transcription factor can nucleate the formation of an oncogenic super-enhancer that involves establishment of additional transcriptional components in adjacent sites (Mansour et al., 2014). Super-enhancers produce relatively high levels of enhancer RNAs (Hnisz et al., 2013), and a recent study showed that inflammation-dependent super-enhancers form domains of coordinately regulated enhancer RNAs (Hah et al., 2015). These results, taken together, suggest that the constituent enhancers of super-enhancers can interact physically and functionally to coordinate transcriptional activity.

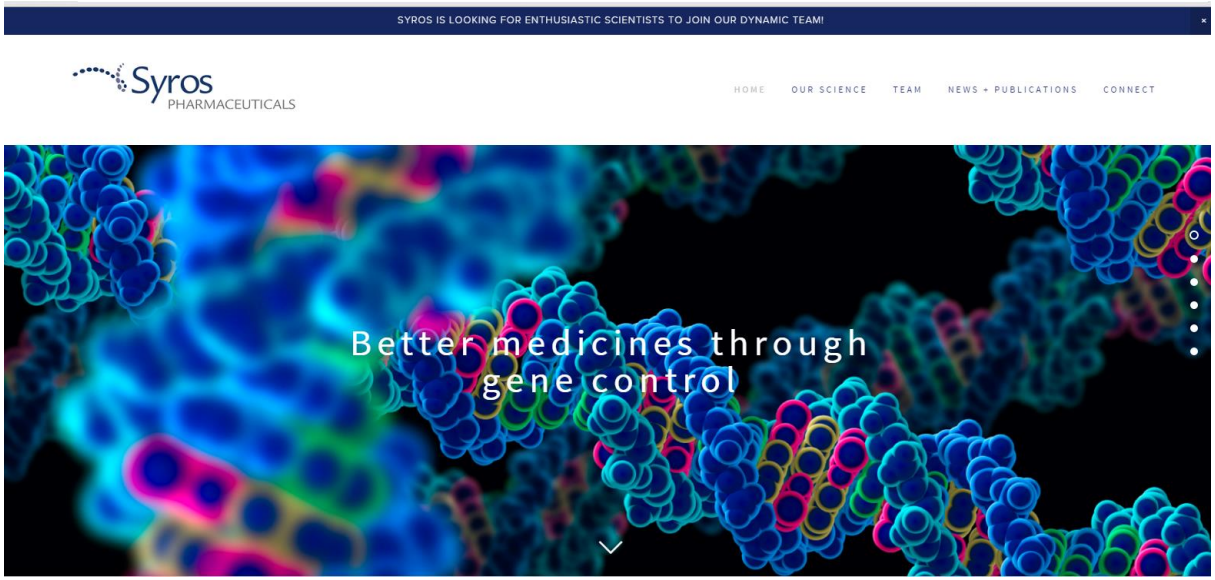
SEs characteristics

- **OCT4 binding in ESC**
- **SE regulates transcription**
- **SE chromatin interaction**
- **SE role in oncogenesis**
- **eRNA linked to SE**
- **inflammation linked to SE**

Our results reveal that SEs are occupied more frequently by terminal transcription factors of the Wnt, TGF-β, and LIF signaling pathways than typical enhancers in ESCs, and genes driven by SEs show a more pronounced response to the manipulation of these pathways than genes driven by typical enhancers (Figure 3). Thus, the clustered enhancer architecture of SEs may have evolved, at least in part, to provide a conduit for these signaling pathways to signal maintenance or change at genes that are key to control of cell identity. Our results also suggest that one reason that tumor cells evolve SEs at key oncogenes is to enhance the connection to oncogenic signaling pathways. The recent report of NOTCH1-driven SEs in T-ALL likely represents another example of this phenomenon (Herranz et al., 2014; Wang et al., 2014). An implication of this model is that therapies that target both oncogenic signaling pathways and super-enhancer components may be especially effective in tumor cells that have signaling and transcriptional dependencies.

Graphical Abstract

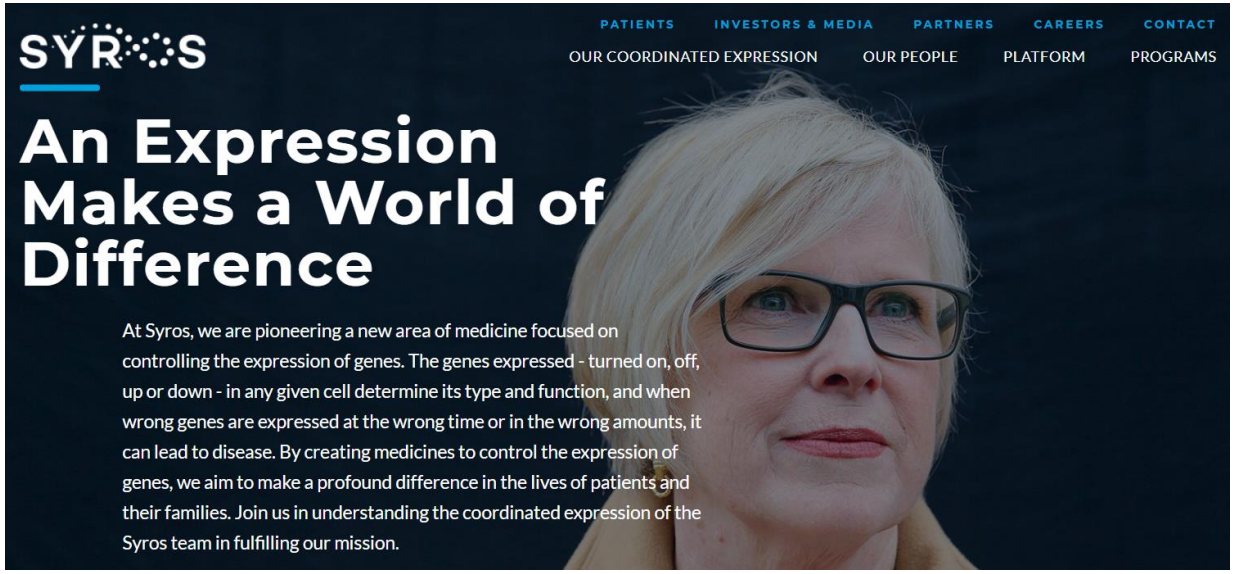




2015

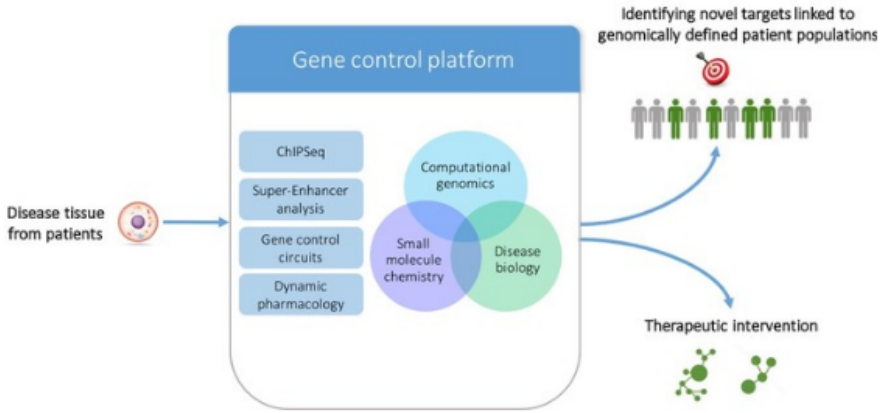
2019

https://d1io3yog0oux5.cloudfront.net/_f40afe575865714f6435a44ff4426eae/syros/db/306/1890/file/gene-control-final.mp4



OUR PLATFORM

Syros is solely focused on gene control. Building on the seminal discoveries of our scientific founders, we have developed what we believe is the first proprietary platform designed to systematically and efficiently analyze non-coding regions of the genome to identify alterations in gene expression programs.

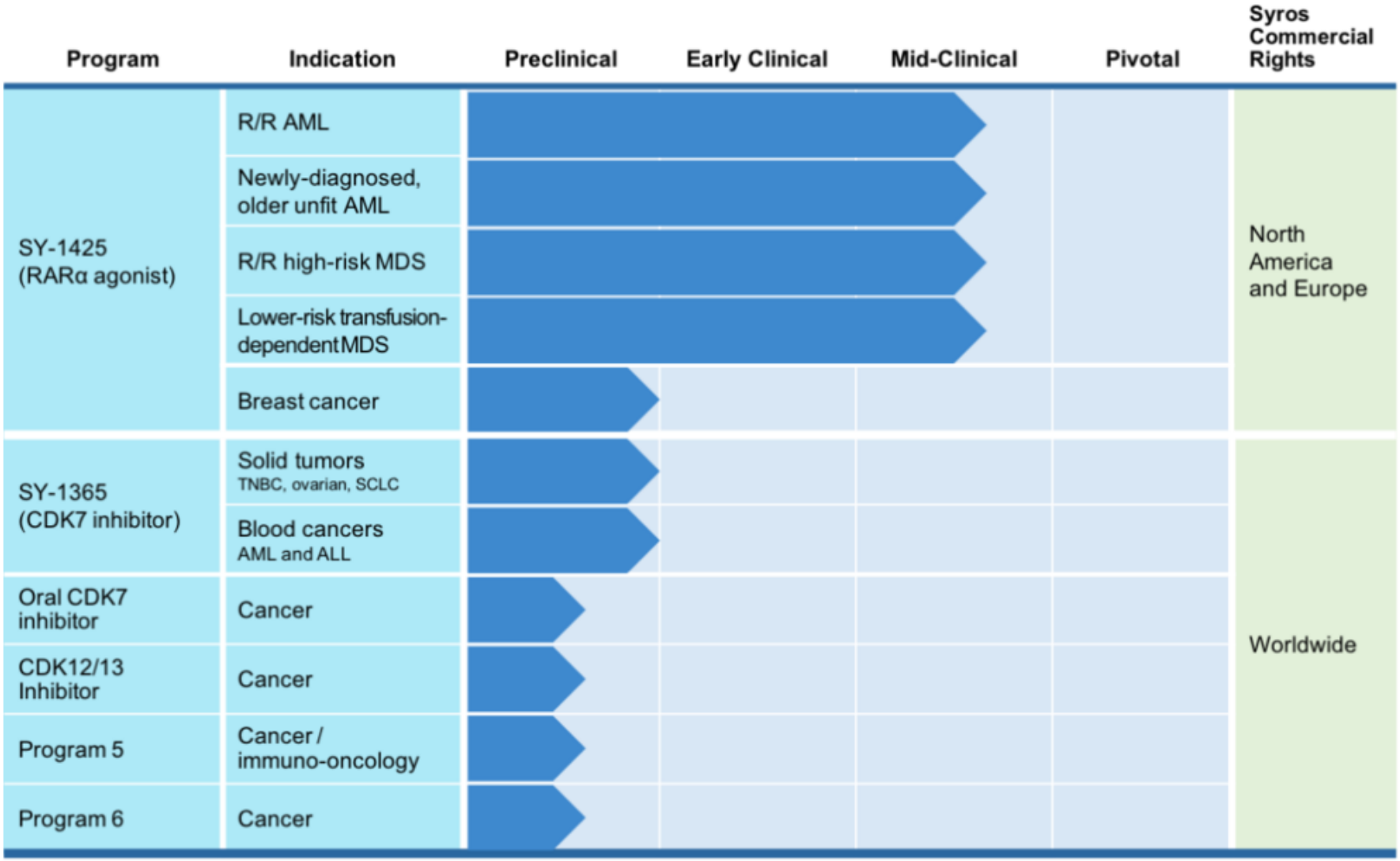


By doing so, we believe our gene control platform will allow us to:

- Identify a wide array of potential new drug targets across a range of diseases
- Provide a new lens for diagnosing and segmenting patients, including those with complex, multi-factorial diseases that have eluded segmentation with other genomic-based approaches
- Advance a new wave of medicines with the potential to influence multiple drivers of disease through a single target, making them less susceptible to drug resistance and providing patients with a more profound and durable benefit than many of today's targeted therapies.

Learn more about our platform

DRUGS DISCOVERY FROM SYROS PHARMACEUTYCAL



*SY-1425 is approved in Japan as Amnolake® (tamibarotene) for patients with relapsed/refractory APL