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

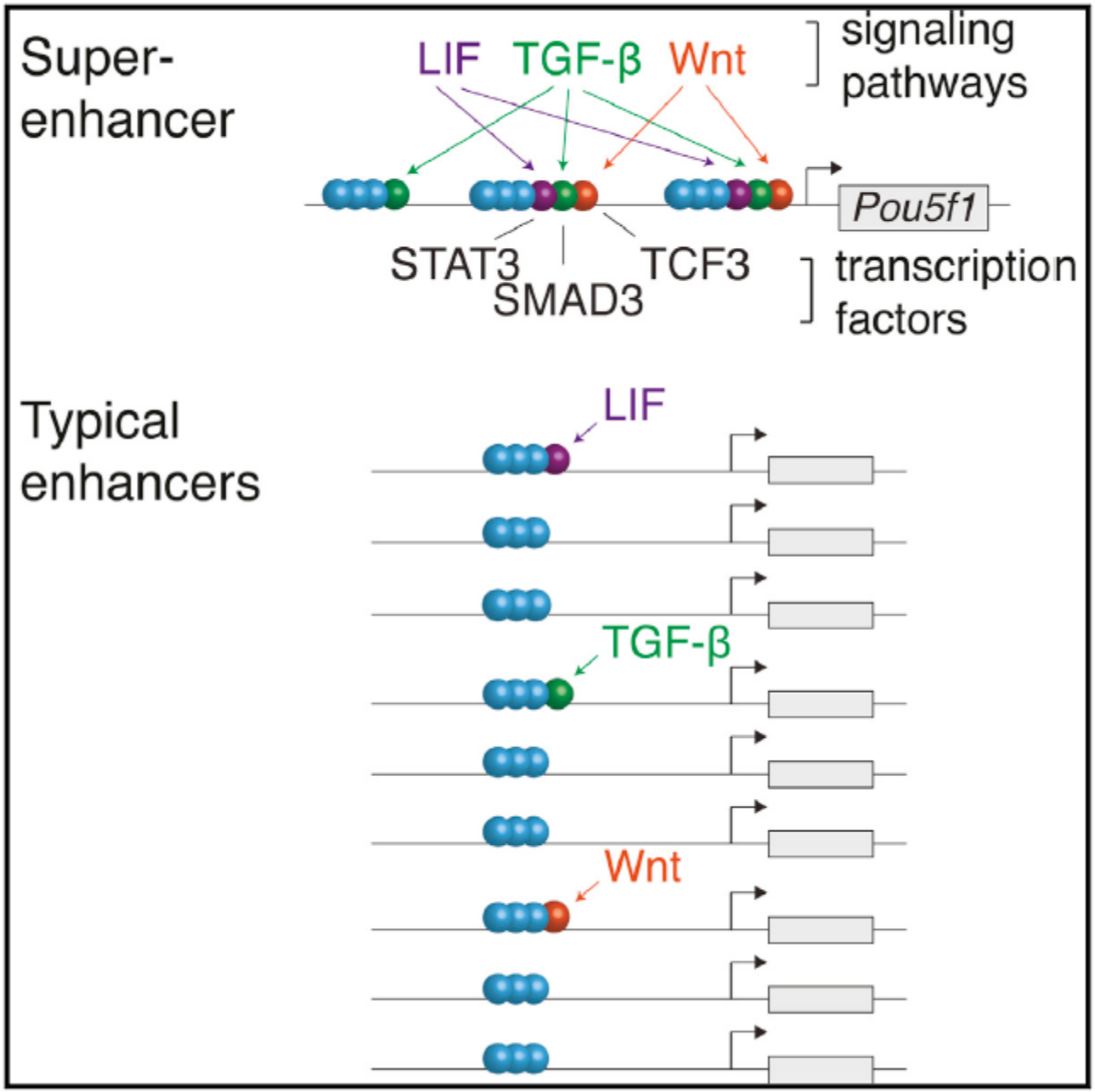
AIM

Super-enhancers (SE)
Functional costituent units

Cancer cells SE target
for oncogenic signalling

CONCLUSION

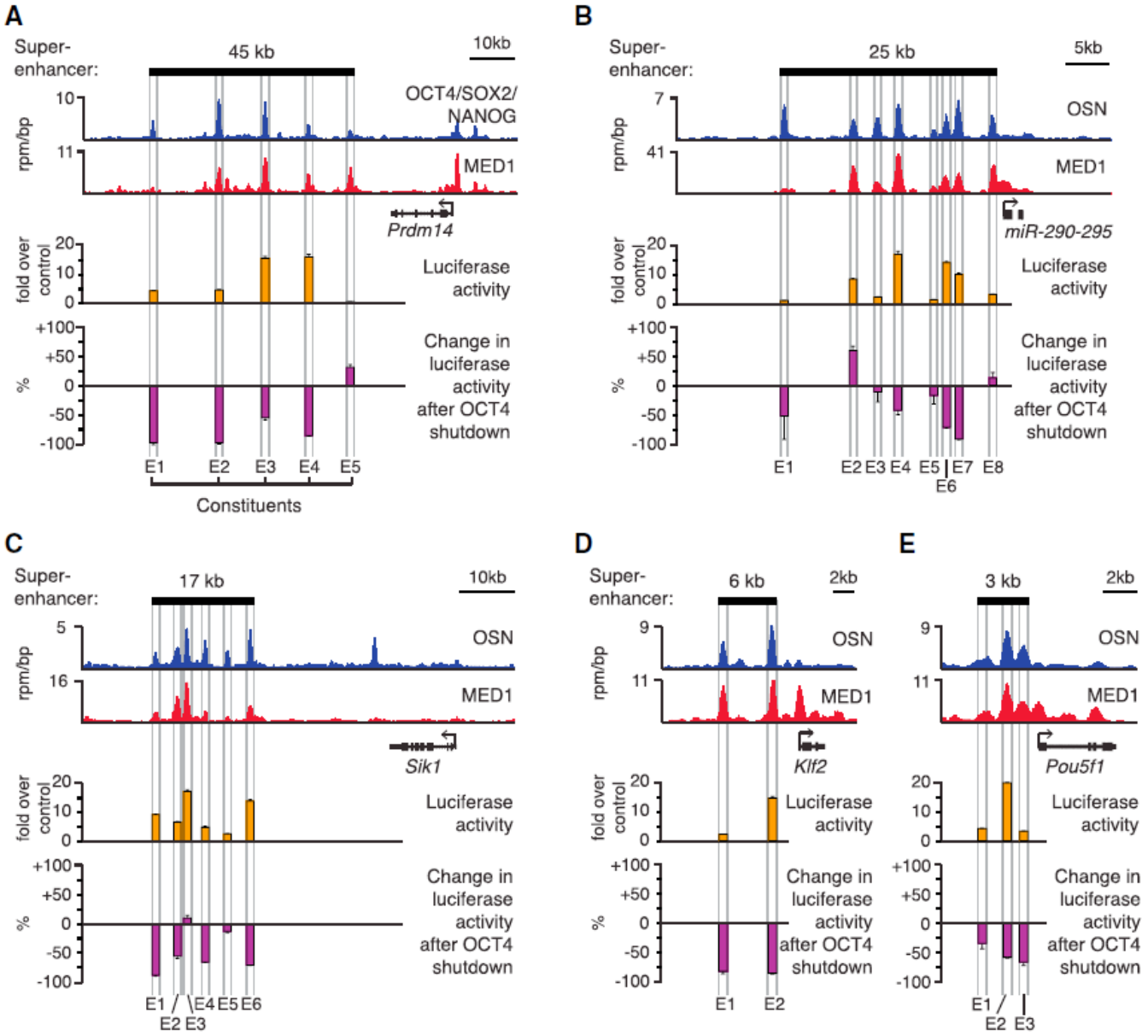
Graphical Abstract



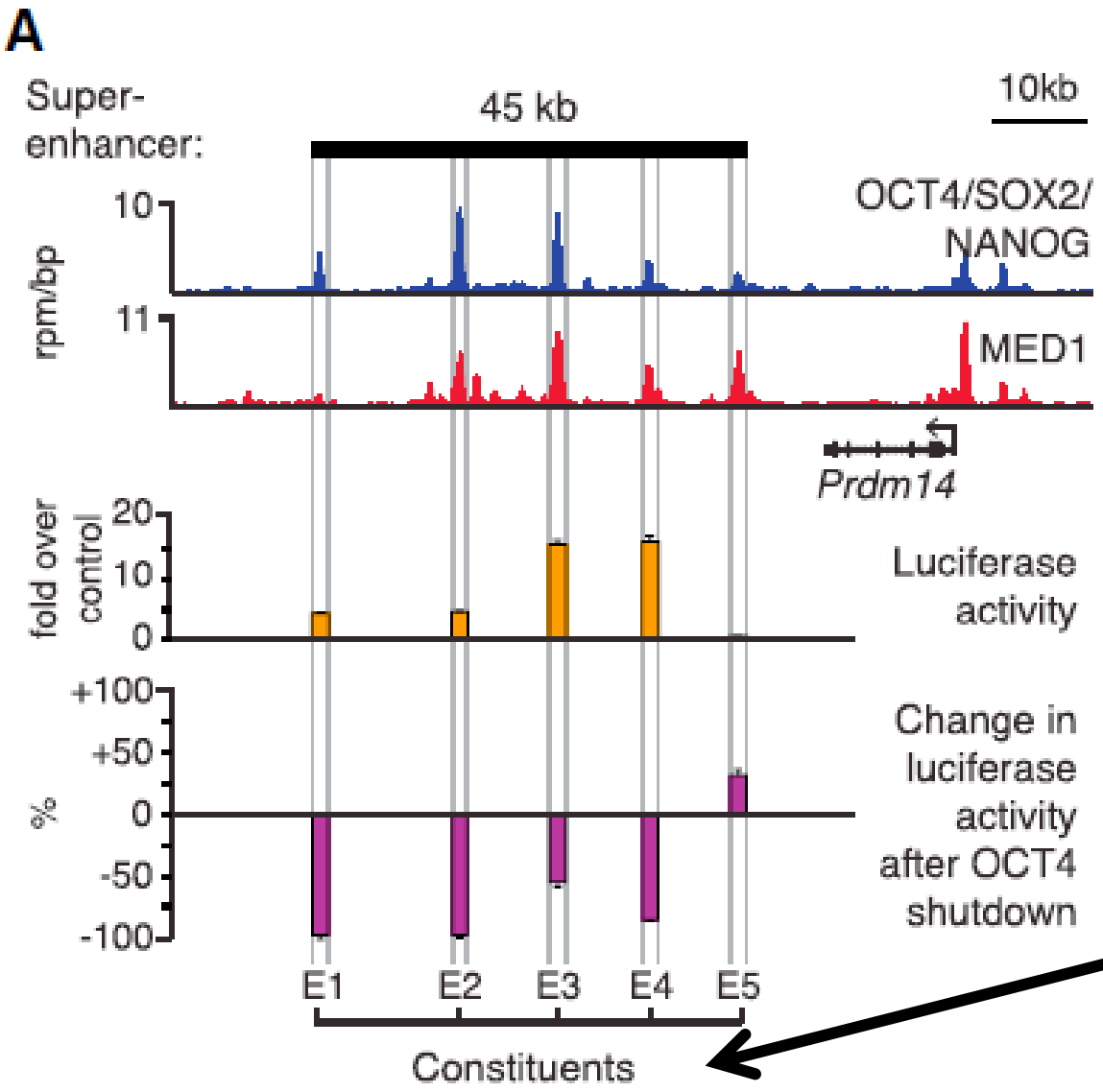
Highlights

- 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

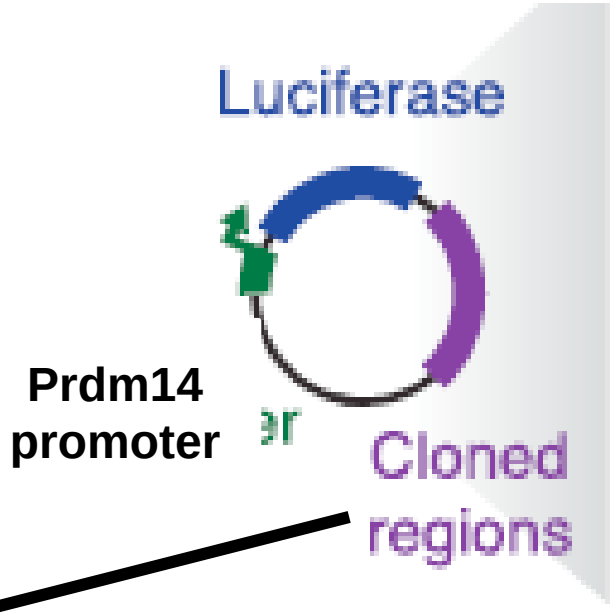
ACTIVITY OF SUPER-ENHANCER CONSTITUENTS IN SEVERAL GENOMIC LOCI



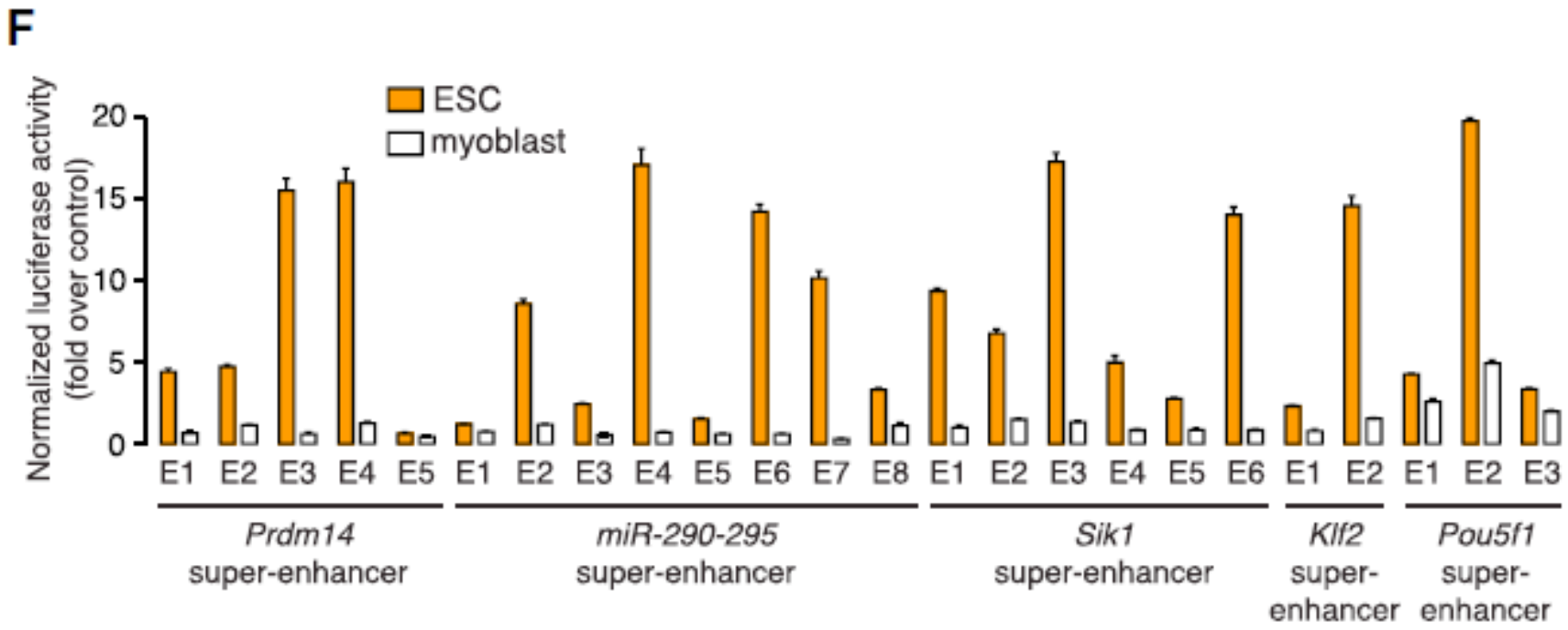
ACTIVITY OF SUPER-ENHANCER CONSTITUENTS



ChIP-Seq profile

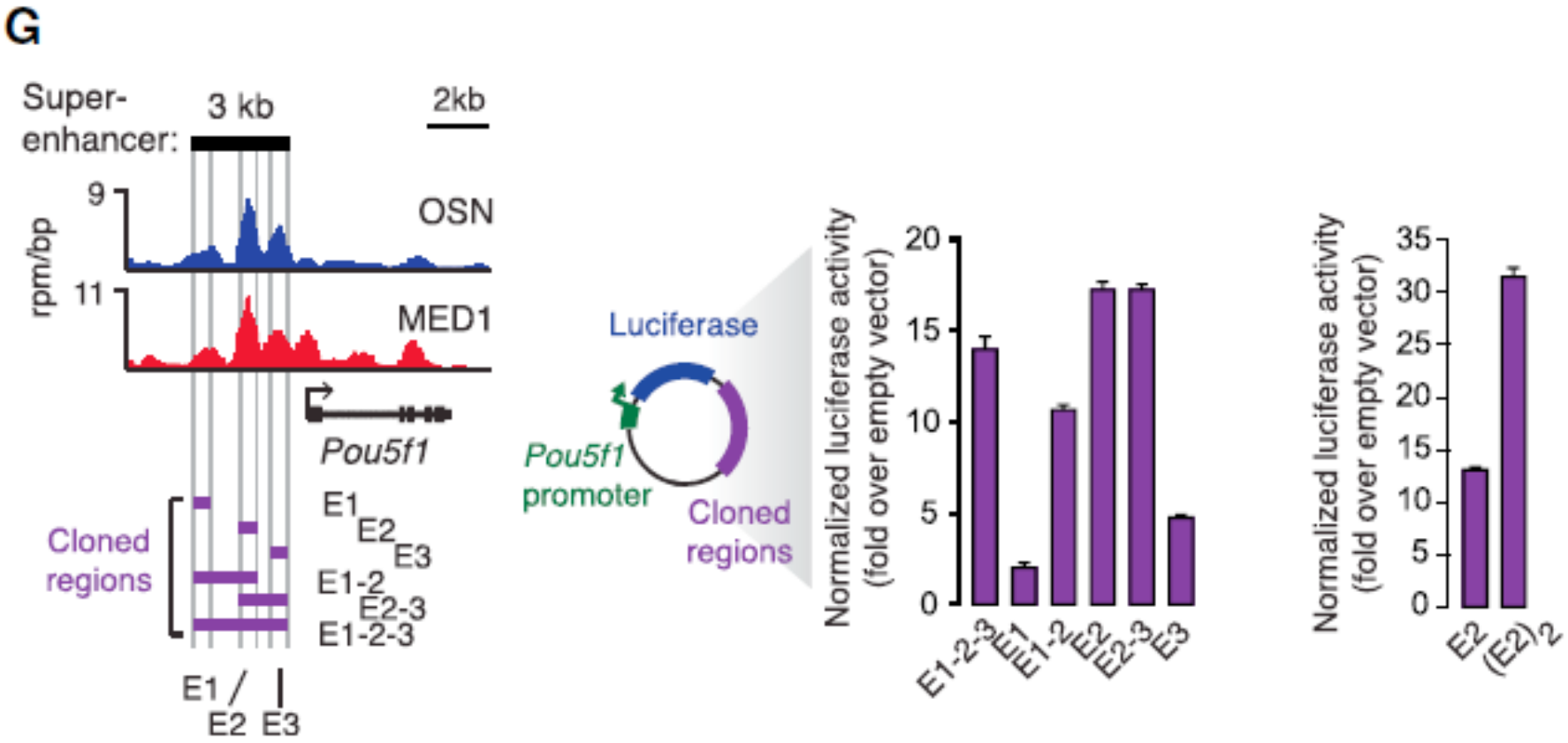


ENHANCER ACTIVITY OF SE CONSTITUENTS IS SPECIFIC FOR ESC, COMPARED TO MYOBLAST

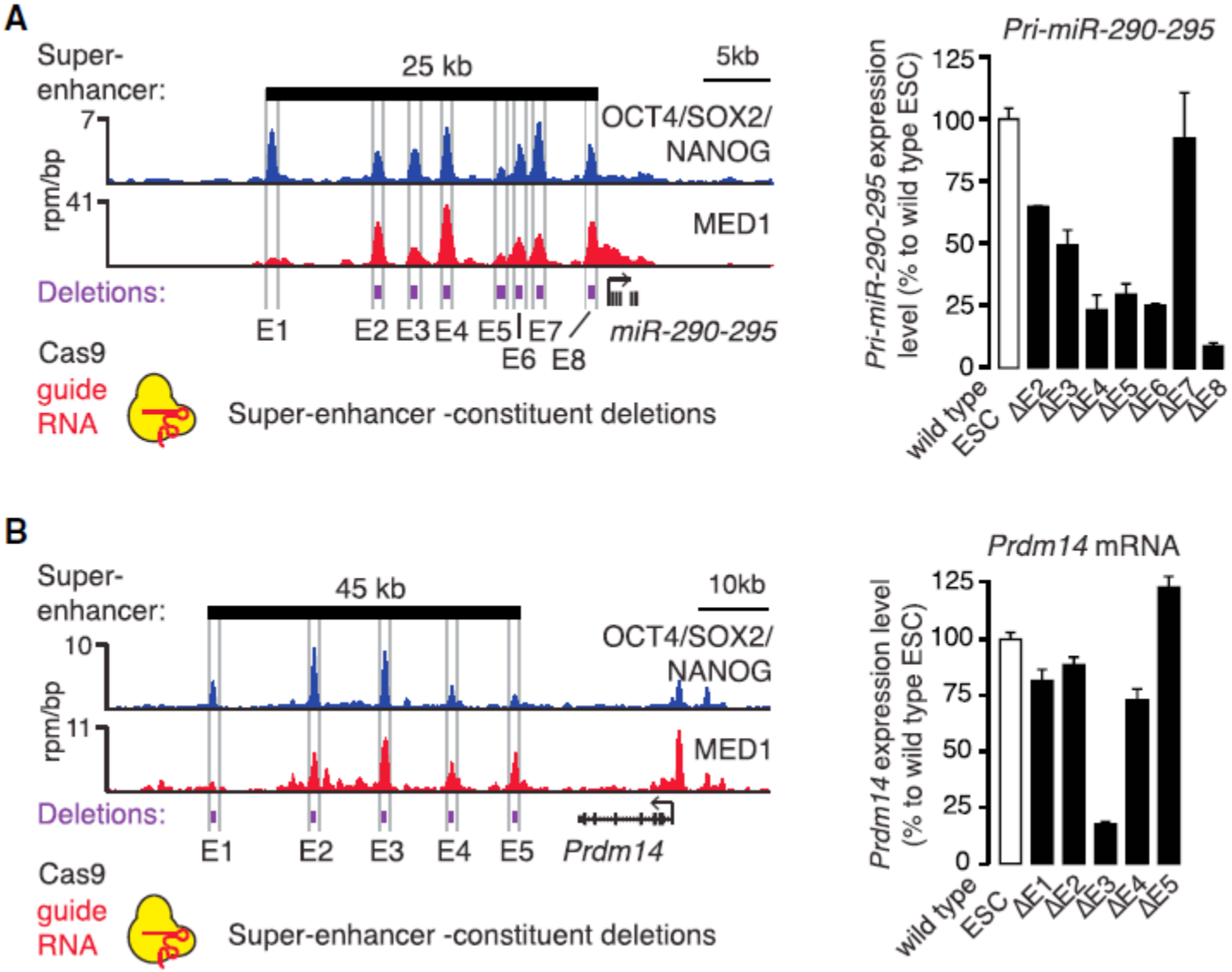


QUESTION: DO "SE CONSTITUENTS" ACT ADDITIVELY, SINERGISTICALLY OR EXERT A COMPLEX INFLUENCE?

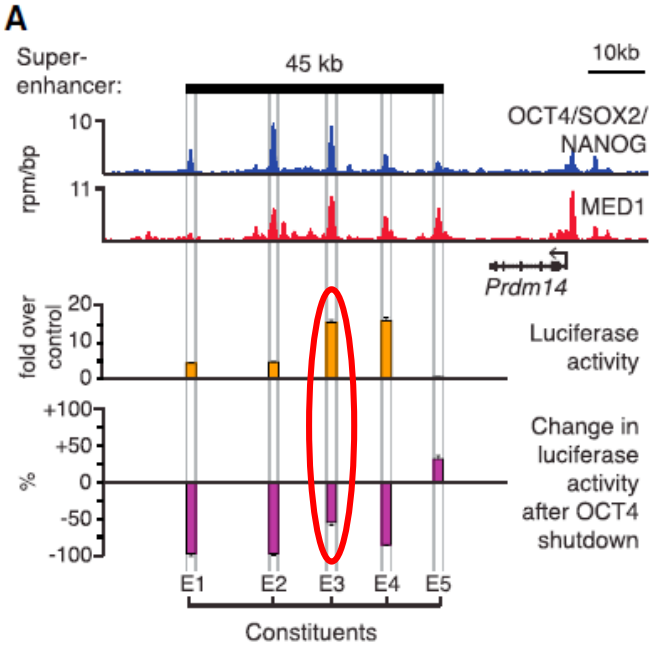
E2 has high activity, E1 and E3 influence E2 activity

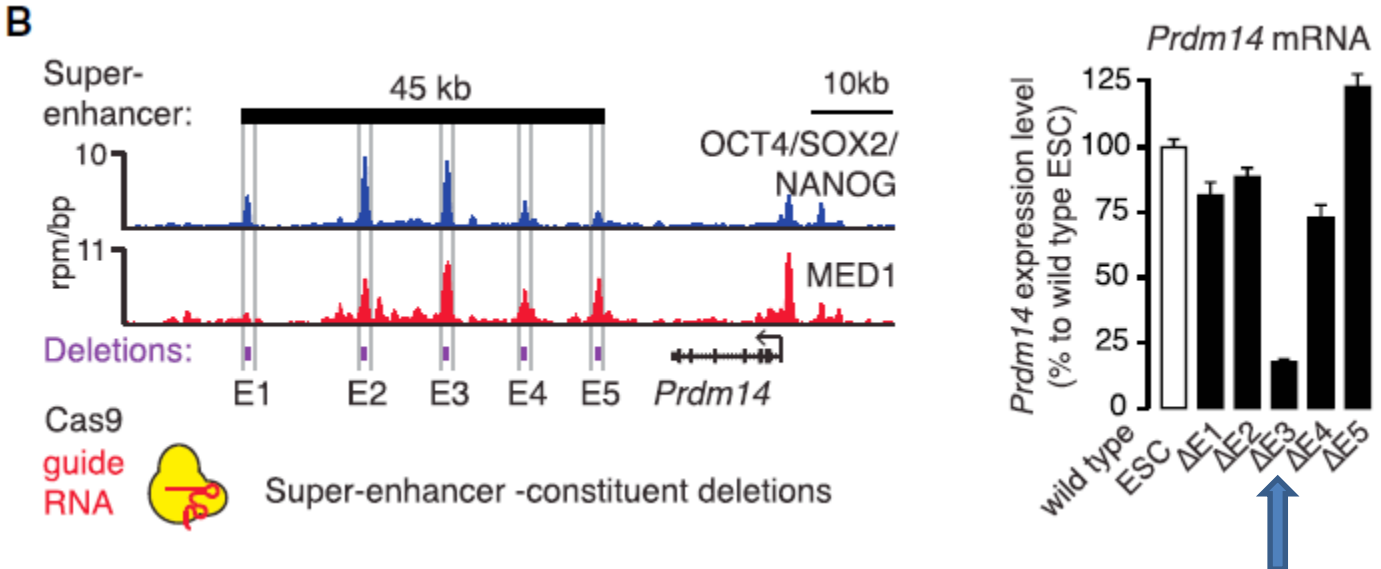


CONTRIBUTIONS OF SUPER-ENHANCER COSTITUENTS TO GENE EXPRESSION IN VIVO

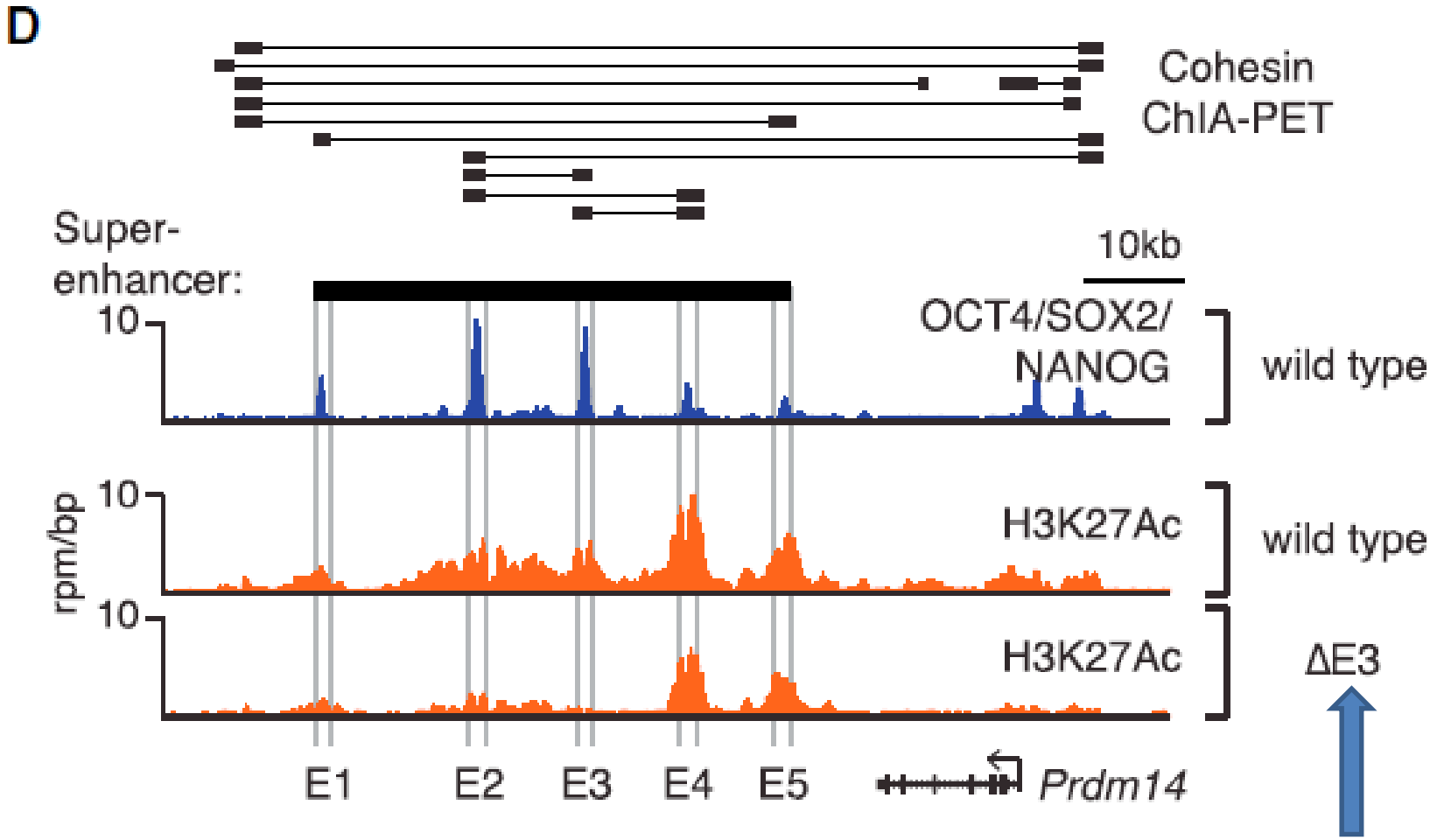


LUCIFERASE ACTIVITY AND DELETION OF SPECIFIC ENHANCER





IDENTIFICATION OF INTERACTION BETWEEN SPECIFIC COSTITUENTS ENHANCERS



DELETION OF E3 INDUCES EPIGENETIC MODIFICATIONS: H3K27Ac reduction

OVERVIEW OF MOLECULAR APPROACHS TO STUDY SUPER-ENHANCERS

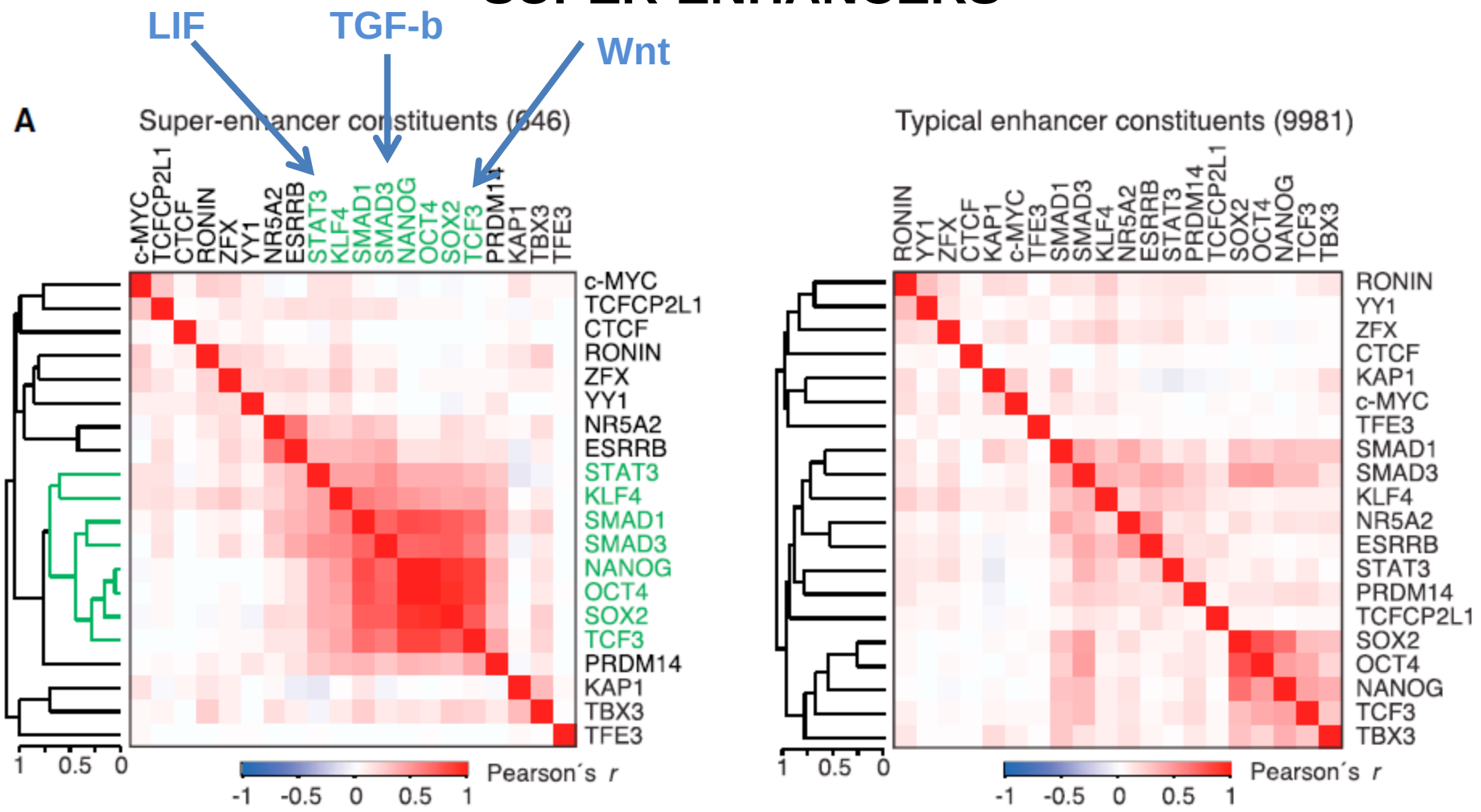
LUCIFERASE ACTVITY WITH VECTOR THAT CONTAINS COSTITUENT ENHANCER → ROLE OF SE IN GENE TRANSCRIPTION REGULATION

CRISPR-CAS9 SYSTEM TO DELETE COSTITUENT ENHANCER

- ROLE OF SE IN GENE EXPRESSION
- ROLE OF SE IN TRANSCRIPION FACTOR BINDING
- ROLE OF SE IN CHROMATIN REMODELING

SUPER-ENHANCERS ENRICHMENT OF TRANSCRIPTION FACTORS PATTERN

SPECIFIC MULTIPLE TRANSCRIPTION FACTORS BIND SUPER-ENHANCERS

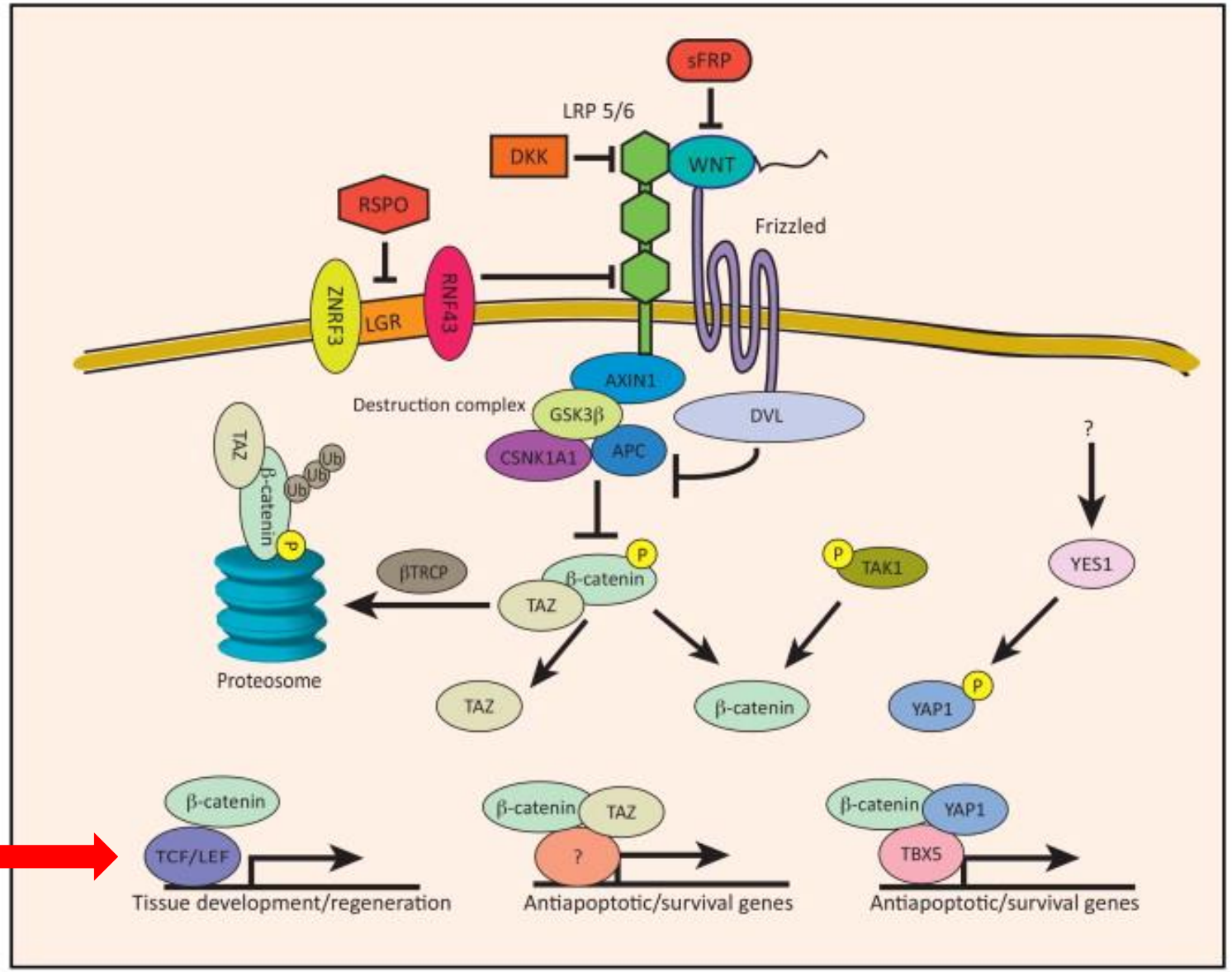


Hierarchical clustering of 20 transcription factor ChIP-seq binding profiles at super-enhancer and typical enhancer constituents. A set of factors with binding profiles similar to OCT4, SOX2, and NANOG is highlighted in green.

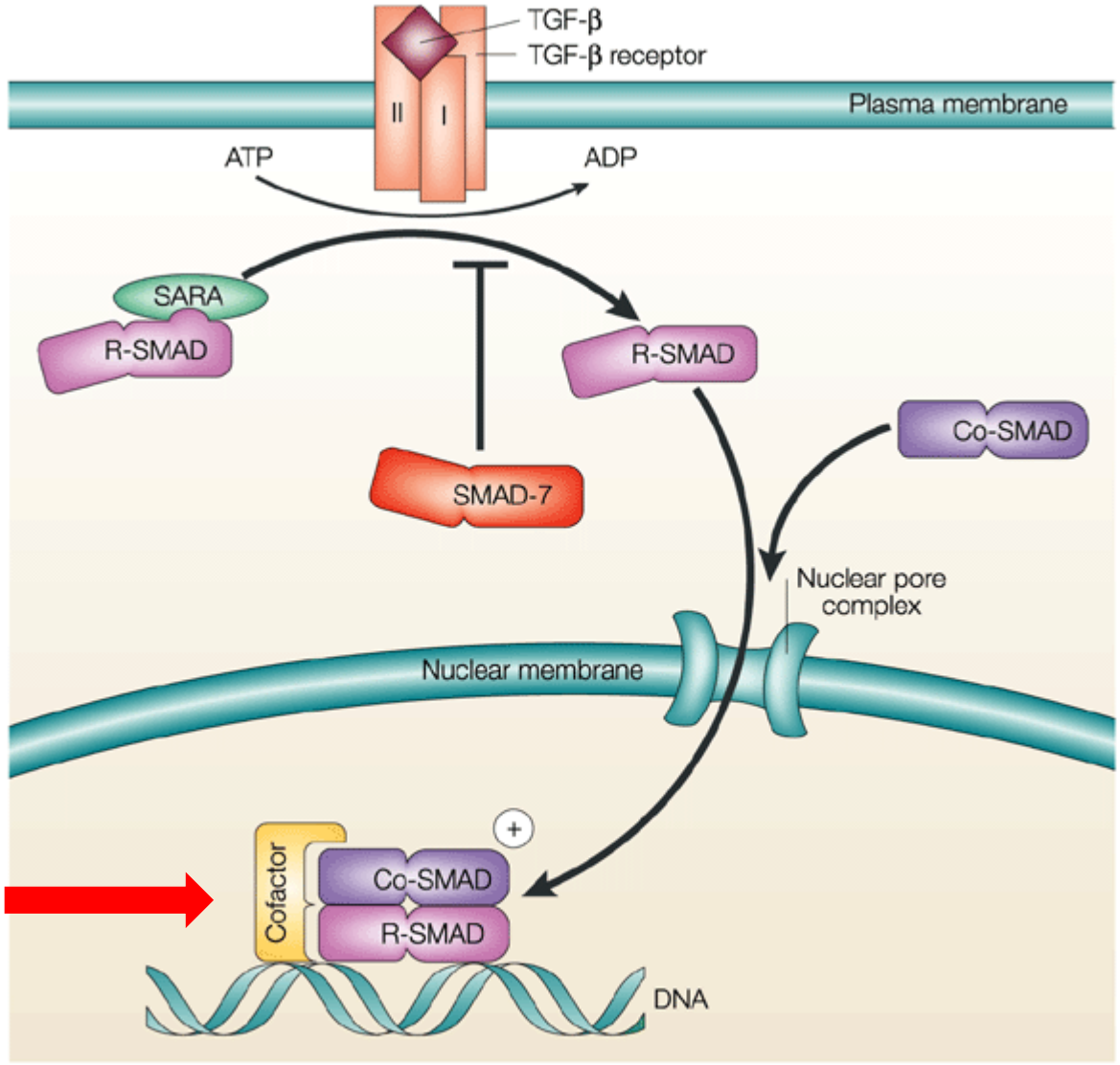
enhancers. An examination of the pattern of transcription factor binding to super-enhancer constituents provided a hypothesis to resolve this conundrum (Figure 3A, Table S3). The terminal TFs of the Wnt (TCF3), TGF-β (SMAD3), and LIF (STAT3) signaling pathways, which play essential roles in transcriptional 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

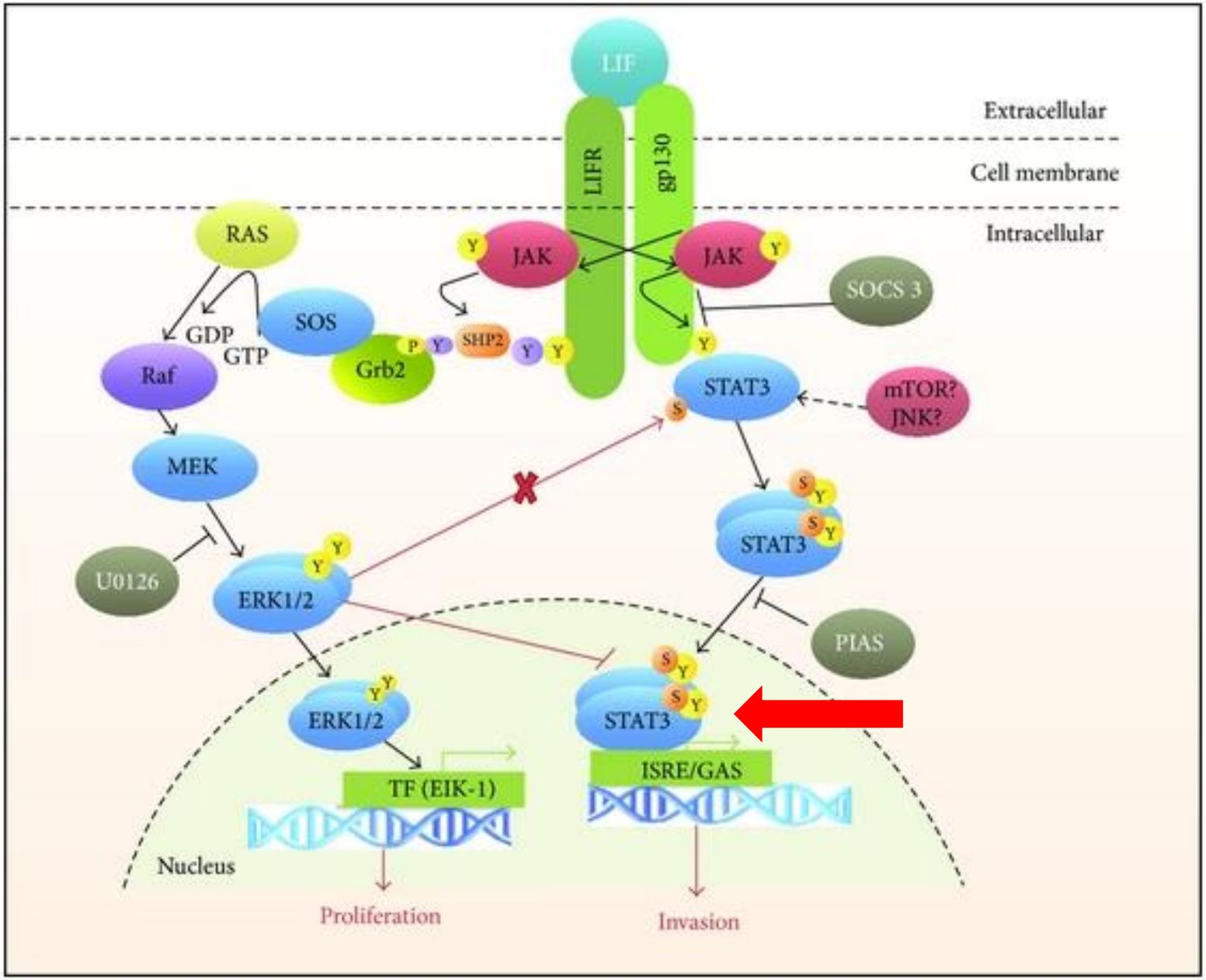
WNT PATHWAY


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

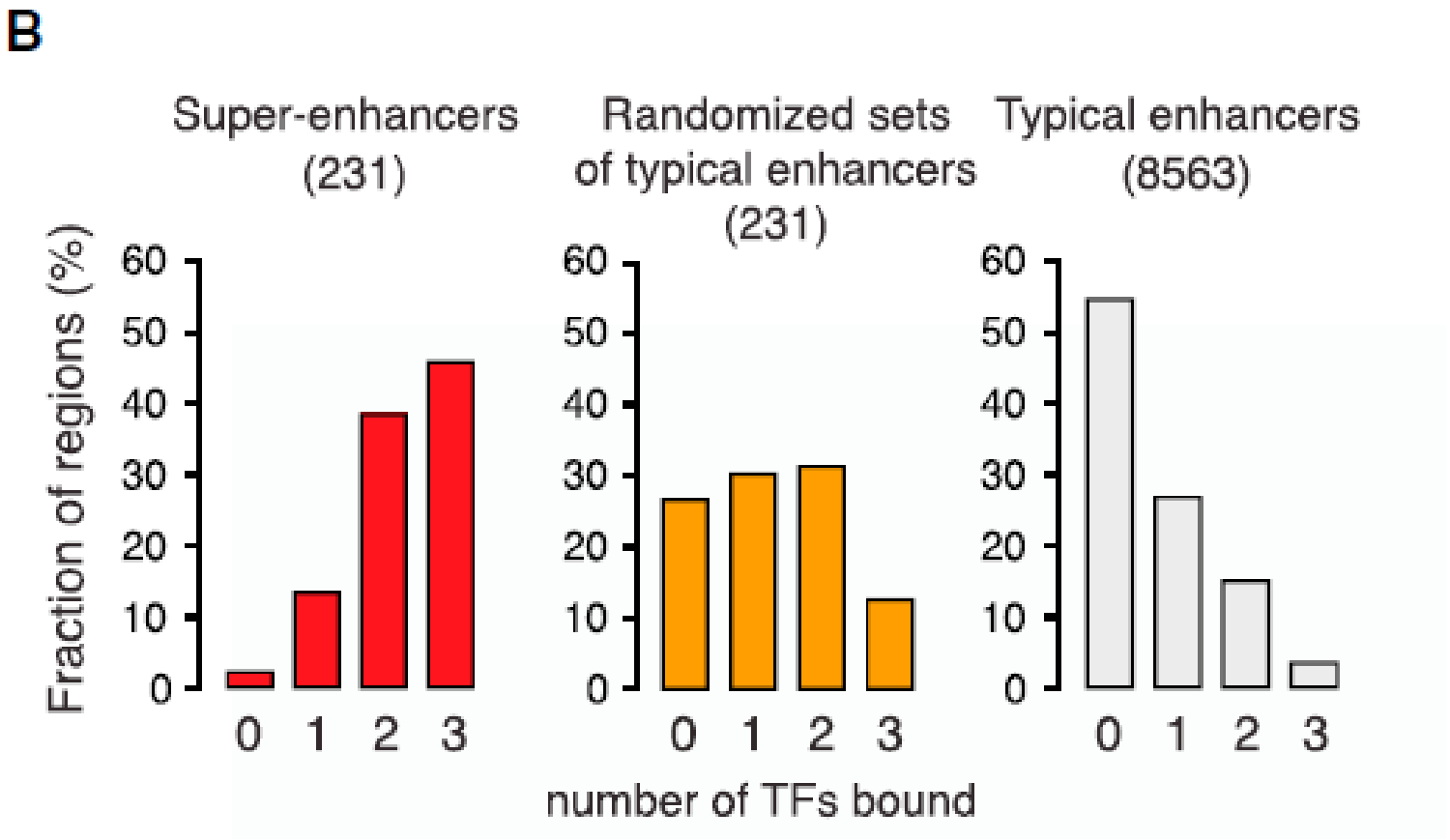
TGF beta PATHWAY



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.

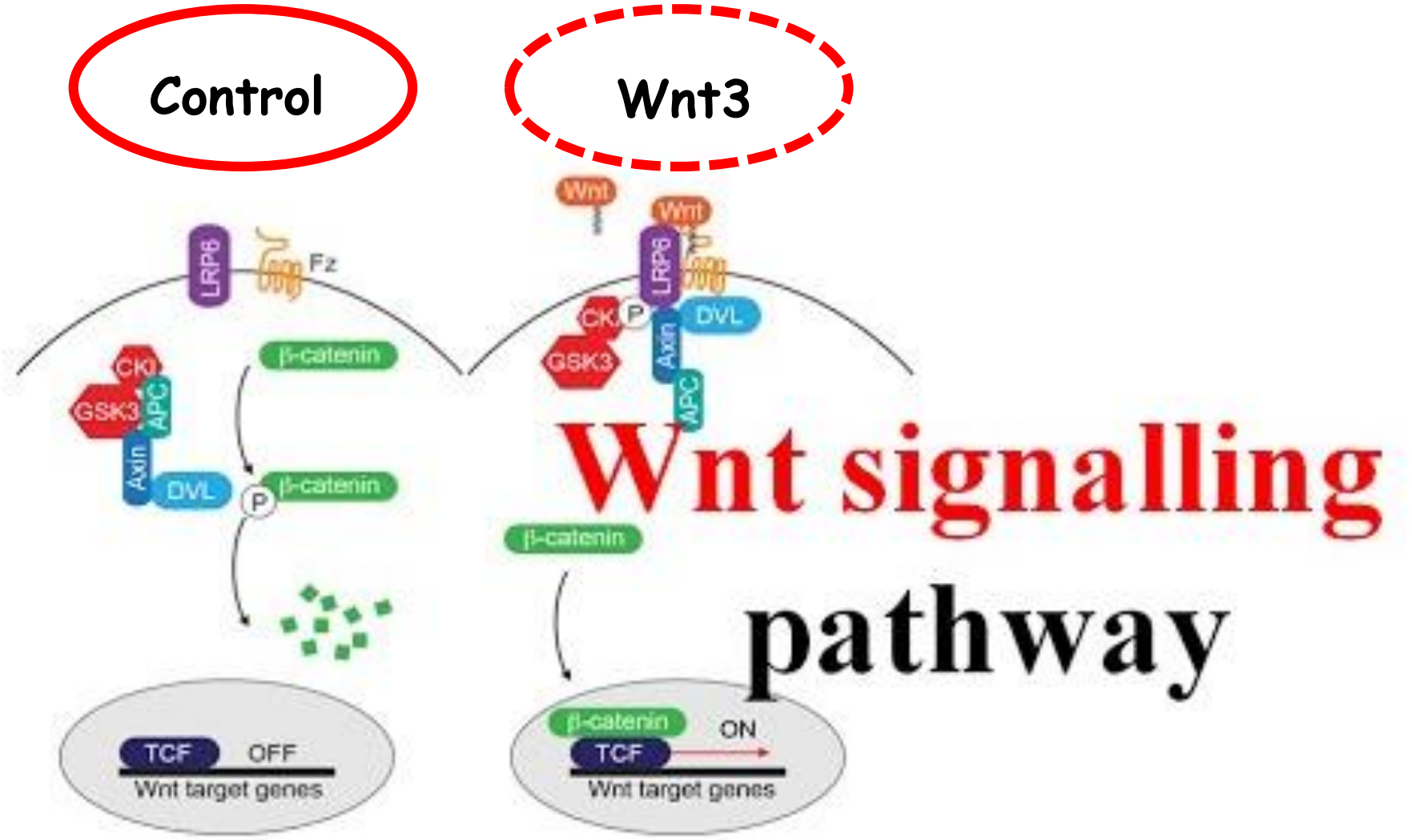
c

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

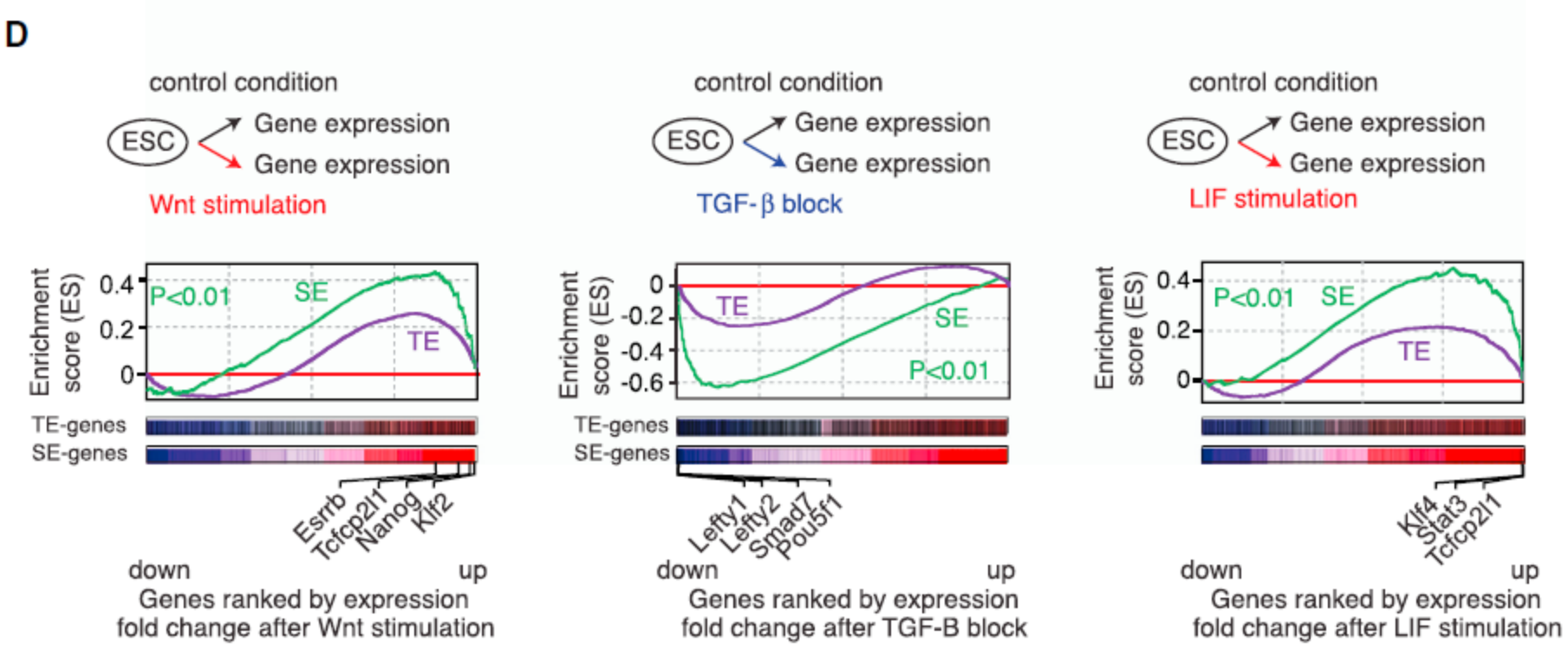
Gene expression analyses

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).

Mouse Embrionic stem cells
RNA-seq

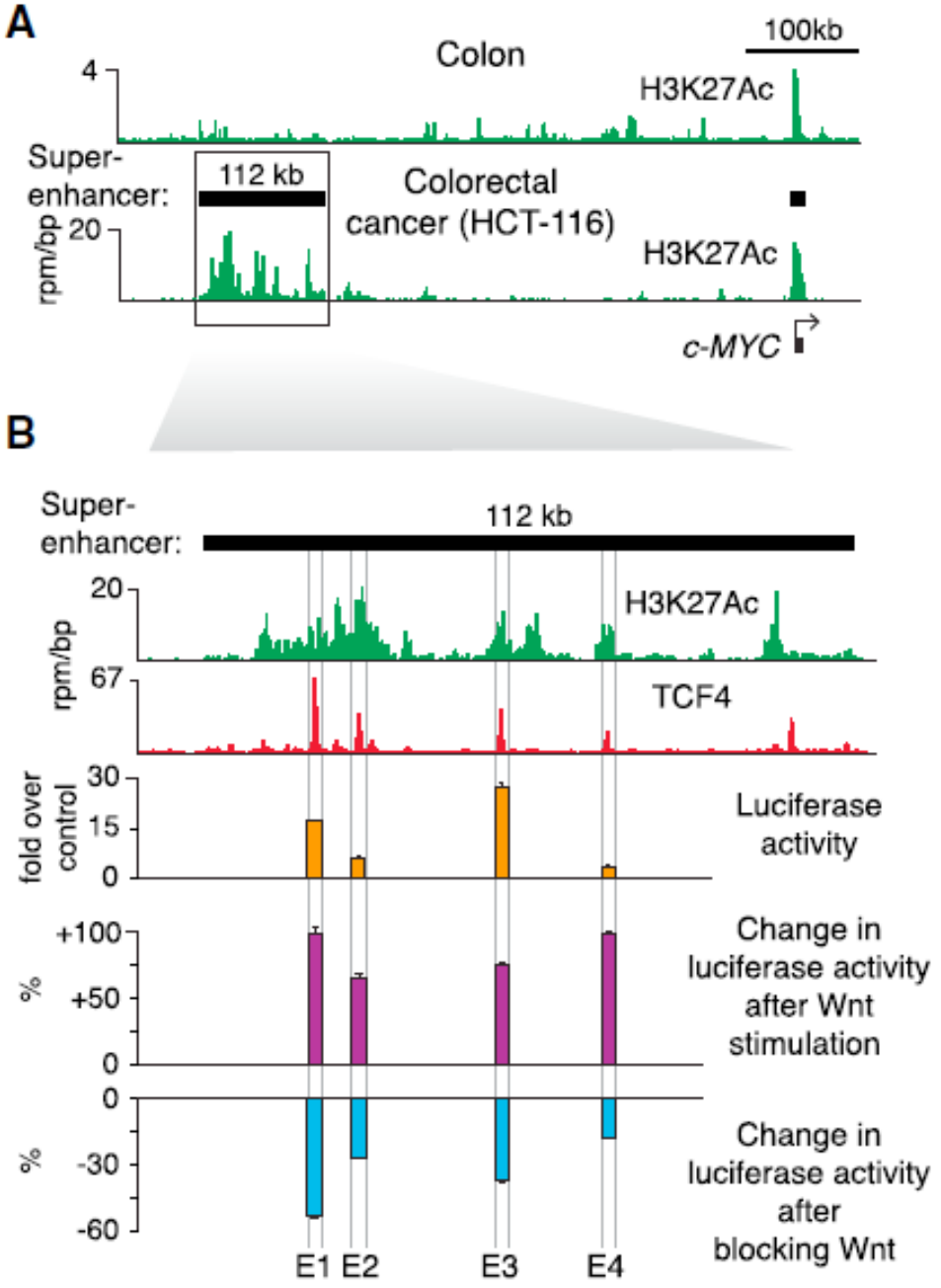


Gene set enrichment analysis (GSEA) of gene expression changes after manipulation of the Wnt, TGF-b, and LIF pathways. “SE-genes” and “TE-genes” indicate genes associated with SEs and typical enhancers, respectively.



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.

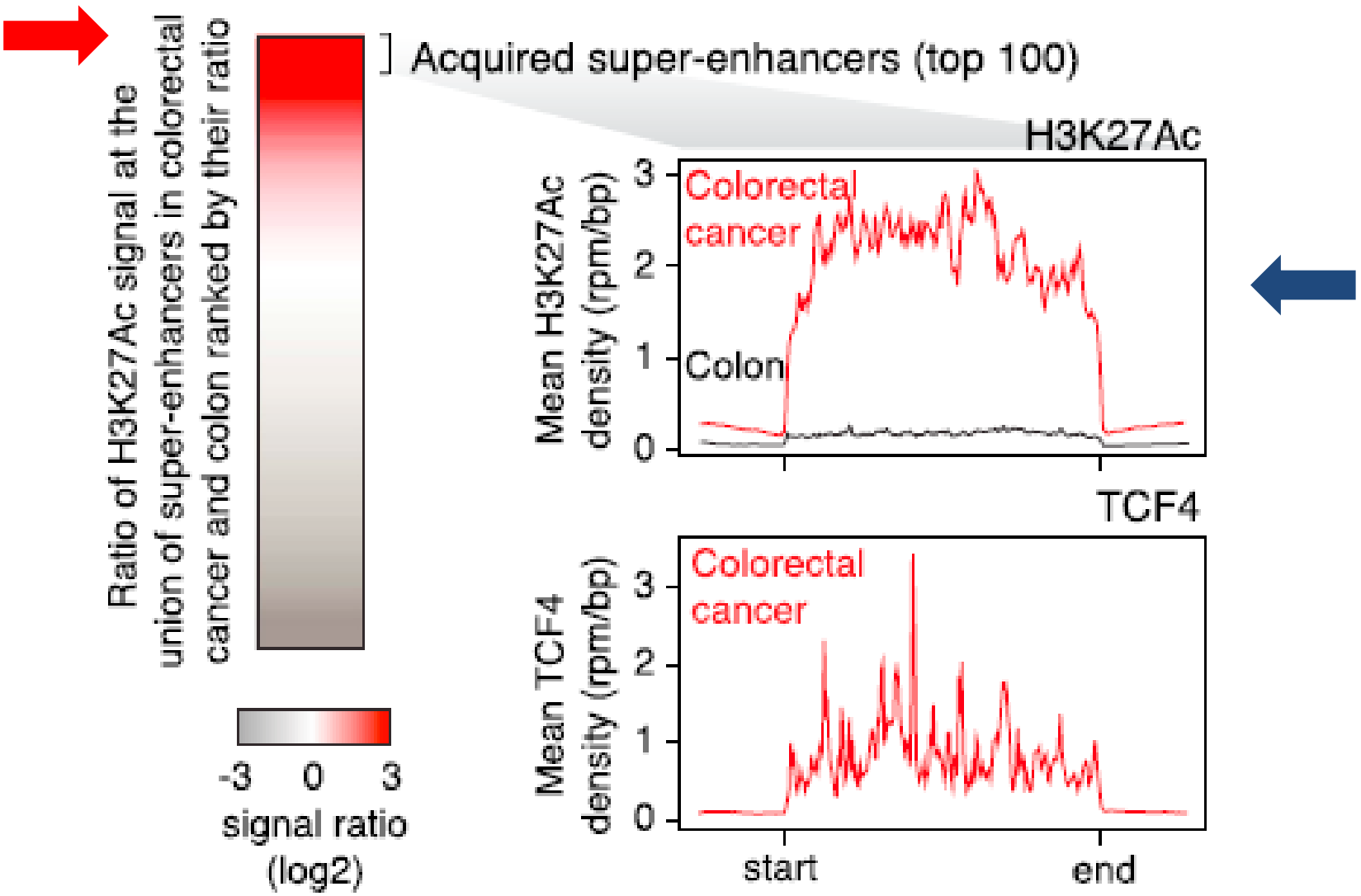
SE function on c-Myc locus



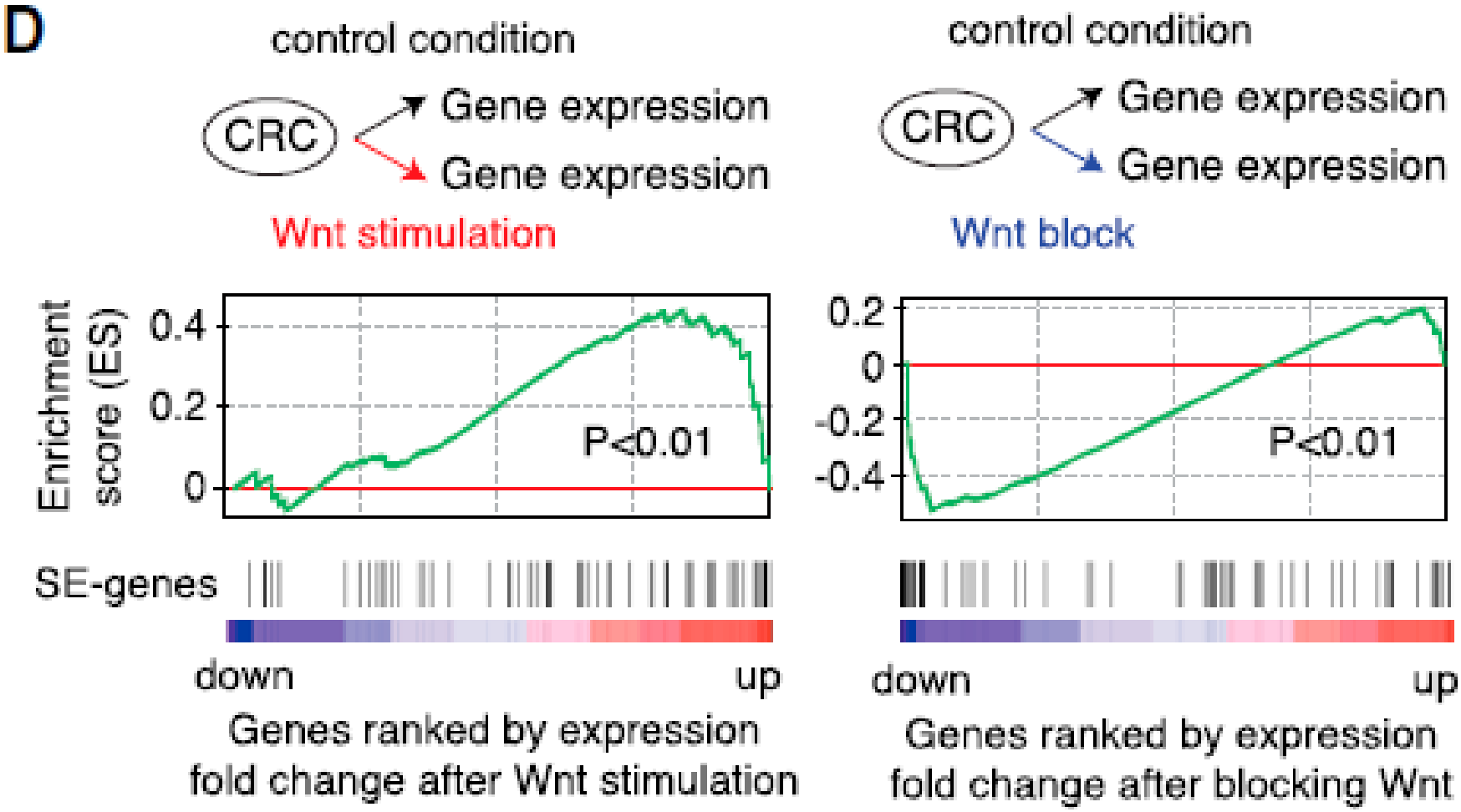
ChIP-seq binding profiles for H3K27Ac at the c-MYC locus in colon and colorectal cancer cells (HCT-116).

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.

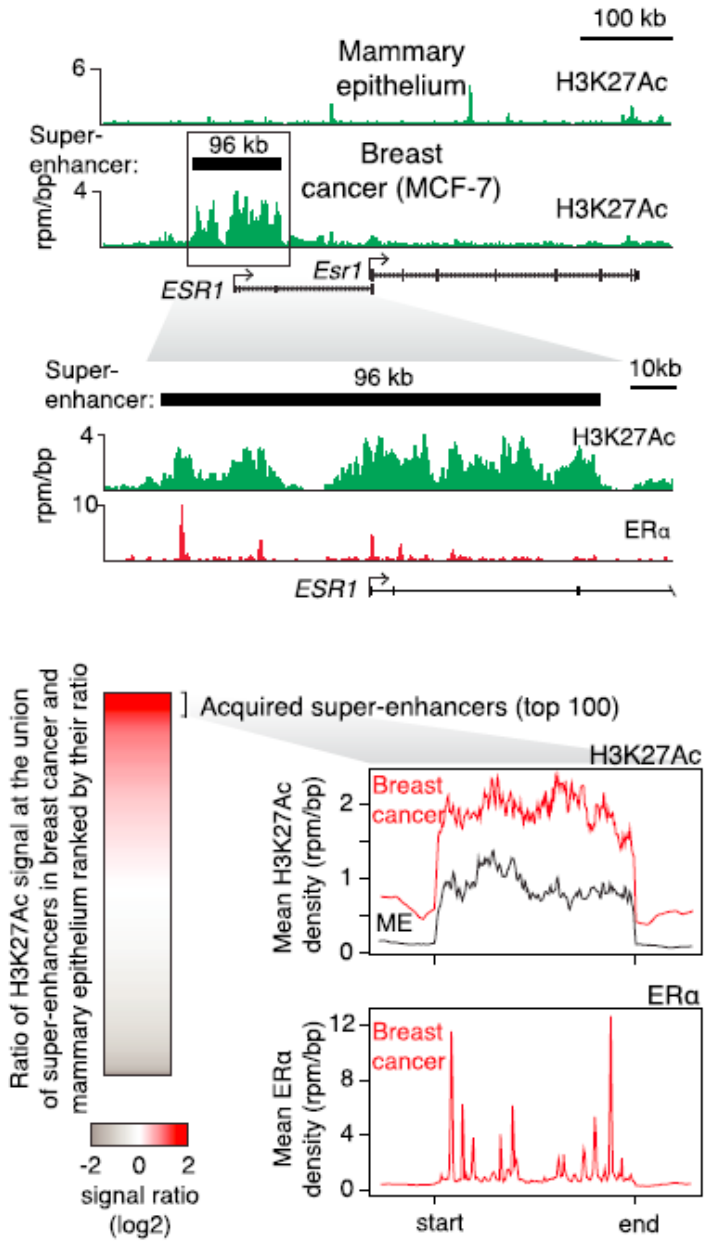
Left: ratio of H3K27Ac in CRC (HCT-116) versus normal colon tissue used densities at the union of SEs identified in the two samples. **Right:** metagene representation of H3K27Ac and TCF4 ChIP-seq densities at the regions corresponding to the top 100 acquired super-enhancers.



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.

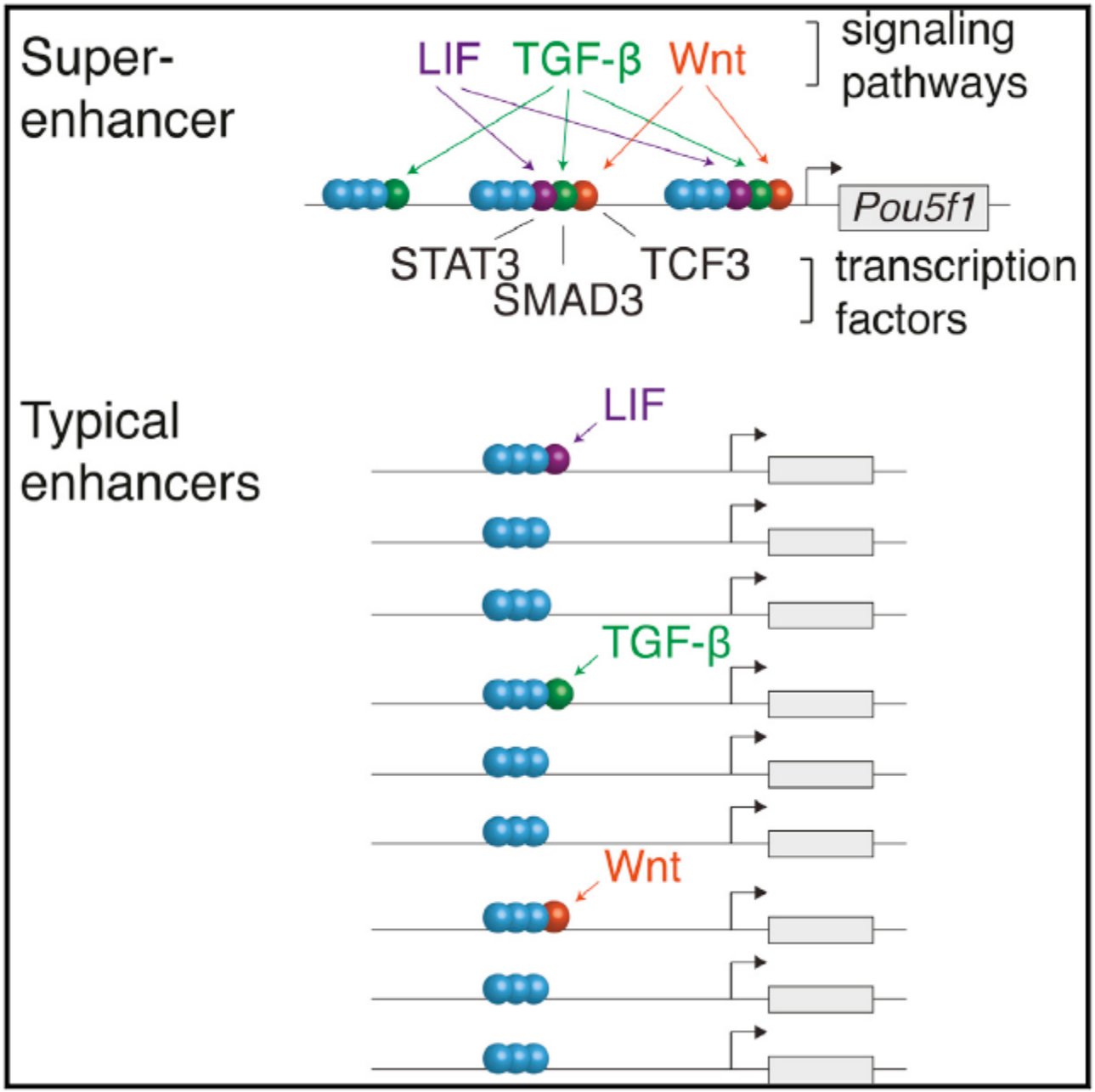
Several lines of evidence argue that the constituent enhancers of at least some super-enhancers can act as an interdependent structural and functional unit to control their associated genes. Our results show that ESC SEs generally consist of clusters of active enhancers that have OCT4-dependent and ESC-specific 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.

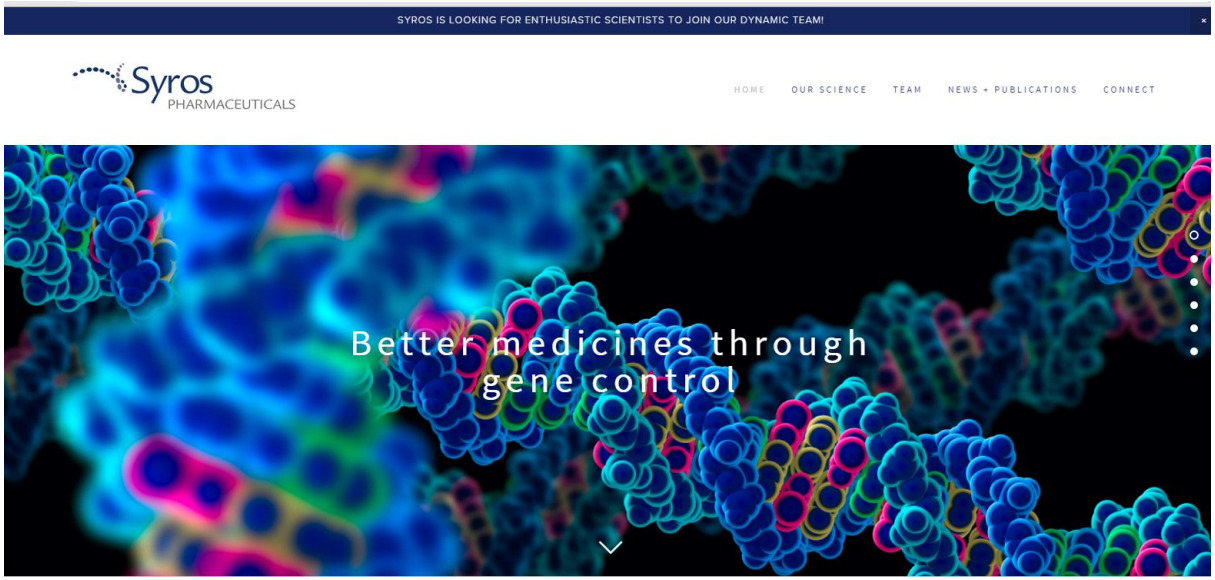
SEs characteristics

- **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

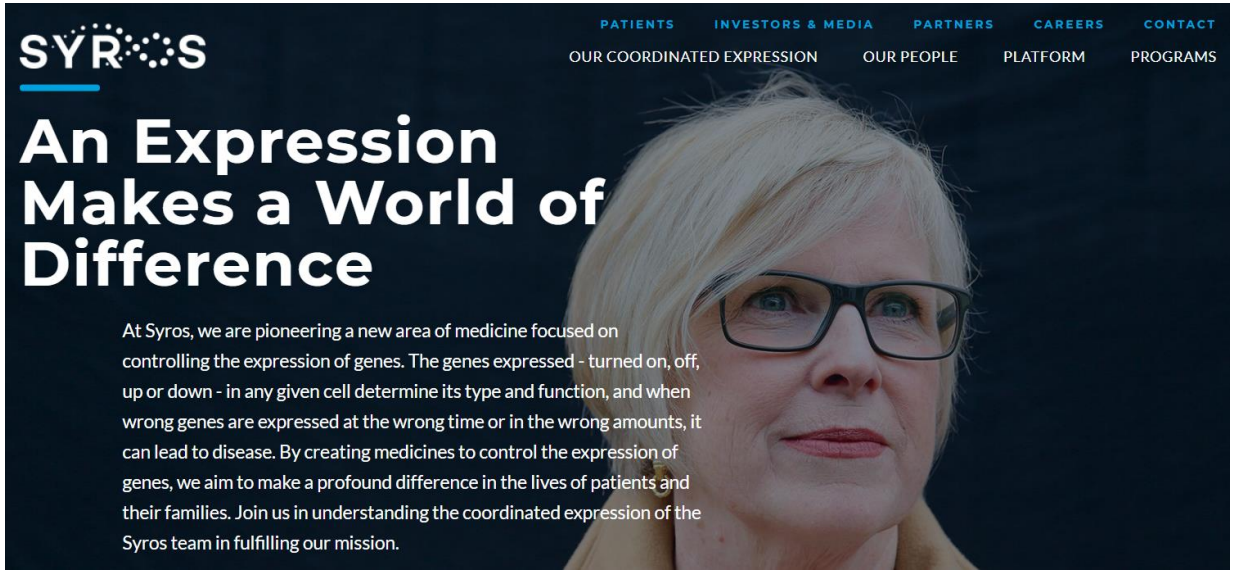




2015

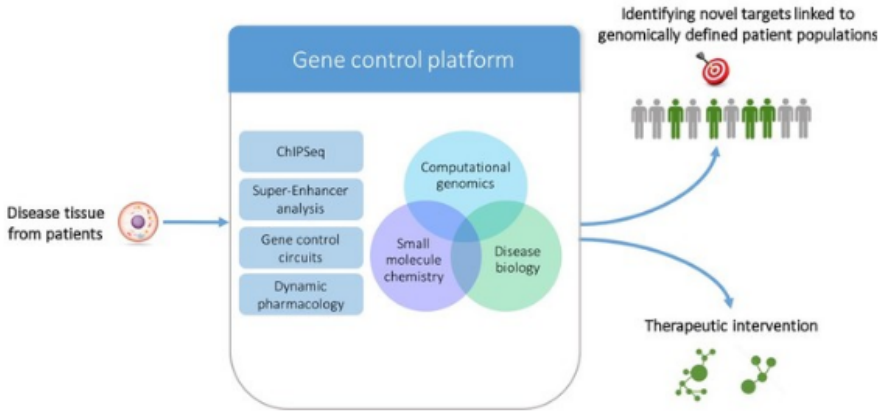
2019

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.

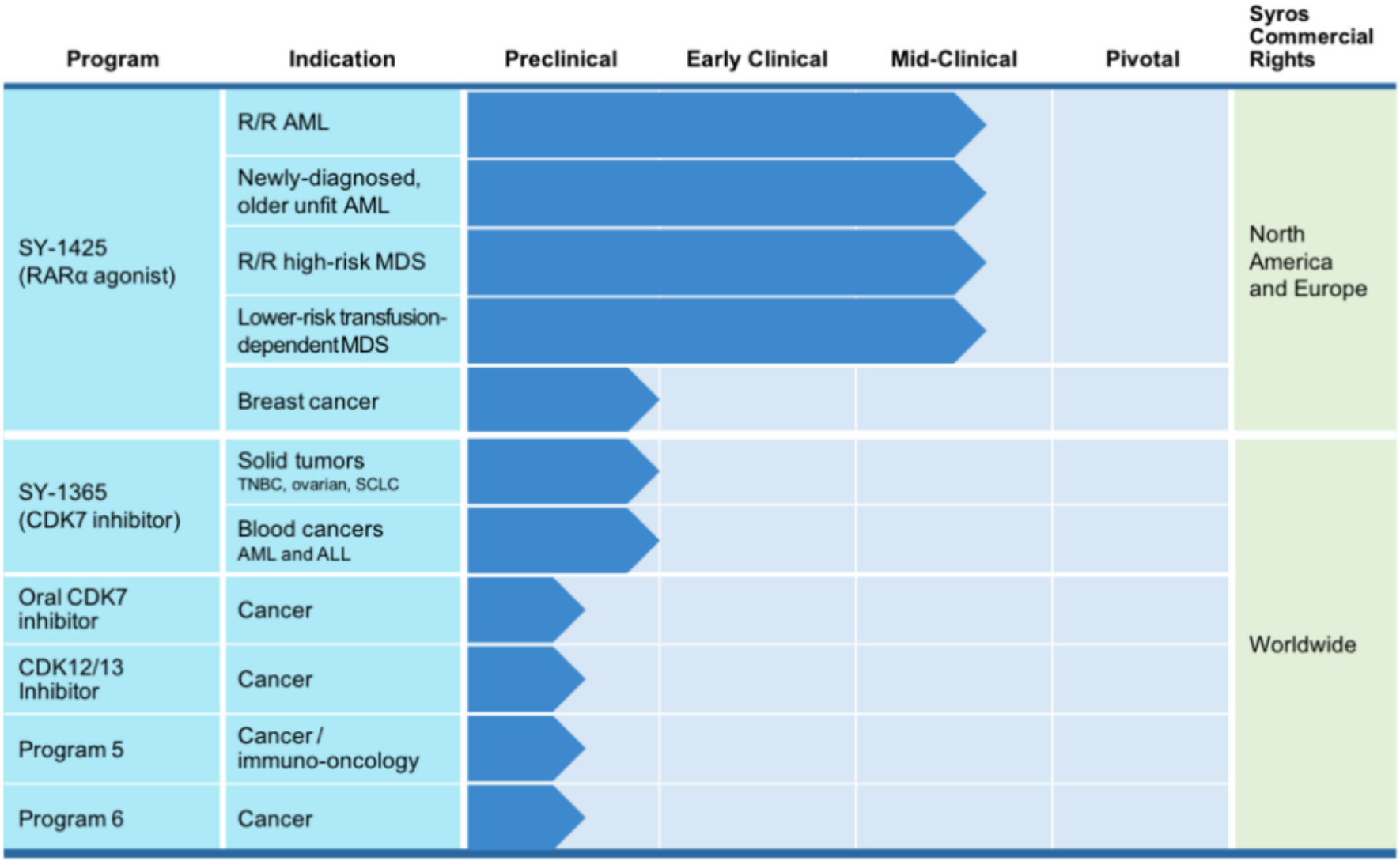


By doing so, we believe our gene control platform will allow us to:

- 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