Super-Enhancers in the Control of Cell Identity and Disease

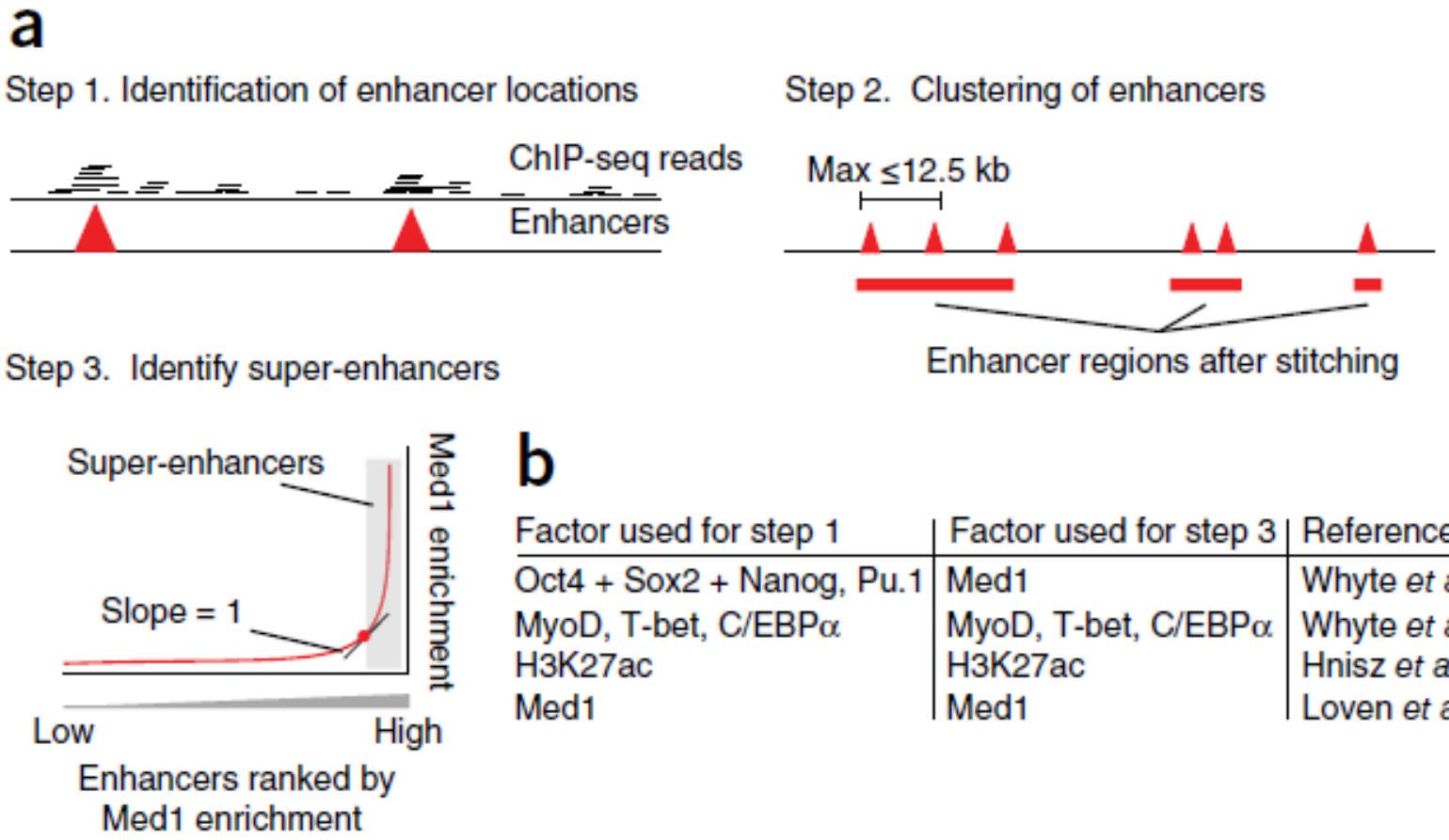
Denes Hnisz,[1,3] Brian J. Abraham,[1,3] Tong Ihn Lee,[1,3] Ashley Lau,[1,2] Violaine Saint-André,[1] Alla A. Sigova,[1] Heather A. Hoke,[1,2] and Richard A. Young[1,2,*]
[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts 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

Super-Enhancers are:

- large clusters of transcriptional enhancers formed by binding of master transcription factors
- are associated with genes that control and define cell identity
- are enriched for disease-associated DNA sequence variation
- are associated with oncogenes that become active during the process of tumor pathogenesis

Superenhancers are defined for the higher enrichment of signals extended to regions

SE are bound with master TFs, Med1 and are marked with H3K27Ac



b

Factor used for step 1	Factor used for step 3	Reference
Oct4 + Sox2 + Nanog, Pu.1	Med1	Whyte *et al.*
MyoD, T-bet, C/EBPα	MyoD, T-bet, C/EBPα	Whyte *et al.*
H3K27ac	H3K27ac	Hnisz *et al.*
Med1	Med1	Loven *et al.*

IDENTIFICATION AND CHARCTERIZATION

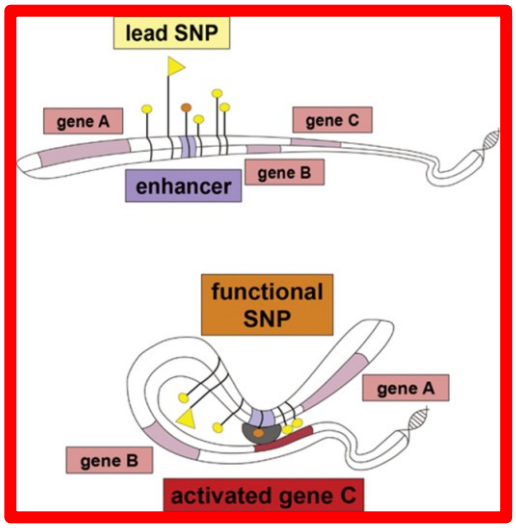
MOLECULAR MECHANISMS

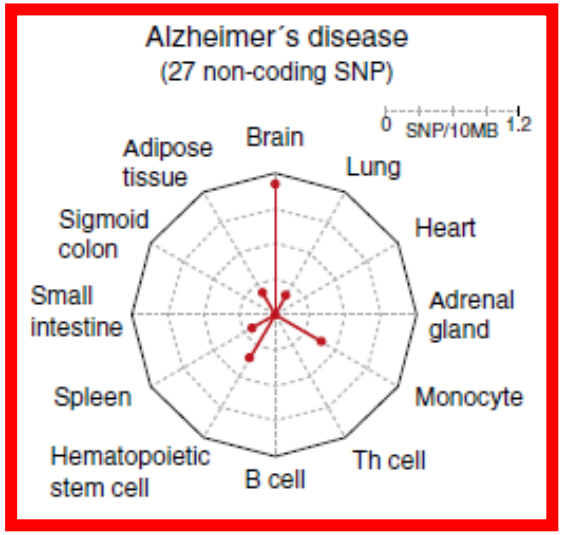
GENOMIC REGULATORY REGIONS

GENE EXPRESSION REGULATION

CELL IDENTITY

BIOLOGICAL FUNCTIONS





TO UNDERSTAND DISEASES

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

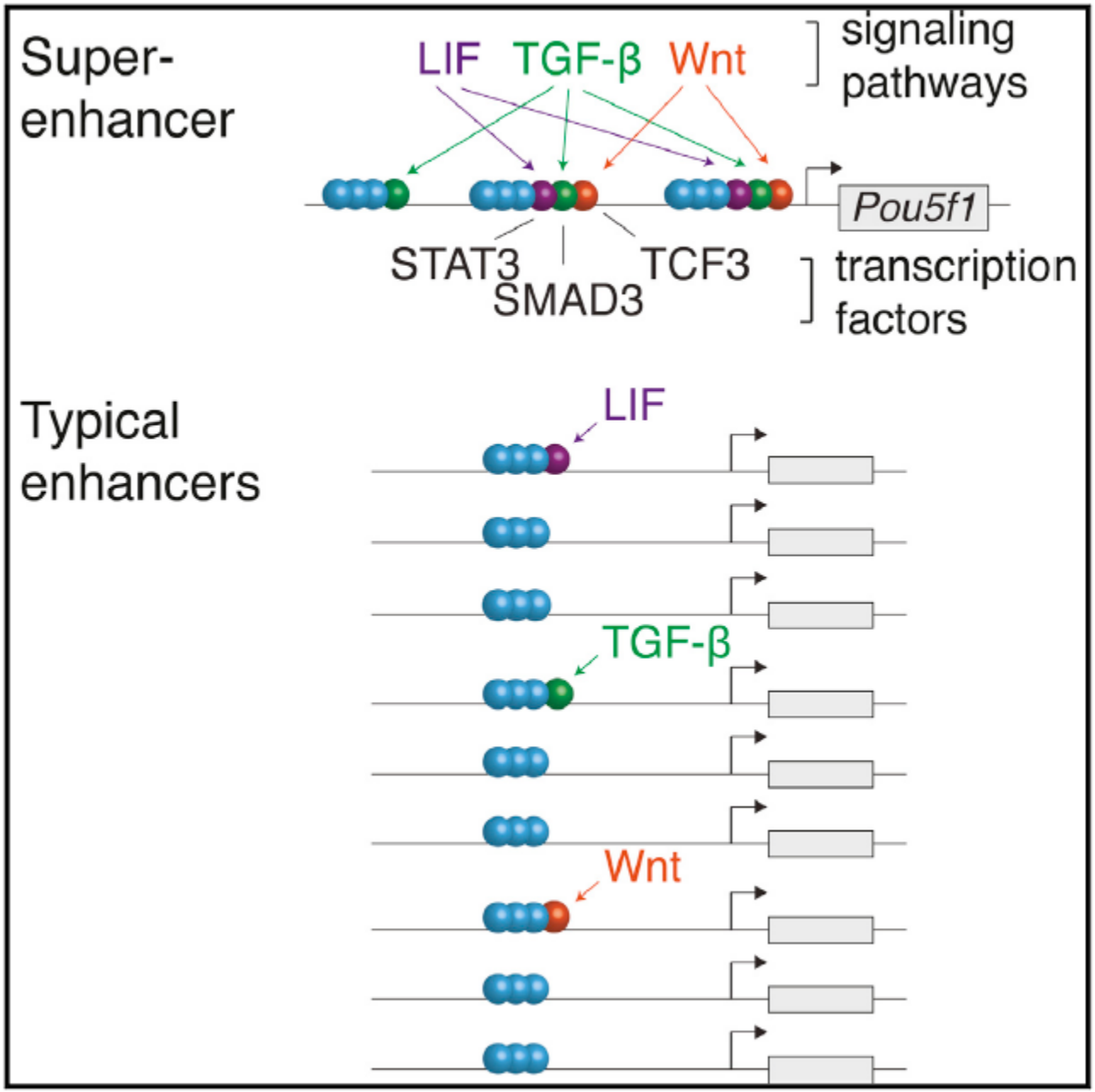
AIM

Super-enhancers (SE)
Functional costituent units

Cancer cells SE target
for oncogenic signalling

CONCLUSION

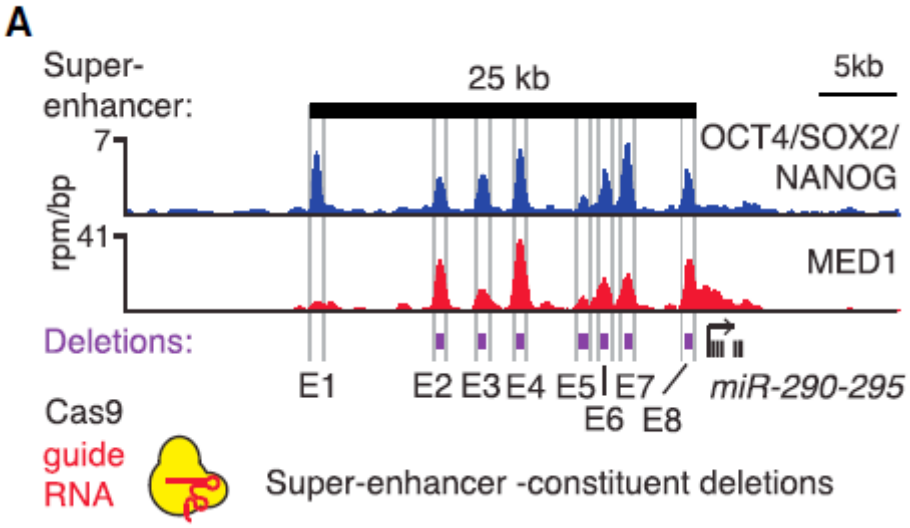
Graphical Abstract



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

METHOD: DELETION OF SPECIFIC GENOMIC REGIONS

TECHNIQUE: CRISPR/CAS9



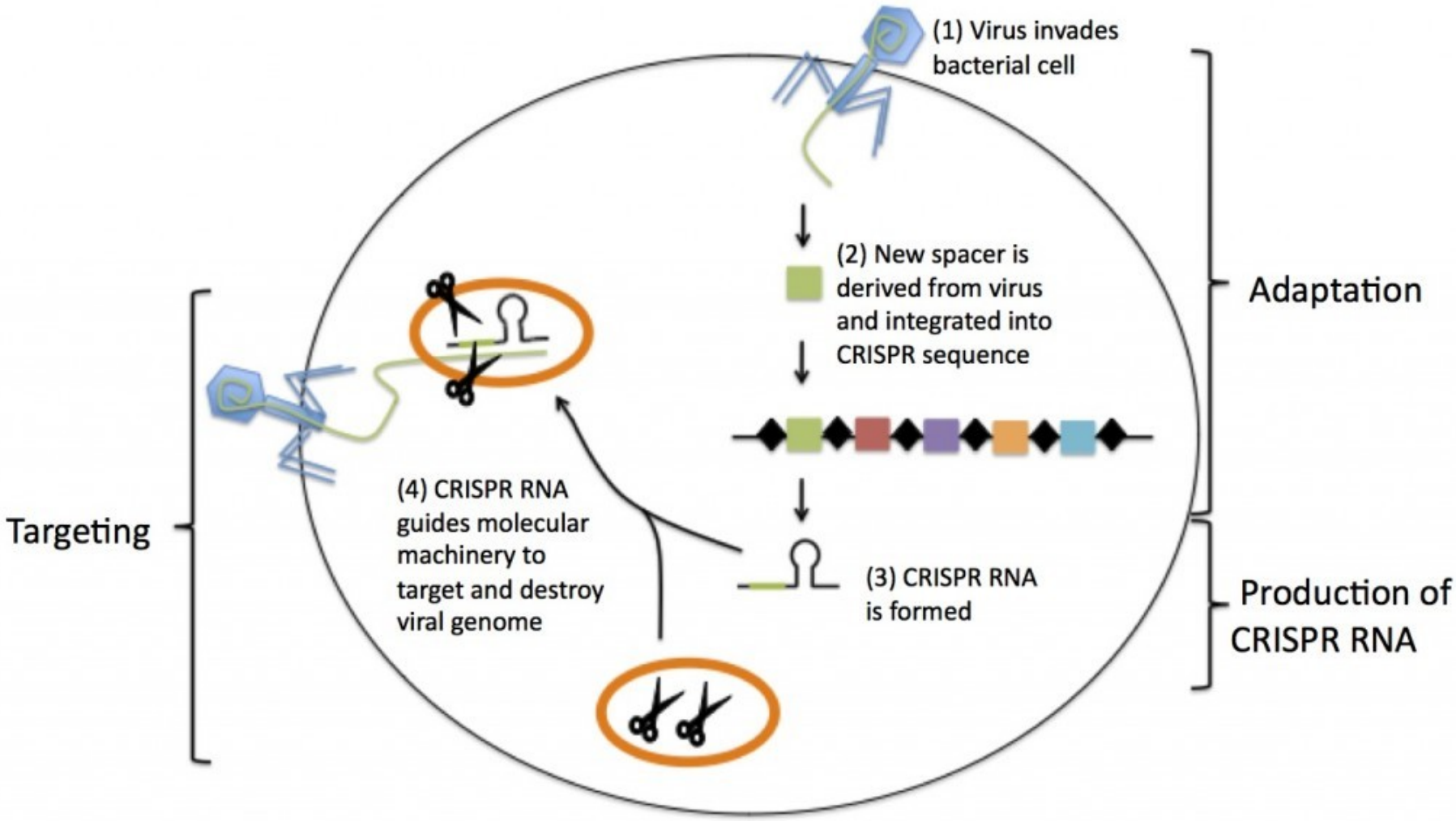
ChIP-Seq DATA used to design STUDY ON COSTITUENTS ENHANCERS FUNCTION

CRISP-Cas9

https://youtu.be/4YKFw2KZA5o

WHAT IS CRISP?

CRISPR is an acronym for **Clustered Regularly Interspaced Short Palindromic Repeat**. This name refers to the unique organization of short, partially palindromic repeated DNA sequences found in the genomes of **bacteria** and other microorganisms.



CRISPR is an acronym for **Clustered Regularly Interspaced Short Palindromic Repeat**. This name refers to the unique organization of short, partially palindromic repeated DNA sequences found in the genomes of **bacteria** and other microorganisms.

While seemingly innocuous, ***CRISPR sequences*** are a crucial ***component of the immune systems*** of these simple life forms. The immune system is responsible for protecting an organism's health and well-being. Just like us, bacterial cells can be invaded by viruses, which are small, infectious agents. ***If a viral infection threatens a bacterial cell, the CRISPR immune system can thwart the attack by destroying the genome of the invading virus***. The genome of the virus includes genetic material that is necessary for the virus to continue replicating. Thus, by destroying the viral genome, the CRISPR immune system protects bacteria from ongoing viral infection.

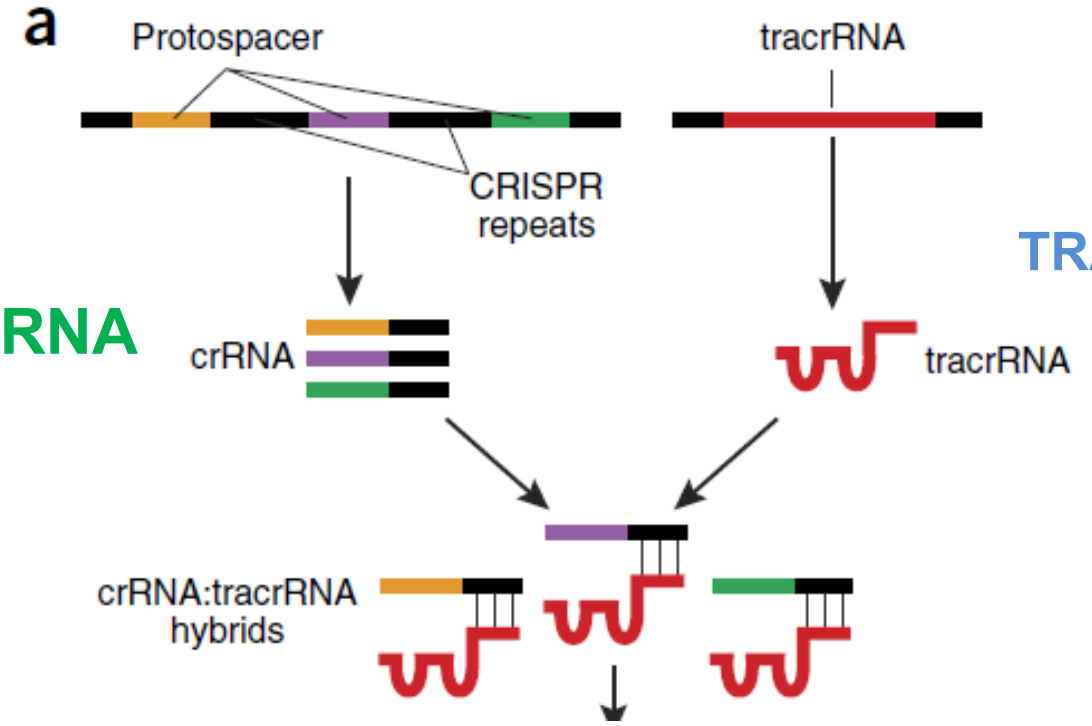
The CRISPR immune system works to protect bacteria from repeated viral attack via three basic steps [5]:

Step 1) **Adaptation** – **DNA from an invading virus** is processed into **short segments** that are inserted into the CRISPR sequence as new spacers.

Step 2) **Production of CRISPR RNA** – CRISPR repeats and spacers in the bacterial DNA undergo transcription, the process of **copying DNA into RNA** (ribonucleic acid). Unlike the double-chain helix structure of DNA, the resulting RNA is a single-chain molecule. **This RNA chain is cut into short pieces called CRISPR RNAs.**

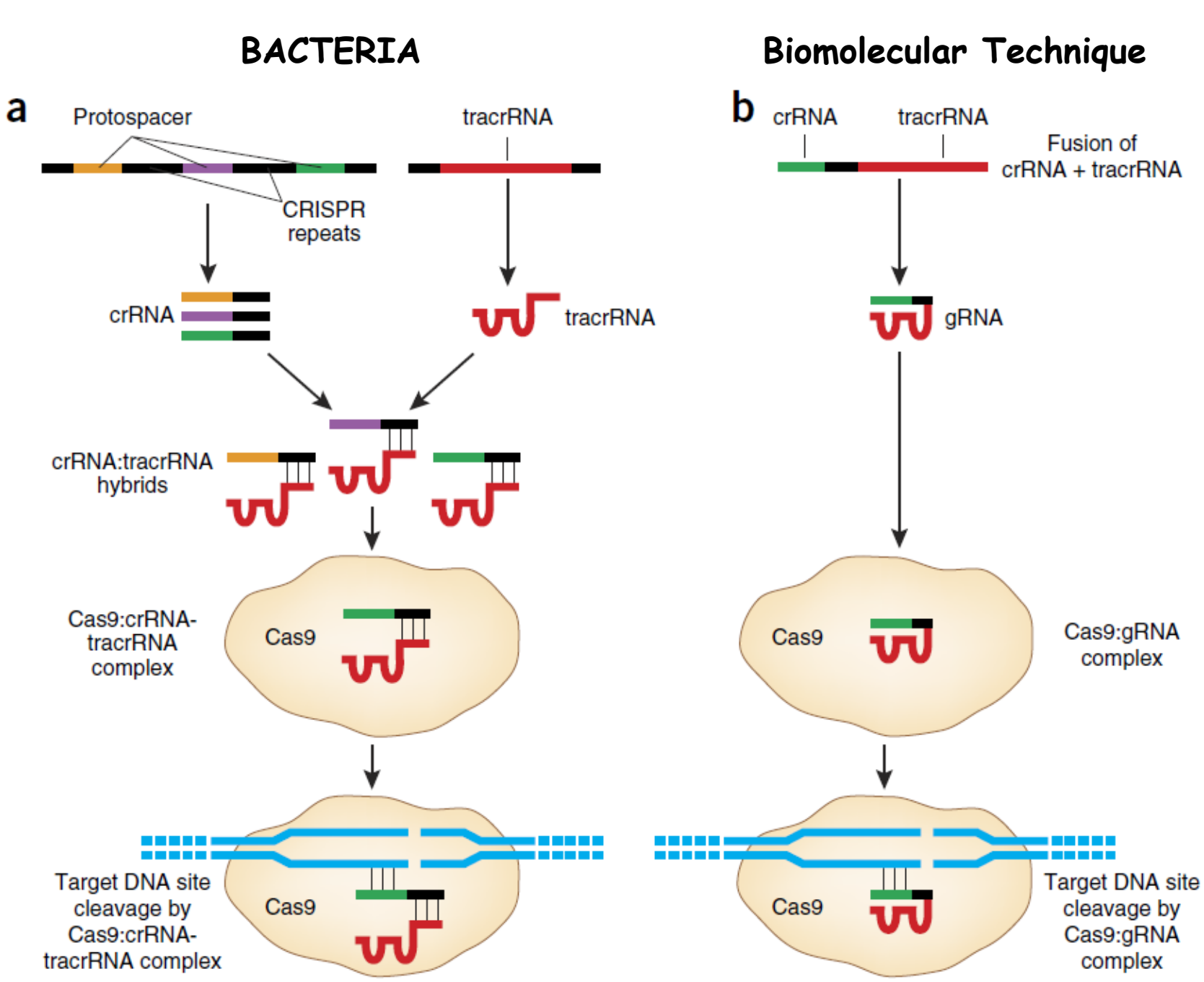
Step 3) **Targeting** – **CRISPR RNAs guide bacterial molecular machinery to destroy the viral material**. Because CRISPR RNA sequences are copied from the viral DNA sequences acquired during adaptation, they are exact matches to the viral genome and thus serve as excellent guides.

as arrays within the bacterial host genome (**Fig. 3a**). Transcripts from the CRISPR repeat arrays are processed into CRISPR RNAs (crRNAs), each harboring a variable sequence transcribed from the invading DNA, known as the "protospacer" sequence, and part of the CRISPR repeat. Each crRNA hybridizes with a second RNA, known as the transactivating CRISPR RNA (tracrRNA)[6], and these two RNAs complex with the Cas9 nuclease[7]. The protospacer-encoded portion

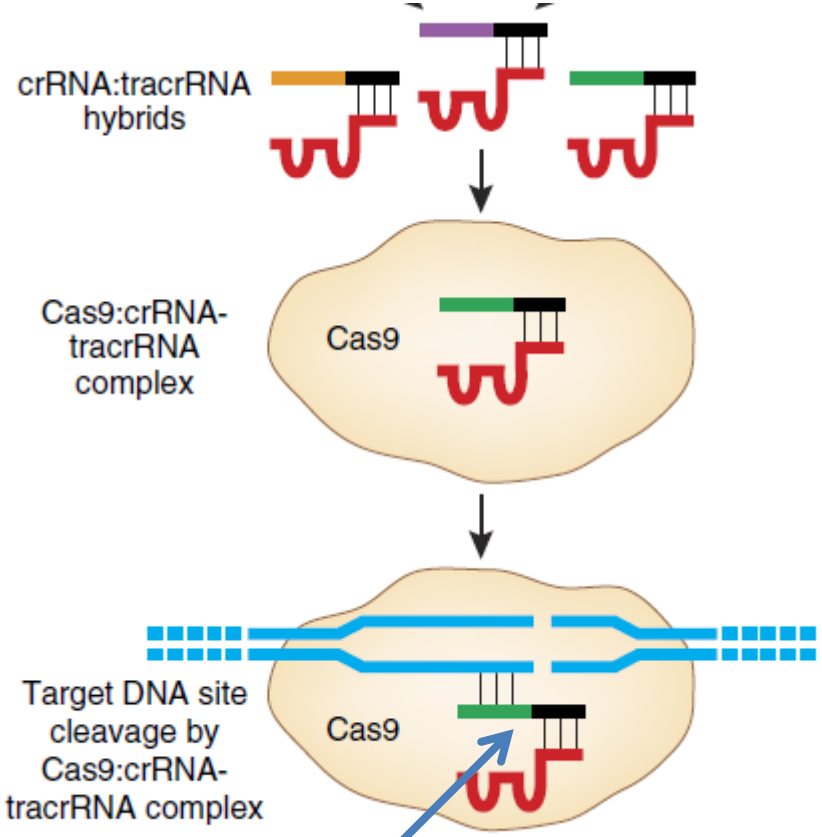


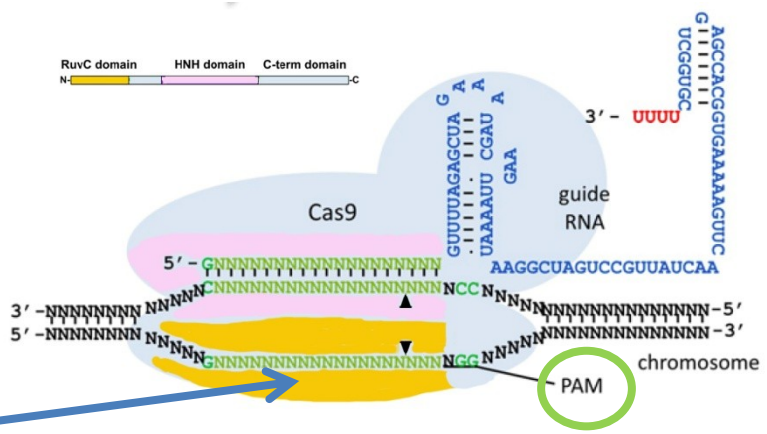
CRISPR RNA

TRANSACTIVATING CRISPR RNA


BACTERIA
Biomolecular Technique
a
Protospacer
tracrRNA
CRISPR repeats
crRNA
tracrRNA
crRNA:tracrRNA hybrids
Cas9:crRNA-tracrRNA complex
Cas9
Target DNA site cleavage by Cas9:crRNA-tracrRNA complex
Cas9
b
crRNA
tracrRNA
Fusion of crRNA + tracrRNA
gRNA
Cas9
Cas9:gRNA complex
Cas9
Target DNA site cleavage by Cas9:gRNA complex

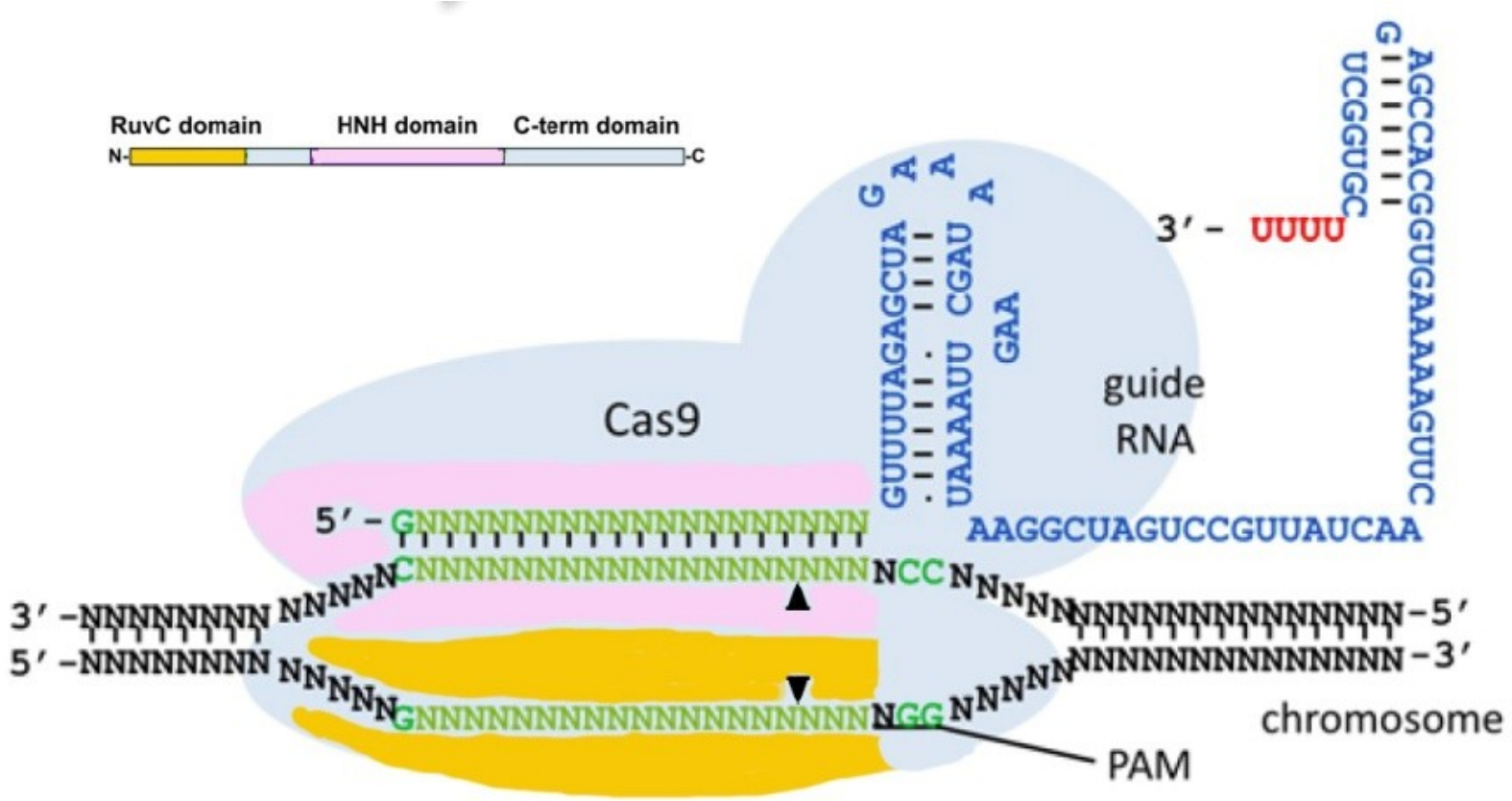
FORMATION OF COMPLEX
crRNA and trancrRNA with protein Cas9



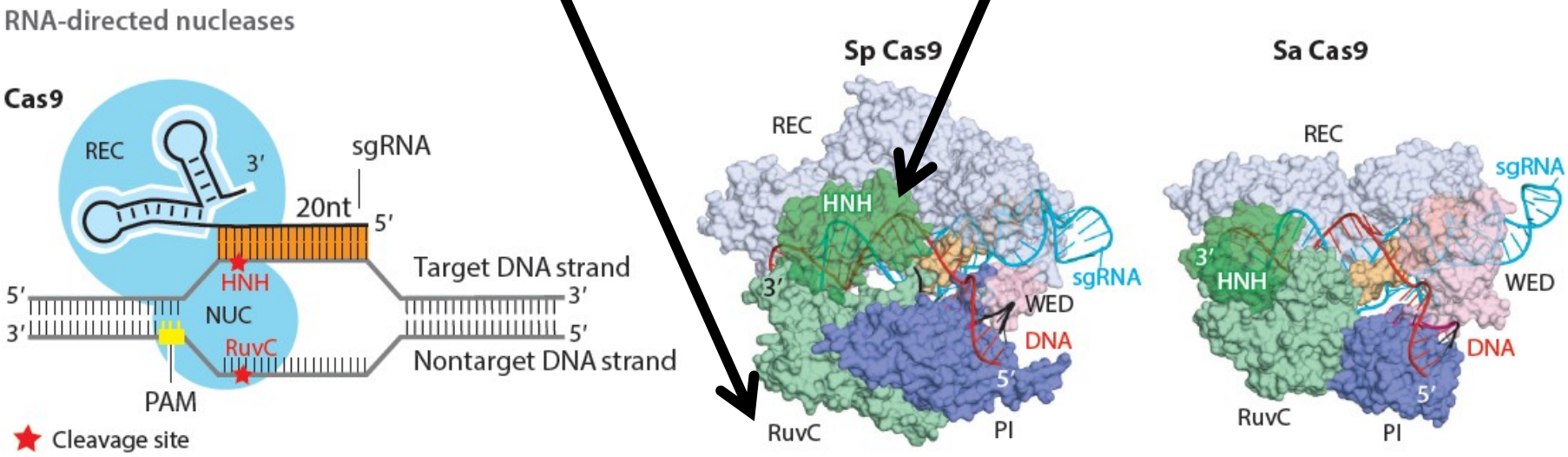


Protospacer of crRNA

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.

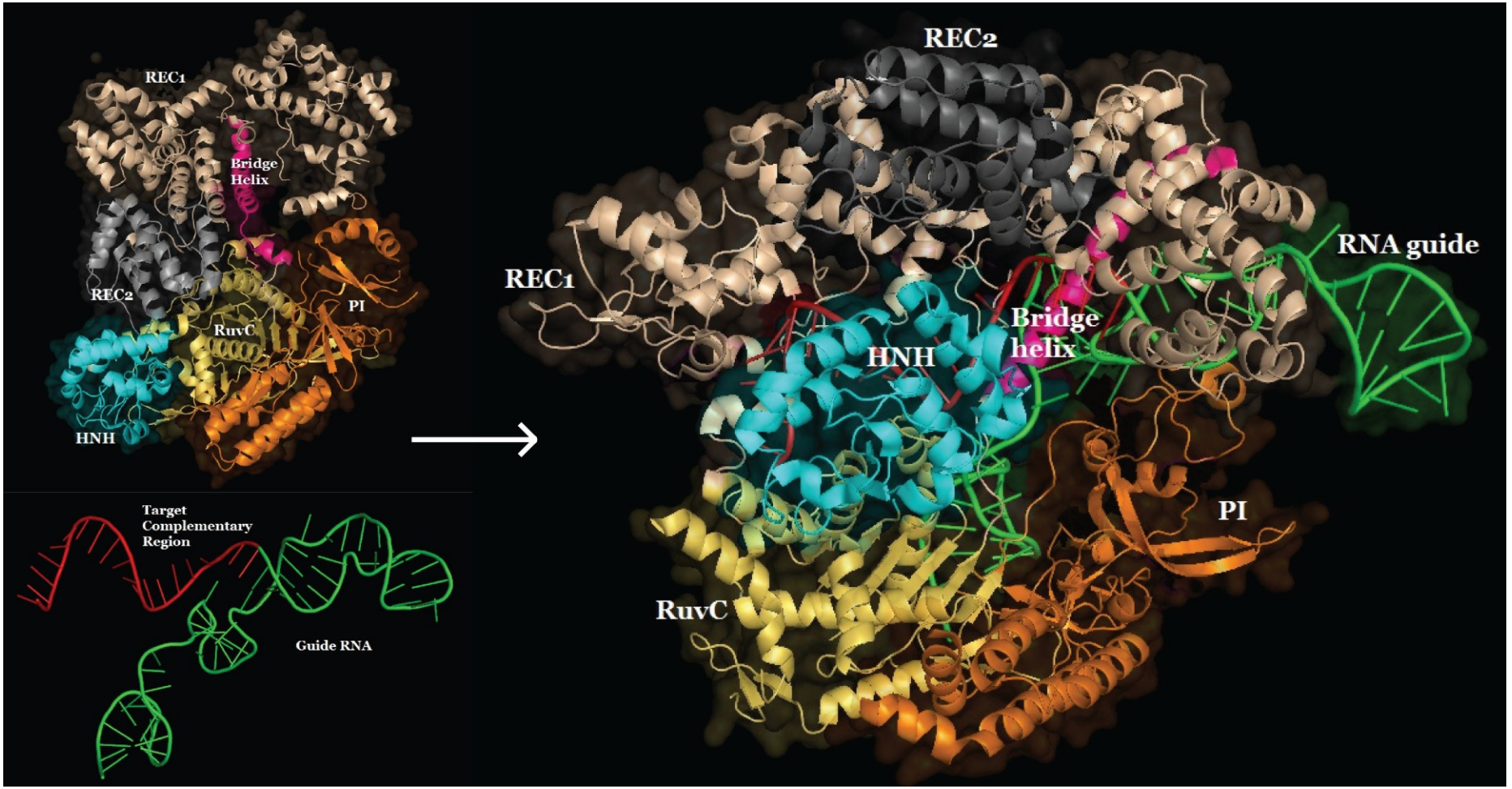


NUCLEASE DOMAINS

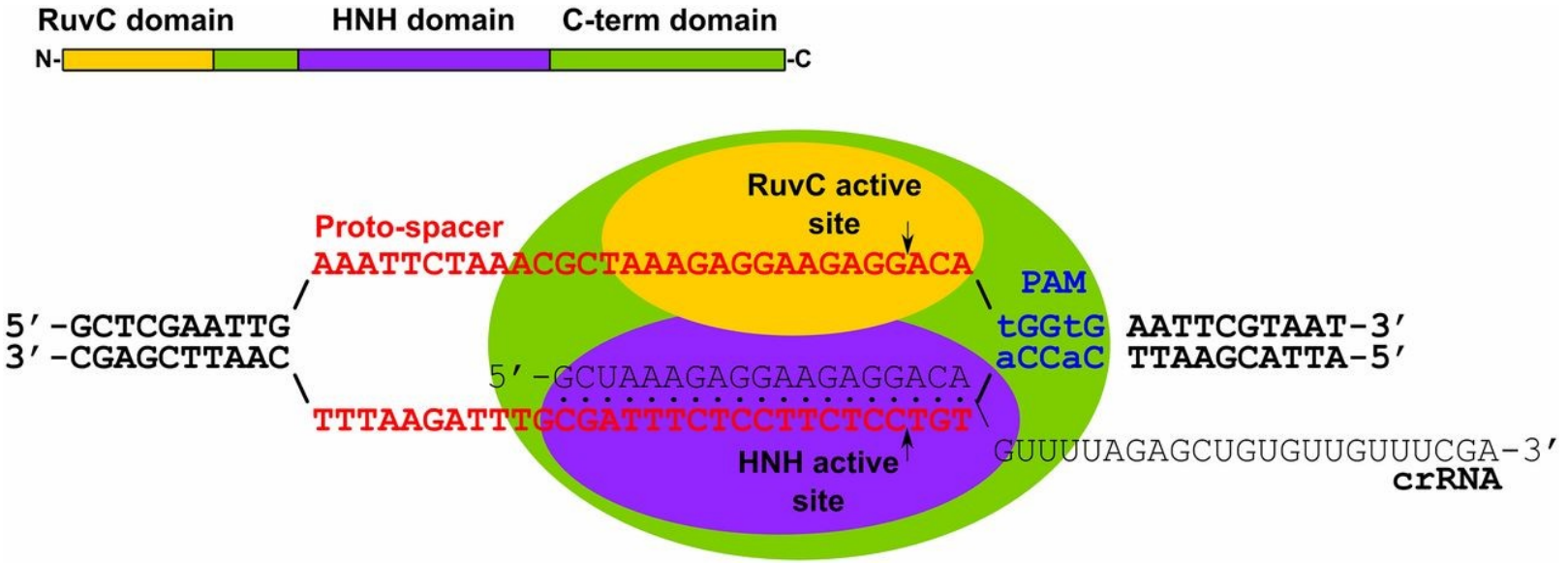


PAM-interacting domain- PI

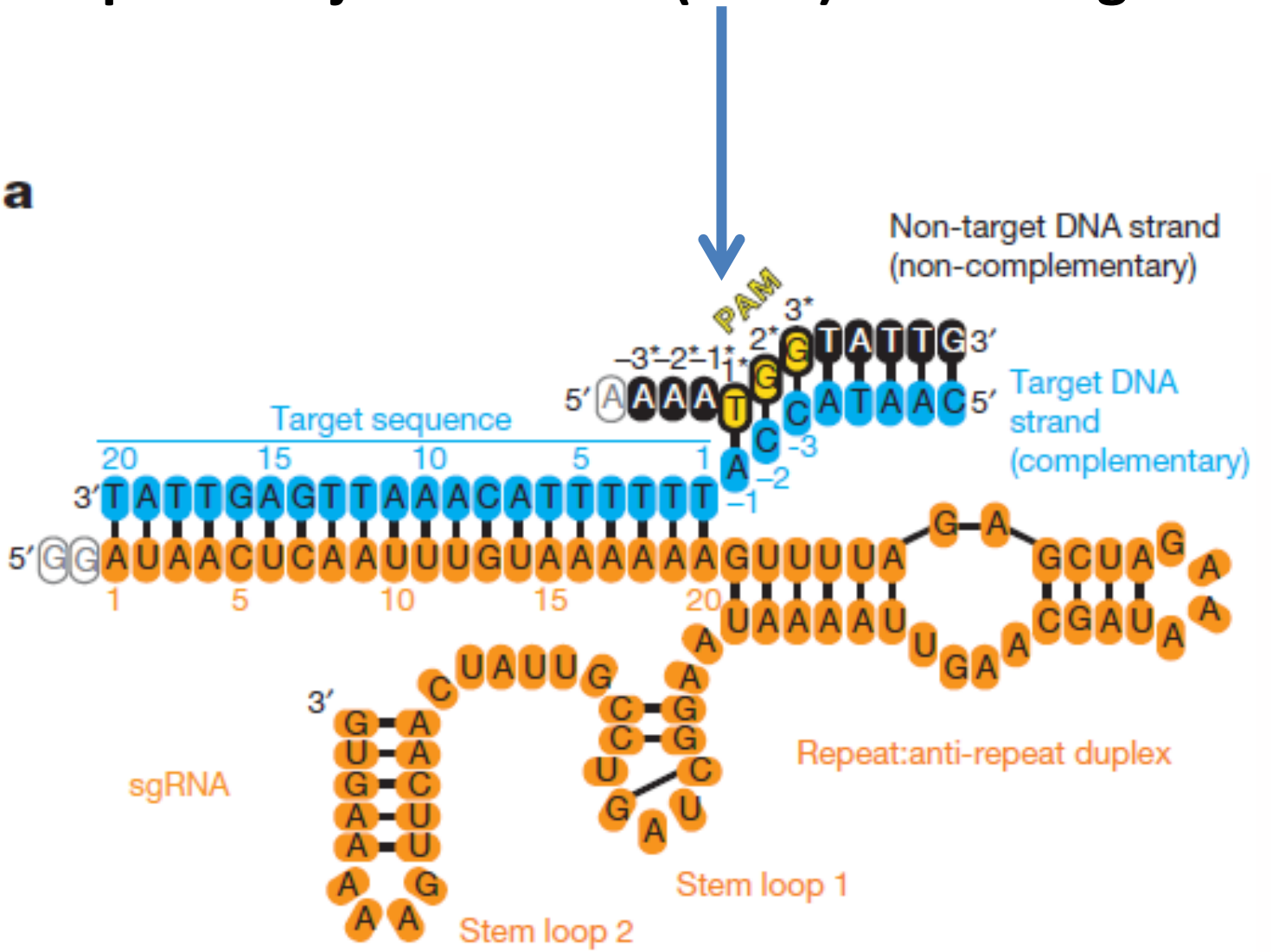
complementary strands, leading to a double-stranded break (DSB). Cas9 is a naturally evolved, RNA-guided nuclease that recognizes its target DNA through approximately 20 nucleotide (nt) base-pairing interaction between a single guide RNA (sgRNA) and its targeted DNA strand. Cas9 also interacts with the protospacer-adjacent motif (PAM) of its DNA target through its PAM-interacting (PI) domain at its C terminus. Cas9 uses its two nuclease domains (HNH and RuvC) to cleave the double-stranded DNA, creating DSBs. The HNH, RuvC, and PI domains, as well as an evolutionarily divergent wedge domain (WED), all reside in the Cas9 nuclease (NUC) lobe. The recognition (REC) lobe of Cas9 contains other regions that interact with the sgRNA–DNA duplex. (*Bottom right*) Crystal structures of Sp Cas9 and Sa Cas9. Crystal structures of *Streptococcus pyogenes* Cas9 (Sp Cas9; Protein Data Bank number 4UN3, 1368 AA) (84) and *Staphylococcus aureus* Cas9 (Sa Cas9 Protein Data Bank number 5CZZ, 1053 AA) (70) were obtained from RCSB Protein Data Bank (**http://www.rcsb.org/pdb/**), compared using PyMOL (PyMOL Molecular Graphics System, Version 1.3, Schrödinger LLC, **https://www.pymol.org/**), and domains are annotated according to References 70, 81, 83, 84. The orientation of the target DNA strand is also shown. (*b*) Cas9 in genomic editing. The DSB generated by Cas9 activates the nonhomologous end


REC1
Bridge Helix
REC2
PI
RuvC
HNH
Target Complementary Region
Guide RNA
REC2
REC1
RNA guide
Bridge helix
HNH
PI
RuvC

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

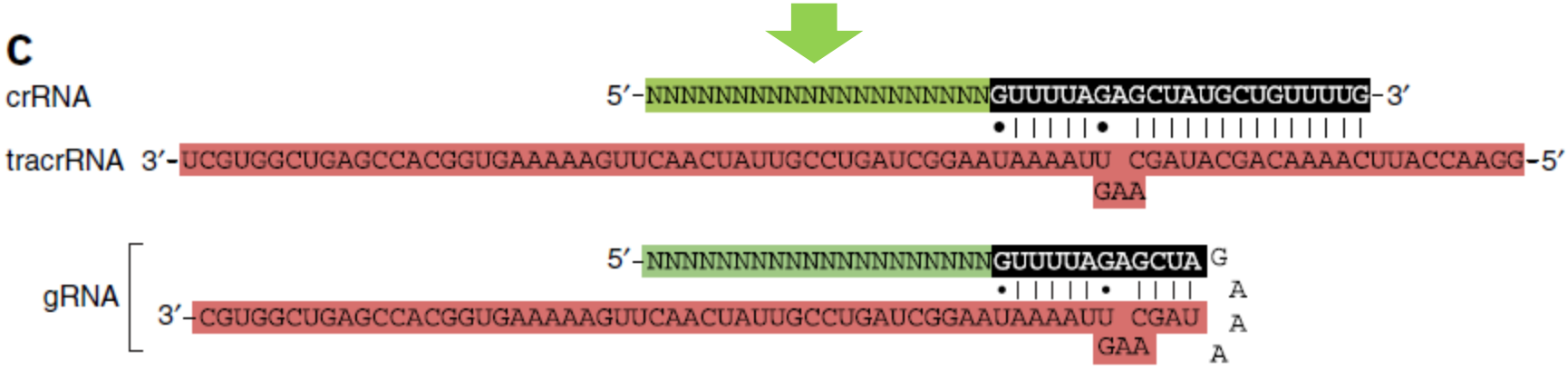


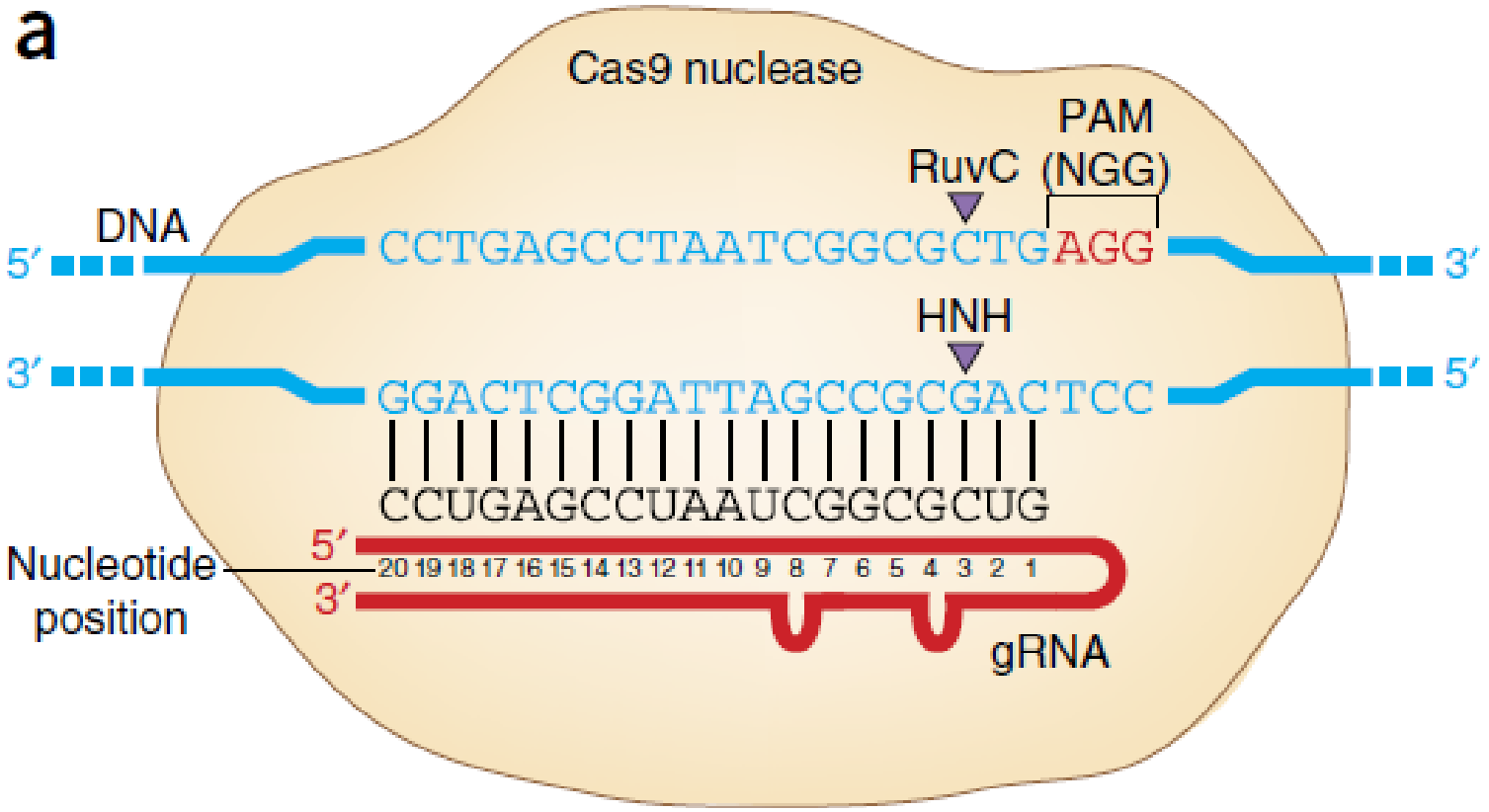
RNA-guided DNA recognition and cleavage strictly require the presence of a **protospacer adjacent motif (PAM) in the target DNA**.

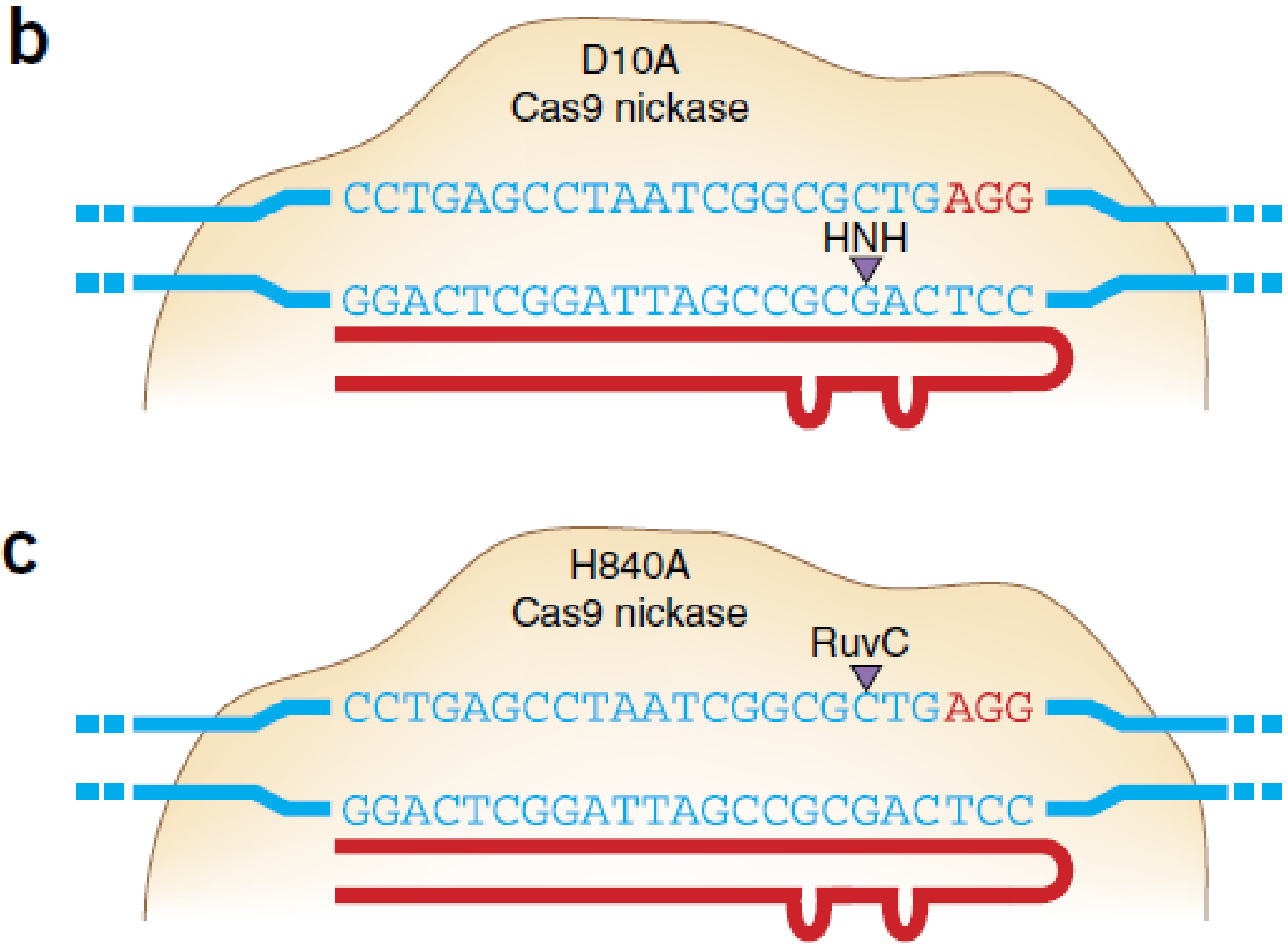


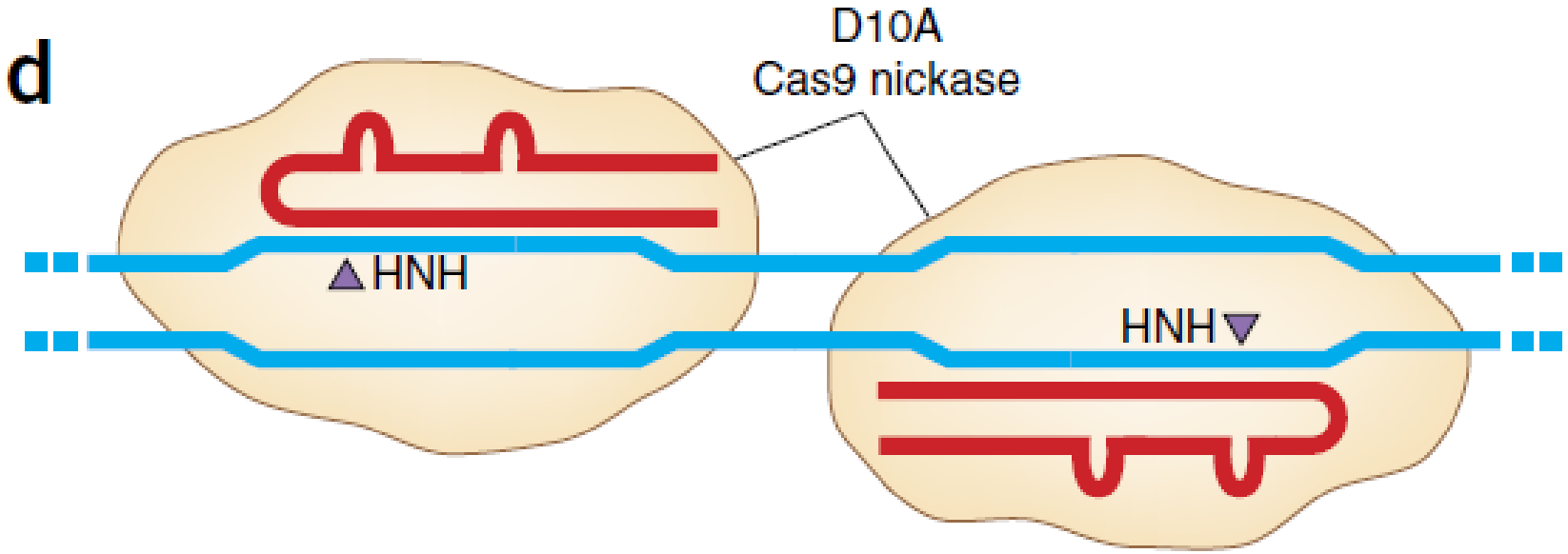
gRNA (guide RNA) is a single sequence of RNA formed with crRNA and tracrRNA

Complementary RNA sequence on target DNA genomic regions




a
Cas9 nuclease
PAM
(NGG)
RuvC
DNA
5′
CCTGAGCCTAATCGGCGCTGAGG
3′
HNH
3′
GGACTCGGATTAGCCGCGACTCC
5′
CCUGAGCCUAAUCGGCGCUG
Nucleotide
position
5′
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
3′
gRNA


b
D10A
Cas9 nickase
CCTGAGCCTAATCGGCGCTGAGG
HNH
GGACTCGGATTAGCCGCGACTCC
c
H840A
Cas9 nickase
RuvC
CCTGAGCCTAATCGGCGCTGAGG
GGACTCGGATTAGCCGCGACTCC


d
D10A
Cas9 nickase
HNH
HNH

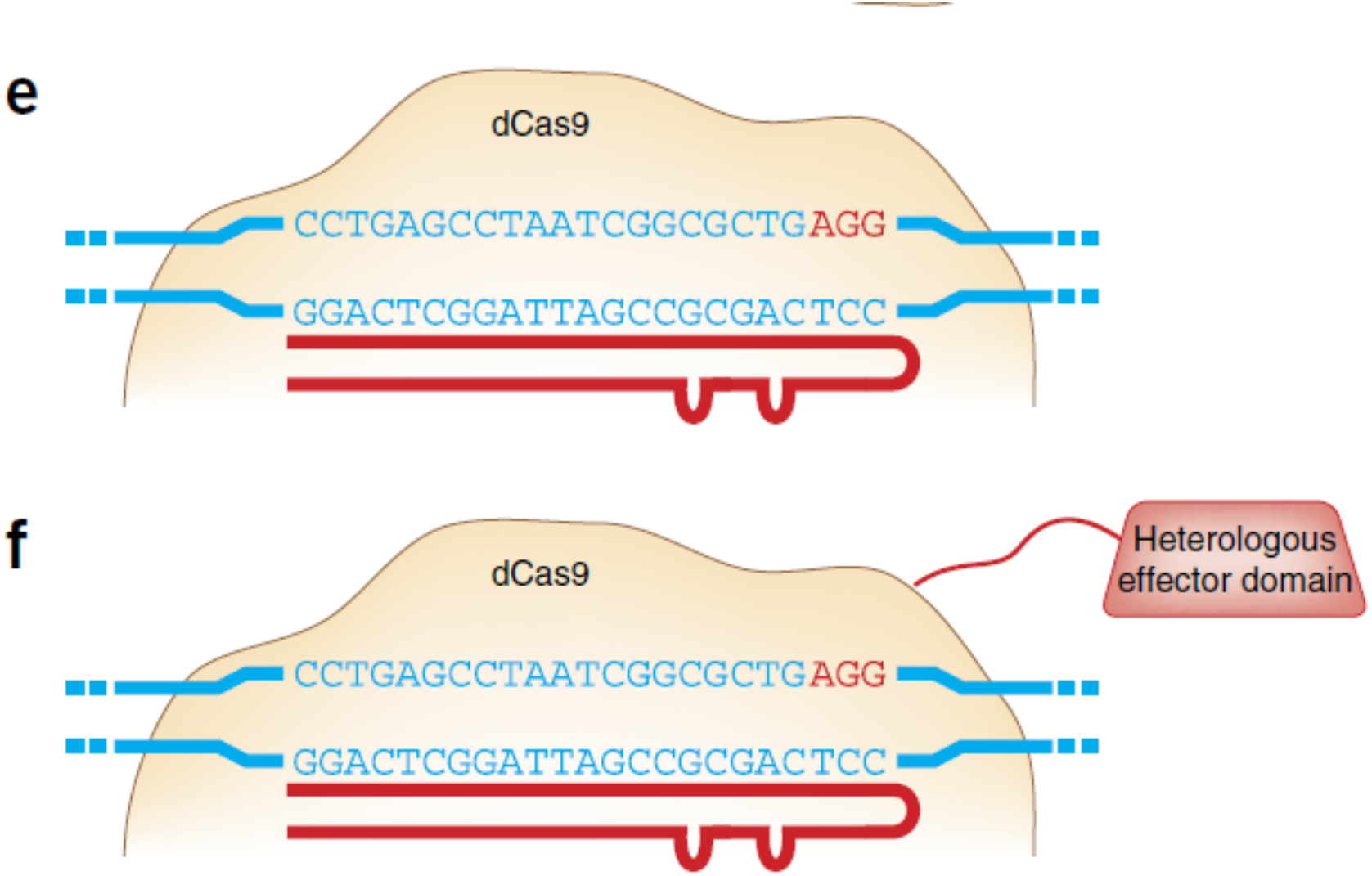
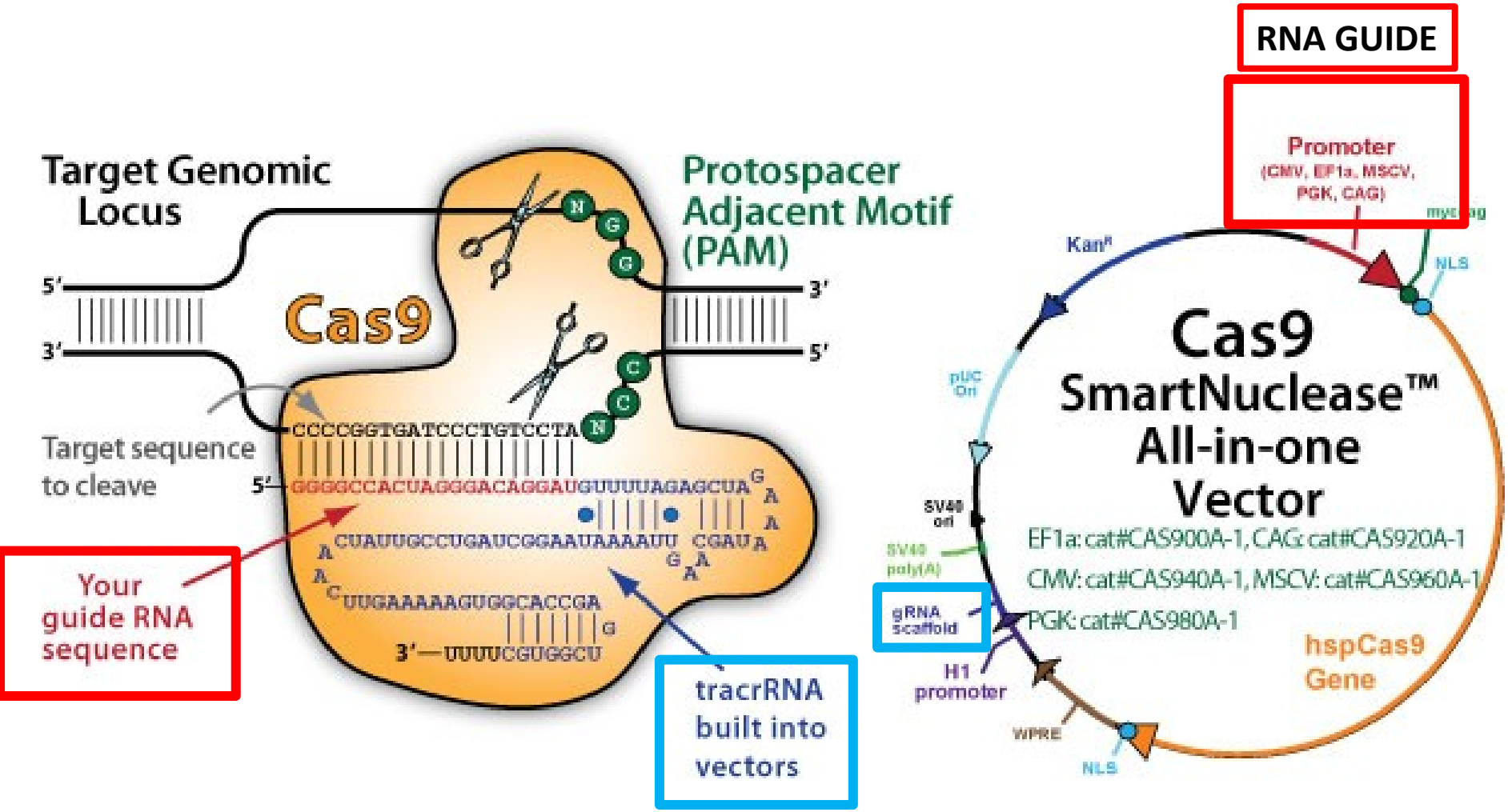

e
dCas9
CCTGAGCCTAATCGGCGCTGAGG
GGACTCGGATTAGCCGCGACTCC
f
dCas9
Heterologous
effector domain
CCTGAGCCTAATCGGCGCTGAGG
GGACTCGGATTAGCCGCGACTCC

Table 1 Published examples of cell types and organisms modified by Cas9

Cell type or organism	Cas9 form	Cell type	Reference numbers
Human cells	Cas9 nuclease	HEK293FT, HEK293T, HEK293, K562, iPSC, HUES9, HUES1, BJ-RiPS, HeLa, Jurkat, U2OS	9,13–16,47, 49–51,54,59, 84,85
	Cas9 nickase	HEK293FT, HEK293T	13,14,47,49
	dCas9 (gene regulation)	HEK293FT, HEK293T	70–72,74,82
	dCas9 (imaging)	HEK293T, UMUC3, HeLa	81
Mouse or mouse cells	Cas9 nuclease	Embryos	14,24–26
	Cas9 nickase	Embryos	47
	dCas9 (gene regulation)	NIH3T3	74
Rat	Cas9 nuclease	Embryos	26,36
Rabbit	Cas9 nuclease	Embryos	27
Frog	Cas9 nuclease	Embryos	28
Zebrafish	Cas9 nuclease	Embryos	17,33,37,60,85
Fruit fly	Cas9 nuclease	Embryos	29,30,61
Silkworm	Cas9 nuclease	Embryos	31
Roundworm	Cas9 nuclease	Adult gonads	32,62–67
Rice	Cas9 nuclease	Protoplasts, callus cells	21,23
Wheat	Cas9 nuclease	Protoplasts	21
Sorghum	Cas9 nuclease	Embryos	23
Tobacco	Cas9 nuclease	Protoplasts, leaf tissue	19,20,23
Thale cress	Cas9 nuclease	Protoplasts, seedlings	19,23
Yeast	Cas9 nuclease	*Saccharomyces cerevisiae*	18
Bacteria	Cas9 nuclease	*Streptococcus pneumoniae*, *E. coli*	8
	dCas9 (gene regulation)	*E. coli*	69,70

HEK, human embryonic kidney; iPSCs, induced pluripotent stem cells; UMUC3, human bladder cancer.

CRISPR-CAS9 system: vector

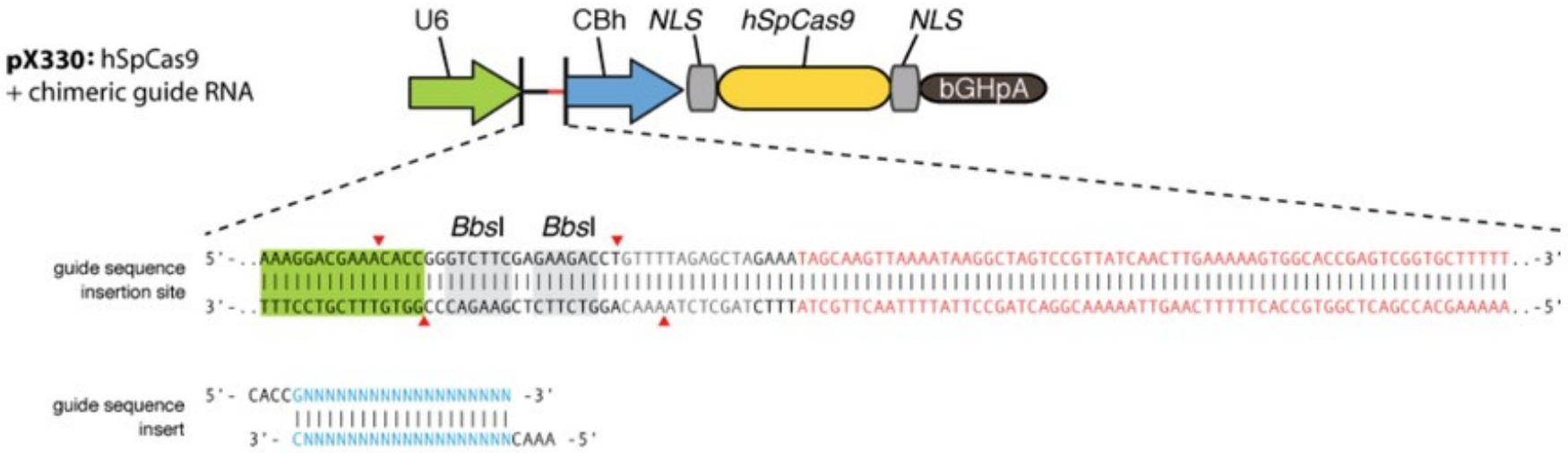


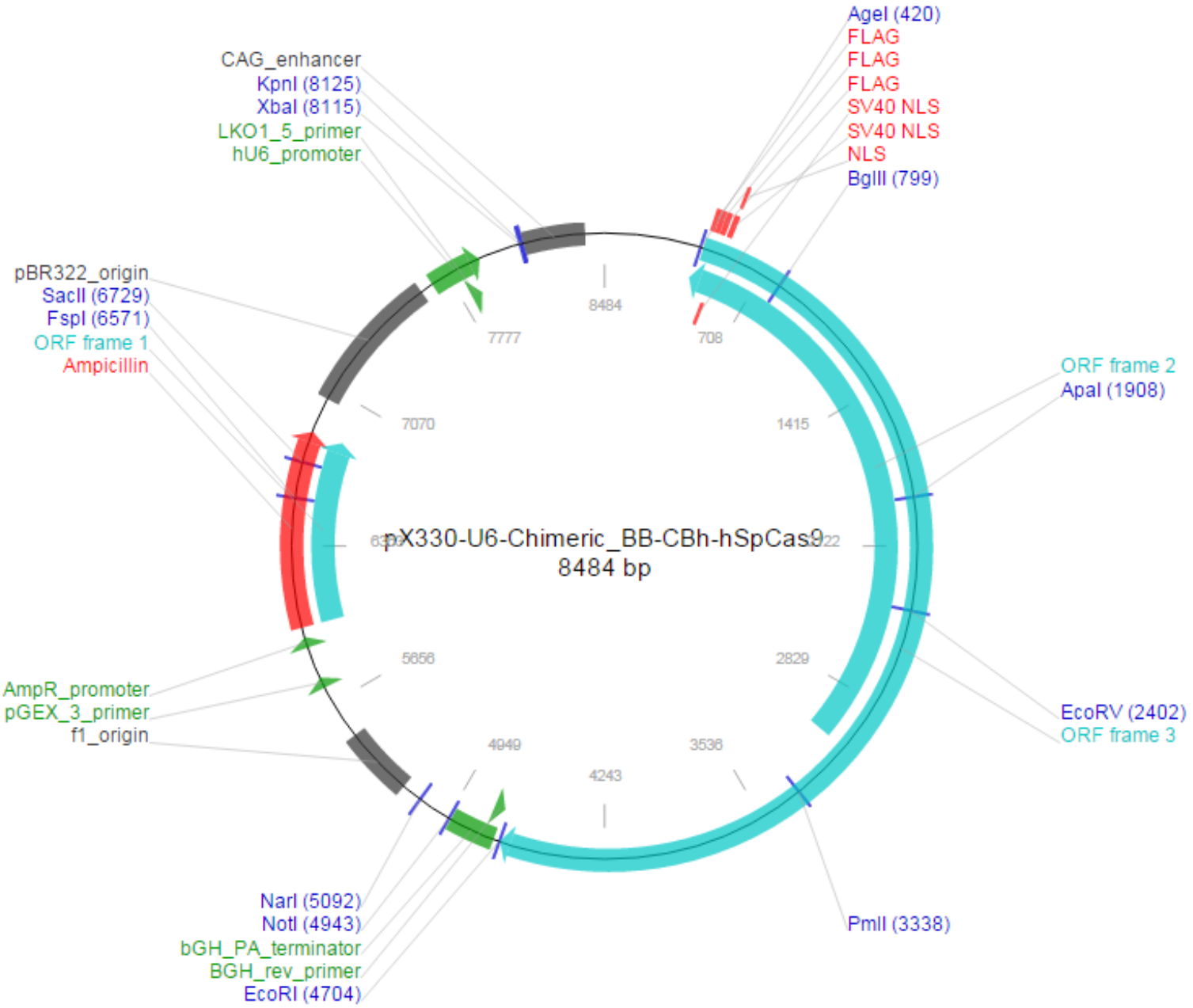
Genome Editing

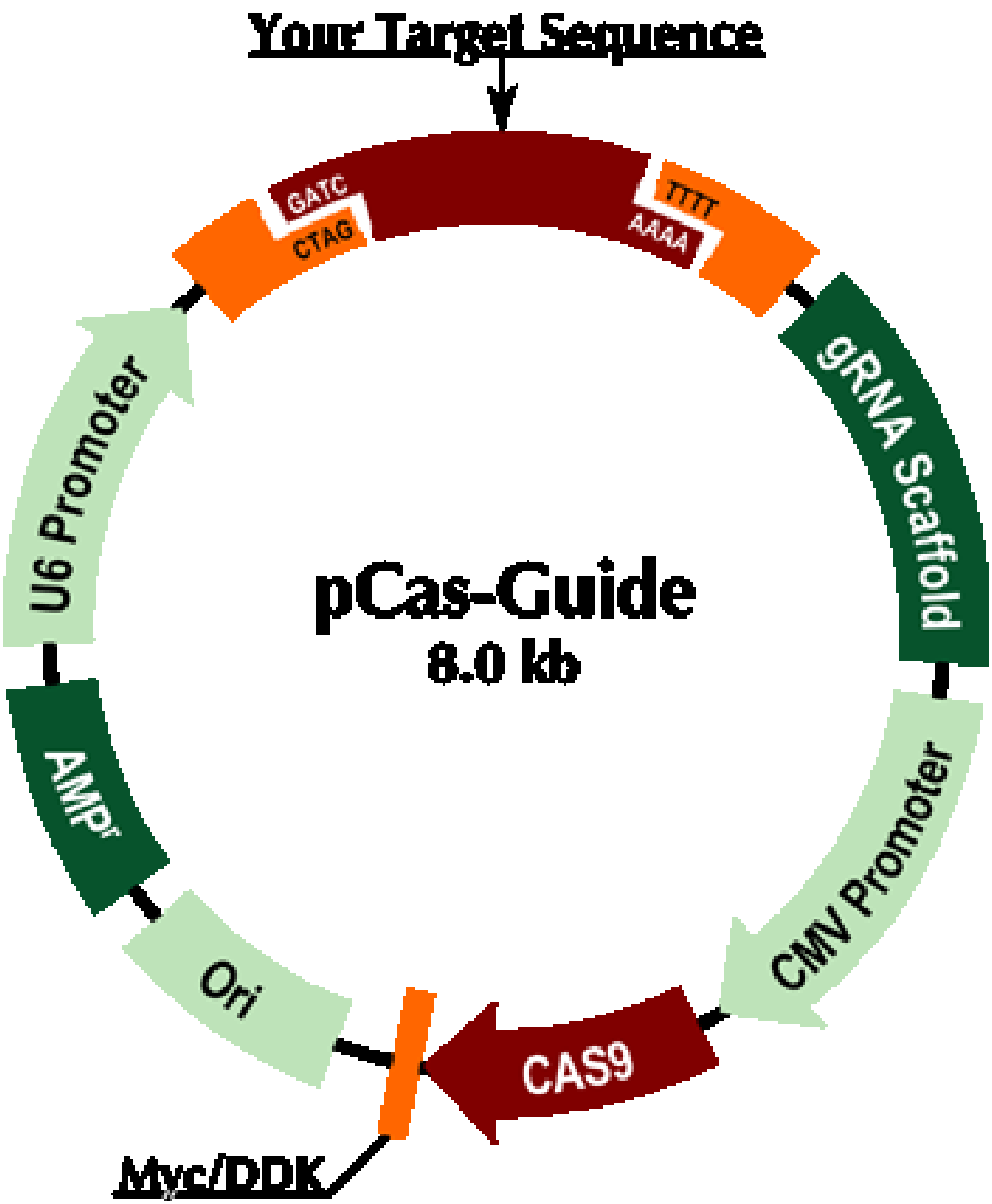
Super-enhancer constituent enhancers (~400 bp) were deleted in V6.5 murine ESCs using the CRISPR/Cas9 system. sgRNAs were cloned into the pX330 vector (Addgene: 42230) containing Cas9. Cells were transfected with two plasmids expressing Cas9 and an sgRNA complementary to each end of the targeted super-enhancer 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.

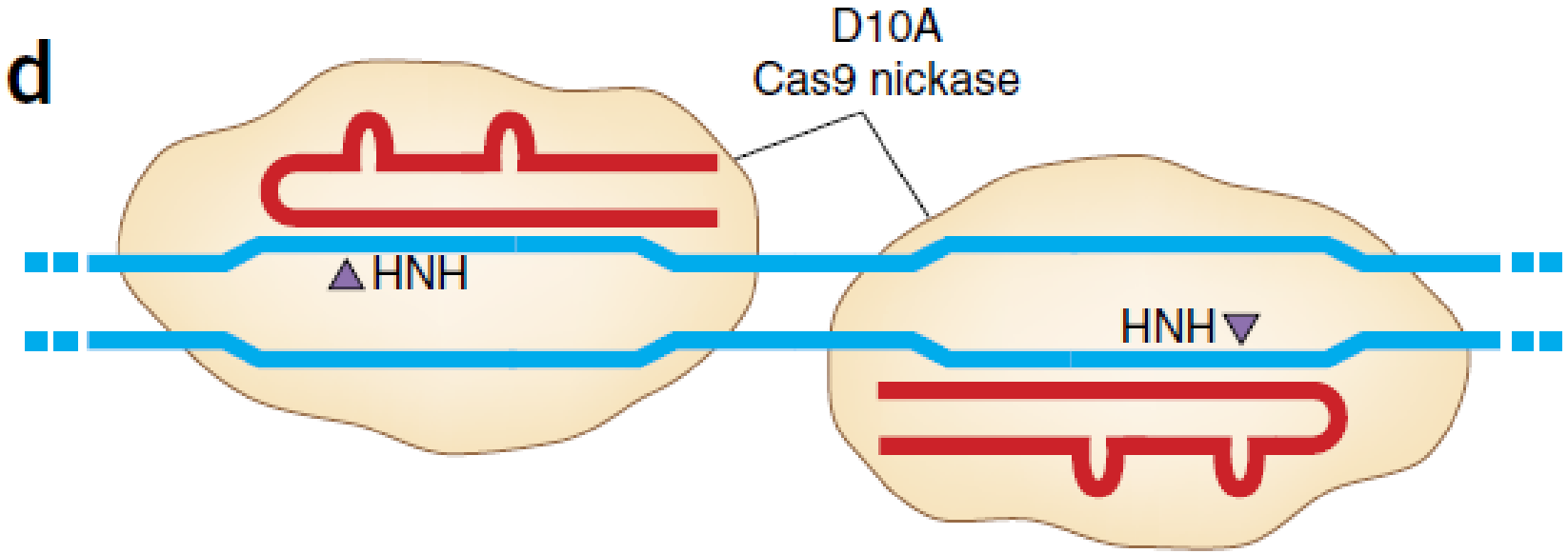



pX330-U6-Chimeric_BB-CBh-hSpCas9
8484 bp
AgeI (420)
FLAG
FLAG
FLAG
SV40 NLS
SV40 NLS
NLS
BglII (799)
ORF frame 2
ApaI (1908)
EcoRV (2402)
ORF frame 3
PmlI (3338)
NarI (5092)
NotI (4943)
bGH_PA_terminator
BGH_rev_primer
EcoRI (4704)
AmpR_promoter
pGEX_3_primer
f1_origin
pBR322_origin
SacII (6729)
FspI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

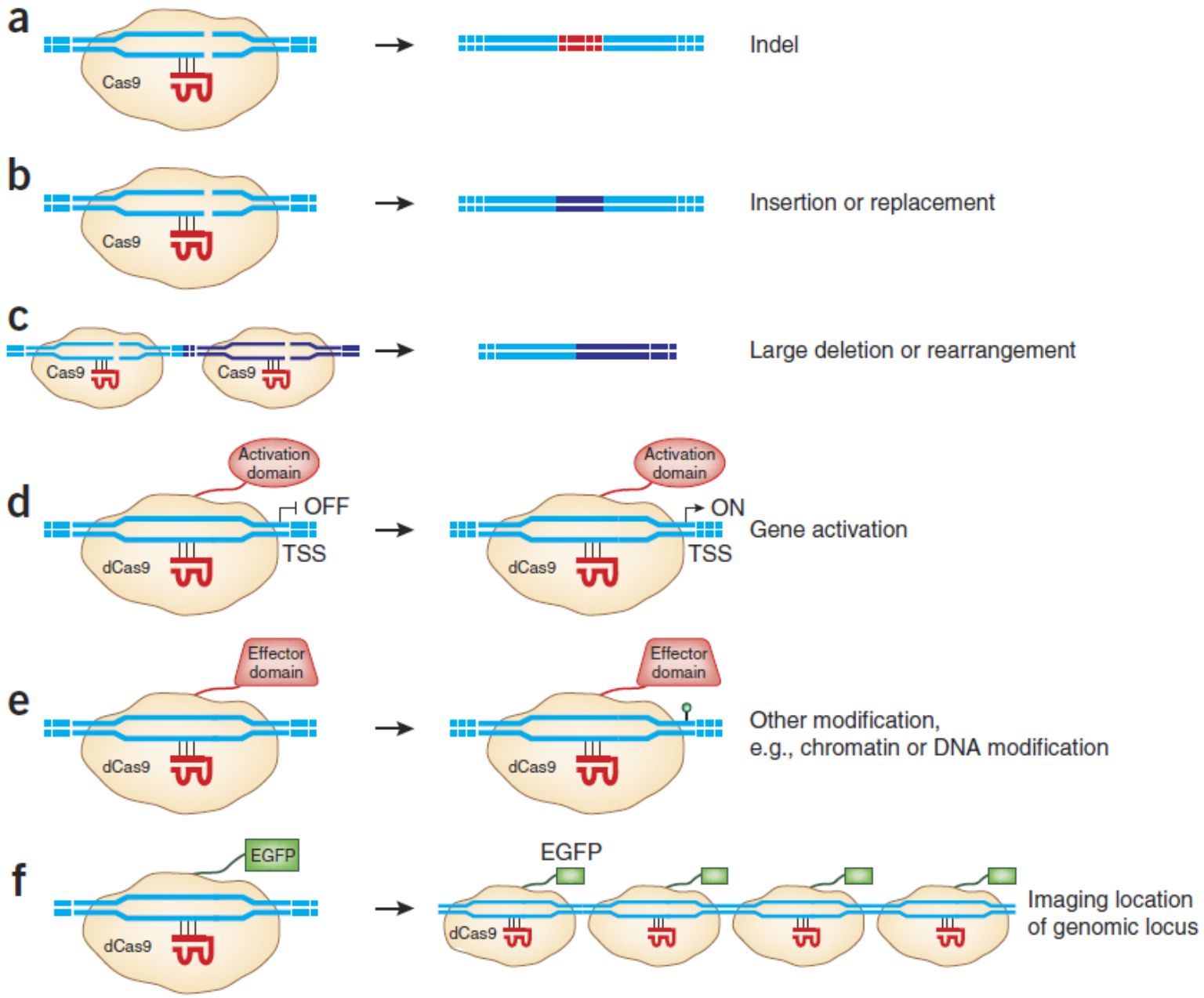

Your Target Sequence
GATC
CTAG
TTTT
AAAA
gRNA Scaffold
CMV Promoter
CAS9
Ori
AMP
U6 Promoter
pCas-Guide
8.0 kb
Myc/DDK

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.

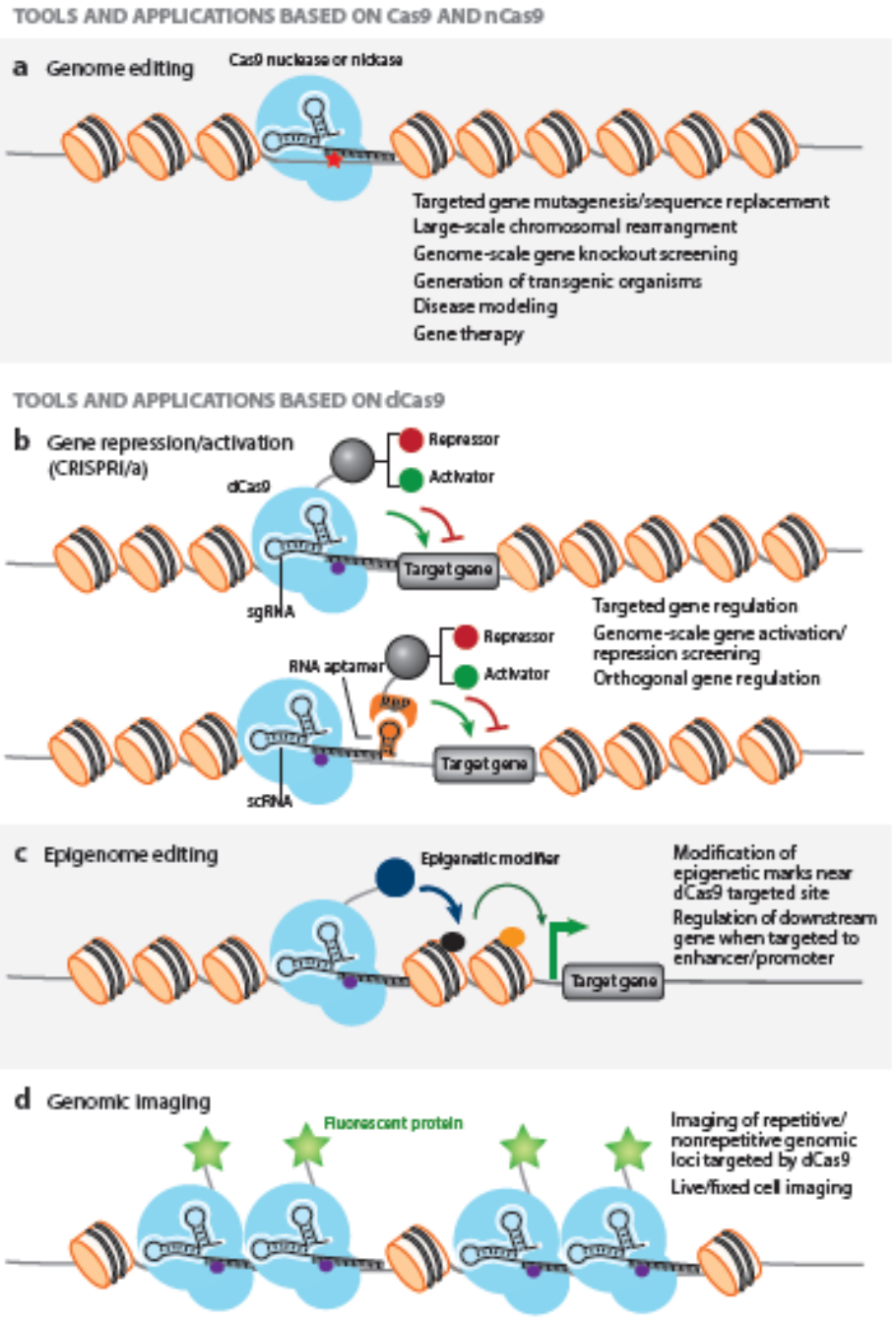

d
D10A
Cas9 nickase
HNH
HNH

OVERVIEW CRISPR-CAS SYSTEM APPLICATION



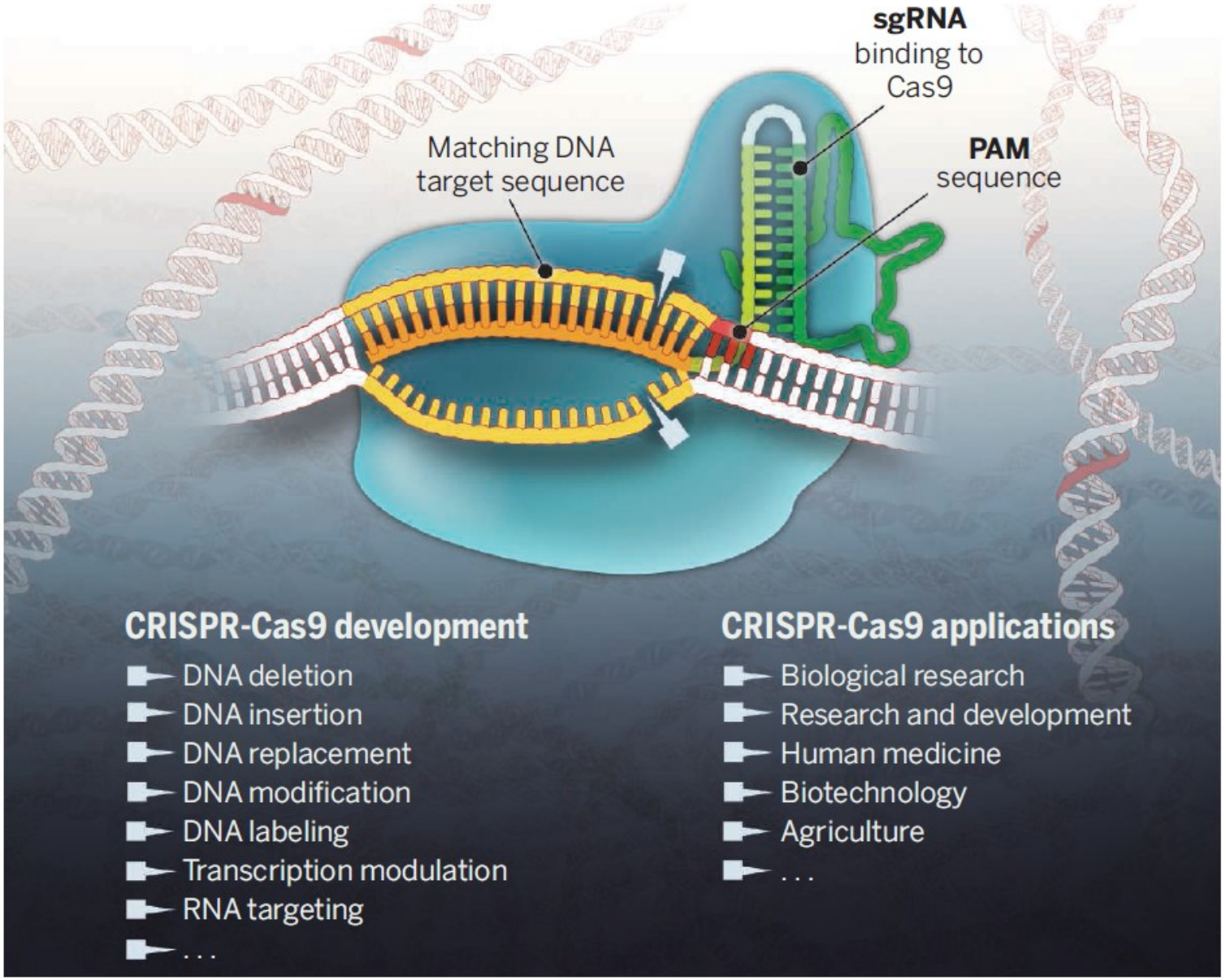
CRISPR/Cas9 in Genome Editing and Beyond

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

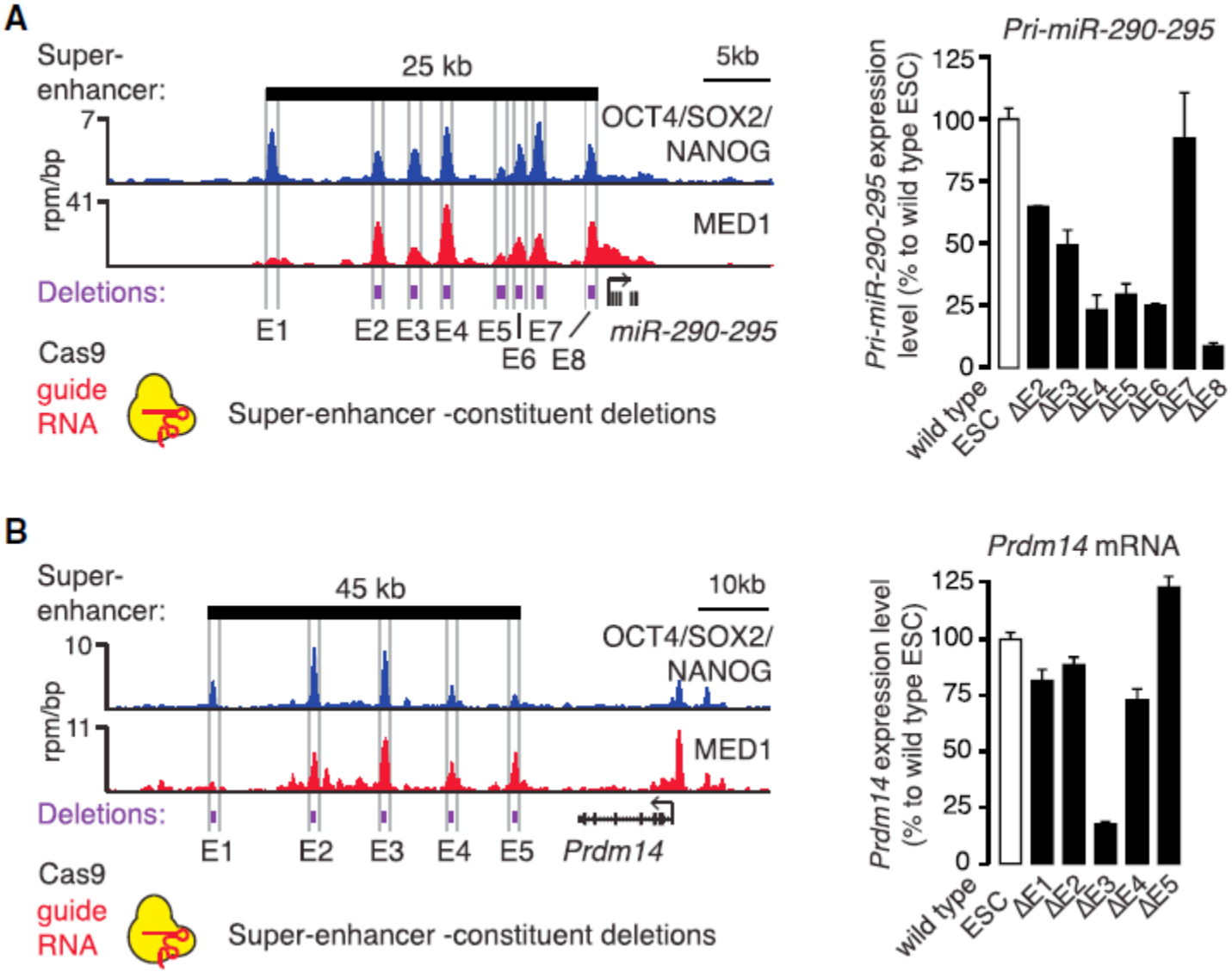


Annu. Rev. Biochem. 2016. 85:22.1–22.38

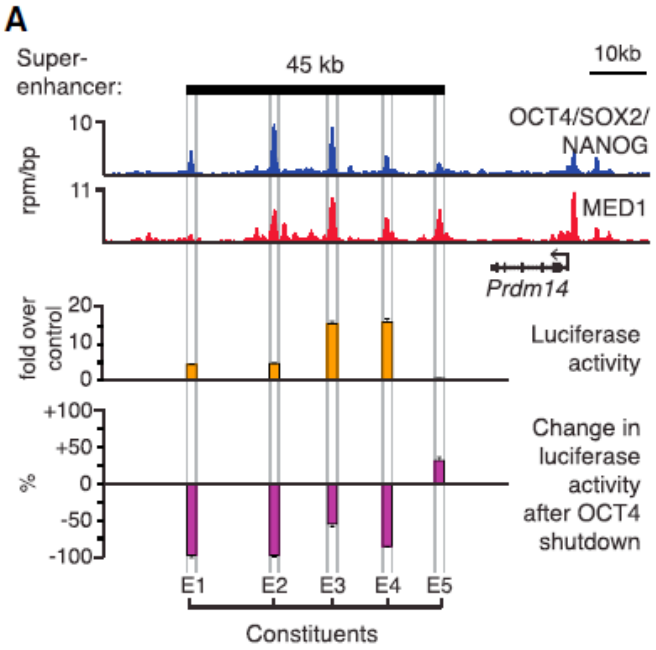
https://youtu.be/UCC2oILE7i0

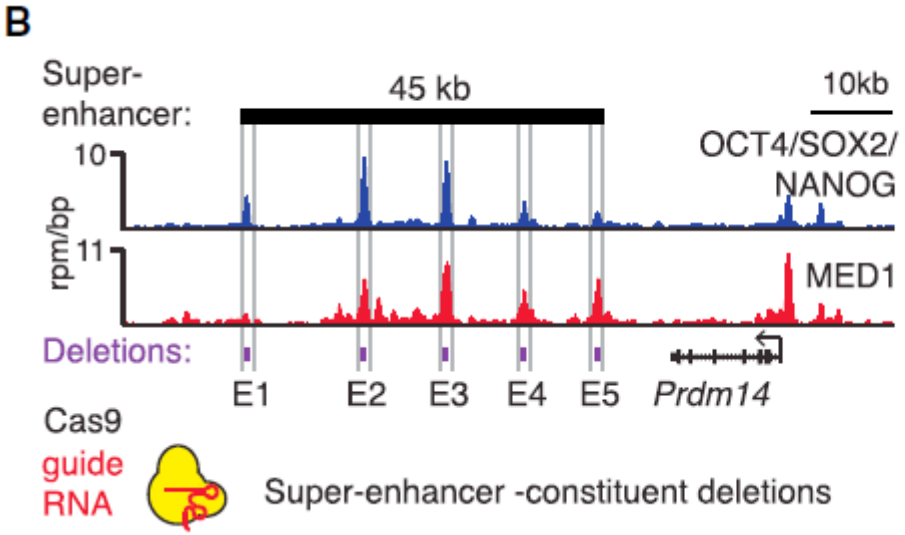


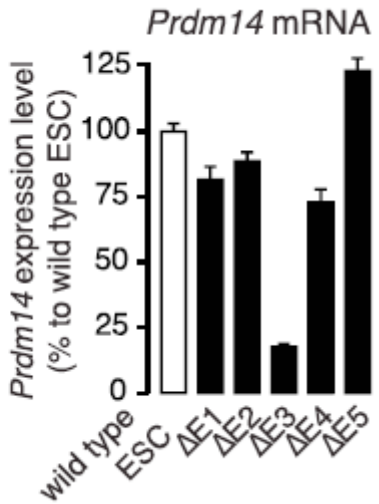
CONTRIBUTIONS OF SUPER-ENHANCER COSTITUENTS TO GENE EXPRESSION IN VIVO



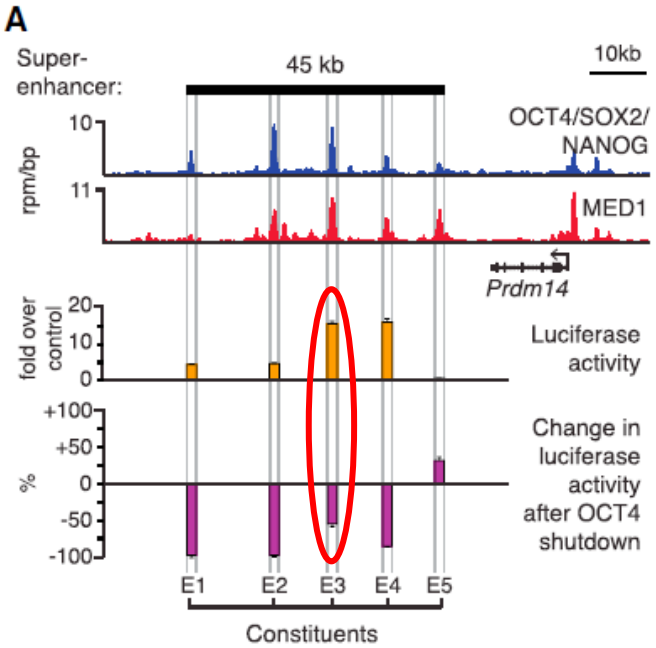
LUCIFERASE ACTIVITY AND DELETION OF SPECIFIC ENHANCER

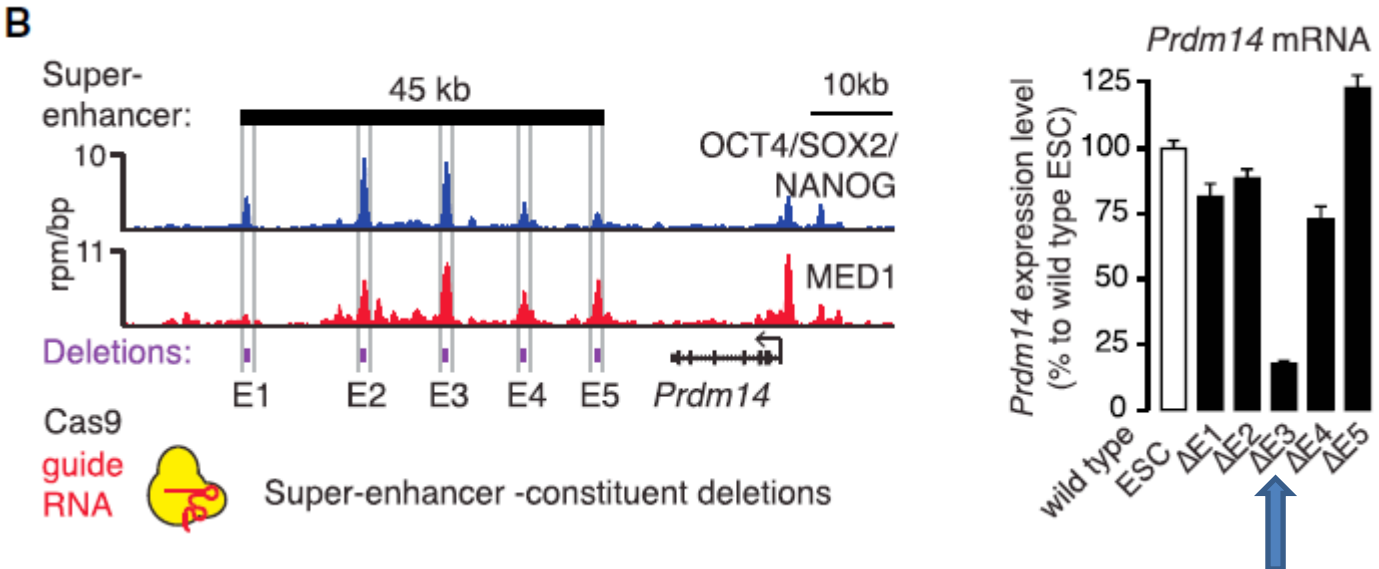




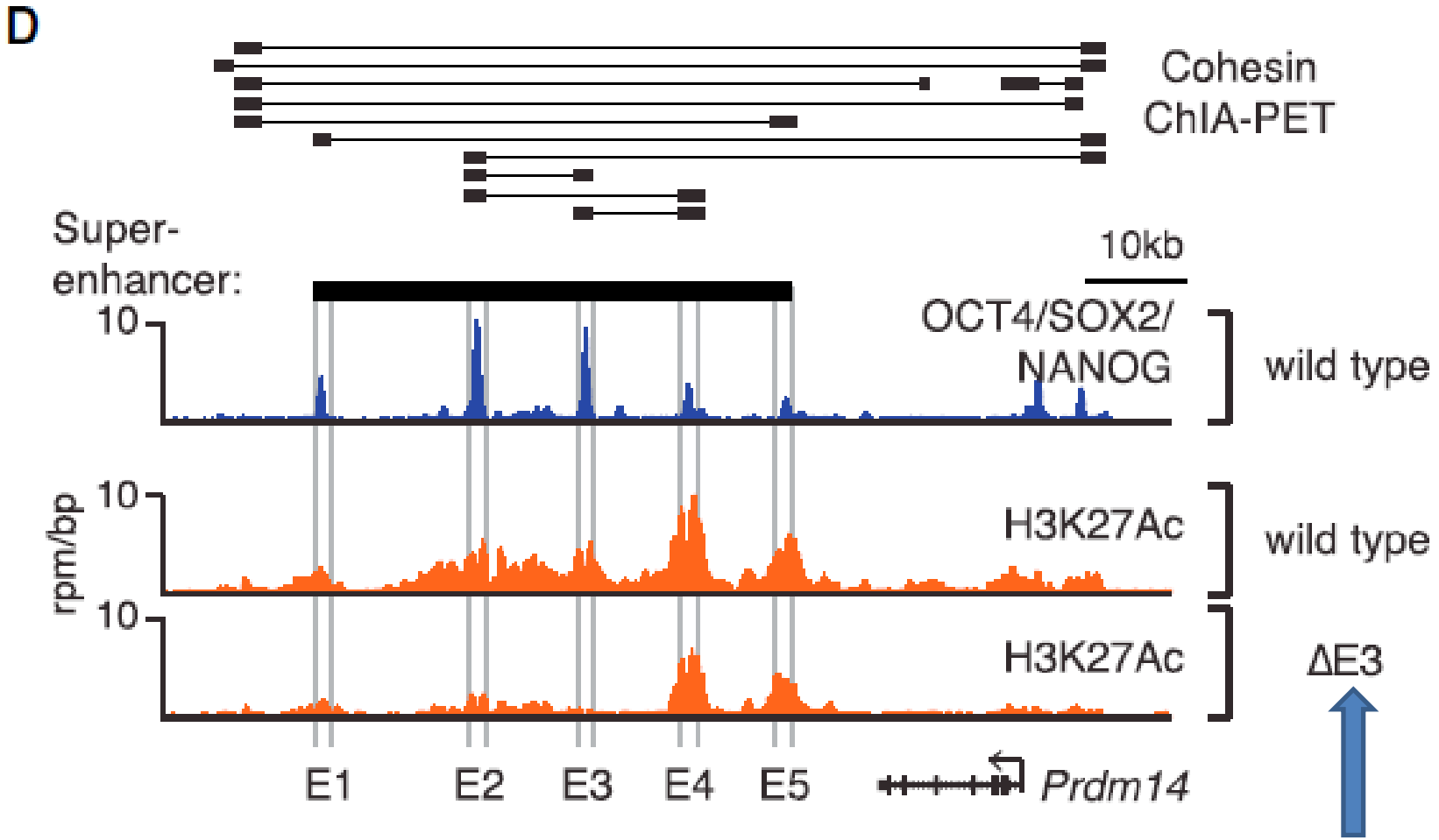


LUCIFERASE ACTIVITY AND DELETION OF SPECIFIC ENHANCER





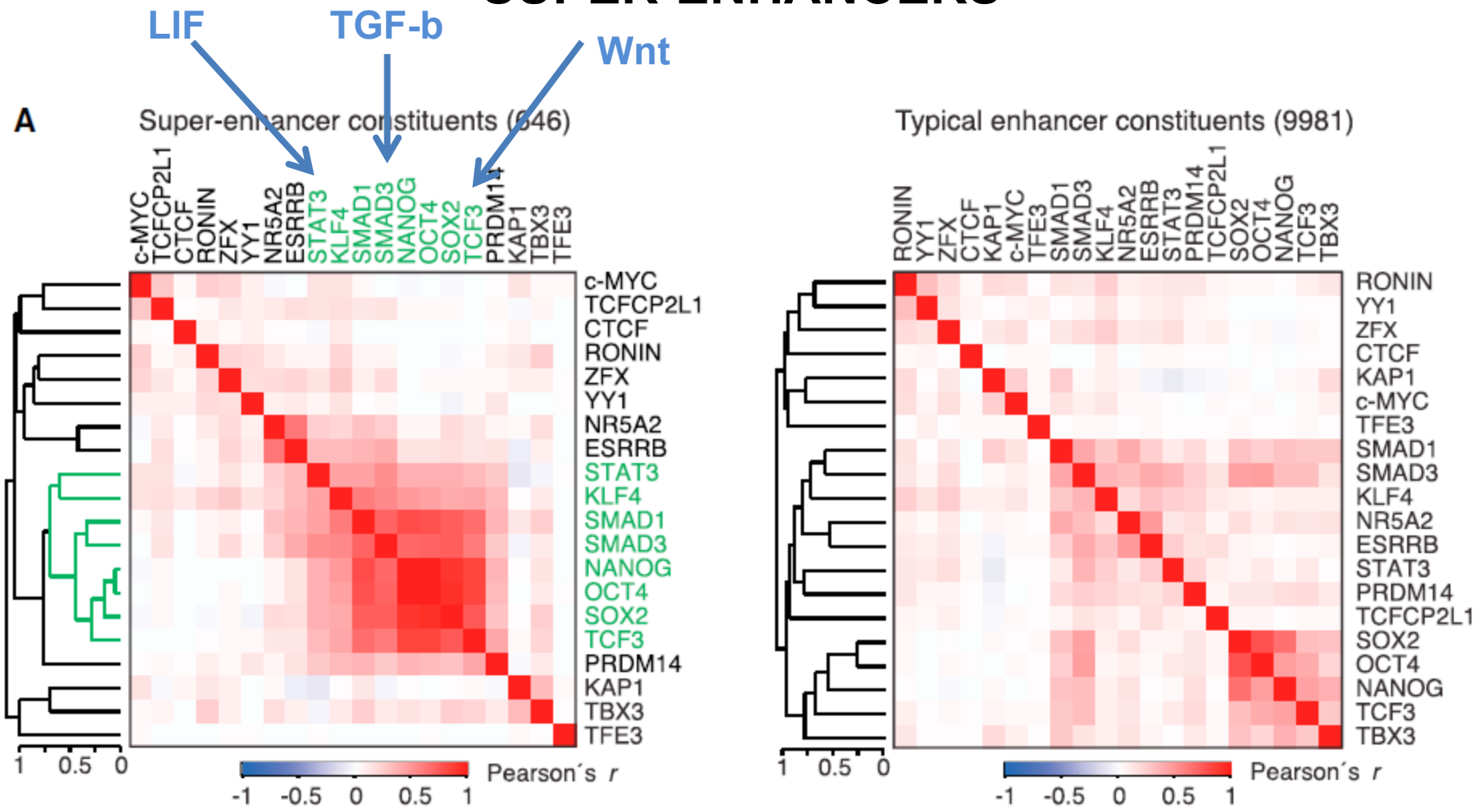
IDENTIFICATION OF INTERACTION BETWEEN SPECIFIC COSTITUENTS ENHANCERS



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.

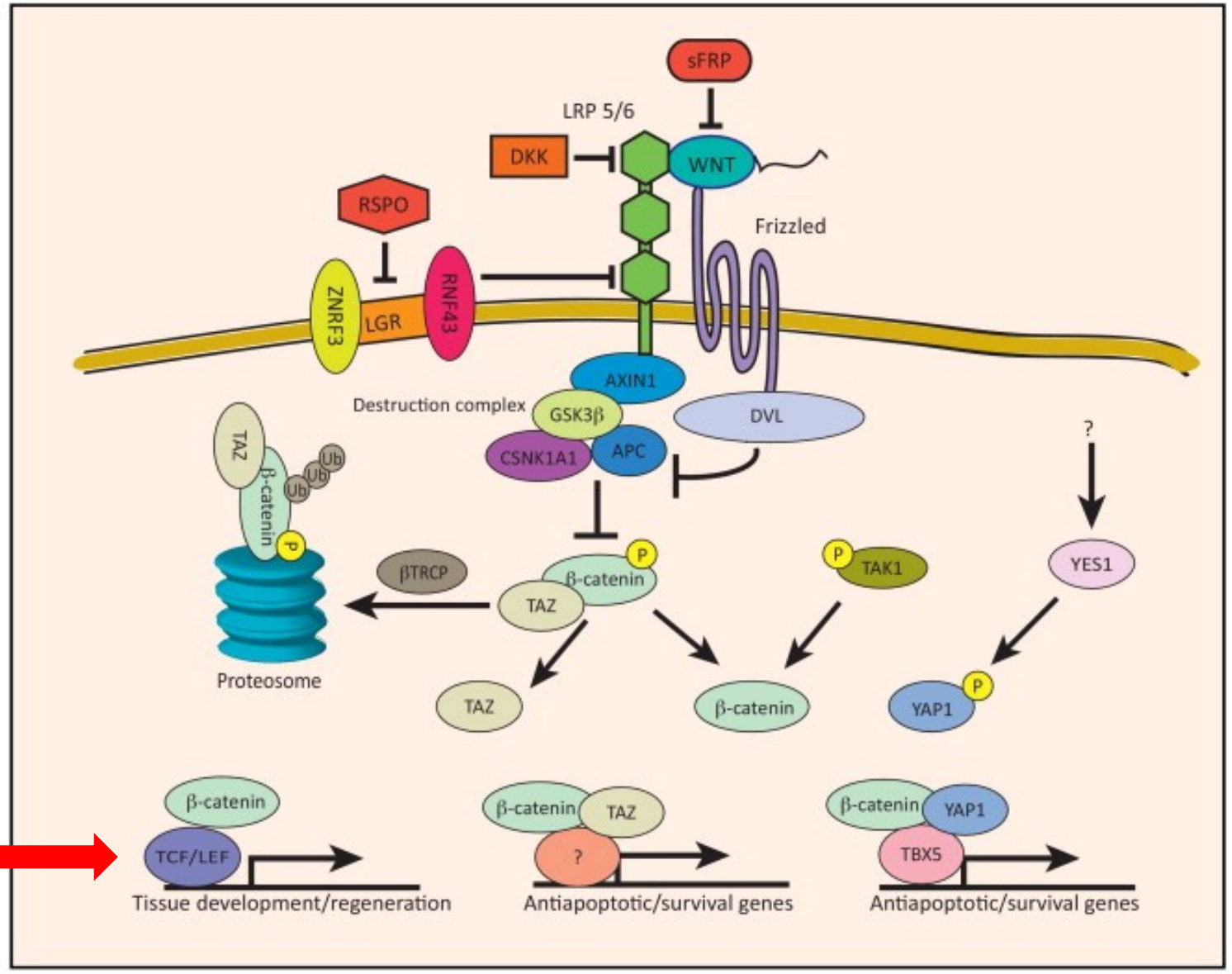
enhancers. An examination of the pattern of transcription factor binding to super-enhancer constituents provided a hypothesis to resolve this conundrum (Figure 3A, Table S3). The terminal TFs of the Wnt (TCF3), TGF-β (SMAD3), and LIF (STAT3) signaling pathways, which play essential roles in transcriptional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

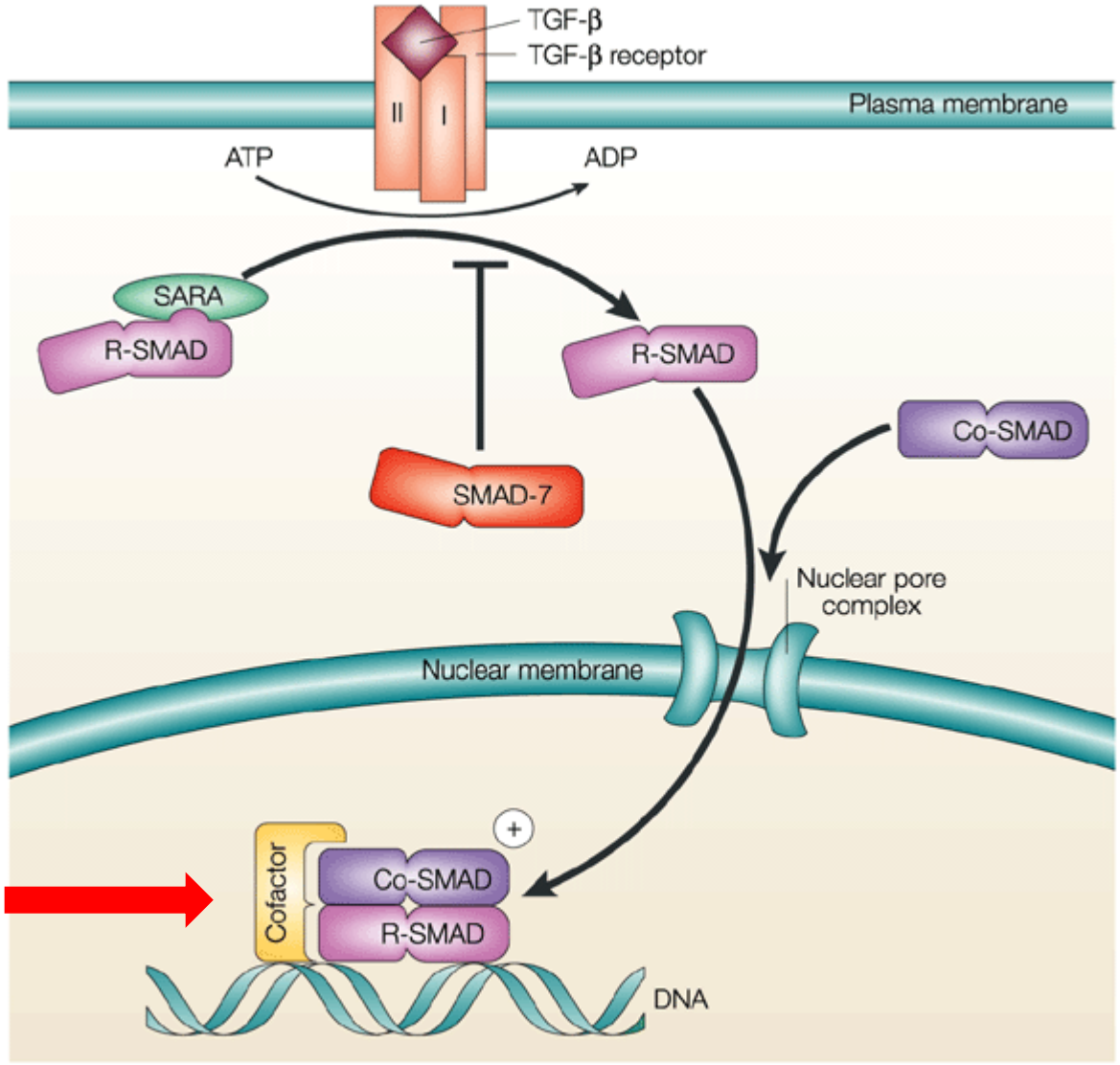
Signal transduction pathways

to activate the transcription factors
And
to change regulatory core

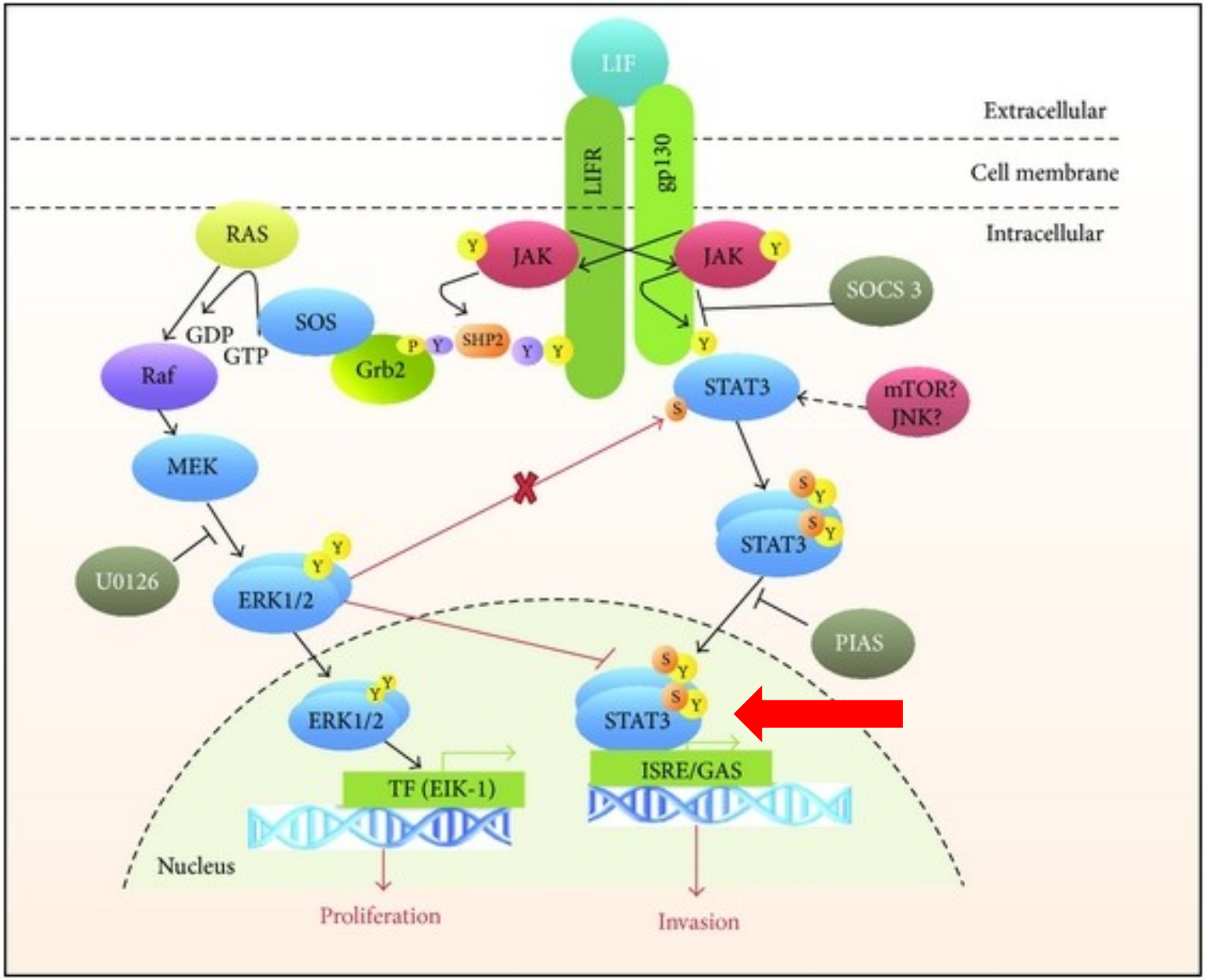
WNT PATHWAY


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

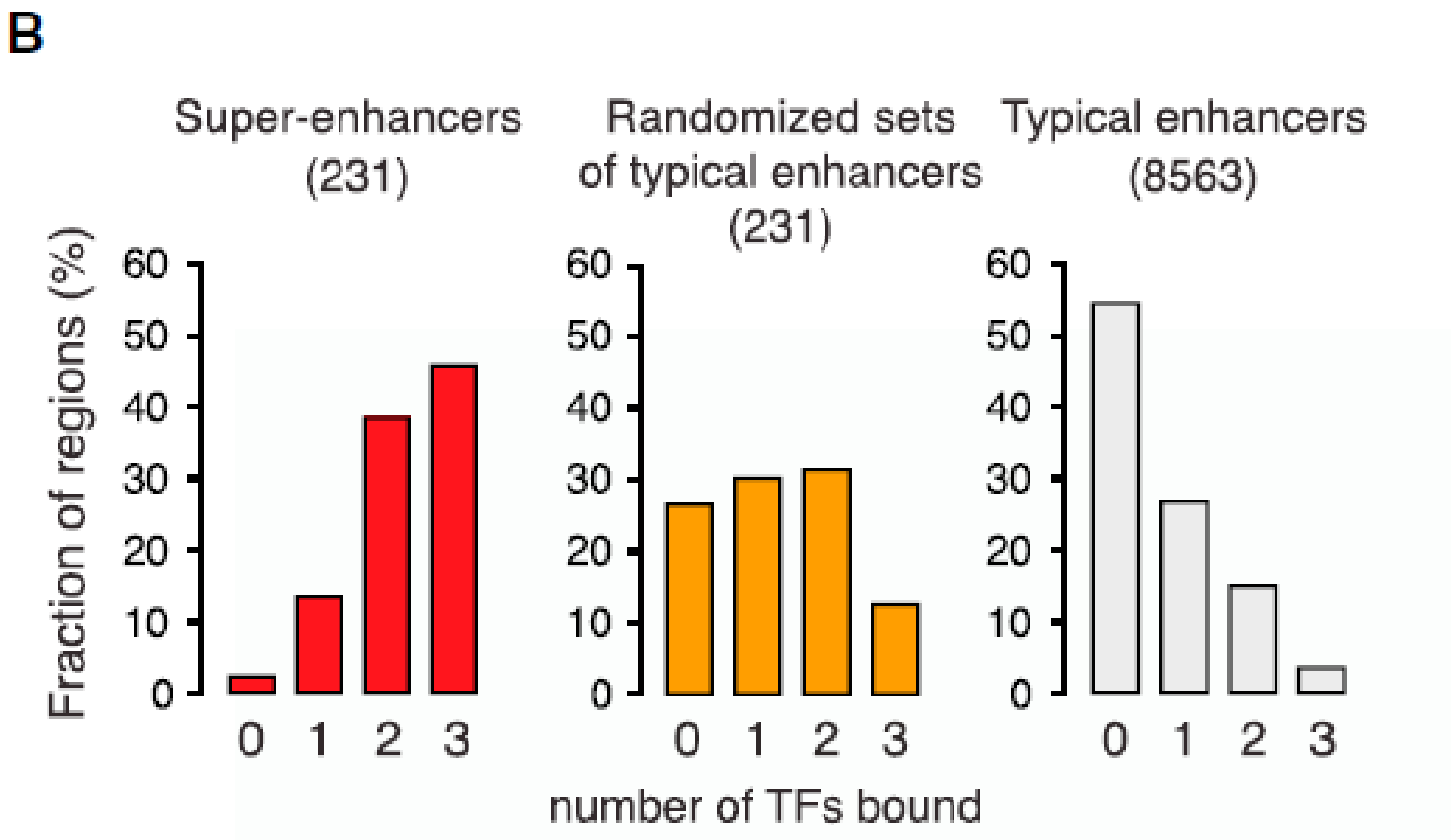
TGF beta PATHWAY



Nature Reviews | Immunology


LIF
Extracellular
Cell membrane
Intracellular
LIFR
gp130
RAS
JAK
JAK
SOCS 3
SOS
GDP
GTP
Raf
Grb2
SHP2
STAT3
mTOR?
JNK?
MEK
U0126
ERK1/2
STAT3
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.

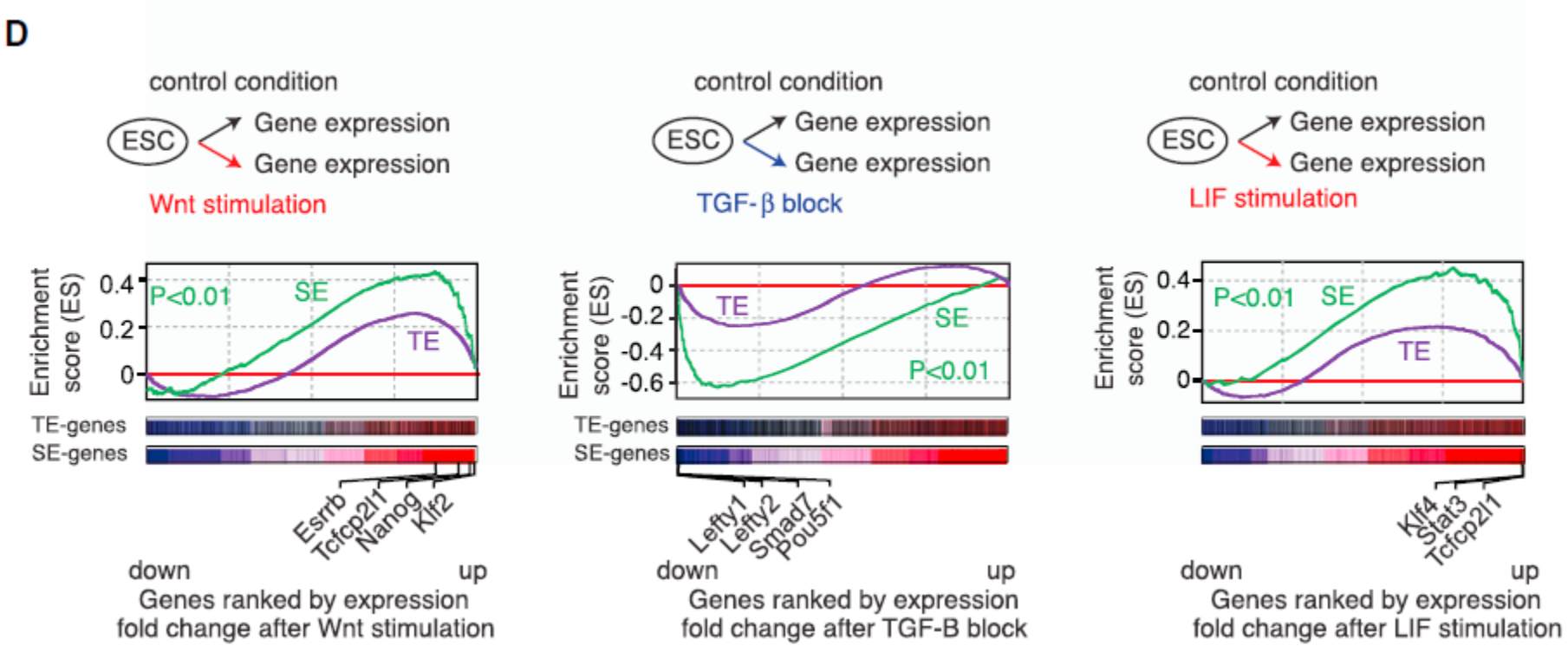
c

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

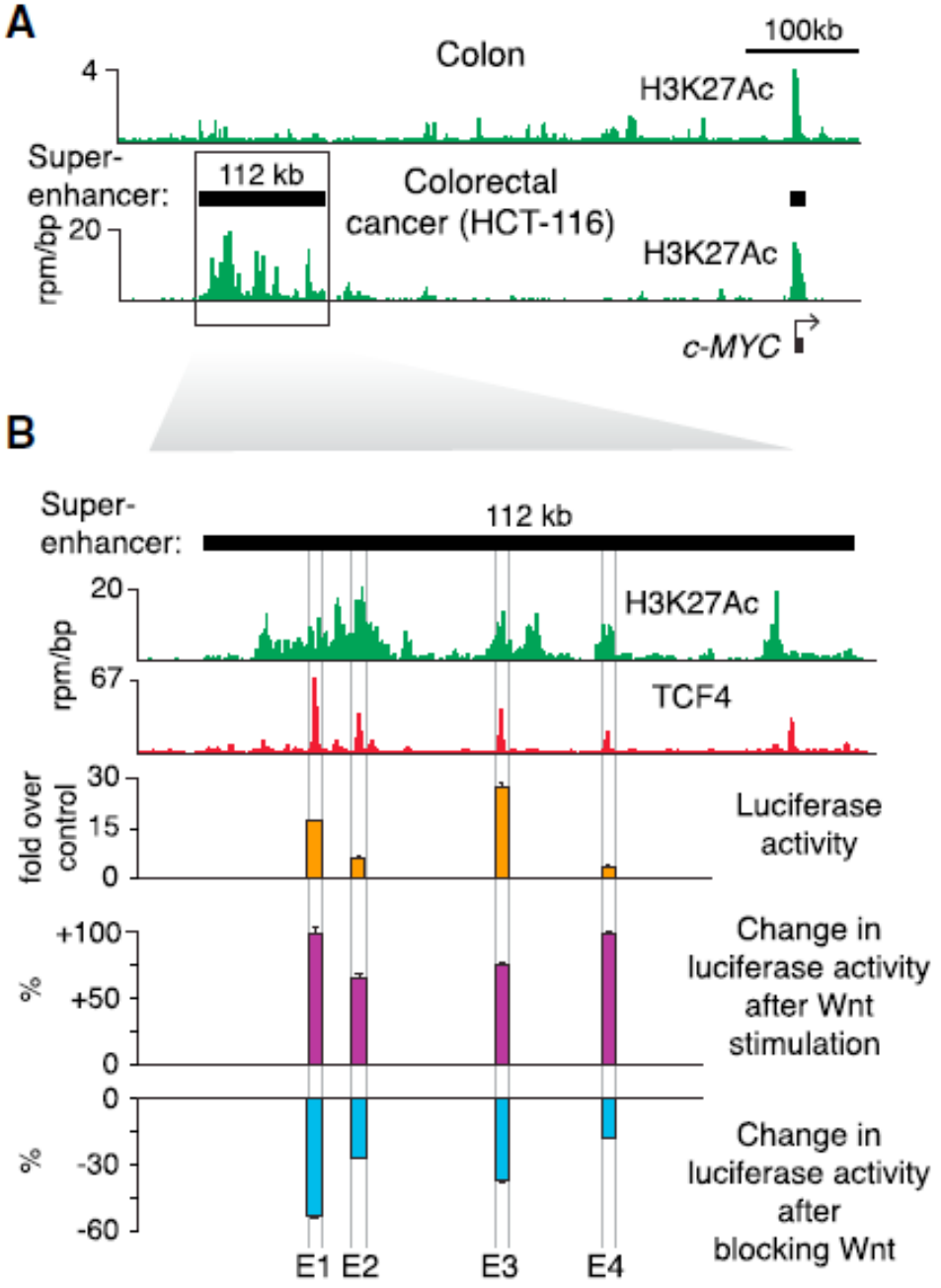
RNA-Seq RPKMs were calculated for two replicates each of murine ESCs treated with LIF for 1h (E-MTAB-1796 Arrayexpress dataset (Martello et al., 2013)) . Reads for each replicate were aligned to the mm9 reference genome using Tophat2 (Trapnell et al., 2009) version 2.0.11, using Bowtie version 2.2.1.0 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.

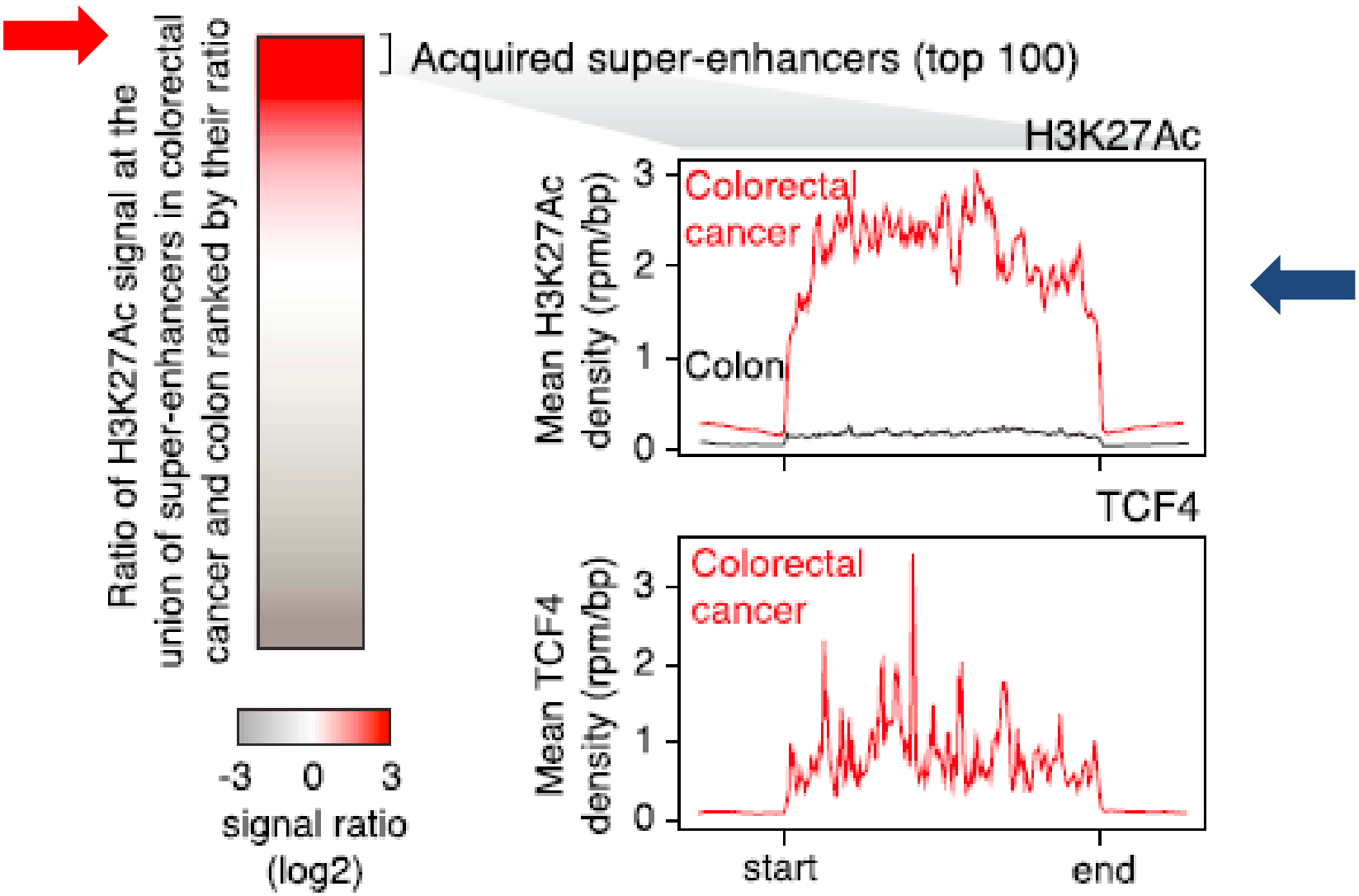
SE function on c-Myc locus



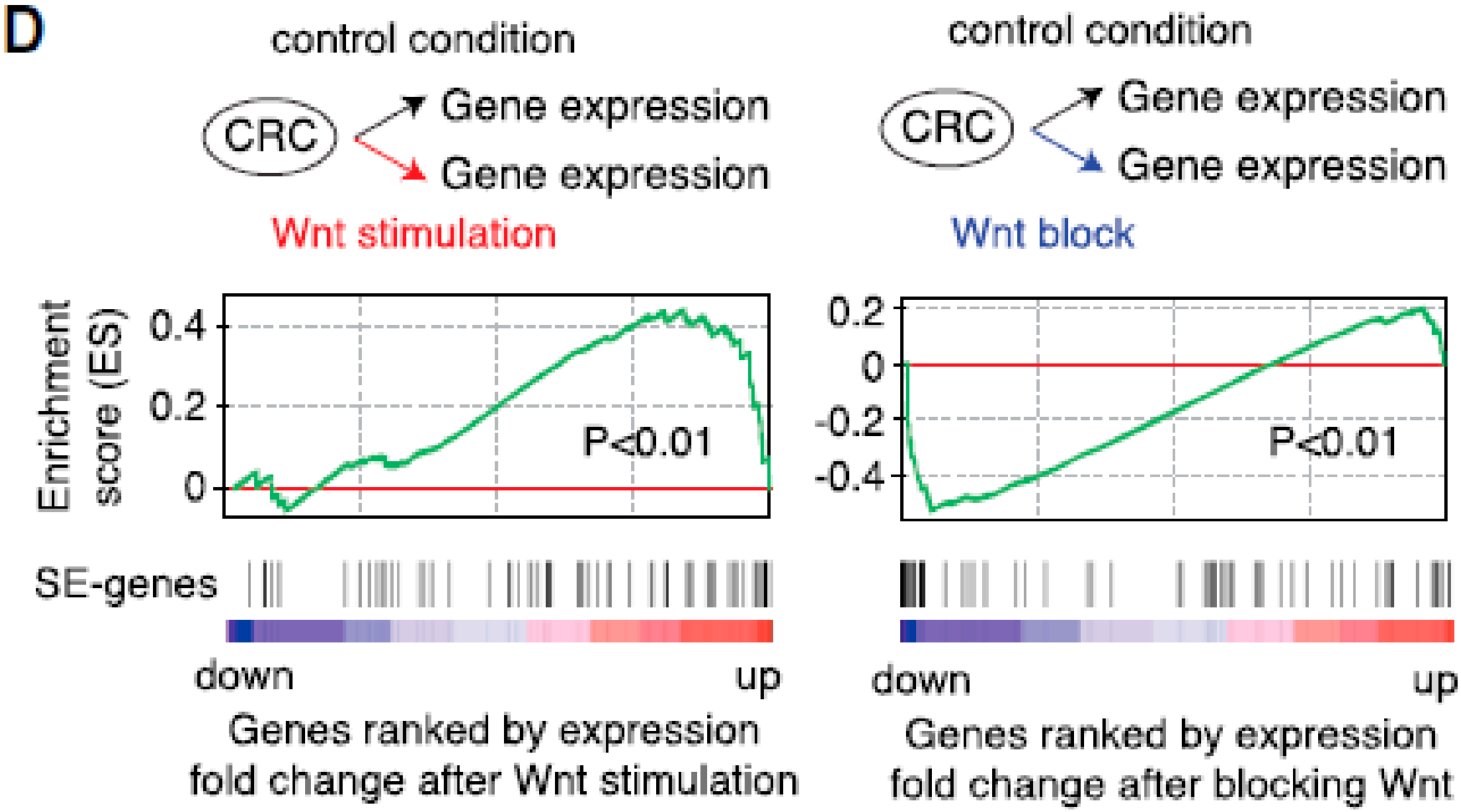
ChIP-seq binding profiles for H3K27Ac at the c-MYC locus in colon and colorectal cancer cells (HCT-116).

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.

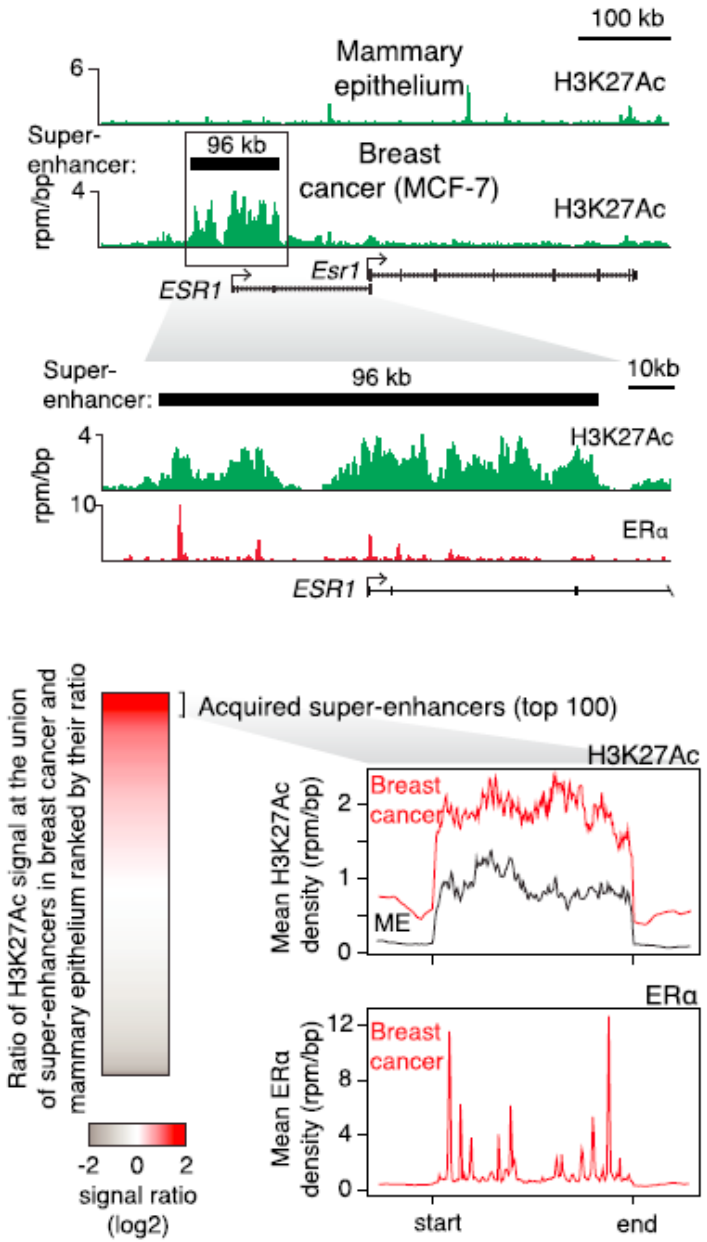
Left: ratio of H3K27Ac in CRC (HCT-116) versus normal colon tissue used densities at the union of SEs identified in the two samples. **Right:** metagene representation of H3K27Ac and TCF4 ChIP-seq densities at the regions corresponding to the top 100 acquired super-enhancers.



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.

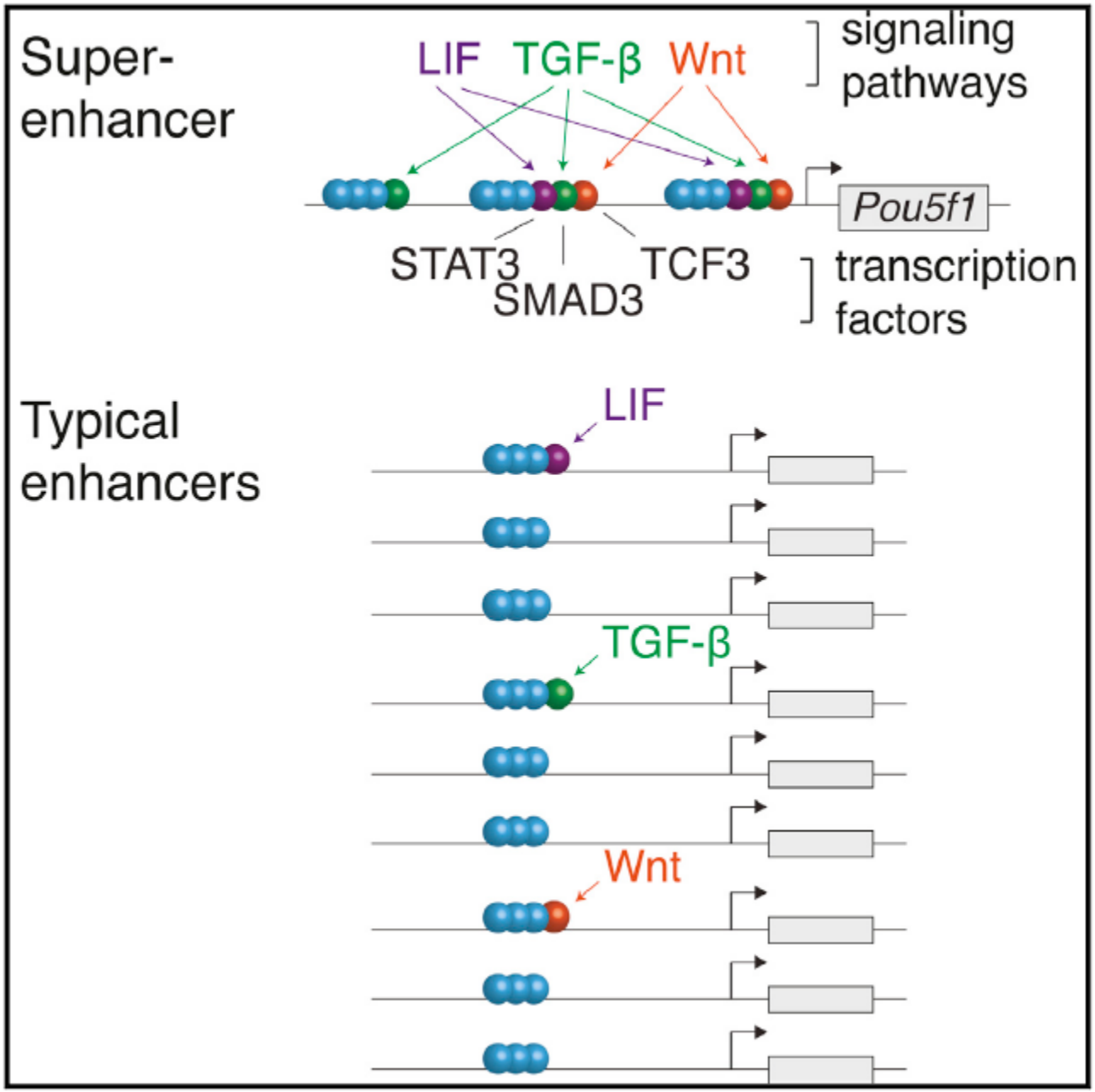
Several lines of evidence argue that the constituent enhancers of at least some super-enhancers can act as an interdependent structural and functional unit to control their associated genes. Our results show that ESC SEs generally consist of clusters of active enhancers that have OCT4-dependent and ESC-specific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

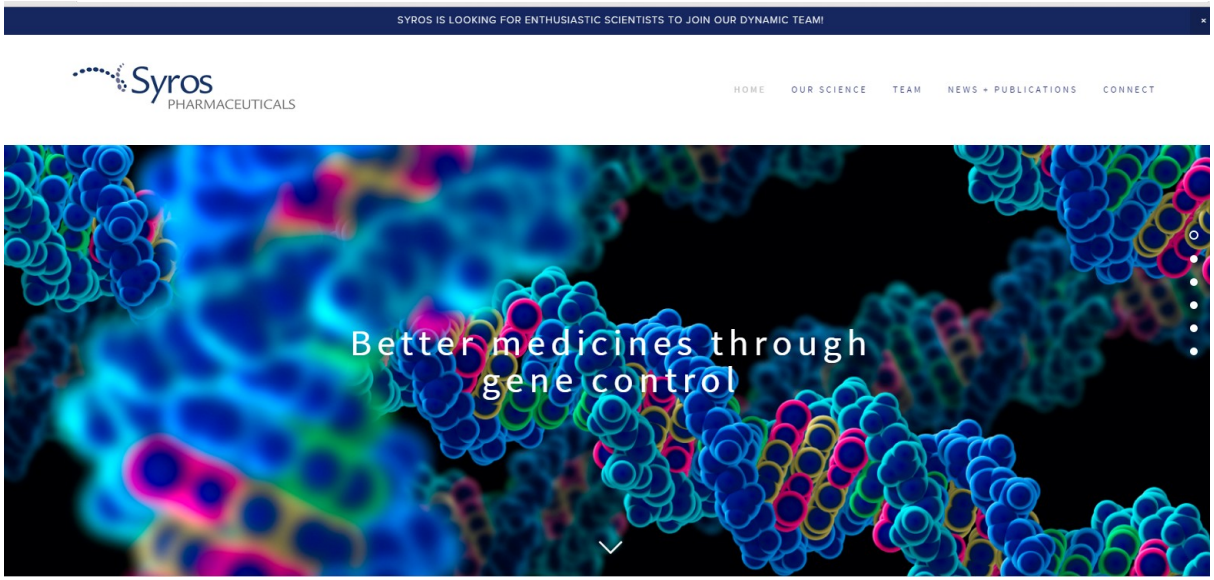
SEs characteristics

- **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.

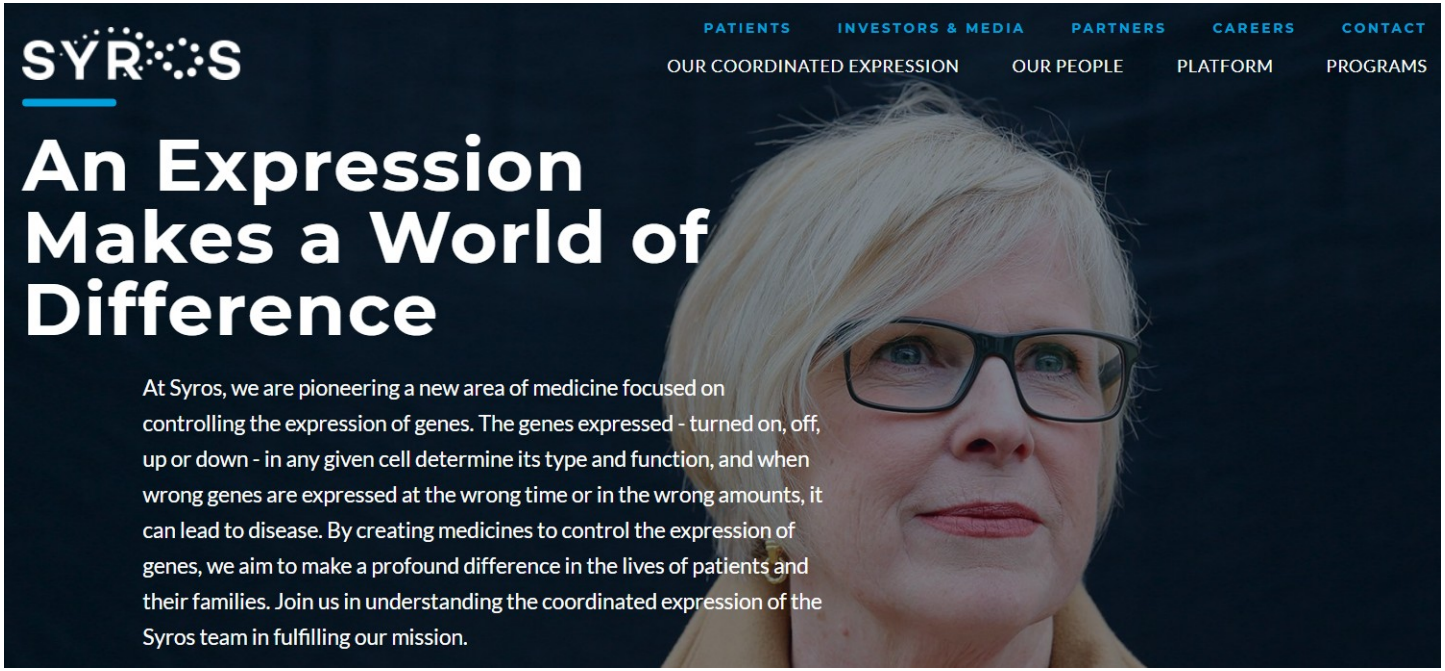
Graphical Abstract




SYROS IS LOOKING FOR ENTHUSIASTIC SCIENTISTS TO JOIN OUR DYNAMIC TEAM!
Syros
PHARMACEUTICALS
HOME
OUR SCIENCE
TEAM
NEWS + PUBLICATIONS
CONNECT
Better medicines through
gene control

2015

2017

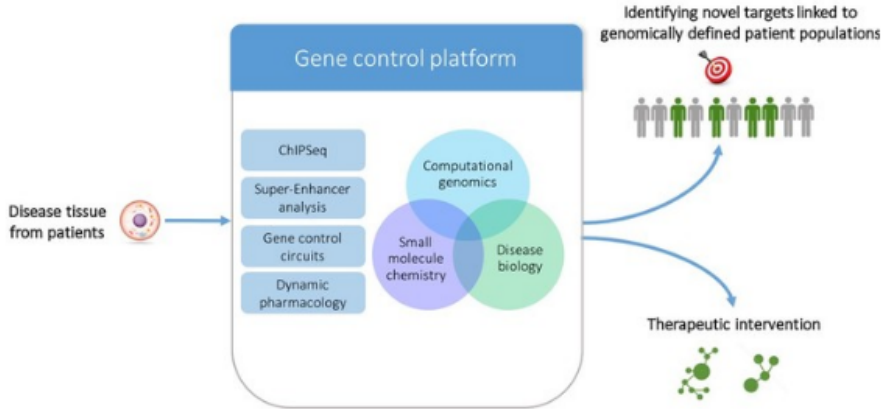

SYROS
PATIENTS
INVESTORS & MEDIA
PARTNERS
CAREERS
CONTACT
OUR COORDINATED EXPRESSION
OUR PEOPLE
PLATFORM
PROGRAMS
An Expression Makes a World of Difference
At Syros, we are pioneering a new area of medicine focused on controlling the expression of genes. The genes expressed - turned on, off, up or down - in any given cell determine its type and function, and when wrong genes are expressed at the wrong time or in the wrong amounts, it can lead to disease. By creating medicines to control the expression of genes, we aim to make a profound difference in the lives of patients and their families. Join us in understanding the coordinated expression of the Syros team in fulfilling our mission.

WE ARE A THERAPEUTICS COMPANY PASSIONATELY COMMITTED TO APPLYING OUR SCIENTIFIC LEADERSHIP IN THE EMERGING FIELD OF GENE CONTROL TO TRANSFORM THE LIVES OF PATIENTS WITH CANCER AND OTHER DISEASES.

Founded by world-leading experts in gene regulation, and led by a seasoned management team, Syros is focused on developing new medicines that control genes that cause disease. Central to our approach is a proprietary drug development platform of integrated gene control assay technologies, computational biology, chemistry, and deep biologic insights that gives us an unprecedented understanding of and control over gene regulation. Our pioneering research and drug development capabilities provide us with the distinctive ability to identify new, important disease and patient-targeted gene control medicines.

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.

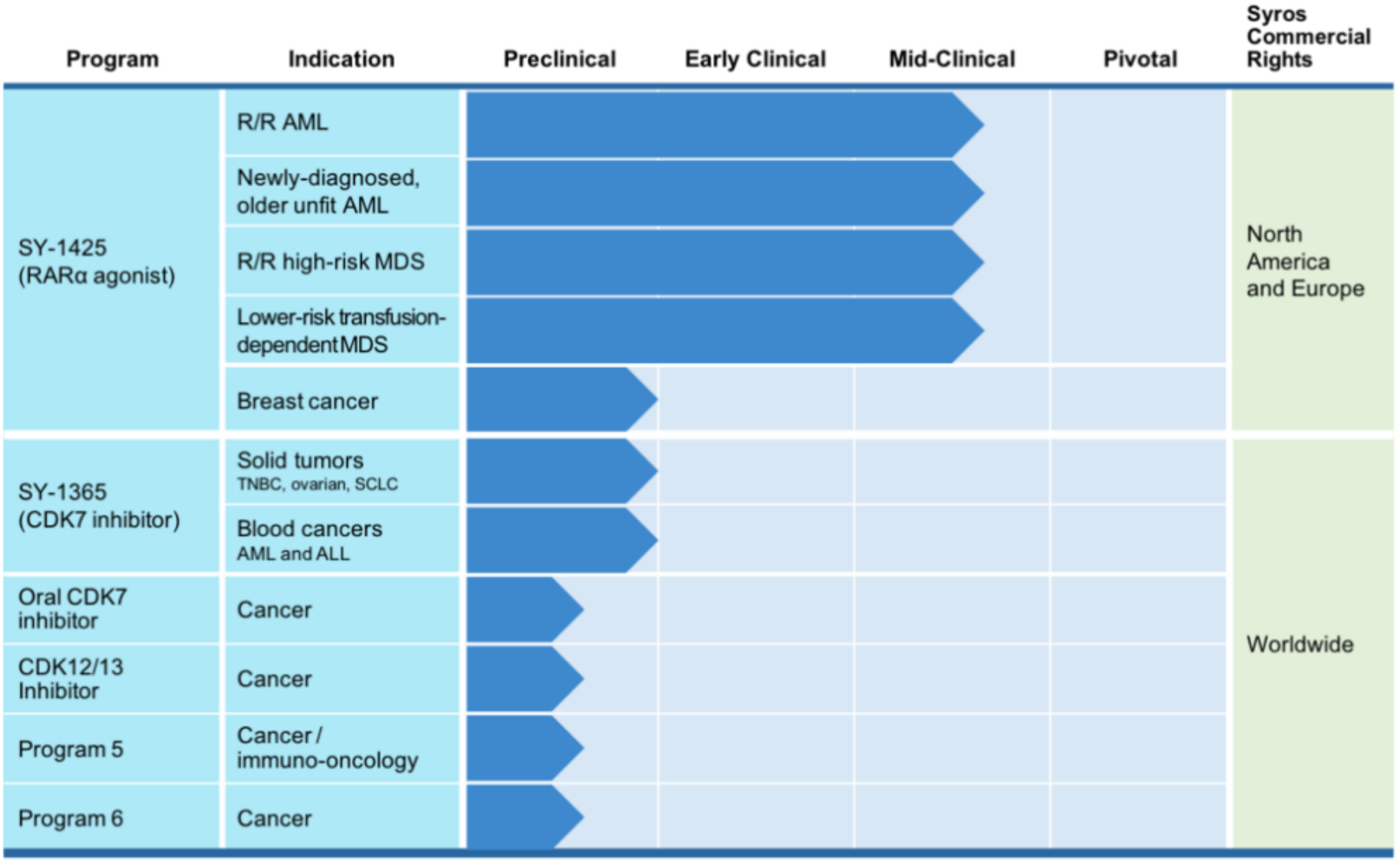


By doing so, we believe our gene control platform will allow us to:

- 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.

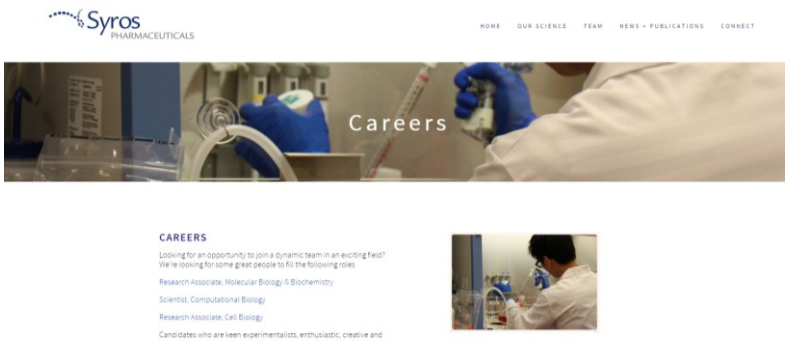
Learn more about our platform

DRUGS DISCOVERY FROM SYROS PHARMACEUTYCAL

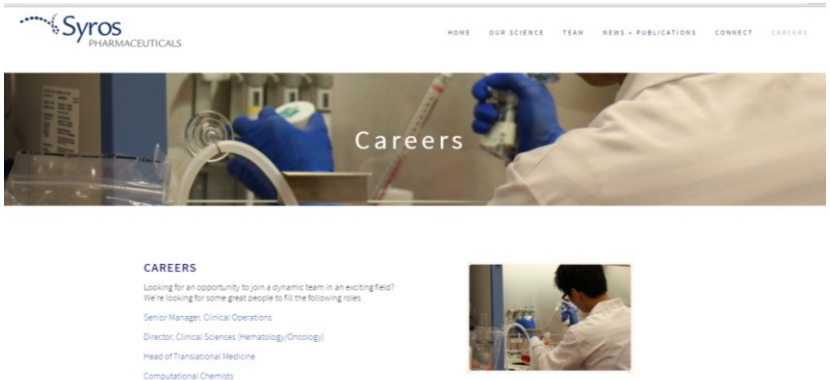


*SY-1425 is approved in Japan as Amnolake® (tamibarotene) for patients with relapsed/refractory APL

2016



2017



2018

What is CRISPR,
its scientific and ethical impact?

https://youtu.be/TdBAHexVYzc