Understanding super-enhancers

Hao Cheng[†], Xiaoyang Dou[†] & Jing-Dong J. Han[*]

Key Laboratory of Computational Biology, CAS Center for Excellence in Molecular Cell Science, Collaborative Innovation Center for Genetics and Developmental Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

Super-enhancers are defined as cluster of enhancers with dense TF binding, which can activate proximal cell-identity gene expression. Here we review the identification, functional significance of super-enhancers, and their relationships with cancer. With the current intense interests in super-enhancers, more super-enhancers will be defined and studied in different cell types and tissues and in different developmental and biological contexts to reveal molecular mechanisms of their functions.

ENHANCERS

Transcription is mainly regulated by transcription factors (TFs) binding at specific DNA sequences to recruit RNA polymerase II initiation or elongation factors. The most studied sites are promoter regions, which harbor transcription initiation sites. There are also some DNA sequences near or far away from promoter regions, which contain multiple transcription factor binding sites. Those DNA sequences are referred to as "enhancers".

The first time the term "enhancer" appeared was in 1981. Julian Banerji et al. found that rabbit β-globin gene expression was enhanced by the SV40 DNA sequence at a distance without orientation (Banerji et al., 1981). Based on the detection of β-globin genes, Grosveld et al. further explored the feature of enhancer by constructing "minilocus" consisting of β-globin genes and an enhancer region (Grosveld et al., 1987). Since then, functional enhancers were detected

in a limited scale for a long time. Recently with the development of high throughput sequencing, more and more enhancers were detected in a genome-wide scale.

Acetyltransferase P300, a coactivator to increase expression of its target gene, has been used to detect enhancer in mouse embryonic forebrain, midbrain and limb tissue (Visel et al., 2009). Meanwhile, histone modification like H3K4me1 has been found to be a unique enhancer mark (Heintzman et al., 2007). According to the studies done in mouse embryonic stem cell (ESC) and adult tissues, enhancer with H3K4me1 can be divided into active and inactive/poised states based on whether it is marked by H3K27ac or not. In addition to P300 and histone modifications, DNase hypersensitive sites also partially overlap with enhancer regions.

WHAT IS A SUPER-ENHANCER?

In *Drosophila*, the *eve* gene is auto-regulated by distal cis-regulatory region, and different cis-sequence within the distal cis-regulatory region function differently in the initiation of different stripes (Harding and Levine, 1988), which gives a prototype of super-enhancers that we will discuss below.

Super-enhancer was first proposed as a large genomic domain constitute of cluster of enhancers occupied by master transcription factors at genes that control pluripotent states of mouse ESC (mESC) (Whyte et al., 2013). It is different from typical enhancer in size, transcription factor binding density and sensitivity to perturbation. Then the definition of super-enhancer was refined as the following.

---

†Contributed equally to this work
*Corresponding author(email: jdhan@picb.ac.cn)

First of all, identify constituent enhancers as master regulator enriched regions, and then constituent enhancers within 12.5 kb are merged into a larger region. Next, normalized Med1 ChIP-seq signals are ranked to generate a curve on the signal to rank plot, and a line with a slope of 1 tangent to the curve is used as a cutoff to separate super-enhancer above the point and typical enhancer below the point of tangency (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). Thereafter, super-enhancers are defined in many cells and tissues by tissue-specific master transcription factors. However, master transcription factors for most cells are not known, alternatively, one can use marks that routinely used to identify typical enhancers, such as H3K27ac, H3K4me1, p300 and DNase hypersensitivity, to define super-enhancers. Among these marks, H3K27ac has been shown to perform the best in mESC. Figure1 shows an example of super enhancers, a super-enhancer over the Sox2 gene, defined with histone mark H3K27ac in different mouse tissues. Furthermore, if these alternative ways are used to define super-enhancer and then master regulator might be inferred in many different cell types based on TF binding profiles or motifs. For cells such as ESCs, myotubes, Th cells, pro-B cells and macrophage with known master regulators, majority of the master TFs were captured by H3K27ac, and candidate master TFs for other cell types and tissues were found through analysis in 86 human cell and tissue samples (Hnisz et al., 2013). As TFs on super-enhancer are highly coordinately regulated, combining these factors in a reasonable way may lead to a more comprehensive detection of super-enhancers (Hnisz et al., 2013).

FUNCTIONS OF SUPER-ENHANCERS

In mESC, a super-enhancer was first defined as cluster of enhancers occupied by OCT4, SOX2 and NANOG and other mediators, in addition, Denes Hnisz et al. found that targeting TFs of Wnt (TCF3), TGF-β(SMAD3) and LIF (STAT3) signaling pathways, which are critical for stem cell state control, bound to individual constitute enhancers inside super-enhancers with similar pattern as OCT4, SOX2 and NANOG. Furthermore, super-enhancer associated genes are sensitive to the perturbation of these signaling pathways. Therefore, super-enhancers may function as a platform for dense transcription factor binding of different singling pathways to determine cell identity (Hnisz et al., 2015). In skin stem cell, pioneer factors such as SOX9 can sense slight changes in microenvironment, and then transduce the signal to establish new super-enhancers. Through such a mechanism or signal transduction, skin stem cells can pursue distinct lineages, repair wounds and exhibit plasticity in transitional states (Adam et al., 2015).

Enhancer RNAs (eRNA) are "linkers" between super-enhancers and the activation of target gene expression. Transcripts coordinately regulated within a super-enhancer unit through eRNAs can be mapped by active RNA polymerase II global nuclear run-on sequencing (GRO-seq) experiments. About 30.6% of typical enhancers and 93.3% of super-enhancers overlap with eRNA in intergenic regions. In response to Toll-like receptor signaling in macrophage, super-enhancer associated genes and super-enhancer associated eRNA show coordinated changes. However, these signal-dependent super-enhancers are different from the cell-identity determined super-enhancers, with the former related with innate immunity and inflammation (Hah et al., 2015).

In summary, super-enhancer can not only control cellular identity by activating cell-type specific signaling pathways, but also respond to stimuli and repress cell maintenance related genes through TF redistribution or eRNA coopera-
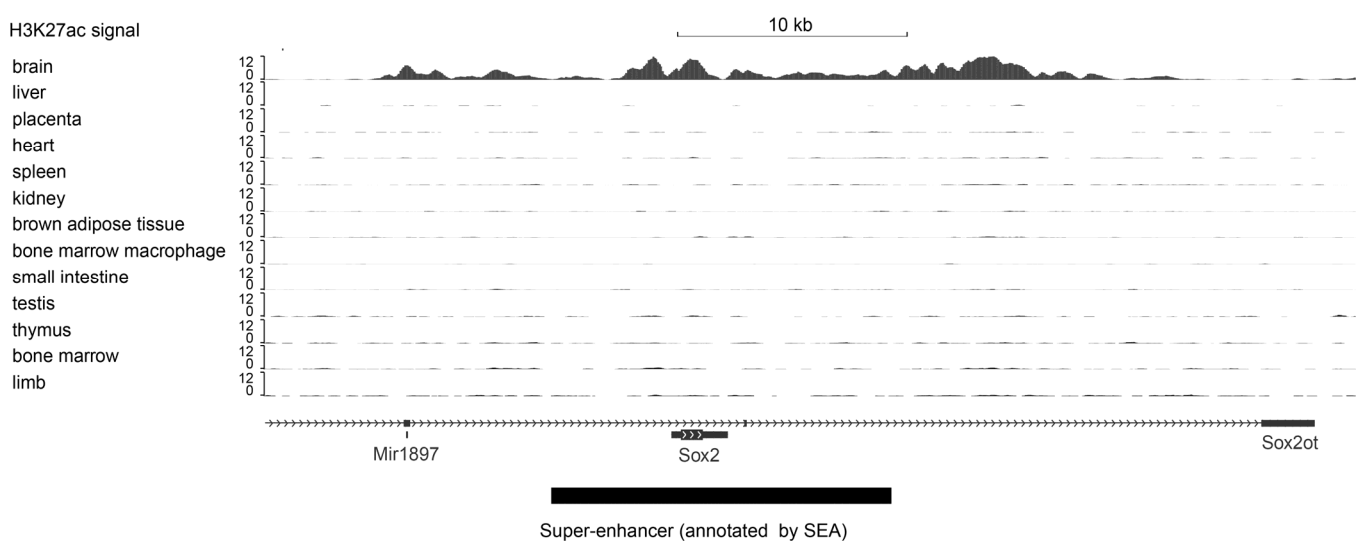
Figure 1   Visualization of H3K27ac signal on a super-enhancer over the Sox2 gene in different mouse tissue annotated by Super-Enhancer Archive (SEA) (Wei et al., 2016).

tion. Super enhancers as dynamic regulatory elements play an import role in cell identity control and response to environment (Hah et al., 2015).

SUPER-ENHANCERS AND CANCER

Super-enhancers were first associated with tumorigenesis when they were defined in a myeloma cell line (Loven et al., 2013). Loven et al. found that BRD4 inhibition by the small molecule JQ1, can selectively repress MYC expression by decreasing BRD4 binding at MYC super-enhancer regions. Further study reveals that some oncogenes, like MYC, acquire specific super-enhancers in multiple myeloma tumor cells compared with healthy cells (Hnisz et al., 2013).

The TAL1 oncogene is overexpressed in many samples of T cell acute lymphoblastic leukemia (T-ALL), but the reason was unknown. Mansour et al. found that one of TAL1 upstream super-enhancers contains a heterozygous 12-bp insertion, which made it possible for transcription factor MYB to bind (Mansour et al., 2014). CBP, CREB-binding protein, was known to promote H3K27ac at enhancer regions. The co-binding of CBP and MYB was found in most of the super-enhancer regions in Jurkat cells.

Another case is for THZ1, a covalent inhibitor of CDK7, while CDK7 is known to promote transcription activation through phosphorylation of RNA Pol II. Chipumuro et al. found that in neuroblastoma cells MYCN was associated with the largest super-enhancer. THZ1 treatment can reduce both mRNA and protein levels of MYCN by decreasing RNA Pol II binding at those super-enhancer regions (Chipumuro et al., 2014).

Taken together, these data suggested that some super-enhancers can promote oncogene overexpression. Disrupting the structure of such super-enhancers or inhibiting the cofactors involved in the formation of super-enhancers provides new routes for cancer therapy.

DIFFERENCE BETWEEN SUPER-ENHANCERS AND OTHER BROAD EPIGENETIC MARKS

Super-enhancers can be defined with histone mark H3K27ac (Hnisz et al., 2013). Several other large epigenetic marks are also found to be related with cell identity or tissue specific functions, such as H3K4me3 broad peaks (Benayoun et al., 2014; Chen et al., 2015) and DNA methylation valleys (large hypo-methylate regions) (Jeong et al., 2014; Xie et al., 2013). Evidence shows that those different broad epigenetic marks marked different sets of genes from those marked by super-enhancers, even though they overlap in some genomic regions. For example, both H3K4me3 broad peaks and super-enhancers marked stem cell regulators in mESC (Benayoun et al., 2014). But further analysis shows that H3K4me3 broad peaks have better discriminative power for cell lineages than super-enhancers. Moreo-

ver, H3K4me3 broad peaks are also found to specifically associate with tumor-suppressor genes, which are conserved in many cell types (Chen et al., 2015). This indicates that different broad epigenetic marks play different roles in regulating gene expression.

SUPER-ENHANCER DATABASE

With research interest on super-enhancers increase, more and more super-enhancers are defined in different cell types. Several tools have emerged to integrate and annotate those super-enhancers. Super-Enhancer Archive (SEA), integrated 83996 super-enhancers in 134 cell types or tissues or diseases in different organisms, such as human, mouse, *Drosophila* and *C. elegans* (Wei et al., 2016). Most of these super-enhancers are computationally identified and some of them are experimentally validated. SEA can support multiple search options, which provide annotations like TF binding sites, CRISPR/Cas9 target sites, SNPs and so on.

Another super-enhancer data base called dbSUPER also integrates super-enhancers in more than 100 cell types (Khan and Zhang, 2016). Different from SEA, users can upload their data to dbSUPER to find and compare super-enhancers. These data can also be easily sent to other tools for further analysis, such as GREAT and Cistrome web-servers.

The above databases provide a useful resource for exploring the function of super-enhancers.

SUMMARY

By definition, super-enhancer is a large cluster of enhancers. Some super-enhancer contains only one constituent enhancer, which means that super-enhancers are not necessarily clusters of enhancers. The current definition of a super enhancer as a large domain of peaks is operational rather than functional. A functional definition of super enhancers still awaits further mechanistic studies of them in many different biological contexts. Moreover, what is the role of each individual enhancer within super-enhancer, and how do they cooperate together to regulate gene expression are still questions to be answered.

With the development of CRISPR/Cas9 technology, it is now possible to investigate the cooperative roles of individual enhancers within super-enhancers by specifically knocking-out an individual enhancer. Also 3D genome technology will depict a more structural and mechanistic view of super-enhancer in regulating gene expression.

Adam, R.C., Yang, H., Rockowitz, S., Larsen, S.B., Nikolova, M., Oristian, D.S., Polak, L., Kadaja, M., Asare, A., Zheng, D., and Fuchs, E. (2015). Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature 521, 366–370.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27, 299–308.

Benayoun, B.A., Pollina, E.A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E.D., Devarajan, K., Daugherty, A.C., Kundaje, A.B., Mancini, E., Hitz, B.C., Gupta, R., Rando, T.A., Baker, J.C., Snyder, M.P., Cherry, J.M., and Brunet, A. (2014). H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell 158, 673–688.

Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X., Shi, X., Wang, Q., and Li, W. (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet 47, 1149–1157.

Chipumuro, E., Marco, E., Christensen, C.L., Kwiatkowski, N., Zhang, T., Hatheway, C.M., Abraham, B.J., Sharma, B., Yeung, C., Altabef, A., Perez-Atayde, A., Wong, K.K., Yuan, G.C., Gray, N.S., Young, R.A., and George, R.E. (2014). CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. Cell 159, 1126–1139.

Grosveld, F., van Assendelft, G.B., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51, 975–985.

Hah, N., Benner, C., Chong, L.W., Yu, R.T., Downes, M., and Evans, R.M. (2015). Inflammation-sensitive super enhancers form domains of coordinately regulated enhancer RNAs. Proc Natl Acad Sci USA 112, E297–302.

Harding, K., and Levine, M. (1988). Gap genes define the limits of antennapedia and bithorax gene expression during early development in *Drosophila*. EMBO J 7, 205–214.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell 155, 934–947.

Hnisz, D., Schuijers, J., Lin, C.Y., Weintraub, A.S., Abraham, B.J., Lee, T.I., Bradner, J.E., and Young, R.A. (2015). Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell 58, 362–370.

Khan, A., and Zhang, X. (2016). dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic Acids Res 44, D164–171.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334.

Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Etchin, J., Lawton, L., Sallan, S.E., Silverman, L.B., Loh, M.L., Hunger, S.P., Sanda, T., Young, R.A., and Look, A.T. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science 346, 1373–1377.

Wei, Y., Zhang, S., Shang, S., Zhang, B., Li, S., Wang, X., Wang, F., Su, J., Wu, Q., Liu, H., and Zhang, Y. (2016). SEA: a super-enhancer archive. Nucleic Acids Res 44, D172–179.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319.