What is the role of superenhancers?

Kevin Stachelek

10 November, 2018

# Enhancers

Enhancers are cis-regulatory elements with lengths of 1e2-1e3 bp that enhance transcription over long genomic distances. Enhancers regulate spatiotemporal gene expression. Enhancers also require transcription factor binding and histone modification for their funciton but how much is unclear (Barakat et al., 2018)

## Can we predict enhancers?

Several groups have predicted ESC enhancers on the basis of histone marks, TF binding, or DNaseI hypersensitivity (Hawkins et al., 2011; Rada-Iglesias et al., 2011; Xie et al., 2013). However, though transcription factor (TF) binding is linked with enhancer activity, no individual TF, histone mark or combination can unequivocally predict enhancer activity (Barakat et al., 2018).

## what is the relationship between TADS and TEs?

There is a class of non protein-coding genomic sequences that have high conservation between species, conserved noncoding elements (CNEs) (Polychronopoulos et al., 2017). CNEs are clustered in hte vicinity of genes with regulatory roles in development and differentiation (Polychronopoulos et al., 2017). In addition, CNEs are organized into genomic regulatory blocks which regularly overlap with TADs (Polychronopoulos et al., 2017). The disruption of CNEs has been found to contribute to developmental diseases and cancer (Polychronopoulos et al., 2017).

## How do transposable elements relate to enhancers?

Many enhancers are found in dense clusters of CNEs. Often these CNEs display higher levels of sequence conservation than the neighboring enhancers even though the core conserved sequence of the enhancer is enough to drive gene expression (Polychronopoulos et al., 2017). Rather than functioning as enhancers themselves, CNEs may serve as resources for regulatory genomic sequences. There is evidence that transposable elements (TEs) have provided a major source of genomic variation for human regulatory elements over evolutionary time (Jacques et al., 2013). There is a wide diversity and slow mutational decay of TEs in mammals. This may have promoted co-option of repetitive sequences for regulatory functions, *exaptation* (Feschotte, 2008).

Frequently CNEs are enriched in open chromatin regions where regulatory sequences are found. Jacques et al. (2013) used DNase I hypersensitivity data sets from ENCODE in normal, embryonic, and cancer cells to show that 44% of open chromatin regions 63% of primate-specific open-chromatin regions are in TEs. While Barakat et al. (2018) found that within SEs, large tracts are nonfunctional, restricted to small subdomains. It is interesting to consider patterns of superenhancer function in the context of TEs. Both represent long tracts of nonfunctional sequence closely associated with functional subdomains.

## Relationship betwene gene expression and promoter-enhancer interaction

One of the key features of enhancers among cis-regulatory elements, is their lack of preference for proximal genomic sequences. Despite typically long-range binding interactions there is a correlation between gene expression, promoter-enhancer interactions and enhancer RNAs (Sanyal et al., 2012).

These long-range interactions have a bias for elements located ~120 kB upstream of the TSS (Sanyal et al., 2012). Binding sites are often not blocked by sites bound to CTCF and cohesin. So many do not demarcate physically insulated gene domains (Sanyal et al., 2012). Only 7% of looping interactions involve the nearest gene. Genomic proximity is not a straightforward predictor for long-range interactions (Sanyal et al., 2012).

# Superenhancers

## Definition of super-enhancers

SEs are dense transcription factor platforms enriched for active chromatin marks (Shin et al., 2016). Compared to traditional enhancers, SEs have increased binding of Mediator, specific histone marks, and lineage-specific TFs (Barakat et al., 2018,@Hay2016,@Whyte2013). Cell identity is also thought to be tightly bound to super-enhancer activity (Hnisz et al., 2013). Disease-associated variation is enriched in super-enhancers of disease-relevant cell types (Hnisz et al., 2013). Cancer cells generate super-enhancers at oncogenes and other drivers (Hnisz et al., 2013).

## Model of Super-enhancer function

One model of super-enhancer activity proposes that SE constituents act additively and independently. Hay et al. (2016) Generated a series of mouse models. They then deleted each of the five regulatory elements of the a-globin super-enhancer individually and in combination. Each constituent enhancer of the a-globin super-enhancer was found to act independently and additively (Hay et al., 2016). Overall, the a-globin super-enhancer specifically impacts hematological phenotype, gene expression, chromatin structure, and chromosome conformation (Hay et al., 2016).

It is also possible to investigate SE function on a global scale by coupling enhancer binding with a funcitonal assay. Barakat et al. (2018) combined chromatin immunoprecipitation (ChIP) with a parallel reporter assay to identify functional enhancers in ESCs, termed Chip-STARR-Seq. By comparing regions bound by NANOG, OCT4, H3K27ac, and H3K4me1 with patterns of functional enhancers it was found that only a small fraction of bound regions exhibited enhancer activity (Barakat et al., 2018).

Out of this set of funcionally verified enhancers, Barakat et al. (2018) also excluded previously predicted ESC enhancers. They then associated their remaining set (The extended enhancer module) with generic cell processes. This extended enhancer module is therefore characterized by reduced binding of pluripotency associated TFs and histone marks.

Alignment of ChIP-STARR-seq data to SEs showed correlation between reads per plasmid (RPP) and H3K27Ac intensity (Barakat et al., 2018).

They then considered the FGFR1 SE specifically. Using Crispr/Cas9 they modified regions of the FGFR1 SE and observed effects via ChIP-STARR-seq, concluding that the FGFR1 SE is composed of small units with enhancer activity (Barakat et al., 2018).

After looking at global patterns of reads per plasmid (RPP) it was found that enhancers contain only a small percentage of active plasmids in their bounds. The fraction is only slightly higher in SEs (Barakat et al., 2018). It therefore appears that enhancer activity is limited to small domains within the SE that frequently overlap with TF binding sites. Further, the observed chromatin signatures at SEs, such as H3K27Ac, might be a consequence of enhancer activity from much smaller units.

Several models of SE activity have been proposed. SE constituents might function: 1) alternatively as independent and additive enhancers (Hay et al., 2016, @Moorthy2017) 2) in a temporal and functional hierarchy (Shin et al., 2016) 3) as interdependent units (Hnisz et al., 2013)

Finally, it is imporant to note that, like many techniques for study of regulatory sequences, Chip-STARR-Seq does not fully recapitulate the endogenous chromatin context. This may be a critical variable.

## Superenhancers targets for Cancer Therapy

Chromatin regulators have been used widely as targets for cancer therapy (Lovén et al., 2013). But why is it that inhibition of chromatin regulators has gene-specific effects in tumor cells? Lovén et al. (2013) studied how inhibition of BRD4 can lead to selective inhibition of myc in multiple myeloma (MM). They found that BRD4 and the SE associated Mediator protein co-occupy 1e3 enhancers associated with active genes (Lovén et al., 2013).

Interestingly, they found that BRD4 and Mediator co-occupy a few very large super-enhancers associated with MYC and other genes. Treatment with JQ1 (BED-bromodomain inhibitor) leads to loss of BRD4 at super-enhancers and transcription elongation defects at super-enhancer genes (Lovén et al., 2013).

## What is the role of Superenhancers in ESC development?

Early studies demonstrated that transcription factors Oct4, Sox2, and Nanog bind enhancers and recruit Mediator to activate gene expression in ESCs (Whyte et al., 2013). Activation follows the formation of enhancer domains at genes that control pluripotent state (Whyte et al., 2013). Reduced levels of Oct4 or mediator cause loss of super-enhancers relative to other genes (Whyte et al., 2013). on the other hand, in differentiated cells superenhancers containing cell-type-specific master TFs are found at genes that define cell identity (Whyte et al., 2013).

## Mechanism of regulation of super-enhancers

### Cohesin

#### Role of Cohesin

Superenhancers are distinguished by the binding of the Cohesin protein, a factor which mediates sister chromatid cohesion during the cell cycle. Cohesin also forms a complex with CTCF and active gene regulatory elements to form long-range interactions between binding sites (Ing-Simmons et al., 2015). Chromosome conformation capture shows that cohesin’s main role in interphase is mediating interactions with chromosome compartments [Ing-Simmons2015]. Cohesin is especially enriched at SEs (Ing-Simmons et al., 2015). Global and local chromosome conformation capture experiments have shown that enhancer elements associate both linearly and in three dimensions (Ing-Simmons et al., 2015).

3D clustering of SEs is facilitated by cohesin. Conditional deletion of cohesin from cells preserves enhancer position and transcription but weakens interactions between enhancers and promoters (Ing-Simmons et al., 2015). Half of all deregulated genes following cohesin deletion lie in the vicinity of enhancer elements, so it is likely that cohesin regulates gene expression through spatial clustering.

# Comparing Enhancers and Superenhancers

## Structural differences enhancers v. superenhancers?

Super-enhancers consist of clusters of enhancers that are occupied by master regulators and Mediator (Whyte et al., 2013). There are also structural features that distinguish SEs from other enhancers. Super-enhancers differ from enhancers by size, and transcription factor density. SEs can also be distinguished by their ability to activate transcription, and sensitivity to perturbation (Whyte et al., 2013).

In the context of ESCs, enhancers and super-enhancers have equivalent regulatory roles by regulation of single or multiple genes. Trans enhancers maintain cell-type-specific expression. They drive cell fate changes during development (Moorthy et al., 2017). Highly expressed genes are associated with individual enhancers, ex. locus control regions (Moorthy et al., 2017). It is predicted that super-enhancers regulate critical cell-identity genes.

The effect of deletings enhancers is variable and results in gene expression changes from 12-92% (Moorthy et al., 2017). Partial deletion that removes only a sub cluster within an SE results in limited change in expression. This demonstrates redundant control by multiple enhancers within a super-enhancer. However, highly expressed genes in ESCs are not necessarily associated with a super-enhancer (Moorthy et al., 2017).

Super-enhancer prediction ignores 81% of active regulatory elements predicted by co-binding of critical pluripotency-associated TFs.

Finally, in certain cases, super-enhancers regulate clusters of paralogous genes (significance?) (Moorthy et al., 2017).

### Superenhancers in Mammary Tissue Development

It has been demonstrated that SEs are critical in development and cell fate specification. A study of SE function in the mammary gland, which is dynamically regulated during pregnancy identified 440 Mammary-specific super-enhancers with half associated with genes activated during pregnancy (Shin et al., 2016). By mutating STAT5 superenhancer binding sites, they showed that the most distal site had the greatest enhancer activity (Shin et al., 2016). While mutation analysis showed that a 1e3 fold induction during pregnancy relied on all enhancers (Shin et al., 2016). Disabling the binding sites of STAT5, NFIB, ELF5 in the proximal enhancer incapacitated the super enhancer (Shin et al., 2016).

Thus, there appears to be a temporal and functional enhancer hierarchy in the context of SEs(Shin et al., 2016).

# TADs

## Definition of TADs

The nuclear organization of chromatin is an active area of research. High-throughput methods of chromosome conformation capture show that interphase chromatin is partitioned into Mb-sized compartments and sub-Mb-sized topological domains (Seitan et al., 2013). Dixon et al. (2012) investigated the three dimensional organization of human and mouse ESCs and terminally differentiated cells using HI-C technology. Megabase-sized chromatin interaction domains are termed ‘topological domains’ (Dixon et al., 2012). These topological domains correlate with sites of heterochromatin constraint and are stable across cell types and conserved across species (Dixon et al., 2012). It is thought that compartmentalization facilitates gene and reg. element matching (Seitan et al., 2013).

## How are TADs demarcated?

The boundaries of topological domains are enriched for the insulator CTCF, housekeeping genes, tRNAs and SINE retrotransposons (Dixon et al., 2012). While non-promoter associated marks, such as H3K4me1 (associated with enhancers) are not enriched or are specifically depleted at boundary regions (Dixon et al., 2012).

## How are TADs formed?

TADs are formed through the action of Cohesin which controls chromosome topology to enable DNA repair and chromosome segregation in cycling cells (Seitan et al., 2013). It also associates with active enhancers and promoters and with CTCF to form long-range interactions for gene regulation (Seitan et al., 2013). Architectural compartments can be maintained in non-cycling mouse thymocytes after genetic depletion of cohesin in vivo (Seitan et al., 2013). It can be show that Cohesin is required for specific long-rang interactions within TADs where cohesin-regulated genes are found (Seitan et al., 2013). It also seems that Cohesin depletion diminishes interactions between cohesin-bound sites by making alternative interactions between transcriptional activation and repression features more prominent–resulting in changes in gene expression (Seitan et al., 2013).

Ultimately, Cohesin-mediated long-range interactions facilitate discrete gene expression states within preexisting chromosomal compartments (Seitan et al., 2013).

Additional work with chromsome conformation capture tehcnology investigated whether variation in chromatin conformation within TADs?

* Giorgetti et al. (2014) developed a polymer model to extract the full repertoire of chromatin conformations within TADs from population-based 3C data.

## How are TADs important in promoter-enhancer interactions?

* Chen et al. (2018) combined genome editing and multi-color live imaging to visualize physical enhancer-promoter interaction in single cells.
* Identify three topological states with different transcription kinetics. Sustained proximity of the enhancer to its target is required for activation (Chen et al., 2018).
* transcription affects 3d topology by enhancing temporal stability of the proximal conformation and spatial compaction (Chen et al., 2018).
* Long range activation results in transcriptional competition at the locus (Chen et al., 2018).

## How does enhancer contact relate to TAD structure?

* The model developed by Giorgetti et al. (2014) predicts physical distances and chromosomal contact variation among cells. It also identifies interactions within single TADs that stabilize boundaries between TADs and allows the identification and validation of structural elements within TADs (Giorgetti et al., 2014).
* Combining polymer model with DNA FISH and RNA FISH in Xic, Giorgetti et al. (2014) dissected the relationship between transcription and spatial proximity to cis-regulatory elements.
* Contacts between potential regulatory elements occur in the context of fluctuating structures rather than stable loops and such fluctuations may contribute to asymmetric expression in Xic (Giorgetti et al., 2014).

## How do TADS mediate enhancer-promoter interaction?

### enhancer-promoter proximity mediated by TADS may be necessary but not sufficient for expression

* Limb-specific Shh expression is regulated by the ZRS enhancer. Williamson et al. (2016) analyzed the higher-order chromatin conformation of Shh in expressing and non-expressing tissues by FISH and 5C.
* Light microscopy identified elevated frequencies of Shh/ZRS colocalization only in the Shh-expressing regions of the limb bud due to enhancer-promoter loop formation.
* in all tissues at all stages Shh-ZRS spatial distance are shorter than the distance to an intervening neural enhancer (Williamson et al., 2016).
* 5C identified a TAD over the Shh/ZRS genomic region and enriched interactions between Shh and ZRS through E111.5 (Williamson et al., 2016).
* Shh/ZRS colocalization correlates with spatiotemporal limb-bud specific Shh expression, but close Shh and ZRS proximity in the nucleus occurs regardless of whether the gene or enhancer is active (Williamson et al., 2016). So expression is not specifically regulated by enhancer-promoter proximity in this case. Rather, constrained chromatin configuration optimizes the opportunity for the active enhancer to locate and instigate the expression of Shh (Williamson et al., 2016).

## How does interTAD arrangement impact expression of border genes?

* The transcriptional activation of Hoxd genes during vertebrate digit development involves modifications in 3D interactions within and around the HoxD gene cluster (Fabre et al., 2015).
* Fabre et al. (2015) used 3D DNA FISH to assess the spatial organization of chromatin at and around the HoxD gene cluster. Although the two TADs are tightly associated, they appear as spatially distinct units.
* Fabre et al. (2015) measured the relative position of genes within a cluster and find that they segregate over long distances, so a physical elongation of the HoxD cluster can occur.
* Fabre et al. (2015) used STORM to find that tissues with distinct transcriptional activity exhibit differing degrees of elongation. The spatial separation of Hoxd genes is strongest in those tissues where they are highly expressed.
* The morphological change of the HoxD cluster in developing digits is associated with its position at the boundary between two TADs (Fabre et al., 2015).
* Variations in fine-scale architecture of gene clusters suggest causal links among spatial configuration, transcriptional activation, and flanking chromatin context (Fabre et al., 2015).
* In the case of HoxD, all enhancers sharing a particular specificity are found within the same TAD (Fabre et al., 2015).
* So a model is suggested wherein reorganization follows a global transition from one set of regulatory contacts to another–between two TADs located on either side of the HoxD locus (Fabre et al., 2015).

## Long

* “increased numbers of TF motifs at an enhancer should be positively correlated with increased nucleosomal eviction, DNA binding, and, consequently, elevated gene-expression output (Long et al., 2016).”
* “at some developmental enhancers, there is evidence for step-wise licensing by lineage-determining TFs, also known as master regulators, or pioneer factors, which directly bind nucleosomal DNA to prime enhancers for activation (Long et al., 2016).”
* “Thus, combinatorial deployment of multiple sub-optimal enhancers may help to promote specificity of regulation without sacrificing signal strength.”
* “One hypothesis is that enhanceosomes may have potential roles in mediating switch-like transcriptional activation of genes requiring strict regulatory inputs from multiple signaling path- ways (Spitz and Furlong, 2012). Consistent with this theory, deeply conserved enhancers tend to fall as clusters near genes encoding developmentally important transcription factors”

## References

Barakat, T.S., Halbritter, F., Zhang, M., Rendeiro, A.F., Perenthaler, E., Bock, C., and Chambers, I. (2018). Functional Dissection of the Enhancer Repertoire in Human Embryonic Stem Cells. Cell Stem Cell 276–288.

Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nature Genetics *50*, 1296–1303.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Fabre, P.J., Benke, A., Joye, E., Nguyen Huynh, T.H., Manley, S., and Duboule, D. (2015). Nanoscale spatial organization of the <i>HoxD</i> gene cluster in distinct transcriptional states. Proceedings of the National Academy of Sciences *112*, 13964–13969.

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. Nature Reviews Genetics *9*, 397–405.

Giorgetti, L., Galupa, R., Nora, E.P., Piolot, T., Lam, F., Dekker, J., Tiana, G., and Heard, E. (2014). Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. Cell *157*, 950–963.

Hay, D., Hughes, J.R., Babbs, C., Davies, J.O., Graham, B.J., Hanssen, L.L., Kassouf, M.T., Oudelaar, A.M., Sharpe, J.A., Suciu, M.C., et al. (2016). Genetic dissection of the -globin super-enhancer in vivo. Nature Genetics *48*, 895–903.

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

Ing-Simmons, E., Seitan, V.C., Faure, A.J., Flicek, P., Carroll, T., Dekker, J., Fisher, A.G., Lenhard, B., and Merkenschlager, M. (2015). Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. Genome Research *25*, 504–513.

Jacques, P.É., Jeyakani, J., and Bourque, G. (2013). The Majority of Primate-Specific Regulatory Sequences Are Derived from Transposable Elements. PLoS Genetics *9*.

Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. Cell *167*, 1170–1187.

Lovén, 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.

Moorthy, S.D., Davidson, S., Shchuka, V.M., Singh, G., Malek-gilani, N., Langroudi, L., Martchenko, A., So, V., Macpherson, N.N., and Mitchell, J.A. (2017). Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. Genome Res *27*, 246–258.

Polychronopoulos, D., King, J.W., Nash, A.J., Tan, G., and Lenhard, B. (2017). Conserved non-coding elements: Developmental gene regulation meets genome organization. Nucleic Acids Research *45*, 12611–12624.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109–113.

Seitan, V., Faure, A., Zhan, Y., McCord, R., Lajoie, B., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A., et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within pre-existing architectural compartments. Genome Research *23*, 2066–2077.

Shin, H.Y., Willi, M., Yoo, K.H., Zeng, X., Wang, C., Metser, G., and Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nature Genetics *48*, 904–911.

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.

Williamson, I., Lettice, L.A., Hill, R.E., and Bickmore, W.A. (2016). <i>Shh</i> and ZRS enhancer colocalisation is specific to the zone of polarising activity. Development *143*, 2994–3001.