Lesson 2.

1) Functional SNP has a role in disease outcome.

Variation in single nucleotide in the genomic regulatory regions may have an impact in gene expression modifying molecular mechanisms that are involved in disease outcome.

2) How SNPs play a FUNCTIONAL role in disease:

Changing consensus sequences for transcription factors binding sites

Changing interaction between transcription factors

Changing epigenetic profiling of specific genomic regions

Changing long range interaction between two genomic regions

3) Pictures that illustrate these concepts

4) How SNPs is involved in disease

5) In this Lesson:

Enhancer Overview

Genomic regulatory network to define cell identity

Genetic variations meaning in cell identity

6) Genomic regulatory regions.

Enhancers are activated through interactions with transcription factors, which recognize and bind to specific DNA sequences within the enhancer region. Bound transcription factors recruit co-regulators, many of which deposit or remove modifications on histones. The way in which each cell type interprets genomic information is therefore closely linked to the organization of its DNA regulatory elements. Enhancers that are active in cell-type-specific epigenomic signatures are typically highly enriched in DNA sequences to which lineage-determining and signal-dependent transcription factors bind. Therefore, the delineation of a particular cell's active enhancer repertoire provides a powerful means of predicting the transcription

factors required for that cell identity. By extension, changes in epigenomic signatures during developmental transitions reflect activation or inhibition of such factors.

Enhancer Characteristitcs:

Enhancers are cis-regulatory elements in proximity of genes

Each cell has a set of enhancers

Enhancers have motifs for sequence-specific transcription factors

Enhancers are marked with epigenetic modifications

Enhancers are in different states of activation

13) Cell type specific Enhancers.

In each cell, the pattern of enhances is specific. In this picture, enhancers are activated for regulation of specific gene. In cell A, enhancer activated gene transcription, while other set of enhancers regulated genes in cell B. A cluster of enhancers, called super-enhancers, can regulate gene in cell A, while other set of SE regulated gene in cell B. Altogether cell type specific enhancers work to define cell identity (morphology and biological function). Enhancers pattern activation in cell type show a regulatory code, genomic regions used for gene expression.

14) Cell type specific Enhancers for gene expressed in different type of cells.

Enhancers contain specific recognition sequences required for binding of transcription factors (TFs) that regulate gene expression in a spatial and temporal fashion.

In the genome, we can find several enhancers associated with one gene. However, a small set of enhancers can be used for gene expression. In this picture, in the specific area of the embryo (marked in red and green) one gene is expressed thanks to transcription factors bound at red enhancer for red area or TFs bound at green enhancer for green area. Therefore, one gene may be regulated by cell-type specific enhancers activation.

15) Cell type specific Enhancers activation during cell development.

Enhancer activation for gene expression regulation depend on stages of cell differentiation. One gene can be expressed thanks to red enhancer in a specific tissue only for first stage of embryo growth, while green enhancer activation controls

gene expressed in green area in second stage of embryo development. Therefore, the enhancer activation depends on specific transcription factors binding and time of cell differentiation.

16) Cell type specific Enhancers selection depend on lineage-determining TFs.

The selection of a large fraction of cis-acting regulatory elements for enhancer activity appears to depend on relatively simple combinations of the lineage-determining transcription factors (LDTFs) for cell type identity. Through cooperative binding to closely spaced recognition motifs, LDTFs initiate binding to otherwise closed chromatin and facilitate chromatin accessibility by recruitment of ATP-dependent nucleosome remodeling complexes. Histone modifiers are recruited to deposit histone marks that demarcate enhancer regions (e.g., monomethylation of histone 3 lysine 4, H3K4me1). Although a large fraction of enhancers requires cooperative binding of TFs, another group of enhancers is activated by sequential binding of TFs at different development stages.

17) Active Enhancers are associated with chromatin states features.

Chromatin states are defined with histone modifications and DNA methylation. We can view regions where there is an active transcriptional activity for gene expression. Histone marks associated with active promoter are H3K4me3, in enhacers H3K4me1 and H3K27Ac. While repressed regions are marked with H3K27me3. So we can distinguish euchromatin and heterochromatin as genomic regions that control gene expression and, therefore, cell activity.

18) Each cell has a specific set of active enhancers.

This picture recapitulates the concept that I explain. If we consider genome without epigenetic modifications, each cells have the same DNA sequence, but the chromatin features signed genomic regions that regulated gene expression. Epigenetic modifications is marked with circles and you can have the perception of cell-type specific regulatory core that is described by enhancers pattern and lineage-determining TFs binding. An addition level of regulation is long-range interactions between genomic regions that form loop and regulatory complexes. How genomic regions interact can control a several set of cells.

19) Long-range interactions form chromatin hubs for gene regulation.

The formation of a chromatin hub by looping of distant sequences has also been detected in the T helper type 2 (TH2) cytokine locus. **Long-range interactions** among multiple enhancer and promoter elements within the locus **increases** significantly upon **TH2 cell differentiation**, suggesting that the three-dimensional organization of poised co-regulated genes is important for rapid activation upon induction. The ability of enhancers to interact with promoters is not limited to genes located in cis on the same chromosome as the enhancer; the olfactory H enhancer has been shown to interact with multiple olfactory receptor genes both on the same and different chromosomes in epithelial tissues, where these genes are specifically expressed.

20) Long-range interactions are mediated by protein complexes.

Enhancers are the platform where TFs find sequence consensus and after this first binding, TFs recruit other cofactors that may form a bridge between two regions or enhancer-promoter. You can view a image for refresh the notion of TFs binding. In the genomic regions we can identify TFs binding using ChIP with antibodies against specific TF, so we view a peak corresponding with the reads, DNA fragments that are bound from our TF of interest. If we compare TFs binding sites with epigenetic profiling we observed TF binding associated with genomic regions marked with histone marks for active chromatin, while no binding is associated with repress chromatin.

TFs associated with active enhancers.

LDTF: lineage-determining TFs

SDTF: signal- determining TFs

CTF: collaborating TFs

focus LDTF.

We have looked as LTDF working for tissue specific gene expression.

Enhancer Activation.

Regulatory genomic regions are active in association with gene transcription, while regulatory genomic regions inactive is associated with silence gene.

Enhancers dynamic activation.

The characteristic features of poised and active enhancers are shown, including the binding of lineage-determining transcription factors (LDTFs) and collaborating

transcription factors (CTFs) to closely spaced recognition motifs (yellow and blue sites, respectively) on the DNA. a | The binding of these factors in concert with nucleosome-remodelling complexes (NRCs) initiates nucleosome displacement to form narrow nucleosome-free regions at poised enhancers. Histone remodeler enzymes deposit the active histone H3 lysine 4 monomethylation (H3K4me1) and H3K4me2 marks, whereas the histone-lysine N-methyltransferase EZH2 (a component of the Polycomb complex) deposits repressive H3K27me3 marks, and histone deacetylase (HDAC)-containing complexes maintain histones in a repressed, deacetylated state. RNA polymerase II (Pol II) is either absent or found at low levels at poised enhancers. In response to various cues, signal-dependent transcription factors (SDTFs) associate with recognition motifs in close association with LDTFs, which results in additional nucleosome displacement, as observed by widening of the DNase I-hypersensitive sites. SDTFs recruit co-activator complexes containing histone demethylase (HDM) complexes that remove H3K27me3 marks, histone acetyltransferases (HATs) that deposit H3K27 acetylation (H3K27ac) marks, and the Mediator complex (MED). The transformation to elongating Pol II results in bidirectional transcription — a hallmark of active enhancers — and the generation of enhancer RNAs (eRNAs), which is closely coupled to enhancer activity.

Enhancers activation upon external stimuli.

During gene transcriptional activation events in response to differentiation or environmental cues [e.g., lipopolysaccharide (LPS) stimulation in macrophages], a group of pre-existing enhancers (highlighted green) bearing mono- or di-methylation of histone lysine 4 (H3K4me1/2) modification is readily activated. A small subset of enhancers is generated de novo in response to LPS stimulation resulting in deposition of H3K4me1/2 enhancer marks. This group of enhancers, highlighted in blue, will be readily activated upon repeated stimulation.

Chromatin transitions to active enhancers involve interactions between cell lineage-determining transcription factors and signal-dependent factors. Enhancers primed by lineage-determining factors frequently require signal-dependent transcription factor binding to gain H3K27ac and become active. Active enhancers can also be selected by interactions between signal-dependent factors and lineage-determining factors.

Environment differentially activates common primed enhancers to drive selection of tissue-specific enhancers. Top panel, LPMs residing in the peritoneal cavity respond to RA, which activates primed RA-responsive enhancers. These in turn induce the expression of direct RA target genes that include transcription factors (e.g.,

GATA6 and RARb) that collaborate with PU.1 to select and activate LPM-specific enhancers. Lower panel, MG residing in the brain respond to TGFb, which activates primed SMAD-responsive enhancers. These in turn induce the expression of direct SMAD target genes that include transcription factors (e.g., Mafb) that collaborate with PU.1 to select and activate MG-specific enhancers

Enhacers Selection

Definition of TFs that regulate enhancer selection

The selection of a large fraction of cis-acting regulatory elements for enhancer activity appears to depend on relatively simple combinations of the lineage-determining transcription factors (LDTFs) for cell type identity. Through cooperative binding to closely spaced recognition motifs, LDTFs initiate binding to otherwise closed chromatin and facilitate chromatin accessibility by recruitment of ATP-dependent nucleosome remodeling complexes. **SDTF and CFT.**

Histone modifiers are recruited to deposit histone marks that demarcate enhancer regions (e.g., monomethylation of histone 3 lysine 4, H3K4me1). Although a large fraction of enhancers requires cooperative binding of TFs, another group of enhancers is activated by sequential binding of TFs at different development stages. A group of nucleosome-avid 'pioneering' TFs, exemplified by Forkhead box A (FoxA) and GATA binding protein (GATA), initiate this sequential binding to condensed chromatin to prime a 'soon-to-be' enhancer. Binding of either LDTFs or pioneering factors permits the binding of other TFs and cofactors upon signal transduction events resulting in cell-type specific, signal-dependent gene expression. Recent findings that enhancers can be transcribed added a new layer of complexity to gene regulation. In addition to mRNAs, ncRNAs represent a highly functional class of molecules because they can possess enzymatic activity (e.g., spliceosomes, ribozyme), have structural roles (e.g., tRNA, ribosomes), and transcriptional functions (e.g., lncRNAs). Understanding how RNAs contribute to enhancer function has thus become an area of active interest.

Cell type-specific enhancer selection and activation

Collaborative interactions between lineage-determining transcription factors (LDTFs) and collaborating transcription factors (CTFs) select enhancers for binding and

activation by signal-dependent transcription factors (SDTFs). Prior to signal-dependent activation, such regions may be poised enhancers or exhibit basal enhancer activity (that is, they are pre-existing enhancers) that is further induced by the binding of a SDTF. The resulting transcription is cell type-specific because the enhancers are selected by the cell type-specific LDTFs. b | SDTFs can direct the selection of latent or de novo enhancers. In these cases, the SDTF functions as an essential CTF to LDTFs to enable concurrent binding of all factors involved. The transcriptional output is cell type-specific because of the requirement for cell type-specific LDTFs for enhancer priming. H3K27ac, histone H3 lysine 27 acetylation.

In activated macrophages, the nuclear factor-κB (NF-κB) subunits p50 and p65, which are signal-dependent transcription factors (SDTFs), and the lineage-determining transcription factor (LDTF) PU.1 collaboratively select de novo enhancers. The subsequent recruitment of histone acetyltransferases (HATs) results in histone acetylation, a mark that is subsequently bound by the bromodomain-containing protein 4 (BRD4)–positive transcription elongation factor-b (P-TEFb) complex, which allows its cyclin-dependent kinase 9 (CDK9) component to phosphorylate the carboxy-terminal domain (CTD) of Pol II. The phosphorylated CTD acts as a docking site for the myeloid/lymphoid or mixed-lineage leukaemia protein 3 (MLL3) and MLL4 histone H3 lysine 4 (H3K4) methyltransferases, which are proposed to deposit H3K4 monomethylation (H3K4me1) and H3K4me2 marks during successive rounds of Pol II elongation.

Example of Enhancers Selection for the definition of cell identity.

During the T cell differentiation epigenome changes because each cell type has a specific biological functions, therefore activates a set of genes. Cell-types specific epigenome derives from antigen exposure and cytokines present in extracellular space. Antigen-bound receptor and cytokines repertoires induce a transduction pathways that remodel chromatin. The chromatin remodeling depend on epigenetic modifications, combinatorial TFs and non coding RNA.

Each step of differentiation is associated with specific TFs that control enhancers activation panel.

On the other hand,

Cytokines and antigen activates a signal transduction pathway where TFs orchestrate gene expression. Between genes important for cell identify, there are master TFs. Master TFs are LDTF that define a specific cell types. However, a complexes TFs cooperate for cell differentiation.

In summary,

T cell signaling and cytokines milieu induce pathway activation that involved cytokines expression in association with cytokines receptors expression, master TFs (or lineage-determining transcription factors) and combinatorial TFs that, altogether, orchestrate cell identity.

Non coding RNA.

A large portion of the human genome is transcribed into RNAs without known protein-coding functions, far outnumbering coding transcription units. Extensive studies of long noncoding RNAs (lncRNAs) have clearly demonstrated that they can play critical roles in regulating gene expression, development, and diseases, acting both as transcriptional activators and repressors. More recently, enhancers have been found to be broadly transcribed, resulting in the production of enhancer-derived RNAs, or eRNAs. Here, we review emerging evidence suggesting that at least some eRNAs contribute to enhancer function. We discuss these findings with respect to potential mechanisms of action of eRNAs and other ncRNAs in regulated gene expression.

Enhancer RNA.

Enhancer-derived RNA (eRNA) functionally contributes to enhancer-mediated coding gene expression. Signal-dependent activation of enhancers leads to increased production of eRNAs, which interact with looping factors (e.g., cohesin complex) and facilitate/stabilize chromosomal looping between the enhancer and promoter(s) of cognate target gene(s). eRNA mediates the loading of RNA PolII, and likely the transcription initiation complex (denoted by TFIID) at the promoter of the target gene. Whether chromosomal looping facilitates RNA PolII loading is unknown.

Signal-induced enhancers formation.

Transcription elongation is required for the deposition of di-methylation of histone lysine 4, or H3K4me2, also a histone mark for an enhancer, during signal-dependent activation of de novo enhancers. The coactivator complexes generate acetylation of histones, which recruits the positive transcription elongation factor (pTEFb, orange) to promote elongation of RNA PolII. Subsequently, the elongating PolII recruits histone methyl transferases, mixed lineage leukemia complexes (MLLs), for di-methylation deposition on H3K4 at enhancers.

ncRNAs collaborate with transcriptional coactivators.

Steroid receptor RNA activator (SRA) was the first ncRNA discovered to act as a transcriptional coactivator; it functions in the transcriptional activator complex steroid receptor coactivator-1 (SRC1) to enhance transcription mediated by steroid hormone receptors.

ncRNAs modulate chromatin loop.

This group of ncRNAs regulates gene activation through a distance by recruiting the Mediator complex to, and con-trolling the phosphorylation of histone H3S10 at its target gene promoter. Intriguingly, ncRNAs might, in specific circumstances, indirectly recruit complexes that exert functions in enhancer–promoter looping.

ncRNAs evict transcriptional repressors.

Example: Xist ncRNA

The Xist gene initiates XCI, and is only expressed when more than one X chromosome is present. At the pre-XCI stage, Xist is silenced by the binding of CCCTC-binding factor (CTCF) at its promoter. At the onset of XCI, a ncRNA transcript Jpx is upregulated from both X chromosomes, which directly binds to CTCF and titrates it away from the Xist promoter. This allows expression of Xist to initiate XCI on one of the X chromosomes.

Enhancer Function

Promoter activation requires that many components of the transcriptional machinery come together in order to assemble the pre-initiation complex, initiate transcription, overcome Pol II pausing and eventually lead to productive transcription elongation.

How does enhancer regulates gene expression?

- Formation of long range interactions
- A cooperation between several genomic regulatory regions is crucial for gene regulation

Chromatin looping

In the nucleus, the genome is organized and partitioned into functional compartments in 3D space, and considerable effort is being directed at understanding enhancer function in the context of 3D chromatin interactions. One strategy is to identify the long-range looping interactions that involve enhancer elements using a variety of

chromosome conformation capture (3C)-based techniques. Genome-wide applications of these techniques to define the chromatin interactomes of human and mouse cells confirmed that the genome is divided into active and inactive compartments. These are further organized into submegabase-sized topologically associated domains (TADs).

Promoters and distal enhancer elements are frequently engaged in multiple long-range interactions and form active chromatin hubs.

Super-enhancers definition.

Super-enhancers were initially defined as large (tens of kilobases in length) genomic loci with an unusually high density of enhancer-associated marks, such as binding of the Mediator complex, relative to most other genomic loci. These regions can also be defined by high-density and/or extended (>3 kb) depositions of the histone mark H3K27ac. Using differences in the density of Mediator complex-binding sites or of H3K27ac marks to distinguish super-enhancers from regular enhancers, most cell types are found to have 300–500 super-enhancers. A substantial fraction of super-enhancers and nearby genes are cell type-specific, and the gene sets that are associated with super-enhancers in a given cell type are highly enriched for the biological processes that define the identities of the cell types.

46) Genetic Variations link with genomic regions.
47) Sequencing approches

Profiling the complete chromatin landscape of all cells in the haematopoietic system is a formidable task. Within the main cell lineages, many cell subtypes have been defined and each one of these must be individually isolated and analysed. Taking into account external influences and the sensitivity of cells to the local environment, distinct cell populations — such as those in different tissues — should be analysed separately13,14. Furthermore, primary cells should be analysed immediately after isolation, as growth in culture following removal from their natural context is likely to influence the underlying chromatin regulatory landscapes. This was recently demonstrated for tissue-resident macrophages, which rapidly lost their tissue-specific chromatin signatures upon isolation and culture13.

The general approach of chromatin studies is to isolate and enrich for genome fragments that are associated with the chromatin feature of interest (for example, nucleosomes, open chromatin, DNA methylation or histone modifications) to create libraries for sequencing (FIG. 1). The challenge of traditional genome-wide studies is

that due to the low efficiency of certain stages of the protocol, a large starting population of cells is needed to produce the large number of unique sequences that are required to adequately sample all regulatory elements

47) Data interpretation. The next steps are to process and normalize the aligned sequences so that the results can be visualized (for example, in a genome browser). Bioinformatic analysis is carried out to interpret the data on a global level and to reconstruct the original chromatin state including: identifying genomic regions that are enriched for a particular chromatin feature (peak calling); characterizing differential regions of enrichment between cell types; annotating the regulatory elements in these regions; finding transcription factor-binding motifs; and reconstructing the regulatory network. Dashed horizontal lines represent the threshold for peak calling.

50) Cell-type-specific enhancers.

Alternative distal regulatory regions (enhancers) involved in lineage specification promote the transcription of the same gene in different cell types. The cell-type-specific transcription factors of B cells (blue) and macrophages (orange) bind in combination with PU.1 to active enhancers marked by histone 3 lysine 4 monomethylation (H3K4me1; green) and H3K27 acetylation (H3K27ac; red). Active promoters are indicated by H3K4me3 (purple) and RNA polymerase II binding.

51) Effect of the tissue environment.

The tissue environment affects the regulatory landscape of a cell through the induction of specific transcription factors, leading to the expression of genes that are likely to be involved in the unique functional pathways of each tissue-specific cell type: Sall1 and Siglech (sialic acid-binding immunoglobulin-like lectin H) for neuronal synapse pruning in microglia (brain-resident macrophages); Clec4f (C-type lectin domain family 4 member F) for erythrocyte recycling in Kupffer cells (liver-resident macrophages); and Car4 (carbonic anhydrase 4) and Chil3 for surfactant clearance in lung macrophages14. LXRα, liver X receptor-α; MEF2C, monocyte enhancer factor 2C; PPARγ, peroxisome proliferator-activated receptor-γ.

52) Association of human chromatin data and susceptibility to immune disease.

Cohort studies are designed to find sources of genetic variation between the control and disease groups, such as the single-nucleotide polymorphism (SNP) shown here located in an activator protein 1 (AP-1)-binding motif within the SMAD3 intron (indicated by the red letter 'T'). This SNP, which is associated with Crohn disease, disrupts the binding of AP-1 to an enhancer that is active in healthy monocytes. By comparing SNPs with chromatin profiles, we can determine whether they are located

in regulatory elements on or near transcription factor-binding sites in the relevant cell type. The resulting disruption of the chromatin state leads to altered gene transcription and provides the mechanism of disease. H3K4me1, histone 3 lysine 4 monomethylation; H3K27ac, H3K27 acetylation; H3K4me3, H3K4 trimethylation.

53)

Macrophages are important effector cells that reside in every tissue of the body [4]. Their diverse functions in different tissue environments as well as their essential roles in health and disease make them an important experimental system to study chromatin priming, signal integration, and cooperative interactions at enhancers. To this end, transcriptomes and primed and active enhancers were compared between macrophages resident in diverse tissues in mice [19–21]. **Different macrophage populations exhibited both a common program of core macrophage gene expression, as well as highly divergent patterns of gene expression that were specific to different tissue environments**. In parallel, each population of tissue macrophages exhibited both common and distinct sets of active enhancers (Figure 4A). Intriguingly, experiments in which specific populations of tissue macrophages were either placed in culture or were transplanted to another anatomic location demonstrated that marked changes occurred in both transcriptomes and enhancer landscapes [21]. These results indicate that **macrophage phenotypes are under constant environmental regulation and that local signals specify the active enhancer repertoire that controls context dependent gene expression**.

Motif enrichment analysis for cell-type-specific enhancers suggested distinct sets of transcription factors that bind to each macrophage enhancer subset. This approach, however, does not establish that the implicated factors participate in **enhancer selection and/or cooperative binding**. Enhancers of all macrophage subsets are highly enriched for the motif to which PU.1 binds, consistent with a requirement for PU.1 for the development of all tissue macrophages.

The observation that mutations in binding sites for C/EBP or AP-1 transcription factors could result in loss of PU.1 binding without mutations in the PU.1 binding site itself suggested that genetic variation between diverse inbred mouse strains could be used to discover **collaborative interactions between PU.1 and unknown lineage-determining factors** for each macrophage type.

Specifically, loss of PU.1 binding at a particular genomic location in one strain compared to another, without a mutation in the PU.1 recognition motif itself, could occur because of a mutation in the recognition motif for a collaborative transcription factor. Thus, characterization of the frequencies of mutations in all potential transcription factor recognition motifs in the vicinity of strain-similar versus strain-specific PU.1 binding could provide the basis for identifying collaborative factors important for PU.1 binding.

We have described enhancer under different expects, so SNP may have an impact in some ways.

SNPs in the genomic regulatory regions may affect:

Enhancer Activation: loss of TFs interaction or TFs recruitment.

Enhancer Selection: loss or association of LTDF

Alteration of timing or specific tissues activation

Long range interaction between genomic regulatory regions

Hnisz et al., 2015. Superenhancers in disease.

In this work authors identified the super-enhancers SE, a cluster of enhancers that regulate genes involved in cell identity and genes that act as oncogene.

Embrionic stem cells are the model that is used for genome-wide analysis.

Chromatin immunoprecipitation assay against Oct4, Nanog and Sox2 is used to identify SE because these TFs are master regulators for ESC, therefore regulate genes important in cell identity. Oct4, Nanog and Sox2 are TFs that maintain the pluripotency state in ESC cells while loss of their expression is associated with cell differentiation.
In the several diapo, I will show you the bioinfomatic analysis of ChIP-Seq for the definition of SE.
There is the characterization of TFs, cofactors and RNAPolII to make a model that show the tridimentional architecture of SE.
The work of ESC is extended to other cell types and cell-type specific SE for one type are 48%, SE common in two o three cell types decreased. This is a heatmap that show SE enrichment.

GENE ASSOCIATED TO SUPERENHANCER IN SEVERAL CELL TYPES: GENE ONTOLOGY

Gene ontology analysis of genes associated with SE show that , for example, in T cell the enrichment of biological functions associated with immune system response.

The **Gene Ontology project** provides controlled vocabularies of defined terms representing **gene product properties**.

One of the main uses of the GO is to perform enrichment analysis on gene sets.

For example, given a set of genes that are up-regulated under certain conditions, an enrichment analysis will find which GO terms are over-represented (or under-represented) using annotations for that gene set.

The results page displays a table that lists **significant shared GO terms (or parents of GO terms)** used to describe the set of genes that users entered on the previous page

P-value is the probability or chance of seeing at least x number of genes out of the total n genes in the list annotated to a particular GO term, given the proportion of genes in the whole genome that are annotated to that GO Term.

In the list of genes associated with SE, there are the lineage-determining TFs for each cell types.

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

Alzheimer's disease is a common form of dementia characterized by progressive neurodegeneration and loss of cognitive functions of the brain, and much of the genetic variation implicated in Alzheimer's disease is associated with amyloid precursor protein, transmembrane proteins, and apolipoprotein E4. The SNP catalog contains 27 SNPs linked to Alzheimer's disease, and five of these occur in the super-enhancers of brain tissue. Thus, 19% (5/27) of all of the Alzheimer's disease SNPs occur in the 1.4% of the genome encompassed by brain tissue superenhancers. Two SNPs occur in the super-enhancer associated with the gene BIN1, whose expression has recently been shown to be associated with Alzheimer's disease risk. Additional variation in the BIN1 super-enhancer, involving a small insertion, was shown to be associated with Alzheimer's disease.
In multiple sclerosis, disease based on T cell disfunction, show an enrichment of 108 SNPs and the major SNPs is associated with T cell.

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

DNA translocation, transcription factor overexpression, and focal amplification occur frequently in cancer, and these mechanisms can account for the ability of cancer cells to acquire super-enhancers. In multiple myeloma, for example, tumor cells often have a translocation that places the 30 IgH super-enhancer adjacent to the MYC gene. Overexpression of the TAL1 transcription factor in acute lymphoblastic leukemia (T-ALL) is associated with super-enhancer formation at another site in the MYC locus.

Hanahan and Weinberg (2011) have proposed that cancer cells acquire a number of hallmark biological capabilities during the multistep process of tumor pathogenesis (Hanahan and Weinberg, 2011). We used these hallmarks as an organizing principle to investigate whether genes that acquire super-enhancers are associated with these biological capabilities in tumor cells. We identified super-enhancers that were acquired by cancer cells (not present in a healthy counterpart) and determined how their associated genes fit into the hallmarks. The results of such analysis with a colorectal cancer line revealed that more than one-third of the super-enhancer genes have functions that are associated with a cancer hallmark. A similar analysis of two additional cancer lines confirmed that a large fraction of genes that acquire superenhancers have hallmark functions.