



**HAL**  
open science

# Discovery and analysis of silencers in *Drosophila* acting as enhancers in other cellular contexts

Alexandre Palagi

► **To cite this version:**

Alexandre Palagi. Discovery and analysis of silencers in *Drosophila* acting as enhancers in other cellular contexts. Molecular biology. Université Côte d'Azur, 2018. English. NNT : 2018AZUR4006 . tel-02509608v2

**HAL Id: tel-02509608**

**<https://theses.hal.science/tel-02509608v2>**

Submitted on 15 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

Découverte et analyse d'inactivateurs de transcription  
chez la Drosophile agissant comme amplificateurs  
dans différents contextes cellulaires

**Alexandre PALAGI**

Laboratoire du Pr. Martha L. Bulyk, Brigham and Women's Hospital, Harvard Medical School  
Laboratoire du Dr. Frédéric Luton, Institut de Pharmacologie Moléculaire et Cellulaire

**Présentée en vue de l'obtention du grade de docteur**  
**en** Interactions moléculaires et cellulaires

**d'**Université Côte d'Azur

**Dirigée par** : Martha L. Bulyk / Frédéric Luton

**Co-encadrée par** : Eric Macia

**Soutenue le** : 16 Mars 2018

**Devant le jury, composé de :**

Martha Bulyk, Professeur, Harvard Medical School

Suzanne Gaudet, Directeur de recherche, Harvard Medical School

Mo Motamedi, Directeur de recherche, Harvard Medical School

Juan Fuxman-Bass, Directeur de recherche, Boston University

Anny Cupo, Directeur de recherche, Université Côte d'Azur

Frédéric Luton, Directeur de recherche, Université Côte d'Azur

Découverte et analyse d'inactivateurs de transcription chez  
la Drosophile agissant comme amplificateurs dans différents  
contextes cellulaires

-

Discovery and analysis of silencers in Drosophila acting as  
enhancers in other cellular contexts

**Jury :**

**Président du jury**

Anny Cupo, Directeur de recherche, Université Côte d'Azur

**Directeurs de thèse**

Martha L. Bulyk, Professeur, Harvard Medical School

Frédéric Luton, Directeur de recherche, Université Côte d'Azur

**Rapporteurs**

Juan Fuxman-Bass, Directeur de recherche, Boston University

Mo Motamedi, Directeur de recherche, Harvard Medical School

**Examineur**

Suzanne Gaudet, Directeur de recherche, Harvard Medical School

# Résumé

Un des enjeux majeurs de la biologie moderne est de comprendre les mécanismes complexes régissant l'expression de gènes d'un organisme en développement. Le modèle actuel du contrôle précis et spatio-temporel de cette transcription repose sur deux types de modules de régulation en cis (cis regulatory modules, CRMs), connus sous la dénomination d'amplificateurs de transcription ou « enhancers » et d'inactivateurs de transcription ou « silencers ». Alors que ces enhancers ont été abondamment étudiés et analysés, seul un relatif petit nombre de silencers a été identifié à ce jour et ces derniers restent jusqu'à présent assez mal compris. Aussi, il s'avère qu'un nombre non négligeable de CRMs jouent par ailleurs un double rôle à la fois d'amplificateurs et d'inactivateurs de transcription en fonction de l'état ou du type cellulaire dans lequel ils se trouvent, rajoutant un niveau supplémentaire de complexité à toute tentative de compréhension de la régulation génique dans différents types cellulaires et tissus. Etant donné que ces éléments à double fonction, que l'on peut appeler « bifonctionnels », sont très mal compris et que la fréquence de ces derniers dans un génome métazoaire typique est totalement inconnue, l'enjeu de mon travail de thèse a été de développer une nouvelle approche permettant d'analyser des centaines de séquences dans le but de détecter une activité de répression spécifique à un tissu, au sein d'embryons de *Drosophila melanogaster*. De façon surprenante, nous avons découvert que tous les éléments ayant une activité de répression transcriptionnelle que nous avons identifiés, s'avèrent aussi avoir une activité d'activation transcriptionnelle dans d'autres contextes cellulaires et présentent certaines caractéristiques que je développerai dans ce manuscrit de thèse. Nos résultats remettent donc en question le paradigme de deux catégories distinctes de CRMs et suggèrent que des milliers, ou plus, d'éléments bifonctionnels restent à être découverts chez la Drosophile et potentiellement  $10^4$ - $10^5$  chez l'humain. Le référencement et la caractérisation de ces éléments devraient s'avérer utiles, si ce n'est cruciaux, afin de comprendre la façon par laquelle ces motifs d'expression sont encodés au sein des génomes d'organismes métazoaires et donc éventuellement chez l'Homme.

Mots clefs : transcription, silencers.

# Abstract

One of the major challenges of modern biology is to understand the complex mechanisms governing the expression of genes in a developing organism. The current model of precise and spatio-temporal control of this transcription is based on two types of cis regulatory modules (CRMs), known as enhancers and silencers. While these enhancers have been extensively studied and analyzed, only a relatively small number of silencers have been identified so far and these remain so far poorly understood. Also, it appears that a significant number of CRMs also play a dual role of both enhancer and silencer depending on the state or the cell type in which they are, adding an additional level of complexity to any attempt to understand gene regulation in different cell types and tissues. Given that these dual function CRMs, which we can call "bifunctional" CRMs, are very poorly understood and that the frequency of these in a typical metazoan genome is totally unknown, the goal of my thesis work was to develop a novel approach to analyze hundreds of sequences for the purpose of detecting tissue-specific repression activity in *Drosophila melanogaster* embryos. Surprisingly, we have found that all the elements with transcriptional repression activity that we have identified, also prove to have transcriptional activation activity in other cellular contexts and have certain characteristics that I will develop in this dissertation. Our results therefore challenge the paradigm of two distinct categories of CRMs and suggest that thousands, or more, of bifunctional elements remain to be discovered in *Drosophila* with potentially  $10^4$ - $10^5$  in humans. The referencing and characterization of these elements should prove useful, if not crucial, to understand the way in which expression regulation is encoded within the genomes of metazoan organisms and therefore in humans.

Key words: transcription, silencers.

# Table of Contents

Acknowledgements .....	1
Introduction .....	2
Regulation of Transcription.....	2
Transcription factors and cis regulatory modules .....	3
Enhancers .....	4
Silencers .....	5
Predictions of CRMs .....	8
Predictions using motifs and conservation .....	8
Predictions from ChIP-seq .....	9
Histone marks and chromatin accessibility .....	9
Spatial proximity between genomic regions .....	10
Experimental identification of enhancers.....	11
Current challenges.....	15
Goals of the dissertation .....	18
References .....	19
Chapter 1: bifunctionality of CRMs .....	31
Author contributions.....	33
Abstract .....	33
Acknowledgments .....	33
Introduction .....	34
Initial library and first experiment.....	35
Screening a library of elements for silencer activity in whole <i>Drosophila</i> embryos .	35
Selection of elements to test for silencer activity in <i>Drosophila</i> embryos.....	35
Promoter competition in sFS-positive elements.....	37
Characterization of silencing activity in the context of distinct enhancers.....	37
All validated silencers act as transcriptional enhancers in other cellular contexts ....	37
Transcription factor compositional complexity at silencers.....	41
Chromatin features of active silencers .....	41
Second library and ongoing analyses .....	43
Second library.....	43
Results and validations.....	43
TF compositional complexity and chromatin features of active silencers .....	43
Current and future experiments and analyses.....	47
Genome editing: CRM knock-out .....	47
Hi-C: mesoderm specific interactions .....	51
Spatiotemporal activity of silencers .....	53
Discussion .....	58
Methods.....	59
Generation of reporter vector pSFSdist.....	59
Design of the candidate silencer libraries.....	59
Performing silencer-FACS-Seq experiments .....	62
Statistical analysis of sFS sequencing reads.....	63
Validation of sFS results .....	64
Assessing CRM bifunctionality.....	65
Downstream analysis of the validated silencers.....	65
Cell sorting and fixation with formaldehyde.....	67
Guide-RNA and primer design.....	67
Cas9 and gRNA preparation and microinjection .....	67
In situ hybridization: probes primer design.....	68
References .....	69
Supplementary figures.....	73

Supplementary References: .....	81
Chapter 2: rare cell purification.....	85
Introduction .....	86
Cell panning approach.....	87
Preliminary results.....	88
pCD8 vector .....	88
Cell panning .....	88
Future directions.....	89
Methods.....	94
Generation of vector pCD8 .....	94
Creation of dpp_VRR:CD8 vector and fly lines .....	94
Cell purification protocol – a work in progress.....	95
Supplementary figure .....	96
References .....	97
Conclusion and Future directions.....	99
Summary .....	99
Limitations.....	100
Future directions.....	101
Concluding remarks .....	102
References .....	103
Annexes and supplementary tables .....	105
Annex 1: full sequence of the pCD8 plasmid.....	105
Annex 2: Protocol for positive panning from Drosophila embryos .....	109
Annex 3 : gRNA list for bifunctional element knockout .....	111
Supplementary table 1 .....	112
Supplementary table 2 .....	132
Supplementary table 3 .....	137
Supplementary table 4.....	148
Supplementary table 5.....	154
Supplementary table 6.....	158
Supplementary table 7.....	159
Supplementary table 8.....	161
Supplementary table 9.....	182
Supplementary table 10.....	187

# Acknowledgements

First of all, I want to thank my PhD advisor, Martha Bulyk, for the great opportunity she offered me to realize my thesis research in her laboratory at Harvard. Martha has been a caring, involved and helpful mentor throughout these three years I spent in Boston. I learned a lot from her and I believe there is still a lot she could teach me, to become a successful scientist.

Then, I want to thank Stephen Gisselbrecht, who prefers to go by “Steve”, my unofficial PhD supervisor and mentor with whom I worked on a daily basis on all these complex and often tedious projects I described in this dissertation. Our discussions have always been enjoyable, and ranged from the subject of French pastries to being able to identify the gender of a fly zooming by with the naked eye, and I will miss our daily discussions and coffee break(s). C’est la vie.

I would like to thank Juan Fuxxman-Bass, Mo Motamedi and Suzanne Gaudet for accepting to be on my defense committee, and for their advice and enthusiasm in my work and this dissertation. Moreover, I want to thank Anny Cupo, Frédéric Luton and Eric Macia, without whom I would not have been able to organize my defense.

The Bulyk lab members, past and present, played a major yet discrete role in these three years by all the conversations we shared, on a very large variety of subjects. More specifically, I want to thank Julia Rogers, Sachi Iiyama, Kian Hong Kock and Katy Weinand for their help during my last weeks in the laboratory as I was writing this dissertation; they provided me with comments, advice and support, without which this thesis would have been much harder to write. I am moreover grateful for the compassion they showed me when my aunt passed away.

My friends have been a source of support through these years spent in graduate school, and I want to specifically thank Laura and Sylvaine for visiting me in Boston before my defense. Hadrien, my oldest and best friend, with whom I learned to walk as a child, never stopped supporting me and kept telling me to be strong.

Moreover, I want to deeply thank Mae, who had to indirectly endure the harsh life of graduate school because of me. She has been there every single day to provide me with support, listened to me and cheered me up when I was feeling down. We are lucky to have found each other.

Last but not least, my family has been instrumental to my scientific success. My grandparents always pushed me and encouraged me to go as far as I could in whichever career path I were to choose and to work hard to get to the top. My parents have invariably supported me, despite the fact that I had to move from Cannes, France to Boston, Massachusetts for many years, and encouraged me every single day though thousands of miles away. Finally, my sister, who also currently is in graduate school and with whom I shared the ups and downs of every PhD student, has always been there to remind me that I was not alone, and that the weather is always better in Provence.

*« Les difficultés commencent : c’est le signe de la réussite. »  
- Marcel Pagnol*

# Introduction

## Regulation of Transcription

At any given time, cells, whether they are prokaryotic or eukaryotic, protozoans or part of a metazoan organism, have to adapt to their environment, maintain their metabolism, divide, react properly to a variety of stresses, and progress through their cellular cycle. In metazoans, these basic functions are followed by the even more complex series of proper cell differentiation necessary for organ creation, function, and maintenance and repair. As every single cell in a metazoan organism contains an essentially identical copy of the same genome, it is within this code that information can be found on how to make RNAs and proteins that the organism need prior to its eventual death. But, as cells drastically differentiate into a large diversity of cell types during development, different sections of their genome are required to perform their specific functions and develop into their specific cell type.

This highly complex process of the regulation of RNA and protein expression is called gene regulation. It allows for all cell types and cellular processes observable in organisms and specifies when and where a given gene should be expressed and is referred to as a spatio-temporal program for gene expression (Lockhart and Winzeler, 2000). The expression of a gene product, essentially a protein, has that include initiation and elongation of transcription (the synthesis of messenger RNA or mRNA), mRNA processing, export and finally protein translation.

For transcription to occur, an important class of proteins is the “General Transcription Factors” category (GTFs), also known as the basal transcriptional factors, which do not actually bind DNA for the most part, but are components of the transcription preinitiation complex that interacts with the RNA polymerase at the promoters of genes (Weinzierl, 1999). These GTFs are a prerequisite for any transcription to occur. A very well-known GTF is the TATA-binding protein (TBP), binding the DNA sequence called the TATA box, about 30 base pairs upstream of the transcription start site (TSS) in about some eukaryotic gene promoters (estimated at 10-20% in humans); its role is to help position the RNA polymerase II over the TSS, with the help of a variety of TBP-associated factors (Kornberg, 2007). Other proteins called Transcription Factors (TFs), one responsible for the finer mechanisms of modulation of gene regulation (Dillon, 2006). Since each step of transcription is tightly controlled, it is therefore crucial to develop an understanding of the regulators involved. The regulation of transcription initiation, which is the main subject of focus of my thesis work, is in an important part controlled by the precise binding of TFs, activating or repressing the gene transcription, to specific stretches of the genome, called *cis*-regulatory elements.

## Transcription factors and cis regulatory modules

According to a common definition, a transcription factor, also called a "Sequence Specific DNA Binding Factor", is a protein able to bind specific DNA sequences, also called motifs or TF binding sites (TFBSs), and thereby control the transcription of often neighboring genes (Karin, 1990; Latchman, 1997). The regulation of transcription is in large part controlled by the binding of transcriptional activators and repressors (both termed as *trans*-factors) to specific DNA sequences referred to as *cis*-regulatory modules (CRMs). A current model of the regulation of transcription is that the main function of a motif is recruiting a given TF to the DNA, which will attract cofactors and initiate transcription (Latchman, 1997).

A TF can perform this function alone or within a protein complex, by promoting or blocking the gene transcription (Lee and Young, 2000; Nikolov and Burley, 1997; Roeder, 1996). As part of the definition of what a TF is, these proteins must contain one or more DNA binding domains (DBDs) which interact and attach to specific sequences of DNA, generally nearby the genes that they regulate (Mitchell and Tjian, 1989; Ptashne and Gann, 1997); thus, proteins that are playing important roles in gene regulation but lack DBDs, such as chromatin remodelers, coactivators, deacetylases, histone acetylases, methylases or even protein kinases, are often not considered as TFs (Brivanlou and Darnell, 2002). This is the definition I will be using in this thesis. Despite their importance, the sequence specificities of most TFs remain unknown, making it difficult to decipher the complex CRM regulatory codes as the CRM identification relies on a priori knowledge of TF motifs.

This highlights the need for universal methods to discover the *cis*-element sequences bound by specific TFs. To remedy this, many methods exist such as DNA microarray-based readout of Chromatin Immune-Precipitation, also known as "ChIP" (discussed later in this thesis), the Systematic Evolution of Ligands by Exponential Enrichment (SELEX-seq) or systems such the Bacterial one-hybrid system (B1H) (Bulyk, 2005) or Yeast one-hybrid (Y1H) (Li and Herskowitz, 1993). These techniques are sometime not sufficient and reliable enough to cover all the TFs (Lee et al., 2002), and therefore other methods were created. For instance, the Bulyk lab created a technology that provides a rapid and high-throughput identification and characterization of the DNA binding sites of many TFs, based on microarray *in vitro* technology and called "Protein Binding Microarrays" (PBMs) (Bulyk, 2007; Mukherjee et al., 2004). In an attempt to bypass the need for known TF binding motifs, the Fuxman-Bass laboratory recently developed a "enhanced Yeast 1-Hybrid" (eY1H) approach that allows for the test of hundreds or thousands of TFs to different regulatory regions, to map genome-scale TF-DNA networks (Fuxman Bass et al., 2016).

In metazoans, DNA regulatory motifs tend to occur within CRMs and regulate the expression of the nearby gene or genes. Moreover, these CRMs can be located far away from the transcription start site: in mammalian genomes for example, clusters of binding sites that regulate expression are usually scattered among dozens of thousands of bases to mega-bases, and can be located upstream or downstream of another gene, or even within intronic regions. Their lengths are also variable, from hundreds up to a couple thousands of base pairs (Davidson and Peter, 2015). CRMs are commonly classified into two distinct groups called transcriptional enhancers and transcriptional silencers (Kolovos et al., 2012), or simply enhancers and silencers.

## Enhancers

Enhancers play crucial roles in gene regulation by activating gene expression in a tissue-specific manner in development, and in adult cells in response to cellular or environmental stimuli. However, it is also important that gene expression not be turned on or up-regulated inappropriately. Silencers, on the other hand, are negative regulatory elements (Ogbourne and Antalis, 1998) that play crucial roles in contributing to the specification of precise gene expression patterns, such as sharp expression domains in a developing organism, by preventing ectopic expression.

A characteristic of enhancers that has been repeatedly shown is that they seem to function independently of the distance and orientation to their target genes, and can function at large distances of several hundred kilobases or even megabases by looping (Amano et al., 2009). In addition, they seem to maintain their functions independently of the sequence context, when placed into heterologous reporter constructs for instance (Arnone and Davidson, 1997). Finally, enhancers are modular and contribute additively and sometimes redundantly to the overall expression pattern of their target genes. This can be recapitulated in reporter assays, as combining multiple sequences in an *in vivo* assay often results in patterns of expression that reflect their combined activity (Arnone and Davidson, 1997). Nucleosomes in proximity to active enhancers generally are identified by post-translational modifications on their histones (Bell et al., 2011; Kouzarides, 2007): the main marks known for active enhancers are the histone H3 lysine 4 monomethylation (H3K4me1) and H3K27 acetylation (H3K27ac). These modifications have been mapped and now used for the prediction of enhancers and it appears that the annotation of elements, in whole genomes, seem to correlate rather well with experimental enhancer experiments (Arnold et al., 2013; Bonn et al., 2012; Heintzman et al., 2007). Despite the spreading use of histone marks for predicting enhancers, there is no current consensus about which marks should be used, but it appears that enhancers are often associated with H3K4me1 and H3K27ac, low levels of H3K4me3 and lack of H3K27me3 (Heintzman et al., 2007; Rada-Iglesias et al., 2011). On the other hand, it has been shown that, in *Drosophila melanogaster*, H3K27ac and H3K79me3 are a combination that seems to predict enhancers (Bonn et al., 2012).

## Silencers

Silencers, as their name implies, suppress gene expression (Maeda and Karch, 2011) and prevent therefore gene expression during differentiation and progression through the cell cycle (Li and Arnosti, 2011). Relative to enhancers, less is known about their underlying mechanisms, but they are commonly categorized according to whether they mediate long-range or short-range repression (Gray and Levine, 1996). In the case of long-range repression, a repressor makes a promoter resistant to the influence of any enhancer, even if those enhancers are located thousands of base pairs from the repressor binding site. Short-range repressors function in a less general manner: they do not interfere with all transcription at a locus but rather block the function of nearby DNA-bound activators, while not acting on more distantly bound activators.

A well-studied example of long-range repression is the Groucho-mediated repression (Chen and Courey, 2000; Mannervik et al., 1999; Parkhurst, 1998). The Groucho protein was first identified in *Drosophila melanogaster* and orthologs have been found in all metazoan organisms as mediators for embryonic segmentation, dorsal-ventral patterning, neurogenesis, and Notch and Wnt signaling (Chen and Courey, 2000). In Humans, the Groucho proteins are called transducing-like cancer-of-split (TLE) proteins and, in yeast, Tup1 appears to be a Groucho homolog (Chen and Courey, 2000).

Groucho acts as a corepressor and does not bind to DNA directly, but is rather recruited by a variety of TFs. It has been suggested that Groucho family proteins are long-range corepressors that silence transcription of linked promoters in a relatively indiscriminate fashion (Barolo and Levine, 1997): for instance, binding Groucho-dependent repressors have been found to block expression regardless of their orientation and distance. Groucho-mediated repression seems to involve large nucleoprotein complexes, called repressosomes, which are defined as clusters of TFs interacting with co-repressors and histone-modifying enzymes that represses transcription (M Gowri et al., 2003) by which long-range repression may take place. Indeed, it has been shown that the Groucho proteins may repress transcription in a long-range fashion, by interacting with histone deacetylases (Chen et al., 1999; Choi et al., 1999; D. Watson et al., 2000; Wu et al., 2001). This correlates with the fact that numerous coactivators have been found to function as histone acetyl transferases (HATs), whereas corepressors have been identified as histone deacetylases (HDACs) (Struhl, 1998), yet, so far, no set of histone modification have been strictly associated with silencers.

Another way by which Groucho proteins repress transcription is by inhibitory interaction between silencer-bound repressors and the basal transcriptional machinery. For instance, the histone deacetylase inhibitor TSA only partially blocks Gal4-Groucho-mediated repression (Chen et al., 1999) and that additional regions of Groucho outside of the HDAC1-interacting GP domain function as repression domains (Fisher et al., 1996). Long-range repression is therefore thought to require the formation of a DNA loop that brings a silencer and its associated repressors and corepressors, in close spatial proximity with a core promoter of a target gene, blocking transcription (Gromöller and Lehming, 2000; Papamichos-Chronakis et al., 2000; Yu et al., 2001; Zaman et al., 2001). Moreover, it is thought that the Groucho repressosome may spread a silent chromosomal state. Studies show that the Groucho/Tup1 superfamily proteins can bind hypoacetylated histone tails (D. Flores-Saaib and J. Courey, 2000; Edmondson et al., 1996), potentially allowing them to spread along the chromatin fiber. By recruiting HDAC1 and, or, other kinds of histone deacetylases, they could create large deacetylated and silenced chromosomal domains.

Long-range repression was initially characterized in yeast while studying the *HMR* and *HML* loci which are responsible for the yeast silent mating type (Loo and Rine, 1995). Each HM locus is maintained in a silent state in the heterochromatin by pairs of silencers that can function, in a long-range fashion in an orientation-independent manner, and involve many different proteins such as ORC, Abf1, Rap1 and the silent information regulators (Sir) Sir1, Sir2, Sir3 and Sir4. These latter proteins do not interact directly with these silencers but are rather recruited by sequence-specific factors. Once recruited, the Sir proteins form a repressosome that remodels the chromatin to a silenced state to repress gene expression (Loo and Rine, 1995). Within the Sir repressosome, Sir2 appears to be a histone deacetylase while Sir3 and Sir4 are proteins able to bind hypoacetylated N-terminal tails of histones 3 and 4 (H3 and H4) (Grunstein, 1998). Combined, these two mechanisms provide a probable explanation for the spread of the repressosome from the silencer on the chromatin and, therefore, of the deacetylated region.

Another well studied mechanism for long-range silencing involves the Polycomb complexes PRC1 and PRC2 which rely on noncoding transcripts from silencing elements for being recruited to their target sites, and silencing gene expression via epigenetic silencing. These proteins have been well known for silencing Hox genes via changes in chromatin marks in *Drosophila melanogaster* developing embryos (Portoso and Cavalli, 2008).

Short-range repression may be a more flexible way to achieve this kind of control when compared to long-range repression. For example, the distance over which a short-range repressor is able to work appears to be dependent on repressor concentration. The repressors associated with short-range silencing may therefore respond to a transcription factor concentration gradient (Hewitt et al., 1999). A perfect example of short-range repression are the repressors that regulate the expression of the pair-rule genes such as *eve* and *hairy* in *Drosophila melanogaster*, which are generally expressed in seven transverse stripes in the early embryo, along the anteroposterior axis (Ingham, 1988). The crucial control of pair-rule gene expression depends on the TFs encoded by the gap genes (for instance *giant*, *hunchback*, *Krüppel*, *knirps*) and by the maternal polarity genes (for instance *bicoid*). These factors work via multiple autonomous enhancers in the pair-rule genes and an individual enhancer often controls a single stripe, independently of the other enhancers involved in the characteristic seven-stripes pattern, which has been generated by an appropriate combination of enhancers within a single locus (Akam, 1989).

The critical ability of repressors to allow multiple enhancers to function autonomously is linked to short-range repression: for instance *Giant* and *Krüppel* are able to silence the activation by *Hunchback* and *Bicoid*, given that the activator and repressor binding sites in the stripe 2 enhancer are spaced by less than a hundred base-pairs, which seems to be the lower limit for this kind of repression (Gray and Levine, 1996; Strunk et al., 2001). CtBPs, or C-terminal Binding Proteins, are a common family of corepressors of many short-range repressors (Turner and Crossley, 2001). First found in early *Drosophila* embryos, CtBP is a corepressor of repressors such as *Giant*, *Krüppel*, *Knirps*, and *Snail*, which are at least partially dependent for their function (Mannervik et al., 1999). It has been suggested that CtBP may function, in part, by recruiting histone deacetylases (Criqui-Filipe et al., 1999; Sundqvist et al., 1998). The differences between long-range and short-range deacetylation may be found in the fact that long-range corepressors can spread along their target sites and recruit histone deacetylases to a large domain, but short-range repressors could lack this ability to spread on the chromatin. Another theory is that long-range and short-range corepressors could interact with different histone deacetylases. For instance, Groucho has been found to bind class I histone deacetylases, whereas CtBP appears to bind both class I and class II histone deacetylases (Bertos et al., 2001).

Another possible mechanism for short-range repression is referred to as “quenching”: this mechanism involves interactions of repressors and their corepressors, with activators bound to nearby sites (Gray and Levine, 1996). By this mechanism, a short-range corepressor may be directed to a specific target by a repressor and then interact with an activator protein which already bound DNA and thus block activation by preventing interaction between the activator and the transcriptional machinery. In *Drosophila*, it has been suggested that, for instance, Krüppel might work through a local quenching mechanism when bound to the *eve* stripe 2 enhancer (Gray and Levine, 1996).

Other general silencing hypotheses exist. For instance, it has been shown that long and short RNA molecules seem to be involved in the inhibition of transcription. Antisense RNAs (asRNAs) for instance are small RNAs targeting promoters and their downstream regions (Janowski and Corey, 2010) that silence gene expression (Janowski and Corey, 2010; Janowski et al., 2005). Finally and though largely contested by the community as little evidence is supporting this model, it has been proposed that miRNAs (for micro RNAs, composed of 20 to 22 nucleotides), initially known for regulating gene expression post-transcriptionally, might actually act at the transcriptional level (initiation and, or, elongation) (Bartel, 2009).

Despite the major role that CRMs play in all the biological functions of any organism and the current methods elaborated to study them, there is still a large discrepancy between the number of motifs occurrences in a genome and the actual in vivo detected active sites, which can be context dependent, (Yáñez-Cuna et al., 2013). Therefore, their role in development, evolution and disease still remains an important question to answer. Being able to identify, characterize and predict the function of CRMs and being able to distinguish between enhancers and silencers is therefore of great interest.

## Predictions of CRMs

### Predictions using motifs and conservation

The seemingly simple model of TFs binding to their preferred binding sequences to regulate gene expression lead to initial attempts to predict enhancers in whole genomes via the computational screening and matching of transcription factor binding motifs (TFBDs). These methods either identify genomic regions that are enriched for TF motif matches (Berman et al., 2002) or look for single matches that are conserved across species and evolutionary preserved (Del Bene et al., 2007; Kheradpour et al., 2007). Other methods use both enrichment and conservation, and some of them try to identify only regions in which TFBSs occur in specific combinations or in a particular order or arrangement (Hallikas et al., 2006; Herrmann et al., 2012; Sinha et al., 2003; Warner et al., 2008).

The relationships between TFs, TFBSs and enhancer activity is however far from simple and is not fully understood as short motifs often appear in the genome, and only a very small proportion of all matches in a genome are bound by the corresponding TF *in vivo* (Wang et al., 2012; Yáñez-Cuna et al., 2012). Also, TFs can bind in a context specific fashion and depend on other proteins (Slattery et al., 2011; Yáñez-Cuna et al., 2012) to do so: for instance, in the early *Drosophila melanogaster* embryos, TFs such as Twist (a master mesoderm regulator) depend on the early zygotic TFs Zelda (also called Vielfaltig) while others depend on Tramtrack. In addition to this, the use of sequence conservation does not actually systematically help as conserved motif matches are not always bound by a TF in a particular cell type or tissue, or even acting as active enhancers (Kheradpour et al., 2007; Yáñez-Cuna et al., 2012). Some recent methods use different machine-learning approaches to try to identify characteristic DNA sequence features in enhancers found experimentally and use these features to predict unidentified enhancers (Burzynski et al., 2012; Kantorovitz et al., 2009; Narlikar et al., 2010). Such methods sometime make use of additional data, such as conservation or the expression of flanking genes and can be rather successful.

A good example such approaches are the is the “PhylCRM” and “Lever” algorithms developed by the Bulyk lab (Warner et al., 2008). The PhylCRM algorithm is able to scan very long genomic sequences for candidate CRMs, by quantifying motif clustering and sequence conservation across many given genomes, using an evolutionary model consistent with the phylogeny of the genomes. Lever systematically identifies the target gene sets that are likely to be regulated by a collection of candidate regulatory motifs (Warner et al., 2008) which can come from various sources such as the literature, online databases or experiments (PBMs, ChIP-seq, SELEX-seq, etc.). This software allows the screening of many gene sets with many motifs, motif combinations and helps with the identification of co-regulated gene sets. While not performing *de novo* motif discovery, it instead evaluates an input collection of motifs for enrichment within PhylCRM-predicted candidate CRMs in the noncoding sequences flanking various input gene sets. When these gene sets correspond to Gene Ontology (GO) categories, the results of Lever analysis allow the unbiased assignment of functional annotations to the regulatory motifs and also to the candidate CRMs that comprise the genomic motif occurrences.

The main function of DNA motifs is to be bound by TFs and recruit them to the DNA and then initiate, or not, transcription by the recruitment of co-factors. Nonetheless, not all the DNA motifs may not be bound or even simply accessible at a given time or cell state (for instance in close chromatin), and therefore other methods were developed to directly find *in vivo* CRMs by looking at the binding of these TFs and cofactors in a genome.

## Predictions from ChIP-seq

Given that functional CRMs are bound by TFs, there has been a rise in the use of genome-wide methods that determine *in vivo* transcription factor binding sites for the prediction of active CRMs, such as ChIP-seq. The ChIP-seq strategy aiming at TFs identifies numerous *in vivo* TFBSs in a given genome, such as in promoters, introns and non-coding intergenic DNA (Spitz and Furlong, 2012). These experiments revealed that TFs show dynamic patterns of binding throughout development, implying clues for spatiotemporal specific regulatory targets (Yáñez-Cuna et al., 2012; Zeitlinger et al., 2007). The caveat in these predictions is that whereas TFBSs detected by ChIP-seq are most of the time validated by other experiments and studies, a large number of them are not functional enhancers (Fisher et al., 2012; Kvon et al., 2012; Li et al., 2008), suggesting that the binding of TFs to DNA does not necessarily imply a role in the regulation of transcription. Another approach is to identify CRMs of transcriptional cofactors, which do not bind to DNA directly but are rather recruited by TFs and play a role in the activation or repression of transcription. Combining cofactor ChIP peaks lead to the mapping of regions of different chromatin/histone modifications, providing some insights on actual regulatory regions (van Bemmelen et al., 2013; Filion et al., 2010; Ram et al., 2011).

It is unclear why a major proportion of binding sites are non-functional, but there are a few possible explanations. First of all, TFs have a general affinity for DNA (Hammar et al., 2012), such that they bind to accessible DNA that does not perfectly match their motifs, inside or outside functional contexts, creating “false positive” (from a functional point of view) peaks in ChIP data. Then, it is well established that enhancers are activated by combinations of TFs, implying that the binding of only one or a few TFs is possibly insufficient to activate transcription (Spitz and Furlong, 2012). Finally, it is also possible that some TFs bind indirectly to CRMs by interacting with other TFs, making it difficult to assess any tissue-specific CRM activity or sequence-specific TF binding (Kvon et al., 2012; Moorman et al., 2006).

## Histone marks and chromatin accessibility

Experiments in the 1970s showed that genomic regions containing expressed genes and silent regions were linked to the degree of compaction of eukaryotic DNA (Axel et al., 1973; Weintraub and Groudine, 1976), as chromatin allows or restricts the access of DNA to TFs. In fact, enhancers are “open” DNA regions by being depleted in nucleosomes and therefore sensitive to endonucleases such as DNase I or micrococcal nuclease (MNase). Many laboratories have made use of these nucleases to map regulatory regions by coupling them with deep sequencing, such as DNase-seq (Boyle et al., 2008) and MNase-seq (Yuan et al., 2005). It has been shown that some of the open DNA regions do not only correspond to active enhancers, but also insulators, silencers and other with yet unknown functions (Arnold et al., 2013; Gray and Levine, 1996; Xi et al., 2007), which would not have been found directly given that the factor involved in their regulation remain uncharacterized. Other methods have been developed recently to study chromatin accessibility, as alternatives or complementary methods to these nuclease-dependent approaches. The Assay for Transposase Accessible Chromatin using sequencing or ATAC-seq (Buenrostro et al., 2013, 2015), for instance, make use of the ability of transposons to incorporate preferentially into nucleosome-free regions of genomes (Buenrostro et al., 2013), to insert sequencing adapters into accessible regions of chromatin and then, using the transposase Tn5. This transposase is mutated and highly efficiently cuts exposed DNA while simultaneously ligating adapters. The adapter-ligated fragments are then isolated, amplified and used for sequencing. This method presents the advantages of being relatively short and requiring 1,000 fold less material than MNase or DNase based methods (Buenrostro et al., 2015).

## Spatial proximity between genomic regions

When playing their role, it has been shown that enhancers move into close spatial proximity of the promoter they regulate via a looping mechanism (Dekker et al., 2002; Petrascheck et al., 2005; Sanyal et al., 2012; Zhang et al., 2013). To capture this information, recent techniques were developed to assess directly any physical contact via the polymerase machinery. This way, interchromosomal and intrachromosomal contacts can be detected by methods such as chromosome conformation capture (3C) (Dekker et al., 2002; Miele et al., 2001; Splinter et al., 2004), circular chromosome conformation capture (4C) (Zhao et al., 2006), chromosome conformation capture carbon copy (5C) (Dostie and Dekker, 2007; Dostie et al., 2006) and Hi-C (Lieberman-Aiden et al., 2009).

With these methods, formaldehyde is used to fix physical contacts within the genome. The genomic DNA is sheared by different means and fragments in close proximity are then ligated. The products of these ligations are chimeric DNA molecules that contain fragment that should, in theory, after sequencing, reflect short to long-range contacts. A similar method to these is the chromatin interaction analysis with paired-end tag (ChIA-PET) (Fullwood et al., 2009), that is a combination of one of the methods above and ChIP, therefore looking for interactions involving a given or a set of TFs or cofactors. Enhancers and their target genes can be found by ChIA-PET by targeting RNA polymerase II (Li et al., 2012).

Although promising, the methods attempt to detect spatial proximity but this does not always reflect functional regulatory relationships (Gibcus and Dekker, 2013; de Laat and Duboule, 2013). Moreover, these methods provide data with low resolution, in the order of several kilobases to megabases (Lieberman-Aiden et al., 2009) even though the resolution improves with greater, but more material consuming, depth of sequencing (Jin et al., 2013; Sexton et al., 2012). Also, it has to be noted that these methods have a high background noise at close distances, a caveat of the approach, and that contacts between regions that are adjacent such as where many enhancers are found relative to promoters, are difficult to detect significantly, in one part due to formaldehyde crosslinking (Poorey et al., 2013; Teytelman et al., 2013) as open genomic regions are more likely to be crosslinked (Lieberman-Aiden et al., 2009).

## Experimental identification of enhancers

The first enhancer was identified more than three decades ago as a short 72-bp sequence of the SV40 virus genome that could enhance the transcription of a reporter gene, in HeLa cells, by several hundred fold (Banerji et al., 1981). Shortly, cellular enhancers were found in animal genomes (Banerji et al., 1983) and since then, many CRMs, in a vast majority (if not all of them) enhancers, have been described, and their biochemical and functional properties have been extensively studied. Different experimental approaches managed to identify enhancers successfully, from old enhancer-trap methods, to image based enhancer testing to more advanced high-throughput methods (Arnold et al., 2013; modENCODE Consortium et al., 2010; Nègre et al., 2011; Schaffner, 2015), but these were not designed to attempt a prediction of functions or expression patterns of genes that are regulated by given CRMs, nor were they trying to identify silencers. Image-based methods rely on the creation of reporter constructs by placing a candidate DNA sequence upstream of a minimal promoter and a reporter gene such as GFP, luciferase or *lacZ*, and using microscopy for detecting the expression accordingly by staining for the protein or more directly by *in-situ* hybridization (Rada-Iglesias et al., 2011; Visel et al., 2009a; Zinzen et al., 2009). These methods provide interesting results about activity patterns in different model organisms such as *Drosophila melanogaster* (Kvon et al., 2012; Manning et al., 2012) or *C.elegans* (Dupuy et al., 2004) but, being tedious, are not suited for large screens.

*Drosophila melanogaster*, also known as “fruit fly” serves as a powerful model organism for investigations of spatiotemporal gene regulation in a developing animal (Davidson, 2001). As a matter of facts, since its introduction into the field of genetic experiments by Thomas Hunt Morgan in 1909, *D. melanogaster* is one the most studied organisms for biological research and more importantly in genetics and developmental biology, and its genome was fully sequenced in 2000 (Adams et al., 2000), containing about 140 million base pairs on 4 chromosomes. This genome has been continuously annotated since 1992 (Gramates et al., 2017), and contains around 14 thousands genes. Moreover, it has been shown that a large number of proteins and pathways are conserved between flies and high order vertebrates and humans, with approximately 75% of known human disease genes having fly homologues (Reiter et al., 2001), including cancer, immune system, neurological disorders, heart diseases, but also visual and auditory systems (Bier, 2005).

The development *D. melanogaster* has been studied and documented for decades, from embryogenesis to adult individuals. The transparency of embryos facilitating the study of their development, tissues such as the mesoderm or brain have been extensively analyzed and mapped. For instance and in regard of transcriptional networks, a core regulatory network has been described for the early mesoderm (Sandmann et al., 2007), identifying TFs such as Twist as orchestrating its development. It has been estimated that there are about 50,000 enhancers in the *D. melanogaster* genome (Pfeiffer et al., 2008) but that only ~1,800 are known so far (Halfon et al., 2008).

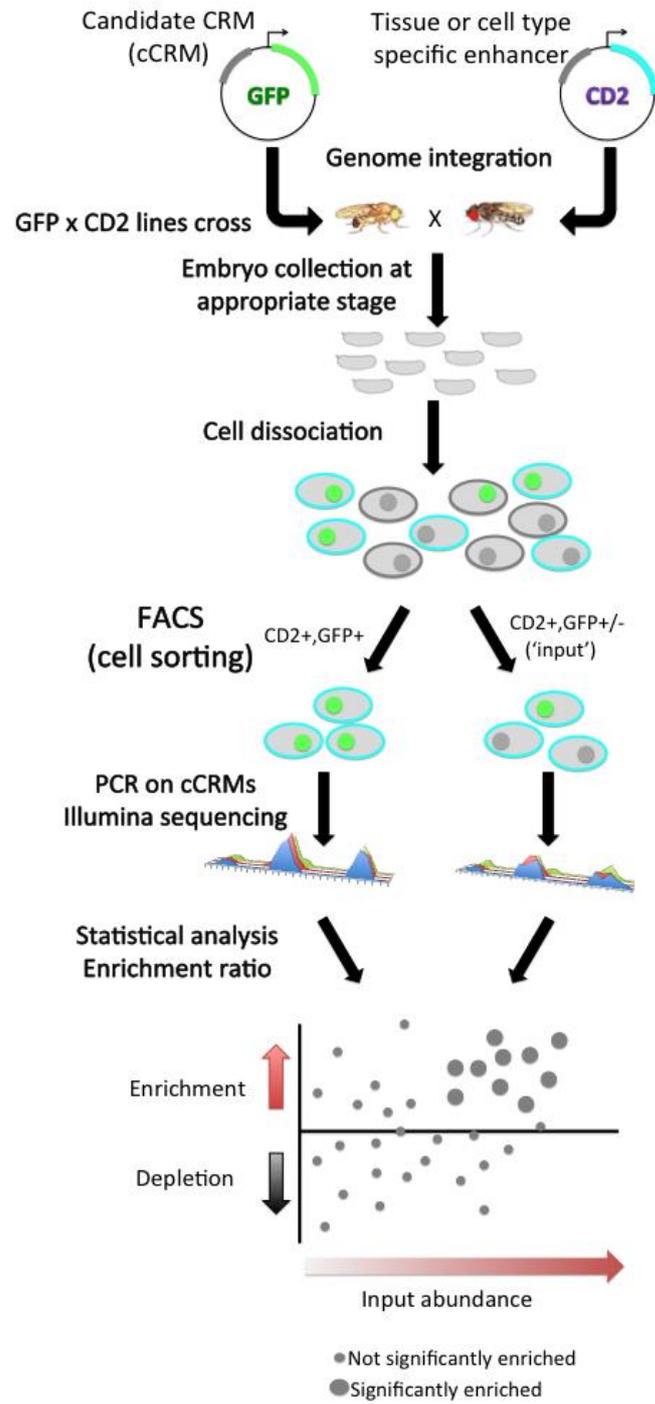
Furthermore, *in vivo* genomic occupancies of numerous TFs and chromatin marks have been profiled by chromatin immunoprecipitation (ChIP)-chip or ChIP-seq (modENCODE Consortium et al., 2010), and the activities of thousands of enhancers have been assayed (Gallo et al., 2011), in *D. melanogaster* embryos. The numerous and advanced ways of manipulating *D. melanogaster* genome, its rapid life cycle, its few needs and the large number of progeny that can easily be generated, along with all the elements discussed before, have made this specie one of, if not the most used and cost effective biological models to this day.

Pursuing the idea of high-throughput and genome-wide screens, different methods have been created, called massively parallel reporter assays (MPRAs), that allow to screen hundreds (Nam and Davidson, 2012) to thousands (Kwasnieski et al., 2012; Melnikov et al., 2012; Patwardhan et al., 2012; Sharon et al., 2012) of candidate CRMs for enhancer activity at once. These MPRAs use deep sequencing and DNA barcoding in experiments where a candidate element is located upstream of a minimal promoter and a reporter gene bearing a heterologous barcode. The cell lines are transformed using these plasmids and expression (RNA levels) of the reporter genes quantified using deep sequencing, allowing to assess the activity of the candidate elements as each candidate is associated to a given barcode by various means. Another similar method, called STARR-seq for Self-transcribing active regulatory region sequencing place the candidate enhancers downstream of a minimal promoter and within the report gene (Arnold et al., 2013), as it was accepted that CRMs and mostly enhancer could work regardless of their position and distance to promoters. In this method, reporter plasmids are transformed into cell lines and deep sequencing is then used, as previously, to detect the RNA levels of the reporter genes and here, therefore, the enhancers themselves.

Nonetheless, it has recently been shown that enhancers are not fully modular and that, in fact, the activity of an enhancer actually depends on its position relatively to the promoter and enhancers within the same locus (Lydiard-Martin et al., 2014), which forces the field into developing methods the role of locus architecture in gene expression and, therefore, stepping away from episomal assays in which a construct works independently from the genome. Moreover, these high-throughput methods do not provide much information about tissue or real cell specificity, in contrast to image based experiments, and are not realized in a genomic context as they are all plasmid based methods. To give a better insight on how gene regulation works in a live tissue and cell specific context, the Bulyk lab developed the enhancer-FACS-seq (eFS) method to address these issues (Gisselbrecht et al., 2013), in *Drosophila melanogaster*.

Similarly to the previous approaches, each candidate CRM (cCRM) is cloned upstream of a reporter gene but, here, integrated into the genome at a single landing site via phiC31-mediated integration. Compared to traditional reporter assays, the main innovation is the use of fluorescence activated cell sorting (FACS) of dissociated cells, instead of microscopy, to screen for tissue-specific enhancers. This approach utilizes a two-marker system: in each embryo, one marker (here, the rat CD2 cell surface protein) is used to label cells of a specific tissue for being sorted by FACS, and the other marker (here, green fluorescent protein GFP) is used as a reporter of CRM activity. Cells are sorted according to their tissue type and then by GFP fluorescence, and the cCRMs are recovered by PCR from double-positive sorted cells, and from total input cells. High-throughput sequencing of both populations then allows measuring the relative abundance of each cCRM in input and sorted populations; one can then assess the enrichment or depletion of each cCRM in double-positive cells versus input as a measure of activity in the CD2-positive cell type being tested. The overview of these process is shown in the figure below.

In the initial report on the enhancer-FACS-seq method, a library of ~500 cCRMs was drawn from a variety of genomic data sources (e.g., TF-bound regions, coactivator-bound regions, DNase I hypersensitive sites, and predictions from the PhylCRM algorithm, by PCR from genomic DNA, and then screened for activity in embryonic mesoderm and in specific mesodermal cell types. The results were validated by traditional reporter gene assay in *Drosophila melanogaster* embryos for 68 cCRMs tested by eFS. The specificity of eFS was excellent among significantly enriched cCRMs, while sensitivity was good where the majority of the CD2-positive cells express GFP. It was found that the known enhancer-associated chromatin marks H3K27ac, H3K4me1, and Pol II are significantly enriched among the enhancers found to be active in mesoderm.



### Overview of the enhancer-FACS-seq method

Although some algorithms have been developed for evaluating the regulatory significance of the DNA motif composition of CRMs, proximal promoter regions have frequently been used as surrogates for regulatory regions in yeast, where most TF binding sites are found (Roth et al., 1998). Also, attempts at creating synthetic enhancer and silencers, or even simply “tuning” CRMs (Barolo, 2016), have always proven a very difficult task, emphasizing our current lack of understanding for the combinatorial complexity of CRMs motif composition. Trying to understand the resulting code that governs expression from regulatory elements remains a major unanswered question in the field.

## Current challenges

Understanding CRMs is currently a major scientific endeavor, as there is an increasing appreciation of their importance not only in development but also in evolution and disease (Carroll, 2008; Dawson and Kouzarides, 2012; Visel et al., 2009b), but unfortunately our knowledge of these elements still remains rather incomplete.

The vast majority of CRMs in any metazoan genome and their spatiotemporal roles are unknown (Yáñez-Cuna et al., 2013) as it is suggested by the number of gene expression patterns that have been linked to specific enhancers for instance (Richardson et al., 2010; Tomancak et al., 2002), the numerous single-nucleotide polymorphisms associated with diseases in non-coding regions of unknown function (International HapMap 3 Consortium et al., 2010) and the large number of genomic regions with enhancer-like chromatin features (ENCODE Project Consortium, 2012; modENCODE Consortium et al., 2010). Despite the fact that tools such as ChromHMM (Ernst and Kellis, 2012) or Spectacle (Song and Chen, 2015) have been created to annotate chromatin states, it still remains a challenge to predict CRMs and their activity from their DNA sequences or from histone marks, and to predict the important parts of an enhancer's sequence.

While enhancers have been extensively studied, silencers happened to be difficult to identify and still remain uncharacterized and relatively poorly understood (Schaffner, 2015). Despite this common treatment of enhancers and silencers as two distinct groups of regulatory elements, a few elements in a variety of eukaryotic systems, as it can be seen in Table 1 below (Bessis et al., 1997; Jiang et al., 1993; Kallunki et al., 1998; Kehayova et al., 2011; Koike et al., 1995; Prasad and Paulson, 2011; Schaeffer et al., 1995; Simpson et al., 1986; Stathopoulos and Levine, 2005; Stroebale and Erives, 2016) have been found to exhibit both activities; *i.e.*, bifunctional elements that can act as either an enhancer or a silencer, depending on the tissue type or cellular conditions, and there has been some evidence suggesting that enhancers and silencers share more properties than initially thought (Raab and Kamakaka, 2010).

The silencer activities of a few known *D. melanogaster* bifunctional CRMs have been described in the *D. melanogaster* embryonic mesoderm (Jiang et al., 1993; Stathopoulos and Levine, 2005; Stroebale and Erives, 2016). The first one is the zen ventral repressor element (VRE) which contributes to specifying the zen expression pattern by repressing its expression ventrally and driving it dorsally, in early stage embryos. When tested in a reporter element upstream of a known enhancer, the eve minimal stripe 2 enhancer (or MSE), the VRE repressed the ventral expression induced by the MSE while extending it dorsally, reproducing the zen expression pattern (Jiang et al., 1993). The other example in *Drosophila* of precise control of expression, but also of the potential bifunctionality of CRMs, is the model proposed of a bifunctional element, a 1.4kb enhancer downstream of the *ind* gene, by Stathopoulos and Levine (Stathopoulos and Levine, 2005). This enhancer was dissected into three modules and two were found to repress gene activity by delimiting patterns of expression. The ventral limits of gene expression in a developing embryo were found to be defined by the Snail and Vnd repressors in one module that can be seen as a silencer, while the dorsal border was defined by another new silencer element. Finally, for nab intronic enhancers have been found to also function as mutual silencers (Stroebale and Erives, 2016) where the distinct nab enhancer fragment analyzed possessed endogenous nab-related enhancer activities and ectopic enhancer activities while also function as interenhancer silencers of ectopic activities and modulators for their nonectopic activities.

Title	Citation	Evidence	Organism
Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties	J. Simpson et al., Nature 323, 551-554 (1986).	Reporter assay	Tobacco
Conversion of a dorsal dependent silencer into an enhancer: evidence for dorsal corepressors	J. Jiang et al., EMBO J 12, 3201-3209 (1993).	Reporter assay	Drosophila
Identification of enhancer and silencer regions involved in salt-responsive expression of Crassulacean acid metabolism (CAM) genes in the facultative halophyte <i>Mesembryanthemum crystallinum</i>	H. J. Schaeffer et al., Plant Mol Biol 28, 205-218 (1995).	Reporter assay	Fungus
Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene	S. Koike, L. Schaeffer, J.-P. Changeux, Proc Natl Acad Sci U S A 92, 10624-10628 (1995).	Reporter assay	Mouse
The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain	A. Bessis et al., Proc Natl Acad Sci U S A 94, 5906-5911 (1997).	Deletion	Mouse
The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during post-natal development	P. Kallunki, G. M. Edelman, F. S. Jones, Proc Natl Acad Sci U S A 95, 3233-3238 (1998).	Deletion	Mouse
Localized repressors delineate the neurogenic ectoderm in the early <i>Drosophila</i> embryo.	A. Stathopoulos, M. Levine, Dev Biol 280, 482-493 (2005).	Reporter assay	Drosophila
A combination of enhancer/silencer modules regulates spatially restricted expression of cadherin-7 in neural epithelium	M. S. Prasad, A. F. Paulson, Dev Dyn 240, 1756-1768 (2011).	Reporter assay	Chicken
Regulatory elements required for the activation and repression of the protocadherin- $\alpha$ gene cluster	P. Kehayova et al., Proc Natl Acad Sci U S A 108, 17195-17200 (2011).	Deletion	Mouse
Integration of Orthogonal Signaling by the Notch and Dpp Pathways in <i>Drosophila</i>	Stroebele, E., and Erives, A., Genetics 203, 219–240 (2016).	Reporter assay	Drosophila

### Table of all known bifunctional elements in the literature

To be included in this list, a study must have demonstrated evidence that a naturally occurring genomic sequence acts as an enhancer and as a silencer in alternate cellular contexts. This includes different tissues, cell types, cultured cell lines, conditions, or developmental ages.

We currently do not understand the mechanisms that allow the same element to have two opposite functions. Many TFs can act as either activators or repressors, depending on the context of the *cis*-element (Ogbourne and Antalis, 1998); however, bifunctionality of a *cis*-element may not require such TFs, since different activators or repressors could bind the same element in different tissues. Understanding how a given gene regulation network functions is still difficult but necessary to reach a complete understanding of any cellular activity, from embryogenesis and complex development to regular metabolism, but also to understand how a disruption in a given network can, or not, lead to a dysfunctional network. Many TFs can act as either as transcriptional activators or repressors, depending on the context of their binding, but the potential bifunctionality of CRMs complicates the prediction of gene expression from sequence and the interpretation of the effects of *cis*-regulatory variation across populations or in evolution, since we can theorize that no such TF is involved in bifunctional elements.

These bifunctional elements complicate the prediction of gene expression from sequence and the interpretation of the effects of *cis*-regulatory variation across population, in evolution, but also our global understand of genotype to phenotype relationships. For instance, being able to predict the functional consequences of mutations in CRMs, which means having a precise understanding of their potential bifunctionality, could be a drastic help in the fight against diseases involving perturbations of gene regulation networks, and potentially cancer.

As I write this dissertation, it remains unknown how general this property might be and how many such bifunctional elements a typical metazoan genome might contain. Screening for bifunctional CRMs requires the ability to assay a *cis*-element for both enhancer activity in one cell type and silencer activity in a different cell type, in a controlled spatiotemporal manner, *i.e.* at a known position within the genome, in a known cell type, at the proper time. However, unlike enhancer assays, no scalable screening technology is currently available to assay silencer activity in a metazoan.

## Goals of the dissertation

The main goal of this dissertation was the discovery and study of silencers and potential bifunctional elements by adapting our previously developed method, enhancer-FACS-seq or “eFS” (Gisselbrecht et al., 2013), into “silencer-FACS-seq” that aimed for the discovery of tissue specific enhancers in developing *Drosophila melanogaster* embryos and more specifically their mesoderm, but also an attempt at creating a method of cellular purification or “panning method”, which aims to reach these small population of cells in a eFS and sFS fashion.

In the first chapter, I will discuss how the eFS method was adapted for the purpose of silencers. In this chapter, I will discuss what candidate elements were chosen for our experiment, the reasoning associated with their selection, and analyze the results from the experiments and validations. I will go over our attempts to try to characterize our newly found silencers, present with the main findings of this work, along with areas of ongoing projects and future studies.

In the second chapter, I will go over the ongoing “cell panning” project by discussing the goals, trials and methods used to purify small or rare population of cells, from dissociated *Drosophila melanogaster*. embryos, for analysis.

Finally, I will end this dissertation with a global discussion on the results of my thesis and present potential future directions for these projects.

## References

- Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195.
- Akam, M. (1989). *Drosophila* development: making stripes inelegantly. *Nature* 341, 282–283.
- Altschuler, S.J., and Wu, L.F. (2010). Cellular Heterogeneity: Do Differences Make a Difference? *Cell* 141, 559–563.
- Amano, T., Sagai, T., Tanabe, H., Mizushima, Y., Nakazawa, H., and Shiroishi, T. (2009). Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* 16, 47–57.
- Arnold, C.D., Gerlach, D., Stelzer, C., Boryń, Ł.M., Rath, M., and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077.
- Arnone, M.I., and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Dev. Camb. Engl.* 124, 1851–1864.
- Axel, R., Cedar, H., and Felsenfeld, G. (1973). Synthesis of globin ribonucleic acid from duck-erythrocyte chromatin in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 70, 2029–2032.
- 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.
- Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33, 729–740.
- Barolo, S. (2016). How to tune an enhancer. *Proc. Natl. Acad. Sci.* 113, 6330–6331.
- Barolo, S., and Levine, M. (1997). hairy mediates dominant repression in the *Drosophila* embryo. *EMBO J.* 16, 2883–2891.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233.
- Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. *Nat. Rev. Genet.* 12, 554.
- van Bommel, J.G., Filion, G.J., Rosado, A., Talhout, W., de Haas, M., van Welsem, T., van Leeuwen, F., and van Steensel, B. (2013). A network model of the molecular organization of chromatin in *Drosophila*. *Mol. Cell* 49, 759–771.
- Berman, B.P., Nibu, Y., Pfeiffer, B.D., Tomancak, P., Celniker, S.E., Levine, M., Rubin, G.M., and Eisen, M.B. (2002). Exploiting transcription factor binding site clustering to identify cis-

regulatory modules involved in pattern formation in the *Drosophila* genome. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 757–762.

Bertos, N.R., Wang, A.H., and Yang, X.J. (2001). Class II histone deacetylases: structure, function, and regulation. *Biochem. Cell Biol. Biochim. Biol. Cell.* *79*, 243–252.

Bessis, A., Champtiaux, N., Chatelin, L., and Changeux, J.-P. (1997). The neuron-restrictive silencer element: A dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 5906–5911.

Bier, E. (2005). *Drosophila*, the golden bug, emerges as a tool for human genetics. *Nat. Rev. Genet.* *6*, 9.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczyński, B., Riddell, A., and Furlong, E.E.M. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* *44*, 148–156.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell* *132*, 311–322.

Brivanlou, A.H., and Darnell, J.E. (2002). Signal transduction and the control of gene expression. *Science* *295*, 813–818.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. *Nat. Methods* *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* *109*, 21.29.1-9.

Bulyk, M.L. (2005). Discovering DNA regulatory elements with bacteria. *Nat. Biotechnol.* *23*, 942–944.

Bulyk, M.L. (2007). Protein binding microarrays for the characterization of DNA-protein interactions. *Adv. Biochem. Eng. Biotechnol.* *104*, 65–85.

Burzynski, G.M., Reed, X., Taher, L., Stine, Z.E., Matsui, T., Ovcharenko, I., and McCallion, A.S. (2012). Systematic elucidation and in vivo validation of sequences enriched in hindbrain transcriptional control. *Genome Res.* *22*, 2278–2289.

Campbell, L.L., and Polyak, K. (2007). Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle Georget. Tex* *6*, 2332–2338.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* *134*, 25–36.

Chen, G., and Courey, A.J. (2000). Groucho/TLE family proteins and transcriptional repression. *Gene* *249*, 1–16.

Chen, G., Fernandez, J., Mische, S., and Courey, A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in *Drosophila* development. *Genes Dev.* *13*, 2218–2230.

Choi, C., Kwon, H.J., and Kim, Y. (1999). Choi CY, Kim YH, Kwon HJ, Kim Y.. The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription. *J Biol Chem* 274: 33194-33197. *J. Biol. Chem.* 274, 33194–33197.

Criqui-Filipe, P., Ducret, C., Maira, S.-M., and Wasylyk, B. (1999). Net, a negative Ras-switchable TCF, contains a second inhibition domain, the CID, that mediates repression through interactions with CtBP and de-acetylation. *EMBO J.* 18, 3392–3403.

D. Flores-Saaib, R., and J. Courey, A. (2000). Analysis of Groucho-histone interactions suggests mechanistic similarities between Groucho- and Tup1-mediated repression. *Nucleic Acids Res.* 28, 4189–4196.

D. Watson, A., Edmondson, D., Bone, J., Mukai, Y., Yu, Y., Du, W., J. Stillman, D., and Dent, S. (2000). Ssn6-Tup1 interacts with class I histone deacetylases required for repression. *Genes Dev.* 14.

Davidson, E.H. (2001). Chapter 2 - Inside the cis-Regulatory Module: Control Logic, and How Regulatory Environment is Transduced into Spatial Patterns of Gene Expression. In *Genomic Regulatory Systems*, (San Diego: Academic Press), pp. 25–62.

Davidson, E.H., and Peter, I.S. (2015). Chapter 1 - The Genome in Development. In *Genomic Control Process*, (Oxford: Academic Press), pp. 1–40.

Dawson, M.A., and Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell* 150, 12–27.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. *Science* 295, 1306–1311.

Del Bene, F., Ettwiller, L., Skowronska-Krawczyk, D., Baier, H., Matter, J.-M., Birney, E., and Wittbrodt, J. (2007). In vivo validation of a computationally predicted conserved Ath5 target gene set. *PLoS Genet.* 3, 1661–1671.

Dillon, N. (2006). Gene regulation and large-scale chromatin organization in the nucleus. *Chromosome Res.* 14, 117–126.

Dostie, J., and Dekker, J. (2007). Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.* 2, 988–1002.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 16, 1299–1309.

Dupuy, D., Li, Q.-R., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stamm, L., Hope, I.A., et al. (2004). A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res.* 14, 2169–2175.

Edmondson, D., Smith, M.M., and Dent, S. (1996). Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev.* 10, 1247–1259.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* 9, 215–216.

Filion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143, 212–224.

Fisher, A., Ohsako, S., and Caudy, M. (1996). The WRPW Motif of the Hairy-Related Basic Helix-Loop-Helix Repressor Proteins Acts as a 4AminoAcid Transcription Repression and Protein-Protein Interaction Domain. *Mol. Cell. Biol.* 16, 2670–2677.

Fisher, W.W., Li, J.J., Hammonds, A.S., Brown, J.B., Pfeiffer, B.D., Weizmann, R., MacArthur, S., Thomas, S., Stamatoyannopoulos, J.A., Eisen, M.B., et al. (2012). DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 109, 21330–21335.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor- $\alpha$ -bound human chromatin interactome. *Nature* 462, 58.

Fuxman Bass, J.I., Reece-Hoyes, J.S., and Walhout, A.J.M. (2016). Gene-Centered Yeast One-Hybrid Assays. *Cold Spring Harb. Protoc.* 2016, pdb.top077669.

Gallo, S.M., Gerrard, D.T., Miner, D., Simich, M., Des Soye, B., Bergman, C.M., and Halfon, M.S. (2011). REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Res.* 39, D118-123.

Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. *Mol. Cell* 49, 773–782.

Gisselbrecht, S.S., Barrera, L.A., Porsch, M., Aboukhalil, A., Estep, P.W., Vedenko, A., Palagi, A., Kim, Y., Zhu, X., Busser, B.W., et al. (2013). Highly parallel assays of tissue-specific enhancers in whole *Drosophila* embryos. *Nat. Methods* 10, 774–780.

Gramates, L.S., Marygold, S.J., Santos, G. dos, Urbano, J.-M., Antonazzo, G., Matthews, B.B., Rey, A.J., Tabone, C.J., Crosby, M.A., Emmert, D.B., et al. (2017). FlyBase at 25: looking to the future. *Nucleic Acids Res.* 45, D663–D671.

Gray, S., and Levine, M. (1996). Transcriptional repression in development. *Curr. Opin. Cell Biol.* 8, 358–364.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8, 172.

Gromöller, A., and Lehming, N. (2000). *Srb7p* is a physical and physiological target of *Tup1p*. *EMBO J.* 19, 6845–6852.

Grün, D., Kester, L., and Oudenaarden, A. van (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* 11, 637.

Grunstein, M. (1998). Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* 93, 325–328.

Halfon, M.S., Gallo, S.M., and Bergman, C.M. (2008). REDfly 2.0: an integrated database of cis-regulatory modules and transcription factor binding sites in *Drosophila*. *Nucleic Acids Res.* *36*, D594-598.

Hallikas, O., Palin, K., Sinjushina, N., Rautiainen, R., Partanen, J., Ukkonen, E., and Taipale, J. (2006). Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* *124*, 47–59.

Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E.G., Berg, O.G., and Elf, J. (2012). The lac repressor displays facilitated diffusion in living cells. *Science* *336*, 1595–1598.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* *39*, 311–318.

Heppner, G.H. (1984). Tumor heterogeneity. *Cancer Res.* *44*, 2259–2265.

Herrmann, C., Van de Sande, B., Potier, D., and Aerts, S. (2012). i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. *Nucleic Acids Res.* *40*, e114.

Hewitt, G.F., Strunk, B.S., Margulies, C., Priputin, T., Wang, X.D., Amey, R., Pabst, B.A., Kosman, D., Reinitz, J., and Arnosti, D.N. (1999). Transcriptional repression by the *Drosophila* giant protein: cis element positioning provides an alternative means of interpreting an effector gradient. *Dev. Camb. Engl.* *126*, 1201–1210.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.

Huang, L., Ma, F., Chapman, A., Lu, S., and Xie, X.S. (2015). Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu. Rev. Genomics Hum. Genet.* *16*, 79–102.

Ingham, P.W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* *335*, 25–34.

International HapMap 3 Consortium, Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* *467*, 52–58.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* *169*, 5429–5433.

Janowski, B.A., and Corey, D.R. (2010). Minireview: Switching on progesterone receptor expression with duplex RNA. *Mol. Endocrinol. Baltim. Md* *24*, 2243–2252.

Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Hardy, D., Shames, D.S., Minna, J.D., and Corey, D.R. (2005). Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. *Nat. Chem. Biol.* *1*, 216–222.

Jiang, J., Cai, H., Zhou, Q., and Levine, M. (1993). Conversion of a dorsal-dependent silencer into an enhancer: evidence for dorsal corepressors. *EMBO J.* *12*, 3201–3209.

- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* *503*, 290–294.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Kallunki, P., Edelman, G.M., and Jones, F.S. (1998). The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 3233–3238.
- Kantorovitz, M.R., Kazemian, M., Kinston, S., Miranda-Saavedra, D., Zhu, Q., Robinson, G.E., Göttgens, B., Halfon, M.S., and Sinha, S. (2009). Motif-blind, genome-wide discovery of cis-regulatory modules in *Drosophila* and mouse. *Dev. Cell* *17*, 568–579.
- Karin, M. (1990). Too many transcription factors: positive and negative interactions. *New Biol.* *2*, 126–131.
- Kehayova, P., Monahan, K., Chen, W., and Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 17195–17200.
- Kheradpour, P., Stark, A., Roy, S., and Kellis, M. (2007). Reliable prediction of regulator targets using 12 *Drosophila* genomes. *Genome Res.* *17*, 1919–1931.
- Koike, S., Schaeffer, L., and Changeux, J.P. (1995). Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 10624–10628.
- Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C., and Teichmann, S.A. (2015). The Technology and Biology of Single-Cell RNA Sequencing. *Mol. Cell* *58*, 610–620.
- Kolovos, P., Knoch, T.A., Grosveld, F.G., Cook, P.R., and Papantonis, A. (2012). Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin* *5*, 1.
- Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 12955–12961.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693–705.
- Kvon, E.Z., Stampfel, G., Yáñez-Cuna, J.O., Dickson, B.J., and Stark, A. (2012). HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. *Genes Dev.* *26*, 908–913.
- Kwasnieski, J.C., Mogno, I., Myers, C.A., Corbo, J.C., and Cohen, B.A. (2012). Complex effects of nucleotide variants in a mammalian cis-regulatory element. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 19498–19503.
- de Laat, W., and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* *502*, 499–506.
- Latchman, D.S. (1997). Transcription factors: an overview. *Int. J. Biochem. Cell Biol.* *29*, 1305–1312.

- Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77–137.
- Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. (2002). Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 298, 799–804.
- Li, J.J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* 262, 1870–1874.
- Li, L.M., and Arnosti, D.N. (2011). Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. *Curr. Biol. CB* 21, 406–412.
- Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98.
- Li, X., MacArthur, S., Bourgon, R., Nix, D., Pollard, D.A., Iyer, V.N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C.L., et al. (2008). Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* 6, e27.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- Lockhart, D.J., and Winzler, E.A. (2000). Genomics, gene expression and DNA arrays. *Nature* 405, 827–836.
- Loo, S., and Rine, J. (1995). Silencing and heritable domains of gene expression. *Annu. Rev. Cell Dev. Biol.* 11, 519–548.
- Lydiard-Martin, T., Bragdon, M., Eckenrode, K.B., Wunderlich, Z., and DePace, A.H. (2014). Locus architecture affects mRNA expression levels in *Drosophila* embryos. *BioRxiv* 005173.
- M Gowri, P., H Yu, J., Shaufli, A., Sperling, M., and K Menon, R. (2003). Recruitment of a Repressosome Complex at the Growth Hormone Receptor Promoter and Its Potential Role in Diabetic Nephropathy. *Mol. Cell. Biol.* 23, 815–825.
- Maeda, R.K., and Karch, F. (2011). Gene expression in time and space: additive vs hierarchical organization of cis-regulatory regions. *Curr. Opin. Genet. Dev.* 21, 187–193.
- Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and Koonin, E.V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1, 7.
- Mannervik, M., Nibu, Y., Zhang, H., and Levine, M. (1999). Transcriptional Coregulators in Development. *Science* 284, 606–609.
- Manning, L., Heckscher, E.S., Purice, M.D., Roberts, J., Bennett, A.L., Kroll, J.R., Pollard, J.L., Strader, M.E., Lupton, J.R., Dyukareva, A.V., et al. (2012). A resource for manipulating gene expression and analyzing cis-regulatory modules in the *Drosophila* CNS. *Cell Rep.* 2, 1002–1013.

Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G., Kinney, J.B., et al. (2012). Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* *30*, 271–277.

Miele, A., Gheldof, N., Tabuchi, T.M., Dostie, J., and Dekker, J. (2001). Mapping Chromatin Interactions by Chromosome Conformation Capture. In *Current Protocols in Molecular Biology*, (John Wiley & Sons, Inc.), p.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* *245*, 371–378.

modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., et al. (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* *330*, 1787–1797.

Moorman, C., Sun, L.V., Wang, J., de Wit, E., Talhout, W., Ward, L.D., Greil, F., Lu, X.-J., White, K.P., Bussemaker, H.J., et al. (2006). Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 12027–12032.

Mukherjee, S., Berger, M.F., Jona, G., Wang, X.S., Muzzey, D., Snyder, M., Young, R.A., and Bulyk, M.L. (2004). Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat. Genet.* *36*, 1331–1339.

Nam, J., and Davidson, E.H. (2012). Barcoded DNA-tag reporters for multiplex cis-regulatory analysis. *PLoS One* *7*, e35934.

Narlikar, L., Sakabe, N.J., Blanski, A.A., Arimura, F.E., Westlund, J.M., Nobrega, M.A., and Ovcharenko, I. (2010). Genome-wide discovery of human heart enhancers. *Genome Res.* *20*, 381–392.

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* *471*, 527–531.

Nikolov, D.B., and Burley, S.K. (1997). RNA polymerase II transcription initiation: a structural view. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 15–22.

Ogbourne, S., and Antalis, T.M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* *331*, 1–14.

Papamichos-Chronakis, M., Conlan, R.S., Gounalaki, N., Copf, T., and Tzamarias, D. (2000). Hrs1/Med3 is a Cyc8-Tup1 corepressor target in the RNA polymerase II holoenzyme. *J. Biol. Chem.* *275*, 8397–8403.

Parkhurst, S.M. (1998). Groucho: making its Marx as a transcriptional co-repressor. *Trends Genet.* *TIG 14*, 130–132.

Patwardhan, R.P., Hiatt, J.B., Witten, D.M., Kim, M.J., Smith, R.P., May, D., Lee, C., Andrie, J.M., Lee, S.-I., Cooper, G.M., et al. (2012). Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.* *30*, 265–270.

- Petrascheck, M., Escher, D., Mahmoudi, T., Peter Verrijzer, C., Schaffner, W., and Barberis, A. (2005). DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res.* *33*, 3743–3750.
- Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.-T.B., Misra, S., Murphy, C., Scully, A., Carlson, J.W., Wan, K.H., Laverty, T.R., et al. (2008). Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 9715–9720.
- Poorey, K., Viswanathan, R., Carver, M.N., Karpova, T.S., Cirimotich, S.M., McNally, J.G., Bekiranov, S., and Auble, D.T. (2013). Measuring chromatin interaction dynamics on the second time scale at single-copy genes. *Science* *342*, 369–372.
- Portoso, M., and Cavalli, G. (2008). The Role of RNAi and Noncoding RNAs in Polycomb Mediated Control of Gene Expression and Genomic Programming. pp. 29–44.
- Pott, S., and Lieb, J.D. (2015). Single-cell ATAC-seq: strength in numbers. *Genome Biol.* *16*.
- Prasad, M.S., and Paulson, A.F. (2011). A combination of enhancer/silencer modules regulates spatially restricted expression of cadherin-7 in neural epithelium. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* *240*, 1756–1768.
- Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* *386*, 569–577.
- Raab, J.R., and Kamakaka, R.T. (2010). Insulators and promoters: closer than we think. *Nat. Rev. Genet.* *11*, 439–446.
- Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* *470*, 279–283.
- Ram, O., Goren, A., Amit, I., Shores, N., Yosef, N., Ernst, J., Kellis, M., Gymrek, M., Issner, R., Coyne, M., et al. (2011). Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* *147*, 1628–1639.
- Reiter, L., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A Systematic Analysis of Human Disease-Associated Gene Sequences In *Drosophila melanogaster*.
- Richardson, L., Venkataraman, S., Stevenson, P., Yang, Y., Burton, N., Rao, J., Fisher, M., Baldock, R.A., Davidson, D.R., and Christiansen, J.H. (2010). EMAGE mouse embryo spatial gene expression database: 2010 update. *Nucleic Acids Res.* *38*, D703-709.
- Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* *21*, 327–335.
- Roth, F.P., Hughes, J.D., Estep, P.W., and Church, G.M. (1998). Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. *Nat. Biotechnol.* *16*, 939–945.
- Rubin, H. (1990). The significance of biological heterogeneity. *Cancer Metastasis Rev.* *9*, 1–20.
- Saliba, A.-E., Westermann, A.J., Gorski, S.A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* *42*, 8845–8860.

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V., and Furlong, E.E.M. (2007). A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* *21*, 436–449.

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

Schaeffer, H.J., Forsthoefel, N.R., and Cushman, J.C. (1995). Identification of enhancer and silencer regions involved in salt-responsive expression of Crassulacean acid metabolism (CAM) genes in the facultative halophyte *Mesembryanthemum crystallinum*. *Plant Mol. Biol.* *28*, 205–218.

Schaffner, W. (2015). Enhancers, enhancers - from their discovery to today's universe of transcription enhancers. *Biol. Chem.* *396*, 311–327.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* *148*, 458–472.

Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and Segal, E. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* *30*, 521–530.

Simpson, J., Schell, J., Montagu, M.V., and Herrera-Estrella, L. (1986). Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties. *Nature* *323*, 551–554.

Sinha, S., van Nimwegen, E., and Siggia, E.D. (2003). A probabilistic method to detect regulatory modules. *Bioinforma. Oxf. Engl.* *19 Suppl 1*, i292-301.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., et al. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* *147*, 1270–1282.

Song, J., and Chen, K.C. (2015). Spectacle: fast chromatin state annotation using spectral learning. *Genome Biol.* *16*, 33.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* *13*, 613–626.

Splinter, E., Grosveld, F., and de Laat, W. (2004). 3C technology: analyzing the spatial organization of genomic loci in vivo. *Methods Enzymol.* *375*, 493–507.

Stathopoulos, A., and Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev. Biol.* *280*, 482–493.

Stroebele, E., and Erives, A. (2016). Integration of Orthogonal Signaling by the Notch and Dpp Pathways in *Drosophila*. *Genetics* *203*, 219–240.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* *12*, 599–606.

Strunk, B., Struffi, P., Wright, K., Pabst, B., Thomas, J., Qin, L., and Arnosti, D.N. (2001). Role of CtBP in Transcriptional Repression by the *Drosophila* giant Protein. *Dev. Biol.* *239*, 229–240.

Sundqvist, A., Sollerbrant, K., and Svensson, C. (1998). The carboxy-terminal region of adenovirus E1A activates transcription through targeting of a C-terminal binding protein-histone deacetylase complex. *FEBS Lett.* *429*, 183–188.

Teytelman, L., Thurtle, D.M., Rine, J., and van Oudenaarden, A. (2013). Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 18602–18607.

Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E., et al. (2002). Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol.* *3*, research0088.

Tsioris, K., Torres, A.J., Douce, T.B., and Love, J.C. (2014). A New Toolbox for Assessing Single Cells. *Annu. Rev. Chem. Biomol. Eng.* *5*, 455–477.

Turner, J., and Crossley, M. (2001). The CtBP family: Enigmatic and enzymatic transcriptional co-repressors.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009a). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* *457*, 854–858.

Visel, A., Rubin, E.M., and Pennacchio, L.A. (2009b). Genomic views of distant-acting enhancers. *Nature* *461*, 199–205.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* *153*, 910–918.

Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* *22*, 1798–1812.

Warner, J.B., Philippakis, A.A., Jaeger, S.A., He, F.S., Lin, J., and Bulyk, M.L. (2008). Systematic identification of mammalian regulatory motifs' target genes and functions. *Nat. Methods* *5*, 347–353.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. *Science* *193*, 848–856.

Weinzierl, R. (1999). *Mechanisms of gene expression : structure, function and evolution of the basal transcriptional machinery* (London : Imperial College Press).

Wu, A.R., Wang, J., Streets, A.M., and Huang, Y. (2017). Single-Cell Transcriptional Analysis. *Annu. Rev. Anal. Chem.* *10*, 439–462.

Wu, J., Suka, N., Carlson, M., and Grunstein, M. (2001). TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. *Mol. Cell* *7*, 117–126.

Xi, H., Shulha, H.P., Lin, J.M., Vales, T.R., Fu, Y., Bodine, D.M., McKay, R.D.G., Chenoweth, J.G., Tesar, P.J., Furey, T.S., et al. (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet.* *3*, e136.

Xie, F., Ye, L., Chang, J.C., Beyer, A.I., Wang, J., Muench, M.O., and Kan, Y.W. (2014). Seamless gene correction of  $\beta$ -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* *24*, 1526–1533.

Yáñez-Cuna, J.O., Dinh, H.Q., Kvon, E.Z., Shlyueva, D., and Stark, A. (2012). Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. *Genome Res.* *22*, 2018–2030.

Yáñez-Cuna, J.O., Kvon, E.Z., and Stark, A. (2013). Deciphering the transcriptional cis-regulatory code. *Trends Genet. TIG* *29*, 11–22.

Yu, X., Li, P., Roeder, R.G., and Wang, Z. (2001). Inhibition of androgen receptor-mediated transcription by amino-terminal enhancer of split., Inhibition of Androgen Receptor-Mediated Transcription by Amino-Terminal Enhancer of split. *Mol. Cell. Biol.* *21*, 21, 4614, 4614–4625.

Yuan, G.-C., Liu, Y.-J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* *309*, 626–630.

Zaman, Z., Ansari, A.Z., Koh, S.S., Young, R., and Ptashne, M. (2001). Interaction of a transcriptional repressor with the RNA polymerase II holoenzyme plays a crucial role in repression. *Proc. Natl. Acad. Sci.* *98*, 2550–2554.

Zeitlinger, J., Zinzen, R.P., Stark, A., Kellis, M., Zhang, H., Young, R.A., and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* *21*, 385–390.

Zhang, Y., Wong, C.-H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E., et al. (2013). Chromatin connectivity maps reveal dynamic promoter–enhancer long-range associations. *Nature* *504*, 306.

Zhao, Z., Tavoosidana, G., Sjölander, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* *38*, 1341–1347.

Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M., and Furlong, E.E.M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* *462*, 65–70.

## **Chapter 1: bifunctionality of CRMs**

This chapter presents the major project of my PhD thesis and will be divided into four parts. The first subchapter will present the results of our initial library and first experiment. The second subchapter will present the work realized for a new library and the associated experiments, and their still ongoing analysis at the time of the writing of this thesis. I will then present current and future experiments and analyses necessary to submit this work for publication. Finally, I will conclude this chapter with a brief discussion about our current results.

**Article in preparation**

**Transcriptional silencers in *Drosophila* serve a dual role as transcriptional enhancers**

\*Stephen S. Gisselbrecht<sup>1</sup>, \***Alexandre Palagi**<sup>1,2</sup>, Julia M. Rogers<sup>1,3</sup>, Jesse V. Kurland<sup>1</sup>, Martha L. Bulyk<sup>1,4†</sup>

<sup>1</sup> Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.

<sup>2</sup> Doctoral School of Life and Health Sciences, University of Nice Sophia Antipolis, 06560 Valbonne, France.

<sup>3</sup> Committee on Higher Degrees in Biophysics, Harvard University, Cambridge, MA 02138, USA.

<sup>4</sup> Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.

\* **Co-first authors**

† **Corresponding author. E-mail: [mlbulyk@genetics.med.harvard.edu](mailto:mlbulyk@genetics.med.harvard.edu)**

## Author contributions

This work was supported by the National Institutes of Health (grant R01 HG005287 and R01 HG009723 to Martha Bulyk). Stephen Gisselbrecht, Alexandre Palagi, and Jesse Kurland performed experiments, Stephen Gisselbrecht and Alexandre Palagi performed data analysis and validations, Martha Bulyk supervised research, Stephen Gisselbrecht, Julia Rogers, Alexandre Palagi, and Martha Bulyk designed the study, and Stephen Gisselbrecht, Alexandre Palagi and Martha Bulyk are working on this manuscript.

## Abstract

A major challenge in biology is to understand how complex gene expression patterns in organismal development are encoded in the genome. While transcriptional enhancers have been studied extensively, few transcriptional silencers have been identified and they remain poorly understood. Here we used a novel strategy to screen hundreds of sequences for tissue-specific silencer activity in whole *Drosophila* embryos. Strikingly, 100% of the tested elements that we found to act as transcriptional silencers were also active enhancers in other cellular contexts. These elements were enriched in highly occupied target (HOT) region overlap (Roy et al., 2010) and specific transcription factor (TF) motif combinations. CRM bifunctionality complicates the understanding of how gene regulation is specified in the genome and how it is read out differently in different cell types. Our results challenge the common practice of treating elements with enhancer activity identified in one cell type as serving exclusively activating roles in the organism and suggest that thousands or more bifunctional CRMs remain to be discovered in *Drosophila* and perhaps  $10^4$ - $10^5$  in human (Heintzman et al., 2009). Characterization of bifunctional elements should aid in investigations of how precise gene expression patterns are encoded in the genome.

## Acknowledgments

We thank Anastasia Vedenko and James Anderson for technical assistance, Luis Barrera for helpful discussions, and Marc Vidal, Suzanne Gaudet, Shamil Sunyaev, Richard Maas, Alan Michelson, Kevin Struhl, and Trevor Siggers for critical feedback on the manuscript. This work was supported by a grant from the National Institutes of Health (R01 HG005287 and R01 HG009723) to Martha Bulyk.

## Introduction

Precise spatiotemporal control of gene expression is mediated by two types of *cis*-regulatory modules (CRMs): transcriptional enhancers and silencers (Ogbourne and Antalis, 1998). Investigations of transcriptional regulation in metazoans have focused primarily on *cis*-regulatory elements that activate gene expression. Transcriptional enhancers play crucial roles in gene regulation by activating gene expression in a tissue-specific manner in development, and in adult cells in response to cellular or environmental stimuli. However, it is also important that gene expression not be turned on or up regulated inappropriately. Transcriptional silencers are active negative regulatory elements (Ogbourne and Antalis, 1998) that play crucial roles in contributing to the specification of precise gene expression patterns, such as sharp expression domains in a developing organism, by preventing ectopic expression. Whereas enhancers have been characterized extensively, silencers are poorly understood and few have been identified across Metazoa.

Unlike for enhancer assays, no scalable screening technology is currently available to assay silencer activity in a metazoan. *Drosophila melanogaster* serves as a powerful model organism for investigations of spatiotemporal gene regulation in a developing animal. The silencer activities of two *Drosophila* CRMs have been described in the embryonic mesoderm (Jiang et al., 1993; Stathopoulos and Levine, 2005). Interestingly while being silencers in the mesoderm, these elements acted as enhancers in other tissues and can be described as bifunctional CRMs (Jiang et al., 1993; Stathopoulos and Levine, 2005). Furthermore, *in vivo* genomic occupancies of numerous TFs and chromatin marks have been profiled by chromatin immunoprecipitation (ChIP) –chip or ChIP-seq (Roy et al., 2010), and the activities of thousands of enhancers (Gallo et al., 2011) have been assayed, in *Drosophila* embryos. Thus, we reasoned that the developing *Drosophila* embryonic mesoderm would serve as a valuable model system in which to develop an approach to screen for silencers and then to inspect those silencers for enhancer activity in other tissues. We therefore applied a two-step strategy to identify silencers in the developing embryonic mesoderm in whole *Drosophila* embryos (Supplementary Fig. 1).

First, we have adapted our previously published technology for highly parallel screening of candidate enhancer sequences in *Drosophila* (Gisselbrecht et al., 2013) to enrich for sequences displaying dominant silencing activity in the embryonic mesoderm. We then individually tested non-promoter sequences that scored positive in this initial screen, to assemble a high-confidence list of validated mesodermal silencers. We found that, although we included several different types of genomic sequence in our candidate silencer library, all of the sequences in which we detected mesodermal silencing activity were enhancers in alternate cellular contexts.

Several genetic and epigenetic features are enriched in the set of validated silencers we have identified, but no combination of these gives enough predictive power to confidently identify silencers in the absence of experimental testing. Indeed, mutating instances of a transcription factor binding site motif significantly enriched in silencers had effects primarily on their activity as enhancers. We suggest that the widely observed distinction between enhancers and silencers may be an oversimplification, and that dual readout of regulatory information from bifunctional CRMs may be a common feature of transcriptional regulation.

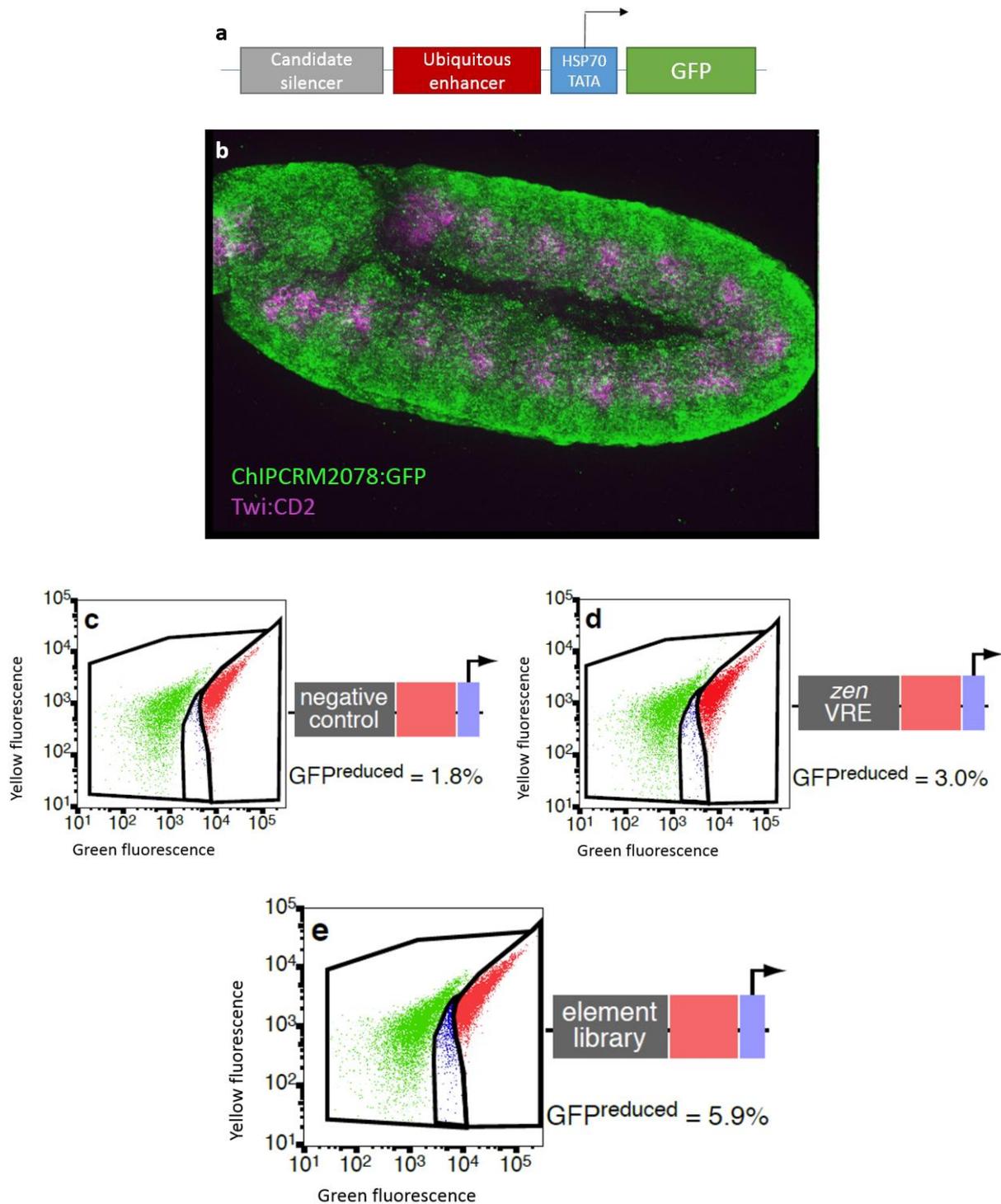
## Initial library and first experiment

### Screening a library of elements for silencer activity in whole *Drosophila* embryos

We adapted our previously published enhancer-FACS-Seq technology, for highly parallel screening of elements for enhancer activity in *Drosophila* embryos (Gisselbrecht et al., 2013), into ‘silencer-FACS-Seq’ (sFS) technology, which enriches for elements that tissue-specifically silence reporter gene expression (see Methods). We generated a reporter vector, pSFSdist, which drives GFP expression under the control of a strong, ubiquitous enhancer and an element from a library of candidate silencers, in a genomically integrated context (Fig. 1a,b, Supplementary Fig. 5). The reporter construct is integrated in a single location in the haploid genome via phiC31 integrase (Gisselbrecht et al., 2013). Flies carrying single insertions from the reporter library are crossed to a strain in which expression of the exogenous marker protein CD2 is driven in a tissue or cell type of interest, and the resulting informative embryos dissociated to produce a single cell suspension. By sorting CD2<sup>+</sup> cells in which GFP expression is reduced from the level driven by the strong ubiquitous enhancer in the absence of silencing activity, we enrich for cells containing silencers active in the cell type of interest, which can be recovered and identified by high-throughput sequencing.

### Selection of elements to test for silencer activity in *Drosophila* embryos

We designed a first library of 576 genomic elements (Supplementary Table 1), chosen to represent a variety of chromatin states or enhancer activities, to test for silencer activity. Since the general features of silencers are unknown, we pursued a few different strategies to identify elements that might correspond to repressive elements. The availability of genome-wide chromatin immunoprecipitation data for well-characterized transcriptional corepressors (Celniker, 2009) provided one source of candidate silencers to test. Next, we reasoned that some CRMs might function as enhancers in one context and as silencers in other contexts, as two such bifunctional CRMs had been identified previously in *Drosophila* (Jiang et al., 1993; Stathopoulos and Levine, 2005). Therefore, since we were designing a library to screen for silencing activity in the mesoderm, we selected CRMs from the REDfly and CAD2 databases (Gallo et al., 2011; Bonn et al., 2012) that exhibited no or highly restricted mesodermal expression at embryonic stage 11. We furthermore filtered out elements associated with genes that show widespread mesodermal expression at this stage. Another potential source of such bifunctional elements that we selected was genomic regions associated with markers of both active and repressed chromatin structure in whole-mesoderm or whole-embryo experiments (Bonn et al., 2012; Rosenbloom et al., 2015; Thomas et al., 2011). All sequences identified from genome-wide ChIP methods were associated with nearby genes (see Methods) and filtered for absence of widespread mesodermal expression. Finally, we included three positive control sequences previously shown to have mesodermal silencing activity, and two types of negative controls: broadly active mesodermal enhancers, and length-matched regions of *E. coli* genomic sequence. Using sFS, we screened our library of genomic elements for silencer activity in embryonic mesoderm. Testing of this library yielded a readily detectable population of mesodermal cells in which green fluorescent protein (GFP) expression is reduced (Fig. 1b,c). Of the 576 sequence elements chosen for inclusion in this library, we detected 372 in cells derived from transgenic flies. We found 79 elements significantly enriched (see Methods) in the reduced-GFP cell population in either of two biological replicate screens (Fig. 2a, Supplementary Fig. 2, Supplementary Table 3).



**Figure 1. Detection of silencer activity in a library of candidate sequence elements.**

(a) Design of the reporter construct, in which a candidate silencer is placed distal to a ubiquitous enhancer driving GFP expression and integrated into the *Drosophila* genome (see Methods). (b) A stage 11 embryo stained with  $\alpha$ -GFP (green) and  $\alpha$ -CD2 (magenta) antibodies shows GFP expression within the mesoderm (marked by *twi*:CD2) and elsewhere. (c-e) FACS of cells prepared from reporter embryos; GFP-expressing cells are characterized by a higher green:yellow fluorescence ratio. Embryos heterozygous for a negative control (non-silencer-containing: c) reporter show a mixture of GFP<sup>+</sup> (red) and GFP<sup>-</sup> (green) mesodermal cells. A known silencer (the *zen* VRE: d) increases the intermediate population of GFP<sup>reduced</sup> (blue) cells, as does the introduction of a library of candidate silencer elements (e).

## **Promoter competition in sFS-positive elements**

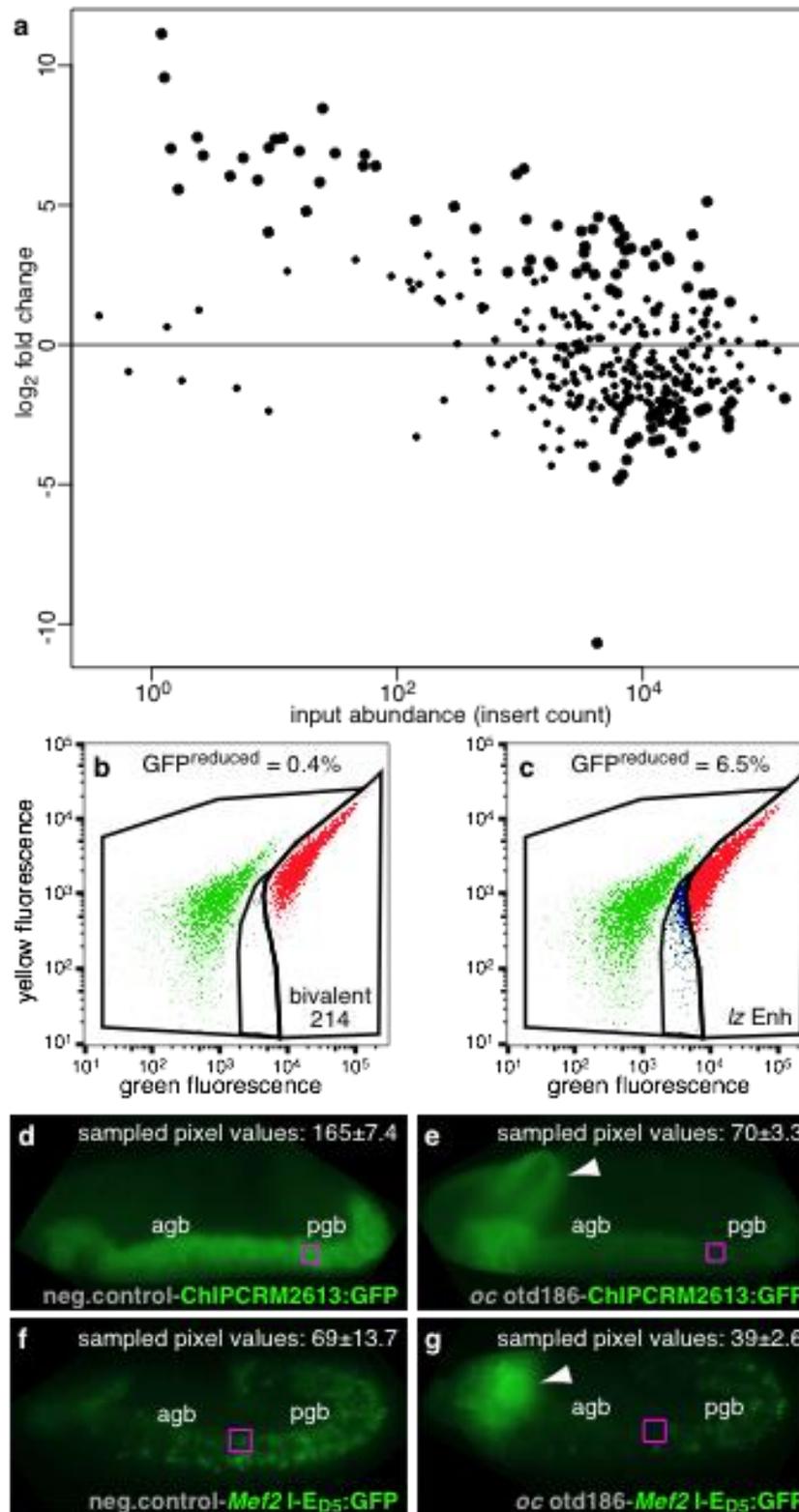
The most enriched feature among elements that were identified as ‘hits’ in our silencer screen is overlap with regions surrounding the transcriptional start sites of genes, which likely reveals the presence of promoter competition. Promoter competition previously has been observed to restrict the enhancer-driven activity of reporter gene promoters (Ohtsuki et al., 1998). Consistent with this interpretation, three of the sequences that caused reduced GFP expression were broadly active mesodermal enhancers that we included in our library as negative controls; all three of these sequences contain promoters. Overall, the set of 41 ‘hits’ that overlapped promoter regions was significantly enriched for mapped instances of the TATA box (Zhu et al., 2009). While these are technical positives in our silencer screen, since our goal was to analyze CRMs that silence gene expression by other means, we omitted any identified ‘hits’ that overlapped promoter regions from further analysis, resulting in a filtered set of 262 detected library elements, of which 38 are genomic regions showing silencer activity in our assay. In the second step of our silencer identification strategy, we then generated pure reporter lines and individually tested each of the 38 non-promoter sFS ‘hits’ for silencer activity by fluorescence-activated cell sorting (FACS) (see Methods). This resulted in a final, high-confidence set of 15 validated, mesodermal silencers (Fig. 2b,c, supplementary figure 5), which we considered a rather low validation rate.

## **Characterization of silencing activity in the context of distinct enhancers**

To confirm that the observed silencing activity does not represent an artifact of the FACS-based assay, we tested the ability of one of the newly discovered silencer elements to suppress activation by two different mesodermally restricted enhancers and visualized reporter gene expression in the resulting embryos. Both constructs showed silencing by the tested element relative to a negative control sequence (Fig. 2d–g). Interestingly, we observed different patterns of silencing in the context of these two mesodermal enhancers, which are active at different times in development; silencing activity was much weaker in the posterior germband at embryonic stage 12 (Fig. 2f,g), a pattern not observed in the earlier embryo (Fig. 2d,e). Thus, the repressive activity of silencers may exhibit complex spatiotemporal regulation similar to that of many enhancers.

## **All validated silencers act as transcriptional enhancers in other cellular contexts**

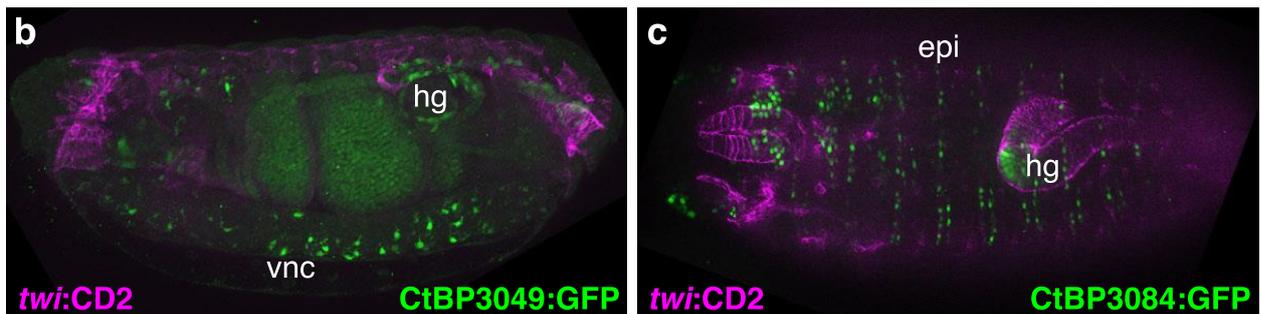
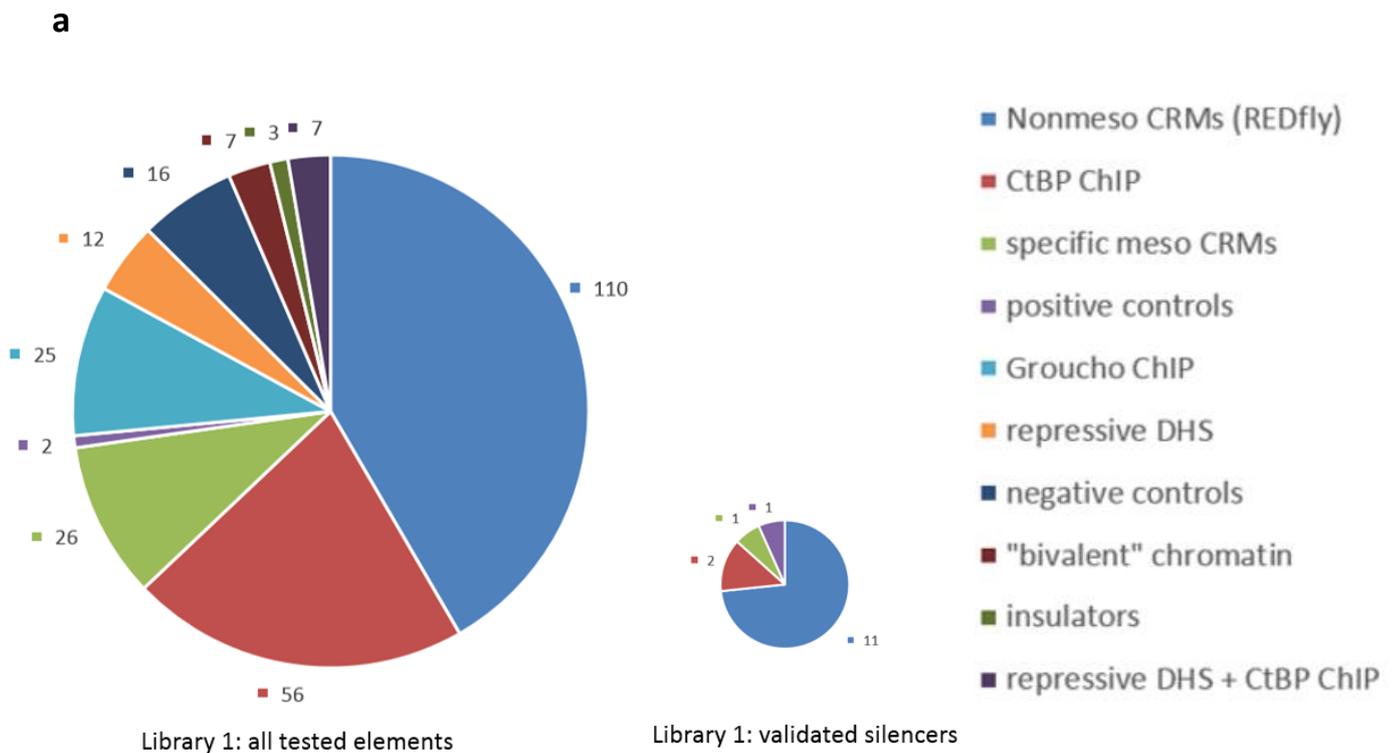
We analyzed these 15 validated mesodermal silencers to determine which genomic features that we explicitly sampled in the design of our element library were predictive of silencer activity. Despite the inclusion in our library of ChIP peaks for two well-known transcriptional corepressors (Groucho, CtBP) (Mani-Telang and Arnosti, 2007; Orian et al., 2007) and for the repressive chromatin mark trimethylated Lys27 of histone H3 (H3K27me3) (Kharchenko et al., 2011), the only screened element type significantly enriched among active mesodermal silencers was non-mesodermal enhancers (Fig. 3a;  $P = 0.0147$ , Fisher's exact test). In fact, all but two of the 15 high-confidence mesodermal silencers were previously reported to have enhancer activity. Testing of the remaining two silencers for enhancer activity revealed that they both also function as non-mesodermal enhancers in the embryo (Fig. 3b,c). Thus, our results suggest that most if not all mesodermal silencers are also enhancers in other cellular contexts.



**Figure 2. Detection and validation of active mesodermal silencers.**

(a) Enrichment ratios for library elements in *twi*:CD2<sup>+</sup>GFP<sup>reduced</sup> cells as compared to input cells. Large points: P<sub>adj</sub> < 0.1 (significantly enriched). (b,c) Examples of results from silencer validation assays. "Bivalent chromatin" element 214 (b) is not significantly enriched in *twi*:CD2<sup>+</sup>GFP<sup>reduced</sup> cells and shows no expansion of the GFP<sup>reduced</sup> population, while the *lz* Crystal Cell Enhancer (c) shows significant enrichment and validates positive. (d,e) Embryos carrying a negative control (d) or validated silencer (*oc otd186*: e) (grey text) modifying the activity of the mesodermal enhancer ChIPCRM2613 (green text) were stained with α-GFP

antibody and photographed with identical exposure times at embryonic stage 7. Silencing of GFP expression is stronger in the posterior germ band (pgb) than in the anterior germ band (agb). **(f,g)** Embryos carrying a negative control **(f)** or validated silencer **(g)** (grey text) modifying the activity of the later-acting mesoderm-specific enhancer *Mef2* I-ED5 (green text) were stained and photographed at embryonic stage 12. At this stage, silencing is stronger in the anterior germ band than in the posterior germ band. Comparable regions of pairs of embryos were sampled (magenta boxes) and fluorescence intensity quantified. Arrowheads in **e** and **g** show the previously described pattern of activity in head segments associated with the *oc otd186* enhancer (Gao and Finkelstein, 1998), in which we detected mesodermal silencer activity.



**Figure 3. Silencer activity is detected only in enhancers.**

(a) Contribution of different sequence types to the library of elements tested for silencer activity (left) and to the population of validated silencers (right). Area is proportional to number of library elements in each class. (b,c) Embryo images showing that newly discovered mesodermal silencers not previously known to be enhancers drive reporter gene expression (GFP, green) outside the mesoderm (marked by *twi:CD2*, magenta). CtBP3049 (b) drives expression primarily in the ventral nerve cord (vnc) and hindgut (hg); CtBP3084 (c) drives expression in epidermal stripes (epi) and hindgut.

## Transcription factor compositional complexity at silencers

We hypothesized that, as these elements are more functionally complex than CRMs that act only as enhancers, they may exhibit a more complex suite of TF interactions across various tissues. Indeed, we observed that validated silencers are strongly enriched for overlap with highly occupied target (HOT) regions, defined by Roy *et al.* as exceeding a TF complexity score threshold of ~10 overlapping, bound factors (Roy *et al.*, 2010) ( $P = 0.00026$ , Fisher's exact test).

To investigate DNA sequence features that might be important for silencer activity, we inspected the enrichment of combinations of DNA binding site motifs for annotated repressors (Supplementary Table 8) among the validated silencers (see Methods). A motif or motif combination was considered significantly enriched if it targets  $\geq 50\%$  of the foreground sequences and has an area under the receiver operating characteristic curve [AUROC]  $\geq 0.65$  and a q-value  $\leq 0.1$  (see methods). Five motif combinations were significantly enriched among mesodermal silencers (Fig. 4a); all shared a core combination of motifs for Snail (*sna*) and Defective proventriculus (*dve*). *Sna* is a well-known repressor of non-mesodermal genes in the developing mesoderm (Nieto, 2002). The *dve* motif is characteristic of the Bicoid group of homeodomain proteins (Noyes *et al.*, 2008), which are not known to be involved specifically in mesodermal repression. Mutation of the *dve* motif sequence 5'-TAATCC-3' in four validated silencers did not detectably impair their silencer activity (Supplementary Fig. 4); however, three of these four mutated elements exhibited consistent alterations of their enhancer activity (Fig. 4b-d, Supplementary Fig. 4), further emphasizing the bifunctional nature of this class of *cis*-regulatory elements.

## Chromatin features of active silencers

Various post-translational modifications of histones have been associated with different functional elements in the genome (Ernst and Kellis, 2010; Fillion *et al.*, 2010; Heintzman *et al.*, 2007; Roy *et al.*, 2010); however, little is currently known about the chromatin features of active silencers (Riel, 2014). We found that mean H3K27me3 ChIP signal is mildly enriched (area under the receiver operating characteristic curve [AUROC]  $\sim 0.60$ ,  $P < 0.03$ , two-tailed Wilcoxon test) (see Methods) among validated mesodermal silencers. Considering additional marks beyond H3K27me3, analysis of a broad panel of whole-embryo ChIP-chip and ChIP-seq data for histone modifications, TFs, coactivators, and corepressors (see Methods) suggested a variety of other marks enriched or depleted among silencers, but none – including H3K27me3 – was statistically significant in this broader analysis after correction for multiple hypothesis testing (Supplementary Table 9). Thus, these results suggest that H3K27me3 may be a potential feature of active, tissue-specific silencers, but it alone is not sufficient to distinguish silencers from other genomic elements.



## Second library and ongoing analyses

### Second library

Given the limited number of silencers identified, we decided to test more element for silencing activity and designed a second library based on the same criteria used for the first library, but automatically filtering for elements that overlapped promoter regions to avoid promoter competition, as previously explained. This second library was composed of a total of 173 elements, including 26 repeats from the first library (supplementary table 2): 14 elements found as acting as silencers in our previous experiments and 12 E.coli genomic DNA negative controls. In designing this second library, we were limited by the datasets available to us, from which we designed our initial library of candidate silencers.

### Results and validations

Of the 173 elements that were included in this second library, we detected 147 elements in cells derived from transgenic flies. We found 15 new elements significantly enriched in the reduced-GFP cell population in either of the two biological replicate screens, excluding cells with no GFP expression as we hypothesized that this could increase the sensitivity of our assay. Indeed, many of the GFP- cells are in fact not carrying any reporter from our library within their genome, as they are the result of the cross between the transformant males, carrying only one copy of our library in their genome, with our homozygous *twi:CD2* virgin females. We hypothesize that processing the GFP- cells with the GFP<sup>reduced</sup> population added a significant amount noise to the data and was the reason for the poor validation rate our method was suffering from. Therefore, we decided to sort and process only the GFP<sup>reduced</sup> cell population (see Methods).

The repeats from the first library validated relatively nicely, with only a few exceptions: we detected 20 out the 26 repeats. 85% of these 20 elements showed the same activity as in the initial screen, as we did not detect 3 of these elements as silencers. To validate these results, we generated pure reporter lines and individually tested 8 elements by FACS, as previously described (see Methods). In the interest of time, we decided to validate 4 elements found as significantly enriched by our assay and 4 that were not (see Methods). All 8 elements validated and therefore we assumed that all 15 elements were *bona fide* silencers, and concluded that the modification to cell sorting fixed the low sensitivity our method was suffering from (Supplementary table 6).

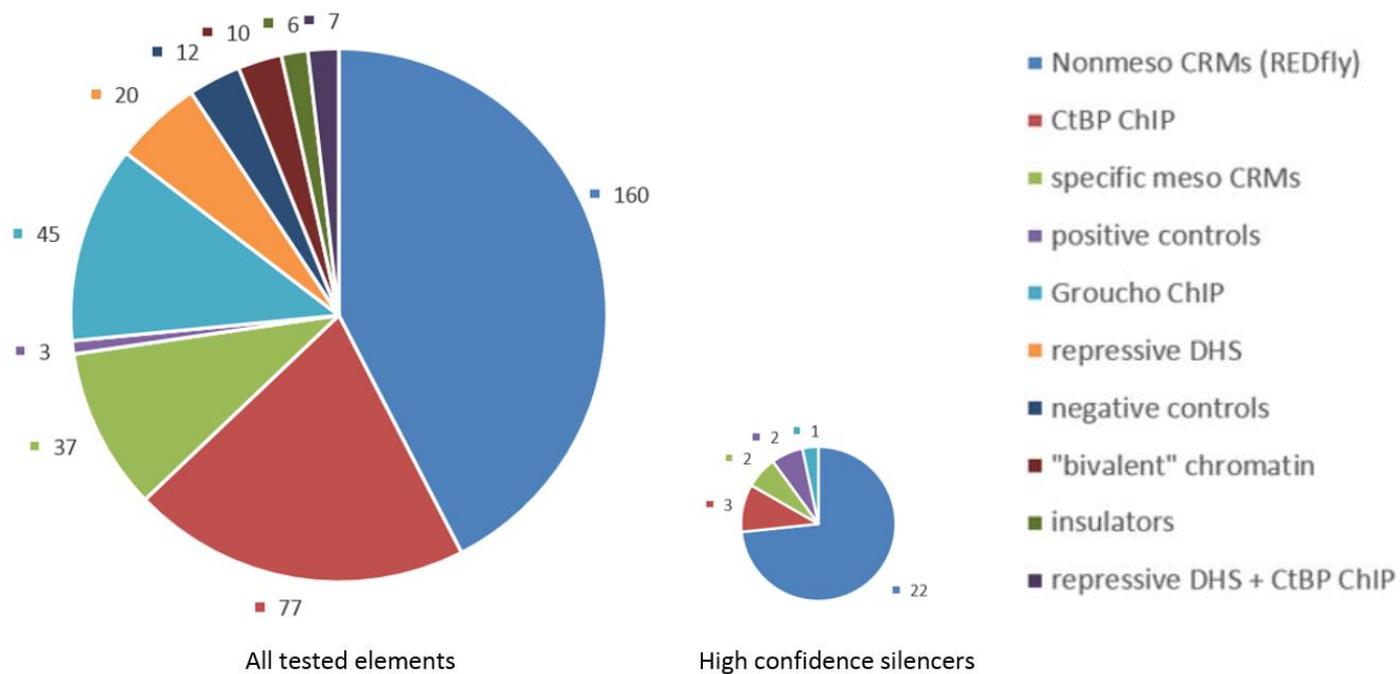
These two libraries, once combined, gave a total of 30 newly identified silencers (Fig. 5, supplementary table 7), with a significant majority of them being already characterized as enhancers in other tissues ( $P = 0.0004393$ , Fisher's exact test).

### TF compositional complexity and chromatin features of active silencers

Combining both libraries, we found out that the validated silencers were still somewhat enriched for overlap with HOT regions (as previously defined), with a complexity score in validated silencers very close to the previously set threshold of 8 (see methods), with a score of  $\sim 7.75$  (AUROC = 0.732,  $P = 2.815e-05$ ) compared to  $\sim 4.09$  in sFS negative elements. We therefore ran a motif finding on these new elements (see Methods) as we previously did on our first set of validated elements, but no combination of motif was significantly detected. As this thesis is being written, an analysis (see Methods) on all validated silencers pooled together is running, and therefore could not be included here.

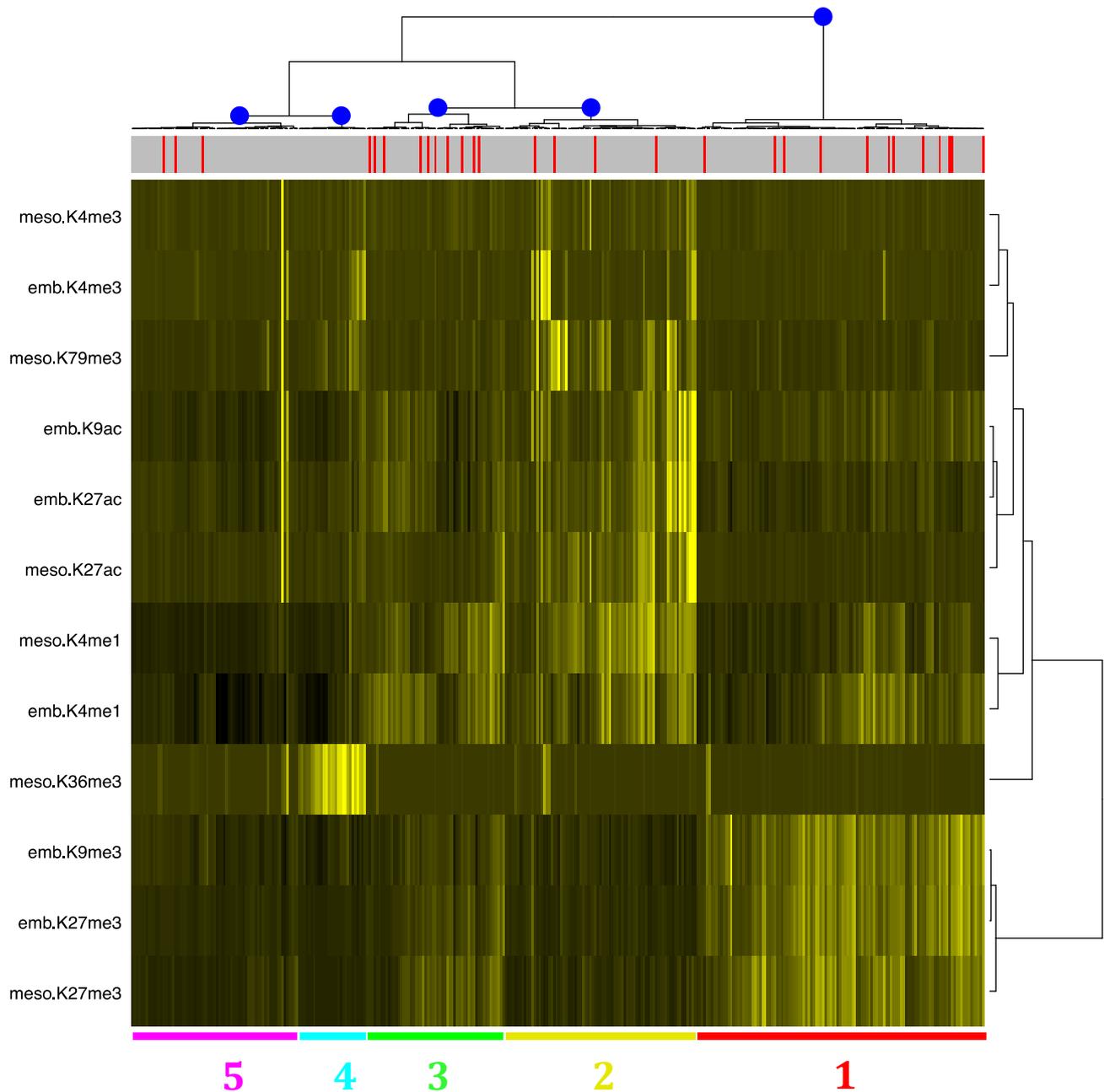
Regarding histone marks, the mild enrichment of the H3K27me3 mark that we observed previously appears to hold up. We found that whole embryo H3K27me3 ChIP signal is mildly enriched (area under the receiver operating characteristic curve [AUROC]  $\sim 0.61$ ,  $P < 0.04$ , two-tailed Wilcoxon test (see Methods)) among validated mesodermal silencers, yet not significantly after correction for multiple hypothesis testing (adjusted  $P \sim 0.16$ ) (see supplementary figure 10). Looking at other additional marks beyond H3K27me3, analysis and biclustering (see mMethods and Fig. 6) of whole-embryo and sorted-mesoderm ChIP-chip and ChIP-seq data for histone modifications (see Methods) revealed that our validated silencers were significantly depleted for the H3K36me1 mark, which is commonly associated with elongating Polymerase II (Ernst and Kellis, 2010), and that a subset of our silencers showed statistically significant enrichment for moderate levels of repressive histone marks (see Fig. 6).

Once again, these results suggest that H3K27me3 may be a potential feature of active, tissue-specific silencers, but it is not sufficient to distinguish silencers from other genomic elements. Therefore, it is obvious to say that further studies, including larger sets of active silencers, are needed to discover significant silencer-associated chromatin marks which together might enable accurate prediction of silencers and offer mechanistic insights on silencer activity.



**Figure 5. Combined libraries and validated silencers**

Contribution of different sequence types to the combined libraries of elements tested for silencer activity (left) and to the population of total high confidence silencers (right). Area is proportional to number of library elements in each class.



**Figure 6. Repressive chromatin marks are associated with silencer activity but are insufficient to identify silencers.**

Mean ChIP-seq signal over each tested library element was calculated for a range of published datasets using chromatin from whole embryos ("emb") or sorted mesodermal cells ("meso") and antibodies to the indicated histone H3 modifications, then Z-transformed and truncated. Biclustering reveals 5 major clusters of elements with broadly similar score profiles (colored bars at right). Aside from cluster 4 (highly enriched for trimethylated H3K36 in the mesoderm, which characterizes regions associated with elongating PolII), all clusters contain a mixture of mesodermal silencers (red in the bar at the top) and nonsilencers (grey). However, clusters 1 and 3, the elements of which show enrichment for histone marks associated with transcriptional repression, are enriched for silencer activity. Interestingly, only cluster 3, with more moderate levels of repressive histone mark enrichment, shows statistically significant enrichment for mesodermal silencers.

## Current and future experiments and analyses

### Genome editing: CRM knock-out

To further assess the role and function of the identified bifunctional elements, we initiated a project, which aims to investigate the effect of the knock-out (KO) of validated bifunctional elements by genome editing. This project is based on the use of the CRISPR/Cas9 technology (Jinek et al., 2012), consisting of the use of an endonuclease, Cas9, and a guide RNA (gRNA). This technology has been adapted from a prokaryotic form of acquired immunity, providing the cell with a resistance to foreign genetic elements carried by bacteriophages (Barrangou et al., 2007).

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats. These repeats were first identified when the *iap* gene was cloned into *E. coli* (Ishino et al., 1987), and are found in ~40% of sequenced bacterial genomes and 90% of sequenced archaea genomes (Grissa et al., 2007). These sequences contain snippets of DNA, the “spacers” in between the repeats, from viruses that previously challenged the bacterium, and it has been shown that bacteria integrated new spacers derived from phage genomic sequences after viral challenge (Barrangou et al., 2007). Briefly, an RNA molecule harboring a spacer sequence helps CRISPR-associated (Cas) proteins to recognize, in a complex also called ribonucleoprotein (RNP), bind and cut exogenous DNA, rendering it inert (Barrangou et al., 2007). This mechanism is based on RNA interference (Makarova et al., 2006) and acts as an acquired immune defense system against such viral encounters for prokaryotes. In 2012, the CRISPR/Cas9 mechanism was detailed and proposed as a programmable tool for genome-engineering and editing (Jinek et al., 2012). Indeed Cas9 can site-specifically cleave double-stranded DNA which results in the activation of the double-strand break (DSB) repair machinery. DSBs can be repaired by the Non-Homologous End Joining (NHEJ) pathway (Overballe-Petersen et al., 2013), an “emergency” response to DSBs, which results in insertions and/or deletions (indels) which disrupt the targeted locus. If a donor template with homology arms to the targeted locus is supplied, the DSB may be repaired by the homology-directed repair (HDR) pathway, which allows for precise replacement and therefore precise genome editing (Overballe-Petersen et al., 2013).

It has then been demonstrated that the CRISPR-Cas9 system is able to specifically target and mutate certain genes in mice using designed guide-RNAs against specific genomic targets on the mouse genome (Wang et al., 2013); shortly after this, the CRISPR/Cas9 system was utilized successfully in humans to correct B-thalassemia in afflicted patients (Xie et al., 2014). The CRISPR/Cas9 system is now seen as a powerful and versatile genome editing tool, with many applications in the fields of biotechnologies, research and medicine (Hsu et al., 2014), and its component are readily accessible from the industry (*e.g* IDT, NEB, etc.).

For our project, guide RNAs (gRNAs) have been designed to remove a whole CRM (see Methods), or each (one at a time) of its specific sub-sections based on genome conservation profiles across the *Drosophila* species. The Cas9 and gRNAs pairs were ordered from IDT (annex 3) and assembled into ribonucleoprotein (RNP) complexes and injected (see Methods) in appropriate fertilized *Drosophila melanogaster* embryos. The effects of the KOs will be assessed by survival rates, microcopy of developing embryos and muscle structures, along with *in-situ* hybridization and RT-qPCR for assessing the expression genes surrounding the previously present CRM.

As a proof of concept for the CRISPR-Cas9 system, we designed and ordered a gRNA from IDT targeting the *white* gene which is responsible for red pigmentation of the adult *Drosophila melanogaster*. This gRNA was incubated with the Cas9-3NLS protein from IDT (see methods) and the resulting RNPs were microinjected posteriorly into wild type fertilized eggs, to edit the nuclei giving rise to the germ band. To our surprise, all of the injected eggs that hatched (21%) and gave viable adults showed mosaic or completely white eyes (*white*- or *w*-) (see figure 7).

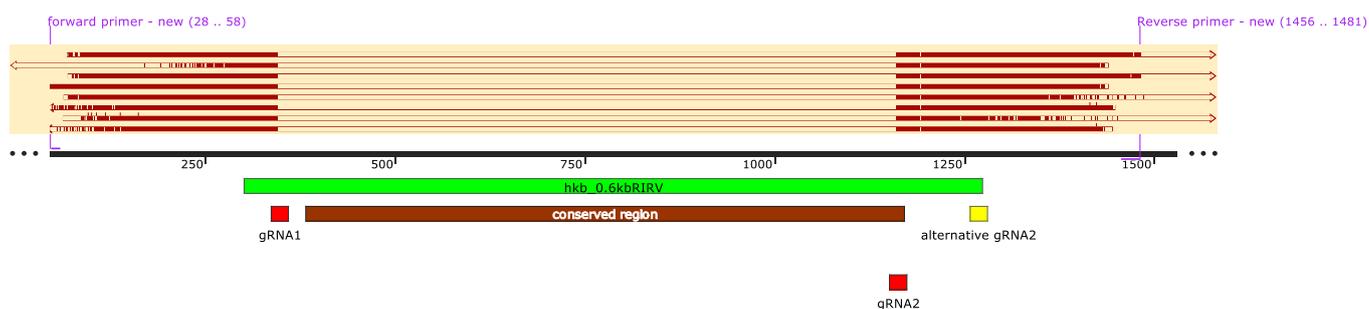
These mutant flies, crossed to *yw* flies, were unsurprisingly carrying at least one mutated and therefore inactive version of the *white* gene: out of the 21 flies crossed individually to *yw* males or females (depending on their respective gender), the F1 progeny (797 flies), from which on average 85% were *w*- (total of 84%, the maximum of 100% per cross, and the minimum of 33% per cross). Such a result, though encouraging, suggest that the RNP mix solution needed to be titrated down, to edit only the nuclei from which originates the germ band and prevent genome editing in other cells as much as possible. Indeed, if knocking out a given bifunctional element is lethal, it is necessary to make sure that the cause of the lethality is linked to the genome editing and knock out, and not to the potential toxicity of the microinjections, or the manner by which the microinjections were realized. We encountered issues we believe are linked to the viscosity of the RNP solution due to the glycerol it contains (see Methods), which sometimes seemed to clog the microinjection needle, and therefore appears as another reason to titrate down this solution in the future.

Nonetheless, we went forward with an initial trial and successfully knocked out the *hkb\_0.6kbRIRV* bifunctional element we previously identified by sFS (see Fig. 8), and established a viable line carrying the knockout over balancer chromosome (TM3). The knockout appears to have taken place precisely where we intended and this line awaits further analyses. An example of analysis is *in situ* hybridization of transcripts from genes located in the vicinity of the knocked-out bifunctional CRM which it was likely to regulate. Probes have already been designed for such genes and, given time constraints, Stephen Gisselbrecht will continue these experiments. Overall, we have identified a method that allows precise sequence knockout in *Drosophila melanogaster*. Though this method needs to be perfected, it has proven effective for our initial proof of concept. Therefore, the Bulyk lab should continue improving this method and use it to reliably dissect the potential complex regulatory networks within which the bifunctional elements we identify may belong.



**Figure 7. Mosaic eyes in Cas9/gRNA injected flies**

Mosaic eyes were observed in flies that were previously microinjected, at the stage of fertilized eggs and posteriorly, with a combination of Cas9 and gRNA directed against *white*. All the flies from this microinjection showed either mosaic (white and red) or totally white eyes.



**Figure 8. Bifunctional CRM knockout by CRISPR/Cas9**

We successfully managed to knock out the hkb\_0.6kbRIRV bifunctional element from *Drosophila melanogaster* using the CRISPR/Cas9 system. This figure shows the results of several sequencing experiments from PCR reactions from adult flies from the line we established, that carries the knockout induced by the gRNAs 1 and 2 (red squares), which surround the observed conserved region across all *Drosophila* species.

## Hi-C: mesoderm specific interactions

In an attempt to reach a potential mechanism that could explain how silencers and bifunctional CRMs function, we initiated a collaboration with the Dekker lab (<http://my5c.umassmed.edu>) from the University of Massachusetts. This collaboration consisted on obtaining Hi-C data from the mesodermal versus non-mesodermal cells from *D. melanogaster* embryos, at the same embryonic stage we did our sFS experiments (~stage 11), and try to observe differential spatial organization of chromosomes between both cell populations. The idea behind this project is to look for overall differential contacts between the developing mesoderm and non-mesodermal tissues, therefore tissue-specific contacts, but more precisely to analyze the behavior of our validated silencers and bifunctional CRMs.

To do so, we used once again our *twi:CD2* line (see Methods), and sorted by FACS mesodermal cells (CD2+) and non-mesodermal cells (CD2-) from homozygous embryos at stage ~11. We proceeded with the fixation of these cells with formaldehyde (see Methods). After several weeks of embryo collection and cell sorting, we managed to accumulate and fix ~22 million CD2+ cells and ~25 million CD2- cells. All these cells were sent, on dry ice, to the Dekker lab for their processing. These cells were processed by the Dekker lab as two biological replicates, each replicate experiment consisting of ~11 million CD2+ and ~12.5 million CD2- cells. At the time this dissertation is being written, these samples were processed by the Dekker lab. The 4bp-restriction enzyme DpnII (GATC cut site) was used for processing our samples, which in theory is appropriate for looking into specific interaction of a small genomic region such as CRM interactions and enhancer-promoter looping, while 6-bp restriction enzymes are more appropriate for large scale structures, as 4-bp restriction enzyme have a higher probability of finding their target restriction site over 6-bp restriction enzyme which allows for more frequent cuts and therefore a higher resolution (Lajoie et al., 2015). Indeed, DpnII in theory provide 16 times as many ligation opportunity when compared to other 6-bp restriction enzymes such as *Eco*r1 or *Hind*III for instance. We have ~80 million reads for one of the CD2+ and CD2- replicate, and ~100 million reads for the other replicate. It is still unclear whether the resolution of the data will be high enough to precisely study the ~1kb elements we tested by sFS and distinguish them from background noise, as there is currently no standard way to assess the resolution of Hi-C data.

Hi-C data processing must go through several steps before analysis. First, sequencing reads must be mapped onto a genome of interest and each mapped read is therefore assigned to one of the restriction fragment. Secondly, basic filtering is applied to remove from further analyses unligated fragments or ligated circularized fragments (self-circling fragment). Then, identical PCR artifacts are removed, such as undigested restriction sites or same paired-end sequence. After filtering, the data is binned into a fixed (or sometimes various) genomic interval size since Hi-C datasets normally cannot be sequenced deep enough to support the maximal data resolution determined by the restriction enzyme used. Once the data is binned, it is then filtered out for low or very noisy signal. Finally, the data must be corrected for mapability, GC content, or fragment length for instance (Yaffe and Tanay, 2011; Hu et al., 2012; Cournac et al., 2012). The resulting product is a genome-wide, binned, interaction matrix from which it is, in theory, possible to identify interaction frequencies between genomic loci with some limitations. Indeed, the measurements are realized on a population of cell, which may lead to an interaction matrix with little position-specific interactions, as such structures may vary due to cell-to-cell heterogeneity and thus may not be identifiable. Moreover, patterns of interaction are not formally defined, and different patterns of interaction can overlap each other, adding another layer of complexity to the analyses. Also, we do not know the impact of the disruption of the cell-cell interactions, prior to FACS, on transcription. Finally, the frequency of interactions measured by Hi-C are not

indicative of the physical distances between the interacting pairs, nor can they be interpreted as frequencies in time.

As stated before, the goal of this study is to try and observe potential differential interactions from the bifunctional CRMs we identified by looking at interactions they make in mesodermal cells to interactions in non-mesodermal cells, as these elements appear to be silencers in the mesoderm while acting as enhancers in other tissues. It is possible that a bifunctional CRM does not interact differently when acting as a silencer or enhancer, but that its activity depends on tissue-specific cofactors, in which case we would not observe any differences between our two datasets. It is also possible that the frequencies of interaction will differ, or that some interactions may be present in one set and absent in the other.

Currently, the reads are being processed and analyzed by Julia Rogers and Stephen Gisselbrecht (Bulyk lab) and no usable data is yet available. Our approach might be limited by its resolution, or the potential heterogeneity within the populations of cells, as the mesodermal cells in a developing embryo lead to several different tissues and cell types, which is even more problematic with all the non-mesodermal cells. Nonetheless, we hope this Hi-C data may help in the characterization of bifunctional elements, by providing tissue specific contacts and more importantly, differential spatial contacts regarding the bifunctional CRMs we identified by sFS, provided that its resolution would be of ~1kb. We are hoping that a mechanism (or several), or clues of a mechanism, governing these elements will arise from this project.

## Spatiotemporal activity of silencers

The activity of the silencers we identified needs to be further assessed in different conditions and in a spatiotemporal manner. Indeed, it is necessary to know whether the silencing effect observed by the sFS method is specific to the enhancer we used and, or, specific to the embryonic stage at which the experiment was realized.

To further address this question, it will be necessary to test the activity of different bifunctional CRMs when placed in front of different mesodermal specific enhancers. We currently have a total of 4 different gateway compatible plasmids derived from the pSFSdist vector, each of them bearing a different enhancer replacing the ubiquitous ChIPCRM2078 enhancer: ChIPCRM2613 drives widespread mesodermal expression (Gisselbrecht et al., 2013), *Mef2* I-ED<sub>5</sub> which drives expression specifically in the fusion competent myoblasts (FCMs) of the developing mesoderm (Duan et al., 2001), ChIPCRM7759 drives expression in the early gastrulating mesoderm (stage ~7/8) (Gisselbrecht et al., 2013), and finally ChIPCRM2497 which drives expression in the late mesoderm (stage ~11/12) (Gisselbrecht et al., 2013). The resulting plasmids are respectively p2613d, pIED5, p7759d and p2497d.

Elements to be tested, *i.e.* one each of a positive control, a negative control, and a newly discovered silencer, will be cloned into all 4 of these vectors. From these, homozygous fly lines will be created and embryos will be analyzed by microscopy. These embryos will need to be collected, fixed, and stained in parallel, and imaged with the same settings (same exposure time, no gamma or contrast correction, see Methods).

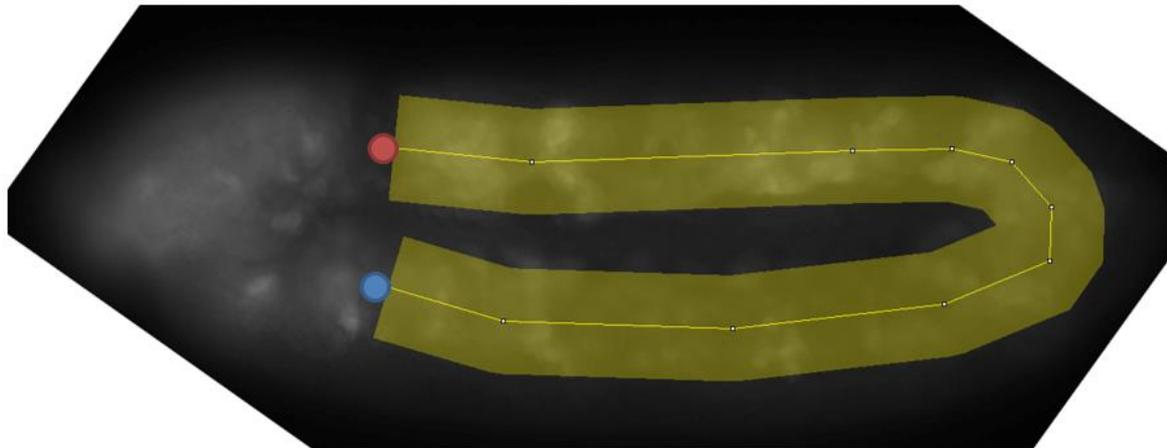
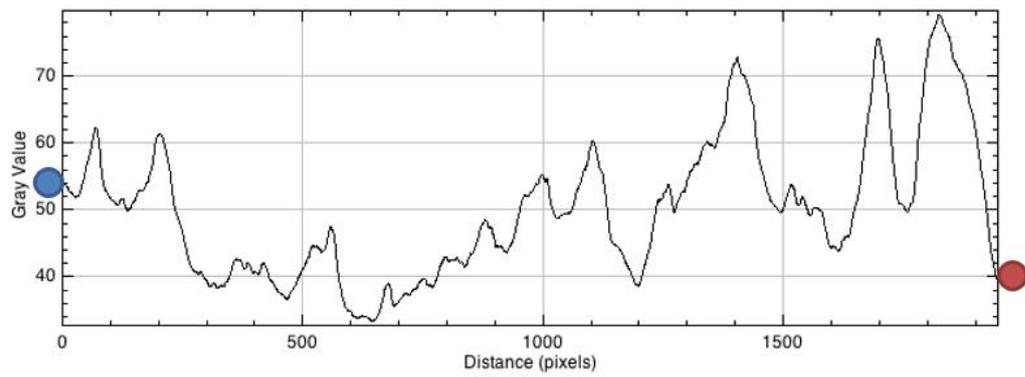
Moreover, we will try to make use of other picture analyzing software such as ImageJ (Schneider et al., 2012), to measure pixel intensity along the mesoderm, following the shape of the mesoderm in a given embryo. As a proof of concept, an example of the intensity diagram provided by ImageJ can be seen in Figure 9 where the intensity represented is the mean intensity measured on a line perpendicular to the tangent of the curve following the mesoderm, in the green channel only. A comparison between 3 embryos for a negative control (line 324) and the *brk*\_NEE-long bifunctional element (line 375) in the p2497d plasmid can be found in Figure 10, where silencing is clearly observable, and in the p7759d plasmid in Figure 11, where no silencing activity is observable. In these examples, the mesoderm was selected from the base of the developing head to the end of the tail of the embryo as shown in Figure 9.

While this approach is promising for measuring the differences in GFP expression between embryos, significance of observations and measurements will need to be addressed. At this stage of this project, several homozygous lines have been created for several bifunctional elements but embryos still need to be collected, processed and analyzed. From what we seem to observe so far, the tested silencers from our high confidence indeed acted as silencers, but some only acted as such with certain enhancers. For instance, the *brk*\_NEE-long element tested in Figure 10 and 11 has an obvious observable silencing activity on the ChIPCRM2497 enhancer (late mesoderm enhancer, stage ~11/12), while no silencing is observable on the ChIPCRM7759 enhancer (gastrulating mesoderm, stage ~8).

These preliminary results seem open the door to a very complex set of interactions and relationships between CRMs, in a spatiotemporal manner, within a large network of gene regulation where CRMs play various roles as enhancer and silencer in tissues and cell types, and at different times, or embryonic stages here.

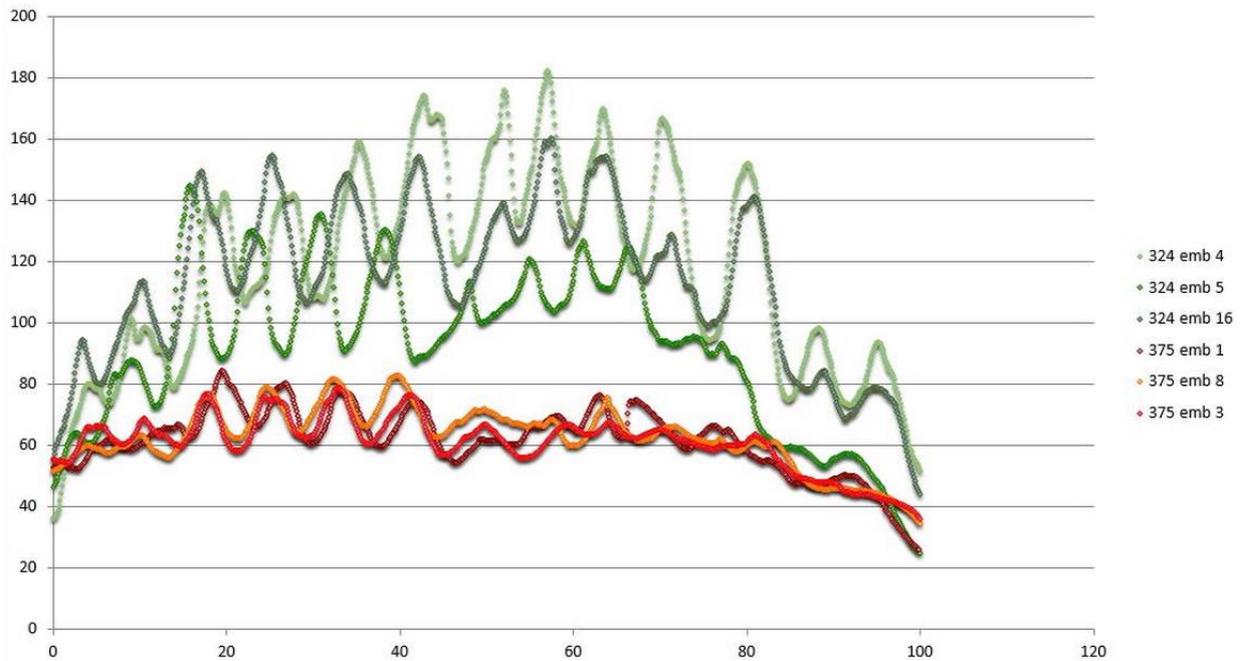
Silencers have been thought to play a key role in the specification of precise gene expression patterns, by enabling, for instance, sharp expression domains in a developing organism by insuring against ectopic expression. For instance, it has been shown that the ventral repression element (VRE) from the *zerknüllt* (*zen*) gene in *Drosophila melanogaster* is required to repress ventral expression of *zen* in the developing embryo (Doyle et al., 1989) and that the deletion of this VRE results in a strong ventral misexpression of *zen*, confirming the idea of the importance of silencers for precise control of expression (Busturia et al., 1997).

Although the significance of our observations needs to be assessed and these silencers tested in different configurations (with different enhancers at different embryonic stages), our preliminary results are interesting, as they seem to show that our high confidence set of silencers did not represent an artifact of the FACS-based assay. These silencers, which appear to be enhancers in other tissues, may act as silencers on other specific enhancers, in a cell-type and embryonic stage specific manner, as key components of the precise spatiotemporal control mechanism(s) of gene expression.



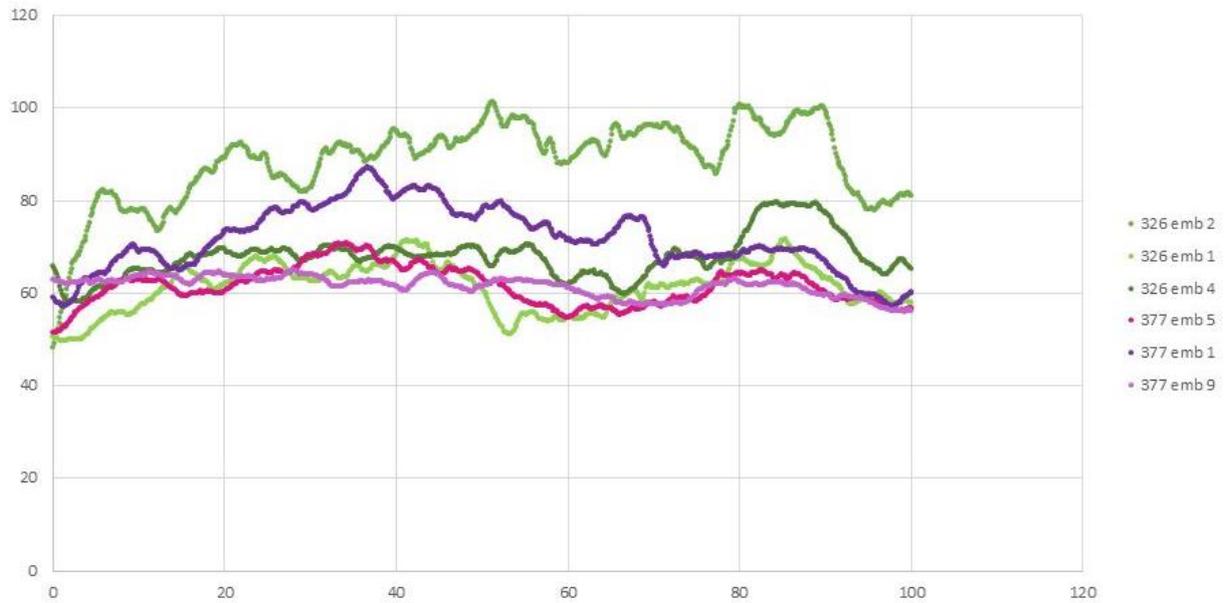
**Figure 9. Example of ImageJ intensity measurement**

This figure presents the pixel intensity diagram provided by ImageJ when selecting the mesoderm from the base of the developing head (blue dot) to the end of the tail of the embryo (red dot). The distance between these two positions is shown on the diagram as pixel values in the X axis. In this example, the embryo analyzed carries the *e\_coreAbdominalCRE* bifunctional CRM upstream of the *ChIPCRM2613* enhancer, and the picture was taken at stage ~8.



**Figure 10. Example of pixel intensity measurements for brk\_NEE-long in p2497d**

Here, the X axis is converted from pixel number to percentage of the measured section of the mesoderm, for a better comparison between embryos. This plot shows the difference of expression induced along the mesoderm for *E.coli* negative control (ecoli\_ctrl15, line 324, shades of green) and the brk\_NEE-long (line 375, shades of red) element upstream of the ChIPCRM2497 enhancer. In this example, the pictures were analyzed for embryos at stage ~11/12.



**Figure 11. Example of pixel intensity measurements for brk\_NEE-long in p7759d**

Here, as in Figure 10, the X axis is converted from pixel number to percentage of the measured section of the mesoderm. This plot shows the difference of expression induced along the mesoderm for *E.coli* negative control (ecoli\_ctrl15, line 326, shades of green) and the brk\_NEE-long (line 377, shades of purple) element upstream of the ChIPCRM7759 enhancer. In this example, the pictures were analyzed for embryos at stage ~8/9.

## Discussion

A major challenge in biology is to understand how complex gene expression patterns in organismal development are encoded in the genome. Precise spatiotemporal control of gene expression is mediated by two types of cis-regulatory modules (CRMs): transcriptional enhancers and silencers. While enhancers have been studied extensively, few silencers have been identified and they remain poorly understood. CRM bifunctionality complicates the understanding of how gene regulation is specified in the genome and how it is read out differently in different cell types. It is unknown how common such dual encoding of regulatory information may be in a typical metazoan genome.

In this chapter, I described our sFS strategy to screen hundreds of sequences for tissue-specific silencer activity in whole *Drosophila* embryos, which enriches for elements that tissue-specifically silence a strong, ubiquitous enhancer that drives GFP reporter expression.

Strikingly, the elements that we found to act as transcriptional silencers were also active enhancers in other cellular contexts. These bifunctional CRMs were enriched in highly occupied target (HOT) region overlap but, so far, did not seem enriched for any particular transcription factor (TF) motif combinations. These results challenge the common practice of treating elements with enhancer activity identified in one cell type as serving exclusively activating roles in the organism and suggest that thousands or more bifunctional CRMs remain to be discovered in *Drosophila* and perhaps  $10^4$ - $10^5$  in human (Heintzman et al., 2009).

The characterization of bifunctional elements should help in investigations on how precise gene expression patterns are encoded in the genome. Nevertheless, this discovery is currently based on a small set of newly discovered bifunctional elements and no underlying mechanism or signature has been identified yet. Our approach is moreover not exempt of limitations: if an element acts as a silencer in only a small fraction of mesodermal or non-mesodermal cells, it might not be detected as significant in our screen. Future screens using other driver lines specific for various mesodermal, or non-mesodermal, subsets might identify additional silencers with more restricted cell-type-specific silencing activity. Our approach is moreover time sensitive and mostly limited by the necessary cell sorting step, which *de facto* limits the size of the library we can screen to be able to detect any significant signal.

Looking at the current analyses and experiments, we hope that the statistical and computational analyses on all our datasets, sFS and potential Hi-C datasets, along with experimental validations such as the CRM knockout project will soon unveil a way to characterize and identify these bifunctional elements.

## Methods

### Generation of reporter vector pSFSdist

Previously (Gisselbrecht et al., 2013) we had blunt-end cloned the 1.8-kb HindIII-SpeI fragment of pPelican (Barolo et al., 2000) (containing a nuclear-localized GFP reporter construct with a gypsy insulator element upstream of the MCS and minimal promoter) into our *Drosophila* transformation vector pWattB to create the cloning intermediate pWBG1i. Here, we amplified the near-ubiquitous enhancer ChIPCRM2078 identified in that study (dm3 coordinates chr3R:7177448-7178447) from OreR genomic DNA with the primers GGGGAATTCATTTTTTGCATGTCCTGCCG and GGGGGTACCGCCGATGACTCAGTGGTTAAG, cloned this into the EcoRI and KpnI sites of pWBG1i, and Gateway-converted the resulting pWBG1i-2078 plasmid by blunt-end cloning the Gateway Reading Frame A cassette into the SphI site (distal to ChIPCRM2078, relative to the Hsp70 promoter driving GFP expression) to create pSFSdist (Supplementary Fig. 5).

### Design of the candidate silencer libraries

Nine categories of elements were included in the candidate silencer libraries:

1. Nonmesodermal enhancers — All annotated *cis*-regulatory modules (CRMs) were downloaded from the REDfly database (Gallo et al., 2011) on January 17, 2014. These were filtered for length  $\leq 1,100$  bp, expression shown in a tissue (*i.e.*, not assayed only in cell culture), lack of mesodermal CRM activity terms, and association with genes that show either no mesodermal expression or sharply restricted mesodermal expression at embryonic stage 11 (when silencing would be assayed). Three additional elements with names containing "NEE" or with the expression term "neurectoderm" were manually added to this set. We removed CRMs entirely contained within other CRMs in our set, and combined overlapping CRMs where this was possible without exceeding 1,100 bp.
2. Restricted mesodermal enhancers — We downloaded all CRMs from REDfly with expression terms "muscle founder cell," "somatic muscle," or "cardioblast," filtered them for length, assessed each CRM and associated gene for restricted expression (on the theory that CRMs associated with genes with widespread mesodermal expression could not have widespread mesodermal silencing activity), and then collapsed redundant and overlapping CRMs as above. We downloaded the CAD2 database (Bonn et al., 2012) and removed anything with source term "REDfly" (as redundant), anything with expression terms M (mesoderm) or S (somatic mesoderm) at stages 9–12 (as unlikely to show widespread mesodermal silencing), and anything with no evidence of expression. For gene-assigned CRMs, we removed anything assigned to a gene with widespread mesodermal expression at stages 10-12. For unassigned CRMs, we associated each window with nearby genes, where "nearby" is defined as any gene overlapping the window, proximal to an intergenic portion of the window, or overlapping a gene which matches one of the first two criteria; we then removed CRMs associated with genes with widespread or ubiquitous expression, or where the only associated genes had no evidence for an expression pattern.
3. Groucho ChIP-chip windows — We downloaded two modENCODE Groucho ChIP-chip datasets (modENCODE\_597 and modENCODE\_623) as binding site .csv files. We filtered the smaller dataset for windows of sequence which overlapped windows in the larger dataset by  $>100$  bp, then length-filtered the resulting common set. We associated each window with

nearby genes as above, then downloaded polypeptide and transcript expression terms from FlyBase for the complete list of 520 genes associated with any GRO window by these metrics. We removed any sequence window associated with a gene that has no associated expression terms (to minimize the chance of including genes with unannotated widespread mesodermal expression) or expression terms containing "ubiquitous," "mesoderm," or "muscle."

4. Positive controls — We included three regions previously shown to have dominant silencing activity in the *Drosophila* embryonic mesoderm: the *zen* VRE (Jiang et al., 1993) and *ind* modules A and BC (Stathopoulos and Levine, 2005).
5. Negative controls — We included two classes of negative control sequence in our library: CRMs associated with genes with widespread mesodermal expression at stage 11, and ~1-kb regions of *E. coli* genomic DNA. For the former, we filtered the CAD2 database (Bonn et al., 2012) for CRMs ( $\leq 1,100$  bp) with well-documented widespread mesodermal expression at st.11 (expression terms M; S,V,C [meaning somatic, visceral, and cardiac mesoderm]; or S,V, in which case we examined the referenced studies to determine if expression appears widespread) and selected additional elements with widespread mesodermal expression from our own prior studies. For the latter, we randomly selected regions of the *E. coli* genome between 900 and 1,100 bp with G+C content similar to *Drosophila* noncoding sequence (between 39% and 43% G+C).
6. "Bivalent" chromatin — We downloaded mapped BiTS-ChIP data (Bonn et al., 2012) showing sequencing reads from isolated mesodermal chromatin immunoprecipitated with total histone H3, H3K4me1, H3K27me3, and H3K27ac. We used MACS (Zhang et al., 2008) separately on each replicate with the --nomodel parameter to identify extended peaks of histone modification enrichment relative to total H3. Using bedtools, we intersected replicates to create high-confidence peak sets, then found the intersection of H3K4me1, H3K27me3, and H3K27ac. After filtering for length  $\geq 200$  and  $\leq 1,100$  bp, we associated windows with nearby genes as above and removed those that have an associated gene with no expression terms or with an expression term including "ubiquitous." We manually assessed expression patterns of genes associated with the remaining sequence windows and removed those with ubiquitous or widespread mesodermal expression.
7. DNase I Hypersensitive Sites (DHSs) with repressive marks — We downloaded DNase Accessibility Regions for st.11 (bdtnpDnaseAccS11) from UCSC Table Browser (in dm3 coords, Apr. 2006) (Rosenbloom et al., 2015; Thomas et al., 2011) and intersected them with BiTS-ChIP H3K27me3 enriched peaks defined above. We filtered for length and expression of associated genes as above, adding a requirement for an expression term including "embryonic." As there remained an unmanageably large number of candidate sequences, we used three criteria to prioritize. We counted candidate sequences associated with each gene associated with any candidate sequence, and chose those that represent unique hits for genes of potential interest. These were highly enriched for intragenic windows, so we also prioritized intergenic windows from the broader list. Finally, we included windows of sequence overlapping CtBP ChIP-chip peaks (see below).
8. CtBP ChIP-chip windows — We downloaded the modENCODE dCtBP ChIP-chip dataset (modENCODE\_607) as a binding site .csv file and filtered for length and expression of associated genes as above. We again identified sequence windows representing unique hits to genes of potential interest; these were also highly enriched for intragenic windows, so we chose all intergenic windows on the filtered list for inclusion, then sorted the unique hits by occupancy score and included the highest-scoring ones.

9. Insulator elements — By changing the configuration of a reporter plasmid (*i.e.*, moving the tested element proximal or distal to the driving enhancer, relative to the promoter), it is possible to distinguish silencer activity from enhancer-blocking insulator activity (Petrykowska et al., 2008). Therefore, for forward-compatibility of our experiments, we included a set of candidate insulator elements in our library. We therefore downloaded BEAF32 ChIP-chip data (Jiang et al., 2009) and modENCODE\_21] and intersected all five datasets to identify the highest-confidence peaks. We similarly downloaded and intersected CTCF ChIP-chip data (modENCODE\_769 and modENCODE\_770), then intersected both of these high-confidence peak sets with each other and with CP190 peaks (modENCODE\_22) to generate a list of "Class I insulators" as defined by Nègre *et al.* (Negre et al., 2010). We then filtered for length (as above) and for non-overlap with H3K4me3 peaks in the BiTS-ChIP data. We also selected 6 insulator elements curated from the literature by Nègre *et al.* for inclusion (see Supplementary Table 1 and 2 for all tested elements).

To ease the identification of tested elements, we appended a 12-nucleotide barcode to each. These were designed by selecting 12-base sequences that each differ from all others by at least three mismatches, then filtering against a large collection of metazoan TF protein-binding microarray data (Hume et al., 2015) (Mariani *et al.*, submitted) for no 8mers with E-scores > 0.35 (Berger et al., 2006) and against a library of *Drosophila* TF PWMs for no PWM scores > 0.8 (Lenhard and Wasserman, 2002). Barcodes passing all of these filters and not containing BmtI restriction enzyme sites were randomly assigned to library elements.

PCR primer design was with MacVector 11.1.2 (MacVector, Apex, NC) for most library elements, starting from the default parameters and then loosening them until a pair was found. Pairs are prioritized by primer quality (GC content, low repetitive content, pair similarity) and by position (attempting to center the target window within the amplified region, except in cases of densely packed or overlapping targets). For dCtBP ChIP peaks, a Primer3-based computational approach was used initially, again starting with very strict parameters and progressively loosening them until a pair was found. This was run on all 92 dCtBP-derived windows, and succeeded on 78; the remaining 14 were designed with MacVector (MacVector, Apex, NC) as above. Forward primers were prepended with the common SEQ1 primer (Gisselbrecht et al., 2013) followed by the barcode for the corresponding window and a BmtI site; reverse primers were prepended with the common SEQ2 primer and the corresponding barcode. The entire library was then amplified in a two-step PCR amplification process and cloned into pDONR and then into pSFSdist, as in (Gisselbrecht et al., 2013).

## Performing silencer-FACS-Seq experiments

The resulting library of pSFSdist reporter constructs was injected into *y w nos:phiC31int; attP40* embryos by Rainbow Transgenic Flies, Inc. (Camarillo, CA). Transgenic male progeny of injected flies were recovered and crossed to *twi:CD2* virgin females to generate populations of informative embryos, exactly as previously described (Gisselbrecht et al., 2013). Since the transgenic males carry the library in their germline at only one construct per haploid genome, half of the embryos resulting from crossing these males with females homozygous for *twi:CD2* are expected to lack GFP.

We previously described a method for isolation of single cells from *Drosophila* embryos, at stage 11, that we modified by including an additional incubation step for staining the cells with commercially available Alexa647-conjugated anti(rat CD2) antibody (AbDSerotec, cat. #MCA154A647) (Gisselbrecht et al., 2013). Briefly, we used the same technique in which the cells are stained on ice with a solution composed of 1:400 dilution of the antibody in Schneider medium +8% FBS and 2  $\mu$ g/mL DAPI. The samples are then washed, filtered with Nytex mesh and the cells processed by FACS. We used the same standard gates as in our previously described method (Gisselbrecht et al., 2013) to isolate viable single cells. The P1 gate (side scatter [SSC-A] vs. forward scatter [FSC-A]) selects cells over debris and yolk granules, the P2 gate (forward scatter amplitude [FSC-A] vs. height [FSC-H]) selects single cells, while the P3 gate (DAPI signal vs. forward scatter plot [FSC-A]) selects live cells. Selection of CD2<sup>+</sup> cells is achieved as previously described (far red [APC-A] signal vs. forward scatter [FSC-A]). In addition to using the preexisting "CD2<sup>+</sup>GFP<sup>+</sup>" gate, we designed two other yellow [PE-A] vs. green [FITC-A] fluorescence gates for the capture of mesodermal cells in which GFP expression is either completely repressed ("CD2<sup>+</sup>GFP<sup>-</sup>") (yellow [PE-A] vs. green [FITC-A] fluorescence) or reduced ("CD2<sup>+</sup>GFP<sup>reduced</sup>"), to distinguish GFP-lacking cells from those in which GFP expression is reduced. As the bulk of GFP<sup>-</sup> cells carrying no GFP transgene will not have the vector sequences used for PCR recovery of library elements (see below), we sorted both low-GFP populations together to capture all cells in which GFP expression is silenced completely or partially.

However, for the second set of experiment implying our second library, we sorted only the cells bearing a reduced GFP expression ("CD2<sup>+</sup>GFP<sup>reduced</sup>"), omitting the cells without any GFP levels, as we hypothesized that the relative higher number of GFP<sup>-</sup> cells in comparison to the GFP<sup>reduced</sup> might induce higher background noise. Indeed we assumed that a significant fraction of the GFP<sup>-</sup> cells were not carrying any of the elements of the library and that, therefore, processing them with the GFP<sup>reduced</sup> cells would decrease the sensitivity of our method. Therefore, we sorted and processed the GFP<sup>reduced</sup> cell population, as defined in the paragraph above.

Figure 1 shows GFP expression profiling (green vs. yellow fluorescence) for 11,031 CD2<sup>+</sup> cells from negative control embryos (c), 13,501 cells from positive control embryos (d), and 16,654 cells from embryos containing the candidate silencer library (e). Figure 2 shows such results for 12,878 CD2<sup>+</sup> cells from validated negative silencer embryos (b) and 17,891 CD2<sup>+</sup> cells from validated positive silencer embryos (c).

Supplementary Fig. 6 shows FACS output for cells isolated from embryos transgenic for a negative control library element in pSFSdist ("MB158" 1-kb *E. coli* genomic sequence), for the absence of inhibition of GFP expression in *twi:CD2*<sup>+</sup> (mesodermal) cells; Supplementary Fig. 7 shows identical output for cells obtained from our library of candidate silencers inhibiting GFP expression in embryos that express CD2 under the *twi* promoter.

Library elements present in each analyzed population were recovered and sequenced exactly as previously described (Gisselbrecht et al., 2013). Briefly, a crude extract of sorted cell genomic DNA serves as template for nested PCR amplification, including 17 cycles with outside primers derived from vector sequence followed by 28 cycles with the SEQ1 and SEQ2 primers present on all library elements. Size-selected PCR products were sonicated and prepared for Illumina sequencing by standard protocols. All finished sequencing libraries were assessed by Agilent 2200 TapeStation and submitted to the Partners Center for Personalized Genetic Medicine for concentration measurement by PicoGreen fluorescence and qPCR, followed by equimolar index pooling and sequencing (50-base paired-end reads) on the Illumina HiSeq 2000.

### Statistical analysis of sFS sequencing reads.

Illumina sequencing reads were filtered by pattern matching (in Perl) for beginning with the SEQ1 or SEQ2 primer sequences, representing reads from one end of a PCR-amplified library element. 15.29% of reads (averaged across all libraries) passed this filter. The next 12 nucleotides of each of these reads were extracted and compared to the list of library barcodes; 98.25% matched. Counts for each library element were pooled from both (paired-end) reads of each library to achieve the final measure of abundance ("insert count") for that element in that library. The second library we designed showed similar numbers: 17.90% of the reads passed the first filter (averaged across all libraries) and 97.86% of these matched our list of library barcodes.

In our previous work (Gisselbrecht et al., 2013) and using the data therein we tested various analysis methods and found that the original DESeq R package (Anders and Huber, 2010) best predicted the results of individual validation of tested windows. We therefore used that package to compare insert counts for inserts recovered from CD2<sup>+</sup> cells in which GFP is reduced or absent to those from "input" cells (sorted CD2<sup>+</sup> or CD2<sup>-</sup> cells without regard for GFP expression status). Each week's sorting was treated as a separate experiment; the CD2<sup>+</sup>GFP<sup>reduced</sup> samples from three days of sorting were treated as biological replicates and compared to six input samples (CD2<sup>+</sup> and CD2<sup>-</sup> each from three days of sorting). As extremely low-abundance regions can give anomalously high enrichment/depletion signals, we filtered for "reliably detected" windows by including only those detected in at least one input sample from every day of sorting. A sample of the enrichment analysis result is shown in Fig. 2a. As a control, we compared recovered insert counts from CD2<sup>+</sup> and CD2<sup>-</sup> cells. As these are sorted from the same population of embryos without regard to reporter activity, they should show no difference except due to experimental noise; Supplementary Fig. 8 shows an example of the distribution of values seen in such a comparison for the first library.

As a further test of the reliability of this method, we compared for both experiments the results of two independent weeks of sorting, for the subset of library elements reliably detected in both weeks, by displaying the results of one analysis colored by the results of the independent experiment (Supplementary Figures 2 and 3). This shows that, while significant depletion calls are highly variable between experiments, significantly enriched library elements are highly concordant. We therefore considered any library element to score positive by sFS if it was significantly enriched (adjusted p-value < 0.1) in the CD2<sup>+</sup>GFP<sup>reduced</sup> cell population in either or both of the independent weeks of experiments.

## Validation of sFS results

We recovered a random sample of library element transgenic fly strains for initial validation by crossing individual transgenic male flies, removed from the population cages used to collect embryos after the end of cell sorting experiments, to virgin females of the second chromosome balancer line *y w; dpp<sup>14</sup> Bl / CyO*. After several days, transgenic males were removed and their inserts recovered and identified by PCR and Sanger sequencing; potentially informative lines were recovered and made homozygous by collecting balanced transgene insertions and self-crossing. We selected 20 of these lines to cover a range of possible outcomes from the described analysis: significantly enriched in the CD2<sup>+</sup>GFP<sup>reduced</sup> population, significantly depleted, or neither. We prepared population cages as for library sorting, using *twi:CD2* virgin females and males of one informative homozygous line for each cage; we also prepared cages in parallel with positive and negative control silencers. (Positive control: *zen* VRE (Jiang et al., 1993); negative control: Ecoli\_control15, 1 kb of *E. coli* genomic DNA.) We then prepared CD2-stained cells as above, and performed analytical flow cytometry using the same equipment we used for preparative FACS. Our key readout of silencer activity was the fraction of CD2<sup>+</sup> cells that fell within a GFP<sup>reduced</sup> gate designed to exclude the majority of both GFP<sup>+</sup> (unsilenced) and GFP<sup>-</sup> cells (non-expressing or non-transgenic cells, as from the rare non-virgin *twi:CD2* female). We measured this fraction for at least two collections of each genotype, typically counting 10<sup>5</sup> events (>10<sup>4</sup> viable cells). A library element was considered validated positive if the range of GFP<sup>reduced</sup> fractions did not overlap that observed for the negative control. Out of the 20 randomly recovered windows tested, 9 scored positive in one or both of the two sFS experiments. Of the 11 sFS negatives, all 11 were negative on individual validation. Nonetheless, only 5/9 positives validated positive by FACS (see Supplementary Table 5).

We therefore decided to individually validate all sFS-positive elements to assemble a high-confidence set of validated mesodermal silencers. However, in our initial exploration of the sFS-positive library elements, we noticed that several sequences included as negative controls scored positive, that these largely overlapped the transcriptional start sites (TSSs) of mesodermally expressed genes, and that there was overall a large and significant enrichment for TSS overlap in the set of sequences scoring positive.

We suspect that this reflects promoter competition (Ohtsuki et al., 1998), an unavoidable artifact of this experimental design. We therefore filtered the 79 sFS-positive windows to remove those likely to contain core promoter elements. Briefly, we compiled a set of TSS positions by extracting them from several transcript annotation files downloaded from FlyBase (Attrill et al., 2016) version 5.57: all start positions from the all-transcript, all-miscRNA, and all-ncRNA files, plus pre\_miRNA start positions from all-miRNA. We assembled coordinates comprising a region of +/- 40 nucleotides around each TSS, and removed those library elements that overlap any of these regions by 10 or more nucleotides. This left 38 sFS-positive non-promoter sequences to validate, 6 of which had already been tested as randomly recovered lines. The second library, however, was filtered prior to the experiment to remove elements overlapping TSSs.

For each of the remaining library elements to validate, PCR product from the original library preparation multiwell plates was purified by agarose gel electrophoresis or AMPure bead purification, BP-cloned into pDONR, sequence-verified, and LR-cloned into pSFSdist. Each resulting plasmid was injected into *y w nos:phiC31int; attP40* embryos (Gisselbrecht et al., 2013), and *white<sup>+</sup>* heterozygous progeny were recovered and crossed to *twi:CD2* virgin females for production of cells for FACS validation as above. See Supplementary Table 5 and 6 for results of all validation FACS experiments.

For the second library however, we tested a total of 8 recovered elements, 4 scored positive in one or both of the two experiments and 4 scored negative: all 8 elements validated by FACS. We therefore assumed that sorting only CD2<sup>+</sup> cells with reduced GFP expression significantly increase the reliability of our method, as stated previously, and assumed that all other elements validated as well.

To verify that an expanded CD2<sup>+</sup>GFP<sup>reduced</sup> cell population indeed represents mesodermal silencing activity of the tested library element, we generated variants of pSFSdist in which the ubiquitously active ChIPCRM2078 is replaced with one of two more specific mesodermal enhancers: ChIPCRM2613 (which drives widespread mesodermal expression (Gisselbrecht et al., 2013)), or *Mef2* I-ED<sub>5</sub>, which drives expression specifically in the fusion competent myoblasts of the developing mesoderm (Duan et al., 2001). Elements to be tested (one each of a positive control, a negative control, and a newly discovered silencer) were Gateway LR-cloned into these vectors and introduced into flies as above; transgenic embryos from homozygous lines were collected, fixed, and stained for GFP expression as previously described (Gisselbrecht et al., 2013). For imaging of fluorescence intensity, embryos were collected, fixed, and stained in parallel, imaged with identical exposure times, and processed for presentation without adjustment of brightness, contrast, or gamma. Distribution of pixel values for the indicated regions (magenta boxes, Fig. 2d-g) were measured in Photoshop and are expressed as Mean±SD.

## Assessing CRM bifunctionality

A large majority of the high-confidence validated silencers (~85%) reported herein were originally included in our libraries of elements to test on the basis of previously characterized enhancer activity (see Fig. 5). From the first set of validated silencers from the first library, the two silencers not previously known to be enhancers were assayed by sFS based on their containing CHIP-chip peaks for the corepressor CtBP. To test the potential enhancer activity of these elements, they were LR-cloned into our pEFS vector (Gisselbrecht et al., 2013). Homozygous lines containing these reporter constructs were generated as above, and embryos were fixed, stained, and imaged as previously described (Gisselbrecht et al., 2013). From the second library, the 5 validated silencers which were not known to be enhancers, were cloned into our pEFS vector as previously stated and these plasmids were sent for microinjection into embryos to Rainbow Transgenic Flies, Inc. (Camarillo, CA). At the time of the writing of this dissertation, the injected flies were pupae and in quarantine, and no pattern of expression for these elements has yet been observed.

## Downstream analysis of the validated silencers

1. Enrichment of input data types. Each tested element belonged to one of nine categories, as described above ("Design of the candidate silencer library"). We compared the prevalence of each category among high-confidence validated silencers to its prevalence among non-TSS-overlapping windows confidently detected in either or both of the two experimental repetitions. Statistical significance of enrichment or depletion was calculated using the fisher.test function in R.

2. Enrichment of histone marks and TF CHIP signal. Mesoderm-specific histone modification CHIP-seq datasets were downloaded from the European Nucleotide Archive (Bonn et al., 2012) or from GEO (Gaertner et al., 2012). Reads mapping to each tested library element (*i.e.*, each non-TSS-overlapping element confidently detected in either or both of the two experiments, each comprising three biological replicates) were counted and, where available,

normalized by dividing by total H3 ChIP read count. Whole embryo histone modification ChIP-seq and ChIP-chip datasets were downloaded from modENCODE as bedfiles. ChIP-chip or ChIP-seq data for individual TFs, coactivators, and corepressors were assembled from modENCODE and other sources (see Supplementary Table 9 and 10). Mean signal over all tested library elements was calculated using bedtools. Enrichment or depletion was measured by calculating the area under the receiver-operator characteristic curve (AUROC), considering high-confidence validated silencers to be "true positives," using the `auROC` function of the `limma` package in R. Statistical significance was assessed using the `wilcox.test` function. Where independent replicates were available, p-values were calculated separately and combined using Fisher's method (Mosteller and Fisher, 1948). P-values were corrected for multiple hypothesis testing using the `p.adjust` function in R with the "fdr" method. TF complexity scores and HOT regions were downloaded from (Roy et al., 2010). Using the previously defined TF complexity score of 8.0 as the cutoff to define a HOT region, tested elements that overlap HOT regions were defined and the enrichment of high-confidence validated silencers in this population was calculated with the `fisher.test` function in R. To generate the heatmap, all histone mark data were Z-transformed, subtracting from each element in each column (i.e. each histone mark dataset) the mean of that column over all library elements and dividing that deviation by the standard deviation over the column; the  $< 0.5\%$  of all Z-scores over 5 were truncated to 5. Truncated Z-scores were biclustered, using 1-Pearson's R as a distance metric and Ward's minimum variance method for clustering.

3. Motif enrichment. We curated a list of 93 repressive TF binding site motifs (see Supplementary Table 4). Gene lists were downloaded from FlyBase (download date: February 3, 2015) with the Molecular Function Gene Ontology term GO:0043565 (sequence-specific DNA binding) and either the Biological Process term GO:0000122 (negative regulation of transcription from RNA polymerase II promoter) or the Biological Process term GO:0045892 (negative regulation of transcription, DNA-templated). These were combined and intersected with the list of *Drosophila* TFs with experimentally determined DNA binding site motifs from CisBP (Weirauch et al., 2014), UniPROBE (Hume et al., 2015), and FlyFactorSurvey (Zhu et al., 2011). For TFs with multiple similar PWMs available, a single representative (learned from ChIP data, where available) was chosen; where a single TF (or its isoforms or heterodimers) gave two unalignable motifs, both were included. We then used the Lever algorithm (Warner et al., 2008) to search for combinations of 1, 2, or 3 motifs enriched among high-confidence validated silencers relative to matched random genomic background sets, as previously described (Gisselbrecht et al., 2013). We consider a motif or motif combination significantly enriched if it targets  $\geq 50\%$  of the foreground sequences and has  $AUROC \geq 0.65$  and  $FDR \leq 0.1$ ; all such combinations are shown in Fig. 4, for the validated elements from the first library we tested.

4. Testing effects of binding site mutations on silencer/enhancer activity. High-quality instances of the *dve* motif were chosen for site-directed mutagenesis in pDONR clones of example high-confidence validated silencers. Mutations were designed to avoid altering or introducing overlapping binding sites for known or suspected regulators; introduced mutations are shown in Supplementary Fig. 4. To assess effects on silencer activity, sequence-validated mutant silencers were LR-cloned into pSFSdist and introduced into flies; silencer activity of the resulting constructs was tested in parallel with their wild type counterparts as described above. To assess effects on enhancer activity, mutant silencers were LR-cloned into pEFS and introduced into flies. Wild type versions of the same elements were LR-cloned into pWattB-*nlacZ* (Busser et al., 2012) and introduced into flies; crossing of the resulting CRM:*LacZ* and CRM<sup>mut</sup>:GFP lines together produces embryos in which the wild type- and mutant-driven expression patterns can be compared directly, which were fixed, stained, and imaged as above.

## Cell sorting and fixation with formaldehyde

We once again used the previously described protocol for isolation of single cells from *Drosophila* embryos, at stage 11, that we modified by including an additional incubation step for staining the cells with commercially available Alexa647-conjugated anti(rat CD2) antibody (AbDSerotec, cat. #MCA154A647) (Gisselbrecht et al., 2013) for the homozygous *twi:CD2* line.

We used the same technique as previously explained in this Methods section and sorted the cells accordingly. Then, we proceeded with the crosslinking of the cells by adding a final concentration of 1% of formaldehyde, incubating 10 minutes at room temperature and gently inverting the tube every minute. The reaction was quenched by adding an excess of glycine (final concentration ~125mM) and incubated for 5 minutes at room temperature, then 15 min on ice to stop all remaining crosslinking. The fixed cells were then washed and resuspended into culture medium (8% FBS in Schneider Medium), and stored at 4°C, protected from light.

## Guide-RNA and primer design

We used the online tool “flyCRISPR” (Gratz et al., 2014) to find CRISPR target sites with the highest stringency settings available, to select 20-nucleotide long sites with no predicted off-target sites. The Protospacer adjacent motif (PAM) was set to be NGG only and the “maximum stringency” was selected to be the maximum available which uses a strict algorithm based on off-target cleavage effects observed in cell lines (Gratz et al., 2014). The gRNA against white (GGCGATACTTGGATGCCCTGCGG) was selected upon these criteria. For our bifunctional CRMs, the DNA sequences provided to “flyCRISPR” were selected after inspection on the UCSC genome browser: for each element, the two gRNAs necessary for a knock-out were tried to be found near the extremities of the tested windows by sFS, in approximate concordance with regions that seemed conserved in the *Drosophila* genre. If two distinct conserved regions seemed visible, a third gRNA was designed to independently knock-out each region. The list of gRNA designed can be found in annex 3, along with primer sets surrounding the window, for quick PCR validation of knock-outs (by obvious change in the size of the PCR product).

## Cas9 and gRNA preparation and microinjection

The Cas9 protein was ordered from IDT (Alt-R S.p. Cas9 Nuclease 3NLS, #1074181) along with the necessary CRISPR RNA (crRNA) (IDT, #1072534) and transactivating (tracrRNA) (IDT #1075928) which respectively specify the DNA target sequence and activate the Cas9 endonuclease. It has to be noted that the Cas9-3NLS endonuclease is delivered in a solution at 50% glycerol. The RNP complexes were assembled according to IDT protocols and to follow the necessary 2:2:1 molar ratios of crRNA:tracrRNA:Cas9. We prepared RNP solutions to have a final concentration of 2ug/ul of Cas9 (final concentration of 10% glycerol), given that it was previously found to be the optimal concentration for microinjection in *Drosophila* (Lee et al., 2014). The microinjection of the RNP complexes was in the same manner as regular microinjection, as previously explained in this Methods section. If the injected eggs led to adult flies, males from those flies should then be crossed to virgin females from an appropriate balancer line, in order for the knock-out to be transmitted to the progeny without concern if it were to be lethal.

## In situ hybridization: probes primer design

The primers for probes for future *in situ* hybridization experiments were designed for genes surrounding the validated bifunctional CRMs we previously identified, 10kb upstream and downstream of the element. A set of primers was designed for each gene, and every set included the T7 promoter sequence on the reverse primer (TAATACGACTCACTATAGGGAGA).

## References

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
- Bessis, A., Champtiaux, N., Chatelin, L. & Changeux, J.P. (1997). The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc Natl Acad Sci U S A* 94, 5906-110
- Busturia, A., Wightman, C. D., and Sakonju, S. (1997) . A silencer is required for maintenance of transcriptional repression throughout *Drosophila* development. *Development* 124:4343-4350.
- Celniker, S., Dillon, L., Gerstein, M., Gunsalus, K., Henikoff, S., Karpen, G., Kellis, M., Lai, E., Lieb, J., MacAlpine, D., Micklem, G., Piano, F., Snyder, M., Stein, L., White, K. & Waterston, R. (2009). Unlocking the secrets of the genome. *Nature* 459, 927-930.
- Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R. & Mozziconacci, J. (2012). Normalization of a chromosomal contact map. *BMC Genomics* 13, 436.
- Doyle, H. J., Kraut, R., and Levine, M. (1989). Spatial regulation of *zerknüllt*: a dorsal-ventral patterning gene in *Drosophila*. *Genes Dev* 3:1518-1533.
- Ernst, J. & Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* 28, 817-25.
- Filion, G., van Bemmel, J., Braunschweig, U., Talhout, W., Kind, J., Ward, L., Brugman, W., de Castro, I., Kerkhoven, R., Bussemaker, H. & van Steensel, B. (2010). Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in *Drosophila* Cells. *Cell* 143, 212-224.
- Gallo, S., Gerrard, D., Miner, D., Simich, M., Des Soye, B., Bergman, C. & Halfon, M. (2010). REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Research* 39, D118-D123.
- Gisselbrecht, S., Barrera, L., Porsch, M., Aboukhalil, A., Estep, P., Vedenko, A., Palagi, A., Kim, Y., Zhu, X., Busser, B., Gamble, C., Iagovitina, A., Singhania, A., Michelson, A. & Bulyk, M. (2013). Highly parallel assays of tissue-specific enhancers in whole *Drosophila* embryos. *Nature Methods* 10, 774-780.
- Gratz, S., Ukken, F., Rubinstein, C., Thiede, G., Donohue, L., Cummings, A. & O'Connor-Giles, K. (2014). Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in *Drosophila*. *Genetics* 196, 961-971.
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8, 172.
- Hartwell S., Goldberg M., Fischer J., Hood L., Aquadro C. (2015). *Genetics: From Genes to Genomes*, 5th Edition, *McGraw-Hill*.
- Heintzman, N., Stuart, R., Hon, G., Fu, Y., Ching, C., Hawkins, R., Barrera, L., Van Calcar, S., Qu, C., Ching, K., Wang, W., Weng, Z., Green, R., Crawford, G. & Ren, B. (2007). Distinct and

- predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* 39, 311-318.
- Heintzman, N., Hon, G., Hawkins, R., Kheradpour, P., Stark, A., Harp, L., Ye, Z., Lee, L., Stuart, R., Ching, C., Ching, K., Antosiewicz-Bourget, J., Liu, H., Zhang, X., Green, R., Lobanenkov, V., Stewart, R., Thomson, J., Crawford, G., Kellis, M. & Ren, B. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108-112.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278.
- Hu, M., Deng, K., Selvaraj, S., Qin, Z., Ren, B. & Liu, J. (2012). HiCNorm: removing biases in Hi-C data via Poisson regression. *Bioinformatics* 28, 3131-3133.
- Huang JD, Dubnicoff T, Liaw GJ, Bai Y, Valentine SA, Shirokawa JM, Lengyel JA, Courey AJ. (1995) Binding sites for transcription factor NTF-1/Elf-1 contribute to the ventral repression of decapentaplegic. *Genes & development* 9.24: 3177-89
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433.
- Jiang, J., Cai, H., Zhou, Q. & Levine, M. Conversion of a dorsal-dependent silencer into an enhancer: evidence for dorsal corepressors. *EMBO J* 12, 3201-9 (1993).
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- Kallunki, P., Edelman, G.M. & Jones, F.S. (1998). The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. *Proc Natl Acad Sci U S A* 95, 3233-8.
- Kehayova, P., Monahan, K., Chen, W. & Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A* 108, 17195-200.
- Kharchenko, P., Alekseyenko, A., Schwartz, Y., Minoda, A., Riddle, N., Ernst, J., Sabo, P., Larschan, E., Gorchakov, A., Gu, T., Linder-Basso, D., Plachetka, A., Shanower, G., Tolstorukov, M., Luquette, L., Xi, R., Jung, Y., Park, R., Bishop, E., Canfield, T., Sandstrom, R., Thurman, R., MacAlpine, D., Stamatoyannopoulos, J., Kellis, M., Elgin, S., Kuroda, M., Pirrotta, V., Karpen, G. & Park, P. (2010). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485.
- Lajoie, B., Dekker, J. & Kaplan, N. (2015). The Hitchhiker’s guide to Hi-C analysis: Practical guidelines. *Methods* 72, 65-75.
- Lee, J., Kwak, S., Kim, J., Kim, A., Noh, H., Kim, J. & Yu, K. (2014). RNA-Guided Genome Editing in *Drosophila* with the Purified Cas9 Protein. *G3&#58; Genes|Genomes|Genetics* 4, 1291-1295.
- Koike, S., Schaeffer, L. & Changeux, J.P. (1995). Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene. *Proc Natl Acad Sci U S A* 92, 10624-8.

Pfeiffer, B., Jenett, A., Hammonds, A., Ngo, T., Misra, S., Murphy, C., Scully, A., Carlson, J., Wan, K., Laverty, T., Mungall, C., Svirskas, R., Kadonaga, J., Doe, C., Eisen, M., Celniker, S. & Rubin, G. (2008). Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences* 105, 9715-9720.

Maurano, M., Humbert, R., Rynes, E., Thurman, R., Haugen, E., Wang, H., Reynolds, A., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutuyavin, T., Stehling-Sun, S., Johnson, A., Canfield, T., Giste, E., Diegel, M., Bates, D., Hansen, R., Neph, S., Sabo, P., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S., Kaul, R. & Stamatoyannopoulos, J. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190-5.

Mani-Telang, P. & Arnosti, D.N. (2007). Developmental expression and phylogenetic conservation of alternatively spliced forms of the C-terminal binding protein corepressor. *Dev Genes Evol* 217, 127-35.

Nieto, M.A. (2002). The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* 3, 155-66.

Noyes, M., Christensen, R., Wakabayashi, A., Stormo, G., Brodsky, M. & Wolfe, S. (2008). Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell* 133, 1277-1289.

Ogbourne, S. & Antalis, T.M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* 331 ( Pt 1), 1-14.

Ohtsuki, S., Levine, M., and Cai, H.N. (1998). Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev* 12, 547-556.

Orian, A., Delrow, J., Rosales Nieves, A., Abed, M., Metzger, D., Paroush, Z., Eisenman, R. & Parkhurst, S. (2007). A Myc Groucho complex integrates EGF and Notch signaling to regulate neural development. *Proceedings of the National Academy of Sciences* 104, 15771-15776.

Overballe-Petersen, S., Harms, K., Orlando, L., Mayar, J., Rasmussen, S., Dahl, T., Rosing, M., Poole, A., Sicheritz-Ponten, T., Brunak, S., Inselmann, S., de Vries, J., Wackernagel, W., Pybus, O., Nielsen, R., Johnsen, P., Nielsen, K. & Willerslev, E. (2013). Bacterial natural transformation by highly fragmented and damaged DNA. *Proceedings of the National Academy of Sciences* 110, 19860-19865.

Prasad, M.S. & Paulson, A.F. (2011). A combination of enhancer/silencer modules regulates spatially restricted expression of cadherin-7 in neural epithelium. *Dev Dyn* 240, 1756-68.

Riel, J.J.G.v. (2014). Identification of epigenomic patterns to annotate regulatory elements in the human genome. *Masters thesis, Utrecht University*.

modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., et al. (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330, 1787-1797.

Schaeffer, H.J., Forsthoefel, N.R., Cushman, J.C. (1995). Identification of enhancer and silencer regions involved in salt-responsive expression of Crassulacean acid metabolism (CAM) genes in the facultative halophyte *Mesembryanthemum crystallinum*. *Plant Mol Biol* 28, 205-18.

Schneider, C., Rasband, W. & Eliceiri, K. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9, 671-675.

Simpson, J., Schell, J., Montagu, M.V., Herrera-Estrella, L. (1986). Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties. *Nature* 323, 551-554.

Stathopoulos, A. & Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev Biol* 280, 482-93.

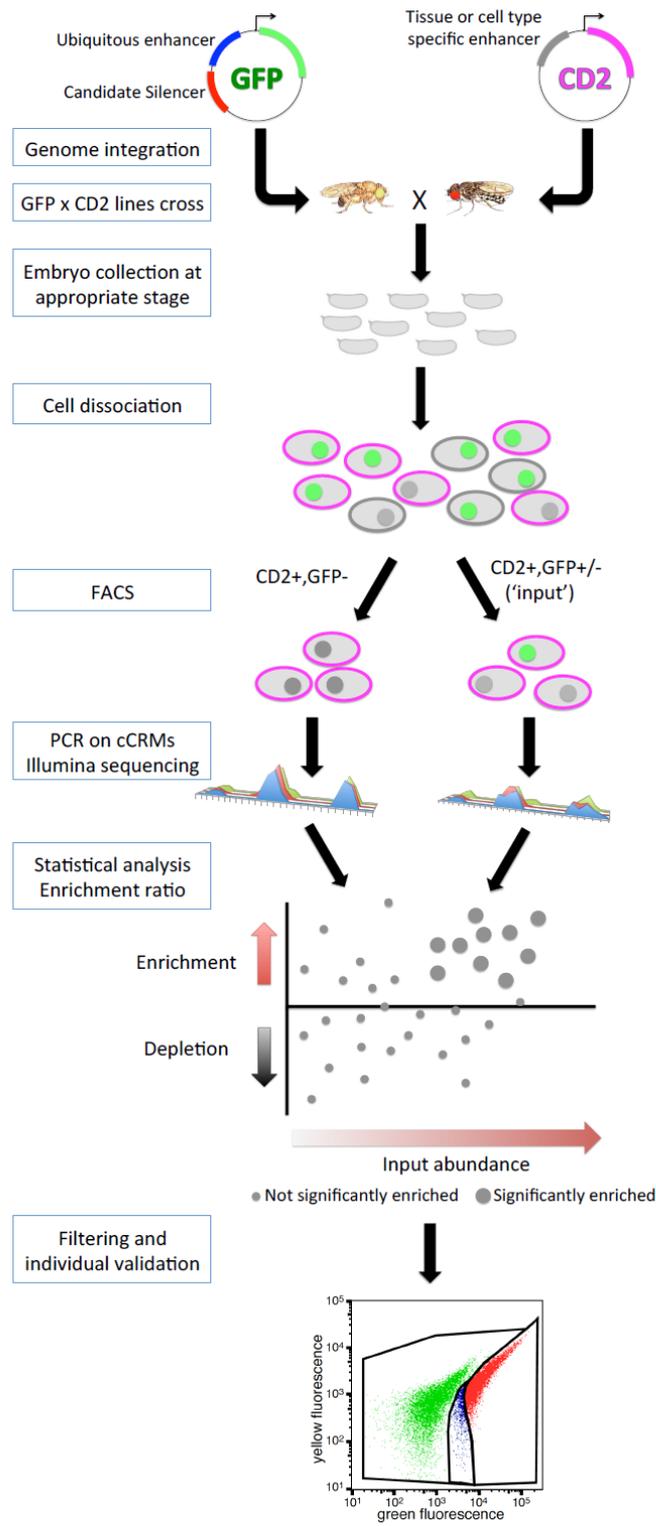
Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.

Xie, F., Ye, L., Chang, J.C., Beyer, A.I., Wang, J., Muench, M.O., and Kan, Y.W. (2014). Seamless gene correction of  $\beta$ -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* 24, 1526–1533.

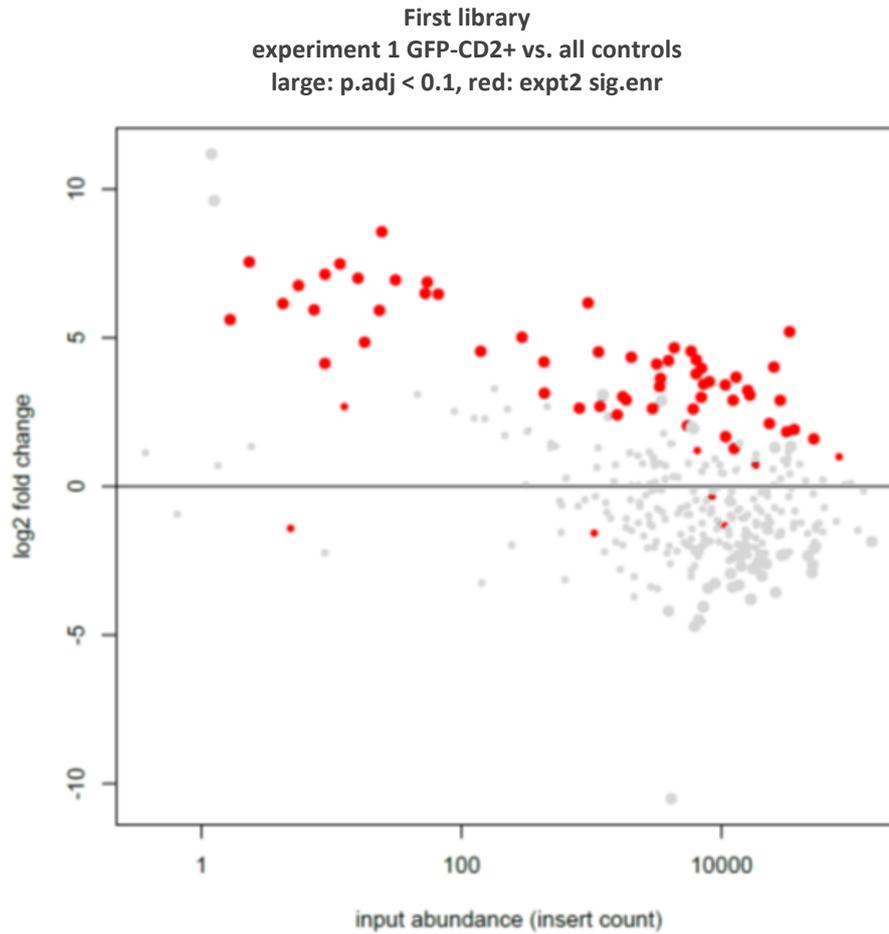
Yaffe, E. & Tanay, A. (2011). Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nature Genetics* 43, 1059-1065.

Zhu, Q. and Halfon, M. (2009). Complex organizational structure of the genome revealed by genome-wide analysis of single and alternative promoters in *Drosophila melanogaster*. *BMC Genomics*, 10(1), p.9.

# Supplementary figures



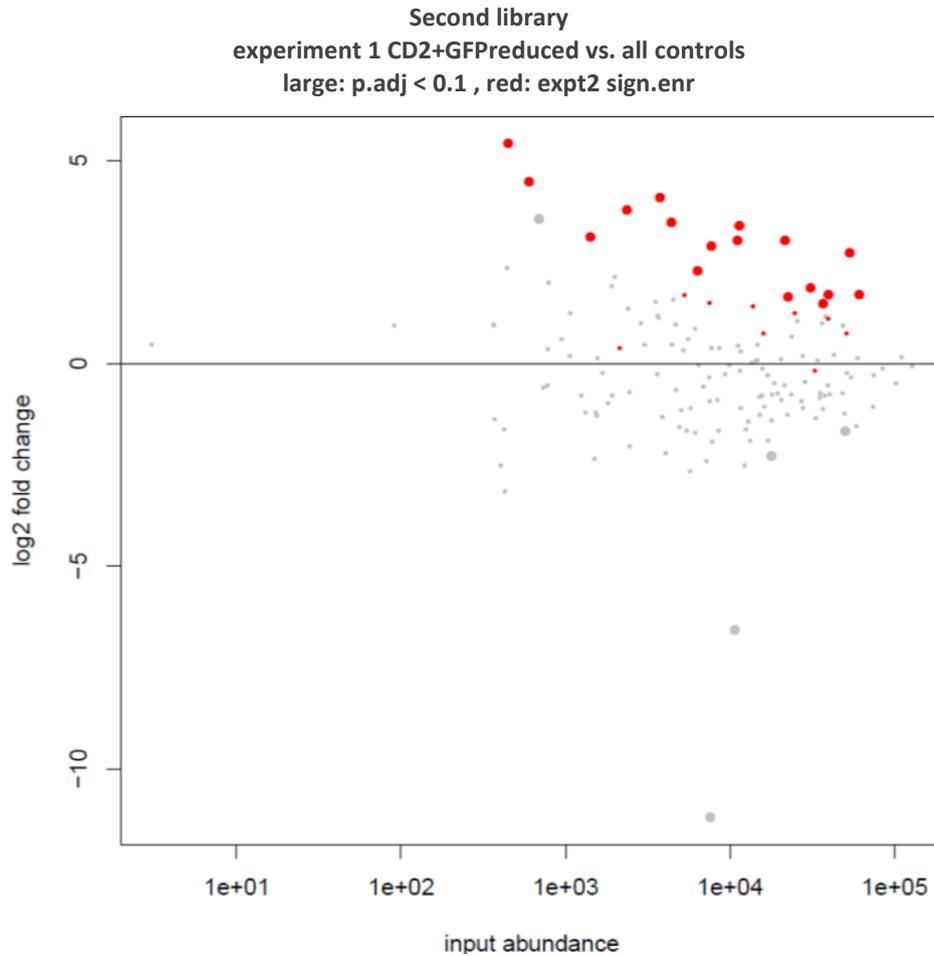
**Supplementary Fig. 1.**  
 Schema of silencer-FACS-seq study design.



**Supplementary Fig. 2.**

Concordance between sFS replicates for the first library.

Enrichment ratios for confidently detected library elements in the CD2+GFP-reduced cell population, as compared to input cells, are plotted vs. input abundance, here for our first library. Positive y-values reflect enrichment in the silencer cell population. Library elements represented by large points were called significantly differentially detected (adjusted p-value < 0.1 according to DESeq; see Methods section) in this experiment; red points were called significantly enriched in an independent repetition of the experiment. Each repetition comprised three biological replicates collected on separate days.

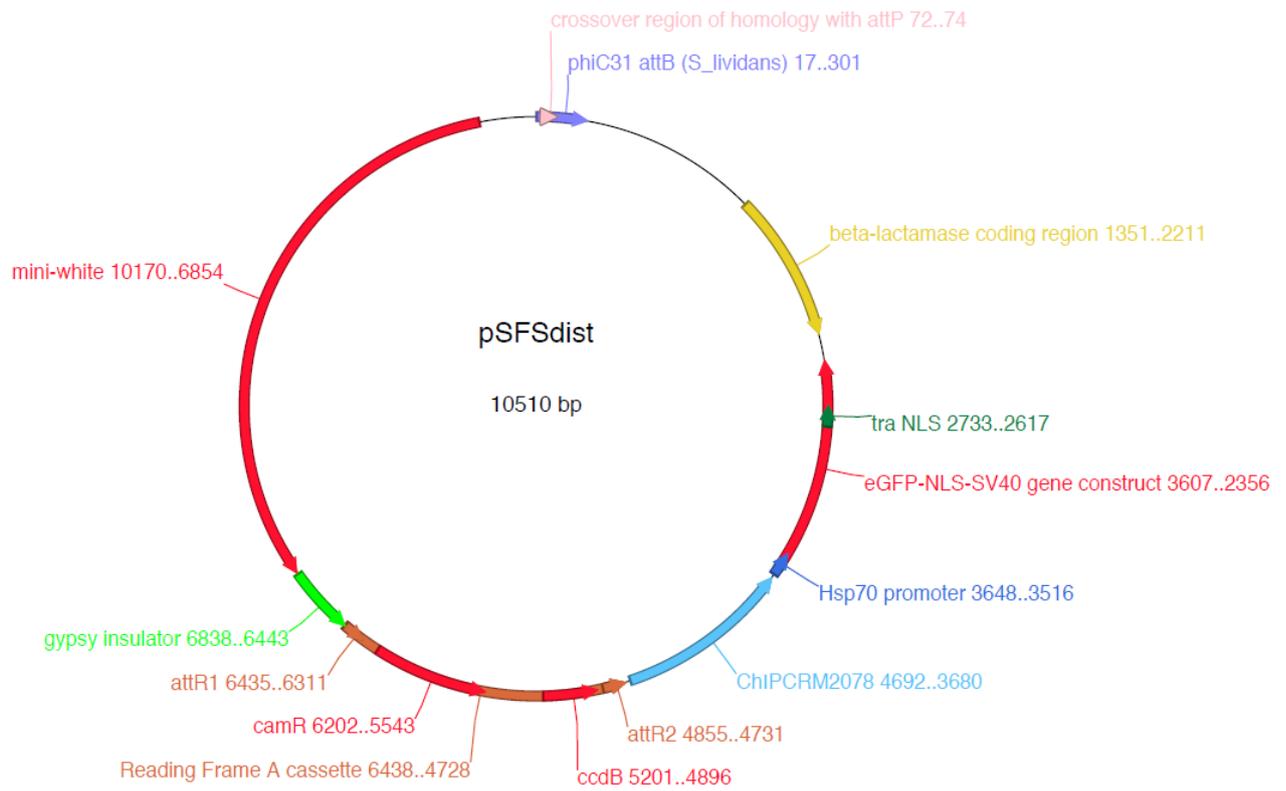


**Supplementary Fig. 3.**

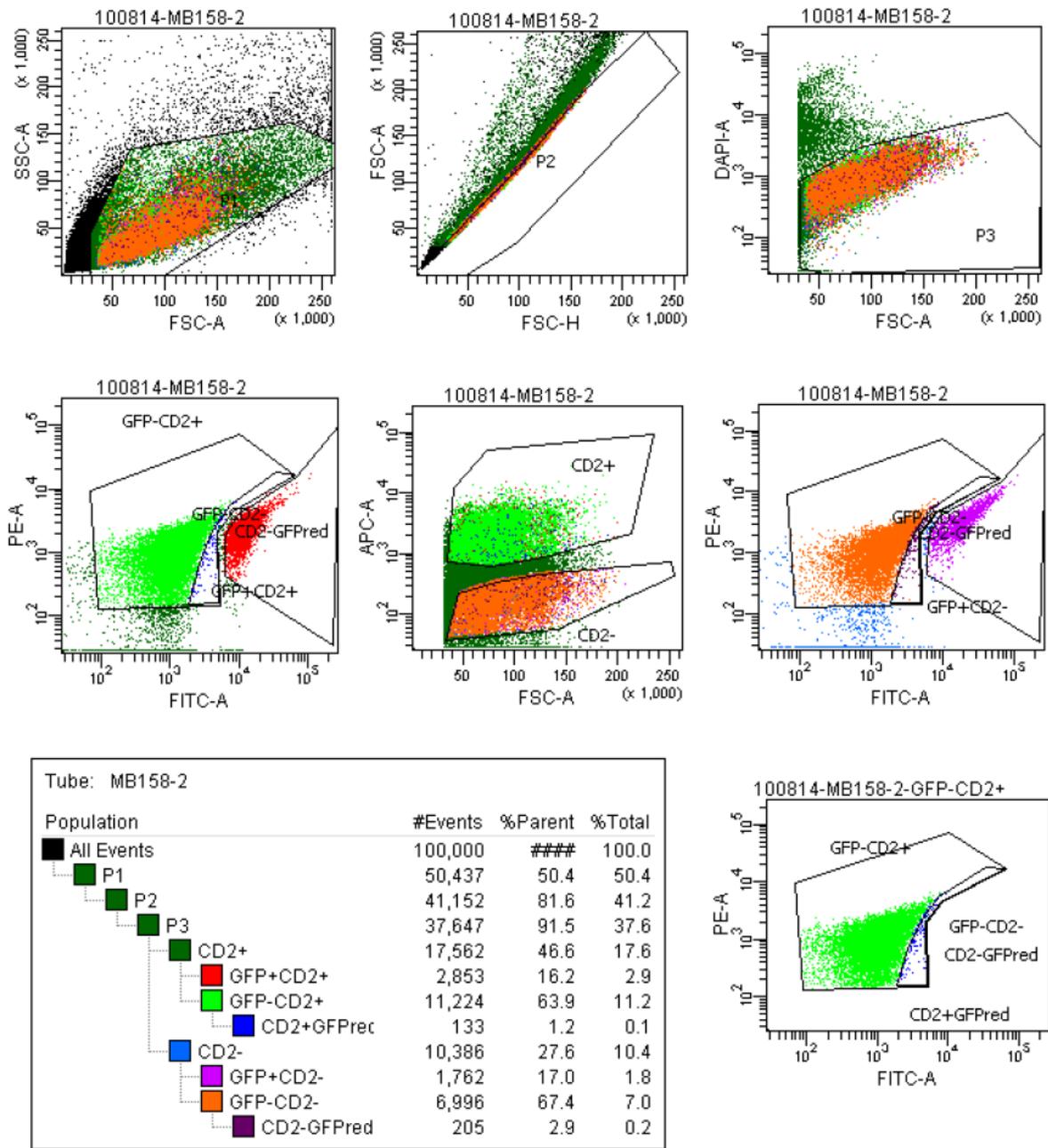
Concordance between sFS replicates for the second library.

As in Supplementary Fig. 2, enrichment ratios for confidently detected library elements in the CD2+GFP<sup>reduced</sup> cell population, as compared to input cells, are plotted vs. input abundance, here for our second library. Positive y-values reflect enrichment in the silencer cell population. Library elements represented by large points were called significantly differentially detected (adjusted p-value < 0.1 according to DESeq; see Methods section) in this experiment; red points were called significantly enriched. Once again, each repetition comprised three biological replicates collected on separate days.





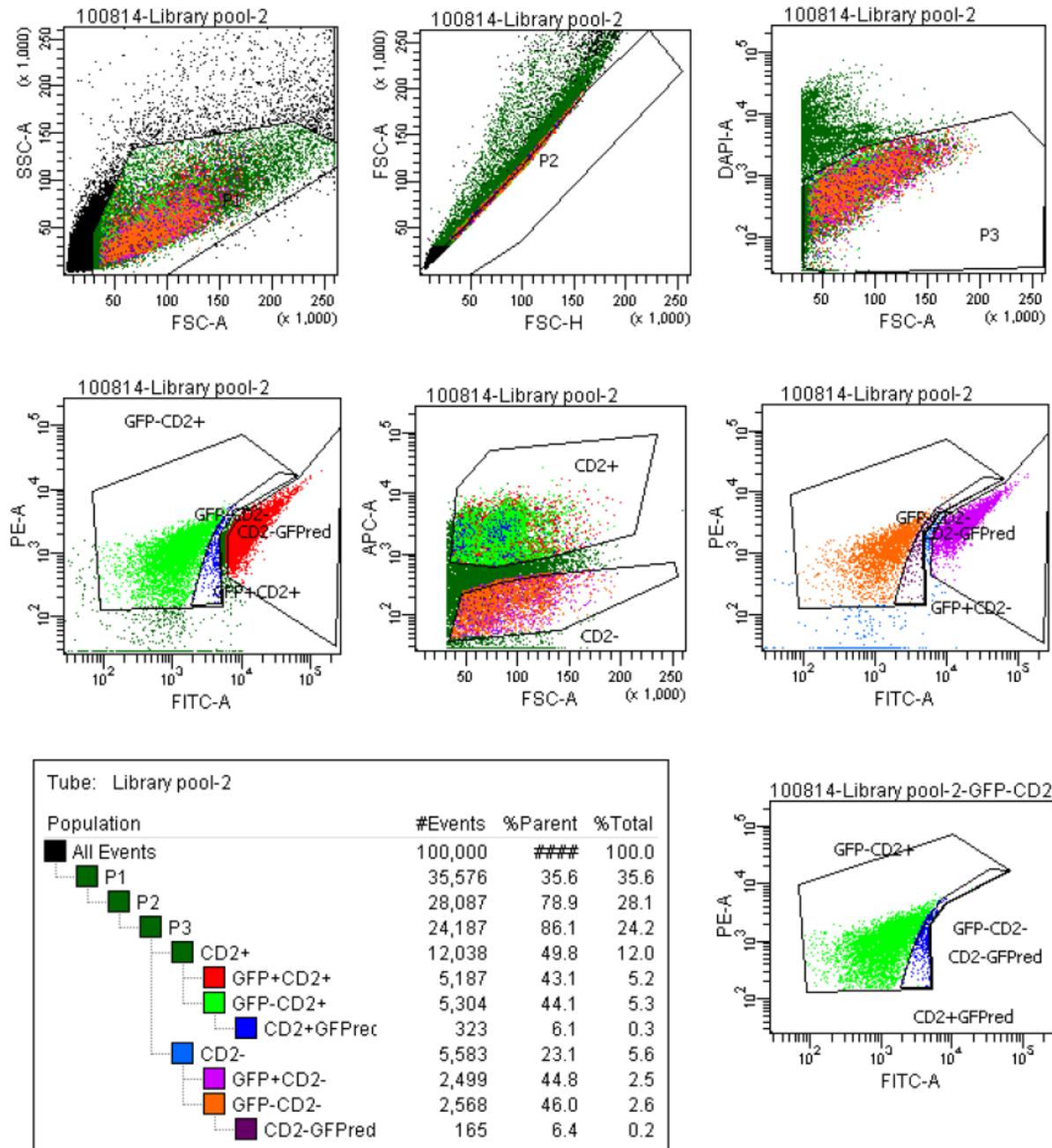
**Supplementary Fig. 5.**  
Map of the pSFSdist plasmid used in this study.



**Supplementary Fig. 6.**

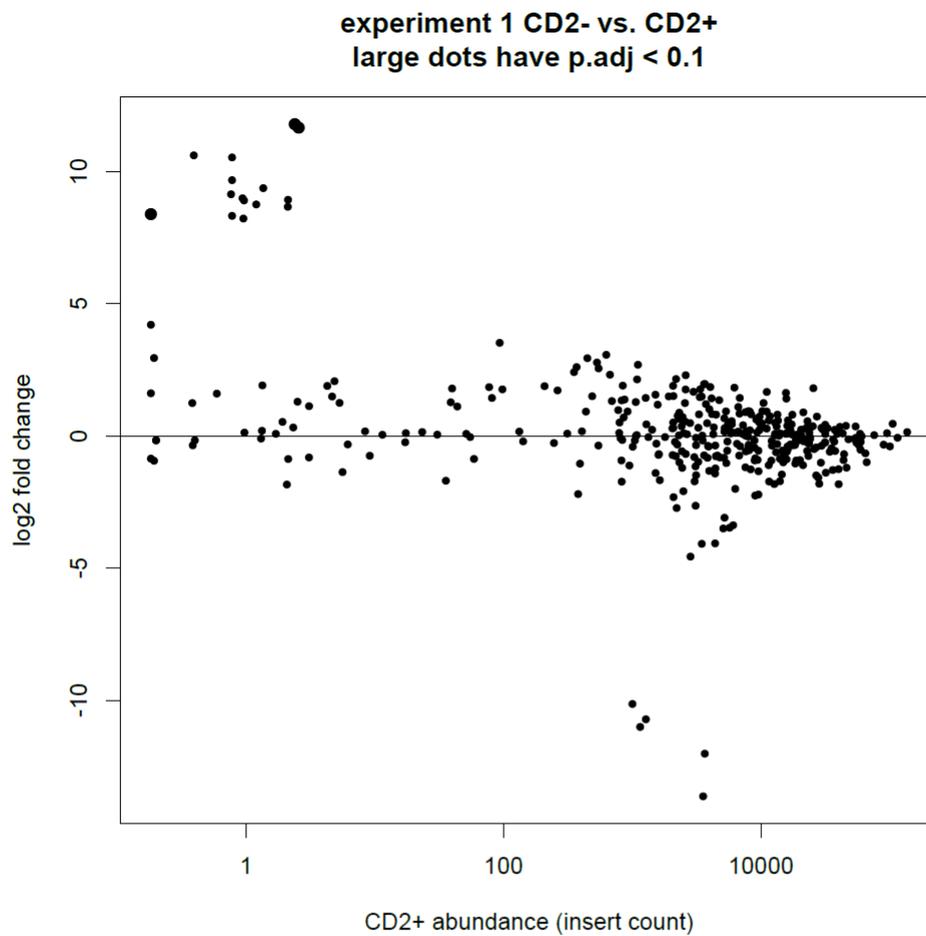
FACS output showing the distribution of GFP expression (detected as the ratio of green [FITC-A] vs yellow [PE-A] fluorescence) in CD2+ and therefore mesodermal cells isolated from embryos made transgenic for a negative control library element in our pSFSdist (“MB158” 1-kb *E.coli* genomic sequence)

FACSDiva Version 6.1.2



**Supplementary Fig. 7.**

FACS output showing the distribution of GFP expression (detected as the ratio of green [FITC-A] versus yellow [PE-A] fluorescence) in CD2+ (mesodermal) cells isolated from embryos transgenic for our library of candidate silencers inhibiting GFP expression.



**Supplementary Fig. 8. : Repeatability of control experiments.**

Enrichment ratios or confidently detected library elements in the CD2+ cell population, as compared to CD2- cells, are plotted versus input abundance. Positive y-values reflect enrichment in the CD2+ cell population. Library elements represented by large points were called significantly differentially detected (adjusted p-value < 0.1 according to DESeq; see Methods section) in this experiment.

## Supplementary References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol* *11*, R106.

Attrill, H., Falls, K., Goodman, J.L., Millburn, G.H., Antonazzo, G., Rey, A.J., Marygold, S.J., and FlyBase, C. (2016). FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*. *Nucleic Acids Res* *44*, D786-792.

Barolo, S., Carver, L.A., and Posakony, J.W. (2000). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. *Biotechniques* *29*, 726, 728, 730, 732.

Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., 3rd, and Bulyk, M.L. (2006). Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat Biotechnol* *24*, 1429-1435.

Bessis, A., Champiaux, N., Chatelin, L., and Changeux, J.P. (1997). The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc Natl Acad Sci U S A* *94*, 5906-5911.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet* *44*, 148-156.

Busser, B.W., Taher, L., Kim, Y., Tansey, T., Bloom, M.J., Ovcharenko, I., and Michelson, A.M. (2012). A machine learning approach for identifying novel cell type-specific transcriptional regulators of myogenesis. *PLoS Genet* *8*, e1002531.

Duan, H., Skeath, J.B., and Nguyen, H.T. (2001). *Drosophila* *Lame duck*, a novel member of the *Gli* superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development. *Development* *128*, 4489-4500.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* *28*, 817-825.

Filion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., *et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* *143*, 212-224.

Gaertner, B., Johnston, J., Chen, K., Wallaschek, N., Paulson, A., Garruss, A.S., Gaudenz, K., De Kumar, B., Krumlauf, R., and Zeitlinger, J. (2012). Poised RNA polymerase II changes over developmental time and prepares genes for future expression. *Cell Rep* *2*, 1670-1683.

Gallo, S.M., Gerrard, D.T., Miner, D., Simich, M., Des Soye, B., Bergman, C.M., and Halfon, M.S. (2011). REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Res* *39*, D118-123.

Gao, Q., and Finkelstein, R. (1998). Targeting gene expression to the head: the *Drosophila* orthodenticle gene is a direct target of the Bicoid morphogen. *Development* *125*, 4185-4193.

Gisselbrecht, S.S., Barrera, L.A., Porsch, M., Aboukhalil, A., Estep, P.W., 3rd, Vedenko, A., Palagi, A., Kim, Y., Zhu, X., Busser, B.W., *et al.* (2013). Highly parallel assays of tissue-specific enhancers in whole *Drosophila* embryos. *Nat Methods* *10*, 774-780.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108-112.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318.

Hume, M.A., Barrera, L.A., Gisselbrecht, S.S., and Bulyk, M.L. (2015). UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions. *Nucleic Acids Res* 43, D117-122.

Jiang, J., Cai, H., Zhou, Q., and Levine, M. (1993). Conversion of a dorsal-dependent silencer into an enhancer: evidence for dorsal corepressors. *Embo J* 12, 3201-3209.

Jiang, N., Emberly, E., Cuvier, O., and Hart, C.M. (2009). Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in *Drosophila melanogaster* links BEAF to transcription. *Mol Cell Biol* 29, 3556-3568.

Kallunki, P., Edelman, G.M., and Jones, F.S. (1998). The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. *Proc Natl Acad Sci U S A* 95, 3233-3238.

Kehayova, P., Monahan, K., Chen, W., and Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A* 108, 17195-17200.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., *et al.* (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485.

Koike, S., Schaeffer, L., and Changeux, J.P. (1995). Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene. *Proc Natl Acad Sci U S A* 92, 10624-10628.

Lenhard, B., and Wasserman, W. (2002). TFBS: Computational framework for transcription factor binding site analysis. *Bioinformatics* 18, 1135-1136.

Mani-Telang, P., and Arnosti, D.N. (2007). Developmental expression and phylogenetic conservation of alternatively spliced forms of the C-terminal binding protein corepressor. *Dev Genes Evol* 217, 127-135.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., *et al.* (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190-1195.

Mosteller, F., and Fisher, R. (1948). Questions and Answers, Answer number 14. *The American Statistician* 2, 30-31.

Negre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., *et al.* (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* 6, e1000814.

- Nieto, M.A. (2002). The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* 3, 155-166.
- Noyes, M.B., Christensen, R.G., Wakabayashi, A., Stormo, G.D., Brodsky, M.H., and Wolfe, S.A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* 133, 1277-1289.
- Ogbourne, S., and Antalis, T.M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* 331 ( Pt 1), 1-14.
- Ohtsuki, S., Levine, M., and Cai, H.N. (1998). Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev* 12, 547-556.
- Orian, A., Delrow, J.J., Rosales Nieves, A.E., Abed, M., Metzger, D., Paroush, Z., Eisenman, R.N., and Parkhurst, S.M. (2007). A Myc-Groucho complex integrates EGF and Notch signaling to regulate neural development. *Proc Natl Acad Sci U S A* 104, 15771-15776.
- Petrykowska, H.M., Vockley, C.M., and Elnitski, L. (2008). Detection and characterization of silencers and enhancer-blockers in the greater CFTR locus. *Genome Res* 18, 1238-1246.
- Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.T., Misra, S., Murphy, C., Scully, A., Carlson, J.W., Wan, K.H., Lavery, T.R., *et al.* (2008). Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc Natl Acad Sci USA* 105, 9715-9720.
- Prasad, M.S., and Paulson, A.F. (2011). A combination of enhancer/silencer modules regulates spatially restricted expression of cadherin-7 in neural epithelium. *Dev Dyn* 240, 1756-1768.
- Riel, J.J.G.v. (2014). Identification of epigenomic patterns to annotate regulatory elements in the human genome. Masters thesis, Utrecht University.
- Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., *et al.* (2015). The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res* 43, D670-681.
- Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., *et al.* (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330, 1787-1797.
- Schaeffer, H.J., Forstheofel, N.R., and Cushman, J.C. (1995). Identification of enhancer and silencer regions involved in salt-responsive expression of Crassulacean acid metabolism (CAM) genes in the facultative halophyte *Mesembryanthemum crystallinum*. *Plant Mol Biol* 28, 205-218.
- Simpson, J., Schell, J., Montagu, M.V., and Herrera-Estrella, L. (1986). Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties. *Nature* 323, 551-554.
- Stathopoulos, A., and Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev Biol* 280, 482-493.
- Thomas, S., Li, X.Y., Sabo, P.J., Sandstrom, R., Thurman, R.E., Canfield, T.K., Giste, E., Fisher, W., Hammonds, A., Celniker, S.E., *et al.* (2011). Dynamic reprogramming of chromatin accessibility during *Drosophila* embryo development. *Genome Biol* 12, R43.

Warner, J., Philippakis, A., Jaeger, S., He, F., Lin, J., and Bulyk, M. (2008). Systematic identification of mammalian regulatory motifs' target genes and functions. *Nature Methods* 5, 347-353.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., *et al.* (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158, 1431-1443.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137.

Zhu, L.J., Christensen, R.G., Kazemian, M., Hull, C.J., Enuameh, M.S., Basciotta, M.D., Brasefield, J.A., Zhu, C., Asriyan, Y., Lapointe, D.S., *et al.* (2011). FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res* 39, D111-117.

## Chapter 2: rare cell purification

This chapter presents a minor and parallel project of my PhD. In the first subchapter, I will present the needs and reason for the development of a reliable method of cell purification or cell panning. Then, I will describe the method that was adapted for use in *Drosophila melanogaster* embryos and will continue with presenting the preliminary results that were obtained so far. I will finally conclude this chapter with a section regarding the future directions of this project.

## Introduction

Metazoans are composed of millions to billions of cells of very different types, creating tissues with various roles. These tissues are used in combination to create organs with very different functions, and allow the organism to live, survive and reproduce. All these functions are the result of very intricate and complex interactions of the building and functioning “units” that are the cells. To understand biology, the field has never ceased to try and reach a deeper understanding of organs, tissues, cell populations and cell types. A relative easy approach that has been used until recently is to analyze the function of cells as a tissue or large population of cells, playing a relative common role in the organ being studied, such as in the experiment I previously described in the first chapter of this dissertation. Nonetheless, cell population assays mask the presence of rare or small subpopulations of cells, by averaging the behavior of the population being analyzed. This can be problematic: for instance, when a population is composed of cells in different cell states, ignoring cell heterogeneity allows models that seem accurate when looking at population trends, but does not reflect any precise individual cell (Altschuler and Wu, 2010). Therefore, cell to cell heterogeneity poses practical challenges for building accurate clinical models, particularly those based on population-averaged measurements, to guide diagnosis and treatment of diseases (Campbell and Polyak, 2007), such as cancer for instance, known for the heterogeneity of the cancerous cells (Heppner, 1984; Rubin, 1990). This is, *de facto*, an issue that needs to be addressed as well for the study of transcription regulation.

Cell type or tissue specific information is a challenge to obtain, and the two main methods currently used are FACS and tissue dissection. Tissue dissection has been successful for studying tissue-specific activity of enhancers (Visel et al, 2009; Blow et al., 2010) and chromatin signatures during mouse development for instance (Soshnikova and Duboule, 2009; Xu et al., 2010), but has not proven to provide cell type-specific signature, since accurate dissection of cells from a heterogeneous tissue is either not possible (e.g. cell of a very small size and, or, interconnected with other cellular or tissue structures) or extremely time consuming. FACS has been used on cells or nuclei from tissues, blood, or even dissociated embryos, to investigate cell type-specific ChIP (Boon et al., 2012) and RNA expression and their related signatures (Christiaen et al., 2008). Nonetheless, cell sorting a sufficient number of cells for subsequent experiments can remain a challenge when dealing with rare populations of cells, and the results do not account for cellular heterogeneity.

This led the birth of single cell based methods to analyze single cell genome and transcriptome, but also proteome and metabolome (Huang et al., 2015; Tsioris et al., 2014; Wu et al., 2017). These methods, such as single cell ATAC-seq methods (Pott and Lieb, 2015) or single-cell RNA-seq methods (Kolodziejczyk et al., 2015) are promising as they should allow to look into cell-to-cell heterogeneity in cancer or in a developing organism, the dynamic of transcription and the regulatory relationships between genes. Nonetheless, these methods suffer inherent limitations: current single cell ATAC-seq approaches, for instance, capture only a small subset of the open chromatin sites in single cells (Pott and Lieb, 2015) and single cell RNA-seq methods have limited sensitivity making it difficult, for instance, to distinguish between technical noise and transcript present in low abundances (about 10 copies per cell) (Grün et al., 2014), leading to an important loss of information. While these methods are improving at a rapid pace and should eventually allow to better understand cellular heterogeneity and create 3D models of tissues, organs or even organisms (Saliba et al., 2014) with information based on gene regulatory networks, no method really exists so far for the discovery of CRMs, enhancers and silencers. In fact, developing such a method remains a challenge, as a single-cell reporter assay would be facing the same issues of cell-type purification and sorting, sensitivity, and cellular heterogeneity.

Thus, studying and identifying CRMs and their roles in rare populations of cells with methods such as eFS or sFS, in which the transcription of reporter gene is regulated by a specific CRM for instance, is strongly limited by the time spent sorting the appropriate rare cell population by FACS, and therefore not realistically feasible. In our previous eFS and sFS experiments, the activities of candidate CRMs were analyzed by looking at the developing mesoderm, a tissue composed of very different cell types that all develop into specific muscles with very different shapes and functions for instance. To investigate the precise activity of a CRM to understand its role, it is therefore necessary to be able to identify in which specific cells this CRM is regulating transcription.

## Cell panning approach

Magnetic nanoparticles or beads coated with antibodies have been developed and used for the separation of cell suspensions from tissues, organ samples or blood. This method was first developed by Miltenyi Biotech (<http://www.miltenyibiotec.com>) and branded as “Magnetic-Activated Cell Sorting” (MACS), but similar methods are available from other companies, such as the ‘DynaBeads’ from ThermoFisher Scientific. In such methods, the antibodies used are specific for certain cell surface markers, either expressed on a population of interest for positive selection, or expressed on undesired cell types for negative selection. After incubation with the antibody-coated beads, the cell suspension is separated by the use of a magnet, attracting the cells that bound to the beads via the specific antibody or antibodies, while unlabeled cells remain undisturbed.

If performing a negative selection, the unbound cells are kept while the bound cells are discarded. On the contrary, if performing a positive selection, only the cells that bound to the beads are kept for further experiments and analyses. The negative selection method requires the use of a cell surface marker, or a set of markers, present on most of the cells but absent on the surface of the population of interest. This can be challenging, especially in the case of rare populations of cells, but has the advantage of leaving the cells of interest untouched and unperturbed. Binding to an antibody might stress the cells and change their transcription profiles, especially in immune cells, since the cell markers used (CD4 or CD8 for instance) are endogenous and may trigger an immune response from the bound cell, or activate signaling pathways and modify their behavior which would result in modified transcription profiles for instance. The positive selection is therefore easier to perform, but does not address this last potential issue.

To purify rare populations of cells from *Drosophila melanogaster* embryos and investigate the regulation of transcription in small or even rare cell populations, we decided to adapt a commercially available kit, from FisherScientific (see Methods), designed for magnetic isolation (positive panning) of CD8+ from human blood, for use in lines that would express the human CD8 cell membrane protein under the control of a specific enhancer, which would drive expression in specific cells from dissociated embryos. This kit has proven effective, for instance, for the positive isolation of CD8 T cells from spleen, for immunotransplantation in mouse, which showed success in curing large lymphoma tumors (Brody et al., 2008).

By doing so, we hypothesized that this method should allow us to discard, prior to FACS sorting, cells of non-interest, *i.e.* cells that do not express the specific cell marker (CD8- cells), while potentially not disturbing their transcriptional profiles, given that this cell marker is exogenous and should not be involved in any signaling pathway. The FACS sorting is nonetheless necessary to ensure the purity of the CD8+ cells (cell type or rare population of interest). This way, we thought it should be possible to work from very large pools of cells from dissociated embryos,

purify in large batches our cells of interest and discard the large number of CD8<sup>-</sup> cells without having to actually sort them.

## Preliminary results

### pCD8 vector

To collect cells from developing ectoderm for our initial proof of concept, we used our gateway compatible pCD8 vector (see Methods) to create a new homozygous fly line (dpp\_VRR:CD8) that would express the human CD8 cell membrane protein in the ectoderm, under the control of the dpp\_VRR enhancer (see Methods). This element was previously reported for inducing expression in the embryonic blastoderm (Huang JD, 1995). We validated the expression of the CD8 protein by microscopy and compared the CD8 expression pattern to our preexisting twi:CD2 line (see Fig. 1. a). We then measured the percentage of cells this construct would be expressed in an embryo and found that about a third of cells carried the CD8 protein on their membrane (see Fig. 1. b). The mesoderm, based on our experiments, represents ~15% of the total cells in an embryo at stage 11 or 12 (supplementary figures 6 and 7 from Chapter 1).

### Cell panning

Using the homozygous dpp\_VRR:CD8 line, we attempted to pull down the CD8<sup>+</sup> cells, by adapting the “Dynabeads FlowComp Human CD8” kit from ThermoFisher Scientific (Catalog number 11362D), initially designed for isolation of flow-compatible human CD8<sup>+</sup> T-cells from peripheral blood mononuclear cells or whole blood. This kit presented the advantages of allowing the dissociation of the beads from the cells prior to cell sorting (and therefore staining), and to be relatively gentle on the samples: the CD8<sup>+</sup> cells are incubated with an antibody that is directed against CD8, and conjugated with a modified biotin with decreased affinity to streptavidin. Once the cells are bound to the antibody, magnetic beads coated with streptavidin (Dynabeads) are used to pull down CD8<sup>+</sup> using a magnet, which attracts the cell-antibody-bead complex and allows the user to discard CD8<sup>-</sup> cells. Then, the CD8<sup>+</sup> are washed and eluted from the beads, by using a release buffer that contains regular biotin. It is by simple competition between the regular biotin and the modified biotin (with lower affinity to streptavidin) that the CD8<sup>+</sup> cells are eluted from the beads. We attempted to adapt the protocol provided by ThermoFisher Scientific to dissociated *Drosophila melanogaster* embryos. This protocol is scalable from  $1 \times 10^7$  to  $5 \times 10^8$  cells and described to work for fewer cells, in which case the volumes of reagents used are the same as for  $1 \times 10^7$  available cells. Moreover, in this protocol, it is supposed that ~20% of the cells are expressing CD8.

A few adjustments needed to be made to use this kit and protocol for transgenic *D. melanogaster* cells, since this kit is designed for sorting human CD8<sup>+</sup> cells from blood. First, capturing all CD8<sup>+</sup> cells required increasing three-fold the amount of antibody directed against CD8<sup>+</sup>, when comparing to the manufacturer’s recommendations. Then, the elution of the CD8<sup>+</sup> cells from the beads proved ineffective with the provided elution buffer, despite trying several different incubation times and increasing the amount of buffer used by many folds. We hypothesized that the concentration of regular biotin needed to be increased as compensation of the increased amount of antibody used in our modified protocol, and therefore tried to elute the cells with a solution saturated in biotin (see Methods).

Although the cytometer was out-of-order at the time of this experiment, pictures were taken under microscope to observe whether cells were present in the elution, and in theory be CD8+. The pictures in figure 3 show that putative CD8+ cells are present in the solution that resulted from the elution with the solution saturated in biotin, which is encouraging, as no cells were observable in the negative control for elution (*i.e.* elution step without biotin, data not shown).

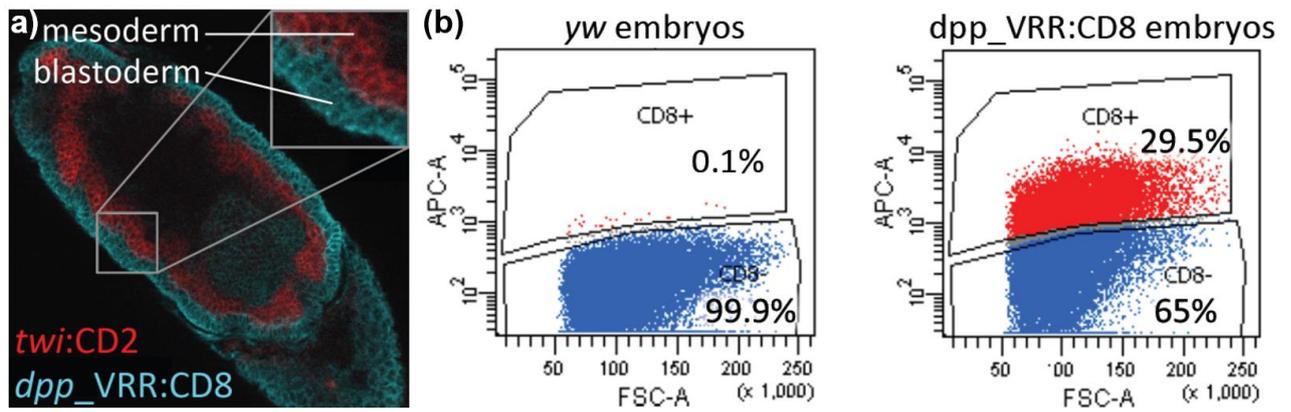
At this stage, this project has proven that it is possible to have the CD8+ cells bind to the provided conjugated antibody with a yield of ~90%, in the dpp\_VRR:CD8 line (Fig. 2) that expresses CD8 in a large fraction of the embryo, ~30% (see Fig. 1.b). This percentage is close to what the manufacturer designed their kit for in human blood and we believe that targeting smaller populations of cells, in other CD8 lines, can be done by adjusting the amount of antibody used.

## Future directions

In order to develop a reliable protocol of cell purification from this project, it is necessary to realize further tests by FACS to know whether the CD8+ are effectively eluted from the magnetic beads by the solution of biotin. The eluted putative CD8+ cells will be stained for the CD8 marker and sorted by FACS and compared to the cells that were not positively selected and eluted, including the potential cells that would remain bound the beads as they can also be sorted. These putative CD8- cells will also be stained for CD8, to assess yield and specificity of the method. The purity of the method will be assessed by this same approach, by specifically looking for unlabeled cells in the eluted fraction.

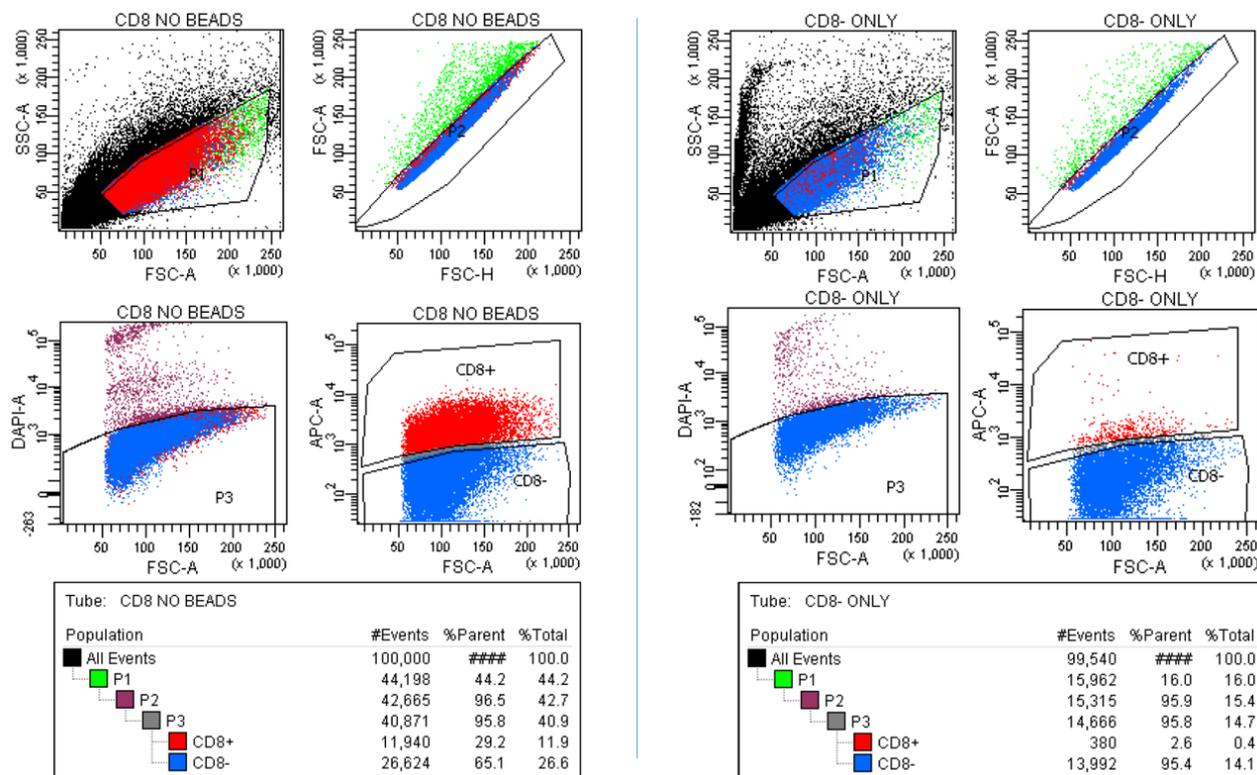
Once this is validated, new drosophila lines will have to be created, to express the CD8 gene under the control of other enhancers, targeting rare cell types, or use endogenous markers. The fusion competent myoblast (FCM) specific enhancer, *Mef2* I-E<sub>D5</sub>, which drives expression specifically in the fusion competent myoblasts (FCMs) of the developing mesoderm (Duan et al., 2001), has already been used and cloned into the pCD8 vector and a homozygous line has been established (Fig. 4). This line would be a great line to start with, as this enhancer drives expression in ~1% of the total cells on an embryo (Duan et al., 2001).

Once a final and reliable protocol is reached, this method of positive cell panning could be used for several approaches. First of all, experiments such as eFS or sFS could be realized on small populations of cells, to identify cell-type specific CRM activity (enhancer, silencer or bifunctional elements). This method could be used for other experiments, such as ATAC-seq, Hi-C, etc, to obtain cell-type specific data, which could then be compared to single cell data available for these same rare populations of cells. It would be indeed interesting to see how data from a “population” based approach such as this method of cell panning compare to single cell data. Nonetheless, it will be necessary to address the question of whether this positive selection method might perturb downstream measurements. This could be done, for instance, by comparing the profiles from a rather small population of cells that remains sortable only by FACS, to cells purified by this method, such as the FCMs (Fig. 4).



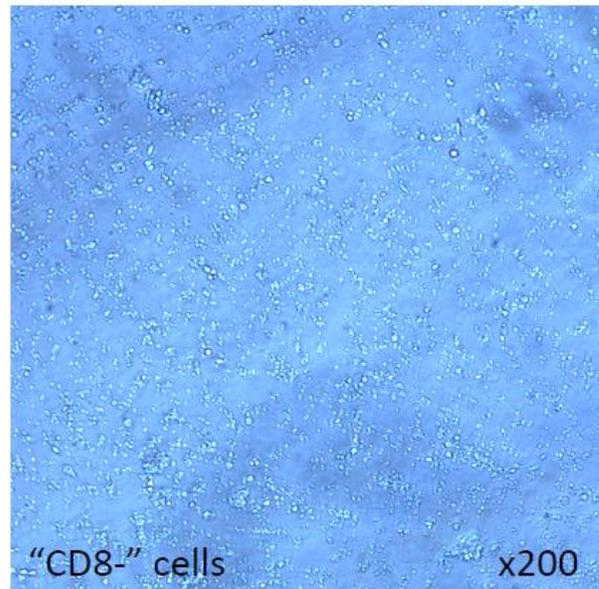
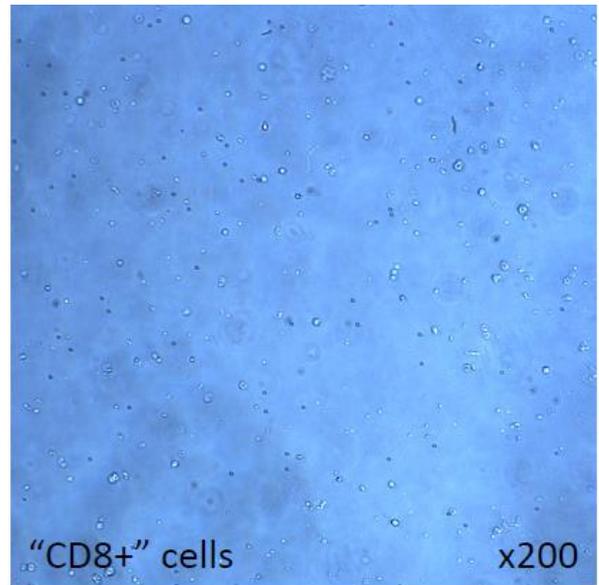
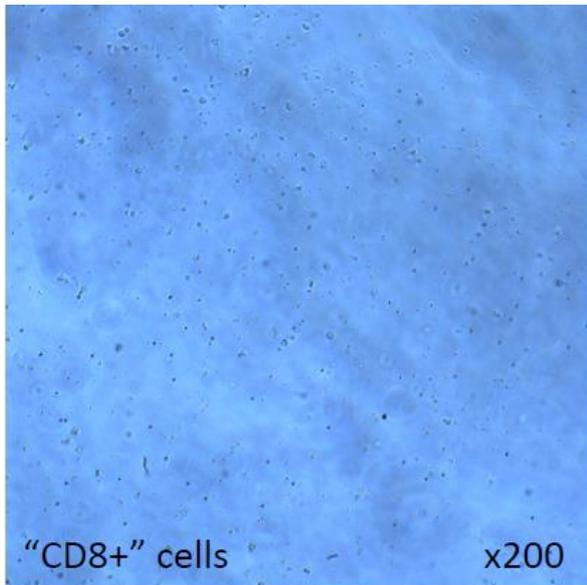
**Figure 1. Analysis of the dpp\_VRR:CD8 line**

(a) dpp\_VRR drives CD8 expression (aqua) in the blastoderm, distinctly outside the mesoderm (marked by twi:CD2, red) at stage ~11. (b) FACS output showing cells from the blastoderm (CD8+) cells (here, red) from dpp\_VRR:CD8 embryos (right), as compared to background levels obtained from wildtype (yw) embryos (left), at stage ~11.



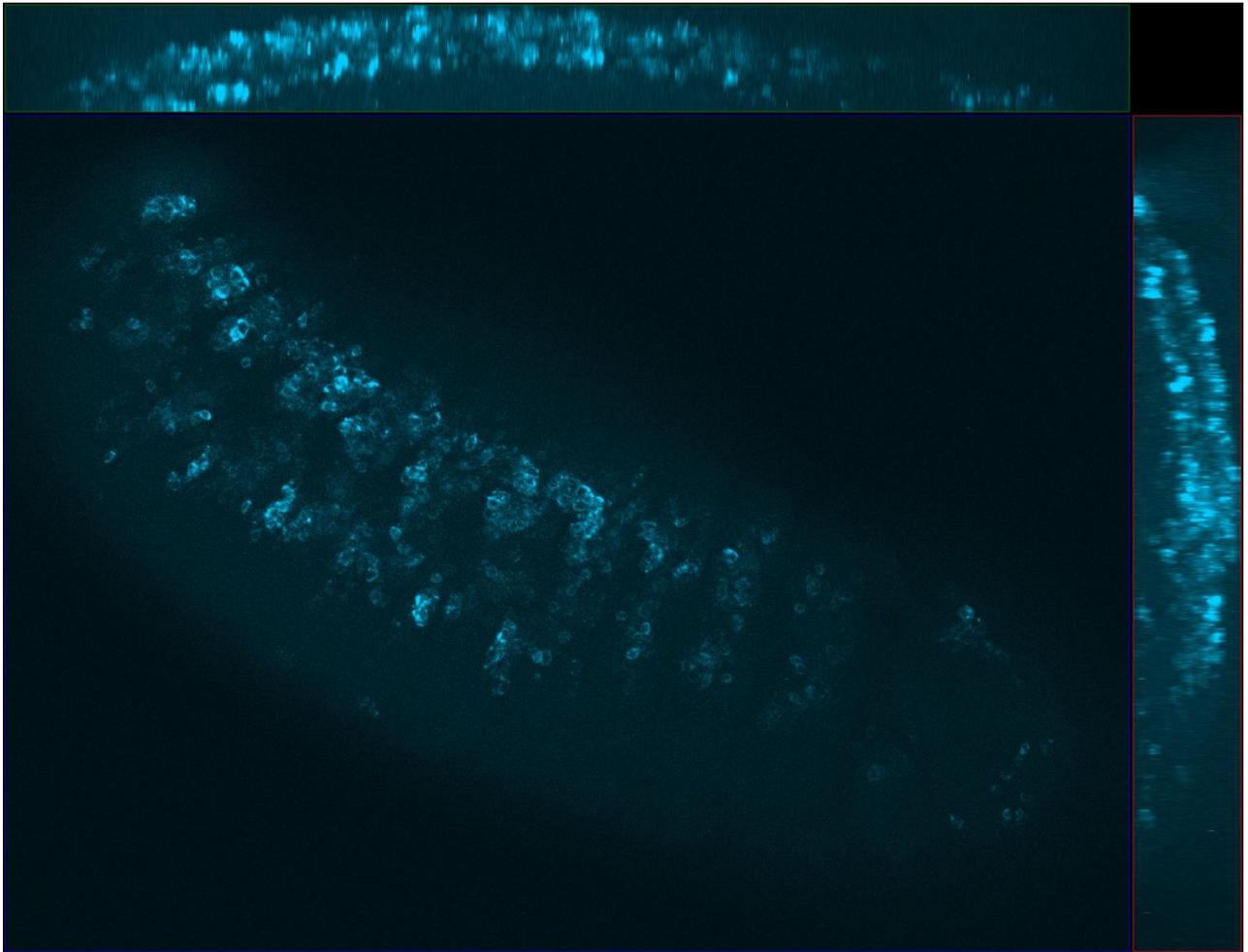
**Figure 2. CD8+ cell retention on beads**

This figure presents the FACS outputs from dpp\_VRR:CD8 cells from dissociated embryos at stage ~11, in two conditions. On the left panel, the cells are not treated and the CD8+ population covers about 30% of the total cells, as expected. On the right, the cells have been incubated with the antibody directed against CD8, conjugated with modified-biotin; the cells presented here are the cells that were not retained by the magnetic beads, after incubation with said beads. The percentage of CD8+ cells retained by the beads is important, ~90%.



**Figure 3. Elution with biotin and microscope observation (x200)**

This figure presents a quick observation of the supposed success of elution of cells from the magnetic beads (top two pictures) with a solution saturated in biotin. The cells were not counted here in this experiment.



**Figure 4. *Mef2* I- $E_{D5}$ :CD8 line**

The *Mef2* I- $E_{D5}$  enhancer drives expression the fusion competent myoblasts within the developing embryonic mesoderm, which represents ~1% of the total cells in an embryo. Here, the expression of CD8 (cyan color) is shown in an embryo at stage 13.

## Methods

### Generation of vector pCD8

To create the pCD8 vector, we PCR amplified the Hsp70 promoter from our pSFS vector and the human CD8 cDNA. These amplicons were used for PCR assembly and then cloned into the pETWN plasmid (Halfon MS, 2008), replacing its *LacZ* cassette. Then, this plasmid was made gateway compatible and turned into our final pCD8 plasmid. The full sequence of this plasmid is available as annex 1 of this dissertation. The detailed map of this plasmid can be below. This plasmid can be inserted randomly into a genome as it contains a P element transposon, with the help of a helper plasmid (Hartwell et al., 2015).

### Creation of dpp\_VRR:CD8 vector and fly lines

The dpp\_VRR element (coordinates chr2L:2456345-2456884 in dm6) was amplified from wild type *Drosophila melanogaster* genome and cloned via Gateway BP Clonase II (Invitrogen) into pDONR221 (Invitrogen) (forward primer GGGGACAAGTTTGTACAAAAAAGCAGGCTCTTCCTATACCTGAATTTTCCACCT, reverse primer GGGGACCACTTTGTACAAGAAAGCTGGGTTAGACACTTGGTTTGTGCGG). Cloning reactions were transformed into *E. coli* Top10 cells (Invitrogen), and plated on LB agar with kanamycin. Plasmids were purified from the resulting colonies, from which the dpp\_VRR element was cloned using Gateway LR Clonase II (Invitrogen) into pCD8. Transformed cells were plated on LB agar with ampicillin, yielding colonies from which the final dpp\_VRR:CD8 plasmid was sequenced and, once the sequence was validated, prepared for embryo injection. Similarly, the *Mef2* I-Ed5 was amplified by PCR and cloned as previously explained (forward primer GGGGACCACTTTGTACAAGAAAGCTGGGTTGCTGGGTATCTGCAAGATGG, reverse primer GGGGACAAGTTTGTACAAAAAAGCAGGCTGGTTTCTTCAGGGGGATCTTT).

This plasmid was injected into posteriorly into syncytial *yw* embryos, along with the helper plasmid p13t/wc carrying the necessary transposase. Surviving males were crossed to excess *yw* virgin females. Transformant male progeny were selected by eye color. These males were then crossed to virgin females from balancer lines to identify within which chromosome the construct was inserted. The lines used for this experiment have the dpp\_VRR:CD8 construct inserted into the 2<sup>nd</sup> chromosome, or the *Mef2* I-Ed5:CD8 construct inserted into the 3<sup>rd</sup> chromosome. Both lines were made homozygous.

Expression of CD8 was validated by crossing males from the CD8 line dpp\_VRR:CD8 to virgin females from the *twi*:CD2 line. Embryos were collected, fixed and stain for both CD2 and CD8 cell membrane proteins and observed under fluorescent microscope (see Fig. 1.a.). For FACS, we used the same method for isolation of single cells from *Drosophila* embryos, described in Chapter 1. We here used the Alexa Fluor 488 conjugated anti(human CD8) (Biolegend, cat. #344716, 1:20 final dilution) for the dpp\_VRR:CD8 line. The cell sorting gates were designed based on the same gates as previously explained for the *twi*:CD2 line, but selecting the CD8+ cells by sorting by designing a special gate to capture these cells described (far red [APC-A] signal vs. forward scatter [FSC-A], see figure 1.b and figure 2 for examples). The *Mef2* I-Ed5:CD8 was stained only for CD8 and observed under microscope (figure 4).

## Cell purification protocol – a work in progress

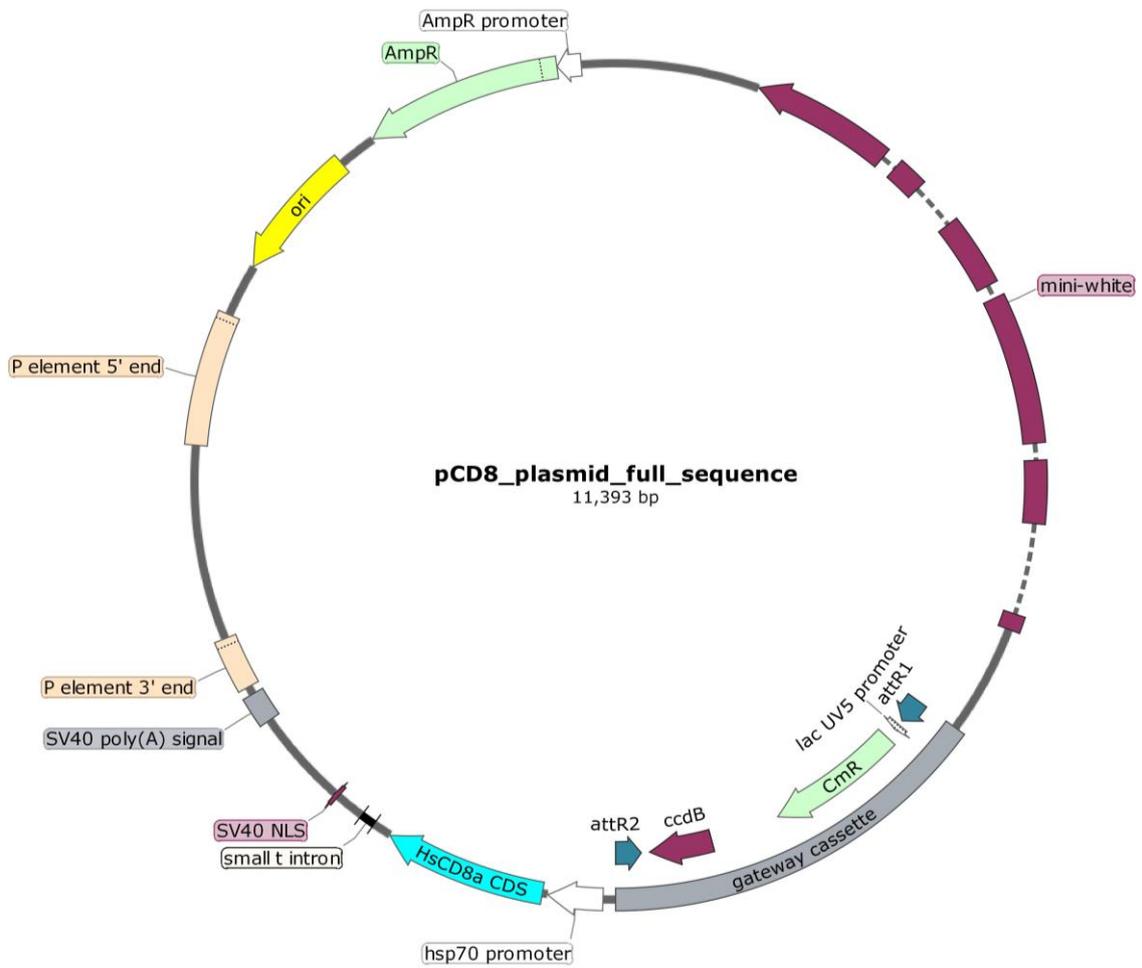
We adapted the previously described method for isolation of single cells from *Drosophila* embryos, at stage 11 (see above and Methods section of Chapter 1 of this dissertation) to include cell purification steps. These steps are adapted from the protocol from ThermoFisher Scientific “Dynabeads FlowComp Human CD8 isolated from PBMC” (available at this address: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/dynaflow\\_hu\\_CD8\\_pbmc\\_man.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/dynaflow_hu_CD8_pbmc_man.pdf)).

To adapt it, it was first necessary to evaluate the amount of CD8<sup>+</sup> cells present in a solution of dissociated cells from whole embryos. On average, we found that ~10% cells survived the dissociation process. Appropriate dilutions of the modified-biotin conjugated anti-CD8 antibody were therefore necessary. It appeared that a 3-fold concentration of this antibody was necessary to pull down the CD8<sup>+</sup> cells after dissociation, when compared to the recommendations of the manufacturer’s protocol (considering the same number of cells).

The isolation buffer used was adapted to match S2 cell culture medium (Na<sub>2</sub>HPO<sub>4</sub>: 0.850g/L, NaCl: 6.2g/L, KCl: 2g/L, KH<sub>2</sub>PO<sub>4</sub>: 0.5g/L, EDTA 2mM, FBS 2%, pH 7.4) and the elution buffer was prepared from this isolation buffer and saturated with D-biotin (0.2mg/ml). The elution buffer was used at a 1:2 final dilution in the solution that contained the CD8<sup>+</sup> cells and the magnetic beads.

Regarding the protocol itself, we only adapted the temperatures and length of incubation: all necessary incubations were done on ice or at 4°C, for 20min. An example of the protocol can be found as annex 2 of this dissertation.

# Supplementary figure



**Supplementary Fig. 1.**  
Map of the gateway compatible pCD8 plasmid.

## References

- Altschuler, S.J., and Wu, L.F. (2010). Cellular Heterogeneity: Do Differences Make a Difference? *Cell* 141, 559–563.
- Blow, M., McCulley, D., Li, Z., Zhang, T., Akiyama, J., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Bristow, J., Ren, B., Black, B., Rubin, E., Visel, A. & Pennacchio, L. (2010). ChIP-Seq identification of weakly conserved heart enhancers. *Nature Genetics* 42, 806-810.
- Bonn, S., Zinzen, R., Perez-Gonzalez, A., Riddell, A., Gavin, A. & Furlong, E. (2012). Cell type-specific chromatin immunoprecipitation from multicellular complex samples using BiTS-ChIP. *Nature Protocols* 7, 978-994.
- Brody, J., Goldstein, M., Czerwinski, D. & Levy, R. (2008). Immunotransplantation preferentially expands T-effector cells over T-regulatory cells and cures large lymphoma tumors. *Blood* 113, 85-94.
- Campbell, L.L., and Polyak, K. (2007). Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle Georget. Tex* 6, 2332–2338.
- Christiaen, L., Davidson, B., Kawashima, T., Powell, W., Nolla, H., Vranizan, K. & Levine, M. (2008). The Transcription/Migration Interface in Heart Precursors of *Ciona intestinalis*. *Science* 320, 1349-1352.
- Duan, H., Skeath, J.B., and Nguyen, H.T. (2001). Drosophila *Lame duck*, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development. *Dev. Camb. Engl.* 128, 4489–4500.
- Grün, D., Kester, L., and Oudenaarden, A. van (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* 11, 637.
- Halfon, M.S., Gallo, S.M., and Bergman, C.M. (2008). REDfly 2.0: an integrated database of cis-regulatory modules and transcription factor binding sites in *Drosophila*. *Nucleic Acids Res.* 36, D594-598.
- Hartwell S., Goldberg M., Fischer J., Hood L., Aquadro C. (2015). *Genetics: From Genes to Genomes*, 5th Edition, *McGraw-Hill*.
- Heppner, G.H. (1984). Tumor heterogeneity. *Cancer Res.* 44, 2259–2265.
- Huang, L., Ma, F., Chapman, A., Lu, S., and Xie, X.S. (2015). Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu. Rev. Genomics Hum. Genet.* 16, 79–102.
- Pott, S., and Lieb, J.D. (2015). Single-cell ATAC-seq: strength in numbers. *Genome Biol.* 16.
- Rubin, H. (1990). The significance of biological heterogeneity. *Cancer Metastasis Rev.* 9, 1–20.
- Saliba, A.-E., Westermann, A.J., Gorski, S.A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42, 8845–8860.

Soshnikova, N. & Duboule, D. (2009). Epigenetic Temporal Control of Mouse Hox Genes in Vivo. *Science* 324, 1320-1323.

Tsioris, K., Torres, A.J., Douce, T.B., and Love, J.C. (2014). A New Toolbox for Assessing Single Cells. *Annu. Rev. Chem. Biomol. Eng.* 5, 455–477.

Visel, A., Blow, M., Li, Z., Zhang, T., Akiyama, J., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E. & Pennacchio, L. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854-858.

Wu, A.R., Wang, J., Streets, A.M., and Huang, Y. (2017). Single-Cell Transcriptional Analysis. *Annu. Rev. Anal. Chem.* 10, 439–462.

Xu, C., Cole, P., Meyers, D., Kormish, J., Dent, S. & Zaret, K. (2011). Chromatin "Prepattern" and Histone Modifiers in a Fate Choice for Liver and Pancreas. *Science* 332, 963-966.

# Conclusion and Future directions

## Summary

Despite the common treatment of enhancers and silencers as two distinct groups of regulatory elements, a few elements in a variety of eukaryotic systems (Bessis et al., 1997; Jiang et al., 1993; Kallunki et al., 1998; Kehayova et al., 2011; Koike et al., 1995; Prasad and Paulson, 2011; Schaeffer et al., 1995; Simpson et al., 1986; Stathopoulos and Levine, 2005; Stroebel and Erives, 2016) have been found to exhibit both activities; *i.e.*, bifunctional elements that can act as either an enhancer or a silencer, depending on the tissue type or cellular conditions. Bifunctionality of *cis*-regulatory modules (CRMs) complicates the prediction of gene expression from sequence and the interpretation of the effects of *cis*-regulatory variation across populations or in evolution. Moreover, it has remained unknown how general this property might be and how many such bifunctional elements a typical metazoan genome might contain.

Screening for bifunctional CRMs requires the ability to assay a *cis*-element for both enhancer activity in one cell type and silencer activity in a different cell type. To perform such experiments *in vitro*, it is necessary to have a candidate cell type for silencer activity; unfortunately, our current state of knowledge does not allow us to predict silencer activity with the information currently available for specific cell lines (Riel, 2014). We have therefore chosen to assay silencer activity *in vivo*, using cells isolated from whole *Drosophila melanogaster* embryos, in an attempt to discover silencers across a range of candidate cell types.

The results of the work presented in this dissertation suggest the possibility that many silencers may also be enhancers in a different cell type. Indeed, many transcription factors (TFs) can act as either activators or repressors, depending on the context of the *cis*-element they bind (Ogbourne and Antalis, 1998). However, bifunctionality of a *cis*-element does not require such TFs, since different activators or repressors could bind the same element in different tissues and therefore cell type-specific silencer activity could be an important contributor to enhancer specificity by preventing inappropriate activation and proper and precise patterns of gene expression.

Pfeiffer *et al.* estimated that there may be over 50,000 enhancers in the *D. melanogaster* genome (Pfeiffer et al., 2008), and Heintzman *et al.* estimated there may be on the order of  $10^5$ - $10^6$  enhancers in the human genome (Heintzman et al., 2009). The observation that the vast majority of complex trait- and disease-associated variants identified from genome-wide association studies (GWAS) map to noncoding sequences, most of which occur within DNase hypersensitive sites (Maurano et al., 2012), emphasizes the importance of understanding these elements. Our detection of mesodermal silencer activity in more than 10% of tested non-mesodermal enhancers suggests that there may be thousands of such bifunctional elements across a range of tissues in *Drosophila*, and perhaps  $10^4$ - $10^5$  in human (Heintzman et al., 2009); since many of the elements we tested could be silencers in a cell type we did not examine or at a later developmental stage, it is possible that these number might be even higher.

## Limitations

Although successful at identifying silencers, the sFS experiment was not able to provide any significant silencer or bifunctional CRM signature. Current available ChIP-seq and ChIP-chip data is limited and we had only access to data from sorted mesoderm or whole embryo. This is a limitation several reasons. First, looking at chromatin marks from whole embryos for the elements we identified may not actually represent the tissue specific histone modifications responsible for an element to be enhancer in a tissue while being a silencer in another.

Indeed, whole embryo data represents an average profile for a specific element across all tissues at a certain time point. In our case, looking at mesoderm specific data seems more appropriate, but not sufficient: to be able to characterize our bifunctional elements, we not only need the mesoderm specific data that was available to us, but also data from all tissues but the mesoderm (non-mesoderm data), so as to be able to actually detect significant differences in chromatin profiles. By doing so, it could be possible to find a signature for the elements we identified.

Moreover, the mesoderm itself is a broad tissue that is composed of several cell types, and it is possible that a given bifunctional element acts as a repressor only a fraction of the mesoderm, *i.e.* in specific sub-mesodermal cell types. Looking for a chromatin mark signature by averaging the profiles within the mesoderm could also be a limitation of the approach we used. Unfortunately, the available data on chromatin marks is rather limited and obtaining such data ourselves would be a very time consuming and extensive project.

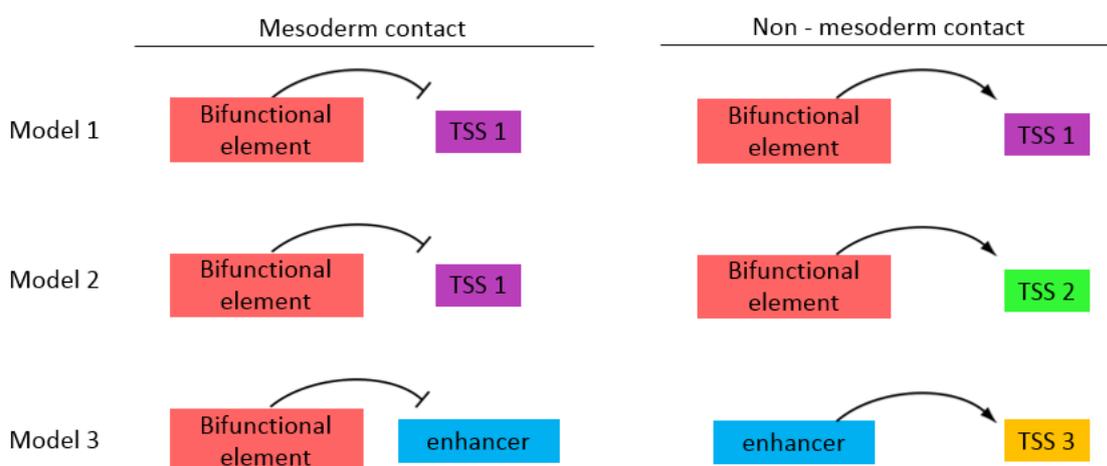
Another limitation in the understanding of the manner in which these bifunctional elements function is that only a relatively small number of TF motifs are known so far in *Drosophila melanogaster*. Our database comprised about a hundred motifs, which is very little in comparison of the *Drosophila* genome. As it appears that these bifunctional elements are enriched for overlap with HOT regions where more than 10 TF should bind (see results in Chapter 1), it is obvious that we need to increase our knowledge of TF binding motifs if we want to understand the transcriptional code responsible for the dual function with our elements. Nonetheless, obtaining such data is significant endeavor that could not be undertaken during this PhD.

## Future directions

The sFS approach could be adapted in future studies to screen for bifunctional elements in mammalian embryonic development or differentiation of adult cells. Moreover, the work presented in this dissertation shows that despite the extensive genome-scale ChIP profiling studies by numerous investigators and consortia, the available data are not sufficient to distinguish the subset of enhancers that are bifunctional and that no signature was found for these bifunctional elements.

Extended efforts in profiling larger sets of tissue-specific chromatin marks might reveal a signature of bifunctional CRMs and the characterization of bifunctional elements should help in elucidating how precise gene expression patterns are encoded in the genome and aid in the interpretation of *cis*-regulatory variation. Spatial chromosomal interaction mapping with techniques such as Hi-C are moreover necessary to try and understand the relationships between CRMs and their target genes, and the mechanisms that control precise spatio-temporal gene expression.

We are currently looking forward to the results of the analysis of the Hi-C data we generated in collaboration with the Dekker laboratory, as we believe that such a potential high-resolution dataset of mesoderm specific interaction mapping will unveil very interesting and crucial information and mechanisms and will lead to new theories about gene expression regulation. Although we were not able to provide any hints of a mechanism for the elements we identified by the sFS method so far, we hope that the HiC data, once analyze will help us understand how the bifunctional elements we identified work and silence their target genes. As shown in figure 1 below, there are three main models that could explain their mechanisms. It is possible that a bifunctional CRM acts as a silencer in the mesoderm and as an enhancer in other tissues, at the same TSS (model 1). It is also possible for a bifunctional CRM to silence expression at a given TSS in the mesoderm and to act as an enhancer in other tissues at other TSSs (model 2). Finally, it is possible that a bifunctional elements indirectly represses transcription in the mesoderm by acting on enhancers in this tissue and not in others (model 3). All these models seem viable and do not exclude one another.



**Figure 1: models of potential function of bifunctional elements**

Here are three different models for the function of bifunctional elements in which transcriptional silencers act directly on transcriptional start sites (TSS) by chromatin contacts (models 1 and 2) or act at distal enhancers (model iii).

By analyzing the HiC data, it should be possible to identify the target regions of our bifunctional CRMs outside the mesoderm (where they seem to act as enhancers) and identify any difference in these contacts in the mesoderm. Combined to the analysis of the HiC data, the CRM knockout project may provide an additional layer of information and, potentially, of validation. Indeed, knocking-out these bifunctional CRMs and *in situ* RNA hybridization should provide us with information about which genes these bifunctional CRMs regulate and in which tissues the target genes are expressed or silenced. As the HiC data should provide us with genome-wide chromosomal contact data, it should be possible to see whether the bifunctional CRMs behave as silencers by directly acting at TSSs (Fig. 1, model 1 and 2), or indirectly at other enhancers (Fig. 1, model 3).

We are moreover expecting additional results from combinatorial motif finding analyses, which may identify a set or sets of motif combinations specific to our validated bifunctional CRMs, despite our limited motif library. Also, and once finalized, the cell-panning approach presented in the second chapter of this dissertation should allow a more precise study of small population of cells, potentially providing us with more detailed data on regulation of transcription. This approach should help us address the question of cell to cell heterogeneity within tissues and obtain more precise transcription profiles and data for experiments such as eFS and sFS, which would, hopefully, lead to actual silencer or bifunctional CRM histone profiles for instance.

## Concluding remarks

Although not yet published and still at the stage of preparation, the work I presented in this dissertation presents an additional level of complexity for transcription regulation, as we now move away from the current binary categorization of CRMs. This seems somehow similar to the evolution of how genetic information was encoded, since the assumption of “One gene, one function” that Beadle and Tatum stated in 1941 is now clearly outdated (Beadle and Tatum, 1941). In the same manner as one gene leads to several different transcripts and thus proteins with sometimes radical different functions, and given the findings presented previously, it seems reasonable to state that they may not be enhancers and silencers, but simply CRMs playing different roles, at different times, in different places of an organism.

## References

- Beadle, G. & Tatum, E. (1941). Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Sciences* 27, 499-506.
- Bessis, A., Champtiaux, N., Chatelin, L. & Changeux, J.P. The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc Natl Acad Sci U S A* 94, 5906-11 (1997).
- Duan, H., Skeath, J.B., and Nguyen, H.T. (2001). Drosophila *Lame duck*, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development. *Dev. Camb. Engl.* 128, 4489–4500.
- Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* 28, 817-25 (2010).
- Filion, G.J. *et al.* Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* 143, 212-24 (2010).
- Gallo, S.M. *et al.* REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Research* 39, D118-23 (2011).
- Gisselbrecht, S.S. *et al.* Highly parallel assays of tissue-specific enhancers in whole Drosophila embryos. *Nat Methods* 10, 774-80 (2013).
- Heintzman, N.D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-8 (2007).
- Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108-12 (2009).
- Jiang, J., Cai, H., Zhou, Q. & Levine, M. Conversion of a dorsal-dependent silencer into an enhancer: evidence for dorsal corepressors. *EMBO J* 12, 3201-9 (1993).
- Kallunki, P., Edelman, G.M. & Jones, F.S. The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. *Proc Natl Acad Sci U S A* 95, 3233-8 (1998).
- Kharchenko, P.V. *et al.* Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* 471, 480-5 (2011).
- Kehayova, P., Monahan, K., Chen, W. & Maniatis, T. Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A* 108, 17195-200 (2011).
- Koike, S., Schaeffer, L. & Changeux, J.P. Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene. *Proc Natl Acad Sci U S A* 92, 10624-8 (1995).
- Pfeiffer, B.D. *et al.* Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9715-20 (2008).

- Maurano, M.T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190-5 (2012).
- Mani-Telang, P. & Arnosti, D.N. Developmental expression and phylogenetic conservation of alternatively spliced forms of the C-terminal binding protein corepressor. *Dev Genes Evol* 217, 127-35 (2007).
- Nieto, M.A. The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* 3, 155-66 (2002).
- Noyes, M.B. *et al.* Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* 133, 1277-89 (2008).
- Ogbourne, S. & Antalis, T.M. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* 331 ( Pt 1), 1-14 (1998).
- Orian, A. *et al.* A Myc-Groucho complex integrates EGF and Notch signaling to regulate neural development. *Proc Natl Acad Sci U S A* 104, 15771-6 (2007).
- Prasad, M.S. & Paulson, A.F. A combination of enhancer/silencer modules regulates spatially restricted expression of cadherin-7 in neural epithelium. *Dev Dyn* 240, 1756-68 (2011).
- Riel, J.J.G.v. Identification of epigenomic patterns to annotate regulatory elements in the human genome. *Masters thesis, Utrecht University* (2014).
- Roy, S. *et al.* Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* 330, 1787-97 (2010).
- Schaeffer, H.J., Forstheofel, N.R. & Cushman, J.C. Identification of enhancer and silencer regions involved in salt-responsive expression of Crassulacean acid metabolism (CAM) genes in the facultative halophyte *Mesembryanthemum crystallinum*. *Plant Mol Biol* 28, 205-18 (1995).
- Simpson, J., Schell, J., Montagu, M.V. & Herrera-Estrella, L. Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties. *Nature* 323, 551-554 (1986).
- Stathopoulos, A. & Levine, M. Localized repressors delineate the neurogenic ectoderm in the early Drosophila embryo. *Dev Biol* 280, 482-93 (2005).
- Stroebele, E., and Erives, A. (2016). Integration of Orthogonal Signaling by the Notch and Dpp Pathways in Drosophila. *Genetics* 203, 219–240.
- Zhu, Q. and Halfon, M. (2009). Complex organizational structure of the genome revealed by genome-wide analysis of single and alternative promoters in *Drosophila melanogaster*. *BMC Genomics*, 10(1), p.9.

# Annexes and supplementary tables

## Annex 1: full sequence of the pCD8 plasmid

>pCD8\_plasmid\_full\_sequence (11,393 bp)

```
ctcgcgcgtttcggatgacggtgaaaacctctgacacatgcagctcccgAGACGGTCACAGCTTGTCTGTAAGCGGA
TGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGGCTG
GCTTA ACTATGCGGCATCAGAGCAGATTG TACTGAGAGTGCACCATATGCGGTGTGAAATA
CCGCACCGAATCGCGCGGAACTAACGACAGTCGCTCCAAGGTCGTCGAACAAAAGGTGAA
TGTGTTGCGGAGAGCGGGTGGGAGACAGCGAAAGAGCAACTACGAAACGTGGTGTGGTGG
AGGTGAATTATGAAGAGGGCGCGGATTTGAAAAGTATGTATATAAAAAATATATCCCGG
TGTTTTATGTAGCGATAAACGAGTTTTTGTATGTAAGGTATGCAGGTGTGTAAGTCTTTTTGGT
TAGAAGACAAATCCAAAGTCTACTTGTGGGGATGTTTCGAAGGGGAAATACTTGTATTCTAT
AGGTCATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGGGCAATAA
ACAGTAAACACGATGGTAATAATGGTAAAAAAAAAAAAACAAGCAGTTATTTCCGATATAT
GTCGGCTACTCCTTGCCTCGGGCCCGAAGTCTTAGAGCCAGATATGCGAGCACCCGGAAGC
TCACGATGAGAATGGCCAGACCCACGTAAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAA
GCGTCTCCAGGATGACCTTGCCCGAAGTGGGGCACGTTGGTGTTCGACGATGTGCAGCTAAT
TTCGCCCCGCTCCACGTCCGCCATTGGTTAATCAGCAGACCCTCGTTGGCGTAACGGAAC
CATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAAGAAG
CCGCCAAAGAGCAGGAATGGTATGATAACCGGCGGACCCACAGACAGCGCCATCGAGGTC
GAGGAGCTGGCGCAGGATATTAGATATCCGAAGGACGTTGACACATTGGCCACCAGAGTG
ACCAGCGCCAGGCAGTTGAAGAAGTGCAGCACTCCGGCCCCGAGTCCGATCATCGGATAG
GCAATCGCCGTGAAGACCAGTGGCACTGTGAGAAAAAGCGGCAATTCGGCAATCGTTTTG
CCCAGAAAGTATGTGTCACAGCGATAAAGTCGACTTCGGGCCCTCCCTCATAAAAACTGGCA
GCTCTGAGGTGAACACCTAAATCGAATCGATTATTAGAAAGTTAGTAAATTATTGAAATG
CAAATGTATTCTAAACATGACTTACATTTATCGTGGCAAAGACGTTTTGAAAGGTCATGTT
GGTCAGGAAGAGGAAGATGGCTCCGTTGATATTCATCACACCCACTTGCCTGAGTTGTTGG
CCCAAAAAGATGAGGCCAATCAAGATGGCAACCATCTGCAAATTAATAATGTTACTCGCAT
CTCATTAATATTCGCGAGTTAAATGAAATTTATTTATCTTCTGCAAACTATAAACTATACA
TCTCATTGAAAAAACTAAGAAGGGTGTGGAATCAGGCAATTCTATCTAAAATCTAGCGA
ATTTGTTTCCAAGAATTGTAAGCGTTATATCATTGTGTTTCCACTGGAACCACTCACCGTTGT
CTGAATAAGTCGCACTTTTACGAGGAGTGGTTCCTTGAGCACCGACAGCCAGGATCGCCAC
AGGACCGCCCGAAGTGCATGAACCAGGTGGCCTTGTAGGTGTACCCATTCTCCGGCTGCT
CCAGTGGCTTCTCCAGATTTTTGGTGGCCAACAAGTGTCCATATCCCGGGCTACTTTGCTA
ATGGCAAATTTGTCGCATATCTTGGCGATCCGATCACGGGACTCGATCTCCCGTCCGGGCA
CAACGGCCAACACCTGTACGTAAGTCCGCCGATTGTAGTTGGTAGGACACTGGGCAC
CCACGCTGGATAGGAGTTGAGATGTTATGTAATACTAGATACCCTTAATAAACACATCGAA
CTCACTAGGAAAAGAAGTCGACGGCTTCGCTGGGAGTGCCCAAGAAAGCTACCCTGCCCT
CGGCCATCAGAAGGATCTTGTCAAAGAGCTCAAACAGCTCGGAAGACGGCTGATGAATGG
TCAGGATGACGGTCTTGCCCTTCTGCGACAGCTTCTTCAGCACCTGGACGACGCTGTGGGC
GGTAAAGGAGTCCAGTCCGAGGTTGGGCTCATCGCAGATCAGAAGCGGCGGATCGGTTAG
AGCCTCGGAGGCGAATGCCAGACGCTTCTTTCTCCGCCGACAGACCTTTCACCCTGCCG
GGCACACCGATGATCGTGTGCTGACATTTGCTGAGCGAAAGCTCCTGGATCACCTGATCCA
CGCGGGCCACTCGCTGCCGATAGGTGAGATGTCGTGGCATCCGCACCATGGCTTGGAAAAT
CAGGTGTTCCCTGGCCGTTAGGGAGCCGATAAAGAGGTCATCCTGCTGGACATAGGCGCAC
CTGGCCTGCATCTCCTTGGCGTCCACAGGTTGGCCATTGAGCAGTCGCATCCCGGATGGCG
ATACTTGGATGCCCTGCGGCGATCGAAAGGCAAGGGCATTGAGCAGGGTCGCTTTCCGGC
ACCGGAACTGCCCATCACGGCCAAAAGTTCGCCCGGATAGGCCACGCCGCAAACTGAGTT
TCAAATTGGTAATTGGACCCTTTATTAAGATTTACACAGATCAGCCGACTGCGAATAGAA
ACTCACCGTTCTTGGACAAATGTTTCCTGGGCGCCGGTATGTGTCGCTCGTTGCAGAATAGT
CCGCGTGTCCGTTGACCAGCTGCCGCCATCCGGAGCCCGGCTGATTGACCGCCCCAAAGA
TGTCCATATTGTGCCAGGCATAGGTGAGGTTCTCGGCTAGTTGGCCGCTCCCTGAACCGGA
GTCCTCCGGCGGACTGGGTGGCCGGAGCGTGCCGTAGTTTTTGGCCTGCCCGAAGCCCTGG
TTAATGCAGCTCTGCGAAGCCGCTCCGCTGTCACCCTGCAATGATAGGGGATCTCAAATAT
```

CAACTACAAGCGTTATGCTCATCTAACCCTGAACAAAAAGTACCCCGAAGTATCCTACGAA  
GTAGGTTTATACTTTTATTTATTTTTTGTGCATCTAGGATCAGCTTAAAATATCTGGTTGTTA  
TATTTTTTGTAAAAAAGAATATAGTCGAAAATGAATGCCCTTAGATGTCTTGATCATGATAT  
GATCTCAAAAATTGTCTTATATAGCGAGAACAGCTACCAGAATAATCTGTTTCGTGTCACT  
ATTTGTTTGTGCAATTGCGGTTTGGGATTTTTGTGGGTTCGAGTTCTCACGCCGAGACAAT  
TTGATGTTGCAATCGCAGTTCCTATAGATCAAGTGAACCTAAGATGTATGCACATGTACTA  
CTCACATTGTTTCAAGATGCTCGGCAGATGGGTGTTTGTCTGCCCTCCGCGAATTAATAGCTCCTG  
ATCCTCTTGGCCATTGCCGGGATTTTTCACACTTTCCCTGCTTACCCACCCAAAACCAAT  
CACCACCCCAATCACTCAAAAAACAACAAAAATAAGAAGCGAGAGGAGTTTTGGCACAG  
CACTTTGTGTTAATTGATGGCGTAAACCGCTTGGAGCTTCGTACGAAACCGCTGACAAA  
ATGCAACTGAAGGCGGACATTGACGCTACGTAACGCTACAAACGGTGGCGAAAGAGATAG  
CGGACGCAGCGGCGAAAGAGACGGCGATATTTCTGTGGACAGAGAAGGAGGCAAACAGC  
GCTGACTTTGAGTGGAATGTCATTTTGTAGTGAGAGGTAATCGAAAGAACCTGGTACATCAA  
ATACCCTTGGATCGAAGTAAATTTAAAACCTGATCAGATAAGTTCAATGATATCCAGTGCAG  
TAAAAAATAAATGTTTTTTTTATCTACTTTCCGCAAAAATGGGTTTTATTAACCTACATA  
CATACTAGAATTATCACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAATAATGATATA  
AATATCAATATATTAATTAGATTTTGCATAAAAAACAGACTACATAATACTGTAAAACAC  
AACATATCCAGTCACTATGGCGGCCGCATTAGGCACCCAGGCTTTACACTTTATGCTTCC  
GGCTCGTATAATGTGTGGATTTTGTAGTTAGGATCCGTCGAGATTTTCAGGAGCTAAGGAAG  
CTAAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAA  
AGAACATTTTGTAGGCATTTTCACTGAGTTGCTCAATGTACCTATAACCAGACCGTTTCACTG  
GATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTA  
TTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGG  
TGAGCTGGTGATATGGGATAGTGTTCACCCTTGTACACCGTTTTCCATGAGCAAACCTGAA  
ACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTC  
GCAAGATGTGGCGTGTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAAT  
ATGTTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAA  
TATGGACAACCTTCTTCGCCCCCGTTTTACCATGGGCAAATATTATACGCAAGGCGACAAG  
GTGCTGATGCCGCTGGCGATTCAGGTTTATCATGCGTTTGTGATGGCTTCCATGTCCGGCAG  
AATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAACGCGTGG  
ATCCGGCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATA  
AGAATATATACTGATATGTATACCCGAAGTATGTCAAAAAGAGGGTATGCTATGAAGCAGC  
GTATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAAT  
ATCTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCTGCTGCGTGCCGAACGCT  
GGAAAGCGGAAAATCAGGAAGGGATGGCTGAGGTTCGCCCGTTTATTGAAATGAACGGCT  
CTTTTGTGACGAGAACAGGGGCTGGTGAAATGCAGTTTAAAGGTTTACACCTATAAAAGAG  
AGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCGGGCGACG  
GATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTGATGATAAAGTCTCCCGTGAACCTTAC  
CCGGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTG  
CCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAA  
AACGCCATTAACCTGATGTTCTGGGGAATATAAATGTGAGGCTCCCTTATACACAGCCAGT  
CTGCAGGTCGACCATAGTACTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTT  
ATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTTCACTTTCTT  
GTACAAAGTGGTGATAATTCGAGCTCGGTACCCGGGGATCCGCGGCCGCTACGTATCTAGC  
GAGCGCCGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAG  
CAAAGTGAACACGTCGCTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCT  
AAACAATCTGCAGTAAAGTGCAAGTTAAAGTGAATCAATTAAGTAACCAGCAACCAAG  
TAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACAAGAAGAGAAC  
TCTGAATAGGGAATTGGGAATTGACACCATGGCCTTACCAGTGACCGCCTTGCTCCTGCCG  
CTGGCCTTGTGCTCCACGCCGCGAGGCCGAGCCAGTTCCGGGTGTGCGCCGCTGGATCGGA  
CCTGGAACCTGGGCGAGACAGTGGAGCTGAAGTGCCAGGTGCTGCTGTCCAACCCGACGT  
CGGGCTGCTCGTGGCTCTTCCAGCCGCGCGGGCGCCGCGCCAGTCCCACCTTCTCCTATAC  
CTCTCCAAAACAAGCCCAAGGCGGCCGAGGGGCTGGACACCCAGCGGTTCTCGGGCAAG  
AGGTTGGGGGACACCTTCGTCCTCACCTGAGCGACTTCCGCCGAGAGAACCAGGGGCTGCT  
ATTTCTGCTCGGCCCTGAGCAACTCCATCATGTACTTCCAGCCACTTCGTGCCGGTCTTCTG  
CCAGCGAAGCCCACCACGACGCCAGCGCCGCGACCACCAACACCGGCGCCACCATCGCG  
TCGCAGCCCCTGTCCCTGCGCCAGAGGCGTGCCGGCCAGCGGCGGGGGGCGCAGTGCAC

ACGAGGGGGCTGGACTTCGCCTGTGATATCTACATCTGGGCGCCCTTGGCCGGGACTTGTG  
GGGTCCTTCTCCTGTCACTGGTTATCACCCCTTTACTGCAACCACAGGAACCGAAGACGTGTT  
TGCAAATGTCCCCGGCCTGTGGTCAAATCGGGAGACAAGCCCAGCCTTTCGGCGAGATACG  
TCTAGGCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAAATTGGACAAA  
CTACCTACAGAGATTTAAAGCTCTAAGGTAATAATAAAATTTTTAAGTGTATAATGTGTTA  
AACTACTGATTCTAATTGTTTGTGTATTTAGATTCCAACCTATGGAAGTGTGAATGGGAG  
CAGTGGTGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTG  
ATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAAGGTAG  
AAGACCCCAAGGACTTTCCCTCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAAT  
AGAAGTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATAACAAGA  
AAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACAT  
ACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAAT  
TGTGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCC  
TTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAC  
CTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTT  
TATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAAATAAAGCA  
TTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGG  
ATCGGGCGAGCTCGAATTGGTCGACCTGCAGCCAAGCTTTGCGTACTCGCAAATTATTA  
AATAAACTTTAAAAATAATTTTCGTCTAATTAATATTATGAGTTAATTCAAACCCACGGA  
CATGCTAAGGGTTAATCAACAATCATATCGCTGTCTCACTCAGACTCAATACGACACTCAG  
AATACTATTCCTTTCCTCGCACTTATTGCAAGCATAACGTTAAGTGGATGTCTCTTGCCGAC  
GGGACCACCTTATGTTATTTATCATGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTC  
GCATATCTGGCTCTAAGACTTCGGGCCCCGACGCAAGGAGTAGCCGACATATATCCGAAATA  
ACTGCTTGTTTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTTTATTGCCCCCTCAAAA  
AGCTAATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTATAGAATACAAGTATT  
TCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAAGACTTACACA  
CCTGCATACCTTACATCAAAAACCTCGTTTATCGCTACATAAAAACACCGGGATATATTTTTTA  
TATACATACTTTTCAAATCGCGCGCCCTCTTATAATTCACCTCCACCACACCACGTTTCGT  
AGTTGCTCTTTCGCTGTCTCCACCCGCTCTCCGCAACACATTCACCTTTTGTTCGACGACC  
TTGGAGCGACTGTCGTTAGTTCGCGCGGATTCCGGTTCGCTCAAATGGTTCGAGTGGTTCAT  
TTCGCTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTGCTCCAATATATT  
TGTATATATTTCCCTCACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATT  
TTTTGTGACCTGTTCCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTT  
GTTCCCTGTGTAGATGCATCTCAAAAAAATGGTGGGCATAATAGTGTGTTTATATATATCA  
AAAATAACAACATAATAATAAGAATACATTTAATTTAGAAAATGCTTGGATTTCACTGGA  
ACTAGAATTAATTCGGCTGCTGCTCTAAACGACGCATTTTCGTAATCCAAAGTACGAATTTTT  
TCCCTCAAGCTCTTATTTTTTACTATTCAAACCTTACTCTGTTTGTGTACTCCCACTGGTATAGC  
CTTCTTTTATCTTTTCTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTC  
GCCTCCTTTTAAATGTCTGACCTTTTGCAGGTGCAGCCTTCCACTGCGAATCATTAAAGTGG  
GTATCACAAATTTGGGAGTTTTACCAAGGCTGCACCCAAGGCTCTGCTCCCACAATTTTCT  
CTTAATAGCACACTTCGGCACGTGAATTAATTTTACTCCAGTCACAGCTTTGCAGCAAAATT  
TGCAATATTTCAATTTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAAATTCGTCC  
GCACACAACCTTTCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAA  
GCTTCGGCTATCGACGGGACCACCTTATGTTATTTATCATGAGGCGAGACCCACGTAGTCC  
AGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACTGGGG  
CACGTGGTGTTCGACGATGTGCAGCTAATTTCCGGGCTCCACGTCCGCCCATTGGTTAAT  
CAGCAGACCCTCGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATAC  
TGGCACCGAGCCCGAGTTCAAGAAGAAGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAG  
CATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC  
CAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG  
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT  
ATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCA  
GCCCCACCGCTGCGCCTTATCCGGTAACATCGTCTTGTAGTCCAACCCGGTAAGACACGAC  
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT  
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA  
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA

ACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAA  
AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA  
ACTCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTA  
AATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTT  
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTT  
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG  
CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCC  
AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATT  
AATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTG  
CCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGT  
TCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCT  
TCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGC  
AGCACTGCATAATTCTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT  
ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC  
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACG  
TTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCC  
ACTCGTGCACCCAACCTGATCTTCAGCATCTTTTACTTTACCAGCGTTTCTGGGTGAGCAAA  
AACAGGAAGGCAAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC  
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA  
TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA  
AAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCG  
TATCACGAGGCCCTTTCGTctcgcgcttcgggatgacggtgaaaacctctgacacatgcagctccgg

## Annex 2: Protocol for positive panning from Drosophila embryos

**This protocol is an example of cell purification for CD8+ cells from the dpp\_VRR:CD8 line. In this example, 2.30.E+07 cells are viable, and 6.90E+06 cells are estimated to be CD8+.**

Overnight embryo collection:

- 1) Maintain flies in population cages at 25° C, feeding them using yeast paste streaked molasses plates.
- 2) Change molasses plates (w/ yeast) one time a day for 2-3 days.
- 3) Change plate, allow flies to prelay for 2 hour on fresh food.
- 4) After 2 hours prelay, change with the new plate with yeast paste, allow flies to lay for 2hours.
- 5) Remove laying plate and allow it to age for 5.5 hour at 25° C or 10 hours at 18°C.

Cell preparation:

- 6) Wash eggs and yeast from plates using water from a squeeze bottle, loosen eggs with camel hair brush and pour the eggs to the 70um stainer ( weight measured), and rinse the materials in the strainer.
- 7) Fill the strainer again with 50% (v/v)bleach to cover the eggs and dechorionate the eggs for 5 minutes.
- 8) During the dechoronation, fill the dounce (VWR 62400-620) with 3.5 mL Schneider medium and 3.5 mL Isolation buffer (see recipe at the end of protocol), and keep on ice.
- 9) After the dechoronation, wash the bleach completely from the eggs in the strainer with water.
- 10) Use a Kimwipe to blot dry the strainer from outside and measure the weight to obtain the wet weight of eggs.
- 11) Brush the eggs into the dounce. Use loose pestle (clearance 0.0035-0.005 inch), gently but firmly dounce to the bottom; give 7 strokes.
- 12) Transfer dounced materials into a 15mL conical centrifuge tube. Spin 40g (500 rpm in our desktop centrifuge at 4°) for 5 min to bring down the tissue cell debris, clumps and vitelline membranes. Single cells and yolk are in the supernatant.
- 13) Transfer supernatant to a clean tube, add 7mL of Isolation Buffer and spin 380g (1500 rpm in our desktop centrifuge) for 10 minutes to bring down single cells and discard supernatant. Repeat once to clean the cells.

Meanwhile, wash beads:

- Resuspend the beads in the vial (vortex >30 sec)
- Transfer the desired volume of beads in 15mL or 50mL tube for all experiments
- Add the same volume of Isolation Buffer or at least 1mL vortex briefly to resuspend the beads
- Place the tube in the magnet for 1min and discard supernatant
- Remove tube from magnet and resuspend the washed beads in the same volume of Isolation Buffer as the initial volume of beads.

- 14) Resuspend the cells in 0.230mL Isolation Buffer to have a concentration in the tube of 10<sup>6</sup> CD8+ cells per mL in each tube.

15) Isolate Cells:

1. Add 34.5  $\mu$ L FlowComp™ Human CD8 Antibody.
2. Mix well and incubate for 10 min on ice.
3. Fill tube up to 15mL with Isolation buffer and centrifuge for 10min at 380g (1500rpm)
4. remove supernatant that contains excess of anti-CD8 antibody.
5. Add 34.5  $\mu$ L resuspended FlowComp™ Dynabeads® and mix well by vortexing.
6. Incubate for 20 min on ice.
7. Place tube in magnet for 3 min, carefully remove supernatant containing the CD8- cells.
8. Add 0.9 mL Isolation Buffer, mix well (or vortex 2–3 sec) and place the tube in the magnet for minimum 3 min.
9. While the tube is still in the magnet, carefully remove the supernatant.
10. Repeat steps 8-9 twice to wash the bead-bound CD8+ cells. These steps are critical to obtain a high purity of isolated cells.

16) Release cells:

1. Resuspend in 0.46 mL Elution buffer and pipette 3-4 times
2. Incubate for 10 min at RT or for 20 min at 4°C under rolling and tilting
3. Pipet 10 times to release the cells and place in a magnet for 1min. Avoid foaming.
4. Transfer supernatant containing the bead-free cells to a new tube and again place on the magnet for 1min to remove any residual beads.
5. Transfer once again the supernatant containing the bead free cells to a new tube.
6. Add 1.8 mL of Isolation Buffer and centrifuge for 10min at 380g (1500rpm).
7. Discard supernatant and resuspend pellet in appropriate staining solution.

17) Shake in a covered ice bucket at 100 rpm for 10 minutes.

18) Add 1 mL (or more to have equal volumes) of primary cell culture medium (8%FBS in Schneider Medium) to each tube. Spin 380g (1500 rpm in our desktop centrifuge) for 10 minutes to bring down single cells and discard supernatant.

Wash once: resuspend in 1 mL culture medium, spin down again and discard supernatant.

Resuspend in 200ul - 1 mL culture medium and filter twice through Nytex into 2063 tubes

18) IF SORTING: Prepare tubes (2063) for receiving sorted cells by marking each tube 1 cm from the bottom and filling with 4 mL of culture medium.

## Annex 3 : gRNA list for bifunctional element knockout

List of gRNAs designed for bifunctional element knockout, along with the PCR primers for PCR validation of said knockout.

<b>Name</b>	<b>oligo type</b>	<b>sequence (5'-3')</b>	<b>PAM</b>
E_0_12h_dCtBP7667.region_3084	gRNA1	ATTAATTTGAGTCGAAATAT	TGG
	gRNA2	GCGGTGGTCTGGTGGTTAGG	TGG
	gRNA3	GAAAGAATCGGTGGAGAAAA	TGG
	PCRfwd	GTGAAAAAGCAAATTTGGGGGT	
	PCRrev	ATGGTTGCTCACAGGCTCAA	
e_coreAbdominalCRE	gRNA1	AATCAAATCACAGGGACTTT	TGG
	gRNA2	TACTTTTACCCCTCAAGTAA	CGG
	PCRfwd	TAATGGATAGGCGTGGTGGC	
	PCRrev	CCAAGCCAAATTCACGGAGG	
gsb_fragIV	gRNA1	CTCAGCCTTTGGACCCTAAG	CGG
	gRNA2	ATGCGTATCTCCGGACTGA	CGG
	PCRfwd	CCCACACTATGGCTAGAATAGCA	
	PCRrev	ACGATTGCCTCTCTTAACTCTCT	
hkb_0.6kbRIRV	gRNA1	CTAAAAGATATCTGCTTTCT	AGG
	gRNA2	CRACTGAAGTTTAGTTACGC	TGG
	PCRfwd	TCCCACGATAGGATTAGTAGTGT	
	PCRrev	TGTA AACATGCATTGGACATGCT	
vnd_NEE	gRNA1	AGATTAGGAGAATTAATTAA	TGG
	gRNA2	GTGTCAATCGTGTGAATGGC	CGG
	PCRfwd	ATGACCTCTGTTTTGTGGGC	
	PCRrev	AAAAGCATACCAAAGATGAGTTGT	

## Supplementary table 1

Genomic coordinates of the first library elements screened by silencer-FACS-Seq. The primers for the negative controls from the *E.coli* genome are shown bottom of this table.

<b>library element name</b>	<b>coordinates (dm3)</b>	<b>coordinates (dm6)</b>
18w_1625	chr2R:15990424-15991391	chr2R:20102919-20103886
ac_0.8	chrX:263085-264110	chrX:369052-370077
Acp65Aa_172	chr3L:6145206-6146175	chr3L:6152106-6153075
amosB	chr2L:18597902-18598899	chr2L:18597902-18598899
amosC	chr2L:18597063-18597969	chr2L:18597063-18597969
amosD	chr2L:18595204-18596246	chr2L:18595204-18596246
Ance_race_533	chr2L:13904341-13905301	chr2L:13904341-13905301
ase_CRM19	chrX:354220-355242	chrX:460187-461209
ato_RE	chr3R:4099748-4100734	chr3R:8274026-8275012
ATP7_5DAG4	chrX:11748537-11749464	chrX:11854504-11855431
bab1_dimorphic_element	chr3L:1084698-1085668	chr3L:1084698-1085668
bcd_49bg-Z	chr3R:2584166-2585243	chr3R:6758444-6759521
btd_R-Ss	chrX:9583710-9584690	chrX:9689677-9690657
btd_Ss-Bg	chrX:9584428-9585527	chrX:9690395-9691494
btl_P[B23]	chr3L:14069010-14069998	chr3L:14075910-14076898
btl_P[B4]	chr3L:14069500-14070582	chr3L:14076400-14077482
C15_350-2	chr3R:17331233-17332270	chr3R:21505511-21506548
cas_csc-1	chr3R:1554785-1555690	chr3R:5729063-5729968
cas_csc-2	chr3R:1553104-1554153	chr3R:5727382-5728431
cas_csc-3	chr3R:1552157-1553136	chr3R:5726435-5727414
cas_csc-5a	chr3R:1548621-1549584	chr3R:5722899-5723862
cas_csc-5c	chr3R:1546450-1547427	chr3R:5720728-5721705
cas_csc-6	chr3R:1545675-1546770	chr3R:5719953-5721048
cas_csc-7a	chr3R:1544440-1545364	chr3R:5718718-5719642

cas_csc-8	chr3R:1541898-1542918	chr3R:5716176-5717196
CG12374_CRM41	chr2R:8655932-8656979	chr2R:12768427-12769474
CG13196_1kb_5'	chr2R:7573891-7574944	chr2R:11686386-11687439
CG13333_link_5'	chr2R:9442853-9443775	chr2R:13555348-13556270
CG16778_tkr-15	chr2R:20988969-20990063	chr2R:25101446-25102540
CG17230_CRM4	chr3R:7297085-7298176	chr3R:11471363-11472454
CG34347_CG11339_EVIII	chr3R:27134344-27135381	chr3R:31308622-31309659
CG3492_CRM17	chr2R:20351984-20353074	chr2R:24464461-24465551
CG42342_3436	chr3R:12377681-12378775	chr3R:16551959-16553053
CG7229_CG7229	chr3L:6778896-6779980	chr3L:6785796-6786880
CG7229_CRM15	chr2R:15150758-15151840	chr2R:19263253-19264335
CG7458_upstreamCRM	chr3L:21947443-21948420	chr3L:21954343-21955320
CG7722_CRM28	chr2R:6824079-6825095	chr2R:10936574-10937590
CG8193_proPO45crystal_cell	chr2R:4932059-4932992	chr2R:9044554-9045487
CG9363_CG9363CRM1	chr3R:5284681-5285769	chr3R:9458959-9460047
CG9571_O-E	chrX:19986450-19987538	chrX:20092417-20093505
cpo_cpoCRM6	chr3R:13777548-13778572	chr3R:17951826-17952850
Cpr47Ee_CG13222-edge	chr2R:7151764-7152722	chr2R:11264259-11265217
crb_Lac-Z	chr3R:20122950-20123916	chr3R:24297228-24298194
CrebA_CrebA-770	chr3L:15523344-15524372	chr3L:15530244-15531272
Crz_380gal4	chr3R:10139892-10140852	chr3R:14314170-14315130
ct_340	chrX:7484874-7485847	chrX:7590841-7591814
ct_ct-3	chrX:7478189-7479230	chrX:7584156-7585197
ct_cutA2	chrX:7468216-7469313	chrX:7574183-7575280
ct_wingmargin_Guss	chrX:7424146-7425126	chrX:7530113-7531093
Cyp6g1_construct5	chr2R:8072880-8073942	chr2R:12185375-12186437
dac_3EE-390	chr2L:16461056-16462067	chr2L:16461056-16462067
dac_RE	chr2L:16465126-16466148	chr2L:16465126-16466148
dap_dap-BB	chr2R:5600855-5601828	chr2R:9713350-9714323
dap_dap-del	chr2R:5601755-5602723	chr2R:9714250-9715218

dap_dap3'2'	chr2R:5596925-5597954	chr2R:9709420-9710449
Ddc_-0.47	chr2L:19119819-19120797	chr2L:19119819-19120797
Ddc_distal_enhancer	chr2L:19120966-19122021	chr2L:19120966-19122021
Ddc_ET	chr2L:19119448-19120508	chr2L:19119448-19120508
Ddc_silencing_element	chr2L:19121524-19122424	chr2L:19121524-19122424
Def_prom	chr2R:5942027-5943010	chr2R:10054522-10055505
dj_dj-promoter	chr3R:2884201-2885169	chr3R:7058479-7059447
djl_-555-+43	chr3R:2885745-2886647	chr3R:7060023-7060925
DII_215	chr2R:20689355-20690348	chr2R:24801832-24802825
DII_304	chr2R:20690263-20691248	chr2R:24802740-24803725
dpn_dl	chr2R:4123455-4124497	chr2R:8235950-8236992
dpp_85.8MX	chr2L:2456545-2457501	chr2L:2456545-2457501
dpp_980-6	chr2L:2453724-2454821	chr2L:2453724-2454821
dpp_BS3.1	chr2L:2480501-2481487	chr2L:2480501-2481487
dpp_construct10	chr2L:2481008-2482099	chr2L:2481008-2482099
dpp_dpoho	chr2L:2482818-2483809	chr2L:2482818-2483809
dpp_P1delta4	chr2L:2453866-2454937	chr2L:2453866-2454937
dpp_VRR	chr2L:2455833-2456926	chr2L:2455833-2456926
dys_M269+dys_V345	chr3R:21925380-21926363	chr3R:26099658-26100641
dys_N283	chr3R:21926337-21927256	chr3R:26100615-26101534
e_A.1	chr3R:17067508-17068463	chr3R:21241786-21242741
e_A.2	chr3R:17067205-17068139	chr3R:21241483-21242417
e_A.3	chr3R:17066757-17067698	chr3R:21241035-21241976
e_A.5	chr3R:17066147-17067114	chr3R:21240425-21241392
e_coreAbdominalCRE	chr3R:17066418-17067406	chr3R:21240696-21241684
E(spl)_m8-0.46	chr3R:21865269-21866269	chr3R:26039547-26040547
Eip71CD_188	chr3L:15503755-15504823	chr3L:15510655-15511723
elav_construct_L	chrX:416654-417626	chrX:522621-523593
ems_elementIV	chr3R:9720474-9721431	chr3R:13894752-13895709
ems_ems_ARFE-subA	chr3R:9723441-9724427	chr3R:13897719-13898705

Est-6_D-511	chr3L:12181127-12182109	chr3L:12188027-12189009
ey_5D11	chr4:730601-731670	chr4:709975-711044
ey_UE0.8	chr4:721768-722862	chr4:701142-702236
ey_UE0.9	chr4:724459-725477	chr4:703833-704851
Fad2_oe1	chr3L:11015797-11016807	chr3L:11022697-11023707
Fas2_540bp_CRM	chrX:4086884-4087943	chrX:4192851-4193910
fkh_salivary_gland_enhancer	chr3R:24418581-24419666	chr3R:28592859-28593944
ftz_-669_-386	chr3R:2688785-2689687	chr3R:6863063-6863965
ftz_5'delta276	chr3R:2689711-2690794	chr3R:6863989-6865072
ftz_Rev413	chr3R:2683243-2684295	chr3R:6857521-6858573
gcm_+3.8_+4.5	chr2L:9577190-9578170	chr2L:9577190-9578170
gsb_fragIV	chr2R:20944064-20945045	chr2R:25056541-25057522
gsb_GLE	chr2R:20946182-20947206	chr2R:25058659-25059683
gt_CE8001	chrX:2324615-2325714	chrX:2430582-2431681
gt_gt1	chrX:2327867-2328788	chrX:2433834-2434755
h_302	chr3L:8662863-8663834	chr3L:8669763-8670734
h_h7AF	chr3L:8658167-8659238	chr3L:8665067-8666138
h_HHRE	chr3L:8700624-8701673	chr3L:8707524-8708573
h_stripe_6+2	chr3L:8659410-8660502	chr3L:8666310-8667402
h_stripe0	chr3L:8680028-8681110	chr3L:8686928-8688010
h_stripe1	chr3L:8663979-8664926	chr3L:8670879-8671826
h_stripe3_ET38	chr3L:8657181-8658199	chr3L:8664081-8665099
hb_0.7	chr3R:4519759-4520743	chr3R:8694037-8695021
hb_distal_minimal	chr3R:4524698-4525688	chr3R:8698976-8699966
hb_HG4-5	chr3R:4530675-4531772	chr3R:8704953-8706050
hb_HG4-6	chr3R:4529043-4530092	chr3R:8703321-8704370
hb_HG4-7	chr3R:4527834-4528862	chr3R:8702112-8703140
hb_lateDm1.0-lacZ	chr3R:4526286-4527384	chr3R:8700564-8701662
hb_matDm0.5-lacZ	chr3R:4522881-4523877	chr3R:8697159-8698155
hh_4075	chr3R:18968885-18969838	chr3R:23143163-23144116

hh_alpha_fragment_(ic-CRE)	chr3R:18970625-18971703	chr3R:23144903-23145981
hh_bar3L2	chr3R:18961205-18962168	chr3R:23135483-23136446
hh_hhf4F	chr3R:18963250-18964276	chr3R:23137528-23138554
hkb_0.6kbRIRV	chr3R:173849-174821	chr3R:4348127-4349099
HLHm5_m5-0.13	chr3R:21854885-21855903	chr3R:26029163-26030181
Hs6st_3748	chr3R:15820869-15821869	chr3R:19995147-19996147
Hsp26_nurse_cell_enhancer	chr3L:9370261-9371323	chr3L:9377161-9378223
kni_223	chr3L:20690490-20691542	chr3L:20697390-20698442
kni_AE20	chr3L:20694209-20695217	chr3L:20701109-20702117
kni_KD	chr3L:20689674-20690594	chr3L:20696574-20697494
kni_proximal_minimal	chr3L:20687475-20688458	chr3L:20694375-20695358
kni_reporter_fragment_EC	chr3L:20699549-20700619	chr3L:20706449-20707519
Kr_delBNc0.8HZ	chr2R:21110473-21111509	chr2R:25222950-25223986
Kr_KrMT	chr2R:21098808-21099895	chr2R:25211285-25212372
lab_1.0	chr3R:2491709-2492721	chr3R:6665987-6666999
lab_HZ550	chr3R:2506791-2507856	chr3R:6681069-6682134
Lip1_prom	chr2L:10701299-10702327	chr2L:10701299-10702327
lz_CrystalCellEnhancer1236-737	chrX:9177203-9178176	chrX:9283170-9284143
lz_LMEE	chrX:9180643-9181594	chrX:9286610-9287561
Mst84Db_-286_+154	chr3R:3190986-3192054	chr3R:7365264-7366332
nab_nab-1	chr3L:4152493-4153571	chr3L:4152493-4153571
nAcRbeta-64B_P-171	chr3L:4428570-4429577	chr3L:4428570-4429577
nerfin-1_fragment14	chr3L:908248-909263	chr3L:908248-909263
nerfin-1_fragment2	chr3L:903575-904640	chr3L:903575-904640
nerfin-1_fragment3	chr3L:904458-905370	chr3L:904458-905370
nerfin-1_fragment4	chr3L:904964-905875	chr3L:904964-905875
nerfin-1_fragment5	chr3L:906392-907315	chr3L:906392-907315
ninaE_proximal_promoter_region	chr3R:15713418-15714371	chr3R:19887696-19888649
nkd_8756	chr3L:19036341-19037409	chr3L:19043241-19044309
nkd_IntE_255	chr3L:19032720-19033620	chr3L:19039620-19040520

nkd_UpE2	chr3L:19047510-19048605	chr3L:19054410-19055505
nos_-708_+20	chr3R:14982247-14983189	chr3R:19156525-19157467
navy_CRM100	chr2R:20162432-20163459	chr2R:24274909-24275936
navy_CRM29	chr2R:20163780-20164851	chr2R:24276257-24277328
oc_intronic_distal	chrX:8536858-8537949	chrX:8642825-8643916
oc_oc7	chrX:8511921-8512824	chrX:8617888-8618791
oc_otd_EHE	chrX:8548837-8549929	chrX:8654804-8655896
oc_otd-186	chrX:8548281-8549211	chrX:8654248-8655178
oc_SBg	chrX:8547554-8548562	chrX:8653521-8654529
otp_C	chr2R:16786416-16787494	chr2R:20898911-20899989
otp_P	chr2R:16785528-16786561	chr2R:20898023-20899056
ovo_del-ap-del-5	chrX:4958967-4959932	chrX:5064934-5065899
ovo_E2	chrX:4914146-4915245	chrX:5020113-5021212
ovo_E3	chrX:4914971-4916045	chrX:5020938-5022012
ovo_E6A	chrX:4917122-4918140	chrX:5023089-5024107
ovo_lacZdel-ap-del-6	chrX:4957750-4958748	chrX:5063717-5064715
pdm2_CRM6	chr2L:12682595-12683690	chr2L:12682595-12683690
per_-603_-449	chrX:2578569-2579555	chrX:2684536-2685522
PH4alphaSG2_SG2-885	chr3R:26315025-26315977	chr3R:30489303-30490255
ple_995bp_wound_response_element	chr3L:6712587-6713662	chr3L:6719487-6720562
ple_WE1	chr3L:6715951-6716856	chr3L:6722851-6723756
pnr_P3	chr3R:11853793-11854843	chr3R:16028071-16029121
pnr_P4	chr3R:11854355-11855347	chr3R:16028633-16029625
Poxn_2	chr2R:11717859-11718948	chr2R:15830354-15831443
Poxn_9	chr2R:11722264-11723254	chr2R:15834759-15835749
prd_cc_repressor	chr2L:12086931-12087941	chr2L:12086931-12087941
prd_deltaQ	chr2L:12085023-12086072	chr2L:12085023-12086072
prd_P1_enhancer	chr2L:12087791-12088820	chr2L:12087791-12088820
prd_PMFE	chr2L:12077955-12078886	chr2L:12077955-12078886
prd_Pstripe_enhancer	chr2L:12088649-12089604	chr2L:12088649-12089604

prd_stripe1_enhancer	chr2L:12089685-12090630	chr2L:12089685-12090630
proPO-A1_crystal_cell	chr2R:13776907-13777945	chr2R:17889402-17890440
proPO-A1_F6	chr2R:13774263-13775214	chr2R:17886758-17887709
Rbp4_pPa104_69	chr3R:14080684-14081665	chr3R:18254962-18255943
repo_-1.1	chr3R:14060749-14061844	chr3R:18235027-18236122
repo_pBJ-111	chr3R:14058681-14059589	chr3R:18232959-18233867
repo_pBJ-145	chr3R:14059539-14060516	chr3R:18233817-18234794
Rh2_-443_+32	chr3R:14724289-14725308	chr3R:18898567-18899586
Rh3_promoter	chr3R:15907857-15908889	chr3R:20082135-20083167
Rh5_promoter	chr2L:12007559-12008510	chr2L:12007559-12008510
Rh6_-555_+121	chr3R:11308239-11309249	chr3R:15482517-15483527
run_neural_6GB	chrX:20559288-20560378	chrX:20688261-20689351
sc_1.1	chrX:289125-290204	chrX:395092-396171
sc_CRM39	chrX:286648-287747	chrX:392615-393714
sens_sensCRM3	chr3L:13395412-13396397	chr3L:13402312-13403297
sev_minimal_enhancer	chrX:10973249-10974269	chrX:11079216-11080236
sev_sev_prom	chrX:10980221-10981320	chrX:11086188-11087287
Sgs4_OP SL	chrX:3143420-3144347	chrX:3249387-3250314
sim_mesectoderm	chr3R:8895383-8896466	chr3R:13069661-13070744
sim_st10	chr3R:8885647-8886704	chr3R:13059925-13060982
slp2_i4753	chr2L:3830134-3831049	chr2L:3830134-3831049
sna_0.25	chr2L:15478141-15479224	chr2L:15478141-15479224
sog_broad_lateral_neurogenic_ectoderm	chrX:15518390-15519344	chrX:15624357-15625311
sog_shadow	chrX:15540621-15541615	chrX:15646588-15647582
Sox15_regionA	chr2R:10098445-10099413	chr2R:14210940-14211908
Sox15_regionB	chr2R:10097816-10098825	chr2R:14210311-14211320
Sox15_regionC	chr2R:10097224-10098222	chr2R:14209719-14210717
SoxN_565	chr2L:8811551-8812587	chr2L:8811551-8812587
SoxN_5830	chr2L:8841052-8842145	chr2L:8841052-8842145
sphinx_1067bp_5'_fragment	chr4:994776-995867	chr4:974150-975241

sqz_sqz-11	chr3R:15000126-15001211	chr3R:19174404-19175489
Sry-alpha_CAHBG	chr3R:25866256-25867179	chr3R:30040534-30041457
ss_E2.0_522	chr3R:12242388-12243423	chr3R:16416666-16417701
ss_E2.0_531	chr3R:12243575-12244533	chr3R:16417853-16418811
ss_P732	chr3R:12218328-12219379	chr3R:16392606-16393657
sv_paxD	chr4:1106414-1107513	chr4:1085788-1086887
sv_SME	chr4:1116254-1117156	chr4:1095628-1096530
ths_Neu4_early_embryonic_enhancer	chr2R:7681709-7682675	chr2R:11794204-11795170
tll_D3	chr3R:26677526-26678474	chr3R:30851804-30852752
tll_K11	chr3R:26674823-26675766	chr3R:30849101-30850044
tll_K7	chr3R:26672825-26673797	chr3R:30847103-30848075
tll_O-E	chr3R:26681267-26682305	chr3R:30855545-30856583
toy_EEP	chr4:1001378-1002361	chr4:980752-981735
trh_trh24	chr3L:374147-375060	chr3L:374147-375060
trh_trh45	chr3L:394619-395621	chr3L:394619-395621
tup_dorsalectoderm	chr2L:18874963-18875896	chr2L:18874963-18875896
vas_96bpEnhancer	chr2L:15061100-15062127	chr2L:15061100-15062127
vas_construct16	chr2L:15073693-15074684	chr2L:15073693-15074684
vg_minimal_boundary_enhancer	chr2R:8776000-8777037	chr2R:12888495-12889532
vg_quadrant_enhancer	chr2R:8783523-8784538	chr2R:12896018-12897033
Vm26Aa_C2	chr2L:5960319-5961345	chr2L:5960319-5961345
Vm32E_-348_-39	chr2L:11171600-11172500	chr2L:11171600-11172500
vnd_743	chrX:486746-487752	chrX:592713-593719
vnd_NEE	chrX:486301-487394	chrX:592268-593361
vvl_484-5prime	chr3L:6782745-6783769	chr3L:6789645-6790669
vvl_587dfr	chr3L:6777942-6779032	chr3L:6784842-6785932
vvl_vvl0.9	chr3L:6758464-6759483	chr3L:6765364-6766383
vvl_vvl1+2	chr3L:6757667-6758693	chr3L:6764567-6765593
vvl_vvlds1.0	chr3L:6816974-6818073	chr3L:6823874-6824973
y_BE1-2	chrX:248439-249451	chrX:354406-355418

y\_BE3  
y\_wing  
Yp1\_oe2  
Yp1\_structure\_8E  
Z600\_z600-lacZ  
Sur\_dSurEN3-SS  
CAD2\_hand\_1  
Hand\_HCH  
CAD2\_Meso-CRM-6028  
CAD2\_sna\_2  
CAD2\_Meso-CRM-6225  
CadN\_Lac-Z  
mib2\_FCenhancer  
Acon\_Lac-Z  
slp1\_5303  
Tg\_cardiac\_enhancer  
numb\_5870  
CG4364\_upstreamCRM  
bib\_5924  
CAD2\_Meso-CRM-965  
CG9416\_GFP  
Atet\_5338  
Kr\_HBg0.6HZ  
Kr\_H\_I  
eve\_MHE  
Ndg\_FCenhancer  
CAD2\_pyr  
CAD2\_Mdr49  
CG32111\_8084  
CAD2\_Meso-CRM-4726

chrX:249145-250233  
chrX:247798-248732  
chrX:9945625-9946666  
chrX:9946630-9947672  
chr3L:15501282-15502351  
chr2L:10197218-10198280  
chr2L:10292609-10293692  
chr2L:10293308-10294379  
chr2L:14266991-14267912  
chr2L:15481938-15482856  
chr2L:17221775-17222718  
chr2L:17662868-17663886  
chr2L:19036808-19037782  
chr2L:21168779-21169764  
chr2L:3827700-3828650  
chr2L:8023372-8024455  
chr2L:9447742-9448771  
chr2L:9710879-9711857  
chr2L:9988389-9989430  
chr2R:13644082-13645146  
chr2R:15258605-15259535  
chr2R:20932312-20933348  
chr2R:21110723-21111783  
chr2R:21113526-21114590  
chr2R:5872394-5873407  
chr2R:6203047-6203992  
chr2R:7620398-7621302  
chr2R:8833522-8834532  
chr3L:12641801-12642834  
chr3L:16760731-16761829

chrX:355112-356200  
chrX:353765-354699  
chrX:10051592-10052633  
chrX:10052597-10053639  
chr3L:15508182-15509251  
chr2L:10197218-10198280  
chr2L:10292609-10293692  
chr2L:10293308-10294379  
chr2L:14266991-14267912  
chr2L:15481938-15482856  
chr2L:17221775-17222718  
chr2L:17662868-17663886  
chr2L:19036808-19037782  
chr2L:21168779-21169764  
chr2L:3827700-3828650  
chr2L:8023372-8024455  
chr2L:9447742-9448771  
chr2L:9710879-9711857  
chr2L:9988389-9989430  
chr2R:17756577-17757641  
chr2R:19371100-19372030  
chr2R:25044789-25045825  
chr2R:25223200-25224260  
chr2R:25226003-25227067  
chr2R:9984889-9985902  
chr2R:10315542-10316487  
chr2R:11732893-11733797  
chr2R:12946017-12947027  
chr3L:12648701-12649734  
chr3L:16767631-16768729

fz2_Lac-Z	chr3L:19139406-19140383	chr3L:19146306-19147283
CAD2_Meso-CRM-4906	chr3L:19184809-19185753	chr3L:19191709-19192653
CAD2_Ket-1	chr3L:2089382-2090424	chr3L:2089382-2090424
CAD2_Ket-2	chr3L:2114510-2115516	chr3L:2114510-2115516
sls_Ket-3_Lac-Z	chr3L:2118030-2119024	chr3L:2118030-2119024
sfl_Lac-Z	chr3L:6515183-6516181	chr3L:6522083-6523081
Doc3_7731	chr3L:8996415-8997429	chr3L:9003315-9004329
CAD2_pnr	chr3R:11850973-11851909	chr3R:16025251-16026187
Ubx_BXD-C	chr3R:12575688-12576663	chr3R:16749966-16750941
tin_tin103A	chr3R:17207742-17208652	chr3R:21382020-21382930
tin_tin103C	chr3R:17208241-17209272	chr3R:21382519-21383550
tin_tinD	chr3R:17209280-17210258	chr3R:21383558-21384536
CAD2_Meso-CRM-2819	chr3R:17222290-17223251	chr3R:21396568-21397529
lbl_SBMs	chr3R:17252278-17253281	chr3R:21426556-21427559
slou_SK16	chr3R:17390942-17391879	chr3R:21565220-21566157
slou_SK19	chr3R:17392337-17393354	chr3R:21566615-21567632
TI_TI287	chr3R:22615901-22616933	chr3R:26790179-26791211
CAD2_Meso-CRM-3418	chr3R:26607382-26608392	chr3R:30781660-30782670
svp_sce	chr3R:8092260-8093272	chr3R:12266538-12267550
CAD2_desat1	chr3R:8269721-8270707	chr3R:12443999-12444985
meso18E_Lac-Z	chrX:19607248-19608272	chrX:19713215-19714239
kirre_-4.6-3.8	chrX:2989425-2990505	chrX:3095392-3096472
kirre_-1.0	chrX:2993145-2994229	chrX:3099112-3100196
org-1_HN39	chrX:8337345-8338317	chrX:8443312-8444284
E0_12_GROAviva_ChIP_chip.region_9	chr2L:3005385-3006375	chr2L:3005385-3006375
E0_12_GROAviva_ChIP_chip.region_25	chr2L:4364826-4365916	chr2L:4364826-4365916
E0_12_GROAviva_ChIP_chip.region_26	chr2L:4367771-4368738	chr2L:4367771-4368738
E0_12_GROAviva_ChIP_chip.region_43	chr2L:6530888-6531887	chr2L:6530888-6531887
E0_12_GROAviva_ChIP_chip.region_46	chr2L:6829880-6830953	chr2L:6829880-6830953
E0_12_GROAviva_ChIP_chip.region_88	chr2L:14403235-14404316	chr2L:14403235-14404316

E0_12_GROAviva_ChIP_chip.region_124	chr2L:17381817-17382775	chr2L:17381817-17382775
E0_12_GROAviva_ChIP_chip.region_125	chr2L:17495387-17496416	chr2L:17495387-17496416
E0_12_GROAviva_ChIP_chip.region_131	chr2L:18598871-18599941	chr2L:18598871-18599941
E0_12_GROAviva_ChIP_chip.region_166	chr2R:5294110-5295135	chr2R:9406605-9407630
E0_12_GROAviva_ChIP_chip.region_179	chr2R:7381681-7382766	chr2R:11494176-11495261
E0_12_GROAviva_ChIP_chip.region_199	chr2R:10852253-10853173	chr2R:14964748-14965668
E0_12_GROAviva_ChIP_chip.region_217	chr2R:14558959-14559942	chr2R:18671454-18672437
E0_12_GROAviva_ChIP_chip.region_233	chr2R:18393902-18394889	chr2R:22506397-22507384
E0_12_GROAviva_ChIP_chip.region_236	chr2R:18402283-18403340	chr2R:22514778-22515835
E0_12_GROAviva_ChIP_chip.region_243	chr2R:19769936-19770987	chr2R:23882413-23883464
E0_12_GROAviva_ChIP_chip.region_247	chr3L:380511-381517	chr3L:380511-381517
E0_12_GROAviva_ChIP_chip.region_258	chr3L:2590366-2591379	chr3L:2590366-2591379
E0_12_GROAviva_ChIP_chip.region_263	chr3L:3854811-3855897	chr3L:3854811-3855897
E0_12_GROAviva_ChIP_chip.region_278	chr3L:7843738-7844664	chr3L:7850638-7851564
E0_12_GROAviva_ChIP_chip.region_279	chr3L:7854680-7855732	chr3L:7861580-7862632
E0_12_GROAviva_ChIP_chip.region_286	chr3L:9415454-9416486	chr3L:9422354-9423386
E0_12_GROAviva_ChIP_chip.region_287	chr3L:9454831-9455827	chr3L:9461731-9462727
E0_12_GROAviva_ChIP_chip.region_311	chr3L:12594129-12595075	chr3L:12601029-12601975
E0_12_GROAviva_ChIP_chip.region_361	chr3L:17356992-17358033	chr3L:17363892-17364933
E0_12_GROAviva_ChIP_chip.region_376	chr3L:18341169-18342254	chr3L:18348069-18349154
E0_12_GROAviva_ChIP_chip.region_377	chr3L:18347618-18348595	chr3L:18354518-18355495
E0_12_GROAviva_ChIP_chip.region_397	chr3L:22451387-22452388	chr3L:22458287-22459288
E0_12_GROAviva_ChIP_chip.region_408	chr3R:2492654-2493749	chr3R:6666932-6668027
E0_12_GROAviva_ChIP_chip.region_415	chr3R:3978032-3979065	chr3R:8152310-8153343
E0_12_GROAviva_ChIP_chip.region_416	chr3R:4005880-4006966	chr3R:8180158-8181244
E0_12_GROAviva_ChIP_chip.region_422	chr3R:5244352-5245319	chr3R:9418630-9419597
E0_12_GROAviva_ChIP_chip.region_439	chr3R:9735607-9736591	chr3R:13909885-13910869
E0_12_GROAviva_ChIP_chip.region_440	chr3R:9755437-9756466	chr3R:13929715-13930744
E0_12_GROAviva_ChIP_chip.region_441	chr3R:9772801-9773843	chr3R:13947079-13948121
E0_12_GROAviva_ChIP_chip.region_444	chr3R:10500128-10501223	chr3R:14674406-14675501

E0_12_GROAviva_ChIP_chip.region_445	chr3R:10501268-10502267	chr3R:14675546-14676545
E0_12_GROAviva_ChIP_chip.region_463	chr3R:12654664-12655640	chr3R:16828942-16829918
E0_12_GROAviva_ChIP_chip.region_467	chr3R:13409319-13410336	chr3R:17583597-17584614
E0_12_GROAviva_ChIP_chip.region_468	chr3R:13414253-13415255	chr3R:17588531-17589533
E0_12_GROAviva_ChIP_chip.region_470	chr3R:13639701-13640639	chr3R:17813979-17814917
E0_12_GROAviva_ChIP_chip.region_471	chr3R:13651536-13652525	chr3R:17825814-17826803
E0_12_GROAviva_ChIP_chip.region_473	chr3R:14821126-14822102	chr3R:18995404-18996380
E0_12_GROAviva_ChIP_chip.region_481	chr3R:16088320-16089304	chr3R:20262598-20263582
E0_12_GROAviva_ChIP_chip.region_482	chr3R:16102202-16103229	chr3R:20276480-20277507
E0_12_GROAviva_ChIP_chip.region_489	chr3R:17326689-17327778	chr3R:21500967-21502056
E0_12_GROAviva_ChIP_chip.region_499	chr3R:17695531-17696528	chr3R:21869809-21870806
E0_12_GROAviva_ChIP_chip.region_506	chr3R:19140180-19141180	chr3R:23314458-23315458
E0_12_GROAviva_ChIP_chip.region_513	chr3R:20421334-20422346	chr3R:24595612-24596624
E0_12_GROAviva_ChIP_chip.region_531	chr3R:22961493-22962480	chr3R:27135771-27136758
E0_12_GROAviva_ChIP_chip.region_562	chrX:3570736-3571779	chrX:3676703-3677746
E0_12_GROAviva_ChIP_chip.region_565	chrX:4083934-4085011	chrX:4189901-4190978
E0_12_GROAviva_ChIP_chip.region_587	chrX:8109215-8110162	chrX:8215182-8216129
E0_12_GROAviva_ChIP_chip.region_590	chrX:8671542-8672579	chrX:8777509-8778546
E0_12_GROAviva_ChIP_chip.region_594	chrX:9643392-9644491	chrX:9749359-9750458
E0_12_GROAviva_ChIP_chip.region_606	chrX:15512939-15514007	chrX:15618906-15619974
E0_12_GROAviva_ChIP_chip.region_620	chrX:19265759-19266850	chrX:19371726-19372817
ind_moduleA	chr3L:15031943-15032964	chr3L:15038843-15039864
ind_moduleBC	chr3L:15032738-15033835	chr3L:15039638-15040735
zen_dorsal_ectoderm	chr3R:2580829-2581762	chr3R:6755107-6756040
CAD2_ttk_early	chr3R:27538548-27539644	chr3R:31712826-31713922
CAD2_blow	chr2R:3472425-3473427	chr2R:7584920-7585922
CAD2_bTub60D	chr2R:20196763-20197766	chr2R:24309240-24310243
CAD2_actin57B	chr2R:16830753-16831785	chr2R:20943248-20944280
CAD2_sns	chr2R:4685108-4686113	chr2R:8797603-8798608
CAD2_htl	chr3R:13875610-13876701	chr3R:18049888-18050979

CAD2_twist	chr2R:18936942-18937936	chr2R:23049419-23050413
CAD2_ttk_late	chr3R:27529652-27530685	chr3R:31703930-31704963
CAD2_RhoL	chr3R:5328991-5330037	chr3R:9503269-9504315
CAD2_Meso-CRM-3775	chr3L:3058767-3059825	chr3L:3058767-3059825
ChIPCRM5405	chr2L:3823411-3824343	chr2L:3823411-3824343
ChIPCRM5432	chr2L:4363360-4364283	chr2L:4363360-4364283
ChIPCRM5792	chr2L:9454132-9455154	chr2L:9454132-9455154
CG7759_33	chr2R:7287136-7288167	chr2R:11399631-11400662
CBP2862	chr3L:320026-321098	chr3L:320026-321098
hth_3	chr3R:6427784-6428745	chr3R:10602062-10603023
ChIPCRM2497	chr3R:12569639-12570593	chr3R:16743917-16744871
ChIPCRM3152	chr3R:21835656-21836646	chr3R:26009934-26010924
ChIPCRM2078	chr3R:7177448-7178448	chr3R:11351726-11352726
rho_NEE_long	chr3L:1461675-1462661	chr3L:1461675-1462661
vn_NEE-long	chr3L:5828505-5829485	chr3L:5835405-5836385
brk_NEE-long	chrX:7190855-7191822	chrX:7296822-7297789
BiTS-ChIP_K4me1+K27me3+K27ac_13	chr2L:5877252-5878345	chr2L:5877252-5878345
BiTS-ChIP_K4me1+K27me3+K27ac_65	chr2L:22430901-22431882	chr2L:22538169-22539150
BiTS-ChIP_K4me1+K27me3+K27ac_73	chr2R:1598304-1599401	chr2R:5710799-5711896
BiTS-ChIP_K4me1+K27me3+K27ac_80	chr2R:3993666-3994721	chr2R:8106161-8107216
BiTS-ChIP_K4me1+K27me3+K27ac_96	chr2R:7245344-7246372	chr2R:11357839-11358867
BiTS-ChIP_K4me1+K27me3+K27ac_106	chr2R:10142814-10143888	chr2R:14255309-14256383
BiTS-ChIP_K4me1+K27me3+K27ac_107	chr2R:10318978-10319971	chr2R:14431473-14432466
BiTS-ChIP_K4me1+K27me3+K27ac_115	chr2R:11818473-11819483	chr2R:15930968-15931978
BiTS-ChIP_K4me1+K27me3+K27ac_134	chr2R:14559686-14560762	chr2R:18672181-18673257
BiTS-ChIP_K4me1+K27me3+K27ac_153	chr2R:20856932-20857931	chr2R:24969409-24970408
BiTS-ChIP_K4me1+K27me3+K27ac_155	chr3L:184534-185555	chr3L:184534-185555
BiTS-ChIP_K4me1+K27me3+K27ac_162	chr3L:1176234-1177274	chr3L:1176234-1177274
BiTS-ChIP_K4me1+K27me3+K27ac_172	chr3L:5892566-5893662	chr3L:5899466-5900562
BiTS-ChIP_K4me1+K27me3+K27ac_190	chr3L:12116996-12117997	chr3L:12123896-12124897

BiTS-ChIP_K4me1+K27me3+K27ac_201	chr3L:14962409-14963507	chr3L:14969309-14970407
BiTS-ChIP_K4me1+K27me3+K27ac_205	chr3L:15718690-15719643	chr3L:15725590-15726543
BiTS-ChIP_K4me1+K27me3+K27ac_212	chr3L:16106832-16107763	chr3L:16113732-16114663
BiTS-ChIP_K4me1+K27me3+K27ac_214	chr3L:16403556-16404528	chr3L:16410456-16411428
BiTS-ChIP_K4me1+K27me3+K27ac_215	chr3L:16722949-16723954	chr3L:16729849-16730854
BiTS-ChIP_K4me1+K27me3+K27ac_219	chr3L:18308482-18309414	chr3L:18315382-18316314
BiTS-ChIP_K4me1+K27me3+K27ac_220	chr3L:18340769-18341697	chr3L:18347669-18348597
BiTS-ChIP_K4me1+K27me3+K27ac_231	chr3L:20784718-20785661	chr3L:20791618-20792561
BiTS-ChIP_K4me1+K27me3+K27ac_236	chr3L:21720852-21721947	chr3L:21727752-21728847
BiTS-ChIP_K4me1+K27me3+K27ac_246	chr3R:224803-225744	chr3R:4399081-4400022
BiTS-ChIP_K4me1+K27me3+K27ac_252	chr3R:1145198-1146137	chr3R:5319476-5320415
BiTS-ChIP_K4me1+K27me3+K27ac_262	chr3R:4094804-4095867	chr3R:8269082-8270145
BiTS-ChIP_K4me1+K27me3+K27ac_320	chr3R:15150586-15151685	chr3R:19324864-19325963
BiTS-ChIP_K4me1+K27me3+K27ac_322	chr3R:16774066-16775162	chr3R:20948344-20949440
BiTS-ChIP_K4me1+K27me3+K27ac_343	chr3R:21836549-21837613	chr3R:26010827-26011891
BiTS-ChIP_K4me1+K27me3+K27ac_365	chr4:854225-855246	chr4:833599-834620
BiTS-ChIP_K4me1+K27me3+K27ac_389	chrX:17520833-17521855	chrX:17626800-17627822
Classl_ins_17	chr2L:7204179-7205202	chr2L:7204179-7205202
Classl_ins_29	chr2L:16719303-16720312	chr2L:16719303-16720312
Classl_ins_34	chr2L:19426381-19427369	chr2L:19426381-19427369
Classl_ins_36	chr2L:20120424-20121485	chr2L:20120424-20121485
Classl_ins_61	chr2R:16554832-16555765	chr2R:20667327-20668260
Classl_ins_77	chr3L:2553597-2554557	chr3L:2553597-2554557
Classl_ins_87	chr3L:6543350-6544404	chr3L:6550250-6551304
Classl_ins_94	chr3L:11782712-11783704	chr3L:11789612-11790604
Classl_ins_103	chr3L:16948111-16949205	chr3L:16955011-16956105
Classl_ins_114	chr3L:21834048-21835060	chr3L:21840948-21841960
Classl_ins_132	chr3R:6500441-6501522	chr3R:10674719-10675800
Classl_ins_133	chr3R:7647743-7648797	chr3R:11822021-11823075
Classl_ins_157	chr3R:16981256-16982186	chr3R:21155534-21156464

Classl_ins_182	chr3R:26427150-26428063	chr3R:30601428-30602341
Classl_ins_196	chrX:3999494-4000444	chrX:4105461-4106411
Classl_ins_198	chrX:4850001-4851012	chrX:4955968-4956979
Classl_ins_202	chrX:9494482-9495522	chrX:9600449-9601489
Classl_ins_207	chrX:11001902-11002920	chrX:11107869-11108887
Classl_ins_209	chrX:11515472-11516391	chrX:11621439-11622358
Classl_ins_210	chrX:13223518-13224484	chrX:13329485-13330451
Classl_ins_212	chrX:14848851-14849839	chrX:14954818-14955806
Classl_ins_215	chrX:15692861-15693858	chrX:15798828-15799825
Classl_ins_229	chrX:18031759-18032839	chrX:18137726-18138806
Classl_ins_244	chrX:20987635-20988713	chrX:21116608-21117686
ins_scs	chr3R:7774471-7775540	chr3R:11948749-11949818
ins_scs-prime	chr3R:7788350-7789392	chr3R:11962628-11963670
ins_1A2_assoc_peak	chrX:255243-256339	chrX:361210-362306
ins_SF1_assoc_peak_center	chr3R:2679760-2680827	chr3R:6854038-6855105
ins_Fab-8	chr3R:12744521-12745503	chr3R:16918799-16919781
ins_Fab-7_minimal_overlap	chr3R:12724486-12725543	chr3R:16898764-16899821
DHS+K27me3_CtBPOverlap_762	chr2R:11732837-11733877	chr2R:15845332-15846372
DHS+K27me3_CtBPOverlap_1612	chr3R:186111-187132	chr3R:4360389-4361410
DHS+K27me3_CtBPOverlap_1813	chr3R:8837370-8838452	chr3R:13011648-13012730
DHS+K27me3_CtBPOverlap_2206	chr4:729184-730270	chr4:708558-709644
DHS+K27me3_CtBPOverlap_2215	chr4:853288-854320	chr4:832662-833694
DHS+K27me3_CtBPOverlap_2267	chrX:482837-483821	chrX:588804-589788
DHS+K27me3_CtBPOverlap_2418	chrX:8665718-8666727	chrX:8771685-8772694
DHS+K27me3_CtBPOverlap_2598	chrX:18206161-18207162	chrX:18312128-18313129
DHS+K27me3_intergenic_16	chr2L:368364-369325	chr2L:368364-369325
DHS+K27me3_intergenic_71	chr2L:4807981-4808897	chr2L:4807981-4808897
DHS+K27me3_intergenic_412	chr2L:18814986-18815972	chr2L:18814986-18815972
DHS+K27me3_intergenic_1490a	chr3L:18393140-18394184	chr3L:18400040-18401084
DHS+K27me3_intergenic_1490b	chr3L:18394914-18396011	chr3L:18401814-18402911

DHS+K27me3_intergenic_1610	chr3R:169435-170404	chr3R:4343713-4344682
DHS+K27me3_intergenic_1932	chr3R:13381473-13382383	chr3R:17555751-17556661
DHS+K27me3_intergenic_2074a	chr3R:20913934-20914991	chr3R:25088212-25089269
DHS+K27me3_intergenic_2074b	chr3R:20917026-20918028	chr3R:25091304-25092306
DHS+K27me3_intergenic_2153	chr3R:26911747-26912841	chr3R:31086025-31087119
DHS+K27me3_one-hit_20	chr2L:419629-420658	chr2L:419629-420658
DHS+K27me3_one-hit_24	chr2L:589607-590595	chr2L:589607-590595
DHS+K27me3_one-hit_75	chr2L:5055795-5056753	chr2L:5055795-5056753
DHS+K27me3_one-hit_103	chr2L:6088117-6089150	chr2L:6088117-6089150
DHS+K27me3_one-hit_202	chr2L:9792224-9793199	chr2L:9792224-9793199
DHS+K27me3_one-hit_348	chr2L:15334541-15335597	chr2L:15334541-15335597
DHS+K27me3_one-hit_372	chr2L:16483697-16484642	chr2L:16483697-16484642
DHS+K27me3_one-hit_402	chr2L:18601807-18602751	chr2L:18601807-18602751
DHS+K27me3_one-hit_877	chr2R:16845185-16846266	chr2R:20957680-20958761
DHS+K27me3_one-hit_1084	chr3L:5874606-5875632	chr3L:5881506-5882532
DHS+K27me3_one-hit_1265	chr3L:12602226-12603228	chr3L:12609126-12610128
DHS+K27me3_one-hit_1324	chr3L:14182120-14183122	chr3L:14189020-14190022
DHS+K27me3_one-hit_1338	chr3L:14592451-14593527	chr3L:14599351-14600427
DHS+K27me3_one-hit_1883	chr3R:12214480-12215427	chr3R:16388758-16389705
DHS+K27me3_one-hit_1994	chr3R:16672486-16673386	chr3R:20846764-20847664
DHS+K27me3_one-hit_2116	chr3R:25517492-25518523	chr3R:29691770-29692801
DHS+K27me3_one-hit_2147	chr3R:26632727-26633769	chr3R:30807005-30808047
DHS+K27me3_one-hit_2225	chr4:1009740-1010779	chr4:989114-990153
DHS+K27me3_one-hit_2281	chrX:2036829-2037748	chrX:2142796-2143715
Ecoli_control3	N/A	N/A
Ecoli_control4	N/A	N/A
Ecoli_control5	N/A	N/A
Ecoli_control7	N/A	N/A
Ecoli_control11	N/A	N/A
Ecoli_control12	N/A	N/A

Ecoli_control15	N/A	N/A
Ecoli_control16	N/A	N/A
Ecoli_control21	N/A	N/A
Ecoli_control22	N/A	N/A
Ecoli_control23	N/A	N/A
Ecoli_control24	N/A	N/A
DHS+K27me3_one-hit_2556	chrX:17207836-17208913	chrX:17313803-17314880
DHS+K27me3_one-hit_2588	chrX:18135231-18136260	chrX:18241198-18242227
DHS+K27me3_one-hit_2606	chrX:19429583-19430513	chrX:19535550-19536480
DHS+K27me3_one-hit_2645	chrX:21244727-21245748	chrX:21373700-21374721
E_0_12h_dCtBP7667.region_2256	chr3L:5548950-5550048	chr3L:5555850-5556948
E_0_12h_dCtBP7667.region_2659	chr3L:14604450-14605479	chr3L:14611350-14612379
E_0_12h_dCtBP7667.region_991	chr2L:22437120-22438207	chr2L:22544388-22545475
E_0_12h_dCtBP7667.region_707	chr2L:15074407-15075441	chr2L:15074407-15075441
E_0_12h_dCtBP7667.region_2182	chr3L:3804155-3805254	chr3L:3804155-3805254
E_0_12h_dCtBP7667.region_4755	chrX:15790373-15791443	chrX:15896340-15897410
E_0_12h_dCtBP7667.region_486	chr2L:9690333-9691430	chr2L:9690333-9691430
E_0_12h_dCtBP7667.region_3919	chr3R:20120318-20121414	chr3R:24294596-24295692
E_0_12h_dCtBP7667.region_258	chr2L:4277263-4278353	chr2L:4277263-4278353
E_0_12h_dCtBP7667.region_2713	chr3L:15950232-15951330	chr3L:15957132-15958230
E_0_12h_dCtBP7667.region_2437	chr3L:9376374-9377373	chr3L:9383274-9384273
E_0_12h_dCtBP7667.region_2793	chr3L:18101233-18102321	chr3L:18108133-18109221
E_0_12h_dCtBP7667.region_3346	chr3R:7412623-7413722	chr3R:11586901-11588000
E_0_12h_dCtBP7667.region_1752	chr2R:16473538-16474572	chr2R:20586033-20587067
E_0_12h_dCtBP7667.region_275	chr2L:4810330-4811384	chr2L:4810330-4811384
E_0_12h_dCtBP7667.region_518	chr2L:10525546-10526620	chr2L:10525546-10526620
E_0_12h_dCtBP7667.region_1104	chr2R:3890223-3891187	chr2R:8002718-8003682
E_0_12h_dCtBP7667.region_1302	chr2R:7788986-7790016	chr2R:11901481-11902511
E_0_12h_dCtBP7667.region_1582	chr2R:12944683-12945617	chr2R:17057178-17058112
E_0_12h_dCtBP7667.region_2748	chr3L:16660796-16661735	chr3L:16667696-16668635

E_0_12h_dCtBP7667.region_3104	chr3R:2022792-2023841	chr3R:6197070-6198119
E_0_12h_dCtBP7667.region_3445	chr3R:9746846-9747838	chr3R:13921124-13922116
E_0_12h_dCtBP7667.region_3446	chr3R:9750854-9751776	chr3R:13925132-13926054
E_0_12h_dCtBP7667.region_3938	chr3R:20592612-20593661	chr3R:24766890-24767939
E_0_12h_dCtBP7667.region_4196	chr3R:26689219-26690188	chr3R:30863497-30864466
E_0_12h_dCtBP7667.region_4427	chrX:5479798-5480878	chrX:5585765-5586845
E_0_12h_dCtBP7667.region_4530	chrX:8727320-8728324	chrX:8833287-8834291
E_0_12h_dCtBP7667.region_3402	chr3R:8681541-8682546	chr3R:12855819-12856824
E_0_12h_dCtBP7667.region_3606	chr3R:12894974-12896064	chr3R:17069252-17070342
E_0_12h_dCtBP7667.region_3031	chr3R:561158-562190	chr3R:4735436-4736468
E_0_12h_dCtBP7667.region_4944	chrX:22015098-22016116	chrX:22613262-22614280
E_0_12h_dCtBP7667.region_3202	chr3R:4553174-4554259	chr3R:8727452-8728537
E_0_12h_dCtBP7667.region_4844	chrX:18704252-18705253	chrX:18810219-18811220
E_0_12h_dCtBP7667.region_3049	chr3R:907088-908163	chr3R:5081366-5082441
E_0_12h_dCtBP7667.region_1981	chr2R:20822326-20823378	chr2R:24934803-24935855
E_0_12h_dCtBP7667.region_1692	chr2R:15090938-15092004	chr2R:19203433-19204499
E_0_12h_dCtBP7667.region_3799	chr3R:17600358-17601412	chr3R:21774636-21775690
E_0_12h_dCtBP7667.region_4154	chr3R:25655952-25657008	chr3R:29830230-29831286
E_0_12h_dCtBP7667.region_2344	chr3L:7339019-7340114	chr3L:7345919-7347014
E_0_12h_dCtBP7667.region_3138	chr3R:2961101-2962189	chr3R:7135379-7136467
E_0_12h_dCtBP7667.region_234	chr2L:3708029-3709016	chr2L:3708029-3709016
E_0_12h_dCtBP7667.region_3785	chr3R:17162159-17163217	chr3R:21336437-21337495
E_0_12h_dCtBP7667.region_4738	chrX:15467073-15468098	chrX:15573040-15574065
E_0_12h_dCtBP7667.region_2057	chr3L:988005-989088	chr3L:988005-989088
E_0_12h_dCtBP7667.region_3325	chr3R:7032200-7033286	chr3R:11206478-11207564
E_0_12h_dCtBP7667.region_3509	chr3R:11149309-11150347	chr3R:15323587-15324625
E_0_12h_dCtBP7667.region_3602	chr3R:12835410-12836498	chr3R:17009688-17010776
E_0_12h_dCtBP7667.region_864	chr2L:19063547-19064615	chr2L:19063547-19064615
E_0_12h_dCtBP7667.region_2824	chr3L:19025236-19026300	chr3L:19032136-19033200
E_0_12h_dCtBP7667.region_3249	chr3R:5388236-5389306	chr3R:9562514-9563584

E_0_12h_dCtBP7667.region_909	chr2L:20029457-20030428	chr2L:20029457-20030428
E_0_12h_dCtBP7667.region_63	chr4:1102343-1103437	chr4:1081717-1082811
E_0_12h_dCtBP7667.region_1790	chr2R:17084978-17086067	chr2R:21197473-21198562
E_0_12h_dCtBP7667.region_2744	chr3L:16633761-16634825	chr3L:16640661-16641725
E_0_12h_dCtBP7667.region_55	chr4:963630-964664	chr4:943004-944038
E_0_12h_dCtBP7667.region_2243	chr3L:5146157-5147235	chr3L:5153057-5154135
E_0_12h_dCtBP7667.region_788	chr2L:17487597-17488678	chr2L:17487597-17488678
E_0_12h_dCtBP7667.region_3439	chr3R:9609748-9610746	chr3R:13784026-13785024
E_0_12h_dCtBP7667.region_2827	chr3L:19054256-19055254	chr3L:19061156-19062154
E_0_12h_dCtBP7667.region_1855	chr2R:18465378-18466346	chr2R:22577873-22578841
E_0_12h_dCtBP7667.region_4460	chrX:6717063-6718123	chrX:6823030-6824090
E_0_12h_dCtBP7667.region_2991	chr3L:23312459-23313549	chr3L:23319359-23320449
E_0_12h_dCtBP7667.region_3696	chr3R:14933397-14934459	chr3R:19107675-19108737
E_0_12h_dCtBP7667.region_1555	chr2R:12448036-12449035	chr2R:16560531-16561530
E_0_12h_dCtBP7667.region_1107	chr2R:3960033-3961115	chr2R:8072528-8073610
E_0_12h_dCtBP7667.region_3770	chr3R:16899442-16900447	chr3R:21073720-21074725
E_0_12h_dCtBP7667.region_3553	chr3R:11836039-11837105	chr3R:16010317-16011383
E_0_12h_dCtBP7667.region_1767	chr2R:16764285-16765340	chr2R:20876780-20877835
E_0_12h_dCtBP7667.region_1028	chr2R:1786069-1787118	chr2R:5898564-5899613
E_0_12h_dCtBP7667.region_3484	chr3R:10498621-10499660	chr3R:14672899-14673938
E_0_12h_dCtBP7667.region_3849	chr3R:18883661-18884717	chr3R:23057939-23058995
E_0_12h_dCtBP7667.region_3937	chr3R:20541751-20542778	chr3R:24716029-24717056
E_0_12h_dCtBP7667.region_3852	chr3R:18955501-18956527	chr3R:23129779-23130805
E_0_12h_dCtBP7667.region_4503	chrX:8002538-8003524	chrX:8108505-8109491
E_0_12h_dCtBP7667.region_3708	chr3R:15213520-15214591	chr3R:19387798-19388869
E_0_12h_dCtBP7667.region_3540	chr3R:11679911-11681008	chr3R:15854189-15855286
E_0_12h_dCtBP7667.region_2952	chr3L:21964840-21965825	chr3L:21971740-21972725
E_0_12h_dCtBP7667.region_3547	chr3R:11787738-11788835	chr3R:15962016-15963113
E_0_12h_dCtBP7667.region_1533	chr2R:12058557-12059556	chr2R:16171052-16172051
E_0_12h_dCtBP7667.region_3084	chr3R:1508921-1509997	chr3R:5683199-5684275

E\_0\_12h\_dCtBP7667.region\_772  
E\_0\_12h\_dCtBP7667.region\_2926  
E\_0\_12h\_dCtBP7667.region\_1246  
E\_0\_12h\_dCtBP7667.region\_3954  
E\_0\_12h\_dCtBP7667.region\_2772  
E\_0\_12h\_dCtBP7667.region\_3525  
E\_0\_12h\_dCtBP7667.region\_3137  
E\_0\_12h\_dCtBP7667.region\_3629  
E\_0\_12h\_dCtBP7667.region\_4575  
E\_0\_12h\_dCtBP7667.region\_3623  
E\_0\_12h\_dCtBP7667.region\_3006  
E\_0\_12h\_dCtBP7667.region\_2635

#### **Negative control**

Ecoli\_control3  
Ecoli\_control4  
Ecoli\_control5  
Ecoli\_control7  
Ecoli\_control11  
Ecoli\_control12  
Ecoli\_control15  
Ecoli\_control16  
Ecoli\_control21  
Ecoli\_control22  
Ecoli\_control23  
Ecoli\_control24

chr2L:16853759-16854754  
chr3L:21504512-21505558  
chr2R:6507041-6508091  
chr3R:20896811-20897745  
chr3L:17363459-17364446  
chr3R:11318064-11319157  
chr3R:2949451-2950400  
chr3R:13385900-13386906  
chrX:9892224-9893283  
chr3R:13271943-13273006  
chr3R:74837-75870  
chr3L:14099332-14100315

#### **Forward primer**

TGATGATGTTGCCGCTGGTC  
CATCAACCAGACCAAAGAAGTCG  
CCAGGCGAAAAAGTTCAGCG  
AGGATGGAGATTATCGTAAAGGGC  
ATTGTCTTTGTCGGATTGACGG  
CTTGCTCAGAAACGATGATGGATG  
TGACAGGATTAGCCAAAACCAGC  
CCCAAACAGTGTATGAATGGTGTG  
GTAGAGTGGCAGGCTTTCGTTG  
GCAAGTATCCCAAAGAAGCCG  
GGGGCAACAGGTGATGTATGTG  
CTGTTGTTGATTACGCCGTCG

chr2L:16853759-16854754  
chr3L:21511412-21512458  
chr2R:10619536-10620586  
chr3R:25071089-25072023  
chr3L:17370359-17371346  
chr3R:15492342-15493435  
chr3R:7123729-7124678  
chr3R:17560178-17561184  
chrX:9998191-9999250  
chr3R:17446221-17447284  
chr3R:4249115-4250148  
chr3L:14106232-14107215

#### **Reverse primer**

CTGTGGAAATCCCGTGACAAATAG  
AAAATGGAAAAACCCGCTCG  
AGCGATGGGTATTCCGTCTTCC  
ATCGCTAACCTGTTGCTGGCTC  
CAACTCTTCACTATGGCATTTC  
CGTCAGTGGCACAAAAAGGAAG  
TCCGTATTTATGACACCGCCC  
AACCCGTGAAAATCTGCCAAC  
TCTGAAGGAAGATGATAAGTGGCG  
AAAGCGAAGCGTTGATGTGG  
GTGGTCGCAATAGTTATCGTTCCG  
CTCCCAGATTTTTTCCGCAC

## Supplementary table 2

Genomic coordinates of the second library elements screened by silencer-FACS-Seq.

<b>Library element</b>	<b>coordinates (dm3)</b>
amosD	chr2L:18595204-18596246
bab1_dimorphic_element	chr3L:1084698-1085668
BiTS-ChIP_K4me1+K27me3+K27ac_13	chr2L:5877252-5878345
BiTS-ChIP_K4me1+K27me3+K27ac_134	chr2R:14559686-14560762
BiTS-ChIP_K4me1+K27me3+K27ac_73	chr2R:1598304-1599401
brk_NEE-long	chrX:7190855-7191822
btd_R-Ss	chrX:9583710-9584690
btd_Ss-Bg	chrX:9584428-9585527
CAD2_htl	chr3R:13875610-13876701
CAD2_Ket-1	chr3L:2089382-2090424
CAD2_Meso-CRM-4726	chr3L:16760731-16761829
CAD2_Meso-CRM-6028	chr2L:14266991-14267912
CAD2_Meso-CRM-6225	chr2L:17221775-17222718
cas_csc-3	chr3R:1552157-1553136
cas_csc-5a	chr3R:1548621-1549584
cas_csc-7a	chr3R:1544440-1545364
CG42342_3436	chr3R:12377681-12378775
CG7722_CRM28	chr2R:6824079-6825095
ChIPCRM2078	chr3R:7177448-7178448
ChIPCRM5792	chr2L:9454132-9455154
Classl_ins_198	chrX:4850001-4851012
cpo_cpoCRM6	chr3R:13777548-13778572
crb_Lac-Z	chr3R:20122950-20123916
dac_RE	chr2L:16465126-16466148
DHS+K27me3_CtBPOverlap_2215	chr4:853288-854320

DHS+K27me3_CtBPOverlap_762	chr2R:11732837-11733877
DHS+K27me3_intergenic_1932	chr3R:13381473-13382383
DHS+K27me3_intergenic_2074b	chr3R:20917026-20918028
DHS+K27me3_intergenic_412	chr2L:18814986-18815972
DHS+K27me3_intergenic_71	chr2L:4807981-4808897
DHS+K27me3_one-hit_103	chr2L:6088117-6089150
DHS+K27me3_one-hit_2281	chrX:2036829-2037748
DHS+K27me3_one-hit_348	chr2L:15334541-15335597
DHS+K27me3_one-hit_402	chr2L:18601807-18602751
DII_304	chr2R:20690263-20691248
dpp_85.8MX	chr2L:2456545-2457501
dpp_BS3.1	chr2L:2480501-2481487
dpp_construct10	chr2L:2481008-2482099
dpp_VRR	chr2L:2455833-2456926
E_0_12h_dCtBP7667.region_1107	chr2R:3960033-3961115
E_0_12h_dCtBP7667.region_1246	chr2R:6507041-6508091
E_0_12h_dCtBP7667.region_1790	chr2R:17084978-17086067
E_0_12h_dCtBP7667.region_2635	chr3L:14099332-14100315
E_0_12h_dCtBP7667.region_2744	chr3L:16633761-16634825
E_0_12h_dCtBP7667.region_3049	chr3R:907088-908163
E_0_12h_dCtBP7667.region_3346	chr3R:7412623-7413722
E_0_12h_dCtBP7667.region_3602	chr3R:12835410-12836498
E_0_12h_dCtBP7667.region_3623	chr3R:13271943-13273006
E_0_12h_dCtBP7667.region_3785	chr3R:17162159-17163217
E_0_12h_dCtBP7667.region_3799	chr3R:17600358-17601412
E_0_12h_dCtBP7667.region_3852	chr3R:18955501-18956527
E_0_12h_dCtBP7667.region_4427	chrX:5479798-5480878
E_0_12h_dCtBP7667.region_4575	chrX:9892224-9893283
E_0_12h_dCtBP7667.region_4755	chrX:15790373-15791443
E_0_12h_dCtBP7667.region_486	chr2L:9690333-9691430

E_0_12h_dCtBP7667.region_4944	chrX:22015098-22016116
E_0_12h_dCtBP7667.region_55	chr4:963630-964664
E_0_12h_dCtBP7667.region_63	chr4:1102343-1103437
E_0_12h_dCtBP7667.region_772	chr2L:16853759-16854754
E_0_12h_dCtBP7667.region_788	chr2L:17487597-17488678
E_0_12h_dCtBP7667.region_909	chr2L:20029457-20030428
e_coreAbdominalCRE	chr3R:17066418-17067406
E0_12_GROAviva_ChIP_chip.region_124	chr2L:17381817-17382775
E0_12_GROAviva_ChIP_chip.region_258	chr3L:2590366-2591379
E0_12_GROAviva_ChIP_chip.region_26	chr2L:4367771-4368738
E0_12_GROAviva_ChIP_chip.region_278	chr3L:7843738-7844664
E0_12_GROAviva_ChIP_chip.region_397	chr3L:22451387-22452388
E0_12_GROAviva_ChIP_chip.region_408	chr3R:2492654-2493749
E0_12_GROAviva_ChIP_chip.region_416	chr3R:4005880-4006966
E0_12_GROAviva_ChIP_chip.region_439	chr3R:9735607-9736591
E0_12_GROAviva_ChIP_chip.region_440	chr3R:9755437-9756466
E0_12_GROAviva_ChIP_chip.region_46	chr2L:6829880-6830953
E0_12_GROAviva_ChIP_chip.region_470	chr3R:13639701-13640639
E0_12_GROAviva_ChIP_chip.region_471	chr3R:13651536-13652525
E0_12_GROAviva_ChIP_chip.region_473	chr3R:14821126-14822102
E0_12_GROAviva_ChIP_chip.region_482	chr3R:16102202-16103229
E0_12_GROAviva_ChIP_chip.region_562	chrX:3570736-3571779
E0_12_GROAviva_ChIP_chip.region_565	chrX:4083934-4085011
E0_12_GROAviva_ChIP_chip.region_590	chrX:8671542-8672579
E0_12_GROAviva_ChIP_chip.region_594	chrX:9643392-9644491
E0_12_GROAviva_ChIP_chip.region_606	chrX:15512939-15514007
E0_12_GROAviva_ChIP_chip.region_620	chrX:19265759-19266850
gsb_fragIV	chr2R:20944064-20945045
gt_CE8001	chrX:2324615-2325714
h_h7AF	chr3L:8658167-8659238

Hand_HCH	chr2L:10293308-10294379
hb_distal_minimal	chr3R:4524698-4525688
hb_HG4-6	chr3R:4529043-4530092
hb_HG4-7	chr3R:4527834-4528862
hb_lateDm1.0- <i>lacZ</i>	chr3R:4526286-4527384
hh_hhf4F	chr3R:18963250-18964276
hkb_0.6kbRIRV	chr3R:173849-174821
hth_3	chr3R:6427784-6428745
ind_moduleA	chr3L:15031943-15032964
ind_moduleBC	chr3L:15032738-15033835
ins_Fab-7_minimal_overlap	chr3R:12724486-12725543
ins_SF1_assoc_peak_center	chr3R:2679760-2680827
kirre_-4.6-3.8	chrX:2989425-2990505
kni_223	chr3L:20690490-20691542
Kr_KrMT	chr2R:21098808-21099895
lbl_SBM5	chr3R:17252278-17253281
lz_CrystalCellEnhancer1236-737	chrX:9177203-9178176
lz_LMEE	chrX:9180643-9181594
Ndg_FCenhancer	chr2R:6203047-6203992
nerfin-1_fragment3	chr3L:904458-905370
nkd_UpE2	chr3L:19047510-19048605
oc_otd-186	chrX:8548281-8549211
ovo_E3	chrX:4914971-4916045
pdm2_CRM6	chr2L:12682595-12683690
per_-603_-449	chrX:2578569-2579555
pnr_P3	chr3R:11853793-11854843
pnr_P4	chr3R:11854355-11855347
prd_P1_enhancer	chr2L:12087791-12088820
repo_-1.1	chr3R:14060749-14061844
rho_NEE_long	chr3L:1461675-1462661

sc_CRM39	chrX:286648-287747
slp1_5303	chr2L:3827700-3828650
sog_broad_lateral_neurogenic_ectoderm	chrX:15518390-15519344
sog_shadow	chrX:15540621-15541615
Sox15_regionC	chr2R:10097224-10098222
SoxN_565	chr2L:8811551-8812587
SoxN_5830	chr2L:8841052-8842145
sphinx_1067bp_5'_fragment	chr4:994776-995867
sqz_sqz-11	chr3R:15000126-15001211
sv_paxD	chr4:1106414-1107513
ths_Neu4_early_embryonic_enhancer	chr2R:7681709-7682675
tin_tinD	chr3R:17209280-17210258
tup_dorsalectoderm	chr2L:18874963-18875896
Ubx_BXD-C	chr3R:12575688-12576663
vnd_743	chrX:486746-487752
vnd_NEE	chrX:486301-487394
vvl_vvl1+2	chr3L:6757667-6758693
y_BE1-2	chrX:248439-249451
y_BE3	chrX:249145-250233
y_wing	chrX:247798-248732

## Supplementary table 3

Results of silencer-FACS-Seq experiments for the first library.

For each element confidently detected in either complete repetition of the sFS experiment (see Materials and Methods for a precise description of confident detection), the source (type of sequence chosen for testing) is shown; mean abundance in input cells across three biological replicates,  $\log_2$  of the fold change (enrichment) in CD2<sup>+</sup>GFP<sup>reduced</sup> cells, and adjusted p-value for enrichment/depletion are shown for two complete experimental repetitions (nd: not detected); whether the element was subsequently annotated as overlapping a transcriptional start site (TSS) is indicated; and the results of validation experiments are shown (nd: not done).

element	source	inputAbundance	Experiment 1 log2FoldChange	padj	inputAbundance	Experiment 2 log2FoldChange	padj	TSS overlap	validation result
18w_1625	Nonmeso CRMs	50732.13456	-0.297378764	0.871800354	31276.68956	-0.160700528	0.995277602	0	nd
Acon_Lac-Z	Specific meso CRMs	11776.15951	-0.534398806	0.881141655	8526.809419	-1.479843321	0.784727605	0	nd
Acp65Aa_172	Nonmeso CRMs	36939.82811	-0.886695317	0.56709847	46204.72616	1.569744892	0.136197399	1	nd
Atet_5338	Specific meso CRMs	33748.8704	0.3565045	0.638448696	41103.37978	-0.079105246	0.993889052	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_106	"Bivalent" chromatin	6260.669278	-2.220898397	0.273643866	1883.818527	-1.14610608	0.874339989	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_107	"Bivalent" chromatin	18474.11973	0.776711833	0.297461012	3077.895276	-0.08679011	0.807961721	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_153	"Bivalent" chromatin	6075.849464	2.536234281	0.001985609	1839.478142	4.495744871	0.000540628	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_155	"Bivalent" chromatin	23314.28543	-1.345061907	0.404360251	62090.80287	-0.487343896	0.917027471	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_201	"Bivalent" chromatin	8142.328604	-1.898205806	0.330192441	6678.396756	-0.893113285	0.77842477	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_205	"Bivalent" chromatin	20309.65208	-0.93036895	0.588898233	26621.86071	-0.742965693	0.96722015	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_214	"Bivalent" chromatin	5998.105506	-1.837685226	0.254756591	1907.255366	-2.32453355	0.917027471	0	-
BiTS-ChIP_K4me1+K27me3+K27ac_219	"Bivalent" chromatin	13247.17625	-2.330666471	0.103480603	5592.562703	0.010495797	0.967071039	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_220	"Bivalent" chromatin	15104.19855	-1.707420728	0.170364813	1974.453621	1.7805079	0.594936531	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_231	"Bivalent" chromatin	6144.781497	-2.240933693	0.22229563	494.5920823	-0.758702309	0.789366919	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_236	"Bivalent" chromatin	2956.284208	-0.210196487	0.927984944	6206.794683	-0.876272375	0.96722015	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_246	"Bivalent" chromatin	76623.06936	-1.250090213	0.317990535	70880.20331	-0.932773553	0.865316701	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_252	"Bivalent" chromatin	21836.90734	-1.346998916	0.3180585	25267.98062	-0.00994452	0.995277602	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_262	"Bivalent" chromatin	22373.67419	-0.247725147	0.909831723	25897.42388	0.184072187	0.77253156	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_343	"Bivalent" chromatin	1827.312606	-1.161346283	0.813778042	5319.489904	-1.9096882	0.840794673	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_389	"Bivalent" chromatin	13405.52939	-1.990846817	0.105389871	15928.48423	-1.240053556	0.77253156	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_65	"Bivalent" chromatin	9528.182621	-1.894526628	0.166339625	8302.960798	-2.040383298	0.44265801	1	nd
BiTS-ChIP_K4me1+K27me3+K27ac_96	"Bivalent" chromatin	50314.70006	-2.70824115	0.029606039	78295.67674	-2.474028649	0.217753773	0	-
C15_350-2	Nonmeso CRMs	2772.385731	-1.268999979	0.773085189	13498.07654	-2.526340754	0.3438437	0	nd
CAD2_Ket-2	Specific meso CRMs	52674.68793	-1.954184698	0.150596149	28821.59749	-2.224882004	0.348404591	0	nd

CAD2_Meso-CRM-2819	Specific meso CRMs	10860.21205	-1.459005211	0.368914552	1.146891281	9.181466875	0.005254219	0	+
CAD2_Meso-CRM-3775	Negative controls	1067.530376	-0.360769954	0.946775187	124.0891439	-1.074591105	0.77253156	0	nd
CAD2_Meso-CRM-4906	Specific meso CRMs	2112.770221	-3.749982564	0.201266734	3416.130767	0.428374223	0.51777849	0	nd
CAD2_Meso-CRM-965	Specific meso CRMs	36632.25984	1.825778319	0.005583975	14671.31008	3.332489023	0.000515588	1	nd
CAD2_actin57B	Negative controls	1156.47524	2.65542879	0.041791735	975.9402164	3.510361713	0.011269571	1	nd
CAD2_bTub60D	Negative controls	3517.347364	-0.921227591	0.925377813	11297.2041	-1.605406174	0.824843969	0	nd
CAD2_hand_1	Specific meso CRMs	45.61165587	3.046865938	0.122791048	93.0272564	0.32722615	0.514431368	0	nd
CAD2_pnr	Specific meso CRMs	25719.28954	1.222560461	0.116106607	19841.68558	1.360440637	0.286329637	1	nd
CAD2_pyr	Specific meso CRMs	5798.271531	-3.075584012	0.145969007	8455.537942	-0.100392578	0.769738145	0	nd
CAD2_sna_2	Specific meso CRMs	10875.69842	-2.039907403	0.22229563	4791.229748	-7.558959333	0.008725821	0	nd
CAD2_sns	Negative controls	937.1756677	6.109425274	1.09E-11	751.1486054	6.976864498	3.38E-14	1	+
CAD2_ttk_early	Negative controls	34766.09931	-0.163630025	0.971635655	22456.42781	-1.499116987	0.660781638	1	nd
CAD2_ttk_late	Negative controls	489.0420081	1.252479407	0.372327237	1706.36855	-0.314075527	0.769738145	0	nd
CAD2_twist	Negative controls	2246.654605	1.014156884	0.359121812	13980.65549	1.178507404	0.281060219	0	nd
CBP2862	Negative controls	5713.016246	-2.106057965	0.359121812	1269.453354	0.663026319	0.441976808	0	nd
CG12374_CRM41	Nonmeso CRMs	18770.47668	-0.78352942	0.669754446	33733.85742	-0.963413173	0.662951478	0	nd
CG13333_link_5'	Nonmeso CRMs	12933.39533	3.592425843	2.82E-07	6585.622766	4.290958562	2.71E-06	1	nd
CG16778_tkr-15	Nonmeso CRMs	4842.665925	-0.777506014	0.854676427	127590.4448	-2.395178247	0.117002613	0	nd
CG17230_CRM4	Nonmeso CRMs	11700.44469	-0.62265638	0.708439789	19181.10106	-1.969130587	0.441976808	0	nd
CG34347_CG11339_EVIII	Nonmeso CRMs	2094.626808	-0.071628445	0.833415143	3484.342227	0.630734441	0.663112564	0	nd
CG4364_upstreamCRM	Specific meso CRMs	54.30860103	6.807936733	0.000613487	30.99419437	7.48956339	0.000135623	1	nd
CG7229_CG7229	Nonmeso CRMs	2953.173845	-3.550567234	0.103480603	239.6423996	-5.882459745	0.993889052	0	nd
CG7229_CRM15	Nonmeso CRMs	4558.284191	-2.257629205	0.239263706	2412.846016	-1.630301674	0.874339989	0	nd
CG7458_upstreamCRM	Nonmeso CRMs	8401.440218	-0.357480036	0.838428709	15205.04054	2.147670884	0.074687828	1	nd
CG7759_33	Negative controls	6313.544127	-4.841339124	0.006363306	1887.211304	-1.64863644	0.96722015	0	nd
CG9416_GFP	Specific meso CRMs	7019.52141	-2.0751848	0.377826386	6663.388203	-1.899317294	0.629000298	0	nd
CG9571_O-E	Nonmeso CRMs	215.4880792	1.646690325	0.331116597	177.4283521	2.97807546	0.281060219	1	nd
CadN_Lac-Z	Specific meso CRMs	1987.457325	-0.474884411	0.96828369	1509.978929	-2.061305883	0.998012994	0	nd
ChIPCRM2497	Negative controls	20699.31968	-3.119047388	0.017108023	26148.59592	-2.131234204	0.325889835	0	nd
ChIPCRM3152	Negative controls	7213.931135	3.382111461	7.35E-05	3582.20821	4.063968269	0.001978837	1	nd
ChIPCRM5405	Negative controls	14154.4692	-0.853126915	0.67166971	2.183799999	0.921066814	0.441976808	0	nd
ChIPCRM5432	Negative controls	143864.3852	-1.918459275	0.084226404	138040.4758	-1.9782794	0.252981049	1	nd
ClassI_ins_157	Insulators	1730.656065	2.973147739	0.02539335	3059.555955	2.172861025	0.033629666	1	nd
ClassI_ins_17	Insulators	7035.455573	-1.300062771	0.532812406	14680.17837	-2.559358541	0.382398502	1	nd
ClassI_ins_182	Insulators	22452.15801	-0.396265884	0.844783761	32053.42924	-1.570653525	0.518954685	1	nd
ClassI_ins_196	Insulators	3160.433084	4.06804298	4.14E-06	16774.72813	4.936583123	1.58E-25	1	nd
ClassI_ins_210	Insulators	8017.093553	3.457664747	1.48E-05	24569.56283	3.589781734	1.96E-11	1	nd
ClassI_ins_212	Insulators	8408.8931	-0.209407776	0.968957928	14401.71686	-0.816572985	0.917027471	1	nd

ClassI_ins_244	Insulators	1653.265586	-2.809920825	0.289723773	3584.094523	-4.147359153	0.375602819	1	nd
ClassI_ins_29	Insulators	28416.88166	2.804928687	2.58E-05	13958.41286	4.01328142	3.60E-09	1	nd
ClassI_ins_34	Insulators	1097.091697	0.566064812	0.569351323	3832.182546	-0.672007228	0.899856604	1	nd
ClassI_ins_36	Insulators	578.9702644	-1.563537457	0.968957928	2106.334751	0.690739829	0.840794673	0	nd
ClassI_ins_61	Insulators	98908.30018	0.050905873	0.902866476	110337.927	0.167242604	0.763372053	1	nd
ClassI_ins_77	Insulators	4140.897366	1.325793558	0.228595776	16803.63727	1.221004981	0.545137287	1	nd
ClassI_ins_87	Insulators	3323.155898	3.293518572	0.001805404	3994.493566	4.378210072	2.59E-05	1	nd
ClassI_ins_94	Insulators	3759.804889	-2.259588127	0.41614141	984.2935752	0.259939086	0.663112564	1	nd
Cpr47Ee_CG13222-edge	Nonmeso CRMs	125368.2669	-0.217871157	0.983249707	81213.74974	0.236650796	0.818617115	1	nd
Cyp6g1_construct5	Nonmeso CRMs	14333.11265	-2.162598536	0.146850629	2747.539735	-1.532604373	0.930975503	1	nd
DHS+K27me3_CtBPoverlap_1612	DHS + CtBP	14548.86574	-1.991623854	0.154958419	12990.06613	-3.227088302	0.275927085	0	nd
DHS+K27me3_CtBPoverlap_1813	DHS + CtBP	3039.227757	0.661803411	0.398311539	4938.300924	1.560179259	0.222330426	1	nd
DHS+K27me3_CtBPoverlap_2206	DHS + CtBP	9161.230063	-1.504942187	0.392583715	7907.276491	-2.171662229	0.5226798	0	nd
DHS+K27me3_CtBPoverlap_2267	DHS + CtBP	15.91357201	6.94187924	0.006067118	1.281679754	10.8211586	0.000306658	0	-
DHS+K27me3_CtBPoverlap_2418	DHS + CtBP	20211.04077	-2.352813172	0.048600125	13240.19223	-1.549061926	0.8294821	0	nd
DHS+K27me3_CtBPoverlap_2598	DHS + CtBP	32361.27674	0.853969903	0.300861563	35016.67192	1.339418913	0.348404591	0	nd
DHS+K27me3_intergenic_1490a	Repressive DHSs	7012.862053	2.886217177	0.000608145	6973.821383	2.411735182	0.018488596	0	-
DHS+K27me3_intergenic_1490b	Repressive DHSs	2.346113532	7.433487833	0.00204646	1.014887402	7.243699904	0.002556024	0	-
DHS+K27me3_intergenic_16	Repressive DHSs	18619.96305	-1.262153072	0.437856863	26867.73989	-1.476649667	0.68895687	0	nd
DHS+K27me3_intergenic_1610	Repressive DHSs	13190.09199	-0.827155731	0.595106972	12402.44452	-0.439768984	0.911086129	0	nd
DHS+K27me3_intergenic_2153	Repressive DHSs	21543.63183	-2.319606505	0.103480603	45383.09396	-3.373494557	0.060216784	0	nd
DHS+K27me3_one-hit_1265	Repressive DHSs	10581.08908	3.356480837	6.19E-05	40725.52059	3.24484261	6.39E-10	1	nd
DHS+K27me3_one-hit_1324	Repressive DHSs	322.8400588	1.739920389	0.308018661	224.938829	2.844722609	0.325355254	1	nd
DHS+K27me3_one-hit_1883	Repressive DHSs	4574.809876	0.237782167	0.833415143	5151.971741	-1.858994551	0.702172921	0	nd
DHS+K27me3_one-hit_1994	Repressive DHSs	8.947333812	-2.374746256	0.76048503	7.805049583	-5.106028696	0.993889052	1	nd
DHS+K27me3_one-hit_202	Repressive DHSs	7616.10566	-1.872096925	0.359121812	10637.75455	-1.466827374	0.874339989	1	nd
DHS+K27me3_one-hit_24	Repressive DHSs	24081.98853	-1.318234546	0.359745989	23858.60008	-0.209523736	0.993889052	0	nd
DHS+K27me3_one-hit_2588	Repressive DHSs	2578.092117	-1.756730172	0.602981694	5353.663695	1.915117896	0.462411999	0	nd
DHS+K27me3_one-hit_2606	Repressive DHSs	9898.310788	0.410016782	0.62773821	1471.109039	1.948373211	0.136197399	0	nd
DHS+K27me3_one-hit_2645	Repressive DHSs	51375.71942	-2.1266589	0.076493709	49075.23181	-1.774034738	0.28297218	1	nd
DHS+K27me3_one-hit_75	Repressive DHSs	30457.29109	-2.337286247	0.065445535	28119.51681	-2.058547103	0.424193668	0	nd
Ddc_-0.47	Nonmeso CRMs	66.53178284	6.403962678	0.001977457	73.27061894	6.555277525	0.000595416	1	nd
Ddc_ET	Nonmeso CRMs	17807.39041	0.562146303	0.518186799	12020.67842	0.09063792	0.917027471	1	nd
Ddc_distal_enhancer	Nonmeso CRMs	4904.627506	-1.661241393	0.527870385	23314.56739	-1.336778825	0.842975364	0	nd
Ddc_silencing_element	Nonmeso CRMs	25428.40705	-0.772065057	0.616726707	9432.336632	1.471816988	0.52889471	0	nd
Def_prom	Nonmeso CRMs	125.4318806	2.267715544	0.320321939	8306.227446	-0.890692663	0.701935751	1	nd
DII_215	Nonmeso CRMs	112474.7463	-1.537143205	0.240461029	61272.7139	-1.020211518	0.660781638	0	nd
Doc3_7731	Specific meso CRMs	177.927909	3.213164028	0.150596149	282.1090362	1.553524208	0.433354104	0	nd

E(spl)_m8-0.46	Nonmeso CRMs	4357.419233	4.575393551	1.78E-09	1967.881727	7.091365097	8.97E-25	1	nd
E0_12_GROAviva_ChIP_chip.region_125	Groucho ChIP	584.074524	-0.657648874	0.773085189	5057.518236	-4.148540337	0.20113555	0	nd
E0_12_GROAviva_ChIP_chip.region_166	Groucho ChIP	2724.756207	-1.046898928	0.858729782	3947.395267	-0.626314161	0.769738145	0	nd
E0_12_GROAviva_ChIP_chip.region_179	Groucho ChIP	15853.23628	3.148509084	1.01E-05	12419.23447	2.648257605	0.002902853	0	-
E0_12_GROAviva_ChIP_chip.region_217	Groucho ChIP	13745.57067	-2.430889395	0.081681307	9488.30916	-0.472557343	0.993889052	0	nd
E0_12_GROAviva_ChIP_chip.region_243	Groucho ChIP	225.759454	2.528637296	0.191817188	280.8403341	3.518805352	0.247793177	1	nd
E0_12_GROAviva_ChIP_chip.region_247	Groucho ChIP	22016.97795	-1.541329595	0.320321939	2472.669172	0.31795834	0.647768377	0	nd
E0_12_GROAviva_ChIP_chip.region_263	Groucho ChIP	7395.950513	-1.914688341	0.229129415	4391.905668	-2.56059446	0.459194941	0	-
E0_12_GROAviva_ChIP_chip.region_279	Groucho ChIP	12244.82662	-1.145343504	0.534987927	18140.85621	-0.853061522	0.948832036	0	nd
E0_12_GROAviva_ChIP_chip.region_286	Groucho ChIP	5814.218377	4.455482081	1.27E-09	5441.015127	5.692584775	2.34E-19	1	nd
E0_12_GROAviva_ChIP_chip.region_311	Groucho ChIP	2769.918196	1.120764116	0.477765466	57279.36523	-1.727074145	0.505974937	0	nd
E0_12_GROAviva_ChIP_chip.region_377	Groucho ChIP	33139.49672	-2.266469397	0.097586157	33479.83002	-0.458254546	0.787029697	0	nd
E0_12_GROAviva_ChIP_chip.region_415	Groucho ChIP	239.8864503	-1.980666373	0.902866476	1433.417612	-11.32923139	0.51777849	0	nd
E0_12_GROAviva_ChIP_chip.region_422	Groucho ChIP	33000.06457	-0.402414002	0.986205568	24419.77718	0.501954413	0.856401969	1	nd
E0_12_GROAviva_ChIP_chip.region_43	Groucho ChIP	16954.04777	-3.847465519	0.003840571	17602.02818	-3.692344124	0.11616671	0	-
E0_12_GROAviva_ChIP_chip.region_441	Groucho ChIP	15717.85987	-2.07943724	0.133177096	17293.59792	-1.862112961	0.521096851	0	-
E0_12_GROAviva_ChIP_chip.region_444	Groucho ChIP	11867.86679	-2.984304585	0.041791735	2319.416493	-2.192610923	0.874339989	0	nd
E0_12_GROAviva_ChIP_chip.region_445	Groucho ChIP	5410.481397	-2.044690712	0.281475579	992.0644174	-2.234058407	0.926424235	0	-
E0_12_GROAviva_ChIP_chip.region_467	Groucho ChIP	5886.472818	-1.469214877	0.433802778	13435.58008	-2.535839755	0.348404591	0	nd
E0_12_GROAviva_ChIP_chip.region_468	Groucho ChIP	3155.678364	-0.152379101	0.986205568	17459.15332	0.134904335	0.769738145	0	nd
E0_12_GROAviva_ChIP_chip.region_489	Groucho ChIP	15375.73573	-1.527961433	0.337262811	8598.2111	0.18920169	0.969055253	0	nd
E0_12_GROAviva_ChIP_chip.region_499	Groucho ChIP	23166.72003	-1.116508698	0.47380203	47619.07532	-0.193799956	0.998012994	1	nd
E0_12_GROAviva_ChIP_chip.region_506	Groucho ChIP	17701.87213	-0.532977432	0.833415143	21621.33189	-0.411266526	0.917027471	0	nd
E0_12_GROAviva_ChIP_chip.region_513	Groucho ChIP	20101.21823	-2.397566069	0.089601128	9749.012982	-3.143920565	0.275927085	0	nd
E0_12_GROAviva_ChIP_chip.region_587	Groucho ChIP	2600.465079	-0.350969678	0.966513888	5795.158338	0.373496849	0.971792972	1	nd
E0_12_GROAviva_ChIP_chip.region_88	Groucho ChIP	7706.827114	-2.036132562	0.397457387	315.7968819	1.227978461	0.738808046	0	nd
E_0_12h_dCtBP7667.region_1028	CtBP ChIP	3665.69473	-1.15993617	0.824621392	13402.8223	-2.180099651	0.52889471	0	nd
E_0_12h_dCtBP7667.region_1104	CtBP ChIP	14926.08481	-1.105235541	0.427331701	6941.999726	-4.056685358	0.12063842	0	nd
E_0_12h_dCtBP7667.region_1302	CtBP ChIP	12075.76981	-0.278118224	0.971635655	18299.39016	-0.641043026	0.96722015	0	nd
E_0_12h_dCtBP7667.region_1533	CtBP ChIP	26223.40904	-3.649957116	0.003333133	6697.508296	-0.644801539	0.93973403	0	nd
E_0_12h_dCtBP7667.region_1555	CtBP ChIP	15239.60054	-1.796805122	0.18600653	48102.43221	-1.87817872	0.368525791	0	-
E_0_12h_dCtBP7667.region_1582	CtBP ChIP	12162.48126	-0.565388961	0.585896003	13946.71145	-4.651277592	0.034450608	0	nd
E_0_12h_dCtBP7667.region_1692	CtBP ChIP	23.32071129	5.822380597	0.019393132	218.4568727	4.150364062	0.082105857	0	-
E_0_12h_dCtBP7667.region_1855	CtBP ChIP	36236.45745	-1.465185079	0.3052858	18444.12058	0.161575124	0.874339989	0	nd
E_0_12h_dCtBP7667.region_1981	CtBP ChIP	5691.697392	0.688943099	0.5461777	6103.184083	-0.278656261	0.867737568	0	nd
E_0_12h_dCtBP7667.region_2057	CtBP ChIP	519.844749	1.336288559	0.595106972	4330.351597	-0.432422952	0.969825082	0	nd
E_0_12h_dCtBP7667.region_2243	CtBP ChIP	3254.614927	-0.834990279	0.968957928	13647.36769	-0.304360201	0.96722015	0	nd
E_0_12h_dCtBP7667.region_2256	CtBP ChIP	1474.768428	0.087861831	0.748194692	4762.294613	-2.379099674	0.615957348	0	nd

E_0_12h_dCtBP7667.region_234	CtBP ChIP	1.263832638	9.57337351	0.001791134	8872.030137	-2.566915097	0.459194941	0	-
E_0_12h_dCtBP7667.region_2344	CtBP ChIP	9799.771335	0.659752911	0.587477854	6168.178033	-1.002912878	0.971792972	0	nd
E_0_12h_dCtBP7667.region_258	CtBP ChIP	1749.039033	-2.082189354	0.5522441	1999.001385	-1.816948276	0.96722015	0	nd
E_0_12h_dCtBP7667.region_2713	CtBP ChIP	6814.143246	-2.247035053	0.191558436	16172.62881	-2.256856937	0.263393647	0	nd
E_0_12h_dCtBP7667.region_2748	CtBP ChIP	46580.56521	-2.392518971	0.065445535	36440.30871	-2.78839065	0.312577017	1	nd
E_0_12h_dCtBP7667.region_275	CtBP ChIP	11501.96151	-1.715885773	0.264587738	19203.74194	-0.927435112	0.77842477	0	nd
E_0_12h_dCtBP7667.region_2772	CtBP ChIP	33024.59593	1.118319362	0.218732079	59890.88534	0.271238751	0.77253156	0	nd
E_0_12h_dCtBP7667.region_2793	CtBP ChIP	7334.81777	-1.697775518	0.526503357	2884.878633	2.031213595	0.625473899	0	nd
E_0_12h_dCtBP7667.region_2827	CtBP ChIP	308.4536358	0.038117818	0.731070839	9541.535929	-2.097402119	0.763372053	0	nd
E_0_12h_dCtBP7667.region_2926	CtBP ChIP	1285.326762	-0.583079469	0.968957928	28253.94327	-2.656577219	0.379154116	0	nd
E_0_12h_dCtBP7667.region_2952	CtBP ChIP	45837.68414	-1.632694757	0.228595776	34464.31083	-1.362786761	0.585652031	1	nd
E_0_12h_dCtBP7667.region_2991	CtBP ChIP	8196.464047	-0.98307614	0.602981694	7073.58385	-1.12961641	0.77842477	0	nd
E_0_12h_dCtBP7667.region_3031	CtBP ChIP	6819.140989	-2.162397673	0.281475579	9331.652415	-1.631870148	0.769738145	0	nd
E_0_12h_dCtBP7667.region_3049	CtBP ChIP	23253.96432	2.052502724	0.006785427	22492.50077	2.180959035	0.029031879	0	+
E_0_12h_dCtBP7667.region_3084	CtBP ChIP	432.6940861	3.023173145	0.105389871	628.9351238	3.706209461	0.02170128	0	+
E_0_12h_dCtBP7667.region_3137	CtBP ChIP	2793.778306	0.023041779	0.731070839	4850.443796	0.569518802	0.51777849	1	nd
E_0_12h_dCtBP7667.region_3202	CtBP ChIP	17664.47194	-2.813777876	0.020397884	12730.4464	-2.824566275	0.333937164	0	nd
E_0_12h_dCtBP7667.region_3249	CtBP ChIP	1310.822206	2.246406732	0.155865332	5452.4453	-1.873573896	0.702172921	0	nd
E_0_12h_dCtBP7667.region_3325	CtBP ChIP	2981.892789	0.964635813	0.66684265	0.769969193	2.622661602	0.154234039	0	nd
E_0_12h_dCtBP7667.region_3402	CtBP ChIP	18218.79533	-1.25857831	0.361366116	13970.90873	-2.079963493	0.590395623	0	nd
E_0_12h_dCtBP7667.region_3439	CtBP ChIP	4022.24969	-4.362383953	0.023163058	358.1559049	-2.141725649	0.911086129	0	nd
E_0_12h_dCtBP7667.region_3445	CtBP ChIP	6335.436035	-0.531758044	0.865949609	17226.48749	-0.515758737	0.911086129	0	nd
E_0_12h_dCtBP7667.region_3446	CtBP ChIP	8432.551507	-0.791562222	0.747982496	20868.26958	-1.181059375	0.977932362	0	nd
E_0_12h_dCtBP7667.region_3484	CtBP ChIP	7889.383254	-3.51486757	0.037985388	16448.70464	-2.471829445	0.186881073	0	nd
E_0_12h_dCtBP7667.region_3509	CtBP ChIP	16339.44668	-2.120973483	0.099386597	10937.43492	-2.474219812	0.57121765	0	-
E_0_12h_dCtBP7667.region_3525	CtBP ChIP	11158.3886	-0.569627485	0.833415143	1331.344092	-0.852064336	0.851205172	0	nd
E_0_12h_dCtBP7667.region_3540	CtBP ChIP	6029.084629	-0.185926687	0.823281392	7030.483561	-1.788792321	0.691660514	0	nd
E_0_12h_dCtBP7667.region_3547	CtBP ChIP	1552.569112	-1.937964533	0.532812406	2047.550272	-2.922261848	0.917027471	0	nd
E_0_12h_dCtBP7667.region_3553	CtBP ChIP	7425.425013	-4.122868117	0.020176094	13268.23402	-3.483177035	0.071878785	0	nd
E_0_12h_dCtBP7667.region_3606	CtBP ChIP	2028.156804	-1.546151485	0.615300876	3139.951795	-3.166460601	0.56464406	0	nd
E_0_12h_dCtBP7667.region_3696	CtBP ChIP	28399.60366	-1.693809354	0.239263706	6456.063286	-0.288142909	0.977932362	0	nd
E_0_12h_dCtBP7667.region_3770	CtBP ChIP	5910.56937	-2.894770753	0.150596149	1474.11988	-4.070008169	0.917027471	1	nd
E_0_12h_dCtBP7667.region_3937	CtBP ChIP	485.2823979	1.383402618	0.477765466	713.0247369	1.066684841	0.777760833	0	nd
E_0_12h_dCtBP7667.region_3938	CtBP ChIP	5114.42411	-1.910873827	0.398956288	3936.40353	0.895388799	0.524693706	0	nd
E_0_12h_dCtBP7667.region_4196	CtBP ChIP	31530.08739	1.795583915	0.056198937	44980.72199	2.536180781	0.000569748	0	-
E_0_12h_dCtBP7667.region_4460	CtBP ChIP	1516.16226	0.611668745	0.589787163	526.4567209	2.469923433	0.510535262	1	nd
E_0_12h_dCtBP7667.region_4503	CtBP ChIP	4590.213581	-1.014377772	0.754081699	11289.33931	0.484060614	0.985470876	0	nd
E_0_12h_dCtBP7667.region_4530	CtBP ChIP	11.63930866	7.399246972	0.004548461	4.321229093	8.756115529	0.002824248	0	-

E_0_12h_dCtBP7667.region_4738	CtBP ChIP	3003.645203	0.804336589	0.616726707	7035.780176	-0.383916216	0.874339989	0	nd
E_0_12h_dCtBP7667.region_4844	CtBP ChIP	10464.85721	0.371993274	0.84037128	9650.847663	0.399088269	0.865316701	0	nd
E_0_12h_dCtBP7667.region_518	CtBP ChIP	632.35223	-3.183421481	0.80733185	851.0091085	-1.799735404	0.874339989	0	nd
E_0_12h_dCtBP7667.region_707	CtBP ChIP	25424.54796	3.933654247	2.36E-10	9218.973228	5.554831893	8.97E-25	1	nd
E_0_12h_dCtBP7667.region_864	CtBP ChIP	7616.105832	0.646769084	0.49699173	9765.947994	-0.588058515	0.866769125	0	nd
E_0_12h_dCtBP7667.region_991	CtBP ChIP	150.7938126	2.171932813	0.283502942	15.14425283	3.081987202	0.173799421	0	nd
Ecoli_control11	Negative controls	14373.79168	-1.159130743	0.526503357	31546.76078	-1.520368592	0.723016534	0	nd
Ecoli_control22	Negative controls	5447.978969	1.983163804	0.04460283	7302.432087	3.070410103	0.000553431	0	-
Ecoli_control23	Negative controls	6725.859428	-2.454104789	0.230586299	10829.31977	-1.82488754	0.49976949	0	nd
Ecoli_control24	Negative controls	5512.009519	-1.014624274	0.569351323	1442.016961	-7.158755206	0.677274607	0	nd
Ecoli_control3	Negative controls	6679.155545	-2.162946824	0.25462274	7224.767626	-1.621917409	0.867737568	0	nd
Ecoli_control5	Negative controls	11583.42715	-2.593263078	0.099250218	5518.903293	0.348097141	0.979402418	0	nd
Eip71CD_188	Nonmeso CRMs	28960.00162	-2.378634952	0.064111401	4411.639305	-0.216160375	0.978966087	1	nd
Est-6_D-511	Nonmeso CRMs	52.81601644	6.407677112	0.001977457	21.63429936	6.839074967	0.007671457	1	+
Fad2_oe1	Nonmeso CRMs	5.552181013	6.695211496	0.006067118	4.847109452	6.370298668	0.024001862	1	nd
Fas2_540bp_CRM	Nonmeso CRMs	5781.189334	-1.060399205	0.747982496	7189.610811	-1.243545741	0.723016534	0	nd
HLHm5_m5-0.13	Nonmeso CRMs	10802.84122	1.627213916	0.103480603	11979.99979	2.349588613	0.034450608	1	nd
Hs6st_3748	Nonmeso CRMs	11676.42272	-2.151221673	0.229129415	3311.938426	-2.954559099	0.656402009	0	nd
Hsp26_nurse_cell_enhancer	Nonmeso CRMs	2765.491889	-2.210793647	0.40679447	3250.975308	-6.12830869	0.163844378	1	nd
Kr_HBg0.6HZ	Specific meso CRMs	15519.85446	-2.163461363	0.120038907	8392.745448	-0.978654084	0.807961721	0	nd
Kr_delBNc0.8HZ	Nonmeso CRMs	16058.15212	-1.056097566	0.528750879	13819.13255	-0.424022722	0.993889052	0	nd
Lip1_prom	Nonmeso CRMs	1124.620287	1.202807872	0.368914552	660.5463572	-0.098842537	0.702172921	1	nd
Mst84Db_-286_+154	Nonmeso CRMs	8435.446877	-0.975861454	0.592337248	8175.865323	-1.245345853	0.947390077	1	nd
PH4alphaSG2_SG2-885	Nonmeso CRMs	40348.21702	0.698281524	0.433802778	41226.46837	-0.899858659	0.967071039	1	nd
Poxn_2	Nonmeso CRMs	1254.982817	-2.253735374	0.640874877	1205.5862	-1.873294348	0.93973403	0	nd
Poxn_9	Nonmeso CRMs	20398.104	-0.268182457	0.967251836	3065.050705	1.653425814	0.521096851	1	nd
Rbp4_pPa104_69	Nonmeso CRMs	2124.593525	-3.05503392	0.210092146	850.6592113	1.314680848	0.647768377	1	nd
Rh3_promoter	Nonmeso CRMs	12.66208275	2.633645876	0.113340764	2.312040337	6.578634576	0.011269571	1	nd
Rh5_promoter	Nonmeso CRMs	18.08746439	4.780349143	0.032890684	49.5511772	5.72870109	0.017534837	1	nd
Rh6_-555_+121	Nonmeso CRMs	4326.791745	0.24765659	0.84037128	1948.203801	-0.708388185	0.874339989	1	nd
Sgs4_OPSL	Nonmeso CRMs	89.20212762	2.454538678	0.200857089	49.71651548	2.929712818	0.286329637	1	nd
Sox15_regionA	Nonmeso CRMs	2360.011333	-0.730838771	0.96828369	4492.775039	0.2379679	0.76750583	0	nd
Sox15_regionB	Nonmeso CRMs	4.339862004	6.039641314	0.003867997	2.49290484	9.093544009	0.006858884	0	-
Z600_z600-lacZ	Nonmeso CRMs	24.62825873	8.46422165	9.02E-06	9.072445656	10.74382656	2.13E-17	1	nd
ac_0.8	Nonmeso CRMs	291.2510727	4.94942675	0.00204646	115.8580661	6.269126678	0.000515588	1	nd
amosB	Nonmeso CRMs	25466.24188	-0.024535731	0.902866476	55665.92221	-0.496481852	0.983843193	0	nd
ato_RE	Nonmeso CRMs	28961.16784	-0.914372282	0.604032255	19431.68015	-1.010292436	0.674358049	0	nd
bcd_49bg-Z	Nonmeso CRMs	2321.96393	-1.34697621	0.778549435	1231.215612	-2.734506766	0.995277602	1	nd

bib_5924	Specific meso CRMs	12192.52373	-3.453087297	0.015140124	13383.55382	-2.862955184	0.3438437	0	nd
brk_NEE-long	Nonmeso CRMs	12699.51764	1.197649651	0.082033809	11507.35938	2.86574562	0.002824248	0	+
btI_P[B4]	Nonmeso CRMs	4.914825216	-1.553053493	0.688670921	4.792029288	8.905689472	0.000859688	1	nd
cas_csc-1	Nonmeso CRMs	32032.05479	0.693882568	0.56709847	20826.956	0.94658891	0.487069312	0	nd
cas_csc-2	Nonmeso CRMs	0.64432281	-0.961667628	0.590922838	4509.506816	2.249781289	0.163844378	0	nd
cas_csc-5c	Nonmeso CRMs	0.370343616	1.041154465	0.531469603	5563.696365	1.369637455	0.382398502	0	nd
cas_csc-6	Nonmeso CRMs	1948.935017	1.059446582	0.426214533	13735.93416	-1.031149831	0.842975364	0	nd
cas_csc-8	Nonmeso CRMs	88337.32151	0.037187899	0.889698896	67032.11111	0.102878784	0.830296122	1	nd
ct_340	Nonmeso CRMs	4252.617198	-10.68246119	4.11E-08	933.0845835	-1.829412855	0.948832036	0	nd
ct_ct-3	Nonmeso CRMs	7.28794074	5.897079201	0.011750425	4.791725361	4.244609265	0.048650774	0	-
ct_wingmargin_Guss	Nonmeso CRMs	11865.54784	-1.070208137	0.599919282	2793.744222	-1.356776656	0.977932362	0	nd
dap_dap-BB	Nonmeso CRMs	52861.49895	-1.411482034	0.293599947	48944.49721	-0.891572371	0.690101954	0	nd
dap_dap-del	Nonmeso CRMs	5969.984667	1.931678043	0.109626008	9636.866414	0.817270327	0.521096851	0	nd
dap_dap3'2'	Nonmeso CRMs	3018.472501	-2.08795146	0.398956288	1218.118004	-0.106253378	0.93973403	0	nd
dj_dj-promoter	Nonmeso CRMs	16423.28144	-1.682632525	0.289723773	196.3296553	0.87488978	0.515024261	1	nd
dpn_dI	Nonmeso CRMs	1833.363198	-1.496719099	0.714429159	1645.422271	-0.782699633	0.80502285	0	nd
dpp_85.8MX	Nonmeso CRMs	1215.869731	3.038617295	0.036393898	18144.75682	1.712861229	0.173799421	0	+
dpp_dpoho	Nonmeso CRMs	3329.978193	0.713949352	0.621961058	14347.9946	-1.79570191	0.659117203	0	nd
dys_N283	Nonmeso CRMs	12184.39204	-2.442314297	0.096748963	8801.032328	-1.136738375	0.913476966	0	nd
e_A.1	Nonmeso CRMs	18363.1154	-1.778613093	0.155865332	39172.1944	-0.946153724	0.769738145	1	nd
e_A.2	Nonmeso CRMs	55238.97007	-2.055111293	0.102772266	42193.99862	-2.066563988	0.433391685	0	-
e_A.3	Nonmeso CRMs	11368.4505	-1.539113815	0.337262811	2589.966883	-0.893994672	0.855260792	0	nd
e_A.5	Nonmeso CRMs	20602.94431	-1.969028472	0.191817188	3945.537345	-0.384820522	0.911086129	0	nd
e_coreAbdominalCRE	Nonmeso CRMs	18507.7074	0.62746079	0.320321939	7911.883493	2.684606651	0.002513616	0	+
elav_construct_L	Nonmeso CRMs	60544.30739	0.146944508	0.731643891	81708.21279	1.153779256	0.135156663	1	nd
ems_elementIV	Nonmeso CRMs	3297.075037	-3.55562423	0.121582954	3098.801023	-0.676573145	0.856401969	0	nd
ems_ems_ARFE-subA	Nonmeso CRMs	49733.51846	-2.959427201	0.023848198	23115.03669	-1.813689721	0.441976808	0	nd
eve_MHE	Specific meso CRMs	7773.333574	-0.608745505	0.891580716	3987.147854	1.418567774	0.459194941	0	nd
ey_5D11	Nonmeso CRMs	18189.6025	-2.843559679	0.029606039	12596.64314	-1.631052058	0.8294821	0	nd
ey_UE0.8	Nonmeso CRMs	884.5262636	-0.520883173	0.833415143	893.3226758	-1.544493274	0.874339989	0	nd
ey_UE0.9	Nonmeso CRMs	12378.6668	2.823141237	0.000194139	9938.437692	3.68034313	2.59E-05	1	+
fkx_salivary_gland_enhancer	Nonmeso CRMs	1334.756458	-0.952338311	0.976073054	1254.393331	-3.121362509	0.993889052	1	nd
ftz_-669_-386	Nonmeso CRMs	28957.59169	-1.626821143	0.22229563	69240.90139	-2.772699112	0.163844378	0	nd
ftz_5'delta276	Nonmeso CRMs	3924.956148	4.145359317	2.43E-07	2948.347489	5.404606423	2.11E-10	1	nd
fz2_Lac-Z	Specific meso CRMs	18585.99822	-1.382855101	0.354680589	22761.51275	-1.083257426	0.917027471	1	nd
gcm_+3.8_+4.5	Nonmeso CRMs	6599.304528	1.131700974	0.161936956	5046.094883	2.131420573	0.088032379	0	-
gsb_GLE	Nonmeso CRMs	18466.88466	1.008724452	0.396382279	37064.77607	0.176949295	0.993889052	0	nd
gsb_fragIV	Nonmeso CRMs	80577.3472	0.922357629	0.239263706	73526.36767	2.371336542	2.17E-05	0	+

gt_gt1	Nonmeso CRMs	30500.40707	0.802089526	0.317990535	26301.34626	-1.397039955	0.639058468	0	nd
h_302	Nonmeso CRMs	6175.413502	-2.700162056	0.116106607	4275.49603	-0.318074559	0.993889052	0	nd
h_HHRE	Nonmeso CRMs	31.12257508	6.861019926	0.002152548	26.27221072	7.658352754	0.000196374	0	-
h_stripe0	Nonmeso CRMs	452.5558073	2.599898334	0.21391744	16006.91853	-2.227293152	0.415492467	0	nd
h_stripe1	Nonmeso CRMs	34507.3169	1.261659137	0.116106607	16788.79525	1.485042634	0.275927085	0	nd
h_stripe3_ET38	Nonmeso CRMs	3431.055451	-0.155273501	0.925377813	3281.881408	-0.627362253	0.76750583	0	nd
h_stripe_6+2	Nonmeso CRMs	2451.352213	0.695881563	0.56709847	13570.12135	-1.092156826	0.77253156	0	nd
hb_matDm0.5-lacZ	Nonmeso CRMs	1.197318925	11.14138061	0.000178031	4825.84231	2.580114126	0.198350306	1	nd
hh_4075	Nonmeso CRMs	3380.057339	3.523875454	0.00015333	7088.645785	2.759801841	0.003025094	0	-
hh_alpha_fragment_(ic-CRE)	Nonmeso CRMs	23683.95767	-1.582593383	0.259685072	35372.65907	-1.208386167	0.50202832	0	nd
hh_bar3L2	Nonmeso CRMs	565.8427631	-0.523381376	0.757721334	7383.565268	-0.994291421	0.865316701	0	nd
hkb_0.6kbRIRV	Nonmeso CRMs	16382.93799	3.027348524	7.35E-05	68334.50987	2.378064304	6.59E-06	0	+
ind_moduleA	Positive controls	1570.138746	2.356180669	0.105389871	11879.76291	2.241737044	0.071878785	0	+
ins_Fab-8	Insulators	8854.737161	0.88334988	0.337964059	30255.22263	-3.005709431	0.157325461	0	nd
kni_AE20	Nonmeso CRMs	35797.04679	-1.301059642	0.3052858	37965.94883	-0.617889203	0.769738145	0	nd
kni_KD	Nonmeso CRMs	12955.50663	0.451338328	0.62773821	4217.877076	1.263238981	0.433391685	0	nd
kni_proximal_minimal	Nonmeso CRMs	2.416890956	1.24558214	0.226064135	16039.57246	-2.546484266	0.433391685	1	nd
kni_reporter_fragment_EC	Nonmeso CRMs	9002.761628	-3.320953946	0.034057607	1732.960163	-2.254314507	1	0	-
lab_1.0	Nonmeso CRMs	5445.637416	-0.725284747	0.795915842	6221.121646	-1.72859767	0.702172921	0	nd
lab_HZ550	Nonmeso CRMs	8.895836807	4.033465645	0.046782335	3.760501477	5.733723666	0.019924591	0	-
lz_CrystalCellEnhancer1236-737	Nonmeso CRMs	3438.932847	2.776287718	0.0072776	5439.011195	-0.263533017	0.855260792	0	+
mib2_FCenhancer	Specific meso CRMs	57422.50823	-1.334438225	0.387390921	34221.77954	-1.661844858	0.590395623	0	nd
nAcRbeta-64B_P-171	Nonmeso CRMs	22061.1381	-2.676615507	0.033308583	44568.5796	-1.76140491	0.408648122	1	nd
nab_nab-1	Nonmeso CRMs	9148.212823	-0.057856846	0.973588568	4858.49372	-0.542633088	0.977932362	0	nd
nerfin-1_fragment14	Nonmeso CRMs	7083.448643	3.891890961	1.21E-07	11420.50945	5.608993009	5.93E-34	1	nd
nerfin-1_fragment4	Nonmeso CRMs	24717.97142	0.504325952	0.616726707	44506.33363	0.169712241	0.673405229	0	nd
nerfin-1_fragment5	Nonmeso CRMs	44622.98015	0.135891054	0.795915842	77630.29185	-1.096623466	0.818617115	0	nd
ninaE_proximal_promoter_region	Nonmeso CRMs	6483.140569	3.66382695	5.40E-07	825.1565772	6.132064447	7.97E-06	1	nd
nkd_8756	Nonmeso CRMs	14156.96012	-2.745335516	0.060950418	27731.91757	-2.195281572	0.51777849	0	nd
nkd_IntE_255	Nonmeso CRMs	13867.06658	-3.392446677	0.019393132	58618.95209	-2.448247442	0.198350306	0	nd
nos_-708_+20	Nonmeso CRMs	40662.11177	-1.299252084	0.274471334	62825.96927	-0.765810959	0.856401969	1	nd
numb_5870	Specific meso CRMs	27061.36296	-1.718656909	0.200857089	39547.57138	-1.869478831	0.44265801	0	nd
nvyr_CRM100	Nonmeso CRMs	33487.88524	5.128244563	3.47E-19	33771.2252	5.977835795	3.62E-159	1	nd
oc_SBg	Nonmeso CRMs	1829.390396	2.830619701	0.036263957	4771.734715	3.263186085	0.002195469	0	+
oc_oc7	Nonmeso CRMs	12013.97543	-2.533120551	0.094799994	3214.84676	-0.419177137	0.977932362	0	nd
oc_otd-186	Nonmeso CRMs	51498.14225	1.53633552	0.031026102	50704.26132	2.309756895	0.000429076	0	+
org-1_HN39	Specific meso CRMs	1398.644023	-1.081889619	0.734089966	1619.457165	-7.663170067	0.51777849	0	nd
otp_P	Nonmeso CRMs	143.2257283	-3.297529271	0.881141655	5185.637653	-0.560606514	0.870386951	0	nd

ovo_E2	Nonmeso CRMs	140.7100904	4.454867535	0.032890684	175.2490961	4.618356615	0.024025987	0	-
ple_995bp_wound_response_element	Nonmeso CRMs	6355.948503	4.203459915	6.85E-09	11818.62921	4.333162312	8.99E-11	1	nd
ple_WE1	Nonmeso CRMs	60306.90853	-1.610667785	0.188608539	87112.21408	-1.661994327	0.348404591	0	nd
prd_Pstripe_enhancer	Nonmeso CRMs	7184.793696	-0.852839611	0.795915842	6223.909808	-0.963818068	0.818959784	0	nd
prd_deltaQ	Nonmeso CRMs	1120.346396	4.485501075	0.000358087	1962.907077	2.964137413	0.058148278	1	nd
prd_stripe1_enhancer	Nonmeso CRMs	15555.93366	-2.125404186	0.151856821	17945.56128	-1.155556872	0.855260792	0	nd
repo_-1.1	Nonmeso CRMs	794.6892587	2.606494931	0.048672947	21978.03525	1.565969144	0.09515724	0	+
repo_pBJ-111	Nonmeso CRMs	3105.443539	-1.881989975	0.457180374	35592.18167	-0.695141711	0.985470876	0	nd
repo_pBJ-145	Nonmeso CRMs	2323.247224	-0.583551446	0.955630731	1217.464814	0.417191076	0.505974937	0	nd
run_neural_6GB	Nonmeso CRMs	1045.407858	-1.606766627	0.96828369	0.376977184	2.762964531	0.021927995	0	-
sev_minimal_enhancer	Nonmeso CRMs	2277.574181	-0.049946454	0.795915842	4394.77233	-1.601509744	0.985470876	1	nd
sim_mesectoderm	Nonmeso CRMs	3050.452275	0.075806515	0.833415143	2433.321475	1.470907898	0.522883573	0	nd
sim_st10	Nonmeso CRMs	9315.057727	-1.176622091	0.518186799	3031.525599	-0.306297165	0.911086129	0	nd
slou_SK16	Specific meso CRMs	6842.647123	-4.649715141	0.005959655	694.1740599	-2.414753858	0.917027471	0	nd
slou_SK19	Specific meso CRMs	2893.117854	-1.366635117	0.821520598	946.3357681	0.641265695	0.593003198	0	nd
slp2_i4753	Nonmeso CRMs	3536.04246	1.731278999	0.112938498	8264.418528	0.568341638	0.867737568	0	nd
sls_Ket-3_Lac-Z	Specific meso CRMs	628.6392803	0.172898295	0.695210371	324.5169595	2.590409079	0.154885936	0	nd
ss_E2.0_522	Nonmeso CRMs	788.7103325	-0.719939077	0.821520598	7041.228142	-4.099597287	0.186388978	0	nd
ss_E2.0_531	Nonmeso CRMs	4007.259011	-2.676112752	0.234304632	33295.28576	-1.677718419	0.590395623	0	nd
ss_P732	Nonmeso CRMs	19120.55915	-2.640966095	0.054150253	9849.998072	-2.547583094	0.453238208	0	nd
ths_Neu4_early_embryonic_enhancer	Nonmeso CRMs	2906.447007	2.558655846	0.019393132	1.201950137	4.793794491	0.023415449	0	+
tin_tin103A	Specific meso CRMs	28841.2942	-1.639598265	0.238499057	8330.777369	-1.532331616	0.769738145	0	nd
tin_tin103C	Specific meso CRMs	6159.325534	1.844548346	0.048076369	4988.316565	1.246504364	0.647768377	0	-
tll_K11	Nonmeso CRMs	28435.74111	0.203995362	0.763141791	47313.56742	-0.665588814	0.874339989	0	nd
tll_K7	Nonmeso CRMs	3671.540875	0.208827288	0.8689282	70329.69963	-1.490639084	0.594936531	0	nd
tll_O-E	Nonmeso CRMs	1.326687722	0.638634821	0.466054805	5584.806669	1.640635993	0.163844378	0	nd
toy_EEP	Nonmeso CRMs	15964.19878	-2.372795826	0.105389871	24254.32481	-1.987527344	0.372915884	0	nd
trh_trh24	Nonmeso CRMs	52982.34031	-1.473387631	0.155865332	72301.05278	-2.097855808	0.348404591	0	nd
trh_trh45	Nonmeso CRMs	21264.08341	-1.388898235	0.303607209	83402.48782	-2.11721191	0.298783415	0	nd
vas_construct16	Nonmeso CRMs	8.998271679	7.058603194	0.006067118	7.271525318	7.85660734	0.010594154	0	-
vg_minimal_boundary_enhancer	Nonmeso CRMs	62042.33719	-0.752454403	0.569351323	121717.2079	-1.010151678	0.639058468	0	nd
vg_quadrant_enhancer	Nonmeso CRMs	18315.84053	-2.113338374	0.120038907	9919.426621	-2.24965962	0.379154116	0	nd
vn_NEE-long	Nonmeso CRMs	13717.40521	1.415152957	0.205684707	8935.156406	1.804105082	0.465739559	0	nd
vnd_NEE	Nonmeso CRMs	431.4183494	4.157019162	0.018732787	1120.250985	4.126656471	0.026534654	0	+
vv1_484-5prime	Nonmeso CRMs	2013.730387	4.265929335	1.01E-05	7820.662951	5.18179902	1.54E-16	1	nd
vv1_587dfr	Nonmeso CRMs	3940.626477	-2.042145532	0.445456502	6732.046897	-0.967778126	0.782475204	0	nd
vv1_vv10.9	Nonmeso CRMs	7005.514263	0.483395307	0.723728155	2272.769085	-4.846360698	0.518954685	0	nd
vv1_vv1ds1.0	Nonmeso CRMs	1.642288214	5.565393301	0.001985609	3.013210483	6.466748971	0.018349621	0	-

zen_dorsal_ectoderm	Positive controls	3654.140805	0.649316525	0.5522441	1489.384011	2.020471457	0.281060219	0	nd
BiTS-ChIP_K4me1+K27me3+K27ac_365	"Bivalent" chromatin	965.7215042	0.808284919	0.635813946	nd	nd	nd	1	nd
CG13196_1kb_5'	Nonmeso CRMs	10817.83753	-1.038153699	0.62773821	nd	nd	nd	1	nd
CG3492_CRM17	Nonmeso CRMs	2991.023346	-1.922266095	0.308018661	nd	nd	nd	0	nd
ClassI_ins_103	Insulators	1.4257136	7.026133648	0.001309368	nd	nd	nd	1	nd
ClassI_ins_132	Insulators	132.5317229	1.98159494	0.228595776	nd	nd	nd	0	nd
Crz_380gal4	Nonmeso CRMs	2.61399287	6.777765335	0.001985609	nd	nd	nd	1	nd
DHS+K27me3_one-hit_372	Repressive DHSs	231.0397621	1.533671374	0.359121812	nd	nd	nd	0	nd
E0_12_GROAviva_ChIP_chip.region_287	Groucho ChIP	4039.687268	2.514639669	0.01367479	nd	nd	nd	1	nd
E0_12_GROAviva_ChIP_chip.region_376	Groucho ChIP	7744.23032	-1.407343542	0.433802778	nd	nd	nd	0	nd
E0_12_GROAviva_ChIP_chip.region_9	Groucho ChIP	2004.601438	-0.456180339	0.867841148	nd	nd	nd	0	nd
E_0_12h_dCtBP7667.region_1752	CtBP ChIP	1539.145129	-3.698317613	0.266168314	nd	nd	nd	0	nd
E_0_12h_dCtBP7667.region_2659	CtBP ChIP	10.08412558	7.364909395	0.003867997	nd	nd	nd	0	-
E_0_12h_dCtBP7667.region_3629	CtBP ChIP	1743.452529	-1.151955619	0.881141655	nd	nd	nd	0	nd
Ecoli_control21	Negative controls	2520.701781	-2.106165709	0.526503357	nd	nd	nd	0	nd
dpp_P1delta4	Nonmeso CRMs	1887.023556	1.232665225	0.188608539	nd	nd	nd	1	nd
dys_M269+dys_V345	Nonmeso CRMs	7533.339676	-1.530422218	0.457378479	nd	nd	nd	0	nd
kirre_-1.0	Specific meso CRMs	1.753309523	-1.280070417	0.788553316	nd	nd	nd	1	nd
nerfin-1_fragment2	Nonmeso CRMs	1799.165257	-4.331673	0.151856821	nd	nd	nd	0	nd
proPO-A1_F6	Nonmeso CRMs	2328.655457	-0.14681	0.96828369	nd	nd	nd	0	nd
tll_D3	Nonmeso CRMs	1074.501927	6.30327236	9.84E-15	nd	nd	nd	1	nd
CAD2_desat1	Specific meso CRMs	nd	nd	nd	1996.978851	-10.97602174	0.154885936	1	nd
CG32111_8084	Specific meso CRMs	nd	nd	nd	3482.152814	-0.139859143	0.930975503	0	nd
ClassI_ins_114	Insulators	nd	nd	nd	1.058340411	8.922731002	0.004494333	1	nd
DHS+K27me3_one-hit_2147	Repressive DHSs	nd	nd	nd	266.1365226	3.752527773	0.368554858	1	nd
DHS+K27me3_one-hit_877	Repressive DHSs	nd	nd	nd	261.6115	-8.203336662	0.995277602	0	nd
E0_12_GROAviva_ChIP_chip.region_131	Groucho ChIP	nd	nd	nd	5159.539014	-1.340957182	0.77842477	0	nd
E0_12_GROAviva_ChIP_chip.region_233	Groucho ChIP	nd	nd	nd	1019.02529	1.878401057	0.244495927	0	nd
E0_12_GROAviva_ChIP_chip.region_481	Groucho ChIP	nd	nd	nd	5.620061535	8.704796472	0.002513616	0	-
E_0_12h_dCtBP7667.region_1767	CtBP ChIP	nd	nd	nd	1975.573307	-2.461709439	0.874339989	0	nd
E_0_12h_dCtBP7667.region_2824	CtBP ChIP	nd	nd	nd	2852.144176	-2.558802746	0.631210449	0	nd
E_0_12h_dCtBP7667.region_3104	CtBP ChIP	nd	nd	nd	3380.476824	-1.535284181	0.995277602	0	nd
Ecoli_control15	Negative controls	nd	nd	nd	2942.184853	-2.985152006	0.685452904	0	nd
Sry-alpha_CAHBG	Nonmeso CRMs	nd	nd	nd	4693.021192	2.372311998	0.041712211	1	nd
TI_TI287	Specific meso CRMs	nd	nd	nd	1680.346258	-3.148013398	0.818617115	0	nd
btI_P[B23]	Nonmeso CRMs	nd	nd	nd	1650.535144	-1.282216929	0.890585268	0	nd
dpp_980-6	Nonmeso CRMs	nd	nd	nd	1093.499714	3.211571273	0.199533149	1	nd
otp_C	Nonmeso CRMs	nd	nd	nd	1835.745654	-3.136265898	0.870386951	0	nd

ovo_del-ap-del-5	Nonmeso CRMs	nd	nd	nd	172.2141091	4.104571661	0.041712211	0	-
prd_cc_repressor	Nonmeso CRMs	nd	nd	nd	2525.350952	-2.50565619	0.807961721	0	nd
sens_sensCRM3	Nonmeso CRMs	nd	nd	nd	0.551054835	1.299462144	0.298783415	0	nd
sfl_Lac-Z	Specific meso CRMs	nd	nd	nd	0.559931735	-Inf	0.77253156	0	nd
svp_sce	Specific meso CRMs	nd	nd	nd	2813.838472	-11.28235168	0.010473714	0	-
vas_96bpEnhancer	Nonmeso CRMs	nd	nd	nd	2794.33553	1.248810409	0.167522481	1	nd

## Supplementary table 4

Results of silencer-FACS-Seq experiments for the second library.

For each element confidently detected in either complete repetition of the sFS experiment (see Materials and Methods for a precise description of confident detection), the source (type of sequence chosen for testing) is shown; mean abundance in input cells across three biological replicates,  $\log_2$  of the fold change (enrichment) in CD2<sup>+</sup>GFP<sup>reduced</sup> cells, and adjusted p-value for enrichment/depletion are shown for two complete experimental repetitions (nd: not detected).

element	repeat	sig.enr	Exp1 CD2+GFPred vs. CD2+			Exp2 CD2+GFPred vs. CD2+		
			Input abundance	log2FoldChange	padj	Input abundance	log2FoldChange	padj
BiTS-ChIP_K4me1+K27me3+K27ac_13	0	0	5553.18709	0.5898193	0.82436539	6166.52588	0.99807734	0.64790826
BiTS-ChIP_K4me1+K27me3+K27ac_134	0	0	17066.1707	-1.913157	0.25833936	30565.1129	-0.7698429	0.8300674
BiTS-ChIP_K4me1+K27me3+K27ac_73	0	0	8633.32878	0.37143904	0.83882888	4293.27731	-0.2806488	0.9113662
CAD2_Ket-1	0	0	5701.73899	-2.6594431	0.18642569	778.169043	-8.8033801	1
CAD2_Meso-CRM-4726	0	0	5456.01196	-1.6654306	0.46338733	4559.82008	0.19041621	0.9113662
CAD2_Meso-CRM-6028	0	0	13279.1935	-1.9172449	0.25833936	13233.3073	-2.0974548	0.67702347
CAD2_Meso-CRM-6225	0	0	2868.65085	0.99946831	0.53548527	9761.50444	0.36144404	0.89219364
CAD2_htl	0	0	16909.5134	-0.3048985	0.99414843	35351.0736	-0.4840614	0.91542276
CG42342_3436	0	0	777.824144	0.35839738	0.85825274	5049.79443	0.6371247	0.85665548
CG7722_CRM28	0	0	36230.1739	0.98548956	0.30391872	24479.7498	1.01198255	0.31014309
ChIPCRM2078	0	0	7583.74579	-11.200406	5.14E-10	7310.50048	-4.4736032	0.85356213
ChIPCRM5792	0	0	22354.9182	-1.2541242	0.29091979	nd	nd	nd
Classl_ins_198	0	0	3514.19476	1.51997179	0.29031158	2062.93214	0.46635989	0.8300674
DHS+K27me3_CtBPOverlap_2215	0	0	15715.6736	-0.1315526	0.93102677	39454.9094	-0.5736355	0.91542276
DHS+K27me3_CtBPOverlap_762	0	0	7526.08195	-0.9438075	0.58340858	3236.49964	-0.0327949	0.91542276
DHS+K27me3_intergenic_1932	0	0	51366.5408	-0.2478255	0.85223445	53874.4128	-0.6746693	0.9113662
DHS+K27me3_intergenic_2074b	0	0	12262.1442	-2.5121973	0.12802238	14312.2885	0.0785103	1
DHS+K27me3_intergenic_412	0	0	3641.44134	1.16769163	0.53548527	7108.06441	-1.055274	0.99724635
DHS+K27me3_intergenic_71	0	0	11489.3817	-0.1775573	0.93102677	15363.268	-2.7618743	0.46101841

DHS+K27me3_one-hit_103	0	0	48601.7115	0.92670262	0.25833936	46690.0427	0.9325274	0.43634827
DHS+K27me3_one-hit_2281	0	0	128075.536	-0.0749699	0.99782889	235980.167	0.00442292	1
DHS+K27me3_one-hit_348	0	0	19433.0745	-0.7505474	0.82436539	42352.2573	-1.0763755	0.67204598
DHS+K27me3_one-hit_402	0	0	110145.629	0.15513259	0.83882888	139588.057	0.34545007	0.77262387
DII_304	0	1	22583.5302	1.62971575	0.02574846	18908.5111	1.9953844	0.0053995
E0_12_GROAviva_ChIP_chip.region_124	0	0	23543.3641	-0.7744512	0.59256434	11454.6123	0.37434413	0.9113662
E0_12_GROAviva_ChIP_chip.region_258	0	0	4060.62605	-2.2005703	0.32728132	5745.8285	-0.155953	0.92027789
E0_12_GROAviva_ChIP_chip.region_26	0	0	3.03168937	0.45912031	0.83882888	6657.24807	-3.0531832	0.94727902
E0_12_GROAviva_ChIP_chip.region_397	0	0	27219.2668	-0.9366836	0.53548527	37351.461	-0.2053201	0.98259655
E0_12_GROAviva_ChIP_chip.region_408	0	0	90.8051081	0.93025518	0.2147916	2094.05414	-1.0588674	0.93774939
E0_12_GROAviva_ChIP_chip.region_439	0	1	6343.46142	2.25998922	0.07599902	7628.4959	3.29851986	8.2998E-06
E0_12_GROAviva_ChIP_chip.region_440	0	0	943.891556	0.59568484	0.85223445	0.4740934	-0.2120508	1
E0_12_GROAviva_ChIP_chip.region_46	0	0	1065.19748	1.24739792	0.51107426	1802.01764	0.48249554	0.8300674
E0_12_GROAviva_ChIP_chip.region_470	0	0	441.669936	2.34866325	0.29031158	22301.8724	-0.1193766	1
E0_12_GROAviva_ChIP_chip.region_473	0	0	16177.4765	-1.0825038	0.57941859	31362.6746	-0.920564	0.85730645
E0_12_GROAviva_ChIP_chip.region_482	0	0	730.745084	-0.6051425	0.93871272	13501.3439	-0.1271852	1
E0_12_GROAviva_ChIP_chip.region_562	0	0	27815.62	-1.0992763	0.51107426	30039.709	-0.5552558	0.9636705
E0_12_GROAviva_ChIP_chip.region_565	0	0	1905.42727	1.90386103	0.12802238	1023.3311	3.01308486	0.17895003
E0_12_GROAviva_ChIP_chip.region_590	0	0	6418.20209	-0.0435306	0.97753196	2995.26562	-0.2921913	0.89219364
E0_12_GROAviva_ChIP_chip.region_594	0	0	3663.98869	1.13311722	0.4927834	7593.75403	1.54915081	0.17895003
E0_12_GROAviva_ChIP_chip.region_606	0	0	428.18951	-3.1479539	0.88472176	nd	nd	nd
E0_12_GROAviva_ChIP_chip.region_620	0	0	17844.1001	-1.4162062	0.25833936	nd	nd	nd
E_0_12h_dCtBP7667.region_1107	0	0	25601.9519	1.04980328	0.26657423	24000.9924	0.70799112	0.64790826
E_0_12h_dCtBP7667.region_1246	0	1	24761.6355	1.23766156	0.22876016	34469.8522	1.6951344	0.01285374
E_0_12h_dCtBP7667.region_2635	0	0	17791.1981	-2.2958588	0.0741017	21290.0625	-0.0131205	1
E_0_12h_dCtBP7667.region_2744	0	0	2395.68819	1.3399362	0.30953547	6991.66335	-0.3464203	0.98047826
E_0_12h_dCtBP7667.region_3346	0	0	35304.3037	-0.7523648	0.63830894	48638.6535	-0.6424179	0.9113662
E_0_12h_dCtBP7667.region_3602	0	0	15660.6528	-0.7996322	0.74923758	11390.558	-1.1147929	0.94727902
E_0_12h_dCtBP7667.region_3785	0	0	20394.7963	0.10658041	0.83882888	20308.6546	-0.2318348	1

E_0_12h_dCtBP7667.region_3799	0	0	54419.889	-0.3597928	0.82436539	90978.5852	-0.4309853	0.9113662
E_0_12h_dCtBP7667.region_3852	0	0	38122.688	1.14058188	0.18642569	37407.0903	0.57082976	0.70645998
E_0_12h_dCtBP7667.region_4427	0	0	7692.81661	0.36672846	0.85825274	5671.63324	-1.2484824	1
E_0_12h_dCtBP7667.region_4575	0	0	403.328476	-2.5205971	0.82458021	nd	nd	nd
E_0_12h_dCtBP7667.region_4755	0	0	5752.1596	-1.1036632	0.59256434	7194.29487	-0.1497562	0.93774939
E_0_12h_dCtBP7667.region_486	0	0	1524.08923	-1.2160888	0.83882888	nd	nd	nd
E_0_12h_dCtBP7667.region_4944	0	0	1246.04928	-0.7807079	0.85825274	3915.34219	-3.4735997	1
E_0_12h_dCtBP7667.region_55	0	0	1546.30253	-1.29975	0.83882888	5209.85608	-0.6275963	1
E_0_12h_dCtBP7667.region_63	0	0	7464.10172	-0.3404933	0.904411	594.004701	3.78659449	0.10267231
E_0_12h_dCtBP7667.region_772	0	0	11615.0431	-1.0950458	0.59256434	11658.7436	-0.6613333	1
E_0_12h_dCtBP7667.region_788	0	0	39232.4671	-0.5303065	0.82436539	24946.2827	-0.0381582	1
E_0_12h_dCtBP7667.region_909	0	0	21374.5309	-0.5423473	0.83882888	4569.83538	-2.4475112	1
Ecoli_control11	1	0	58694.5929	-1.5586874	0.13133978	38401.4891	-2.5057924	0.14348868
Ecoli_control12	0	0	40285.5098	-0.7788515	0.59256434	80638.6357	-0.6486661	0.90501279
Ecoli_control15	1	0	34088.0173	0.05796053	0.86892959	56080.8286	-0.276022	0.92027789
Ecoli_control16	0	0	6893.84129	-0.5772463	0.88362044	7678.75079	-0.4609299	0.99724635
Ecoli_control21	1	0	102193.115	-0.4889988	0.63830894	103903.928	-0.1645455	0.98259655
Ecoli_control22	1	0	3618.86711	-0.2628566	0.93735001	22371.7119	-0.1966921	1
Ecoli_control23	1	0	6120.86239	-1.7190329	0.49760301	16147.6307	-1.0509785	0.89219364
Ecoli_control24	1	0	50768.25	-1.7132693	0.12802238	31389.6628	-3.1238049	0.06000452
Ecoli_control3	1	0	73857.3305	-1.0656348	0.32728132	91333.0875	-1.4679481	0.22987452
Ecoli_control4	0	0	36574.8307	-1.123794	0.3393797	93747.8022	-1.6538628	0.1842483
Ecoli_control5	1	0	59468.6423	0.12589708	0.83882888	82114.5998	-0.2548484	1
Ecoli_control7	0	0	23748.8385	0.6648025	0.65280828	44164.4073	1.11091078	0.1842483
Hand_HCH	0	0	1500.14462	-2.350304	0.65280828	nd	nd	nd
Kr_KrMT	0	0	1921.55458	-0.7867705	0.83019067	924.484907	0.92647889	0.81926714
Ndg_FCenhancer	0	1	51101.3556	0.72916506	0.53548527	45625.2583	1.35653457	0.02481976
Sox15_regionC	0	0	13581.2084	0.02109948	0.86038246	34743.2091	-0.2663826	1
SoxN_565	0	0	14683.1808	0.06457821	0.961218	11106.5979	-1.7691782	0.9113662

SoxN_5830	0	0	4917.97983	-1.5652484	0.48913942	4770.98437	0.19165029	0.91542276
Ubx_BXD-C	0	0	10682.9884	-6.5657724	1.32E-06	25696.5299	-5.1869732	0.01151925
amosD	0	0	2451.87361	-0.7179024	0.83882888	15812.6166	-1.6332659	0.77262387
bab1_dimorphic_element	0	0	20558.0326	-0.9078809	0.59256434	19021.7869	0.09206049	0.98259655
brk_NEE-long	1	1	447.536383	5.418255	3.19E-06	10643.8053	3.37921825	2.9141E-06
btd_R-Ss	0	0	7203.70017	-2.419553	0.18642569	20274.2849	-0.9633086	0.82108452
btd_Ss-Bg	0	0	786.521101	1.98133855	0.25833936	2380.66984	0.8568495	0.82108452
cas_csc-3	0	0	3870.99919	-1.3294431	0.53960702	7658.46892	-2.6478827	0.94727902
cas_csc-5a	0	0	11611.4825	0.2995822	0.83882888	18953.7404	-0.6111587	0.98259655
cas_csc-7a	0	0	84584.4353	-0.1349536	0.93102677	87049.4055	-0.5925327	0.89219364
cpo_cpoCRM6	0	0	4671.42055	0.95291882	0.59256434	39112.451	-0.5647426	0.9113662
crb_Lac-Z	0	1	5271.40758	1.67362769	0.20826761	4562.04442	2.62928302	0.00490148
dac_RE	0	0	1323.04115	-1.2043966	0.82436539	nd	nd	nd
dpp_85.8MX	1	1	39437.9564	1.08907067	0.27336708	54943.3111	2.20799388	2.8702E-10
dpp_BS3.1	0	0	28565.2263	-0.4521698	0.83882888	18635.3897	-2.6888487	0.27730893
dpp_VRR	0	0	1994.78292	2.13000839	0.13093387	2969.71486	2.06715809	0.21297785
dpp_construct10	0	0	7779.7781	-1.9420728	0.31341971	9632.90249	-0.897554	0.98259655
e_coreAbdominalCRE	1	1	2130.12436	0.36951991	0.83882888	3811.45184	2.59455652	0.01029836
gsb_fragIV	1	1	60969.1537	1.69881373	0.00149751	86028.0335	2.00364215	3.6425E-18
gt_CE8001	0	0	1062.64478	0.18897389	0.86038246	2313.03963	0.30368768	0.85356213
h_h7AF	0	0	2453.6326	-2.0503783	0.32728132	1359.55501	-0.4233885	0.9113662
hb_HG4-6	0	0	1562.0818	0.11799599	0.86038246	12229.3131	-0.0565444	1
hb_HG4-7	0	0	37611.0221	-0.8034646	0.53548527	26537.4147	-0.3785016	0.93774939
hb_distal_minimal	0	0	6131.99837	0.85777415	0.3393797	62968.9912	0.81157593	0.43634827
hb_lateDm1.0-lacZ	0	1	7498.32142	1.48286367	0.20826761	4008.53209	2.68267242	0.00471382
hh_hhf4F	0	0	17956.4461	-0.7559391	0.76465641	9168.10873	-0.5685835	1
hkb_0.6kbRIRV	1	1	1416.78801	3.10030476	0.00457588	7360.08373	2.81655884	0.00022694
hth_3	0	0	12521.3939	-1.6317807	0.3393797	18564.1946	-0.8979293	0.89219364
ind_moduleA	1	1	2354.79858	3.78441531	5.25E-05	1449.75668	5.31228843	7.0733E-11

ind_moduleBC	0	1	4399.63633	3.46903508	2.08E-05	2179.56647	3.02448989	0.01440073
ins_Fab-7_minimal_overlap	0	0	4435.22908	0.44535508	0.83882888	6153.47233	0.96402486	0.64790826
ins_SF1_assoc_peak_center	0	0	8541.90892	-1.6624975	0.36022063	2235.83223	0.82343559	0.84325086
kirre_-4.6-3.8	0	0	1809.33676	-0.9775094	0.82436539	6508.83478	0.69521166	0.8050209
kni_223	0	0	9884.40651	-0.0511796	0.93163123	20825.3992	0.06072803	0.98259655
lbl_SBM5	0	0	11163.3141	0.41769828	0.7885808	35964.6786	0.96518807	0.50316876
lz_LMEE	0	0	33094.1821	-1.3551476	0.32728132	18022.4202	-0.4355868	0.98259655
nerfin-1_fragment3	0	1	32821.5217	-0.1913071	0.8709211	52070.3555	1.2100981	0.08783068
nkd_UpE2	0	0	14658.5856	0.45265719	0.83882888	nd	nd	nd
oc_otd-186	1	1	53426.1362	2.73107345	6.05E-08	50367.324	2.6712196	8.3632E-22
ovo_E3	0	0	5046.23636	-1.165776	0.62839416	5620.30797	-0.5549806	0.98259655
pdm2_CRM6	0	0	363.22373	0.95164896	0.60827011	531.921121	-2.3957255	0.98047826
per_-603_-449	0	0	5189.9125	0.32872897	0.85825274	16496.4945	0.78156047	0.60270871
pnr_P3	0	1	21554.1589	3.02528494	4.52E-08	18808.1444	3.6089497	2.6433E-18
pnr_P4	0	1	11389.6734	3.37981321	9.60E-08	6063.38812	3.28464161	4.0626E-05
prd_P1_enhancer	0	0	365.704053	0.9195016	0.821507	nd	nd	nd
repo_-1.1	1	0	14359.8402	0.08180605	0.93102677	7921.95892	1.64336854	0.17462038
rho_NEE_long	0	1	36791.6068	1.45523919	0.02609037	19738.2441	1.57541044	0.07728697
sc_CRM39	0	0	423.784254	-1.633322	0.82436539	nd	nd	nd
slp1_5303	0	0	3010.76342	0.46315847	0.83882888	24044.6186	0.15753437	1
sog_broad_lateral_neurogenic_ectoderm	0	1	13759.8675	1.40337124	0.25833936	21014.3608	2.5598644	1.546E-05
sog_shadow	0	1	30933.7773	1.86622747	0.00205099	38446.6735	2.08998279	4.0626E-05
sphinx_1067bp_5'_fragment	0	0	9323.22989	-0.267641	0.93871272	8154.7798	-1.0065251	1
sqz_sqz-11	0	0	4683.55145	-0.6470979	0.82436539	6061.39434	0.28075656	0.91542276
sv_paxD	0	0	777.669254	-0.5492737	0.88267888	nd	nd	nd
ths_Neu4_early_embryonic_enhancer	1	1	11133.3086	3.01299236	3.25E-05	55423.1108	2.35099842	4.4907E-14
tin_tinD	0	0	369.179038	-1.3807279	0.83019067	nd	nd	nd
tup_dorsalectoderm	0	1	7708.80015	2.89290808	0.00076569	1351.0371	5.50656152	3.1749E-11
vnd_743	0	1	690.163594	3.55512633	0.00205099	501.02327	2.85918525	0.44307581

vnd_NEE	1	1	600.062042	4.46262929	0.00101304	1526.33339	4.09344898	4.0626E-05
y_BE1-2	0	0	4525.97615	1.57569391	0.31341971	7667.65242	0.77841741	0.75741851
y_BE3	0	0	1675.5719	-0.242862	0.9284889	0.8071724	-1.4569829	0.91542276
y_wing	0	0	12900.7008	-1.4255168	0.55166804	25229.2684	0.88480684	0.43634827
E0_12_GROAviva_ChIP_chip.region_278	0	0	nd	nd	nd	9626.98633	-0.9880647	1
E0_12_GROAviva_ChIP_chip.region_416	0	0	nd	nd	nd	1337.64355	-2.6059644	0.98259655
E0_12_GROAviva_ChIP_chip.region_471	0	0	nd	nd	nd	4173.57367	1.04922058	0.56201023
E_0_12h_dCtBP7667.region_1790	0	0	nd	nd	nd	1305.07142	2.63128042	0.14363933
E_0_12h_dCtBP7667.region_3049	1	0	nd	nd	nd	1.97585386	1.29974453	0.77262387
E_0_12h_dCtBP7667.region_3623	0	0	nd	nd	nd	775.516748	-2.9906667	0.98259655
lz_CrystalCellEnhancer1236-737	1	0	nd	nd	nd	1.67213974	0.68180579	0.8300674
wvl_vvl1+2	0	0	nd	nd	nd	58.5323199	4.15672789	0.91542276

## Supplementary table 5

Results of FACS silencer validation experiments for the first library.

Percentage of CD2<sup>+</sup> cells falling in the GFP<sup>reduced</sup> gate is shown for all replicates of all validation experiments, as described in the Methods section.

### Key

random.validation: Percentage of CD2<sup>+</sup> cells falling in the GFP<sup>reduced</sup> gate in embryos from homozygous silencer reporter lines randomly recovered from the tested library.

sFS.positive.validation: Percentage of CD2<sup>+</sup> cells falling in the GFP<sup>reduced</sup> gate in embryos from heterozygous silencer reporter lines generated to test non-TSS-overlapping sFS-positive library elements. (Elements were tested in groups and compared to a negative control tested in parallel.)

random.validation

<b>library element</b>	<b>repl.1</b>	<b>repl.2</b>	<b>repl.3</b>	<b>repl.4</b>	<b>repl.5</b>	<b>repl.6</b>
Ecoli_control15	0.64%	0.51%	0.65%	0.44%	0.21%	0.37%
CAD2_sns	17.54%	16.68%				
ey_UE0.9	10.83%	9.66%				
oc_otd-186	7.58%	6.19%				
gsb_fragIV	2.64%	2.50%				
Est-6_D-511	2.57%	2.35%				
E0_12_GROAviva_ChIP_chip.region_263	0.62%	0.48%				
BiTS-ChIP_K4me1+K27me3+K27ac_96	0.66%	0.37%				
kni_reporter_fragment_EC	0.54%	0.44%				
E0_12_GROAviva_ChIP_chip.region_441	0.54%	0.39%				
ovo_E2	0.48%	0.40%				
e_A.2	0.39%	0.47%				
tin_tin103C	0.34%	0.48%				
E0_12_GROAviva_ChIP_chip.region_43	0.55%	0.39%	0.27%			
vvl_vvlds1.0	0.32%	0.46%				
E0_12_GROAviva_ChIP_chip.region_445	0.32%	0.42%				
svp_sce	0.40%	0.28%				
E_0_12h_dCtBP7667.region_3509	0.34%	0.33%				
BiTS-ChIP_K4me1+K27me3+K27ac_214	0.29%	0.36%				
h_HHRE	0.47%	0.16%				
E_0_12h_dCtBP7667.region_1555	0.26%	0.29%				

## sFS.positive.validation

group	library element	repl.1	repl.2	repl.3	repl.4
1	Ecoli_control15	0.7	0.8		
1	E_0_12h_dCtBP7667.region_1692	0.7	0.7		
1	E_0_12h_dCtBP7667.region_234	1.2	0.6		
1	E_0_12h_dCtBP7667.region_4530	0.7	0.6		
1	E_0_12h_dCtBP7667.region_4196	1.6	0.7		
1	dpp_85.8MX	8.7	4.2		
1	gcm_+3.8_+4.5	1.9	1		
2	Ecoli_control15	0.3	0.5	0.6	
2	h_HHRE	0.4	0.4		
2	DHS+K27me3_CtBPoverlap_2267	0.5	0.9		
2	repo_-1.1	0.6	0.7	1.3	1.1
2	ct_ct-3	0.3	0.4		
2	e_coreAbdominalCRE	1	1.7		
2	hh_4075	0.3	0.3		
3	Ecoli_control15	1.2	1	0.8	0.7
3	lab_HZ550	1.1	1	0.9	1.1
3	oc_SBg	11.5	10.8	9.3	9
3	Sox15_regionB	0.5	0.5	0.7	0.8
3	ths_Neu4_early_embryonic_enhancer	8.4	8.2	5.8	5.5
3	vas_construct16	0.7	0.6	0.4	0.7
3	E_0_12h_dCtBP7667.region_3084	3.7	3.6	1.9	2.1
4	Ecoli_control15	1	0.9	1.6	1.4
4	ovo_del-ap-del-5	0.5	0.6	1.4	1.3
4	CAD2_Meso-CRM-2819	1	1.3	3.2	2.9
4	lz_CrystalCellEnhancer1236-737	1.5	1.9	4.8	5.1
4	E0_12_GROAviva_ChIP_chip.region_481	1	1.1	1.5	2.1
4	DHS+K27me3_intergenic_1490a	0.8	0.8	1.3	1.4
4	ind_moduleA	13.1	13.8	18.2	18.3

5	Ecoli_control15	1.9	1.6	1.4	1.1
5	hkb_0.6kbRIRV	11.1	10.8	12	10.2
5	vvl_vvlds1.0	1.2	1.4	0.8	1
5	tin_tin103C	1.2	1.3	1.1	0.9
5	brk_NEE-long	9.7	9.8	10.1	9.9
5	E0_12_GROAviva_ChIP_chip.region_179	0.7	0.7	0.7	0.6
5	Ecoli_control22	1.6	1.8	1.4	1.3
6	Ecoli_control15	1	1	1	1
6	E_0_12h_dCtBP7667.region_3049	1.6	1.5	1.5	1.8
6	DHS+K27me3_intergenic_1490b	0.6	0.8	0.8	0.9
6	ovo_E2	0.8	0.8	0.8	0.9
6	run_neural_6GB	1	1.3	0.6	0.7
6	vnd_NEE	14.6	15	15.7	16.1
6	E_0_12h_dCtBP7667.region_2659	1.3	1.6	1.5	1.4

## Supplementary table 6

Results of FACS silencer validation experiments for the second library.

<b>element</b>	<b>call</b>	<b>sFS call</b>	<b>repl.1</b>	<b>repl.2</b>	<b>repl.3</b>	<b>repl.4</b>
Ecoli_control4			1.80%	1.20%	1.90%	1.50%
DHS+K27me3_one-hit_2281	(-)	-	2.10%	2%	2.50%	
crb_Lac-Z	+	+	6.20%	5%	5.80%	
CRM_1807	+	+	5.20%	6.60%	5.30%	
hb_lateDm1.0-lacZ	+	+	7.40%	5.80%	4.70%	
slp1_5303	(-)	-	2.30%	2.20%	1.80%	
CAD2_htl	-	-	1.40%	1.60%	1.60%	
rho_NEE_long	+	+	6.60%	5.70%	5.50%	
E_0_12h_dCtBP7667.region_4755	-	-	1.60%	1.90%	1.40%	

## Supplementary table 7

All validated silencers and their coordinates, along with the category they are associated with.

id	coordinates (dm3)	source
dpp_85.8MX	chr2L:2456545-2457501	Nonmeso CRMs (REDfly)
tup_dorsalectoderm	chr2L:18874963-18875896	Nonmeso CRMs (REDfly)
Ndg_FCenhancer	chr2R:6203047-6203992	specific meso CRMs
ths_Neu4_early_embryonic_enhancer	chr2R:7681709-7682675	Nonmeso CRMs (REDfly)
Dll_304	chr2R:20690263-20691248	Nonmeso CRMs (REDfly)
gsb_fragIV	chr2R:20944064-20945045	Nonmeso CRMs (REDfly)
E_0_12h_dCtBP7667.region_1246	chr2R:6507041-6508091	CtBP ChIP
ind_moduleA	chr3L:15031943-15032964	positive controls
ind_moduleBC	chr3L:15032738-15033835	positive controls
nerfin-1_fragment3	chr3L:904458-905370	Nonmeso CRMs (REDfly)
rho_NEE_long	chr3L:1461675-1462661	additional NEEs
CAD2_Meso-CRM-2819	chr3R:17222290-17223251	specific meso CRMs
hkb_0.6kbRIRV	chr3R:173849-174821	Nonmeso CRMs (REDfly)
hb_lateDm1.0- <i>lacZ</i>	chr3R:4526286-4527384	Nonmeso CRMs (REDfly)
pnr_P3	chr3R:11853793-11854843	Nonmeso CRMs (REDfly)
pnr_P4	chr3R:11854355-11855347	Nonmeso CRMs (REDfly)
repo_-1.1	chr3R:14060749-14061844	Nonmeso CRMs (REDfly)
e_coreAbdominalCRE	chr3R:17066418-17067406	Nonmeso CRMs (REDfly)
crb_Lac-Z	chr3R:20122950-20123916	Nonmeso CRMs (REDfly)
E0_12_GROAviva_ChIP_chip.region_439	chr3R:9735607-9736591	Groucho ChIP
E_0_12h_dCtBP7667.region_3049	chr3R:907088-908163	CtBP ChIP
E_0_12h_dCtBP7667.region_3084	chr3R:1508921-1509997	CtBP ChIP
vnd_NEE	chrX:486301-487394	Nonmeso CRMs (REDfly)
vnd_743	chrX:486746-487752	Nonmeso CRMs (REDfly)
oc_SBg	chrX:8547554-8548562	Nonmeso CRMs (REDfly)

oc_otd-186	chrX:8548281-8549211	Nonmeso CRMs (REDfly)
lz_CrystalCellEnhancer1236-737	chrX:9177203-9178176	Nonmeso CRMs (REDfly)
sog_broad_lateral_neurogenic_ectoderm	chrX:15518390-15519344	Nonmeso CRMs (REDfly)
sog_shadow	chrX:15540621-15541615	Nonmeso CRMs (REDfly)
brk_NEE-long	chrX:7190855-7191822	additional NEEs

## Supplementary table 8

PWMs for repressive TFs used in motif enrichment analysis.

MOTIF 0:

>ab_M4542_1.01		sd	1.5								
0.286026201	0.288503254	0.242950108	0.195227766	0.321041215	0.221258134	0.412147505	0.002169197	0	0.973969631	0	
0	0.648590022	0.281995662	0.054229935	0.117136659	0.373101952	0.305856833	0.225596529	0.312364425	0.340563991		
0.362255965	0.46637744	0.42516269	0.436813187	0.474285714							
0.344978166	0.407809111	0.364425163	0.37527115	0.260303688	0.422993492	0.041214751	0.982646421	0.986984816	0.002169197		
0	0	0.062906725	0.707158351	0.811279826	0.427331887	0.277657267	0.249457701	0.281995662	0.30802603	0.25813449	
0.234273319	0.277657267	0.268980477	0.211538462	0.194285714							
0.176855895	0.199566161	0.190889371	0.229934924	0.114967462	0.262472885	0.449023861	0	0	0	0.989154013	
0.997830803	0.017353579	0.004338395	0.034707158	0.086767896	0.151843818	0.078091106	0.151843818	0.130151844	0.221258134		
0.247288503	0.162689805	0.180043384	0.181318681	0.142857143							
0.192139738	0.104121475	0.201735358	0.199566161	0.303687636	0.093275488	0.097613883	0.015184382	0.013015184	0.023861171		
0.010845987	0.002169197	0.271149675	0.006507592	0.09978308	0.368763557	0.197396963	0.36659436	0.340563991	0.249457701		
0.180043384	0.156182213	0.093275488	0.125813449	0.17032967	0.188571429						
mu=15.466654	sigma=3.670248	threshold=9.961282									

MOTIF 1:

>cwo_M4719_1.01		sd	1.5								
0.5	0.1	0	0.9	0	0	0	0	1	0	0	0.7
0	0	1	0	1	0	0	0.1	0	1	0.8	0.1
0.5	0	0	0	0	0.9	0	0.9	0	0	0.1	0.2
0	0.9	0	0.1	0	0.1	1	0	0	0	0.1	0
mu=19.413216	sigma=2.527889	threshold=15.621382									

MOTIF 2:

>al_M4552_1.01		sd	1.5								
0.189655172	0.208443272	0.080645161	0.004608295	0.976958525	0.97235023	0	0.011520737	0.627272727			
0.351724138	0.263852243	0.5	0	0.023041475	0.02764977	0	0.006912442	0.015151515			
0.148275862	0.321899736	0.020737327	0	0	0	0.023041475	0.055299539	0.327272727			
0.310344828	0.205804749	0.398617512	0.995391705	0	0	0.976958525	0.926267281	0.03030303			
mu=9.665823	sigma=2.418970	threshold=6.037368									

MOTIF 3:

>aop_M4558_1.01		sd	1.5								
0.526315789	0	0	0	0	1	0.947368421	0.263157895	0.052631579			

	0.157894737	0.894736842	1	0	0	0	0	0	0	0.315789474			
	0.315789474	0.105263158	0	1	1	0	0	0	0.736842105	0			
	0	0	0	0	0	0.052631579	0	0	0.631578947				
MOTIF 4:	mu=14.280921	sigma=1.851768	threshold=11.503268										
	>B-H1_M1847_1.01	sd	1.5										
	0.19047619	0	1	1	0.380952381	0.047619048	0.047619048						
	0.19047619	0	0	0	0	0.333333333	0						
	0	0	0	0	0	0.952380952							
	0.619047619	1	0	0	0.619047619	0.619047619	0						
MOTIF 5:	mu=9.569882	sigma=1.538576	threshold=7.262018										
	>B-H2_M1848_1.01	sd	1.5										
	0.285714286	0	1	1	0.238095238	0.142857143	0						
	0.095238095	0	0	0	0	0.047619048	0						
	0.047619048	0	0	0	0	0.095238095	1						
	0.571428571	1	0	0	0.761904762	0.714285714	0						
MOTIF 6:	mu=9.655945	sigma=1.681040	threshold=7.134385										
	>bin_M4585_1.01	sd	1.5										
	0.45	0	1	1	0	0.95							
	0	0	0	0	0	1	0						
	0.5	0	0	0	0	0	0						
	0.05	1	0	0	0	0	0.05						
MOTIF 7:	mu=11.797096	sigma=1.260249	threshold=9.906723										
	>Blimp-1_M4586_1.01	sd	1.5										
	0.75	0	0	0	0	0.75	0	0	0	0.083333333	0		
	0	0.916666667	0	0.083333333	0	1	0.083333333	0.916666667	0	0.083333333	0	0.666666667	
	0.25	0	0	0	0	0	0.166666667	0	0	0.083333333	0	0	
	0	0.083333333	1	0.916666667	1	0	0	0.083333333	1	0.833333333	0.916666667	0.333333333	
MOTIF 8:	mu=18.230035	sigma=2.613552	threshold=14.309707										

	>bowl_M4588_1.01	sd	1.5																	
	0.363636364	0	0.583333333	0	0.416666667	1	0	0												
	0.636363636	1	0.083333333	0	0	0	0	0.727272727												
	0	0	0.333333333	1	0.083333333	0	1	0												
	0	0	0	0	0.5	0	0	0.272727273												
MOTIF 9:	mu=12.025484	sigma=1.360199	threshold=9.985186																	
	>brk_M1892_1.01	sd	1.5																	
	0.1	0	0	0	0.1	0.1	0													
	0.5	0.4	0	0	1	0	0.8	0.5												
	0.4	0	1	1	0	0.9	0	0.1												
	0	0.6	0	0	0	0	0.1	0.4												
MOTIF 10:	mu=12.000991	sigma=2.038048	threshold=8.943918																	
	>btd_M2117_1.01	sd	1.5																	
	0.133333333	0.333333333	0	0	0	0	0	0.266666667	0	0.2										
	0.333333333	0.666666667	1	0	1	1	0.933333333	0.6	0.666666667	0.233333333										
	0.033333333	0	0	0.933333333	0	0	0	0	0	0.1										
	0.5	0	0	0.066666667	0	0	0.066666667	0.133333333	0.333333333	0.466666667										
MOTIF 11:	mu=13.810290	sigma=2.279288	threshold=10.391358																	
	>C15_M4608_1.01	sd	1.5																	
	0.177328844	0.14365881	0	0.967452301	1	0.23681257	0.109988777	0.634118967												
	0.231200898	0.134680135	0	0	0	0.089786756	0.071829405	0												
	0.242424242	0.086419753	0	0	0	0.06285073	0.346801347	0.365881033												
	0.349046016	0.635241302	1	0.032547699	0	0.610549944	0.471380471	0												
MOTIF 12:	mu=7.414993	sigma=1.921533	threshold=4.532694																	
	>ci_M4711_1.01	sd	1.5																	
	0.545454545	0	0.727272727	0.181818182	0.090909091	0.727272727	0	0.272727273	0	0.909090909	0									
	0	0	0.272727273	0.818181818	0.909090909	0.181818182	0.909090909	0.727272727	1	0	0.818181818									
	0.363636364	1	0	0	0	0	0	0	0.090909091	0.181818182										

	0.090909091	0	0	0	0	0.090909091	0.090909091	0	0	0	0			
MOTIF 13:	mu=16.009337	sigma=2.669297	threshold=12.005391											
	>cic_M4710_1.01	sd	1.5											
	0.058823529	0	0	1	0	0	0	1						
	0.588235294	0.611111111		1	0	0	0	0.555555556		0				
	0	0	0	0	0	0	0	0.444444444		0				
	0.352941176	0.388888889		0	0	1	1	0	0					
MOTIF 14:	mu=12.610787	sigma=1.123827	threshold=10.925046											
	>croc_M2416_1.01	sd	1.5											
	0.444444444	0.636363636		0.454545455	0.666666667	0.666666667	0	1	1	1	0	1	0.066666667	0.4
	0.333333333	0.142857143		0.25										
	0.222222222	0.090909091		0.181818182	0.083333333	0	0	0	0	0	0.2	0	0.066666667	0.2
	0.142857143	0.166666667		0.181818182	0.166666667	0.333333333	0	0	0	0	0	0.266666667	0.133333333	0.2
	0	0		0.181818182	0.166666667	0.333333333	0	0	0	0	0	0.266666667	0.133333333	0.2
	0.357142857	0.5												
	0.333333333	0.272727273		0.181818182	0.083333333	0	1	0	0	0	0.8	0	0.6	0.266666667
	0.357142857	0.083333333												
MOTIF 15:	mu=13.942857	sigma=2.301460	threshold=10.490667											
	>ct_M1897_1.01	sd	1.5											
	0	0.05	0.45	1	0.85	0								
	0.3	0.1	0	0	0	0.95								
	0	0	0.55	0	0.05	0.05								
	0.7	0.85	0	0	0.1	0								
	mu=8.033671	sigma=1.978360	threshold=5.066130											
MOTIF 16:	>CTCF_M2196_1.01	sd	1.5											
	0.396424816	0.158254469		0.049421661	0.068349106	0.130389064	0	0.833333333	0.001577287	0	0.092008412	0.443743428		
	0.048370137	0.124079916		0.117245005	0.166666667									
	0.075709779	0.004206099		0.950052576	0.001577287	0.868559411	1	0.163512093	0.479495268	0.893796004	0.100946372			
	0.045741325	0.184016824		0.434279706	0.115141956	0.211882229								
	0.406414301	0.796004206		0	0.864879075	0	0	0.002103049	0	0	0.000525762	0.375394322	0.412197687	
	0.201366982	0.603049422		0.460567823										
	0.121451104	0.041535226		0.000525762	0.065194532	0.001051525	0	0.001051525	0.518927445	0.106203996	0.806519453			
	0.135120925	0.355415352		0.240273396	0.164563617	0.160883281								
MOTIF 17:	mu=15.071818	sigma=3.781289	threshold=9.399885											
	>Aef1_M4550_1.01	sd	1.5											
	0.5	0	1	1	0	1	1	0	1	0.75	0.166666667			
	0.166666667	1	0	0	1	0	0	1	0	0.25	0.75			

	0.33333333	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0.08333333	
MOTIF 18:	mu=18.433967	sigma=1.492901	threshold=16.194615										
	>Clk_cyc_M4713_1.01	sd	1.5										
	0.5	0	0.9	0	0	0	0	0.55					
	0.5	1	0	1	0	0	0	0.15					
	0	0	0	0	1	0	1	0.25					
	0	0	0.1	0	0	1	0	0.05					
MOTIF 19:	mu=13.365529	sigma=1.335658	threshold=11.362042										
	>D_M4739_1.01	sd	1.5										
	0.029411765	0	0.117647059	0.764705882	0	0	0	0	0.029411765	0.235294118	0.323529412		
	0.235294118	0.823529412	0.529411765	0	0.117647059	0	0	0	0.235294118	0.323529412	0.029411765		
	0.264705882	0.029411765	0	0	0.088235294	1	0	0.088235294	0.176470588	0.058823529			
	0.470588235	0.147058824	0.352941176	0.235294118	0.882352941	0.911764706	0	1	0.647058824	0.264705882	0.588235294		
MOTIF 20:	mu=11.302874	sigma=2.600774	threshold=7.401713										
	>dar1_M4631_1.01	sd	1.5										
	0.314141414	0.261700096	0.334572491	0.270446097	0.125464684	0.011152416	0.847583643	0.008364312	0.689591078	0.003717472			
	0.030669145	0.031598513	0.552973978	0.204123711	0.270883055								
	0.337373737	0.316141356	0.161710037	0	0.846654275	0.977695167	0.101301115	0.983271375	0	0.986988848	0.958178439		
	0.910780669	0.18866171	0.388659794	0.354415274									
	0.171717172	0.265520535	0.304832714	0.712825279	0.019516729	0	0.040892193	0	0.220260223	0	0.002788104		
	0.001858736	0.017657993	0.151546392	0.136038186									
	0.176767677	0.156638013	0.198884758	0.016728625	0.008364312	0.011152416	0.010223048	0.008364312	0.090148699	0.00929368			
	0.008364312	0.055762082	0.24070632	0.255670103	0.238663484								
MOTIF 21:	mu=14.862728	sigma=3.639479	threshold=9.403509										
	>Dfd_M4728_1.01	sd	1.5										
	0.217610063	0.113207547	0.028930818	0.895597484	0.902515723	0	0.037735849	0.606918239					
	0.316981132	0.335220126	0.003144654	0.091823899	0.095597484	0.054716981	0.034591195	0.055345912					
	0.241509434	0.132704403	0.062264151	0.008176101	0.001886792	0.094968553	0.57672956	0.303773585					
	0.223899371	0.418867925	0.905660377	0.004402516	0	0.850314465	0.350943396	0.033962264					
MOTIF 22:	mu=6.693361	sigma=2.769185	threshold=2.539583										



	0.206955923	0.999269006	0.000243665	0.000243665	0.999269006	0.000243665	0.028824201	0.087239583				
MOTIF 27:	mu=11.736539	sigma=1.485983	threshold=9.507564									
	>E-spl-m3_M4838_1.01	sd	1.5									
	0.090909091	0.181818182	0.090909091	0	0.818181818	0	0	0	0			
	0.045454545	0	0.136363636	1	0	0.954545455	0	0.045454545	0			
	0.181818182	0.818181818	0.636363636	0	0.136363636	0	1	0	1			
	0.681818182	0	0.136363636	0	0.045454545	0.045454545	0	0.954545455	0			
MOTIF 28:	mu=13.756692	sigma=2.616989	threshold=9.831209									
	>E-spl-m5_M4840_1.01	sd	1.5									
	0	0.052631579	0	0	1	0	0	0.052631579	0	0	0	
	0	0	0	1	0	1	0	0.105263158	0	0.684210526	0.5	
	0	0.947368421	1	0	0	0	1	0	1	0.105263158	0.166666667	
	1	0	0	0	0	0	0	0.842105263	0	0.210526316	0.333333333	
MOTIF 29:	mu=19.172168	sigma=2.014500	threshold=16.150418									
	>E-spl-m7_M4841_1.01	sd	1.5									
	0	0	0	0.052631579	0	0.894736842	0	0	0	0	0.055555556	
	0.105263158	0.052631579	0.052631579	0.052631579	0.052631579	1	0	1	0	0.052631579	0	0.888888889
	0.263157895	0.052631579	0.947368421	0.894736842	0	0	0	0	1	0	1	0.055555556
	0.631578947	0.894736842	0	0	0	0.105263158	0	0	0.947368421	0	0	
MOTIF 30:	mu=18.715605	sigma=2.805909	threshold=14.506741									
	>E-spl-m8_M4768_1.01	sd	1.5									
	0	0.066666667	0.133333333	0	0.8	0	0	0	0	0		
	0.266666667	0	0.066666667	1	0	1	0	0.066666667	0	0.785714286		
	0	0.933333333	0.733333333	0	0.133333333	0	1	0	1	0.071428571		
	0.733333333	0	0.066666667	0	0.066666667	0	0	0.933333333	0	0.142857143		
MOTIF 31:	mu=16.268979	sigma=2.660259	threshold=12.278591									
	>E-spl-mbeta_M4842_1.01	sd	1.5									
	0.133333333	0	0	1	0	0	0	0.071428571				

0.066666667	0.125	1	0	1	0	0.5625	0			
0.8	0.8125	0	0	0	1	0	0.928571429			
0	0.0625	0	0	0	0	0.4375	0			
mu=13.833749		sigma=2.139235		threshold=10.624897						
MOTIF 32:										
>E-spl-mdelta_M4843_1.01 sd 1.5										
0.173913043	0.086956522	0.217391304	0	0.782608696	0	0	0	0	0	0
0.043478261	0	0.043478261	1	0	0.956521739	0	0.260869565	0		
0.086956522	0.913043478	0.565217391	0	0.130434783	0	1	0	1		
0.695652174	0	0.173913043	0	0.086956522	0.043478261	0	0.739130435	0		

mu=13.313383		sigma=2.481484		threshold=9.591157						
MOTIF 33:										
>E-spl-mgamma_M4844_1.01 sd 1.5										
0.130434783	0.083333333	0.083333333	0	0.75	0	0	0	0	0	0
0.173913043	0	0	1	0	1	0	0.083333333	0	0.666666667	
0	0.916666667	0.833333333	0	0.208333333	0	1	0	1	0.285714286	
0.695652174	0	0.083333333	0	0.041666667	0	0	0.916666667	0	0.047619048	

mu=16.023282		sigma=2.576089		threshold=12.159148						
MOTIF 34:										
>E-spl-mgamma_M4845_1.01 sd 1.5										
0.428571429	0	0	0	1	1	0				
0	0	0	1	0	0	0				
0	1	0	0	0	0	1				
0.571428571	0	1	0	0	0	0	0			

mu=12.903830		sigma=0.395347		threshold=12.310809						
MOTIF 35:										
>EcR_M4752_1.01 sd 1.5										
0.642857143	0	0	1	0	0	0	0			
0	0	0	0	0.928571429	1	0.142857143	0.5			
0	0	1	0	0.071428571	0	0	0			
0.357142857	1	0	0	0	0	0.857142857	0.5			

mu=13.052781		sigma=1.347945		threshold=11.030864						
MOTIF 36:										

>en_M4764_1.01	sd	1.5									
0.197452229	0.238319592	0.18430792	0.198371577	0.772020725	1	0	0.08919319	0.612045366	0.280343008		
0.27388535	0.263839812	0.264248705	0.048482605	0	0	0	0.016425499	0.252638522			
0.221771859	0.242245779	0.171354552	0	0.227979275	0	0	0.174315322	0.359796637	0.298812665		
0.306890562	0.255594817	0.380088823	0.753145818	0	0	1	0.736491488	0.011732499	0.168205805		
mu=7.449052			sigma=1.911944		threshold=4.581135						
MOTIF 37:											
>esg_M4767_1.01	sd	1.5									
0	1	0	0	0	0	0	0.181818182				
1	0	1	0.909090909	0	0	0	0.181818182				
0	0	0	0	0	1	0					
0	0	0	0.090909091	1	0	0	0.636363636				
mu=12.556829			sigma=1.362288		threshold=10.513397						
MOTIF 38:											
>eve_M1001_1.01	sd	1.5									
0.1950859	0.106149549	0.745347352	0.803280667	0.084324907	0.130553413	0.40331548	0.228660074	0.186386104			
0.348430915	0.07863752	0.066943592	0.07737813	0.168119877	0.189006695	0.108822606	0.290683925	0.283672485			
0.193144763	0.022985163	0.120765464	0.027624531	0.07257912	0.314751186	0.304819283	0.273755811	0.249091513			
0.263338421	0.792227768	0.066943592	0.091716672	0.674976096	0.365688706	0.183042632	0.206900191	0.280849897			
mu=3.329769			sigma=2.706932		threshold=-0.730629						
MOTIF 39:											
>fd64A_M4783_1.01	sd	1.5									
0.15	0.35	0.2	1	0.95	1	0	1				
0	0.25	0.15	0	0.05	0	0.85	0				
0.2	0.3	0	0	0	0	0	0				
0.65	0.1	0.65	0	0	0	0.15	0				
mu=9.918080			sigma=1.874029		threshold=7.107037						
MOTIF 40:											
>Fer3_M4790_1.01	sd	1.5									
0.052631579	0.105263158	0.421052632	0.736842105	0	0.947368421	0	0.263157895	0	0		
0	0.105263158	0.210526316	0	1	0	0.052631579	0.421052632	0.052631579	0		
0.789473684	0.210526316	0.368421053	0.105263158	0	0.052631579	0.789473684	0.263157895	0	0.894736842		

	0.157894737	0.578947368	0	0.157894737	0	0	0.157894737	0.052631579	0.947368421	0.105263158		
MOTIF 41:	mu=11.340378	sigma=2.909501	threshold=6.976127									
	>fkh_M4791_1.01	sd	1.5									
	0.148148148	0.111111111	0.518518519	0.407407407	0.296296296	0	0.962962963	1	1	0	0.888888889	
	0.148148148	0.037037037	0.074074074	0.111111111	0.074074074	0.518518519	0.037037037	0	0	0.814814815	0	
	0.148148148	0	0.407407407	0.259259259	0.481481481	0	0	0	0	0	0	
	0.555555556	0.851851852	0	0.222222222	0.148148148	0.481481481	0	0	0	0.185185185	0.111111111	
MOTIF 42:	mu=11.499229	sigma=2.492014	threshold=7.761209									
	>foxo_M4792_1.01	sd	1.5									
	0	0.15	0.4	0.4	0.8	0.8	1	0	1			
	0.3	0.6	0	0	0.05	0.2	0	0.9	0			
	0.15	0.15	0.55	0	0	0	0	0	0			
	0.55	0.1	0.05	0.6	0.15	0	0	0.1	0			
MOTIF 43:	mu=10.229988	sigma=2.319180	threshold=6.751218									
	>ftz_M4800_1.01	sd	1.5									
	0.097982709	0.043227666	1	1	0	0	0.778097983					
	0.198847262	0.028818444	0	0	0	0.014409222	0					
	0.161383285	0.074927954	0	0	0.129682997	0.397694524	0.221902017					
	0.541786744	0.853025937	0	0	0.870317003	0.587896254	0					
MOTIF 44:	mu=8.314127	sigma=1.968776	threshold=5.360963									
	>Gsc_M1006_1.01	sd	1.5									
	0.207447972	0.070500152	0.823900567	0.800901891	0.078336912	0.084960697	0.101355776	0.148879872				
	0.325343183	0.046787444	0.041431848	0.045410278	0.041721597	0.693803489	0.577458665	0.348326325				
	0.141865663	0.081909491	0.064978072	0.085262818	0.098594838	0.089316725	0.132772464	0.299807246				
	0.325343183	0.800802914	0.069689513	0.068425014	0.781346653	0.131919088	0.188413095	0.202986557				
MOTIF 45:	mu=4.826003	sigma=3.299439	threshold=-0.123156									
	>h_M2123_1.01	sd	1.5									
	0.029411765	0.088235294	0.029411765	0.647058824	0.029411765	0.058823529	0.029411765	0.029411765	0.205882353	0.029411765		

0.058823529	0.205882353	0.911764706	0.029411765	0.882352941	0.029411765	0.294117647	0.029411765	0.5	0.882352941		
0.882352941	0.5	0.029411765	0.294117647	0.029411765	0.882352941	0.029411765	0.911764706	0.205882353	0.058823529		
0.029411765	0.205882353	0.029411765	0.029411765	0.058823529	0.029411765	0.647058824	0.029411765	0.088235294	0.029411765		
MOTIF 46:	mu=11.249869	sigma=4.162214	threshold=5.006547								
>hb_M4821_1.01	sd	1.5									
0	1	0.942857143	1	1	1	1	0.285714286	0.4	0.382352941	0.383333333	
0.328571429	0	0	0	0	0	0	0.471428571	0.128571429	0.279411765	0.35	
0.057142857	0	0	0	0	0	0	0.114285714	0.328571429	0.235294118	0.216666667	
0.614285714	0	0.057142857	0	0	0	0	0	0.128571429	0.142857143	0.102941176	0.05
MOTIF 47:	mu=12.247652	sigma=1.854851	threshold=9.465376								
>HGTX_M4827_1.01	sd	1.5									
0.325301205	0.285714286	0.00172117	0.839931153	0.967297762	0.012048193	0.103270224	0.767641997				
0.134251291	0.056798623	0.018932874	0.024096386	0.030981067	0.010327022	0.015490534	0.012048193				
0.17383821	0.08777969	0.030981067	0.051635112	0.00172117	0.089500861	0.333907057	0.199655766				
0.366609294	0.569707401	0.948364888	0.084337349	0	0.888123924	0.547332186	0.020654045				
MOTIF 48:	mu=7.224892	sigma=2.868667	threshold=2.921892								
>hkb_M4828_1.01	sd	1.5									
0.09375	0.03125	0.90625	0	0	0	0	0.34375				
0.0625	0.96875	0.09375	1	0	1	1	0.75	0.5625			
0.09375	0	0	0	1	0	0	0	0			
0.75	0	0	0	0	0	0	0.25	0.09375			
MOTIF 49:	mu=14.900081	sigma=2.194723	threshold=11.607997								
>Hmx_M4847_1.01	sd	1.5									
0.087671233	0	0.98630137	0.961643836	0	0	0.216438356					
0.128767123	0	0.01369863	0.038356164	0	0.106849315	0					
0.134246575	0.068493151	0	0	0.189041096	0.167123288	0.734246575					
0.649315068	0.931506849	0	0	0.810958904	0.726027397	0.049315068					
MOTIF 50:	mu=8.339892	sigma=2.397002	threshold=4.744389								

>Hr46_M1361_1.01	sd	1.5												
0.22138278	0.00005916		0.00005375	0.99446356	0.00542894	0.00005375	0.0000769	0.20214802						
0.22138278	0.10062707		0.00542894	0.00005375	0.99446356	0.99983874	0.4076438	0.36165462						
0.2900641	0.00005916	0.99446356	0.00542894	0.00005375	0.00005375	0.00005375	0.0000769	0.24468311						
0.26717033	0.89925461		0.00005375	0.00005375	0.00005375	0.00005375	0.00005375	0.5922024	0.19151425					
mu=10.789819			sigma=1.482775			threshold=8.565656								
MOTIF 51:														
>Hsf_M4858_1.01	sd	1.5												
0	0.4	0	1	1	0.4	0.2	0.2	0.2	0	0	0.6	0	1	
0	0.6	0	0	0	0	0.4	0	0	1	0	0.4	0	0	
0.4	0	1	0	0	0.2	0.2	0	0	0	0.4	0	1	0	
0.6	0	0	0	0	0.4	0.2	0.8	0.8	0	0.6	0	0	0	
mu=18.858010			sigma=1.469780			threshold=16.653340								
MOTIF 52:														
>ind_M4862_1.01	sd	1.5												
0.047619048	0.428571429		0	0	1	1	0	0	0.904761905					
0.428571429	0.19047619		0.571428571	0	0	0	0	0	0					
0.095238095	0.285714286		0	0	0	0	0.047619048	0.380952381	0.095238095					
0.428571429	0.095238095		0.428571429	1	0	0	0.952380952	0.619047619	0					
mu=11.116169			sigma=1.705847			threshold=8.557399								
MOTIF 53:														
>kni_M4876_1.01	sd	1.5												
0.076923077	0	0	0.153846154	0.038461538	0	0	0.807692308	0.230769231	0.384615385	0	0.153846154			
0.269230769	0.038461538		0	0.692307692	0.307692308	1	0.192307692	0.038461538	0.230769231	0	0	0.076923077		
0.461538462	0	1	0.115384615	0	0	0	0.153846154	0.346153846	0	0.038461538	0.038461538			
0.192307692	0.961538462		0	0.038461538	0.653846154	0	0.807692308	0	0.192307692	0.615384615	0.961538462			
0.730769231														
mu=13.409592			sigma=2.825261			threshold=9.171700								
MOTIF 54:														
>Kr_M2126_1.01	sd	1.5												
0.138117284	0.158179012		0.116512346	0.962962963	0.833333333	0.008487654	0	0	0	0	0	0.087191358		
0.24537037	0.162808642													
0.325617284	0.145833333		0.297839506	0	0.166666667	0.902777778	0.908179012	1	0.155864198	0	0.087191358			
0.30787037	0.24845679		0.334876543											
0.183641975	0.168209877		0.05632716	0	0	0	0	0	0	0	0	0.178240741	0.219907407	0.168981481

	0.352623457	0.527777778	0.529320988	0.037037037	0	0.088734568	0.091820988	0	0.844135802	1	0.912808642
	0.426697531	0.286265432	0.333333333								
	mu=13.902670	sigma=2.773848	threshold=9.741897								
MOTIF 55:	>lbe_M4891_1.01	sd	1.5								
	0.119205298	0.046357616	0.966887417	1	0	0.109271523	0.821192053				
	0.304635762	0	0.033112583	0	0.347682119	0.231788079	0.039735099				
	0.102649007	0.043046358	0	0	0.082781457	0.268211921	0.139072848				
	0.473509934	0.910596026	0	0	0.569536424	0.390728477	0				
	mu=6.906401	sigma=2.116203	threshold=3.732097								
MOTIF 56:	>Lim1_M4895_1.01	sd	1.5								
	0.147940075	0.044007491	0.957865169	1	0.004681648	0.029026217	0.731273408				
	0.258426966	0.051498127	0	0	0.007490637	0.012172285	0.001872659				
	0.092696629	0	0.028089888	0	0	0.078651685	0.24906367				
	0.500936633	0.904494382	0.014044944	0	0.987827715	0.880149813	0.017790262				
	mu=8.743815	sigma=2.490507	threshold=5.008054								
MOTIF 57:	>dm_Max_M4738_1.01	sd	1.5								
	0.375	0.375	0	1	0	0.041666667	0	0	0.041666667	0.208333333	
	0.041666667	0.333333333	1	0	1	0	0	0	0.125	0.041666667	
	0.5	0.166666667	0	0	0	0.958333333	0	1	0.833333333	0.208333333	
	0.083333333	0.125	0	0	0	0	1	0	0.541666667		
	mu=14.544813	sigma=2.164634	threshold=11.297862								
MOTIF 58:	>mirr_M4929_1.01	sd	1.5								
	0	0	0	0.303180212	0.445229682	0.268551237					
	0	0	0	0.144169611	0.199293286	0.231095406					
	0	1	0	0	0.065724382	0.111660777					
	1	0	1	0.552650177	0.28975265	0.38869258					
	mu=6.557403	sigma=0.969995	threshold=5.102410								
MOTIF 59:	>Mnt_M4920_1.01	sd	1.5								
	0	1	0	0	0	0					
	1	0	1	0	0	0					
	0	0	0	1	0	1	0.888888889				

	0	0	0	0	1	0	0.111111111													
	mu=14.105703		sigma=1.117217		threshold=12.429877															
MOTIF 60:	>net_da_M4935_1.01		sd 1.5																	
	0.454545455	0.272727273	0	1	0	0	0	0	0	0.045454545										
	0.181818182	0.590909091	1	0	0.636363636	1	0.045454545	0	0.181818182											
	0.318181818	0.136363636	0	0	0.090909091	0	0	1	0.318181818											
	0.045454545	0	0	0.272727273	0	0.954545455	0	0.454545455												
	mu=12.242896		sigma=1.880825		threshold=9.421658															
MOTIF 61:	>oc_M4939_1.01		sd 1.5																	
	0.179487179	0.021978022	0.941391941	0.904761905	0	0.005494505	0.048832272	0.121212121												
	0.313186813	0	0.047619048	0.095238095	0	0.93040293	0.709129512	0.390572391												
	0.141025641	0.043956044	0	0	0.093406593	0	0.038216561	0.2996633												
	0.366300366	0.934065934	0.010989011	0	0.906593407	0.064102564	0.203821656	0.188552189												
	mu=8.702774		sigma=2.811466		threshold=4.485575															
MOTIF 62:	>odd_M2128_1.01		sd 1.5																	
	0.466666667	0.666666667	0	0.8	0	0	1	0	0.733333333	0.333333333										
	0.333333333	0.333333333	0.933333333	0	0	0	0	0	1	0.2	0.266666667									
	0	0	0.066666667	0.2	1	0	0	1	0	0.066666667	0.4									
	0.2	0	0	0	1	0	0	0	0	0										
	mu=16.124337		sigma=1.638387		threshold=13.666756															
MOTIF 63:	>ovo_M4960_1.01		sd 1.5																	
	0.376693767	0.296296296	0.113022113	0.44963145	0.014742015	0.01965602	0.007371007	0.004914005	0.054054054	0.987714988										
	0.071253071	0.233415233	0.095823096	0.167487685																
	0.06504065	0.024691358	0.029484029	0.383292383	0.985257985	0.98034398	0	0	0	0.007371007	0.152334152									
	0.223587224	0.4004914	0.248768473																	
	0.25203252	0.503703704	0.054054054	0.049140049	0	0	0.985257985	0	0.002457002	0.004914005	0.213759214									
	0.27027027	0.093366093	0.307881773																	
	0.306233062	0.175308642	0.803439803	0.117936118	0	0	0.007371007	0.995085995	0.943488943	0	0.562653563									
	0.272727273	0.41031941	0.275862069																	
	mu=13.720333		sigma=3.060606		threshold=9.129424															
MOTIF 64:	>pan_M1915_1.01		sd 1.5																	
	0	0	0	0	0.098591549	0.154929577	0.098591549	0	0	0	0	0.154929577	0.478873239	0.309859155						

0	1	0	0.281690141	0.605633803	0.14084507	0.563380282	0.422535211	0	0	0	0.126760563	0	0
0.478873239	0	1	0.563380282	0.042253521	0.267605634	0	0.211267606	0	0	0.154929577	0.577464789		
0.295774648	0												
0.521126761	0	0	0.154929577	0.253521127	0.436619718	0.338028169	0.366197183	1	1	0.845070423	0.14084507		
0.225352113	0.690140845												
mu=14.826421	sigma=2.421125	threshold=11.194734											
MOTIF 65:													
>pho_M4974_1.01	sd	1.5											
0.284653465	0.115023474	0.1002331	0.076923077	0.020979021	0	0	1	0.023310023	0.153846154	0.27972028	0.055944056		
0.076923077	0.151658768	0.211586902											
0.163366337	0.070422535	0.566433566	0.792540793	0	1	1	0	0	0.20979021	0.013986014	0.186480186		
0.135198135	0.208530806	0.241813602											
0.361386139	0.488262911	0.144522145	0.102564103	0.892773893	0	0	0	0	0	0.121212121	0.137529138	0.060606061	
0.475524476	0.376777251	0.188916877											
0.190594059	0.32629108	0.188811189	0.027972028	0.086247086	0	0	0	0	0.976689977	0.515151515	0.568764569		
0.696969697	0.312354312	0.263033175	0.35768262										
mu=13.260887	sigma=3.297621	threshold=8.314457											
MOTIF 66:													
>pnr_M2200_1.01	sd	1.5											
0.454545455	0.390103567	0.248561565	0.232451093	1	0	0	0	0.962025316	0.077100115	0.692750288			
0.126582278	0.223245109	0.609896433	0.100115075	0	0	1	0	0.001150748	0	0.00575374			
0.3141542	0.31990794	0	0	0	0	1	0.036823936	0	0.207134638				
0.104718067	0.066743383	0.141542002	0.667433832	0	1	0	0	0	0.922899885	0.094361335			
mu=13.599606	sigma=2.349751	threshold=10.074980											
MOTIF 67:													
>retn_M0111_1.01	sd	1.5											
0.227052445	0.364833245	0.954294957	0.003226175	0.180784555	0.601594601	0.500975127	0.400462448						
0.269126365	0.112105499	0.000149714	0.022005608	0.422972015	0.071491874	0.146431311	0.148802511						
0.203401331	0.244555038	0.000360946	0.000879235	0.060177951	0.188591594	0.170128912	0.221987261						
0.300419859	0.278506218	0.045194383	0.973888982	0.336065479	0.138321931	0.18246465	0.22874778						
mu=4.099024	sigma=2.032783	threshold=1.049849											
MOTIF 68:													
>rn_M4988_1.01	sd	1.5											
0.284140969	0.279661017	0.31779661	0.247881356	0.394067797	0.112288136	0.980932203	0.955508475	1	1	0.993644068			
0.953389831	0.48940678	0.362288136	0.247881356	0.315677966	0.292372881	0.283898305	0.343220339	0.368644068	0.408898305				
0.355932203	0.332627119	0.255874674	0.207843137										
0.36123348	0.311440678	0.271186441	0.343220339	0.146186441	0.491525424	0.004237288	0.01059322	0	0	0			
0.046610169	0.283898305	0.302966102	0.343220339	0.277542373	0.319915254	0.315677966	0.330508475	0.273305085	0.254237288				
0.294491525	0.201271186	0.373368146	0.341176471										
0.259911894	0.190677966	0.220338983	0.194915254	0.332627119	0.171610169	0.008474576	0	0	0	0	0		
0.082627119	0.19279661	0.245762712	0.224576271	0.201271186	0.243644068	0.156779661	0.205508475	0.222457627	0.211864407				
0.245762712	0.174934726	0.266666667											

0.094713656	0.218220339	0.190677966	0.213983051	0.127118644	0.224576271	0.006355932	0.033898305	0	0	0.006355932
0	0.144067797	0.141949153	0.163135593	0.18220339	0.186440678	0.156779661	0.169491525	0.152542373	0.11440678	
0.137711864	0.220338983	0.195822454	0.184313725							
mu=12.182066	sigma=2.903127	threshold=7.827375								
MOTIF 69:										
>run_M4991_1.01	sd	1.5								
0.419354839	0.903225806	0.935483871	0	0	0.225806452	0	0.967741935	0.709677419		
0.096774194	0	0	1	0.967741935	0.032258065	1	0	0		
0	0.064516129	0.064516129	0	0.032258065	0.741935484	0	0	0.290322581		
0.483870968	0.032258065	0	0	0	0	0	0.032258065	0		
mu=13.529779	sigma=2.331814	threshold=10.032058								
MOTIF 70:										
>scrt_M5003_1.01	sd	1.5								
0.46743295	0.204697987	0.07165109	0.987538941	0	0	0.003115265	0	0.009345794	0.087227414	0.112149533
0.38317757	0.38317757									
0.22605364	0.617449664	0.819314642	0	0.99376947	0.956386293	0.00623053	0	0	0.003115265	0.009345794
0.442367601	0.171339564									
0.256704981	0.110738255	0.034267913	0.009345794	0	0	0.012461059	0.987538941	0	0.105919003	0.856697819
0.068535826	0.22741433									
0.049808429	0.067114094	0.074766355	0.003115265	0.00623053	0.043613707	0.978193146	0.012461059	0.990654206	0.803738318	
0.021806854	0.105919003	0.218068536								
mu=16.109866	sigma=3.555741	threshold=10.776254								
MOTIF 71:										
>sens_M5007_1.01	sd	1.5								
0.318965517	0.303609342	0.343220339	0.139830508	0.73940678	0.985169492	0.972457627	0	0.012711864	0.822033898	
0.055084746	0.455508475	0.177966102	0.06122449	0.452970297						
0.310344828	0.284501062	0.25	0.292372881	0.171610169	0	0.002118644	0	0.951271186	0	0.834745763
0.021186441	0.802721088	0.297029703								
0.219827586	0.218683652	0.279661017	0.139830508	0.044491525	0	0	0.01059322	0	0.008474576	0.088983051
0.398305085	0.764830508	0.01814059	0.086633663							
0.150862069	0.193205945	0.127118644	0.427966102	0.044491525	0.014830508	0.025423729	0.98940678	0.036016949	0.169491525	
0.021186441	0.137711864	0.036016949	0.117913832	0.163366337						
mu=13.391161	sigma=3.757074	threshold=7.755550								
MOTIF 72:										
>shn_M5008_1.01	sd	1.5								
0	0	0	1	0.111111111	0	0	0	0.125		
0	0	0	0	0	0.666666667	1	1	0.888888889	0.125	
0.555555556	1	1	0	0	0	0	0	0		
0.444444444	0	0	0	0	0.888888889	0.333333333	0	0	0.111111111	0.75
mu=18.430350	sigma=1.949648	threshold=15.505877								
MOTIF 73:										
>slp1_M0663_1.01	sd	1.5								

0.148384371	0.343158793	0.083876382	0.999359858	0.915696854	0.999359858	0.000213378	0.999359858	0.437060066	0.332663323
0.253628962	0.02776782	0.000213378	0.000213378	0.000213378	0.000213378	0.912222651	0.000213378	0.186478177	0.222445557
0.148384371	0.517642556	0.000213378	0.000213378	0.000213378	0.000213378	0.000213378	0.000213378	0.185492009	0.222445557
0.449602289	0.111430823	0.915696854	0.000213378	0.083876382	0.000213378	0.087350584	0.000213378	0.190969739	0.222445557
mu=10.722530		sigma=2.262950		threshold=7.328105					
MOTIF 74:									
>slp2_M5018_1.01	sd	1.5							
0.333333333	0	1	1	1	0	1	0.611111111		
0	0.055555556	0	0	0	1	0	0.166666667		
0.666666667	0	0	0	0	0	0	0.166666667		
0	0.944444444	0	0	0	0	0	0.055555556		
mu=12.595841		sigma=1.477250		threshold=10.379965					
MOTIF 75:									
>sna_M5021_1.01	sd	1.5							
0.212093023	0.051959891	0.962625342	0.001823154	0.004557885	0.000911577	0	0.092980857	0.09206928	0.427529626
0.241567912	0.278030994	0.268978444	0.23752495						
0.586046512	0.919781222	0.016408387	0.991795807	0.987237922	0	0.005469462	0.174111212	0.372835005	0.268915223
0.481312671	0.346399271	0.375820056	0.423153693						
0.147906977	0.001823154	0.010027347	0.003646308	0.000911577	0	0.965360073	0	0.034639927	0.211485871
0.206927985	0.207122774	0.162674651							0.121239745
0.053953488	0.026435734	0.010938924	0.002734731	0.007292616	0.999088423	0.029170465	0.732907931	0.500455789	0.09206928
0.155879672	0.16864175	0.148078725	0.176646707						
mu=13.811644		sigma=3.344246		threshold=8.795275					
MOTIF 76:									
>Sox15_M5027_1.01	sd	1.5							
0.047619048	0.857142857	0	0	0	0	0.19047619	0.142857143		
0.666666667	0	0	0	0	0	0.095238095	0.238095238		
0	0	0	0	1	0	0.095238095	0.047619048		
0.285714286	0.142857143	1	1	0	1	0.619047619	0.571428571		
mu=10.584262		sigma=1.962432		threshold=7.640614					
MOTIF 77:									
>tgo_ss_M5055_1.01	sd	1.5							
0.2	0	0	1	0	0.1	0	0.9		
0	0	1	0	0.9	0	1	0		
0.8	0	0	0	0.1	0.9	0	0.1		
0	1	0	0	0	0	0	0		

mu=14.268010      sigma=1.939614      threshold=11.358589										
MOTIF 78:										
>tgo_trh_M5059_1.01	sd	1.5								
0.238095238	0.047619048	0	0	1	0.047619048	0	0	0.857142857	0	
0.19047619	0.142857143	0	1	0	0.952380952	0	0	0.095238095	0.428571429	
0.523809524	0.761904762	0.095238095	0	0	0	1	0	0.047619048	0.19047619	
0.047619048	0.047619048	0.904761905	0	0	0	0	1	0	0.380952381	
mu=14.553419      sigma=2.508099      threshold=10.791271										
MOTIF 79:										
>dys_tgo_M4748_1.01	sd	1.5								
0.52173913	0.608695652	0.391304348	0	0	0	0	0	0.913043478	0	0.130434783
0.086956522	0	0	0	1	0	0	0	0.043478261	0.695652174	0.173913043
0.304347826	0.217391304	0.347826087	0	0	1	0	1	0.043478261	0.043478261	0.043478261
0.086956522	0.173913043	0.260869565	1	0	0	1	0	0	0.260869565	0.652173913
mu=14.394711      sigma=2.241800      threshold=11.032011										
MOTIF 80:										
>tin_M1924_1.01	sd	1.5								
0.085436893	0	1	0	0	0	0.108737864	0.900970874	0.368932039	0.475728155	
0.584466019	1	0	1	0	0.366990291	0.066019417	0	0.266019417	0.155339806	
0.304854369	0	0	0	0	0	0.825242718	0.099029126	0.365048544	0.104854369	
0.025242718	0	0	0	1	0.633009709	0	0	0	0.26407767	
mu=13.162619      sigma=2.062487      threshold=10.068888										
MOTIF 81:										
>Trl_M1884_1.01	sd	1.5								
0.056338028	0.084507042	0.169014085	0	0.042253521	0.014084507	0.183098592	0.070422535	0.098591549	0.126760563	
0.309859155	0.323943662	0.084507042	0.901408451	0	0.985915493	0	0.76056338	0.183098592	0.464788732	
0.126760563	0.154929577	0.422535211	0.042253521	0.197183099	0	0.028169014	0.098591549	0.154929577	0.14084507	
0.507042254	0.436619718	0.323943662	0.056338028	0.76056338	0	0.788732394	0.070422535	0.563380282	0.267605634	
mu=7.965974      sigma=3.150079      threshold=3.240856										
MOTIF 82:										
>usp_M5083_1.01	sd	1.5								
0.19047619	0	0	1	0	0	0	0	0.19047619		
0.047619048	0	0	0	1	1	0.285714286	0.571428571	0		

	0.33333333	0	1	0	0	0	0	0	0.285714286										
	0.428571429	1	0	0	0	0	0.714285714	0.428571429	0.523809524										
MOTIF 83:	mu=13.087672	sigma=1.121727	threshold=11.405082																
	>vnd_M5090_1.01	sd	1.5																
	0.011986301	0.729452055	0.017123288	0.003424658	0.001712329	0.321917808	0.786815068	0.344178082											
	0.584760274	0.066780822	0.974315068	0.004280822	0.163527397	0.179794521	0.010273973	0.211472603											
	0.001712329	0.020547945	0.005993151	0.035958904	0.014554795	0.476883562	0.054794521	0.344178082											
	0.401541096	0.183219178	0.002568493	0.956335616	0.820205479	0.02140411	0.148116438	0.100171233											
MOTIF 84:	mu=7.835463	sigma=2.816287	threshold=3.611032																
	>Vsx1_M1011_1.01	sd	1.5																
	0.322643538	0.143724162	0.015039639	0.964759964	0.978334784	0.006494691	0.011006577	0.705052561	0.246711813										
	0.227363058	0.143724162	0.011366715	0.008430696	0.008809687	0.00243753	0.010575003	0.005197992	0.082948581										
	0.3425947	0.048808082	0.00998106	0.018208333	0.003407067	0.007698188	0.016919943	0.278181104	0.644336372										
	0.107398704	0.663743594	0.963612586	0.008601007	0.009448462	0.983369591	0.961498478	0.011568343	0.026003234										
MOTIF 85:	mu=10.307768	sigma=3.067908	threshold=5.705905																
	>Vsx2_M1013_1.01	sd	1.5																
	0.143060837	0.000322165	0.99909201	0.99909201	0.000302663	0.000302663	0.895661157	0.091197822											
	0.048003802	0.064755155	0.000302663	0.000302663	0.000302663	0.000302663	0.000344353	0.181941924											
	0.048003802	0.000322165	0.000302663	0.000302663	0.000302663	0.000302663	0.103650138	0.499546279											
	0.760931559	0.934600515	0.000302663	0.000302663	0.99909201	0.99909201	0.000344353	0.227313975											
MOTIF 86:	mu=11.170827	sigma=2.140773	threshold=7.959669																
	>CHES-1-like_snw_prim	sd	1.5																
	0.33765309	0.303076748	0.286023877	0.242090491	0.496843913	0.345249556	0.20037772	0.537414724	0.01873036	0.921424381									
	0.925301795	0.976316921	0.006982374	0.953833343	0.894458512	0.383696157	0.185356644	0.144489359	0.101382458	0.209927536									
	0.215955868	0.149825462																	
	0.162574851	0.123163264	0.186785865	0.192126734	0.138521353	0.075989067	0.241083741	0.012000903	0.081109568	0.066096176									
	0.037171484	0.005559191	0.946143603	0.006081319	0.018259554	0.179078927	0.34940349	0.126066794	0.189338709	0.241305721									
	0.203231669	0.20170226																	
	0.194621139	0.131538419	0.163533012	0.200413259	0.179700233	0.23646922	0.359547009	0.442252791	0.008145674	0.001920677									
	0.004114961	0.006387352	0.002989982	0.016468412	0.025629314	0.074724176	0.135407872	0.311617494	0.116012025	0.326785576									
	0.166670829	0.397313213																	
	0.305150921	0.442221568	0.363657247	0.365369517	0.1849345	0.342292156	0.198991529	0.008331582	0.892014399	0.010558767									
	0.03341176	0.011736536	0.043884041	0.023616926	0.06165262	0.362500741	0.329831993	0.417826353	0.593266808	0.221981167									
	0.414141634	0.251159066																	

	mu=12.584244	sigma=3.851293	threshold=6.807305																
MOTIF 87:	>CHES-1-like_snw_sec	sd	1.5																
	0.355958217	0.164614473	0.121492567	0.38612324	0.412840422	0.23629575	0.450289568	0.776833298	0.087013383	0.011310211									
	0.988603856	0.004358941	0.009478461	0.004359321	0.019688632	0.805750544	0.430217278	0.188648426	0.24398479	0.251125765									
	0.531516561	0.173733912	0.176500359																
	0.427264942	0.290268897	0.25629959	0.212561019	0.271704751	0.10597402	0.161580536	0.016398459	0.086404094	0.013044796									
	0.002568426	0.985692144	0.000814299	0.989649056	0.66167156	0.032739359	0.250488026	0.190657318	0.072542053	0.381157102									
	0.12767027	0.455660865	0.339793489																
	0.126435847	0.269162448	0.253490609	0.164240975	0.145459719	0.221655307	0.28252387	0.175735912	0.177803381	0.974351182									
	0.006213604	0.001352677	0.985615452	0.004072166	0.12194641	0.132252841	0.13206911	0.294144673	0.377661578	0.127728571									
	0.213634571	0.20850557	0.209833292																
	0.090340994	0.275954182	0.368717234	0.237074766	0.169995108	0.436074922	0.105606026	0.031032331	0.648779141	0.001293811									
	0.002614114	0.008596238	0.004091788	0.001919457	0.196693398	0.029257256	0.187225585	0.326549582	0.305811578	0.239988562									
	0.127178598	0.162099653	0.27387286																
MOTIF 88:	mu=14.935922	sigma=3.838537	threshold=9.178116																
	>jumu_bml	sd	1.5																
	0.30034486	0.35033655	0.29704267	0.23201877	0.15588586	0.96752335	0.10084446	0.0166085	0.00543079	0.16087107									
	0.2074585	0.17748645	0.21142616	0.24636601	0.117816	0.01978135	0.79917454	0.00008374	0.99434535	0.25997962									
	0.23625963	0.23018752	0.27147656	0.22070308	0.64491858	0.00247129	0.00013996	0.98231779	0.00002154	0.31125212									
	0.25593701	0.24198948	0.22005462	0.30091213	0.08137956	0.01022402	0.09984104	0.00098997	0.00020232	0.26789718									
MOTIF 89:	mu=7.755305	sigma=2.478217	threshold=4.037980																
	>slou_bml	sd	1.5																
	0.303646632	0.216112711	0.117463264	0.437256276	0.717859366	0.044395795	0.000981968	0.842353561	0.291081128	0.208047495									
	0.220492709	0.291721646	0.457659583	0.230561922	0.093342328	0.026131586	0.050998868	0.012886752	0.187466848	0.295233448									
	0.280301169	0.294653497	0.068451512	0.135709893	0.037950159	0.00136772	0.057501063	0.030759529	0.312186892	0.264480201									
	0.19555949	0.197512146	0.356425641	0.19647191	0.150848147	0.928104899	0.890518102	0.114000158	0.209265132	0.232238856									
MOTIF 90:	mu=4.929028	sigma=2.752555	threshold=0.800195																
	>ttk-PA_SOLEXA	sd	1.5																
	0.358	0.355	0.535	0.342	0.039	0.468	0.702	0	0	0	0.064	0.46	0.356	0.349					
	0.268	0.254	0.167	0.338	0.764	0.454	0.113	1	1	1	0.841	0.523	0.211	0.323	0.287				
	0.189	0.184	0.068	0.133	0.149	0	0.003	0	0	0	0.019	0.146	0.183	0.195					
	0.185	0.206	0.229	0.187	0.048	0.078	0.182	0	0	0	0.159	0.394	0.183	0.139	0.17				
MOTIF 91:	mu=11.924584	sigma=2.783903	threshold=7.748729																
	>ttk-PF_SOLEXA	sd	1.5																

0.685	0.999	0	0	0.999	0	0.769	0.306	0.231	0.306	0.304
0.301	0	0	0	0	0.544	0.074	0.348	0.344	0.328	0.32
0.005	0	1	1	0.001	0.001	0.061	0.009	0.189	0.211	0.235
0.009	0	0	0	0	0.455	0.096	0.338	0.235	0.155	0.141
mu=11.294729		sigma=1.903760		threshold=8.439089						
MOTIF 92:										
>Ubx_bmlsd	1.5									
0.267693023	0.214841827	0.155433382	0.364361216	0.704926555	0.093800313	0.277999318	0.57114856	0.408627926	0.263939767	
0.202318035	0.24466797	0.360040753	0.379231079	0.069971892	0.068113007	0.080448615	0.111904894	0.198900565	0.243647114	
0.317297839	0.387572167	0.090578532	0.109743489	0.179126744	0.032824271	0.029009364	0.124916979	0.185453658	0.238822578	

## Supplementary table 9

Enrichment of published ChIP signals at validated silencers for the first library. ChIP-chip and/or ChIP-seq signal for various chromatin proteins was mapped onto the coordinates of the tested library elements, and enrichment/depletion (area under the receiver-operator characteristic curve [AUROC]) at validated silencers and its significance (two-tailed Wilcoxon test p-value corrected for multiple hypothesis testing) are shown.

### Keys

sorted.mesoderm.histone:	Area under the receiver-operator characteristic curve (AUC), and p-value of enrichment/depletion of histone post-translational modifications, as measured by ChIP-seq in sorted mesodermal cells, at validated silencers. Adjusted.p-value shows p-values after correction for multiple hypothesis testing. See Methods for sources of data.
whole.embryo.histone:	Area under the receiver-operator characteristic curve (AUC), and p-value of enrichment/depletion of histone post-translational modifications, as measured by ChIP-chip or ChIP-seq in whole embryos, at validated silencers. Adjusted.p-value shows p-values after correction for multiple hypothesis testing.
TF.coact.corepr:	Area under the receiver-operator characteristic curve (AUC), and p-value of enrichment/depletion of TF, coactivator, or corepressor localization, as measured by ChIP-chip or ChIP-seq in whole embryos, at validated silencers. Adjusted.p-value shows p-values after correction for multiple hypothesis testing.

**sorted.mesoderm.histone**

<b>test</b>	<b>AUC</b>	<b>p-value</b>	<b>adjusted.p-value</b>
H3K27me3 rep1	0.66298	0.03501	
H3K27me3 rep2	0.5784946	0.3104	
JZ H3K27me3	0.6368664	0.07671	
H3K27me3.combined.p		0.02769432	0.181474
H3K4me1 rep1	0.6150538	0.1368	
H3K4me1 rep2	0.6086022	0.1603	
H3K4me1.combined.p		0.1056967	0.38548208
H3K4me3 rep1	0.5360983	0.6416	
H3K4me3 rep2	0.4835637	0.833	
H3K4me3.combined.p		---	1
H3K27ac rep1	0.4061444	0.225	
H3K27ac rep2	0.3465438	0.04693	
H3K27ac.combined.p		0.05861179	0.28314923
H3K36me3 rep1	0.6064516	0.1693	
H3K36me3 rep2	0.5339478	0.6617	
H3K36me3.combined.p		0.3572532	0.70249664
H3K79me3 rep1	0.5259601	0.7368	
H3K79me3 rep2	0.5453149	0.5601	
H3K79me3.combined.p		0.7779375	0.95650196

## whole.embryo.histone

<b>test</b>	<b>embryo ages</b>	<b>data source</b>	<b>AUC</b>	<b>p-value</b>	<b>adjusted.p-value</b>
H3K4me1	4-8	modENCODE_778	0.6952381	0.01154	0.17887
H3K4me3	0-12	modENCODE_622	0.7210445	0.004236	
H3K4me3	4-8	modENCODE_790	0.5929339	0.2296	
H3K4me3.combined.p				0.007718	0.159505416
H3K9ac	4-8	modENCODE_822	0.5963134	0.2131	0.600554545
H3K9me3	0-12	modENCODE_621	0.6436252	0.06322	
H3K9me3	4-8	modENCODE_802	0.6411674	0.06788	
H3K9me3.combined.p				0.02768429	0.181474
H3K27ac	4-8	modENCODE_835	0.5023041	0.9778	1
H3K27me3	0-12	modENCODE_919	0.6571429	0.04209	
H3K27me3	4-8	modENCODE_811	0.6442396	0.0621	
H3K27me3.combined.p				0.01815787	0.181474

## TF.coact.corepr

test	embryo ages	data source	AUC	p-value	adjusted.p-value
Kr	0-8	modENCODE_898	0.5751152	0.3318	0.702253333
Sin3A	0-12	GSM569791	0.4764977	0.7624	0.956501961
Su(var)3-9	0-12	GSM636838	0.4562212	0.5722	0.840945455
Pcl	0-8	GSM569800	0.4958525	0.9588	1
D	0-8	GSM628262	0.6436252	0.06322	
D	0-8	GSM621330	0.7176651	0.004859	
D.combined.p				0.00279172	0.11292463
BEAF-32	0-12	modENCODE_21	0.4215054	0.3104	0.702253333
CP190	0-12	modENCODE_22	0.4261137	0.3398	0.702253333
CTCF	0-12	modENCODE_769	0.4371736	0.4171	
CTCF	0-12	modENCODE_770	0.5010753	0.9905	
CTCF.combined.p				1	1
CtBP	0-12	modENCODE_607	0.503533	0.9651	1
en	0-12	modENCODE_3184	0.564977	0.4013	
en	0-12	modENCODE_625	0.5050691	0.9493	
en.combined.p				0.7486039	0.956501961
Trl	0-12	modENCODE_23	0.4583717	0.5913	0.840945455
Gro	0-12	modENCODE_623	0.6457757	0.05937	0.283149231
hairy	0-8	modENCODE_2574	0.6525346	0.0485	0.273363636
hkb	0-8	modENCODE_2575	0.6353303	0.08008	0.330997333
jumu	0-8	modENCODE_2576	0.6801843	0.01976	0.181474
mod(mdg4)	0-12	modENCODE_24	0.5490015	0.5271	0.840945455
run	0-12	modENCODE_617	0.5447005	0.5641	0.840945455
sbb	0-12	modENCODE_609	0.5751152	0.3318	0.702253333
sens	4-8	modENCODE_2577	0.6466974	0.05778	
sens	4-8	modENCODE_978	0.6064516	0.1687	
sens	4-8	modENCODE_979	0.6362519	0.07805	
sens.combined.p				0.02584075	0.181474
Su(Hw)	0-12	modENCODE_27	0.5400922	0.6051	
Su(Hw)	0-12	modENCODE_901	0.5437788	0.5722	
Su(Hw).combined.p				0.7134683	0.956501961
ttk	0-12	modENCODE_615	0.4749616	0.7473	0.956501961
Ubx	3-8	modENCODE_612	0.6070661	0.1663	
Ubx	3-8	modENCODE_613	0.6374808	0.0754	
Ubx.combined.p				0.06744626	0.29869058

zfh1	0-12	modENCODE_604	0.5447005	0.5641	0.840945455
pan.ave	0-8	modENCODE_4074	0.6196621	0.1218	0.411484211
Su(H) repl 1	0-8	modENCODE_5017	0.3050691	0.01167	
Su(H) repl 3	0-8	modENCODE_5017	0.6980031	0.01042	
Su(H).combined.p				1	1
hairy.ave	0-8	modENCODE_4982	0.569278	0.3708	0.702496636
Pc.rep1	5-13	GSE55257	0.6138249	0.1411	
Pc.rep2	5-13	GSE55257	0.5889401	0.2503	
Pc.combined.p	5-13	GSE55257		0.1533966	0.47552946
ph.rep1	5-13	GSE55257	0.5763441	0.3239	
ph.rep2	5-13	GSE55257	0.5087558	0.9113	
ph.combined.p	5-13	GSE55257		0.6553376	0.902909582
Su(z)12.rep1	5-13	GSE55257	0.5081413	0.9176	
Su(z)12.rep2	5-13	GSE55257	0.5145929	0.8517	
Su(z)12.combined.p	5-13	GSE55257		0.974176	1
HP1a	6-20	GSE56101	0.66851	0.02927	0.181474
pho (PhoN ab)	6-12	E-TABM-525 (ArrayExpress)	0.5072115	0.9298	
pho (PhoZ ab)	6-12	E-TABM-525 (ArrayExpress)	0.4041896	0.2312	
pho.combined.p	6-12	E-TABM-525 (ArrayExpress)		1	1
bab1	0-12	modENCODE_628	0.4589862	0.5968	0.840945455
chinmo	0-12	modENCODE_608	0.4211982	0.3085	0.702253333
disco	0-8	modENCODE_2572	0.5797235	0.3029	0.702253333
Dll	0-12	modENCODE_606	0.5529954	0.4939	0.840945455
exd	0-8	modENCODE_3183	0.4362519	0.4103	0.748194118
ftz-f1	0-12	modENCODE_624	0.5754224	0.3298	0.702253333
GATAe	0-8	modENCODE_2573	0.506298	0.9366	1
inv	0-12	modENCODE_605	0.5066052	0.9334	
inv	0-12	modENCODE_619	0.5554531	0.474	
inv.combined.p				0.803221	0.957686577
kn	0-12	modENCODE_618	0.5210445	0.7868	0.956501961
nej	4-8	modENCODE_855	0.4528418	0.5428	0.840945455
Stat92E	1-12	modENCODE_616	0.506298	0.9366	1
all.pub.HDAC	0-12	GSE20000	0.3683564	0.08867	0.34359625
sna	2-4	doi: 10.1101/gad.1509607 (Young lab data download page)	0.5852982	0.2868	0.702253333
twi	2-4	doi: 10.1101/gad.1509607 (Young lab data download page)	0.6057559	0.1864	0.55032381
dl	2-4	doi: 10.1101/gad.1509607 (Young lab data download page)	0.6223994	0.1261	0.411484211

## Supplementary table 10

Enrichment of published ChIP signals at validated silencers for the first and second library combined . ChIP-chip and/or ChIP-seq signal for various chromatin proteins was mapped onto the coordinates of the tested library elements, and enrichment/depletion (area under the receiver-operator characteristic curve [AUROC]) at validated silencers and its significance (two-tailed Wilcoxon test p-value corrected for multiple hypothesis testing) are shown.

### Keys

sorted.mesoderm.histone:	Area under the receiver-operator characteristic curve (AUC), and p-value of enrichment/depletion of histone post-translational modifications, as measured by ChIP-seq in sorted mesodermal cells, at validated silencers. Adjusted.p-value shows p-values after correction for multiple hypothesis testing. See Methods for sources of data.
whole.embryo.histone:	Area under the receiver-operator characteristic curve (AUC), and p-value of enrichment/depletion of histone post-translational modifications, as measured by ChIP-chip or ChIP-seq in whole embryos, at validated silencers. Adjusted.p-value shows p-values after correction for multiple hypothesis testing.

### sorted.mesoderm.histone

histone.mark	AUC	p.value	adj.p
meso.K27ac	0.3765346	0.02765	0.15932
meso.K27me3	0.5891961	0.1117	0.26808
meso.K36me3	0.4982385	0.9757	0.9939
meso.K4me1	0.546493	0.4073	0.69822857
meso.K4me3	0.5004804	0.9939	0.9939
meso.K79me3	0.5361375	0.5196	0.75853333

### whole.embryo.histone

histone.mark	AUC	p.value	adj.p
emb.K27ac	0.468026	0.5689	0.75853333
emb.K27me3	0.615245	0.03983	0.15932
emb.K4me1	0.615245	0.03983	0.15932
emb.K4me3	0.5056048	0.9211	0.9939
emb.K9ac	0.5575958	0.3045	0.609
emb.K9me3	0.594534	0.0918	0.26808