Experimental proposal
Rodrigo Arzate-Mejia
Zürich November 2020

In situ functional validation of regulatory elements

Background. Regulatory elements, such as promoters, enhancers, silencers and insulators are vastly abundant in the genome of eukaryotes. In particular, mammalian genomes contain hundreds of thousands of genomic regions that, based on data from DNAse-seq, ATAC-seq and/or ChIP-seq experiments, are suggested to be regulatory elements, however, just a minority has been functionally validated *in vitro* and even less *in vivo*.

The advent of CRISPR-Cas9 technology for genome and epigenome editing have allowed to interrogate the activity of regulatory elements in their native context in an unprecedent manner. In particular, the use of the inactive form of the Cas9 (termed death Cas9, dCas9) in isolation or fused to activator (VP64) or repressor (KRAB) domains allows the activation or inactivation of regulatory elements by means of targeted chromatin remodeling (VP64 and KRAB) or by direct competition with the binding sites of Transcription Factors (TFs).

Experimental strategies for the functional validation of regulatory elements identified by ATAC-seq. An ATAC-seq experiment can identify genomic regions that are more accessible to the activity of the Tn5 enzyme. Such regions, termed Tn5 hypersensitive sites (THSS) are thought to represent regulatory elements based on their position in respect with genes, their overlap with Transcriptional Start Sites (TSS) as well as their enrichment for TF binding motifs (TFBMs) and of "active" histone-posttranslational modifications such as H34me2, H3K4me3 and H3K27ac.

To date, there are at least two major strategies to interrogate the potential regulatory activity of a THSS. One of such strategies is to test the regulatory activity of the DNA sequence at the THSS out of its chromosomal context employing reporter plasmids that test the capacity of a given DNA sequence to enhance, silence, induce or insulate the transcription of a reporter gene (Luciferase or GFP). In contrast, direct manipulation of the THSS in its chromosomal context allow to study the endogenous activity of such THSS and its effect in gene expression. To manipulate the activity of THSS in its chromatin context the system CRISPR-Cas9 has been extensively employed, in particular, the use of the inactive form of the Cas9 (dCas9).

The most employed strategy to inactivate regulatory elements identified by ATAC-seq is the use of heterochromatinization of a THSS by directing the dCas9-KRAB complex to the THSS of interest (Figure 1, left). Single or multiple sgRNAs spanning the THSS can be used to attract dCas9-KRAB complexes. The KRAB domain then recruit chromatin remodelers, DNMTSs, histone methyltransferases and the HP1 histone which overall results in chromatin compaction, H3K9me3 and gain of DNA methylation, which overall greatly reduces the probability of TF binding and therefore of Tn5 cut. Since ATAC-seq detect regulatory sequences with "active" histone-posttranslational modifications it is likely that the great majority of THSS could positively influence transcription constitutively or under specific stimulus or developmental stages. Finally, the CRISPR-dCas9 system was recently repurposed to inhibit the binding of TFs to their cognate motifs by steric hinderance and to inactivate regulatory elements (Figure 1, right).

Finally, while the CRISPR-dCas9 system allow for targeted transcriptional activation by recruitment of co-activators via the VP64 domain to a genomic site of interest, the use of dCas9-KRAB system remains standard to functionally validate THSSs identified from ATAC-seq data.

Experimental strategies for the functional validation of Differentially Accessible Regions (DARs) between juvenile and adult mouse spermatogonia. The chromatin accessibility landscape of spermatogonial cells is different between postnatal and adult stages. In particular, it was found that the transition to the adult stage is accompanied by the gain of THHSs at gene bodies, intergenic regions and putative promoter regions. Furthermore, THHSs more accessible in adult spermatogonia are enriched in binding motifs for the TF c-fos, which is also induced in the adult spermatonogial stage. Importantly, c-fos controls de expression of the cell cycle regulator Cyclin A2 (He at al., 2008) in spermatogonial cells and has been proposed to be important for the self-renewal capacity of such stem cells. Therefore, the increase in chromatin accessibility at thousands of THHSs with motif binding sites for c-fos and the concomitant transcriptional induction of c-fos in adult spermatogonial cells, suggest that c-fos binding could result in chromatin opening and activation of enhancers as has been previously shown in mouse neuronal cells.

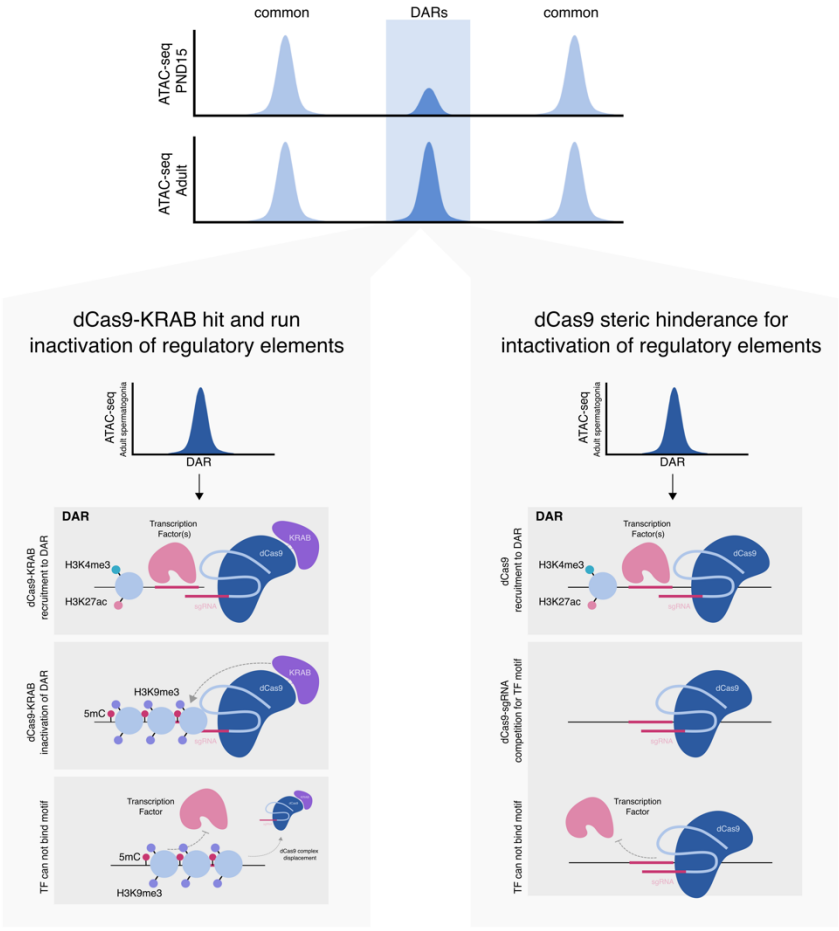
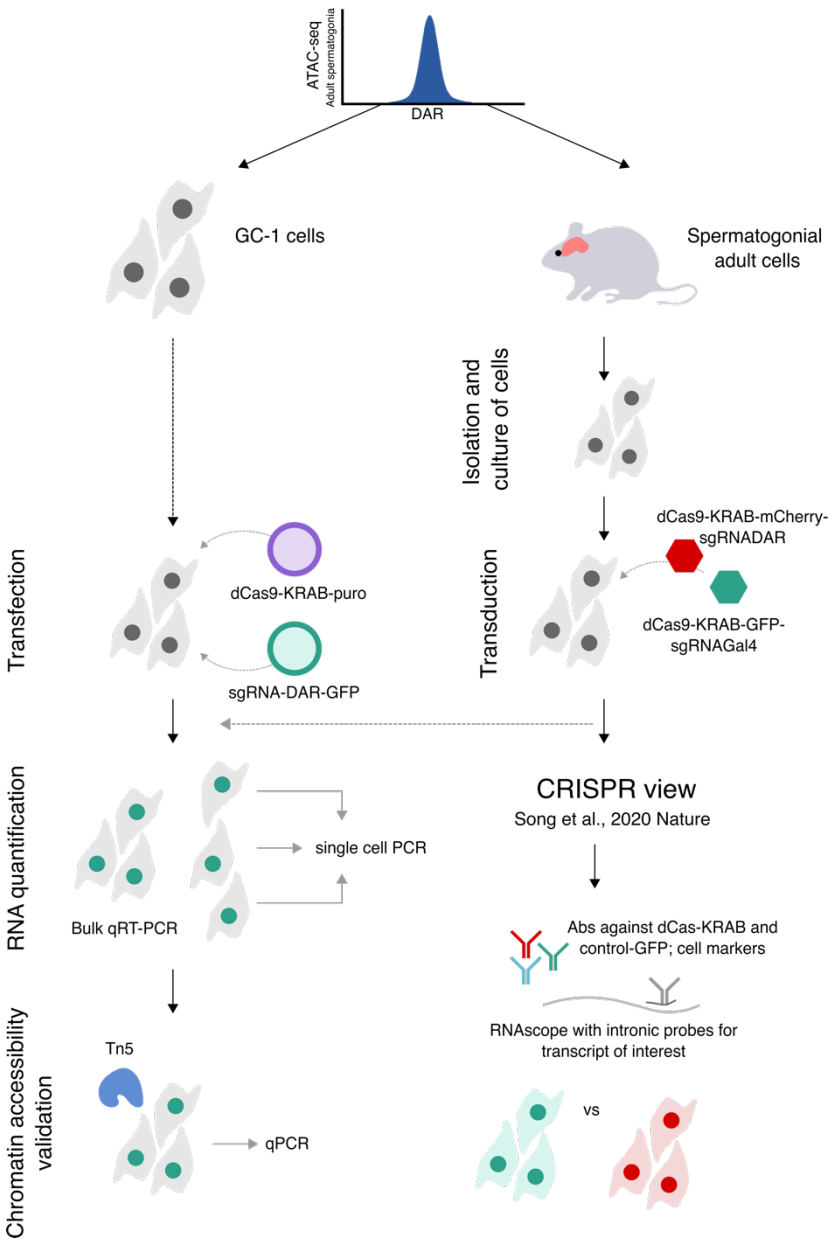


Figure 1. Strategies to inactivate DARs.

Based on these observations I suggest to perform CRISPRi (CRISPR inactivation) experiments of DARs with more accessibility in adult spermatogonia. The two general CRISPRi strategies are show in Figure 1. In either case, the dCas9 complex will be directed to the target DAR by a sgRNA directed against the c-Fos binding motif. In the case of the dCas9-KRAB strategy, the inactivation of the DAR is a consequence of the heterochromatinization of the THSSs via a hit and run mechanism and therefore does not require the continuous presence of the dCas9-KRAB complex at the genomic site of interest. In the case of using just dCas9, the inactivation of the DAR is achieved by the competition of the sgRNA for the TF motif of the protein of interest. As for this proposal, precluding the binding of c-fos to a DAR could cause its inactivation.

Below I describe different experimental approximations to inactivate DARs (more accessible in adult):



Inactivation of DARs in GC-1 cells.

Advantages

- Running in the lab.
- Previous experience in transfecting this specific cell line.
- Amount is no a limiting factor.

Disadvantages

- No pure population of spermatogonia cells.
- Not sure of the matching developmental time with our ATAC-seq datasets.

Suggested route

Given the availability and experience working with this cell line it might be strategic to test the inactivation system here. I suggest to use a dCas9-KRAB-puro plasmid which expresses at high levels the fusion protein and select transfected cells with puro (rutine in this kind of assays in cell lines). sgRNAs directed against c-fos motifs of the DARs of interest will be transfected in a separate plasmid encoding a GFP protein directed to the nuclei. GFP+ cells that survived selection then can be used to quantify transcription of the gene associated with the DAR. If desired, an additional control could be done to confirm the loss of accessibility by performing a Tn5 transposition reaction and then qPCR with the isolated DNA. We have all reagents needed to perform this experiment in the lab. We just need to choose the DARs of interest but I suggest to take a look at the enhancers that regulate c-fos.

Additional controls: stain for markers of spermatogonial cells.

Additional experiment

Hyphotesis: c-fos is important to promote opening of chromatin in spermatogonial cells. Repression of c-fos by targeting dCas9-KRAB directly to its promoter will result in decreased accessibility.

Experiment: target c-fos promoter with dCas9-KRAB, validate reduction and perform Tn5 reaction. Purify DNA and run qPCR against candidate DARs.

Expected results: inactivation of c-fos induces reduction in accessibility at selected THHSs.

Inactivation of DARs in adult spermatogonia cells. Given that the ATAC-seq data was derived from mouse spermatogonia, the best way to provide functional evidence on the relevance of DARs between developmental stages is to inactivate DARs or induce them in primary cultures of adult spermatogonia. Hue M. La et al., 2018 Nat Com (https://www.nature.com/articles/s41467-018-04827-z) applied a method to culture spermatogonial cells using MEF feeders (https://www.jove.com/t/50017/serial-enrichment-spermatogonial-stem-progenitor-cells-sscs-culture).

Cultures of primary cells can be used to perform CRISPR view a recently published method in Nature (October 2020; https://www.nature.com/articles/s41586-020-2825-4). This method was designed to inactivate regulatory elements such, as THSSs or in our case DARs, in primary culture cells that are limited by cell number. Briefly, cells are transduced with plasmids encoding dCas9-KRAB-mCherry-sgRNAs_DAR and as a control dCas9-KRAB-GFP-sgRNAs_Gal4. dCas9 alone can also be transduced in case we are interested in specifically inactivating the binding of a TF. Cells transduced at a low multiplicity will be transduced by on average a single viral particle. Then, a single cell culture will have both, cells with the target sgRNA, and other cells with the control one. Effects in expression will be evaluated via RNAscope and further selection of cells can be achieved via immunostaining with cell specific markers.

To me, CRISPR view is the ideal method, however, it requires extensive optimization, starting from making the primary cell cultures to work up the optimization of RNAscope, however, the Karayannis Group can do the technique, so we can learn from them.