# Crispr-cas9 history and engineering

A modular platform using engineered single short guide RNA to allow programming of CRISPR specificity, permitting high efficacy gene induction for analysis of gene function.

The ability to control gene expression has been the key method in elucidating their respective functions, pathways, and regulatory elements; paving a way for future therapeutic applications.

The two main approaches of determining gene function involve the analysis of loss-of-function (LOF) and gain-of-function (GOF) mutations. LOF involves a mutation in an allele where partial or full loss in genetic function occurs. GOF involves the introduction of a mutation which generates a new allele associated with a new function. The problem with GOF screening approaches is that they're hindered by a requirement for large comprehensive cDNA library overexpression systems which rarely encompass the full spectrum of isoform variation. Viral expression vectors are not large enough to allow these to be cloned. LOF screening is the predominant way of analysing gene function, using techniques such as Transcription-activator-like effector nucleases (TALENs) RNA interference and Zinc finger nucleases (ZFNs). However, these are difficult to construct on a genome wide scale, unlike CRISPR-Cas9.

## A brief history of CRISPR-Cas9

In 1987, Ishino et al observed the presence of CRISPR repeats within bacterial genomes, but it wasn't until 2006 that Makarova proposed for its use as an adaptive immune system. Cas9 or CRISPR associated protein 9 is an endonuclease, guided by RNA and associated with CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). The system functions by

interrogating and cleaving foreign DNA from bacteriophages by unwinding the foreign DNA and checking its complementation to a 20 base pair spacer region on the guide RNA. If the DNA substrate is complementary to the guide RNA, cleavage of the DNA occurs (Heler R, 2015). (Jinek M, 2012) discovered that by inactivating Cas9's two catalytic domains, its DNA cleavage ability is disrupted – thereby creating catalytically dead or dCas9. This provides a platform for an RNA-guided transcript activator (dCas9-activator) using a single short guided RNA (sgRNA).

## Engineering CRISPR-Cas9

In their article: Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex, Konermann et al develop a system using programmable DNA binding proteins for engineering synthetic transcription factors for the modulation of endogenous gene expression. This allowed GOF screening and was successful in turning on tens of thousands of individual genes in parallel.

To allow rational engineering of the CRISPR-Cas9 system, the structure of the Cas9-sg-RBA-target DNA tertiary complex had to be elucidated. To do this, crystallographic studies were performed. Optimal anchoring positions were determined for the activation domains. The team settled on the addition of protein interacting RNA aptamers to the tetraloop and stemloop 2 to facilitate the recruitment of effector domains to the Cas9, as illustrated in figure 1.

Fusion of the dCas9 to transcriptional activation domains converts the Cas9 nuclease into an dCas9-activator. Linking the dCas9 to domains of proteins

involved in transcriptional activation and allowing CRISPR to target promotor sequences regulating transcription of particular genes provides a means of modulating natural gene expression. The efficacy of this system is low – causing at most a fivefold increase in activation. Tiling the promoter region with several sgRNAs can produce a substantial transcriptional activation.

Konermann et al overcame this low efficiency by turning CRISPR sgRNA into a modular platform which assembles multiple different transcriptional activators. The addition of the protein interacting RNA aptamers attracts RNA binding proteins. The complex can be used to target the transcription activation domains of different transcription factors, creating a system termed the synergistic activator mediator (SAM) by its authors. Astonishingly, this complex can induce more than 100-fold activation of genes.

Parallels can be drawn with the cells natural mechanisms of gene regulation; enhancers can turn on gene expression by generating long non-coding RNAs (lncRNAs) which act as modular scaffolds, recruiting cellular machinery similarly to CRISPR. Konnermann’s findings appear to mimic the lncRNAs by orchestrating the use of multiple proteins to have them work in cohesion.

## Current Applications

The authors displayed the applications of this response by creating a library of sgRNAs, thereby allowing individual activation of over 23, 000 genes. Their experiments were centred around melanoma cancer cells. PLX-4720 is a common drug treatment, capable of killing these cancerous cells. The experiment involved activation of individual genes to establish which ones

would provide resistance to the killing effects of the PLX-4720 treatment. Drug resistance was determined by calculating the relative frequency of sgRNAs in melanoma cells post drug treatment. sgRNAs were correspondent to the genes involved in known drug-resistance pathways. This verified that the SAM technique could identify biologically significant outcomes of varied gene expression. It was determined that 13 genes whose altered gene expression produced a state of drug resistance.

## Potential applications

The significance of the findings of Konnermann et al are a new and improved programmable targeting system for DNA – by which RNA sequences can be engineered to determine specificity. Through this, single sgRNA-mediated gene upregulation can be performed. This next generation of CRISPR expands the Cas9 toolbox, further engineering may take advantage of the modular nature of this system. The scaffolding allows variation in the use of aptamers, for recruitment of specific effectors It has been proposed to replace the MS2 stem loops with PP7 elements to recruit repressive elements as opposed to activators, thereby opening the possibility of bidirectional transcript control. Further research is required to determine off target effects of CRISPR and validate experiments to confirm effects of altered gene expression. This will require a detailed understanding of regulatory elements and further experiments with gene sub libraries. Future applications will involve positive and negative selection screens to determine genetic elements in cells.