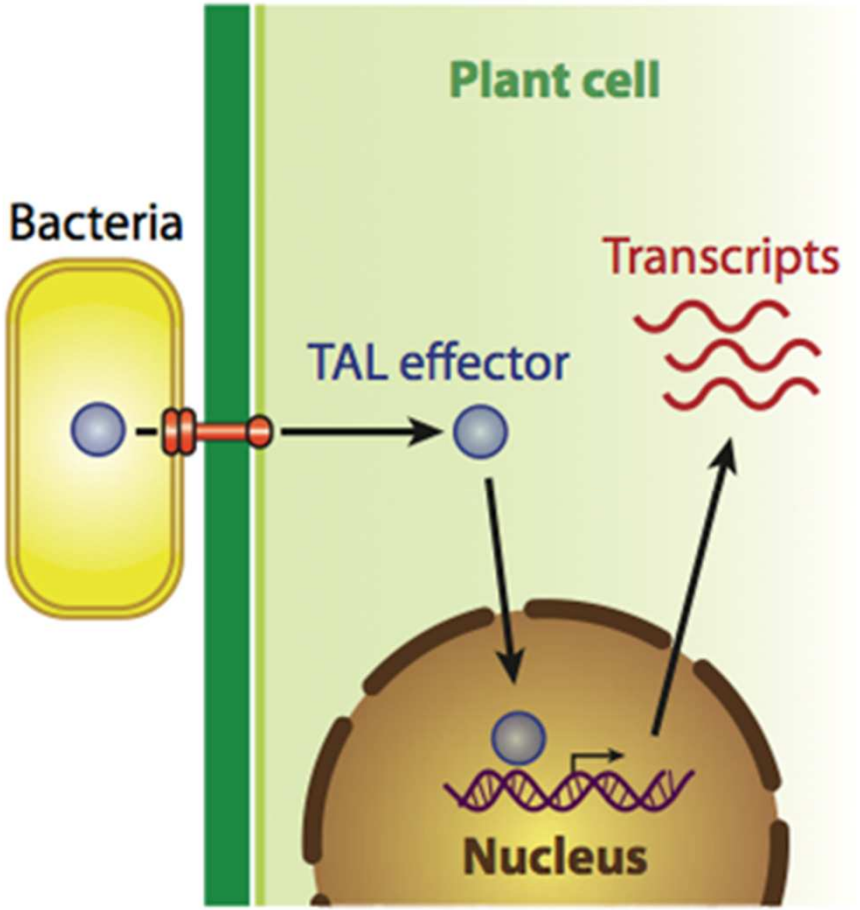
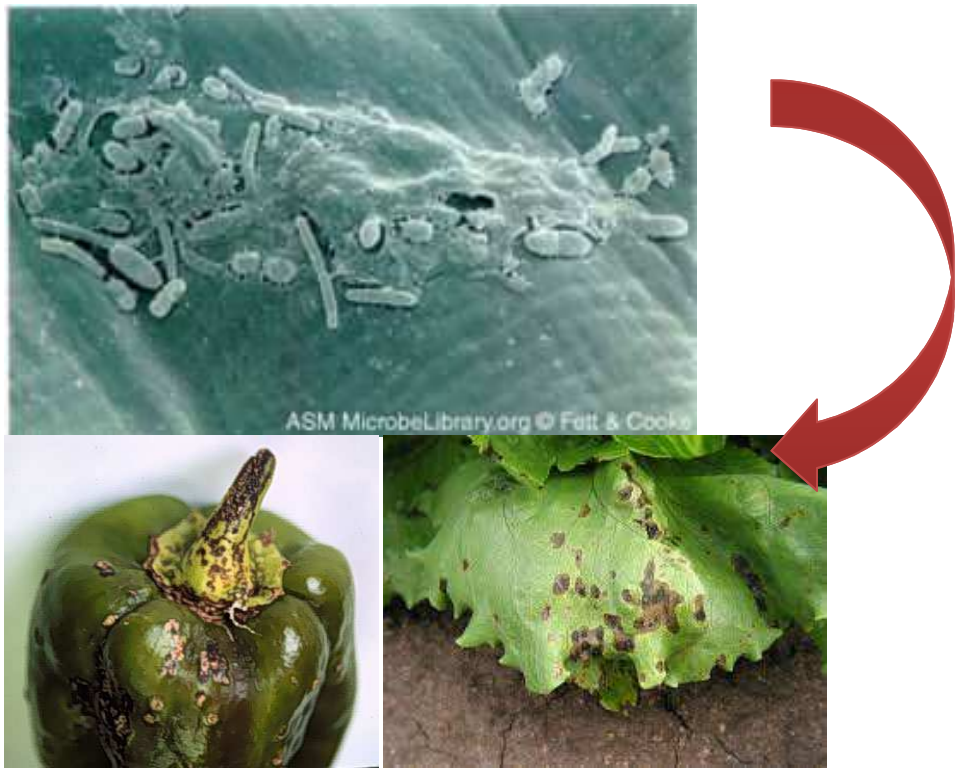
TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS (TALEs) AND TALE-BASED TECHNOLOGIES FOR GENOME ENGINEERING

Macromolecole Biologiche, 03-12-2015

Transcription activator-like effectors (TALEs)

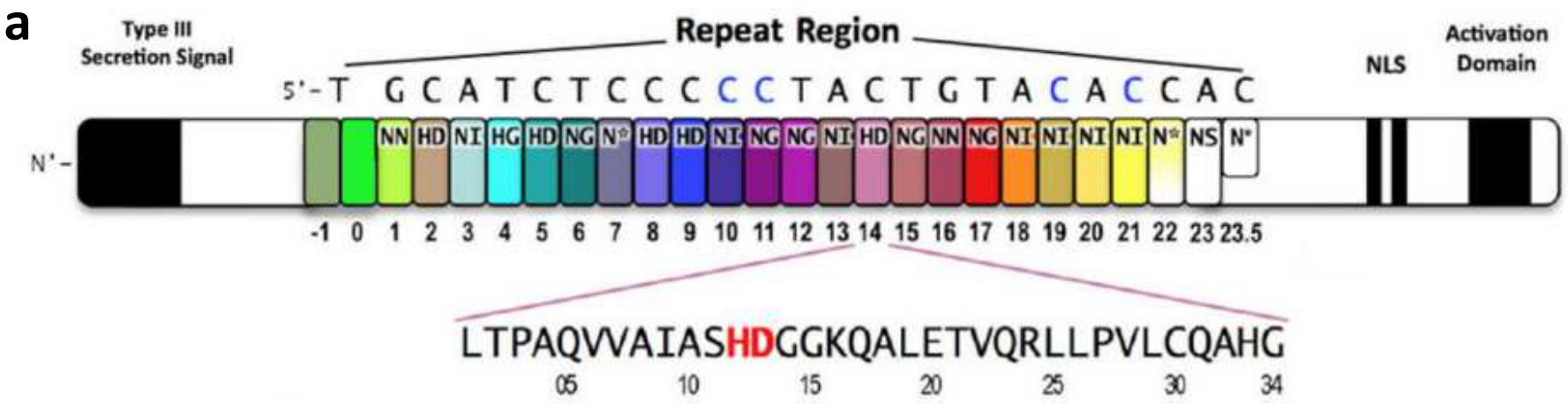


The first TAL effector identified was AvrBs3 from *Xanthomonas Campestris*, a pathogen of pepper.



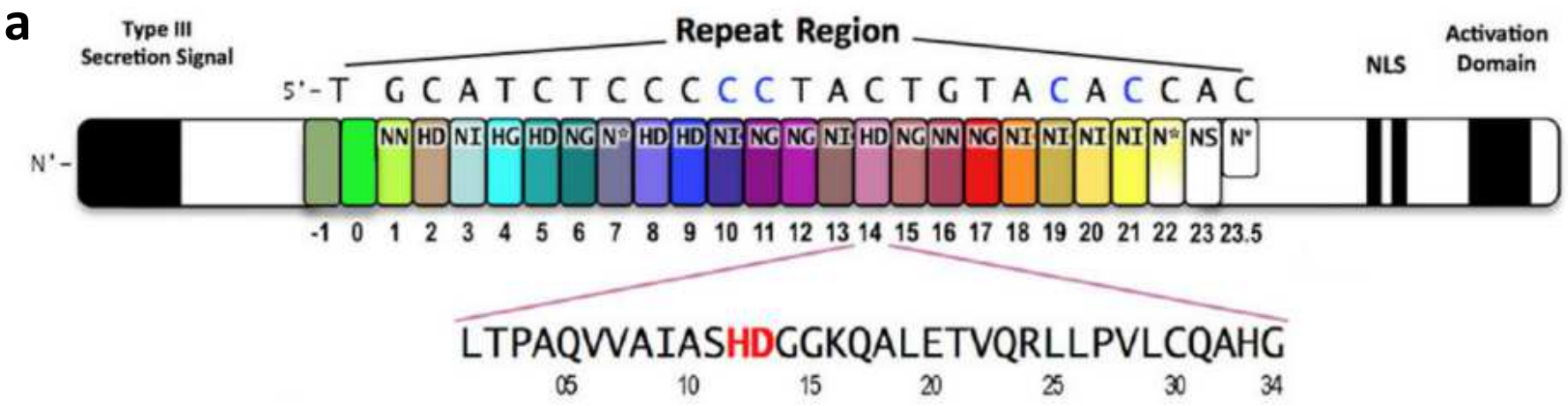
Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function
Boch and Bonas, *Annu. Rev. Phytopathol.* 2010.

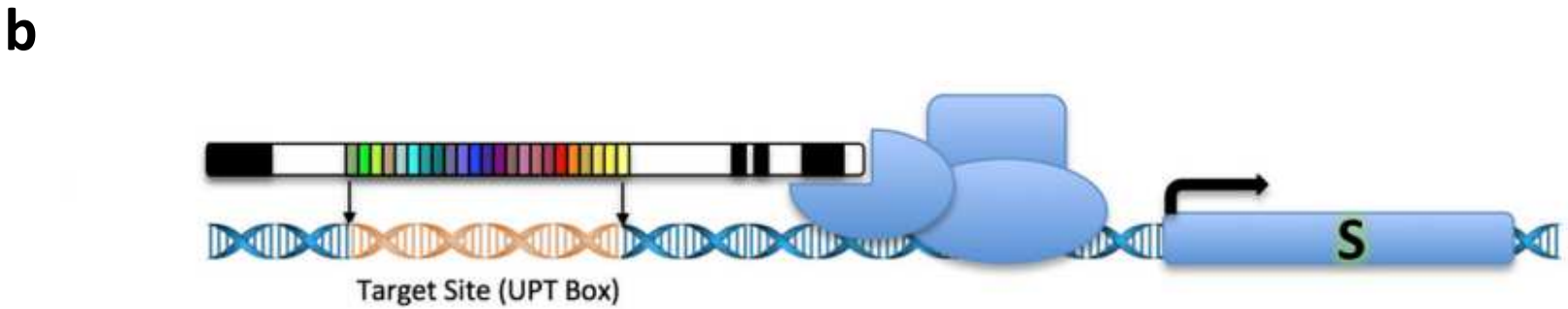
Tell me a tale of TALEs: protein structure



a) TAL effectors contain N-terminal signals for bacterial type III secretion, variable numbers of tandem repeats that specify the target nucleotide sequence, nuclear localization signals, and a C-terminal region that is required for transcriptional activation.

Tell me a tale of TALEs: protein structure and function

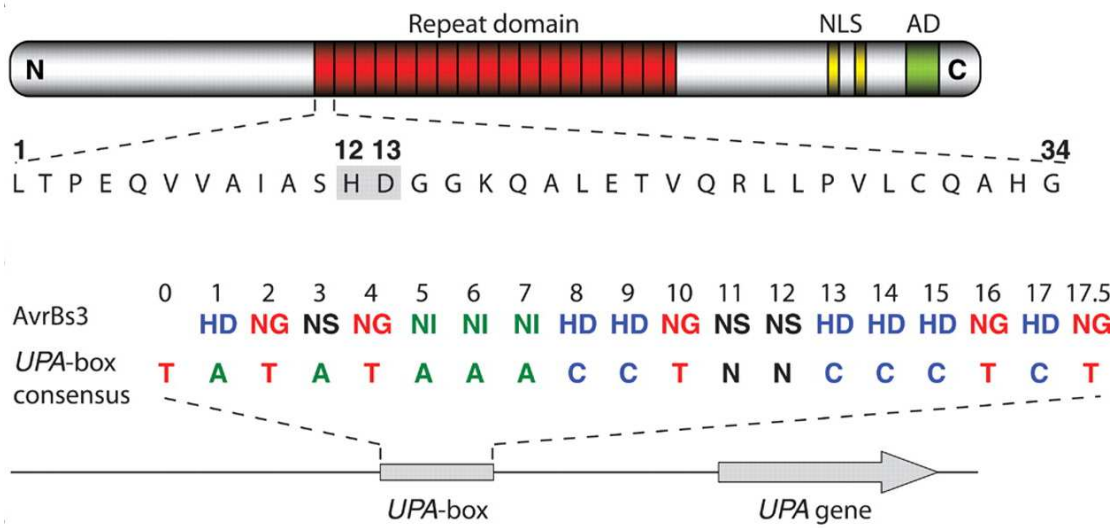




a) TAL effectors contain N-terminal signals for bacterial type III secretion, variable numbers of tandem repeats that specify the target nucleotide sequence, nuclear localization signals, and a C-terminal region that is required for transcriptional activation.

b) TAL effectors are translocated into the plant nucleus, where they bind to target sites located in the 5' promoter regions of genes that are subsequently activated.

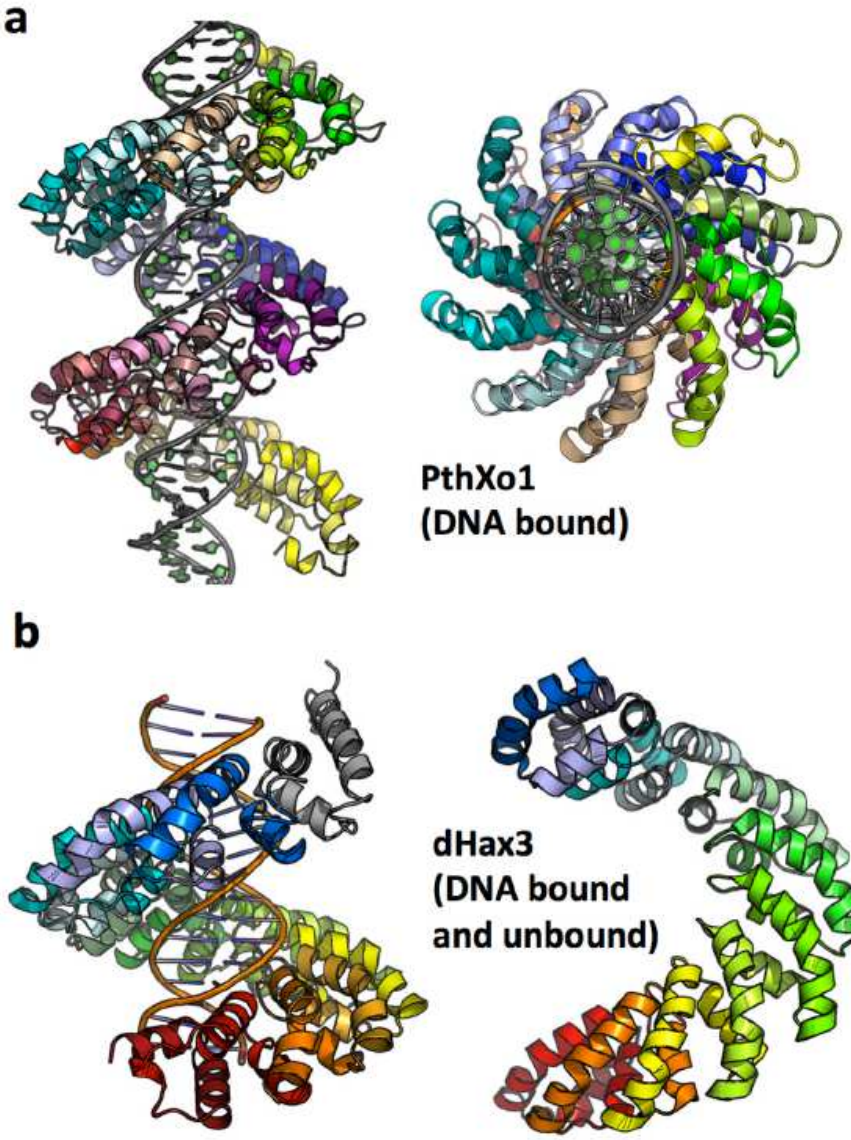
Breaking the code of TALE-DNA binding specificity



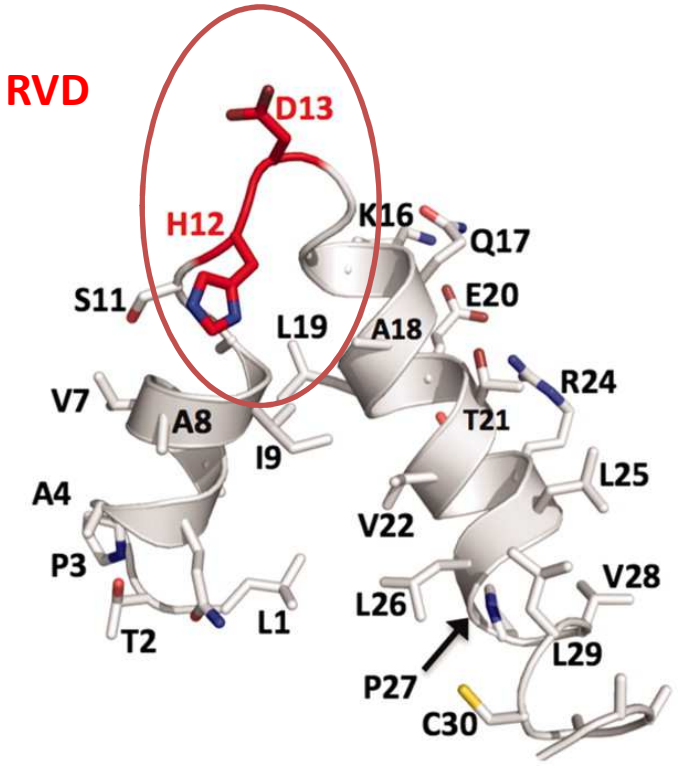
THE TALE CODE

Di-Amino acid		Nucleotide bound
NI	=	A
NG	=	T
HD	=	C
NN	=	G/A

The structural basis for the DNA recognition "code"

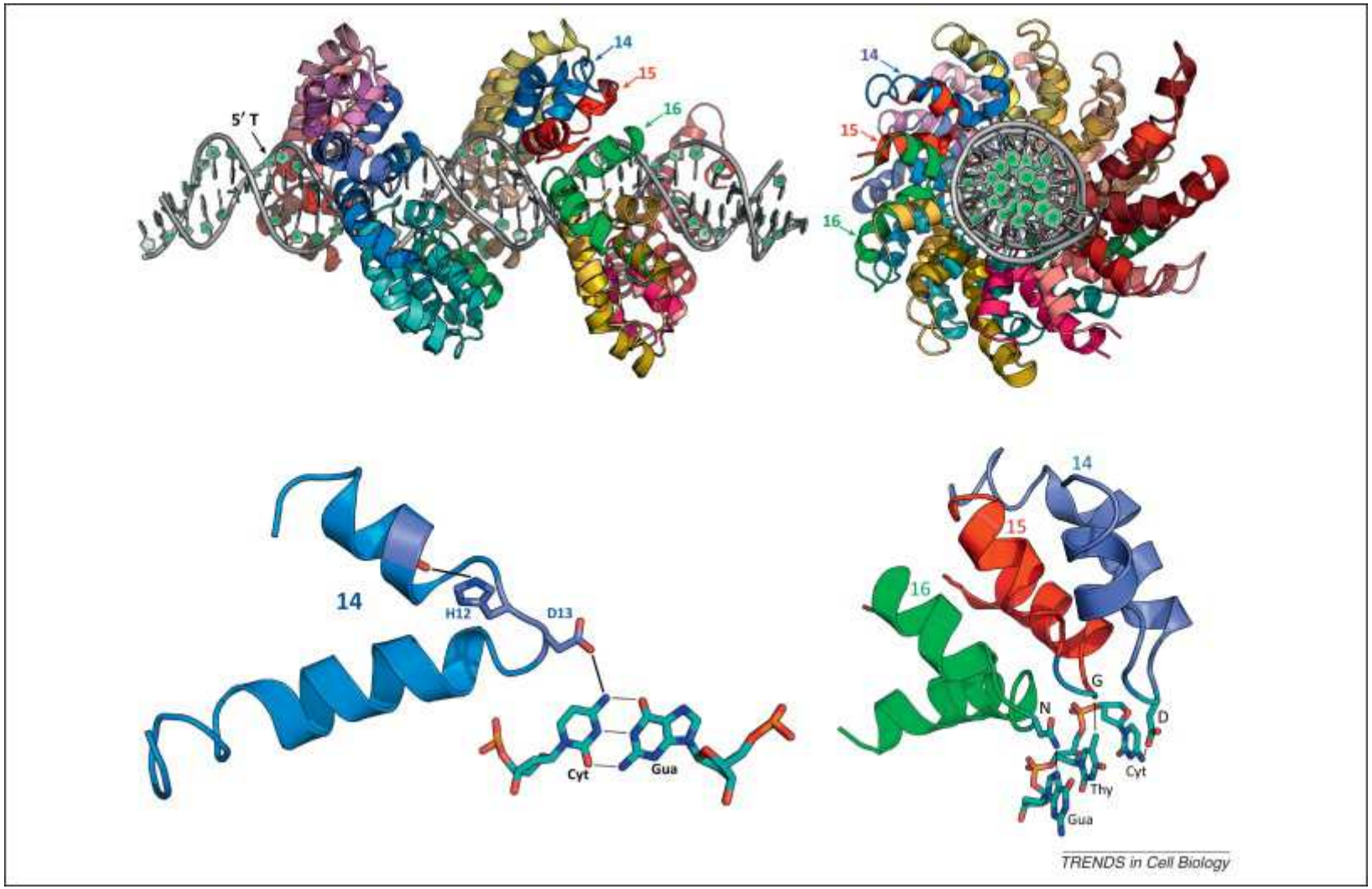


Each TAL repeat forms a left-handed, two-helix bundle, in which the two hypervariable residues in each repeat (RVD) are found at the end of the loop that connects the two helices.



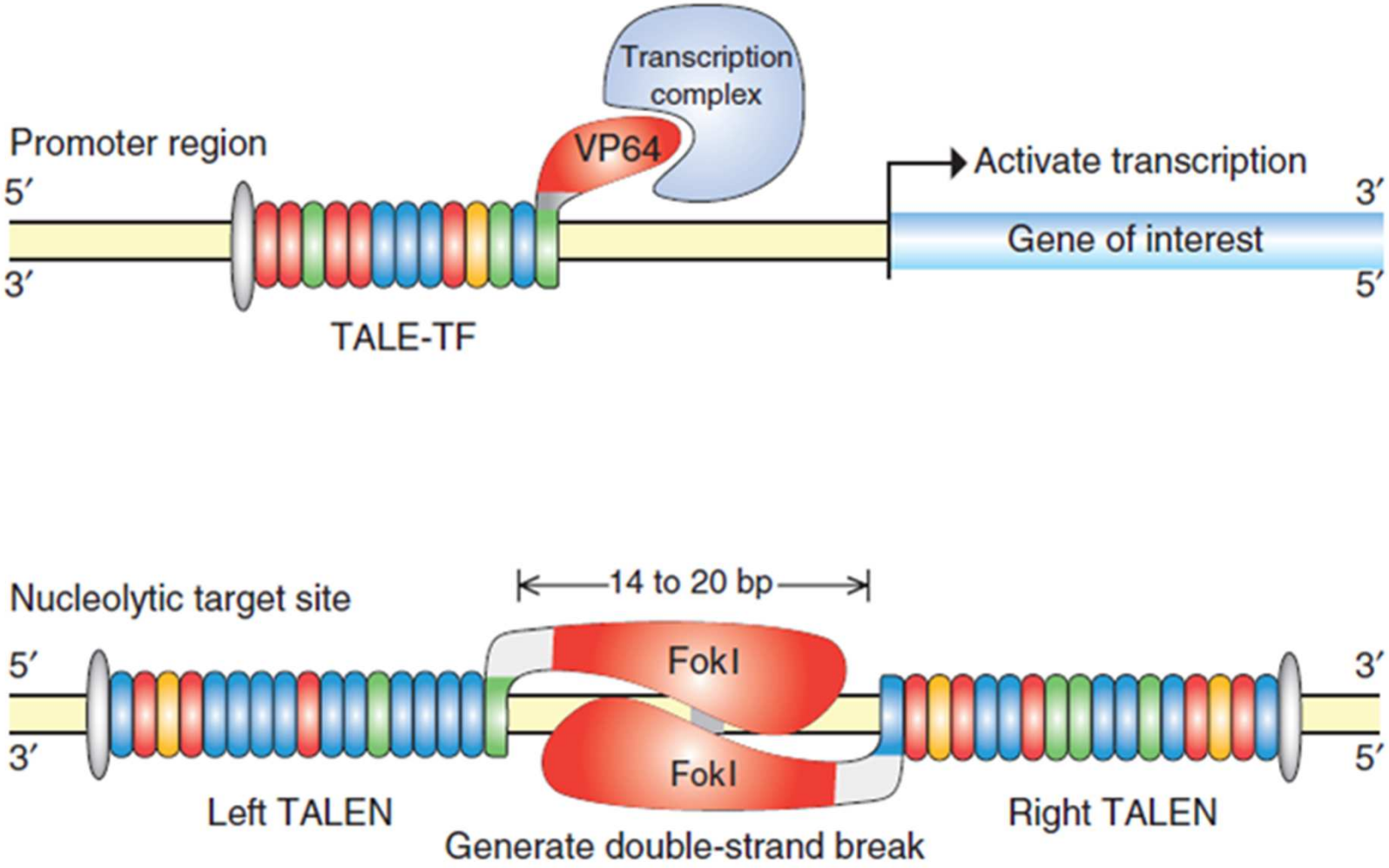
TAL effectors: function, structure, engineering and applications. *Curr Opin Struct Biol. 2013*

Recognition mechanism



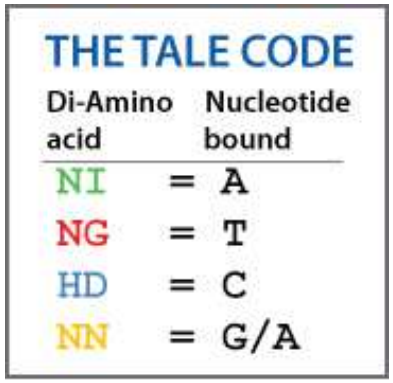
The sequence-specific contacts between the effector and the DNA are formed solely by the second residue of each RVD (at position 13 in each repeat) to atoms on the major groove edge of each base on a single contiguous strand of the DNA target.

TALE-derived technologies: TALE-TF and TALEN



GENOME EDITING APPLICATIONS

Genome modifications	Description	Genome editing tools
Gene tagging	Add a fusion tag (e.g. luciferase, GFP) to track an endogenous promoter activity or an endogenous protein expression and location	TALEN
Gene mutagenesis	Introduce point mutations to an endogenous gene	TALEN or CRISPR
Gene knockout	Introduce deletions or insertions (e.g. a selection marker) to knockout an endogenous gene	TALEN or CRISPR
Gene activation	Activate an endogenous gene expression	TALE-TF or CRISPR-TF
Gene repression	Repress an endogenous gene expression	TALE-R or CRISPR-R
Safe harbor knockin	Knockin an exogenous ORF or other genetic element to safe harbor sites of human or mouse genome	TALEN or CRISPR



Published in final edited form as:
Nat Commun. ; 3: 968. doi:10.1038/ncomms1962.

Comprehensive Interrogation of Natural TALE DNA Binding Modules and Transcriptional Repressor Domains

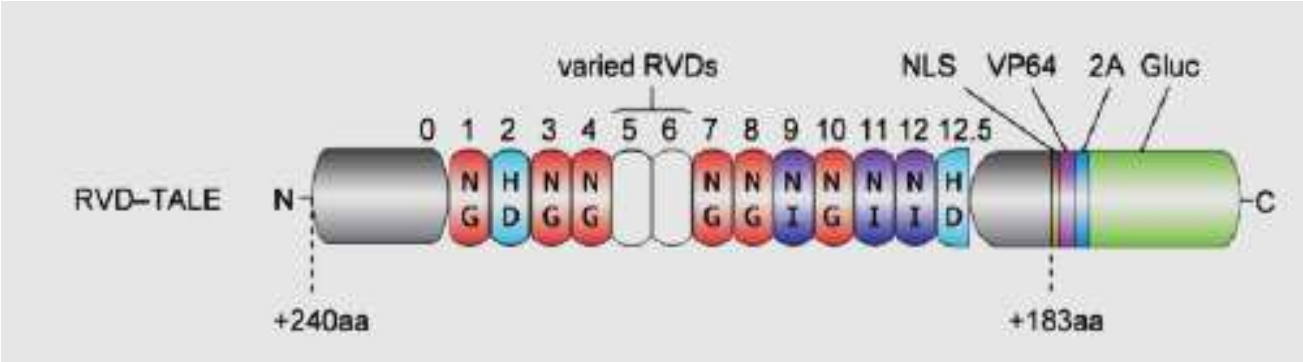
Le Cong[1,2], **Ruhong Zhou**[3], **Yu-chi Kuo**[1], **Margaret Cunniff**[1], and **Feng Zhang**[1,*]

Objectives:

- To identify a more specific guanine-binding RVD with higher biological activity, evaluating a total of 23 naturally occurring RVDs from the set of known *Xanthomonas* TALE sequences

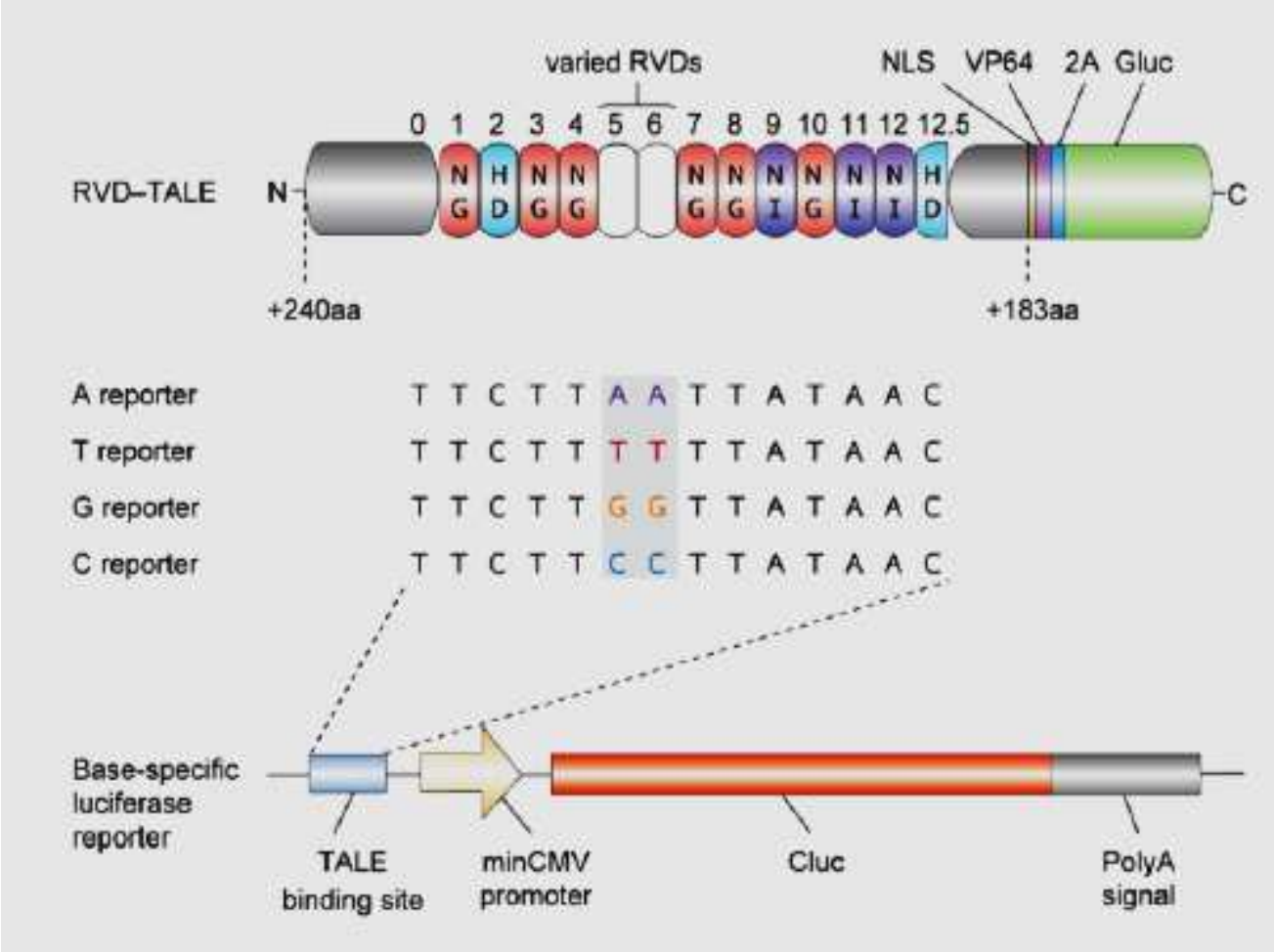
- To develop a TALE DNA binding domain fused with a repression domain in order to target transcriptional repression od endogenous mammalian gene expression.

Experimental approach: screening of 23 RVD-TALE



Comparison of each RVD-TALE's ability to activate transcription from each of the four base-specific *Cypridina* luciferase reporter (Cluc) plasmids with A, G, T, and C substituted in the 6th and 7th positions of the TALE-binding site.

Experimental approach: screening of novel RVD-TALE

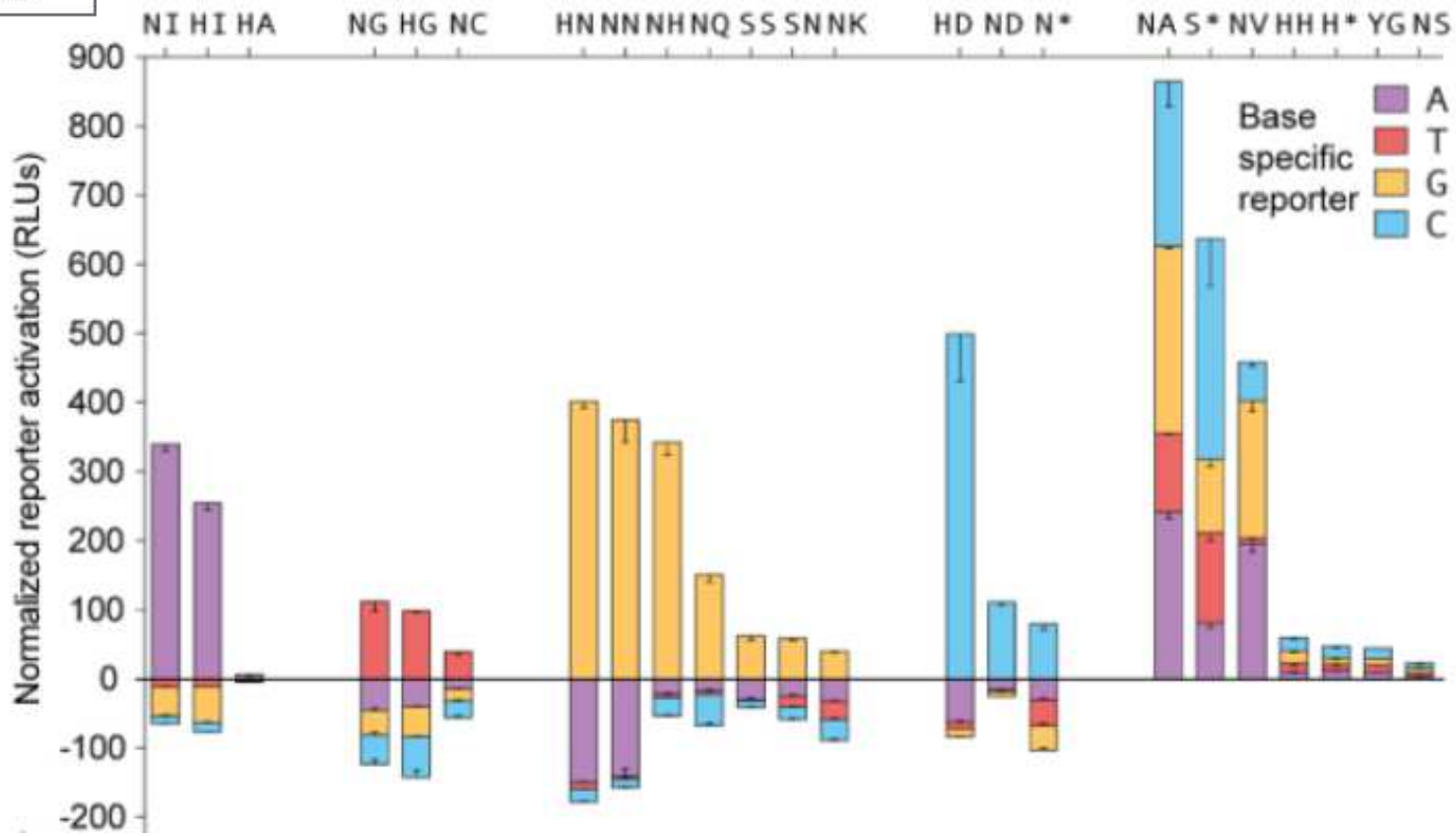


Comparison of each RVD-TALE's ability to activate transcription from each of the four base-specific *Cypridina* luciferase reporter (Cluc) plasmids with A, G, T, and C substituted in the 6th and 7th positions of the TALE-binding site.

Experimental approach: screening of novel RVD-TALE

THE TALE CODE

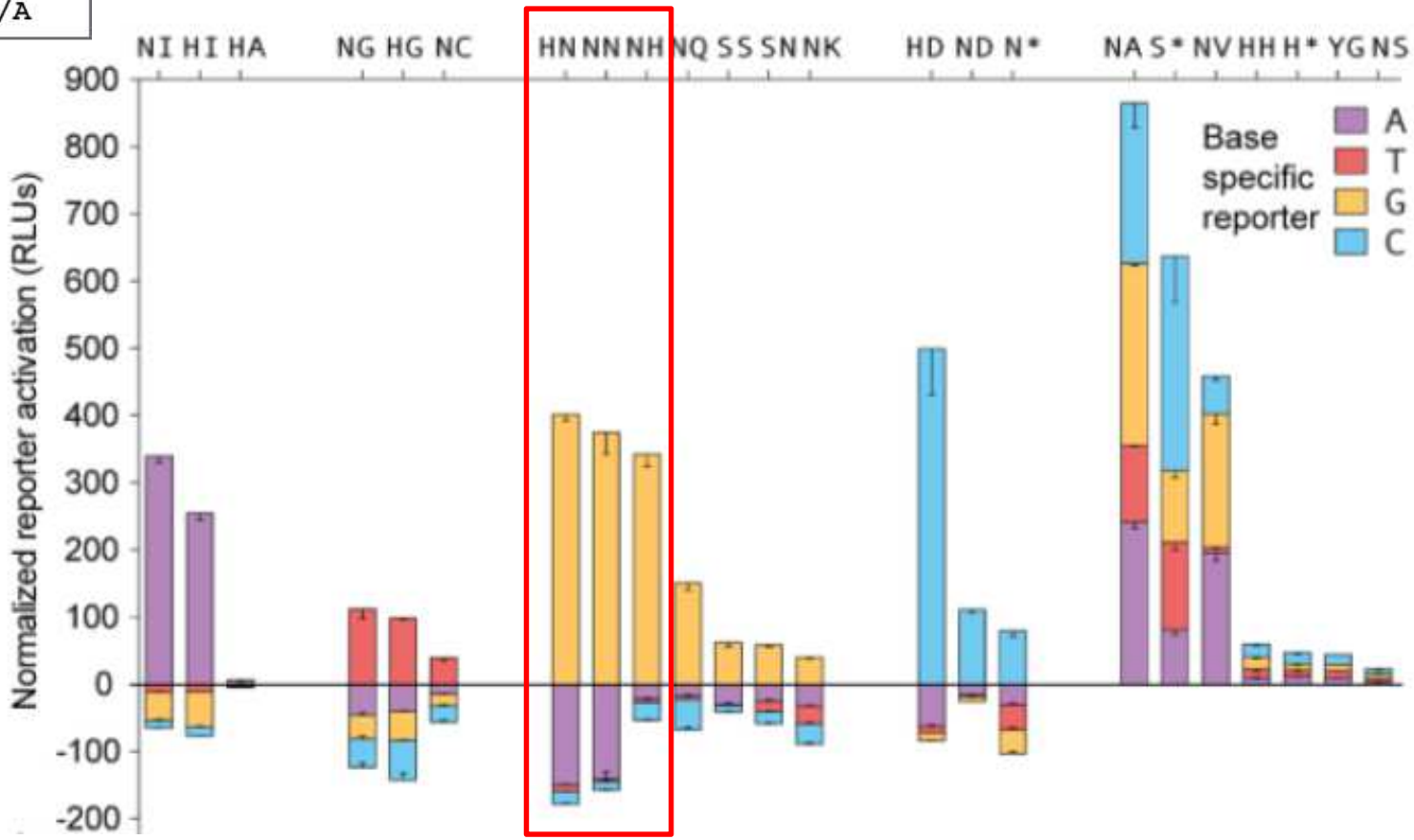
Di-Amino acid	Nucleotide bound
NI	= A
NG	= T
HD	= C
NN	= G/A



Interestingly, the NH-TALE exhibited significantly higher specificity for the G-reporter than the NN-TALE, suggesting that NH might be a better RVD for targeting guanines.

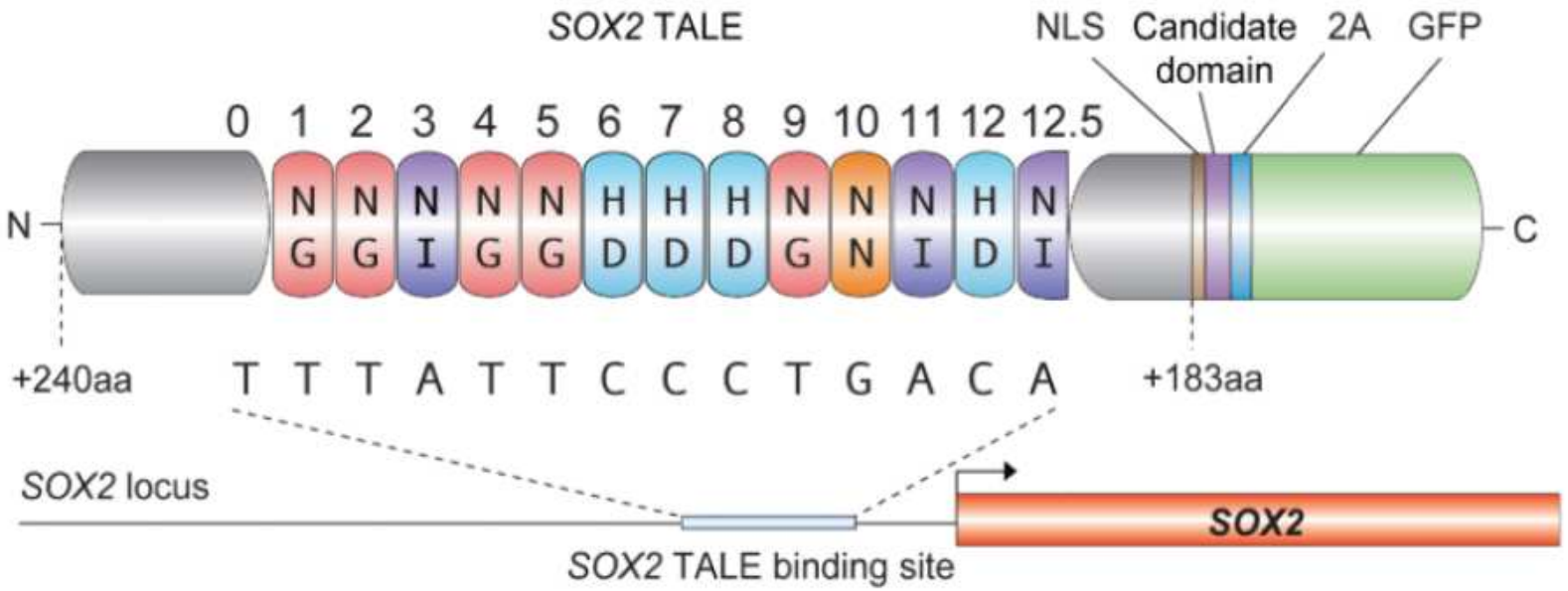
Experimental approach: screening of novel RVD-TALE

THE TALE CODE	
Di-Amino acid	Nucleotide bound
NI	= A
NG	= T
HD	= C
NN	= G/A



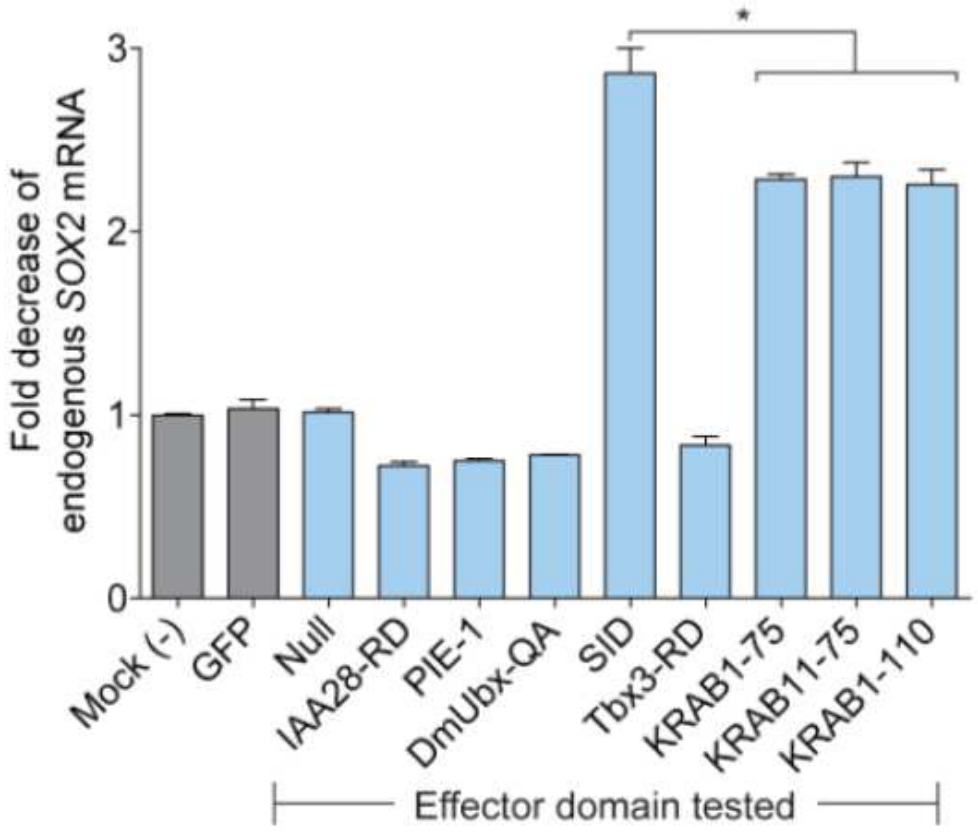
Interestingly, the NH-TALE exhibited significantly higher specificity for the G-reporter than the NN-TALE, suggesting that NH might be a better RVD for targeting guanines.

Development of a TALE transcriptional repressor

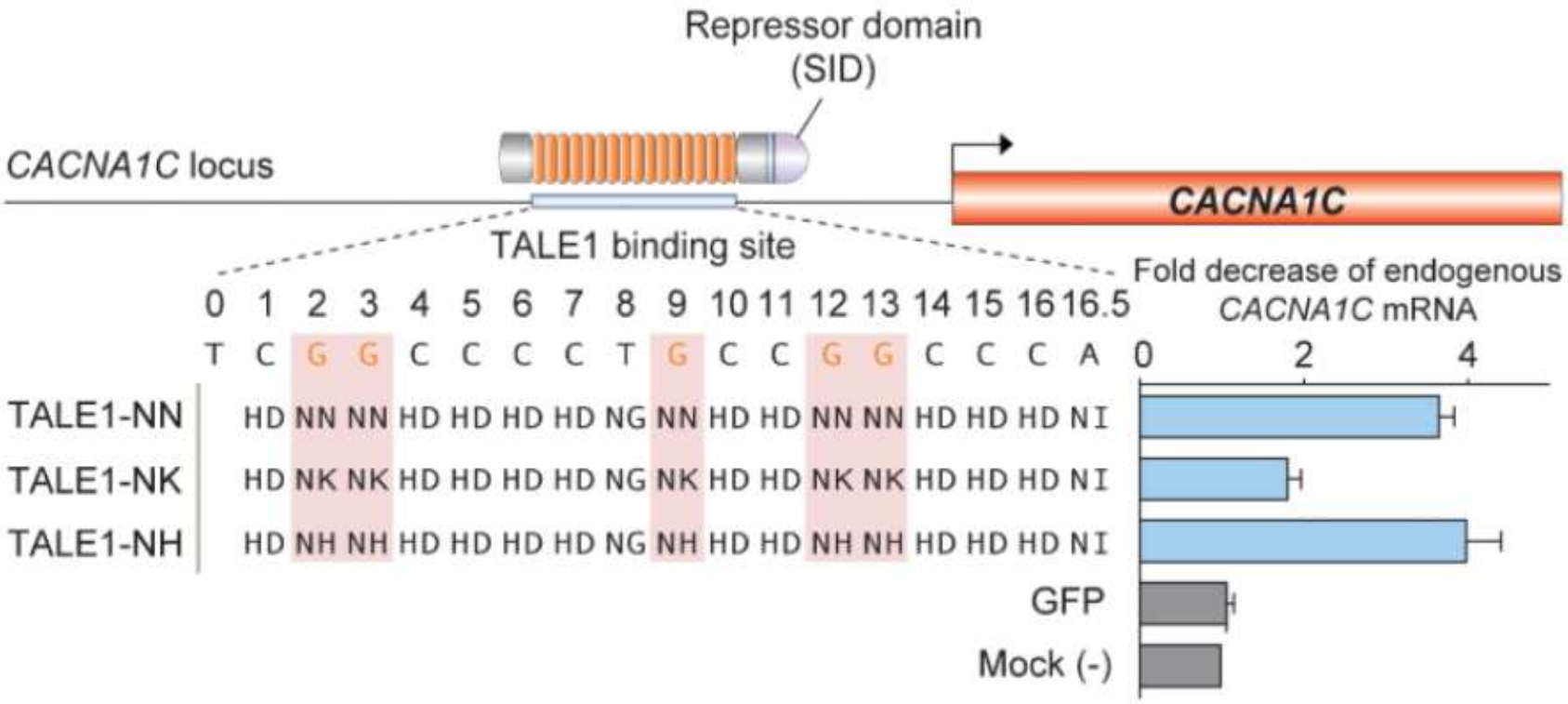


Development of a TALE transcriptional repressor

Candidate domains	Organismic origins
PIE-1	*C. elegans*
Ubx-QA	*D. melanogaster*
IAA28-RD	*A. thaliana*
SID	*H. sapiens*
Tbx3-RD	*H. sapiens*
KRAB	*H. sapiens*



Development of a TALE transcriptional repressor



Conclusions

- TALEs can be easily customized to recognize specific sequences on the endogenous genome, and can be used to target a diverse range of effector domains to specific genomic loci.

Conclusions

- TALEs can be easily customized to recognize specific sequences on the endogenous genome, and can be used to target a diverse range of effector domains to specific genomic loci.

- TALE-RVD screening results identified a set of novel RVDs with useful activity and specificity. Among them, the RVD NH demonstrated increased specificity for recognizing guanine.

Conclusions

- TALEs can be easily customized to recognize specific sequences on the endogenous genome, and can be used to target a diverse range of effector domains to specific genomic loci.

- TALE-RVD screening results identified a set of novel RVDs with useful activity and specificity. Among them, the RVD NH demonstrated increased specificity for recognizing guanine.

- Screening of six different repressor domains across a range of host species resulted in the identification of two repressor domains capable of repressing mammalian transcription. The SID domain achieved the strongest level of repression at endogenous genomic loci.

ZFNs or TALEs protein?

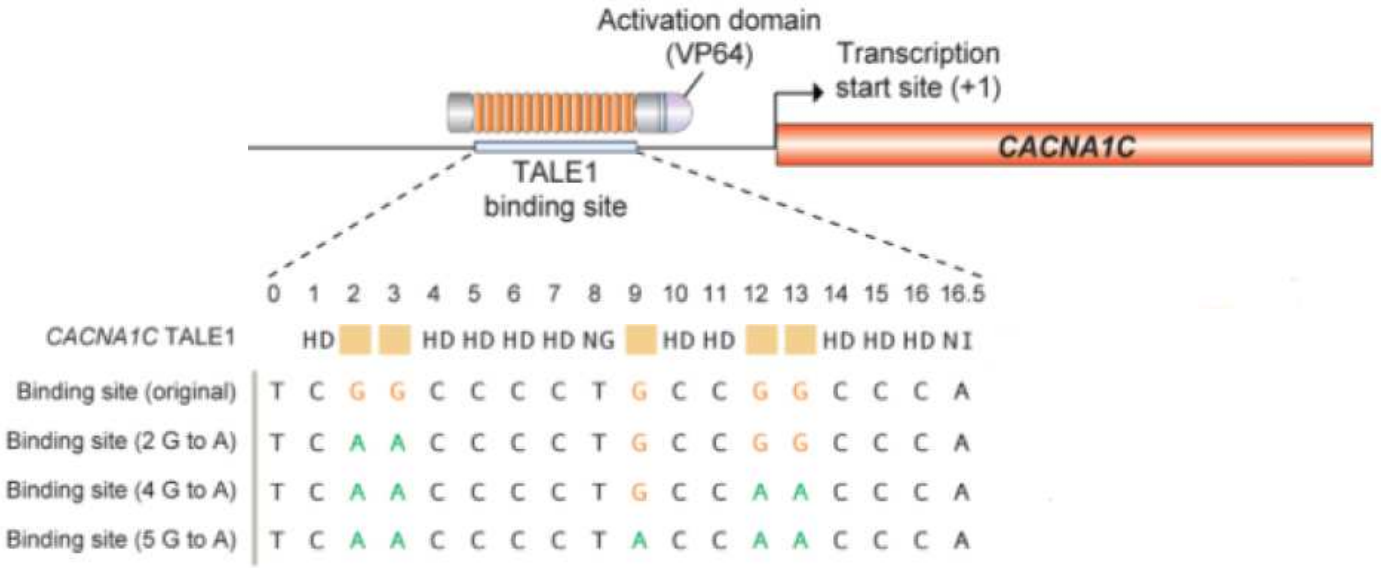
ZFPs		TALEs	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Small size: efficient delivery into the cells	Many off-targets	High specificity to the target	Sensitive to DNA methylation of the targeted region
	Time consuming		
Engineered from human proteins might be less susceptible to adverse immunoreactions		Successfully used in combination with catalytic domain of multiple enzymes	Engineered from bacterial backbones might elicit immunoresponses
			Big size might complicate delivery
Successfully used in combination with catalytic domain of multiple enzymes			

ZFN or TALEs protein?

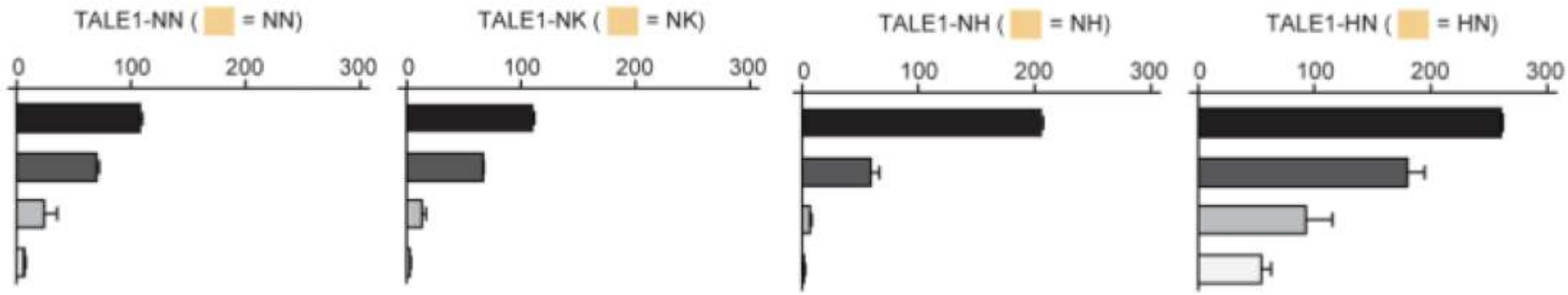
ZFPs		TALEs	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Small size: efficient delivery into the cells	Many off-targets	High specificity to the target	Sensitive to DNA methylation of the targeted region
	Time consuming		
Engineered from human proteins might be less susceptible to adverse immunoreactions		Successfully used in combination with catalytic domain of multiple enzymes	Engineered from bacterial backbones might elicit immunoresponses
			Big size might complicate delivery
Successfully used in combination with catalytic domain of multiple enzymes			

Obviously, this is not the end of the story….

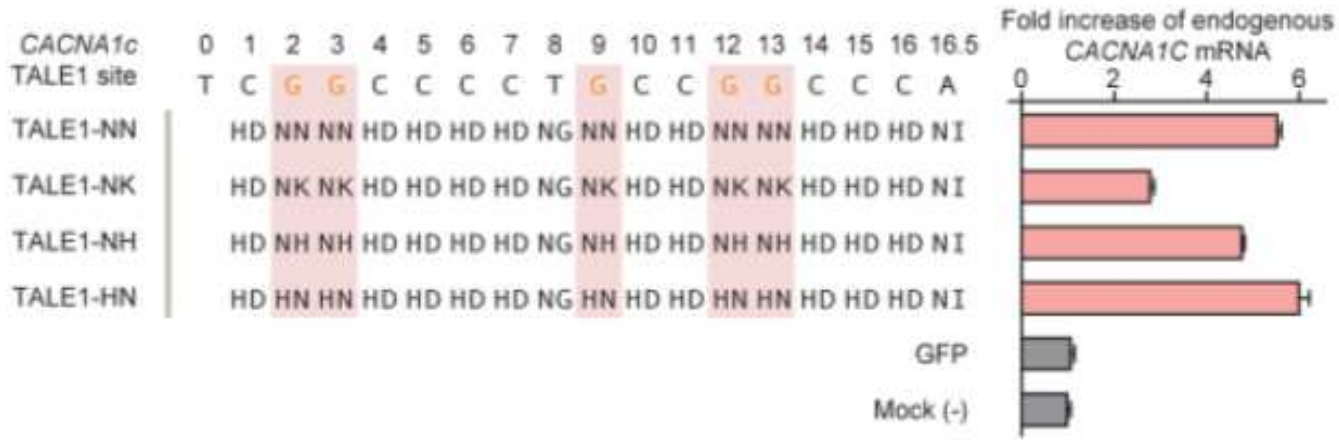
Characterization of guanine-specific repeat-variable diresidues (RVDs)



Normalized reporter activation (RLUs) by:



Evaluation of guanine-binding RVDs at endogenous genome loci



Using RT-qPCR, the performance of NN, NK, NH, and HN for targeting endogenous genomic sequences was compared. In particular was tested the ability of NN-, NK-, NH-, and HN-TALEs to activate *CACNA1C* transcription by targeting an endogenous target site.