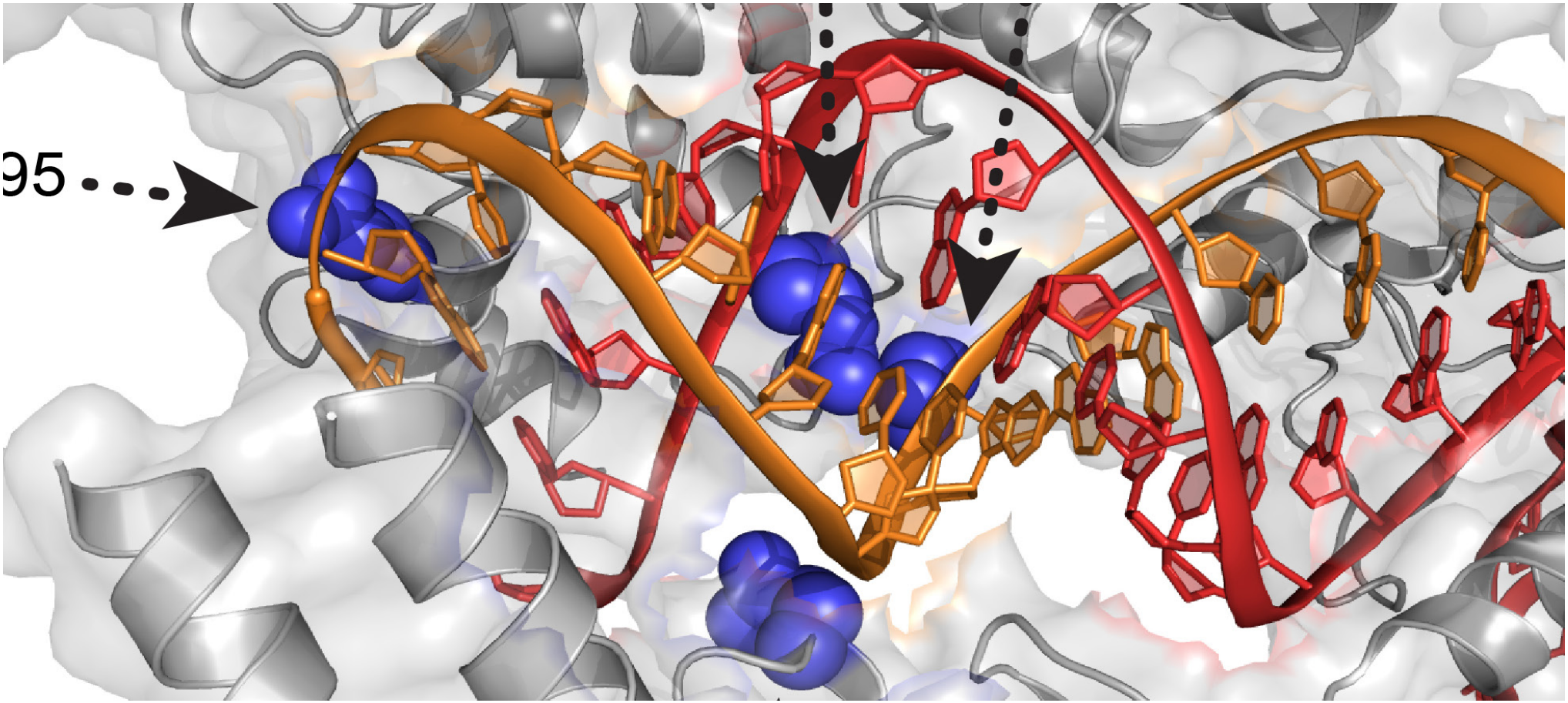
Enzyme engineering

TJC 8.3.16 Valeria Eckhardt



Overview

Methods

- Rational design
- Evolutionary methods

Applications

- Specific Cas9: SpCas9-HF1

Discussion

- Enzyme engineering for Gene-Directed Enzyme Pro-drug Therapy

Rational design

Requirement:

Well known structure and mechanism of the enzyme of interest

NMR/ X-ray christ. Peptidomimetics

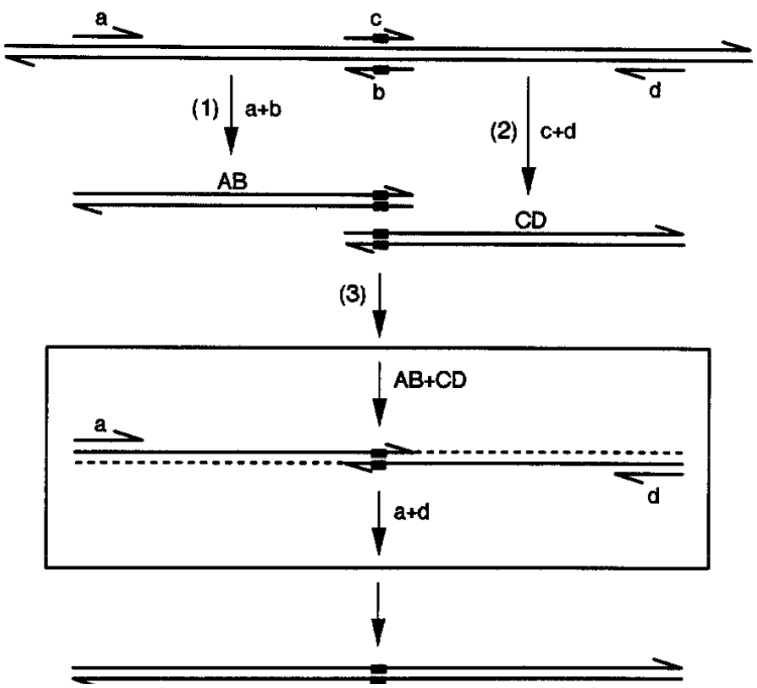
Homology modeling

Methods:

Site-directed mutagenesis for introduction of specific amino acids into a target gene

- Overlap extension
- Whole plasmid single round PCR

Overlap extension



Ho et al., 1988

1. PCR

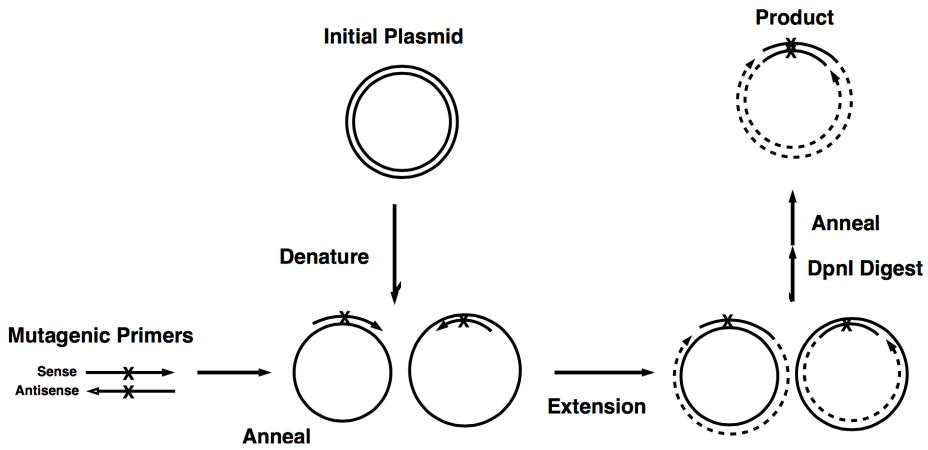
4 primers, 2 PCRs take place

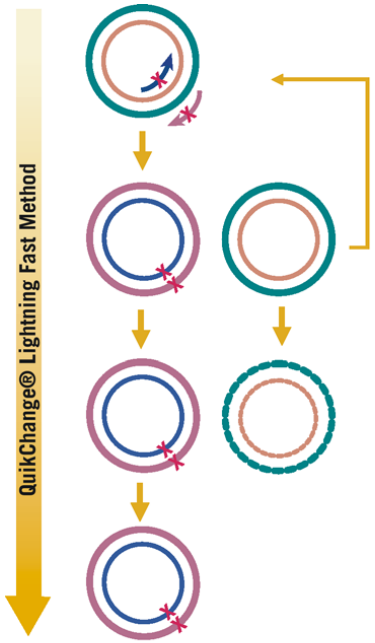
2 double-stranded fragments,
containing the desired mutagenic codon

2. PCR

non-mutated primer set amplifies mutagenic DNA

Whole plasmid single round PCR





1. **Mutant Strand Synthesis**
   Perform thermal cycling to:
   - Denature DNA template
   - Anneal mutagenic primers containing desired mutation
   - Extend and incorporate primers with our exclusive *Pfu*-based DNA Polymerase Blend
   - **Total reaction time: 1 hour***

2. **Faster *Dpn* I Digestion of Template**
   - Digest parental methylated and hemimethylated DNA with NEW *Dpn* I enzyme
   - **Total reaction time: 5 minutes**

3. **Transformation**
   - Transform mutated molecules into competent cells for nick repair
   - **Total reaction time: 1.5 hours**

Evolutionary methods

Evolution is based on

1. Variation (library, the larger the better)
2. Increase in fitness by variation
3. Inheritability

Requirement:

availability of a suitable selection scheme (best high throughput) that favors the desired protein properties

Directed evolution

1. Generate variability

Genetic environment stable, only modify GOF

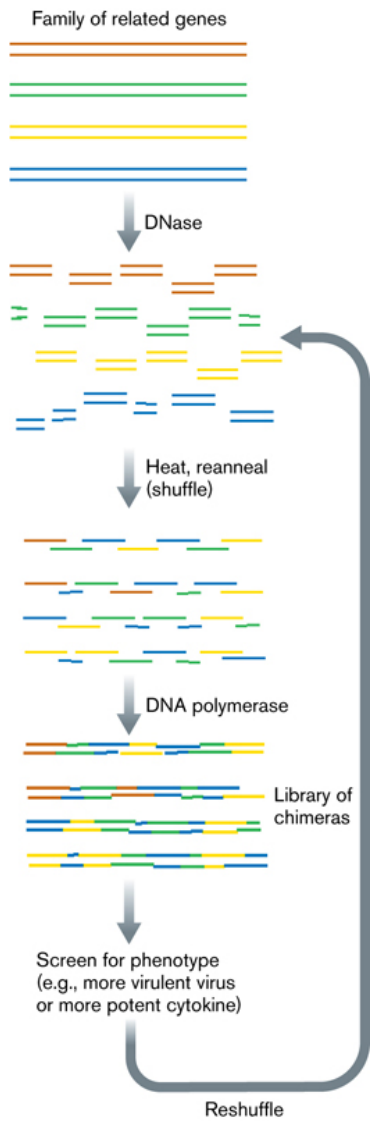
- Random point mutations i.e. by error prone PCR (more G,C)
- Indels: by transposons
- Mimick recombination: **DNA shuffling**
- Systematically randomize by saturation mutagenesis: replace a single amino acid within a protein with each of the natural amino acids (structure partially known)

Directed evolution

Mimick recombination: DNA shuffling

With restriction enzymes:

1. Digest family of related genes
2. DNA ligase to join fragments
3. Chimeric genes



With DNAse1

Annealing where homology

Series of chimeric genes with shuffled components

Directed evolution

2. Increase in fitness by variation

In vivo evolution

Transform each cell with one plasmid containing a different member of the library.

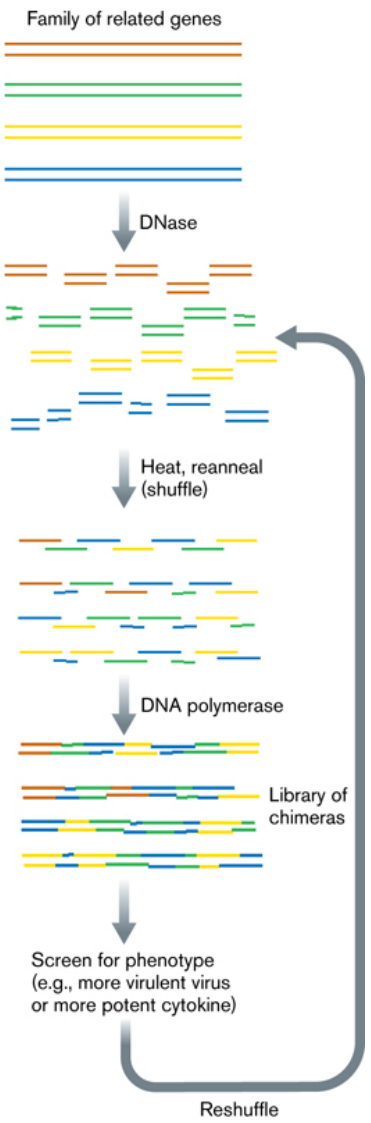
1 PCR product per cell

Selection systems

Protein function directly coupled to survival of the gene

Selection for binding activity

- FACS
- Phage display



Series of chimeric genes with shuffled components

Library is cloned into host cell

Selection for enhanced new gene function

FACS

- Cells present a scissile bond (i.e. Arg- Val linkage), which can be cleaved by a surface-displayed enzyme or not.
- The scissile bond links a fluorophore and a quencher (diminishes fluorophore intensity).

- Scissile bond not cleaved: No fluorescence emission
  The fluorophore emission is quenched by the quenching fluorophore.
- Scissile bond cleaved: fluorescence occurs
  Fluorophore and quenching fluorophore are then separated

(Antikainen & Martin, 2005).

Phage display

Library is subcloned into a bacteriophage vector coding for a phage coat protein.

1 PCR product per phage vector

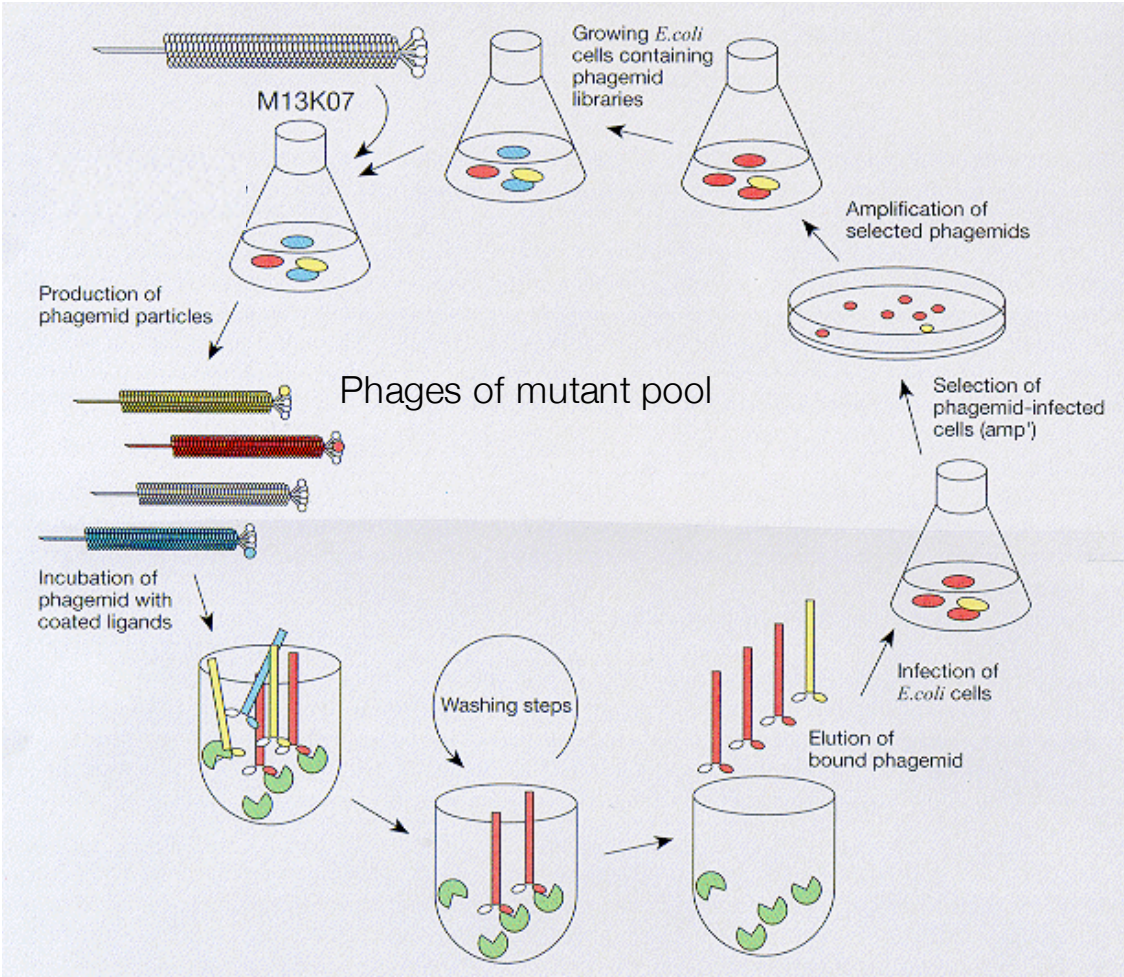
Transduce into E.coli.

Each phage of the mutant pool expresses a different enzyme displayed on the coat protein of the phage surface.

Those that remain can be eluted, used to produce more phage to produce a phage mixture that is enriched with relevant (i.e. binding) phage.

The repeated cycling of these steps is referred to as 'panning'.

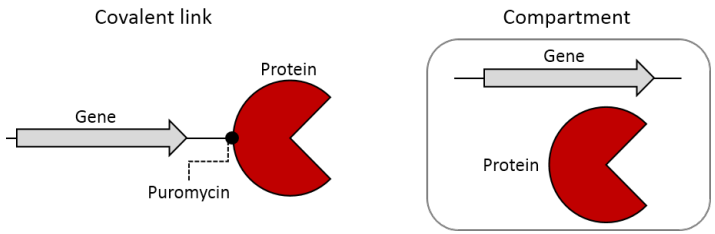
Phage eluted in the final step can be used to infect a suitable bacterial host, from which the phagemids can be collected and the relevant DNA sequence excised and sequenced to identify the relevant, interacting proteins or protein fragments.



Evolutionary methods

Evolution is based on

1. Variation (library, the larger the better)
2. Increase in fitness by variation
3. Inheritability



Genotype-phenotype link: isolate DNA and functional enzyme

- mRNA display where the mRNA gene is linked to the protein at the end of translation by puromycin
- protein and its gene can be co-localised by compartmentalisation in living cells

Applications of enzyme engineering

- Food industry with a variety of food-processing enzymes
  - Lactase: production of low-lactose milk
  - Protease: meat tenderization
  - Pectinase: clarification of wine

- Industry: In Petroleum biorefining new biocatalysts are required

- Environmental applications of enzyme engineering: design microorganisms to eliminate environmental pollutants

Specific Cas9: SpCas9-HF1

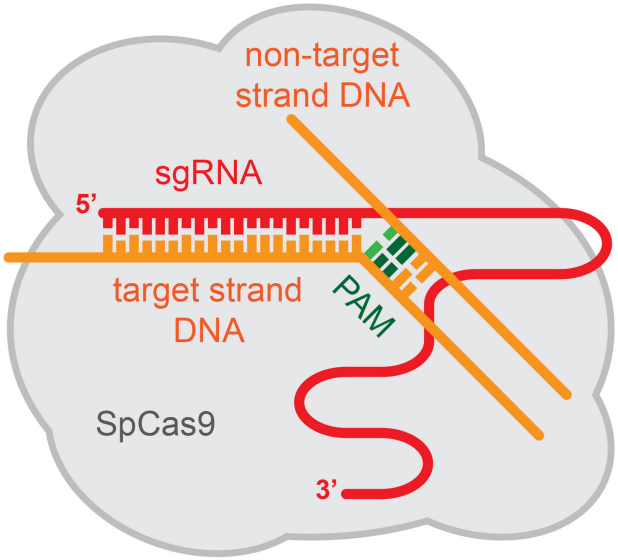
High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects

Benjamin P. Kleinstiver[1,2]*, Vikram Pattanayak[1,2]*, Michelle S. Prew[1], Shengdar Q. Tsai[1,2], Nhu T. Nguyen[1], Zongli Zheng[3] & J. Keith Joung[1,2]

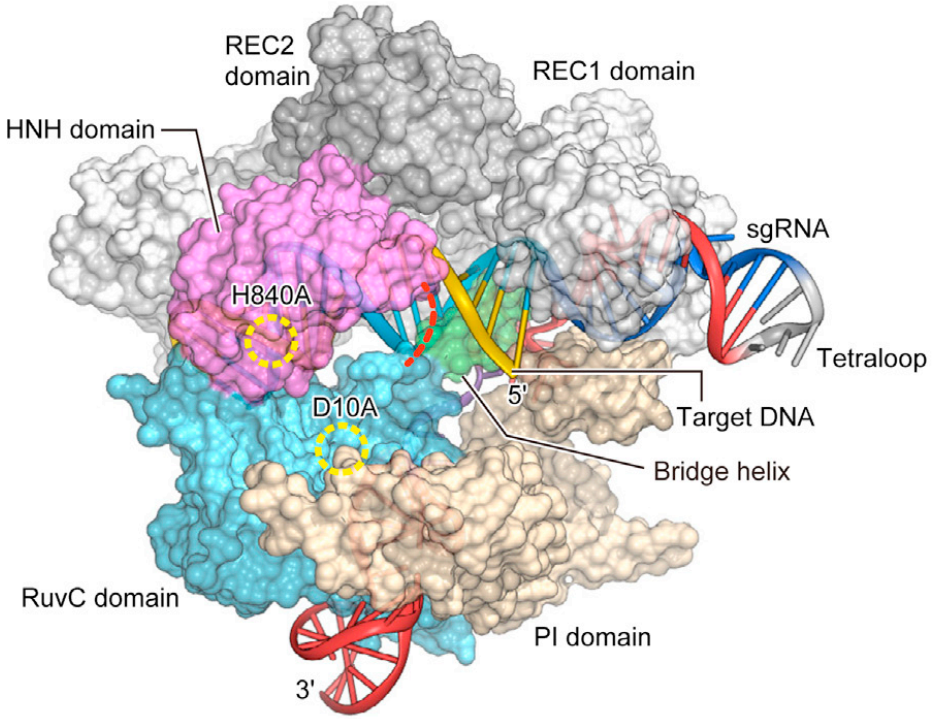
Approach:
Rational design of a more specific SpCas9 enzyme
Requirement:
Well known structure and mechanism of the enzyme of interest

Specific Cas9: SpCas9-HF1

Structure and mechanism of the SpCas9



Kleinstiver et al., 2016



Nishimasu et al. 2014
Anders et al., 2014

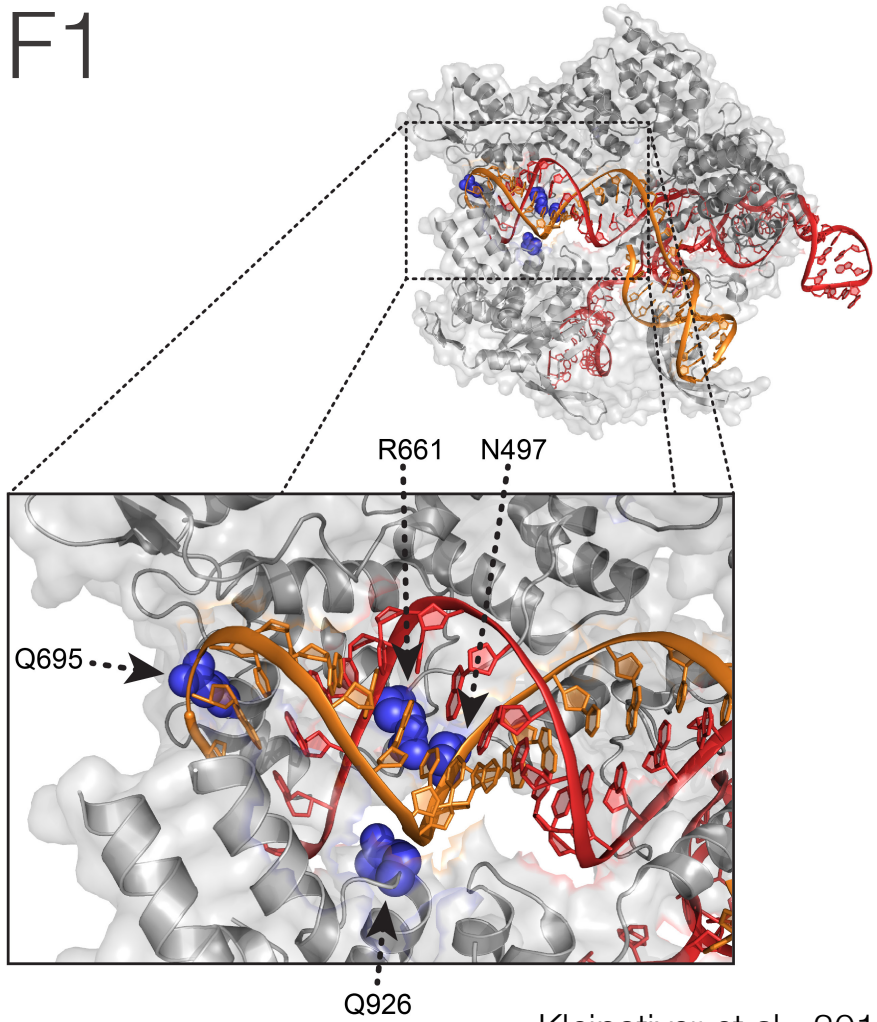
Specific Cas9: SpCas9-HF1

Excess energy hypothesis

SpCas9–sgRNA complex with more E than needed for optimal recognition of target DNA

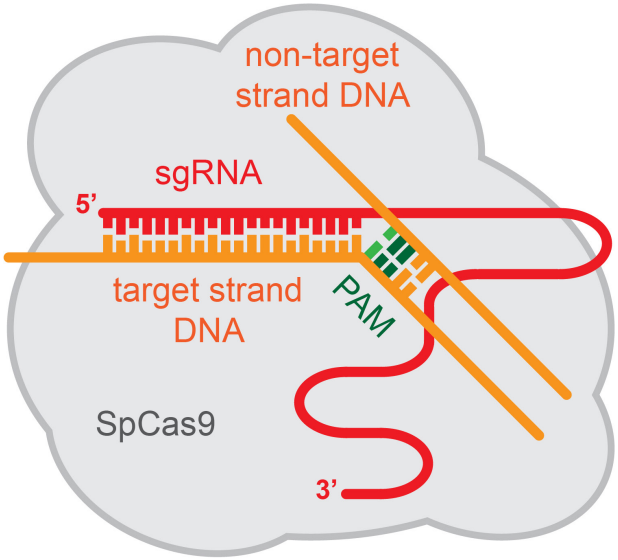
enables cleavage of mismatched off-target sites

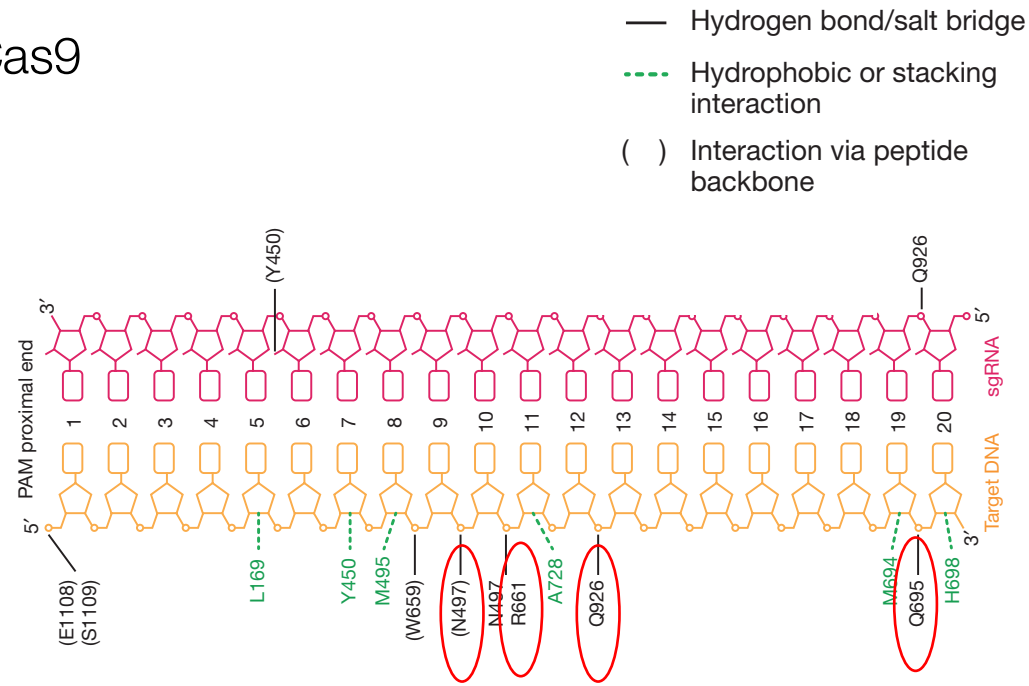
Direct hydrogen bonds made by four SpCas9 residues to the phosphate backbone of the target DNA strand.



Kleinstiver et al., 2016

Structure and mechanism of the SpCas9

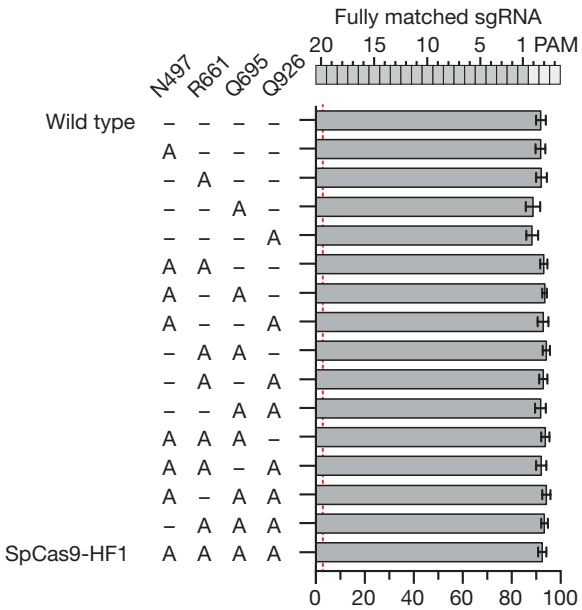




Direct hydrogen bonds made by four SpCas9 residues to the phosphate backbone of the target DNA strand.

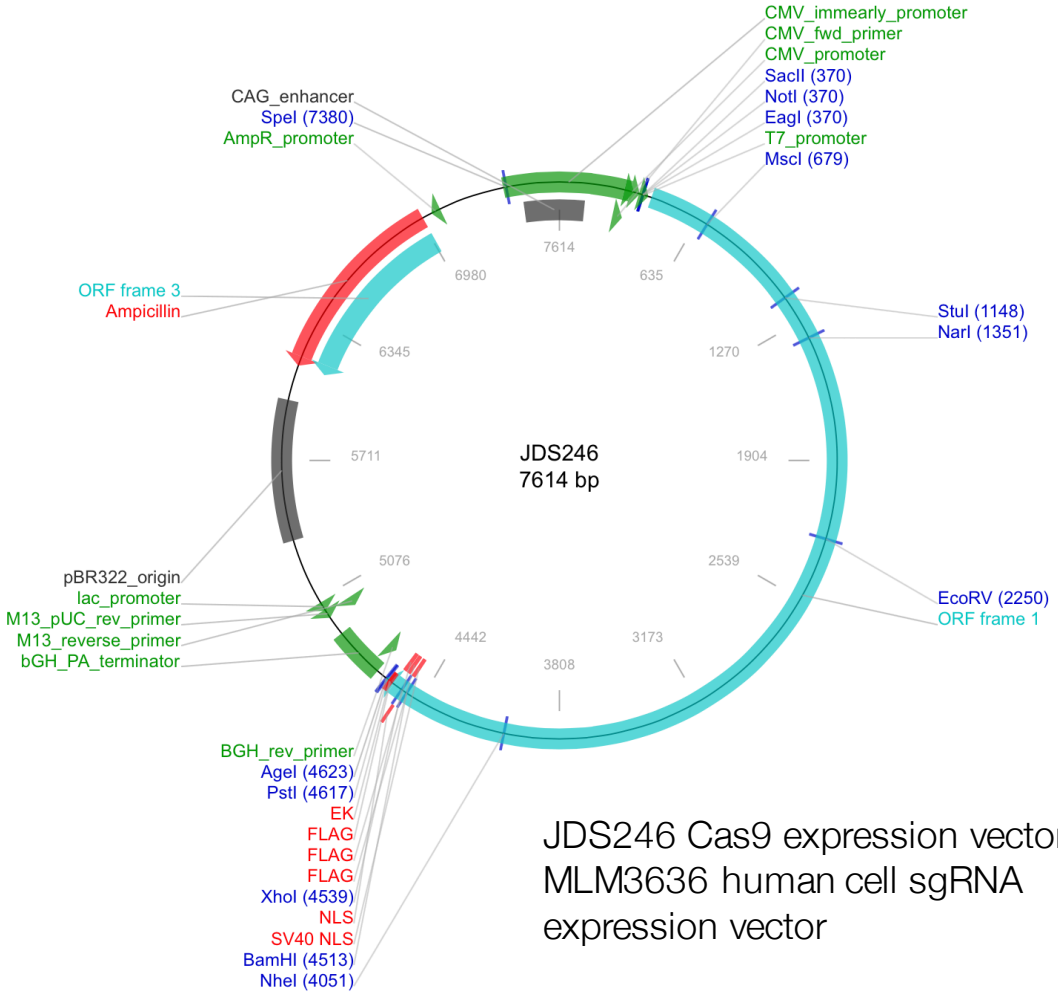
Kleinstiver et al., 2016

Alteration of SpCas9 DNA contacts



SpCas9 expression plasmids containing amino acid substitutions generated by standard PCR and molecular cloning into JDS246 (addgene)

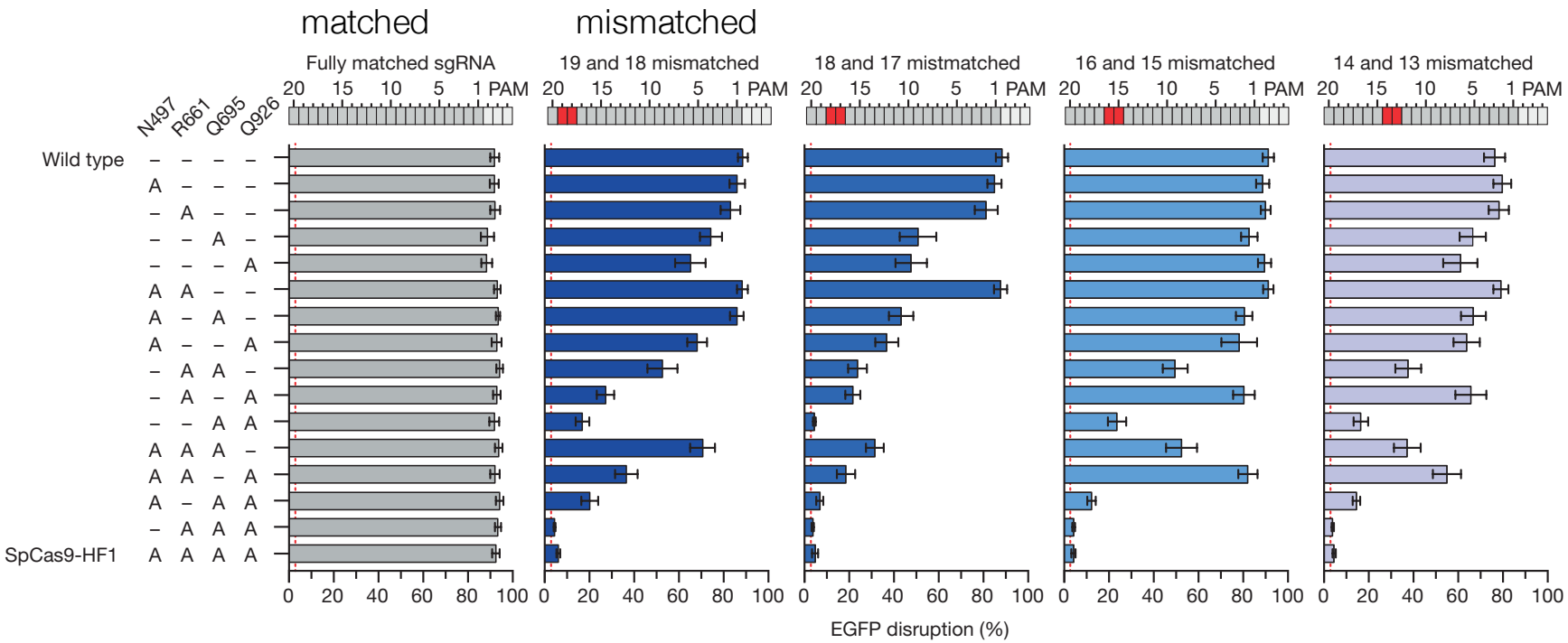
JDS246 Cas9 expression vector
MLM3636 human cell sgRNA expression vector



JDS246 Cas9 expression vector
MLM3636 human cell sgRNA expression vector

Alteration of SpCas9 DNA contacts

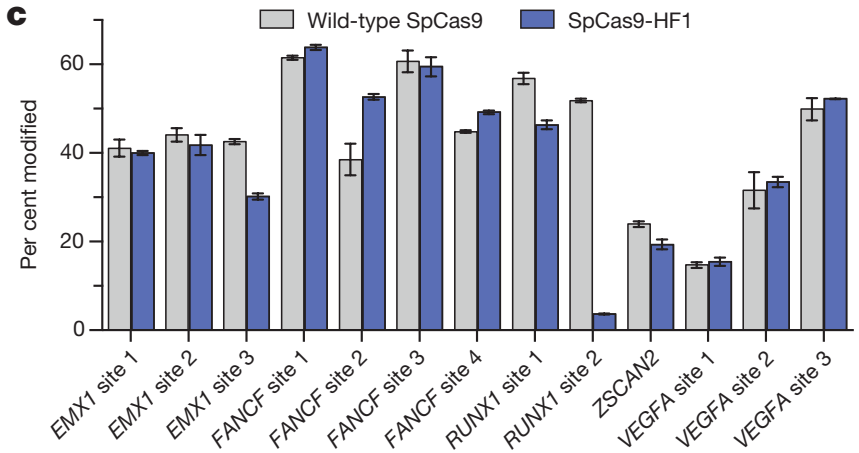
EGFP disruption assay: U2OS.EGFP cells + EGFP-targeted sgRNA
52h post transfection FACS for loss of EGFP expression
Background EGFP loss was determined using negative control transfections

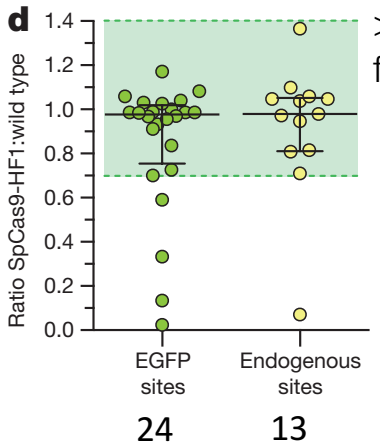


Alanine substitution of one or all of these residues did not reduce on-target cleavage efficiency of SpCas9 with this EGFP-targeted sgRNA.
One of the triply substituted variants and the quadruple substitution variant showed minimal EGFP disruption

SpCas9-HF1 retains high on-target activities

T7 endonuclease I assays: detect indels
Amplify ~100 ng DNA of on-target sites with Phusion Hot-Start Flex DNA Polymerase.
Kit to purify PCR product
Denaturation and annealing of ~200 ng of the PCR product
Digestion with T7 endonuclease I (recognizes and cleaves non-perfectly matched DNA = indel)
Purify digestion products
Quantify using a QIAxcel capillary electrophoresis instrument (Qiagen, replace traditional, labor-intensive gel analysis)
Quantify frequencies induced by Cas9-sgRNA complexes.



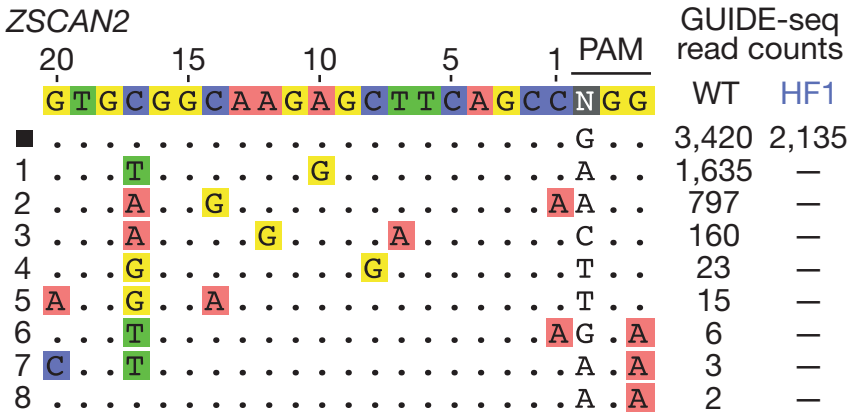


SpCas9-HF1 exhibited > 70% of the on-target activities for 86% (32/37) of the sgRNAs.

Genome-wide specificity of SpCas9-HF1

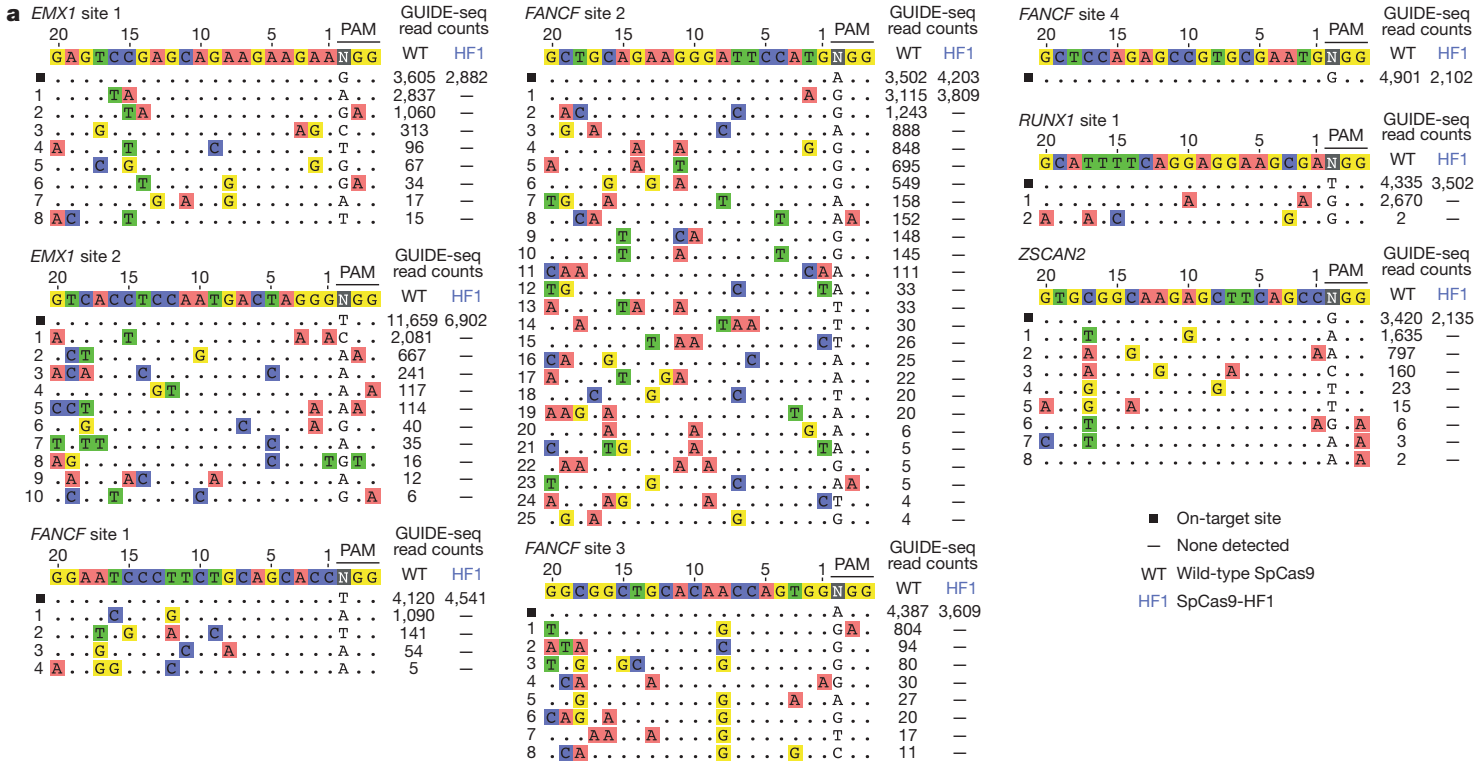
GUIDE-seq:
Transfect U2OS cells with Cas9 and sgRNA plasmids and GUIDE-seq dsODN

GUIDE-seq relies on the integration of a short dsODN tag into DNA breaks
Enables amplification and sequencing of adjacent genomic sequence
The number of tag integrations at any given site provides a quantitative measure of cleavage efficiency.
Tag-specific amplification and library preparation and before high-throughput sequencing on an Illumina MiSeq



Positional profiles for potential genomic off-target sites were restricted to sequences containing five or fewer mutations relative to the on-target site and to sequences containing NGG PAMs. 2 to 25 per sgRNA

Genome-wide specificity of SpCas9-HF1



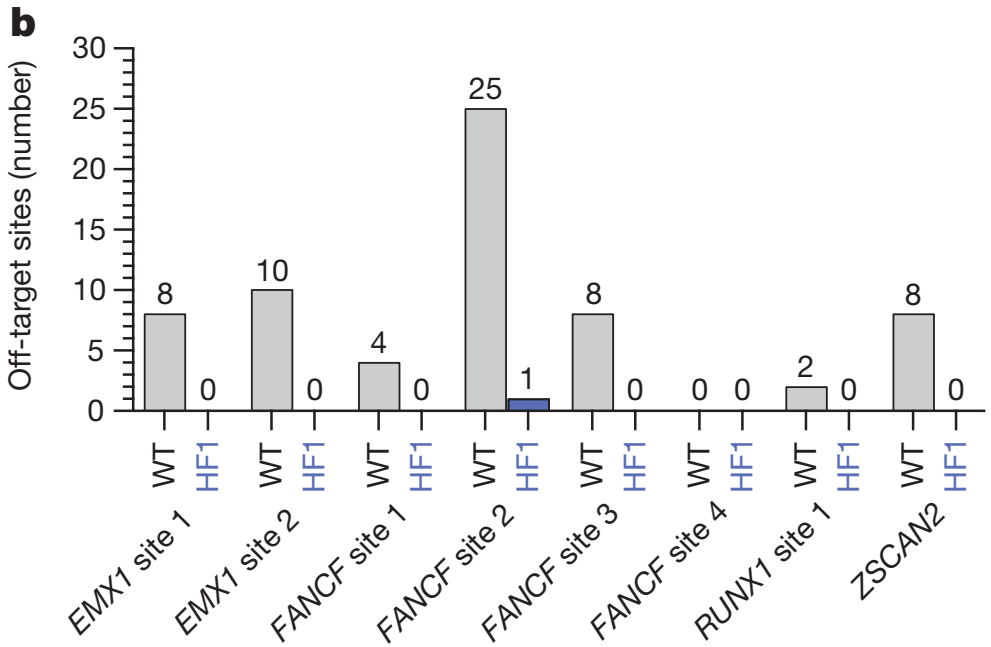
wt Cas9 no off-target

2 – 25 off target sites (up to 5bp mismatch)

wtCas9: 7of the 8sgRNAs induced cleavage at multiple off-target sites

SpCas9-HF1, shows complete absence of GUIDE-seq detectable off-target events, despite one

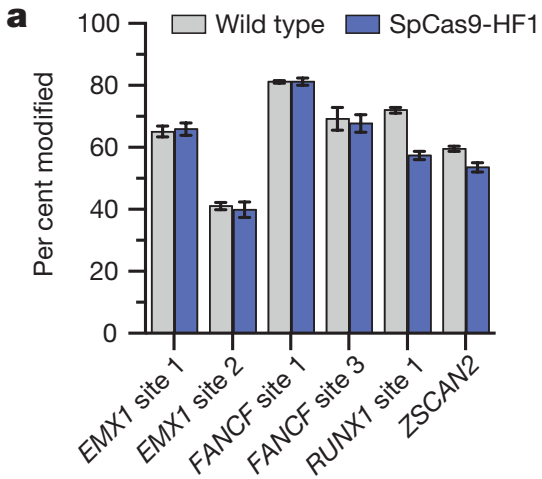
Genome-wide specificity of SpCas9-HF1



2 – 25 off target sites (up to 5bp mismatch)
wtCas9: 7of the 8sgRNAs induced cleavage at multiple off-target sites
SpCas9-HF1, shows complete absence of GUIDE-seq detectable off-target events, despite one

Genome-wide specificity of SpCas9-HF1

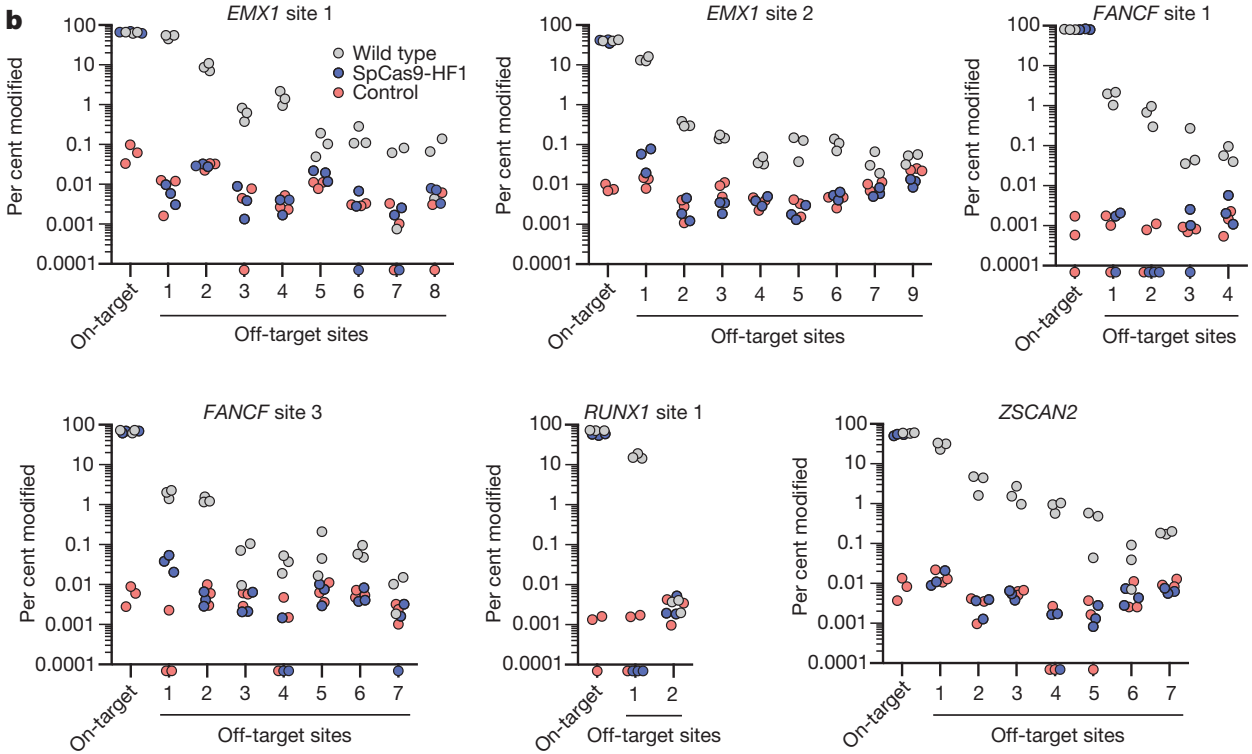
Confirm GUIDE-seq:
Transfect human cells only with sgRNA- and Cas9- encoding plasmids (without the GUIDE-seq tag).
Targeted amplicon sequencing = deep sequencing of off-target sites identified by GUIDE-seq
More directly measure frequencies of indel mutations induced by wtSpCas9 and SpCas9-HF1.



Comparable frequencies of indel mutations
induced by wtSpCas9 and SpCas9-HF1
sgRNA and Cas9 work

Genome-wide specificity of SpCas9-HF1

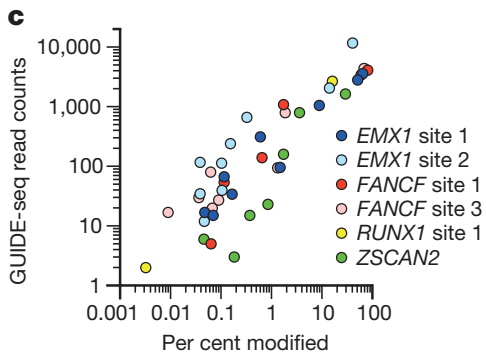
Confirm GUIDE-seq with Targeted amplicon sequencing



Wt SpCas9 shows indel mutations at 35 of the 36 off-target sites

SpCas9-HF1 = negative control

Correlation

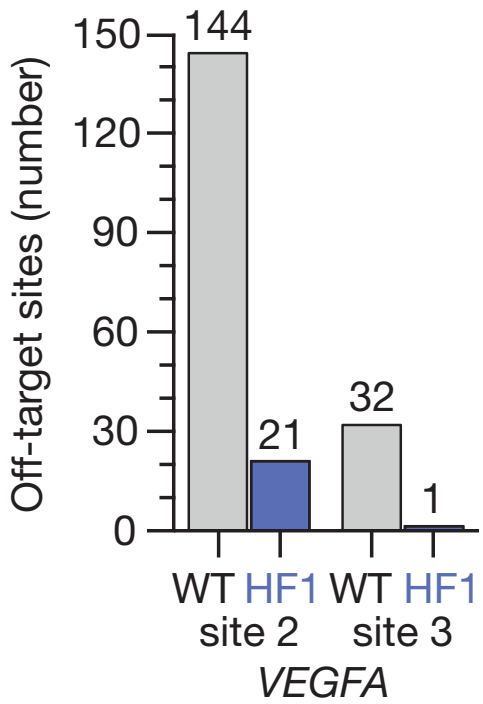


negative control experiments: Cas9 plasmids were co-transfected with a U6-null plasmid

Genome-wide specificity of SpCas9-HF1

Can SpCas9-HF1 reduce genome-wide off-target effects of sgRNAs designed against atypical repetitive sequences with many off-target sites?

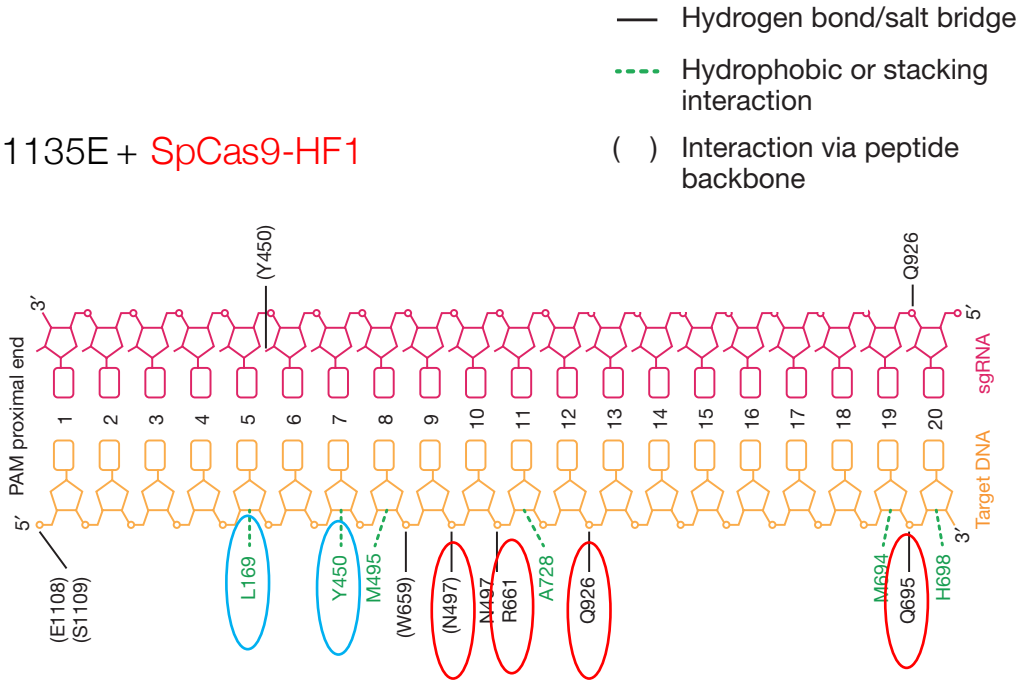
GUIDE-seq experiments:
SpCas9-HF1 is highly effective at reducing off-target sites of these sgRNAs (without impairment in on-target activity)



Refining the specificity of SpCas9-HF1

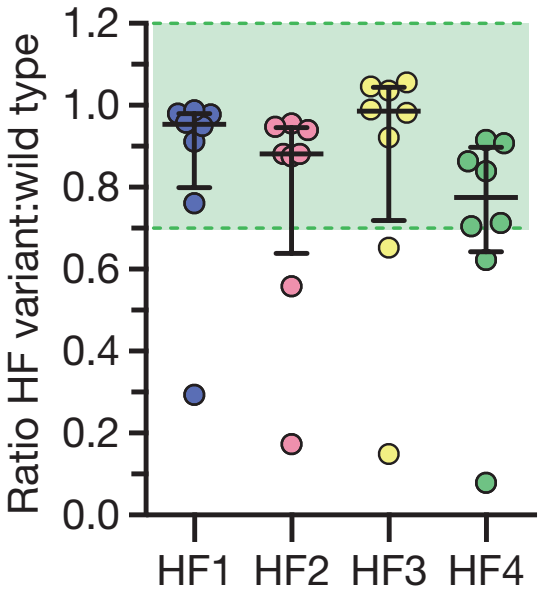
SpCas9-HF1
SpCas9-HF2: Aspartate – Glutamate substitution D1135E + SpCas9-HF1
SpCas9-HF3: Leucine L169A + SpCas9-HF1
SpCas9-HF4: Tyrosine Y450A + SpCas9-HF1

Base stacking interaction, 120° shift
Hydrophobic interaction



Direct hydrogen bonds made by four SpCas9 residues to the phosphate backbone of the target DNA strand.

Refining the specificity of SpCas9-HF1



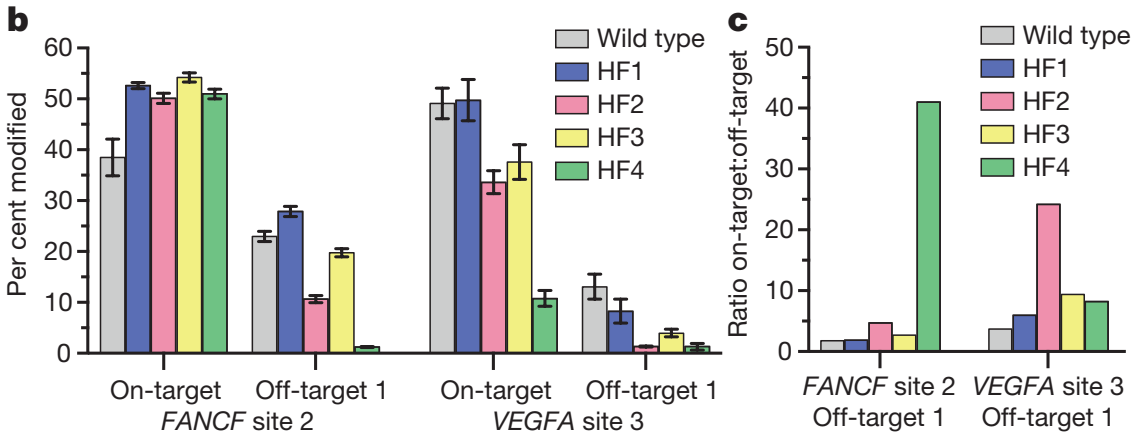
SpCas9-HF2-4 retained >70% of the activities observed with wild-type SpCas9 with the same six out of eight EGFP-targeted sgRNAs

Refining the specificity of SpCas9-HF1

Can SpCas9-HF2-4 reduce indel frequencies at two off-target sites that remained susceptible to modification by SpCas9-HF1?

Check by T7 endonuclease I assay

Specificity amongst variants



Better on-target activity
SpCas9-HF-4 and 2 reduced indel mutation frequencies to near back- ground level
SpCas9-HF-4 and 2 highest specificity

Discussion of of SpCas9-HF1-4

SpCas9-HF1 reduces all or nearly all genome-wide off-target effects to undetectable levels
Tested by GUIDE-seq and targeted next-generation sequencing

Most robust effects obtained with sgRNAs designed against standard, non-repetitive target sequences

Introducing substitutions at other non-specific DNA contacting residues can further reduce some of the very small number of residual off-target sites

Further biochemical experiments and structural characterization to define the mechanism of SpCas9-HF1 specificity

Mechanism of target recognition: active cleavage complex through conformational changes or kinetics that might be affected by the substitutions in our SpCas9-HF1 variant

Refine specificity by substitutions in residues that contact the non-target DNA strand

Extended enhancement of specificity to other naturally occurring and engineered Cas9 orthologues

Discussion Enzyme engineering

Rational design

+ opting straight forward for improvements
- Structure and mechanism need to be known

Directed evolution

+ no need to understand structure and mechanism
- high-throughput assay is required
- often highly specific for particular activity and not applicable to other DE experiments

Semirational approaches to combine both

Huge variability of applications in food industry, environmental applications, industry, medicine

With improving technologies (higher resolution of enzyme structures, high throughput) more efficient enzyme engineering possible

Possible application of enzyme engineering?

Gene-Directed Enzyme Prodrug Therapy

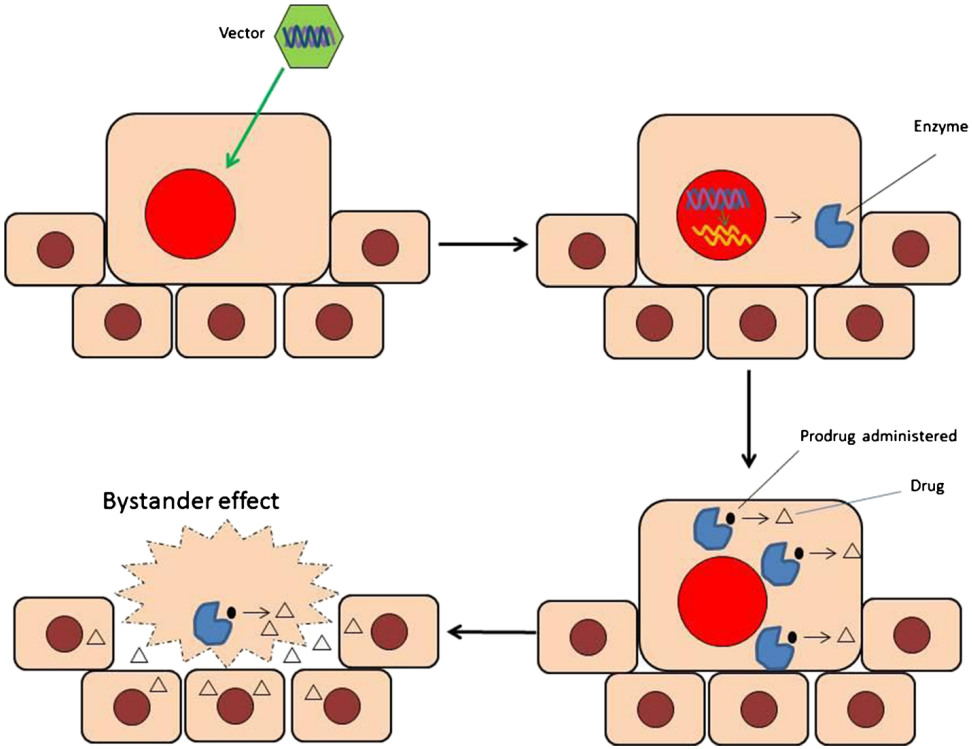
Jin Zhang,[1,4] Vijay Kale,[2] and Mingnan Chen[3,4]

Received 25 June 2014; accepted 22 September 2014; published online 23 October 2014

GDEPT utilizes transgenes which encode enzymes that can convert pro-drugs into active therapeutic metabolites.

Three key features in cancer therapy
1. cell- specific gene delivery and expression
2. Controlled conversion of pro-drugs to drugs in target cells
3. Expanded toxicity to the target cells neighbors (passive diffusion)

Gene-Directed Enzyme Pro-drug Therapy



1. The enzyme-coding gene is cloned into a vector and transfected into a tumor cell.

2. The gene is transcribed into an mRNA which later is translated into the enzyme inside the tumor cell.

3. A pro-drug is administered systemically and absorbed by the same cell

4. the pro-drug can then be converted to a cytotoxic drug by the enzyme inside the cell.

Enzyme-gene expression controlled by tumor cell-specific promoters, enzymatic reaction precisely in tumor cells

Gene-Directed Enzyme Pro-drug Therapy

Herpes Simplex virus system

Enzyme	Prodrug	Drug	Mechanism
Thymidine kinase	Ganciclovir (GCV)	Ganciclovir triphosphate	Inhibit DNA polymerase, GT incorporate into replicating cell DNA leading to replication failure and cell death.

Ganciclovir = nucleoside substitute, can be phosphorylated by Thymidine kinase
Incorporated into replicating DNA
No additional nucleosides can bind
Replication terminated

Enzyme should have high catalytic activity, so that tumor cells can convert pro-drugs even at low substrate concentration

Structure-guided Engineering of Human Thymidine Kinase as PET reporter gene by Cambpell et al., 2012

Enzyme engineering

Questions?