

Hands on CRISPR

Author: Fábio Ferreira

Hands on CRISPR

Let's do some CRISPR!!!



EnGen sgRNA Synthesis Kit

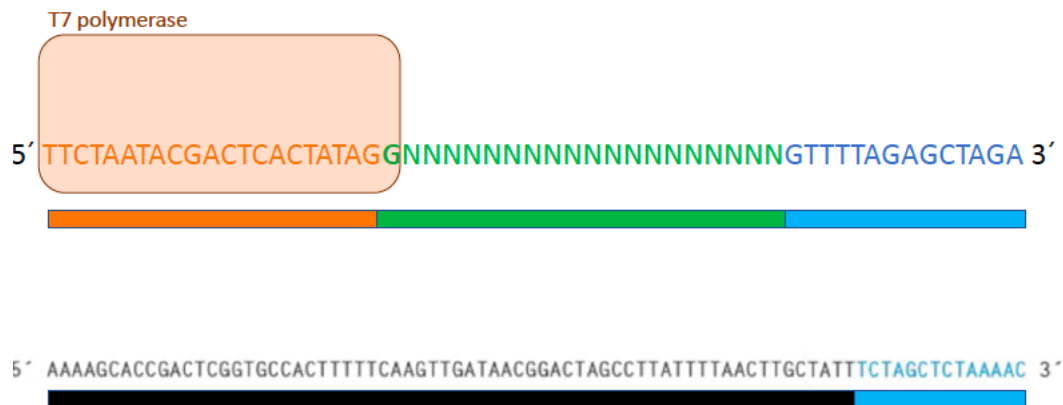
Hands on CRISPR

EnGen sgRNA Synthesis Kit

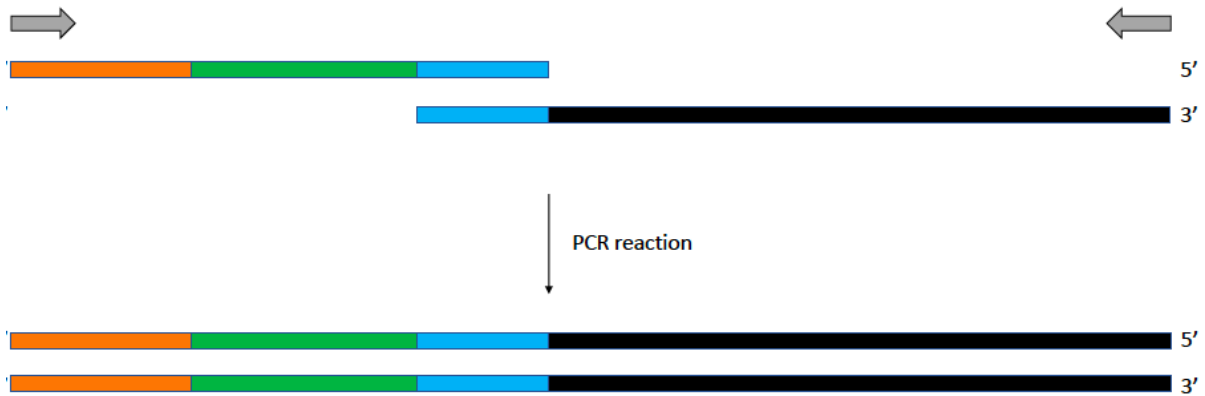
1. Select 20 nucleotide target sequence (not including the PAM (NGG) sequence). Use of a target DNA selection program is recommended.
2. Check input sequence for presence of "G" at the 5' end. If there are no "G's" at the 5' end, add one "G".
3. To the 5' end; append T7 promoter sequence: **TTCTAATACGACTCACTATA**
4. To the 3' end; append 14 nucleotide overlap sequence: **GTTTTAGAGCTAGA**
5. Check complete oligo sequence:

5' **TTCTAATACGACTCACTATAG**(N)₂₀**GTTTTAGAGCTAGA** 3'

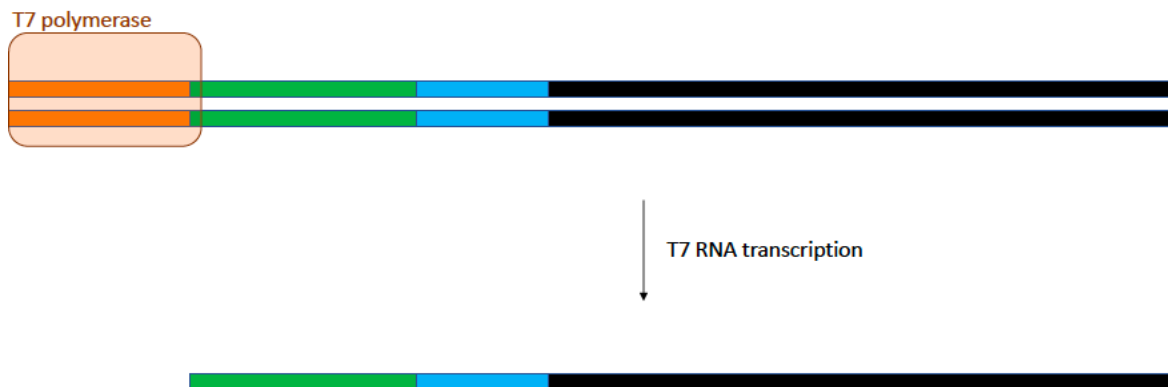
Hands on CRISPR



Hands on CRISPR



Hands on CRISPR



Some online tools

- Zhang lab (overview): <https://zlab.bio/guide-design-resources>
- Benchling: <https://www.benchling.com/crispr/> (favorite)
- CRISPick: <https://portals.broadinstitute.org/gppx/crispick>
- CHOPCHOP: <https://chopchop.cbu.uib.no/>
- CRISPRscan: <https://www.crisprscan.org/> (many species)
- CRISPy: http://staff.biosustain.dtu.dk/laeb/crispy_scoeli/ (*S. coelicolor*)



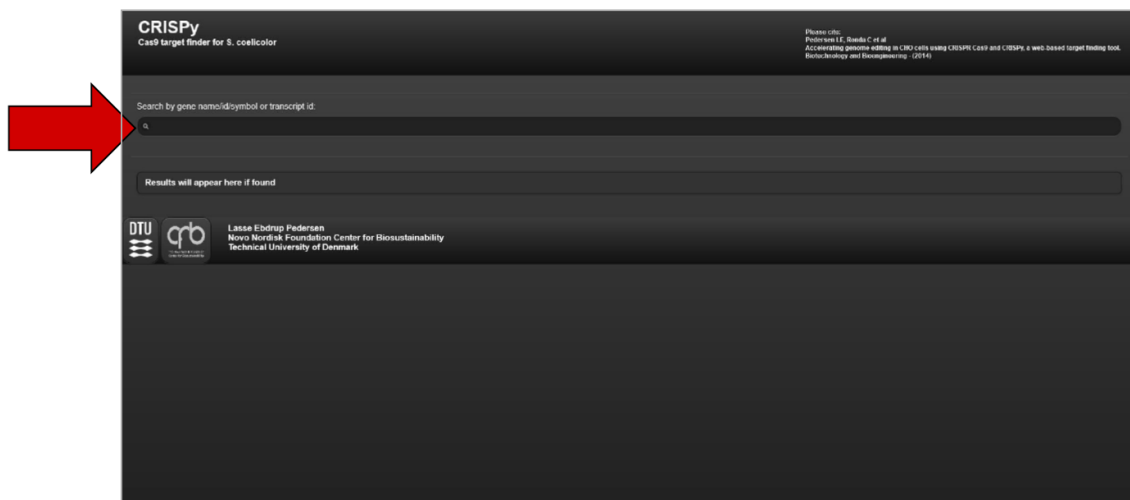
Design sgRNA: Tools and parameters

EXAMPLE 1- *S. coelicolor*

To find *S. coelicolor* target sequence use CRISPy - a specific Cas9 target finder for *S. coelicolor*:

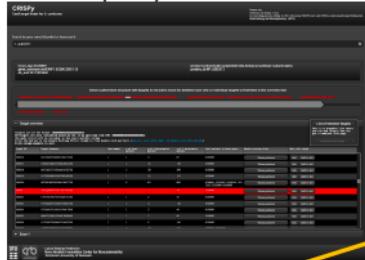
- a. actIORF1 (ACT1)
- b. actVB (ACT5)

CRISPy: http://staff.biosustain.dtu.dk/laeb/crispy_scoeli/ (*S. coelicolor*)

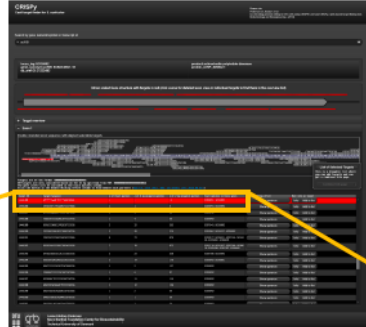


Design sgRNA: Tools and parameters

act1ORF1 (ACT1)



actVB (ACT5)



Target ID	Target Sequence	Exon number	# of Exact matches	# of 1 bp mismatch matches	# of 2 bp mismatch matches	Exact matches in these genes
1446187	GTCCCGGAGCATTCCCTGGTCGG	1	1	1	15	SC05092
1446188	TGGCGTCCCGGAGCATTCCCTGG	1	1	4	36	SC05092

Design sgRNA: Tools and parameters

actVB

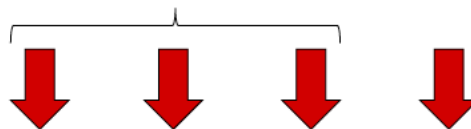
Target sequence

GTCCCGGAGCATTCCCTGGTCGG

Guide sequence

GTCCCGGAGCATTCCCTGGT (+ CGG, PAM)

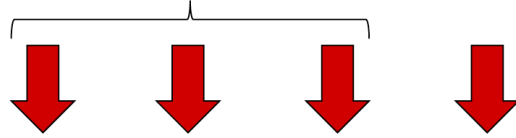
Off-targets?



Target ID	Target Sequence	Exon number	# of Exact matches	# of 1 bp mismatch matches	# of 2 bp mismatch matches	Exact matches in these genes
1446187	GTCCCGGAGCATTCCCTGGTCGG	1	1	1	15	SC05092
1446188	TGGCGTCCCGGAGCATTCCCTGG	1	1	4	36	SC05092



Off-targets?



Target ID	Target Sequence	Exon number	# of Exact matches	# of 1 bp mismatch matches	# of 2 bp mismatch matches	Exact matches in these genes
1446187	GTCCCGGAGCATTCCCTGGTCGG	1	1	1	15	SC05092
1446188	TGGCGTCCCGGAGCATTCCCTGG	1	1	4	36	SC05092

Design sgRNA: Tools and parameters

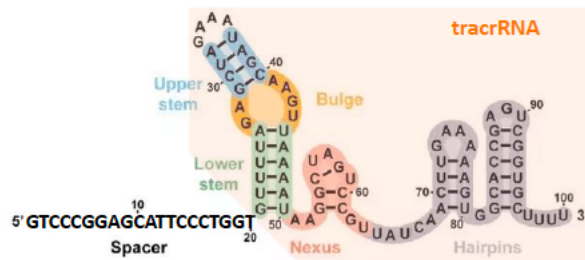
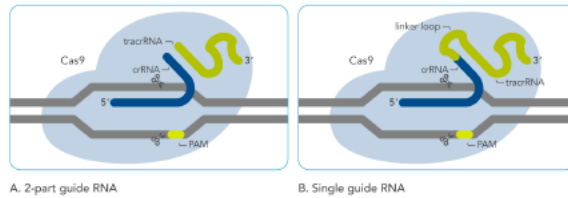
actVB

Target sequence

GTCCCGGAGCATTCCCTGGTCGG

Guide sequence

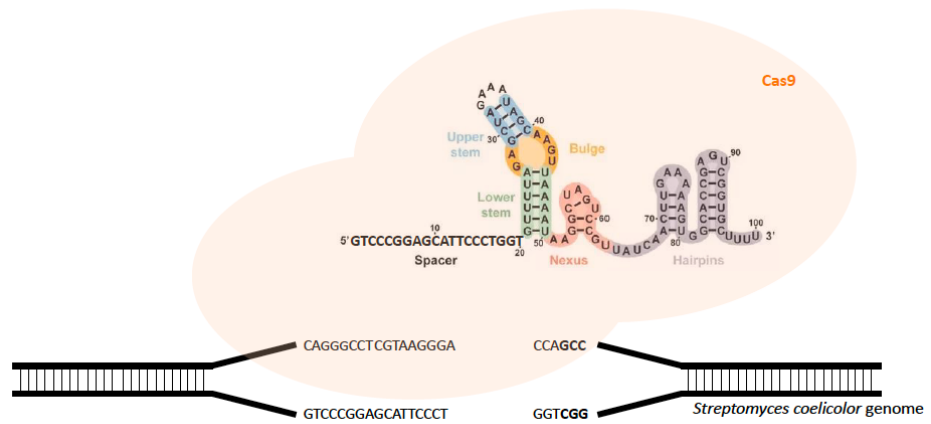
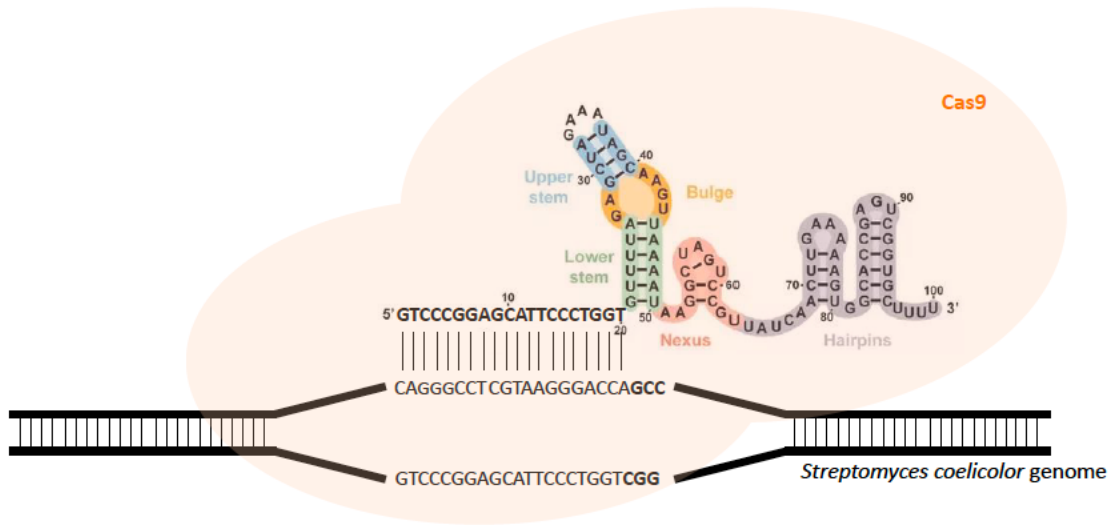
GTCCCGGAGCATTCCCTGGT (+ CGG, PAM)



actVB

target sequence GTCCCGGAGCATTCCCTGGTCGG

Guide sequence GTCCCGGAGCATTCCCTGGT (+ CGG, PAM)



Hands on CRISPR

slide

5' TTCTAATACGACTCACTATAG**G**NNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGA 3'

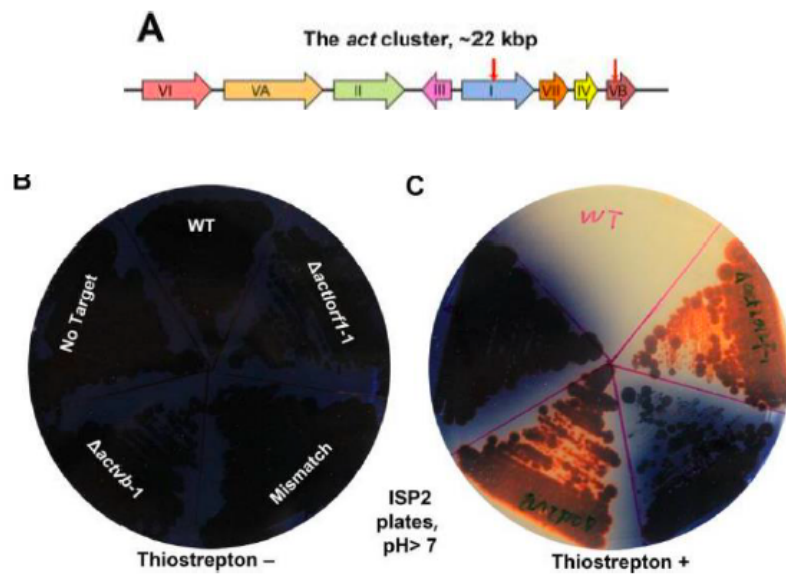
sgHL3:

TTCTAATACGACTCACTATAG**G**TCCCGGAGCATTCCCTGGTGTTTTAGAGCTAGA

sgHL4:

TTCTAATACGACTCACTATAGCCAGGCAGACCACTGCCAGGTTTTAGAGCTAGA

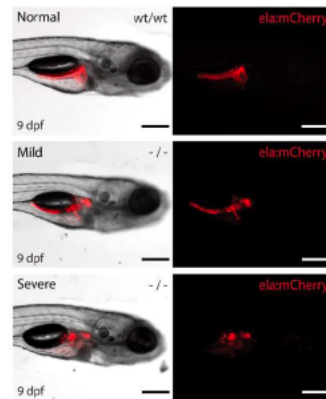
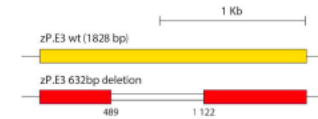
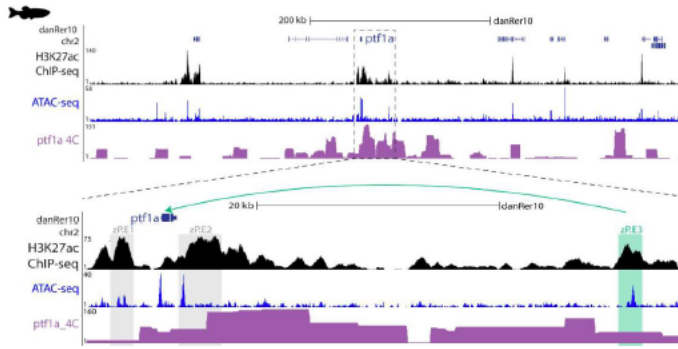
IMAGENS A B C - Tong et al, 2015



Tong et al., 2015

Zebrafish example

A non-coding genomic region from Zebrafish (*Danio rerio*)



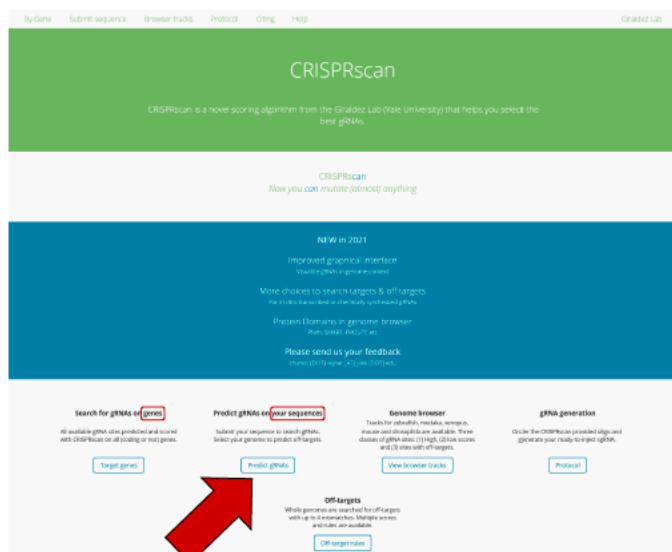
Bordeira-Carriço, Teixeira, Duque... Bessa, *BioRxiv* (2020)

Design sgRNAs

CRISPRscan (many species):
<https://www.crisprscan.org/>

Design sgRNAs

CRISPRscan:
<https://www.crisprscan.org/>



Design sgRNAs

The screenshot shows the CRISPRscan web interface. At the top, there are navigation tabs: "By Gene", "Submit sequence", "Browser tracks", "Protocol", "Citing", "Help", and "CRISPRscan". Below the tabs is a text input area containing a DNA sequence. A dropdown menu shows "Zebrafish - Danio rerio" and "Cas9 - NSG". There are buttons for "Get sgRNAs", "OK", and "Example".

Below the sequence input is a genomic browser track showing a scale from 28,746,400 to 28,750,000. A table below the track lists CRISPRscan results:

CRISPRscan score	Locus	Target sequence	Off-targets	CFD	All	Seed
69	2:29750346-29750368 (-)	TGTGTGAGTAGTTCGACATTCGG	1.08	0	0	
51	2:29749829-29749851 (+)	AGTGGACTGAAGTCCAGTCCCTGG	3.11	0	0	
46	2:29749284-29749306 (+)	GTGAATGTTGTTGACTACAAGG	7.20	0	0	
46	2:29749855-29749877 (+)	GCATGACCTGCTCGGGCTCCAGG	1.79	0	0	
45	2:29749873-29749895 (-)	GTCAGAAGCTTGTCAAGGCTCGG	4.03	0	0	
43	2:29750295-29750317 (+)	AGAGTGTGAGATGATAGCTATGG	5.68	0	0	
40	2:29749767-29749789 (+)	GACATCAGGACATAACGACAGG	2.92	0	0	
40	2:29750162-29750184 (+)	AGTGGCTAAATAGTCAGATGG	6.40	0	0	
40	2:29749618-29749640 (-)	AGAGGCAGATGACAGCAAAAAGG	57.86	0	0	

On the right side, a "Site type" panel shows details for the selected site (gG1BNGG), including the genome sequence, gRNA sequence, and off-targets (top 30 shown out of 118).

Design sgRNAs

The screenshot shows the Cas9 analysis tool. It displays a sequence alignment between a scaffold and a genome. The scaffold sequence is "GGTCATGACGGTTCGGGGCGG" and the genome sequence is "GGTCATGACGGTTCGGGGCGGAGG". The alignment shows a mismatch at the end of the sequence.

Below the alignment, a table summarizes the mismatch analysis:

chromosome coordinate	mismatch with target	all	seed	CFD
KN150225.1:13638[.....] +	GGTCATGACGGTTCGGGGCGG → 0 mismatch	→ yes	yes	1.00
21:5126754[.....X.....] +	GGTCATGACGGACGAGGCGG → 2 mismatches in seed	→ yes	no	0.75
22:17959653[X.....X.....] +	TGTTATGACTGTCGGGACGG → 4 mismatches	→ no	no	0.61

Additional information: "mismatch with target", "all: up to 2", "seed: up to 2 outside seed", "CFD: up to 4". Total CFD score: 4.14.

All
 According to [Hsu et al.](#), *Nature Biotechnology* 2013, potential off-targets can have a maximum of 2 mismatches with the sgRNA.

Seed
 With the method published by [Cong et al.](#), *Science* 2013, potential off-targets must match perfectly in their seed (12 nt 3' of the PAM sequence) and a maximum of 2 mismatches in the rest of the sgRNA. This rule is more stringent than the *All* method and therefore less off-targets are found.

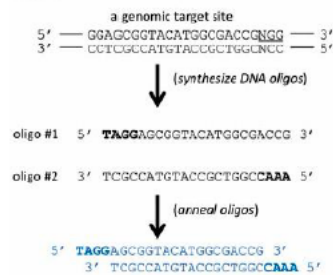
CFD (Cutting Frequency Determination)
[Doench et al.](#), *Nature Biotechnology* 2016 measured the cutting efficiency of potential off-targets and integrated them into the CFD score. Potential off-targets with up to 4 mismatches are scored with Doench et al. matrix.

Design sgRNAs



Clone sgRNAs

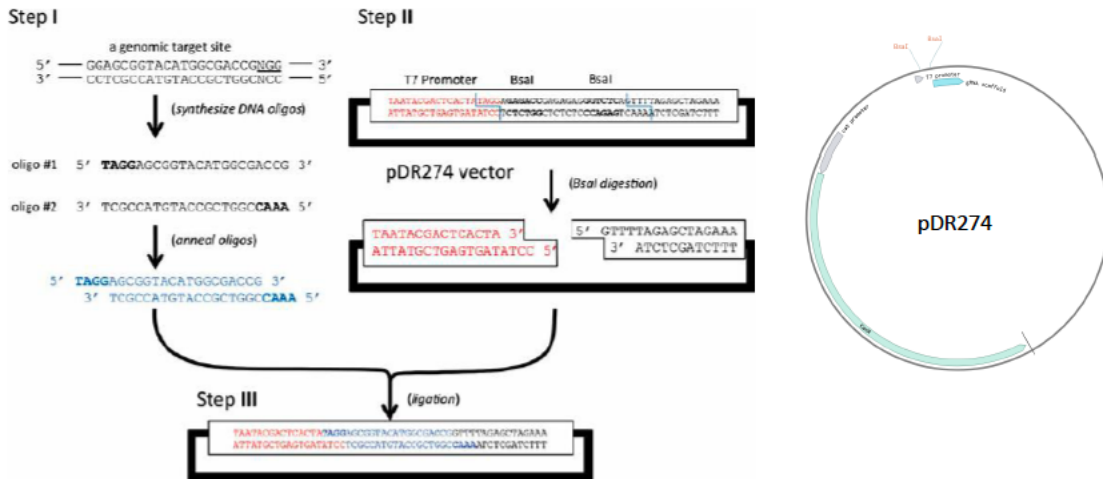
Step 1



sgRNA name	Sequence (with overhangs)
Ptf1aE3_Sg1_F	TAGGGAATGTTGTTTGACTACA
Ptf1aE3_Sg1_R	AAACTGTAGTCAAACAACATTCCC
Ptf1aE3_Sg3_F	TAGGCGAAAACGTTTTAGCAGA
Ptf1aE3_Sg3_R	AAACTCTGCTAAAACGTTTTCGCC
Ptf1aE3_Sg4_F	TAGGTGTGAGGTAGTTCGCATT
Ptf1aE3_Sg4_R	AAACAATGCGAACTACCTCACACC



Clone sgRNAs



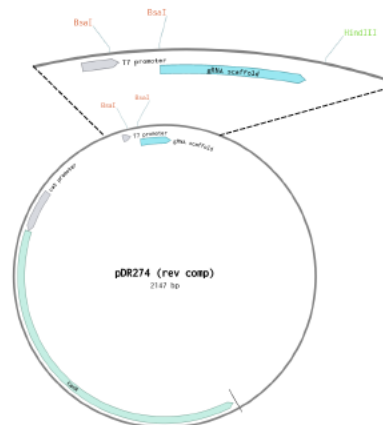
Clone sgRNAs

Cleave DNA in vitro

DNA Template (zebrafish *ptf1a* enhancer) – PCR product (~1800bp)

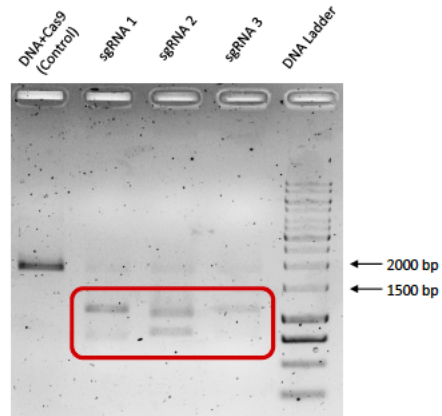
+
 sgRNA (synthesized with T7 DNA polymerase)
 - includes spacer and tracrRNA

+
 Cas9 protein



Cleave DNA in vitro

Check cleavage efficiency in an agarose gel:



In silico activity Protocol

Design the target for the sgRNA - Target-specific oligo design

This kit contains the *S. pyogenes* Cas9 Scaffold Oligo within the EnGen 2X sgRNA Reaction Mix, *S. pyogenes*. Target-specific oligos are designed by the user as follows:

1. Select 20 nucleotide target sequence (not including the PAM (NGG) sequence). Use of a target DNA selection program is recommended. We recommend [ChopChop](#).
2. Check input sequence for presence of "G" at the 5' end. If there are no "G's" at the 5' end, add one "G" (making it a total of at least one G at the 5' end).
3. To the 5' end; append T7 promoter sequence: TTCTAATACGACTCACTATA
4. To the 3' end; append 14 nucleotide overlap sequence: GTTTTAGAGCTAGA
5. Check complete oligo sequence: 5'
TTCTAATACGACTCACTATAG(N)20GTTTTAGAGCTAGA

Design the respective sequences to synthesize one sgRNA to target the "red" sequence of the partial (Bull's) sequence - Template for oligo design and sequence to mutate

The challenge!

Design the Genomic target region that should have the following configuration:

Target(20N) NGG

Design the sequence of the oligo to order, required for the sgRNA transcription, having the following configuration:

T7Promoter (G) Target(20N) Linker

Design the full template for the sgRNA sequence transcription, that must have the following configuration:

T7Promoter (G) Target(20N) Scaffold

In silico Components:

T7 promoter

TTCTAATACGACTCACTATA(**G**)

Linker

GTTTTAGAGCTAGA



Scaffold (5>3)

GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCAC
CGAGTCGGTGCTTTT

Sequence to mutate:

TCGGTCCC GTGTATCCATGGCCCAGCCGGCACTTGT CAGACTGT CACAGATCGGGGGTGAGATA
TTTCTCCTGTTCCAACACGTGATGATTTTCGCTTGCAATGGCTCGTGGCCAACATAGCGAGTATGG
ACTCCAGGACGTCCTGATCCCGGATGCCTGGGGTTTTTACCAGTTTCGT CAGAGTCCTCAGGATA
CTACGCCCACTTCTATCGCCAAGTTTCATGAGAAACACTCGCCTTTGGATTAATCTCAGCGACC
CAGGGCACGCAAGAGGGGTGCGTCGT CGAAGCCAGGTCACCACTTACA ACTTGGGGGGGCCTA
GTGGACGGGGTATTTCGATCCTTCCGTAGCTCAA ACTTTATAGAGT CACAGACACCTCCC ACTGAC
CTTTGTAGTTGGTGCCGGTCCACTACAGGCTACAAGGCTCTGAGTAGTGGCGGGTGGGGTACTC
GAGAACACGTTAGTCAGCTCCGGCGTTCCACTTCAGCCTAGCCTAAGACAAATAACCCTTGCAG
TACCACTAACAAGGAGCGCCGGGGCGGGATATTTTAAATTA AAAAAACA ACTTACCCCTACGAA
CGATGTGAACGAGCGTAAAATTGCATGGTGACCCGGGCGCCAAGTCTCTCTTGACTGCCCGCGA
GCCTTGGGCGTTCGTCTTTACTGAGTCTCGGGTTACGTCCTGCCTTCTATTGTCCCAAATTAGAAC
GACTATCGGCCTGATCACAAGATTACCATCTCTGGAATTCGGATTCGCAAAGTAGTTTTTATTCGG
CAATTGATGAGAGCAGGGCTGTAGATGCAAGCTCGTTCAGCATGTCCCTCACGTACCTCCTTGG
CGTATGGCTGGTCACCTAGGTACGGTGGCTGATGAAAGGGAGGCTTAGCTGGGTGCAGAGGTA
TATAAGGGTCGTCCGGTATCCTTTGTTTCAGCGGAAGCAACGCGGACAGGCATCAGACTCATTG
CGGTAGGGAAGCCCGCATCAAGAATTATATGTCCTGTGTCGCCGAGCTCTCTGAGCGTGTTCGG
AGCCTGTGGTTCTCGGTGTCATCCGTGGACACCTAGAAGAGCGGTGGTAGGTGCGGTAGGATG
GCTTTAGTACACGTGAGAGACTGAAAGATCTATAAGCTCGGGCTTTAAACGAAACCGTCAACCAA
GTCGCTCGAGCGTAAAATATAGATTCCCATCGGGCGCCAAGTTTGTTCAGCGGTAGCTACGGG
GAGGAAGGTTTAAAGCTAGCATTCTCAGTCTCGTAGGTACGTTTTCAACTCGCTTCACGCTGTG
TTGGTTTGAAGGCATGCAGTATGTATGTATGATCAACTAAATGCTCGATAAAAGGAGATGTTATG
AGGCTTTTCGCTGATGCCTGTTAGGTTGGATGGATCCGGCAGACAGCAGCGATTTCATGACTT
ACTCCACATCTGACTCTACGCCTAGCTTCACATGCATGCTATTAGAGCTAGGCCACTGGCTGGGA
TACGCCGTCCCAATGGTACATAGTGGCCAGTAAGATCATCGTTCGTATTCTATGCATGGTGACCT
CTCTCAGGTTAGTTCGATTCTACCTGCCAAATTTGGTTGGGCTGTTGATCGTGAGTCCACCGCAT
CCCCCCCCAAAACAATCGGGTTTAGCTTCTAAACTAGAGCTTCATCAGGAACCGTCGGGATGGTTA
CGTAATCGAGTGT CAGTGC GTGTCTCTGTGCGGCTCCACCCTGATAAGTGC GTCTTCGAGCCG
GCGATTTCGTCGGTTCGCGTTCGAGGGAGCCCAGTGATAGAAGTGTAACGTACCCCTGGGTAAGTTG
TCACTTGGTTCCGCGTTTTTTTAAAGCTGTATTTAAAGGTGTGAGTGTGCAATCCTCCAGATTAGCT
AGCATTCTGACTGGAGGCTCTACCGTACGCCCAAGAGTTACGTTACGGTTCGCTAGAAATGGG
TCACCCGTGT

TTCTAATACGACTCACTATAGTGCCTGTTAGGTTGGATGGATGG GTTTTAGAGCTAGA

Please order mine: TTCTAATACGACTCACTATAGTGCCTGTTAGGTTGGATGGATGG
GTTTTAGAGCTAGA

