**GENETICS** 

# OFF TARGET SITES OF THE CRISPR CAS BABIES

14.12.2018

Unfortunately we don't know the genome of the Chinese parents which would be necessary for any prediction of genome mutations in the twins. So far we only know one confirmed off target site from Dr. He's Hongkong slides that reports 5 mismatches.

target	C A G A A T T G A T A C T G A C T G T A T G G	
off -target	C A G A A T T T A T A C C C A C A C T A G G G	

reported off target site

At least I would expect for statistical reasons, some 1 bp mismatch, some 2bp, some 3bp..., but not just one 5bp off target site.

So we could <u>check several</u> online services with the know sgRNA sequence that Dr. He used.

#### **CRISPOR**

Position/ Strand @	Guide Sequence + PAM + Restriction Enzymes 9 + Variants 9 Only G- Only G- Only A-	Specificity Score 2	Predicted Efficiency Show all score	<u> </u>	Frame	Off-targets for 0-1-2-3-4 mismatches + next to PAM	Genome Browser links to matches sorted by CFD off-target score 99 exons only 10 chr3 only
21 / fw	CAGAATTGATACTGACTGTA TGG .g. Enzymes: HpyCH4III Cloning / PCR primers	77	38	48		0-0-1-13-136 0-0-0-3-3 150 off-targets	4:Intergenic:RP11-6N13.1- RNASSP189 4:exon:RP11-871F6.3 3:Intergenic:UBR3-RNU6-1006P show all

CRISPOR finds one region with 2 mismatches (intergenic) and 13 with 3 mismatches (7 intergenic hits, 5x introns (ST8SIA6, CTD-2532D12.5, CNTN5, C2D3, RABGAP1) and 1x exon of ACACA). This doesn't match the reported off target sequence.

#### **Cas-OFFinder**

Bulge Type	Target	Chromosome	Position	Direction	Mismatches	Bulge Size	
RNA	CTRNA: CAGAATTGATACTGACTGTATGGNGG DNA: CAGAATTGATACTCTGTAgGCTGG	chr8	135886978	+	2	2	
RNA	CTRNA: CAGAATTGATACTGACTGTATGGNGG DNA: BAGAATTGATACTGACTATGGTGG	chr4	41673033		2	2	
RNA	CTRNA: CAGAATTGATACTGACTGTATGGNGG DNA: BAGAATTGATACTGACTATGGTGG	chr4	41673033		2	2	
RNA	CTRNA: CAGAATTGATACTGACTGTATGGNGG DNA: CAGACTTTACTGACTaTATGGAGG	chr13	27086698	-	2	2	
RNA	Crrna: CAGAATTGATACTGACTGTATGGNGG DNA: CAGAATTTAgTGAgTGTATGGGGG	chr22	32835334	+	2	2	

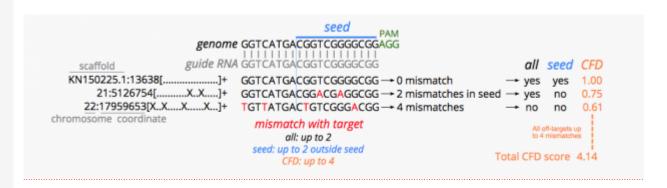
Cas-OFFinder identifies 5 regions with 2 mismatches (1x intergenic, 2x sites in LIMCH1, 1x intergenic, 1 in BPIFC).

CCtop (yes, I was waiting for that name :-)

Coordinates	strand MM		target_seq		PAM distance		gene name	gene id
chr7:4557658-4557680	+	4	CAATATTT [TTACTGACTGTA]	GGG	2763	-	MRPL49	ENSORLG00000001995
chr8:11092944-11092966	-	4	CAGAGTTA [ATGGTGACTGTA]	AGG	1274	-	rab26 member ras oncogene family	XLOC_022613
chr13:23409646-23409668	+	4	AAGAACTG[ATAAAGACTGTA]	TGG	96411	-	SI	ENSORLG00000012288
chr17:23133141-23133163	-	4	CACAATCT [ATACTGACTATA]	GGG	3233	I	LINGO3	XLOC_009548
chr5:26674537-26674559	+	3	CAGAATGG [ATACAGACAGTA]	AGG	21340	-	SULF2	XLOC_019364
chr7:24324097-24324119	-	4	CAGCACTG [ATATTGACTGAA]	AGG	5382	-	acap3a	ENSORLG00000016028
chr11:5882914-5882936	-	4	AAGAATTA [ATACTGTCTCTA]	AGG	4234	-		ENSORLG00000002725
chr11:14821058-14821080	+	4	CAGCACTA [ATACTGACTGTC]	AGG	0	E	ENSORLG00000006093	XLOC_002395
chr3:33197680-33197702	-	4	CAGCACTG [ATACTGACAGTG]	GGG	NA	-	NA	NA
chr15:14877318-14877340	+	4	CAGAAGTC [ATACTGACAGTC]	GGG	0	E	fbxo5	ENSORLG00000003909

COS finds only one intergenic region with 3 mismatches.

### **CRISPRSCAN**



has the best explanation of what it does but cannot search from sgRNA to off-target as many other online tools.

## **OFFspotter**

Chroth	Stranki	Start Φ	End ♥	Given query \$	Actual genomic hit •	Number of mismatches	Pre-mRNA (Unspliced)	mRNA (SUTH)	mRNA (CDS) •	mRNA (SUTR)	lincRNA (Unspliced) ©	lincRNA (Spliced)	GC a	External links
3	-	46414940	46414962	CAGAATTGATACTGACTGTA	CAGAATTGATACTGACTGTA-TGG	0	ENSC00000160791: (CCR5 - chemokine (C-C motif) receptor 5 (gene/pseudogene)) ENST00000343801		ENSG00000160791; (CCR5 - chemokine (C-C motif) receptor 5 (gene/pseudogenel) ENST00000292303				35.0%	UCSC Genome Browser
							ENST00000292303 ENST00000445772		ENST00000343801 ENST00000445772					
13		54263120	54263142	CAGAATTGATACTGACTGTA	*AGAATTGATACTG*CTGTA-TGG	2							30.0%	UCSC Genome Browser
8	-	139024559	139024581	CAGRATTGRTACTGACTGTA	EAEAETTGATACTGACTGTA-AGG	3					ENSG00000253288: (RP11-238K6.1) ENST00000518973		25.0%	UCSC Genome Browser
2	٠	170946127	170946149	CAGAATTGATACTGACTGTA	ARGARITOTINCAGROTGTA-GGG	3		-		-			30.0%	UCSC Genome Browser
13	٠	88793606	88793628	CAGAATTGATACTGACTGTA	CAEARTTGATGCTGACTGTA-TGG	3	-		-		-		35.0%	UCSC Genome Browser

is spotting 1 intergenic region with 2 mismatches, 3 with 3 mismatches (2 intergenic, 1 in-RP11-238K6.1).

Taken these five different predictions together, I can't make conclusion but think that we need to resequence the families with much higher coverage (150-200x fold coverage) and even compare the mutations with the phenotype of the children.

CC-BY-NC Science Surf accessed 03.11.2025 ☐