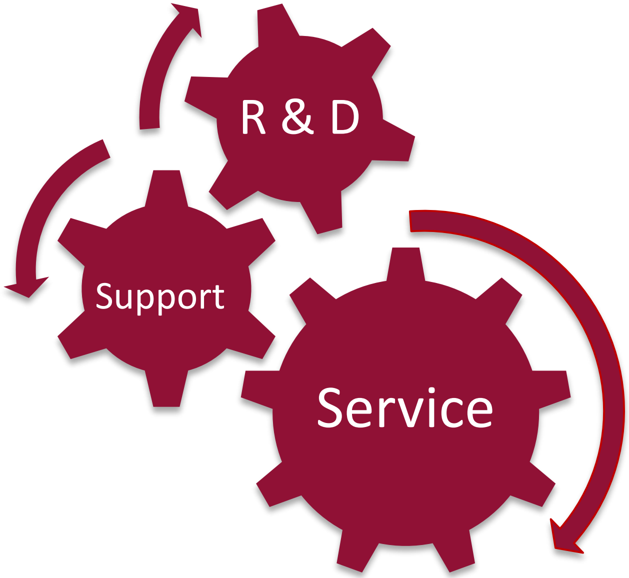
High Efficiency Genome Editing Guided by Targeted Deep Sequencing

Shondra Pruett-Miller, Ph.D.
Director, Center for Advanced Genome Engineering
Cell and Molecular Biology Department
Assistant Member

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

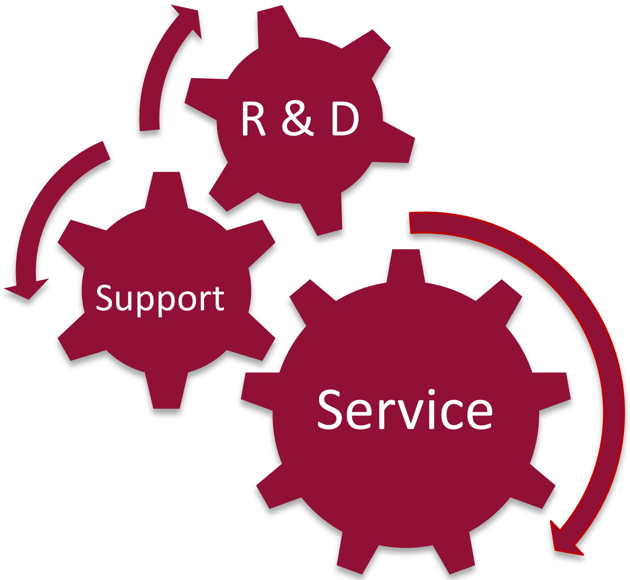
CAGE Mission



- ✓ Provide gene editing technology for investigator-initiated projects
- ✓ Perform R&D to keep SJCRH investigators at the forefront of emerging technologies
- ✓ Support key initiatives such as the gene therapy and cellular therapy programs

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

CAGE Mission



- ✓ Provide gene editing technology for investigator-initiated projects
- ✓ Perform R&D to keep SJCRH investigators at the forefront of emerging technologies
- ✓ Support key initiatives such as the gene therapy and cellular therapy programs

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

CAGE

Center for Advanced Genome Engineering

St. Jude Children's Research Hospital

Design and Validate gRNAs
(weeks)

>2000 gRNAs

Create custom cell lines
(3-6 months)

110 custom cell lines

Generation of animal models
(3-4 months)

17 animal models

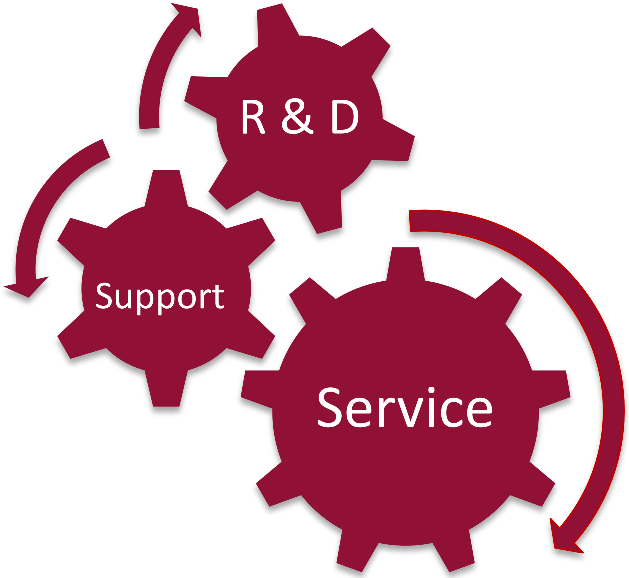
Facilitate High-throughput genetic screens

2 custom, pooled libraries

In our first year, with 5.5 FTE

On track to complete 192 cell lines with 7 FTE this year!!!

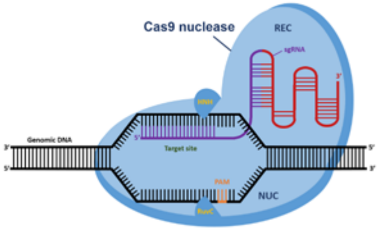
CAGE Mission

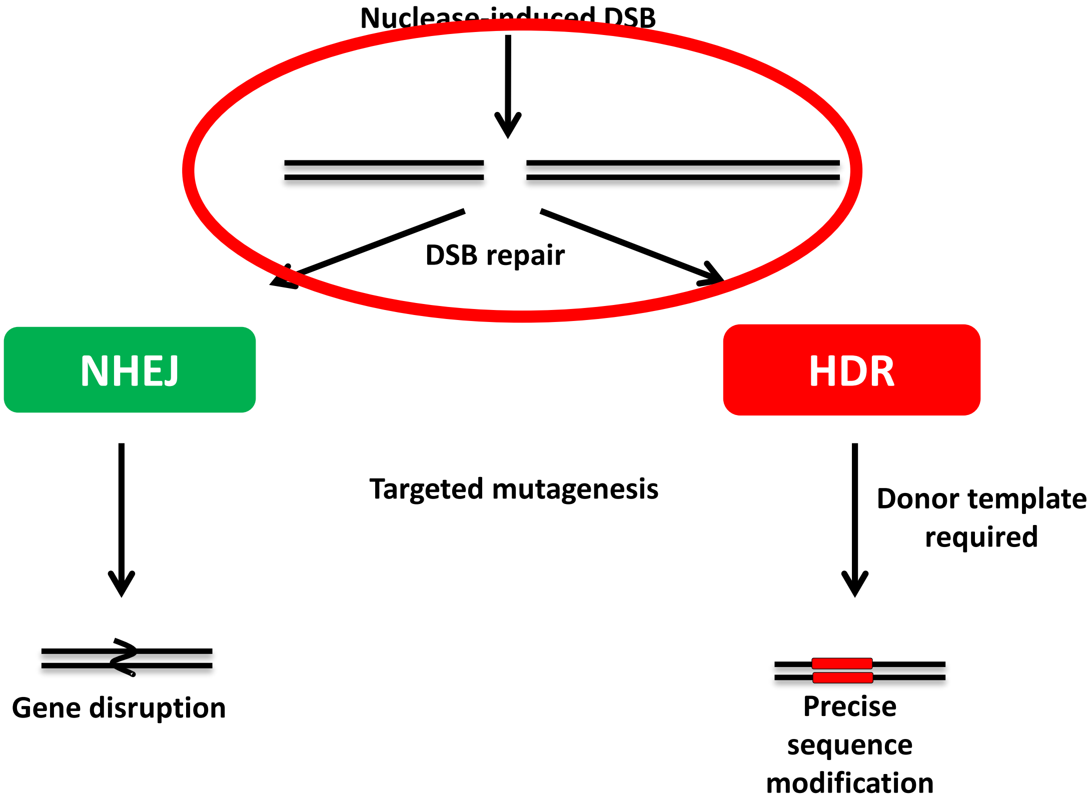


- ✓ Provide gene editing technology for investigator-initiated projects
- ✓ Perform R&D to keep SJCRH investigators at the forefront of emerging technologies
- ✓ Support key initiatives such as the gene therapy and cellular therapy programs

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

It’s All about that Break





Nuclease-induced DSB

DSB repair

NHEJ

HDR

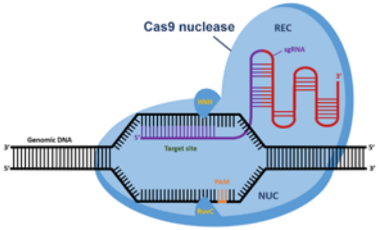
Targeted mutagenesis

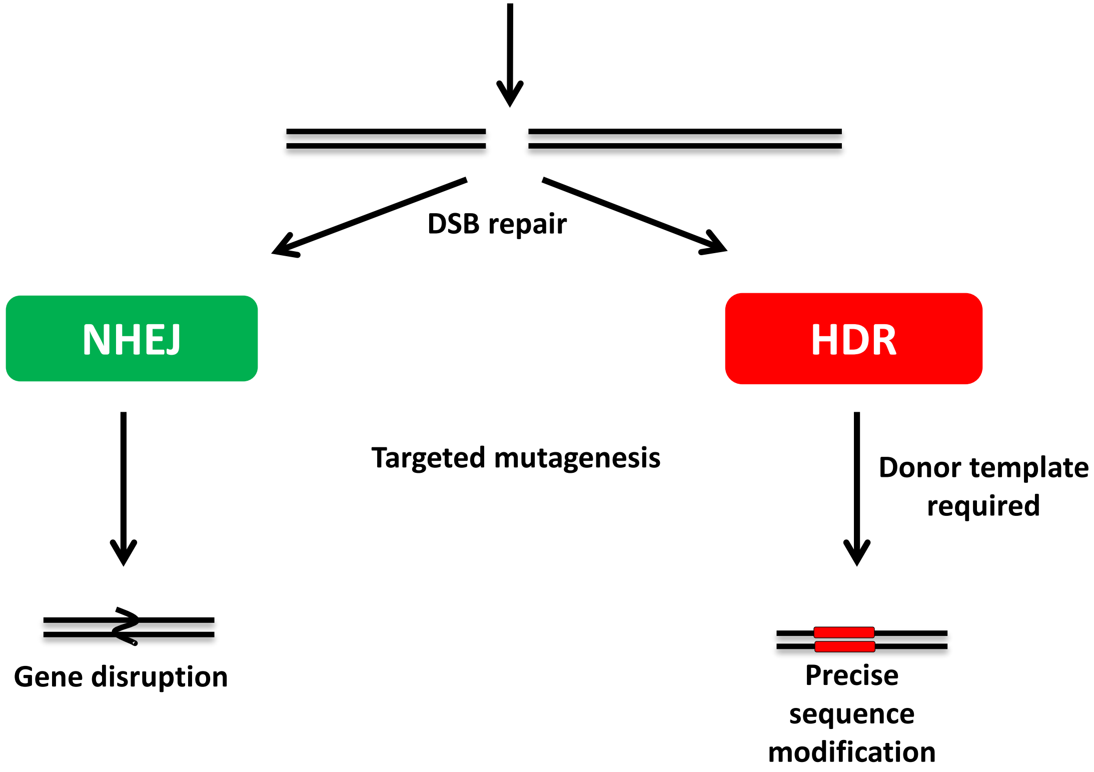
Donor template required

Gene disruption

Precise sequence modification

It’s All about that Break





Nuclease-induced DSB

DSB repair

NHEJ

HDR

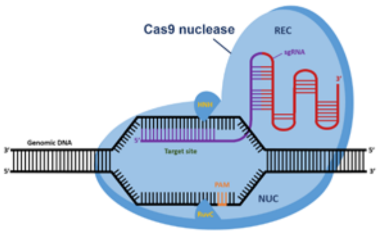
Targeted mutagenesis

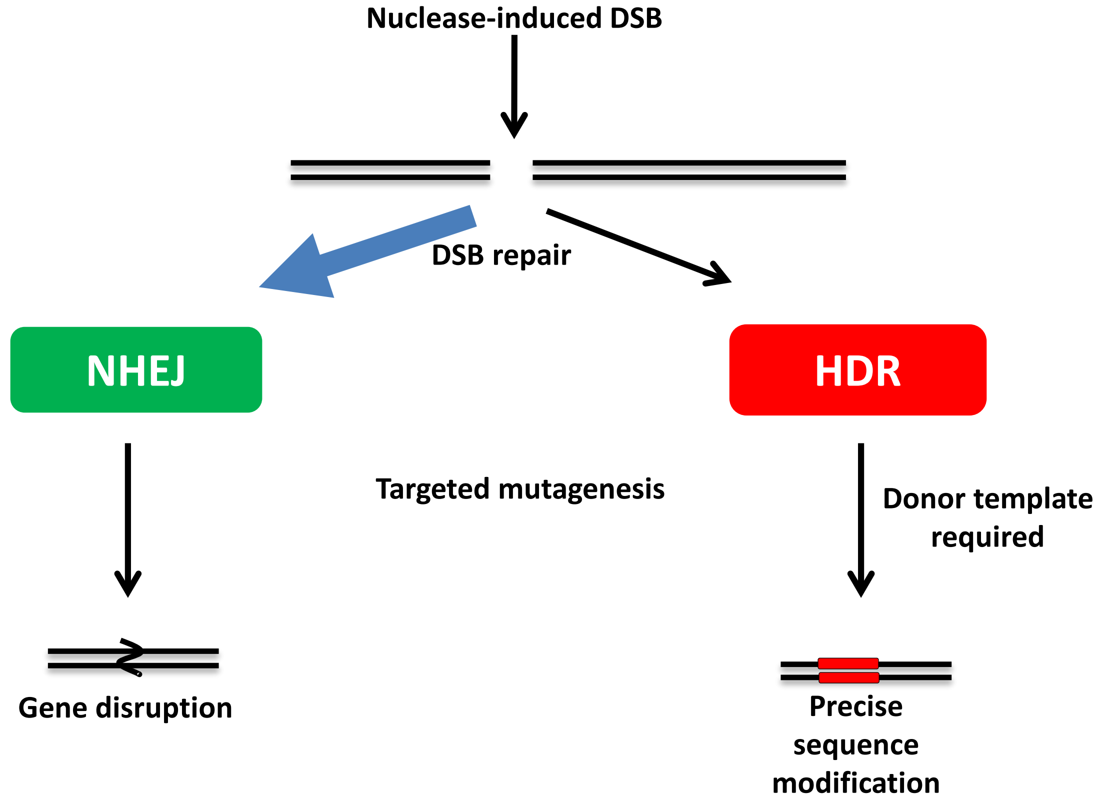
Donor template required

Gene disruption

Precise sequence modification

It’s All about that Break





Nuclease-induced DSB

DSB repair

NHEJ

HDR

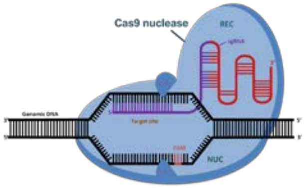
Targeted mutagenesis

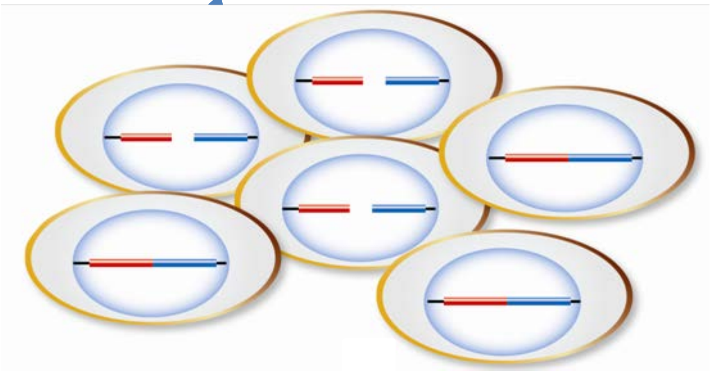
Donor template required

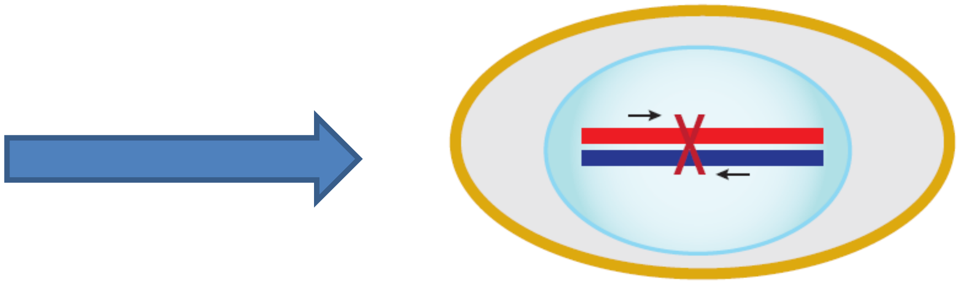
Gene disruption

Precise sequence modification

Assaying for activity

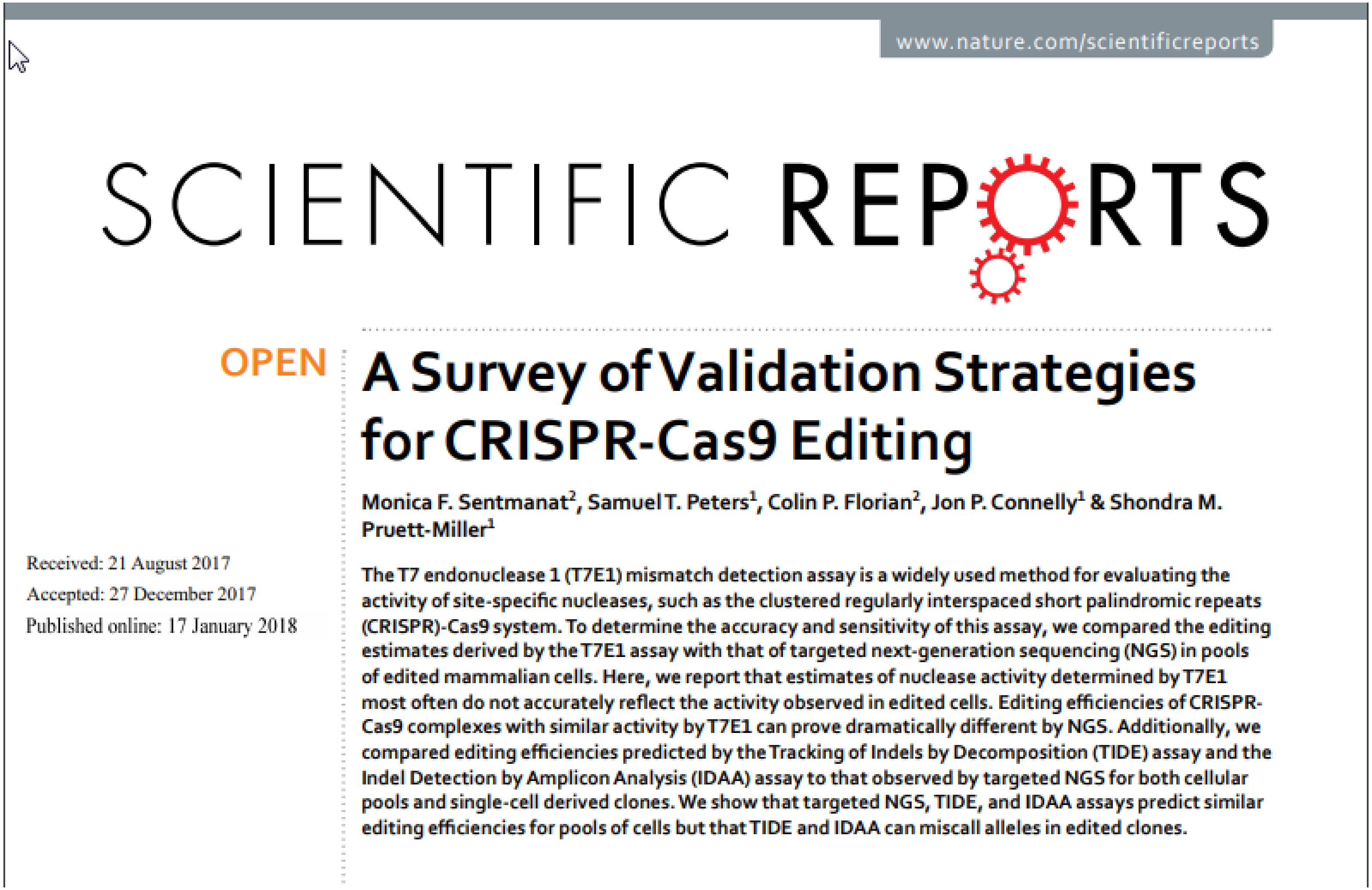






PCR amplify the region of interest

IDAA



www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

A Survey of Validation Strategies for CRISPR-Cas9 Editing

Monica F. Sentmanat[2], Samuel T. Peters[1], Colin P. Florian[2], Jon P. Connelly[1] & Shondra M. Pruett-Miller[1]

Received: 21 August 2017

Accepted: 27 December 2017

The T7 endonuclease 1 (T7E1) mismatch detection assay is a widely used method for evaluating the activity of site-specific nucleases, such as the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system. To determine the accuracy and sensitivity of this assay, we compared the editing estimates derived by the T7E1 assay with that of targeted next-generation sequencing (NGS) in pools of edited mammalian cells. Here, we report that estimates of nuclease activity determined by T7E1 most often do not accurately reflect the activity observed in edited cells. Editing efficiencies of CRISPR-Cas9 complexes with similar activity by T7E1 can prove dramatically different by NGS. Additionally, we compared editing efficiencies predicted by the Tracking of Indels by Decomposition (TIDE) assay and the Indel Detection by Amplicon Analysis (IDAA) assay to that observed by targeted NGS for both cellular pools and single-cell derived clones. We show that targeted NGS, TIDE, and IDAA assays predict similar editing efficiencies for pools of cells but that TIDE and IDAA can miscall alleles in edited clones.

gRNA Activity Assay Summary

	T7E1	TIDE	NGS	IDAA
Identify active gRNAs	Yes	Yes	Yes	Yes
Show indel sizes	No	Yes	Yes	Yes
Show sequence identities	No	No	Yes	No
Accurately portray editing rates – pools of cells	No	Yes (low)	Yes	Yes (low)
Accurately portray editing rates – clones	No	Mostly	Yes	Sometimes
Throughput	Low	Medium	High	Medium

Sentmanat et al., Sci Reports. 2018

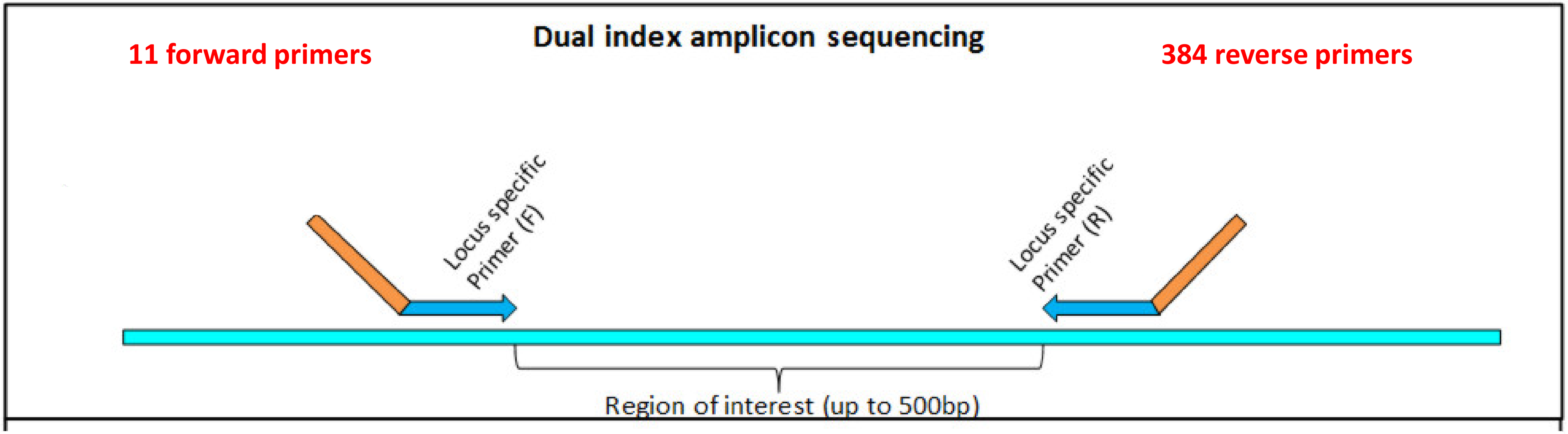
gRNA Activity Assay Limitations

	T7E1	TIDE	NGS	IDAA
Barrier to entry (cost, expertise, etc)	Low	Low	High	Low
PCR based bias	Yes	Yes	Yes	Yes
Detect gross chromosomal abortions	No	No	No	No
Amplicon size	300bp-900bp	~750bp	500bp-600bp	200bp-900bp

Sentmanat et al., Sci Reports. 2018

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

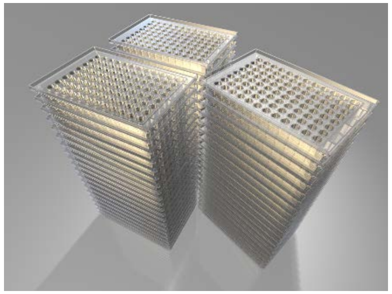
Dual index, 2-step PCR sequencing



11 forward primers

384 reverse primers

Library Prep for NGS

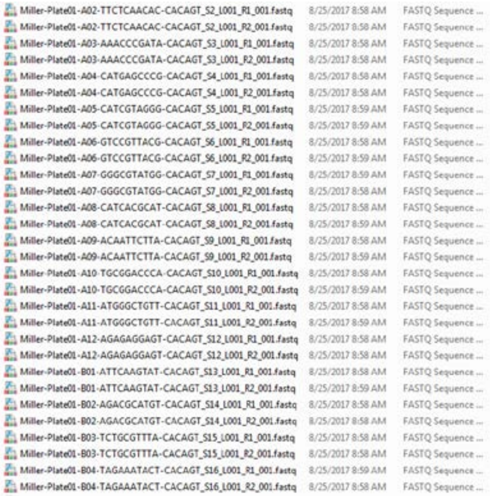


2-step PCR

Unique primers
for each project

Pool samples

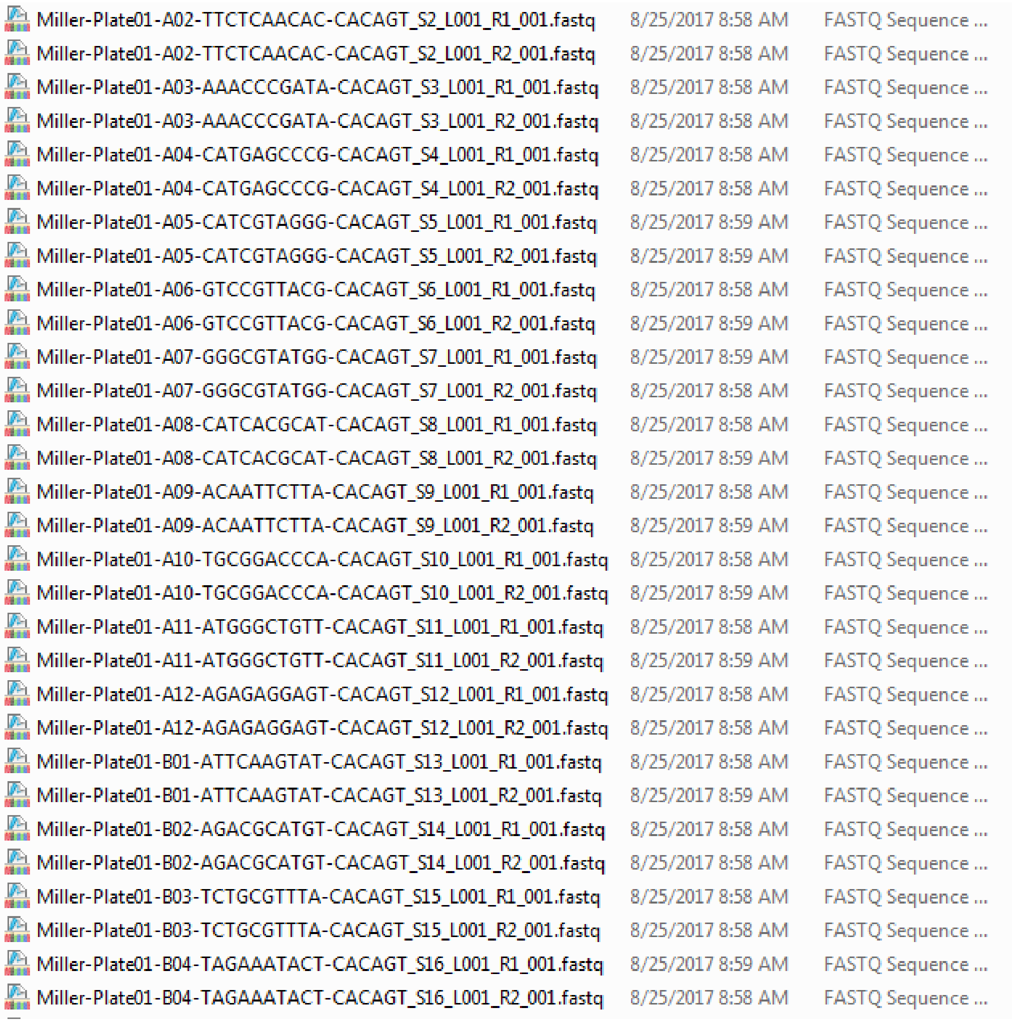
miSeq



```
Miller-Plate01-A02-TTCTCAACAC-CACAGT_S2_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A02-TTCTCAACAC-CACAGT_S2_L001_R2_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A03-AAACCCGATA-CACAGT_S3_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A03-AAACCCGATA-CACAGT_S3_L001_R2_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A04-CATGAGCCCG-CACAGT_S4_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A04-CATGAGCCCG-CACAGT_S4_L001_R2_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A05-CATCGTAGGG-CACAGT_S5_L001_R1_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A05-CATCGTAGGG-CACAGT_S5_L001_R2_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A06-GTCCGTTACG-CACAGT_S6_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A06-GTCCGTTACG-CACAGT_S6_L001_R2_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A07-GGGCGTATGG-CACAGT_S7_L001_R1_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A07-GGGCGTATGG-CACAGT_S7_L001_R2_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A08-CATCACGCAT-CACAGT_S8_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A08-CATCACGCAT-CACAGT_S8_L001_R2_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A09-ACAATTCTTA-CACAGT_S9_L001_R1_001.fastq      8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A09-ACAATTCTTA-CACAGT_S9_L001_R2_001.fastq      8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A10-TGCGGACCCA-CACAGT_S10_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A10-TGCGGACCCA-CACAGT_S10_L001_R2_001.fastq     8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A11-ATGGGCTGTT-CACAGT_S11_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A11-ATGGGCTGTT-CACAGT_S11_L001_R2_001.fastq     8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-A12-AGAGAGGAGT-CACAGT_S12_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-A12-AGAGAGGAGT-CACAGT_S12_L001_R2_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B01-ATTCAAGTAT-CACAGT_S13_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B01-ATTCAAGTAT-CACAGT_S13_L001_R2_001.fastq     8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-B02-AGACGCATGT-CACAGT_S14_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B02-AGACGCATGT-CACAGT_S14_L001_R2_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B03-TCTGCGTTTA-CACAGT_S15_L001_R1_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B03-TCTGCGTTTA-CACAGT_S15_L001_R2_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
Miller-Plate01-B04-TAGAAATACT-CACAGT_S16_L001_R1_001.fastq     8/25/2017 8:59 AM    FASTQ Sequence ...
Miller-Plate01-B04-TAGAAATACT-CACAGT_S16_L001_R2_001.fastq     8/25/2017 8:58 AM    FASTQ Sequence ...
```

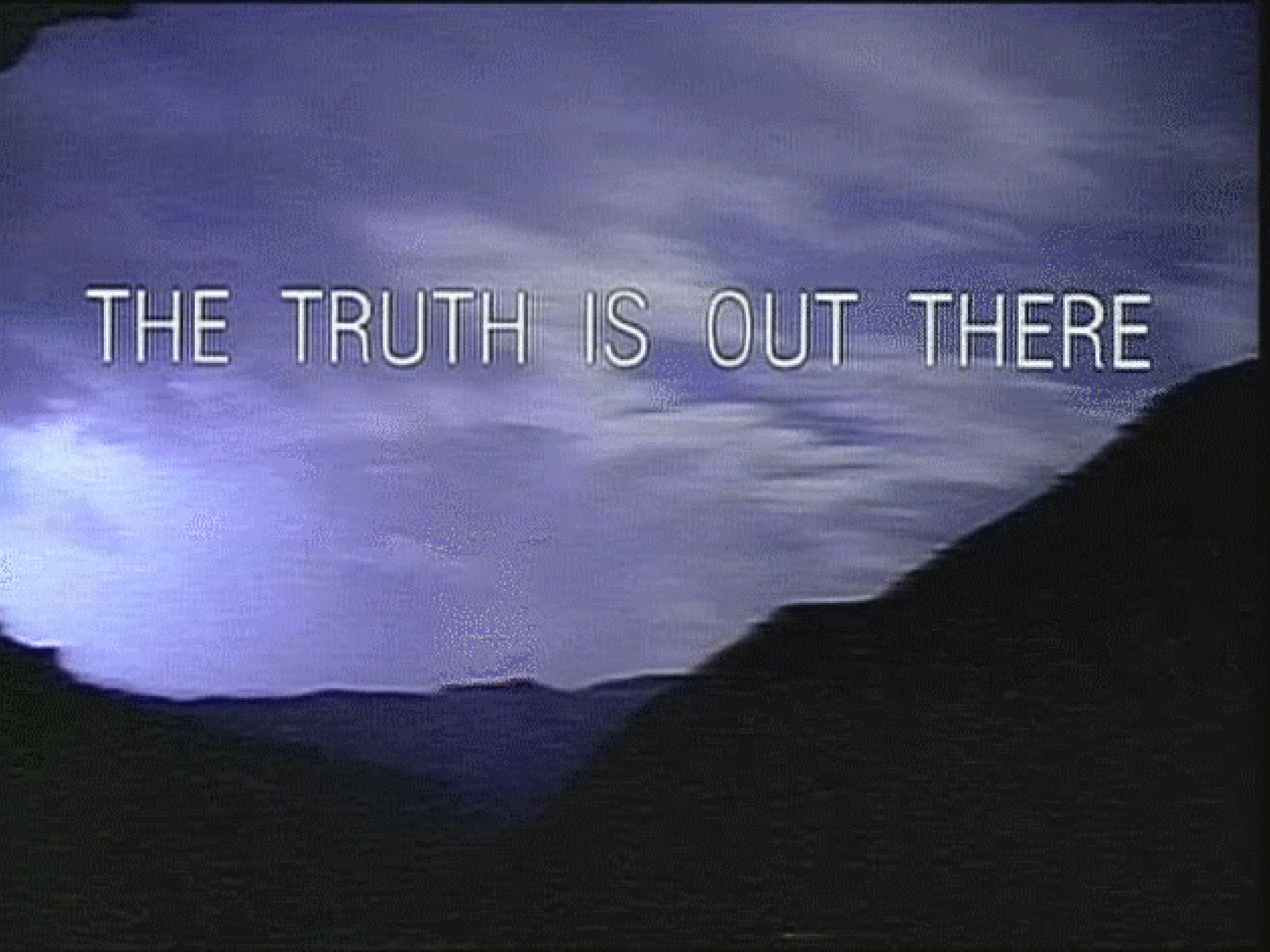
Lots of data

Targeted NGS output

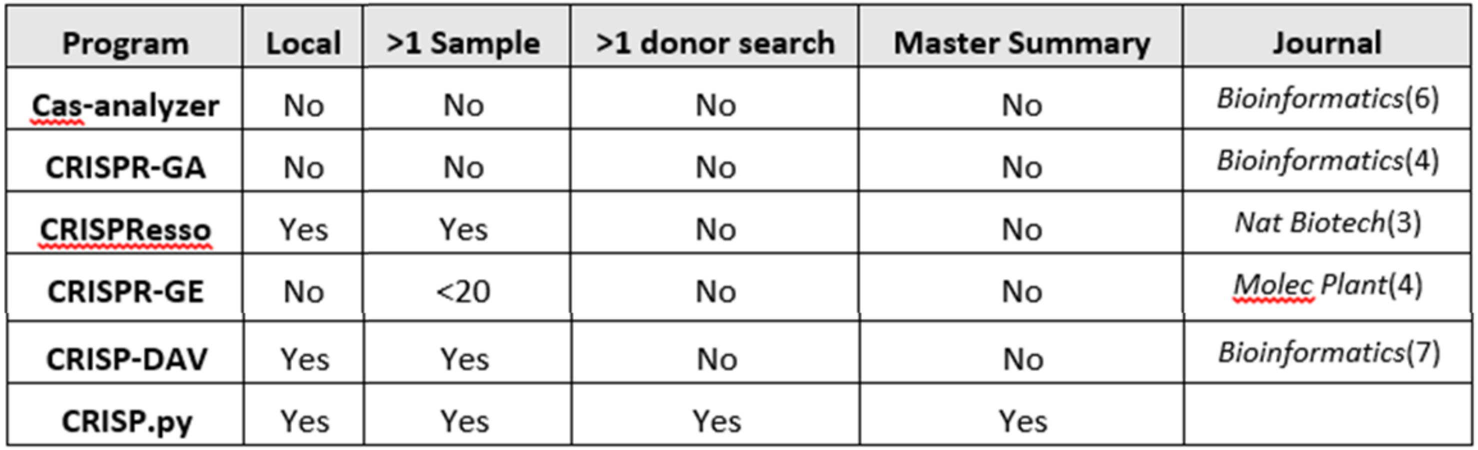


Miller-Plate01-A02-TTCTCAACAC-CACAGT_S2_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A02-TTCTCAACAC-CACAGT_S2_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A03-AAACCCGATA-CACAGT_S3_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A03-AAACCCGATA-CACAGT_S3_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A04-CATGAGCCCG-CACAGT_S4_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A04-CATGAGCCCG-CACAGT_S4_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A05-CATCGTAGGG-CACAGT_S5_L001_R1_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A05-CATCGTAGGG-CACAGT_S5_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A06-GTCCGTTACG-CACAGT_S6_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A06-GTCCGTTACG-CACAGT_S6_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A07-GGGCGTATGG-CACAGT_S7_L001_R1_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A07-GGGCGTATGG-CACAGT_S7_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A08-CATCACGCAT-CACAGT_S8_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A08-CATCACGCAT-CACAGT_S8_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A09-ACAATTCTTA-CACAGT_S9_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A09-ACAATTCTTA-CACAGT_S9_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A10-TGCGGACCCA-CACAGT_S10_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A10-TGCGGACCCA-CACAGT_S10_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A11-ATGGGCTGTT-CACAGT_S11_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A11-ATGGGCTGTT-CACAGT_S11_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-A12-AGAGAGGAGT-CACAGT_S12_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-A12-AGAGAGGAGT-CACAGT_S12_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B01-ATTCAAGTAT-CACAGT_S13_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B01-ATTCAAGTAT-CACAGT_S13_L001_R2_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-B02-AGACGCATGT-CACAGT_S14_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B02-AGACGCATGT-CACAGT_S14_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B03-TCTGCGTTTA-CACAGT_S15_L001_R1_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B03-TCTGCGTTTA-CACAGT_S15_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...
Miller-Plate01-B04-TAGAAATACT-CACAGT_S16_L001_R1_001.fastq	8/25/2017 8:59 AM	FASTQ Sequence ...
Miller-Plate01-B04-TAGAAATACT-CACAGT_S16_L001_R2_001.fastq	8/25/2017 8:58 AM	FASTQ Sequence ...

Tells you gRNA activity, size of indels and sequence identity

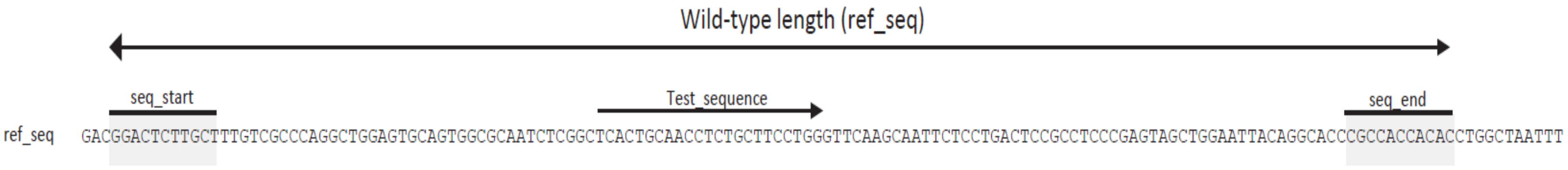

THE TRUTH IS OUT THERE

How does CRIS.py compare?



Program	Local	>1 Sample	>1 donor search	Master Summary	Journal
Cas-analyzer	No	No	No	No	*Bioinformatics*(6)
CRISPR-GA	No	No	No	No	*Bioinformatics*(4)
CRISPResso	Yes	Yes	No	No	*Nat Biotech*(3)
CRISPR-GE	No	<20	No	No	*Molec Plant*(4)
CRISP-DAV	Yes	Yes	No	No	*Bioinformatics*(7)
CRISP.py	Yes	Yes	Yes	Yes	

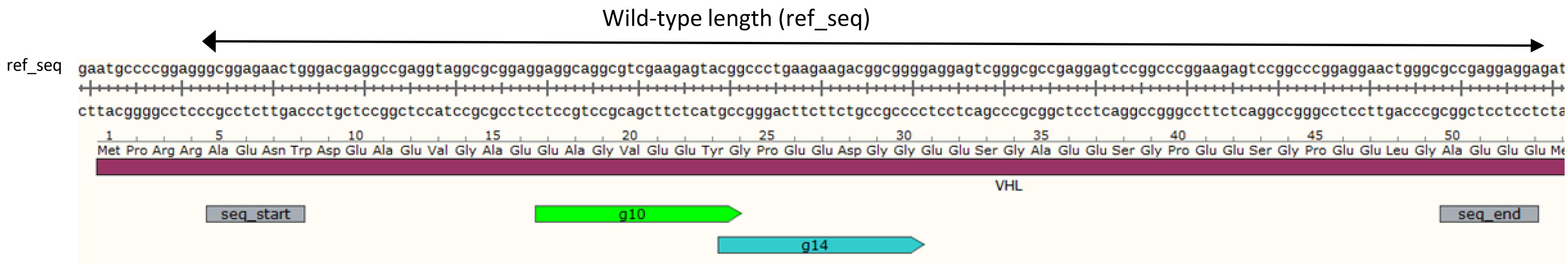
CRIS.py: A Versatile and High-throughput analysis program



Search sequence for

1. Length
2. Exact sequence matches

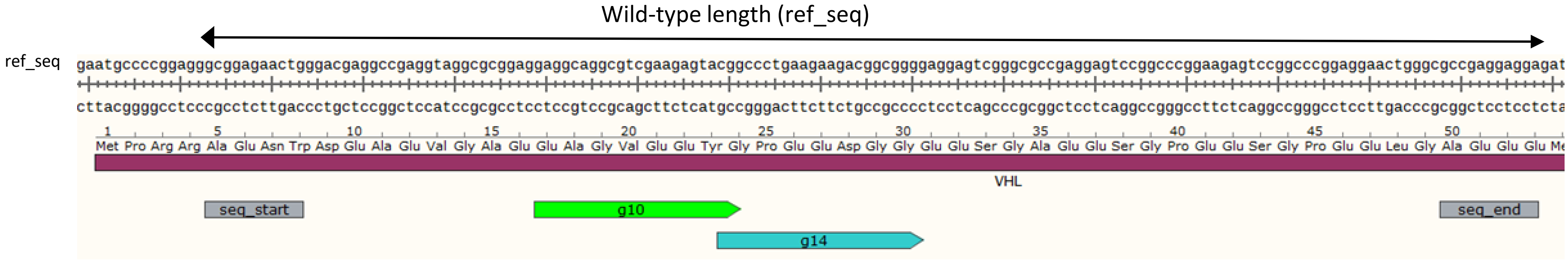
CRIS.py input



```
def get_parameters():
    #Note to user- Change text inside of quote marks '' for your experiment
①   ID = 'Locus_1'
②   ref_seq = str.upper('agggaatgccccggagggcggagaactgggacgaggccgaggtaggcgcggaggaggcaggcgtcgaagagtacggccctgaagaagacggcgggga
③   seq_start = str.upper('GCGGAGAACTG')
④   seq_end = str.upper('GCCGAGGAGGA')
⑤   fastq_files = '*.fastq'
⑥   test_list = [
                str('g10'),    str.upper('GAGGCAGGCGTCGAAGAGTACGG'),
                str('g14'),    str.upper('CGGCCCTGAAGAAGACGGCGGGG')]

    return ID,ref_seq,seq_start,seq_end,fastq_files,test_list
```

CRIS.py Output – Master Summary (.CSV)



Fastq File Name

User inputted name

# of reads

Test_Sequences

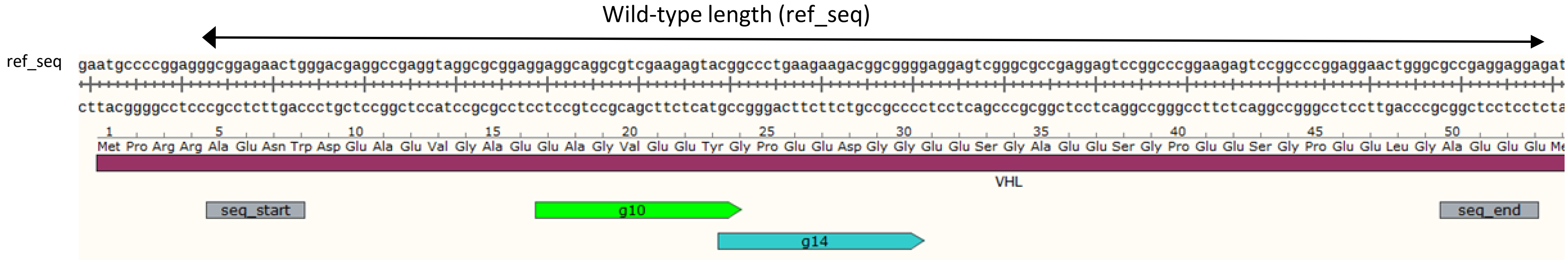
% of reads not WT_length

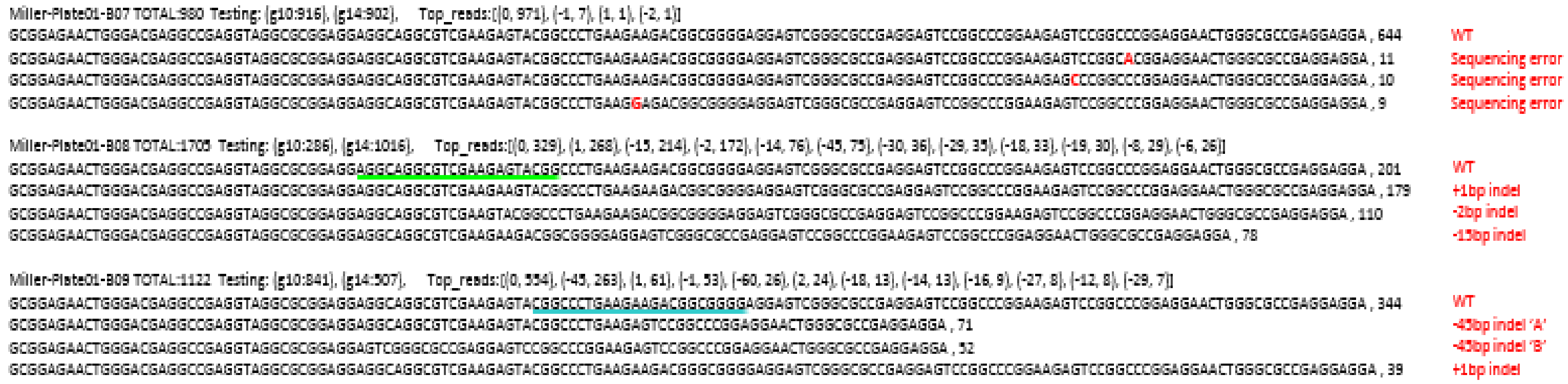
Top indels and frequencies

QC Checks

Name	Sample	Total	g10	g14	Total_indel	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)	SNP_test	raw_wt_counter
Miller-Plate01-B07	Neg	980	916 (93.5%)	902 (92.0%)	0.9	0	971 (99.1%)	-1	7 (0.7%)	1	1 (0.1%)	-2	1 (0.1%)	0.99	975 (1.1)
Miller-Plate01-B08	g10 + Cas9	1705	286 (16.8%)	1016 (59.6%)	80.7	0	329 (19.3%)	1	268 (15.7%)	-15	214 (12.6%)	-2	172 (10.1%)	0.95	301 (1.1)
Miller-Plate01-B09	g14 + Cas9	1122	841 (75.0%)	507 (45.2%)	50.6	0	554 (49.4%)	-45	263 (23.4%)	1	61 (5.4%)	-1	53 (4.7%)	0.98	894 (1.1)

CRIS.py Output – Sequence File (.TXT)





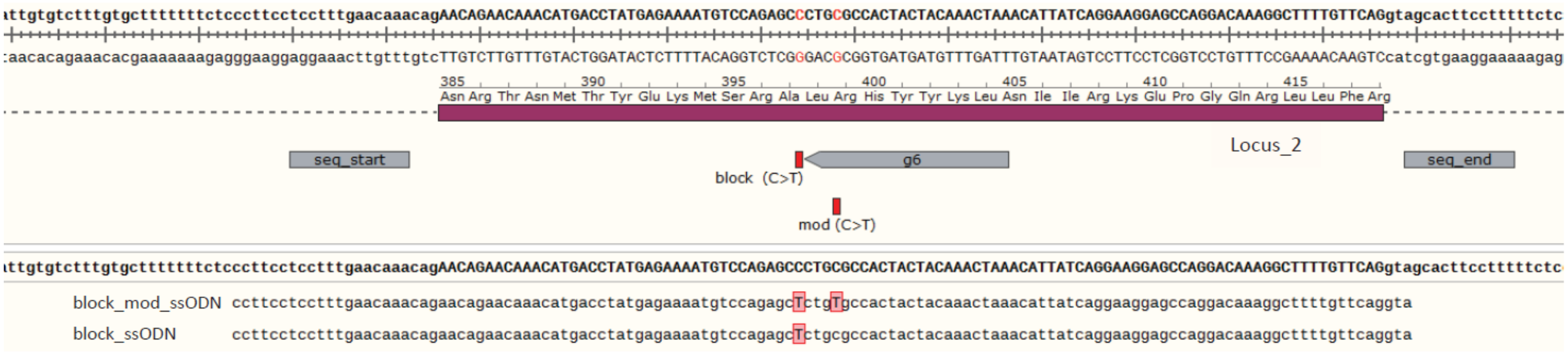
```
Miller-Plate01-B07 TOTAL:980  Testing: [g10:916], [g14:902],     Top_reads:[(0, 971), (-1, 7), (1, 1), (-2, 1)]
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 644      WT
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCACGGAGGAACTGGGCGCCGAGGAGGA , 11       Sequencing error
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGCCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 10       Sequencing error
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGGAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 9        Sequencing error

Miller-Plate01-B08 TOTAL:1705  Testing: [g10:286], [g14:1016],     Top_reads:[(0, 329), (1, 268), (-15, 214), (-2, 172), (-14, 76), (-45, 75), (-30, 36), (-29, 35), (-18, 33), (-19, 30), (-8, 29), (-6, 26)]
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 201      WT
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 179     +1bp indel
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 110        -2bp indel
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 78                      -15bp indel

Miller-Plate01-B09 TOTAL:1122  Testing: [g10:841], [g14:507],     Top_reads:[(0, 554), (-45, 263), (1, 61), (-1, 53), (-60, 26), (2, 24), (-18, 13), (-14, 13), (-16, 9), (-27, 8), (-12, 8), (-29, 7)]
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 344      WT
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 71                                                          -45bp indel 'A'
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 52                                                          -45bp indel 'B'
GCGGAGAACTGGGACGAGGCCGAGGTAGGCGCGGAGGAGGCAGGCGTCGAAGAGTACGGCCCTGAAGAAGACGGGCGGGGAGGAGTCGGGCGCCGAGGAGTCCGGCCCGGAAGAGTCCGGCCCGGAGGAACTGGGCGCCGAGGAGGA , 39     +1bp indel
```

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

CRIS.py allows multiple HDR events to be analyzed concurrently

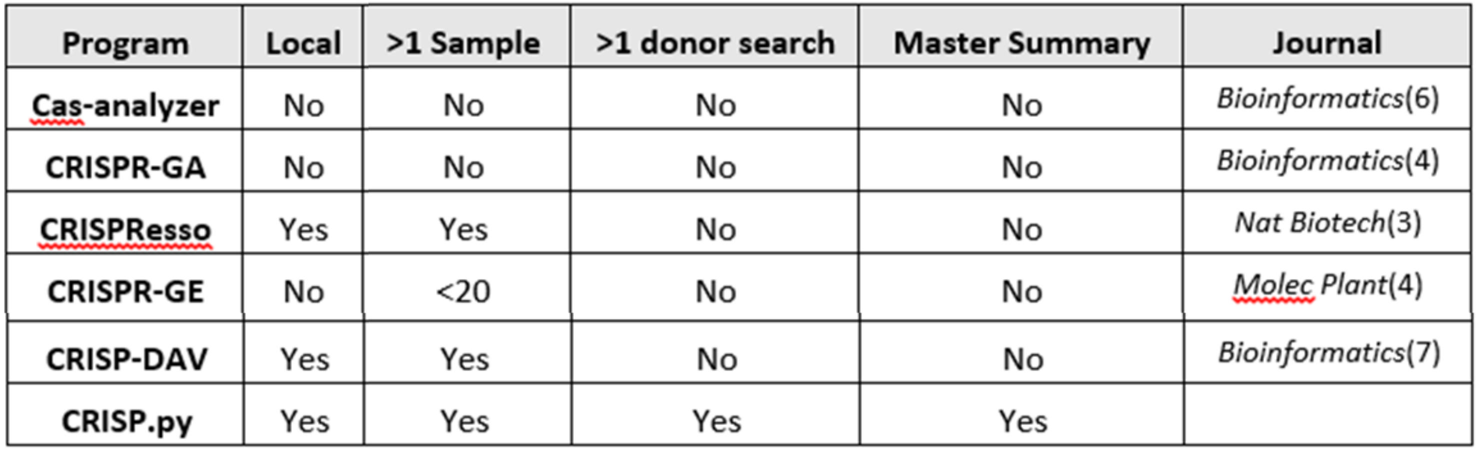


Name	Sample	Total	g6	block_mod_ssODN	block_ssODN	Total_indel	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)	SNP_test	raw_wt_counter
Miller-Plate07-H09	Neg	1760	1743 (99.0%)	0 (0.0%)	0 (0.0%)	0.1	0	1758 (99.9%)	-1	2 (0.1%)	NA		NA		1	1775 (1.0)
Miller-Plate07-H10	gRNA+ssODNs	1998	443 (22.2%)	267 (13.4%)	208 (10.4%)	52.8	0	944 (47.2%)	-7	187 (9.4%)	-2	118 (5.9%)	-6	85 (4.3%)	0.99	448 (1.0)

```
Miller-Plate07-H09 TOTAL:1760  Testing: (g6:1743), (block_mod_ssODN:0), (block_ssODN:0),     Top_reads:[(0, 1758), (-1, 2)]
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 1677  WT
TCCTTTGAACAAACGGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 3     Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAGGGCTTTTGTTCAGGTAGCACTTCCTTT , 3     Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGTCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2     Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCGCTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2     Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTATTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2     Sequencing error

Miller-Plate07-H10 TOTAL:1998  Testing: (g6:443), (block_mod_ssODN:267), (block_ssODN:208),     Top_reads:[(0, 944), (-7, 187), (-2, 118), (-6, 85), (-8, 83), (-1, 76), (-5, 62), (-11, 51), (-4, 45), (1, 41), (-13, 41), (-3, 29)]
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 424     WT
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCTCTGTGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 245     Block_mod
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCTCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 189     Block
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 168          -7bp indel
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 78       -2bp indel
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 66          -6bp indel
```

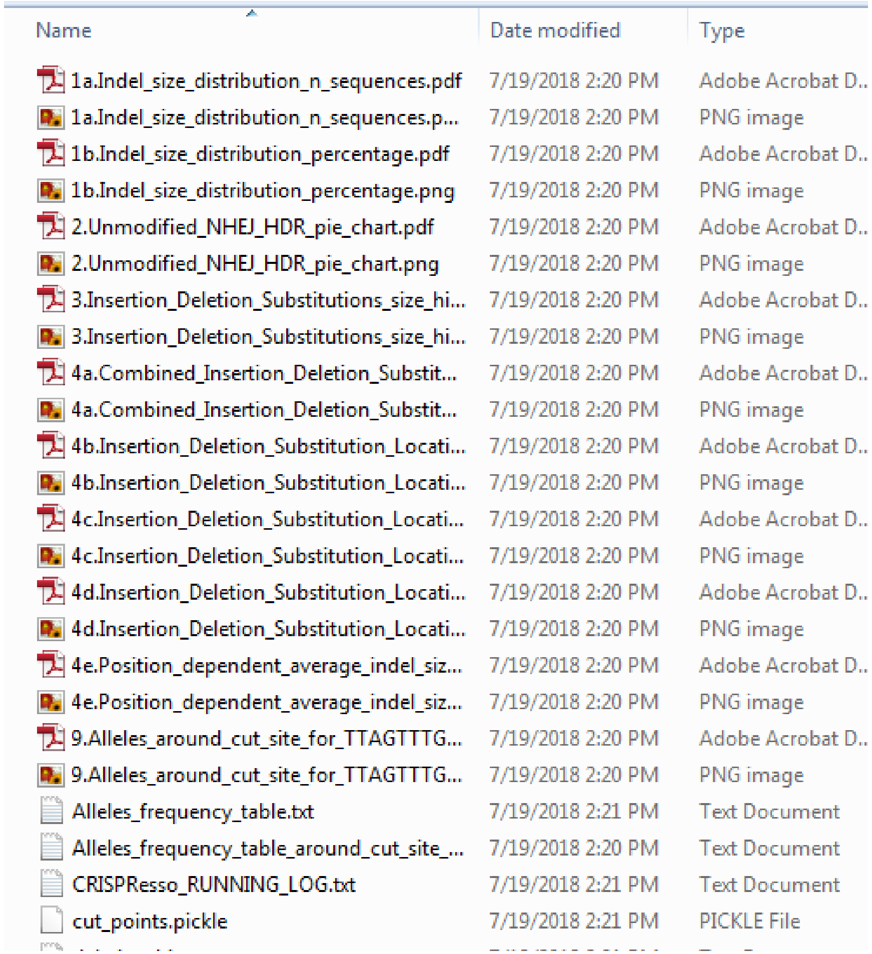
How does CRIS.py compare?



Program	Local	>1 Sample	>1 donor search	Master Summary	Journal
Cas-analyzer	No	No	No	No	*Bioinformatics*(6)
CRISPR-GA	No	No	No	No	*Bioinformatics*(4)
CRISPResso	Yes	Yes	No	No	*Nat Biotech*(3)
CRISPR-GE	No	<20	No	No	*Molec Plant*(4)
CRISP-DAV	Yes	Yes	No	No	*Bioinformatics*(7)
CRISP.py	Yes	Yes	Yes	Yes	

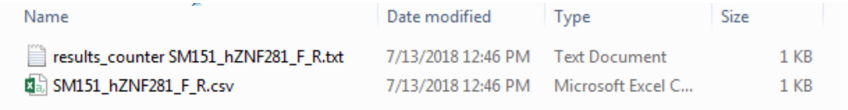
Output comparison

CRISPresso



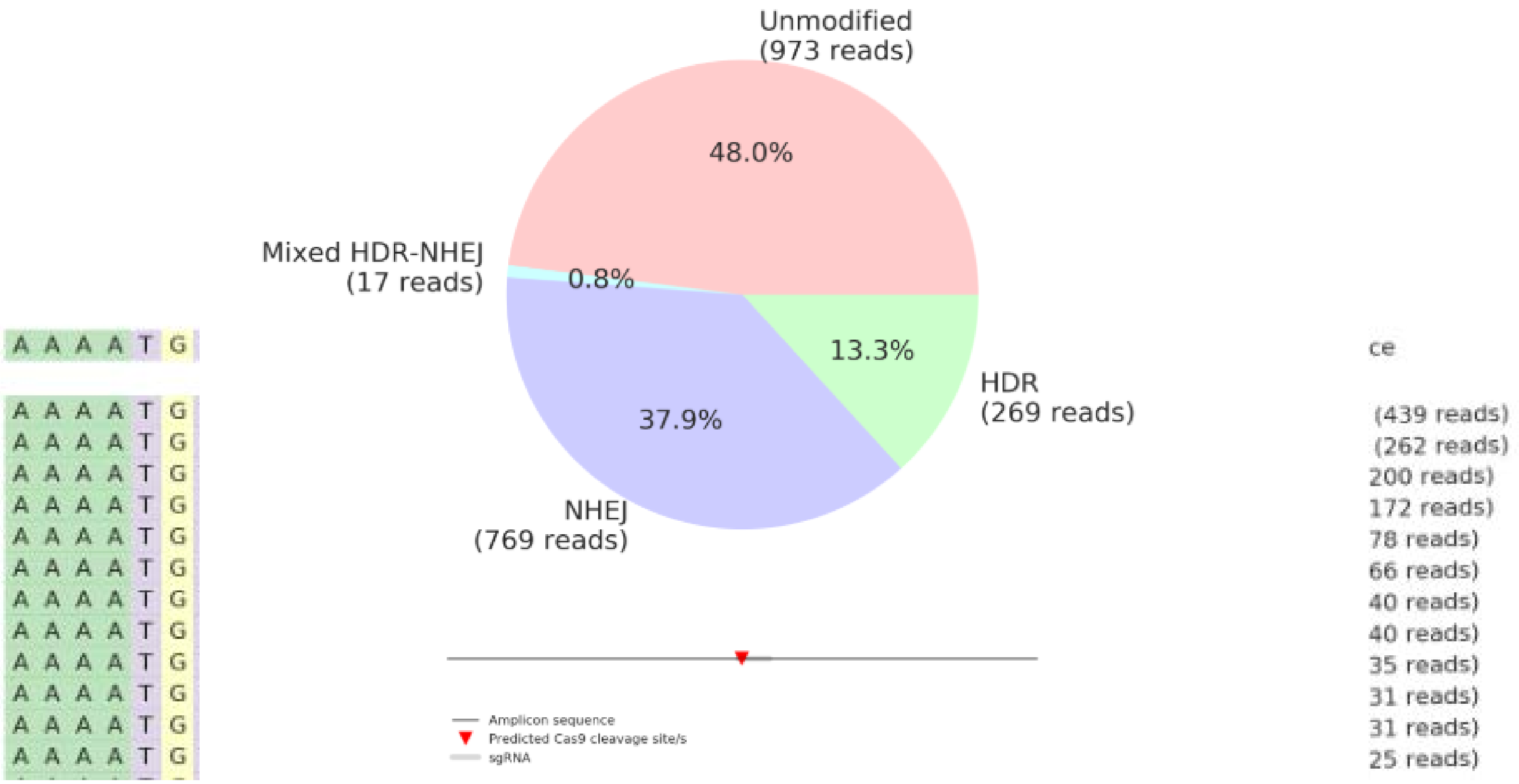
Name	Date modified	Type
1a.Indel_size_distribution_n_sequences.pdf	7/19/2018 2:20 PM	Adobe Acrobat D..
1a.Indel_size_distribution_n_sequences.p...	7/19/2018 2:20 PM	PNG image
1b.Indel_size_distribution_percentage.pdf	7/19/2018 2:20 PM	Adobe Acrobat D..
1b.Indel_size_distribution_percentage.png	7/19/2018 2:20 PM	PNG image
2.Unmodified_NHEJ_HDR_pie_chart.pdf	7/19/2018 2:20 PM	Adobe Acrobat D..
2.Unmodified_NHEJ_HDR_pie_chart.png	7/19/2018 2:20 PM	PNG image
3.Insertion_Deletion_Substitutions_size_hi...	7/19/2018 2:20 PM	Adobe Acrobat D..
3.Insertion_Deletion_Substitutions_size_hi...	7/19/2018 2:20 PM	PNG image
4a.Combined_Insertion_Deletion_Substit...	7/19/2018 2:20 PM	Adobe Acrobat D..
4a.Combined_Insertion_Deletion_Substit...	7/19/2018 2:20 PM	PNG image
4b.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	Adobe Acrobat D..
4b.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	PNG image
4c.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	Adobe Acrobat D..
4c.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	PNG image
4d.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	Adobe Acrobat D..
4d.Insertion_Deletion_Substitution_Locati...	7/19/2018 2:20 PM	PNG image
4e.Position_dependent_average_indel_siz...	7/19/2018 2:20 PM	Adobe Acrobat D..
4e.Position_dependent_average_indel_siz...	7/19/2018 2:20 PM	PNG image
9.Alleles_around_cut_site_for_TTAGTTTG...	7/19/2018 2:20 PM	Adobe Acrobat D..
9.Alleles_around_cut_site_for_TTAGTTTG...	7/19/2018 2:20 PM	PNG image
Alleles_frequency_table.txt	7/19/2018 2:21 PM	Text Document
Alleles_frequency_table_around_cut_site_...	7/19/2018 2:20 PM	Text Document
CRISPResso_RUNNING_LOG.txt	7/19/2018 2:21 PM	Text Document
cut_points.pickle	7/19/2018 2:21 PM	PICKLE File

CRIS.py

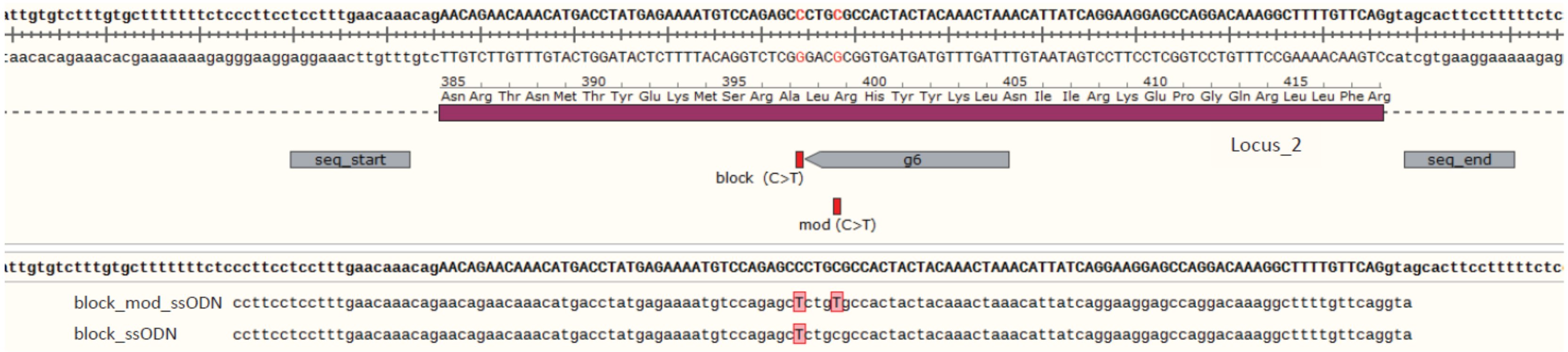


Name	Date modified	Type	Size
results_counter SM151_hZNF281_F_R.txt	7/13/2018 12:46 PM	Text Document	1 KB
SM151_hZNF281_F_R.csv	7/13/2018 12:46 PM	Microsoft Excel C...	1 KB

CRISPResso


Unmodified
(973 reads)
48.0%
Mixed HDR-NHEJ
(17 reads)
0.8%
13.3%
HDR
(269 reads)
37.9%
NHEJ
(769 reads)
A A A A T G
ce
(439 reads)
(262 reads)
200 reads)
172 reads)
78 reads)
66 reads)
40 reads)
40 reads)
35 reads)
31 reads)
31 reads)
25 reads)
Amplicon sequence
Predicted Cas9 cleavage site/s
sgRNA

CRIS.py



```
ttgtgtctttgtgctttttttctcccttcctcctttgaacaaacagAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGgtagcacttcctttttctc
aacacagaaacacgaaaaaaagagggaaggaggaaacttgtttgtcTTGTCTTGTTTGTACTGGATACTCTTTTACAGGTCTCGGGACGCGGTGATGATGTTTGATTTGTAATAGTCCTTCCTCGGTCCTGTTTCCGAAAACAAGTCcatcgtgaaggaaaaagag

ttgtgtctttgtgctttttttctcccttcctcctttgaacaaacagAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGgtagcacttcctttttctc
      block_mod_ssODN  ccttcctcctttgaacaaacagaacagaacaaacatgacctatgagaaaatgtccagagcTctgTgccactactacaaactaaacattatcaggaaggagccaggacaaaggcttttgttcaggta
      block_ssODN      ccttcctcctttgaacaaacagaacagaacaaacatgacctatgagaaaatgtccagagcTctgcgccactactacaaactaaacattatcaggaaggagccaggacaaaggcttttgttcaggta
```

Name	Sample	Total	g6	block_mod_ssODN	block_ssODN	Total_indel	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)	SNP_test	raw_wt_counter
Miller-Plate07-H09	Neg	1760	1743 (99.0%)	0 (0.0%)	0 (0.0%)	0.1	0	1758 (99.9%)	-1	2 (0.1%)	NA		NA		1	1775 (1.0)
Miller-Plate07-H10	gRNA+ssODNs	1998	443 (22.2%)	267 (13.4%)	208 (10.4%)	52.8	0	944 (47.2%)	-7	187 (9.4%)	-2	118 (5.9%)	-6	85 (4.3%)	0.99	448 (1.0)

```
Miller-Plate07-H09 TOTAL:1760  Testing: (g6:1743), (block_mod_ssODN:0), (block_ssODN:0),     Top_reads:[(0, 1758), (-1, 2)]
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 1677   WT
TCCTTTGAACAAACGGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 3      Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAGGGCTTTTGTTCAGGTAGCACTTCCTTT , 3      Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGTCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2      Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCGCTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2      Sequencing error
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTATTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 2      Sequencing error

Miller-Plate07-H10 TOTAL:1998  Testing: (g6:443), (block_mod_ssODN:267), (block_ssODN:208),     Top_reads:[(0, 944), (-7, 187), (-2, 118), (-6, 85), (-8, 83), (-1, 76), (-5, 62), (-11, 51), (-4, 45), (1, 41), (-13, 41), (-3, 29)]
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 424     WT
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCTCTGTGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 245     Block_mod
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCTCTGCGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 189     Block
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 168         -7bp indel
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCTGCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 78        -2bp indel
TCCTTTGAACAAACAGAACAGAACAAACATGACCTATGAGAAAATGTCCAGAGCCCACTACTACAAACTAAACATTATCAGGAAGGAGCCAGGACAAAGGCTTTTGTTCAGGTAGCACTTCCTTT , 66          -6bp indel
```

Overall compare

	Total reads	WT	NHEJ	HDR #1	HDR #2	HDR-NHEJ
CRIS.py	1963	22.4%	51.9%	13.4%	10.4%	0.6%
CRISPresso	2028	48.0%	37.9%	13.3%		0.8%

CRIS.py enables cell ploidy determinations
7 Copies in N2A cells!!!

- ☐ 2C4 (-1)
- ☐ 2C4 (-2)
- ☐ 2D2 (1)
- ☐ 2D2 (-5)
- ☐ 2D2 (1)
- ☐ 2D2 (-1)
- ☐ 2D2 (-1)
- ☐ 2D2 (-2)
- ☑ 2G6 (-5) - 252 ex...
- ☑ 2G6 (14) - 177 ex...
- ☑ 2G6 (1) - 139 exa...
- ☑ 2G6(-4) - 139 exa...
- ☑ 2G6 (-4) - 129 ex...
- ☑ 2G6 (1) - 128 exa...

```
ctggatgctgtggagcttcattgaaagatggctactggccttggcgtcctggt cttgggctctctgccgcatctctctg
gacctacgacacctcgaagtaactttctaccgatgaccggaaccgcaggacca gaacccgagagacggcgtagagagac
   5           10          15          20           25          30
Trp Met Leu Trp Ser Phe Ile Glu Arg Trp Leu Leu Ala Leu Ala Ser Trp Ser Trp Ala Leu Cys Arg Ile Ser Leu
```

```
Original Sequence              ctggatgctgtggagcttcattgaaagatggctactggccttggcgtcctggt cttgggctctctgccgcatctctctg
2G6 (-5) - 252 exact match     CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGGCGTCCT--- --TGGGCTCTCTGCCGCATCTCTCTG
2G6 (14) - 177 exact match     CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGG--------- -----GCTCTCTGCCGCATCTCTCTG
2G6 (1) - 139 exact match      CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGGCGTCCTGGTGCTTGGGCTCTCTGCCGCATCTCTCTG
2G6(-4) - 139 exact match      CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGGCGTCC---- CTTGGGCTCTCTGCCGCATCTCTCTG
2G6 (-4) - 129 exact match     CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGGCGTC----T CTTGGGCTCTCTGCCGCATCTCTCTG
2G6 (1) - 128 exact match      CTGGATGCTGTGGAGCTTCATTGAAAGATGGCTACTGGCCTTGGCGTCCTGGTTCTTGGGCTCTCTGCCGCATCTCTCTG
```

sample	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)
2G6	-4	356 (27.9%)	1	352 (27.6%)	-5	332 (26.1%)	-14	229 (18.0%)

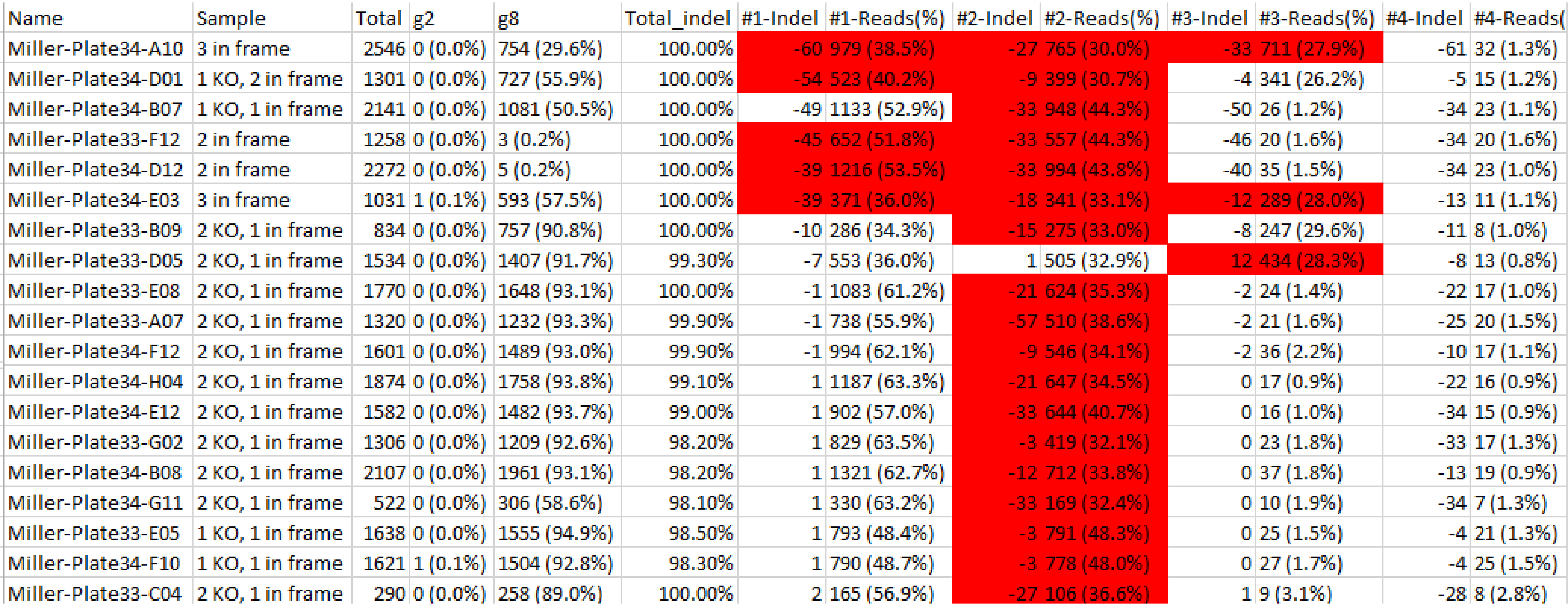
St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

CRIS.py allows easy identification of Heterogeneity within a cell population

	Total	WT_g15	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)
Diploid	149	0 (0.0%)	1	77 (51.7%)	-10	71 (47.7%)	-21	1 (0.7%)	NA	
	622	0 (0.0%)	-4	354 (56.9%)	-10	265 (42.6%)	-11	2 (0.3%)	-5	1 (0.2%)
	748	0 (0.0%)	-4	386 (51.6%)	-2	360 (48.1%)	-3	2 (0.3%)	NA	
	944	0 (0.0%)	-17	527 (55.8%)	-1	413 (43.8%)	-2	2 (0.2%)	-18	2 (0.2%)
	1520	0 (0.0%)	-10	779 (51.2%)	-1	732 (48.2%)	-11	5 (0.3%)	-2	4 (0.3%)
	2419	0 (0.0%)	-2	1258 (52.0%)	-1	1151 (47.6%)	-3	10 (0.4%)	NA	
	2656	0 (0.0%)	-20	1520 (57.2%)	-1	1126 (42.4%)	-21	6 (0.2%)	-2	4 (0.2%)
	2884	3 (0.1%)	-17	1518 (52.6%)	-10	1341 (46.5%)	-45	7 (0.2%)	-11	7 (0.2%)
	2657	10 (0.4%)	-22	1336 (50.3%)	-10	1297 (48.8%)	0	10 (0.4%)	-11	8 (0.3%)
Triploid	765	0 (0.0%)	1	531 (69.4%)	-1	231 (30.2%)	0	2 (0.3%)	-2	1 (0.1%)
	782	0 (0.0%)	-2	467 (59.7%)	-1	311 (39.8%)	-3	4 (0.5%)	NA	
	2131	0 (0.0%)	-7	754 (35.4%)	-1	733 (34.4%)	-2	635 (29.8%)	-3	7 (0.3%)
	2721	0 (0.0%)	-17	1051 (38.6%)	-4	908 (33.4%)	-1	756 (27.8%)	-5	3 (0.1%)
Tetraploid	253	0 (0.0%)	-28	107 (42.3%)	-1	74 (29.2%)	1	71 (28.1%)	-2	1 (0.4%)
	4471	0 (0.0%)	-17	3319 (74.2%)	-10	1134 (25.4%)	-18	10 (0.2%)	-11	4 (0.1%)
Homozygous clones	225	0 (0.0%)	-1	225 (100.0%)	NA		NA		NA	
	265	0 (0.0%)	-1	263 (99.2%)	-2	2 (0.8%)	NA		NA	
	484	0 (0.0%)	-4	483 (99.8%)	-5	1 (0.2%)	NA		NA	
	521	0 (0.0%)	-2	520 (99.8%)	-3	1 (0.2%)	NA		NA	
	1088	0 (0.0%)	-4	1086 (99.8%)	-5	2 (0.2%)	NA		NA	
	1447	0 (0.0%)	-17	1440 (99.5%)	-18	6 (0.4%)	-19	1 (0.1%)	NA	
	2729	25 (0.9%)	-1	2626 (96.2%)	6	33 (1.2%)	-2	33 (1.2%)	0	26 (1.0%)
	199	0 (0.0%)	-21	129 (64.8%)	-19	67 (33.7%)	-23	1 (0.5%)	-22	1 (0.5%)
	298	0 (0.0%)	-24	298 (100.0%)	NA		NA		NA	

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

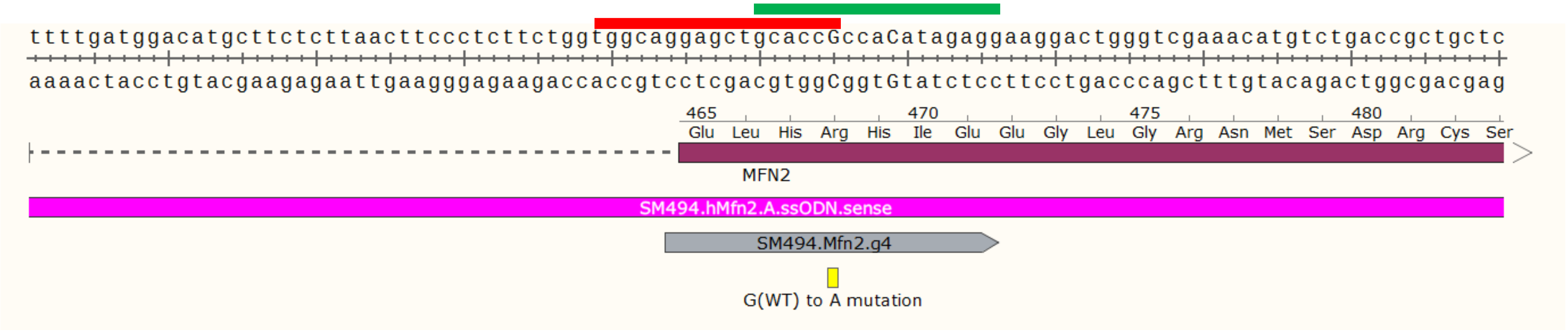
CRIS.py Master File is sortable and allows screening of hundreds of samples



Name	Sample	Total	g2	g8	Total_indel	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(
Miller-Plate34-A10	3 in frame	2546	0 (0.0%)	754 (29.6%)	100.00%	-60	979 (38.5%)	-27	765 (30.0%)	-33	711 (27.9%)	-61	32 (1.3%)
Miller-Plate34-D01	1 KO, 2 in frame	1301	0 (0.0%)	727 (55.9%)	100.00%	-54	523 (40.2%)	-9	399 (30.7%)	-4	341 (26.2%)	-5	15 (1.2%)
Miller-Plate34-B07	1 KO, 1 in frame	2141	0 (0.0%)	1081 (50.5%)	100.00%	-49	1133 (52.9%)	-33	948 (44.3%)	-50	26 (1.2%)	-34	23 (1.1%)
Miller-Plate33-F12	2 in frame	1258	0 (0.0%)	3 (0.2%)	100.00%	-45	652 (51.8%)	-33	557 (44.3%)	-46	20 (1.6%)	-34	20 (1.6%)
Miller-Plate34-D12	2 in frame	2272	0 (0.0%)	5 (0.2%)	100.00%	-39	1216 (53.5%)	-33	994 (43.8%)	-40	35 (1.5%)	-34	23 (1.0%)
Miller-Plate34-E03	3 in frame	1031	1 (0.1%)	593 (57.5%)	100.00%	-39	371 (36.0%)	-18	341 (33.1%)	-12	289 (28.0%)	-13	11 (1.1%)
Miller-Plate33-B09	2 KO, 1 in frame	834	0 (0.0%)	757 (90.8%)	100.00%	-10	286 (34.3%)	-15	275 (33.0%)	-8	247 (29.6%)	-11	8 (1.0%)
Miller-Plate33-D05	2 KO, 1 in frame	1534	0 (0.0%)	1407 (91.7%)	99.30%	-7	553 (36.0%)	1	505 (32.9%)	12	434 (28.3%)	-8	13 (0.8%)
Miller-Plate33-E08	2 KO, 1 in frame	1770	0 (0.0%)	1648 (93.1%)	100.00%	-1	1083 (61.2%)	-21	624 (35.3%)	-2	24 (1.4%)	-22	17 (1.0%)
Miller-Plate33-A07	2 KO, 1 in frame	1320	0 (0.0%)	1232 (93.3%)	99.90%	-1	738 (55.9%)	-57	510 (38.6%)	-2	21 (1.6%)	-25	20 (1.5%)
Miller-Plate34-F12	2 KO, 1 in frame	1601	0 (0.0%)	1489 (93.0%)	99.90%	-1	994 (62.1%)	-9	546 (34.1%)	-2	36 (2.2%)	-10	17 (1.1%)
Miller-Plate34-H04	2 KO, 1 in frame	1874	0 (0.0%)	1758 (93.8%)	99.10%	1	1187 (63.3%)	-21	647 (34.5%)	0	17 (0.9%)	-22	16 (0.9%)
Miller-Plate34-E12	2 KO, 1 in frame	1582	0 (0.0%)	1482 (93.7%)	99.00%	1	902 (57.0%)	-33	644 (40.7%)	0	16 (1.0%)	-34	15 (0.9%)
Miller-Plate33-G02	2 KO, 1 in frame	1306	0 (0.0%)	1209 (92.6%)	98.20%	1	829 (63.5%)	-3	419 (32.1%)	0	23 (1.8%)	-33	17 (1.3%)
Miller-Plate34-B08	2 KO, 1 in frame	2107	0 (0.0%)	1961 (93.1%)	98.20%	1	1321 (62.7%)	-12	712 (33.8%)	0	37 (1.8%)	-13	19 (0.9%)
Miller-Plate34-G11	2 KO, 1 in frame	522	0 (0.0%)	306 (58.6%)	98.10%	1	330 (63.2%)	-33	169 (32.4%)	0	10 (1.9%)	-34	7 (1.3%)
Miller-Plate33-E05	1 KO, 1 in frame	1638	0 (0.0%)	1555 (94.9%)	98.50%	1	793 (48.4%)	-3	791 (48.3%)	0	25 (1.5%)	-4	21 (1.3%)
Miller-Plate34-F10	1 KO, 1 in frame	1621	1 (0.1%)	1504 (92.8%)	98.30%	1	790 (48.7%)	-3	778 (48.0%)	0	27 (1.7%)	-4	25 (1.5%)
Miller-Plate33-C04	2 KO, 1 in frame	290	0 (0.0%)	258 (89.0%)	100.00%	2	165 (56.9%)	-27	106 (36.6%)	1	9 (3.1%)	-28	8 (2.8%)

KO of gene is lethal

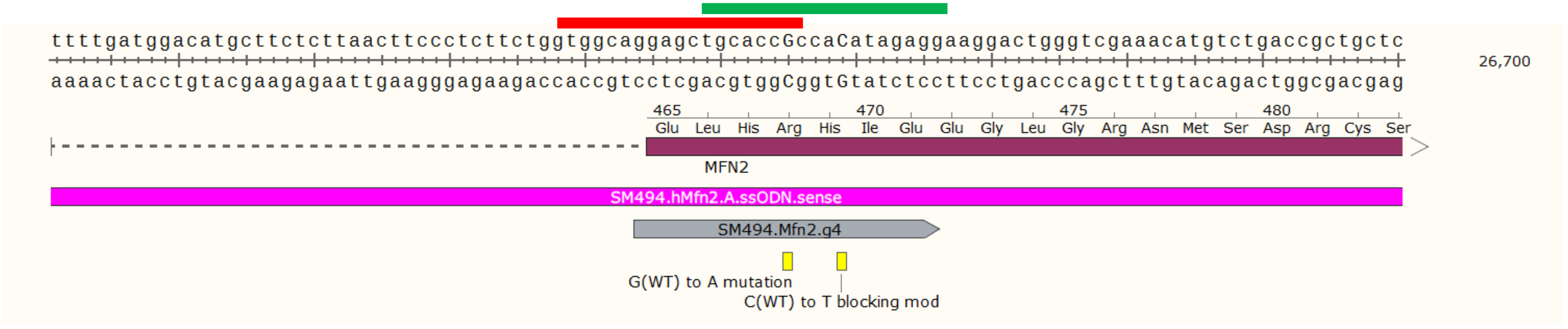
1 blocking modification may not be enough



Total	%Targeted plus NHEJ	%Targeted_precise	
8328	3	0.3	

Only 10% of targeted events contain no indels

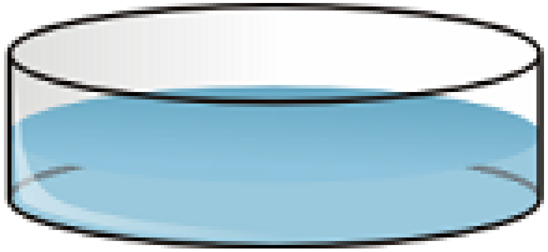
2 blocking modifications



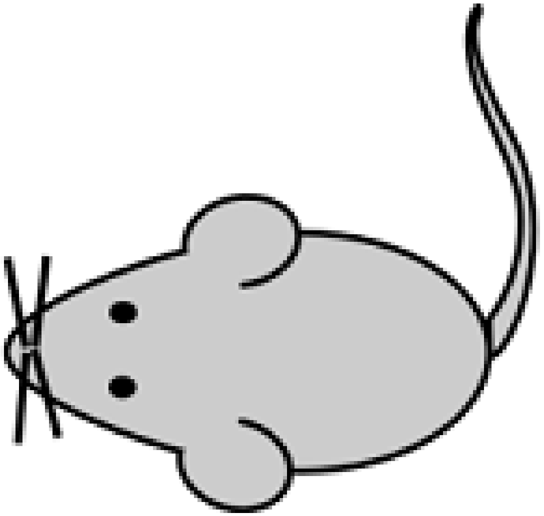
Total	%Target_including NHEJ	%Target_Precise	
7074	7.9	5.8	

74% of targeting events contain no indels

CRIS.py can be used on any DNA sample



In Vitro



In Vivo

Genotyping pups to know which to breed

KNOCKOUTs

Pups	Total	WT_g9	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)
A1	8677	0 (0.0%)	-18	4691 (54.1%)	-3	3915 (45.1%)	-4	25 (0.3%)	-19	23 (0.3%)
A2	8714	0 (0.0%)	-12	4690 (53.8%)	1	3976 (45.6%)	-13	22 (0.3%)	0	19 (0.2%)
A3	9470	4506 (47.6%)	0	4793 (50.6%)	1	3868 (40.8%)	-3	770 (8.1%)	-1	29 (0.3%)
A4	6592	6 (0.1%)	-60	3462 (52.5%)	1	1306 (19.8%)	-25	1024 (15.5%)	-3	742 (11.3%)
A5	7578	7047 (93.0%)	0	7524 (99.3%)	-1	44 (0.6%)	-2	5 (0.1%)	-3	3 (0.0%)
A6	6500	507 (7.8%)	-45	3312 (51.0%)	-2	2043 (31.4%)	-3	577 (8.9%)	0	545 (8.4%)

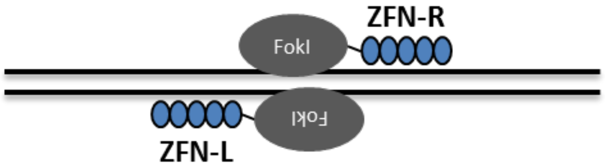
Genotyping pups to know which to breed

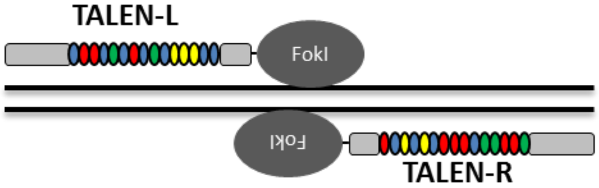
KNOCKINs

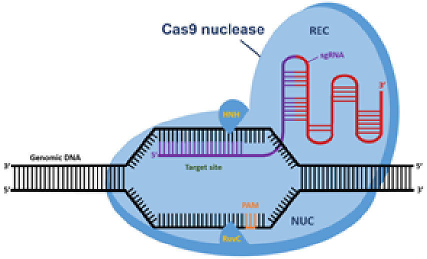
	Pup	Total	Mut and Blocking Mod	#1-Indel	#1-Reads(%)	#2-Indel	#2-Reads(%)	#3-Indel	#3-Reads(%)	#4-Indel	#4-Reads(%)
→	3	3806	3567 (93.7%)	0	3796 (99.7%)	-61	5 (0.1%)	-1	5 (0.1%)	NA	
→	18	3721	2510 (67.5%)	0	2609 (70.1%)	1	1110 (29.8%)	-1	2 (0.1%)	NA	
	38	4920	2554 (51.9%)	0	4916 (99.9%)	-1	4 (0.1%)	NA		NA	
→	60	112	54 (48.2%)	0	99 (88.4%)	2	13 (11.6%)	NA		NA	
→	6	4656	1614 (34.7%)	0	2308 (49.6%)	1	1531 (32.9%)	-12	814 (17.5%)	-1	2 (0.0%)
→	9	5151	1087 (21.1%)	-11	2816 (54.7%)	-7	1169 (22.7%)	0	1162 (22.6%)	-8	2 (0.0%)
→	56	4856	411 (8.5%)	0	2865 (59.0%)	1	1984 (40.9%)	-1	4 (0.1%)	-30	1 (0.0%)
	63	3797	239 (6.3%)	-7	1947 (51.3%)	0	788 (20.8%)	1	551 (14.5%)	19	505 (13.3%)
	43	6200	113 (1.8%)	0	6066 (97.8%)	1	62 (1.0%)	-1	51 (0.8%)	-12	11 (0.2%)
	52	4408	18 (0.4%)	0	4198 (95.2%)	-7	114 (2.6%)	1	33 (0.7%)	19	29 (0.7%)
	42	4660	10 (0.2%)	0	4629 (99.3%)	1	14 (0.3%)	-12	8 (0.2%)	8	5 (0.1%)
	17	3490	2 (0.1%)	0	2462 (70.5%)	8	581 (16.6%)	1	446 (12.8%)	-1	1 (0.0%)
	32	2598	3 (0.1%)	0	2578 (99.2%)	1	6 (0.2%)	-11	6 (0.2%)	-1	3 (0.1%)
	35	4377	3 (0.1%)	0	2105 (48.1%)	-12	1892 (43.2%)	-7	376 (8.6%)	-1	3 (0.1%)
	44	4546	4 (0.1%)	0	4511 (99.2%)	1	24 (0.5%)	-11	5 (0.1%)	-1	4 (0.1%)
	51	3658	3 (0.1%)	0	2441 (66.7%)	-1	1190 (32.5%)	-7	17 (0.5%)	1	5 (0.1%)
	54	3872	4 (0.1%)	0	3863 (99.8%)	-1	9 (0.2%)	NA		NA	

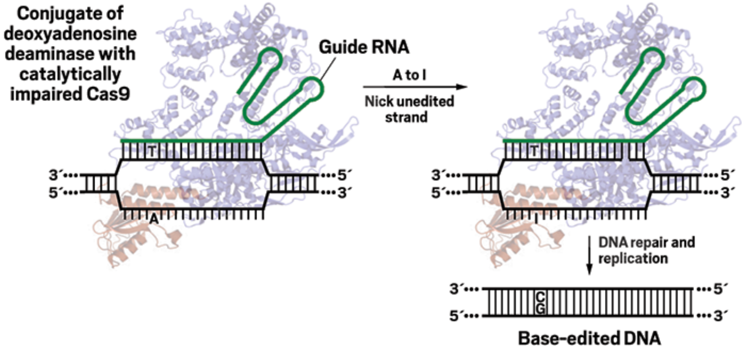
St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

CRIS.py is platform agnostic





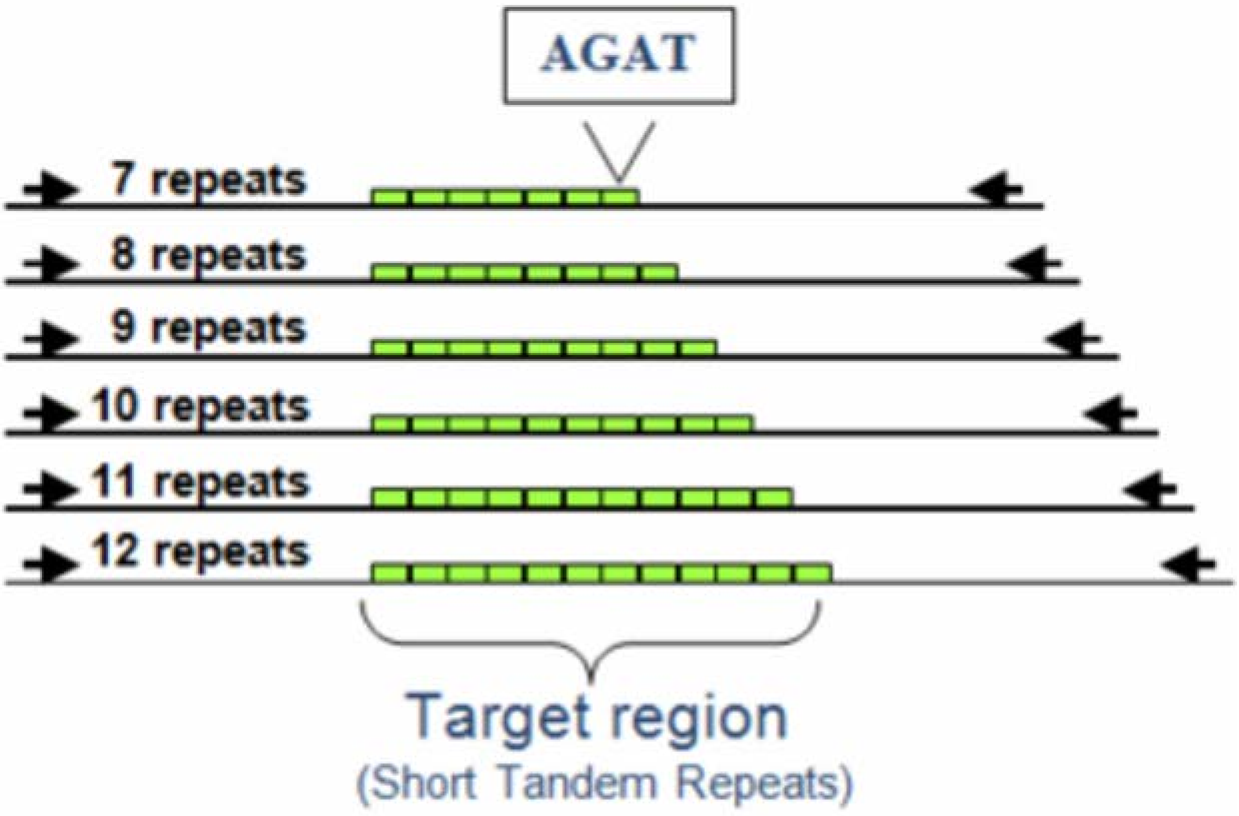




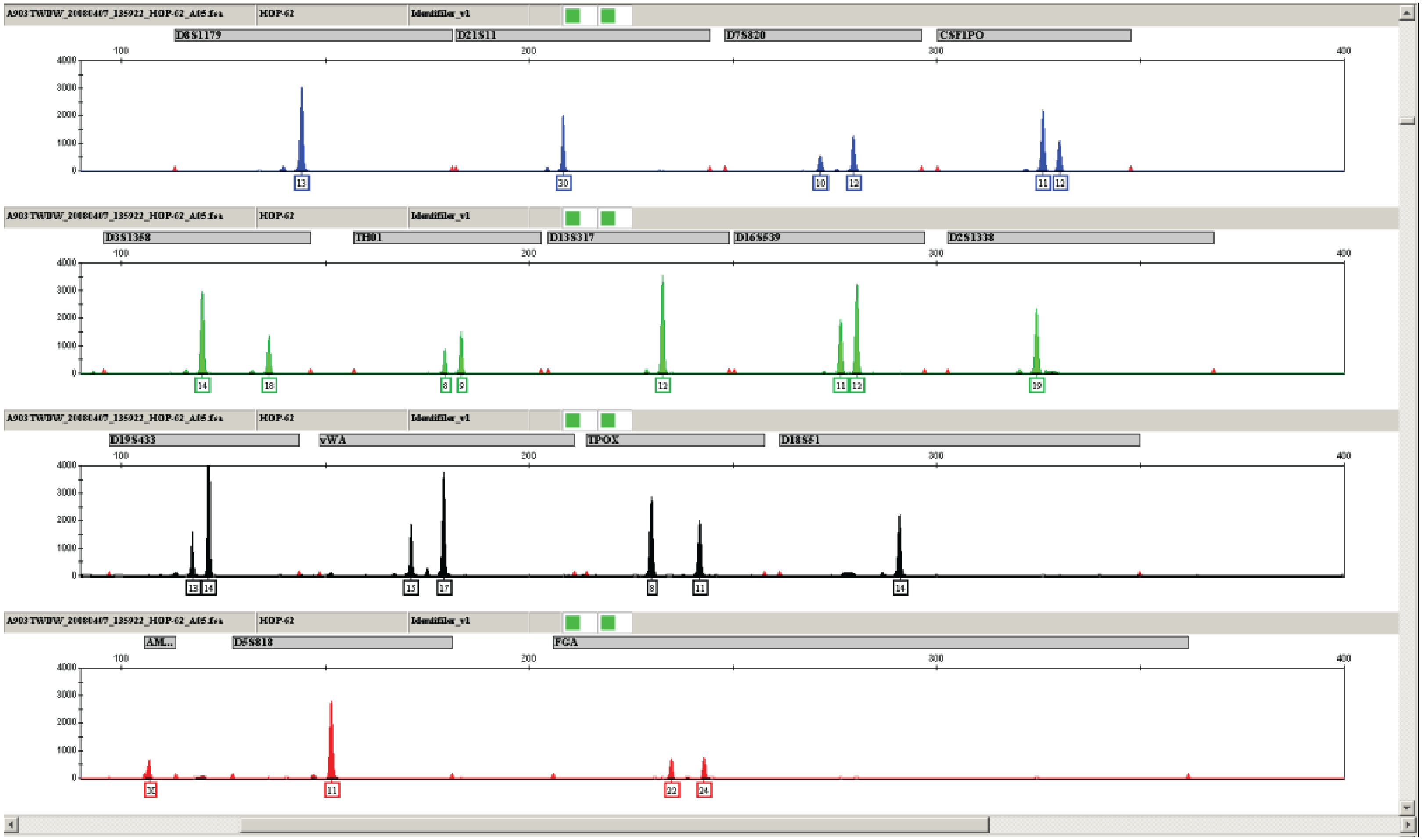
CRIS.py for cell line identity

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

Short Tandem Repeats (STR)



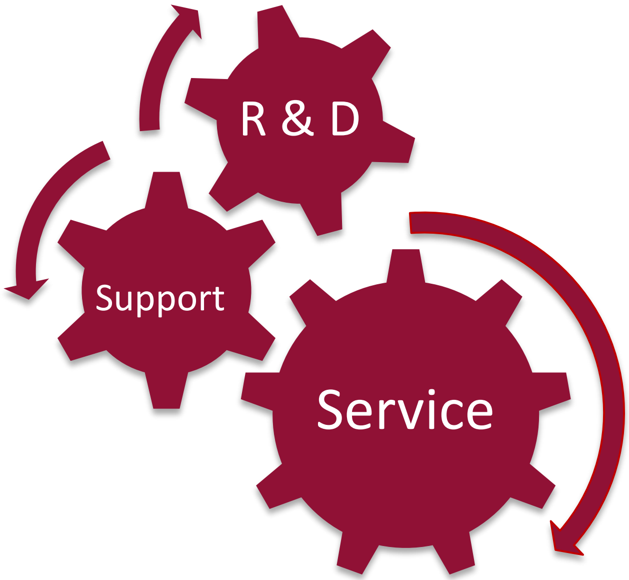
Traditional STR profiling



Identical or Fraternal

	AMELX		CSF1PO		D13S317		D3S1358		D7S820		D8S1179		FGA		TH01		TPOX	
PC	12	13.5	10	12	13	13	16	17	9.25	10.25	13	14	13	14	6	9.75	8	9
HC	12	12	10	12	9	9	15	15	10.25	12.25	14	12	14	16	9.75	9.75	8	12
CC	12	12	10	10	9	13	15	16	9.25	10.25	13	14	14	14	9.75	9.75	12	9
LC	12	13.5	12	12	9	13	15	16	9.25	10.25	13	14	13	16	6	9.75	9	8
DC	12	13.5	12	12	9	13	15	16	9.25	10.25	13	14	13	16	6	9.75	9	8

CAGE Mission



- ✓ Provide gene editing technology for investigator-initiated projects
- ✓ Perform R&D to keep SJCRH investigators at the forefront of emerging technologies
- ✓ Support key initiatives such as the gene therapy and cellular therapy programs

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

St. Jude Children’s Research Hospital Collaborative Research Consortium on Novel Gene therapies for Sickle Cell Disease (SCD)


Normal red blood cells


Sickled red blood cells


David Liu, PhD
The Broad Institute


Keith Joung, MD, PhD
Harvard


Mitch Weiss, MD, PhD
Dept. of Hematology


Shengdar Tsai, PhD
Dept. of Hematology


Jon Tisdale, MD
NHLBI

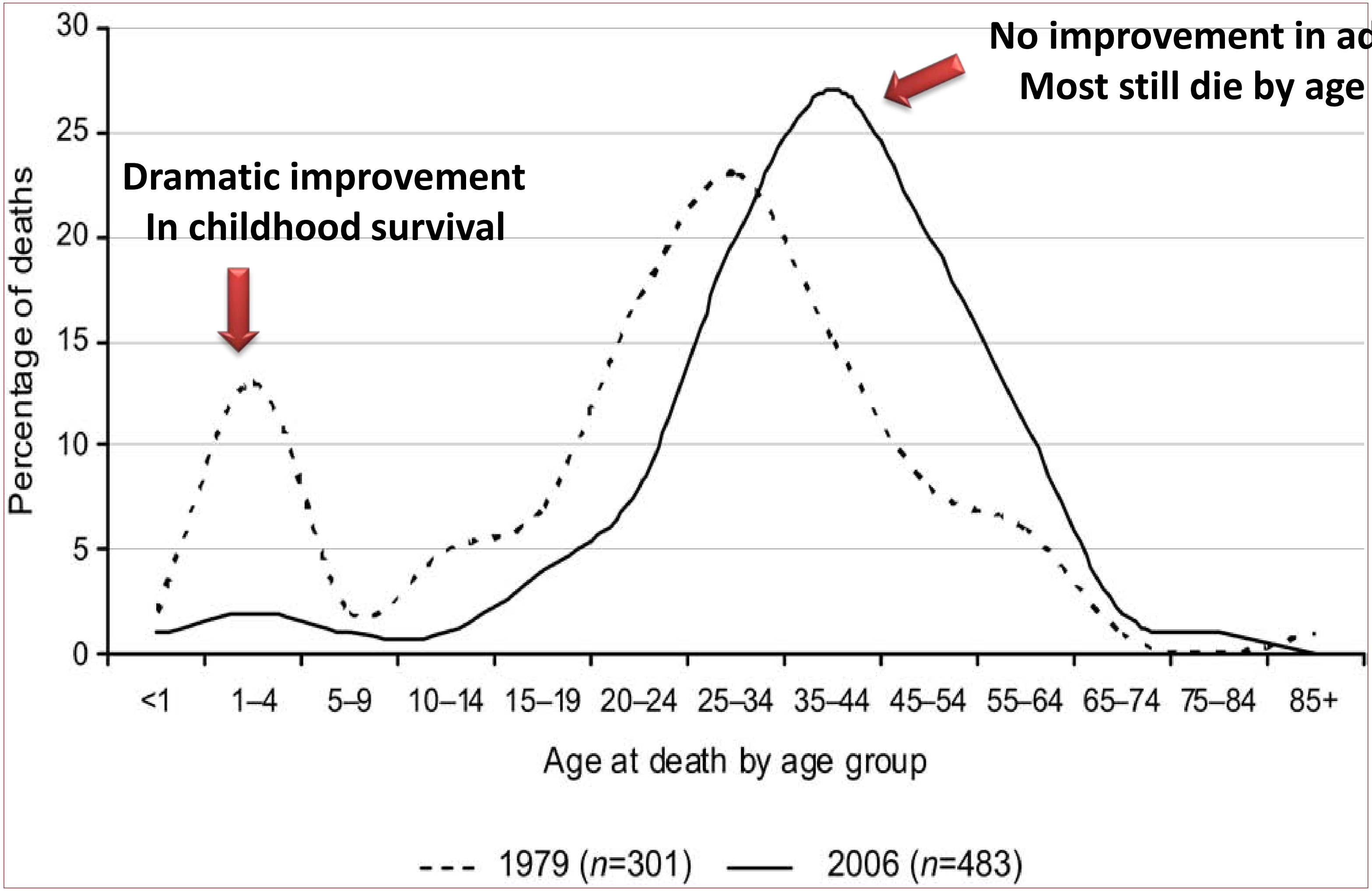

Dan Bauer, MD, PhD
Boston’s Children’s Hospital


Shondra Pruett-MIller, PhD
Dept. of Cell and Molecular Biology


Byoung Ryu, PhD
Dept. of Hematology

St. Jude Children’s Research Hospital
ALSAC • Danny Thomas, Founder

Historical Progress in SCD



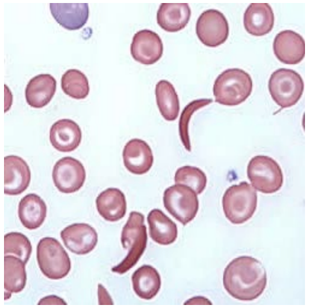
No improvement in adults
Most still die by age 45

Dramatic improvement
In childhood survival

Hassell, K. Am J Prev Med. 2010 Apr;38(4 Suppl):S512-21.

Competition between HDR and NHEJ

Sickle cell disease



DSB+ template

..CCT-GAG-GAG..

CCT-GTG-GAG

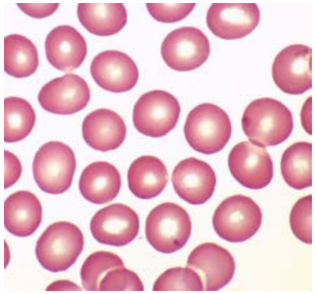
Pro-Val-CTC

β^s

HDR

CCT-GAG-GAG

Pro-Glu-Glu



β^{wt}

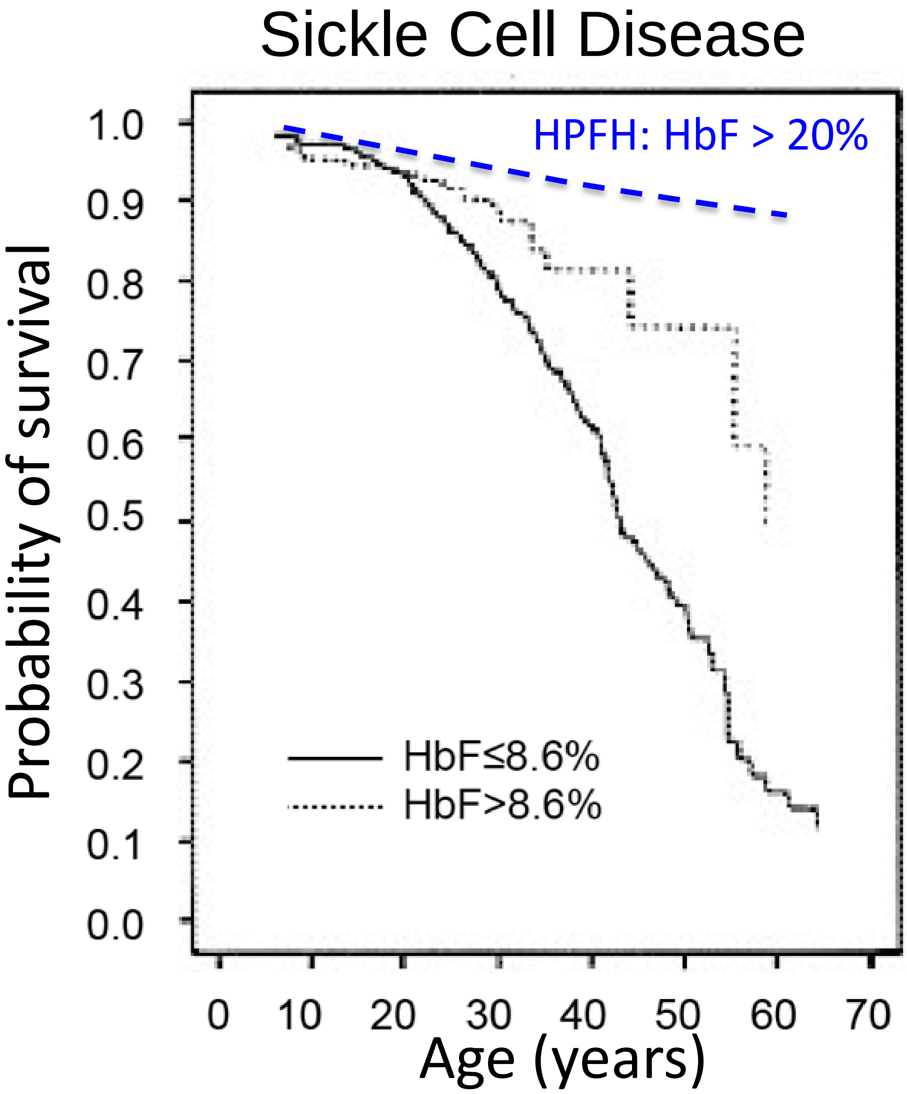
NHEJ

CCT-XXXX…STOP

Pro-

β thal

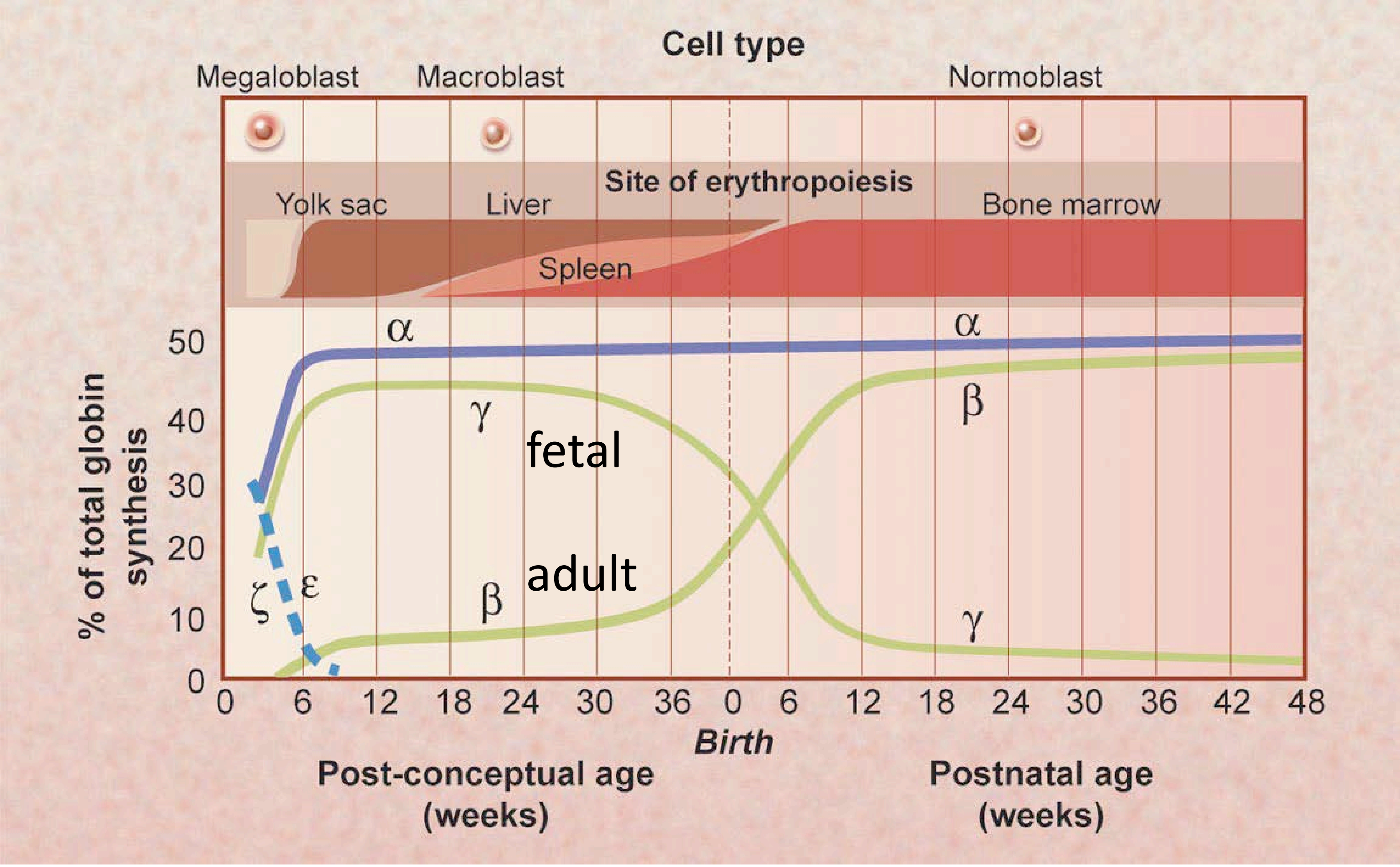
HbF ameliorates SCD



- High HbF improves survival
- Adult HbF levels are genetically determined (HI > 0.9)
- Extreme cases: HPFH (HbF > 20%)
- HPFH is caused by mutations in the extended β globin locus

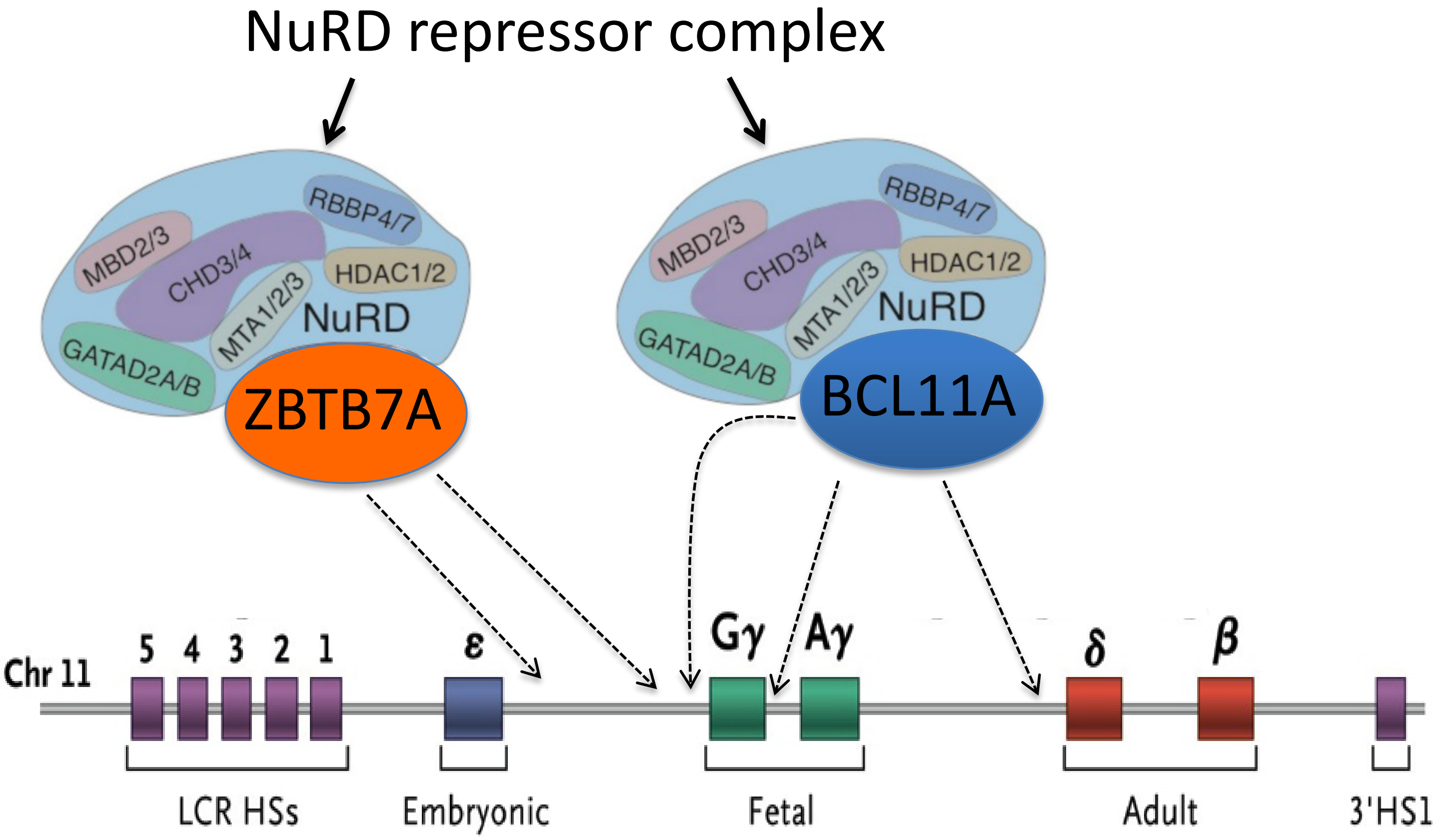
Platt et al. NEJM, 1994.

Hemoglobin switching at birth



Schechter, Blood 2008

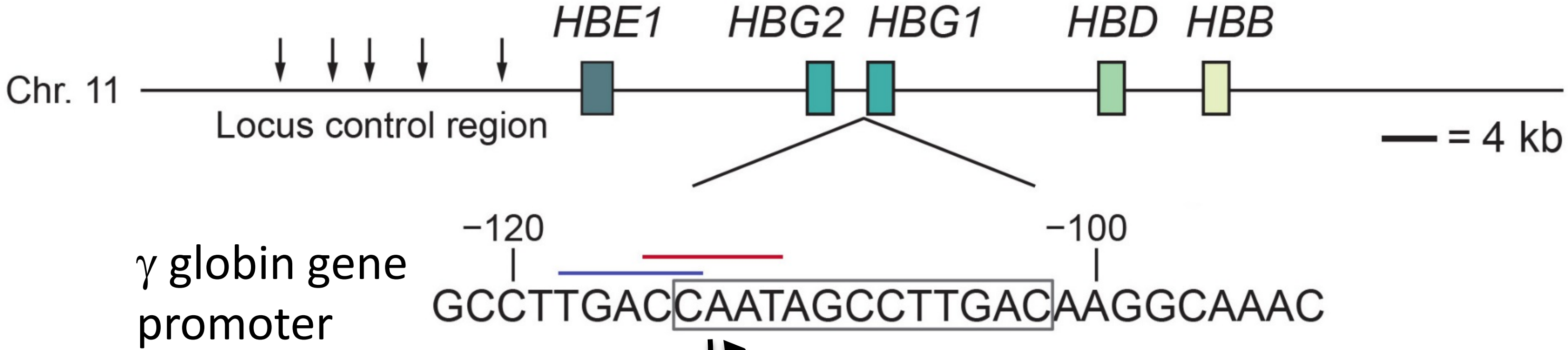
Mechanism of postnatal γ globin repression



NuRD repressor complex

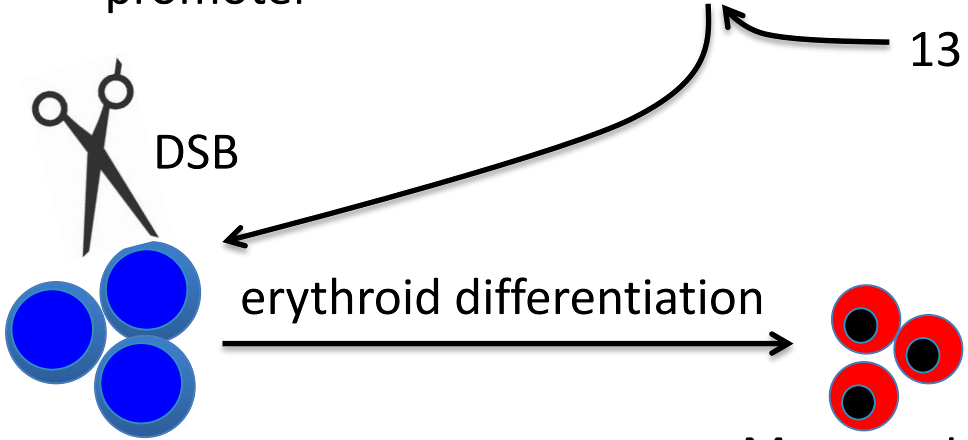
St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

NHEJ to induce HbF for β-hemoglobinopathies



γ globin gene promoter

13 nt deleted region in HPFH



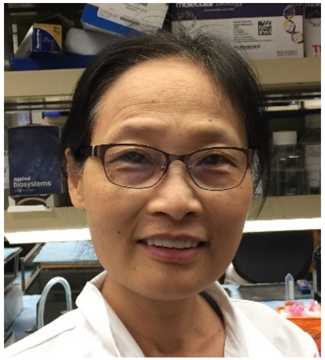
DSB

erythroid differentiation

CD34$^+$ cells (normal or SCD)

Measure HbF

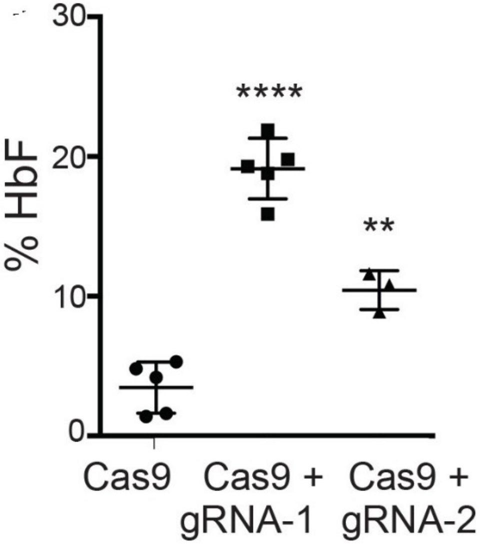
Liz Traxler

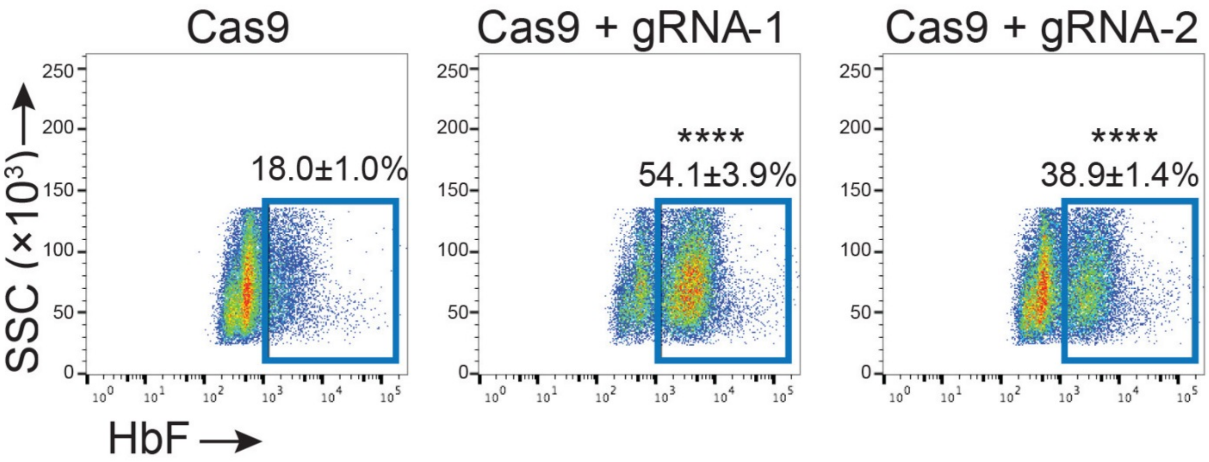


Yu Yao

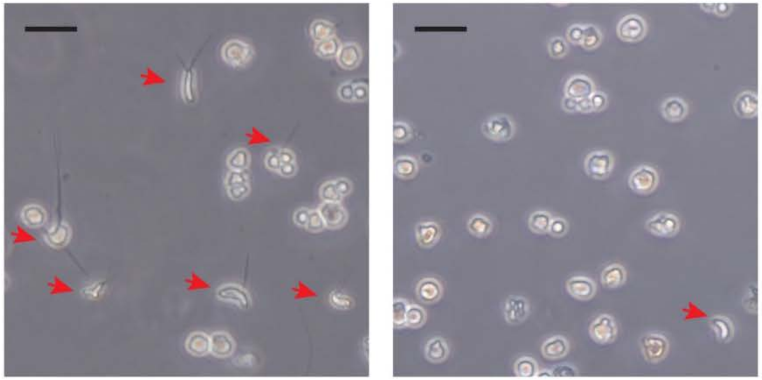
Traxler et al, Nat Med 2016

NHEJ to induce HbF for β-hemoglobinopathies



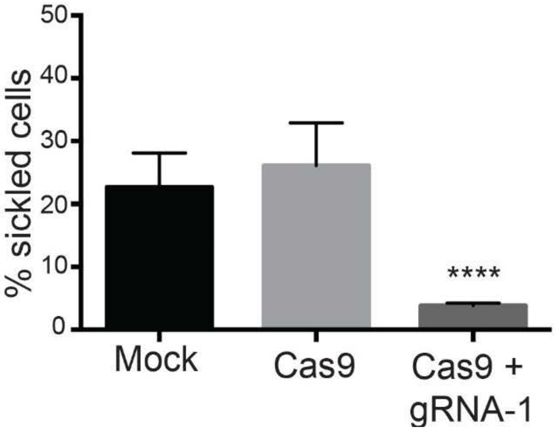


SCD cells
+
hypoxia



Cas9

Cas9 + gRNA-1



CAGE
Center for Advanced Genome Engineering
St. Jude Children's Research Hospital

- Patrick Connelly, Ph.D.
- Sam Peters, B.S.
- Sadie Sakurada, B.S.
- Shaina Porter, Ph.D.
- Rachel Levine, Ph.D.
- Baranda Hansen, M.S.
- Jamie Papizan, Ph.D.

GEiC
Genome Engineering & iPSC Center
Washington University in St. Louis

- Monica Sentmanat, Ph.D.
- Colin Florian, B.S.
- Amber Neilson, B.S.
- Zhihong Zhang, B.S.
- Yi-Hsien Chen, Ph.D.
- Elena Artimovich, Ph.D.
- Julia Wang, B.S.
- Monika Minner, M.S.

Funding:

St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder

DORIS DUKE
CHARITABLE FOUNDATION

TEAM
ALSAC


St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder
Finding cures. Saving children.