Defining Genome Editing Technologies for Therapeutics

Shengdar Q. Tsai, Ph.D.
Assistant Member
Department of Hematology
St. Jude Children's Research Hospital

St. Jude Children's Research Hospital

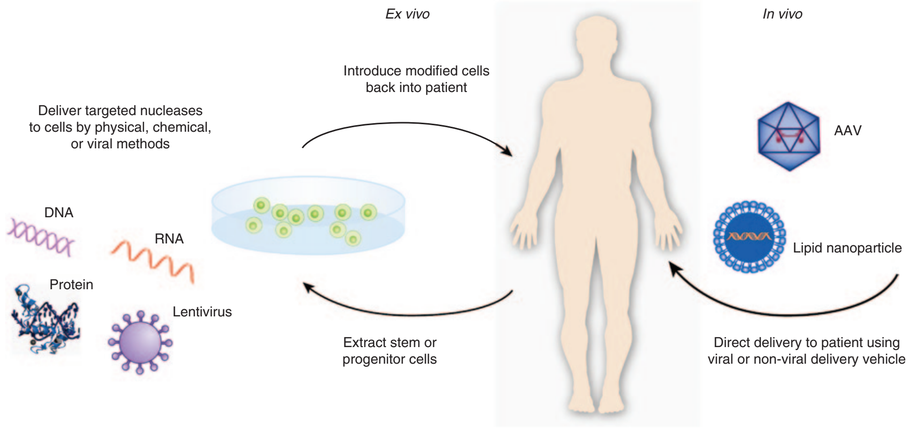
April 23, 2018
National Institute of Standards Technology

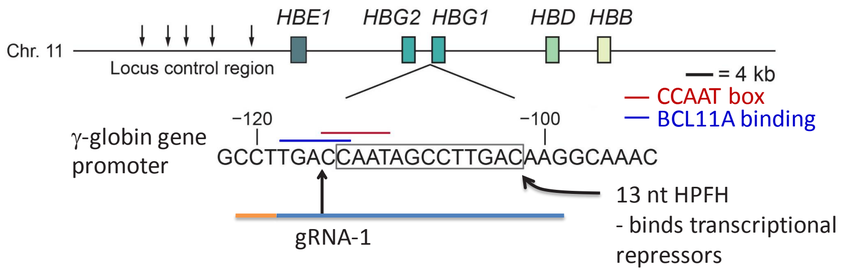
Potential Conflicts of Interest Disclosure

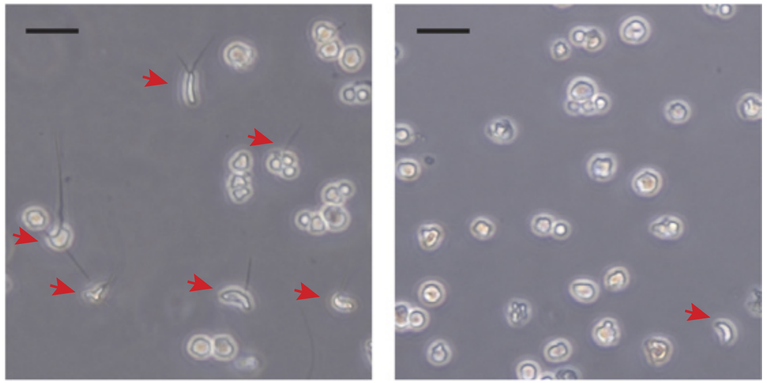
Shengdar Q. Tsai is a scientific co-founder of Monitor Biotech.

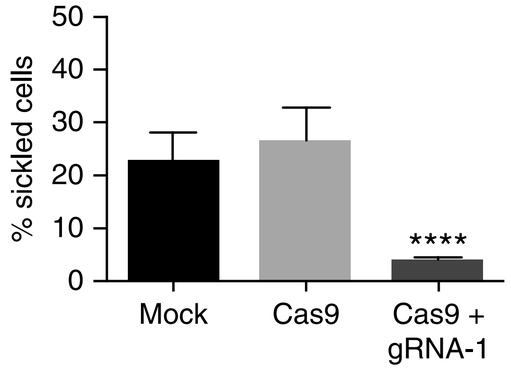
1. **Introduction**
2. State-of-the-art
3. Challenges
4. Summary

Genome editing for therapeutics









Exciting promise of genome editing is the potential to develop curative genetic therapies.

Maeder & Gersbach *Mol Ther.* 2016
Traxler *Nat. Med.* 2016.

Safety considerations: What about off-target effects?

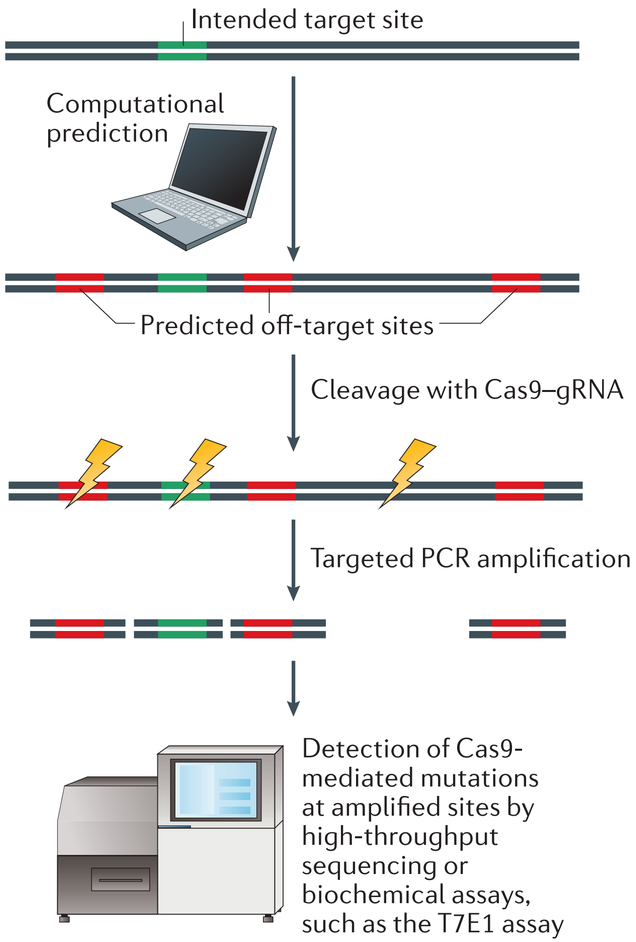
- Typically, hundreds of millions or more cells modified by therapeutic gene editing
- Even low-frequency **off-target mutations** may be relevant if they induce a cellular growth advantage
- Important to address, particularly for **therapeutic applications**

Defining **where** off-targets may occur enables critical safety monitoring

- Unlike gene therapy, no vector integrations in gene editing for easy tracking
- **Defining off-target locations is important**, even if we cannot currently interpret the function of many off-target sequences
- Methods for defining off-targets as comprehensively as possible enable **monitoring for clonal expansion** of cells harboring specific unintended edits
- However, new methods for assessing functional impact (**risk**) are urgently needed.

1. Introduction
2. **State-of-the-art**
3. Challenges
4. Summary

In silico prediction



Principle

- Computationally predict sites based on sequence similarity, etc.

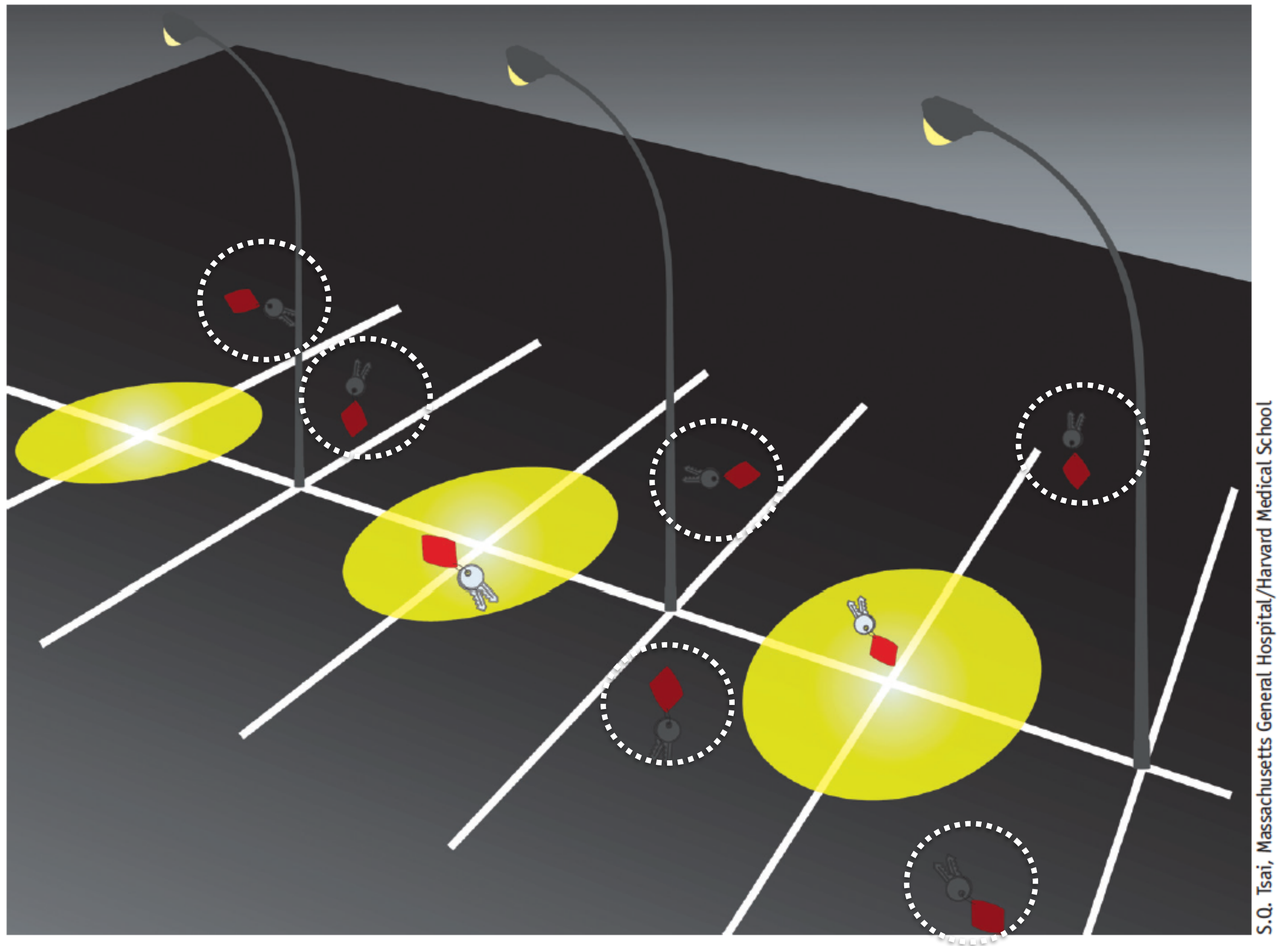
Advantages

- Easy

Disadvantages

- Biased by assumptions

The Streetlight Effect



S.Q. Tsai, Massachusetts General Hospital/Harvard Medical School

Whole genome sequencing is broad but shallow

- not practical to sequence large numbers of genomes
- significant limitation: it will miss off-target effects with low-frequency (i.e. sensitivity is poor)

Strategies for Defining Gene Editing Nuclease Genome-wide Activity

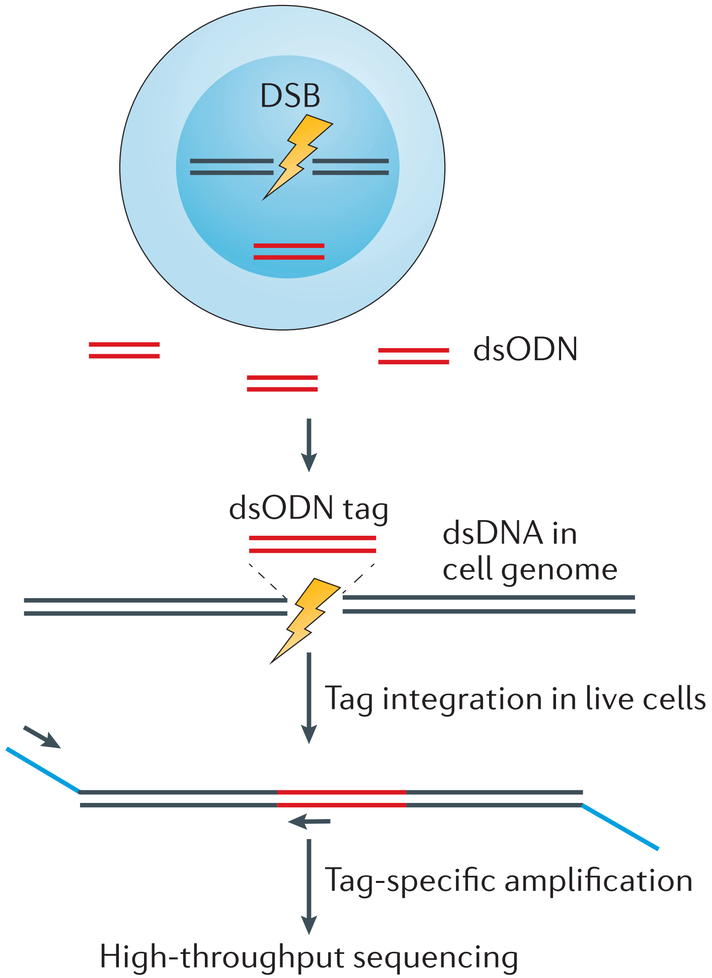
Cell-based Methods

- Integration deficient lentiviral (**IDLV**) capture
- High-throughput genome-wide translocation sequencing (**HTGTS**)
- Breaks labeling, enrichment on streptavidin and next-generation sequencing (**BLESS**)
- Genome-wide unbiased identification of DSBs enabled by sequencing (**GUIDE-seq**)

In vitro Methods

- Digested genome sequencing (**Digenome-seq**)
- selective enrichment and identification of adapter-tagged DNA ends by sequencing (**SITE-Seq**)
- Circularization for *in vitro* Reporting of Cleavage Effects by Sequencing (**CIRCLE-seq**)

GUIDE-seq: Genome-wide unbiased identification of DSBs enabled by sequencing



Principle

- Optimized tag integration into DSBs followed by tag-specific amplification and sequencing

Advantages

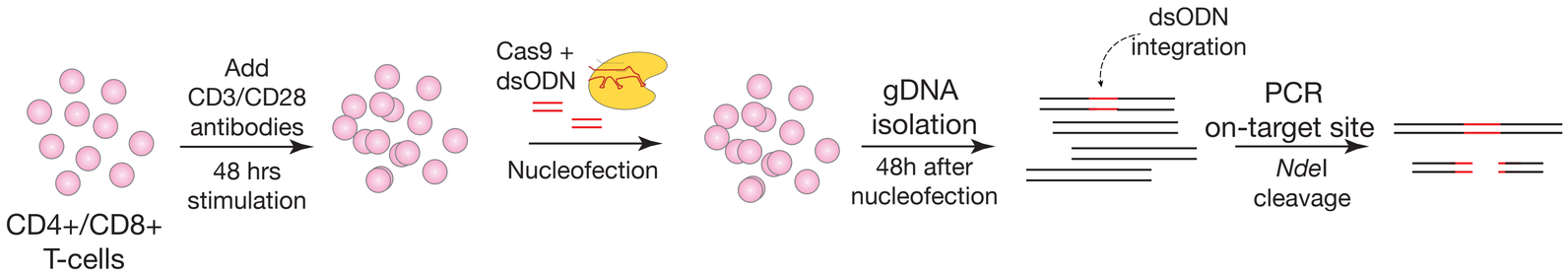
- Quantitative and unbiased method
- Can identify background or fragile sites

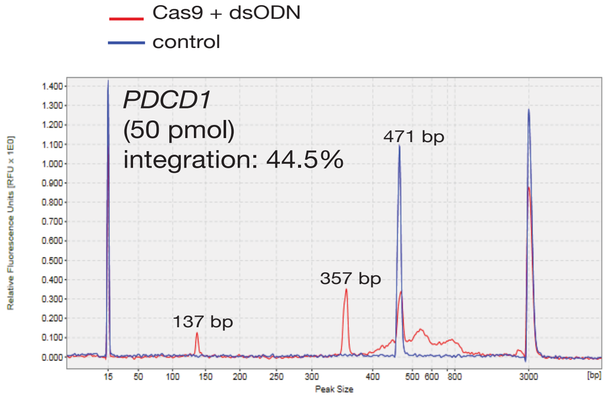
Disadvantages

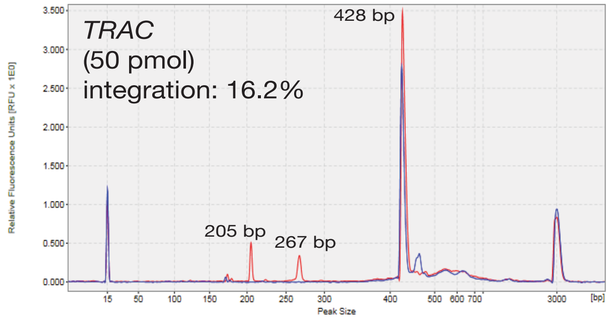
- Requires transfection of DNA tag, limiting use in some primary cells

Tsai et al. *Nat Rev Genetics* 2016

GUIDE-seq in Human Primary T-cells







GUIDE-seq

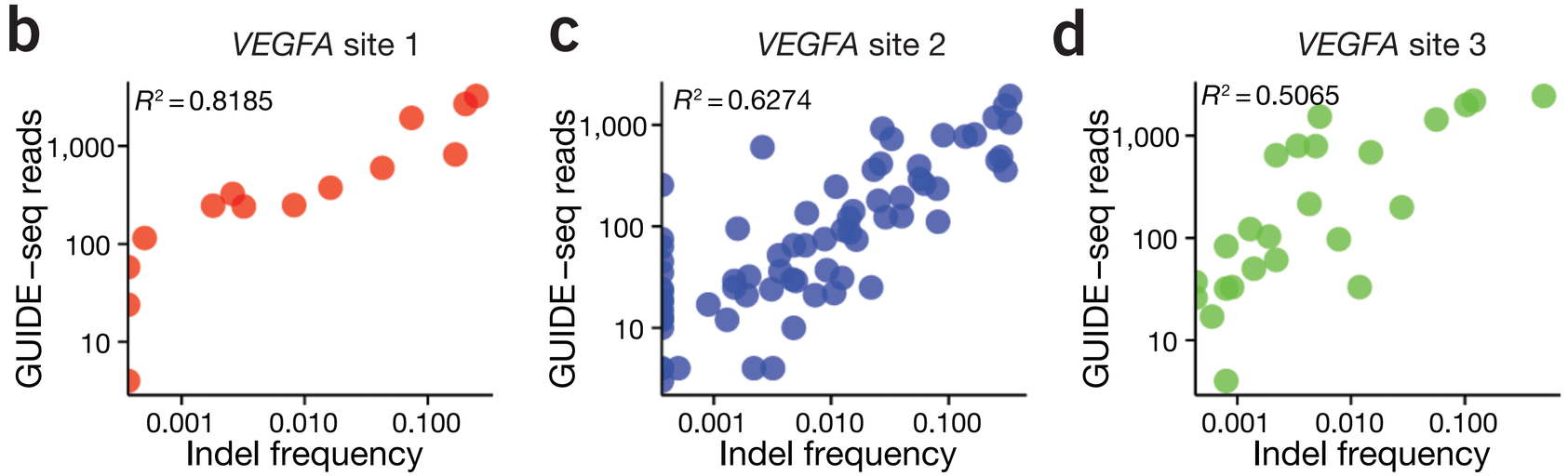
PDCD1

1									10										20		23	
G	G	G	C	G	G	T	G	C	T	A	C	A	A	C	T	G	G	G	C	N	G	G	Reads
·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	5098
A	A	·	G	·	C	G	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	17

TRAC

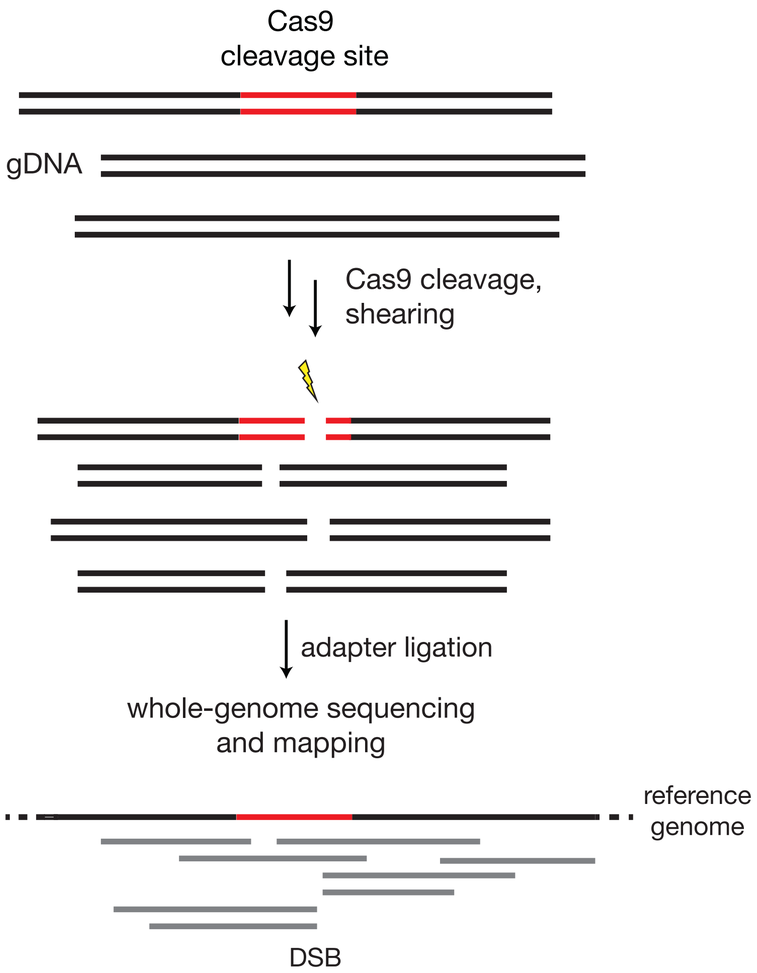
1									10										20		23	
C	T	T	C	A	A	G	A	G	C	A	A	C	A	G	T	G	C	T	G	N	G	G	Reads
·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1506
T	·	G	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	234

GUIDE-seq is quantitative but has limits to sensitivity



- Tag integration proportional to mutagenesis frequency
- Increasing sensitivity will require linear scaling of input genomes and sequencing

Digenome-seq: Digested Genome Sequencing



Principle

- *In vitro* cleavage of genomic DNA, whole genome sequencing, and identify sites with uniform ends

Advantages

- Simple, PCR-free

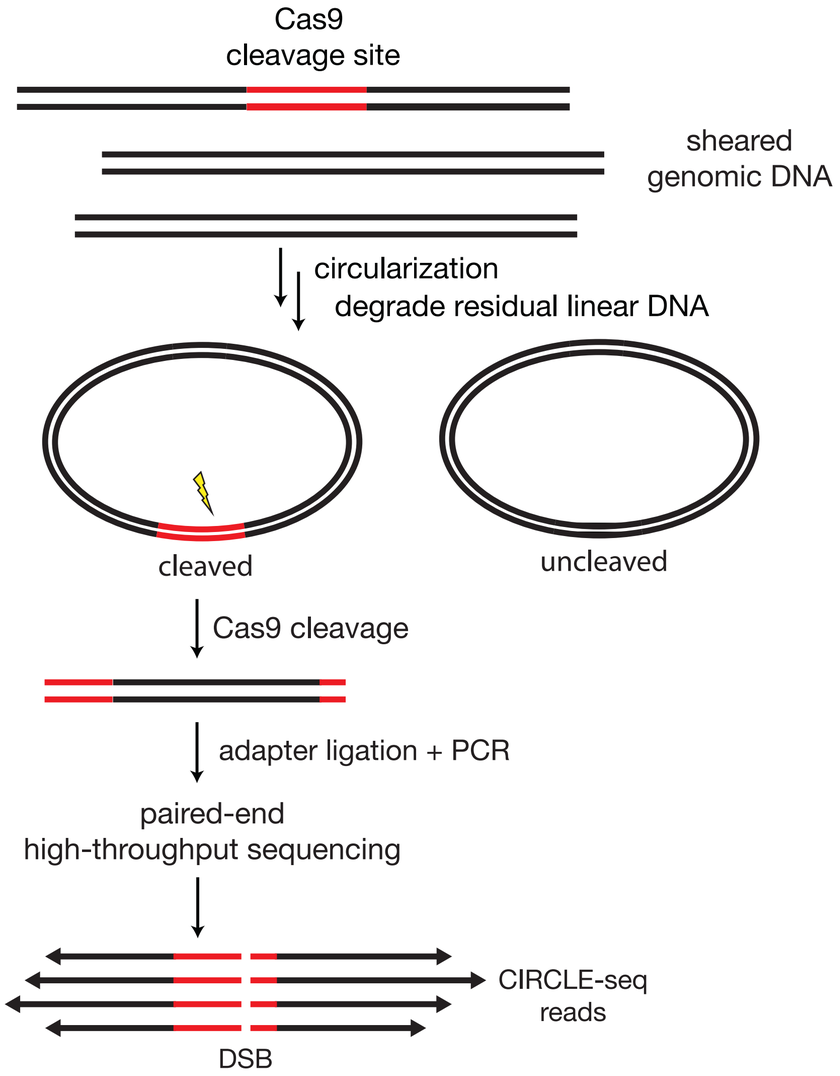
Disadvantages

- High number of sequencing reads required and high background
- Candidate sites need to be confirmed in cells

Tsai et al. *Nat Rev Genetics* 2016.

Standing out from a crowd

CIRCLE-seq: Circularization for *in vitro* reporting of cleavage effects by sequencing



Principle

- Selective sequencing of nuclease-cleaved genomic DNA

Advantages

- Low background
- High sensitivity
- Reference-free

Disadvantages

- Candidate sites need cellular confirmation

Tsai et al. *Nature Methods 2017*

CIRCLE-seq improves signal to noise while using 100-fold less reads

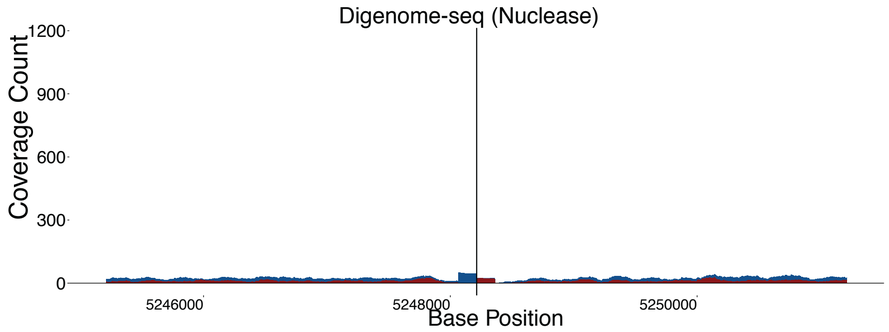
Digenome-seq

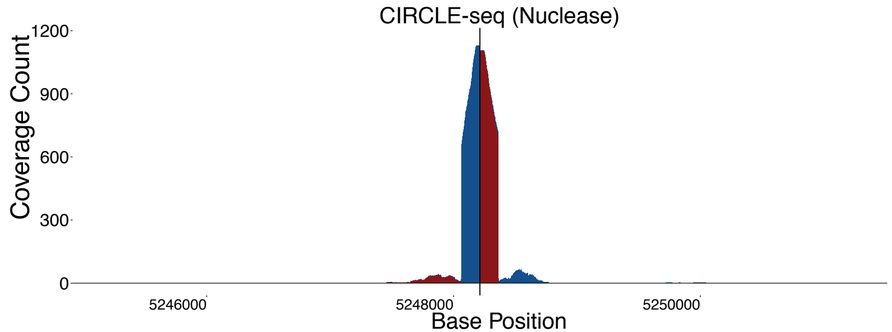
CIRCLE-seq

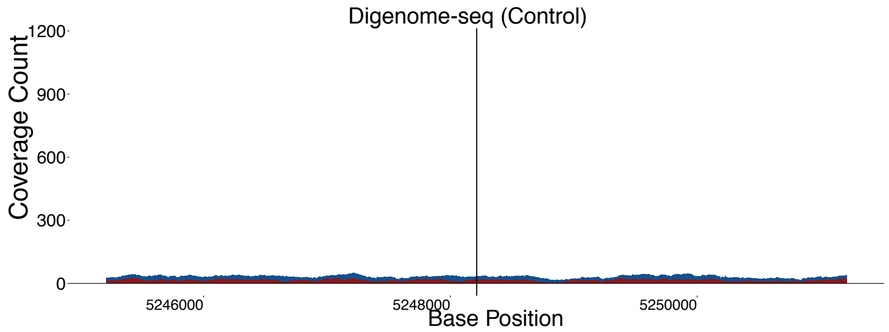
HBB

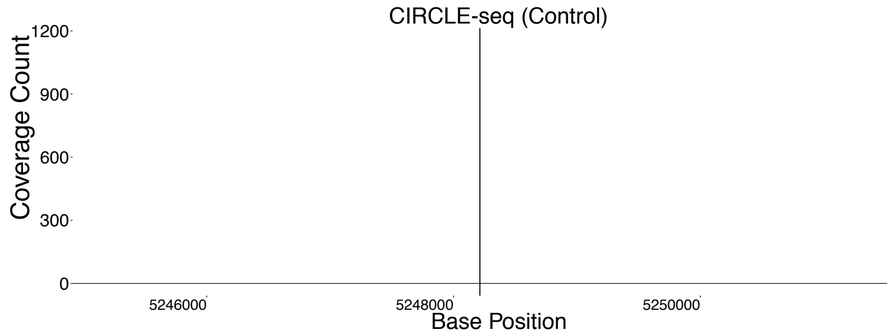
Nuclease

Control









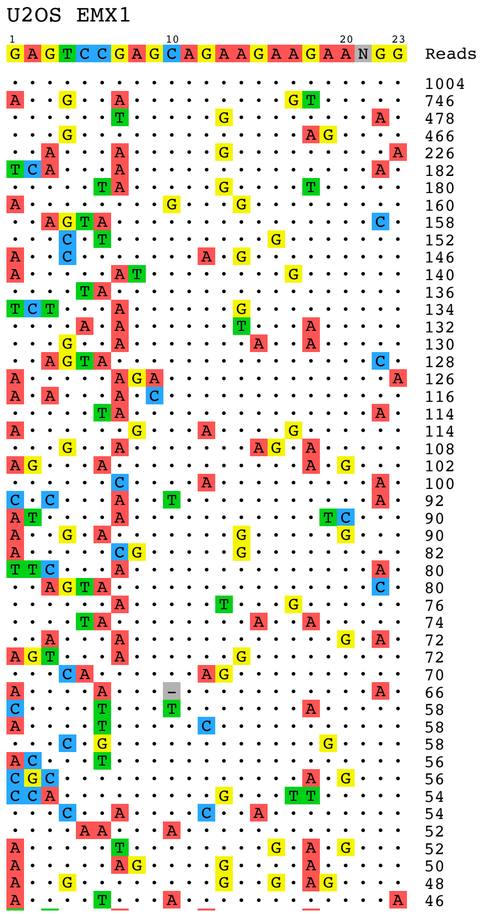
400 million reads

4 million reads

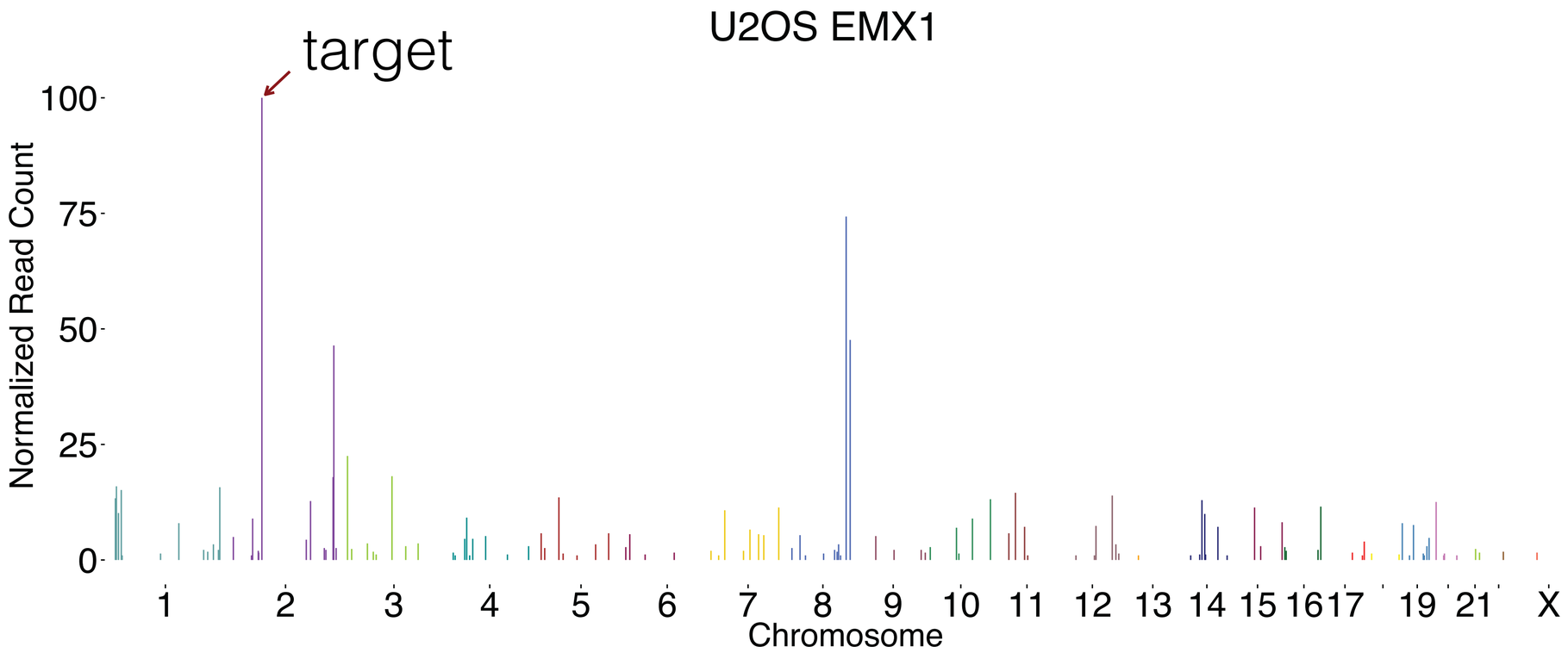
St. Jude Children's Research Hospital®
ALSAC • Danny Thomas, Founder
Finding cures. Saving children.

Tsai et al. *Nature Methods 2017.*

CIRCLE-seq genome-wide nuclease activity profiles



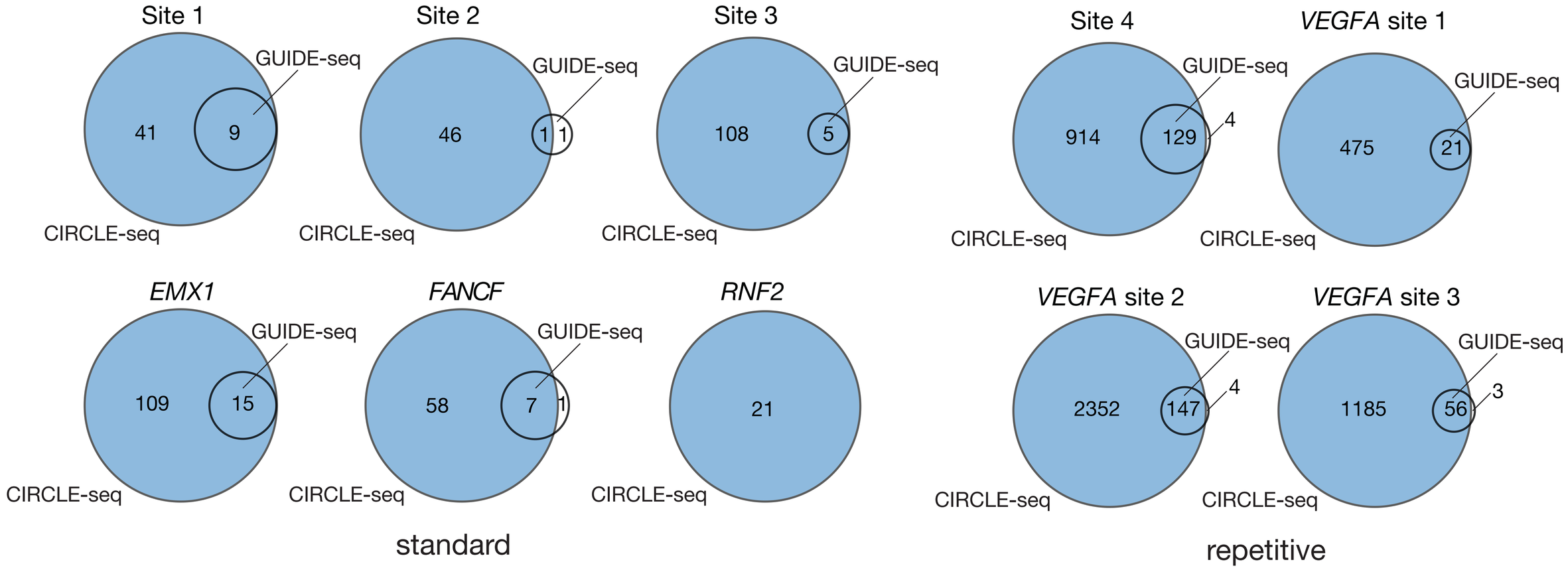
(cont.)



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

Tsai et al. *Nature Methods 2017*

CIRCLE-seq detects nearly all sites previously detected by GUIDE-seq



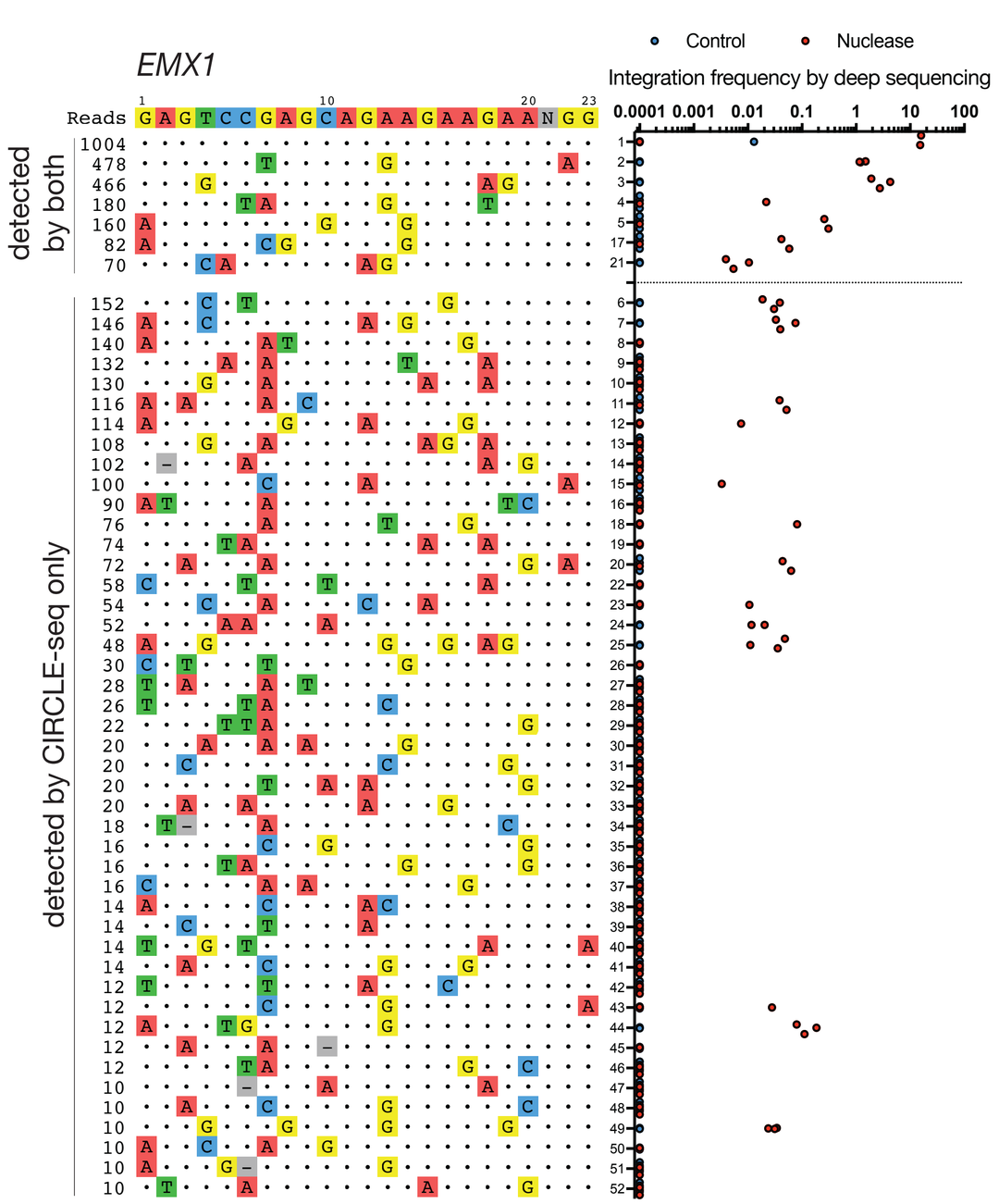
St. Jude Children's Research Hospital
ALSAC • Danny Thomas, Founder
Finding cures. Saving children.

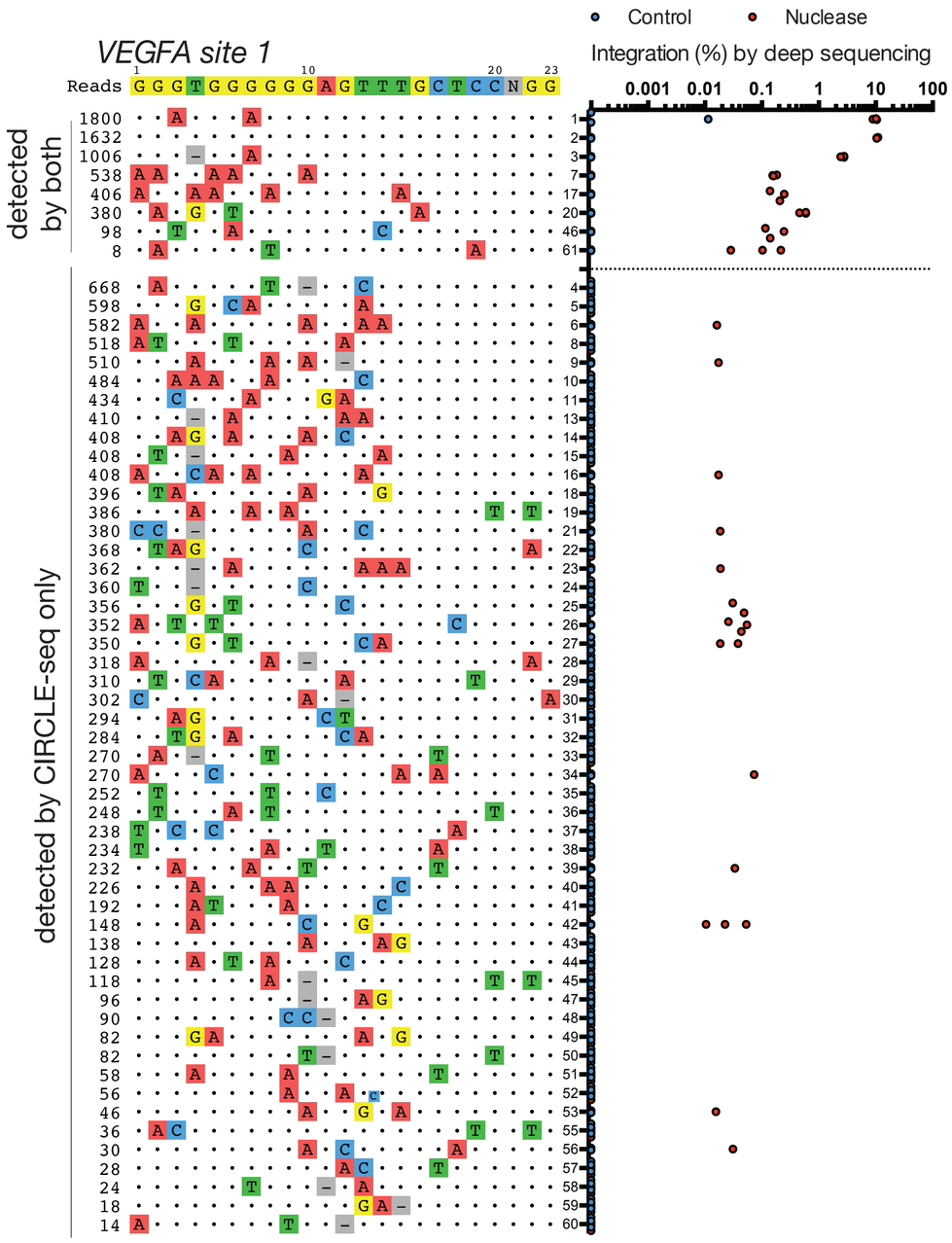
Tsai et al. *Nature Methods 2017*

Targeted tag sequencing strategy can overcome standard NGS limits of detection

- Background indel rate of modern high-throughput sequencing around 0.1%
- Tag integration is unambiguous evidence of DSB repair
- Proxy for mutagenesis frequency

Tsai et al. *Nature Methods 2017.*





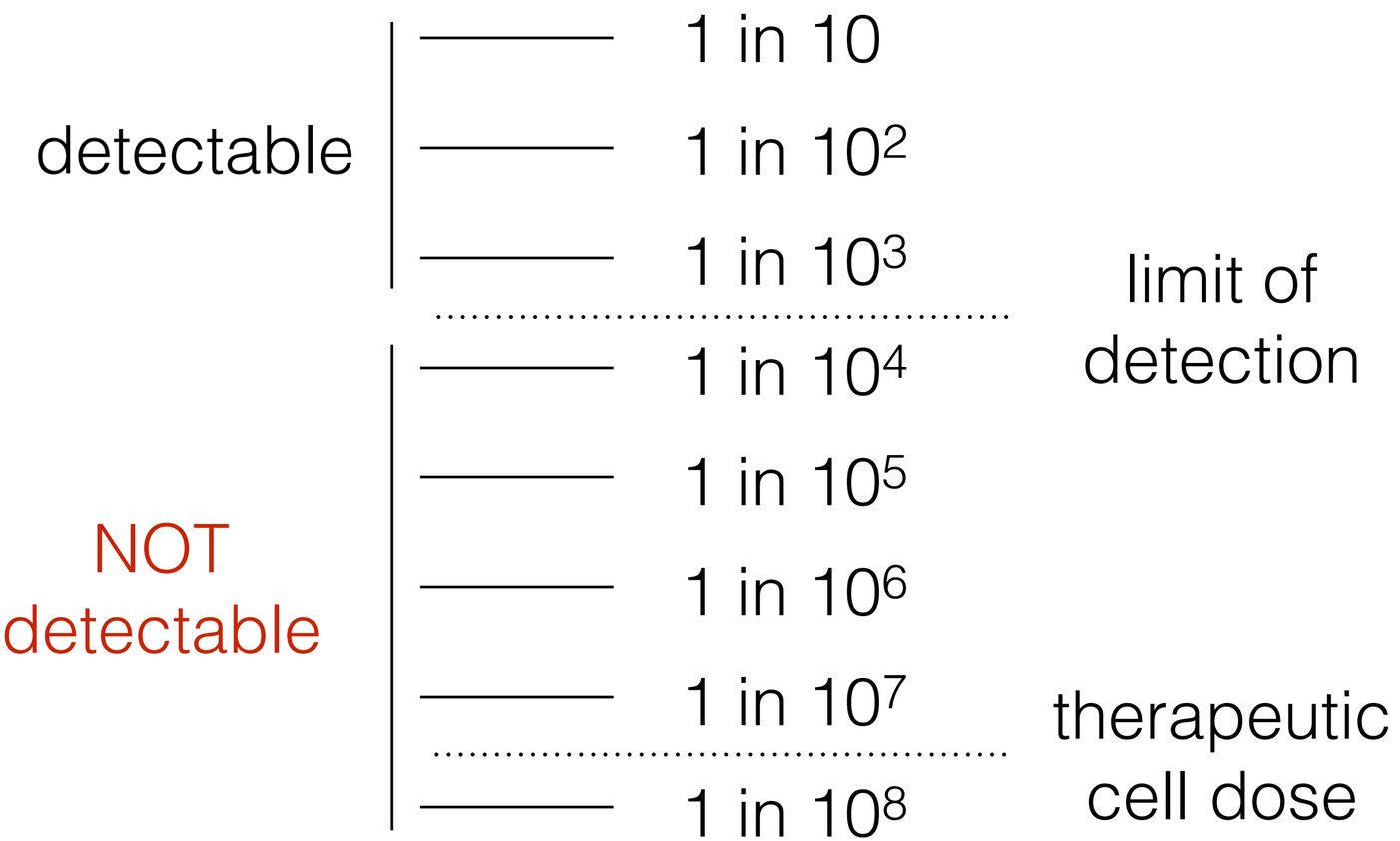
~25% of sites detected only by CIRCLE-seq confirmed as *bona fide* cellular sites.

Tsai et al. *Nature Methods 2017*

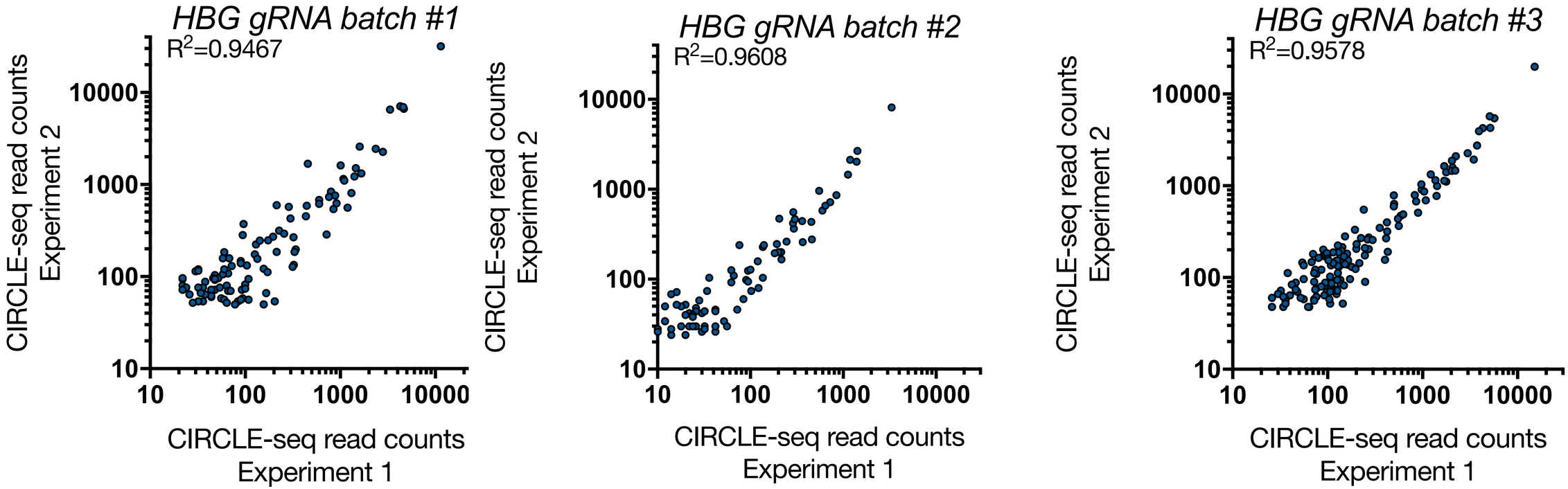
1. Introduction
2. State-of-the-art
3. **Challenges**
4. Summary

Technical limitations of standard high-throughput sequencing

- Error-rate obscures low-frequency mutations
- Short read length prevents mapping through repetitive regions

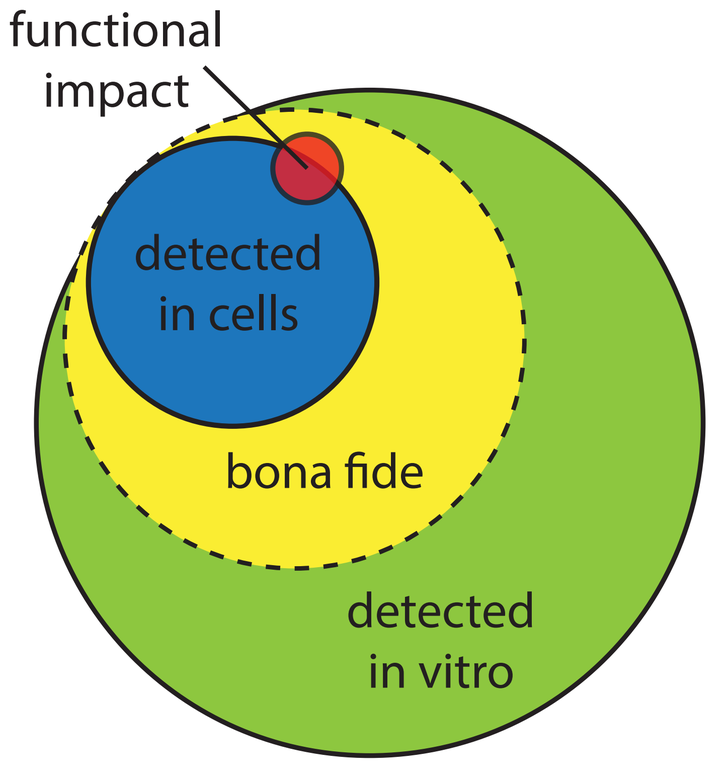


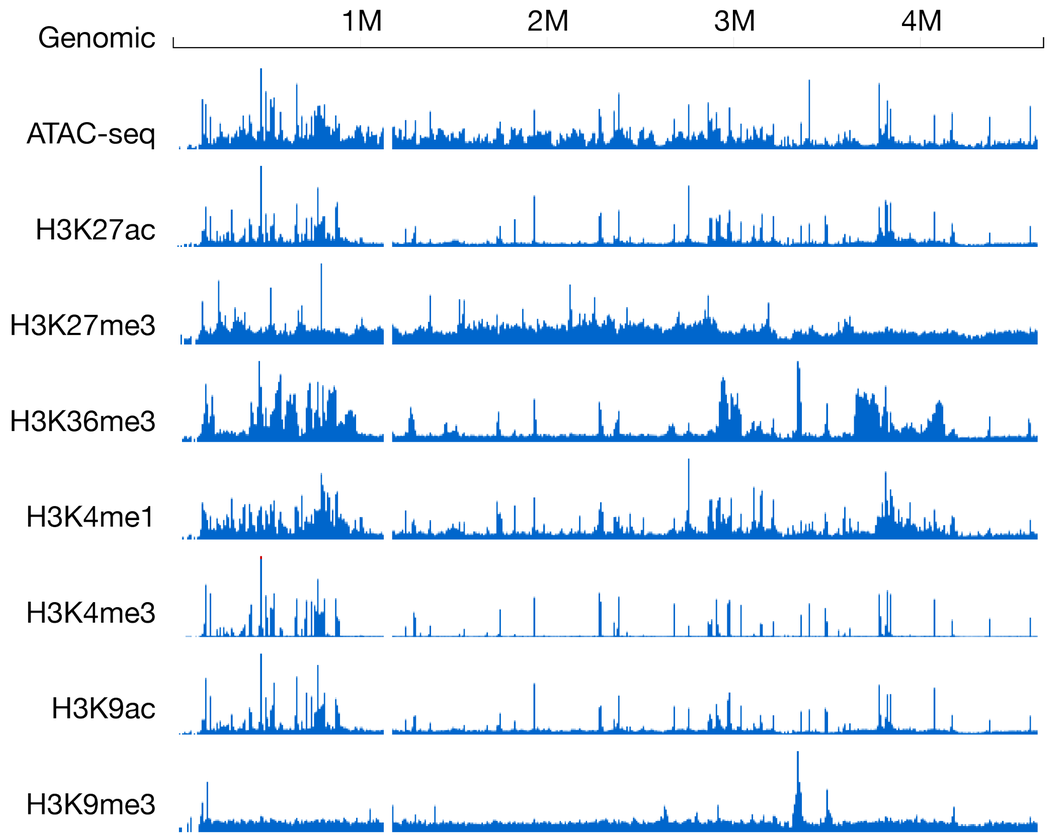
What is technical reproducibility of measurements?



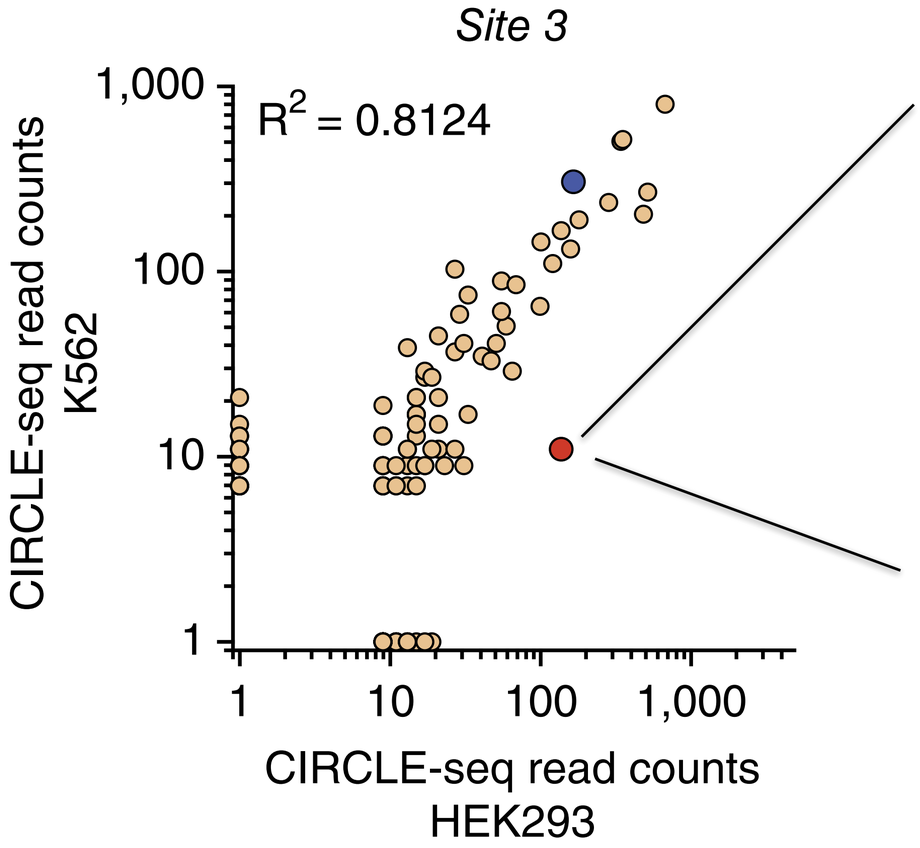
unpublished

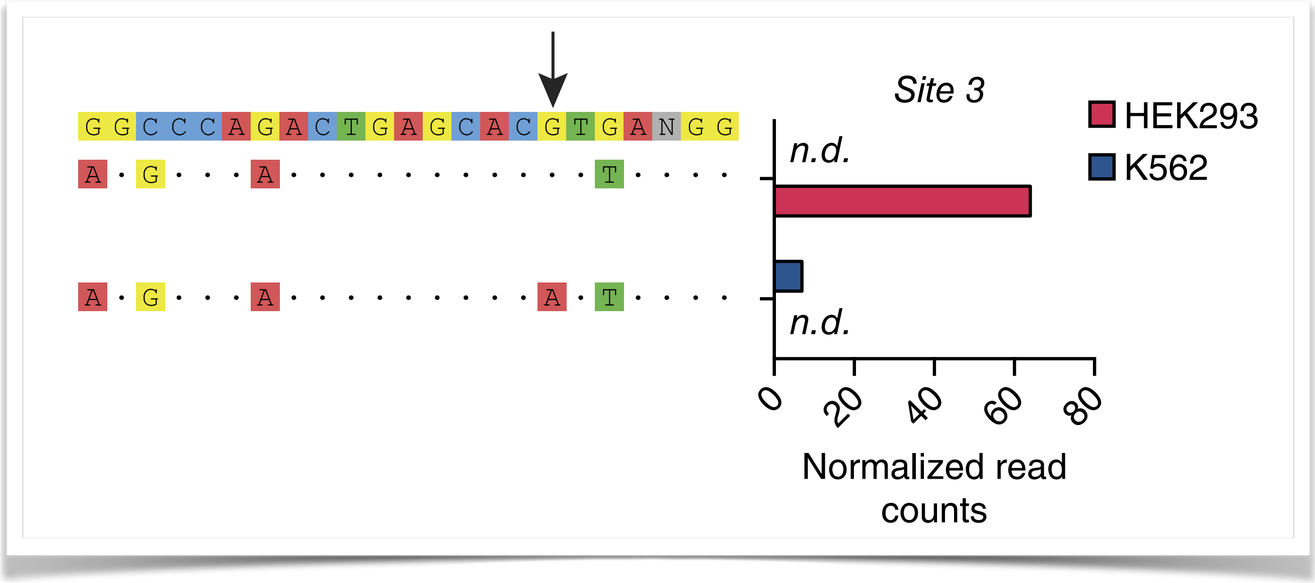
What determines relationship between cellular and *in vitro* measurements?





What is the impact of human genetic variation on genome editing activity?





Could be tested in genome-in-a-bottle reference cell lines.

Tsai et al. *Nature Methods 2017.*

Towards Developing Genome Editing Measurement Standards

- Reference standards for validating low-frequency mutation detection.
- Challenge and benchmark detection methods at common targets/nucleases/cells.
- Define reproducibility and limits of detection.

1. Introduction
2. State-of-the-art
3. Challenges
4. **Summary**

Summary

- Important to empirically define genome-wide activity of gene editing for therapeutics
- GUIDE-seq and CIRCLE-seq are two complementary methods to define genome-wide genome editing activity
- Glass half full: fortunate to have many sensitive and unbiased genomic methods to define nuclease specificity
- NIST consortium has opportunity to develop standards to define genome editing measurements and increase confidence in safety of promising genome editing therapeutics

Thank you

Tsai Lab*
Cicera Lazzarotto
Nikolay Malinin

Weiss Lab
Jean-Yves Metais
Yu Yao
Sagar Keriwala

Joung Lab
Yanfang Fu
Nhu Nguyen
Ved Topkar
Jeff Sander

Aryee Lab
Jose Malagon-Lopez
Vishal Thapar