# CRISPResso2

**Khalid Akram**
3rd year Ph.D Biotechnology,
Genome Engineering Laboratory,
University of Westminster, London.

# The journey so far

- 1992 – 2017 Software Engineer / Java / C/C++ / Perl / HTML/CSS, SQL, full stack developer

- 2018 – embarked on BSc Biological Sciences (came across CRISPR, **sickle cell disease** cure, edited mosquito DNA to change **eye colour**, Dr He Jianku, met Dr Kalpana)

- 2020 – as a 2nd year BSc student in first iteration of **The Gene Editors of The Future** - deleting Q Arm of chromosome 14

- 2021-2022 – final year project: developed a CRISPR based tool for **detection and diagnosis** of Influenza A virus

- 2022 – present – PhD Biotechnology. Working with CRISPR edited MCF-7 breast cancer cells. Characterizing the effects of the loss of the **ZFP36L1 gene** – expresses an RNA binding protein.

# What is CRISPResso2?

- CRISPresso2 is a software tool used for analyzing CRISPR/Cas9 (and other types of CRISPR) genome editing experiments.
- Provides an integrated platform for **quantifying** the **efficiency** and **outcomes** of CRISPR-based modifications in genomic DNA.
- Helps researchers evaluate the **extent** of gene editing by analysing sequencing data, particularly focusing on indel (insertions and deletions) frequencies and types.
- Useful for assessing the accuracy and effectiveness of CRISPR experiments.

# What questions can it answer?

Q1. Did my CRISPR experiment work?

Q2. What was the efficiency of the edit?

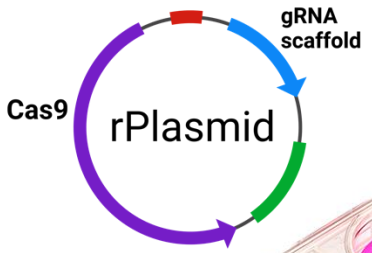
Q3. What **kind** of edits were made across the cell population? (insertions / deletions / mutations).

# Developers of CRISPResso2

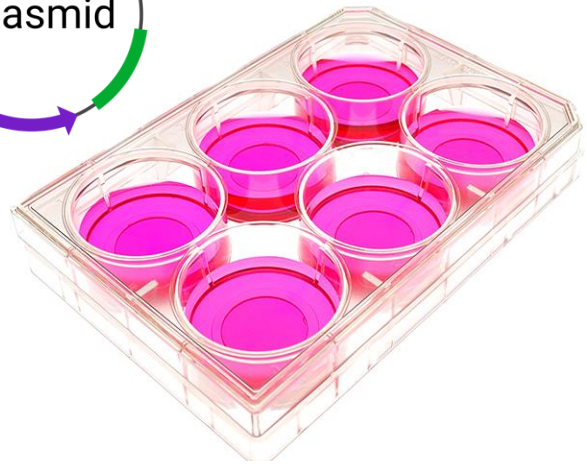
**Luca Pinello**
**Ph.D., Associate Professor**
**Massachusetts General Hospital and HMS**
(CRISPResso)

**Kendell Clement**
**Assistant Professor**,
Biomedical Informatics
University of Utah,
(CRISPResso2)

# Where does CRISPResso2 fit into the CRISPR workflow?




## 1. Plasmid mediated CRISPR transfection



## 2. DNA Extraction

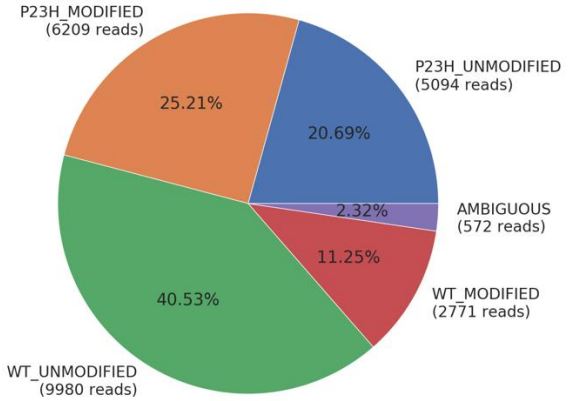
## 3. Sequencing (e.g. Ampliseq)

Header Sequence Quality

```
@HWI-ST227:389:C4WA2ACXX:7:1204:2272:59979
GGAGGAAGGTCCTCGCTCCTCTTTCATATAAGGGAAATGGCTGAAT
+
FFFFHHHHHHJIJJJJJJJJIJJJIGIGIGGIJJIJIJJJJJJIII
@HWI-ST227:389:C4WA2ACXX:7:1205:15214:42893
GAGGATCCCAGGGAGGAAGGTCCTCGCTCCTCTTTCATCTAAGGGA
+
12BAFB?A:3<AE1@<FF;1*@EG*)?0?DBD>9BF9B*?######
@HWI-ST227:389:C4WA2ACXX:8:2208:2467:44624
AAAGAGGAGAGAGGACCATCCTCCCTGGGATCCTCAGAAGTCTACT
+
BDDA:DB?2AA@FC>F?EEGC<FED>GFD;?GBB?<?F99*/9?9?
```

## 4. *.fastq* format files

## 5. Analysis using CRISPResso2




## 6. Reports & Visualisation

# Processing of .fastq files by CRISPResso2

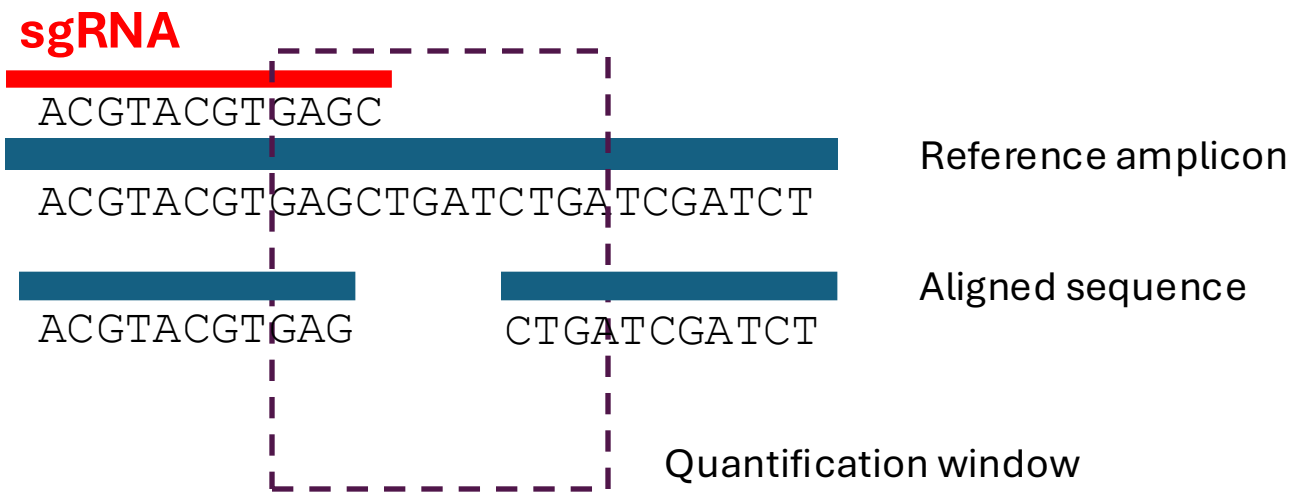
## 1. Data

Header Sequence Quality

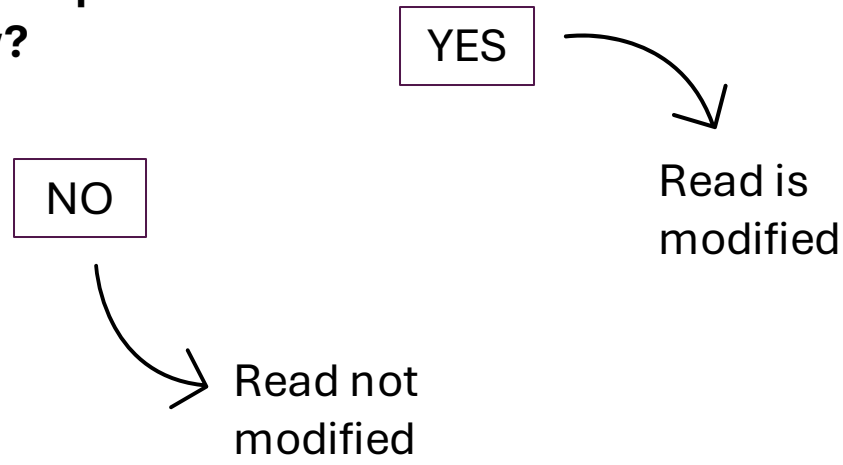
```
@HWI-ST227:389:C4WA2ACXX:7:1204:2272:59979
GGAGGAAGGTCCTCGCTCCTCTTTCATATAAGGGAAATGGCTGAAT
+
FFFFHHHHHHJIJJJJJJJIJJJIGIGIGGIJJIJIJJJJJJIII
@HWI-ST227:389:C4WA2ACXX:7:1205:15214:42893
GAGGATCCCAGGGAGGAAGGTCCTCGCTCCTCTTTCATCTAAGGGA
+
12BAFB?A:3<AE1@<FF;1*@EG*)?0?DBD>9BF9B*?######
@HWI-ST227:389:C4WA2ACXX:8:2208:2467:44624
AAAGAGGAGAGAGGACCATCCTCCCTGGGATCCTCAGAAGTCTACT
+
BDDA:DB?2AA@FC>F?EEGC<FED>GFD;?GBB?<?F99*/9?9?
```

Sequencing reads (.fastq)

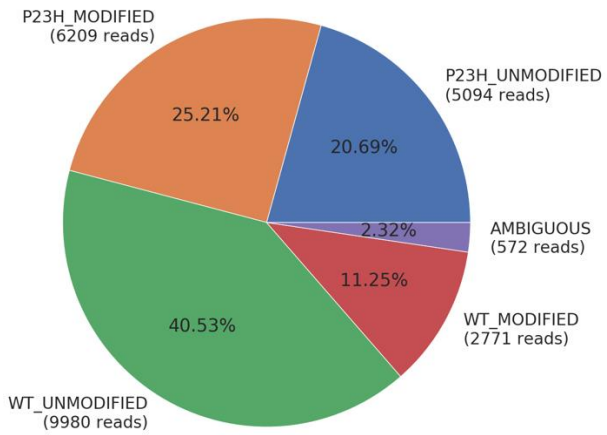
## 2. Alignment




## 3. Are edits present within the quantification window?




## 4. Summary of Editing

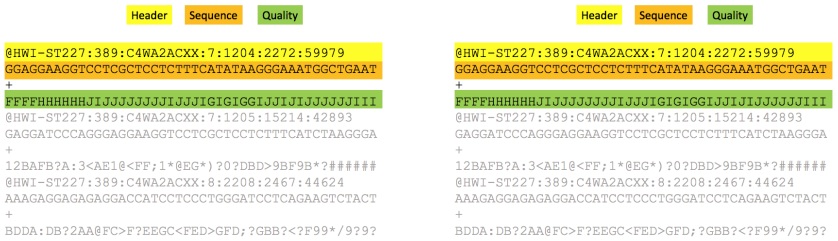


# Inputs to CRISPResso2

**Editing Type**

- Cas9
- Cpf1
- Base editors
- Prime editors
- Custom

**Data**




Fastq File 1



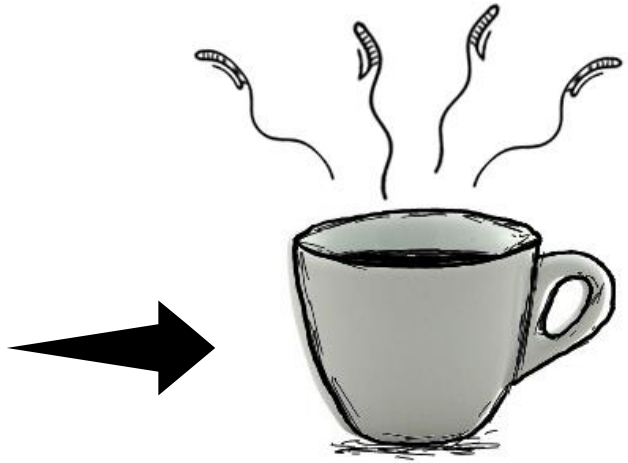

Fastq File 2

AATGTCCCCCAATGGGAAGTTCATCTGGCACTGCCCACAGGTGAG
GAGGTCATGATCCCCTTCTGGAGCTCCCAACGGGCCGTGGTCTGG
TTCATCATCTGTAAGAATGGCTTCAAGAGGCTCGGCTGTGGTT

Amplicon

GTGCGGAGCCACTTCGAGCAGC

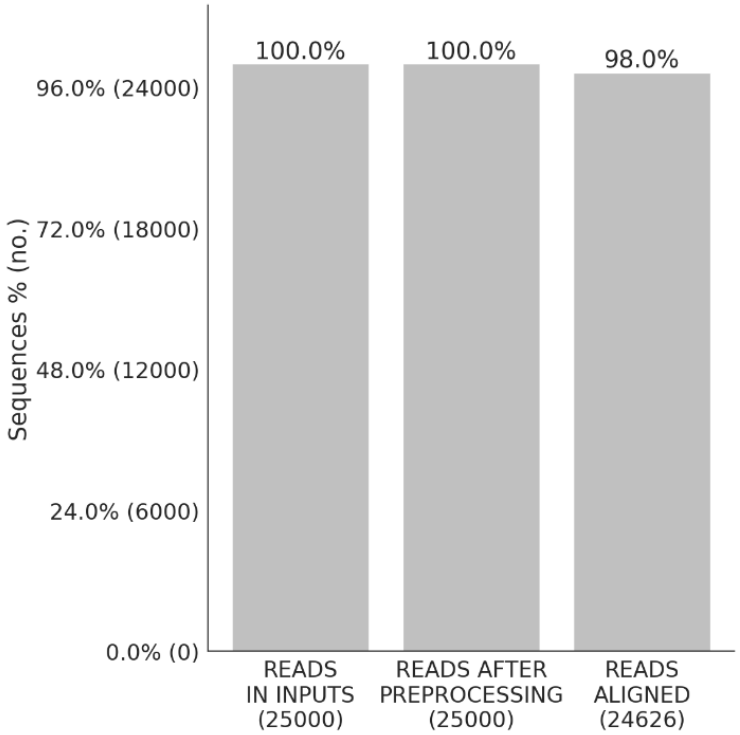
sgRNA



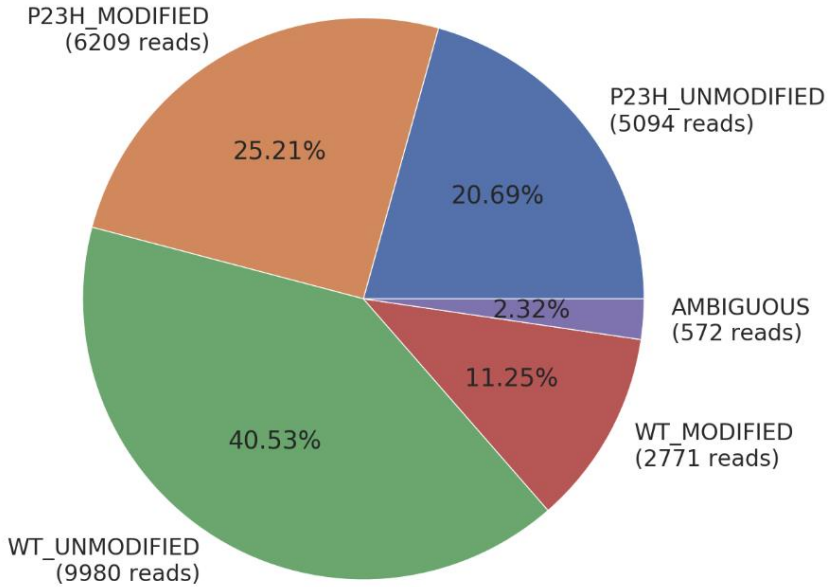
CRISPResso2

# Outputs from CRISPResso2

CRISPResso2

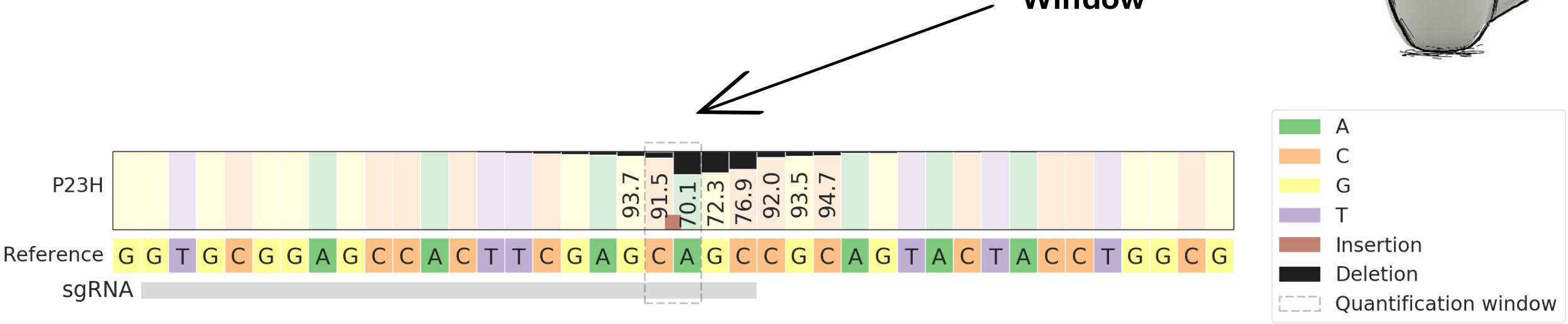



Mapping Statistics



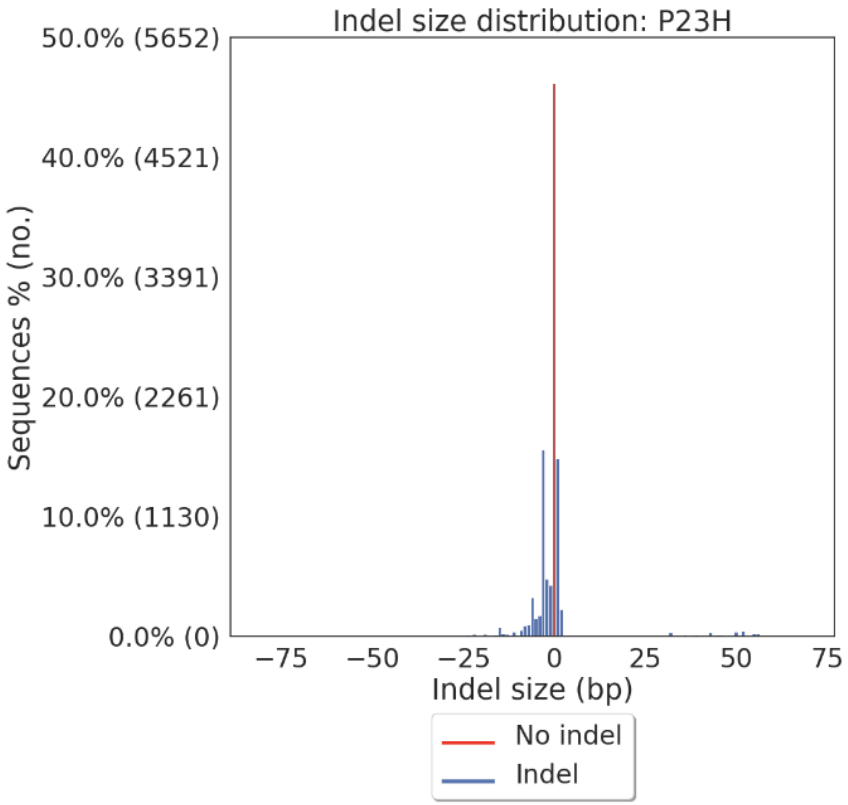

Quantification of Editing

# Outputs from CRISPResso2

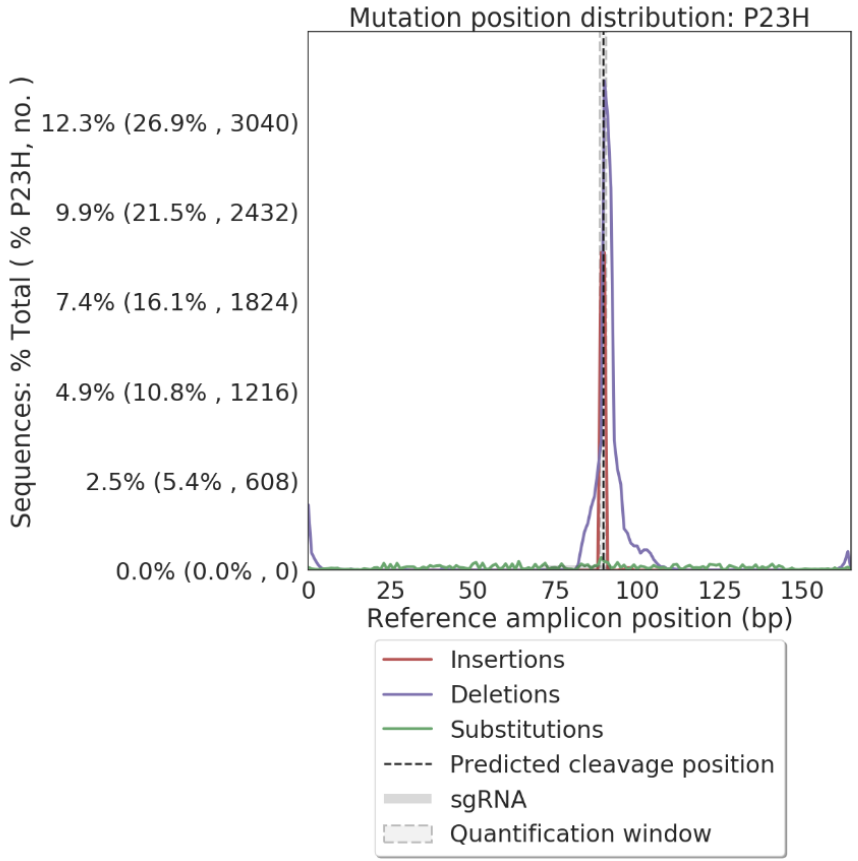



Distribution of nucleotides in the proximity of the sgRNA and cutting site

# Outputs from CRISPResso2

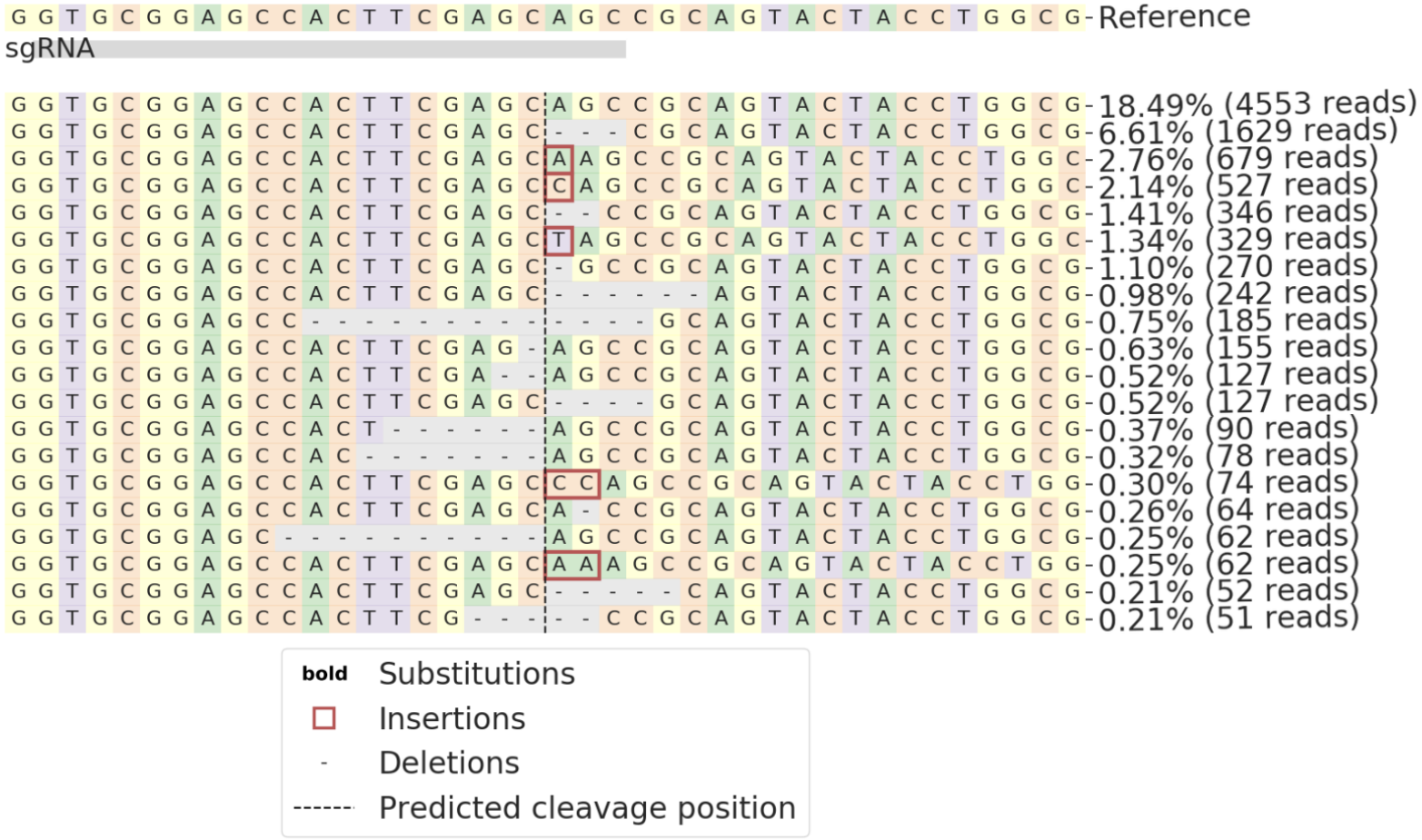



Distribution alleles with and without indels




Frequency of insertions, deletions, and substitutions across the amplicon, including modifications outside of the quantification window.

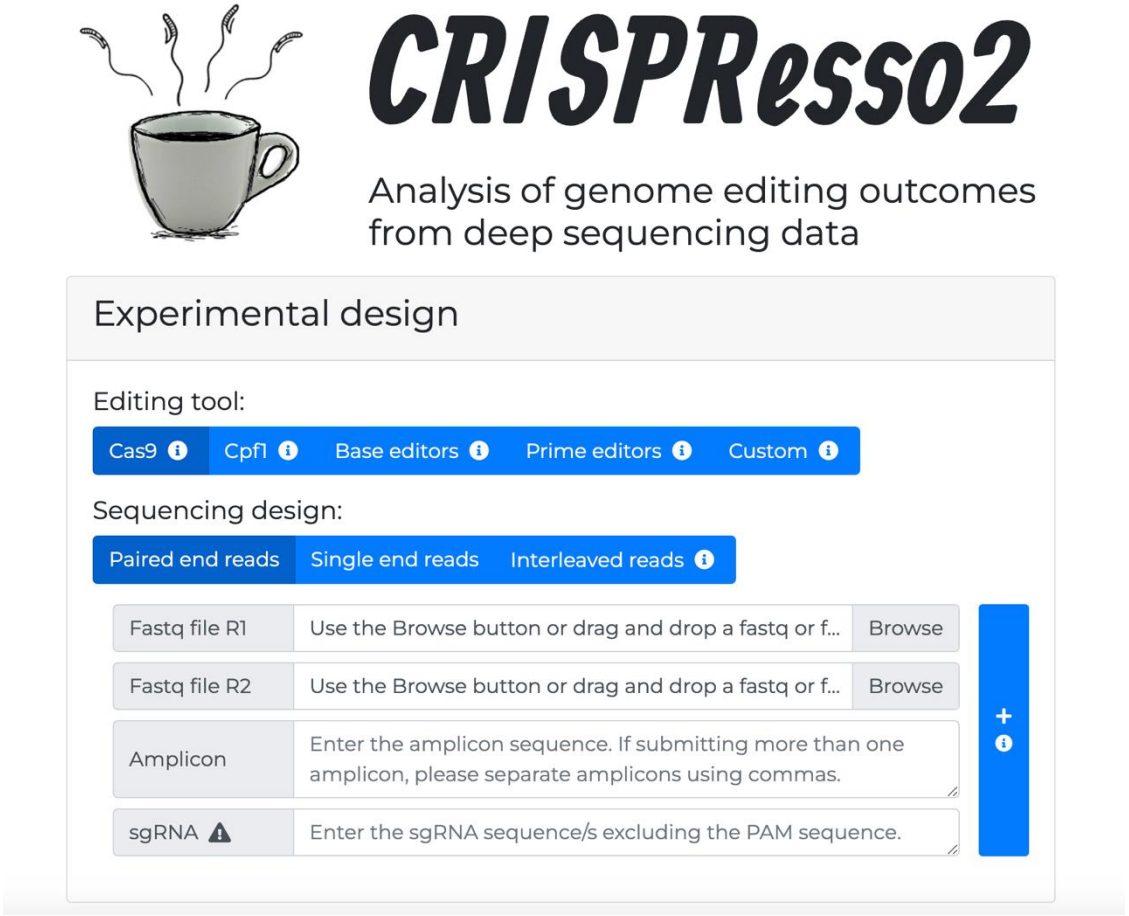
# Outputs from CRISPResso2




Visualization of the distribution of identified alleles around the cleavage site for the sgRNA GTGCGGAGCCACTTCGAGCAGC. Substitutions are in bold. Red rectangles highlight inserted sequences. Horizontal dashed lines indicate deleted sequences. The vertical dashed line indicates the predicted cleavage site.

# How to use CRISPResso 2 Online (Web)

**http://crispresso2.pinellolab.org/**




**Pros:**

- No installation
- Easy to use
- Great interface

**Cons:**

- Can't use if offline
- Slow if uploading large files
- Slow if many people are using site simultaneously
- URL is **http** rather than **https** – security concerns
- Sometimes site is unreachable

# How to use CRISPResso 2 locally (on own computer)

From official documentation:
**https://docs.crispresso.com/installation.html**

CRISPResso can be installed using the conda package manager Bioconda, or it can be run using **the Docker containerization** system.

**Running via docker is probably the easiest way, as you don't have to worry about which packages to install**

**Docker**

CRISPResso can be used via the Docker containerization system. This system allows CRISPResso to run on your system without configuring and installing additional packages. To run CRISPResso, first download and install docker: https://docs.docker.com/engine/installation/.

Next, Docker must be configured to access your hard drive and to run with sufficient memory. These parameters can be found in the Docker settings menu. To allow Docker to access your hard drive, select 'Shared Drives' and make sure your drive name is selected. To adjust the memory allocation, select the 'Advanced' tab and allocate at least 4G of memory.

To run CRISPResso, make sure Docker is running, then open a command prompt (Mac) or Powershell (Windows). Change directories to the location where your data is, and run the following command:

```
docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso -h
```

**Pros:**

- *FAST!!*
- Works offline
- Large files are not an issue
- Docker containers make running it a breeze.
- Platform independent

**Cons:**

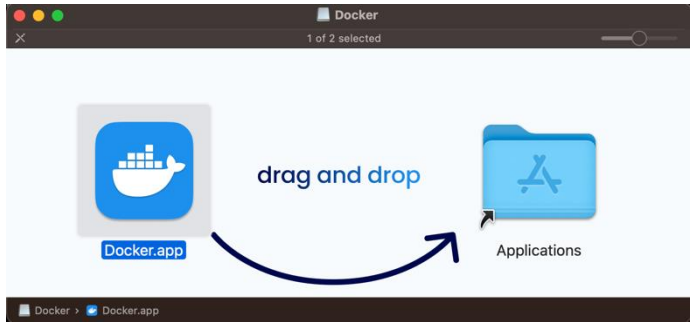
- Need to be slightly more tech savvy
- Command line interface can be daunting for some
- Install Docker
- Some command line knowledge required

# Getting Docker up and running (on Apple Mac)

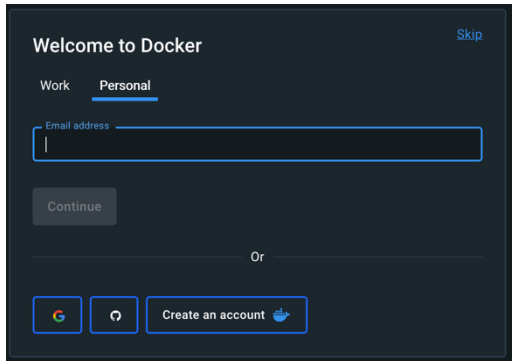
**https://docs.docker.com/desktop/**

**https://docs.docker.com/desktop/install/mac-install/**

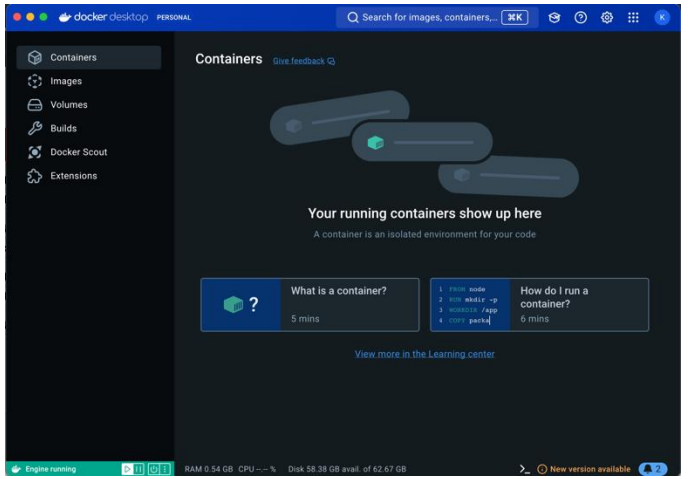
1. Download the installer (Docker.dmg for Mac)
2. Double click then copy to Applications folder



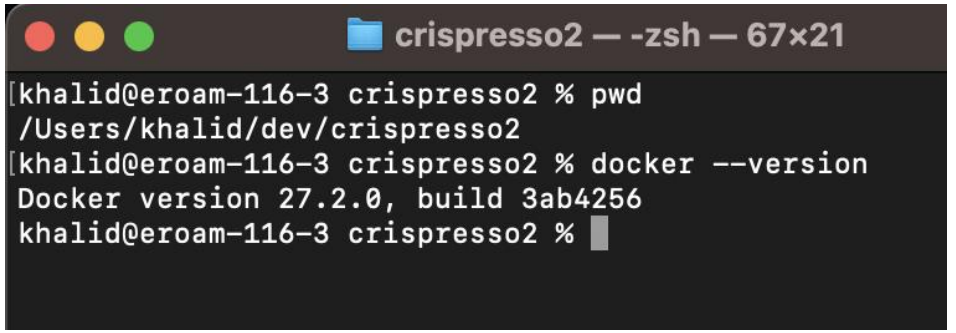

3. Login using email or create a new account




4. Home screen after sign up and logging in




5. Check Docker is running and version via terminal using **`docker --version`** command.



# Run CRISPresso2 via Docker

1. Execute the following command to download latest version of Crispresso2 container

```
docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso -h
```

2. Data to be processed should be copied into working directory (e.g. `/dev/crispresso2`)

3. Run crispressso2 with arguments as follows

```
docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso -
r1 nhej.r1.fastq.gz -r2 nhej.r2.fastq.gz -a
AATGTCCCCCAATGGGAAGTTCATCTGGCACTGCCCACAGGTGAGGAGGTCATGATCCCCTTCTGGAGCTCCCAA
CGGGCCGTGGTCTGGTTCATCATCTGTAAGAATGGCTTCAAGAGGCTCGGCTGTGGTT
```

4. If executed successfully a **directory** and **HTML** file will be created as follows:

   a) `CRISPResso_on_nhej.r1_nhej.r2/`
   b) `CRISPResso_on_nhej.r1_nhej.r2.html`

5. Open HTML file 4(b) in browser to view reports

# Sample command line output

```
khalid@eroam-116-3 crispresso2 % docker run -v ${PWD}:/DATA -w /DATA -i pinellolab/crispresso2 CRISPResso -r1 MCF7-L1W_S2_L001_R1_001.fastq.gz -r
2 MCF7-L1W_S2_L001_R2_001.fastq.gz -a CCTATGCTTGGACCTAGGTGTCATAACTTACTTTAAATATGTATGTTTGGTTTTCATTCATATTGACAGTACTACCTCTCAGTTTTCTTTCAGATATTGTTTTGTAT
TTACCCATGAAGACATTGTTTTTTGGACTCTGCAAATAGGACATTTCAAAGATGAGTGAAAAAAAATTGGAAACAACTGCACAGCAGCGGAAATGTCCTGAATGGATGAATGTGCAGAATAAAAGATGTGCTGTAGAAGAAAGAA
AGGTATGTTGTTCATTGACTATTCTTTTGGGTGAGAAATTTAATTTATATTTGACTGTGCAAAGAGTCAGTTGTTACTTGTAAACTTCAAGTCATTGTTTAGGTCAGAGTTGCTGTTGTCTAAATGCACCAGGACCTAGTTGTTG
AAAGGGTAAACTGGAATAAACTTTAATTGGGTTTACAAAATGAGAATTCTTACTGTATATTTTCTCTTTTTCGGGTTGACTTTACCAGT -g GGAAACAACTGCACAGCAG
WARNING: The requested image's platform (linux/amd64) does not match the detected host platform (linux/arm64/v8) and no specific platform was req
uested
INFO  @ Mon, 07 Oct 2024 18:30:25:
        Creating Folder CRISPResso_on_MCF7-L1W_S2_L001_R1_001_MCF7-L1W_S2_L001_R2_001

INFO  @ Mon, 07 Oct 2024 18:30:25:
        Computing quantification windows

INFO  @ Mon, 07 Oct 2024 18:30:25:
        CRISPRessoPro not installed

INFO  @ Mon, 07 Oct 2024 18:30:25:
        Added 0 guides with flexible matching
       Original flexiguides: ['None']
       Found guides: []
       Mismatch locations: []

INFO  @ Mon, 07 Oct 2024 18:30:26:
        Processing sequences with fastp...

INFO  @ Mon, 07 Oct 2024 18:30:39:
        Done!

INFO  @ Mon, 07 Oct 2024 18:30:39:
        Done!

INFO  @ Mon, 07 Oct 2024 18:30:40:
        Aligning sequences...

INFO  @ Mon, 07 Oct 2024 18:30:40:
        Processing reads; N_TOT_READS: 0 N_COMPUTED_ALN: 0 N_CACHED_ALN: 0 N_COMPUTED_NOTALN: 0 N_CACHED_NOTALN: 0

INFO  @ Mon, 07 Oct 2024 18:30:49:
        Processing reads; N_TOT_READS: 10000 N_COMPUTED_ALN: 841 N_CACHED_ALN: 6731 N_COMPUTED_NOTALN: 447 N_CACHED_NOTALN: 1981

INFO  @ Mon, 07 Oct 2024 18:30:57:
        Processing reads; N_TOT_READS: 20000 N_COMPUTED_ALN: 1474 N_CACHED_ALN: 13479 N_COMPUTED_NOTALN: 845 N_CACHED_NOTALN: 4202

INFO  @ Mon, 07 Oct 2024 18:31:04:
        Processing reads; N_TOT_READS: 30000 N_COMPUTED_ALN: 2023 N_CACHED_ALN: 20265 N_COMPUTED_NOTALN: 1189 N_CACHED_NOTALN: 6523

INFO  @ Mon, 07 Oct 2024 18:31:10:
        Processing reads; N_TOT_READS: 40000 N_COMPUTED_ALN: 2525 N_CACHED_ALN: 27023 N_COMPUTED_NOTALN: 1534 N_CACHED_NOTALN: 8918
```

# Post analysis steps

- Interpret allele frequencies and mutation types.

  **Assess Editing Efficiency**: Examine the percentage of edited alleles compared to the wild-type sequence. Low editing efficiency might indicate issues with sgRNA design, Cas9 delivery, or cell type.

  **Mutation Types**: Determine the predominant mutation types (insertions, deletions, substitutions) and whether they match the expected outcome (e.g., frameshift mutations, knockout efficiency).

  **On-target and Off-target Analysis**: Review any off-target effects or mutations in other regions if genome-wide CRISPR sequencing was performed.

- Validate your results with sequencing and functional assays.

  **Functional Assays**: If the study involves gene knockout or specific genetic modifications, functional assays can determine if the edited cells behave according to the hypothesis (e.g., loss of a phenotype, change in metabolic activity, or altered growth patterns).

- Troubleshoot any inefficiencies (redesign sgRNA, optimize delivery).

  **Low Editing Efficiency**: If CRISPResso2 shows a low rate of edits, troubleshoot by: redesigning the guide RNA (sgRNA) for better targeting. Optimizing the transfection or delivery method of Cas9 and the guide RNA (e.g., improving Cas9 or sgRNA concentration, delivery method). Ensuring that the Cas9 variant being used matches the goal (wild-type Cas9 for cutting, Cas9 nickase for precise editing, etc.).

- Isolate specific clones if necessary.

  **Clone Isolation**: If aiming for a homogeneous population of cells with specific mutations (especially for knockout studies or disease modelling), proceed with **clonal isolation**.

- Document and report results.

# Acknowledgements


## Developers of CRISPResso2

Luca Pinello Ph.D., Associate Professor
Kendell Clement, Assistant Professor

## Funding and Support

Quentin Hogg Trust, University of Westminster (https://quintinhoggtrust.org/)

## Genome Engineering Laboratory

(University of Westminster)

Dr Kalpana Surendranath, Associate Professor, University of Westminster
Dr John Murphy, Professor, University of Westminster
Dr Munuse Savash, Ph.D, University of Oxford
Dr Varsha Tieote, Ph.D, University of Westminster
Dr Nadine Solaiman, Ph.D, University of Westminster
Dr Ahmed Sidali, Ph.D, University of Westminster

All the M.Sc. and B.Sc. students who have contributed so much to the laboratory and the Gene Editors of the Future program.

## Gene Editors of The Future

(University of Westminster)

Big thank you to all the students from iterations 1-3 of this program

# Any questions?

# Contact


Email: **k.akram@westminster.ac.uk**
LinkedIn: https://www.linkedin.com/in/khalidakram/
Insta: @geneeditors