Intro to CRISPR-Cas9

... And Trends in CRISPR Mediated Genetic Modification

Sarah Davis, 2019

Table of Contents

1. Why present this topic?
2. What is it?
3. How does it work?
4. What is the significance?
5. Can CRISPR really be used effectively on humans in the future?
6. What happened in November?
7. How can we apply computing science to this system?

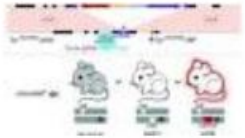
Why present this topic?

Top stories



Scientists engineer new CRISPR platform for DNA targeting

MIT News · 1 day ago



CRISPR/Cas9 used to control genetic inheritance in mice: Technology offers powerful new genetic tools f…

ScienceDaily · 6 hours ago



CRISPR explained: The revolutionary tool that's transforming genetic engineering

CNET · 12 hours ago

snipped by me, January 23rd 2019

Why present this topic?

- ➔ LOTS of buzz surrounding it
- ➔ Tall claims of what it is and what it can do
- ➔ Genome editing will likely only increase in popularity with time
- ➔ Ethically questionable uses
- ➔ Computing science can improve it

First, a little background

What is it CRISPR-Cas9?

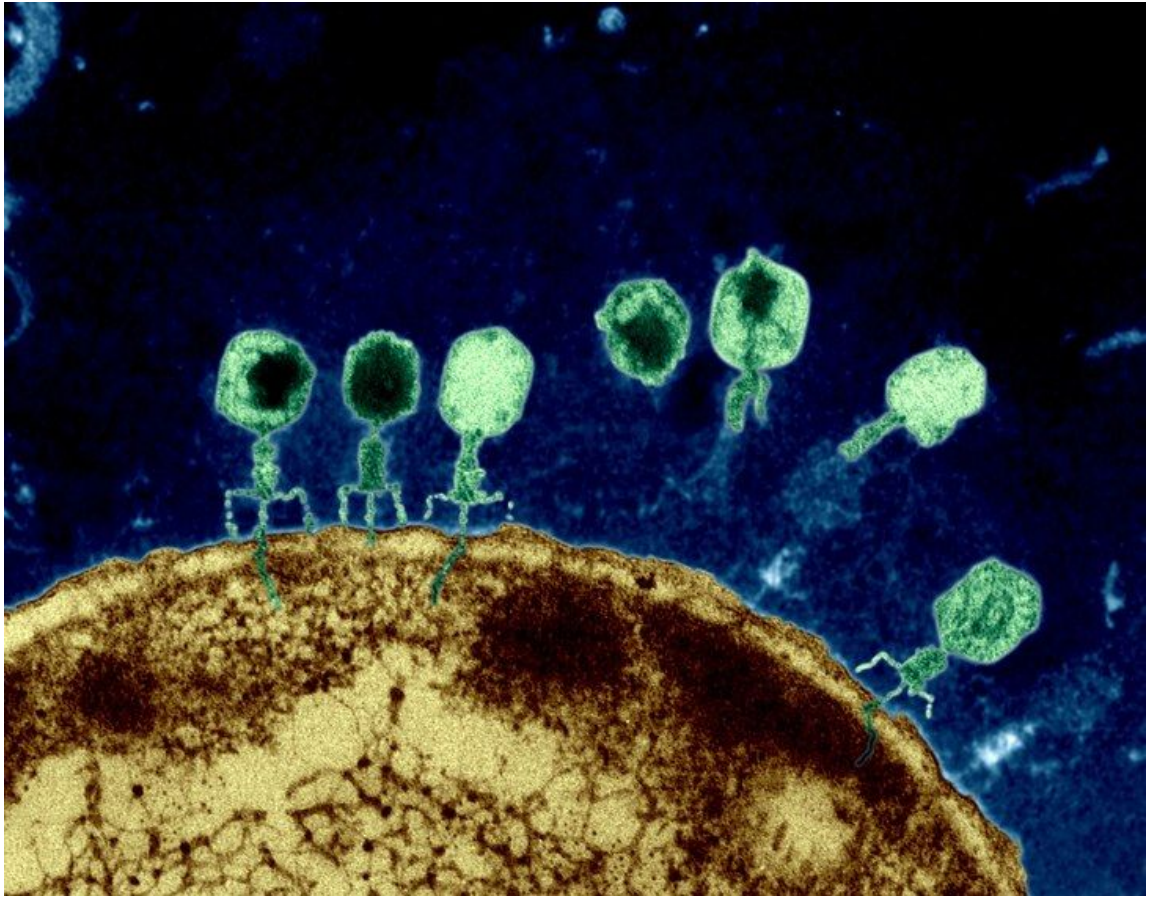


Image from Eye of Science / Science Source

What is CRISPR-Cas9?

- ➔ A DNA cutting mechanism harnessed from bacterial cell defense systems
- ➔ Allows geneticists and medical researchers to remove, add, or alter sections of DNA

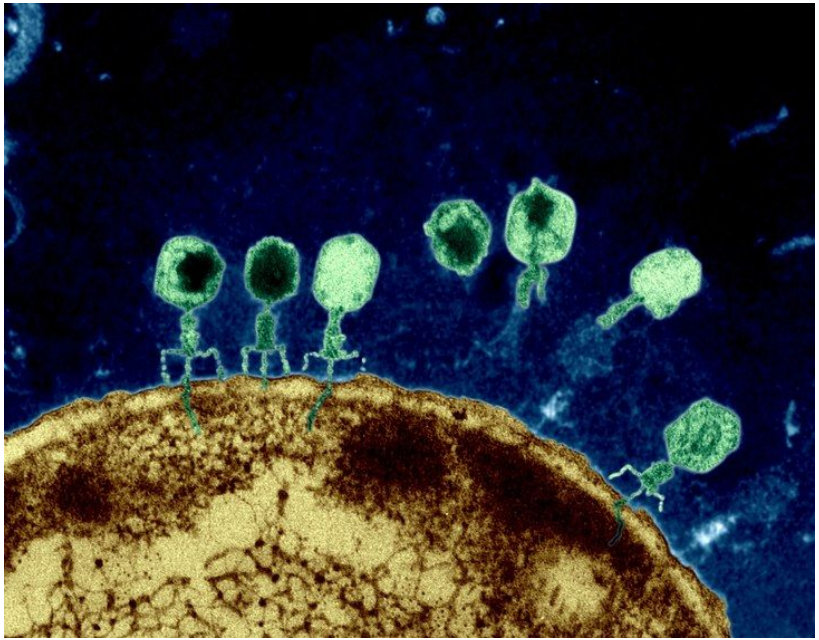
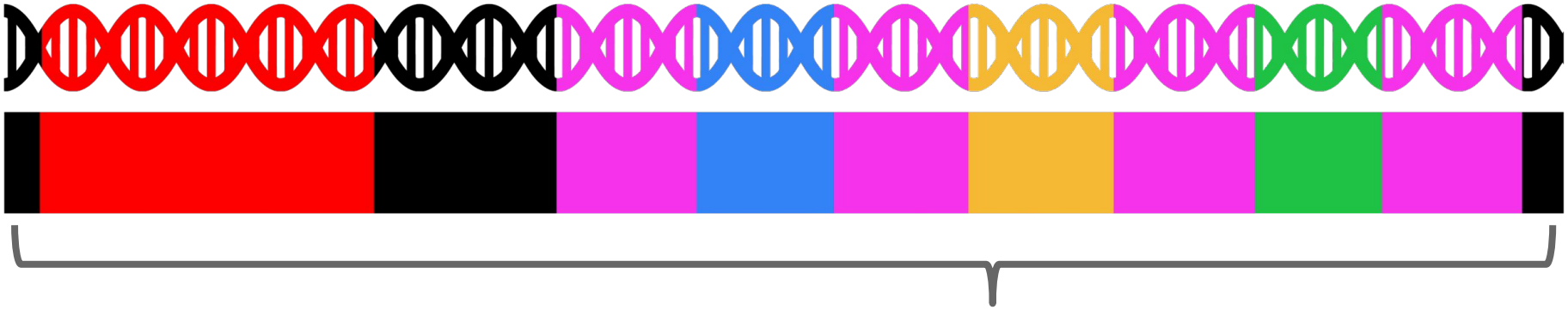


Image from Eye of Science / Science Source

What is CRISPR-Cas9?

- ➔ CRISPR stands for “Clustered Regularly Interspaced Short Palindromic Repeats”

- ➔ Cas9 refers to CRISPR Associated protein 9



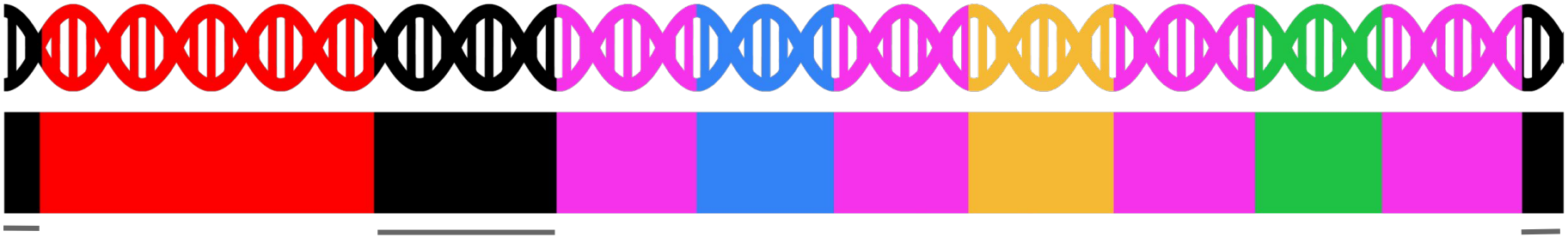
Other DNA

Cas genes

CRISPR repeats

Spacer DNA

E.coli DNA



Other DNA



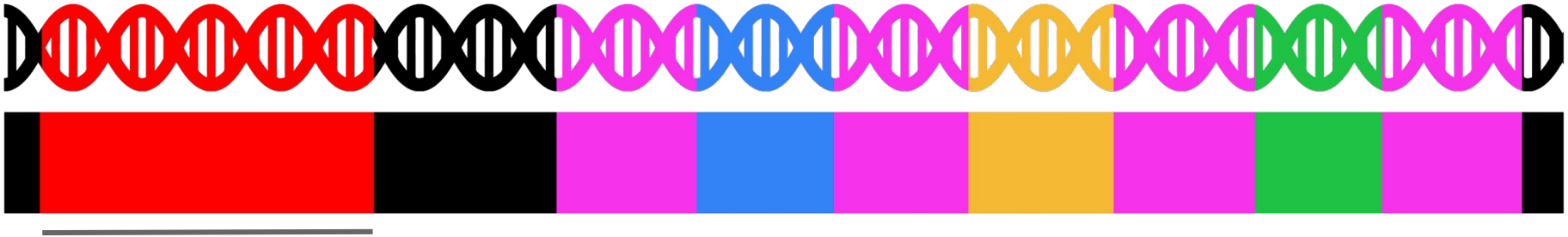
Cas genes

CRISPR repeats

Spacer DNA

Other DNA:

- Filler / miscellaneous (not important to CRISPR system)



Other DNA



Cas genes



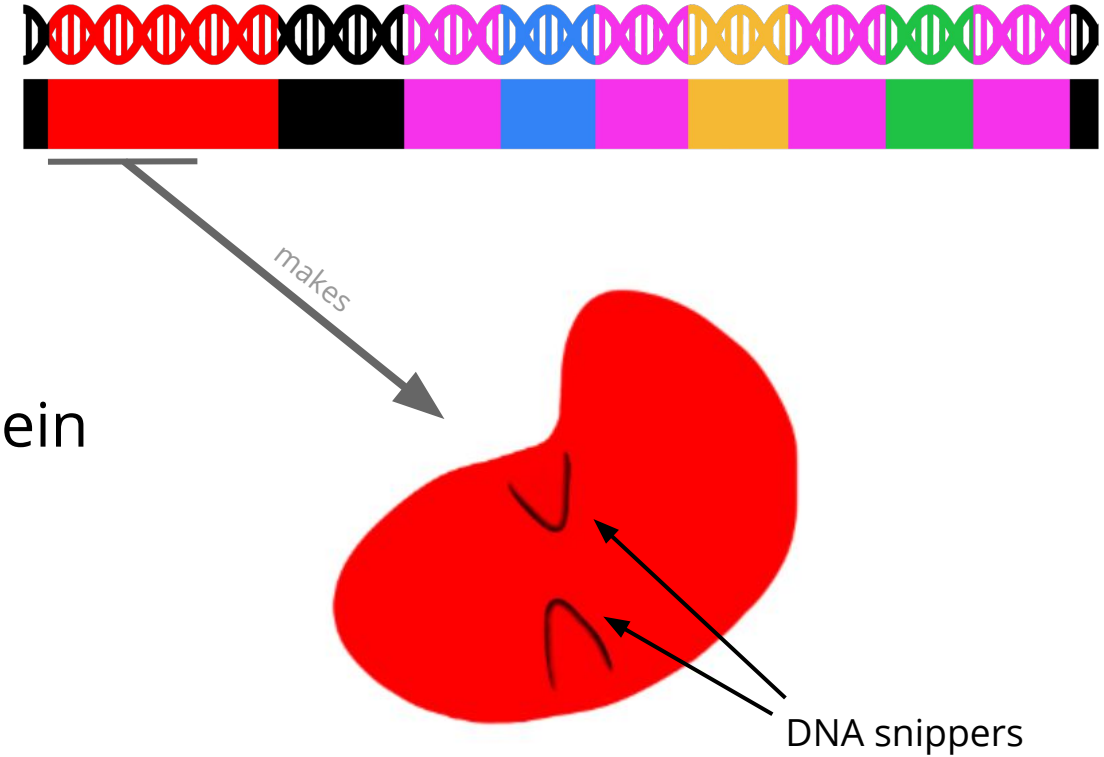
CRISPR repeats

Spacer DNA

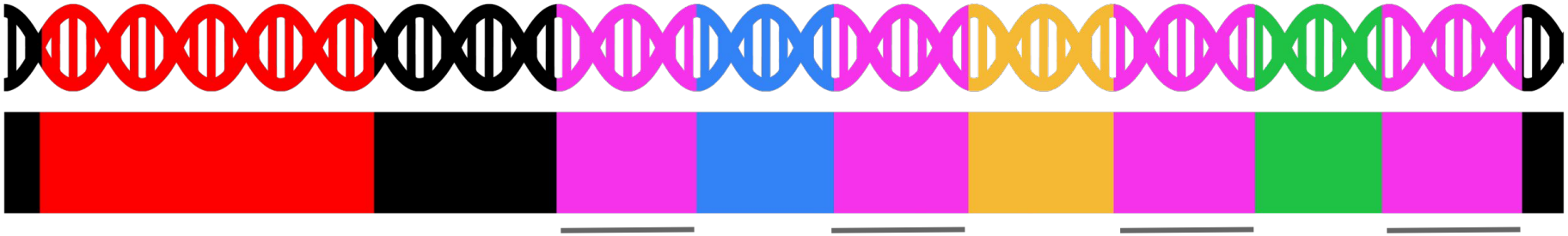
Cas Genes:

- Codes for proteins that cut and unwind DNA

Mechanism



1. Make a Cas protein



Other DNA



Cas genes



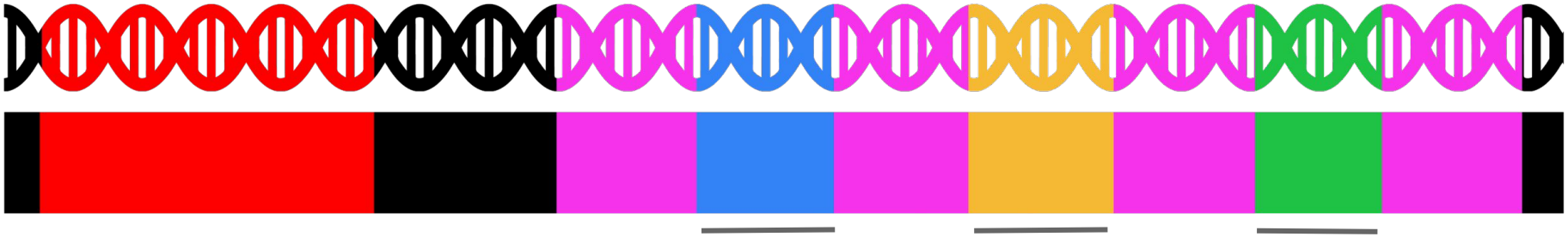
CRISPR repeats

ACTGCAGT
TGACGTCA

Spacer DNA

CRISPR repeats:

- ❑ Identical sequences
- ❑ Codes for a product that sticks to Cas protein



Other DNA



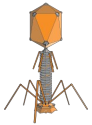
Cas genes

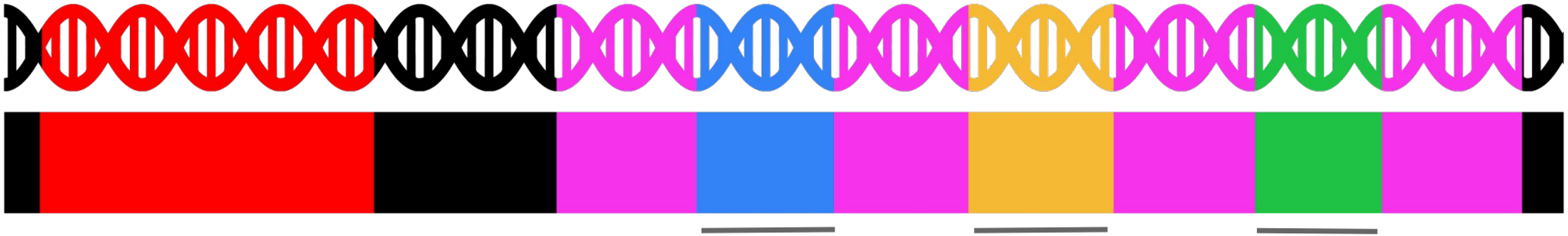


CRISPR repeats

ACTGCAGT
TGACGTCA

Spacer DNA





Other DNA



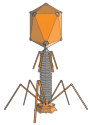
Cas genes



CRISPR repeats

ACTGCAGT
TGACGTCA

Spacer DNA



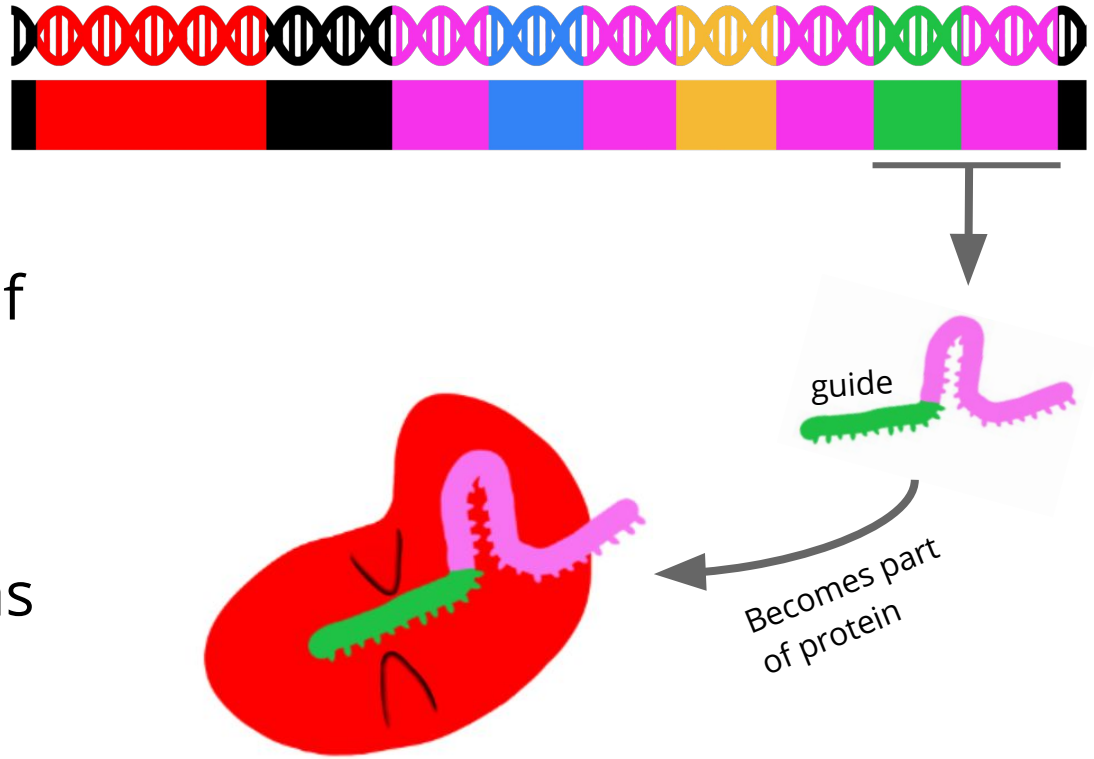
Spacer DNA:

- Contains sequences copied from viruses
- Allows bacteria to recognize viruses it had encountered before
  - Like immunization!

Mechanism

2. Make a “guide” out of some CRISPR DNA and some spacer DNA

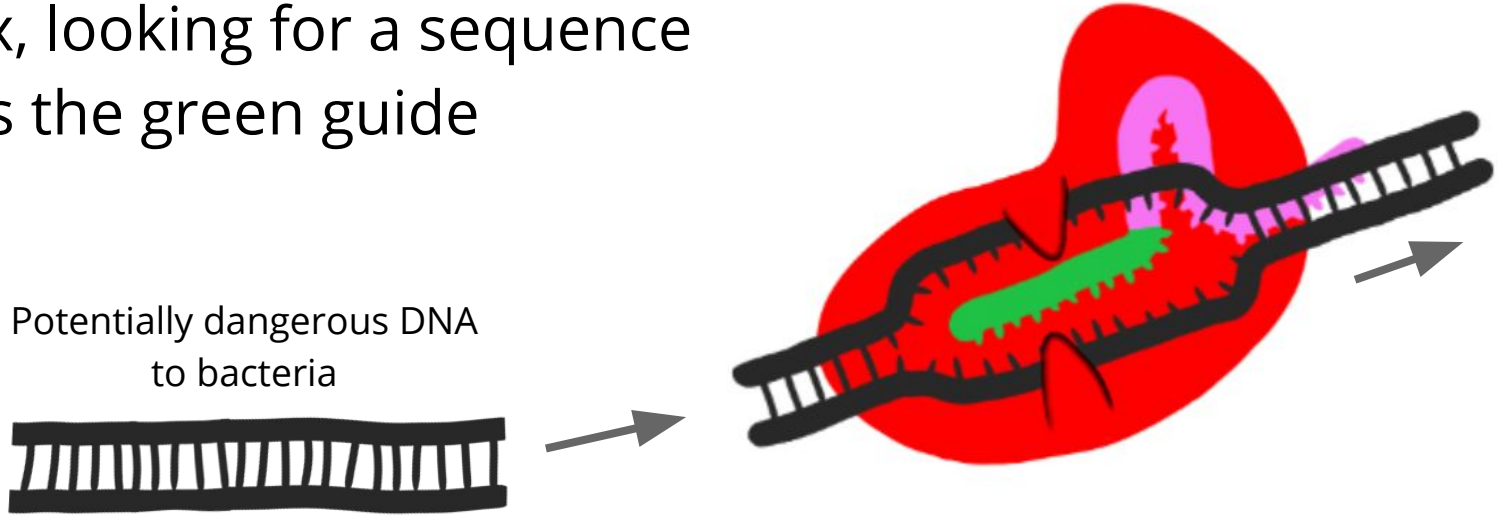
3. Guide attaches to Cas protein



Mechanism

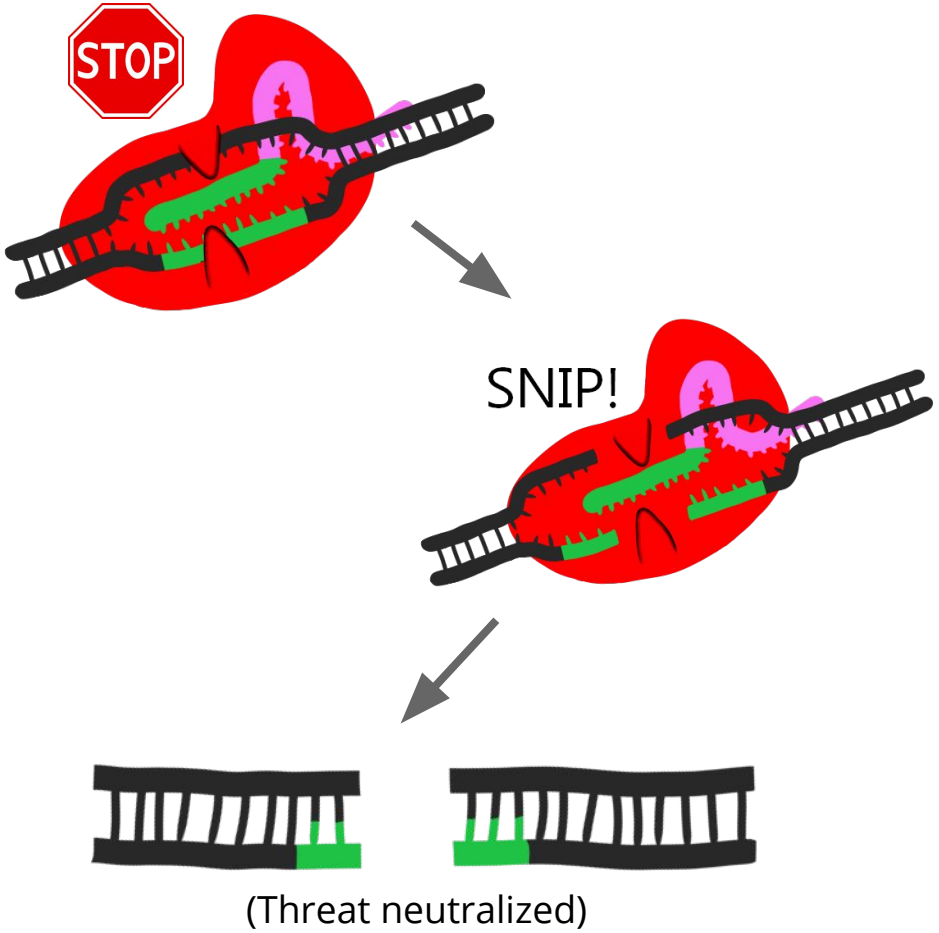
4. Feed DNA through protein complex, looking for a sequence matches the green guide

Potentially dangerous DNA to bacteria



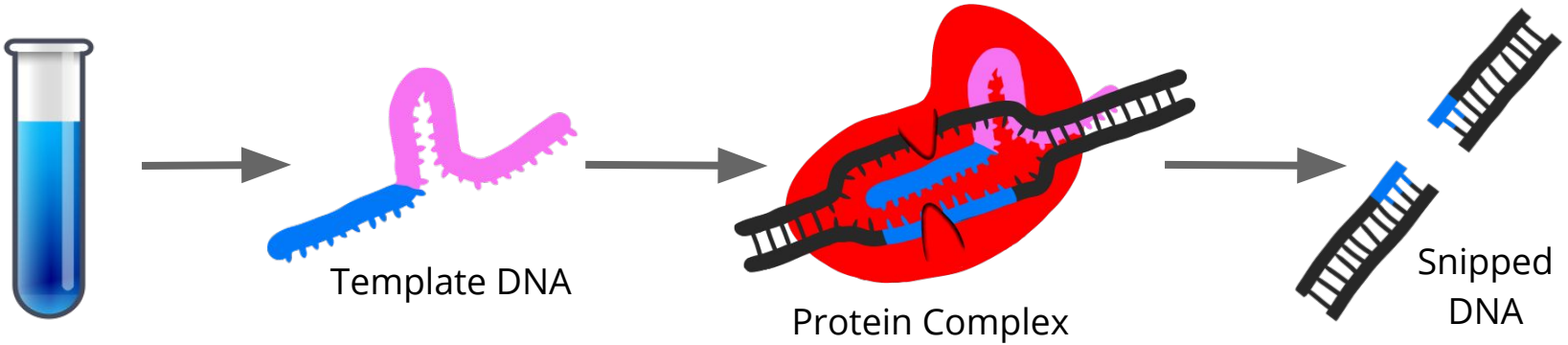
Mechanism

5. If / when the protein finds the matching green sequence, cuts DNA



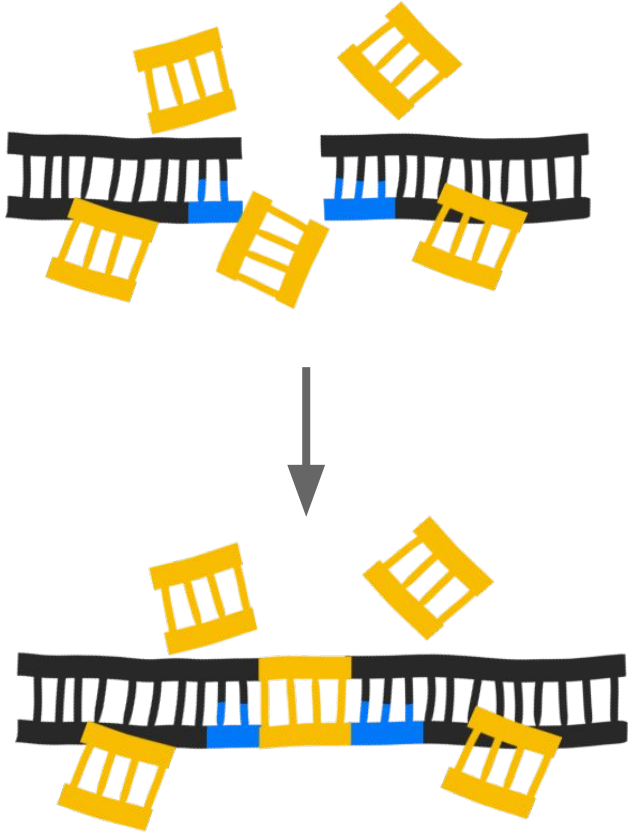
How does this apply to other organisms?

- ➔ Scientists synthesize their own “template” DNA copy in the lab
- ➔ Release protein complex into host genome, ready to snip when it finds the blue sequence



One last step...

- ➔ Flood area with DNA you want to incorporate into the genome!
- ➔ If you're lucky, it will be incorporated into the host genome



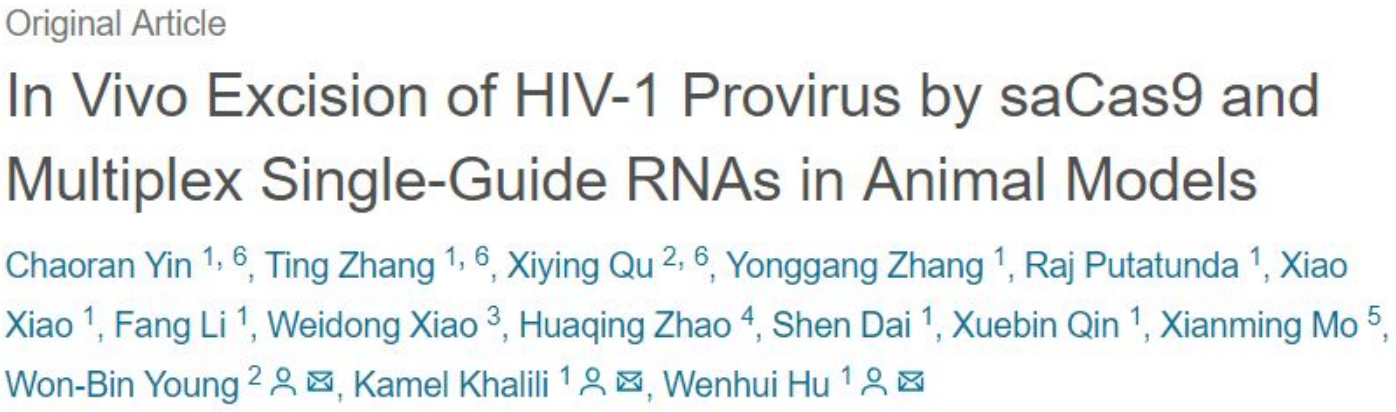
What is the significance?

Biological Significance

- ➔ There are many possible CRISPR applications
  - ◆ Everything alive has DNA, some things are aren't alive have DNA as well
- ➔ Most CRISPR articles currently being published seek to understand the mechanisms more thoroughly
- ➔ However, there are a few papers with tangible, concrete results...

Problems We're Applying CRISPR To

- ➔ HIV
  - ◆ Can suppress HIV transcription and delete HIV structural/functional genes in animal models



Original Article

In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models

Chaoran Yin [1, 6], Ting Zhang [1, 6], Xiying Qu [2, 6], Yonggang Zhang [1], Raj Putatunda [1], Xiao Xiao [1], Fang Li [1], Weidong Xiao [3], Huaqing Zhao [4], Shen Dai [1], Xuebin Qin [1], Xianming Mo [5], Won-Bin Young [2] ✉, Kamel Khalili [1] ✉, Wenhui Hu [1] ✉

Problems We're Applying CRISPR To

- ➔ Muscular dystrophy
  - ◆ Duchenne's MD: Fatal disease caused by striated muscles, used CRISPR to correct DMD mutations in animal models



RESEARCH ARTICLE | GENETICS

CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice

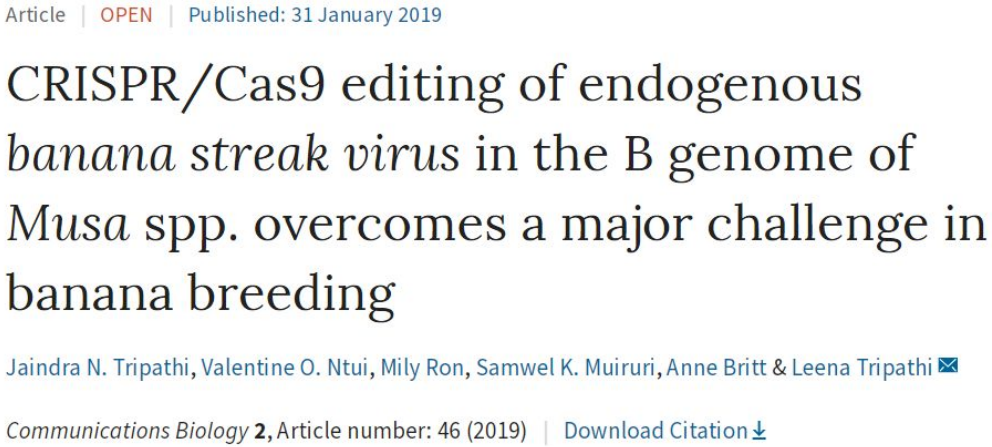
Yu Zhang[1,2,3,*], Chengzu Long[1,2,3,*,†,‡], Hui Li[1,2,3], John R. McAnally[1,2,3], Kedryn K. Baskin[1,2,3], John M. Shelton[4], Rhonda …

+ See all authors and affiliations

Science Advances 12 Apr 2017:
Vol. 3, no. 4, e1602814
DOI: 10.1126/sciadv.1602814

Problems We're Applying CRISPR To

- Various agriculture diseases
  - Certain foods are very susceptible to viruses and fungal infections
  - Edited out viral genes in bananas



Article | OPEN | Published: 31 January 2019

CRISPR/Cas9 editing of endogenous *banana streak virus* in the B genome of *Musa* spp. overcomes a major challenge in banana breeding

Jaindra N. Tripathi, Valentine O. Ntui, Mily Ron, Samwel K. Muiruri, Anne Britt & Leena Tripathi

Communications Biology **2**, Article number: 46 (2019) | Download Citation

Can CRISPR really be used effectively on humans?

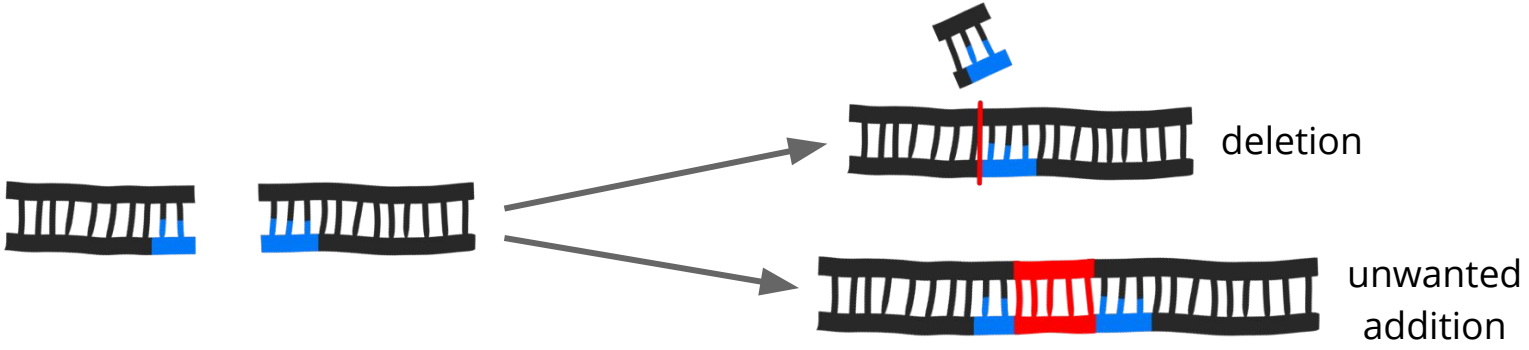


Sarah Davis, 2019

Immediate Problems

1. CRISPR has been shown to induce unwanted mutations

    A. When both strands of DNA are broken, the cell goes into panic mode
        - A sort of “sloppy” repair mechanism is activated
        - Might delete sections of DNA / insert incorrect sequences



Immediate Problems

1. CRISPR has been shown to induce unwanted mutations

    A. When both strands of DNA are broken, the cell goes into panic mode
        - A sort of “sloppy” repair mechanism is activated
        - Might delete sections of DNA / insert incorrect sequences
    B. CRISPR has been known to target any sequence that matches its guide

Immediate Problems

1. CRISPR has been shown to induce unwanted mutations

    A. When both strands of DNA are broken, the cell goes into panic mode
        - A sort of “sloppy” repair mechanism is activated
        - Might delete sections of DNA / insert incorrect sequences
    B. CRISPR has been known to target any sequence that matches its guide
        - What happens if that sequence happens to be found elsewhere on the genome?

Immediate Problems

2. CRISPR is not effective as previously thought

    A. A recent paper found that 85% of blood samples tested carried an antibody against the Cas9 protein
        - This greatly reduces effectiveness, as protein is destroyed by human immune system
    B. There is a lot of randomness in the cell, further decreasing effectiveness
        - Cas9 protein not guaranteed to reach its target simply due to physical barriers
        - Many endogenous proteins need access to DNA too, could knock off Cas9

Immediate Problems

3. People have around 30 trillion human cells as adults

    A. Germline vs. Somatic



Germline Mutation:
Will be passed on to future offspring, and will likely affect all the cells in the body



Somatic Mutation:
Occurs in a single cell, only tissues derived from that cell will be affected

Immediate Problems

3. People have around 30 trillion human cells as adults

    A. Germline vs. Somatic

        a. Somatic: Attempting CRISPR gene therapy on an adult is difficult, as one would have to attempt to change every affected cell
        b. Germline: CRISPR can be used on an embryo, but this makes the induced mutation a permanent part of their genome
            i. Mutation would be passed along to their children

This brings up all sorts of ethical issues...

Ethical Concerns

Ethical Concerns (As Outlined by OmicsOnline)

- ➔ Balance of risks and benefits in terms of mutagen effects
  - ◆ Previously discussed
- ➔ Threat to the ecosystem
  - ◆ CRISPR has a biologically active product - the spread of a new phenotype may be hard to control (new invasive species?)





Ethical Concerns (As Outlined by OmicsOnline)

- ➔ Balance of risks and benefits in terms of mutagen effects
  - ◆ Previously discussed
- ➔ Threat to the ecosystem
  - ◆ CRISPR has a biologically active product - the spread of a new phenotype may be hard to control (new invasive species?)
- ➔ Regulation for Consumers
  - ◆ Should people be able to patent new transgenic organisms?



A genetically engineered salmon, top, and a regular one. Paul Darrow for The New York Times

Ethical Concerns (As Outlined by OmicsOnline)

- ➔ Application of CRISPR-Cas9 to the human germline
  - ◆ No way to practice informed consent with embryo
  - ◆ Should it be allowed to therapeutic reasons?
  - ◆ Who is liable for possible damage to future generations?
- ➔ Genome editing for enhancement
  - ◆ Should parents be allowed to choose what they want their baby to look like?
  - ◆ Socially, there would be a problem if some population has an advantage over others because of edited genetics

Recent Controversy

- ➔ Dr. He Jiankui became widely known in November of 2018 for announcing he had created the first human germline genetically engineered babies
  - ◆ Two twin girls, named Lulu and Nana
- ➔ These girls were carried to term, with seemingly no health defects thus far

Recent Controversy

- ➔ Dr. He used CRISPR to disable a gene called CCR5
  - ◆ CCR5 is exploited by HIV to infect human cells
- ➔ He’s data wasn’t published or peer reviewed, so the details of the experiment are still somewhat unclear
- ➔ However, according to the slides he prepared for the Human Genome Editing Summit:
  - ◆ The edit didn’t go to completion
  - ◆ The mutations the girls did receive were completely novel, never before tested

Recent Controversy

- ➔ Dr. He was nearly universally condemned for his actions
- ➔ The university he worked for cut ties with him not too long ago



How can computing science be applied?

Where does computing science come in?

- ➔ As we saw previously, a CRISPR experiment is not a guaranteed success
- ➔ Experiments are starting to generate a lot of data about CRISPR effectiveness in different conditions / places

Target gene: POLR2A

Each row represents one screening experiment

Screen ID	Cell line	Cas-9 variant	Screen type	Screen	Performance	Hit
GCR1671 (Wang et al.)	KBM7	hSpCas9	negative selection	viability		Yes
GCR2031 (Wang et al.)	Jiyoye	hSpCas9	negative selection	viability		Yes
GCR1788 (Wang et al.)	Raji	hSpCas9	negative selection	viability		Yes
GCR1630 (Wang et al.)	K562	hSpCas9	negative selection	viability		Yes

Search Results for “Machine Learning CRISPR”

A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action

Shiran Abadi, Winston X. Yan, David Amar, Itay Mayrose

Version 2 ▾ Published: October 16, 2017 • https://doi.org/10.1371/journal.pcbi.1005807

Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation

Max A Horlbeck[1,2,3,4], Luke A Gilbert[1,2,3,4], Jacqueline E Villalta[1,2,3,4,8], Britt Adamson[1,2,3,4], Ryan A. Pak[1,5], Yuwen Chen[1,2,3,4], Alex P Fields[1,2,3,4], Chong Y Park[1,5], Jacob E Corn[5,6], Martin Kampmann[1,2,3,4,7], Jonathan S Weissman[1,2,3,4,†]

In Silico Meets *In Vivo*: Towards Computational CRISPR-Based sgRNA Design

Guo-hui Chuai [1, 3], Qi-Long Wang [2, 3], Qi Liu [1]

Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets

4 Author(s) Michael K. K. Leung ; Andrew Delong ; Babak Alipanahi ; Brendan J. Frey View All Authors

Where does computing science come in?

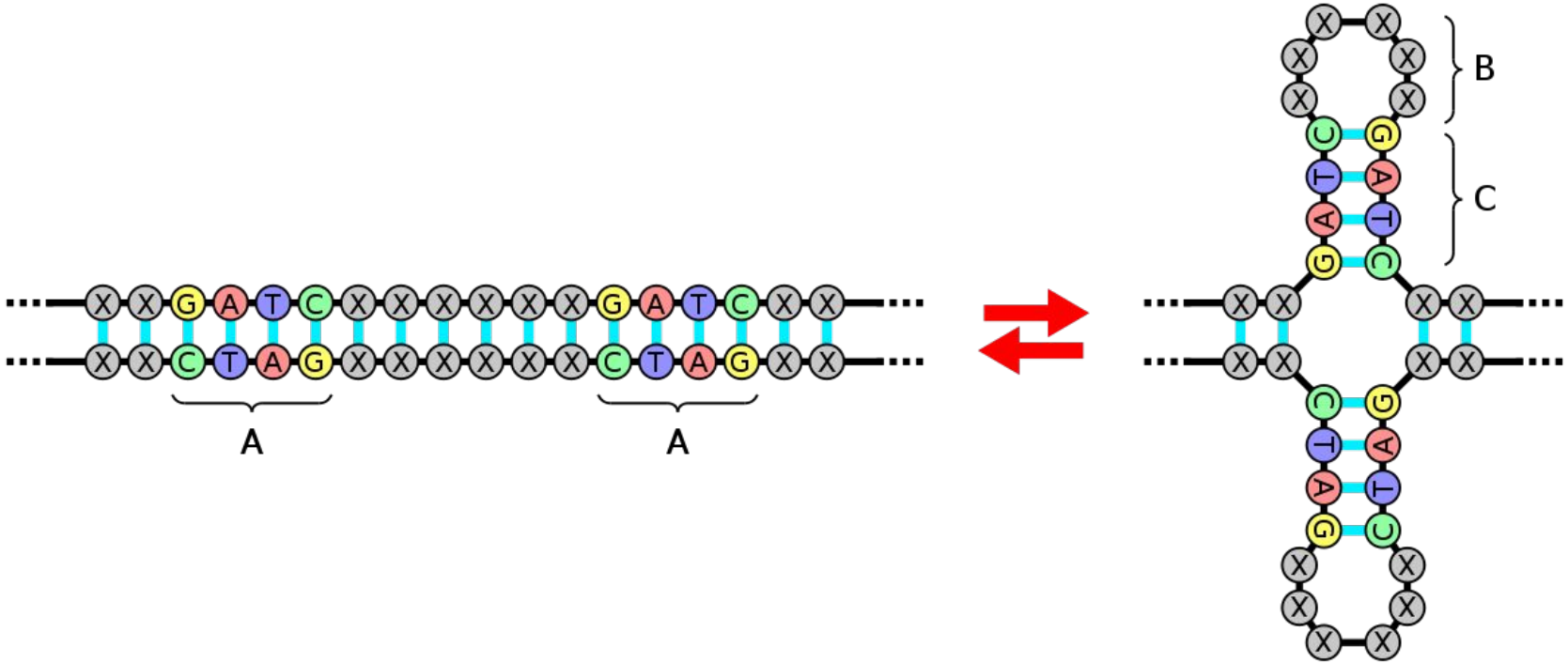
- ➔ Machine learning has great potential to improve the efficacy of CRISPR-Cas9 screenings
  - ◆ Training classifiers to predict best possible combinations of experimental decisions and environments
- ➔ In doing this, CRISPR becomes safer and more reliable

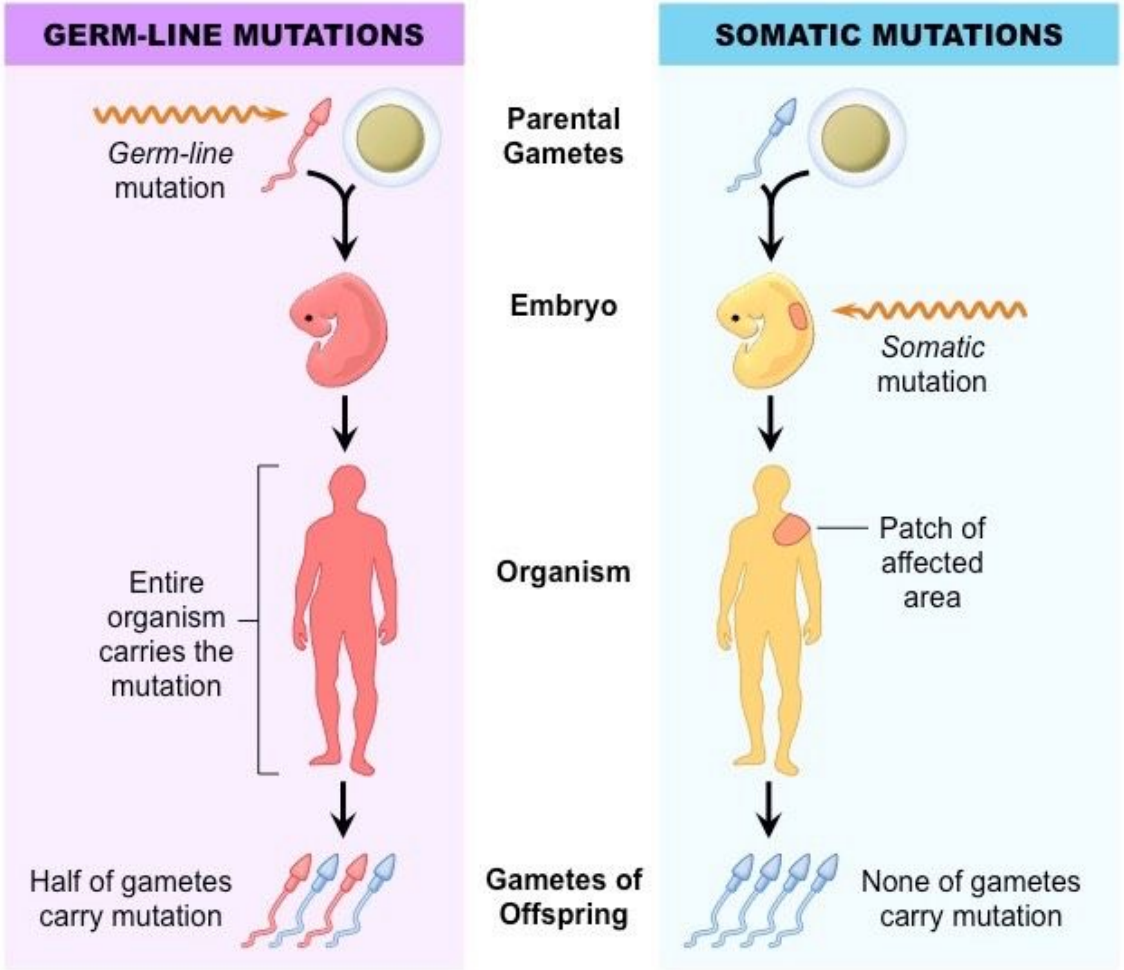
Takeaways

- ➔ CRISPR has potential, but humans are far from being experts at it
- ➔ It is possible CRISPR is not viable as a gene editing tool on a large scale
- ➔ People generally shouldn't go against the advice of the greater scientific community
- ➔ The fusion of machine learning and gene editing tools is just beginning

Thank you to Dr. Mike Harrington of the Department of Biology!

Palindromic DNA





Sarah Davis, 2019