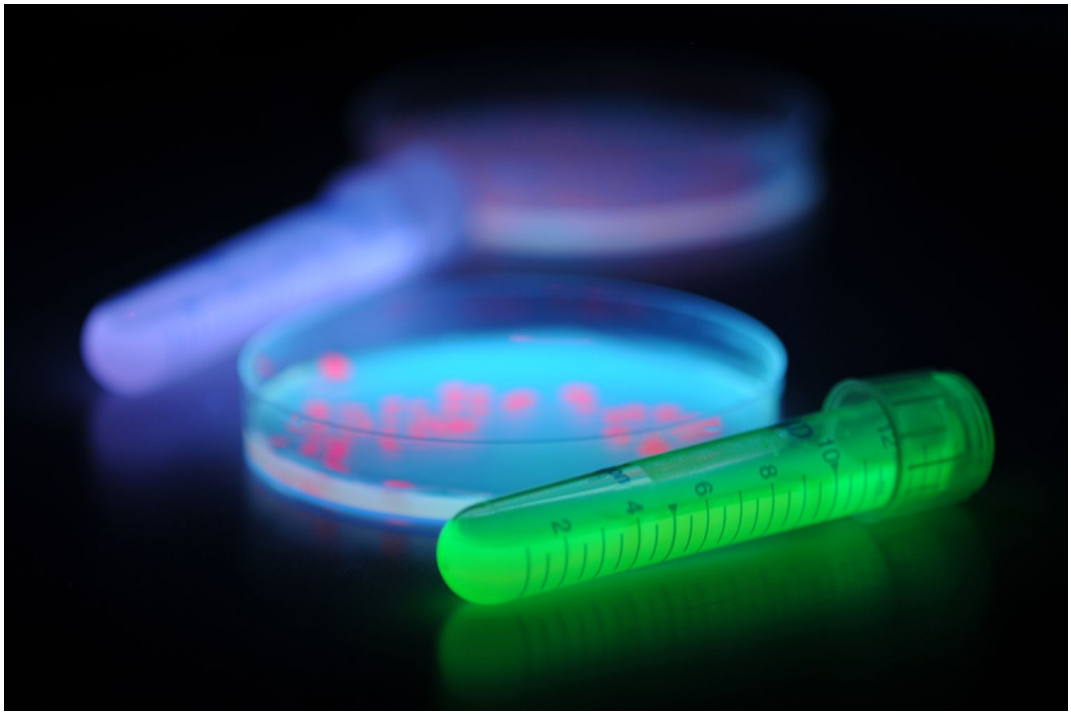
The Use of CRISPR-Cas9 for the Purposes of Bioluminescence

Swan Portevin, Lenn Jakobsen and Ian Brouwer
Gymnasium Kirschgarten

The revolutionary CRISPR-Cas9 gene editing system has sparked a scientific renaissance, offering unprecedented precision and efficiency in genetic manipulation. In the realm of fluorescent proteins, this technology has opened new avenues for exploration and innovation and can among other things be used to study the development of organ systems with unparalleled clarity and specificity. This paper aims to explain how the production of such fluorescent proteins can be achieved especially with the help of CRISPR-Cas9 in prokaryotic cells. We will also give a brief overview of the scientific history of CRISPR-Cas9, explain how it operates and discuss its possibilities in an interview with Prof. Dr. Markus Affolter leader of a research group at Biozentrum Basel investigating the development of cellular networks with the help of fluorescent proteins.



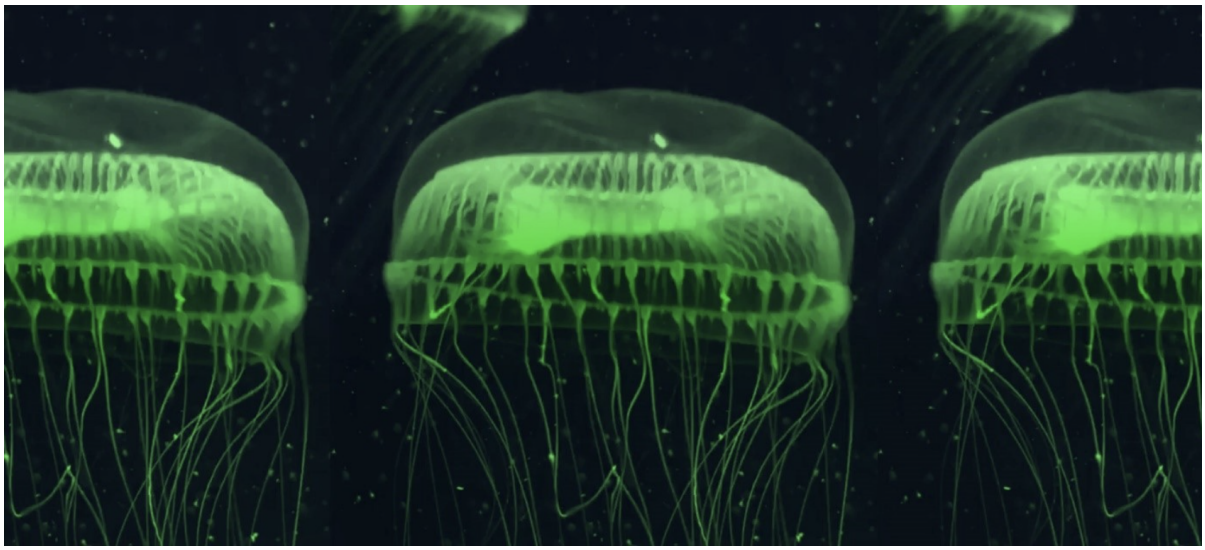
Introduction

Ever since the 2020 Nobel Prize in Chemistry has been awarded to Emmanuelle Charpentier and Jennifer A. Doudna for developing the use of CRISPR-Cas9 as a gene-editing technique[1] the technology has remained highly controversial and of great public interest. While first attempts at making inheritable changes to the human genome by Chinese scientist He Jiankui[2] have made headlines in 2019, we cannot help but think the public's knowledge of CRISPR-Cas9 might be limited to its applications in the potential treatment of genetic diseases and the modification of the human genome in general. In this paper we therefore wish to present a much overlooked application of CRISPR-Cas9 to show the diversity of its possible uses.

Fluorescent proteins are often essential for labelling, organising and tracking cells and even proteins[3][4]. This is especially crucial in the study of dynamic systems like the development of cellular networks such as blood vessels or organs. While the labelling of bacteria and other prokaryotic cells can be achieved by inserting a plasmid responsible for producing a fluorescent protein like pGLO which will be discussed in a later section, the process becomes more complicated for eukaryotic cells as the transcription of DNA is carried out in the nucleus[5], thus sometimes making a simple insertion of DNA into the cell unsuitable as the DNA would have to enter both the cell and the nucleus[6]. However, methods of nucleic acid insertion into eukaryotic cells do exist and are widely used[7], among the most notable being stable and transient transfection, which make use of viral vectors[7]. However, compared to these methods editing the cells genome with CRISPR-Cas9 offers certain advantages, for example by making it possible to exploit pre-existing regulatory structures.

Description of engineering technique

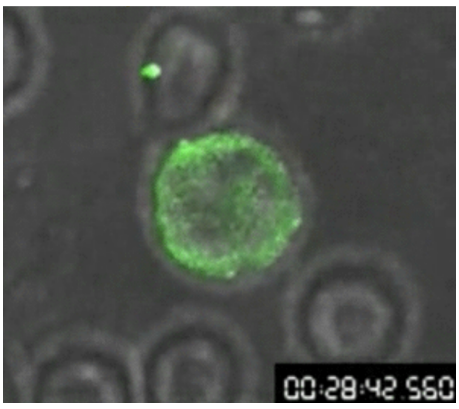
pGLO– The Fluorescent Plasmid that Caused a Scientific Revolution



The Jellyfish Aequorea victoria Glows After Exposure to UV Light[a]

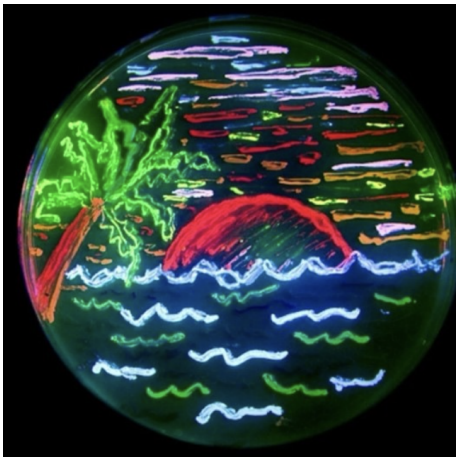
pGLO is a genetically engineered plasmid, a type of small DNA-molecule that can independently replicate and has its own genome[8]. It resides symbiotically within the cell and is comparable to mitochondrial plasmids except the fact that it is created by scientists in the lab. pGLO's defining feature is that it causes its host to glow under the spectrum of blue light through ultraviolet radiation (UV)[9]. It owes this defining trait to a protein it causes the synthesis of called the Green Fluorescent Protein (GFP). This protein absorbs the light of the spectrum mentioned above and then emits a green fluorescence with a wavelength of 509nm[9].

The pGLO plasmid contains not only the gene needed to produce GFP, but also usually two others. The first of the two, araC is a regulatory gene that affects the synthesis of GFP. The second is the gene Bla that produces the enzyme beta-lactamase, which makes the pGLO resistant against the beta-lactam family of antibiotics, the family of antibiotics that includes penicillin[10]. GFP also is very well tolerated in many organisms[11]. These traits of the plasmid make it ideal to be inserted into cells that need tracking for research purposes. This can include transmission of diseases between cells. An example of a use-case is studying the HIV-virus and how it passes between cells[10].



Cell Transfer of HIV with pGLO[a]

The scientists Osamu Shimomura credited with the discovery of GFP in 1962, Marty Chalfie who managed to produce GFP in organisms such as flatworms and Roger Tsien who created early GFP mutants, were all awarded the Nobel Prize in 2008 for the discovery and development of the Green Fluorescent Protein[12].



A Beach with Mutand pGLO[b]

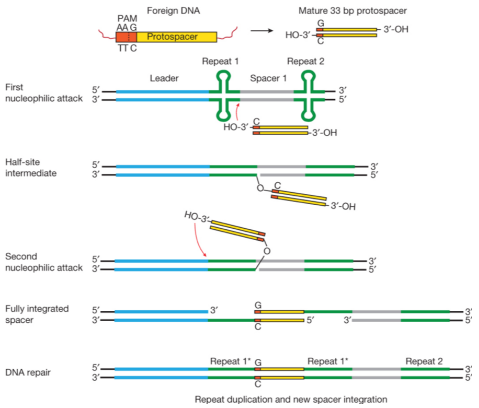
CRISPR-Cas9: The next scientific revolution

CRISPR-Cas9 gene editing was first discovered as a defence mechanism in the adaptive immune system of bacteria[13]. The term "adaptive" immune system refers to the immune functions that use previous encounters with an attacker to defend against future attacks. More specifically CRISPR-Cas9 is a form of protection against insertion of alien DNA into the bacteria, as for example used by bacteriophages as a means to reproduction[14]. It achieves this by saving parts of the inserted DNA and then recognising and destroying it in the future. In this section we first want to have a closer look at how each of these steps is achieved in nature and will then explain how they can be combined with DNA repair mechanisms to add genes to produce fluorescent proteins.

To remember a part of foreign DNA the bacteria makes use of arrays of so called clustered regularly interspaced palindromic repeats (short CRISPR-arrays) these are sections of the bacteria's genome made out of multiple identical sequences known as palindromic repeats following a single promoter sequence[15]. The most important two properties of palindromic repeats are first of all that they're an equal number of nucleotides away from each other and secondly, as the name suggests, that they are read the same on both strands (a palindrome is a word that reads the same forwards as backwards), that means you obtain the same sequence of nucleotides regardless of the strand you read, as long as you read both strands from 3' to 5' or 5' to 3'. This second property of "palindromicity" helps endonucleases identify them[16].

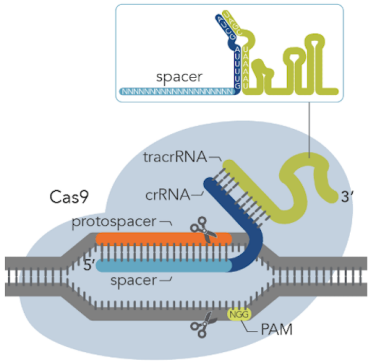
The DNA-sequences between these palindromic repeats are as mentioned before all of equal length and known as spacers. These spacers are at the core of the memory process of the defence mechanism.

When a piece of DNA is inserted into the bacteria, a part of this DNA is integrated into this cluster of palindromic repeats as a new spacer. The sequence to become a spacer is known as a protospacer and is cut out of the inserted DNA molecule by the collective action of the two enzymes Cas1 and Cas2[13][17]. It is then inserted between the promoter and the adjacent palindromic repeat. In this process the palindromic repeat is actually unwound so that each of the strands is on one side of the new spacer, this way after the repair and ligation of the CRISPR-array there will be a new palindromic repeat between the promoter and the new spacer, creating an new, elongated, fully functional CRISPR-array.



DNA Acquisition in CRISPR-Cas Systems[c]

During an attack the CRISPR-array is transcribed into RNA. This RNA strand that contains all spacers and palindromic repeats of the CRISPR-array is known as pre-CRISPR RNA (pre-crRNA)[18][19]. The transcription of a separate gene outside of the CRISPR-array produces so called trans-activating CRISPR RNA (tracrRNA)[18]. This tracrRNA contains sequences that are complementary to parts of the palindromic repeat[19], causing tracrRNA to hydrogen-bond to the palindromic repeats in the pre-crRNA. An enzyme called RNaseIII now cuts the pre-crRNA's palindromic repeats thereby cutting the tracrRNA bonded to it[18][19]. The resulting fragments containing one spacer, one part of a repeat and the part of the tracrRNA that is bonded to it are known as cr:tracrRNA. This cr:tracrRNA bonds to the endonuclease Cas9[20] and is now commonly referred to as guide RNA (gRNA). The Cas9 enzyme is a single polypeptide containing a domain capable of identifying the PAM: the PAM interacting domain (PI).



The gRNA-Cas9 Complex in Action[d]

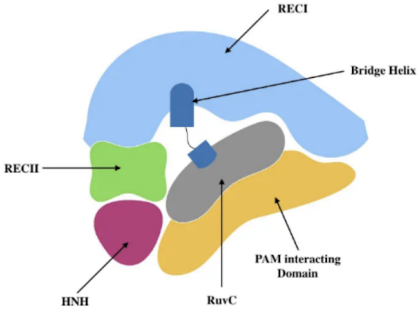


Diagram of Cas9 Domains[e]

To distinguish between the memorised, foreign DNA-sections and the identical spacer in its own DNA, the bacteria makes use of a so called Protospacer Adjacent Motif (short PAM). A PAM is a short sequence of nucleotides that is exploited as a sort of beacon for the Cas1 and Cas2 enzymes[13]. The enzymes will read through the intruding DNA until they find a PAM on the complementary strand and will then choose the sequence before the PAM (in the 3' direction on the strand not containing the PAM).

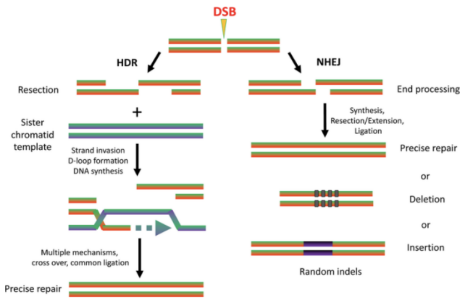
From this follows that the defence mechanism can only be effective against DNA containing the PAM, therefore the PAM should be short and common. It furthermore can't be contained in the palindromic repeats, or the spacer, as that would make a distinction between the alien DNA and the bacteria's own spacers more difficult. The latter is solved by choosing the sequence before the PAM as the protospacer, this way the protospacer can't contain the PAM as it would have stopped earlier and the sequence before that PAM would have been chosen instead.

The former is a mere matter of choosing a palindromic sequence without the PAM as a palindromic repeat. The canonical PAM i.e. used by the well-known bacteria Streptococcus pyogenes is 5'-NGG-3', where N can be any nucleotide[21].

If the PI identifies a PAM in a DNA molecule, it will unwind the two DNA strands in the 3' direction of the strand not containing the PAM[19], which then allows the spacer part of the gRNA to hydrogen bond to the strand not containing the PAM, if that strand is complementary to the gRNA's spacer part[22]. This condition is equivalent to checking if the sequence of nucleotides on the strand not containing the PAM matches the one of the spacer in the bacteria's genome. Should this be the case, the gRNA, as mentioned above, bonds to the DNA strand that does not contain the PAM which causes the Cas9 enzyme to cut both DNA strands, 4 Nucleotides in the 3' direction on the strand not containing the PAM[22]. This is known as a double strand break (DSB).

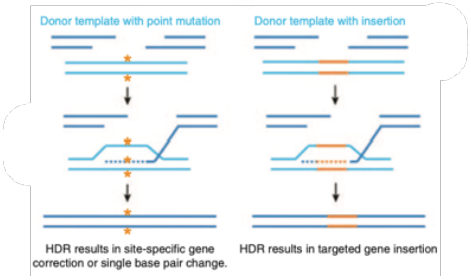
The synthesis of an artificial gRNA molecule therefore allows to cut DNA 4 nucleotides ahead of one specified PAM without affecting the DNA around the other PAMs. With sufficient knowledge of the DNA one wants to edit, this technology can therefore be used to cause a double strand break at a desired location.

There are two mechanisms to repair DSBs[23] the most common one being Non-homologous end joining (NHEJ)[20]. This mechanism is highly efficient[20] and therefore takes place in the majority of DSBs. It is however error prone and often causes insertions or deletions leading to frameshifts[20]. The second mechanism homologous recombination, also known as homology directed repair (HDR), is less efficient as NHEJ but is less error prone[24] and can be used to manipulate the reparation of the DNA.



Mechanisms of HDR and NHEJ[f]

During HDR the strands are repaired with the aid of a repair template, a piece of DNA that is mostly identical to both sides of the cleaved DNA, in nature this could for example be a sister chromatid. As there are multiple different exact HDR mechanisms[24] using both single- and double-stranded repair templates[25], we will not explain the details of the specific mechanisms here, but will rather focus on their broad similarities and their use in gene editing.



Gene-Editing with HDR[g]

The strands of the broken DNA with the 5' end at the DSB are resected by nucleases[24] creating an unpaired "overhanging" part on each strand. The template is then unwound around the position[24] of the DSB and the overhanging parts of the broken DNA repair on the unwound strand of the template. The exact way this reparation happens varies, as mentioned, greatly with the specific mechanism. In some only one strand is repaired directly from the template (i.e. SDSA) while in others both strands are repaired with the template. For our purposes it is only important that it is possible to insert a gene by providing a repair template (for example as a plasmid) that resembles both sides of the broken DNA, but contains the additional genetic information at the position of the DSB. After the targeted DNA has been repaired with this template it will contain the additional gene. Deletions and replacements are also possible[26] but rarely needed for our purpose, the production of fluorescent proteins. In practice, the following steps need to be taken to create fluorescent proteins in a procaryotic cell:

1. Choose a suitable position in the cell's DNA, it should be frequently expressed or regulated by the genes that you want to regulate your fluorescent protein production too, it also needs to be adjacent to a PAM.
2. Synthesise a suitable cr:tracrRNA capable of bonding to the nucleotides adjacent to the specific PAM chosen, alternatively a single molecule called single guide RNA (sgRNA) in all other aspects equivalent to cr:tracrRNA can be used[27][28].
3. Synthesise a repair template that includes a gene for the production of a fluorescent protein (i.e. the green fluorescent Protein GFP) at the position of the anticipated DSB but is otherwise homologous to the target DNA. The longer the added gene, the longer the homologous parts of the template need to be[29]. It is advisable to alter all PAMs in the template to a silent mutation, to avoid cleavage of the template by the sgRNA-Cas9-complex.

Further steps can be taken to increase the efficiency, like for example chemically inhibiting the genes involved in NHEJ to increase the fraction of DSBs repaired by HDR.

Documentation and pictures of research institutions visited

Markus Affolter[30] is a Swiss Developmental Biologist and Professor at the Biozentrum University of Basel, Switzerland. He studied biology at the ETH Zürich and at the Laval University in Canada and got his PhD in 1988. He works in Biozentrum Basel where he became a Professor of Developmental Biology in 2005.

Professor Markus Affolter[h]

-What motivates you to work in your field?

I have always been interested in how animals develop and have therefore gone in this research field (developmental biology)

-How did you get into this field?

As a student at ETH (1978 to 1980), and then a Bac, Master and PhD student in Quebec city, Canada, I followed the discovery of the Homeobox in Basel in 1984, and then absolutely wanted to join the lab of Walter Gehring in Basel, which I did in 1988.

-What technology did you use before CRISPR-Cas9?

Reverse genetics with transposable elements, then zinc fingers and TALEN, then CRISPR.

-What impact did CRISPR-Cas9 have on your work?

A huge impact, since anything was now possible, all kinds of genome modification, and this fast and cheap and simple.

-How do you use CRISPR-Cas9 in your lab and how might that differ from the techniques other labs might employ?

We use genome editing in order to tag endogenous genes (proteins) with short peptide tags, and then manipulate these tagged proteins with small nanobodies (see our papers on pubmed)[31].

-What do you hope your research can one day be used for?

To modify the function of proteins directly and in disease.

-Where do the limitations of CRISPR-Cas9 limit your research and what do you hope for in future gene-editing technologies?

For what we want to do, there are no limitations at the moment.

Discussion

History of CRISPR-Cas9

CRISPR-Cas9 was first discovered in 1993 by Francisco Mojica[32] who named it CRISPR locus. In the year 2000 he realized that what seemed like disparate repeat sequences weren't random but shared a common set of features. It led him to the correct hypothesis that CRIPR is an adaptive immune system. Another group of scientists, led by Pourcel, published a paper in 2005 with the same results[33]. In March 2006 Alexander Bolotin was working on the bacteria Streptococcus thermophilus and found an unusual CRISPR locus[34]. The CRISPR array differed from previously reported systems, by containing novel cas genes instead of some of the known cas genes, including one for an enzyme now known as Cas9. One year later Philippe Horvath and Danisco France researched how CRISPR-Cas9 would respond to attacks from bacteriophages[35]. They were able to show that CRISPR-Cas9 was indeed an adaptive immune response in which new phage DNA sequences were being added to the genome to counter the next phage attack. Thus proving that CRISPR-Cas9 can modify its user's own genome to fight viruses. They furthermore discovered that the enzyme used to inactivate the phages' attack was likely Cas9. A vital discovery was made in 2010 by Sylvain Moineau, he demonstrated that CRISPR-Cas9 caused double strand breaks fairly precisely 3 nucleotides upstream a protospacer adjacent motif[36]. In 2011 Emmanuelle Charpentier and her team were sequencing small RNA on Streptococcus pyogenes and discovered a new type of RNA, that they called tracrRNA that bound to the crRNA and Cas9 enzyme to locate the CRISPR-Cas9 system's target[37]. Later that year the successful cloning of an entire CRISPR-Cas locus by Siksnys et al. from the university of Vilnius Lithuania suggested that all the required components of the CRISPR-Cas9 system were known[38]. This led two members of the University of California, Emmanuelle Charpentier and Jennifer Doudna to use the protein Cas9 to target specific regions of the genome and change it to their will[39]. Making them the first to successfully use Cas-9 in a eukaryotic cell granting Charpen and Doudna the 2020 Nobel Prize in Chemistry[1].

Technical Discussion

Long before CRISPR-Cas9 became a viable gene-editing tool, there were already known methods that enabled the editing of genes[40]:

- Meganucleases,
- zinc finger
- Nucleases (ZFNs)
- TALENs

They all could target specific regions of the genome and separate it into two fragments. However, these technologies are not accurate and in comparison. CRISPR outperformed these earlier gene editing methods due to its exceptional efficiency and user-friendliness. A group of scientists calculated the cutting efficiency of CRISPR, TALEN and ZFN based URR (urea reduction ratio)[41] editing data at 96h. The three methods were ranked the following[42]:

- CRISPR-Cas9 with a score of 4.84
- TALEN with a score of 3.89
- ZNF with a score of 3.38

However, comparing single aspects can lead to false conclusions. So, to make sure, they compared further aspects such as the efficiency and specificity of SpCas9 and TALEN at the E6 and E7 promoter sites, which will not be discussed further here. CRISPR-Cas9 outperformed TALEN and it left no doubt that CRISPR-Cas9 is now the way to go for the editing of any eukaryotic genome.

Despite seeming like unrelated fields at first glance, progress in fluorescence microscopy has been strongly dependent on gene technology. For example it wasn't long after Martin Chalfie et al. discovered in 1992 that GFP can be expressed in organisms other than the jellyfish it was originally extracted from[11] that Stanley Falkow et al. developed a modified form known as EGFP that exhibited better folding at 37°C in 1996[43].

Ethical Discussion

Regarding the ethical aspects of CRISPR-Cas9, there are many arguments that need to be considered. Its use in the human genome is of controversy that our project does not touch on. Instead, other interesting techniques come into question. The use of CRISPR-Cas9 in the creation of bioluminescence for research purposes will therefore be the focus of this section.

There are some risks in the use of CRISPR-Cas9 in gene editing to create bioluminescent organisms. These include creating animals that will most likely never be able to mate as they aren't accepted by their own kind and have a decreased chance of being able to be released into the wild. Furthermore, there is some concern with influencing the natural state of a species by releasing too many gene-edited organisms capable of producing fertile offspring into the wild.

Fluorescent Monkey[a]

The positives of the use of gene edited organisms in research are considerable. They often provide much more ethical ways of doing studies, as many things that traditionally required the death of an organism can now be done in vivo. This leads to a reduction in animal suffering with little additional effort. It simplifies lab procedures in many cases such as removing the need for time intensive dissections, thereby increasing the efficiency of scientific research.

To conclude, the use of gene altering technologies, such as CRISPR-Cas9 or pGLO, to create bioluminescence can have both positive and negative consequences. It often simplifies research processes while frequently barely altering standard procedures in a laboratory. The future looks bright with the use of bioluminescent plasmids with better life quality awaiting more lab animals by the day and many newly study-able processes being investigated. Thus, we firmly believe that the use of bioluminescence in gene editing is of great value to the scientific community as a whole.

Summary

The significance of bioluminescence in biological research can hardly be overstated. It has become absolutely crucial for labelling, organizing, and tracking organisms, cells and even molecules. This is essential for studying dynamic systems like the development of cellular networks.

The same can be said about gene technology, which is in the eyes of many one of the most interesting developments in modern science and is often hoped to one day be useful for curing genetic disorders, developing resilient and nutritious crops, and advancing environmental protection efforts.

We hope to have shown in this paper that the overlap between these fields, despite being frequently overlooked in both the potential applications and the ethical discussion of gene technology, has yielded some of the most useful techniques in the field of biology. Among these of course being the traditional plasmid pGLO that has been used to produce GFP since shortly after its first extraction but still remains relevant even in modern scientific research.

Significant further progress has since been made and is to be expected with the introduction of CRISPR-Cas9 setting a new standard of precision and efficiency in the field of gene-editing and allowing for new possibilities in the creation of bioluminescence.

References

[1] Nobel Prize in Chemistry 2020 Press Release

[2] BBC: China jails 'gene-edited babies' scientist for three years

[3] Fluorescent labeling and modification of proteins

[4] ScienceDirect: biochemistry genetics and molecular biology fluorescent labelling

[5] LibreTexts Biology: Transcription in Eucaryotic cells

[6] The Scientist: Infographic possible mechanisms of gene transfer in eukaryotes

[7] Transfection types, methods and strategies: a technical review

[8] Application Note for the "pGLO™ Bacterial Transformation Kit Extension" by BIO RAD

[9] Agilent: Excitation and Emission of Green Fluorescent Proteins

[10] ScienceDirect: beta-lactam drugs

[11] Nature Chemistry: Green fluorescent protein

[12] Nobel Prize in Chemistry 2008 Press Release

[13] Molecular Cell: Adaptation in CRISPR-Cas Systems

[14] National Library of Medecin: Bacteriophages

[15] CRISPRloci, comprehensive and accurate annotation of CRISPR–Cas systems

[16] ScienceDirect: Palindromic DNA

[17] Allosteric regulation in CRISPR/Cas1-Cas2 protospacer acquisition mediated by DNA and Cas2

[18] Leibniz Forschungsinstitut fuer molekulare Pharmakologie: Introduction to CRISPR/Cas

[19] ResearchGate: Mechanism of the CRISPR-Cas9 type II system

[20] addgene/CRISPR Guide: ko-generation

[21] Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease

[22] Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing

[23] The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway

[24] addgene/Blog: HDR

[25] addgene/CRISPR Guide: HDR

[26] Rapid and Efficient Gene Deletion by CRISPR/Cas9

[27] Integrated DNA Technologies: How to design sgrna sequences

[28] GeneWatch: Gene-Editing Techniques

[29] Integrated DNA Technologies: What is HDR

[30] Wikipidia: Markus Affolter

[31] Markus Affolter's Publications

[32] Mojica, F.J.M., D ez-Villase or, C.S., Garc a-Mart nez, J.S., and Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J Mol Evol 60, 174–182.

[33] CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies

[34] Bolotin, A., Quinquis, B., Sorokin, A.,and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151, 2551–2561.

[35] Barrangou, Rodolphe, Christophe Fremaux, Hélène Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. "CRISPR Provides Acquired Resistance against Viruses in Prokaryotes." Science 315 (2007): 1709–12.

[36] The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA

[37] The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA

[38] The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli

[39] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.

[40] Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing

[41] Urea Reduction Ratio

[42] The comparison of ZFNs, TALENs, and SpCas9 by GUIDE-seq in HPV-targeted gene therapy

[43] ScienceDirect: FACS-optimized mutants of the green fluorescent protein (GFP)

Images

[a] University of California: How glow-in-the-dark jellyfish inspired a scientific revolution

[b] National Geographic: Shining for Science

[c] Doudna Lab: DNA Acquisition in CRISPR-Cas systems

[d] Integrated DNA Technologies: Genome editing with CRISPR-Cas9

[e] geneticeducation.co: Cas9 Structure and Function

[f] ResearchGate: Current updates of CRISPR/Cas9-mediated genome editing and targeting within tumor cells: an innovative strategy of cancer management, Khaled S. Allemailem et al.

[g] Creative Biolabs: Mechanisms of Gene Editing

[h] University Basel: Markus Affolter