COSMOS

ADDING VALUE TO CAMELINA AND CRAMBE OIL

What is the CRISPR/Cas system and how is it used in the COSMOS project?

The CRISPR/Cas system has been making headlines throughout the media in the recent years as a new genome editing tool to speed up the breeding of crops and animals and even as a tool to cure diseases. But what is the CRISPR/Cas system actually and what are the genetic effects? This brochure explains what the CRISPR/Cas system is and how this tool is used in the COSMOS project.

Authors: Niels de Beus[1], Jean-Luc Dubois[4], Robert van Loo[2], Heiko Keller[3], Frans Krens[2], Stephan Piotrowski[1], Nils Rettenmaier[3], Pia Skoczinski[1], Jan Springer[2], Tobias Wagner[3]

[1]Nova-Institut GmbH, [2]Wageningen University & Research, [3]IFEU - Institut für Energie- und Umweltforschung Heidelberg GmbH, [4]ARKEMA

The COSMOS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 635405.

www.cosmos-h2020.eu/

A primer in genetics

In general, the CRISPR/Cas system is a tool which allows for directed changes in the genome of cells. The genome is a set of long DNA molecules (the chromosomes). The DNA contains information about how all things within living organisms are regulated. The genome consists of over 40,000 genes (how many exactly depends on the specific organism).

The DNA of the genes consists of four different nucleotides (A, C, G and T) and one could view the gene as a sentence in the genetic language that forms a piece of genetic programme. A gene can code for a specific function. Changes in the letters can occur during the copying of the DNA in cell division or under influence of UV-light, cosmic radiation and chemicals. Such changes in the DNA-letters ('mutations') are rare, but despite being rare they are the driving force of evolution.

Most changes in the letters of genes are neutral and do not affect their function. If they do, such a change can either promote the chances of survival or having offspring or decrease these chances. When the chances of survival and having offspring decrease – for example if the changes cause a crucial process like the digestion of food not to function because a basic digestive enzyme is no longer working, such a mutation is called 'harmful' (for the fitness of the organism). And, if a mutation improves the success rate of having offspring such a mutation is called 'beneficial' (for the fitness of the organism).

All breeding of plants and animals depends on the presence of mutations. This presence of mutations within species is called genetic variation. Although the extent of genetic variation is very large in our crops and farm animals, genetic variation is sometimes lacking in key functions which could improve their performance; the natural mutations in the genes of interest just did not yet occur or were suppressed due to evolutionary pressures that are no longer present in our cropping or animal husbandry systems.

This is where the possibility of being able to induce the necessary genetic variation would be of great help. Various methods exist to improve the rate of mutations. Traditionally, gamma radiation and chemicals have been used to induce a higher mutation frequency than naturally occurs. These forms of mutation breeding have been used for over 100 years and mutation breeding is therefore also seen as a traditional, safe method of breeding.

Legally, mutation breeding received a status as genetic modification, but when the legislation on genetic modification was introduced, mutation breeding was directly excluded from strict regulations for the use and release of genetically modified organisms (GMOs) since mutation breeding had a long history of safe use.

A disadvantage of the traditional mutation breeding tools is the fact that they cause random mutations, that were not intended and may cause harmful consequences for the organism. Therefore, the mutation rate needs to be rather low to avoid too many mutations in other genes than the targeted gene. Therefore, very large sets of individuals of the crop species have to be analysed to find the desired mutation.

The CRISPR/Cas System Basics

In contrast to traditional mutation breeding, the CRISPR/Cas system is a targeted mutation tool specifically and only addressing the target gene. How is such a targeted mutation in the gene of interest induced by the CRISPR/Cas system? Before answering this question, it helps to explain the natural origin of the CRISPR/Cas system.

Where does it come from? A Brief history of the tool

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats and was identified in bacteria as key part of their defence system against viruses, while Cas stands for CRISPR associated system. This Cas refers to a large variety of different CRISPR associated proteins (e.g.

Cas3, Cas6, Cas8, Cas9 etc.). These endonucleases are necessary CRISPR/Cas system components for DNA cutting with Cas9 (CRISPR associated protein 9) being the most prominent and commonly used one. Therefore, Cas9 is used in the following for describing the basic principle of the CRISPR/Cas system.

The DNA of bacterial viruses is different from the bacterial DNA and some bacteria can recognise some DNA sequences of viruses using CRISPR as a 'library' of virus sequences in their own DNA and then attack those DNA sequences with a DNA-cutting enzyme.

Upon infection by foreign DNA (from a virus for example), the first response of bacteria is to integrate part of the virus DNA in the CRISPR region (the acquired part is called the spacer). Once the viral DNA is integrated in the CRISPR region, its DNA is 'transcribed' into RNA (called crRNA) with the same sequence of letters and little RNA pieces ('guide RNA') are attached to a DNA cutting enzyme (one example is Cas9). The little mRNA sequence can find the viral DNA sequence and then the Cas9 can cut it up.

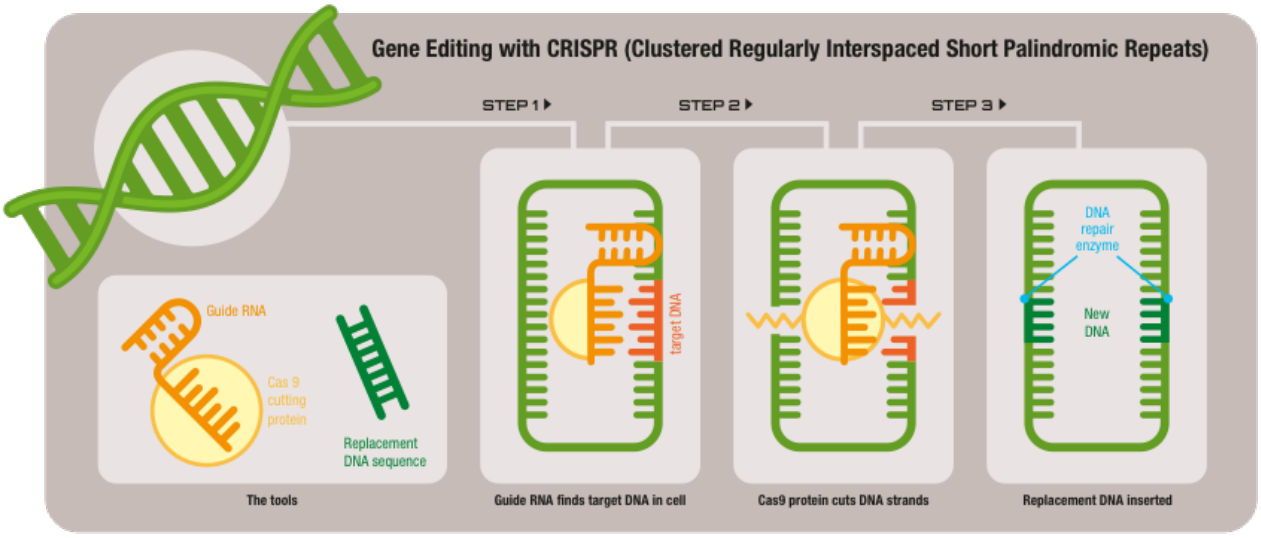

Figure 1: The basic principle of gene editing with the CRISPR/Cas system

The cut-up DNA of the virus is not functional anymore and the virus attack is efficiently blocked. In bacteria, the CRISPR/Cas system thus provides acquired resistance to viral infections and it was discovered that this system can also work in plants and animals and even humans.

The trick is that by designing the DNA sequences that lead to guide-RNA it can target the Cas9 to a gene of interest. The Cas9 enzyme causes a break in the DNA of the gene of interest and that is the start of the induction of a mutation. Breaks in DNA occur periodically and organisms have tools to repair such breaks.

Every now and then an error is made in the repair. Such error could be that a little stretch of DNA is either inserted or deleted. The effect of this is that the DNA transcription to RNA and the translation of RNA to proteins will lead to a different protein (shorter or longer and often with totally different amino acids). In many cases that protein will be non-functional then.

This is used to knock-out genes that deliver unwanted traits to our crops and farm animals and with some more effort could be used also to optimise the desirable functions of genes.

Since the Cas9 enzyme in the CRISPR/Cas system is directed by a specific guide RNA with a sequence of nucleotides (the 'letters') matching only the target gene, it can be designed in such a way that Cas9 will only make such breaks in the target gene. This is different to the traditional

mutation breeding tools that make mutations not only in the target gene but also in many other genes.

So, with the CRISPR/Cas system there is a tool that can be used to knock-out undesirable genes in crops and farm animals. In medicine, the technique can even be used to knock-out genes thatcause diseases, or correct a naturally occurring mutation that today results in a disease.

As mentioned above, various Cas proteins exist and form the basis for the classification of different bacterial CRISPR/Cas systems into three types; type I, II and III (see Figure). During the production of the crRNA, type I and type III systems use Cas6 to break the long strand of crRNA into multiple smaller crRNAs. Type II systems do not use Cas6 to break down the long strand but another cutting mechanism called RNase III. In this system, an additional small RNA code is produced which is complementary to the repeat sequence. This is known as the tracrRNA. Once the crRNA has matured, it associates with Cas proteins. Different systems associate with different Cas proteins:

- Type I systems associate with multiple Cas proteins to form a complex which recognizes foreign DNA (since the crRNA is compatible with the DNA) and starts DNA degradation.
- Type II systems rely on association of both crRNA and tracrRNA with Cas9, a single multifunctional protein.
- Type III systems also require association with multiple Cas proteins, however they do not target DNA, but foreign RNA, enabling them to target RNA based viruses.

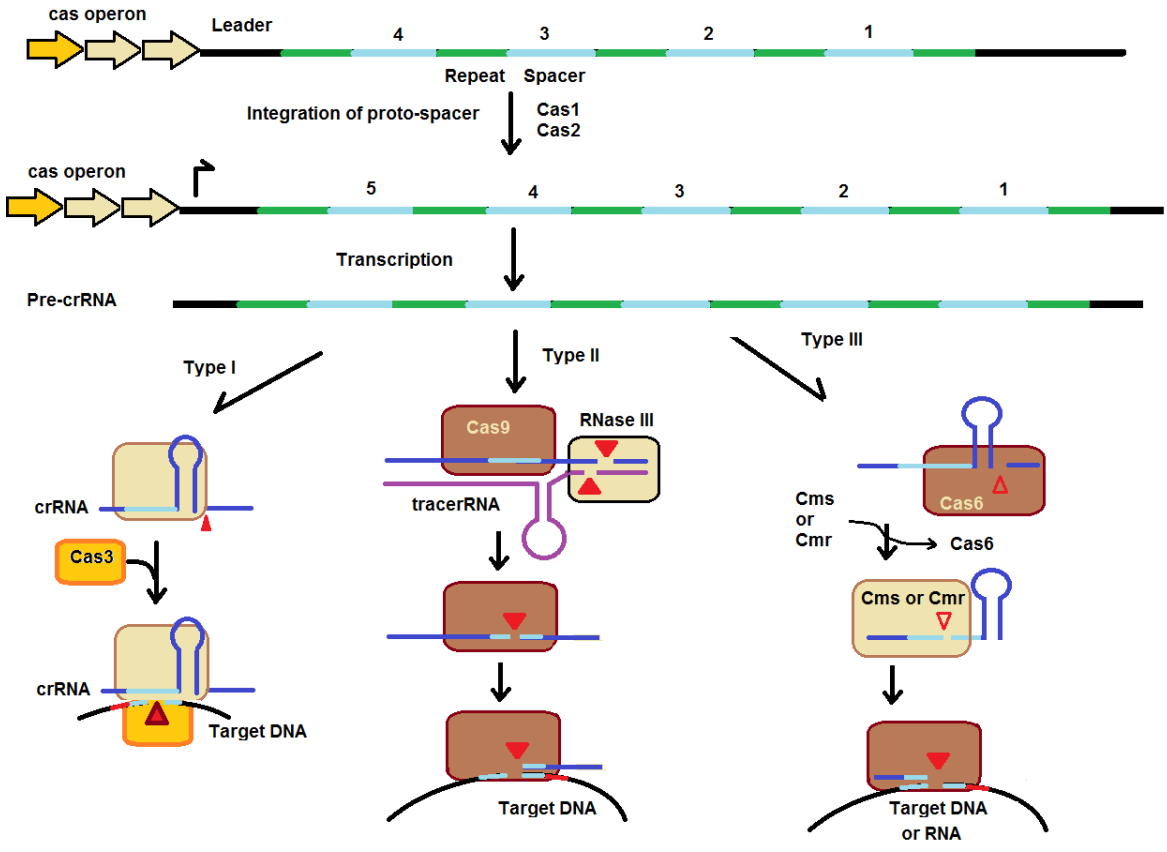

Figure 2: Models of the CRISPR/Cas systems (Makarova et al. 2011)[1]

[1] Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J. and Koonin, E.V. 2011: Evolution and classification of the CRISPR-Cas systems, Nature Reviews Microbiology, 9(6):467-477. doi:10.1038/nrmicro2577.

All three systems require a match between the CRISPR sequence and the infecting sequence in order to create the response. During the immune response, the DNA strands are broken and the repair mechanism can lead to the loss or insertion of some DNA-nucleotides and thereby leads to insertions or deletions in the targeted gene that makes the (viral) gene non-functional. The removed part can be integrated at the CRISPR region again, to enhance immunity.

The type II system is simplified and used as a genome editing tool, mainly because it relies on few additional proteins (only 1, the multifunctional Cas9).

Similar to the bacterial defence system, the gene editing tool consists of a crRNA component, a tracrRNA component and a Cas9 component. The crRNA and the tracrRNA can be combined to create a single guide RNA (sgRNA). The CRISPR system can be introduced into the cell using various methods; by using a plasmid, by virus infection or by delivering the enzymes and RNA into the cell. In some methods, the DNA is integrated into the host genome while in other methods this introduction does not take place. Depending on the interpretation of the genetically modified organism (GMO) regulations, integration of the DNA into the genome can result in a plant which is considered a GMO. However, introducing it by delivering the enzyme and RNA into the cell does not lead to integration but transient expression, resulting in a non-GMO (since the genetic code of the plant has not been altered).

Since the Cas9 enzyme in the CRISPR system is directed by a specific guide RNA (crRNA + tracrRNA) with a sequence of nucleotides (the 'letters') that matches the target gene, it can be designed in such a way that Cas9 will make such breaks specifically in the target gene. In contrast, traditional mutation breeding tools do not have a specific target gene and result in non-specific mutations within various genes of the organism.

Despite the fact that the guide RNA is designed to only match the target, it has been found that off-target mutations (additional mutations in other genes than the target gene) can occur, but are rare. Prior screening for potential off-target sites should be used to avoid off-target mutations.

In short, by introducing the CRISPR/Cas system and a targeted RNA-guide, it can be used to create targeted mutations in genes in any species (plants, animals and also humans, e.g. for gene therapy applications). The RNA-guide needs to be designed in such a way that it is unique to the target gene. CRISPR/Cas9 has already been successfully applied in a broad range of crop species including citrus fruits, maize, poplar, potato, rice, sorghum, soya bean, tomato and wheat, in animal genetics (e.g. to reduce chicken egg white protein allergenicity) and human therapeutics development (e.g. to knock-out cancer genes).

Prospects for the use of the CRISPR/Cas system and barriers to overcome

The system described above is considered a knock-out system, since it renders specific genes non-functional. Currently, most applications focus on knock-outs. However, in future applications of the CRISPR/Cas system, it is possible to not just knock-out a targeted gene, but to replace a little part of the gene in order to improve the gene's function.

In this application of the CRISPR/Cas system, two cuts in the target genes are created to remove part of the DNA and insert a replacement DNA sequence. The ends of the replacement DNA have to be very similar to the original removed DNA for the repair process to work. This DNA-repair-method is called Homology Dependent Repair (HDR).

The development of the CRISPR/Cas system as a molecular biology technique has therefore resulted in a powerful tool to edit genomes which is easily accessible and cheaper compared to alternatives, as fewer tries are required to obtain the desired mutation which ultimately reduces costs. This gives researchers hope to eliminate diseases from plants and even humans, but also wipe out pathogens and create more productive crops. Since the system targets a specific

sequence, it can be used to modify the desired part of the genome. Care should be taken to design the CRISPR guide RNA in such a way that off-site changes are avoided and prior analysis of potential off-site mutation should be carried out; although the CRISPR/Cas system is highly specific, a low probability of off-site changes in similar gene sequences is still present and demand additional checks on such off-site sequences.

In general, the CRISPR/Cas system is a tool to modify every organism from bacteria over fungi to plants but has gained prominent interest related to efficient and specific plant/crop improvement. In contrast to traditional mutation breeding tools that cause random mutations in non-target genes, the CRISPR/Cas system provides a very high precision in the position of mutations as it targets genes of interest only, resulting in less 'collateral' damage than traditional mutation breeding. These 'collateral' damages in traditional mutation breeding have never been proven to be a safety risk. Further, after identifying a CRISPR/Cas mutant, in most cases a back-crossing programme will be carried out to introgress the desired mutations into the background of various elite wild types not carrying any off-target mutations. In the process of the back-crossing, only the desired mutation is introgressed and the off-target mutations will be lost after a few cycles of backcrossing.

Still, despite the technical ease of use and precision of inducing mutations in target genes only, transparency about the application of the CRISPR/Cas system is important and breeders should inform the public about how the CRISPR/Cas technology is used. Society will only obtain the benefit of this new CRISPR/Cas technology when it is put to its best use and trust is built with the general public that this technology is beneficial and safe.

The legal debate

In the EU, there are strict regulations regarding the deliberate release of genetically modified organisms (GMOs), including every organism from bacteria over fungi to plants. The EU Directive 2001/18[2] includes a list of genetic modification and plant-breeding processes that are considered to result in GMOs. Furthermore, an amendment to this directive (Directive (EU) 2015/412[3]) gives Member States the possibility to restrict or prohibit the cultivation of GMOs in their territory through a so-called 'opt-out'. Plants that fall within the scope of EU GMO legislation are subjected to these very strict regulations, practically making it impossible to grow them in the field in most EU Member States.

Many Member States have made use of the 'opt-out', including Germany and France. Scientific research is permitted only under specific conditions, preventing the release of the GMO into the environment. Plants that fall outside the scope can be grown without restriction. The EU Directive considers all organisms and offspring which contain or contained transgenes (genes originating from other species, foreign DNA), as GMOs and these are therefore subjected to regulation. Since the EU Directive focusses on specific processes to induce mutations, it is considered a **process-based legislation**, meaning that when foreign genes are used during the process to induce a mutation within the organism, although these transgenes are no longer present in the 'product' organism, it is subjected to the regulation. The opposite of the process-based legislation is the **product-based legislation.** Product-based legislation is based on novel features of a product regardless of the method of origin. This legislation is applied in Canada for example.

Since in the EU there is process-based legislation, it would be no problem one might think: the CRISPR/Cas system is either on the list of techniques or it is not on the list of techniques. However, the CRISPR/Cas gene editing tool was developed in 2012, therefore discussion is ongoing if it should be included in the list or not, and therefore be subject to strict regulation, or not. Furthermore, techniques which induce *random* mutations in the DNA are not subject to strict regulation. This includes for example EMS (Ethyl Methane Sulfonate), which causes random mutations which are then screened for desired effects. Techniques such as EMS are non-specific

[2] *http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:JOL_2015_068_R_0001*
[3] *http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32001L0018*

causing off-target mutations. However, such techniques are not considered harmful or risky, due to a history of safe use.

Looking at the final effect of the CRISPR/Cas system in breeding, CRISPR/Cas can be regarded as at least as safe as traditional mutation breeding methods as it creates the same functional mutation while causing much fewer side effects (much less 'collateral' damage by mutations in non-targeted genes). Some 'off-target' mutations with the CRISPR/Cas system have been reported[4,5]. In cases where the guide RNA apparently was not totally unique, with a proper design of the guide RNA, part of such off-target mutations can be avoided. Their frequency will be orders of magnitude lower than with classical mutation breeding methods and the occurrence of off-target mutations itself is not seen as an unacceptable risk (since traditional techniques are not subject to strict regulation).

However, currently, the CRISPR/Cas system has to be introduced in the crop species or farm animal. The currently most used systems make use of genetic modification to bring in the gene that codes for the CRISPR/Cas system (i.e. the Cas gene plus the nucleotide sequence for the guide RNA) as this system does not naturally occur in plants or animals. After successful introduction of the CRISPR/Cas system it starts to induce the desired mutations in the target gene. Once the plants with the desired mutations have been found, these plants will be further multiplied to obtain large amounts of seeds and in this process the plants are either self-pollinated or crossed with wild type plants and after that self-pollinated to obtain stable mutation lines.

In this process, the breeder ensures that the desired mutation passes to the next generation while the CRISPR/Cas system itself is not transmitted to the next generation. After the mutations have been induced, the CRISPR/Cas system is no longer needed and can be 'crossed out' by classical plant breeding methods.

The final product of the CRISPR/Cas system is then e.g. a plant that has the desired mutations and does not have the CRISPR/Cas system anymore. This CRISPR/Cas-mutant cannot be functionally or molecularly distinguished from naturally occurring mutations or mutations induced using the approved traditional mutation breeding methods. In a product-based regulation system, a CRISPR/Cas mutant would then be seen as completely similar to these other types of mutants and could be used and released as a normal variety without any specific restrictions other than the standard laws concerning safety of products that apply to any product.

In regions that use a product-based regulation this is the actual situation (e.g. Canada where the novel plant trait regulations apply that are the same for transgenic plants or plants from traditional mutation breeding or the CRISPR/Cas system).

The situation is different in regions that apply a process-based regulation, such as in the EU. Since the CRISPR/Cas system can be introduced by using transgenes, proponents of an inclusion of the technology in the list according to EU Directive 2001/18 argue that the system is then an offspring of a transgene which under the current EU regulation is considered GMO and therefore subject to the regulation, even though these transgenes can be removed later through crossing. However, other introduction methods can be used which do not result in the incorporation of transgenes in the plant DNA, or not using transgenes at all (when the enzyme and RNA sequence are delivered in the cell), this further complicates the legal discussion and interpretation. Furthermore, the CRISPR/Cas technology cannot only be used for the very limited and targeted modifications as intended in the COSMOS project. Within the COMSOS project, all modifications are based on best practices and of course follow any regulation required. With the decreasing costs and the ease of use of such technologies, accessibility to genetic modification increases. Regulators may opt to use GMO regulation to control accessibility, but such a decision would not be based on scientific knowledge.

---

[4] Gupta, R.M. and Musunuru, K. 2014: Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9, J. Clin. Invest., 2014 Oct;124(10):4154-61.

[5] Kosicki, M., Tomberg, K. and Bradley, A. 2018: Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements, Nature Biotechnology, 2018, 36(8):765-711,

However, the modifications made by the CRISPR/Cas system cannot be traced back to the technique when the system is removed through crossing, or if systems are used that leave no foreign DNA sequences behind. This creates a problem with enforceability, especially regarding imported goods since other countries may have regulations where such goods require no labelling.

Although several experts expected that the CRISPR/Cas system and other gene editing tools (TALEN and Zinc finger nucleases) would not fall under the EU Directive 2001/18, on July 25th, 2018, the European Court of Justice decided that organisms (including crops, plants, bacteria etc.) obtained by genome editing techniques (mutagenesis) are GMOs and are subject to the obligations laid down by the GMO directive.

This decision caused negative reactions in industry and university research with a common sense in calling this ruling a throwback in innovation and future food security leaving the European Union behind in technology development and also in economic terms. It also sheds light on the flaws of the underlying regulatory regime[6].

Nevertheless, gene editing using the CRISPR/Cas system is an innovative, promising and efficient tool that will be further improved and applied but not with the hoped and expected non-GMO regulated background which would have included less money and approving efforts.

Application in COSMOS (Camelina & crambe Oil crops as Sources for Medium-chain Oils for Specialty oleochemicals)

In the project, the CRISPR/Cas tool is used to introduce mutations in specific genes involved in fatty acid and glucosinolate biosynthesis in oil crops. The aim is to improve the seed oil and seed meal quality of the two industrial oil crops camelina and crambe. This entails the creation of high levels of mono-unsaturated fatty acids, optimisation of fatty acid profiles and the reduction of Anti-Nutritional Factors (ANF) such as glucosinolates.

The creation of high mono-unsaturated fatty acid profiles means that the financial yield per hectare can be increased. The reduction of ANF enables dual use of the crops: the seeds can be pressed for the oil while the press cake, which contains the protein, can be used as feed, leading to an increased availability of domestic feed.

[6] Purnhagen, K.P., Kok, E., Kleter, G., Schebesta, H., Visser, R.G.F., and Wesseler, J. 2018: EU court casts new plant breeding techniques into regulatory limbo, Nature Biotechnology, 2018, 36(8):799–800, *https://www.nature.com/articles/nbt.4251.*

Glossary

Genes	A specific sequence in the DNA having a function in cell or organism performance after activation
DNA	Double stranded molecule containing genetic information; can be thought of as a storage place for the genetic information of the cell.
RNA	Single stranded molecule containing genetic information; contrary to DNA, RNA is actively used to translate the DNA into activity, functionality and protein.
Mutations	Changes in the genetic sequence of the DNA
Gene editing	Technique to change the gene expression or the genetic sequence of a Cell/organism
Transgenic	A cell which contains genes originating from another species
Crossing	Combination of two plants where genetic material is exchanged, a plant's way of mating
Back-crossing	Crossing of a hybrid with one of its parents or an individual genetically similar to its parent, in order to achieve offspring with a genetic identity which is closer to that of the parent
Introgression	The movement of a gene (gene flow) from one species into the gene pool of another by the repeated backcrossing of an interspecific hybrid with one of its parent species
Plasmid	A plasmid is a small DNA molecule within a cell that is physically separated from the genomic DNA and can replicate independently