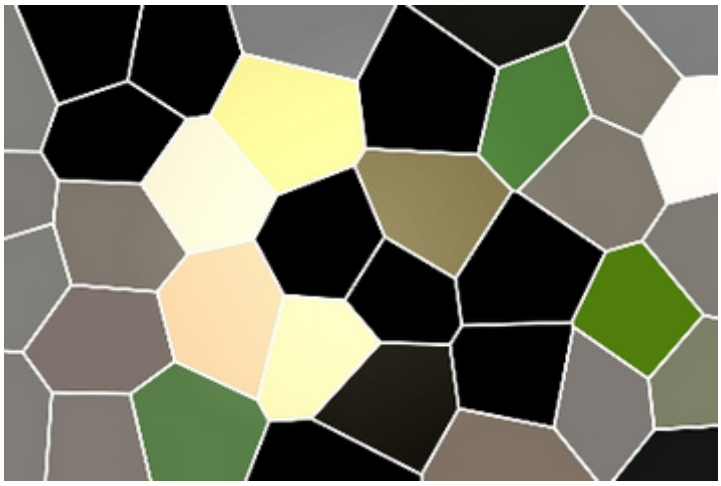
Veille citoyenne d'information sur les OGM et les semences

Genetically modifying a plant is far from being harmless (follow-up)

Par Eric MEUNIER

Publié le 06/06/2017, modifié le 09/03/2026



A new scientific article (june 2017) tells about hundreds of unexpected mutations in mice genetically modified with Crispr-Cas9. A result that do confirm the two articles Inf'OGM published during the summer 2016 and republished here as an updated version.
In a previous article [1], Inf'OGM pointed out the issues of mutations and epigenetic mutations (epimutations) occurring at each step prior to any technique of genetic modification. We will now look at the unintended effects or limits of the subsequent steps: selection of cells permanently or temporarily modified, full plant regeneration from modified cells and crossing of this genetically modified plant with an Elite variety (a step usually called backcrossing).

Modifying the genome of a plant means working in vitro on plant cells which need to have been prepared and cultivated before receiving the material which will produce the desired genetic modification(s).

Selecting and regenerating modified cells is not harmless

Due to the low, if not very low, efficiency of cells' transformation [2], the use of genetic modification techniques requires to select the few cells which have been successfully transformed. This selection is done by using selection markers: usually antibiotic or herbicide resistance genes. Genetically modified cells are placed in a solution containing antibiotics or herbicide and only the

modified cells resistant to the antibiotics or herbicide survive. Genes which allow growth in the presence of a molecule usually toxic or genes allowing cells to use a source of carbon usually not metabolised… or any other selection marker which could remain or are suppressed by the end of the modification processes [3]. But suppressing a selection marker is done through imprecise and unreliable techniques which can therefore potentially induce cell toxicity and other chromosomal rearrangements [4], leave genetic footprints [5] and lead to recombination sites (where genetic sequences can be exchanged between two strands of DNA) with unpredictable effects [6]. These suppression techniques are moreover not available in all plant species.

GM plants must then be regenerated from the selected transformed cells. This needs another round of cells' cultures which require synthetic hormones and might induce once again mutations and epimutations with possible chromosomal rearrangements [7].

Whether it is the genetic modification tool in itself, the upstream or downstream steps, all these manipulations induce mutations and/or epimutations. But, we are often told that mutations and epigenetic mutations will not be present in commercialised plants. Why? The next step, called backcrossing, allow supposedly to get rid of all those unintended effects…

Backcrossing theory

Here's the general principle: a company wishing to market a genetically modified plant (be it legally considered as a GMO or not) doesn't modify one of its "Elite" variety's genome, part of its germplasm collection but one more common variety. Once the modification is obtained, the company crosses this common genetically modified variety with the "Elite" variety. It will then cross the progeny with the initial Elite variety until off-springs are considered (on a statistical basis) nearly similar (named "near-isogenic varieties" and now "negative segregants" for its applications to GMOs) to the Elite variety. The difference between the Elite variety and the near-isogenic variety is supposedly almost only the modified sequence. GNIS, the French organism in charge of seeds certification, explains that, at the last crossing (which could be the 14th), "*the part* [of the genome coming from] *the Elite line is 96,88%, it is therefore estimated that the obtained line is close enough to the Elite line. We're close to obtaining an isogenic variety, being different from the Elite line only by one gene* [the one containing the modification]" [8]. It has to be noted that 96.88% similarity leaves a lot of space for mutations, epimutations and possible chromosomal rearrangements. For plants with a large genome such as bread wheat, with a 17 billion base pairs genome, 3.12% represents more than 500 million base pairs difference…

The limits of the theory

The underlying theory of this unintended effects "cleaning" thanks to backcrossing lies on the following hypothesis: that the unintended effects that appeared during the previous steps are far enough from the region where the required targeted genetic modification occurred. However, the closer to the targeted trait, the bigger the probability they will be transmitted, along with the genetic modification, during each backcrossing step. From a minimum of seven, the number of backcrossings needed could rise to fourteen in order to try to get rid of the unintended effects. But, beside this issue of the unintended effects' proximity with the genetically modified sequence, this theoretical cleaning is undermined by two other biological phenomenons. First, genetic sequences can exist and evolve as a block which may be important. In this "linkage disequilibrium" situation, as these blocks are transmitted unaltered to their off-spring, the unintended effects contained in the block will be inherited with the genetically modified sequence at each backcross. Second, some genetic sequences are able to insert themselves when gametes are generated. Called "selfish DNA", those sequences may be present in a bigger number of gametes than

expected and, from generation to generation, they will always be in the progeny. In cases where unintended effects are located in or near those sequences, it will not be easy to get rid of them [9].

These considerations imply that rigorous testing is needed on a case by case basis. It is to be reminded here that the legislation covering transgenic plants is the only one requiring such testings on a case by case basis with the most exhaustive combination of analysis (even though it does have some deficiencies). It has also to be reminded that genome sequencing, and more particularly 'next generation of sequencing' (NGS), cannot fully answer the question of the presence of unintended effects due to the current limits of this kind of analyses (see box below).

Through those two articles, we discussed the mutations and epimutations as part of the unintended effects that biotechnologies trigger. It therefore appears that the claims about the ability to detect and remove all the unintended effects due to new techniques are more about beliefs than "sound science". Considering the observation of potential effects, the limits of backcrossing and sequencing, sincere doubts can be expressed about the ability of breeders to effectively remove all the unintended effects and to efficiently, cost-effectively and rapidly detect the remaining mutations and epimutations through sequencing. We therefore wonder why the HCB's Scientific Committee ignored those points of risks sources in its interim report, published in February 2016, on the impacts of the new techniques of genetic modification?

This consideration has just been reinforced by a recent article published in june 2017 and telling about hundreds of unexpected mutations in mice genetically modified by Crispr-Cas9 [10]. Those hundreds of mutations were unnoticed as, according to the scientists, the algorithm used are focusing only on the "likely off-target sites for a given gRNA, but these algorithms may miss mutations". The scientists therefore consider that only whole-genome sequencing would allow the detection of all the off-target mutations, despite the fact that this technique is neither perfect (see box below)...

Sequencing and the related software tools? Anything but certainty!

On April 0216 the 7th, during the hearing organized by the French Parliamentary Office for Science and Technology Assessment, André Choulika, head of Cellectis, talking about off-target effects claimed "*to resequence entirely* [the plant genome] *is very important* […] *because you are required to provide the entire sequence in the application dossier*". Except that as it has been shown, results of sequencing are not thoroughly reliable.
"Next generation of sequencing" (NGS) is now rather cheap and fast. But "several problems" are inherent to sequencing as well as reading and using the results. First of all, several steps of sequencing call into question the reliability of the final result. DNA must be properly extracted, cut into pieces which are latter sequenced thanks to different platforms and methods. However, platforms and methods are quite different from one another in terms of limits and reliability of the results, which mean "deep sequencing" have to be performed [11].
The sequences must then be read to stick the sequences in order to achieve a complete genome. The obtained sequence is then compared to other ones considered as "reference" which are kept in databases already containing several mistakes [12]. All these steps give rise to an important lack of precision in the final results regarding the detection of unintended effects and therefore, on the risk assessment. The uncertainties concerning the sequences, as well as their assembly and comparisons, increase with polyploid genomes and the repeated sequences [13]. Finally, mutations

can greatly differ in their biological importance…. [14]. And In most of the cases epimutations are not monitored in the case of gene editing.

Many articles sum up the difficulties found at each step, compare the methods, platforms [15] and associated software [16], discuss the gold standards and other standards to be implemented in order to improve the reliability of the entire process [17]. As outlined by numerous authors, all those techniques and steps are in progress but not yet mature, they are evolving and need many checks for case-by-case assessment by very well trained people.
As also pointed out by several scientists, one of the present challenges for molecular biology is to determine how to face the accumulation of so many results (one those famous “big data”), some of which contain mistakes. Besides, facing scepticism towards any sequencing result, the minimum requests from reviewers are so high that more and more researchers must now provide “deep” sequencing results to support the credibility of their raw results [18].

---

[2] « Less is more : strategies to remove marker genes from transgenic plants », Yau, Y.Y. et al, (2013), BMC Biotechnology.

[3] « Alternatives to Antibiotic Resistance Marker Genes for In Vitro Selection of Genetically Modified Plants – Scientific Developments, Current Use, Operational Access and Biosafety Considerations », Breyer et al (2014) Critical Reviews in Plant Sciences, Vol 33, Issue 4, 286-330
« Suitability of non-lethal marker and marker-free systems for development of transgenic crop plants : present status and future prospects », Manimaran et al (2011) Biotechnol Adv. 29(6), 703-14
« Effects of antibiotics on suppression of Agrobacterium tumefaciens and plant regeneration from wheat embryo », Han, S-N. et al, (2004), Journal of Crop Science and Biotechnology 10, 92-98.

[4] Some of those systems can remain as extra-chromosomic circular elements during several generations (as for wheat).

[5] which could be used to identify the transforming event

[6] Biotechnology 13., *ibid.*

[7] « Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications », Neelakandan, A.K. et al, (2012), Plant Cell Reports 31, 597-620
« Regeneration in plants and animals : dedifferentiation, transdifferentiation, or just differentiation ? », Sugimoto, K. et al, (2011), Trends in Cell Biology 21, 212-218.

[8] http://www.gnis-pedagogie.org/biotechnologie-amelioration-reproduction-retrocroisement.html

[9] « Distorsions de ségrégation et amélioration génétique des plantes (synthèse bibliographique) », Diouf, F.B.H. et al , (2012), Biotechnologie Agronomie Société Et Environnement, 16(4), 499-508
« Quantitative trait locus mapping can benefit from segregation distortion », Xu, S. (2008), Genetics, 180(4), 2201-2208
« Genetic map construction and detection of genetic loci underlying segregation distortion in an intraspecific cross of Populus deltoides », Zhou, W et al, (2015), PLoS ONE, 10(5), e0126077

[10] « Unexpected mutations after CRISPR–Cas9 editing in vivo », K.A. Schaefer et al, (juin 2017), Nature Methods, 14(6), 547-548

[11] « Next generation sequencing technology : Advances and applications », Buermans, H.P.J. et al, (2014), Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1842(10), 1932-1941
« Next-generation sequencing platforms », Mardis, E.R. (2013), Annual Review of Analytical Chemistry, 6(1), 287-303
« Applications of next-generation sequencing. Sequencing technologies - the next generation », Metzker, M.L. (2010),

Nature Reviews Genetics, 11(1), 31-46.

[12] « Next-generation sequence assembly : four stages of data processing and computational challenges », El-Metwally, S. et al, (2013), PLoS Comput Biol 9, e1003345
« Systematic comparison of variant calling pipelines using gold standard personal exome variants », Hwang, S. et al, (2015), Scientific reports 5, 17875
« Sequence assembly demystified », Nagarajan, N. et al, (2013), Nat Rev Genet 14, 157-167.« Improving the quality of genome, protein sequence, and taxonomy databases : a prerequisite for microbiome meta-omics 2.0 », Pible, O. et al, (2015). Proteomics 15, 3418-3423
« Theoretical analysis of mutation hotspots and their DNA sequence context specificity », Rogozin, I.B. et al, (2003), Mutation Research/Reviews in Mutation Research 544, 65-85

[13] « Sequencing technologies and tools for short tandem repeat variation detection », Cao, M.D. et al, (2014), Briefings in Bioinformatics

[14] « Open chromatin reveals the functional maize genome », Rodgers-Melnick, E. et al, (2016). Proceedings of the National Academy of Sciences 113, E3177-E3184
« Evolutionary patterns of genic DNA methylation vary across land plants », Takuno, S. et al, (2016), Nature Plants 2, 15222.

[15] « Systematic comparison of variant calling pipelines using gold standard personal exome variants », Hwang, S., et al, (2015), Scientific reports 5, 17875
« Principles and challenges of genome-wide DNA methylation analysis », Laird, P.W. (2010), Nature Reviews Genetics 11, 191-203
« Performance comparison of whole-genome sequencing platforms », Lam, H.Y.K. et al (2012), Nat Biotech 30, 78-82
« Low concordance of multiple variant-calling pipelines : practical implications for exome and genome sequencing », O'Rawe, J. et al, (2013), Genome Medicine 5, 1-18.

[16] « Next-generation sequence assembly : four stages of data processing and computational challenges », El-Metwally, S. et al, (2013), PLoS Comput Biol 9, e1003345

[17] « Rapid evaluation and quality control of next generation sequencing data with FaQCs », Lo, C.-C. Et al, (2014), BMC Bioinformatics 15, 1-8

[18] « Droplet barcoding for massively parallel single-molecule deep sequencing », Lan, F. et al, (2016), Nat Commun 7

Adresse de cet article : https://infogm.org/genetically-modifying-a-plant-is-far-from-being-harmless-follow-up/