Risk assessment of T-DNA borders from *Agrobacterium tumefaciens* in cisgenic crops

E. Jacobsen and J. Schaart

Plant Breeding, Wageningen University and Research Centre, PO Box 16, 6700 AA Wageningen, The Netherlands

 Agrobacterium tumefaciens mediated insertion of DNA	1.	Introduction	1
 3. Are T-DNA border like sequences present in non-GM plants?	2.	Agrobacterium tumefaciens mediated insertion of DNA	2
 4. Does translation of T-DNA borders lead to novel amino acid sequences? 5. Are the remainders of the T-DNA borders transcribed in GM plants? 6. Is the produced mRNA stable? 7. Do transcribed T-DNA borders lead to functional proteins? 8. Risks for allergenicity 9. The baseline 10. Conclusions 11. References 	3.	Are T-DNA border like sequences present in non-GM plants?	3
 5. Are the remainders of the T-DNA borders transcribed in GM plants?	4.	Does translation of T-DNA borders lead to novel amino acid sequences?	4
 6. Is the produced mRNA stable? 7. Do transcribed T-DNA borders lead to functional proteins? 8. Risks for allergenicity 9. The baseline 10. Conclusions 11. References 	5.	Are the remainders of the T-DNA borders transcribed in GM plants?	5
 7. Do transcribed T-DNA borders lead to functional proteins?	6.	Is the produced mRNA stable?	6
8. Risks for allergenicity 9. 9. The baseline 7. 10. Conclusions 9. 11. References 11.	7.	Do transcribed T-DNA borders lead to functional proteins?	6
9. The baseline 7 10. Conclusions 7 11. References 1	8.	Risks for allergenicity	6
10. Conclusions	9.	The baseline	7
11. References	10.	Conclusions	9
	11.	References	11

1. Introduction

Cisgenic plants are genetically modified plants that do not contain foreign genes. The plants have been enriched with genes from sexually compatible plants. ^{1,2,3,4,5} The introduced genes are flanked by their native promoters and terminators in natural orientation, and contain their native introns.

The Netherlands Commission on Genetic Modification (COGEM) has discussed the safety of cisgenic plants, and has concluded that the risks of cisgenic plants are similar to the risks of conventionally bred plants. However, in view of food and feed safety, the COGEM recommended investigating whether T-DNA borders from Agrobacterium tumefaciens may lead to allergenicity or toxicity.⁶ When A. tumefaciens is used for genetic modification, so-called T-DNA borders repeats may co-insert with the cisgenes. Remnants of the T-DNA borders are found back as two short DNA sequences of at maximum only three and 22 basepairs, respectively, which flank the cisgene in the cisgenic plant. The COGEM experts are of the opinion that it is very unlikely that these short DNA sequences will lead to a risk, for three reasons. Firstly, it is very unlikely that this border insertion will give rise to a new functional protein. Secondly, the chance is negligibly small that, if a functional protein would be produced, that this protein would lead to an ecological advantage of the plant. Thirdly, the genomes of plants are dynamic, and in nature numerous small or larger rearrangements of the genome occur. Potential risks of insertion of short 'foreign' DNA sequences therefore do not exceed the risks of naturally occurring risks or risks

of conventional breeding. For these reasons, the very unlikely risks are, according to the COGEM, within the baseline of conventional breeding.⁷

The COGEM discussed this in a meeting with two experts from RIKILT and SIR, working in the field of feed and food safety. These experts did neither falsify nor confirm the scientific reasoning of the COGEM experts, but instead provided the legal answer that cisgenic plants are genetically modified, as is the case for transgenic plants, and therefore should be subjected to a full risk analysis regarding food and feed safety.⁷ This statement is a correct legal answer for the present GMO regulation in the EU. However, it does not answer the scientific question on biosafety of T-DNA borders, nor does it answer the question whether a distinction between transgenic and cisgenic plants in the EU Regulation would make sense from a biosafety point of view.

As a result of above-mentioned, the COGEM recommended a deeper investigation in the safety of T-DNA borders.⁷

This report is written in view of this recommendation, and provides an answer to the question: 'Do T-DNA borders, that may be present in cisgenic plants, lead to a risk in view of feed and food safety, compared to risks of plants from conventional breeding?'

2. Agrobacterium tumefaciens mediated insertion of DNA

For introduction of recombinant DNA into plants different technologies can be applied, such as ballistic methods⁸ and *Agrobacterium tumefaciens* mediated insertion.⁹ Ballistic methods frequently lead to multiple insertions and to unpredictable rearrangements in the genome. *A. tumefaciens* mediated insertion leads more often to one insertion only, and shows no or hardly any rearrangements in the genomic DNA of the plant.^{10,11} These single copy lines can easily be selected among the transformants obtained. Therefore, for the far majority of transformations of plants, *A. tumefaciens* is used nowadays.¹²

A. tumefaciens is a natural pathogen of many plants, especially dicots, and causes crown gall disease in more than 140 plant species. For causing crown galls, A. tumefaciens transfers a part of its natural plasmid DNA to the nucleus of the plant cell, and inserts that DNA into the genomic DNA of the plant cell. In natural situations this inserted DNA contains genes that cause crown gall formation, the typical symptom of infection by A. tumefaciens. For genetic modification of plants, plant breeders and researchers have replaced the crown gall inducing genes from A. tumefaciens by other genes. In case of transgenesis, foreign genes are used as genes of interest. In case of cisgenesis, genes are used from plants that are sexually compatible with the recipient plant.²

The DNA that is inserted into the plant is named T-DNA (<u>Transfer-DNA</u>). This T-DNA is flanked in the circular Ti plasmid of *A. tumefaciens* by the left T-DNA border and the right T-DNA border. A T-DNA border usually consists of 25 basepairs (Table 1). The borders are oligonucleotide sequences in a direct repeat orientation. The left border and right border are very similar or identical, but in the plasmid upstream the

right border there is an enhancer that makes this right border the initiation site. The left borders functions as termination site in the transfer process. A single stranded copy of the T-DNA (T-strand) is formed that is transmitted to the plant cells and incorporated into the plant genome. The incorporated T-strand contains 3 nucleotides from the right border repeat at the 5' end and (upto) 22 nucleotides from the left border repeat at the 3' end. These remainders are not essential, and have no biological function (Fig. 1).

Table 1. DN	NA sequences	of two differen	nt generally used	l T-DNA l	eft borders
	A sequences	of two unferen	n generany useu		

Source	Sequence
Nopaline LB repeat (25 bp)	TGG * CAGGATATATTGTGGTGTAAAC (\rightarrow T – DNA)
Octopine-TL LB repeat (25 bp)	$CGG^{*}CAGGATATATTCAATTGTAAAT({\scriptstyle \rightarrow}T\text{-}DNA)$

*= nick, general site of integration

It may happen that the termination reaction at the left border repeat is skipped or that the initiation reaction occurs at the left border repeat instead of the right border repeat. As a result a part of the backbone of the Ti plasmid (binary vector) remains attached to a border, and may be co-inserted into the plant's genome. When GM crop plants are meant for introduction onto the market, these plants have to be evaluated for presence of such backbone DNA. Plants that contain backbone DNA from *A. tumefaciens* are usually discarded, and not submitted for approval for introduction onto the market. Cisgenic plants never contain backbone DNA with foreign genes. If they would harbour such foreign genes, the plants are not cisgenic by definition, as no foreign genes are allowed in cisgenic plants.¹ Therefore absence of backbone DNA must always be ascertained before plants are labelled as 'cisgenic'.

Figure 1. Schematic representation of a circular Ti-plasmid from *A. tumefaciens* and T-DNA transfer. Only part of the 25 bp long left border and right border are integrated into the plant genome.



3. Are T-DNA border like sequences present in non-GM plants?

The three remaining basepairs of the right border are self-evidently very common in living organisms, including plants. On the average a sequence of three basepairs is

present once in approximately $4^3 = 64$ nucleotides. In this view, this remainder is not unique in plants at all, but very common. It is present numerous times (>> 100 000 times) in each plant genome. Therefore risk of insertion of this triplet is negligible compared to the risks caused by natural rearrangements and mutations within plant genomes.

Because of this negligible risk of the right border, this report focuses only on the remainder of the left border.

Rommens et al¹³ aligned the T-DNA border sequence with the whole genome sequences of *Arabidopsis* and rice (*Oryza sativa*) and detected in both plant genomes T-DNA border like sequences. That encouraged looking for border like sequences in their crop of interest, i.e. potato. They isolated putative border like sequences from potato accessions, using a wide variety of border-specific degenerated primers. They amplified fragments and sequence analyzed these, and detected border like sequences in potato. In order to underline that these borders are takes from the plant's own genome, they named these plant DNA borders P-DNA borders, rather than T-DNA borders. Subsequently, Rommens et al. showed that these borders were functional for effective transfer of DNA from *Agrobacterium* binary plasmids to the genome of individual plant cells.^{13, 14}

Conner et al.¹⁵ showed that border like sequences are present in numerous plants. These researchers have designed borders, based on sequence information from the crop plant species themselves, and designed P-DNA borders too.

It can be concluded that T-DNA-like sequences are present in natural plant genomes. This has been proven for many crops¹⁵

4. Does translation of T-DNA borders lead to novel amino acid sequences?

We performed a bioinformatics study, using the published whole genome sequences of *Arabidopsis thaliana* and rice, similar to the mentioned studies of Rommens et al¹³. However, we performed this at the amino acid level. We translated the whole genome sequence of the plant into amino acid sequences, using three reading frames, and two directions, leading to six amino acid sequences per whole genome. Then we aligned the translated remainder of the left border (Table 3) with the whole genome amino acid sequences. Note that we did not align the border sequences with sequences from real proteins or Open Reading Frames in these plants. As the T-DNA can insert not only in genes but also in between genes, we took the whole genome. The aim was to study whether the borders would lead to novel amino acid sequences.

For Arabidopsis this resulted in 53 perfect hits for the remainder of the nopaline left border, and 75 perfect hits for the remainder of the octopine left border (Table 2). For rice the number of perfect hits was approximately 3 times higher.

Table 2. Number of perfect matches of the remainder of the nopaline border and ocotopine border in Arabidopsis and rice, if both the borders and the whole genome sequences are translated into amino acids sequences.

Remainder of the left border of	Arabidopsis	Rice
Nopaline	53	151
Octopine	75	263

Firstly, it is remarkable that the number of hits is much higher compared to the number of hits found by Rommens et al., although the same genomes were used. This difference is caused by the phenomenon that different nucleotide triplets may lead to the same amino acid. As example, the first reading frame of the remainder of the nopaline left border comprises 21 nucleotides, translated into seven amino acids. However, this sequence of seven amino acids can also be coded by 255 other nucleotide sequences. Rather than comparing one border DNA sequence, we actually looked at 256 DNA sequences. As a result, the number of hits after translation was much higher.

Secondly, the number of hits for rice is approximately 3 times higher compared to Arabidopsis. This is caused by the genome size. The genome of rice (390 Mb) is 3.4 times larger than the genome of Arabidopsis (115 Mb).¹⁶

Arabidopsis and rice have relatively small genomes. Most crops have larger genomes. As a consequence, the number of hits for border like sequences at the amino acid level, will be higher for these crops too.

It can be concluded that translation of T-DNA borders does not lead to new novel amino acid sequences.

5. Are the remainders of the T-DNA borders transcribed in GM plants?

The remainders of a T-DNA borders are too small for coding for a gene with a promoter, a coding sequence, a stop codon and a terminator. Genes vary strongly in size, but usually plant genes including their promoters and introns comprise several thousands of basepairs. Three or 22 basepairs is therefore far too small for being a complete gene. Moreover it does not contain a combination of a stop and start codon. Therefore a border is not a gene, and therefore cannot cause an anti-nutritional protein on its own.

However, it is possible that the T-DNA is inserted in a promoter or in an exon or intron, and that the border is transcribed into RNA as part of a larger stretch of mRNA.¹⁷

The remainders of the T-DNA borders can in theory be expressed in two directions, i.e. starting from the insert, and reading towards the DNA flanking the T-DNA. During this transcription, the remainder of the border would be transcribed too.

However, in case of cisgenesis the T-DNA between the borders consists of complete genes only, including their stop codons and terminators. Transcription of the cisgene starts within the T-DNA, and stops also within that T-DNA. Transcription does not proceed to the border and the flanking DNA. Therefore the chance that the border will be transcribed starting from the T-DNA and proceeding outwards is negligibly small.

The other direction is possible, i.e. from outside the T-DNA, reading the border and towards the T-DNA. This may be the case, if the T-DNA is inserted into a gene. The result of this insertion is very likely a gene that can be regarded as a combination of a part of the original gene, and a new part, consisting of the remainder of the border, and a stretch of the T-DNA, until a stopcodon is encountered. This may lead to a peptide.

It can be concluded that transcription of the border starting from the T-DNA going outwards to the flanking DNA, is highly unlikely. However, transcription starting from flanking DNA, covering the border, and proceeding into the T-DNA is possible.

6. Is the produced mRNA stable?

Transcription, starting from the flanking DNA, and reading the border and part of the T-DNA, will lead to mRNA. The mRNA however has after its stop codon probably no terminator sequence nor a polyadenylation site. Such sequences contain important information for correct processing of a stable mRNA molecule.^{18,19} Therefore, if the mRNA would be produced, it is highly likely that it will be unstable.

7. Do transcribed T-DNA borders lead to functional proteins?

If this unstable RNA were translated, this will lead to a fusion protein. Insertion of a T-DNA into a coding sequence will result in a chimeric protein which, depending on the position of insertion, will most likely lose its characteristic functional domains. This fusion protein is very likely a non-functional and unstable protein.

COGEM has already indicated that the chance is negligible that this fusion protein will provide an ecological advantage to the plant.⁶ We agree on this.

8. Risks for allergenicity

We searched in Allergen Online whether the remaining sequence of the right border is part of known allergens. We used Allergen Online, version 9.0, updated in January 2009 (http://www.allergenonline.org/index.shtml). Allergen Online provides access to a peer reviewed allergen list and sequence searchable database intended for identifying proteins that may present a potential risk of allergenic cross-reactivity. It contains a comprehensive list (1386 sequence entries) of unique proteins of known and putative allergenic proteins (food, airway, venom/salivary and contact). A number of the allergenic wheat gliadins or glutenins may also cause celiac disease and those are listed if there is evidence of IgE binding in the database too.

Table 2. Translated remainders of left border sequences. The DNA sequences are read from the flanking DNA inwards into the T-DNA. This leads to three reading frames per border.

Source	DNA sequence	Reading	Peptide
		frame	
Nopaline	CAGGATATATTGTGGTGTAAAC	1	GlnAspIleLeuTrpCysLys
	NCAGGATATATTGTGGTGTAAAC	2	GlyTyrIleValValter
	NNCAGGATATATTGTGGTGTAAAC	3	ArgIleTyrCysGlyValAsn
Octopine	CAGGATATATTCAATTGTAAAT	1	GlnAspIlePheAsnCysLys
	NCAGGATATATTCAATTGTAAAT	2	GlyTyrIleGlnLeuter
	NNCAGGATATATTCAATTGTAAAT	3	ArgIleTyrSerIleValAsn

The six amino acids sequences, ranging from 5 to 7 amino acids (Table 2) were compared with the 1386 amino acid sequences of known and putative allergens, mentioned in comprehensive database of Allergy Online.

There appeared to be no allergen at all in the database that contained any of these amino acids sequences shown in Table 2.

This indicated that it is very unlikely that, in case the remainder of the left border were translated into a (fusion) protein, it would act as an epitope for allergenicity.

9. The baseline

The baseline for cisgenesis in plants is conventional plant breeding, and natural genome plasticity in non-gm crop plants and their wild relatives.

In conventional plant breeding, frequently alleles from wild germplasm are introgressed and domesticated into varieties. These varieties are marketed for e.g. feed and food purposes. Along with the wanted alleles, numerous unknown alleles are co-inherited into these varieties. These unknown alleles may be expressed, and may produce proteins. However, the far majority of these alleles is neither characterised nor known, and we are not aware of their functions nor of their possibly antinutritional effects. However, the associated risks are accepted, and belong the socalled baseline. Mankind is living with these risks for hundreds of years, and has accepted the risks for food and feed associated with conventional breeding. Classical breeding and its associated risks are regarded as the reference.

Since the discovery of X-ray induced mutations in barley nearly 80 years ago,²⁰ plant breeders and geneticists have realized how DNA mutations can be induced for widening the genetic variation in their germplasm. During the past seventy years, mutation breeding led to more than 2250 plant varieties.^{21,22}70 % of these varieties were released as directly induced mutants, and the other 30 % from crosses with induced mutants. The use of chemical treatments was relatively infrequent, but

gamma rays were most frequently used (64 %), followed by X-rays (22 %). Error! Bookmark not defined.

The FAO/IAEA website contains a database of plant varieties derived from induced mutations (http://www-infocris.iaea.org/MVD/default.htm). This list has been composed on the basis of official information from plant breeders and authorities. The composers of the database mention that the list is far from complete, as frequently it is not published how new varieties have been obtained. In spite of that, the list contains now (March 2009) already 2570 released plant varieties. In reality the number of induced mutant varieties is much larger. If spontaneous mutants were also included, the list would further expand strongly.

The induced mutant varieties have been developed in 175 plant species, including rice, wheat, barley, cotton, rapeseed, sunflower, grapefruit, apple, banana and many other species. They are released in Europe, Asia, North America, South America and Australia. Dozens of these varieties are grown at large scales. Many millions of people eat and use products of these varieties. The mutant varieties and those originating from translocation events or back crosses, usually are not molecularly characterized by DNA sequencing, nor compared with their parents at the DNA sequence level. Due to the treatments, many chromosomal breakages may have occurred, which sometime are perfectly repaired, but sometimes the repair is not perfect, leading to e.g. gene disruption, chromosomal arrangements, deletions, insertions, and so on. Sometimes cytogenetic or genetic marker studies are performed to locate introduced chromosomal parts in the recipient genome. However, the plant breeders of mutant varieties usually do not know the number of mutations or changes, the kind of mutations, nor the number of rearrangements in their varieties. Neither they know whether new Open Reading Frames or fused genes have been created by the mutations, nor whether expression levels of genes have changed due to mutation, or due to introgression of DNA from wild germplasm. However, the cultivars generally have been phenotyped thoroughly by the breeders and compared phenotypically to their parents and contemporary cultivars in view of their commercial value. In spite of the absence of molecular characterization, the cultivars, either from induced mutation, spontaneous mutation or introgression breeding, have been widely accepted, grown and used.

These cultivars from conventional breeding and mutation breeding belong to the baseline.

The question we have to answer is whether the risks for food and feed of the remainder of the left T-DNA border exceed the accepted risk level of the baseline, or whether the risk from the border is lower than the risks of conventional plant breeding.

In case of cisgenesis, established plant varieties are used, with a known history of safe use. To such a variety one or several well-characterised cisgenes are added, and in addition a small remainder of the left border. The level of uncertainty, and the level of associated risks, is far below the accepted risks in conventional breeding and mutation breeding. Compared to conventional breeding and mutation breeding, the risks for food and feed of the T-DNA borders are negligibly small.

10. Conclusions

When *A. tumefaciens* is used for insertion of genes during cisgenesis, remainders of the right and left T-DNA borders may be present in the cisgenic plant. The remainder of the right border is only 3 basepairs. These basepairs occur numerous times in each plant genome, and do not cause any risk beyond the accepted baseline risks of conventional breeding and natural mutation. The remainder of the left border usually consists of 22 or less basepairs. This study focuses on the risks for food and feed of the 22 basepairs of the remainder of the left border. If plants do contain backbone DNA from the plasmid of *A. tumefaciens*, these plants are not cisgenic.⁴

Studies from Rommens and Conner have revealed that border like sequences are commonly present in crops. When translating whole genome DNA sequences to amino acid sequences, the peptides encoded by border like sequences are even more frequently present.

The remainder of the left border is too small to code for a gene. In that perspective, the borders do not pose a risk on its own.

If the T-DNA is inserted into an existing gene, the border may be attached to the existing gene, and be transcribed, leading to a fusion mRNA. Because the border is not followed by a terminator sequence or polyadenylation site, it is unlikely that the fusion mRNA is stable. It is even more unlikely that the fusion protein will have a biological function.

Comparison of the amino acids sequences from the border with amino acids sequences of know allergens in the comprehensive database of Allergy Online, shows that the peptides translated from the border sequences do not match with any known allergen. Therefore it is extremely unlikely that the border sequence will act as an epitope, or will induce to allergy.

Therefore it can be concluded that the chance that T-DNA borders will lead to an epitope or another anti-nutritional peptide is negligibly small.

In conventional breeding many unknown alleles are inserted into varieties for feed and food, without knowing the possibly anti-nutritional effects of these unknown alleles, and unknown intergenic regions. The risks associated with these unknown alleles and intergenic regions are accepted as baseline risks. Border-like sequences probably are already present in the varieties, or can be introgressed from wild relatives by means of conventional breeding.

Mutation breeding has led to more than 2570 varieties. Millions of people in Europe eat and use products of these varieties. The plant breeders of mutant varieties usually do not know the number of mutations or changes, the kinds of mutations, nor the number of rearrangements or gene disruptions in these mutant varieties. Neither they know whether new Open Reading Frames or fused genes have been created by the mutations, nor whether expression levels of genes have changed due to mutation, or due to introgression of DNA from wild germplasm. The associated risks are generally accepted, and belong to the baseline. In the case of cisgenesis, known cultivars with a history of safe use are used. These cultivars are enriched with one or a few cisgenes, and the remainder of the well-known left border. The level of uncertainty in case of cisgenesis is well below the level of uncertainty in conventional breeding and mutation breeding. Comparison of the low level of uncertainty of T-DNA effects compared to the higher uncertainty levels of conventional breeding and mutation breeding, further underlines that the risks for food and feed of the T-DNA borders are negligible small compared to the baseline risks of conventional breeding and mutation breeding.

Acknowledgements

Drs. E. Severing (Wageningen University UR) was involved in the bioinformatics of the T-DNA borders on the risk analysis for allergenicity. We appreciated the critical remarks from Prof. Dr. P.J.J. Hooykaas made during the writing process.

11. References

¹ Jacobsen E., Schouten H.J., 2008. Cisgenesis, a new tool for traditional plant breeding, should be exempted from the regulation on genetically modified organisms in a step by step approach. Potato Research. doi:10.1007/s11540-008-9097-y.

² Schouten H.J., Jacobsen E., 2008. Cisgenesis and intragenesis, sisters in innovative plant breeding. Trends in Plant Science 13: 260-261.

³ Schouten, H.J., Krens, F.A., Jacobsen, E., 2006. Do cisgenic plants warrant less stringent oversight? Nature Biotechnology 24: 753.

⁴ Schouten, H.J., Krens, F.A., Jacobsen, E., 2006. Cisgenic plants are similar to traditionally bred plants. EMBO Reports 7: 750-753.

⁵ Jacobsen, E., Schouten, H.J., 2007. Cisgenesis strongly improves introgression breeding and induced translocation breeding of plants. Trends In Biotechnology 25: 219-223.

⁶ COGEM, 2006.Vereenvoudiging van regelgeving bij genetische modificatie met planteigen genen, cisgenese, een reële optie? CGM/060428-05.

⁷ COGEM, 2008. Signalerende brief 'cisgenese en voedselveiligheidsbeoordeling', 28 oktober 2008. CGM/081028-04.

⁸ Stomp AM, Weissinger AK, Sederoff RR, 1989.Ballistic transformation of conifer. US Patent number 5122466.

⁹ Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA, 1983. A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid. Nature 303: 179-180.

¹⁰ Gheysen G, Villarroel R, Montagu M van, 1991. Illegitimate recombination in plants: a model for T-DNA integration. Genes & Dev. 5: 287-297

¹¹ Tzfira T, Li J, Lacroix B, Citovsky V., 2004. Agrobacterium T-DNA integration: molecules and models. Trends Genet. 20, 375-83.

¹² Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA, 1983. A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303, 179 – 180.

¹³ Rommens CM, Humara JM, Ye J, Yan H, Richael C, Zhang L, Perry R, Swords K, 2004. Crop improvement through modification of the plant's own genome. Plant Physiol 135, 421-31.

¹⁴ Rommens CM, Haring MA, Swords K, Davies HV, Belknap WR. 2007. The intragenic approach as a new extension to traditional plant breeding. Trends in Plant Science 12: 397-403.

¹⁵ Conner AJ, Barrell PJ, Baldwin SJ, Lokerse AS, Cooper PA, Erasmuson AK, Nap JP, Jacobs JME, 2007. Intragenic vectors for gene transfer without foreign DNA. Euphytica 154, 341-353.

¹⁶ <u>http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/G/GenomeSizes.html</u>

 $\frac{17}{2}$ Brunaud et al., 2002. T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites. EMBO reports 3, 1152–1157.

¹⁸ Green P.J., 1993. Control of mRNA stability in higher plants. Plant Physiology 102: 1065-1070.

¹⁹ Hoof A. van, Green P.J., 1996. Premature nonsense codons decrease the stability of phytohemagglutinin mRNA in a position-dependent manner. The Plant Journal 10: 415-424.

²⁰ Stadler LJ, 1928. Mutations in barley induced by X-rays and radium. Science 68: 186-187.

²¹ <u>Maluszynski M, Nichterlein K, Van Zanten L, Ahloowalia BS (2000). Officially</u> released mutant varieties – The FAO/IAEA Database. Mutation Breeding Review 12: 1-84.

²² Ahloowalia BS, Maluszynski M, Nichterlein K. (2004) Global impact of mutationderived varieties. Euphytica 135 (2): 187 – 204.