TWN Biotechnology & Biosafety Series 9

Genetically Engineered Cells and Organisms: Substantially Equivalent or Different?

by Terje Traavik, Kaare M Nielsen
and David Quist

TWN
Third World Network

Genetically Engineered Cells and Organisms: Substantially Equivalent or Different?

Terje Traavik, Kaare M Nielsen and David Quist

TWN
Third World Network
Penang, Malaysia

Genetically Engineered Cells and Organisms: Substantially Equivalent or Different?

is published by
Third World Network
131 Macalister Road
10400 Penang, Malaysia.

© Norsk institutt for genøkologi (GenØk), Tromsø,
and Tapir Academic Press, Trondheim, 2009

Printed by Jutaprint
2 Solok Sungei Pinang 3, Sg. Pinang
11600 Penang, Malaysia.

ISBN: 978-967-5412-07-3

Contents

Chapter 1. **Introduction** 1

Chapter 2. **Lack of Precision in Recombinant DNA Techniques** 3
- 2.1 *The '-omes' and the '-omics'* 6

Chapter 3. **Changes in the Genome** 8
- 3.1 *Observations from studies of genetically modified (GM) plants* 9
- 3.2 *Why do DNA rearrangements occur?* 11

Chapter 4. **Changes in the transcriptome** 13
- 4.1 *Example of new transcripts originating from a plant transgene* 13
- 4.2 *Examples of the activity of the 35S CalMV plant promoter in mammalian cells* 14
- 4.3 *Example of upregulation of an endogenous gene under the influence of a transgene promoter* 15
- 4.4 *Does 'transvection' occur during transgenesis in mammalian cells?* 16

Chapter 5. **Changes in the Proteome** 18
- 5.1 *An α-amylase inhibitor-1 gene transferred from common bean to pea* 18
- 5.2 *Production of recombinant protein in milk* 19

Chapter 6. **Changes in the Metabolome** 21

Chapter 7. **Changes in the Epigenome** 23

Chapter 8. **Changes in the Interactome** 25

Chapter 9. **Concluding Remarks** 27

Resources 28

References 29

Acknowledgement

This paper was originally published as a chapter in the book *Biosafety First – Holistic Approaches to Risk and Uncertainty in Genetic Engineering and Genetically Modified Organisms*, 2007, Terje Traavik and Lim Li Ching (eds.), Tapir Academic Press, Trondheim, ISBN: 9788251921138. It is reprinted here with permission.

Chapter 1

Introduction

The dynamic and interconnected regulation of the genome is now slowly being revealed. The genome does not function in a constant, stable and linear fashion, but is instructed by and fine-tunes its activities according to networks of signals received from the external ecosystem and the internal environment of the organism. The genomic signal pathways may be modified by ecosystem variation as well as by physiological changes in the organism. Thus, the chromatin structure, the genome, the epigenome, the transcriptome, the proteome, the metabolome, and the interactome are interlinked and intertwined in various ways with information transfer in multiple directions.

Integration of foreign DNA into an established genome may have unanticipated side-effects, e.g. in terms of chromatin changes, genome instability, unexpected protein products from the transgene(s), and influence on overall organismal gene expression patterns, in quantitative as well as qualitative terms, of the recipient organism. In this paper we discuss and exemplify, from a precautionary point of view, the changes that may occur in modified genomes and the consequences they may have. We structure the discussion as follows:

- Lack of precision in recombinant DNA techniques
- Changes in the genome
- Changes in the transcriptome

- Changes in the proteome
- Changes in the metabolome
- Changes in the epigenome
- Changes in the interactome
- Concluding remarks

Chapter 2

Lack of Precision in Recombinant DNA Techniques

Genetic engineering (GE) techniques are presented by many as a tool for the safe and predictable production of genetically modified organisms (GMOs). The intended change in gene expression in GMOs is, however, often not simply a matter of transcription and translation of the inserted recombinant DNA sequences, as symbolised by the Central Dogma model. While achieving a stable, single-copy recombinant DNA insertion is the aim of the genetic engineer, it is not the norm.

Available methods for transfer of gene constructs into cells are inefficient and imprecise. Insertional mutagenesis is a default consequence of recombinant DNA insertions. The resulting phenotypic consequences of the insertion events are largely determined by the characteristics of the gene transfer vector and the location of and number of copies inserted per cell.

While many emphasise the precision of recombinant DNA techniques, none of the currently available methods permit predetermination of **where** in the recipient cell-DNA the gene construct will be inserted, or the number of copies that will be inserted into GMOs of commercial relevance. The specific locations of the inserts may nevertheless substantially influence the functions of the inserted DNA as well as its effects on the cell's own genes. For instance, within the same transformed/transfected mammalian cell culture we will find cells with quite different characteristics.

These, in principle indefinite number of variants, arise due to varying insertion sites and number of full or partial DNA copies. In addition to full vector copies, a number of rearranged or truncated versions, some of them quite small, may be inserted into some cells. These aberrant versions can still influence the integrity and functions of the recipient genome, and they may go undetected by conventional testing.[1] Impacts arising from uncharacterised insertions cannot be predicted from characterised insertions.

Furthermore, if the characterised inserts are identical between, for example, two recombinant maize lines (events), but the insertion sites are different, one **cannot** extrapolate any biosafety conclusions from one line (event) to the other. The context of the insert would obviously be different, as would be the genes that may be affected directly or indirectly and therefore also the resulting plant phenotype.

The integration of foreign DNA (transgene) in a new host genome may influence any of the gene expression control processes. New gene products may also arise and the transgene product may also vary in its properties. For instance, read-through transcription, initiated somewhere in the insert and ending outside it, or initiated in adjacent regions and ending in the insert, may be sources for novel RNAs and recombinant proteins.[2]

The consequences of insertion may, as earlier stated, vary considerably according to the exact insertional locations and/or construct organisation. This is valid for the expression of the inserted transgene as well as for changes in the recipient organism's own genes and their expression levels. The insertion may have effects by introducing a change in chromatin structure, the topography as well as the proteins binding to the DNA (Recillas-Targa, 2006), or by inducing changes in DNA methylation patterns and other epigenetic characteristics. Furthermore,

cis-acting regulatory DNA motifs may be present in the insert, or may arise from the 'new' sequences created by integration that can alter the expression level of genes adjacent or even distant to the insert.

Genetically engineered cell cultures may be used to produce recombinant products **under contained laboratory conditions**. This implies that the product that the gene is coding for (e.g. insulin) is extensively purified before it is taken out of the laboratory, while the genetically engineered cells and DNA are destroyed inside the laboratory. Such applications of genetic engineering may, in principle, be made safe. However, when recombinant cells are developed and placed in the open environment, changes in the gene expression levels and small metabolite contents will vary according to changing ecosystem conditions.

Under the influence of given sets of ecosystem variables, the recombinant organisms may over time expose phenotypic traits that have environmental or consumer health implications. 'Consumers' may include a number of wildlife species in addition to humans and domestic animals. From biosafety/risk assessment/regulatory points of view it is hence imperative to reveal whether, compared to its unmodified counterpart, a GMO has experienced changes in the interacting regulatory parts, its *'interactome'*: the genome, epigenome, transcriptome, proteome, and metabolome working as overlapping layers of information involved in cellular function (Box 2.1). Only when minimal changes are observed will it be justified to claim 'substantial equivalence'.

Box 2.1
The '-omes' and the '-omics'

Genome: 1) The entire collection of genetic material in an organism, virus or organelle. 2) The haploid set of chromosomes (DNA) of a eukaryotic organism.

Genomics: The study and development of genetic and physical maps, large-scale DNA sequencing, gene discovery, and computer-based systems for managing and analysing genomic data.

Proteome: The full complement of proteins that are found in a particular cell or tissue under a particular set of circumstances. May include information on their relative or absolute abundance.

Proteomics: The study of the structure and expression of proteins, and of the interactions between proteins.

Interactome: The complete collection of all physical protein-protein interactions that can take place within the cell.

Interactomics: The study and construction of comprehensive sets of protein-protein interactions.

Transcriptome: The full complement of expressed gene transcripts, including alternative splice variants that are found in a particular cell or tissue under a particular set of circumstances. This may include information on the relative or absolute abundance of transcripts.

Transcriptomics: The study of the full complement of expressed gene transcripts. Several techniques have been developed for parallel analysis of the expression of thousands of genes, most notably cDNA microarrays and oligonucleotide arrays.

Metabolome: The assembly of substrates, metabolites, and other small molecules that is present in a population of cells.

Metabolomics: Study of the structure and distribution of all metabolites (small molecules), particularly organic compounds.

Functional genomics: A whole spectrum of approaches, under development, to ascertain the biochemical, cellular and/or physiological properties of each and every gene product and its regulation. These include near-saturation mutagenesis (i.e. screening hundreds of thousands of mutants to identify genes that affect traits as diverse as embryogenesis, immunology and behaviour), high throughput reverse genetics (methods to systematically and specifically inactivate individual genes), and elaboration of genetic tools.

NOTES:

[1] It is a common phenomenon for transgene constructs to integrate in multiple places in the genome, and for very small parts of the construct to integrate independently of full-sized versions (for recent comprehensive reviews, see Filipecki & Malepszy, 2006; Latham et al., 2006).

[2] Abortive transcription from read-through might, for example, produce novel short and double-stranded (ds)RNA molecules. A risk factor emerging from the production of novel dsRNA is the potential to induce gene silencing either locally, or on other genes. The same dsRNA can have different effects at different concentrations, in some cases showing non-specific effects at concentrations lower than those needed to induce silencing (Zhao et al., 2001). It should also be appreciated that any new RNA transcript may undergo a large series of modifications that result in 'a family' of different RNA molecules, all derived from the same original source. The family members do not necessarily give rise to the same proteins or even proteins with similar functions.

Chapter 3

Changes in the Genome

The whole purpose of a transgenesis process is of course to change the genome of the recipient organism. There are a number of possible, unpredictable consequences of DNA insertions in GMOs. They may be sorted into the following categories:

1. Genome destabilisation
2. Chromatin changes with consequences for transgene as well as genome gene expression
3. *De novo* methylation of the transgene or spread of the transgene methylation pattern to endogenous genes, i.e. epigenomic effects
4. Introduction of new regulatory elements, e.g. promoters, enhancers and dehancers, known or hidden splice sites, start codons, terminators, etc. These may cause:
   a. Unpredictable, environment-dependent level of transgene expression, and
   b. Unpredictable, environment-dependent influence on expression pattern of recipient genome in terms of:
      i. Signal transduction-dependent *promoter* effects
      ii. Signal transduction-dependent *enhancer/silencer* effects
      iii. Signal transduction-dependent effects of transferred DNA methylation patterns
5. Activation of endogenous mobile elements (‘jumping genes’). Once activated, they may engage in:

6. Reinsertion at new chromosomal loci
7. Horizontal gene transfer to other individuals or species
8. Unanticipated and unpredictable changes in gene products, e.g. by post-translational modifications
9. Silencing or over-expression of genes.

Some prominent uncertainties are related to the fact that the recipient organism receives a new promoter/enhancer. These elements govern the gene expression levels of their attached transgenes, but after insertion, they may also change the gene expression and methylation patterns in the recipient chromosome(s) over long distances up- and downstream from the insertion site. Promoters/enhancers function in response to signals received from the internal or external environment of the organism. For a GMO this may result in unpredictability with regard to:

- The chromatin organisation and contents of the recipient genome
- Altered expression of a large number of the organism's own genes
- Altered influence of geographical, chemical (i.e. *xenobiotics*) and ecological variables of the environment
- Transfer of vector sequences within the chromosomes of the organism, and vertical and/or horizontal gene transfer to other organisms.

Few published studies have been devoted to the clarification of such putative changes in GMOs.

3.1 Observations from studies of genetically modified (GM) plants[3]

Agrobacterium-mediated gene transfer to plants can result in insertion site mutations of the T-DNA, leading to truncations, interspersions, or other complex rearrangements of the

recombinant DNA. Superfluous T-DNA integration frequently accompanies *Agrobacterium*-mediated transformation, where whole and partial copies of the transgenes become integrated.

For example, a molecular analysis of *Agrobacterium*-transformed *Arabidopsis thaliana* plants revealed that 80% of the transformants had a single insertion event; of these, only 22% contained a single copy of the transgene (the desired number for stable integration and expression in transgenic lines), and the remainder of these single-insertion events contained incomplete T-DNAs, tandem T-DNAs, or T-DNA fragments. These results indicate that even relatively simple T-DNA insertions undergo large- or small-scale rearrangements during the transformation process.

Plants transformed via particle bombardment methods are often more likely than *Agrobacterium*-mediated transformed plants to demonstrate complex integration patterns. The majority of integrated DNA is arranged either as multiple copies of the intact transgene, or as multiple copies with interspersed plant genomic DNA. Further, short recombinant DNA fragments may frequently integrate along with intact or rearranged multimers.

In a study of transgenes integrated into two lines of transgenic oat, 50 of the 82 transgene fragments identified (61%) were 200 bp or shorter. One study even reported the presence of bacterial DNA at a particle bombardment insertion site. As with *Agrobacterium*-mediated transformation, simple single copy insertion events tend to be the exception, and complex and errant integration the rule.

Given the complex transgene integration locus patterns accompanying transformation, developing a transgenic plant line requires careful selection of stable and high expressing transformation events for product development. However, the initial transformation process is not the only step where the transgenes

might undergo significant rearrangement. Tissue culture is a common means to produce sufficient transgenic germplasm for further product development. During this process, undesirable tissue culture-induced genetic rearrangements, termed *somaclonal variation,* can occur in both conventional and transformed lines.

Further along the development of the transgenic plant line is selective crossbreeding with elite crop germplasm for high agronomic performance. This process involves a number of introgressive hybridisations (introgression and subsequent backcrossing) to produce plants homozygous for the recombinant trait in the elite crop line. During this process, the complex nature of the recombinant DNA integration loci can lead to deviations in the expected Mendelian patterns of inheritance.[4] For instance, these irregular patterns have been observed during inheritance in lettuce (McCabe & Mohapatra, 1999), rice, maize, and barley. Subsequent selection procedures of the GM material may also introduce further genomic reorganisations (Hernandez et al., 2003).

3.2 Why do DNA rearrangements occur?

In plants, exogenous DNA transfer (e.g. with *A. tumefaciens* pathogenesis) elicits a wound response that activates nucleases and DNA repair enzymes. The transferred DNA is thus either degraded or used as a substrate for DNA repair, resulting in its potential rearrangement and incorporation in the genomic DNA (Takano et al., 1997). Furthermore, specific transforming plasmid structure and construct properties can enhance recombination events all along the transformation process. Indeed, some genetic elements can act as hotspots and undergo recombination at high frequency. This is, for example, the case for the 3′ end of the CaMV 35S promoter, which contains an imperfect palindrome of 19 bp.

Illegitimate recombination can also occur in the borders of the Ti plasmid of *Agrobacterium tumefaciens*, especially in the right border that contains an imperfect palindromic sequence of 11 bp. The 3′ end of the *nos* terminator is also theoretically highly prone to recombination (Kohli et al., 1999). Hotspots for recombination may lead to tandem transgene repeats with interspersed plant DNA sequences in a single genetic locus. Presence of several inserts may also result from multimerisation in the plasmid before transformation or from multiple insertions.

A number of transgenic and genomic rearrangements have been reported for already commercialised transgenic crop plant varieties.

NOTES:

[3] For further information and references, see the recent review by Latham et al., 2006.

[4] Given the likelihood of transgene re-assortment during one or more of these steps in the production of a transgenic line, arriving at a stable and well-performing transgenic line requires the careful selection from many transformation events brought through development. Technical dossiers on commercial crop lines invariably suggest the stability of the inserted construct. Yet how robust are these analyses? Documentation of transgene locus structure (organisation and copy number) and stability through inheritance in the scientific literature (as well as in applications for commercial approval) almost always rely on Southern blot analysis to demonstrate transgene copy number and integrity of the single-copy inserts. However, recent studies have determined that Southern blot analysis often lacks sufficient resolution to accurately determine copy number or transgene organisation, and may have difficulties in detecting small rearrangements or solitary fragments (Hoebeeck et al., 2007).

Chapter 4

Changes in the Transcriptome

The intention of a transgenic process is to have the transgene expressed. Hence, the intended change is to add one transcript to the transcriptome of the GMO. However, the inherent lack of insertion precision may lead to the expression of additional, unintended transcripts as well.

Although only a small number of published studies have been designed to reveal transcriptome aberrations in GMOs, there are published studies that exemplify the following:

1. Qualitative transcriptome changes, due to inefficient terminator motifs in a transgenic plant variety
2. Quantitative transcriptome changes, due to the influence of the transgene regulatory sequences on endogenous genes located close or distant to the insertion site.

4.1 Example of new transcripts originating from a plant transgene

New evidence suggests that the *nos* terminator sequence used in a number of transgenic plant varieties is a recombination hotspot, prone to read-through, and may contain a cryptic *cis*-acting splice sequence that could generate novel RNA molecules and proteins at any place it is inserted into the genome (Rang et al., 2005).

The Roundup Ready (RR) soybean varieties derive from a soybean line into which a gene coding for glyphosate-resistant enol-pyruvylshikimate-3-phosphate-synthase (EPSPS) was introduced. The insert and the flanking regions in RR soybean have recently been characterised. It was shown that a further 250-bp fragment of the *epsps* gene is localised downstream of the introduced *nos* terminator of transcription, derived from the nopaline synthase gene of *Agrobacterium tumefaciens*. At least 150 bp of this DNA region is transcribed in the RR soybean variety.

Transcription of the additional fragment depends on whether read-through events ignore the *nos* terminator signal located upstream. The data indicate that the read-through product is further processed, resulting in four different RNA variants from which the transcribed region of the *nos* terminator is completely deleted. Deletion results in the generation of open reading frames which might code for (as yet unknown) EPSPS fusion proteins. The *nos* terminator is used as a regulatory element in several other transgenic plants intended for food production. This implies that read-through products and transcription of RNA variants might be a common feature in such plants.

4.2 Examples of the activity of the 35S CaMV plant promoter in mammalian cells

In most of the transgenic crop plants commercialised, the transcription of the transgene is governed by the 35S promoter taken from the Cauliflower Mosaic Virus (CaMV). CaMV is a DNA-containing para-retrovirus that replicates by means of reverse transcription. It was earlier assumed that the 35S promoter exclusively functions in plants, and that it would therefore not represent a food/feed safety issue if the transgene under the control of such a promoter would transfer horizontally. The following quote is representative of this assumption: ‘There have

also been (scientifically unfounded) concerns that the strong plant virus promoter used to express transgenic DNA might be active in mammalian cells' (Gasson & Burke, 2001).

There have now been published studies indicating that the 35S CaMV promoter has the potential for transcriptional activation in mammalian systems, in addition to studies in different yeast species. First, 35S promoter activity was demonstrated in human fibroblast cell cultures, thereafter in hamster cells, and very recently 35S promoter activity was established in human enterocyte-like cells (Myhre et al., 2006). Such cells line the surface of human intestines, and are hence highly relevant to whether uptake of transgenic DNA from the gastro-intestinal tract may have effects on the host if unintentionally taken up. However, no published studies have investigated 35S CaMV activity *in vivo*, and this is hence an obvious area of omitted research. This example illustrates how safety assumptions/claims made in the absence of experimental investigation on the issue can be misleading.

4.3 Example of upregulation of an endogenous gene under the influence of a transgene promoter

X-Scid is a disease linked to a defective gene on the X chromosome that leads to a total breakdown of the immune system due to lack of T cells. Victims are known in the media as 'bubble boys', having to live their short lives within totally contained plastic cubicles, since every kind of innocent infection will kill them.

A gene therapy protocol was developed in order to cure, or at least alleviate, the symptoms of X-Scid victims. Bone marrow cells were taken from the patient and grown in culture. The cells were transfected with a vector that contains a healthy copy of the defective gene. The vector was a deletion mutant of MLV

(murine leukaemia virus), with the transgene under control of a strong promoter. After having the bone marrow cells controlled for expression of the transgene, and observing a lack of any unwanted phenotypic characteristics, the cells were returned to the patient. The rationale was that the transferred healthy gene, following integration into the genomes of the bone marrow cells, should produce the proteins that make production of T cells possible, and hence provide the patient with a functional immune system.

In an initial series of 11 treated patients, the strategy seemed to work according to plan, until a tragic setback was recognised: one of the treated patients developed a highly aggressive type of cancer. It turned out that in treated cells from this patient, the gene transfer vector had integrated into a genomic location next to the *Lmo2* gene. This gene encodes a protein product that is known to be cancer causing when over-expressed. In the present case, the strong promoter of the gene therapy vector had forced the *Lmo2* gene to over-express. In a commentary article in *New Scientist* these events were dubbed 'Gene therapy's worst nightmare'. Yet what was observed was an illustration of the known insertion site unpredictability of current recombinant DNA techniques.

4.4 Does 'transvection' occur during transgenesis in mammalian cells?

A relevant question to ask is whether known, unknown or hidden DNA motifs in the gene vector, including its plasmid backbone sequences, may act as transcriptional enhancers and hence influence transcription of endogenous genes, whether integrated in the host genome or present on an un-integrated vector. Transcriptional enhancers are relatively short (30–500 bp), *cis*-acting DNA sequences usually comprised of several binding sites for transcription factor (TF) activator proteins. The hallmark of

enhancers is their ability to communicate with promoters, often activating genes over a large distance. Some enhancers are able to activate promoters in *trans*, i.e. when the enhancer is on a different genomic entity than the promoter.

Recent studies (D'Aiuto et al., 2006) have demonstrated that a CMV (human cytomegalovirus) enhancer can increase the activity of its cognate promoter in *trans,* in the absence of factors that physically bring the enhancer into close proximity of the promoter. A process like this is called *transvection*. Interestingly, the authors also provided evidence that the CMV enhancer may activate other promoters in the modified host genome. Because such transactivation effects may result in unwanted or unexpected transcriptional activation of endogenous genes, these findings are important for conception of the range of transcriptional effects expected in various genetic engineering and gene therapy approaches.

Chapter 5

Changes in the Proteome

Inherent in a recombinant organism is one or more intended proteomic changes, namely the expression of the transgenic protein(s) that will confer the desired new trait or property.

As earlier indicated, integration of foreign DNA may lead to additional quantitative and qualitative differences in the expressed proteins in a modified cell. Some cellular processes may lead to unexpected protein products from any given gene sequence. All these processes also apply to transgenes as well. Unfortunately, there are few published studies that have systematically compared the proteomes of GMOs to their unmodified counterparts. There are, however, two examples that illustrate the profound and unpredictable differences in the biological functions of a recombinant protein when it is being post-translationally modified, i.e. glycosylated, in its new host organism.

5.1 An α-amylase inhibitor-1 gene transferred from common bean to pea

It was recently shown that expression of a recombinant plant protein (α-amylase inhibitor-1, αAI) from the common bean in a non-native host plant, i.e. transgenic pea, led to the synthesis of a structurally modified, probably aberrantly glycosylated form, of this inhibitor (Prescott et al., 2005). Employing models

of inflammation, it was demonstrated that consumption of the modified αAI and not the native form predisposed the mice to antigen-specific CD4+ Th2-type inflammation. Furthermore, consumption of the modified αAI concurrently with other heterogeneous proteins promoted immunological cross priming, which then elicited specific immunoreactivity of these usually non-immunogenic proteins. This investigation demonstrated that recombinant expression of non-native proteins in plants may lead to the synthesis of structural variants with altered immunogenicity. The frequency at which alterations in structure and immunogenicity of recombinant proteins in new hosts occur is most often not known.

5.2 Production of recombinant protein in milk

The European Medicine Agency (EMEA)'s decision in February 2006 to approve a recombinant product containing antithrombin-α, had been eagerly awaited because it would be the first drug produced in a transgenic farm animal to reach the market. The active ingredient, *human antithrombin-α*, is produced by and purified from the milk of transgenic goats. GTC Biotherapeutics has been developing Atryn since 1993, principally for treating patients suffering from hereditary antithrombin deficiency, a rare condition affecting one person in every 3–5,000, that puts them at increased risk of deep vein thrombosis.

The decision of EMEA was, however, based on a lack of appropriate data to allay concerns about Atryn's immunogenicity. As pointed out by an anonymous editorial commentator in *Nature Biotechnology* (2006, 24: 368), the EMEA decision '*rather skirts around some of the underlying issues that transgenic protein producers have to face*'. These issues have been discussed in this paper.

Of particular concern are different and unpredictable post-translational modifications compared to native proteins. In the case of Atryn, this really seems to matter. Compared with a conventional antithrombin-α product, Atryn's serum half-life was reduced seven- to ten-fold, necessitating infusion of the protein rather than a one-off injection.

One of EMEA's main concerns with Atryn was, however, its potential immunogenicity. The underlying problem is that it is extremely difficult to produce 'nature-identical' proteins in milk from transgenic animals. For instance, in cows, sheep and goats, glycosylated proteins typically contain N-glycolylneuraminic acid (NGNA), a modification which is virtually absent in native human proteins. Furthermore, the high concentration of protein produced in milk, around a gram per litre, overrides the glycosylation capacity of the mammary gland. Only rabbits and chickens have human-like glycosylation patterns. The *Nature Biotechnology* commentator concluded: '*Thus, if immunogenicity of milk-produced proteins turns out to be a generic problem, then a whole class of transgenic production methods may turn out to have a limited future. Chicken milk, anyone*?'

Chapter 6

Changes in the Metabolome

Unintended effects of transgenesis are closely related to changes in the metabolite levels. One of the major challenges is how to analyse the overall metabolite composition of GMOs in comparison to their unmodified counterparts. Metabolomics offer one possible solution.

The quality of crop plants is a direct function of the metabolite content. The metabolome determines the flavour, aroma and texture of crops, their storage properties, nutritional values and performance in the field. Genetic (metabolic) engineering has the potential to improve plant properties. However, problems may arise from such approaches because the organismal metabolism forms a large interconnected network. '*Just as the flap of a butterfly wing might cause a hurricane, changes in the flux of one branch might lead to unexpected changes in other parts of the network*' (Memelink, 2005).

A number of unexpected changes following genetic engineering have been seen in experimental studies with, for instance, *Arabidopsis* sp. and tomatoes (e.g. Romer et al., 2000; Hemm et al., 2003). Field trials with transgenic wheat lines have demonstrated how profoundly the environment affects the metabolome of transgenic as well as unmodified varieties, but have also demonstrated important differences between a transgenic wheat line and its parental, unmodified counterpart (Baker et al., 2006).

Potatoes produce a number of toxic secondary metabolites, which are divided into two groups: the sesquiterpenes and the glycoalkaloids (PGAs). Whereas PGAs are largely produced and present in toxic quantities in both the foliage and 'green' potatoes, it is well documented that the levels of PGAs and sesquiterpenes are affected by biotic and abiotic stress. The development of GM potato varieties has made it prudent to ascertain whether there may be changes in the amounts or types of these secondary metabolites, either as a direct effect of the transgene or due to its interactions with environmental variables.

One such study has been published by Matthews et al. (2005). Transgenic potato lines were exposed, along with non-transgenic lines, to a range of biotic and abiotic stresses and a range of environmental conditions in the field and store. Following stress, a comparison was made of levels of potato glycoalkaloid and sesquiterpene levels between the two groups. Significant differences were observed in the levels of both glycoalkaloid and sesquiterpene levels between transgenic and control material and between infected and non-infected material. The study did, however, also illustrate the profound impact that environmental parameters may have on the metabolome of transgenic as well as unmodified potatoes.

Chapter 7

Changes in the Epigenome

Epigenetic changes[5] can be induced in cells during the transgenesis process, and can become inherited in the consecutive generations (Filipecki & Malepszy, 2006). It is, however, difficult to ascertain whether epigenetic imprinting is due to the transgenesis or cell regeneration techniques. It is known from a number of organisms that an inserted DNA fragment may both transfer its own methylation pattern to the surrounding DNA and have its own pattern changed by the surrounding recipient DNA.

The transgenesis process may induce mutagenic stress-related mechanisms described as 'programmed loss of cellular control'. According to Filipecki and Malepszy (2006), this may lead to (i) genetic changes such as polyploidy, aneuploidy, chromosome rearrangements, somatic recombination, gene amplifications, point mutations, and excisions and insertions of retrotransposons, and (ii) epigenetic changes, including DNA methylation and histone modifications.

Regulation of gene expression by induced changes in DNA methylation is a very potent regulatory mechanism. DNA methylation is based on the existence of 'the 5th base'. Transgenesis may induce methylation changes in both directions:

- DNA *hypomethylation* leading to
  * Gene activation
  * Chromosome instability
- DNA *hypermethylation* leading to
  * Gene silencing
  * Chromatin remodelling
  * RNA-associated silencing

In recombinant plants, DNA methylation changes may occur in both directions, but *hypomethylation* has been more frequently reported. Already in 1996 it was clearly demonstrated that different epigenetic expression states might arise in transgenic plants regenerated from the same material (Matzke & Matzke, 1996), and that these states are stably inherited by the following generations.

As pointed out earlier, the influence of the environment on the initiation and persistence of epigenomic programmes cannot be overestimated, but this is an area of omitted research. In spite of a considerable number of peer-reviewed articles concerning epigenetic consequences of transgenesis in model organisms such as *Arabidopsis*, the epigenomes of marketed, transgenic crop plants are virtually unknown.

NOTE:

[5] Epigenetics was introduced by Conrad Waddington in 1942 as the study of the processes by which genotype gives rise to phenotype. In 1987, Robin Holliday redefined epigenetics as: 'Nuclear inheritance which is not based on differences in DNA sequence'. Epigenetics encompasses heritable changes in DNA or its associated proteins except mutations in gene sequence. Many investigators in the field of epigenetics focus on histone modifications and DNA methylation, two molecular mechanisms that are often linked and interdependent.

Chapter 8

Changes in the Interactome

The concepts and technologies of classical molecular biology have dominated genetic engineering approaches during the last 50 years. This has favoured methods that have approached complex processes by separation and isolation of single pathways and molecules. Nonetheless, biologists have continually been aware that a fundamental characteristic of all biological organisations is that functional units never exist in isolation. Biological complexity is based on synergistic cooperation achieved by interactions between the components of the cell (Uhrig, 2006). Proteins are essential for almost all biological processes. They operate entirely on the basis of interactions with other molecules, i.e. other proteins, nucleic acids, lipids, or low molecular metabolites and other compounds.

Only rarely is the protein monomer the functionally active form, as most often assumed when using transgenes. Comprehensive knowledge of protein interactions is therefore an important source of information to functionally annotate proteins and to understand and model processes on a genome-wide level. That the transgenic protein product provides the intended function and trait (e.g. insecticidal effects or herbicide tolerance in plants) does not preclude that it contains additional active domains that become evident in its new genomic, biological and environmental host context. Such ‘novel’ domains may be inherent in the amino acid chain, or arise as a result of alternative folding due

to host-specific post-translational modifications. The recombinant protein may therefore engage in complex formations with endogenous proteins and other cellular components when present in novel environments. This may, in turn, lead to activation or inhibition of cellular processes, or even create new intracellular processes. To what extent this occurs is unknown, since the studies needed for clarification are rarely conducted.

Chapter 9

Concluding Remarks

As stated by Haslberger (2006), there is a general need for a holistic and integrated basis for assessment of the properties and effects of GMOs. This conclusion was also drawn by a recent World Health Organisation (WHO) report (2005). Lack of knowledge concerning the putative and unpredictable changes in the contents of GMOs discussed in this paper has won increasing acceptance during recent years. A fact that has been reflected in a number of expert committee reports from international organisations such as WHO, the Food and Agriculture Organisation (FAO), and the Organisation for Economic Cooperation and Development (OECD). Many of the risk issues identified here that lack answers were identified before the first transgenic plants were commercially grown in 1996. The application of the modern '-omics' techniques can contribute to reveal many risk-relevant differences in composition between recombinant organisms and their isogenic, parental counterparts under relevant environmental conditions.

Resources

Campbell NA and Reece JB. (2005) *Biology*, Seventh Edition. Pearson-Benjamin Cummings (ISBN 0-321-26984-5, International Edition).

Doerfler W. et al. (2001) 'Foreign DNA integration – Perturbations of the genome – oncogenesis'. *Annals of the New York Academy of Sciences* 945:276-288.

Gibson G and Muse SV. (2004) *A Primer in Genome Science*. Second Edition. Sinauer Associates, Inc. Publishers (ISBN 0-87893-232-1).

Karp GE. (2005) *Cell and Molecular Biology. Concepts and Experiments*. John Wiley & Sons Inc. (ISBN 0-471-65665-8, Wiley International Edition).

References

Baker JM et al. (2006) ‘A metabolic study of substantial equivalence of field-grown genetically modified wheat’. *Plant Biotechnol J* 4: 381-392.

D’Aiuto L et al. (2006) ‘Evidence of the capability of the CMV enhancer to activate in transgene expression in mammalian cells’. *DNA and Cell Biology* 25: 171-180.

Filipecki M. and Malepszy S. (2006) ‘Unintended consequences of plant transformation: a molecular insight’. *J Appl Genet* 47: 277-286.

Gasson M and Burke D. (2001) ‘Scientific perspectives on regulating the safety of genetically modified foods’. *Nat Rev Genet.* 2:217-22.

Ghavididel A. et al. (2005) 'A skeleton of the human interactome'. *Cell* 122 :830-832.

Haslberger AG. (2006) ‘Need for an “integrated safety assessment” of GMOs, linking food safety and environmental considerations’. *J Agric Food Chem* 54: 3173-3180.

Hemm, MR et al. (2003) ‘The *Arobidopsis ref2* mutant is defective in the gene encoding CYP83A1 and shows both phenylpropanoid and glucosinolate phenotypes’. *Plant Cell* 15: 179-194.

Hernandez et al. (2003) ‘A Specific Real-Time Quantitative PCR Detection System for Event MON810 in Maize YieldGard® Based on the 3’-Transgene Integration Sequence’. *Transgenic Res*. 12: 179-189.

Hoebeeck J, Speleman F and Vandesompele J. (2007) ‘Real-time quantitative PCR as an alternative to southern blot or fluorescence in situ hybridization for detection of gene copy number changes’. *Methods Mol Biol* 353: 205-226.

Kohli A et al. (1999) ‘Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination’. *Plant J*. 17:591-601.

Latham JR, Wilson AK, Steinbrecher RA. (2006) ‘The mutational consequences of plant transformation’. *J Biomed Biotechnol.* 2006:1-7.

Matthews D et al. (2005) ‘Toxic secondary metabolite production in genetically modified potatoes in response to stress’. *J. Agric. Food Chem.* 2005, 53, 7766-7776.

Matzke MA and Matzke AJM. (1996) ‘Stable epigenetic states in differentiated plant cells: implications for somaclonal variation and gene silencing in transgenic plants’. In: Russo et al., eds. *Epigenetic mechanisms of gene regulation*. Cold Spring Harbor, NY: Cold Spring Harbor Press: 377-392.

McCabe, M. S., U. B. Mohapatra, et al. (1999). ‘Integration, expression and inheritance of two linked T-DNA marker genes in transgenic lettuce’. *Molecular Breeding* **5**(4): 329-344.

Memelink J. (2005) ‘Tailoring the metabolome without a loose stitch’. *Trends in Plant Science* 10: 305-307.

Myhre MR et al. (2006) ‘The 35S CaMV plant virus promoter is active in human enterocyte-like cells’. *Eur Food Res Technol* 222, 185-193.

Prescott VE et al. (2005) ‘Transgenic expression of bean α-amylase inhibitor in peas results in altered structure and immunogenicity’. *J. Agricult Food Chem*. 53-9023-9030.

Rang A. et al. (2005) ‘Detection of RNA variants transcribed from the transgene in Roundup Ready soybean’. *Eur Food Res Technol* 220:438-443.

Recillas-Targa F. (2006) ‘Multiple strategies for gene transfer, expression, knockdown, and chromatin influence in mammalian cell lines and transgenic animals’. *Mol Biotechnol* 34: 337-354.

Romer S. et al. (2000) ‘Elevation of the provitamin A content of transgenic tomato plants’. *Nat Biotechnol* 18: 666-669.

Takano M et al. (1997) ‘The structures of integration sites in transgenic rice’. *Plant J*. 11:353-61, 1997.

Uhrig, JF. (2006) ‘Protein interaction networks in plants’. *Planta* 224: 771-781, 2006.

WHO (2005) ‘Study on modern biotechnology, human health and development’. http://www.who.int/foodsafety/biotech/who_study/en/index.html

Zhao, Z., Y. Cao, et al. (2001) ‘Double-stranded RNA injection produces nonspecific defects in Zebrafish’. *Develop. Biol*. 229: 215-223.

Other titles in the TWN Biotechnology & Biosafety Series

No. 1 Fatal Flaws in Food Safety Assessment: Critique of the Joint FAO/WHO Biotechnology & Food Safety Report *by Mae-Wan Ho and Ricarda A Steinbrecher (68 pages)*

No. 2 Gene Technology in the Etiology of Drug-resistant Diseases *by Mae-Wan Ho, Terje Traavik, Orjan Olsvik, Tore Midtvedt, Beatrix Tappeser, C Vyvyan Howard, Christine von Weizsäcker and George C McGavin (112 pages)*

No. 3 SURVIVAL, PERSISTENCE, TRANSFER – An update on current knowledge on GMOs and the fate of their recombinant DNA *by Beatrix Tappeser, Manuela Jäger and Claudia Eckelkamp (48 pages)*

No. 4 HORIZONTAL GENE TRANSFER – The Hidden Hazards of Genetic Engineering *by Mae-Wan Ho (32 pages)*

No. 5 SLIPPING THROUGH THE REGULATORY NET: 'Naked' and 'free' nucleic acids *by Mae-Wan Ho, Angela Ryan, Joe Cummins and Terje Traavik (32 pages)*

No. 6 'GOLDEN RICE' – An Exercise in How Not to Do Science *by Mae-Wan Ho (24 pages)*

No. 7 Genetic Engineering and Omitted Health Research: Still No Answers to Ageing Questions *by Terje Traavik and Jack Heinemann (36 pages)*

No. 8 Potential Socio-Economic, Cultural and Ethical Impacts of GMOs: Prospects for Socio-Economic Impact Assessment *by Elenita C Daño (32 pages)*

The integration of foreign DNA into an established genome, through genetic engineering, may have unanticipated side-effects in the recipient organism. Despite this, genetically modified organisms (GMOs) are often claimed to be 'substantially equivalent' to their unmodified counterparts. This paper examines whether this claim is justified, by looking at whether a GMO, compared to its unmodified counterpart, has experienced changes in the interacting regulatory parts, its 'interactome': the genome, epigenome, transcriptome, proteome and metabolome, which work as overlapping layers of information involved in cellular function. Examples are given, from a precautionary point of view, of the changes that may occur in modified genomes and the consequences they may have.

Professor Kaare M Nielsen *is a Professor of Pharmaceutical Microbiology at the Department of Pharmacy, the University of Tromsø, Norway. Nielsen holds a PhD in Biology from the Norwegian University of Science and Technology, Trondheim, and has completed postdoctoral studies in Germany, the Netherlands, and the USA. Nielsen is an adviser to GenØk-Centre for Biosafety. His research group focuses on understanding the molecular mechanism behind, and the ecological effects and evolutionary impacts of, horizontal transfer of genes into various organisms.*

Dr David Quist *is a senior scientist at GenØk-Centre for Biosafety in Tromsø, Norway. He received his PhD from the University of California, Berkeley, where he studied the fate of transgenic DNA in traditional varieties of maize in Oaxaca, Mexico. His current work focuses on the detection and characterisation of transgenes in varying biological and ecological contexts, risk research in emerging biotechnologies, and eco-epidemiological approaches to monitoring transgene flow.*

Professor Terje Traavik *is the author of more than 200 scientific articles and book chapters. He was Professor of Virology at the University of Tromsø, Norway, 1983–2003. Originally a medical and molecular virologist, Traavik later crossed into molecular and cellular cancer research. In 1997, he initiated and became the Scientific Director of GenØk-Centre for Biosafety, and since 2003 he has also been Professor of Gene Ecology at the University of Tromsø. He is currently a member of the Norwegian Biotechnology Advisory Board and adviser on environmental safety of GM vaccines for the European Medicines Agency (EMEA).*

BIOTECHNOLOGY & BIOSAFETY SERIES

is a series of papers published by the Third World Network. It is aimed at deepening public understanding of ecological and safety aspects of new biotechnologies, especially genetic engineering.