Review of the report by Velimirov *et al* (2008): "Biological effects of transgenic maize NK603xMON810 fed in long term reproduction studies in mice"

Food Standards Australia New Zealand
January 2009

1. INTRODUCTION

In November 2008, the Austrian Ministry of Health, Family and Youth publicly released the results of a study it had commissioned to investigate long term reproductive effects in mice following dietary exposure to a genetically modified (GM) corn. The study report titled "*Biological effects of transgenic maize NK603 x MON810 fed in long term reproduction studies in mice*" by Velimirov *et al* (2008) was posted on the Government website.
The study investigated the possible reproductive and long-term effects of NK603xMON810 corn, using three experimental systems:

1. a conventional multigenerational study (MGS);
2. a life-term (feeding) study (LTS); and
3. reproductive assessment by continuous breeding (RACB)

The corn crop used in this study was a conventionally bred cross of the two independent GM corn lines, MON810 and NK603. The MON810xNK603 crop is known as a "stacked" line because it contains both the insect-protection and herbicide-tolerant traits. With regard to traits that have been stacked by conventional breeding, FSANZ considers that once an independent GM crop line has undergone assessment and is approved, it may be used in any conventional plant breeding program without further assessment.

The authors' evaluation of the results concluded that statistically significant reductions in fertility indices were evident in animals in the RACB study, especially in the third and fourth litters. No such effect was evident in the MGS study, and no treatment-related effects were reported in the LTS study. The purpose of the current review was to examine the scientific basis for the conclusions reported by Velimirov *et al* and consider whether these findings could have any impact on the previously determined safety of food derived from corn lines NK603 and MON810.

2. BACKGROUND

FSANZ previously received separate applications seeking approval for corn lines NK603 and MON810 in the *Australia New Zealand Food Standards Code*. Neither NK603 nor MON863 corn is grown in Australia or New Zealand but may be found in imported processed foods.

Insect resistant corn line MON810 (Application A346) was approved by FSANZ in 2000 (assessment report available at http://www.foodstandards.gov.au/_srcfiles/A346%20FA.doc). The insect resistance trait is conferred by the presence in the corn of the *cry1A(b)* gene derived from the soil bacterium, *Bacillus thuringiensis subsp. kurstaki*. The gene codes for a Bt protein that is toxic to the larvae of certain species of Lepidopteran insects. The safety assessment of MON810 corn did not identify any public health and safety concerns.

Glyphosate tolerant corn line NK603 (Application A416) was approved by FSANZ in 2002 (assessment report available at http://www.foodstandards.gov.au/_srcfiles/A416_FAR.pdf). The glyphosate-tolerance trait is conferred by the presence in the corn of the *epsps* gene derived from a common soil bacterium, *Agrobacterium sp*. strain CP4 (*cp4 epsps*).

The expressed CP4 EPSPS enzyme has a reduced affinity for the herbicide compared with the native corn enzyme, and thus confers glyphosate tolerance to the whole plant. The safety assessment concluded that food derived from corn line NK603 was as safe for human consumption as food from other commercial corn varieties.

The safety assessments of corn lines MON810 and NK603 followed the standard assessment protocol developed by FSANZ specifically for GM foods. At the time, FSANZ evaluated all available safety data including an acute oral toxicity study in mice of the novel insecticidal protein present in MON810 corn, and an acute oral toxicity study in mice of the novel protein in NK603 corn. In addition, feeding studies using broiler chickens, pigs and rats did not show any adverse effects when fed diets containing NK603 corn. Both lines have been approved for food use in a number of overseas countries including Canada and the United States.

The protocol used by FSANZ to evaluate GM food safety is based on scientific principles developed at the international level by expert bodies such as the FAO/WHO, the OECD and Codex. Similar assessment protocols for GM foods are used by food regulators in Canada, Japan, the United States and by the European Food Safety Authority (EFSA) who provide risk assessments of foods and feeds to the European Commission (EC).

Opposition to GM crops

FSANZ is aware that the Austrian government has imposed numerous moratoria on corn line MON810 and certain other GM crops over a number of years and for a variety of reasons. On 8 May 2008, based on favourable conclusions issued by the EFSA, the EC forced Austria to lift its moratorium on MON810 corn with respect to importation for human food and animal feed uses. Austria has extended its moratorium on the cultivation of MON810 corn in that country.

3. REPRODUCTION STUDIES

3.1 Test diets

In all three experiments, the test diets consisted of 33% corn, either NK603xMON810 corn (GM), the near isogenic line (ISO, non-GM), or an unrelated commercial non-GM corn grown in Austria (REF-A). Corn line NK603xMON810 was produced by crossing NK603 and MON810 inbred GM lines using conventional plant breeding.

Due to the lack of availability of a non-GM parental line, the authors initially chose one commercially-available non-GM line as a comparator, a NK603-near isogenic line (trade name DKC 26-75). The GM and ISO corns were grown in Canada in 2005 and 2007. As slight GM contamination of ISO harvested in 2005 was detected, a second substantially equivalent non-GM cultivar (Sarastro) was chosen and grown in Austria (designated REF A).

All three corns (i.e. GM, ISO and REF A) were incorporated into standard laboratory mouse feed at a level of 33%. The authors stated that in order to prevent heat-inactivation of any

constituents, diets were presented to mice as a meal rather than as pellets to avoid the high temperature and pressure usually needed to produce pelleted rodent feed. Compositional analysis on an unspecified number of samples indicated that the three corns and their respective diets were equivalent in terms of crude nutrients, vitamins, minerals and trace elements. Additional analyses on the whole diets indicated their equivalence with regard to fatty acids, amino acids, microbial contamination and pesticide residues.

3.1.1 Evaluation

The study report did not indicate how many batches of diet were prepared, the stability of the diets in the form of a meal, or how many samples of corn/diet from each treatment were actually analysed. Based on the presentation of data on the composition of the diet, it appears that only one sample from each diet was analysed. This represents an inadequate sampling regime, given that the authors' repeatedly stated the importance of examining the potential effect of transgenic feed on physiological and metabolic processes. In addition, the use of a single dietary concentration precludes the establishment of a dose-response relationship, which is a key study design variable supporting the identification of a treatment-related effect, and thus substantially limits the discriminative power of this study.

3.2 Mouse strain

A pathogen-free, out-bred mouse strain, OF1, was used for all three experiments. Although the source was not specified in the study report, the OF1 mouse strain was originally established at Charles River Laboratories, France. While claiming the use of this strain was to "ensure good breeding success and provide a diverse genetic background for potential feed impacts beyond fertility parameters", the authors suggested less variable results could have been obtained if an inbred strain was used. The lack of an identified source for the animals used impedes the use of historical control data for this mouse strain which would otherwise assist in the identification of natural variation in the parameters of interest.

3.3 Multigeneration study (MGS)

3.3.1 Materials and methods

Mice and their offspring were exposed through all phases of life to the test or control diets for four parental generations. The F0 parents were exposed from birth and 18-24 pairs mated at seven weeks of age. Males were killed one week after mating. Offspring (F1) were weaned after three weeks and mated at seven weeks. The whole process was repeated to generate F2 and F3 parents/offspring and F4 offspring.

For parental animals, feed consumption (unspecified time) and bodyweight were recorded (females: at and 1 week after mating and 1, 2 and 3 weeks after delivery; males: at and 1 week after mating). The following reproduction parameters were recorded: deliveries; litter size; pup number at birth and weaning (per pair and per group); and pup deaths/group (birth to weaning). The following pup parameters were recorded: litter weight (0, 1, 2, 3, 4, 5, 6, 7, 14 and 21 days after birth); individual pup weights (0, 7, 14 and 21 days after birth); individual female and male pup weights (4 and 5 weeks after birth); and bodyweight gain from birth to weaning.

Liver, spleen, kidney and testes weights from five F2, F3 and F4 pups/sex/group were recorded at approximately five weeks of age. The intestine, kidneys, liver, pancreas, spleen, lung and testes from five F3 pups/sex/group were histopathologically examined at seven weeks of age. The small intestines from these same mice were examined immunohistochemically for the presence of $CD3^+$ lymphocytes, $CD20^+$ B-lymphocytes and macrophages.

The liver, pancreas and spleen from five F3 pups/sex/group were examined by electron microscopy for the following morphometric parameters: area and perimeter of nuclei; area of nucleoli and nucleolar components; and number of nuclear pores. It was unclear from the study report whether these were the same F3 mice used for the previous tissue analysis. Results were statistically analysed using a number of methods which are summarised in Table 1. Statistical comparisons were made between GM and ISO groups, and ISO and REF A groups; differences were considered statistically significant at $p<0.05$.

3.3.2 Results

Five dams died during the study (two from the ISO group and three from the GM group) but the deaths were unrelated to treatment. There was no treatment-related effect on food consumption or bodyweight. Graphically-presented data illustrated that there was a statistically significant reduction in food consumption in proceeding generations in both GM and ISO groups. The bodyweight of dams was successively significantly lower over the four generations in all diet groups.

Table 1: Summary of statistical tests performed on MGS data

Statistical test	Parameters
Kolmogorov-Smirnov-test	Test for normal distribution of data
t-test	Feed consumption, bodyweight, litter weight, pup numbers, pup losses and pup bodyweight gain; organ weights; CD3 staining; ultrastructural data
χ^2-test	Frequency of deliveries per group
Levene test	Homogeneity of variances
2-way ANOVA & Duncan's post hoc test	Further comparison of intra-group differences

There were no significant differences in any reproduction parameters between treatment groups. Similar to the inter-generational reduction in feed consumption and bodyweight, significant decreases in litter size were observed in successive generations in both the control and GM groups (illustrated graphically). In addition, the number of non-deliveries per group increased with successive generations (5.6/0, 4.2/8.3, 20.8/12.5 and 27.3/36.4% in ISO/GM groups in the F0, F1, F2 and F3 generations, respectively), and the litter size decreased (see Table 2).

Despite finding no statistical difference between the GM and ISO groups, and a large variation in the results, the authors stated that pup losses *were always higher in the GM group* and over all generations *about twice as many pups were lost in the GM group as compared to the ISO group* (14.59% vs 7.4%). These claims are inconsistent with statements in the report that all data showed a high degree of variation. There was no attempt by the authors to evaluate whether or not the findings were treatment-related.

3.3.3 Evaluation

Careful examination of the data presented in the reproduction studies shows major deficiencies in the interpretation of results. An inspection of the tabulated results for various measures of fertility in the mice revealed several systematic calculation errors with regard to (i) average litter size at birth, (ii) average litter size at weaning, and (iii) pup losses per group (Table 2). These calculation errors have lead the authors to incorrectly report significant differences between ISO and GM groups that are simply not supported by the data. FSANZ has revised the calculations and, once corrected, there appears to be no significant differences in litter weight or individual pup weight at birth or after weaning between the GM and ISO control groups over four generations. The corrected values have been highlighted in Table 2.

In addition, the authors analysed litter sizes and pup weights by dividing litters into those with ≤ 8 pups, and those with > 8 pups (i.e. small and large litters, respectively). When litters were categorised in this manner, and based on the incorrectly reported values, the authors found significant differences in individual pup weights between the GM and ISO groups in the small, but not large, litters.

Table 2: Selected litter parameters

Parameter	F0	F1	F2	F3
ISO (control)				
Deliveries	17/18 pairs	23/24 pairs	19/24 pairs	16/22 pairs
Total pups at birth	185	198	202	145
Total pups at weaning	151	192	191	142
Average litter size at birth[1]	10.28±0.980 **[10.88]**	8.25±0.778 **[8.61]**	8.42±1.025 **[10.63]**	6.59±1.146 **[9.06]**
Average litter size at weaning[2]	8.39±0.936 **[8.88]**	8.00±0.766 **[8.35]**	7.96±0.928 **[10.05]**	6.45±1.040 **[8.88]**
Average pup losses/group[3]	2.06±0.683 **[2.00]**	0.26±0.157	0.58±0.289	0.19±0.136
GM				
Deliveries	18/18 pairs	22/24 pairs	21/24 pairs	14/22 pairs
Total pups at birth	184	189	208	125
Total pups at weaning	138	167	183	115
Average litter size at birth[1]	10.22±0.629	7.88±0.779 **[8.59]**	8.92±0.875 **[9.90]**	5.68±1.10 **[8.93]**
Average litter size at weaning[2]	7.67±0.792	6.96±0.786 **[7.59]**	7.63±0.85 **[8.71]**	5.23±1.03 **[8.21]**
Average pup losses/group[3]	2.61±0.837 **[2.56]**	1.00±0.510	2.95±0.631 **[1.19]**	0.71±0.322
REF A (control)				
Deliveries	16/18	21/24	22/24	20/22
Total pups at birth	138	194	230	199
Total pups at weaning	121	176	209	189

Parameter	F0	F1	F2	F3
Average litter size at birth[1]	7.67±1.042 **[8.63]**	7.46+0.736 **[9.24]**	9.20±0.735 **[10.45]**	9.05±0.774 **[9.95]**
Average litter size at weaning[2]	6.72±0.928 **[7.56]**	6.77±0.705 **[8.38]**	8.36±0.709 **[9.5]**	8.59±0.732 **[9.45]**
Average pup losses/group[3]	1.06±0.322	0.78±0.281 **[0.86]**	0.91±0.266 **[0.95]**	0.50±0.212

Results are expressed as the mean ± one standard; bolded values indicate a computational error in the original study report, with the correct value contained in parentheses; 1 = total pups at birth ÷ deliveries; 2 = total pups at weaning ÷ deliveries; 3 = (total pups at birth – total pups at weaning) ÷ deliveries

Disregarding the inaccuracies in the data, any apparent differences between GM and non-GM groups were in any case inconsistent between generations, with the GM group having a higher mean pup bodyweight in the F1 and F4 generations, but lower in the F2 generation. In the absence of a consistent pattern of change, there is no scientific justification for the authors to conclude that their observations were diet-related.

The authors also compared their results from the two non-GM diets ISO and REF A, and found a number of statistically significant differences:

- significantly lower (p=0.001) feed consumption in the REF A than the ISO group in the F1 generation;
- feed consumption was significantly lower in proceeding generations (as seen in the GM group);
- significant differences in bodyweight;
- in small litters, significant lower (p<0.05) individual pup weights in the ISO than the REF A group in the F3 generation; and
- at 5 week of age, significantly heavier (p=0.021) offspring in the F1 and F2 generations in the REF A *versus* ISO group.

It is noted that the number of parameters that were statistically different between the two non-GM control groups exceeded that between the GM and ISO groups. While the two non-GM corns were grown in different geographical locations (Canada and Austria) and derived from different lines, compositional analysis indicated their dietary equivalence with regard to crude nutrients, vitamins, minerals and trace elements. Additional analyses of the whole diets indicated they were also equivalent in terms of constituent fatty acids and amino acids. Given these similarities, the above statistical differences observed between mice in the two control groups clearly reflect normal biological variability.

These results, showing statistical differences between mice consuming control diets that are compositionally equivalent, highlight two important points relevant to this study. Firstly, intrinsic variability (noise) within and between any experimental groups should be expected. Secondly, a robust interpretation of a toxicological study should not be based solely on statistical differences.

Organ weight analysis revealed a number of statistically significance but inconsistent differences between the GM and ISO control group:

- significantly increased (p=0.004) relative spleen weight in five week old F2 males; and
- significantly lower relative kidney weight in F3 and F4 females, and F2 males (p=0.045, 0.029 & 0.000 respectively).

In the absence of any differences in absolute organ weights (data not provided in the study report), inconsistency over generations, and the absence of any histopathological abnormalities, none of these differences can be interpreted as treatment-related. It is noted that statistical comparisons between the two control groups revealed significant differences in relative testes (F3 and F4), liver (F3 males and females) and kidney (F4 females) weights.

Immunohistochemical staining of the intestine revealed significantly lower ($p<0.05$) $CD3^+$ intraepithelial lymphocytes in the second intestinal segment of males in the GM group *versus* the ISO control (0.67 ± 0.068 *versus* 1.13 ± 0.070 per 0.1 mm^2, respectively), but in no other regions of the small or large intestine.
However, significantly lower ($p<0.05$) $CD3^+$ intraepithelial lymphocytes were also detected in the REF A control (0.30 ± 0.106 *versus* 0.67 ± 0.068 per 0.1 mm^2, respectively) in the same segment. In females, significantly higher $CD3^+$ lymphocytes were measured in the 3rd intestinal segment of the GM group relative to the ISO control (1.25 ± 0.031 *versus* 0.57 ± 0.056 per 0.1 mm^2, respectively), which was considered to be incidental in nature.

There were no significant differences in the number of $CD20^+$ cells or macrophages between the groups. This part of the experiment was limited in value because only a small number of F3 mice were examined.

A number of significant differences in morphometric parameters ascertained by electron microscopy between the GM and ISO groups were reported. However, significant differences were also detected between the two control groups and therefore such differences are not considered treatment-related.

3.3.4 Conclusion

Under the conditions of this study, *ad libitum* feeding of NK603xMON810 corn to mice at a dietary level of 33% had no adverse effect on reproductive or toxicological parameters through four parental generations and their offspring.

3.4 Life-long feeding study (LTS)

Ten F1 female mice/group obtained from the MGS were continued on treatment for 22 months. Feed consumption and bodyweight were recorded at weekly and fortnightly intervals, respectively. The number of deaths during the study was recorded. At death or sacrifice, mice were subjected to a gross necropsy and a histopathological evaluation (details of methods and organs/tissues examined were not specified). Standard haematological, clinical chemistry and urinary endpoints were not investigated. Data were statistically analysed by “descriptive statistics” and the Kaplan-Meier survival test.

3.4.1 Results

There was no significant difference in survival, with an average life span of 16.3, 15.7 and 17.0 months in the ISO, REF A and GM groups, respectively. Only two mice per group remained alive at the end of 22 months, with leukaemia reported to be the common cause of death. The authors speculated this to have been *triggered by the activation of a murine retrovirus inherent in the genome of the test mice*.

There was no significant difference in food consumption between any of the diet groups. It was stated that bodyweights could not be compared because of the confounding effect of cancer. No necropsy or histopathology data were provided in the study report.

3.4.2 Evaluation

In attempting to dismiss the lack of any findings, the authors state that long term studies "*have limitations as the majority of outbred or inbred strains develop different forms of cancer*" and therefore "*diet related differences could be masked and not really assessed*". These comments are extraordinary in view of the level of resources (both animal and human) required for such long-term studies and noting that previously, suggestions for these types of studies to be compulsory in GM food assessments have been expressed.

The low group size used over such a long duration of dosing, and the use of only female animals, represents an inadequate study design. Specifically, OECD Test Guideline 452[1] recommends that for rodents, at least twenty animals/sex/group should be used in a long-term study of this type.

The overall study design should have addressed many of the expected and potentially confounding issues, such as the onset of age-related disease, that beset long term animal studies. For example, it is common practice to compare body weight gain in animals over the first 12-months of a long-term study when the greatest treatment-related differences commonly manifest. As indicated above, the use of sufficient numbers of animals per group would also address some of the problems encountered in the Austrian study.

Most importantly, for the majority of strains of laboratory animals, the historical control incidence of cancer is known and can therefore be used to aid in the interpretation of results. This basic information was apparently not considered in the study design and methodology of the Austrian study, and therefore the results are extremely limited in scientific value.

3.5 Reproductive assessment by continuous breeding (RACB)

3.5.1 Materials and methods

Twenty-four breeding pairs of mice per group were commenced on one of the three test diets one week prior to cohabitation and then for a period of 20 weeks. During this time, each pair was allowed to deliver four litters at approximately 3-4 weeks apart. Feed intake (unspecified time) and bodyweight were recorded for parental mice (the latter at and 1 week after mating, and at 1, 2 and 3 weeks after mating). The following reproduction parameters were analysed: deliveries/group; litter size distribution; number of pups/pair and /group at birth and weaning; pup losses/group (birth to weaning); and birth interval. The following pup parameters were recorded: litter weight (0, 1, 2, 3, 4, 5, 6, 7, 14 and 21 days after birth); pup weight (0, 7, 14 and 21 days after birth); and bodyweight gain from birth to weaning. Data were statistically analysed using the relevant tests described in Table 1 for the MGS data.

[1] http://oberon.sourceoecd.org/vl=3728241/cl=21/nw=1/rpsv/cgi-bin/fulltextew.pl?prpsv=/ij/oecdjournals/1607310x/v1n4/s37/p1.idx

3.5.2 Results

No significant differences in feed consumption or bodyweight were observed between the GM and ISO groups. Selected parameters for the four litters are summarised in **Table 3**. A marginal decline in number of litters in the GM group over succeeding litters was observed, however this was not significantly different to the ISO group (23, 23, 22 and 20 litters per 24 pairs over the 1st, 2nd, 3rd and 4th litters, respectively, *versus* 24 in the ISO control for all litters). There were no significant differences in any parameters in the 1st or 2nd litters in the GM or ISO groups.

3.5.3 Evaluation

The authors claim that females in the GM group always had smaller litters compared to females of the ISO group. The 3rd and 4th litters of the GM group were significantly smaller than the ISO control group. While the authors interpret this finding as a treatment-related effect, careful scrutiny of their results does not support this view.

Firstly, in tabulated results, several calculation errors have been made with regard to the size of the 4th litter at birth and weaning in the GM group (refer to Table 3). While the absence of individual litter data in the study report precludes an independent statistical analysis, by observation, the corrected numbers for the GM litters appear to be similar to those in the ISO group.

Secondly, the sizes of the 3rd and 4th litters (birth and weaning) of the GM group are actually higher than that of the 1st litter, which was not significantly different to the control. It would appear therefore that the statistical difference in the 3rd and 4th litters is based on the larger (>11) litter size in the control group. It is worth noting that the size of the 3rd and 4th litters of the ISO control group exceeds the normal litter size (of 9-10)[2] for this particular mouse strain and was significantly higher than the 1st litter. On this basis, the significant differences can not be reasonably interpreted as treatment-related.

An obvious difference between the GM and ISO groups was that pup deaths in the GM group were actually lower than the ISO group in the 1st, 3rd and 4th litters, reaching statistical significance in the 3rd litter; this difference was not discussed by the authors.

Unlike the MGS, no statistical comparisons were made between the two control groups. Based on results from the MGS, a number of statistically significant differences would have been expected, reflecting natural variability. However, the authors presented no discussion on the absence of a comparison between non-GM control groups in the RACB.

The authors reported that the average weight of the 3rd litter of the GM group was significantly lower (p=0.026, 0.04 and 0.031, respectively) than the ISO control group at birth, day two and at weaning; no differences were reported for the 1st and 2nd litters. They also stated that the average weight of the 4th litter (at birth) of the GM group was significantly lower than the ISO control group. Comparing average litter weights is of questionable value because it is dependent on litter size, which varied between the GM and ISO control groups.

[2] http://www.criver.com/SiteCollectionDocuments/OF1-MICE.pdf

Using the approach of dividing litters into small and large (as in the MGS), the average pup weights in the ISO control were significantly lower than the small 4th litter of the GM group. The authors note that *individual pup masses in litters with a high number of pups are commonly lower*. Therefore, the converse should follow: that pup weights should be higher in litters with a low numbers of pups. However, as the authors chose to differentiate litters into large and small, this should have controlled for this effect.

Table 3: Selected litter parameters

Parameter	**1st litter**	**2nd litter**	**3rd litter**	**4th litter**
ISO (control)				
Deliveries	24/24	24/24	24/24	24/24
Total pups at birth	216	260	286	273
Total pups at weaning	200	241	254	235
Average litter size at birth[1]	9.00±0.614	10.83+0.473	11.92+0.496	11.38±0.462
Average litter size at weaning[2]	8.33±0.56	10.04+0.480	10.58+0.454	9.79±0.525
Average pup losses/group[3]	0.67+0.305	0.79+0.289	1.33+0.433	1.58±0.371
GM				
Deliveries	23/24	23/24	22/24	20/24
Total pups at birth	189	245	213	197
Total pups at weaning	187	226	207	173
Average litter size at birth[1]	8.22±0.590	10.65+0.649	9.68+0.688*	8.21±1.077* **[9.85]**
Average litter size at weaning[2]	8.13±0.560	9.83+0.550	9.06+0.820 **[9.41]**	7.21±0.985* **[8.65]**
Average pup losses/group[3]	0.09±0.060	0.83+0.375	0.12+0.081* **[0.27]**	1.00±0.376 **[1.2]**

Results are expressed as the mean ± one standard; bolded values indicate a computational error in the original study report, with the correct value contained in parentheses; 1 = total pups at birth ÷ deliveries; 2 = total pups at weaning ÷ deliveries; 3 = (total pups at birth – total pups at weaning) ÷ deliveries

The authors reported that the Levene test for the homogeneity of variance revealed a significantly higher (p=0.000) variation in the number of pups at delivery and weaning in the GM *versus* the ISO control group. While the authors concluded that this high variability requires further scrutiny, they provided no scientific rationale for why this should be considered necessary. Rather, the difference is entirely consistent with the large variability already apparent in this study and further investigation is not justified on the basis that no treatment-related effects were evident.

3.5.4 Conclusion

Under the conditions of this study, *ad libitum* feeding of NK603xMON810 corn to mice at a dietary level of 33% had no adverse effect through one parental generation and four of their successive litters.

4. ANALYSIS OF GENE EXPRESSION

With the aim of identifying any differences in gene expression with different diets, animals in the multigeneration study were used for DNA microarray analyses. Samples of intestinal tissue were obtained from twelve F3 males/diet group at seven weeks of age. A segment of small intestine was dissected from each animal to include a variety of cell lineages including epithelial, immune and endothelial cells.

Total RNA was extracted from the tissue samples and then used to generate labelled cDNA probes. Triplicate samples were generated for each group.

Pairwise comparisons of data included GM vs ISO, ISO vs A REF, and GM vs ISO and A REF, based on the rationale that ISO and A REF diets were both non-GM diets. Analysis of data used commercial software. As this was considered a pilot study, a minimum two-fold change, representing either up or down regulated genes, was applied as a cut-off point.

The list of differentially expressed genes was compared to a publicly available database of classified mouse genes, and where possible genes were allocated to biological processes and pathways. According to the methods used, the comparisons indicated the following:

GM vs ISO: 440 genes differentially expressed. Of these, those showing at least a 2 fold change in expression revealed 43 genes up-regulated and 98 down-regulated. The genes identified were found to be involved in a wide range of biological processes including protein biosynthesis, signalling and protein metabolism.

ISO vs REF A: 1016 genes differentially expressed. Of these, those showing at least a 2 fold change in expression revealed 182 genes up-regulated and 4 down-regulated. Identified genes were found to be involved in a wide range of biological processes including signalling pathways, circadian cycle and T-cell activation pathway.

GM vs ISO and REF A: 2374 genes differentially expressed. Of these, 421 genes showed at least a 2 fold change in expression. Identified genes were found to be involved in a wide range of biological processes particularly protein biosynthesis, protein metabolism and modification.

Further gene expression analysis focussed on 45 differentially regulated genes found between the GM and combined non-GM groups using commercial TaqMan® low density arrays (TLDA) and quantitative real time polymerase chain reaction (qRT-PCR). After statistical treatment of the data, eleven genes were confirmed. The same gene analyses were apparently used on intestinal samples obtained from animals in the RACB experiment, however no technical details were provided in the report. Three genes were identified from these samples.

In addition to the microarray screening, the same set of total RNA samples used in the microarrays on MGS animals were sent to another laboratory specialised in microRNA (miRNA) profiling. The authors state that the chip targeted all 460 major mature miRNAs of mouse contained in the latest version of the murine miRBase 10 database.

4.1 Results

The authors briefly discussed the cellular function, if known, of several of the eleven genes identified in their screening process. This information however did not allow the authors to draw conclusions on the significance of the differences. Apparent differences observed between ISO and GM fed mice detected by microarray analysis occured in several biological processes and pathways. It was noted that although quantitative changes were apparently observed at the level of gene expression, these results provided no information on whether corresponding changes in protein expression could also be demonstrated. The results of the miRNA pairwise comparisons (GM/ISO and ISO/REF A) showed no significantly different expression of miRNAs in samples obtained from animals in either the GM, ISO or REF A diet groups.

4.2 Evaluation

From the experimental details provided, it is not possible to evaluate the approach used in the microarray profiling, the integrity of the samples, the validity of the *in silico* analysis or the statistical methods used in the study. It is therefore not possible to draw any meaningful conclusions regarding comparative changes in gene expression in intestinal tissue obtained from animals fed diets containing either GM, ISO or REF A corn.

The microarray screening data showed many more differentially expressed genes between the two control groups ISO and REF A, than was found when GM and ISO groups were compared. The authors were therefore unable to speculate on the relevance of the data in terms of the two novel traits in the GM corn, nor to any other parameter. In addition, their results provide no information on actual changes in protein expression levels.

Use of the microarray screening approach to detect meaningful changes at the level of gene expression in tissues from animals fed on different diets is considered technically ambitious. Although described as a pilot study, the relevance of these results to an assessment of food safety has not in any case been demonstrated.

Furthermore, in terms of a food safety assessment, a difference between a GM and non-GM counterpart detected in any parameter, including gene expression, is not in itself an indication that the GM food is less safe than the conventional form. Differences may in fact be neutral with respect to food safety, as aptly demonstrated by the detection of compositional differences between conventional cultivars grown in different environments. On the basis of these DNA microarray results, it is not possible to draw any conclusions regarding changes in gene expression that could have an impact on the safety of NK603xMON810 corn.

5. DISCUSSION

The conclusions drawn by Velimirov *et al* concerning a reproductive effect in mice fed a diet containing GM corn are not supported by the evidence presented in this report. Not only can significant flaws in the experimental design be identified, but there is also a general lack of detail and transparency in the presentation of information on methodology and the statistical evaluation of data. These deficiencies appear to have contributed to erroneous interpretations of the results.

For example, an inspection of tabulated results presented in the report for various measures of fertility in the mice reveals several systematic errors that have lead the authors to their conclusion. The following example is used to illustrate the source of these errors:
Fertility in toxicology studies refers to the number of live pups per pregnant dam. Thus, if 20 of 20 mice on the control diet all had litters, but only 16 of 20 mice on the GM diet had litters, the appropriate divisor for each group is 20 and 16, not 20 as used by the investigators. By erroneously dividing the number of pups delivered by the number of dams in each study the authors confound a measure of fertility with a measure of mating success (or fecundity) which is the proportion of mated dams which become pregnant.

Interpreting data from reproduction studies needs to take these two quite separate concepts of fertility and fecundity into account. In simple terms, fertility is a measure of 'actual' reproduction, the ability to deliver live offspring from a successful mating, whereas fecundity is a measure of 'potential' reproduction (or ability to reproduce, or get pregnant). Thus, fertility in mice is determined by dividing the number of delivered pups by the number of pregnant mice. Instead, Velimirov *et al* used the total number of paired mice (pregnant and non-pregnant) in their calculation of fertility.

Fecundity in mice is usually estimated by considering the number of pairings (male/female) that result in a pregnancy. Fecundity in the strain of mice used in the Austrian experiments can be estimated by combining the data from the two large reproduction studies (MGS & RACB), as only one dietary concentration of corn (33%) was used. This is done by adding the total number of pairings of mice on the non-GM control diet and the total number of pregnancies. Performing these calculations, the ratio of pregnant dams:pairings was 250:272 or 91.9% for the animals on the control diet. Similarly, for mice on the GM diet, the ratio of pregnant dams:pairings was 163:184 or 88.6%.

Combining the data from the two reproduction studies to examine fertility, the ratio of delivered pups:pregnant dams was 2526:250 yielding 10.1 pups per pregnancy for the control mice, compared with 1550:163 or 9.5 pups per pregnancy for the GM fed mice. For these parameters, the magnitude of the difference is considered small enough to show no difference in fertility or fecundity between mice on GM or non-GM control diets.

Conventional toxicological assessments rely primarily on scientific judgement and contextual interpretations of data rather than statistical analysis, which is generally used only as a supplemental screening tool to detect possible treatment-related effects. It is important to recognise that when many individual biological measurements are compared using statistical analysis, there will be some significant differences which are found to be unrelated to treatment.

For example, in the current study, this is demonstrated by the statistically significant differences between the two control groups (ISO and REF A) in the MGS with regard to food consumption, bodyweight, litter sizes, pup weights and relative testes, liver and kidney weights. Clearly, differences such as these observed in independent control groups, are attributable to normal biological variation. Moreover, given the large intra-group variation observed throughout this study, the results for the MGS and RACB are strongly indicative of inherent variability in this mouse strain, rather than to significant physiological differences due to diet alone.

While the authors acknowledge the inherent variability reflected in their results, they also show a willingness to interpret any differences between GM and control diet groups as biologically significant. Consequently their interpretations are more conclusive than a robust scientific analysis of the results would reasonably allow. In our view, none of the findings can reasonably be attributed to the nature and composition of the mouse diets, and the authors' statements on this point are largely speculative.

The authors also engage in contradictory discussion on the merits of the outbred mouse strain used in the study, and claim that the RACB is a better assay compared with the MGS, because it generated results that were statistically different between the treatment groups.
In addition, the authors claimed that the MGS was less physiologically stressful with the potential to mask the impact of dietary factors. In fact, no biologically significant effects occurred in either experiment and different parameters were analysed in each system. MGS is an established and validated test method widely used to study the reproductive effects of a range of chemical substances.

The absence of treatment-related differences between the diet groups in a study with such extensive scope does not support the general tenet that effects of toxicological significance occurred due to a GM diet. Given the comprehensive nature of the investigations, there is no justification for further animal studies to be undertaken.

6. CONCLUSION

After evaluating the data presented, FSANZ is unable to confirm that reproduction in mice was adversely affected following incorporation of GM corn NK603xMON810 into the diet. Based on the available evidence, FSANZ concludes that any differences between dietary groups detected in this study merely reflect normal biological variability. FSANZ therefore reaffirms its previous conclusions on the safety of food derived from NK603 and MON810 corn lines. In particular, the authors' recommendations to examine reproductive performance, supplemented with genomic, proteomic and metabolomic studies, on a routine basis to inform the safety assessment of a GM food are not supported by the findings in this study.