

EVALUATION OF HORIZONTAL GENE TRANSFER MONITORING EXPERIMENTS CONDUCTED IN NEW ZEALAND BETWEEN 2004 AND 2009

Jack A. Heinemann^{1,2*}, Brigitta Kurenbach^{1,2} and Nikki Bleyendaal¹

¹Centre for Integrated Research in Biosafety and the School of Biological Sciences, University of Canterbury, Christchurch New Zealand

²GenØk – Centre for Biosafety, Tromsø, Norway

*Corresponding author. Phone +64 3 364 2500 email jack.heinemann@canterbury.ac.nz.

Abstract

In 2002, the Environmental Risk Management Authority of New Zealand (ERMANZ or the Authority) approved an application by the company AgResearch, Ltd. to create and dispose of genetically modified (GM) bovine. As part of its risk management strategy, the Authority imposed a requirement for monitoring soil microorganisms for uptake of transgenes by horizontal gene transfer (HGT). HGT is routinely considered in biosafety risk assessment because it may be a process that underpins eventual adverse effects to human health or the environment.

While granting approval to the company to make GM bovine, the Authority considered that HGT-related risks were potentially non-negligible and therefore approval was contingent upon meeting regulatory controls that made the risk negligible through risk mitigation. ERMENZ's requirements placed upon AgResearch the burden to conduct a monitoring effort capable of delivering the risk mitigating activity that the Authority sought. Using colony hybridisation and PCR, AgResearch monitored antibiotic resistance phenotypes in soil bacteria cultured from samples of soil taken from offal pits containing the carcasses of GM bovine and surrounding control sites between 2004 and 2009 in an attempt to determine if any of the antibiotic resistance in soil bacteria was caused by the uptake of transgenes originally from the GM animals.

The Centre for Integrated Research in Biosafety (INBI) at the University of Canterbury has reviewed AgResearch's reports of these monitoring efforts released to GE Free New Zealand (in Food and Environment) under the Official Information Act.

In summary, AgResearch undertook a challenging project at the forefront of theory and practice in microbial science. However, their experiments suffered from a design that was incapable of detecting HGT with the sensitivity necessary to detect bacteria that might cause the adverse effects of concern to the Authority, including but not restricted to bacteria developing antibiotic resistance because they acquired a resistance gene used in the production of GM bovine. Notably, the sampling depth in all but one year was in the range of 2-6 m above the soil interface with the carcasses. Importantly, no study confirmed that the samples were taken from soil in contact with carcasses.

Moreover, the suitability of control sites and the efficacy of the sampling were not demonstrated. Not just the design but the standards of follow-up on observations and determining causes of negative results (e.g. particularly from routine molecular work such as sequencing and PCR) was below what we would expect, and what we would expect to be sufficient for assurance that risk management controls were met. INBI finds that these experiments were irretrievably flawed for providing baseline data for future soil analysis, effectively monitoring HGT as a risk management strategy or influencing the assessment of the risk of HGT in future applications. We suggest ways AgResearch could have chosen to improve experimental designs and lead to more confidence-building outcomes.

Introduction

In 2002, the Environmental Risk Management Authority of New Zealand (ERMANZ, or the Authority) approved an application from AgResearch, Ltd. to develop and later dispose of genetically modified/engineered bovine (ERMANZ 2002). In reaching this decision, the Authority meticulously described the different risk scenarios that it considered. Based on then available information, some of these risk scenarios could not be effectively categorised as either low risk, low likelihood, or both. The Authority suggested that unintended non-negligible harms might arise from horizontal gene transfer (HGT): *"There are potentially non-negligible risks to the environment that are not related to the ability of the cattle to escape. These risks include unintended insertion of viral cell receptors and creation of new viral reservoirs, and adverse effects arising as a result of HGT"* (p. 47 ERMENZ, 2002, emphasis added).

Generally, the kinds of adverse effects that were described by the Authority included: (a) an increase in potential exposure of humans or animals to antibiotic resistant microorganisms; (b) the possibility that newly

introduced resistance might compromise antibiotic therapy in humans or animals; (c) the possibility that newly introduced resistance might increase the potential for therapy failure through an increase in antibiotic resistance in the soil reservoir; (d) unintended or unexpected new traits that might arise because all possible biochemical activities of these genes could not be described in advance; and (e) unintended or unexpected new traits that might arise in microorganisms through transfer of parts of the genes and gene constructs (including promoter and other DNA sequences that can influence in species-specific ways that genes are expressed or their products are modified).

When the Authority considered the potential for adverse effects arising from the transfer of transgenes or genetic elements from the disposal of GM bovine in unlined offal pits, it readily acknowledged that the GM bovine work would have “*significant uncertainty as to the magnitude and likelihood of the adverse effect arising*” (p. 21 ERMANZ 2002). Consequently, the Authority imposed a risk management plan to insure that the risk of adverse effects would be negligible. The plan had two fundamental requirements for managing risks associated with HGT^{1,2} (ERMANZ 2002):

1. Limits on the kinds of DNA that could be used in the making of GM bovine (controls 9.1 and 9.2); and
2. “*Micro-organisms shall be tested for the presence of the introduced genetic modifications at the disposal sites. If HGT is detected, genetic modification and disposal of cattle shall be immediately halted*” (p. 58 ERMANZ, 2002).

The Authority concluded that “[w]ith these controls in place, the combined non-negligible risks referred to above are considered to be low, even after taking account of uncertainty” (p. 48 ERMANZ, 2002). The Authority directly tied monitoring of soil microorganisms to its risk assessment, and encouraged monitoring to be “*as extensive as possible*” (p. 21 ERMANZ, 2002), saying of many additional controls that “[i]n general, these restrictions are aimed at removing classes of risk associated with HGT, viral and prion diseases, and antibiotic resistance” (p. 46 ERMANZ, 2002).

There is no question that HGT was a dominating issue in the Authority’s³ deliberation. The uncertainties about the likelihood and magnitude of potential adverse effects from HGT were so significant that the Authority required the project to stop should there be simply *the detection* of HGT. We believe that this placed upon AgResearch the reasonable expectation that it would and could conduct a monitoring effort capable of delivering the risk mitigating activity sought by the Authority.

“*The Committee’s view is that every reasonable opportunity should be taken to monitor developments such as this for the occurrence of adverse effects and for information on the significance of pathways such as HGT*” (p. 21 ERMANZ, 2002).

“*The applicant will monitor for HGT at the disposal sites and in the event of HGT being detected the project will be halted and a remediation plan developed (control 6.4)*” (p. 25 of ERMANZ, 2002, emphasis added).

The Authority allowed AgResearch to design, conduct and supervise the monitoring of HGT, and this latitude created a potential conflict of interest for AgResearch when set against its funding criteria and overall goal of delivering commercially applicable research results from the development of GM bovine.

In 2010, the Centre for Integrated Research in Biosafety (INBI) was approached by the civil society organisation GE Free New Zealand (in Food and Environment) to review the gene transfer monitoring work conducted by or for AgResearch. GE Free had received the summaries of the AgResearch work through a

¹ “In assessing risks, the impact of the containment regime was considered (section 45(1)(a)(iii) of the HSNO Act) in relation to the ability of the organisms or any heritable material to escape from containment; the ability to meet the requirements of s45A(2) in regard to removing or destroying material; the management of risks and other factors” (p. 14 ERMANZ, 2002).

² See sections 3.4-3.5 of ERMANZ, 2002.

³ Horizontal gene transfer is mentioned 40 times, and appears in a major heading, in the 62 page decision (ERMANZ, 2002).

request under the Official Information Act. The work was described in a series of reports (dated from November 2004 to December 2009):

- “Microbial characterisation of soils from offal pits” (2004),
- “Monitoring offal pits for puromycin resistance bacteria and the presence of genes encoding resistance to puromycin” (2007),
- “Monitoring offal pits for the presence of puromycin resistant bacteria and the presence of genes encoding resistance to puromycin” (2008),
- “Monitoring offal pits for the presence of puromycin resistant bacteria and the presence of genes encoding resistance to puromycin” (2009a)⁴, and
- “Interim Report Monitoring offal pits for puromycin resistance bacteria and the presence of genes encoding resistance to puromycin” (2009b).

The reports were conducted to demonstrate (in AgResearch’s words): “*that the method of disposal of GM animal carcasses is safe for the environment*” (p. 3 of 2004 report) and form part of AgResearch’s requirements for field testing transgenic bovine (Goven *et al.* 2008).

Specifically, INBI was tasked to provide an independent evaluation of the:

- experimental design of the study series for its suitability for monitoring all relevant risk pathways of HGT;
- methodology of the studies for their suitability for detecting HGT in the soil of offal pits;
- scope of the HGT monitoring effort with regard to its ability to detect potential adverse effects during the timeframe of monitoring and/or before any products of HGT would amplify to the point of being able to cause an adverse effect.

If possible, INBI was to provide advice on how the experimental approach could be improved.

As part of our public service and specifically our formation mission, INBI accepted this task pro bono. Work was initiated in April 2010. On 20 April 2010, INBI submitted supplementary questions to AgResearch (Supplementary Material) with a promise to provide AgResearch with an advanced draft of our final report, to allow AgResearch to register with us any errors of fact before the report was released. AgResearch accepted and released answers to our supplementary questions under the Official Information Act. A draft of this report was then provided to AgResearch on 23 November 2010⁵.

Experimental approaches

HGT is a powerful evolutionary force, observable both in nature and in the laboratory, and is of course even used to create GMOs (e.g. microbes and plants) (Heinemann 1991; Kidwell 1993; Heinemann 1997; Syvanen & Kado 1998; Syvanen & Kado 2002; Heinemann 2003; Hehemann *et al.* 2010; Knight, Bailey & Foster 2010; Moran & Jarvik 2010). HGT provided the foundation techniques for the earliest experiments in molecular biology and microbial genetics and has been studied for nearly a century (Griffith 1928; Avery, MacLeod & McCarty 1944; Lederberg & Tatum 1946; Hershey & Chase 1952; Cavalli-Sforza 1957; Stroun, Anker & Auderset 1970; Heinemann & Kurenbach 2009).

Nevertheless, the difficulty of monitoring HGT should not be underestimated. While the biochemistry of HGT pathways has been studied for many decades, and the effects of HGT are well described in some outcomes [e.g. the evolution of antibiotic resistance (Amábile-Cuevas, Cardenas-Garcia & Ludgar 1995; Anonymous 1995; Levy 1998), recombination of novel viruses (Gibbs & Weiller 1999), rate of change of the *Escherichia coli* chromosome (Ochman, Lawrence & Groisman 2000)], capturing transfer in nature in real time has been a considerable challenge. To do it for risk assessment is a formidable task with its own additional challenges. The AgResearch approach to monitoring HGT described below is not unique and contributes to a series of attempts that have never resulted in confirmed positive detections without experimental intervention (Conner, Glare & Nap 2003). So why have dedicated studies of transgene transfer using soil microcosms generally yielded negative results?

There are two possibilities (Heinemann 2003). The first is that HGT is exceedingly rare. We can effectively discount this, because the biochemical pathways by which genetic material moves during HGT are environmentally and phylogenetically ubiquitous, and very familiar biological events are mediated by HGT (Heinemann 2003; Heinemann & Traavik 2004). These include the movements of gene vectors such as plasmids, viruses and transposable elements that are observed in both prokaryotes (e.g. bacteria) and

⁴ This report appears not to be for only internal use because it has the phrase on the title page: “Client report number: SC11 0/2009/32”.

⁵ AgResearch replied with comments on 14 December 2010. We thank AgResearch for their cooperation and suggestions for improving clarity.

eukaryotes (e.g., animals, plants and fungi): “[i]t can therefore be concluded that eukaryotes possess the same capacity and similar mechanisms for effective HGT as prokaryotes do, and laboratory experiments have shown that these mechanisms are functional” (p. 132 de la Cruz & Davies 2000).

The second possibility is that negative results are due to experimental bias:

- either the experimental techniques are not sensitive enough to detect HGT at levels that can eventually result in environmentally relevant effects; and/or
- the experiments have design flaws.

Both of these last two reasons are sufficient to explain current experience (Heinemann & Traavik 2004).

Importantly, the operational definitions of HGT are inconsistent particularly in the biosafety-relevant literature. Lately HGT has come to be understood by some as transfers of genetic material that either transform the phenotype of a recipient *organism* or at least result in large fragments of DNA being irreversibly incorporated into chromosomes. This technical definition both ignores the many pathways of gene transfer that have informed study of HGT for decades and is far too restrictive for use in risk assessment, risk management and monitoring. First, even small changes of a recipient genome due to uptake of DNA can change the function of altered genes. Second, not all changes to the ecosystem are caused by changes in organisms. HGT can result in the creation of novel gene vectors such as viruses and plasmids that do not necessarily amplify through bacterial reproduction. Finally, transferred DNA does not have to be stably retained in order for it to change the traits of an organism. Even transient exposure to DNA may result in heritable epigenetic changes (e.g., Al-Kaff *et al.* 2000; Heinemann 2009). These changes are not maintained by the transferred DNA.

The effect of unjustifiably restricted definitions is to both limit what can qualify as evidence of HGT and significantly bias the results both qualitatively and quantitatively in favour of non-detection. Fortunately, ERMENZ has a more responsible definition of HGT: “HGT is defined as the transfer of genetic material from one organism to another organism outside the context of parent to offspring (i.e. vertical) reproduction” (p. 1 ERMENZ 2006). This definition captures all HGT phenomena and therefore all pathways through which HGT might contribute to an adverse effect. Also of significance, the 2002 amendment to the Hazardous Substances and New Organisms Act of New Zealand clarified the need of risk assessors to be mindful of the full range of genetic material that might contribute to potential adverse effects (Heinemann 2004).

Adverse effects arising from HGT are only possible if HGT occurs. To rule out HGT altogether is difficult, probably impossible with existing technology (Heinemann & Traavik 2004). Nevertheless, this is how AgResearch attempted to address the risk of HGT. They chose to conduct a monitoring effort that would test whether or not HGT was occurring. If it was not detected, then the conclusion would be that no adverse effect could subsequently arise. The burden was therefore to conduct a monitoring effort capable of proving that the negative result (no detection of HGT) was robust and of sufficient scale to inspire confidence that HGT was not happening at frequencies (transfer/time) that might ultimately lead to an adverse effect. This was not their only option, as we discuss at the end of the report.

The methodology used by AgResearch started with taking soil samples at varying depths from ofal pits and control sites after disposal of GM bovine carcasses (Table 1). The soil was suspended in liquid medium and then diluted for the purpose of identifying microorganisms that displayed a target phenotype and evidence of DNA originating from the discarded carcasses. This methodology was only capable of monitoring a small subset of relevant microorganisms (essentially only soil bacteria) and only those that are culturable under aerobic conditions, that is, only the estimated 0.1-1% of the bacteria in soil (Daniel 2004). While all reports list how soil samples were diluted in the procedure and normalised to number of bacteria per gram of soil, none made mention of the actual weight of soil examined. Thus, the implied scale of the experiment in terms of bacteria/gram was significantly higher than the actual number of grams of soil or number of bacteria surveyed.

The reports did not draw on a completely uniform methodology. Differences of significance are noted in the next section.

Table 1. Reported sampling depths

Year	Offal Pit (metres)	Control site (metres)
2004	Various to 5.5	Various to 2.4
2005	Not included in report	Not included in report
2007	0.2-0.3 ^A	0.2-0.3
2008	0.2-0.3	0.2-0.3
2009a	0.2-0.3	0.2-0.3
2009b	0.15 ^B	0.15

^A The reasons for shallow sampling were reportedly: 1) the claim that the studies described in the 2004 report

established a lack of change in resistance phenotypes as a function of sample depth, and 2) difficulty in getting deeper samples.

^B No justification was given for the change in depth to half that of the studies 2007 - 2009a.

Report summaries

2004

Soil samples for this report were collected in May 2004. Total CFUs/g soil (see Glossary) were 10^3 - 10^6 for controls, and 10^4 - 10^7 for pits. The most important data are those that come from bacteria displaying the target phenotype. The ability to survive on media supplemented with 100µg/ml kanamycin was demonstrated for 0-22% of aerobically culturable bacteria for control samples and 0-13% for offal pit samples. Survival on 125 µg/ml puromycin was reported for 0.6% and 1.6% of aerobically culturable bacteria from control and offal pit samples, respectively. Probing for puromycin and kanamycin (*nptII*) resistance genes by colony hybridisation involved a maximum of 512 bacteria from control sites and 648 bacteria from pit sites (Table 2).

The report indicated that the composition of the microbial community differed between control and offal pit sites, determined by DGGE (see Glossary). However, this variability was neither explained nor could it be distinguished from a simple anecdote because of the sampling scale and experimental design. There were no *confirmed* transfers of large fragments of DNA unambiguously derived from the dead animals or the microbes that inhabited the animals prior to death. The researchers noted that there were problems in efficiently recovering DNA for the molecular survey. Importantly, they also noted that the use of culturable bacteria introduced a significant bias and recommended a metagenomic approach in future experiments.

Researchers noted a high frequency of puromycin resistance genes, and a surprisingly high frequency of kanamycin resistance was seen from bacterial isolates originating in the offal pits, but there was no explanation for this unexpected phenotype and no follow-up on this observation. Since it is not uncommon for mobile genetic elements to collect multiple antibiotic resistance genes, it is possible that a selection for kanamycin resistance (for example, chemotherapy or use of a chemically similar prophylactic during animal lifetime) might co-select for puromycin through linkage (Heinemann 1999; Heinemann, Ankenbauer & Amabile-Cuevas 2000). *The high frequency of puromycin resistance in this set of experiments is noteworthy since this is the only time in the 2004-2009 time period that AgResearch sampled to depths where the soil could have been in contact with the carcasses.*

Confusingly, additional data appended to the end of the 2004 report (and titled "Offal pit summary for 2005") referred to probing of resistant colonies in the next year (2005). These data had no context in the report and it is not clear whether it revisited previous isolates from 2004 or new isolates obtained in 2005. Nearly 34% (176) of the 520 colonies described in this section of the 2004 report gave a qualified positive result when using the probe for the puromycin resistance gene. These positive detections were described as "faint". However, colony hybridisation is not quantitative and the intensity of the band was not shown to relate to the concentration of targets. Nearly 32% of these 176 "positive" colonies then produced some PCR product using the puromycin resistance gene specific primers, although none was of expected size (Figure 1).

While not discussed in the 'Offal pit summary for 2005' of the 2004 report, the data from this brief report were later referenced and interpreted in the "Background" section of the 2007 report (see below) in this way: "*PCR failed to amplify a segment of the puromycin gene from any of the bacteria.*" However, such a conclusion is not appropriately extrapolated from the data in the section called 'Offal pit summary for 2005'. No PCR amplicon was of anticipated size, but as none of the products were sequenced this result also does not show that the amplicons did not contain DNA of bovine origin. Nothing is known about the potential structure of a transgene taken up by a putative recombinant soil bacterium, and therefore whether the anticipated amplicon is the expected amplicon in actual field samples. Moreover, AgResearch did not report that the actual sequence of the target transgene in the GM bovine, which may have changed during transformation. Therefore, the procedures used for monitoring should have been able to detect significant sequence variants of the targeted genes rather than only the sequence originally used to transform the GM bovine (Figure 1).

Only 2% of 1130 colonies probed positive for the *nptII* gene and these were reported to be even fainter than for puromycin resistance gene target blots. These faint positives were dismissed as background binding to protein. Since *nptII* was not relevant to the AgResearch experiments (that is, *nptII* was apparently not part of the transgene construct used in the GM bovine), this result could be noteworthy. In effect, *nptII* represents a *de facto* negative control: what is the background of *nptII* in the offal pits when there has been no history of anthropogenic introduction of *nptII* genes? This makes the puromycin result potentially more interesting in contrast since that recombinant sequence has been introduced into the soil by human activities.

Several potential causes of negative results were then listed, including an undescribed potential problem with the probe. There was no subsequent discussion of any attempts to verify that the negative results were meaningful or caused by poor reagents, again leaving the possibility that true positives were present but not confirmed.

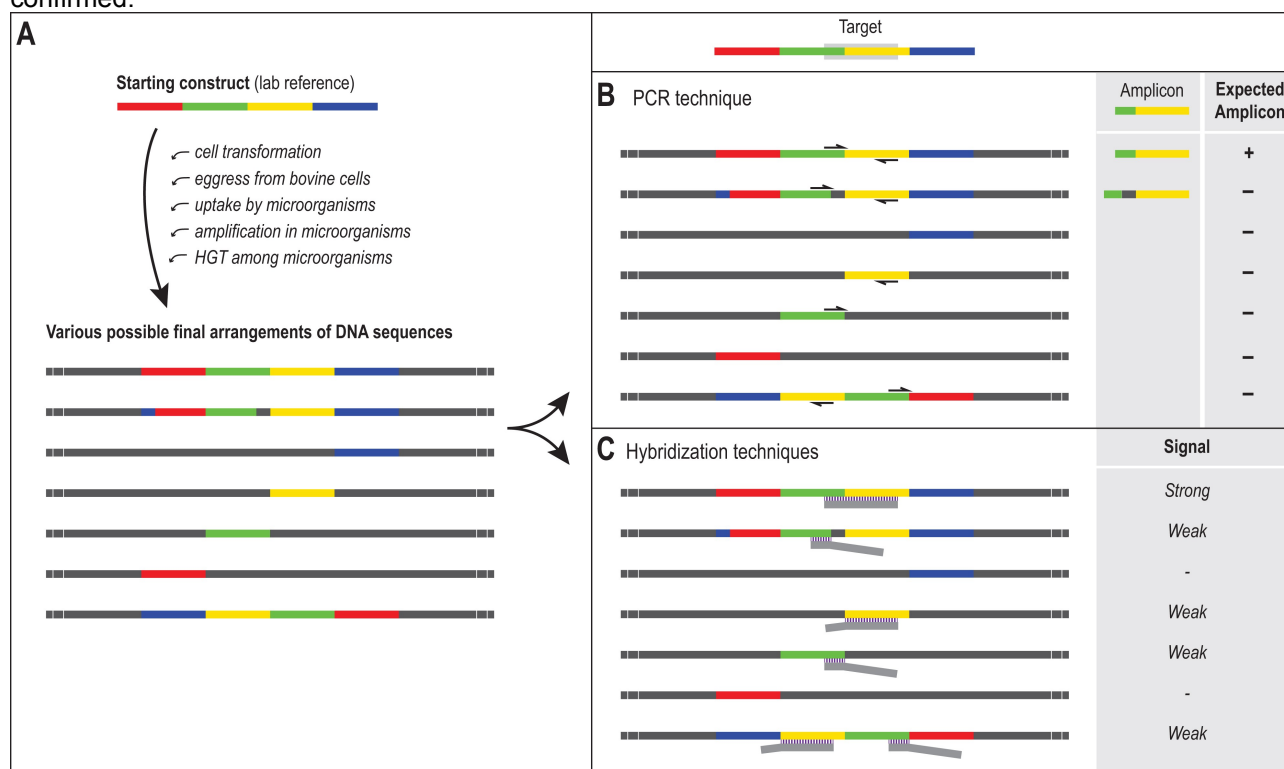


Figure 1. Two different techniques were used to detect target DNA sequences in soil bacteria.

The targets were parts of two different antibiotic resistance genes, one that confers resistance to kanamycin and the other to puromycin.

A (top) is a stylized version of all or part of the construct developed in the laboratory and then introduced into animal cells (middle), for amplification into adult animals. The final structure of the insert was not described in the reports from AgResearch, so there is no way to confirm that the DNA sequence in the animals was true to the laboratory reference material. Shown (bottom) are possible changes to parts of the sequence after the hypothetical events of DNA escaping from dead animal cells and into the soil and transfer to soil bacteria and or after multiple transfers among soil bacteria. The black lines represent animal or bacterial chromosomal DNA, and the red, blue, green and yellow are arbitrary to assist with visualization of possible changes in the DNA sequence relative to laboratory standard material.

B (top) illustrates the effects of potential sequence changes on use of the polymerase chain reaction (PCR) technique in detection experiments. The key reagent in a PCR is the primer (black half arrows), single-stranded short DNA strands that prime DNA synthesis (see Glossary). Two primers define the size of the amplified DNA (amplicon) that is copied enough times to be visible above the background of starting DNA material. The reaction is sensitive to the relative spacing between the primers and also the orientation of the primers, and is absolutely dependent upon both primers binding tightly to their target DNA sequence. The gray arrows (bottom) follow the different hypothetical sequence variants of the original target to the PCR and illustrate how false negative results can be obtained.

C illustrates the use of hybridization techniques for detecting putative targets. Two single-stranded DNA probes are made that match the same position on either of the two strands of target DNA. These probes find their target by forming hydrogen bonds with the target, in the same fashion that DNA duplexes are normally bound together. Small changes in number or order of nucleotides in the target region may reduce the strength of binding (leading to a weaker signal), but careful controls can be done to reduce the chances of false negatives. Only significant or complete loss of target will produce a false negative result. Therefore, the many weak signals generated by hybridization techniques may have been detecting target DNA that was unable to be confirmed by PCR, but which could have been confirmed by sequencing metagenomic DNA. Use of metagenomic DNA would also have extended AgResearch's ability to detect HGT beyond just culturable soil bacteria.

2007

Soil samples for this report were collected in December 2006. In this report, kanamycin and puromycin resistances were monitored, but kanamycin resistant isolates were not followed through to the hybridisation

analysis. Total calculated CFUs/g of soil from control samples ranged from 10^4 - 10^7 , and pit samples from 10^3 - 10^7 (Table 2). Probing for the puromycin resistance gene involved a maximum of 338 bacteria from control sites and 349 bacteria from pit sites.

Table 2. Calculated total scale of AgResearch survey to inform risk assessment and risk management of HGT.

Report	Number of bacteria screened by probe		Maximum equivalent of soil surveyed (g) ($\times 10^{-7}$)
	Control	Pit	
2004	512	648	6
2007	338	349	3
2009a	249	250	2
2009b	251	251	3
Cumulative total weight of soil surveyed 2004-2009			1.4 μ g

^A Based on estimated 2×10^9 bacteria per gram of soil. Sum of control and pit.

Using colony hybridisation to determine whether the displayed resistance phenotype was associated with the specific target DNA sequence, up to 41% of the CFUs first selected on puromycin were presenting as positive for the target puromycin resistance DNA sequence (range 18.2-40.6%) from control sites and 17-20% from pit sites. Up to 5% of the CFUs first selected on cycloheximide were presenting as positive for the puromycin resistance target sequence (range 0.7-4.5%) from control sites and about 3% from pit sites.

The highest proportion of CFUs with positive probe detections came from plates selecting puromycin resistance. The use of puromycin as a screening agent on the plates had a concentrating effect (up to 40% of the CFUs were positive for the target DNA sequence) compared to cycloheximide.

The positive detections of all putative resistance genes were largely dismissed as non-specific binding between the probe and *Pseudomonas* bacteria (because 4 of 85 (5%) chosen at random were *Pseudomonas*). AgResearch asserts that *Pseudomonas* commonly cause false positive detections. This could have easily been confirmed using other probes (and does not seem to have been). Even if *Pseudomonas* can non-specifically bind to probes, it might be that these *Pseudomonas* isolates were also in possession of the target DNA sequence.

AgResearch then used PCR to test for the presence of an anticipated amplicon that corresponded with the laboratory version of the puromycin resistance gene. It is not made clear how many of the 85 colonies from above were screened to yield 13 positives. On the benefit of the doubt, we assume that all were screened and that 15% (13 of 85) of those that returned a positive by hybridisation returned a positive by PCR.

An attempt was made to sequence the amplicons, but only from 3 of the 13 PCR positives (23%, representing 4% of those that probed positive). According to the report, no useful DNA sequence information was obtained. Only cloning vector DNA was detected, indicating that AgResearch either failed to clone the amplified target DNA or failed to isolate clones with the amplified target DNA. The implication was that the amplicons were artefacts, however that conclusion cannot be drawn from the data provided because the cause of the failure to generate confirmed sequence from the clones was not properly determined.

The size of the amplicons was not consistent with expectations based on the knowledge of the gene as preserved in the laboratory. But once again, there is no particular reason to expect that the gene will be preserved faithfully after transfer to soil bacteria (Figure 1). Moreover, there is no way to determine the number of horizontal transfers that these genes might have made between exiting the bovine and entering the bacteria ultimately isolated by AgResearch and thus no basis for extrapolating what the structure of the gene might be at the time of AgResearch's experiments.

The poor amplification from most of the 85 probe-positive colonies could be explained by non-specific inhibitors of the PCR (Heinemann, Sparrow & Traavik 2004). Simple and routine controls could have been used to exclude this possibility but no controls of the type are reported.

The final conclusion that: "there was no evidence of any puromycin resistance gene similar to the one used in the transgenic cattle in the offal or control soil bacteria, and no increase in puromycin resistant bacteria in offal pits compared to control soil" is in our opinion provisional and far overstates the power of the data. The conclusion implies that these experiments were competent to detect these genes had they been present. The "gene survey" described in the 2007 report was, in practical terms, neither suited to establishing the basis of this conclusion nor suitable for providing a scientifically valid reason to accept it. This report does not

describe the kinds of control soils and technique verification, sensitivity and pre-testing that would have been necessary to establish confidence in the conclusion.

2008

Soil samples described in this report were collected in December 2007. Total CFUs calculated were 10^5 - 10^6 per gram of soil for control and pit. Three of the 10 plates (listed in table 1 of the report) from the pits were plated at too small a dilution factor to get an accurate count. This potentially significantly underestimates the titre from 20% of the puromycin samples and only from the key sampling environment, the ofal pit. Thus it is not possible to conclude, as AgResearch did (p. 4), that there “*was no increase in puromycin resistance [sic] bacteria from ofal pit samples (ACU) compared to control paddock soil (CP)*”.

The data handling in this section is well outside accepted norms, both because of the failure to report raw numbers (which could then be subject to analysis by independent scientists) and because of the arbitrary way in which samples were excluded from analysis. First, the averages of CFUs from the pits with GM bovine appear to be miscalculated and in favour of understating the number of resistant bacteria. The averages of the ACU series appear to be divided by a factor of 5 (for the 5 plates counted), but in fact only have 4 and 3 entries in the series because they used the wrong dilution factor. We calculate the correct averages from this series and find that the actual CFUs in the ofal pits are 5.46 and 6.93 (bacteria per g (\log_{10})) instead of 5.36 and 6.71, respectively (Table 3). Second, plates ACU1 and ACU2 had too many bacteria applied to enumerate CFUs on cycloheximide medium. Plate ACU2 also had too many puromycin tolerant/resistant CFUs to count at the dilution used (table 1 of 2008 report). These samples should have been repeated using larger dilution factors. Too many to count could be 10x or 100x more than are possible to enumerate on a single plate. Substituting the conservative factor of 10x too many, then based on the single highest enumerated count (from ACU1), the average CFUs of bacteria per g (\log_{10}) jumps from 5.36 (2×10^5) to 6.28 (2×10^6), potentially significantly higher than the 5.70 in the controls (CP plates).

Table 3. Miscalculated CFUs in 2008 report

Sample ID	^A Numbers as reported	Numbers corrected	Numbers with conservative substitution for missing data	^A Numbers as reported	Numbers corrected	Numbers with conservative substitution for missing data
	Puromycin medium			Cycloheximide medium		
ACU1	5.92	5.92	5.92	OG	OG	8.0
ACU2	OG	OG	6.92	OG	OG	8.0
ACU3	4.98	4.98	4.98	6.89	6.89	6.89
ACU4	5.29	5.29	5.29	7.00	7.00	7.00
ACU5	4.41	4.41	4.41	6.89	6.89	6.89
Average	5.36 ^B	5.46	6.28	6.71	6.93	7.65

^A From table 1 (pg. 5) of 2008 report. Values are in bacteria per g (\log_{10}). OG = overgrown (or too many to count).

^B This is the result if the sum of the series were divided by 5 instead of 4. Dividing by 5 understates the average because there is no fifth value in the series.

Probing for the puromycin resistance gene involved a maximum of 249 bacteria from control sites and 250 bacteria from pit sites. Of these 499 colonies, 62 (12.4%) were presenting as positive by hybridisation for the respective target sequences. DNA from these 62 was then amplified by PCR, but no product was reported. A negative result from PCR can have many causes, only one of which is that the target sequence is truly not present (Heinemann, Sparrow & Traavik 2004). *No expected controls for the PCR are mentioned.* For example, when mixing DNA from the isolates and the laboratory positive control, was an amplicon made? If yes, then the PCR had no contaminants and was operating as expected. *Such controls should be standard operating procedure and should be described.* The appropriate follow-up would have been to isolate, clone and sequence the genomic DNA from a set of those presenting as positives by hybridisation.

The final conclusion that: “[t]here was no evidence of any puromycin resistance gene similar to the one used in the transgenic cattle in the ofal or control soil bacteria, and no increase in puromycin resistant bacteria in ofal pits compared to control soil” is an overstatement based on the data presented and, in our opinion, is not supported by the data in this report. As stated, in our opinion the conclusion is unsubstantiated.

2009a

Soil samples described in this report were collected in December 2008. Total CFUs calculated from control plates were approximately: 10^5 - 10^6 , pit: 10^6 - 10^7 per gram of soil. AgResearch claims that the much higher

number of puromycin tolerant/resistant bacteria for pit samples was not statistically significantly different. However, neither statistical analysis nor raw data was provided to allow an independent confirmation.

Probing for the puromycin resistance gene involved a maximum of 251 bacteria from control sites and the same number from pit sites. Of these, 20% (51) and 11% (28) of the CFUs from the control and pit sites, respectively, produced evidence of hybridisation with the puromycin gene-specific probe. Two of the colonies (one from a control and one from a pit site) returning a positive result for the puromycin resistance gene probe also produced an amplicon using gene specific PCR primers, but the amplicons were of different size and neither was of the size anticipated from the starting (laboratory) material.

Each of these amplicons was cloned and then sequenced. The sequences reportedly did not have “significant” matches to laboratory starting material⁶. Caution should be applied when attempting to use significant matches to confirm HGT. Recombination following gene transfer does not have to result in retention of tracts of nucleotides large enough to be judged to be quantitatively significantly different to be relevant (Heinemann 2000; Heinemann & Billington 2004; Heinemann & Traavik 2004). Significant changes can be evidence of HGT, but lack of significant changes is not evidence of the absence of HGT.

Determining these sequences was a considerable advance in the experimental approach over that described in the previous two reports, but still falls short of some obvious reporting standards. First, neither the original sequence used in the transgenic cows nor those determined by sequencing have been reported. This makes it impossible for others to determine “significance”. Second – and critically – AgResearch should report if the region amplified by the primers in the PCR was indeed the DNA that was hybridising to the probe (Figure 1). This could have been tested easily by Southern blotting of the two PCR fragments using the same probe as above. If we take into account that the structure of the putative recombinant DNA could easily be changed (over the number of transfers or number of years that would be required for a recombinant DNA or bacterium to travel to the position near the top of the pit being sampled), the probe may be a better indicator of a potential positive result than arbitrarily chosen primers. The only definitive experiment would have been to clone the genomic DNA that was being detected by the probe, and then sequence that clone.

The final conclusion that: “[t]here was no evidence of any puromycin resistance gene similar to the one used in the transgenic cattle in the offal or control soil bacteria, and no significant increase in puromycin resistant bacteria in offal pit soil compared to control soil” is an overstatement based on the data presented. Once again, the conclusion was an extrapolation beyond what we regard as proper scientific evidence. The point of the experiment was to establish the source of resistance in a bacterium (i.e., due to part or all of a transgene), not the overall frequency of resistance. The overall frequency might be an issue relevant to a hypothetical remediation plan⁷ (e.g., a consideration of whether additional numbers of resistant bacteria would matter), but is not sufficient to dismiss HGT as the cause of a resistance phenotype.

2009b

Soil samples described in this report were collected in November 2009. Total CFUs calculated from plates of control and pit samples was 10^6 - 10^7 per gram of soil (Table 2). No further data were provided.

Summary evaluation

Overall, AgResearch’s methodology was far too narrow to detect HGT at a sensitivity that would trigger a remediation plan for the purposes of preventing potential adverse effects. This is because the AgResearch approach was limited to detection of:

1. DNA transfer only;
2. DNA fragments of very large size only (e.g., a substantial portion of an entire gene);
3. DNA transfer events that are stably incorporated into organismal genomes but not mobile genetic elements that may not amplify in bacterial culture or in culturable bacteria;
4. phenotype changing HGT events; and
5. only expected phenotypes.

In choosing to address the risk of HGT by monitoring if HGT was occurring in the soils, AgResearch accepted for itself an unnecessarily difficult task because the amount of biological material that must be screened to detect HGT at relevant levels is enormous. Our previous calculations put the burden on them to sift through the genetic material of literally tons of soil (Heinemann & Traavik 2004), which they did not do

⁶ There were no described criteria for setting a threshold of significance.

⁷ As required under approval GMD002232 (ERMANZ, 2002).

(see below). This is effectively impossible to do, but because it is impossible to do is not an excuse for not properly monitoring HGT for the purposes of risk management when indeed there are better ways to do it (see below).

Was AgResearch monitoring consistent with the approval requirements?

Recall from above that there were two key approval requirements to manage the risks of HGT at offal sites. The first was to limit the range of DNA sequences that could be used in the making of GM bovine and the second was to monitor soil microorganisms for HGT.

We have found that AgResearch only monitored a very small subset of soil microorganisms. It monitored only bacteria and not fungi or other microorganisms, and only aerobically culturable bacteria. The Authority explicitly referred to other kinds of microorganisms in its decision: "*It is now widely accepted that HGT occurs, most significantly within the prokaryote kingdom, but instances of nucleic acid transfer between prokaryotes and eukaryotes are documented in the scientific literature*" (p. 24 ERMANZ, 2002). By design of experiments, AgResearch ignored up to an estimated 99.9% of bacteria and all other kinds of microbes.

Most disconcerting is the possible disconnect between AgResearch's understanding of the purpose of monitoring and the intent of the Authority in issuing the control. The prominence of HGT in the risk assessment leaves us in no doubt that the Authority intended HGT risks to be taken seriously by AgResearch and that the monitoring conducted was to be for risk mitigation (p. 47 ERMANZ, 2002). In its reply to INBI's 23 November draft report (provided to AgResearch for comment), AgResearch expressed what appears to us to be a different intention when it says: "*We have not undertaken risk related research... ERMA has assessed the risk of HGT to be at very low or negligible levels and required a monitoring programme. We have undertaken an annual monitoring activity*", referring to page 19 of the ERMANZ decision on the GM bovine application 02028 (ERMANZ 2002). It goes on to say: "*This was a monitoring study intended to reveal the occurrence of significant (readily detectable) events.*"

There are a number of differences between what the Authority specified and what AgResearch appeared to have done. First, there is no reason to believe that the Authority considered "significance" as purely a quantitative increase in the population of antibiotic resistant bacteria. The uncertainty of the risks were both qualitative (as in the creation of a new trait which would cause significant adverse effects) and extended beyond just counting the number of antibiotic resistant bacteria.

"The Committee was initially limited in its ability to fully assess risks because of the lack of specificity of a number of the elements that could be involved in the genetic modification process, and lack of evidence regarding absence or presence of potential adverse effects. This constraint has been overcome to a satisfactory extent by restricting the scope of the approved organism, requiring monitoring of micro-organisms, and limiting the period of approval to 7½ years" (p. 48 ERMANZ, 2002, emphasis added).

If AgResearch thought from the outset that the risk was negligible regardless of whether or not it was to monitor for HGT, then this could explain why it designed experiments for only "*readily detectable*" events, moreover only those that would appear after short periods of time (see below), only in bacteria and only in bacteria that were culturable under aerobic conditions⁸. It is our opinion that AgResearch has raised considerable doubt over whether the experiments it conducted were intended to, and therefore could, deliver either risk mitigation or inform the Authority about the potential for adverse effects arising through HGT in New Zealand soils as was required⁹: *In our view, the monitoring effort was far too superficial to address the Authority's concerns.*

Sampling depth

AgResearch reported constructing 27 offal pits (as called in the reports listed above, or "holes" as described in subsequent information, see Supplementary Material) that were 7 m deep (Figure 2). ERMANZ imposed a requirement (control 1.4 of ERMANZ 2002) that the offal pits conform to MAF Regulatory Standard

⁸ In addition, page 19 of the decision is only about HGT from milk disposed of on site, not in the offal pits containing dead GM bovine and therefore does not relate specifically to the ERMANZ assessment of risk at offal pit sites or its requirement to monitor offal pits.

⁹ We consider it noteworthy that the results of the monitoring may have formed part of the portfolio of evidence that ERMANZ has relied upon to update its thinking on HGT (see <http://www.ermanz.govt.nz/BertDocs/Supporting%20Documents%20for%20EandR.pdf>). If ERMANZ has, then it becomes critically important that the experiments conducted by AgResearch were intended to inform ERMANZ that monitoring was no longer required for risk mitigation.

154.03.06 (MAF 1999). This standard requires that the carcasses remain a minimum of 2 m below ground level¹⁰. Therefore, it appears that on only one occasion (2004 report) has AgResearch attempted to sample for antibiotic resistant bacteria, and by extension for HGT, in soil that might have been in contact with GM bovine (Figure 2A). AgResearch further reported, that as pits subsided with time, additional soil was added to the top (Supplementary Material). The research reports failed to note or exclude that the testing might have been further removed from relevant soil by the addition of fresh soil in particular pits (Figure 2B).

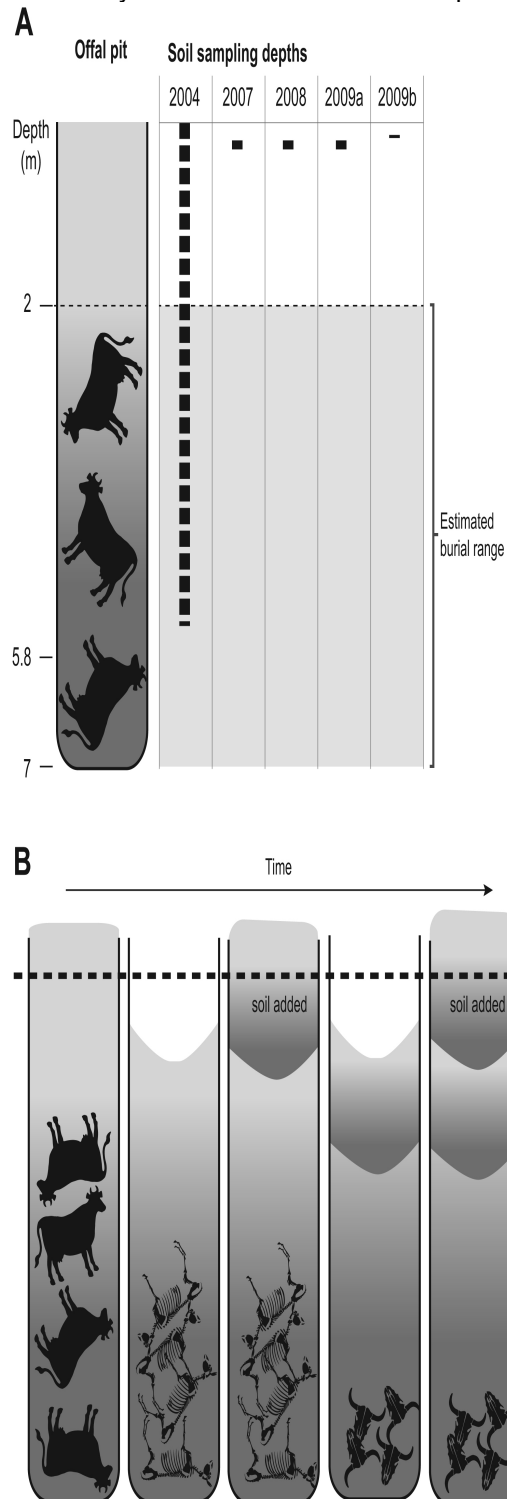


Figure 2. Sampling depth and effect on experimental findings.

(A) AgResearch sampled soil from offal pits (cylinders in figure) at varying distances from the surface (depth, in metres). The reports did not specify the how close the carcasses came to the surface in the actual pits sampled. The carcasses may have been between 5.8 and 2 m from the surface (they had to be a minimum of 2 m deep to comply with ERMENZ control 1.4). There was no indication of whether any

¹⁰ "...carcasses and tissues may be buried on-site to a minimum depth of two meters provided that conditions are favourable (an assessment shall take into account drainage, soil type and land use)."

soil sampled was in contact with the carcasses, but it is possible that it was for samples taken in 2004. Depth and year at which samples were taken are shown as black bars.

(B) Soil (grey) subsidence in the pits over time was compensated through the addition of fresh soil (dark grey). The reports made no mention of whether soil was added to pits prior to sampling. Since subsidence takes time, samples taken before the addition of fresh soil to the pits were in most years both likely to have been well above the interface with carcasses (which were about a minimum of 1.7 m lower than sample depth) and to have provided too little time for the appearance of a target population of detectable size. Samples taken after the pits were topped up would have been in fresh soil never in contact with the carcasses.

AgResearch reports a variety of sampling depths. But only in the 2004 report was sampling beyond 30 cm, and in the 2009a report, sampling was to the depth of only 15 cm (Figure 2A). It is difficult to believe that there would have been a physical problem to prevent access to layers deeper than one or two shovel blades. In the single case where sampling was deeper (data in 2004 report), the age of the sampled pit(s) was at most just over a year and could have been as little as weeks or months. This could be a very short time with regard to detecting a population of bacteria displaying a new trait. Pre-testing in relevant conditions would have been necessary to establish appropriate timeline data. In any case, the majority of carcasses were below the layer of soil extracted.

Therefore, for all but one sampling exercise, the soil was a minimum of 1.7 m from contact with the microbes that would have been in contact with carcasses (Table 1). Because of compensation for subsidence and because AgResearch may have buried the carcasses much deeper, the samples could have been up to 5.8 m from the soil layer in first contact with carcasses¹¹. If the average soil bacterium is about 1 µm (1 one-millionth of a metre) in diameter, within the time between filling the pit and sampling, a recombinant bacterium would have to migrate a minimum distance 1.7 million times its size in a dry pit, or the gene would have to transfer a minimum of 1.7 million times, against gravity directly toward the surface in order to have the potential to be detected.

We can only conclude that the soils examined by AgResearch were irrelevant to the task of monitoring HGT because they were not in contact with the carcasses, and were so far away on the scale of microorganisms as to have been equivalent to sampling in an arbitrary and distant locations.

AgResearch repeatedly claimed that taking deeper samples was unjustified based on earlier findings of a lack of change in resistance phenotypes as a function of sample depth. For example, the 2007 report justifies a depth of 30 cm based on the results of sampling described in 2004. The 2008 report repeats this claim, this time referring to the 2007 data. The 2009a report justifies the sampling depth based on a non-existent finding from the 2008 report that there was no relationship between sampling depth and resistance. These claims are erroneous. The only time AgResearch reported measuring this relationship was in 2004. By repeating the claim that there was no relationship between resistance and depth in multiple years, AgResearch potentially creates the false impression that the 2004 results were replicated when there is no indication that they were.

The strength of the data in the 2004 report for establishing the validity of the sampling depth is questionable. First, the 2004 data was based on a “*preliminary analysis*” (AgResearch words). Second, it described only a small number of samples and not the same combination of antibiotics used in later studies. Third, AgResearch reported higher numbers of bacteria at all depths for offal pit samples than control samples, and lower numbers of bacteria at increasing depth in controls. Thus, the indications are that parameters relevant to HGT may change with depth. Fourth, these data were taken at the earliest time point of the experiment and may not have been representative of the time necessary for a recombinant bacterium to establish a population of a detectable size. In any case, it is not clear what the relevance of bacterial resistance frequency per vertical metre is for determining sampling depth since the target is putative recombinant bacteria, not estimates of numbers of bacteria not susceptible to the antibiotics being used.

Inconsistent sampling and plating

The use of different antibiotics in the different reports creates the impression that the experiment series was not carefully planned from the beginning and this creates unnecessary questions about the purpose and intent of the project. For example, only in 2004-5 did AgResearch use probes (Figure 1) for the *nptII* kanamycin resistance gene. In the 2007 report (and in subsequent communications, see Supplementary Materials), AgResearch said that *nptII* was not relevant to the transgenes used in the GM bovine experiments and monitoring for it only added costs. This explanation leaves unanswered why then

¹¹ That was the depth we suggested in the draft report provided to AgResearch for factual review and it was not contested.

kanamycin resistance was monitored on plates in both 2007 and 2008. This odd pattern might be explained as historical artefacts in the timeline of protocol development, or as aborted attempts to enrich the baseline descriptions of the sampled soils. However, if the inconsistencies were due to protocol development, why repeat irrelevant measurements over multiple years? The failure to follow up on kanamycin resistance phenotypes and the limited scale of the research raises doubts about whether these studies were meant to provide usable baseline information.

Sampling power was too small

The total survey in each year amounted to a tiny fraction of the soil bacteria. AgResearch failed to answer when asked how it determined that its sampling power was adequate (Supplementary Materials). We can only surmise that no effort was made to perform a power analysis.

When you take into account that there are on the order of 2×10^9 (or 2 billion) bacteria per gram of soil (Heinemann & Traavik 2004), then the total cumulative scale of the multiyear AgResearch survey sums to the number of bacteria (total) in an estimated 1.4 μg of soil (Table 2).

To put this in context, about 4 million bacteria (1 μm in diameter) could fit on the head of a pin (of 2 mm diameter). The number of bacteria subjected to hybridisation analysis by AgResearch from 2004 (~3000) is 0.08% of those that would fit on the head of a pin. In contrast, the volume of soil in just one pit (offal pit diameter 0.9 m and depth 7 m, Supplementary Materials) is 4.45 m^3 . This equates to a mass of ~6 tons of soil with an estimated 1×10^{16} (10,000,000,000,000,000) potentially exposed bacteria¹².

Experiments were not appropriate for the problem

As stated repeatedly in the AgResearch reports, the purpose of the monitoring was to detect any occurrence of transgenes transferring from the GM animals to soil bacteria. For example: “*There is some concern regarding the possibility of horizontal transfer of puromycin resistance encoding gene from transgenic cattle carcasses buried at Ruakura to soil bacteria*” (p. 6-7 2008 report). Plating and hybridisation experiments alone are not sufficient to determine if the source of the detected puromycin resistance genes was indeed the GM bovine or if the occurrence of resistant bacteria was due to other unrelated factors.

The task was not to determine if the frequency of resistance phenotypes was overall higher or lower in the offal pits, or whether a phenotypic change would result from the transfer. These latter measurements might help to raise confidence in the findings or even focus the search, but they do not in themselves deliver the answer. Even if the frequency of puromycin resistance was higher in control sites than in offal pits that observation would not be sufficient to dismiss horizontal transgene transfer as the cause of some phenotypic resistance in the offal pits.

Indeed, it is difficult to see how the control sites differed from the sampled sites in any of the experiments because samples were taken so far from the majority of the biological material that was the source of the transgene. This further exposes the lack of analysis determining what sites would make suitable controls. The selection of control sites based on the criteria of “similar soil type” and similar use (Supplementary Material) is both vague and not obviously related to how these soils serve as controls. First, how similar were the soils? Considering that in the 2004 report, soil at depth (nearer the actual interface with the carcasses) was seen to be compositionally different to soil at surface, this criterion does not seem to have been met. Second, use as a paddock for grazing is hardly a use similar to a grave. Instead, controls should have been samples from equivalent offal pits filled with non-GM animal carcasses and matched for depth and soil composition.

Another dimension that is critical to the estimation of risk is time. In this case, the important time period is between the creation of a recombinant bacterium and the appearance of either a population of recombinant bacteria large enough to detect or large enough to cause an adverse effect. This is the lag time (see figure 2

¹² In response to this section of the 23 November advanced draft of this report, AgResearch said about their work that it only “was a monitoring study *intended* to reveal the occurrence of significant (readily detectable) events” (emphasis added). Although the Authority has operated on the standard of “reasonableness” in setting its requirements, we could not find verification that the controls were *intended* to be set at the limit of “readily detectable”. Indeed, in recalling the wording of the approval, we find that the Authority wanted monitoring to be “as extensive as possible” (p. 21 ERMENZ, 2002), and *any detection* (p. 25 ERMENZ, 2002), not just HGT at significant readily detectable levels, would be considered important enough to suspend approval for further work until an acceptable remediation plan was activated. In our view, “readily detectable” is a much lower standard than “as extensive as possible” (within reason).

of Heinemann & Traavik 2004). There is no evidence that AgResearch even considered the time dimension, much less discussed how it might affect their experiments.

The natural production of other antibiotics in some soils might decrease the lag time of a burgeoning recombinant population through co-selection. We assume that the AgResearch puromycin resistance gene encodes the N-acetyl-transferase enzyme. It would be premature to assume that this activity had no unknown selective advantage in some environments. There are many reasons why a transgene could be selected in the environment even if the environment did not have the *expected* selective force (for a review see Heinemann, Ankenbauer & Amábile-Cuevas 2000). For example, high frequencies of ampicillin resistance are found in the bacteria of the sediment of the Atlantic Ocean off the coast of New York not because the Atlantic Ocean has clinically relevant concentrations of the antibiotic, but because it is co-selected with heavy metal resistance. Resistance genes also can contribute to bacterial fitness even in the absence of antibiotics (for a review see Heinemann, Ankenbauer & Amábile-Cuevas 2000).

Setting aside the fundamental failures to scope the experimental parameters and perform preliminary power experiments to evaluate the usefulness of the chosen monitoring scale, how well did AgResearch do on the basic molecular analysis? Unfortunately, here there were also problems. The molecular series was composed of largely unfinished experiments. Negative results were left unchallenged when simple and routine controls could have added confidence to the findings. A serious disappointment was to leave description of potential recombinants to hybridisation and PCR amplification, with confirmed sequencing of only two PCR amplicons.

The necessary and missing experiment was to clone a statistically meaningful number of genome fragments from bacteria known to be independent of one another and that probed positive for the target DNA. Each of these clones should have been sequenced and subjected to careful comparisons with the starting transgene (e.g. taking into account events as described in figure 1 of Heinemann & Traavik 2004).

The impression we have from careful reading of an experiment series spanning 6 years, using significant public resources and unhindered by any detectable oversight from external and independent review, is that the experiments appear to be designed and executed well short of the Authority's wish for monitoring to be "*as extensive as possible*" (p. 21 ERMENZ, 2002).

Conclusions not justified by the data

The recurrent conclusion of the reports is that the experiments have produced "*no evidence*" of HGT. Given the context, that these experiments were performed for the purpose of increasing our knowledge of the risks and safety of transgenes used in GM animals¹³, this claim comes with the responsibility of demonstrating that the negative result (that is, the absence of detectable gene transfer) has scientific meaning. It is difficult for us to imagine scientists designing experiments less likely to find a target recombinant bacterium even if it did exist. The small scale of the survey, the distance from the transgene source from which the samples were taken, and the uncritical approach to the negative results make the conclusion drawn unsupported from the data¹⁴.

With the exception of the 2004 report, a final critical omission was the failure to discuss the limits of the detection experiments. AgResearch reports only emphasised the number of surveyed bacteria. We found this emphasis conveniently distracting for the uncritical reader who might not notice the absence of any effort to put these sample numbers into the context of the microbial population size, or consider the kinds of microbes (e.g. unculturable bacteria and eukaryotic microbes) that were not part of the monitoring.

¹³ "As well as providing an assurance on the effectiveness of controls, the information is potentially valuable for future applications" (p. 21 ERMENZ 2002). See also pp. 21-22 of the ERMENZ Evaluation and Review Report on application GMD02028: "It is the stated purpose of the Authority to examine, with every GMO application, but especially outdoor developments or field trial, the practicality of monitoring for possible effects or other factors. Information of this type is essential in providing a platform for examining future approvals (especially releases) and for providing a positive assurance of the efficacy of containment and other controls."

Contained outdoor trials of GMOs provide opportunities for research to help address some of the uncertainties discussed above. The applicant has already indicated that they will be investigating HGT around GM cattle burial sites" (http://www.ermanz.govt.nz/BertDocs/GMD02028_GMD02028%20EnR%20final.pdf).

¹⁴ However, this does not imply that there is evidence of HGT, only that it is incorrect to suggest that these experiments were qualified to address the scientific problem.

Could AgResearch have done better?

Having chosen this particular experimental approach to meet its regulatory obligations, could AgResearch nevertheless have done better? Setting aside for a moment that their approach was arguably flawed, AgResearch was working in a very difficult area and certainly was not alone in choosing this experimental design. However, prior to the launch of these experiments was the appearance of literature that already was describing the flaws of this approach (e.g. Heinemann 2003; Heinemann & Traavik 2004). There is no evidence in the reports listed above that this or similar literature was even considered as late as December 2009.

Had the AgResearch experimental approach been the first of its kind, then perhaps this approach, flaws and all, would have been a state of the art effort. It was not, and was even less innovative than work conducted in the 1990s (e.g., Gebhard & Smalla 1998; Nielsen, van Elas & Smalla 2000). Benefiting from some 20 years of environmental monitoring and access to the peer-review standard of such work, we believe that AgResearch could have conducted better experiments even in a flawed design and better still worked to reduce the flaws in the design.

Notably, more work mapping the possible risk pathways of HGT could have inspired a more diverse and sophisticated approach that was not necessarily more costly considering that it would more likely have produced publishable and confidence-building outcomes.

The AgResearch experiments were conducted without any known selective force *in situ* to help amplify the population of any putative recombinant bacteria. Incorporating selection in a preliminary microcosm experiment would have provided more useful information about how large the sample sizes had to be from the pits in order to detect any target population. A series of microcosms could have been set up with known and varying numbers of input recombinant bacteria without selection (to determine if populations would amplify in the absence of any *known* selection for the puromycin resistance gene), and then with varying amounts of puromycin to accelerate the formation of a detectable number of recombinants. Depending on the results of these range-finding experiments, microcosms could be set up using a known number of GM animal cell inputs and replicates with varying amounts of puromycin to accelerate formation of any potential population of recombinant bacteria.

Field or microcosm work could also have played a role in estimating the statistical power needed for the final monitoring effort. Finally, a metagenomic approach is necessary (but may not be sufficient) to reveal what is happening in soil because only a very small proportion of soil bacteria are culturable, and AgResearch effectively only followed a small number of this small proportion, and ignored all other kinds of soil microbes (for example the eukaryotic microbes) that may be as or even more important in this instance.

The descriptive work attempted by AgResearch between 2004 and late 2009 was conducted at a scale that renders it effectively anecdotal for providing environmental baseline data and irrelevant for evaluating risk. Still, relatively small changes to the approach, significant improvement in the follow through on normal experimental controls and adherence to normal standards of data reporting and analysis would have made this effort appear far more professional and would have instilled more confidence that the risks related to HGT, which the Authority indicated might be important, had been addressed.

Final word

AgResearch is a biotechnology provider that has indicated its hope to deliver new or less costly therapeutics through its GM bovine programme, and so provide public health benefits¹⁵. That programme is regulated by ERMENZ under the Hazardous Substances and New Organisms Act of New Zealand, legislation which can foster biotechnology research and development by ensuring that it is conducted safely¹⁶ and retains the confidence of New Zealanders. The lure of new technological solutions to address pressing human needs can be hard to resist. However, we believe that it is important to emphasise that the envisaged potential of these technologies should not tempt us to unduly put our environment at risk. This is because many technologies at the forefront of promise eventually fail or develop more slowly than expected, and because tradeoffs need to be carefully assessed once the potential of a development is well understood. All biotechnology companies have an obligation in the short term to ensure that they do not undermine public trust in an industry that may one day provide safe and effective GM-based therapeutics or other products in the long term.

¹⁵ <http://www.agresearch.co.nz/transgenic/transgenic-letter.asp>

¹⁶ "All persons exercising functions, powers, and duties under this Act...shall take into account the need for caution in managing adverse effects where there is scientific and technical uncertainty about those effects" (s7 of the Hazardous Substances and New Organisms Act 1996).

Acknowledgements

We thank Joanna Goven and Simon Terry for helpful comments and Matt Walters for assistance with graphics. We acknowledge the in-kind support from the University of Canterbury and GenØk – Norwegian Centre for Biosafety for this critic and conscience of society activity. The authors are grateful to both AgResearch and GE Free New Zealand for their cooperation in this science shop activity of the Centre for Integrated Research in Biosafety.

Bibliography

- Al-Kaff, N.S., Kreike, M.M., Covey, S.N., Pitcher, R., Page, A.M. & Dale, P.J. 2000, 'Plants rendered herbicide-susceptible by cauliflower mosaic virus-elicited suppression of a 35S promoter-regulated transgene', *Nature Biotechnology*, vol. 18, no. 9, pp. 995-999.
- Amábile-Cuevas, C.F., Cardenas-Garcia, M. & Ludgar, M. 1995, 'Antibiotic resistance', *American Scientist*, vol. 83, pp. 320-329.
- Anonymous 1995, 'Report of the ASM task Force on Antibiotic Resistance', *Antimicrobial Agents and Chemotherapy*, vol. 39, no. Supplement, pp. 2-23.
- Avery, O.T., MacLeod, C.M. & McCarty, M. 1944, 'Studies on the chemical nature of the substance inducing transformation of Pneumococcal types', *Journal Experimental Medicine*, vol. 79, pp. 137-158.
- Cavalli-Sforza, L.L. 1957, 'Bacterial genetics', *Annual Review of Microbiology*, vol. 11, pp. 391-418.
- Conner, A.J., Glare, T.R. & Nap, J.-P. 2003, 'The release of genetically modified crops into the environment Part II. Overview of ecological risk assessment', *The Plant Journal*, vol. 33, pp. 19-46.
- Daniel, R. 2004, 'The soil metagenome – a rich resource for the discovery of novel natural products', *Current Opinion in Biotechnology*, vol. 15, pp. 199-204.
- de la Cruz, F. & Davies, J. 2000, 'Horizontal gene transfer and the origin of species: lessons from bacteria', *Trends in Microbiology*, vol. 8, pp. 128-133.
- ERMANZ 2002, 'Environmental Risk Management Authority Decision GMD002232', Wellington.
- ERMANZ 2006, *Risk assessment of horizontal gene transfer from GM plants to bacteria and human cells*, ERMA New Zealand, ER-GI-05-1 07/06.
- Gebhard, R. & Smalla, K. 1998, 'Transformation of Acinetobacter sp. strain BD413 by transgenic sugar beet DNA', *Applied and Environmental Microbiology*, vol. 64, pp. 1550-1554.
- Gibbs, M.J. & Weiller, G.F. 1999, 'Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus', *Proceedings National Academy Science USA*, vol. 96, pp. 8022-8027.
- Goven, J., Hunt, L., Shamy, D. & Heinemann, J.A. 2008, *Animal biopharming in New Zealand. Drivers, scenarios and practical implications*, Constructive Conversations, University of Canterbury, 12.
- Griffith, F. 1928, 'The significance of Pneumococci types', *Journal Hygiene*, vol. 27, pp. 113-159.
- Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M. & Michel, G. 2010, 'Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota', *Nature*, vol. 464, pp. 908-912.
- Heinemann, J.A. 1991, 'Genetics of gene transfer between species', *Trends in Genetics*, vol. 7, pp. 181-185.
- Heinemann, J.A. 1997, *Assessing the risk of interkingdom DNA Transfer*, Norwegian Biotechnology Advisory Board, Oslo.
- Heinemann, J.A. 1999, 'How antibiotics cause antibiotic resistance', *Drug Discovery Today*, vol. 4, pp. 72-79.
- Heinemann, J.A. 2000, 'Horizontal transfer of genes between microorganisms', in J Lederberg (ed.), *Encyclopedia of Microbiology*, 2 edn, Academic Press, San Diego, CA, pp. 698-706.
- Heinemann, J.A. 2003, 'Is horizontal gene transfer the Cinderella of genetics?', *New Zealand Bioscience*, vol. 12, pp. 51-54.
- Heinemann, J.A. 2004, 'Challenges to regulating the industrial gene: Views inspired by the New Zealand experience.', in K Dew & R Fitzgerald (eds.), *Challenging Science: Science and Society Issues in New Zealand*, Dunmore, Dunedin.
- Heinemann, J.A. 2009, *Hope not Hype. The future of agriculture guided by the International Assessment of Agricultural Knowledge, Science and Technology for Development*, Third World Network, Penang.
- Heinemann, J.A., Ankenbauer, R.G. & Amábile-Cuevas, C.F. 2000, 'Do antibiotics maintain antibiotic resistance?', *Drug Discovery Today*, vol. 5, pp. 195-204.
- Heinemann, J.A. & Billington, C. 2004, 'How do genomes emerge from genes?', *ASM News*, vol. 70, pp. 464-471.
- Heinemann, J.A. & Kurenbach, B. 2009, 'Horizontal Gene Transfer', in M Schaechter (ed.), *Encyclopedia of Microbiology*, 3 edn, Elsevier, San Diego, pp. 1-10.

- Heinemann, J.A., Sparrow, A.D. & Traavik, T. 2004, 'Is confidence in monitoring of GE foods justified?', *Trends in Biotechnology*, vol. 22, pp. 331-336.
- Heinemann, J.A. & Traavik, T. 2004, 'Problems in monitoring horizontal gene transfer in field trials of transgenic plants', *Nature Biotechnology*, vol. 22, pp. 1105-1109.
- Hershey, A.D. & Chase, M. 1952, 'Independent functions of viral protein and nucleic acid in growth of bacteriophage', *Journal General Physiology*, vol. 36, pp. 39-56.
- Kidwell, M.G. 1993, 'Lateral transfer in natural populations of eukaryotes', *Annual Review Genetics*, vol. 27, pp. 235-256.
- Knight, C.J., Bailey, A.M. & Foster, G.D. 2010, 'Investigating Agrobacterium-mediated transformation of *Verticillium albo-atrum* on plant surfaces', *PLoS ONE*, vol. 5, p. e13684.
- Lederberg, J. & Tatum, E.L. 1946, 'Gene recombination in *Escherichia coli*', *Nature*, vol. 158, p. 558.
- Levy, S.B. 1998, 'The challenge of antibiotic resistance', *Scientific American*, vol. 278, no. (March), pp. 32-39.
- MAF 1999, 'Containment standard for field testing of farm animals', in 154.03.06,
- Moran, N.A. & Jarvik, T. 2010, 'Lateral transfer of genes from fungi underlies carotenoid production in aphids', *Science*, vol. 328, pp. 624-627.
- Nielsen, K.M., van Elsas, J.D. & Smalla, K. 2000, 'Transformation of *Acinetobacter* sp. strain BD413(pFG4DnptII) with transgenic plant DNA in soil microcosms and effects of kanamycin on selection of transformants', *Applied and Environmental Microbiology*, vol. 66, pp. 1237-1242.
- Ochman, H., Lawrence, J.G. & Groisman, E.A. 2000, 'Lateral gene transfer and the nature of bacterial innovation', *Nature*, vol. 405, pp. 299-304.
- Stroun, M., Anker, P. & Auderset, G. 1970, 'Natural release of nucleic acids from bacteria into plant cells', *Nature*, vol. 227, pp. 607-608.
- Syvänen, M. & Kado, C.I. (eds.) 1998, *Horizontal Gene Transfer*, 1 edn, Chapman and Hall, New York.
- Syvänen, M. & Kado, C.I. (eds.) 2002, *Horizontal Gene Transfer*, 2 edn, Academic Press, San Diego.

Glossary

amplicon The accumulated product amplified by a polymerase chain reaction (PCR). Will be of a defined number of nucleotides.

colony forming units (CFU) The number of colonies that appear on the surface of media is taken as an approximation of the number of culturable bacteria first applied to the media. Provided that the bacteria were well separated, then each subsequent "colony" formed by their reproduction should be clones of the original bacterium.

cycloheximide antibiotic (against fungi)

denaturing gradient gel electrophoresis (DGGE) The process whereby a small sample of DNA (or RNA) is applied to an electrophoresis gel containing a denaturing agent. It is used to discern differences in DNA sequences and identify particular gene mutations.

g⁻¹ per gram

genome A collection of genetic material and genes that normally reproduce in synchrony. An organismal genome is dependant upon successful reproduction of the entire organism for its reproduction.

horizontal gene transfer (HGT) The reproduction of genetic material uncoupled from organismal reproduction.

hybridisation

m metre

µg microgram is 1 one-millionth of a gram (10⁻⁶ grams)

mg milligram is 1 one-thousandth of a gram (10⁻³ grams)

µm micrometre is 1 one-millionth of a metre (10⁻⁶ metres)

kanamycin antibiotic

plating Samples of bacteria are said to be "plated" when they are applied to solid media, usually in a dish called a Petri plate.

polymerase chain reaction (PCR) The exponential amplification of DNA fragments through thermal cycling, whereby repeated heating and cooling of the reaction denatures the double-stranded DNA and allows for enzymatic replication and annealing.

primer A strand of nucleic acid that serves as a starting point for DNA synthesis.

probe A fragment of DNA or RNA of variable length (often 100-1000 bases long), used to detect the presence of nucleotide sequences (the DNA target) complementary to the sequence in the probe. The probe hybridizes to single-stranded nucleic acid whose base sequence allows probe-target base pairing due to complementarity between the probe and target.

puromycin antibiotic

transformation Any change in phenotype that might result from a change in genotype or differentiation. Most commonly used in bacterial genetics to refer to a change in genotype due to uptake of DNA.

vertical reproduction The concomitant reproduction of genetic material and cell/organism.