

Monitoring and modeling horizontal gene transfer

Kaare M Nielsen^{1,2} & Jeffrey P Townsend³

Monitoring efforts have failed to identify horizontal gene transfer (HGT) events occurring from transgenic plants into bacterial communities in soil or intestinal environments. The lack of such observations is frequently cited in biosafety literature and by regulatory risk assessment. Our analysis of the sensitivity of current monitoring efforts shows that studies to date have examined potential HGT events occurring in less than 2 g of sample material, when combined. Moreover, a population genetic model predicts that rare bacterial transformants acquiring transgenes require years of growth to out-compete wild-type bacteria. Time of sampling is therefore crucial to the useful implementation of monitoring. A population genetic approach is advocated for elucidating the necessary sample sizes and times of sampling for monitoring HGT into large bacterial populations. Major changes in current monitoring approaches are needed, including explicit consideration of the population size of exposed bacteria, the bacterial generation time, the strength of selection acting on the transgene-carrying bacteria, and the sample size necessary to verify or falsify the HGT hypotheses tested.

Whereas most HGT events taking place in bacterial communities are deleterious, rare positively selected events can enable bacteria to adapt rapidly and efficiently to changing environments. The societal impact of HGT events is exemplified by instances of bacterial acquisition of virulence traits and antimicrobial drug resistance^{1,2}. Based on evidence that transgenes can persist in soil for years after harvest and in mouse feces after digestion³, HGT of transgenes into bacteria has been identified as a potential risk associated with the use of transgenic plants^{4,5}. The process of HGT from transgenic plants into bacteria has been studied both under defined experimental conditions and through screening of samples from sites where transgenic plants have been introduced or their products consumed³. Several experimental laboratory studies have demonstrated that some bacterial species can take up fragments of plant transgenes under highly optimized conditions^{6–9}. However, agricultural field trials and human feeding trials have not yielded conclusive evidence for HGT from transgenic plants to bacteria under natural conditions^{10,11}.

Here, we assess the feasibility of monitoring HGT processes through field or gut sampling of transgenic plant-exposed bacteria. How likely

are these studies to detect HGT if it occurs? As shown below, it is enormously difficult to detect rare HGT events and the resulting bacterial transformants within large heterogeneous microbial communities. The large number of bacteria present in most natural environments (e.g., leaf material contains $\sim 10^6$ bacteria per gram, and soil and intestine $\sim 10^9$ bacteria per gram) quickly saturates sampling efforts and, hence, obscures conclusions that can be drawn from current monitoring designs. Moreover, the majority of bacteria in soil and intestine are nonculturable, thereby severely limiting the possibility of unambiguously identifying transgene-carrying bacteria and characterizing the HGT event at the molecular level. To illustrate this, we examine experimental designs from published studies, using the two most comprehensive studies as examples, demonstrating that they fail to resolve the issue of whether recombinant genes present in transgenic plants can spread horizontally to exposed microbial communities in agricultural soil¹⁰ or in the intestine¹¹. To enhance future efforts, we develop monitoring strategies based on population genetic considerations of HGT occurring in bacterial populations.

Monitoring HGT of a plant marker gene in soil

The presence of antibiotic resistance marker genes in transgenic plants has generated concern that such genes, if horizontally transferred to bacteria, may impede treatment of bacterial infections¹². The most frequently used plant marker gene is the neomycin phosphotransferase type II (*nptII*) gene encoding resistance to some aminoglycosides, including kanamycin. The most extensive published study monitoring HGT of any plant marker gene from transgenic plants to soil microbial communities was done after a field trial of transgenic sugar beet (carrying the *nptII* marker gene) in Germany¹⁰. Samples of soil bacteria taken at various time points were selected on kanamycin-containing medium and the colonies further analyzed by PCR and dot-blot analysis. No transformants were identified by this method. How likely were these studies to detect an HGT event, if it had occurred?

Monitoring potential horizontal transfer of the *nptII* marker gene into soil microbes is problematic. A minority of soil bacteria display high levels of intrinsic resistance to most commercially produced antibiotics, including kanamycin. A culture-based screening for kanamycin-resistant soil bacteria yields $\sim 10,000$ resistant bacteria per gram of soil¹³. This high prevalence of resistance therefore quickly saturates any screen for potential new bacterial *nptII* recombinants arising from HGT from transgenic plants. A general observation of kanamycin-resistant colonies in soil is not indicative of the number of bacteria transformed to kanamycin resistance by the plant transgene. This is because several mechanisms can reduce bacterial susceptibility to kanamycin, including intrinsic resistance and at least ten other known kanamycin-resistance genes¹⁴. Further genotypic analysis of the resistant colonies, for example, by hybridization or PCR, is neces-

¹Department of Pharmacy, Faculty of Medicine, University of Tromsø, N9037 Tromsø, Norway, ²Norwegian Institute of Gene Ecology, Research Park, N9294 Tromsø, Norway, ³Department of Plant & Microbial Biology, University of California, Berkeley, California 947200-3102, USA. Correspondence should be addressed to K.M.N. (knielsen@farmasi.uit.no).

Published online 31 August 2004; doi:10.1038/nbt1006

Box 1 Modeling bacterial growth

From Hartl and Clark²³ the current frequency p of a mutant starting at frequency p_0 can be modeled deterministically with Malthusian relative fitness m per generation over (t) generations as

$$\frac{p}{(1-p)} = \frac{p_0 e^{mt}}{(1-p_0)} \quad (1)$$

Immediately after a single HGT event, the frequency of the transgene in a haploid bacterial population is $1/N$ where N is the number of individuals in the population of interest. Thus, for the spread of a transgene subsequent to HGT,

$$\frac{p}{(1-p)} = \frac{e^{mt}}{(N-1)} \quad (2)$$

Solved for p , this yields

$$P = \frac{e^{mt}}{(N-1 + e^{mt})} \quad (3)$$

which was used to generate **Figure 1**.

Then, as long as sample size $n \ll N$ (which is realistic for any currently feasible monitoring study), the probability that a spreading transgene is not detected when sampling is performed at random is $\sim 1 - (1-p)^n$, which was used with p from equation 3 to generate **Figure 2a-c**.

sary to confirm putative HGT events. It seems unlikely that transgenes will ever allow unambiguous selection of bacterial transformants on laboratory media by eliminating all growth of nontransformed bacteria. The current methodology for detecting HGT events thus relies on individual analysis of bacterial colonies, an approach that is time-consuming, expensive and generally limited to processing, at most, several thousand samples.

The study of Gebhard and Smalla¹⁰ screened 4,000 kanamycin-resistant bacteria for potential uptake of the plant marker gene. The 4,000 resistant bacteria represent the amount of kanamycin-resistant bacteria found naturally in 0.4 g of soil, since approximately one of 10,000 culturable bacteria per gram of soil harbors a kanamycin-resistant phenotype¹³. To obtain the 4,000 resistant isolates, the researchers screened $\sim 10,000$ -fold more bacteria on selective media. However, despite this large sampling effort, the experimental design would not detect an *nptII* transformant unless the transfer rate of the *nptII* marker gene was at least one per 4×10^7 exposed, culturable, and selectable bacterium recovered from soil. This rate of HGT is unrealistically high; natural transformation in soil typically occurs at rates below 10^{-7} transformants per recipient¹⁵. Much lower frequencies are expected between transgenic DNA and bacteria under natural conditions. Incorporating a lower transformation frequency (10^{-10}) and prevalence of transformants in the experimental design leads to the observation that at least 100 g of soil sample would have to be analyzed to detect a single transformant among the 10^6 naturally kanamycin-resistant bacteria recovered from the soil sample. This approach would require the use of at least 3,000 agar plates to separate the individual kanamycin-resistant isolates for further molecular analysis. Screening of HGT from transgenic plants into culturable bacteria will therefore, in most cases, be limited to a total soil sample size of 1 gram, since the molecular analysis of individual bacterial colonies in most research laboratories is limited to $\sim 10,000$ isolates. The limited sample size that can be realistically processed makes it exceedingly difficult to accommodate in the experimental design the expected highly variable spatial and temporal distribution of potential transformants in soil. Thus, field monitoring of HGT events occurring in the culturable bacterial fraction is severely limited by existing methodologies for identifying rare transformant genotypes within large bacterial communities.

To detect potential HGT events occurring in the nonculturable fraction of soil bacteria, the researchers isolated DNA from the bacterial

fraction of soil (DNase-treated), and the transgene was assayed with PCR. The bacterial fraction was also plated on kanamycin-containing agar, and the DNA extracted from the resulting lawn of kanamycin-resistant bacteria was targeted for PCR with several transgene-specific primer sets. A 50- μ l PCR solution with 400 ng isolated DNA screens $\sim 2.5 \times 10^7$ bacterial genomes. However, a successful PCR amplification of DNA extracted from soil typically requires >50 copies of the transgene to be present per reaction¹⁶, or one transformant per 5×10^5 bacteria in the sample; the implied rate of HGT is thus higher than experimentally measured rates of HGT in soil¹⁵.

Furthermore, the DNA extraction is biased by a number of factors including biased enrichment (if DNA is extracted after plating), fragmentation, uneven cell lysis, and the presence of DNA from fungi (a major component

of soil biomass) in DNA samples extracted directly from soil. This bias results in a 10- to 100-fold reduction of the sensitivity of the PCR analysis, effectively limiting detection to only those HGT events generating $>2 \times 10^{-5}$ to 2×10^{-4} transformants per total number of bacterial genomes analyzed. It is important to note that increasing the number of samples analyzed by PCR will not increase the likelihood of detecting HGT events if the true prevalence of transformants is below the target copy number required to achieve a successful PCR amplification. A final complication is that even if positive results are obtained through amplification, it remains challenging to trace the PCR signals back to proven bacterial transformants and to rule out signals arising from decaying plant residues^{10,11}. In conclusion, despite the large number of bacteria and genomes analyzed by cultivation and PCR-based analysis in published agricultural field studies^{10,16,17}, the combined amount of soil material examined totals effectively <2 g. Thus, the low sample size analyzed in these studies is insufficient to provide a basis for determination of HGT frequencies of transgenes in agricultural fields. What can be concluded is that HGT from transgenic plants does not occur in the monitored agricultural fields at exceptionally high frequencies.

Monitoring HGT of an herbicide resistance gene in the intestine

A recent study in the UK examined potential horizontal transfer of the 5-enoylpyruvylshikimate-3-P synthase (*epsps*) gene from soybean meal to the microbial flora of human intestine from both healthy volunteers and ileostomists¹¹. The volunteer subjects were fed soy burgers and soy shakes. Because the *epsps* gene does not confer a readily selectable bacterial phenotype, HGT monitoring was based on PCR analysis of bacteria sampled from the intestine and passed through several rounds of subculturing. Whereas indications were found that HGT had occurred in the microflora of three out of seven ileostomists before the feeding study, it was concluded that gene transfer did not occur during the feeding experiment. How likely were these studies to detect an HGT event, if it had occurred? The DNA sample analyzed per PCR analysis was 1 μ g, comprising $\sim 6.3 \times 10^7$ bacterial genomes. A copy number of 80 or more was required for successful amplification. Thus, the PCR monitoring assumed a transformant prevalence higher than 1 per 7.9×10^5 bacteria for identification of HGT events.

This experimental design is particularly limited, however, by the multiple rounds of subculturing performed before bacterial DNA extraction. During the subculturing process, the bacteria (present in

0.5 gram of sample) were enriched 200-fold by growth in Luria-Bertani medium and then diluted 200-fold, and the process was repeated over seven rounds. The transgene was detected by PCR of DNA from the culture during the first rounds of subculturing, albeit the potential bacterial transformants could not be recovered on solid medium. Performing one or more rounds of subculturing makes the analysis effectively uninformative for drawing conclusions about the presence of bacterial transformants in the original intestinal sample. During competitive growth *in vitro*, the initial bacterial species composition and distribution is rapidly altered and only a minor uncharacterized fraction of the bacteria dominates in the enriched culture (the original sample was diluted 10^{16} -fold in this analysis). No a priori expectation exists that potential transformants will be among enriched bacterial species that reproduce best during the >100 generations passed during subculturing. Thus, conclusions cannot be drawn from this study regarding the possible transformability of intestinal bacteria, though it can be concluded that the subculturing did not reveal culturable transformants responding to the specific enrichment, which were present at a prevalence higher than 1 per 5×10^5 culturable bacteria in the initial sample.

Monitoring diverse bacteria in soil or in the human intestine

Less than 10% of the bacterial species present in soil or human intestine grow on laboratory media and can be analyzed via culture-based techniques. Assuming a bacterial density of 10^8 bacteria per gram of soil or intestine content, culture-based techniques have examined potential HGT events in only 0.004% of the bacterial population present in a single gram of soil and, as a result of subculturing steps, an even smaller population in the human intestine. Culturable bacteria subject to further detailed analysis are usually obtained from agar plates plated with diluted bacterial suspensions. Analysis of highly diluted bacterial suspensions may lead to a less complete view of the bacterial diversity present in the original sample. This is because in the serial dilution process, rare bacterial species will be diluted out and only the most abundant culturable bacterial species will be present in the higher dilutions used for plating and therefore have the chance of being picked for further analysis.

As discussed above, a culture-independent PCR-based screen for HGT events does not substantially enhance the number of bacteria analyzed per reaction. Assuming an optimal bacterial species abundance of 50 cells per species per PCR sample, $\sim 5 \times 10^5$ different species are targeted per PCR analysis. However, a biased DNA extraction efficiency^{18,19} and uneven species abundance distribution result in considerably fewer species being present at concentrations higher than the 50–100 genomes (sensitivity limit) required per PCR analysis. Effectively, the number of species analyzed is reduced to several hundred dominant members of the larger microbial populations.

Modeling the effect of selection on transformants

Published monitoring studies have implicitly relied on detecting bacterial transformants within their first generations. Most sampling has been done following transgene exposure and no considerations of transformant growth dynamics have been presented. However, since many bacteria have large population sizes and relatively long generation times in nutrient-limited soil environments, it may take years or several decades for the transformants to reproduce and out-compete nontransformed cells of the population at a scale that is detectable by monitoring programs.

Beyond an initial stochastic period, the frequency of a horizontally transferred, advantageous transgene in a bacterial population increases with time until it is fixed. Thus, the later samples are taken,

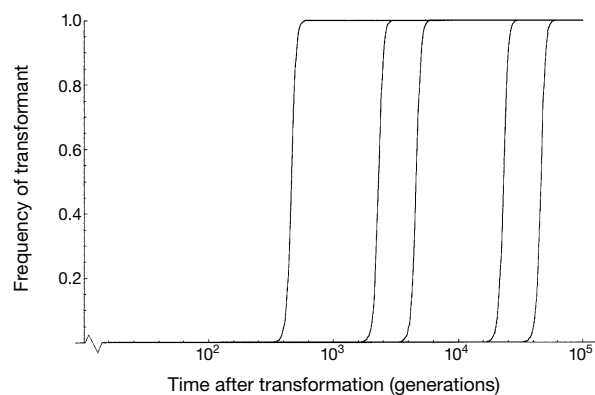


Figure 1 Time to fixation (in bacterial generations) of a single HGT event occurring in a bacterial population of $N = 10^{10}$. Selection coefficients (from left to right) are 0.05, 0.01, 0.005, 0.001 and 0.0005. Bacterial generation time in natural environments ranges from less than an hour to weeks or even longer. Thus, assuming a generation time of 30 minutes, the x-axis spans from 6 hours to 6 years. Assuming a generation time of 8 hours, the x-axis spans from 9 hours to 90 years. Assuming a generation time of one week, the x-axis spans from 20 hours to 2,000 years.

the larger the proportion of bacteria that will carry a positively selected transgene. Population genetic modeling of selection and growth can provide an estimate of a realistic time frame for monitoring efforts and sampling time. Here we model bacterial growth deterministically assuming a single HGT event has taken place and that the transgene is on the way to fixation, that is, it will eventually be present in all members of the bacterial population (see **Box 1**). The selection coefficient conferred by the transferred genetic material will then determine the rate of increase in frequency of rare transformants in a large bacterial population (**Fig. 1**). In the laboratory, selection coefficients (m) as small as 0.01 are frequently measured, and over evolutionary time, selection coefficients as small as the inverse of the effective population size (here, this would be as small as 10^{-10}) are important for determining the genetic content of organisms. Attributing this range of selection coefficients to transgene-harboring bacteria, we calculate that between less than a year to more than 100 years will be needed for transformants carrying positively selected transgenes to multiply to proportions that are detectable within a bacterial population (**Fig. 1**).

Detecting HGT events: sample size and sampling time

The probability of detecting rare HGT events in a large bacterial population depends in part on the sample size (n) of the monitoring effort, where n is the number of bacteria examined for possession of the transgene. If the frequency of the transgene in the population at a given time (**Fig. 1**) is p , then the probability of detection is $1 - (1 - p)^n$. **Figure 2a** shows the number of samples needed to detect a positively selected HGT event, with 95% probability, in a bacterial population of 10^{10} , within the range of a feasible monitoring study (<100,000 bacterial isolates or targets genomes). Clearly, a strongly positively selected transgene and later sampling time drastically reduce the number of bacterial samples needed. Conversely, an early sampling time combined with a weakly selected transgene make virtually any sampling effort futile. The probability of detecting an HGT event with a sample size of 5,000 and the probability of detecting an HGT event when sampling after 1,000 generations are illustrated in **Figure 2b** and **2c**, respectively. As seen, sample size considerations are strongly dependent on the time of sampling and the selective value of the transgene in the transformant.

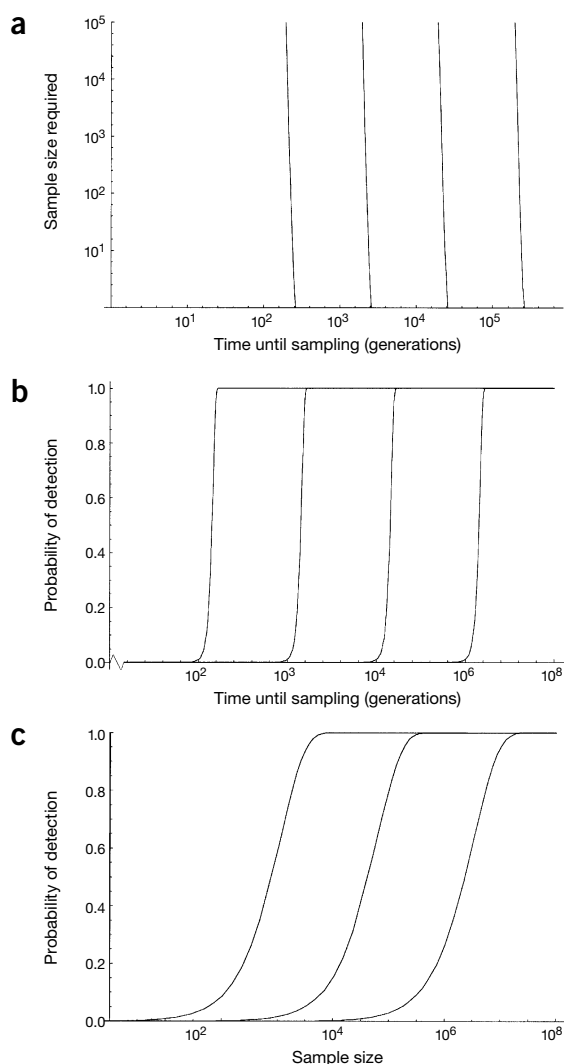


Figure 2 The probability of detection of an HGT event into a bacterial population of 10^{10} . The figure shows how bacterial sample-based monitoring depends upon sample size, sampling time and selection coefficient. **(a)** The sample size necessary to have a 95% probability of detecting a HGT event over increasing time with selection coefficients (from left to right) of 0.1, 0.01, 0.001 and 0.0001. **(b)** The probability of detection of an HGT event into a bacterial population over time when the sample size is limited to 5,000 bacteria with selection coefficients (from left to right) of 0.1, 0.01, 0.001 and 0.0001. **(c)** The probability of detecting bacterial transformants with increasing sample size, assuming 1,000 generations have passed since the initial HGT event (with selection coefficients of 0.016, 0.012 and 0.008).

in and, thus, released from most bacteria (see the Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.ad.jp/kegg/genes.html>, for EC 2.5.1.19). The protein coding sequence of the recombinant counterpart would therefore present little added risk if released from a transgenic plant³. However, increasingly novel recombinant DNA constructs, such as those engineered for production of fine chemicals, pharmaceuticals and vaccines, are now being developed²⁰. These novel constructs have few natural counterparts that will allow a comparative risk assessment to be made. Such recombinant DNA may thus require increased attention in an HGT context, attention particularly devoted to determining whether positive selection of the transgene would be likely in specific bacterial populations or broader bacterial communities. We emphasize that the overwhelming majority of HGT events in nature are known to be deleterious to the bacterial transformant^{21,22}. A deleterious effect is likely to be the outcome of the majority of HGT events occurring from transgenic plants as well, but rare beneficial events cannot be a priori excluded.

Concluding remarks

The above calculations illustrate that real-time monitoring of HGT processes in complex microbial communities such as soil and the intestines has been severely limited by methodological constraints and has relied on weak assumptions on the scale and nature of the process. Little information on the potential for HGT of recombinant DNA into bacterial communities can be drawn from these studies because of the lack of defined testable hypotheses and the necessarily limited sample size analyzed (<2 g).

We suggest several major changes that are needed to advance investigations of HGT processes in natural bacterial communities. First, the current focus on HGT frequencies is largely irrelevant (assuming a few HGT events do occur) as a basis to understand the effects of HGT of recombinant DNA into bacterial communities. An observable effect will be seen only after growth and amplification of the initially produced transformants. Thus, the focus of future research and monitoring efforts must be shifted to the identification of bacterial genetic compositions and environmental conditions that facilitate positive selection of bacterial transformants carrying horizontally acquired transgenes. Second, additional information and more detailed hypotheses on transgene selection coefficients, the population size and the generation time of the exposed microbial populations are therefore essential to enable realistic monitoring strategies of dynamic bacterial populations. Lastly, in the interest of producing quality data that would be informative in biological risk assessment, it is crucial to clearly communicate the inherent limitations in the methods and experimental design used. This requires careful consideration of the detection limit, the sensitivity of the methods used and the sampling size and time necessary to resolve the hypotheses presented.

Assessing the long-term impact of HGT events

Much of the current debate on the relevance and risk of HGT to transgenic plant biosafety is focused on the expected low (if any) probability of HGT from transgenic plants. However, realistic estimates of HGT rates are rarely obtainable, nor are they predictive of the long-term effects of infrequent HGT events. Risk is composed of both a probability (frequency of HGT) and an effect component (effect of transgene on survival of bacterial transformant). Cautious consideration of the latter is rarely made. Since rare HGT events cannot be excluded within the current understanding of the process, we argue that risk assessment should be based on the effect component. Most horizontally transferred transgenes are likely to cause negative effects on transformant survival. Transgenes that do not affect transformant survival (neutral or near neutral selection) are also likely to be lost from the bacterial population over time either through loss of transformants or due to mutational inactivation and deletions. To evaluate potential risk-relevant effects, it will be necessary to identify those transgenes that will experience positive selection if horizontally transferred into bacteria, acknowledging the temporal and spatial variation in selection in complex environments such as soil.

Since most transgenes in commercial use originate from soil, empirical evidence is already available on their genetic effects when exposed to bacteria³. For instance, the *epsps* gene is naturally present

ACKNOWLEDGMENTS

K.M.N. received financial support from the Research Council of Norway. J.P.T. thanks the Miller Institute for Basic Research in Science for funding and administrative support. We thank M. Choi for comments on the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at <http://www.nature.com/naturebiotechnology/>

- Maynard Smith, J., Feil, E.J. & Smith, N.H. Population structure and evolutionary dynamics of pathogenic bacteria. *Bioessays* **22**, 1115–1122 (2000).
- Ochman, H., Lawrence, J.G. & Groisman, E.A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).
- Nielsen, K.M. An assessment of factors affecting the likelihood of horizontal transfer of recombinant plant DNA to bacterial recipients in the soil and phytosphere. *Coll. Biosafety Rev. (Italy)* **1**, 96–149 (2003).
- Conner, A.J., Glare, T.R. & Nap, J.-P. The release of genetically modified crops into the environment, part II. Overview of ecological risk assessment. *Plant J.* **33**, 19–46 (2003).
- Nielsen, K.M., Bones, A., Smalla, K. & van Elsas, J.D. Horizontal gene transfer from transgenic plants to terrestrial bacteria—a rare event? *FEMS Microbiol. Rev.* **22**, 79–103 (1998).
- Ceccherini, M.T. *et al.* Degradation and transformability of DNA from transgenic leaves. *Appl. Environ. Microbiol.* **69**, 673–678 (2003).
- De Vries, J., Meier, P. & Wackernagel, W. The natural transformation of the soil bacteria *Pseudomonas stutzeri* and *Acinetobacter sp.* by transgenic plant DNA strictly depends on homologous sequences in the recipient cells. *FEMS Microbiol. Lett.* **195**, 211–215 (2001).
- Gebhard, F. & Smalla, K. Transformation of *Acinetobacter sp.* strain BD413 by transgenic sugar beet DNA. *Appl. Environ. Microbiol.* **64**, 1550–1554 (1998).
- Kay E. *et al.* *In situ* transfer of antibiotic resistance genes from transgenic (transplasmic) tobacco plants to bacteria. *Appl. Environ. Microbiol.* **68**, 3345–3351 (2002).
- Gebhard, F. & Smalla, K. Monitoring field releases of genetically modified sugar beets for persistence of transgenic plant DNA and horizontal gene transfer. *FEMS Microbiol. Ecol.* **28**, 261–272 (1999).
- Netherwood, T. *et al.* Assessing the survival of transgenic plant DNA in the human gastrointestinal tract. *Nat. Biotechnol.* **22**, 204–209 (2004).
- European Food Safety Authority. Opinion of the scientific panel on genetically modified organisms on the use of antibiotic resistance genes as marker genes in genetically modified plants. *The EFSA J.* **48**, 1–18 (2004).
- Hanscke, R.B. & Schmidt, F.R.J. Screening of soil bacteria for plasmids carrying antibiotic resistance. *Biol. Fertil. Soils* **9**, 257–260 (1990).
- Shaw, K.J., Rather, P.N., Hare, R.S. & Miller, G.H. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiol. Rev.* **57**, 138–163 (1993).
- Nielsen, K.M., van Elsas, J.D. & Smalla, K. Transformation of *Acinetobacter sp.* BD413(pFG4ΔnptII) with transgenic plant DNA in soil microcosms and effects of kanamycin on selection of transformants. *Appl. Environ. Microbiol.* **66**, 1237–1242 (2000).
- Badosa, E., Moreno, C. & Montesinos, E. Lack of detection of ampicillin resistance gene transfer from Bt176 transgenic corn to culturable bacteria under field conditions. *FEMS Microbiol. Ecol.* **48**, 169–178 (2004).
- Paget, E., Lebrun, M., Freyssinet G. & Simonet, P. The fate of recombinant plant DNA in soil. *Eur. J. Soil. Biol.* **34**, 81–88 (1998).
- Frostegaard, Å. *et al.* Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* **65**, 5409–5420 (1999).
- Martin-Laurent, F. *et al.* DNA extraction from soils: old bias for new microbial diversity analysis methods. *Appl. Environ. Microbiol.* **67**, 2354–2359 (2001).
- Ma, J.K.-C., Drake, P.M. & Christou, P. The production of recombinant pharmaceutical proteins in plants. *Nat. Rev. Genet.* **4**, 794–805 (2003).
- Elena, S.F. *et al.* Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. *Genetica* **102–103**, 349–358 (1998).
- Remold, S.K. & Lenski, R.E. Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. *Nat. Genet.* **36**, 423–426 (2004).
- Hartl, D.L. & Clark, A.G. *Principles of population genetics*, edn. 3 (Sinauer Associates Inc. Sunderland, MA, USA, 1997).