Genetically modified plants and the 35S promoter: assessing the risks and enhancing the debate.

R. Hull, S.N. Covey and P. Dale
John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Abstract
The 35S promoter, derived from the common plant virus, cauliflower mosaic virus (CaMV), is a component of transgenic constructs in more than 80% of genetically modified (GM) plants. Alarming reports have suggested that the 35S promoter might cause accidental activation of plant genes or endogenous viruses, promote horizontal gene transfer, or might even recombine with mammalian viruses such as HIV, with unexpected consequences. In this article, we discuss the properties of CaMV and the 35S promoter and the potential risks associated with the use of the promoter in GM plants, concluding that any risks are no greater than those encountered in conventional plant breeding.

Introduction
In a recent article, Ho *et al*. (1999) suggested that the widespread use of the 35S promoter of cauliflower mosaic virus (CaMV) in transgenic plants is "a recipe for disaster". Ho et al. (1999) base their arguments on three considerations, a) that the CaMV 35S promoter has a hotspot for recombination (Kohli *et al*., 1999); b) that the 35S promoter has several domains with different tissue specificities and c) that the 35S promoter is very efficient and can function in a wide range of organisms, not only plants but also bacteria and animals. From this they deduce that the 35S promoter could recombine to activate dormant viruses, create new viruses and "cause cancer by the overexpression of normal genes". As scientists who have worked on CaMV for up to 25 years and have contributed much to the understanding of its molecular biology we wish to put these scenarios in the correct context.

The first plant promoter
Over 15 years ago, CaMV was one of several plant genetic systems being studied for its potential use in plant transformation (Hull, 1983; 1984; 1985). As part of these studies, much basic research went on into understanding the genetic organisation of CaMV and the means by which its genes are expressed and regulated. The CaMV genome was the first significant piece of plant DNA to be completely sequenced (Frank *et al*., 1980) and the two CaMV promoters, the 35S and 19S promoters, were the first plant promoters identified (Covey *et al*., 1981; Hull and Covey, 1983c; Odell *et al*., 1995). Because of the latter discovery, and the finding that the 35S promoter was active in directing heterologous expression of plant genes in a variety of plants, its use in the development of GM plants for research and agronomic applications became widespread.

The virus
CaMV can infect a wide range of crucifers (see Schoelz and Bourque, 1999) and is commonly found in cabbages, cauliflowers, oilseed rape, mustard and other brassicas in temperate countries (Tomlinson, 1987). A survey of a local market, as part of a risk assessment exercise for the Ministry of Agriculture, Fisheries and Food (the UK regulatory authority for biosafety of genetic manipulation of plant pests) in the late 1980s, showed that about 10% of the cauliflowers and cabbages were infected with CaMV. The virus is transmitted in nature by aphids (see Schoelz and Bourque, 1999) and it might be expected that organically grown crucifers, on which the aphids have not been controlled by insecticides, would have higher rates of infection. Infection early in the plant growth might affect the quality, especially of cauliflowers, but later infections show leaf symptoms but little other overt effects. The virus infects most cells of the plant and produces about 10^5 particles per cell. Each particle contains one molecule of the viral genome, an 8 kbp circular double-stranded DNA with one copy each of the two promoters.

The replication cycle of the virus has two phases (see Hohn, 1999), the first in the nucleus where the viral genome is uncoated, forms a minichromosome and is transcribed to give two RNA species, the 35S RNA (using the 35S promoter) and the 19S RNA (using another promoter). These RNAs pass to the cytoplasm where the 35S RNA acts as a template for reverse transcription as well as a template for translation of some gene products; the 19S RNA is the template for the translation of just one gene product. Various unencapsidated replication intermediates are found in infected cells (Hull and Covey, 1983a) which are estimated to give a further 10^4 CaMV molecules per cell. CaMV was the first plant virus shown to involve reverse transcription in its replication (Hull and Covey, 1983b; Pfeiffer and Hohn, 1983) but has now been joined by more than a thirty other plant viruses which have the same replication mechanism.

Reverse transcribing elements

There is a wide range of reverse transcribing elements from animals, plants, bacteria and fungi (Hull, 1999). These elements are grouped together on common features of replication which use the enzyme reverse transcriptase. It is considered likely that they had a common ancestor. There are two basic types of reverse transcribing elements. Reverse transcribing viruses express gene products which enable them to move between hosts. These viruses fall into two basic groups, the retroviruses which encapsidate RNA and whose replication involves integration of the viral sequence into the host genome, and pararetroviruses which encapsidate DNA and whose normal replication does not involve integration of the viral sequence. In contrast, retroelements are reverse transcribing sequences integrated into the host genome but which do not produce gene products necessary for easy horizontal movement between hosts.

Retroviruses

The retrovirus family comprises a large number of viruses which infect vertebrates and cause a range of diseases including tumors, leukemia and immunodeficiency. Retroviruses express a gene product, the integrase protein, which facilitates the integration of the viral genome into the host chromosomes where it can stay dormant for a considerable time. The virus particles, which are horizontally transmitted, contain RNA transcribed from this integrated genome. On entry into a new host this RNA is reverse transcribed to give the DNA which is then integrated.

Plant pararetroviruses

There are two groups of plant pararetroviruses, the caulimoviruses and the badnaviruses (Hull, 1995; Lockhart *et al*., 1995) (the International Committee on Taxonomy of Viruses has a more complex classification system but for this article it is best to consider these viruses in these two groups). The caulimovirus group includes viruses which infect groundnuts, soybeans and cassava as well as brassicas. Crops infected by badnaviruses include banana, cocoa, citrus, yams, pineapple and sugarcane. Although there are differences in genome organisation between these two virus groups their replication and the expression of their genes are very similar. The DNA in the virus particles is transcribed in the nucleus (using the 35S promoter) to give RNA which is replicated by reverse transcription to form the DNA molecules which are encapsidated in the virus particles.

Animal pararetroviruses

The one group of animal pararetroviruses, the hepadnaviruses, contains the human-infecting hepatitis B virus and several viruses which infect ground squirrels, woodchucks and ducks. These viruses have a very different genome organisation to plant pararetroviruses, and although they replicate by reverse transcription, there are major differences from retroviruses and plant pararetroviruses in the details of the replication mechanism (Mason *et al*., 1987; compare Seeger 1999 for hepadnaviruses with Hohn, 1999 for caulimoviruses).

Retrotransposons.

Retrotransposons are similar to animal retroviruses in that they are integrated into the host chromosomes and their replication is by mechanisms basically the same as those of retroviruses using a promoter analogous to the 35S promoter. As noted above, they lack the gene products which enable them to spread horizontally between hosts as do the retroviruses. However, there is some recent evidence that they can move very infrequently between hosts (Jordan et al., 1999) by an unknown mechanism. Plant genomes contain large numbers of retrotransposons (Bennetzen and Kellog, 1997). For instance, up to 50% of the maize genome is made up of retrotransposons (SanMiguel *et al*., 1996) though many of these have mutated and are likely to be inactive.

Relationships between reverse transcribing elements

As noted above, all these elements have a common mode of replication involving reverse transcription of RNA to form DNA. However, functionally reverse transcribing elements fall into two types, the retroviruses, plant pararetroviruses and retrotransposons forming one type and the animal pararetroviruses forming the other type. Within these types, each virus or element is a highly coordinated structure and perturbation of the sequence leads to loss of infectivity or functionality. This is exemplified by the attempts to use cauliflower mosaic virus as a gene vector (see Hull, 1983). In most cases, it is only when closely related sequences are added to or exchanged within the genome that viability is retained.

There are many instances of plant and animal cells containing more than one type of reverse transcribing virus or element. For instance, there are several different human retroviruses and no instance of recombination has been found between them. This indicates that there are many constraints on natural recombination.

Now to consider the specific points raised by Ho *et al*. (1999).

1. The integrated 35S promoter can recombine with dormant viruses and also create new viruses. We need to consider the situation firstly in plants and then in animals which may eat those plants

For plants:

a) As noted above, there are more than 10^5 copies of the 35S promoter in each cell of a plant naturally infected by CaMV, in contrast to the one to a few copies of the 35S promoter in each cell of transformed plants. All these other pararetroviruses have similar numbers of copies of their genomes in their hosts but differ markedly in sequence from CaMV, even in the 35S promoter region. Also noted above is that plants contain large numbers of retrotransposons.

In spite of these high numbers of both 35S promoter and retrotransposons no cases of natural recombination leading to new viruses have been found in spite of intensive research on these virus groups.

b) There is uncertainty concerning the stage of transformation at which the recombination described by Kohli *et al.* (1999) occurred. They did not distinguish between recombination taking place during the process of transformation and recombination taking place after the sequences had been integrated. There is accumulating evidence of rearrangements of DNA during transformation (e.g. De Groot *et al.*, 1994; Register *et al.*, 1994). In most cases, these rearrangements result in the non-functioning of the transgene and are selected out in the early stages of analysis of the properties of the transformed lines. Furthermore, the construct used for transformation by Kohli et al (1999) had three copies of the 35S promoter, one in inverse orientation in relation to the other two. The presence of repeated sequences in transformed integrants, and especially inverse repeats, also tends to lead to gene silencing (see Kooter et al. 1999), a condition which would be selected against in developing the transgenic line.

c) Assuming that the integrated 35S promoter does have recombinational properties, for it to effect the activation of a dormant virus or create a new virus, the whole promoter would have to be either excised and reinserted precisely at the new site or its 3' end linked precisely with another gene. The first case means that there would have to be two recombination "hotspots", that identified by Kohli *et al.* (1999) just downstream of the TATA box and an upstream one which would enable the other promoter elements to be included in the excised fragment. There is no evidence for such an upstream "hotspot". Excision of the promoter would also be required for any potential effects on animals.

      It is important to note that genetic recombination is a normal feature of conventional plant breeding and of all natural populations (Hayward *et al.*, 1993). Genetic variation is obtained by recombination following hybridisation between genetically different plants which may be from different species and genera. There are also irregular sources of recombination by minor and major chromosomal rearrangements and from the movement of transposable elements, which can move from one part of the genome to another. Thus, recombination and hotspots for recombination are not unique features of the CaMV 35S promoter

For animals

d) The scenario suggested here by Ho *et al.* (1999) is that the 35S promoter would recombine with hepadnaviruses such as human hepatitis B virus. This is based on the suggestion of a "close relationship between CaMV and hepadnaviruses". Hepadnaviruses have been classed as pararetroviruses as they have a DNA genome and they do not involve integration in their replication cycle. However, as noted above, their replication cycle differs significantly from that of caulimoviruses. Furthermore, there is no sequence similarity between the "hotspot" on the CaMV 35S promoter and hepadnavirus sequences important in replication.

e) Various problems would have to be overcome for the CaMV 35S promoter in transgenic food plants to interact with human (or other animal) viruses or any other sequences. Firstly, as noted above, the promoter would have to excise in an appropriate manner. If the transgenic food was cooked the DNA would be denatured and be very unlikely to renature in an operable form. The next major problem is that the DNA containing the promoter would be exposed to nucleases both from the plant cells when they were disrupted and in the animals gut. Less than 5% administered DNA survives up to 7 hours in an animal gut and that DNA is cleaved into very small pieces (Schubbert *et al.*, 1994). The DNA would then have to pass into the gut cells and integrate precisely to activate the animal sequence. As noted above, the consumption of CaMV-infected vegetables would result in the ingestion of vastly more copies of the 35S promoter than the consumption of transgenic plants containing the promoter. Brassicas are not the only crops which contain pararetrovirus sequences. All banana varieties (and other Musa varieties), the worlds fourth most important agricultural commodity, have multiple copies of the sequences of banana streak badnavirus integrated into their genomes (Harper *et al.*, 1999; Ndowora *et al.*, 1999). In spite of exposure of humans to these pararetroviruses there is no evidence of any ill effects from them even in countries such as Uganda where bananas are the staple diet and HIV is rife.

2. That recombination of the 35S promoter would lead to overexpression of "normal" genes.

Once again, one has to consider both plants and animals separately.

Plants

f) The same arguments apply here as those under a) - c) above. However, if we assume that the 35S promoter does recombine to enhance the expression of a "normal" gene leading to the plant producing a harmful substance what are the factors to be considered? Recombination is likely to be a very rare event and to be in one or a few cells. For it to have any impact it would have to be in germline cells so that it could be fully expressed in the progeny. Also to have any impact it would have to be in an early generation of bulking up of seed of that line. Recombination would result

3

in the loss of the transgenic character which would be recognised in these early generations. By affecting a "normal" plant gene it would probably affect the plant phenotype. Thus, the normal procedures used over many years for the selection of new plant varieties from conventional breeding programmes would identify this significant change.

g) Plants contain many secondary metabolites which have evolved to provide defense mechanisms against herbivores. The standard tests for carcinogens using rodents indicate that a significant proportion of these secondary metabolites are "carcinogens". For instance, out of 28 tested from coffee, 19 were carcinogenic in rodents and 35 out of 63 natural plant products were carcinogens (Ames and Gold, 1997). Since these potentially carcinogenic compounds have not been implicated in causing human cancers, it is considered that humans are more adapted to them. Thus, the overexpression of normal genes is very unlikely to cause cancers

Animals

h) The same arguments as in d) and e) for the 35S promoter recombining with an animal gene sequence also apply here. It must be remembered that animals other than humans also eat brassicas and there are no reports of any poisoning or disease which could be related to CaMV infection.

Conclusions

The paper by Ho *et al.* (1999) highlights some of the basic features of the controversy over genetically modified crops. The transgenic situation has to be compared with the natural situation not with a utopian one. It is well known that viral sequences recombine naturally and that the vast majority of these recombinants are unsuccessful. Very occasionally new viruses arise, this being one of the major ways by which viruses evolve. However, this recombination is between viruses which occur in plant cells at very much higher concentrations than those of transgenic sequences, whether they be promoters or the transgenes themselves. Furthermore, plants are full of noxious substances which have evolved over many millenia to protect the plant against herbivores and pathogens. Plant breeding and selection has, and is, continuing to minimize(d) the effects that these substances have on humans.

From the arguments above, there is no evidence that the CaMV 35S promoter will increase the risk over those already existing from the breeding and cultivation of conventional crops. . There is no evidence that the 35S promoter, or other retroelement promoters, will have any direct effects, in spite of being consumed in much larger quantities than would be from transgenes in GM crops. Furthermore, there are compelling arguments to support the view that there would be no more risks arising from potential recombination than there are from existing non-transgenic crops.

References

Ames, B.N. and Gold, L.S. (1997). Environmental pollution, pesticides, and the prevention of cancer: Misconceptions. FASEB Journal 11: 1041-1052

Bennetzen J.L. and Kellog, E.A. (1997). Do plants have a one-way ticket to genomic obesity? The Plant Cell 9: 1509-1514.

Covey. S.N., Lomonossoff, G.P. and Hull, R. (1981). Characterisation of cauliflower mosaic virus DNA sequences which encode major polyadenylated transcripts. Nucl. Acids Res. 9: 6735-6747

De Groot, M.J., Offringa, R., Groet, J., Does, M.J., van Hooykaas, P.J. and dan Elzen, P.J. (1994). Non-recombinant background in gene targetting: illegitimate recombination between *hpt* gene and defective 5' deleted *nptII* gene can restore kan[r] phenotype in tobacco, Plant Mol. Biol. 25: 721-733.

Frank, A., Guilley, H., Jonard, G., Richards, K.E. and Hirth, L. (1980). Nucleotide sequence of cauliflower mosaic virus DNA. Cell 21: 285-294.

Harper, G., Osuji, J.G., Heslop-Harrison, J.S. and Hull, R. (1999). Integration of banana streak badnavirus into the *Musa* genome: molecular and cytological evidence. Virology 255: 207-213.

Hayward, M.D., Bosemark, N.O. and Romagosa, I. (Eds) (1993). Plant breeding: Principles and Prospects. pp.550. Chapman and Hall, London

Ho, M.-W., Ryan, A. and Cummins, J. (1999). Cauliflower mosaic virus promoter – a recipe for disaster. Microb. Ecol. Health Dis. 10: 33-59.

Hohn, T. (1999). Caulimoviruses - molecular biology. In Encyclopedia of Virology. Eds A. Granoff and R.B. Webster. Pp.1281-1285. Academic Press, San Diego

Hull, R. (1983). The current status of plant viruses as potential RNA/DNA vector systems. In Plant Biotechnology. Society for Experimental Biology Seminar Series 18. Eds. S.H. Mantell and H. Smith. Cambridge University Press. pp 299-312

Hull, R. (1984). Is cauliflower mosaic virus still a potential vector for genes into plants? Heredity 53: 573.

Hull, R. (1985) The potential of plant viral nucleic acids in gene transfers. Swiss Biotech 3: 35.

Hull, R. (1995). Genus Caulimovirus. In Virus Taxonomy, 6th report of the International Committee on Taxonomy

of Viruses. Eds F.A. Murphy et al. Pp. 189-192. Springer Verlag, Wien.

Hull, R. (1999). Classification of reverse transcribing elements: a discussion document. Archives of Virology 144: 209-214.

Hull, R. and Covey, S.N. (1983a). Characterisation of cauliflower mosaic virus DNA forms isolated from infected turnip leaves. Nucleic Acids Research 11: 1881-1895

Hull, R. and Covey, S.N. (1983b) Does cauliflower mosaic virus replicate by reverse transcription? Trends in Biochemical Sciences 8: 119-121

Hull, R. and Covey, S.N. (1983c). Replication of cauliflower mosaic virus DNA. Sci. Prog. Oxf. 68: 403-422.

Jordan, I.K., Matyunira, L.V. and McDonald, J.F. (1999) Evidence for the horizontal transfer of long terminal repeat retrotransposon. Prtoceedings of the National Academy of Sciences, USA 96: 12621-12625

Kohli, A., Griffiths, S., Palacios, N., Twyman, R.M., Vain, P., Laurie, D.A. and Christou, P. (1999). Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology medaited recombination. The Plant Journal 17: 591-601.

Kooter, J.M., Matzke, M.A. and Meyer, P. (1999). Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends in Plant Sciences 4: 340-347

Lockhart, B.E.L., Olszewski, N.E. and Hull, R. (1995). Genus Badnavirus. In Virus Taxonomy, 6th report of the International Committee on Taxonomy of Viruses. Eds F.A. Murphy et al. Pp. 185-188. Springer Verlag, Wien.

Mason, W.S., Taylor, J.M. and Hull, R. (1987). Retroid virus genome replication. Advances in Virus Research 32: 35-95.

Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R. Olszewski, N.E. and Lockhart, B. (1999). Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. Virology 255: 214-220.

Odell, J.T., Nagy, F. and Chua, N.-H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313, 810-812.

Pfeiffer, P. and Hohn, T. (1983). Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell 33: 781-789.

Register, J.C., Peterson, D.J., Bell, P.J., et al. (1994) Structure and function of selectable and non-selectable transgenes in maize after introduction by particle bombardment. Plant Mol. Biol. 25: 951-961

SanMiguel, P., Tikhonov, A., Jin, Y.-K., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Avramova, Z. and Bennetzen, J.L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765-768.

Schoelz, J.E. and Bourque, J.E (1999). Plant pararetroviruses (Caulimoviridae). In Encyclopedia of Virology. Eds A. Granoff and R.B. Webster. Pp.1275-1281. Academic Press, San Diego

Schubbert, R., Lettman, C. and Doefler, W. (1994). Ingested foreign (phage M13) DNA survives transiently in the gastronomic tract and enters the bloodstream of mice. Mol. Gen. Genet. 242: 495-504.

Seeger, C. (1999). Hepadnaviruses (Hepadnaviridae): Hepatitis B viruses: Molecular biology. In Encyclopedia of Virology. Eds A. Granoff and R.B. Webster. Pp.645-650. Academic Press, San Diego

Tomlinson, J.A. (1987). Epidemiology and control of vegetable diseases. Ann. Appl. Biol. 110, 661-681.