Gene shuffling improves the function of *Tobacco mosaic virus* movement protein

R.L. Toth, S. Chapman & G. Pogue

In a major collaborative venture between SCRI and Large Scale Biology Corporation (LSBC), a Californian-based biotechnology company, we are exploring the use of viral vectors for the expression of foreign proteins in plants. Vectors based on plant viruses offer an alternative to stable transformation for the production of biopharmaceuticals in plants and much research has focused on their use for vaccine production, through the expression either of short pathogen-derived peptides as fusions to viral coat proteins or of complete proteins[1]. The speed and ease with which high levels of foreign gene expression can be achieved has also led to their extensive use as research tools for virological and cell biological studies, and for *in planta* analysis of protein functions[2]. Virus expression vectors based on the *Tobacco mosaic virus* (TMV) genome have proved to be powerful tools for expression of pharmaceutically-relevant proteins and have speeded plant functional genomics through either over-expression of plant genes or gene silencing. Despite the utility of the present vector systems, the inclusion of additional genetic load, in the form of foreign genes, reduces the replication and movement efficiencies of the vectors when compared with their wild-type counterparts and these reductions can negatively effect host range and expression levels. Previous improvements in levels of foreign protein expression, genetic stability and infection phenotypes of TMV-based vectors have been made through the inclusion of extra promoter elements to drive foreign gene expression and the creation of hybrid tobamovirus genomes[2]. These approaches have suffered from their reliance on the available genetic diversity of virus isolates and did not seek to evolve optimal viral components to compensate for deficiencies present in the expression vectors.

We are using various techniques to improve TMV as a vector for expression of useful foreign proteins. TMV has a positive-sense, single-stranded, RNA genome of 6396 nucleotides and encodes four protein products. One of these, a 30kDa protein (30K or movement protein), is essential for cell to cell movement of the virus. It has been shown to associate with plasmodesmata, intercellular cytoplasmic channels, to bind single-stranded nucleic acids *in vitro* and to interact with plant cytoskeletal elements. Further to its coding function, the 30K gene contains an RNA structure, from which assembly of viral RNA and coat protein into virus particles is initiated, and a subgenomic promoter sequence that in the natural virus directs the synthesis of the mRNA encoding the coat protein. The multifunctional nature of this viral gene, and its requirement for viral movement, makes it an ideal target for improvement through mutagenesis and DNA shuffling.

The technique of DNA shuffling[3] is a PCR-based method for fragmentation and reassembly of mutated copies of a gene of interest to produce a population of recombined gene sequences. This *in vitro* method mimics, in an accelerated fashion, the natural evolutionary process of genetic recombination. Many of the shuffled gene sequences produced will be nonfunctional or debilitated in function, but a proportion will have an improved function, which can be identified using an appropriate screen. The technique has proved to be extremely successful in evolving enzymes with novel specificities or enhanced activities and more recently in altering the cell tropisms and processing characteristics of retroviruses[4-6]. We have performed DNA shuffling to evolve the 30K gene and its product to compensate for losses in activity

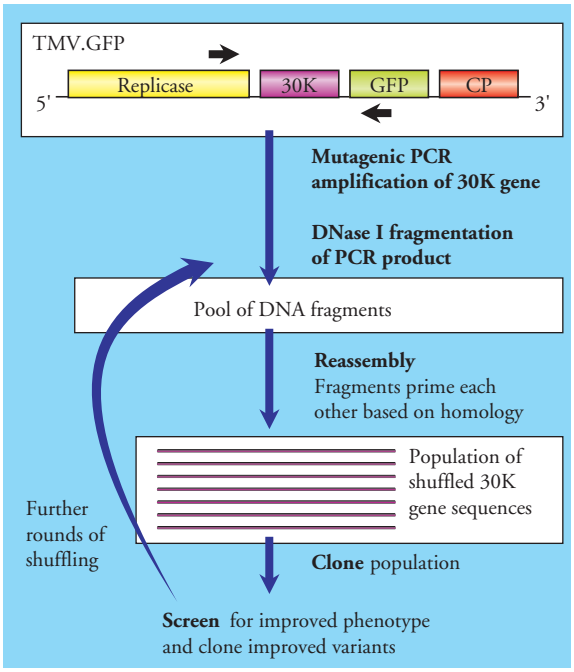

Figure 1 Shuffling of the TMV 30K gene.

incurred by foreign gene insertion. A previously described TMV-based vector that carries the gene for green fluorescent protein (TMV.GFP, Fig. 1) was used to permit an easy visual screen for improvements in cell to cell movement.

Shuffling of the movement protein gene

A portion of the TMV genome encompassing the subgenomic promoter for the 30K gene and the 30K gene itself was subjected to DNA shuffling (Fig. 1). The population of shuffled fragments was cloned into TMV.GFP and inoculated to leaves of the highly susceptible experimental host *Nicotiana benthamiana* as reassembled virus particles. The shuffled population produced about 20% of the number of fluorescent infection foci produced by unshuffled virus and the size of the infection foci, generally smaller, was more disparate for the population than the control, indicating that the majority of the population was disfunctional or disabled (Fig. 2A). Visual screening of approximately 5000 fluorescent infection foci identified 70 lesions that appeared larger than the control lesions, indicating faster cell to cell movement. To confirm the improved phenotypes, homogenates prepared from the variant lesions were inoculated to half-leaves of the more restrictive host *N. tabacum* cv. Xanthi nn with controls on the opposing half-leaf (Fig. 2B). A total of 53 variants showed faster cell to cell movement and the shuffled portions of their genomes were cloned back into the progenitor plasmid through RT-PCR.

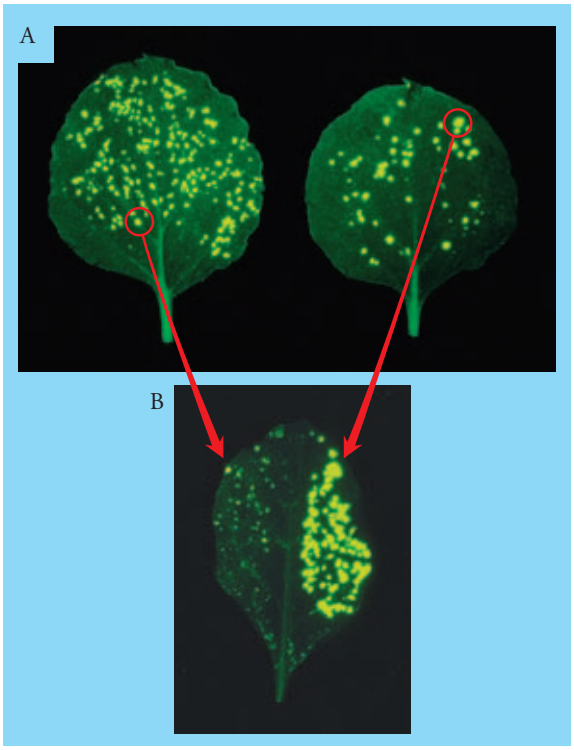

Figure 2 Screening shuffled populations for improved movement phenotype. A; *N. benthamiana* leaves inoculated with unshuffled control virus (on the left) or 30K shuffled population (on the right). B; Individual lesions passaged to *N. tabacum* cv. Xanthi. Unshuffled control on left side of leaf; improved variant on right side of leaf.

Clones produced from the 53 variants were used in a second round of shuffling to generate recombinants with new patterns of mutations and phenotypes. The population obtained was screened on *N. benthamiana* and *N. tabacum* cv. Xanthi nn as before. Variants showing faster cell to cell movement than the best clone from the first round were selected and 37 clones obtained. These clones were input into a third round of shuffling in the same way. The population produced was screened on the restrictive production host *N. tabacum* cv. MD609 to facilitate discrimination of differences in cell to cell movement rate. After secondary screening, three clones showing faster cell to cell movement than the best clone from the second round were obtained.

Shuffled clones are improved for cell to cell and systemic movement

The cell to cell and long-distance movement rates of the progenitor, unshuffled, clone (*unshuff*) and the best clone from each of the three rounds of shuffling (*shuff1*, *shuff2* and *shuff3*) were assessed to evaluate

Inoculum	Mean area of fluorescent infection foci (mm²)	
	N. tabacum cv. MD609	*N. benthamiana*
unshuff	0.219	0.172
shuff1	0.309	0.554
shuff2	0.816	0.660
shuff3	0.967	0.913

Table 1 Fluorescent infection foci area measurement. The areas of fluorescent foci on inoculated leaves of *N. tabacum* cv. MD609 and *N. benthamiana* were measured at 5 dpi or 3 dpi, respectively. Analysis of variance gave values of 0.0599 and 0.0683 for the least significant difference at the 5% level for measurements of foci on *N. tabacum* cv. MD609 and *N. benthamiana*, respectively.

more critically the visually discriminated improvements. Rates of cell to cell movement were investigated through area measurements of multicellular, fluorescent, infection foci produced by the four clones on *N. benthamiana* and *N. tabacum* cv. MD609 (Table 1). Statistical analysis of the data showed that, on both hosts, *shuff1* produced significantly larger lesions than the progenitor construct and that the two clones from subsequent rounds of shuffling produced larger lesions than the best clone from the previous round of shuffling.

To test whether the mutations that improved cell to cell movement also affected the speed at which uninoculated leaves became infected, the time at which systemic fluorescence developed was monitored for the four clones on the two different hosts. All the shuffling progeny produced systemic fluorescence at the same time on the experimental host *N. benthamiana*, one day faster than the progenitor. On the production host *N. tabacum* cv. MD609, there were more marked differences in the timing of systemic GFP expression (Fig. 3). The shuffling progeny produced systemic fluorescence at least 13 days in advance of the progenitor and the clones *shuff2* and *shuff3* produced systemic infections one to two days in advance of *shuff1*. By 20 dpi, when *shuff1* had only systemically infected 2-4 leaves, *shuff2* and *shuff3* had systemically infected 4-5 or 5-6 leaves respectively. Furthermore, a higher proportion of the surfaces of systemically infected leaves displayed green fluorescence in *shuff2* and *shuff3* infections than in *shuff1* infections. Although the 30K protein has been proposed to have a role in systemic movement, it seems more probable that the enhanced long-distance movement is a consequence of the faster cell to cell move-

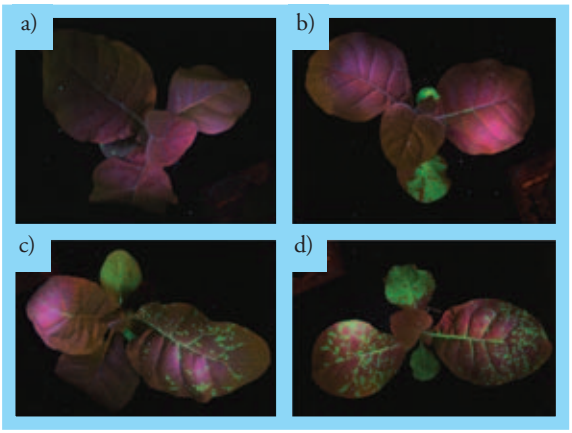

Figure 3 Systemic infections of *N. tabacum* cv. MD609 plants 10 days post-inoculation with *unshuff* (a), *shuff1* (b), *shuff2* (c) and *shuff3* (d). Plants were inoculated with reassembled virus and maintained at 28°C prior to observation of green fluorescence under UV illumination.

ment, possibly allowing the virus to evade host responses more effectively.

Detection of 30K in protein extracts prepared from protoplasts inoculated with transcripts of *shuff1*, *shuff2* and *shuff3*, showed that they accumulated higher levels of 30K protein than *unshuff*, indicating that they contained mutations that stabilised the movement protein. Whether the mutations affect other aspects of 30K protein functionality is the topic of current studies in our laboratory. Whatever the mechanisms, the shuffling process has produced mutations that can compensate for the reduced levels of 30K expression that result from the introduction of a foreign gene. In addition, the mutations not only improve movement on permissive hosts but also broaden the host range to previously restrictive hosts, such as *N. tabacum* cv. MD609.

Effects of mutations on wild-type virus

Given the high mutation rate reported for TMV and susceptibility of RNA viruses to recombination, it seemed unlikely that if the changes found in the improved vectors could confer a selective advantage to unmodified wild-type virus that they would not have evolved naturally. To test this hypothesis and to address possible environmental concerns of the escape from contained facilities or future field use of such improved vectors, the mutated regions from the *shuff2* and *shuff3* clones and the equivalent region from the progenitor GFP-expressing vector were engineered back into a wild-type TMV cDNA clone. The cell to cell movement rates of virus produced from these clones was assessed by measuring the areas of necrotic

lesions induced on *N. tabacum* cv. Xanthi NN, which has the hypersensitivity resistance gene N. The mean lesion areas at 2 dpi for the *unshuff*, *shuff2* and *shuff3* derivatives were 0.566, 0.581 and 0.573 mm^2, respectively. Analysis of variance gave a value of 0.0571 for the least significant difference at the 5% level, indicating no statistical difference in the rates of cell to cell movement. Further, all the viruses produced systemic symptoms on the tobacco cultivar MD609 at 4 dpi and accumulated to similar levels in systemically infected tissue (data not shown). Thus, the 30K mutations found in the improved shuffled clones are only beneficial in a virus vector background, and compensate for losses in movement brought about by the insertion of a foreign gene. The ineffectivenes of the mutated sequences in a wild-type background is in accord with previous experiments that have shown that the level of 30K expression does not limit the rate of cell to cell movement of wild-type TMV infections.

Conclusions

The purpose of this work was to evolve the TMV 30K gene and its product to compensate for losses in activity induced by foreign gene inclusion. The sequences of plant RNA viruses, such as TMV, do not simply encode proteins, but also have roles in replication, subgenomic mRNA transcription, directing assembly and as translational enhancers. Their base composition may be constrained further by packaging requirements and effects on RNA stability. The successful application of shuffling to improvement of the multifunctional 30K gene encourages us to believe that, with the development of appropriate screens, shuffling other portions of the genome should make improvement of desirable vector traits, such as genetic stability and high levels of foreign gene expression, easier than through rational design. The improved vectors produced in this study will be of utility for *in planta* production of therapeutics due to the increased biomass of infected material and will provide more robust tools for basic plant research.

References

[1] Giddings, G., Allison, G., Brooks, D. & Carter, A. (2000). *Nature Biotechnology* 18, 1151-1155.

[2] Lacomme, C., Pogue, G.P., Wilson, T.M.A. & Santa Cruz, S. (2000). Plant viruses. In *Genetically Engineered Viruses* (Eds. Ring, C.J.A. & Blair, E.D.) BIOS Scientific Publishers, Oxford, UK, pp. 59-105.

[3] Stemmer, W.P.C. (1994). *Nature* 370, 389-391.

[4] Patten, P.A., Howard, R.J. & Stemmer, W.P.C. (1997). *Current Opinions in Biotechnology* 8, 724-733.

[5] Soong, N.-W. et al. (2000). *Nature Genetics* 25, 436-439.

[6] Powell, S.K. et al. (2000). *Nature Biotechnology* 18, 1279-1282.