

DNA Sequencing

Jo Badge

1. Introduction

Although the question of whether viruses can be classified in any way into a phylogenetic organization is still hotly debated, the fact that nucleotide sequence data can provide useful information on the grouping together of certain unknown viruses is becoming recognized. It has been shown that, within the *Potyviridae*, which is the largest and most sequenced group of the plant viruses to date, the coat proteins (CPs) of strains of the same virus have an amino acid sequence identity of 90% or more, compared to that of distinct viruses that have an amino acid sequence identity in the range of 38–71% (**I**). Nucleotide sequencing of viral genomes may therefore prove to be a useful tool in their identification.

The method of sequencing chosen will depend on the type of virus. Viral genomes can be sequenced directly, although many researchers prefer to clone the genome to ensure it is kept in a robust form. Double-stranded DNA viruses can be cloned directly by digestion with appropriate restriction enzymes. However, because the majority of plant viruses are positive-sense single-stranded RNA, another approach is needed. The genome can be reverse-transcribed and the resulting first-strand cDNA made into double-stranded cDNA and cloned directly into a blunt cut vector for sequencing. A library of clones containing viral cDNA inserts can then be sequenced. A miniprep quantity of DNA can be sequenced by double-stranded DNA sequencing. This involves the alkali denaturation of the template, which is then ethanol-precipitated, in order to maintain the denatured state of the DNA. The template is then annealed to a primer, which is a short oligonucleotide sequence complementary to the template or to the multiple-cloning site of the vector used. A DNA polymerase (often *Klenow* or T7) is used to extend the complementary strand of the template DNA from