

Untitled

Schematic

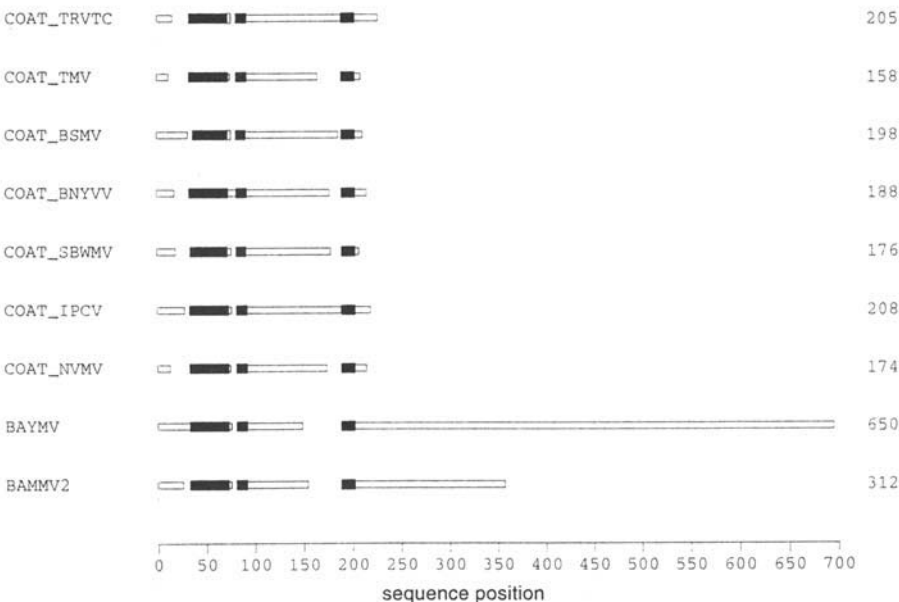


Fig. 6. A scheme of the conserved blocks detected by MACAW program in the capsid proteins of rod-shaped plant viruses and in nonstructural proteins of bymoviruses. Numbers at the end of each line indicate the position of the rightmost amino acid in the protein.

expression-type patterns. The PROSITE database is the largest existing collection of such patterns (36). Such patterns have only limited sensitivity and specificity. As mentioned in the description of the BLA program, the specificity may be enhanced through combining BLAST and PROSITE search, i.e., detecting patterns that are conserved in the alignments produced by BLAST. However, even with this additional filter, PROSITE patterns are generally more useful as a concise presentation of the information contained in a multiple alignment than as a means of detecting new relationships between sequences.

More powerful in detecting new relationships between proteins are methods that make use of the information contained in a multiple alignment to produce a profile or position-dependent weight matrix, which is then utilized for database screening (40–43). The Motif Search Tool (MoST) method (43) converts an alignment block into a position-dependent weight matrix, using the Dirichlet