B

```
COAT_TRVTC      21 WVEKHNWSDILRRLTKIKFALQADRDMIPGI- 8aa-PVDENTRFPS
COAT_TMV        17 WADPIELINLCTNALGNQFQTQQARTVVQRQ- 7aa-SPQVTVRFPd
COAT_BSMV       33 WVHVEAWNKFLDNLRGINFSVASSRSQVAEY- 8aa-PADVDRRFAG
COAT_BNYVV      17 KFMTDRWARVSDVVSVIKQSHAMDLSKAANL-19aa-FVSPMTRFPq
COAT_SBWMV      19 ATHAYIRLSTLmsqieswqatrasvlthlgv-12aa-FFSRTKRFGa
COAT_IPCV       33 WIKSDRWDQLLADLRAVNFEVNSSRSEVASI- 8aa-PAAVSARFPG
COAT_NVMV       15 CATWYKRTDLLDTIRKIKKGDLSITAQVVAA-10aa-EWGHATRFPD
BAYMV           38 FITVRDLHSSLKSALDVDLDTEGGRNAVLDL- 8aa-LVRREKRFPA
BAMMV           11 FISDRFVEFLAMLVLLASALEQdykthdard-13aa-VVARERRFPN

COAT_TRVTC    168   VVQRTFEKEYSLRW
COAT_TMV      139   YNRSSFESSSGLVW
COAT_BSMV     177   YTRKTFERELALEW
COAT_BNYVV    163   WTRDKFEDRFKLPW
COAT_SBWMV    159   YTQDSFEAKYNLKW
COAT_IPCV     179   YDQFLFESTFSVNW
COAT_NVMV     149   YTRTTIENKLGIVW
BAYMV         145   YTRNIFENDNHLVW
BAMMV         126   YNRARFEVNHKVMW
```

Fig. 7. Detection of conserved motifs in protein sequences. **(A)** Results of the database screening, with a position-dependent weight matrix derived from the most conserved block in the alignment of rod-shaped virus capsid proteins using MoST. Sequences added at given iteration are shown in boldface. A segment from bymovirus RNA-2-encoded protein is italicized. **(B)** Conserved motifs in rod-shaped virus capsid proteins and in bymovirus nonstructural proteins detected by MACAW.

ity between these proteins and the products of nonstructural BL1 genes of the two-component geminiviruses ***(48)***. In BLAST searches of the database with individual geminiviruses, this similarity could be detected only when the capsid protein of beet curly top geminivirus was used as a query (data not shown).

3. Concluding Remarks

The field of computer analysis of protein sequences has already become very diverse, and it is impossible to cover even the most widely used and important methods in a brief text. Here we presented only a small set of relatively straightforward, statistically reliable techniques that allow a researcher to rapidly progress from an uncharacterized protein sequence to a meaningful multiple alignment and/or conserved motif(s) useful both for the purpose of classification and for experimental design. What has to be emphasized is the interaction between basic methods for database screening in search of pairwise sequence similarity (e.g., BLAST), multiple alignment construction methods (e.g., MACAW), and motif analysis methods (e.g., MoST), and iterative application of each of these approaches.

Extracting information from the sequences of viral proteins in general, and capsid proteins in particular, is an especially challenging task, because these sequences show conservation only within a few functionally impor-