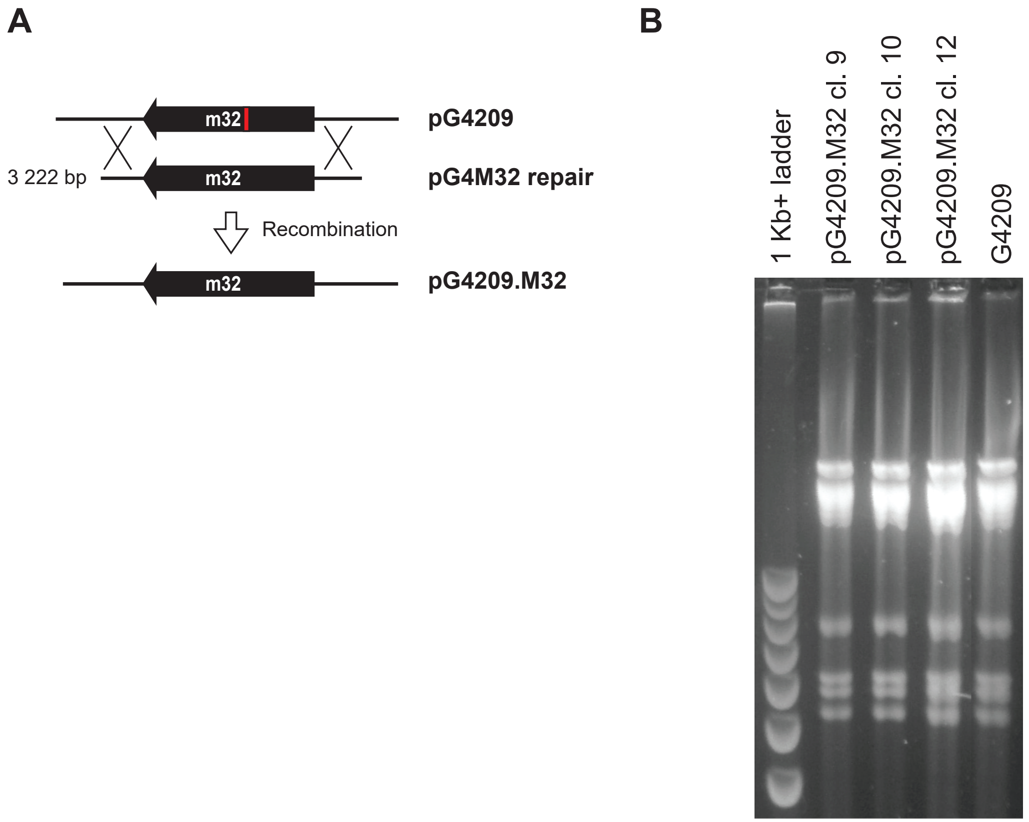
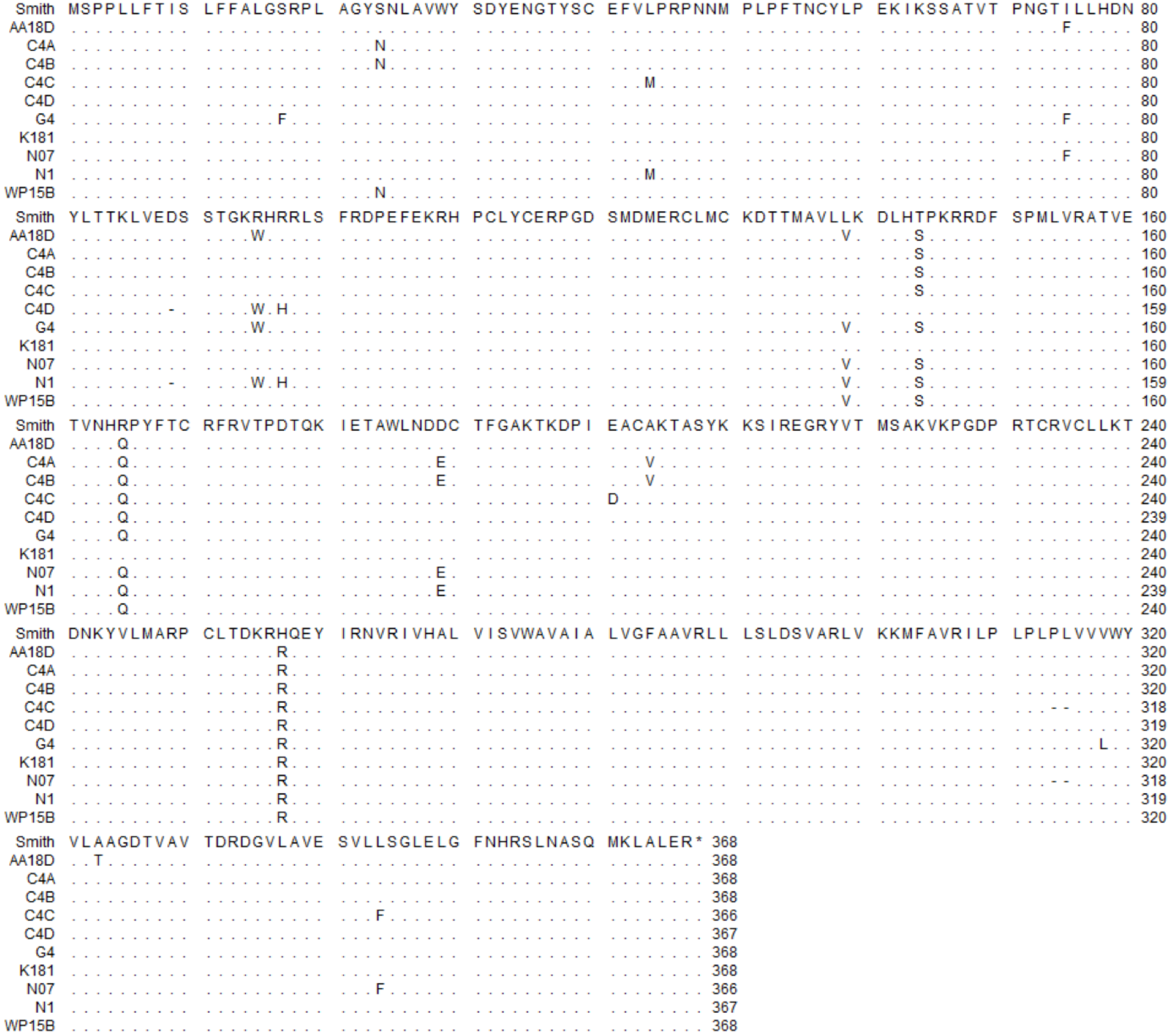


Supplementary Figure 1. Repair of M32. A. Schematic of two-step allele replacement-based repair of the single G to A substitution at 40 711 (indicated by red line). Recombination between repair plasmid pG4M32repair and pG4209 leads to repair of the indicated mutation and production of pg4209.M32. **B**. HindIII RFLP of three repaired clones, Cl9, Cl10 and Cl12 of G4209.M32. Repair was confirmed by sequence analysis.



```
 Smith MSPPLLFTIS LFFALGSRPL AGYSNLAVWY SDYENGTYSC EFVLPRPNNM PLPFTNCYLP EKIKSSATVT PNGTILLHDN 80
 AA18D .......... .......... .......... .......... .......... .......... .......... ....F..... 80
   C4A .......... .......... ...N...... .......... .......... .......... .......... .......... 80
   C4B .......... .......... ...N...... .......... .......... .......... .......... .......... 80
   C4C .......... .......... .......... .......... ...M...... .......... .......... .......... 80
   C4D .......... .......... .......... .......... .......... .......... .......... .......... 80
    G4 .......... ......F... .......... .......... .......... .......... .......... ....F..... 80
  K181 .......... .......... .......... .......... .......... .......... .......... .......... 80
   N07 .......... .......... .......... .......... .......... .......... .......... ....F..... 80
    N1 .......... .......... .......... .......... ...M...... .......... .......... .......... 80
 WP15B .......... .......... ...N...... .......... .......... .......... .......... .......... 80
 Smith YLTTKLVEDS STGKRHRRLS FRDPEFEKRH PCLYCERPGD SMDMERCLMC KDTTMAVLLK DLHTPKRRDF SPMLVRATVE 160
 AA18D .......... ....W..... .......... .......... .......... ........V. ...S...... .......... 160
   C4A .......... .......... .......... .......... .......... .......... ...S...... .......... 160
   C4B .......... .......... .......... .......... .......... .......... ...S...... .......... 160
   C4C .......... .......... .......... .......... .......... .......... ...S...... .......... 160
   C4D ........-. ....W.H... .......... .......... .......... .......... .......... .......... 159
    G4 .......... ....W..... .......... .......... .......... ........V. ...S...... .......... 160
  K181 .......... .......... .......... .......... .......... .......... .......... .......... 160
   N07 .......... .......... .......... .......... .......... ........V. ...S...... .......... 160
    N1 ........-. ....W.H... .......... .......... .......... ........V. ...S...... .......... 159
 WP15B .......... .......... .......... .......... .......... ........V. ...S...... .......... 160
 Smith TVNHRPYFTC RFRVTPDTQK IETAWLNDDC TFGAKTKDPI EACAKTASYK KSIREGRYVT MSAKVKPGDP RTCRVCLLKT 240
 AA18D ....Q..... .......... .......... .......... .......... .......... .......... .......... 240
   C4A ....Q..... .......... ........E. .......... ...V...... .......... .......... .......... 240
   C4B ....Q..... .......... ........E. .......... ...V...... .......... .......... .......... 240
   C4C ....Q..... .......... .......... .......... D......... .......... .......... .......... 240
   C4D ....Q..... .......... .......... .......... .......... .......... .......... .......... 239
    G4 ....Q..... .......... .......... .......... .......... .......... .......... .......... 240
  K181 .......... .......... .......... .......... .......... .......... .......... .......... 240
   N07 ....Q..... .......... ........E. .......... .......... .......... .......... .......... 240
    N1 ....Q..... .......... ........E. .......... .......... .......... .......... .......... 239
 WP15B ....Q..... .......... .......... .......... .......... .......... .......... .......... 240
 Smith DNKYVLMARP CLTDKRHQEY IRNVRIVHAL VISVWAVAIA LVGFAAVRLL LSLDSVARLV KKMFAVRILP LPLPLVVVWY 320
 AA18D .......... ......R... .......... .......... .......... .......... .......... .......... 320
   C4A .......... ......R... .......... .......... .......... .......... .......... .......... 320
   C4B .......... ......R... .......... .......... .......... .......... .......... .......... 320
   C4C .......... ......R... .......... .......... .......... .......... .......... ...--..... 318
   C4D .......... ......R... .......... .......... .......... .......... .......... .......... 319
    G4 .......... ......R... .......... .......... .......... .......... .......... .......L.. 320
  K181 .......... ......R... .......... .......... .......... .......... .......... .......... 320
   N07 .......... ......R... .......... .......... .......... .......... .......... ...--..... 318
    N1 .......... ......R... .......... .......... .......... .......... .......... .......... 319
 WP15B .......... ......R... .......... .......... .......... .......... .......... .......... 320
 Smith VLAAGDTVAV TDRDGVLAVE SVLLSGLELG FNHRSLNASQ MKLALER* 368
 AA18D ..T....... .......... .......... .......... ........ 368
   C4A .......... .......... .......... .......... ........ 368
   C4B .......... .......... .......... .......... ........ 368
   C4C .......... .......... ...F...... .......... ........ 366
   C4D .......... .......... .......... .......... ........ 367
    G4 .......... .......... .......... .......... ........ 368
  K181 .......... .......... .......... .......... ........ 368
   N07 .......... .......... ...F...... .......... ........ 366
    N1 .......... .......... .......... .......... ........ 367
 WP15B .......... .......... .......... .......... ........ 368
```

Supplementary Figure 2. Conservation of m133 sequence in MCMV. The annotated m133 ORF is present in all MCMV strains and is predicted to encode a protein of between 366 and 368 AA. Sequences are highly conserved at the AA level and range from 100% sequence identity (C4A versus C4B) and 97.01% sequence similarity (C4C versus G4).

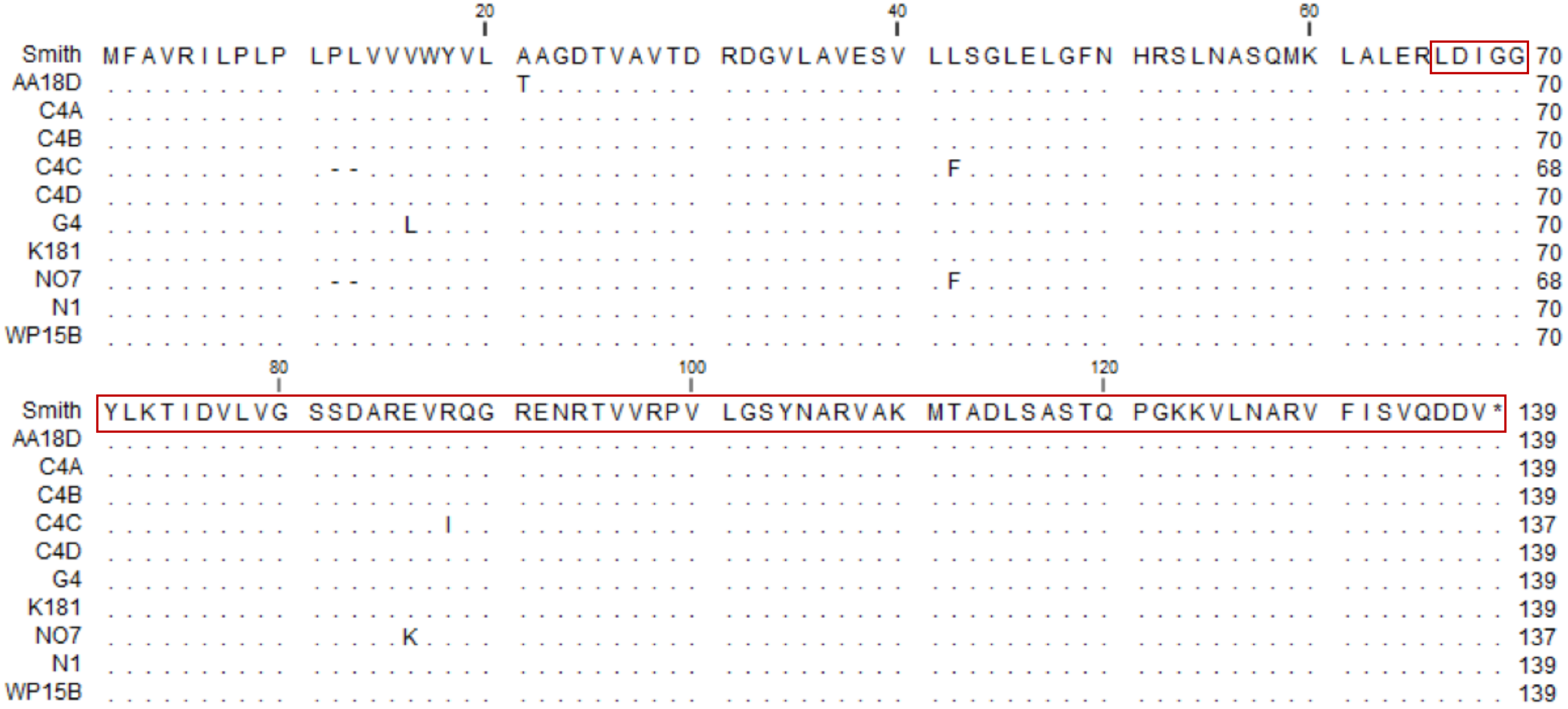
```
Smith MSDVQGHDNG GLTKGFTHAQ TARFFAYVGP GHGGDRQSPA LFHVPVSRNS GHPEDRDGVV ERRLHLRRQN ERSDRGLRQN GQLQEIHTRG 90
AA18D .......... ..S....L.. .......... .......... .......... .......... ....Y..... .......... .......... 90
  C4A .......... .......L.. .......... .......... .......... .......... ...V...... .......... .......... 90
  C4B .......... .......L.. .......... .......... .......... R......... ...V...... .......... .......... 90
  C4C .......... .......L.. .......... .......... .......... .......... ....Y..... .....C.... ......D... 90
  C4D .......... .......... .......... .......... .......... .......... K...Y..... .......... .......... 90
   G4 .......... ..S....L.. .......... .......... .......... .......... ....Y..... .......... .......... 90
 K181 .......... .......... .......... .......... .......... .......... .......... .......... ......D... 90
  NO7 .......... ..S....L.. .......... .......... .......... .......... ...V...... .......... ......D... 90
   N1 .......... ..S....L.. .......... .......... .......... .......... ...V...... .......... ......D... 90
WP15B .......... ..S....L.. .......... .......... .......... .......... ....Y..... .......... .......... 90

Smith KIRDHERQGE TRRPAHLPRM PAEDGQQVRA DGASMPDGQA APGIHPKRAH RARVGDQRVG GRDRPGGLRR RAPAIKPGQR RPAGEKNVRG 180
AA18D .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
  C4A .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
  C4B .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
  C4C .......... .....Y...V .......... ...P...... .......... .......... .......... .......... .......... 180
  C4D .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
   G4 .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
 K181 .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
  NO7 .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
   N1 .......... .........V .......... .......... .......... .......... .......... .......... .......... 180
WP15B .......... .........V .......... .......... .......... .......... .......... .......... .......... 180

Smith TYTAVTVTAG RGVVRARRRR HGRRDGPRRG AGRGERPALR ARAGVQPSIP ERVPNETGAG EVALPRNGRV PRRALSLTNH PPALTCARFC 270
AA18D .......... ......H... .......... .......... .......... ........T. .......... .......SD- ---------- 259
  C4A .......... .......... .......... .......... .......... .......... .......... .......... .......... 270
  C4B .......... .......... .......... .......... .......... .......... .......... .......... .......... 270
  C4C ......--.. .........* ---------- ---------- ---------- ---------- ---------- ---------- ---------- 198
  C4D .......... .......... .......... .......... .......... .......... .......... .......... .......... 270
   G4 .......... .V........ .......... .......... .......... .......... .......... .......... .......... 270
 K181 .......... .......... .......... .......... .......... .......... .......... .......... .......... 270
  NO7 ......--.. .......... .......... .......V.. .......... .......... .......... ......D--- ---------- 255
   N1 .......... .......... .......... .......... .......... .......... .......... .........- .......... 269
WP15B .......... .......... .......... .......... .......... .......... .......... .......... .......... 270

Smith LSGWI 275
AA18D ----- 259
  C4A ..... 275
  C4B ..... 275
  C4C ----- 198
  C4D ..... 275
   G4 ..... 275
 K181 ..... 275
  NO7 ----- 255
   N1 ..... 274
WP15B ..... 275
```

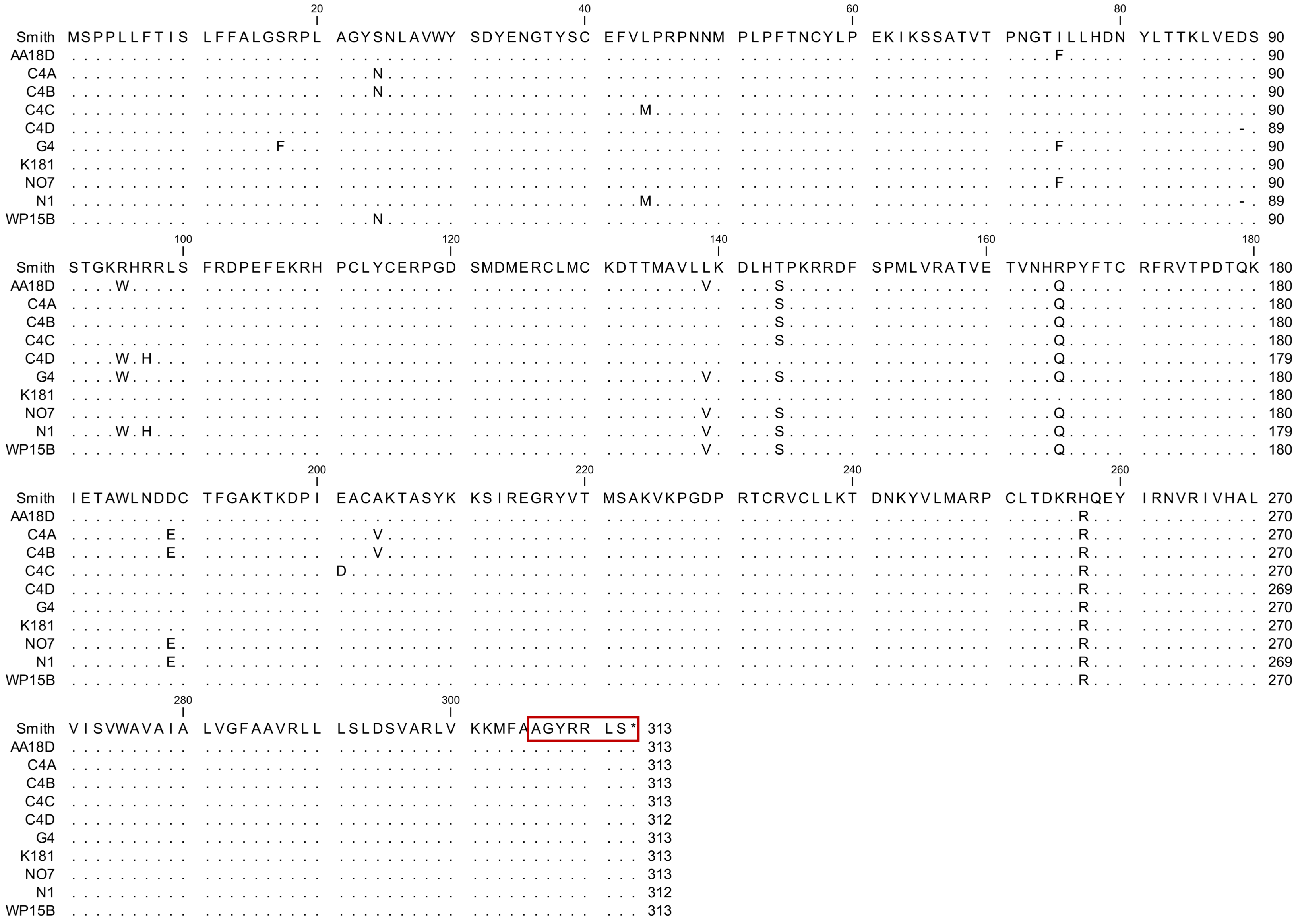
Supplementary Figure 3. Conservation of m132.1 sequence in MCMV. The annotated gene m132.1 is present in all MCMV strains. Across the sequenced MCMV strains, m132.1 shows sequence length polymorphism with the smallest protein, encoded by C4C, predicted to be 197 AA. Most m132.1 ORFs are predicted to encode a protein of 275 AA. Sequence conservation is most pronounced across the amino terminus of the protein as length variation is due to truncation of the carboxy terminus in AA18D, C4C and NO7. Overall AA sequence similarity remains high from 99.64% (C4A versus C4B and G4 versus WP15B) to 69.45% (C4B versus C4C and C4C versus Smith).



```
                                    20                  40                  60
Smith   MFAVRILPLP LPLVVVWYVL AAGDTVAVTD RDGVLAVESV LLSGLELGFN HRSLNASQMK LALERLDIGG 70
AA18D   .......... .......... T......... .......... .......... .......... .......... 70
C4A     .......... .......... .......... .......... .......... .......... .......... 70
C4B     .......... .......... .......... .......... .......... .......... .......... 70
C4C     .......... .--....... .......... .......... .F........ .......... .......... 68
C4D     .......... .......... .......... .......... .......... .......... .......... 70
G4      .......... .....L.... .......... .......... .......... .......... .......... 70
K181    .......... .......... .......... .......... .......... .......... .......... 70
NO7     .......... .--....... .......... .......... .F........ .......... .......... 68
N1      .......... .......... .......... .......... .......... .......... .......... 70
WP15B   .......... .......... .......... .......... .......... .......... .......... 70

                 80                  100                 120
Smith   YLKTIDVLVG SSDAREVRQG RENRTVVRPV LGSYNARVAK MTADLSASTQ PGKKVLNARV FISVQDDV* 139
AA18D   .......... .......... .......... .......... .......... .......... ......... 139
C4A     .......... .......... .......... .......... .......... .......... ......... 139
C4B     .......... .......... .......... .......... .......... .......... ......... 139
C4C     .......... .......I.. .......... .......... .......... .......... ......... 137
C4D     .......... .......... .......... .......... .......... .......... ......... 139
G4      .......... .......... .......... .......... .......... .......... ......... 139
K181    .......... .......... .......... .......... .......... .......... ......... 139
NO7     .......... .....K.... .......... .......... .......... .......... ......... 137
N1      .......... .......... .......... .......... .......... .......... ......... 139
WP15B   .......... .......... .......... .......... .......... .......... ......... 139
```

Supplementary Figure 4. Conservation of sgg1.1 sequence in MCMV. The novel two exon sgg1.1 gene is conserved across all MCMV strains sequenced to date, encoding a predicted protein of between 136 to 138 AA. Sequences are highly conserved at the AA level and range from 100% sequence identity (Smith, WP15B, N1, K181, C4D, C4B and C4A) to 96% similarity (e.g. NO7 versus AA18D). Red box denotes boundary of exon 2.



```
                                20                   40                   60                   80
 Smith  MSPPLLFTIS LFFALGSRPL AGYSNLAVWY SDYENGTYSC EFVLPRPNNM PLPFTNCYLP EKIKSSATVT PNGTILLHDN YLTTKLVEDS 90
 AA18D  .......... .......... .......... .......... .......... .......... .......... ....F..... .......... 90
   C4A  .......... .......... ...N...... .......... .......... .......... .......... .......... .......... 90
   C4B  .......... .......... ...N...... .......... .......... .......... .......... .......... .......... 90
   C4C  .......... .......... .......... .......... ...M...... .......... .......... .......... .......... 90
   C4D  .......... .......... .......... .......... .......... .......... .......... .......... ........-. 89
    G4  .......... ......F... .......... .......... .......... .......... .......... ....F..... .......... 90
  K181  .......... .......... .......... .......... .......... .......... .......... .......... .......... 90
   NO7  .......... .......... .......... .......... .......... .......... .......... ....F..... .......... 90
    N1  .......... .......... .......... .......... ...M...... .......... .......... .......... ........-. 89
 WP15B  .......... .......... ...N...... .......... .......... .......... .......... .......... .......... 90
                 100                  120                  140                  160                  180
 Smith  STGKRHRRLS FRDPEFEKRH PCLYCERPGD SMDMERCLMC KDTTMAVLLK DLHTPKRRDF SPMLVRATVE TVNHRPYFTC RFRVTPDTQK 180
 AA18D  ...W...... .......... .......... .......... ........V. ...S...... .......... ....Q..... .......... 180
   C4A  .......... .......... .......... .......... .......... ...S...... .......... ....Q..... .......... 180
   C4B  .......... .......... .......... .......... .......... ...S...... .......... ....Q..... .......... 180
   C4C  .......... .......... .......... .......... .......... ...S...... .......... ....Q..... .......... 180
   C4D  ...W.H.... .......... .......... .......... .......... .......... .......... ....Q..... .......... 179
    G4  ...W...... .......... .......... .......... ........V. ...S...... .......... ....Q..... .......... 180
  K181  .......... .......... .......... .......... .......... .......... .......... .......... .......... 180
   NO7  .......... .......... .......... .......... ........V. ...S...... .......... ....Q..... .......... 180
    N1  ...W.H.... .......... .......... .......... ........V. ...S...... .......... ....Q..... .......... 179
 WP15B  .......... .......... .......... .......... ........V. ...S...... .......... ....Q..... .......... 180
                                200                  220                  240                  260
 Smith  IETAWLNDDC TFGAKTKDPI EACAKTASYK KSIREGRYVT MSAKVKPGDP RTCRVCLLKT DNKYVLMARP CLTDKRHQEY IRNVRIVHAL 270
 AA18D  .......... .......... .......... .......... .......... .......... .......... ......R... .......... 270
   C4A  ........E. .......... ...V...... .......... .......... .......... .......... ......R... .......... 270
   C4B  ........E. .......... ...V...... .......... .......... .......... .......... ......R... .......... 270
   C4C  .......... .......... D......... .......... .......... .......... .......... ......R... .......... 270
   C4D  .......... .......... .......... .......... .......... .......... .......... ......R... .......... 269
    G4  .......... .......... .......... .......... .......... .......... .......... ......R... .......... 270
  K181  .......... .......... .......... .......... .......... .......... .......... ......R... .......... 270
   NO7  ........E. .......... .......... .......... .......... .......... .......... ......R... .......... 270
    N1  ........E. .......... .......... .......... .......... .......... .......... ......R... .......... 269
 WP15B  .......... .......... .......... .......... .......... .......... .......... ......R... .......... 270
                 280                  300
 Smith  VISVWAVAIA LVGFAAVRLL LSLDSVARLV KKMFAAGYRR LS* 313
 AA18D  .......... .......... .......... .......... ... 313
   C4A  .......... .......... .......... .......... ... 313
   C4B  .......... .......... .......... .......... ... 313
   C4C  .......... .......... .......... .......... ... 313
   C4D  .......... .......... .......... .......... ... 312
    G4  .......... .......... .......... .......... ... 313
  K181  .......... .......... .......... .......... ... 313
   NO7  .......... .......... .......... .......... ... 313
    N1  .......... .......... .......... .......... ... 312
 WP15B  .......... .......... .......... .......... ... 313
```

Supplementary Figure 5. Conservation of sgg1 sequences in MCMV. The canonical two exon sgg1 gene is conserved across all MCMV strains sequenced to date, encoding a predicted protein of 311 or 312 AA. Sequences are highly conserved at the AA level and range from 100% sequence identity (C4A versus C4B) and 97.12% sequence similarity (N1 versus Smith). Red box denotes boundary of exon 2.