The murine norovirus (MNV) core subgenomic RNA promoter consists of a stable stem-loop that can direct accurate initiation of RNA synthesis

Muhammad Amir Yunus[a,b*], Xiaoyan Lin[c*], Dalan Bailey[a,d], Ioannis Karakasiliotis[a,e], Yasmin Chaudhry[f], Surender Vashist[f], Guo Zhang[a], Lucy Thorne[f], C. Cheng Kao[c] and Ian Goodfellow[a,f#]

Section of Virology, Faculty of Medicine, Imperial College London, London, UK[a]; Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, Pulau Pinang, Malaysia[b]; Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, USA[c]; School of Immunity and Infection, University of Birmingham, Birmingham, UK[d]; Laboratory of Molecular Virology, Hellenic Pasteur Institute, Athens, Greece[e]; Division of Virology, Department of Pathology, University of Cambridge, Addenbrookes Hospital, Cambridge, UK[f]

Running title: Norovirus subgenomic RNA synthesis

*These authors contributed equally to the work

Address correspondence to: Ian Goodfellow, ig299@cam.ac.uk

Word count abstract: 214

Word count text: 5892

Abstract

All members of the *Caliciviridae* family of viruses produce a subgenomic RNA during infection. The subgenomic RNA typically encodes only the major and minor capsid proteins, but in murine norovirus (MNV), the subgenomic RNA also encodes the VF1 protein that functions to suppress host innate immune responses. To date, the mechanism of norovirus subgenomic RNA synthesis has not been characterized. We have previously described the presence of an evolutionarily conserved RNA stem-loop structure on the negative-sense RNA, the complementary sequence of which codes for the viral RNA-dependent RNA polymerase (NS7). The conserved stem-loop is positioned 6 nucleotides 3' of the start site of the subgenomic RNA in all caliciviruses. We demonstrate that the conserved stem-loop is essential for MNV viability. Mutant MNV RNAs with substitutions in the stem-loop replicated poorly until they accumulated mutations that revert to restore the stem-loop sequence and/or structure. The stem-loop sequence functions in a non-coding context as it was possible to restore the replication of a MNV mutant by introducing an additional copy of the stem-loop between the NS7 and VP1 coding regions. Finally, *in vitro* biochemical data would suggest that stem-loop sequence is sufficient for the initiation of viral RNA synthesis by the recombinant MNV RdRp, confirming that the stem-loop forms the core of the norovirus subgenomic promoter.

Importance

Noroviruses are a significant cause of viral gastroenteritis and it is important to understand the mechanism of norovirus RNA synthesis. Herein we describe the identification of an RNA stem-loop structure that functions as the core of the norovirus subgenomic RNA promoter in cells and in vitro. This work provides new insights into the molecular mechanisms of norovirus RNA synthesis and the sequences that determine the recognition of viral RNA by the RNA dependent RNA polymerase.

55 **Introduction**

56 Noroviruses, members of the *Caliciviridae* family of small positive-sense
57 RNA viruses are a major cause of viral gastroenteritis in the developed world (1,
58 2). Despite their impact, noroviruses remain poorly characterized as an in-depth
59 understanding of the molecular mechanisms of norovirus genome translation and
60 replication have been hampered by the inability to culture human norovirus (3).
61 Murine norovirus (MNV) is a model system with which to understand the
62 norovirus life cycle (4) as it replicates in cultured cells (5) and a number of
63 tractable reverse genetics systems are available (6-9). Studies of MNV have
64 therefore led to a number of significant advances in the understanding of the
65 molecular mechanism of norovirus genome translation and replication (Reviewed
66 in 10).

67

68 The norovirus RNA genome typically encodes three open reading frames
69 (11). MNV also encodes an additional ORF in the region overlapping the VP1
70 coding region (12, 13) (Fig. 1A). All members of the *Caliciviridae* family of small
71 positive-sense RNA viruses synthesize a shorter than genome length,
72 subgenomic, RNA (sgRNA) which directs the translation of the major and minor
73 structural proteins, VP1 and VP2 respectively (10). The MNV sgRNA also
74 encodes the VF1 protein, an antagonist of the innate immune response (12).
75 Production of a sgRNA during the viral life cycle is a common feature of many
76 positive-sense RNA viruses (14) and it is often used to control the expression of
77 the viral proteins.

The mechanism of calicivirus sgRNA synthesis is not well-understood. Initial *in vitro* biochemical approaches using rabbit hemorrhagic disease virus (RHDV) suggested that a sequence upstream of the first nucleotide of the subgenomic RNA is required for sgRNA synthesis (15). This sequence acts in the complement of the RHDV genomic RNA. Bioinformatic analysis also confirmed the presence of evolutionarily conserved and stable RNA stem-loop structures upstream of the start of the sgRNA of caliciviruses (16). Within the region coding for the viral RNA-dependent RNA polymerase (NS7, RdRp) in all caliciviruses analyzed, a ~24-50 nt stem-loop structure could be predicted on the negative-sense genomic RNA, precisely 6 nt 3' of the sgRNA start site. This stem-loop structure in the MNV minus-strand RNA will be referred to as Sla5045 (for stem-loop anti-sense 5045). We previously observed that the stable structure of Sla5045 is essential for the recovery of viable MNV in cultured cells (16). In the current study we present results consistent with the hypothesis that Sla5045 forms the core of the norovirus sgRNA promoter.

Materials and Methods

Cell lines and plasmid constructs. The murine leukemia macrophage cell line (RAW264.7) was grown and maintained in Dulbecco modified Eagle medium (DMEM; Gibco) with 10% (v/v) fetal calf serum (FCS), penicillin (100 SI units/ml) and streptomycin (100 µg/ml) and 10 mM HEPES buffer (pH 7.6). Baby hamster kidney cells (BHK-21) engineered to express T7 RNA polymerase (BSR-T7 cells,

101 obtained from Karl-Klaus Conzelmann, Ludwig Maximillian University, Munich,
102 Germany) were maintained in DMEM containing 10% FCS, penicillin (100 SI
103 units/ml), streptomycin (100 µg/ml) and 1.0 mg/ml geneticin (G418). All cells
104 were maintained at 37°C with 10% CO_2.

105 The full length MNV-1 cDNA clone, pT7:MNV3'Rz, contains the MNV-1
106 genome under the control of truncated T7 polymerase promoter (7). Mutant
107 derivatives of pT7:MNV3'RZ were made as previously describe (7, 16). These
108 include: (i) a frameshift in the NS7 region of ORF1 (pT7:MNV 3'Rz F/S), (ii) the
109 m53 mutations to destabilize the Sla5045 (pT7:MNVm53 3'Rz) and (iii) the m53r
110 mutations to restore the Sla5045 (pT7:MNVm53r 3'Rz). Other full length MNV-1
111 cDNA clones carrying m53 mutant suppressors (m53SupA, m53SupG and
112 m53SupH) in m53 backbone construct were generated by overlapping PCR
113 mutagenesis using pT7:MNVm53 3'Rz as a template. All primer sequences used
114 in this work and protocols will be made available upon request. Mutant
115 suppressors were also engineered into pT7:MNV3'Rz, producing WTSupA,
116 WTSupB and WTSupC cDNA constructs. Additional synonymous mutations at
117 position 4922 in the WT and m53 cDNA constructs with adenyltate and guanylate
118 substitutions were also generated by overlapping mutagenesis PCR. The
119 luciferase reporter-expressing MNV replicons used in this study were as reported
120 previously (17). Briefly, the luciferase gene was fused to VP2 in the WT, POL^{FS},
121 m53 and m53r constructs and separated by the foot and mouth disease virus 2A
122 protease (FMDV 2A). The expression of luciferase reporter gene in these
123 constructs is under the control of the MNV TURBS sequence (18, 19).

124 Translation of the luciferase reporter protein occurs as a VP2 fusion protein and
125 is co-translationally cleaved at the specific FMDV 2A cleavage site to release
126 both luciferase-2A and VP2.

127 Sla5045Dup contains a second copy of the WT or m53 stem-loop (Sla2)
128 inserted outside of the NS7 coding region in the intergenic region between Orf1
129 and Orf2 was used to generate constructs WT/WT, WT/m53, m53/WT and
130 m53/m53 by overlapping PCR mutagenesis. A second panel of Sla5045Dup
131 mutant constructs was also generated based on the m53/WT construct as
132 detailed in the text.

133

134 **Reverse genetics and virus yield determination**. Recoveries of full-length
135 infectious MNV-1 cDNA clones were performed using the established reverse
136 genetics system (7). Typically, BSR-T7 cells were infected with fowlpox virus
137 expressing T7 RNA polymerase at a multiplicitiy of infection (MOI) of 0.5 plaque
138 forming units (PFU) per cell before transfection with 1 µg of cDNA was carried
139 out using Lipofectamine 2000 (Invitrogen) according to the manufacturer's
140 instructions. Note that the levels of T7 RNA polymerase normally expressed in
141 BSR-T7 cells are not sufficient to drive MNV recovery. Where indicated, the
142 RNA-mediated reverse genetics system was used. In this case, RNA transcripts
143 were produced using *in vitro* transcription reactions which consisted of 200 mM
144 HEPES pH7.5, 32 mM magnesium acetate, 40 mM DTT, 2 mM spermidine, 7.5
145 mM of each NTP (ATP, UTP, GTP and CTP), 40 unit of RNAse inhibitor
146 (Promega), 250 ng of linearized DNA template and 50 µg/ml of T7 RNA

147 polymerase. The reactions were incubated at 37°C for 2-7 h, treated with DNase
148 I (New England Biolabs) and the RNA was purified by precipitation using lithium
149 chloride and resuspended in RNA storage solution (Ambion) (9). Prior to
150 transfection into BSR-T7 cells, post-transcriptional enzymatic capping was
151 performed on the purified RNA transcripts using a ScriptCap system (Epicentre),
152 according to manufacturers instructions. Typically one microgram of capped RNA
153 was transfected using Lipofectamine 2000 according to the manufacturers'
154 instructions. The yield of infectious was determined 24 h post transfection of
155 cDNA or capped RNA, using 50% tissue culture infectious dose ($TCID_{50}$). Note
156 that BSR-T7 cells, a BHK cell derivative can support only a single cycle of virus
157 replication, possible due to the lack of a suitable receptor for virus reinfection.
158
159 **RNA purification and genome copy determination**. All RNA purifications from
160 infected and transfected cells were performed using the GenElute Mammalian
161 Total RNA Miniprep Kit (Sigma Aldrich) according to the manufacturer's
162 instructions. RNA was quantified by spectrophotometry and qualitatively
163 inspected by on agarose gels stained with ethidium bromide.
164
165 The expression level of genomic RNA for WT and m53SupA MNV were
166 quantified using one step RT-qPCR. Viral genome copies were determined using
167 the MESA Blue qPCR Assay (Eurogentech) performed in parallel with control
168 RNAs of known concentration using the primers 404F
169 (GGAGCCTGTGATCGGCTCTATCTTGGAGCAGG) and 491R

170 (GCCTGGCAGACCGCAGCACTGGGGTTGTTGACC) to amplify the viral
171 genomic RNA.
172 Where described, strand-specific RT-qPCR was performed using tagged
173 RT primers to detect either positive or negative-sense genomic RNA as
174 previously described (20). In addition, a set of primers was designed to amplify
175 nucleotides 5473-5422 to facilitate the simultaneous quantification of both
176 genomic and subgenomic RNAs. Control RNAs for the genomic or subgenomic
177 RNAs, of positive or negative-sense polarity, were generated by *in vitro*
178 transcription and used as standards to facilitate quantification.
179
180 **Western blotting**. Samples for western blot analysis were obtained by lysing
181 cells in radioimmunoprecipitation assay buffer (RIPA - 50 mM Tris–HCl pH 8.0,
182 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 and 0.1% SDS). Protein
183 concentration in the lysates were quantified using the bicinchoninic acid (BCA)
184 protein assay (Pierce). Equal amounts of protein per samples were separated by
185 SDS-PAGE were transferred to a polyvinylidene fluoride (PVDF) Immobilon-P
186 transfer membrane (Millipore) for detection using Enhance Chemiluminescence
187 (ECL) reagents by semi-dry transfer. The blocking buffer (PBST) contained 0.1%
188 Tween 20 in PBS containing 5% w/v milk powder and the rabbit anti-MNV NS7 or
189 VP1 serum. Antibody binding was detected using a secondary, horseradish
190 peroxidase-conjugated secondary antibody (source). The signal was detected by
191 enhanced chemiluminescence kit (ECL; Amersham Biosciences).
192

193 **Luciferase reporter assay**. Cell lysates used to detect luciferase activity were
194 prepared by lysing the transfected cells in 1 X Reporter Lysis Buffer (Promega)
195 according to the manufacturers instructions. The protein concentration in each
196 lysates was standardized and total protein (30 μg) from each cell lysate was
197 analyzed. The luciferase assays used a 1:500 dilution of coelentrazine
198 (Promega) in PBS as a substrate. Luciferase activity was measured in an auto-
199 injector luminometer (FLUOstar Omega, BMG Labtech).
200
201 **Northern blot assay**. Total RNAs were purified from infected cells and
202 denatured by glyoxylation according to the manufacturers instructions (Ambion).
203 Denatured RNAs along with the relevant control RNA generated by *in vitro*
204 transcription to produce full-length MNV genomic and subgenomic RNAs of both
205 polarities, were separated by agarose electrophoresis. Denatured RNAs were
206 transferred to Hybond N+ nylon membranes using mild alkaline conditions and
207 the viral RNA detected using the NorthernMax northern blotting system (Life
208 Technologies). RNA probes used to detect the positive and negative-sense viral
209 RNAs were generated by incorporating ^{32}P-UTP or CTP using the Maxiscript *in*
210 *vitro* labeling system, according to the manufacturer's instructions (Ambion). The
211 RNA probe used to detect the negative-sense viral RNA consisted of nts 1-200 of
212 the positive sense sgRNA, whereas the positive sense viral RNA was detected
213 using a probe complementary to nts 7161-7382.
214

215 **5' rapid amplification of cDNA ends (RACE)**. Total RNA preparations were
216 subjected to cDNA synthesis using a reverse primer complementary to
217 sequences ~700 bases downstream (on positive polarity) of Sla5045. The
218 resulting cDNAs were treated with ribonuclease to degrade the viral RNA
219 template prior. The purified cDNAs were then used as a template for RACE
220 assay using 5' RACE according to the manufacturer's instructions (Life
221 Technologies). Purified cDNAs were first subjected to the 3' end poly(C) tailing
222 reaction catalyzed by terminal transferase followed by PCR using an anchor
223 primer that anneals to the poly(C) 3'end of cDNA and the MNV specific reverse
224 primer. Sequences of the PCR products were determined using the BigDye
225 terminator kit and analyzed using the Vector NTI software (Invitrogen).
226
227 **Primer extension analysis.** Primer extension analysis was performed on RNA
228 isolated from either infected cells or cells transfected with *in vitro* transcribed
229 capped RNAs generated from MNV cDNA containing clones. The ^{33}P γ-ATP
230 labeled primer 5129R (GGTTGAATGGGGACGGCCTGTTCAACGG) was
231 annealed to 10μg of total RNA was and extended using Superscript III according
232 to the manufacturers' instructions (Life Technologies). Reactions were stopped
233 by the addition of formamide containing dye followed by separation on 6% urea-
234 PAGE. The extended products were visualized following exposure to
235 autoradiography film and compared to samples isolated form infected cells and
236 control RNA produced by *in vitro* transcription.
237

238 **Recombinant RdRp expression and purification.** The cDNA encoding MNV
239 NS7 (nts 3537 to 5069) was subcloned in pET-15b. The construct contains an N-
240 terminal six-histidine tag to facilitate protein purification. The plasmid expressing
241 MNV NS7 was transformed into *E. coli* Rosetta cells, which were grown at 37°C
242 until the optical density at OD_{600} was between 0.6 and 0.8.
243 Isoproplythiogalactoside was added to a final concentration of 1 mM induce
244 protein expression for 16-18 h at 16°C. The cells were harvested by
245 centrifugation and resuspended in 40 ml lysis buffer (100 mM TrisCl, pH7.9, 300
246 mM NaCl, 10% glycerol, 15 mM imidazole, 5 mM β-mercaptoethanol, 0.1% Triton
247 X-100, protease inhibitor cocktail (Sigma). The recombinant protein was purified
248 by Ni-NTA (Qiagen) according to the manufacturer's instructions. The elute
249 protein was dialyzed in buffer containing 100 mM TrisCl, pH 7.9, 300 mM NaCl,
250 10% glycerol, 5 mM β-mercaptoethanol and stored in aliquots at -80°C.
251 Recombinant GII.4 NS7 was expressed and purified as previously
252 described (17). Briefly, *E. coli* Top 10 cells transformed with pBAD GII.4NS7
253 plasmid were cultured at 37°C until an OD_{600} of around 0.6 was reached. L-
254 arabinose was added to a final concentration of 0.02% to induce protein
255 expression. The cells were grown at 37°C for 5 h and then harvested and
256 suspended in lysis buffer (50 mM HEPES pH 7.9, 150 mM NaCl, 10% glycerol,
257 10 mM Imidazole, proteinase inhibitor cocktail (Sigma)). The protein was purified
258 by Talon metal affinity resin (Clontech Laboratories) and the elute protein
259 dialyzed in buffer containing 100 mM TrisCl, pH 7.9, 300 mM NaCl, 10% glycerol,
260 5 mM β-mercaptoethanol and stored in aliquots at -80°C. All recombinant

261 proteins were quantified by using the Bradford method followed SDS-PAGE and
262 staining with Coomassie brilliant blue to examine their purity.
263
264 **RNA synthesis assay.** RNA synthesis by the recombinant RdRps was assayed
265 as previously described (21) with minor modifications. The 20 µl reaction
266 contained 20 mM sodium glutamate (pH 7.4), 12.5 mM dithiothreitol, 4 mM
267 $MgCl_2$, 1 mM $MnCl_2$, 0.5% Triton X-100 (v/v), 0.2 mM GTP, 0.1 mM ATP and
268 UTP, 3.3 nM α-^{32}P-CTP (MP Biomedicals), 50 nM MNV RNA template, 25 nM
269 recombinant RdRp. The reaction was incubated at 30°C for 2 h, and then
270 stopped by the addition of EDTA (pH 8.0) to final concentration of 10 mM. The
271 RdRp products were subjected to electrophoresis in 24% polyacrylamide gel
272 containing 7.5 M urea and 0.5X TBE. The radiolabeled RNA products were
273 visualized and quantified by using a PhosphoImager (Typhoon 9210; Amersham
274 Biosciences) and ImageQuant software.
275
276 **Results**
277 **The stability of Sla5045 contributes to norovirus infectivity and**
278 **subgenomic RNA synthesis.**
279 Sequence analyses of published noroviruses genomes indicate that the
280 stem-loop sequence of Sla5045 is highly conserved (data not shown and Fig.
281 1B). Several positions in the Genogroup II (GII) or V (GV), representing human
282 and murine noroviruses respectively, appear to show some variation in
283 sequence, although in all cases a complex stem-loop structure is retained (Fig.

284 1B). We initially sought to test the functional significance of the stability of
285 Sla5045. Three nucleotide substitutions were introduced to the MNV genome,
286 resulting in a construct named m53 (Fig. 1C). M53 was previously shown to have
287 a severely reduced recovery of infectious virus (16). However, it was not
288 determined whether the inability to produce detectable virus was due to a defect
289 in RNA encapsidation or sgRNA production, or both. To further validate the role
290 of Sla5045 in the norovirus life cycle and to examine whether the m53 mutations
291 affect viral RNA synthesis, we inserted the previously described m53
292 substitutions into the MNV luciferase replicon, Mflc-R (17). The replicon consists
293 of a full-length MNV genome expressing a *Renilla* luciferase-FMDV 2A peptide-
294 VP2 fusion protein in the VP2 coding region. Transfection of BSRT7 cells with
295 enzymatically capped, *in vitro* transcribed RNA from the Mflc-R cDNA leads to
296 luciferase expression in the absence of infectious virus production. As luciferase
297 is expressed from ORF3 and occurs via a conserved termination-reinitiation
298 mechanism of translation at the translational termination signal of ORF2 (18, 19),
299 it functions as a highly sensitive marker for the production of sgRNA. The
300 introduction of the m53 mutations resulted in luciferase levels similar to a
301 replication-defective replicon containing a frame-shift in the NS7 coding region
302 (Fig. 1D). A construct containing compensatory mutations to m53, named m53r
303 (Fig. 1C), that restored the stability of Sla5045 was found to restore luciferase
304 expression levels to nearly the level of the WT MNV replicon (Fig. 1D). The
305 introduction of m53r mutations also resulted in the recovery of infectious virus
306 (Fig. 1E). These data are consistent with the hypothesis that Sla5045 plays a role

307 in viral RNA synthesis and does not simply affect a late stage in the viral life
308 cycle such as encapsidation.
309
310 We analyzed the production of sgRNA in cells transfected with capped
311 RNAs of m53 and m53r using a primer extension assay. A radiolabelled primer
312 complementary to nucleotides 5121-5129 was annealed to RNA isolated from
313 transfected cells and extended using reverse transcriptase, generating a product
314 of 105 nts corresponding to the start of the sgRNA. Whilst sgRNA was detected
315 in infected or cells transfected with WT RNA, it was absent in m53 but restored to
316 almost wild-type levels in the m53r (Fig. 1F). Taken together, these data confirm
317 that the stability of Sla5045 contributes to MNV sgRNA synthesis, as well as that
318 the production of luciferase and the presence of infectious virus correlate with
319 sgRNA synthesis.
320
321 **Rapid reversion and suppression of Sla5045 disruption occur in cell**
322 **culture.**
323 The replication cycle of many positive-sense RNA viruses are error-prone
324 and the reversion or phenotypic suppression of the introduced mutations can
325 occur even during low levels of viral replication. Such reversion or suppression
326 can become apparent after repeated passage of samples obtained by reverse
327 genetics recovery. To examine whether it was possible to isolate revertants or
328 suppressors of MNV containing the m53 mutations, named MNVm53, clarified
329 tissue culture supernatants from three independent reverse genetics recoveries

330 were serially passaged in RAW264.7 cells until cytopathic effect (CPE) was
331 observed. In all cases CPE was observed after 3 passages in permissive cells.
332 Passages performed with similar samples obtained using the POL^{FS} construct
333 failed to result in CPE. Individual viral isolates were isolated by limiting dilution
334 and the sequence of the NS7 coding region was determined. In all cases we
335 observed additional changes either in Sla5045 or sequences elsewhere in the
336 viral genome (Fig. 2). In total we identified 12 different independently-derived
337 viable viruses that could be divided into two groups; phenotypic revertant
338 mutants (Rev 1-9) that contained mutations that phenotypically restored the
339 base-pairing of Sla5045 and suppressor mutants (Sup A-C) that retained the
340 m53 mutations but had additional changes elsewhere. All Rev isolates contained
341 mutations that appeared to restore the stability of Sla5045 by either reintroducing
342 the WT sequence (Rev 2-6 and 8), or by introducing a stabilizing mutation on the
343 opposite side of the stem (Fig. 2B). In contrast the suppressor mutants retained
344 the m53 mutations but had additional changes in the NS7-coding region. Of the
345 three Sup mutations identified, SupA and Sup C had synonomous mutations that
346 did not alter the corresponding amino acid (data not shown), while SupB resulted
347 in a threonine to alanine substitution at residue 439 of NS7.
348

349 **Both coding and non-coding suppressor mutations in the NS7 coding**
350 **region complement the defect in MNVm53**

351 To confirm that SupA-C can suppress the defect in MNVm53 replication,
352 individual mutations were introduced into full-length cDNA clones from WT MNV

353 or m53 and the effect on virus recovery examined in BSRT7 cells. The BHK-
354 derived BSRT7 cells are permissive for viral replication but not viral infection,
355 with the viral yields being indicative of a single cycle of viral replication (7). All
356 three suppressor mutations were viable in the m53 background, confirming their
357 ability to restore the replication defect of MNVm53 (Fig. 3A). SupA, with an
358 A4922G substitution on the negative-sense RNA, was consistently better at
359 restoring the replication defect of MNVm53 than the SupB or C. improving viral
360 yields by more than 10-fold when compared to the virus with both the m53 and
361 either the SupB or SupC mutations. All three Sup mutations were also viable in
362 the WT cDNA background, with SupA producing viral yields that were
363 comparable to WT MNV by 48 h after transfection (Fig. 3A and 3B). Sequence
364 analysis indicated that all viruses were stable for at least three passages in cell
365 culture (data not shown). Given that Sup A and C had synonomous mutations,
366 these results strongly suggest that the suppression of the m53 defect acted on
367 the viral RNA.
368
369 The replication of MNVm53SupA was further characterized by one-step
370 growth curve analysis (Fig. 3B-D). The production of infectious virus (Fig. 3B),
371 viral protein (Fig. 3C) and viral RNA (Fig. 3D) were all reduced in MNVm53SupA
372 when compared to WT MNV. The mutation introduced in SupA results in a silent
373 U4922C mutation on the positive-sense RNA (Fig. 4A). This position
374 corresponded to an alanine at codon 461 in NS7 and enables additional silent
375 mutations to be introduced. We examined whether substitutions to the other

376 three nucleotides at this position would also suppress the m53 defect as well as
377 affect WT MNV. All 4 mutations were well tolerated in a WT MNV background
378 whereas only the substitutions equivalent to SupA, A4922G on the minus strand,
379 or A4922U on the minus-strand RNA, restored replication in the MNVm53
380 backbone (Fig. 4B). All viable viruses could be repeatedly passaged in cell
381 culture and the sequences introduced remained stable (data not shown). These
382 data suggest that there is a sequence specific requirement for the ability of
383 position 4922 to suppress the defect in MNVm53.
384
385 **Sla5045 function at a different location in the MNV genome**
386 Due to the location of Sla5045 in the NS7 coding region, mutational
387 analysis was limited to synonymous changes that do not alter the NS7 protein
388 sequence. To facilitate additional mutational analysis, we added a second copy
389 of Sla5045, named Sla2, after the ORF1 termination codon. Sla2 is located within
390 a noncoding intragenic region introduced between the coding sequence of NS7
391 and VP1 in construct Sla5054Dup (Fig. 5A). The cDNA of Sla5045Dup was
392 engineered to contain either a WT or m53 version of Sla5045 or in the region
393 containing Sla2. *In vitro* transcribed and capped RNA was then transfected into
394 cells and the effect on virus yield examined 24 hours post transfection. RNA-
395 mediated reverse genetics recovery was used to significantly improve the yield of
396 recovered viruses and to enhance our ability to detect viruses that replicate
397 poorly. As expected, RNA generated from the WT MNV cDNA clone resulted in
398 ~10^5 infectious units per ml, whereas m53 and a polymerase mutant failed to

399 produce any detectable virus (Fig. 5B). Sla5045Dup containing Sla5045 WT/Sla2
400 WT (WT/WT) reproducibly yielded viable virus whereas the WT/m53 construct,
401 containing a mutated Sla2 did not (Fig. 5B). In constructs where Sla5045
402 contained the m53 mutations, a WT Sla2 restored replication whereas a m53
403 Sla2 did not (Fig. 5B). These results demonstrate that only stem-loop sequences
404 positioned at Sla2 function as a sgRNA promoter in the context of the
405 Sla5045Dup cDNA constructs. Importantly however, given that the yield of virus
406 obtained from the Sla5045Dup constructs WT/WT and m53/WT was significantly
407 lower than that of the WT virus (Fig. 5B), it appears that the sgRNA promoter
408 functions less well in the Sla2 position.
409
410 Western blot analysis was used as an indirect measure of viral RNA
411 transfection efficiency as we have previously observed that viral antigen
412 expression after transfection of RNA is due to translation of the capped RNA only
413 and not due to viral replication (9). Similar levels of the NS7 protein were
414 detectable in all cases, although the Sla5045Dup construct that contained
415 Sla5045 WT/Sla2 WT produced a protein detected by the anti-NS7 antisera with
416 higher molecular weight (identified by an asterisk in Fig. 5B). We suspect that
417 this was due to the insertion of an additional stable stem-loop structure in close
418 proximity to the NS7 stop codon resulting in suppression of translational
419 termination as this has been widely documented in the literature (Reviewed in
420 22). This hypothesis is also supported by the absence of the extended product
421 when the Sla5045 was disrupted by the introduction of the m53 mutations. The

422 lack of NS7 expression in the POL^{FS} polymerase mutant was due to the
423 introduction of a frame-shift mutation leading to the production of a truncated
424 NS7 protein that is rapidly degraded (9).
425
426 Although the titers of the Sla5045Dup viruses recovered were low, some
427 of the introduced mutations were stable and could be passaged in cell culture,
428 indicating that they are viable (Fig. 5C). Sequence analysis and 5' RACE
429 confirmed that the introduced mutations were stable for up to 3 passages in cell
430 culture and that the start of the sgRNA produced during authentic norovirus
431 replication in cell culture is position 5052, as has been predicted previously (Fig.
432 5D) (4).
433
434 **The sequence, positioning and stability of Sla5045 are essential for**
435 **norovirus replication.**
436 Our data indicates that the Sla5045 is essential for norovirus replication
437 and that its function can be complemented by an additional copy in *cis* in a non-
438 coding region upstream of the capsid-coding sequence. To examine the
439 sequence requirement further we made constructs that contained a mutated m53
440 within the NS7 coding region along with various mutant forms in the second copy
441 present in the non-coding region. The construct Sla5045Dup m53/Sla2 m53r
442 contained the m53r derivative of Sla5045 that is predicted to compensate for the
443 disrupted base pairs in m53 (Fig. 6A). As expected, the Sla5045Dup m53/Sla2
444 m53r produced viable virus after reverse genetics recovery (Fig. 6B). The

445 resultant virus could be passaged in cell culture (Fig. 6C) and that stably
446 maintain both copies of Sla5045 (data not shown). In contrast Sla5045Dup
447 m53/Sla2 WT+8, where the spacing between the stem-loop and the predicted
448 initiation site of the sgRNA was increased from 6 to 8 nts was not viable.
449 Significant alterations to the base stem sequences (Sla5045Dup m53/Sla2 SL+
450 or SL-) or disruption of the top stem sequence (Sla5045Dup m53/Sla2 S2) also
451 prevented replication (Fig. 6A-C). In contrast, the introduction of two nucleotide
452 changes in the terminal loop region of Sla5045 in the construct Sla5045Dup
453 m53/Sla2 TL-Dis had no effect on the ability of Sla5045 to complement the defect
454 in the m53 mutant backbone did not produce viable virus (Fig. 6A-C).
455
456 **Sla5045 functions as a template for the norovirus RNA polymerase NS7 *in***
457 ***vitro.***
458 Results from characterization of Sla5045 are all consistent with the
459 hypothesis that Sla5045 forms the core of the promoter for MNV sgRNA
460 synthesis and is used to direct sgRNA synthesis. However, it is unclear whether
461 it functions directly in RNA synthesis. To determine whether this is the case, we
462 examined whether RNAs that contain Sla5054 can direct RNA-dependent RNA
463 synthesis by the recombinant MNV NS7 protein *in vitro* (Fig. 7A). An RNA
464 containing Sla5045 with a 5' overhang that contains the initiation nucleotide 5052
465 was chemically synthesized in a RNA we termed the MNV proscript, MNVps. The
466 term proscript is used to denote that the RNA contains both a putative promoter
467 and the template sequence (23) (Fig. 7B). Should the MNV proscript direct

468 accurate initiation of RNA synthesis from nt 5052, an 11-nt product should result.
469 Based on comparison with RNAs of known lengths, the products of MNVps were
470 11-nt in length. We also observed a 10-nt RNA that is presumed to be a
471 premature termination product that has been previously observed with other
472 recombinant RdRps. (Fig. 7C; 24). To ascertain that the RNA was initiated from
473 the expected cytidylate at nt 5052, we tested an RNA that is identical to MNVps
474 except that nt 5052 was substituted with an adenylate named IM. IM failed to
475 direct the synthesis of either the 10- and 11-nt RNA products (Fig. 7C). These
476 data confirm that Sla5045 can accurately direct *de novo* initiation of the norovirus
477 sgRNA synthesis.

478 In the *in vitro* RNA synthesis reactions we also detected an RdRp product
479 of ca. 60-nt, longer in length than the input proscript. This RNA is likely formed by
480 viral RdRp using the 3' terminal nucleotide of the template RNA to form a primer-
481 extended product (25). This activity is common to numerous viral RdRps (26, 27).
482 Notably, IM generated a higher abundance of the primer-extended RNA products
483 than did MNVps. Since *de novo* initiated RNA product generated from MNVps
484 was more abundant that those from primer extension, we conclude that the
485 recombinant MNV RdRp has a propensity for *de novo* initiated RNA synthesis
486 from the MNV sgRNA.

487 To examine further whether the MNV Sla5045 can direct a genotype-
488 specific RNA synthesis by the MNV RdRp, we performed *in vitro* RNA synthesis
489 reaction with proscript containing the comparable sequence from a human GII.4
490 norovirus named GII.4Ps (Fig. 7B). Accurate *de novo* initiated RNA synthesis

491 from GII.4Ps should give rise to an 8-nt product (Fig. 7D). The recombinant MNV
492 RdRp did produce RNAs of 8-nt and a 7-nt from GII.4Ps (Fig. 7D). However, in
493 more than six independent assays, the amount of the RNA product was 27% of
494 the amount produced from the MNV RdRp. In contrast, the recombinant GII.4
495 RdRp produced more than 30-fold the amount of products from the GII.4Ps than
496 from the MNVps (Fig. 7D). The MNV RdRp reproducibly produced a reduced
497 amount of the primer-extension product from the GII.4ps (13%, Fig. 7D).
498 Furthermore, the GII.4 RdRp also produced a reduced level of primer-extension
499 products from the MNVps (Fig. 7D). These results suggest that MNV RdRp can
500 specifically recognize the MNVps.
501
502 **Negative-sense subgenomic RNA is produced during MNV replication**
503 The presence of Sla5045 on the negative-sense genomic RNA and the
504 ability of Sla5045 to function *in vitro* as a template for priming by the RdRp would
505 fit with the hypothesis that MNV uses a process of internal initiation for the
506 generation of sgRNA. However, negative-sense sgRNA is readily detected in
507 cells infected with feline calicivirus (28) and in Norwalk replicon containing cells
508 (29). Northern blot and strand-specific RT-qPCR analysis of RNA isolated from
509 MNV infected cells also confirmed the presence of a negative-sense sgRNA
510 intermediate (Fig. 8A). Whilst we observed similar levels of negative-sense
511 genomic and sgRNA produced in infected cells, the levels of sgRNA were
512 typically >26 fold higher than the genomic RNA levels (Fig. 8B), fitting with
513 previous observation on FCV and Norwalk virus (28, 29). These data are

514 consistent with the hypothesis that newly synthesized MNV sgRNA may function
515 as a template for further rounds of replication via a negative-sense sgRNA
516 intermediate.
517
518 **Discussion**
519 In this study we used a combination of genetic and biochemical
520 approaches to characterize a stem-loop RNA sequence that we predicted to be
521 the MNV sgRNA promoter. A mutant MNV that has substitutions that disrupted
522 the predicted secondary structure of the stem-loop was poorly infectious while
523 nucleotide substitutions that restored the stem-loop structure allowed the virus to
524 regain infectivity. Suppressor mutations found after repeated passages in cell
525 culture also had restored Sla5045 structure, although additional suppressor
526 mutants were also isolated outside of the Sla5045. Examination of a second copy
527 of the sla5045 in a non-coding region of the MNV genome revealed that the
528 stem-loop structure functions in *cis*. Using purified recombinant viral RdRp, we
529 demonstrate an RNA containing Sla5054 can direct the genotype-specific initiate
530 RNA synthesis from Sla5045.
531
532 A change in the codon of viral RdRp NS7 was found in SupB, which
533 suppressed the changes in m53. Further analysis of this nucleotide position
534 indicated that the effect was mediated in a non-coding context as changing this
535 codon (ACU) to GUU to code for valine, or to GAU to code for aspartic acid, also
536 restored replication to the m53 (data not shown). There are several explanations

537 for our ability to isolate second site suppressor mutations outside of Sla5045; the
538 first is that Sla5045 may form higher order RNA structures with other regions of
539 the viral genome, including the positions identified as second site suppressors,
540 and that the second site suppressors stabilize these interactions. Our
541 bioinformatic analysis has failed to identify any obvious interactions between the
542 regions (data not shown). However, accurate predictions of long- or medium-
543 range RNA-RNA interactions are not currently available and we cannot at this
544 point rule out that long-range RNA-RNA interaction contribute to Sla5045
545 function. In addition, the introduction of the m53 mutations could result in
546 promiscuous base-pairing of Sla5045 with other sequences flanking the start of
547 the sgRNA. The suppressor mutations then function to disrupt the 'off-target'
548 RNA-RNA interactions, favoring the restoration of the Sla5045 structure.
549 Distinguishing between these two possibilities will be challenging, however our
550 observation that the addition of Sla5045 to the 3' end of mini-genome RNA
551 encoding a reporter gene in the antisense orientation is not sufficient to drive
552 RNA synthesis and reporter gene expression in infected cells (data not shown),
553 suggests that additional RNA sequences may contribute to RNA synthesis
554 directed by core promoter that consists of Sla5045.
555
556 The production of a sgRNA is conserved in all members of the
557 *Caliciviridae* yet the mechanism of its synthesis is poorly understood. In contrast,
558 the synthesis of sgRNA is well characterized in numerous other RNA viruses
559 (Reviewed in 14). For viruses that possess only a single genomic RNA and

560 generate a single sgRNA, one commonly used mechanism for sgRNA synthesis
561 is the premature termination of negative-strand RNA synthesis that is used by the
562 viral replicase as a template to produce the positive-sense sgRNA. A hallmark of
563 the premature termination mechanism is the production of a truncated negative-
564 sense sgRNA during virus replication. Low levels of negative-sense sgRNA have
565 been identified in cells stably replicating the Norwalk virus replicon (29), in
566 purified replication complex from feline calicivirus infected cells (28) and as now
567 in MNV infected cells (Fig 8). However, our data and those from previous
568 publications (15) would suggest an alternative model for MNV sgRNA synthesis.
569 This mechanism is used by plant viruses (23, 30, 31) and alphaviruses (32, 33)
570 and involves the binding of the viral replicase complex to a specific sequence
571 and/or structure present in negative-sense genomic RNA. Our data is consistent
572 with Sla5045 being required as an RNA structure needed for MNV infectivity and
573 to function as a template for RNA synthesis by the RdRp. It is also important to
574 note that in the plant viruses, it is not clear which subunit(s) of the viral replicase
575 can specifically recognized the sgRNA promoter and our results demonstrate that
576 the norovirus RdRp is sufficient for recognition of the sgRNA core promoter. In
577 addition we observed that a genotype-specific interaction between the RdRp and
578 the subgenomic proscript affected the amount of RNA synthesis (Fig. 7D). Our
579 data would however indicate that as negative-sense sgRNA is present in MNV
580 infected cells, the newly synthesized MNV sgRNA may function as a template for
581 additional rounds of RNA replication by the viral RdRp via a dsRNA intermediate.

582 The precise initiation site of the calicivirus sgRNA has been identified
583 previously in a number of studies. The human norovirus sgRNA initiation site has
584 been previously confirmed using a helper virus system to drive the production of
585 viral genomic RNA in cells (34). In the related vesivirus FCV, 5'RACE was also
586 used to identify the specific sgRNA initiation site (35). To our knowledge, our
587 work provides the first confirmation of the start site of norovirus sgRNA during an
588 authentic infectious norovirus replication cycle.

589 Our data would indicate that the positioning of the core sgRNA promoter is
590 critical for its function as in the Sla5045Dup constructs, only the sgRNA promoter
591 sequences in Sla2 region functioned (Fig. 5). In addition, increasing the distance
592 between the stem-loop structure and the sgRNA initiation site from 6 to 8 nts,
593 also debilitated virus replication (Fig. 6). These data are in agreement with our
594 previous observations using bioinformatics analysis of all calicivirus genomes
595 that indicate an absolute conservation of the spacing i.e. invariably 6 nts (16).
596

597 Overall our data demonstrate that norovirus sgRNA synthesis relies on a
598 sequence and genotype-specific interaction of the viral RdRp with a stem-loop
599 sequence on the minus-strand RNA. These observations add to our growing
600 understanding of the norovirus life cycle and the molecular mechanisms used by
601 caliciviruses to control viral genome translation and replication.
602
603

604 **Acknowledgements**

605 The authors would like to thank Stanislav Sosnovtsev for critical comments on
606 the work. This work was supported by grants from the Wellcome Trust (Ref:
607 WT097997MA) and BBSRC (BB/I012303/1) to IG, from the Indiana Economic
608 Development Corporation to CK (AI073335) and from the Research University
609 Grant, USM, to MAY (1001/CIPPT/811233). IG is a Wellcome Senior Fellow.

610

611 **References**

612 1. **Glass RI**, **Parashar UD**, **Estes MK**. 2009. Norovirus gastroenteritis. N.
613 Engl. J. Med. **361**:1776–1785.
614 2. **Hall AJ**, **Lopman BA**, **Payne DC**, **Patel MM**, **Gastañaduy PA**, **Vinjé J**,
615 **Parashar UD**. 2013. Norovirus disease in the United States. Emerging
616 Infect. Dis. **19**:1198–1205.
617 3. **Duizer E**, **Schwab KJ**, **Neill FH**, **Atmar RL**, **Koopmans MPG**, **Estes MK**.
618 2004. Laboratory efforts to cultivate noroviruses. J. Gen. Virol. **85**:79–87.
619 4. **Karst SM**, **Wobus CE**, **Lay M**, **Davidson J**, **Virgin HW**. 2003. STAT1-
620 dependent innate immunity to a Norwalk-like virus. Science **299**:1575–
621 1578.
622 5. **Wobus CE**, **Karst SM**, **Thackray LB**, **Chang K-O**, **Sosnovtsev SV**,
623 **Belliot G**, **Krug A**, **Mackenzie JM**, **Green KY**, **Virgin HW**. 2004.
624 Replication of Norovirus in cell culture reveals a tropism for dendritic cells
625 and macrophages. PLoS Biol. **2**:e432.
626 6. **Ward VK**, **McCormick CJ**, **Clarke IN**, **Salim O**, **Wobus CE**, **Thackray LB**,
627 **Virgin HW**, **Lambden PR**. 2007. Recovery of infectious murine norovirus
628 using pol II-driven expression of full-length cDNA. Proc. Natl. Acad. Sci.
629 U.S.A. **104**:11050–11055.
630 7. **Chaudhry Y**, **Skinner MA**, **Goodfellow IG**. 2007. Recovery of genetically
631 defined murine norovirus in tissue culture by using a fowlpox virus
632 expressing T7 RNA polymerase. J. Gen. Virol. **88**:2091–2100.
633 8. **Arias A**, **Ureña L**, **Thorne L**, **Yunus MA**, **Goodfellow I**. 2012. Reverse
634 genetics mediated recovery of infectious murine norovirus. J Vis Exp
635 e4145.
636 9. **Yunus MA**, **Chung LMW**, **Chaudhry Y**, **Bailey D**, **Goodfellow I**. 2010.
637 Development of an optimized RNA-based murine norovirus reverse
638 genetics system. J. Virol. Methods **169**:112–118.
639 10. **Thorne LG**, **Goodfellow IG**. 2014. Norovirus gene expression and
640 replication. J. Gen. Virol. **95**:278–291.
641 11. **Clarke IN**, **Lambden PR**. 2000. Organization and expression of calicivirus
642 genes. J. Infect. Dis. **181 Suppl 2**:S309–16.
643 12. **McFadden N**, **Bailey D**, **Carrara G**, **Benson A**, **Chaudhry Y**, **Shortland**

644 **A**, **Heeney J**, **Yarovinsky F**, **Simmonds P**, **Macdonald A**, **Goodfellow I**.
645 2011. Norovirus regulation of the innate immune response and apoptosis
646 occurs via the product of the alternative open reading frame 4. PLoS
647 Pathog. **7**:e1002413.
648 13. **Thackray LB**, **Wobus CE**, **Chachu KA**, **Liu B**, **Alegre ER**, **Henderson**
649 **KS**, **Kelley ST**, **Virgin HW**. 2007. Murine noroviruses comprising a single
650 genogroup exhibit biological diversity despite limited sequence divergence.
651 J Virol **81**:10460–10473.
652 14. **Sztuba-Solińska J**, **Stollar V**, **Bujarski JJ**. 2011. Subgenomic messenger
653 RNAs: mastering regulation of (+)-strand RNA virus life cycle. Virology
654 **412**:245–255.
655 15. **Morales M**, **Bárcena J**, **Ramírez MA**, **Boga JA**, **Parra F**, **Torres JM**.
656 2004. Synthesis in vitro of rabbit hemorrhagic disease virus subgenomic
657 RNA by internal initiation on (-)sense genomic RNA: mapping of a
658 subgenomic promoter. J. Biol. Chem. **279**:17013–17018.
659 16. **Simmonds P**, **Karakasiliotis I**, **Bailey D**, **Chaudhry Y**, **Evans DJ**,
660 **Goodfellow IG**. 2008. Bioinformatic and functional analysis of RNA
661 secondary structure elements among different genera of human and
662 animal caliciviruses. Nucleic Acids Res. **36**:2530–2546.
663 17. **Subba-Reddy CV**, **Yunus MA**, **Goodfellow IG**, **Kao CC**. 2012. Norovirus
664 RNA Synthesis Is Modulated by an Interaction between the Viral RNA-
665 Dependent RNA Polymerase and the Major Capsid Protein, VP1. J Virol
666 **86**:10138–10149.
667 18. **Meyers G**. 2007. Characterization of the sequence element directing
668 translation reinitiation in RNA of the calicivirus rabbit hemorrhagic disease
669 virus. J Virol **81**:9623–9632.
670 19. **Napthine S**, **Lever RA**, **Powell ML**, **Jackson RJ**, **Brown TDK**, **Brierley I**.
671 2009. Expression of the VP2 protein of murine norovirus by a translation
672 termination-reinitiation strategy. PLoS ONE **4**:e8390.
673 20. **Vashist S**, **Ureña L**, **Goodfellow I**. 2012. Development of a strand specific
674 real-time RT-qPCR assay for the detection and quantitation of murine
675 norovirus RNA. J. Virol. Methods.
676 21. **Yi G**, **Deval J**, **Fan B**, **Cai H**, **Soulard C**, **Ranjith-Kumar CT**, **Smith DB**,
677 **Blatt L**, **Beigelman L**, **Kao CC**. 2012. Biochemical study of the
678 comparative inhibition of hepatitis C virus RNA polymerase by VX-222 and
679 filibuvir. Antimicrob. Agents Chemother. **56**:830–837.
680 22. **Firth AE**, **Brierley I**. 2012. Non-canonical translation in RNA viruses. J.
681 Gen. Virol. **93**:1385–1409.
682 23. **Siegel RW**, **Adkins S**, **Kao CC**. 1997. Sequence-specific recognition of a
683 subgenomic RNA promoter by a viral RNA polymerase. Proc. Natl. Acad.
684 Sci. U.S.A. **94**:11238–11243.
685 24. **Kim Y-C**, **Russell WK**, **Ranjith-Kumar CT**, **Thomson M**, **Russell DH**,
686 **Kao CC**. 2005. Functional analysis of RNA binding by the hepatitis C virus
687 RNA-dependent RNA polymerase. J. Biol. Chem. **280**:38011–38019.
688 25. **Kao CC**, **Yang X**, **Kline A**, **Wang QM**, **Barket D**, **Heinz BA**. 2000.
689 Template requirements for RNA synthesis by a recombinant hepatitis C

690 virus RNA-dependent RNA polymerase. J Virol **74**:11121–11128.
691 26. **Ranjith-Kumar CT**, **Gajewski J**, **Gutshall L**, **Maley D**, **Sarisky RT**, **Kao**
692 **CC**. 2001. Terminal nucleotidyl transferase activity of recombinant
693 Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA
694 synthesis. J Virol **75**:8615–8623.
695 27. **Tomar S**, **Hardy RW**, **Smith JL**, **Kuhn RJ**. 2006. Catalytic core of
696 alphavirus nonstructural protein nsP4 possesses terminal
697 adenylyltransferase activity. J Virol **80**:9962–9969.
698 28. **Green KY**, **Mory A**, **Fogg MH**, **Weisberg A**, **Belliot G**, **Wagner M**, **Mitra**
699 **T**, **Ehrenfeld E**, **Cameron CE**, **Sosnovtsev SV**. 2002. Isolation of
700 enzymatically active replication complexes from feline calicivirus-infected
701 cells. J Virol **76**:8582–8595.
702 29. **Chang K-O**, **Sosnovtsev SV**, **Belliot G**, **King AD**, **Green KY**. 2006.
703 Stable expression of a Norwalk virus RNA replicon in a human hepatoma
704 cell line. Virology **353**:463–473.
705 30. **Adkins S**, **Siegel RW**, **Sun JH**, **Kao CC**. 1997. Minimal templates
706 directing accurate initiation of subgenomic RNA synthesis in vitro by the
707 brome mosaic virus RNA-dependent RNA polymerase. RNA **3**:634–647.
708 31. **Sivakumaran K**, **Chen M-H**, **Roossinck MJ**, **Kao CC**. 2002. Core
709 promoter for initiation of Cucumber mosaic virus subgenomic RNA4A. Mol.
710 Plant Pathol. **3**:43–52.
711 32. **Levis R**, **Schlesinger S**, **Huang HV**. 1990. Promoter for Sindbis virus
712 RNA-dependent subgenomic RNA transcription. J Virol **64**:1726–1733.
713 33. **Ou JH**, **Rice CM**, **Dalgarno L**, **Strauss EG**, **Strauss JH**. 1982. Sequence
714 studies of several alphavirus genomic RNAs in the region containing the
715 start of the subgenomic RNA. Proc. Natl. Acad. Sci. U.S.A. **79**:5235–5239.
716 34. **Asanaka M**, **Atmar RL**, **Ruvolo V**, **Crawford SE**, **Neill FH**, **Estes MK**.
717 2005. Replication and packaging of Norwalk virus RNA in cultured
718 mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **102**:10327–10332.
719 35. **Neill JD**. 2002. The subgenomic RNA of feline calicivirus is packaged into
720 viral particles during infection. Virus Res. **87**:89–93.
721

722

723

724

725

726

727

728

729 **Figure legends**:

730

731 **Fig. 1. The stability of Sla5045 is required for norovirus infectivity,**
732 **replication and sgRNA synthesis.**

733 **A**) A schematic of the murine norovirus (MNV) genome, highlighting the four
734 open reading frames and Sla5045. **B**) Genetic conservation of the Sla5045
735 sequences on the norovirus negative-sense viral RNA. The position of the
736 potential norovirus subgenomic RNA initiation site is indicated with an arrow. The
737 position of the VP1 initiation codon on the positive-sense RNA is highlighted in
738 bold and sequence variation observed in different isolates is shown as bold italic
739 text. Genogroup II noroviruses (GII); the sequence of the GII isolate
740 Hu/GII.4/MD-2004/2004/US is shown (DQ658413) with variation between the
741 following isolates shown: FJ595907.1, AB447425.1, HM635151, HM635164,
742 JF262610.1 and JF262592. Genogroup V (GV) murine norovirus; the sequence
743 of the CW1 isolate (DQ285629) is shown with variation between the following
744 isolates displayed; JN975491 and JN975492. Note that for simplicity, only
745 sequences showing variation in the stem-loop sequence were used in the
746 analysis and shown in the figure. C) The structure of Sla5045 and the mutations
747 introduced to generate the mutant m53 and m53r as described previously (16). In
748 the case of m53r, the mutations shown are in addition to those present in m53.
749 D) Luciferase replicon based analysis of the effect of Sla5045 disruption.
750 Luciferase levels represent the mean of triplicate independent samples. E) Virus
751 yield after reverse genetics recovery of full-length cDNA constructs containing

752 either MNV wild type (WT), a polymerase frame shift in the NS7 (POL^{FS}), m53 or
753 m53r. The virus titres represent the mean values obtained 24 hours post
754 transfection of viral cDNA expression constructs. The detection limit of the assay
755 is indicated with a dotted line. In all cases where shown, error bars represent the
756 SEM. F) Primer extension-mediated detection of sgRNA synthesis in MNV
757 infected cells (Inf) or cells transfected with capped RNAs of WT, POL^{FS}, m53 or
758 m53r clones. *In vitro* transcribed sgRNA is was used as a positive control
759 (sgRNA) and produces a product 3 nts longer due to the addition of three 5' G
760 nts by T7 RNA polymerase. Note that 1000 fold less RNA was used for the
761 reaction using RNA from infected cells.
762
763
764 **Fig. 2: Sequence analysis of m53 phenotypic revertants and second site**
765 **suppressor mutants.**
766 A) Positions mutated in the m53 mutant are shown in bold, with the any changes
767 observed in the phenotypic revertants highlighted in grey. The nucleotide
768 sequences shown are for the negative-sense RNA genome. An asterisk indicates
769 that the identified sequence change results in a coding change in the viral NS7
770 protein (T439A). Revertant mutants are defined as those that phenotypically
771 restore the base pairing of Sla5045. Suppressors are those that retain the
772 mutations in Sla5045 but have additional mutations outside of the region. B)
773 Location of reversion mutations isolated after repeated passage of m53 in cell

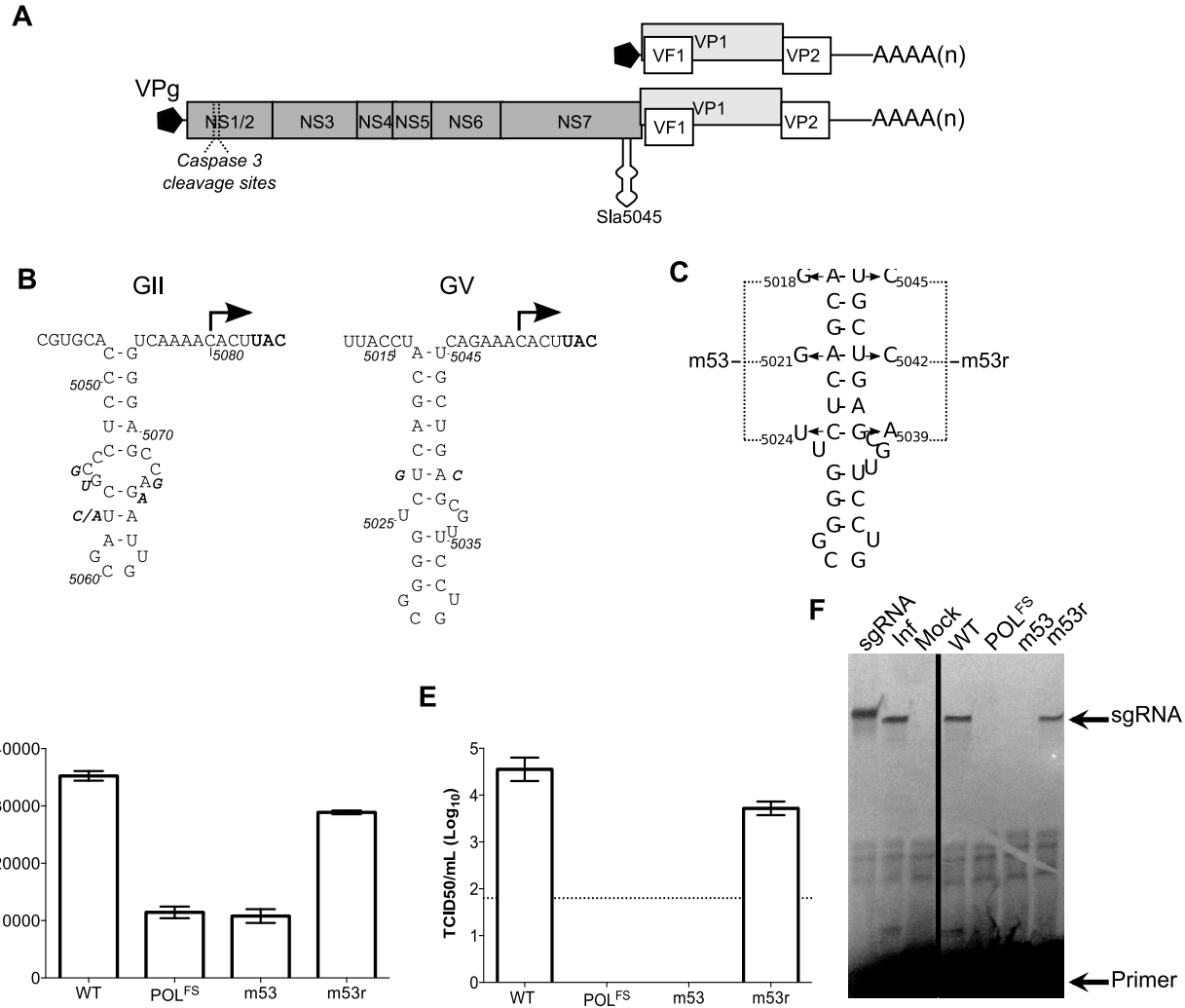
774 culture. Note that the suppressor mutations are not shown as they lie outside of
775 the region containing Sla5045.
776
777 **Fig. 3: Characterization of m53 suppressor mutations.**
778 A) Virus yield following reverse genetics rescue of cDNA constructs containing
779 with WT or m53 MNV cDNA with or without the additional second site mutations
780 as detailed in Table 1. Reverse genetics recovery was performed as described in
781 the materials and methods then the virus yield at 24 hours post transfection
782 determined by TCID50. The dotted line represents the detection limit of the assay
783 with error bars representing the SEM. Single step growth curve analysis of
784 m53SupA replication in comparison to the WT MNV. RAW264.7 cells were
785 infected at a multiplicity of infection of 4 TCID50 per cell and samples harvest at
786 the indicated time post infection. Infectious virus was released by freeze-thaw
787 and titrated. Infectious virus titres are shown in panel B, protein expression levels
788 of NS7 and VP1 in Panel C and viral genomic RNA levels in panel D. In panel D
789 only samples from the exponential growth phase are shown. Error bars represent
790 the SEM.
791
792 **Fig. 4: Sequence changes at position 4922 can compensate a defect in**
793 **Sla5045.**
794 A) Schematic illustration of the region of the MNV genome mutated in m53 and
795 m53SupA. The sequence of the region shown is represented in the positive
796 polarity to highlight the NS7 reading frame with the negative polarity sequence

797 shown below. The amino acids encoded by the region are highlighted, as well as
798 the position of the SupA mutation (in italics) and the mutations introduced in m53.
799 B) Virus yield following reverse genetics rescue of cDNA constructs containing
800 with wild-type (WT) or m53 MNV cDNA with various changes at position 4922.
801 Note that the nucleotide sequences shown represent the sequence of the anti-
802 sense RNA genome with G representing the previously isolated SupA
803 suppressor mutation and A representing the nucleotide present in the wild-type
804 cDNA construct. Reverse genetics recovery was performed as described in the
805 materials and methods then the virus yield at 24 hours post transfection
806 determined by TCID50. The dotted line represents the detection limit of the assay
807 with error bars representing the SEM.
808
809 **Fig. 5: Duplication of Sla5045 enables mutagenesis in a non-coding context**
810 A) Schematic illustration of the reconstruction of Sla5045Dup stem-loop
811 duplication constructs. The position of the authentic Sla5045 on the anti-sense
812 RNA genome is shown and the introduced additional Sla5045 shown as Sla2.
813 For simplicity, only wild-type copies of the stem-loop are shown. The positions of
814 the NS7 stop and VP1 start codons on the corresponding positive sense viral
815 RNA are also shown for reference. B) Virus yield following reverse genetics
816 rescue of capped *in vitro* transcribed RNA of wild type (WT), polymerase frame-
817 shift (POL^{FS}), m53 or Sla5045 stem-loop duplication (Sla5045Dup) constructs.
818 The Sla5045Dup constructs are shown using the nomenclature copy1/copy2 as
819 described in the text. Reverse genetics recovery was performed as described in

820 the materials and methods using *in vitro* transcribed, enzymatically capped RNA
821 and the virus yield at 24 hours post transfection determined by TCID50. A
822 western blot analysis for NS7 expression is shown below the virus yield data.
823 The asterisk indicates a greater than full length NS7 product observed in cells
824 transfected with SLa5045Dup WT/WT. C) Virus yield assay following a single
825 passage of virus recovered following RNA transfection of wild type MNV (WT) or
826 various Sla5045 stem-loop duplications. The dotted line represents the detection
827 limit of the assay with error bars representing the SEM. D) 5'RACE analysis of
828 WT MNV or Sla5045 stem-loop duplications. Viral RNA was isolated from
829 infected cells and subjected to 5' RACE analysis. Samples were prepared with or
830 without reverse transcription (RT) and then subjected to PCR as described in the
831 materials and methods. The wild-type MNV virus was included as a control.
832 Samples were then sequenced and aligned as shown. The lower case polyG
833 tract is introduced as a result of the 5' RACE methodology.
834
835 **Fig. 6. Mutational analysis of Sla5045 in a non-coding context.**
836 A) Schematic illustration of Sla5045 mutants characterized in the Sla5045Dup
837 cDNA backbone. The Sla5045 is shown in the 3'-5' direction with the site of
838 subgenomic RNA initiation shown in italics. Mutations introduced are highlighted
839 in bold. B) Virus yield following reverse genetics rescue of capped *in vitro*
840 transcribed RNA of WT, polymerase frame-shift (POL^{FS}), m53 or Sla5045 stem-
841 loop duplication (Sla5045Dup) constructs. The Sla5045Dup constructs are
842 shown using the nomenclature copy1/copy2 as described in the text. Reverse

843 genetics recovery was performed as described in the materials and methods
844 using *in vitro* transcribed, enzymatically capped RNA and the virus yield at 24
845 hours post transfection determined by $TCID_{50}$. C) Virus yield assay following a
846 single passage of virus recovered following RNA transfection the various
847 Sla5045 stem-loop duplications. The dotted line represents the detection limit of
848 the assay with error bars representing the SEM
849
850 **Fig. 7: Sla5045 can direct for the initiation of RNA synthesis by the MNV**
851 **RdRp *in vitro.***
852 A) Recombinant RdRps from MNV and the human GII.4 norovirus used in this
853 study. The RdRps was expressed and purified as described in the materials and
854 methods. The gel images were from SDS-PAGE stained with Coomassie Blue.
855 B) Schematics of the synthetic MNV proscript, MNVps, and the GII.4 RNA
856 proscript, GII.4Ps, used for the *in vitro* RNA synthesis assays. The initiation
857 nucleotide used to initiate MNV subgenomic RNA synthesis is in bold and
858 denoted with a bent arrow. A proscript containing a mutated initiation site named
859 IM is used to control for specificity. C) Denaturing PAGE analysis of *in vitro* RNA
860 synthesis assays containing the MNV NS7 with either MNVps or the initiation
861 mutant (IM). The positions of primer extension (PE) and *de novo* (DN) initiation
862 products are highlighted with arrows. The lengths of the RNAs were assigned by
863 comparison to RNAs of 8 to 51-nt that were labeled at the 5' terminus with a ^{32}P.
864 D) Virus genotype-specific RdRp-proscript interaction results in higher level of

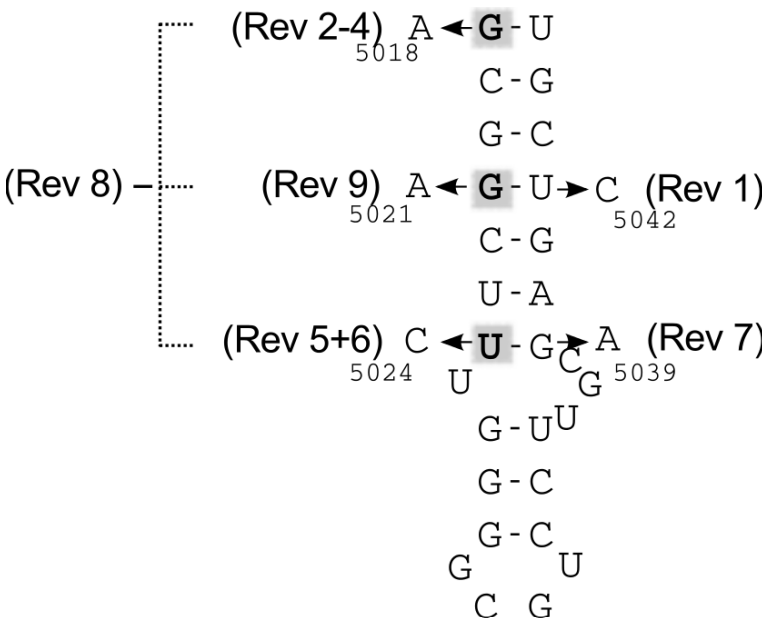
865 RNA-dependent RNA synthesis. The results are representative of six
866 independent experiments.
867
868 **Fig. 8: Negative-sense subgenomic RNA is produced during MNV**
869 **replication.** A) Northern blot analysis of RNA isolated from either mock infected
870 (mock) or MNV infected (Inf) cells. 1.5μg of RNA harvested 10 hours post
871 infection of cells with a multiplicity of infection of 5 TCID50 per cell, was
872 denatured by glyoxylation prior to agarose gel electrophoresis. Positive and
873 negative-sense RNA was detected using strand specific RNA probes as
874 described in materials and methods. Control *in vitro* transcribed RNAs were
875 generated representing the positive- (+) or negative-sense (-) viral RNAs were
876 used as a control. B). Quantitative real-time PCR analysis of the RNA sample
877 shown in panel A. The RNA was analyzed using a strand-specific RT-qPCR
878 assay designed to detect only the genomic RNA or both genomic and
879 subgenomic RNAs simultaneously. The genome copy number is shown as
880 genome equivalent per 25ng of total RNA and was determined by comparison to
881 *in vitro* transcribed control RNAs. The data show was represent the mean and
882 standard deviation of 4 independent repeats.
883

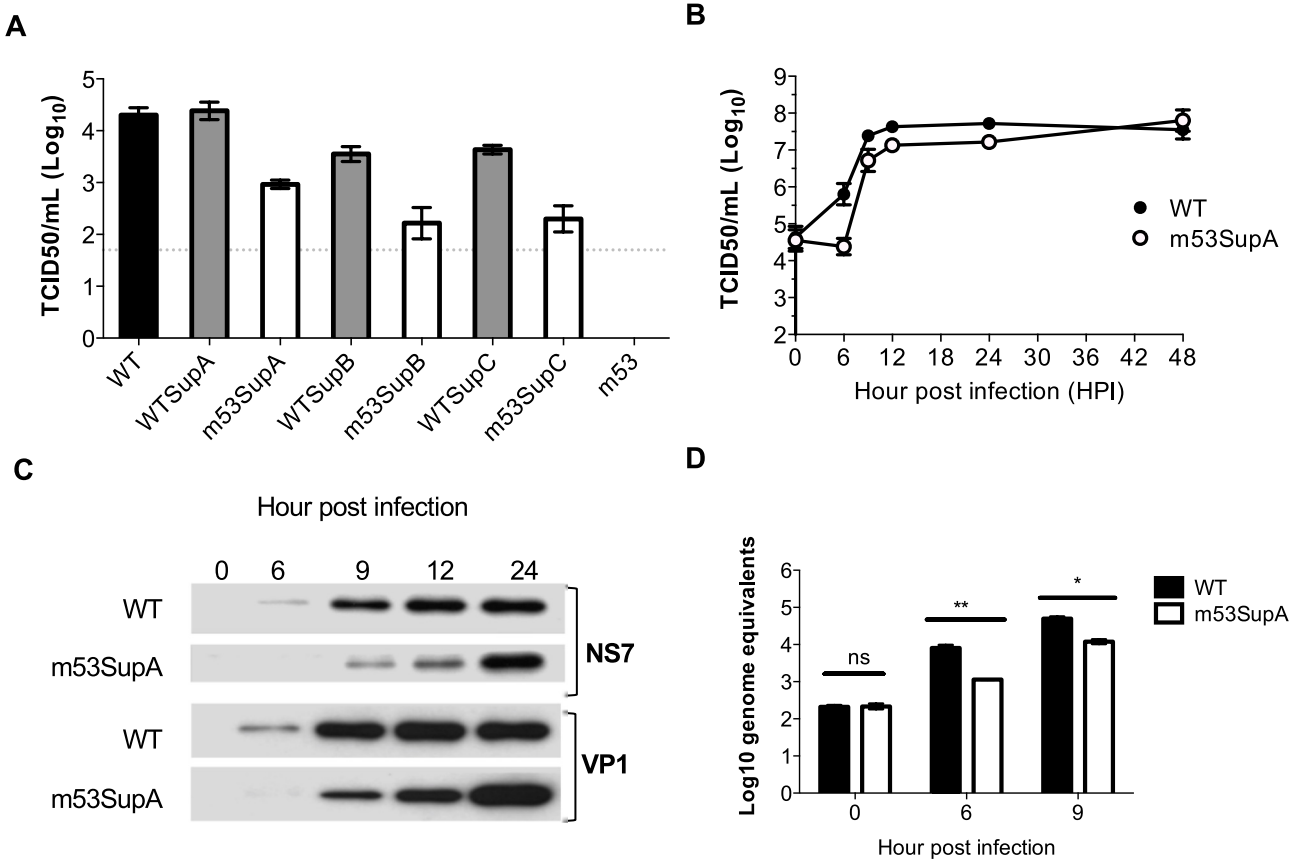

A
VPg
NS1/2
NS3
NS4
NS5
NS6
NS7
Caspase 3
cleavage sites
VF1
VP1
VP2
AAAA(n)
Sla5045
B
GII
GV
CGUGCA UCAAAACACUUAC
5080
5050
5070
5060
UUACCU CAGAAACACUUAC
5015
5045
5025
5035
C
m53
m53r
5018
5021
5024
5045
5042
5039
D
Relative Luminescene Units
40000
30000
20000
10000
0
WT
POL^FS
m53
m53r
E
TCID50/mL (Log10)
5
4
3
2
1
0
WT
POL^FS
m53
m53r
F
sgRNA
Inf
Mock
WT
POL^FS
m53
m53r
sgRNA
Primer

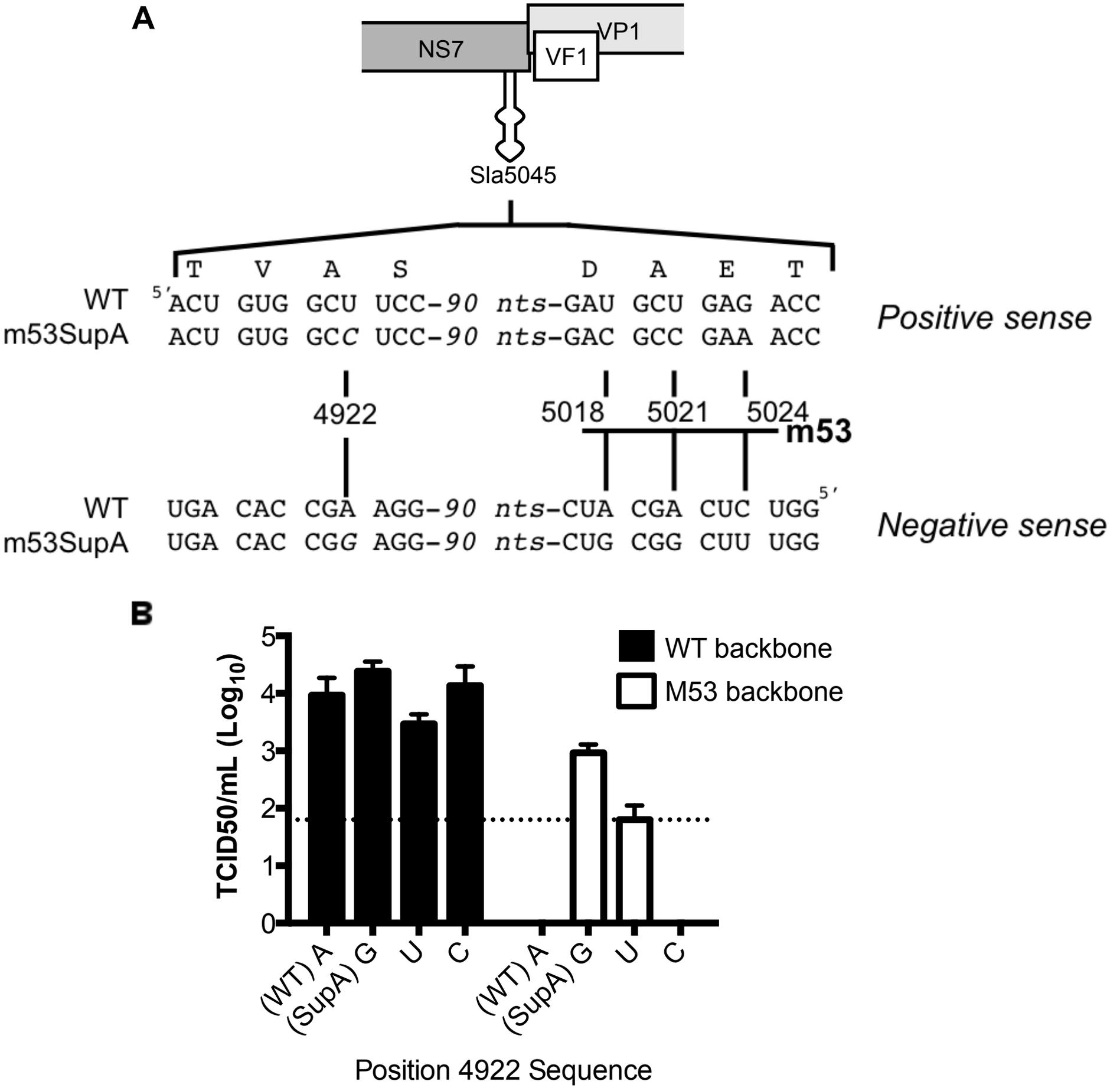
A

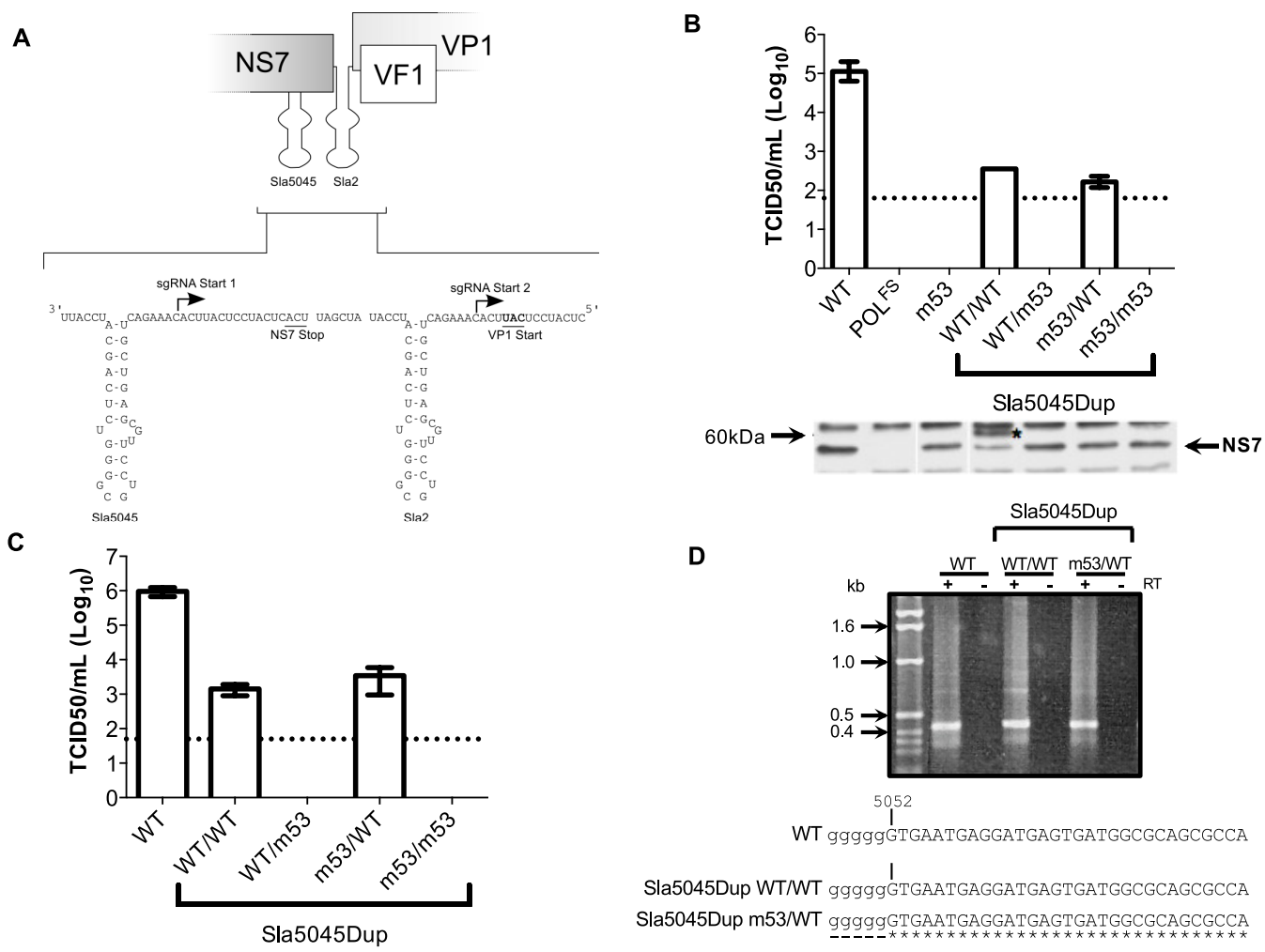
Virus	Nucleotide position within Sla5045						Changes outside of Sla5045
	5018	*5021*	*5024*	*5039*	*5042*	*5045*	
WT	A	A	C	G	U	U	
m53	**G**	**G**	**U**	G	U	U	
Rev1	**G**	**G**	**U**	G	C	U	G4910A
Rev2	A	**G**	**U**	G	U	U	
Rev3	A	**G**	**U**	G	U	U	G4859A
Rev4	A	**G**	**U**	G	U	U	U4967C
Rev5	**G**	**G**	C	G	U	U	
Rev6	**G**	**G**	C	G	U	U	U4967C
Rev7	**G**	**G**	**U**	A	U	U	
Rev8	A	A	C	G	U	U	
Rev9	**G**	A	**U**	G	U	U	
SupA	**G**	**G**	**U**	G	U	U	A4922G
SupB	**G**	**G**	**U**	G	U	U	U4914C*
SupC	**G**	**G**	**U**	G	U	U	U4967C

B

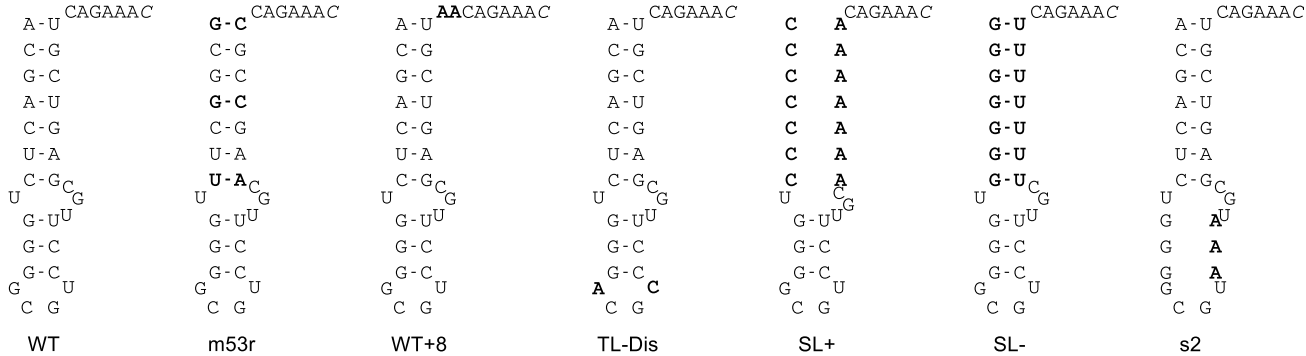



A
TCID50/mL (Log10)
0
1
2
3
4
5
WT
WTSupA
m53SupA
WTSupB
m53SupB
WTSupC
m53SupC
m53
B
TCID50/mL (Log10)
2
3
4
5
6
7
8
9
0
6
12
18
24
30
36
42
48
Hour post infection (HPI)
WT
m53SupA
C
Hour post infection
0
6
9
12
24
WT
m53SupA
WT
m53SupA
NS7
VP1
D
Log10 genome equivalents
0
1
2
3
4
5
6
ns
**
*
0
6
9
Hour post infection
WT
m53SupA

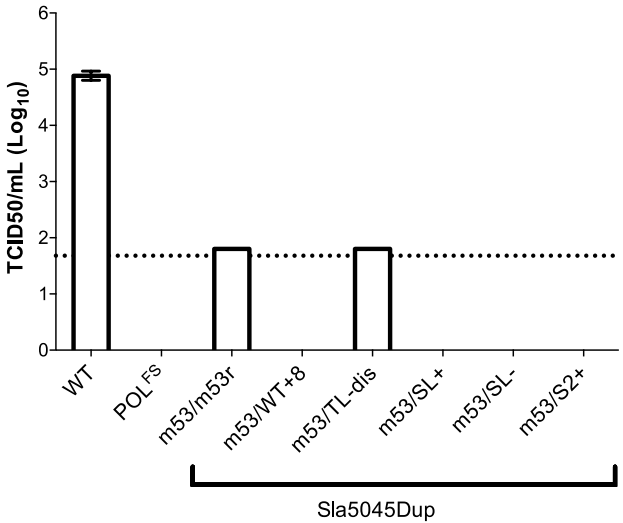

A
NS7
VP1
VF1
Sla5045
T V A S D A E T
WT 5′ACU GUG GCU UCC-90 nts-GAU GCU GAG ACC
m53SupA ACU GUG GCC UCC-90 nts-GAC GCC GAA ACC
Positive sense
4922
5018 5021 5024
m53
WT UGA CAC CGA AGG-90 nts-CUA CGA CUC UGG 5′
m53SupA UGA CAC CGG AGG-90 nts-CUG CGG CUU UGG
Negative sense
B
TCID50/mL (Log10)
0
1
2
3
4
5
WT backbone
M53 backbone
(WT) A
(SupA) G
U
C
(WT) A
(SupA) G
U
C
Position 4922 Sequence


A
NS7
VF1
VP1
Sla5045
Sla2
sgRNA Start 1
sgRNA Start 2
3' UUACCU CAGAAACACUUACUCCUACUCACU UAGCUA UACCU CAGAAACACUUACUCCUACUC 5'
NS7 Stop
VP1 Start
Sla5045
Sla2
B
TCID50/mL (Log10)
WT
POL FS
m53
WT/WT
WT/m53
m53/WT
m53/m53
Sla5045Dup
60kDa
NS7
C
TCID50/mL (Log10)
WT
WT/WT
WT/m53
m53/WT
m53/m53
Sla5045Dup
D
Sla5045Dup
WT
WT/WT
m53/WT
kb
RT
1.6
1.0
0.5
0.4
5052
WT ggggggGTGAATGAGGATGAGTGATGGCGCAGCGCCA
Sla5045Dup WT/WT ggggggGTGAATGAGGATGAGTGATGGCGCAGCGCCA
Sla5045Dup m53/WT ggggggGTGAATGAGGATGAGTGATGGCGCAGCGCCA

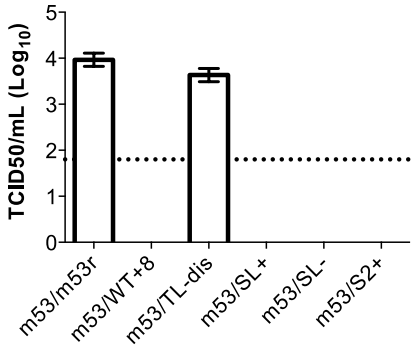
A

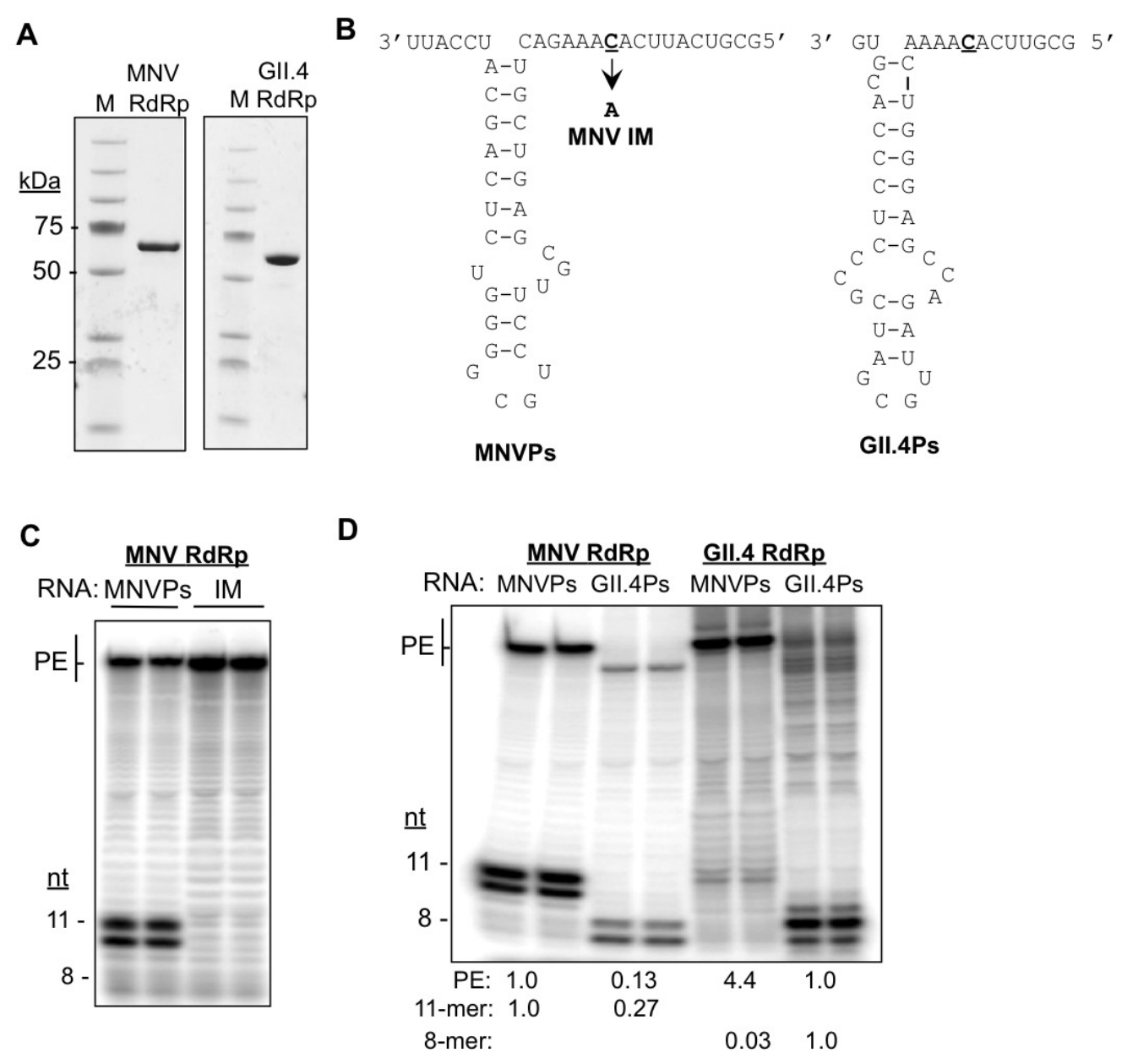

WT
m53r
WT+8
TL-Dis
SL+
SL-
s2

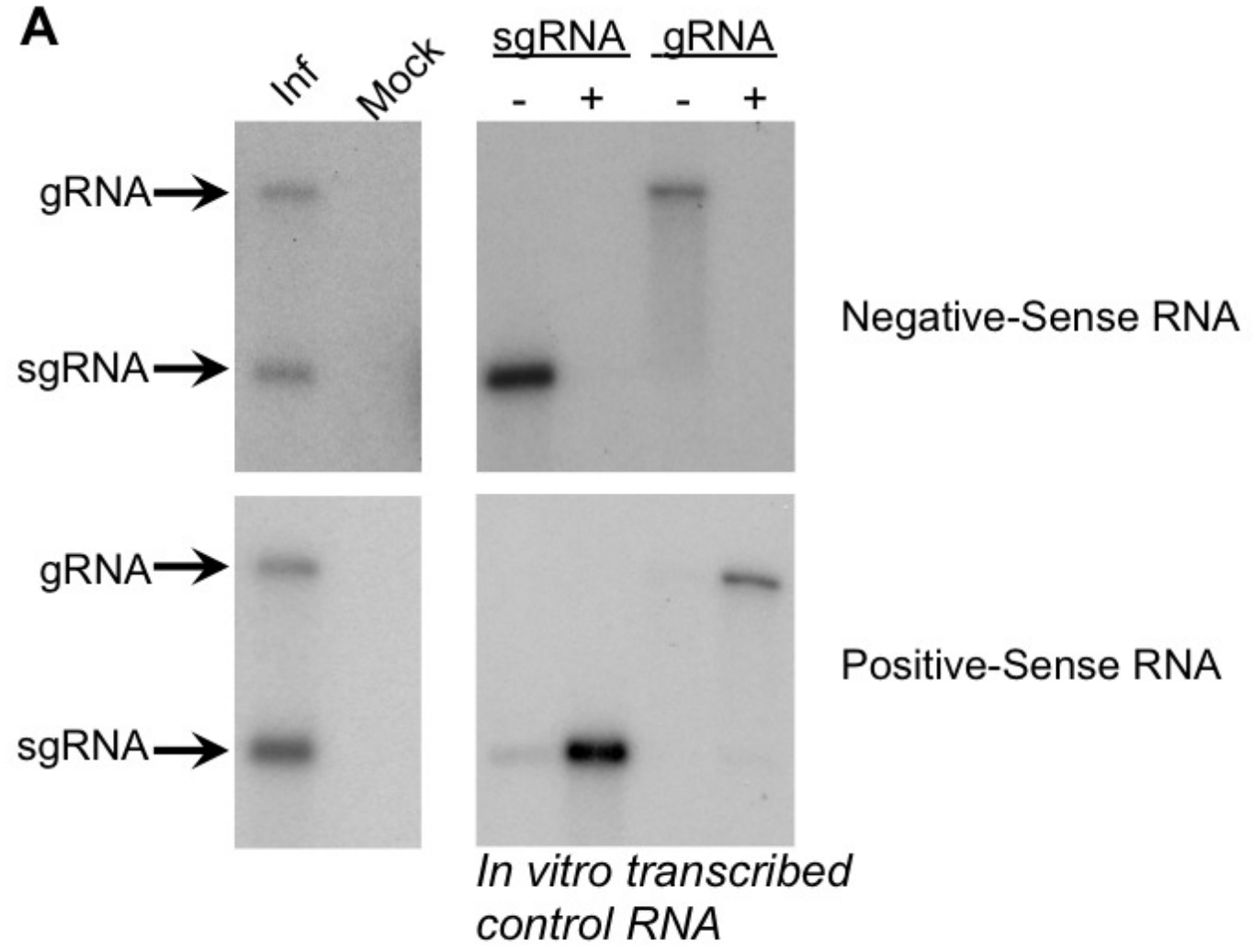
B

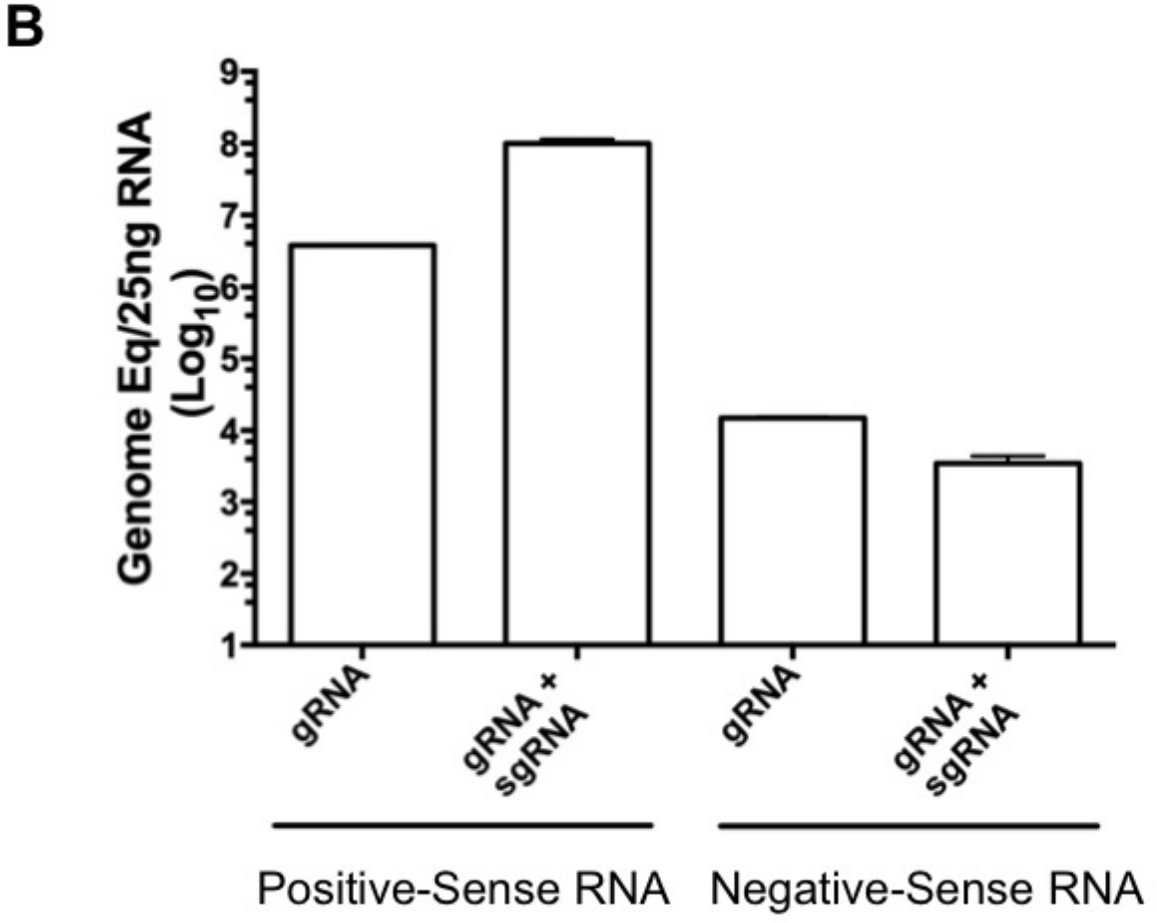

TCID50/mL (Log10)
WT
POL FS
m53/m53r
m53/WT+8
m53/TL-dis
m53/SL+
m53/SL-
m53/S2+
Sla5045Dup

C


TCID50/mL (Log10)
m53/m53r
m53/WT+8
m53/TL-dis
m53/SL+
m53/SL-
m53/S2+


A
MNV
GII.4
M RdRp
M RdRp
kDa
75
50
25
B
3′ UUACCU CAGAAACACUUACUGCG 5′
A
MNV IM
MNVPs
3′ GU AAAACACUUGCG 5′
GII.4Ps
C
MNV RdRp
RNA: MNVPs IM
PE
nt
11 -
8 -
D
MNV RdRp
GII.4 RdRp
RNA: MNVPs GII.4Ps MNVPs GII.4Ps
PE
nt
11 -
8 -
PE: 1.0 0.13 4.4 1.0
11-mer: 1.0 0.27
8-mer: 0.03 1.0


A
Inf
Mock
sgRNA
gRNA
-
+
-
+
gRNA
sgRNA
Negative-Sense RNA
gRNA
sgRNA
Positive-Sense RNA
In vitro transcribed
control RNA


B
Genome Eq/25ng RNA
(Log10)
9
8
7
6
5
4
3
2
1
gRNA
gRNA + sgRNA
gRNA
gRNA + sgRNA
Positive-Sense RNA
Negative-Sense RNA