## Supplementary Files

### Supplementary Material and Methods

#### Synthesis of 5'-(4-fluorosulfonylbenzoyl)guanosine (FSBG)

Guanosine monohydrate (875 mg, 2.90 mmol) was co-evaporated twice with anhydrous DMF and subsequently dissolved in DMPU with gentle warming. The clear solution was cooled in an ice bath, and 4-(fluorosulfonyl)benzoyl chloride (812 mg, 3.65 mmol) was added. After 15 minutes the mixture was warmed to room temperature and stirred for another 4 hours. Petroleum ether 40/60 (50 mL) was added and a white precipitate formed. The organic layer was decanted and the residue triturated twice with a 1/1 mixture of ethyl acetate/diethyl ether (2 x 50 mL). The residue was re-crystallized from MeOH/water and further purified by C18-RP-HPLC (Phenomenex Gemini C18, pore size 110Å, particle size 5  $\mu$ m, 150 x 21.2 mm, gradient 20 – 50% Acetonitrile in 0.1 % aqueous TFA, 20 mL/min) to yield the title compound as a white solid (232 mg, yield 17%) (Supplementary Figure 5).

#### Supplementary tables

 Table S1: Virus genome used for the bioinformatics analyses.

Virus name	Species	(Sub)family	Acronym	Accession number
Gill-associated virus	Gill-associated virus	Roniviridae	GAV	AF227196
Yellow head virus	to be established	Roniviridae	YHV	EU487200
Cavally virus	Alphamesonivirus 1	Mesoniviridae	CAVV	HM746600
Casuarina virus	to be established	Mesoniviridae	CASV	NC_023986
Dak Nong virus	to be established	Mesoniviridae	DKNV	AB753015.2
Hana virus	to be established	Mesoniviridae	HanaV	JQ957872
Nse virus	to be established	Mesoniviridae	NseV	JQ957874
Meno virus	to be established	Mesoniviridae	MenoV	JQ957873
SARS coronavirus Frankfurt 1	Severe acute respiratory syndrome-related coronavirus	Coronavirinae	SARS-CoV	AY291315
Rabbit coronavirus HKU14	Betacoronavirus 1	Coronavirinae	RbCoV_HKU14	JN874560
Murine hepatitis virus strain 2	Murine coronavirus	Coronavirinae	MHV-2	AF201929
Human coronavirus HKU1	Human coronavirus HKU1	Coronavirinae	HCoV_HKU1	AY884001
Betacoronavirus Erinaceus/VMC/DEU/2012	to be established	Coronavirinae	EriCoV	KC545383

Virus name	Species	(Sub)family	Acronym	Accession number
Bat coronavirus (BtCoV/133/2005)	Tylonycteris bat coronavirus HKU4	Coronavirinae	BtCoV/133/2005	DQ648794
Bat coronavirus HKU5-1	Pipistrellus bat coronavirus HKU5	Coronavirinae	BtCoV_HKU5	EF065509
MERS coronavirus EMC/2012	to be established	Coronavirinae	MERS-CoV	JX869059.2
Bat coronavirus HKU9-10-2	Rousettus bat coronavirus HKU9	Coronavirinae	BtCoV_HKU9	HM211101
Bat coronavirus CDPHE15/USA/2006	to be established	Coronavirinae	BtCoV_CDPHE15	KF430219
Human coronavirus NL63	Human coronavirus NL63	Coronavirinae	HCoV-NL63	AY567487
Miniopterus bat coronavirus HKU8	Miniopterus bat coronavirus HKU8	Coronavirinae	BtCoV_HKU8	EU420139
Rhinolophus bat coronavirus HKU2	Rhinolophus bat coronavirus HKU2	Coronavirinae	BtCoV_HKU2	EF203064
Bat coronavirus 1A	Miniopterus bat coronavirus 1	Coronavirinae	BtCoV_1A	EU420138
Alpaca respiratory coronavirus	Human coronavirus 229E	Coronavirinae	ACoV	JQ410000
Bat coronavirus (BtCoV/512/2005)	Scotophilus bat coronavirus 512	Coronavirinae	BtCoV/512/2005	DQ648858
Porcine epidemic diarrhea virus	Porcine epidemic diarrhea virus	Coronavirinae	PEDV	KC140102
Rousettus bat coronavirus HKU10	to be established	Coronavirinae	BtCoV_HKU10	JQ989271
Mink coronavirus strain WD1127	to be established	Coronavirinae	MCoV	HM245925
Feline coronavirus UU2	Alphacoronavirus 1	Coronavirinae	FCoV_UU2	FJ938060
Infectious bronchitis virus	Avian coronavirus	Coronavirinae	IBV	KC008600
Bottlenose dolphin coronavirus HKU22	Beluga whale coronavirus SW1	Coronavirinae	BdCoV_HKU22	KF793824
Sparrow coronavirus HKU17	to be established	Coronavirinae	SpCoV_HKU17	JQ065045
Munia coronavirus HKU13- 3514	Munia coronavirus HKU13	Coronavirinae	MuCoV_HKU13	FJ376622
Common-moorhen coronavirus HKU21	to be established	Coronavirinae	CMCoV_HKU21	JQ065049
Bulbul coronavirus HKU11- 934	Bulbul coronavirus HKU11	Coronavirinae	BuCoV_HKU11	FJ376619.2

Virus name	Species	(Sub)family	Acronym	Accession number
Thrush coronavirus HKU12- 600	Thrush coronavirus HKU12	Coronavirinae	ThCoV_HKU12	FJ376621
White-eye coronavirus HKU16	to be established	Coronavirinae	WECoV_HKU16	JQ065044
Night-heron coronavirus HKU19	to be established	Coronavirinae	NHCoV_HKU19	JQ065047
Wigeon coronavirus HKU20	to be established	Coronavirinae	WiCoV_HKU20	JQ065048
Porcine torovirus	Porcine torovirus	Torovirinae	PToV_SH1	NC_022787
Breda virus	Bovine torovirus	Torovirinae	BRV-1	AY427798
White bream virus	White bream virus	Torovirinae	WBV	DQ898157
Fathead minnow nidovirus	to be established	Torovirinae	FHMNV	GU002364.2
Ball python nidovirus	to be established	Torovirinae	BPNV	NC_024709
Possum nidovirus	to be established	Arteriviridae	WPDV	JN116253
Simian hemorrhagic fever virus	Simian hemorrhagic fever virus	Arteriviridae	SHFV-LVR	AF180391
Simian hemorrhagic fever virus	to be established	Arteriviridae	SHFV-krtg2	JX473847
Simian hemorrhagic fever virus	to be established	Arteriviridae	SHFV-krtg1	JX473848
Simian hemorrhagic fever virus	to be established	Arteriviridae	SHFV-krc1	HQ845737
Porcine reproductive and respiratory syndrome virus	Porcine reproductive and respiratory syndrome virus	Arteriviridae	PRRSV-2	JX138233
Porcine reproductive and respiratory syndrome virus	Porcine reproductive and respiratory syndrome virus	Arteriviridae	PRRSV-1	GU737264.2
Lactate dehydrogenase- elevating virus	Lactate dehydrogenase-elevating virus	Arteriviridae	LDV-C	L13298
Lactate dehydrogenase- elevating virus	Lactate dehydrogenase-elevating virus	Arteriviridae	LDV-P	U15146
Equine arteritis virus	Equine arteritis virus	Arteriviridae	EAV	DQ846750

Table S2. Multiple sequence alignment of NiRAN of nidoviruses (see a separate file in fasta format).

# Supplementary figures

JPRED_R PSIPRED_R JPRED_M JPRED_C PSIPRED_C JPRED_C JPRED_A PSIPRED_A		00000000000 000000000000 000000000000			
	10 20 30 1234567890123456789012345678	40 50 60 901234567890123456789012345	70 80 567890123456789012345678	90 100 11 901234567890123456789	.0 120 130 012345678901234567890
<u>Roniviridae</u> GAV YHV Mesoniviridae	HTSGIFCKPRASSSTLIINGKNRH) HTSGIFCKPRASSSTLIINGKNRH)	KTKHSLTREIKIAQTLS KTKHSLKR <mark>B</mark> IKTAQALS	HIPEAIQFKKDSHGY HIPEAVQFHKDEHGY	Y <mark>R</mark> .EISQFSLADVLHG.FANQ Y <mark>R</mark> .EISQFSLADVLHG.FANQ	)IEPDFLAKYTNERNIKVSKTT 2IEPDFLAKYTNERNIKVSNVT
CAVV CASV DKNV HanaV NseV	HTIDISCN.KTSISVIDEVNNVNN HTIDISCN.KTSTSYIDEANRIVNN HTIDISCN.KTSLSVIDEYNQTVNN YTIDISCN.KTSTSVIDSKNNIVNN YTIDISCN.KTSTSVVNCBOOIVN	MIKQH. IVKBYKIYEMLIN. NIKAK. VAKBYNVYEMLIN. KIKEH. IVKEYEIYETLIN. MIKNN. IVKBYNIYEMLIN. MIKNN. VKYBOLVYDTLIN.	QYPNLFLIEHKLVNFTIPHL QYPNIFLIKHKIVHDTIPHL QDSNLFLIKHKLVHATIPHL QYSDLFLIEHKMVTSTIPHL OESDLFLIEHOLETOPIPYL	LRYNMTALSFADLYGL.IKEE LRYNNTALSFADLYGL.IKEE LRYNMTALSFADLFGL.IKEE LRYNMTALSFADLFGL.IKDE KRLDMTALSVADLYGV.IKDE	NWHPIYDTLPQVTYHKINDDL NWHPIYDTLPQVTYHKIDPDL NWHPIYDTLPQVTYHKISNDL NWHPIYDTLPQVTYHNISNDL NWHPIYDTLPQVTYHNIDNDL NWHPIYDTLPQVKYHKLDODL
MenoV Coronavirinae SARS-CoV RbCoV HKU14	YTLDISCN.KTSVSYITPNSEIVNV VAGFAKFL.KTNCCRFQEKDEEGNLLDSYFV VAGIGLHL.KUNCCRFORL.DESGNKMDRFFV	KVKPNIEKEHTIYEMLTKT KRHTMSNYQHEETIYNLVKD	SKIDLFLIKHTFMRQSINYL CPAVAVHDFFKFRVDGDMVPHI CRVVAEHDFFTFDVEGSRVPHI	MRHNMTALSLADLYGV.IKSE SRORLTKYTMADLYYA.LRHF VRKDLTKYTMLDLCYA.LRHF	NWKPLYDTLPQVTYHKISPTL DEGNCDTLKEILVTYNCCDDD FDRNDCSLLCDILSMYAGCEOS
MHV-2 HCoV_HKU1 EriCoV BtCoV/133/2005	RAGIGLYY.KVNCCRFQRADEDGNTLDKFFV RAGIGLYY.KVNCCRFQRI.DDDGNKLDKFFV VAGIGKHY.KTNTCRFVEL.DDQGHKLDSYFV VAGIGKYY.KTNTCRFVQV.DDEGHKLDSYFI	KRTNLEVYNKEKECYELTKE. KRTNLEVYNKEKTYYELTKS. KRHTMENYELFRCYDLLKA. KRHTMSNYELEKRCYDLLKA.	. CGVVAEHEFFTFDVEGSRVPHI . CGVVAEHDFFTFDIDGSRVPHI . CDSVAAHDFFTFDIDGSRVPHI . CDAVAHDFFVFDVDKTKTPHI	VRKDLSKYTMLDLCYA.LRHF VRKNLSKYTMLDLCYA.LRHF VRQRLTEYTMMDLVYA.LRHF VRQSLTEYTMMDLVYA.LRHF	DRNDCSTLKEILLTYAECDES DCNDCSVLCEILCEYADCKES DQNNCEVLKTILVRYGCCEES DQNNCEVLKSILVKYGCCEQS
BtCoV_HKU5 MERS-CoV BtCoV_HKU9 BtCoV_CDPHE15 HCoV-NL63	VAGIGKYY.KTNTCRFVEV.DDEGHRLDSFFV VAGIGKYY.KTNTCRFVEL.DDGGHRLDSYFV VAGFGLHL.KNNCCRYQEL.DAEGNQLDSYFV VACIGGFL.KVNCVRFRNTDMHDAFYV VSFLGKCL.KMNCVRFRNADLRDGYFV	RHTMENYELEKRCYDLVKD RHTMENYELEKHCYDLLRD RHTESNYLLEQRCYEKLKD. RCTKSVMDHEQSIYDKLAK RCTKSVMEHEQSMYNLLNF	CDAVAVHDFFIFDVDKVKTPHI CDAVAPHDFFIFDVDKVKTPHI CGVVARHDFFKFNIDGEMTPHV CDAVAKHDFFTWKEGRSIYGNV SGALAEHDFFTWKDGRVIYGNV	VRQRLTEYTMMDLVYA.LRHF VRQRLTEYTMMDLVYA.LRHF SRERLTKYTMADLVYS.LRHF CRCNLTKYTMMDLCFA.LRNF SRHNLTKYTMMDLVYA.MRNF	DQNNCEVLKSILVKYGCCDAS DQN.SEVLKAILVKYGCCDVT DNNNCDMLKEILVLRGCCTED DERNCQTLKEILVLTGACDES FDEQNCDVLKEVLVLTGCCDNS
BtCoV_HKU8 BtCoV_HKU2 BtCoV_1A ACoV BtCoV/512/2005	VACIGKFV.KVNCVRFKNADKHDAFYV VACIGKFL.KVNCVRLKNLDKHDAFFVV VASIGKFV.KVNCVRFKNLDKHDAFFIV ASFIGKNL.KSNCVRFKNADKDDAFYIV VACISKFL.KVNCVRLKNLDKHDAFYIV	KRCTKSVMEHEQSIYDALKD. RCTKSVMEHEQSMYNKLSG. RCTKSVMEHEQSIYDILKY. KRCTKSVMEHEQSMYNLLKG. KCTKSVMEHEQSIYNLLSD.	CGAVSPHDFFVWKDGRSVYGNI SNALAVHDFFTWKDGRSIYGNV SGALAIHDFFLWKDGRAIYGNI CNAVAKHDFFTWHEGRTIYGNV CGAVAKHDFFTWKEGRSVYGNV	ARHDLTKYTMMDLVHA.LRNF CRQDLSKYTMMDLCYA.LRNF CRQDLTKYTMMDLVNA.LRNF SRQDLTKYTMMDLCFA.LRNF CRQDLTEYTMMDLCYA.LRNF	DEKNCETLKEILVISGACDSS DERNCETLKEILVLTGCCDQS DEKNCEVLKEILVLTGACDSS DEKDCEVLKEILVLTGCCGTD DENNCETLKKILVVVGACDES
PEDV BtCoV_HKU10 MCoV FCoV_UU2 TBV	VACLGKFL.KVNCVRLKNLDKHDAFYV VACLGKFL.KTNCVRFRNKLHNDAYFV VACIGKFL.KTNCSRFKNLDAHDAYIV VACIGKFL.KTNCSRFRNLDKHDAYYV SAGMFINLKEPNGBFOFUPDTEDGN.LEVCOSFUV	RCTKSVMEHPOSIYSRLEK RCPKSVMEHPOSIYDILKD RCRKSVMDHEQVCYNALKH KRCTKSVMDHEQVCYNDLKD CTTSSVMEHPOVCYNDLK	CGAVAEHDFFTWKDGRAIYGNV SGAIATHDFFVWKDGRMIYGNI SNALASHDFFEYSEGRHVFGNV SGVVAEHDFFSYKEGRCEFGNV SEVTADHDFFVENKN IYN	CRKDLTEYTMMDLCYA.LRNF SRQDLTKYTMMDLCYA.LRNF CRRNLTKYTMMDLCYA.LRNF ARKDLTKYTMMDLCYA.IRNF SROBLTKYTMMDLCYA.IRNF	DENNCDVLKSILIKVGACEES DEKNCETLKEILVITGACDQS DEKNCDVLKEILVLTNCCDST DEKNCEVLKEILVTLGACNES DEKNCEVLKEILVTLGACNES
BdCoV_HKU22 SpCoV_HKU17 MuCoV_HKU13 CMCoV_HKU21	TAGMYASL.KHNCARFQELDENDDEIDSFFV TSGIFLST.KTNCARFKTVRSALPLPN.KGEVDLYFV TSGIFLST.KTNCARFKTQRCNLPIPY.KGLVDLYFV TSGIFLST.KTNCARFKTQRCNLPIPY.SGVVDLYFV	KOTTPHNFEHEEKCYLDLK. KOCAQKVFEIEEKCYNALSAELYTTDI SOCSLSVFETEACYNAFDKALITTEI NOCSTSSFELEEKCYNLFSSEFKSTDI	. ADCVAVHDFFRF.EGMYNI ETFGVLAKTEFFKF.DKIPNV DTFGVLAKTEFFKF.DKIPNV DTFGVLAKTEFFKF.DKIPNV	CRORLTKYTMMDLCYA.FRHF NROYLTKYTLLDLAYA.LRHI NROYLTKYTLLDLAYA.IRHI NRHNLTKYTLLDLAYA.IRHI	DPNDCDVLKEILVVKGCCEWD STSK.DVIQEILTTMCGTPED STSR.DVIKEILITICGTPED LSTSK.DVIQDILITMCGTPQT
MCOV_HKU12 WECOV_HKU12 WECOV_HKU16 NHCOV_HKU19 WiCOV_HKU20 Toroviringe	ISGIFLSI. KINCSRFKIIRONLELPT. NGSVELIFY NSGIFLSI. KINCSRFKIIRONLELPT. DKAVELYFY TAGIFKSI. KNNCRFKSINLEVR. IGNVDLFFY NVGIFTNI. KINCARHRVA	KOCSQQVFEIBERCYNLLADNIKSTPI KOCSQQVFEIBERCYNLLADNIKSTPI KOCSQUVFEIBEACYNKFDDKLKSTDI KOCTENVFRABEQCYNMLPKSIVSTDI KOCDEQQFRK <mark>B</mark> EYFYSVLPQHF	STFGVLARTEFFKF.DKIPNV STFGVLARTEFFKF.DKIPNV KTFGVLAKTDFFKF.DKIPNV DKFSCVAYHDFFKF.DGVNNV .KGDIVPQHDFFKF.DGTPNV	NRQYLIKYIILDLAYA.LRHI NRQYLIKYILLDLAYA.LRHI NRQYLIKYILLDLAYA.LRHI VRRHLIKYILLDLVYA.LRHI VRQYLIKYILLDLVYA.LRHI	SISK.DVIRELLIIICGITEE "STSK.DVIREULIITMCGTEE "STSK.DVIREILIITMCGTSED "STSQ.EIIQEILTTMCGTSEE "SDSV.ELLREILQTHCGTKDD
PTOV_SH1 BRV-1 WBV FHMNV BPNV Arteriviridae	NAGIVKVN.KSNTHSVE	RVKDQHEFALSRTSF. RVKDQHEFALARTAF. KDTGTAHEIVAKQLLAKG KTKQTAHELATATVLTKQG KTGTEEDLRHEYNQYLALRDL	.LPSIIPHHMVSY.NNEWFL LPSIIPHHMVHQ.NGEWFL LP.VLQHARFNH.DGTDYL LP.VLKHDRINY.NKQDIL .IA.MPEHKLIQLENGSYIL	VRGPTTQWSLGDLVYA.IWL. VRGPTTQWSLGDLVYA.IWL. IRYYTTPYSLGDLVYA.YMV. VRYYTTPYSLGDLVYA.YQL. IRGPVTEKSLGDLVYS.HLH.	GDQDYLAECGFVFNPSRD. .GDQAYLDECGFVFNPSRD. .GDFKHMLIALDITDE. .GDLQHMQLVLNLDDH. .NQTEDAVEIPDK.
WPDV SHFV-LVR SHFV-krtg2 SHFV-krtg1 SHFV-krc1 PRRSV-2 PRRSV-1 LDV-C LDV-P	I.TIVPRS.RTFIC. GAFN) K.IITHHA.RTRAF. SSIDF A.IVEHKS.RTRAF. NGCDI X.IVDHHS.RTRAF. ADIDI K.IVDHS.RTVAF. ADIDI K.IVVHS.RTFTL. GPVNI K.IKYHS.RTFSI. GDVNI K.IVRYHS.RTFSI. GDVNI	RVLP. TKBEINN. RVVS. PDBAMRTARL. RAVS. PBAANTIRL. NAVS. PEANRTIRL. VIVD. ANKLDRINRL. VAS. EVBLKDAVEH. VTS. EVBVKKSTE. RVMG. REBYGRTVCK. RVMG. FDBYRTMCK.	TPAGSYDLVEG CM SPSPQFIAFSSDDK. FLL SPTPQFVVAQLSDG. YLI SPTPQFVVARLADG. YLI SIEPQPVVATITDGK. FVI . NQCFVAPTVDGK. VVL . QGHAVVANLCSG. VVL . QGHCLVANLVDG. VVV . PGHLLVAKLTDG. VVV	INPHYP.SLIDKLLA.AIH. LRHPP.SLIDVLTK.GLD. MRHPP.SLIDVITK.GFD. MRHPP.SLIDVITK.GFD. LRHPP.SLIDVITK.GFD. LSAVP.SLIDVLTK.GLD. LSAVP.SLIDVLLKPGD. MRHPP.SLIDVLLKPGD. MRHPP.SLIDVLLKPGD. MRHPP.SLIDVLLKPGD.	
EAV motifs	KGFVALHS.RTMFL.AA	₩FVCDE型FTKTPK	DTLLGYVRACPGYWFI	FRTHRSLIDAYWD.SME. ************************************	
	preA <sub>N</sub>	A <sub>N</sub>		BN	

JPRED_R PSIPRED_R JPRED_M PSIPRED_C JPRED_C JPRED_A PSIPRED_A PSIPRED_A							
	140 12345678901234567	150 160 7890123456789012345	170 180 56789012345678901	190 2345678901234567	200 210 89012345678901	220 2 23456789012345678	30 901 genome coordinates (nt)
<u>Roniviridae</u> GAV YHV Mesoniviridae	WLC	NHKSCNILKMDIFDYTYT NHKSCNILKMDIFDYTYT	CYTKAHSFALQACSTYI CYTKAQSFALQACTTYI	NFDITPDNICPEG.VY NFDITPDNICPEG.VY	FETFRIGT.RDPIK	ALNAVFYCIER.HWFFS ALNAVTYCIER.HWFSA	GLS 12389-12946 GLS 12428-12985
CASV CASV DKNV HanaV NseV MenoV	LLKIKS.HTPSPQHTCC LFKIKQ.HTPSPQHTCC LIKIKS.HTPSPQHTCC LLKIKQ.HTPSPQHTCC LKKIKE.HTPCPQHTCC LNNIQQ.HSSAPQHTCC	CMLCRRFLVEFGLLLHKL CMLCRRFLAEFGLLLHKL CMLCRRFLVEFGLLHKL CMLCRRFLAEFGLLLHKL CLLCRRFLAEFGLLMHKL CMLCRRFLAEFGLLMHKL	NYKVFETTRAILT.HY NFKVYETTTNIYA.HY NYKVFETTTRLLQ.GY NYKVFETTTRKMLT.HY NFKVYETTTTLR.SY NKQVFSTINFMFQ.HY	DFVLTADNVDLNG.IL DFVLTADNIDLNG.IL EFVLTADNIDLNG.IL DFVLTADNIDLNG.IL EFVLTADNVDLNG.IL DFVLTSDNVDLNG.IL	FEDYKLKK.STIAH FEDYILREEVLVNI FEDYIPRA.EYFVC FEDYVPKKYIDINI FEDYLQLD.EPRNV FEDYTKTE.YTRTL	DVKSQLR.IMQ.PYYHA DVKSQLR.IMQ.PYYHT DVKYQLR.IMQ.PYYHT DVTSQLR.IMQ.PYYHY ECVDQLR.KMQ.PYYHS ELKDQLR.KMQ.PYYHT	LYS 7925-8518 LYS 7815-8411 LYS 8098-8691 LYS 7916-8512 LYS 7864-8457 LYS 7777-8373
Coronavirinae SARS-CoV RbCoV_HKU14 MHV-2 HCoV_HKU1 EriCoV BtCoV/133/2005	YFN.KKDWYDFVEI YFT.QKDWYDFVEI YFQ.KKDWYDFVEI YFS.KKDWYDFVEI YFD.NKLWFDFVEI YFD.NKLWFDFVEI	NPDILRVYANLGERVRQS NPDIINVYKKLGPIFNRA NSDIINVYKKLGPIFNRA NPDIINIYKKLGPIFNRA NPDVIRVYHKLGELVRRA NPSVIGVYHKLGERIROA	LLKTVQFCDAMRDAGIY LVNTAEFADALVEAGLY LLNTAKFADTLVEAGLY LLNTVSFADTLVKVGLY MLSTVKFCDHMVKSGLY MLNTVKMCDHMVKSGLY	/GVLTLDNQDLNGNWY /GVLTLDNQDLNGMWY /GVLTLDNQDLYGQWY /GVLTLDNQDLYGQWY /GVLTLDNQDLNGKWY	FGDFVQVA.PGCGV FGDYVITA.PGCGV FGDFVKTV.PGCGV FGDFIQTA.PGFGV FGDFVVTQ.PGAGV	PIVDSYYSLLM.PILTI AVADSYYSYMM.PILAM AVADSYYSYMM.PMLTM AVADSYYSYMM.PMLTM AIVDSYYSYLM.PVLSM AIVDSYYSYLM.PVLSM	TRA 13494-14120 CCHA 13618-14244 ICHA 13538-14164 ICHV 13603-14229 ITDA 13689-14315 ITDA 13669-14295
BtCoV_HKU5 MERS-CoV BtCoV_HKU9 BtCoV_CDPHE15 HCoV-NL63 BtCoV_HKU8	YFD.NKLWFDFVEI YFE.NKLWFDFVEI YFD.RKDWYDPVEI YFE.NKLWFDPVEI YFD.SKGWYDPVEI YFD.NKNWYDPVEI	NP NVISVYHKLGERIRQA NPSVIGVYHKLGERVRQA NPDIIRVYHKLGETVRKA NEDIHRVYAKLGVVVARA NEDIHRVYASLGKIVARA NEDIHRVYAKLGCVVANA	VINTVKFCDQMVKSGI ILNTVKFCDHMVKAGI VISAVEMADAMVEQGI MLNCVKLCDAMVKAGI MLKCVALCDAMVAKGV MLKCVALCDAMVAKGV	/GVLTLDNQDLNGKWY /GVLTLDNQDLNGKWY IGVLTLDNQDLNGKWY /GVLTLDNQDLNGKFY /GVLTLDNQDLNGNFY /GVLTLDNQDLNGNFY	FGDFVITQ.PGAGV FGDFVITQ.PGSGV FGDFIEGP.AGAGV FGDFVPSL.EGMGV FGDFVVSL.PNMGV	AIVDSYYSYLM.PVLSM AIVDSYYSYLM.PVLSM AVMDTYYSLAM.PIYTM PLCTSYYSYMM.PIMGM PCCTSYYSYMM.PIMGI PLATSYYSYLM.PVMGM	Image: 13790-14416           ITDC         13538-14161           ITDC         13538-14161           ITDM         12974-13600           ITNC         12625-13239           ITNC         12532-13146           ITNC         13028-13642
BtCoV_HKU2 BtCoV_1A ACoV BtCoV/512/2005 PEDV BtCoV_HKU10	YFD.NKVWYDPVE YFD.NKSWYDPVE YFD.NKSWYDPVE YFD.NKLWFDPVE YFD.NKVWFDPVE YFD.NKFWFDPVE	NEDLHRVYALLGQRVANA NEDIHRVYAKLGDVIANA NEDIHRVYAALGTVVANA NEDVHRVYAKLGTIVARA NEDIHRVYALLGTIVSRA NEDLHRVYALLGKIVANA	MLKCVKLCDEMVTKGV MLKCVALCDAMTEKGI MLKCVALCDEMVLRGV MLKCVKYCDAMVEQGI MLKCVKFCDAMVEQGI MLKCVRLGDAMVKHGV	/GVLTLDNQDLNGNFY /GVLTLDNQDLNGNFY /GVLTLDNQDLNGNFY /GVLTLDNQDLNGDFY /GVVTLDNQDLNGDFY /GVVTLDNQDLNGDFY	FGDFVDVM.PGMGI FGDFVTSI.PGVGV FGDFVTSV.FGMGI FGDFVTSV.KGMGV FGDFTCSI.KGMGI	PCCTSYYSYMM.PIMTM PVCTSYYSYMM.PAMGM PYCTSYYSYMM.PVMGM PICTSYYSYMM.PVMGM PICTSYYSYMM.PVMGM PLCTSYYSYMM.PVMGM	ITNC 12539-13153 IANC 13141-13755 ITNC 12625-13239 ITNC 12743-13357 ITNC 12743-13357 ITNC 12709-13223 ITNC 12713-13327
FCOV_UU2 IBV BdCoV_HKU22 SpCoV_HKU17 MuCoV_HKU13 CMCoV_HKU11	FFD.NPDWIDPVEI FFE.NKDWFDPVEI HPKWFEENKDWYDPIEI YFD.QPNWYDPVEI WFG.ENWFDPIEI WFG.ELWYDPIEI	NEAIHEVYARLGHIVANA NPKYYAMLAKMGPIVANA NPDWFSLISRLGPIFQRA NPSFYKEFHKLGDILNRC NPTFYNEFHKLGVLINRC	MIKCVALCDAMVEKGY LINAVEF GNIMVEKGY LINAVEF GNIMVEKGY VINANKF ASACIDAGI VINANKF ASACIDAGI VINANKF ASACIDAGI UNANKF ASACCPD VGLV	IGIIILDNQDLNGNFI IGIIILDNQDLNGNFY IGVIILDNQDLNGNFY IGVVILDNQDLNGNFY IGILTPDNQDLLGQIY IGILTPDNQDLLGQIY	FGDFVSTI.GGGGC FGDFQKTV.PGAGV FGDFKKVL.PGCGV FGDFIITQ.PGNGC	ACVISIISIAM.PLMGM ACVISIYSYMM.PLMGM PVFDTYYSYMM.PLIAM PVTTSYYSYMM.PCLTA VDLASYYSYLM.PIMSM CDMSSYYSYLM.PIMSM VDLESYYSYLM PLMSM	1252         1252         12006           TISC         1242         13054           TIDA         12445         13083           CDA         12480         13091           TITHM         11520         12167           TITHM         11668         12335           TITHM         11167         11804
BuCoV_HKU11 ThCoV_HKU12 WECoV_HKU12 WECoV_HKU19 WiCoV_HKU20 Toroviringe	WFGDSWFDPIE WFD.EQWYDPIE WFG.DLWYDPIE WFV.DGWYDPIE FFE.GVWYDPIE	NPTFYREFHKLGSVLNRC NQTFYREFHKLGSILNRC NPTFYREFHKLGSILNRC NPTFYDEFHKLGSLINNC NVHFYTEFHKLGEIINKC	VINANAFAKACSELGI VINANTFAKACADAGL VINANAFAKAAADSGV VVMANKFADTCKTVGL VLKANTFVSTVNKLGF	/GILTPDNQDLLGQIY /GILTPDNQDLLGQIY /GILTPDNQDLLGQIY /GILTADNQDLGQIY .GVLTPDNQDLGGXIY	FGDFIITQ.PGNGC FGDFIITQ.QGNGC FGDCILTE.PGNGC FGDFVVTQ.PGNGC FGDFVVTQ.SK.S.VGC	VDLSSYYSYLM.PIMSM VDLSSYYSYLM.PIMSM IDLSSYYSYLM.PIMSM IEMDAYLSYIM.PSMSM VDMESYYSYLM.PAMSI	IIII         IIII         IIII         IIII         IIIII         IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
PTOV_SH1 BRV-1 WBV FHMNV BPNV Arteriviridae	EFFEI EFLDI TCLDI RVLDI GKPE	DAKQRSHLAFLLEPA DANQRSYLANLLEPA PGNYSSYY.NFKEQLRNK PGFYSEYH.FFKSEIRKV ASEQLANELAQQLEPA	ILSFCDIFSYIKKFGL ILSFCEIFHCVKGCQVI LASVIPNVNKILAAEL LTKCIPNVNIILKANLI FDNLDKLRDWCYKYDV	YKLTLDNMDLCGNLY YKITLDNLDLKGQLY LAITLDNLDLNGFLY MAITLDNIDLNGFLY QLPITLDNIDLDGQLY	FGDYPCPNV FGDYPCPNK FGDYPTNGK FGDYPQKEI FGDMGTSSH	VDNQSALFVLA.EVWSM VDNQSALFVLA.EVWSM VTNYHVVSCMR.QIATF PSNRHVIEAIR.QLAVF .NIDIALSDLM.RLWSI	ITRK 14139-14663 ITRK 14292-14816 CSL 14666-15193 CAL 15435-15962 ITNR 18564-19085
WPDV SHFV-LVR SHFV-krtg2 SHFV-krtg1 SHFV-krc1 PRRSV-2 PRRSV-1 LDV-C LDV-P EAV			GSCYVTPTVP AT.CRI AQ.YQ' AE.HI' AE.WQI ASP.KL IAP.GI AD.LI' AD.LI' AE.FFI CVYALP.TISI	<i>I</i> I LK N P G D T D G C HE & 2P L H S P G D Q G I D G Y L & A Q H G P G D Q G I D G Y L & <i>A</i> Q H G P G D Q G I D G Y L & A Q H G P G D Q G I D G Y L & A C H G P G N T G I D G Y L & A C H G P G N T G I D G T L & P G H A A G N M G Y D G S V & S P T H G P G N T G I H F V & S P T H G P G N T G I H F V & S P T G Y S P G D V A V T G E R &	FEADHSKESR           FEADHSKEAI           FEADHSKDLV           FEADHSKEEL           FEADHSKEEL	VFAQDIATALRNEACH %LSNQIISACAARRGDA KFSAEIIAACSARRGDA KFSAEIIAACSARRGDA TLSAQIIQACDMRRGDA LLSQIIQACDMRRGDA ELSEQIITACSIRRGDA RLSQIITACSIRRGDA RLTADLVHAFQGFHGAS	TTIG 2889-3233 PGC 6587-6937 PRY 6412-6759 PAY 6405-6752 PPAC 6412-6762 PEI 7350-7694 PPNL 7388-7732 PSL 6902-7243 PAL 6837-7178 YSY 5468-5839
motifs					****		

motifs

\*\*\*\*\*\*

C<sub>N</sub>

**Supplementary Figure 1:** The nidovirus-wide NiRAN MSA encompassing conserved motifs. Virus names and accession numbers are listed in Table S1. Fully and partially conserved residues are depicted in white font with red background and red font, respectively. Sequence motifs are indicated by stars. Secondary structure predictions are shown on the top of the MSA. The name of each prediction indicates software used (Jpred 3 or PSIPRED) and which family-specific NiRAN MSA (R, *Roniviridae*; M, *Mesoniviridae*; C, *Coronaviridae*; A, *Arteriviridae*) were used to produce it. The plot was generated with ESPript.



**Supplementary Figure 2:** Sequence variation, domain organization, and secondary structure of NiRAN-RdRp-containing proteins of nidovirus families. For each family, the similarity density plot obtained for the MSA of proteins including the NiRAN and RdRp domains is shown. To highlight the

regional deviation of conservation from that of the MSA average, areas above and below the mean similarity are shaded in black and gray, respectively. Sequence motifs of NiRAN and RdRp are labelled. Uncertainty in respect to the domain boundary between NiRAN and RdRp is indicated by dashed horizontal lines. Domain boundaries used for all bioinformatics analyses are indicated by dashed vertical lines. Below each similarity density plot predicted secondary structure elements are presented in gray for  $\alpha$ -helices and black for  $\beta$ -strands.



**Supplementary Figure 3:** Pairwise MSA-based HMM-HMM comparison of NiRANs of different origins. Each MSA of NiRAN was converted into an HMM profile, all possible pairs of different HMMs were aligned using HH-align. The label at the left and top of each plot specifies the group of viruses used as query and target in HMM-HMM comparison, respectively. Below each dot-plot the confidence (%) of the target being homologous to the query and the E value of the top local hit are shown in black and green, respectively. The four plots highlighted with grey background are also presented in Fig. 3.

В

с	F * S O V V V					
	GLDQDKVLDL	ΤΝΜΙΚΥΟΡΤΕ	LSSKDKAKAR	QLAHLLLDLA	NPVEAVNQLN	5 0
	LRAPHIFPGD	V G R R <b>T F A D S K</b>	(D K)G F V A L H S R)	TMFLAARDFL	FNIKFVCDEE	100
	FTKTPK <b>DTLL</b>	GYVRACPGYW	FIFRRTHRSL	IDAYWDSMEC	VYALPTISDF	150
	DVSPGDVAVT	GERWDFESPG	GGRAKRLTAD	LVHAFQGFHG	A S Y S Y D D K <b>V A</b>	200
	AAVSGDPYRS	DGVLYNTRWG	ΝΙΡΥSΥΡΤΝΑ	LEATACYRAG	CEAVTDGTNV	250
	IATIGPFPEQ	QPIPDIPKSV	LDNCADISCD	AFIAPAAETA	LCGDLEKYNL	300
	STQGFVLPSV	FSMVRAYLKE	EIGDAPPLYL	PSTVPSKNSQ	AGINGAEFPT	350
	KSLQSYCLID	DMVSQSMKSN	LQTATMATCK	RQYCSKYKIR	SILGTNNYIG	400
	LGLRACLSGV	ΤΑΑΕΩΚΑGΚ	GSPIYLGKSK)	FDPIPAPDKY	CLETDLESCD	450
	RSTPALVRWF	ATNLIFELAG	QPELVHSYVL	NCCHDLVVAG	SVAFTKRGGL	500
	SSGDPITSIS	NTIYSLVLYT	QHMLLCGLEG	YFPEIAEKYL	DGSLELRDMF	550
	KYVRVYIYSD	DVVLTTPNQH	YAASFDRWVP	HLQALLGFKV	DPKKTVNTSS	600
	PSFLGCRFKQ	VDGKCYLASL	Q D R V T R <b>S L L Y</b>	HIGAKNPSEY	YEAAVSIFKD	650
	SIICCDEDWW	TDLHRRISGA	ARTDGVEFPT	IEMLTSFRTK	QYE	

**Supplementary Figure 4:** (**A**) FSBG and (**B**) GTP structures indicating the spatial separation of the points of attack in FSBG and GTP. Asterisks mark the positions of the nucleophilic attack. (**C**) Mass spectrometry analysis of FSBG-linked EAV nsp9 identified seven unique, modified peptides (outlined) located either in vicinity of the NiRAN (dark gray background) or within the C-terminal RdRp domain (light gray background). Residues carrying the sulfonylbenzoyl modification are colored in red. Sequence or structural motifs are indicated by dashed lines above the sequence in the order preA<sub>N</sub>, A<sub>N</sub>, B<sub>N</sub>, C<sub>N</sub>, A<sub>R</sub>, and E<sub>R</sub>. See also Fig. 2A.

Α



**Supplementary Figure 5:** NMR analysis of 5'-(4-fluorosulfonylbenzoyl)guanosine. (**A**) <sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>)  $\delta$  10.70 (s, 1H), 8.38 – 8.12 (m, 4H), 7.93 (s, 1H), 6.52 (broad s, 2H), 5.75 (d, *J* = 4.8 Hz, 1H), 5.75 (broad s, 2H), 4.65 (dd, *J* = 11.9, 3.6 Hz, 1H), 4.59 – 4.42 (m, 2H), 4.34 (t, *J* = 5.1 Hz, 1H), 4.25 – 4.12 (m, 1H). (**B**) <sup>13</sup>C NMR (75 MHz, DMSO-d<sub>6</sub>)  $\delta$  163.92, 156.63, 153.77, 151.20, 136.22, 135.72, 130.97, 128.98, 104.16, 87.13, 81.06, 72.98, 70.17, 65.53. Corresponding peaks and atoms are indicated by numbers.