

Table S1. Coverage of the bacteriophage ϕ R1-37 particle proteins by tryptic peptides identified by mass spectrometry.

Gene	Genome location	MW (Da)	Mascot score	The peptides in the predicted gene products identified by Mascot analysis are highlighted (different colors are used to differentiate between neighboring peptides) .
g045	11548:12651	42,530	388	MMSLTSLLEDDNDVLYKNPLMKVKKPKYFYGVSNYKAIQEILSSGVIDASQYRLRDTKAYRANDAVYFTDDLGRAFNWAYTKRSDAFNDLSGKDAFVFPVNPDIWSEVYFDEADFYIAVLGVLVYKMNQGEKLEKKAIEETKISVDENNNLNIRTKVDIHANPLANKVWELAFNTAGKIIISGKVSTIATPLFDILRNADIDLNDLVKISYGVKANTYKNI PMRYNKPESFNKALNMFDKLYSKDKDEFVKWVEEMKLDILAEMPPYQLKAKNIIAKKISHQYQKMLLKFIEFCSSFAYGKLNVMGYFSGDR EALSLEVQOKKIEHISDRYLKTTEQDHAINDYMRQFNKNSLGIKFTLLQNKGV
g046	12654:13259	24,028	50	MVTQSKLFPKPTDKMRRIPYIFISHMANFSLKFNPIAPYKLVYVAGYNLDKVESYFLENLEKKQLPYHFYVCFIKGDWKVFTGAPINFRSPLLTELSSAGYIQTAYTESIVVAVQDNFSLRVPDQLHDI IASKIIVIPYKQLLRNKDYLDVSVWFDEAFNKKQYSDMERYTYLRNKYPFQVTEMKFFDRVMFNIACRTYS
g047	13259:13720	17,751	322	MTVGVDSISYEGFDSSKYATFPESIADSKINPLLLLNVLSDKLVTNLDSBEIILQDPTCTFKDMLHFFKPYTYPVFPVPEKRYPEMTAKRLYGSADFWEILLLVNMMSTVMEYDKETINYIPNDKLTKITKQVQLHKGIRTVQEDDLASKDLL
g048	13800:14435	24,079	112	MILT EEDVLNEIGFYSQDGRVTELEGISSFVSKVYRKLNSIHQLYELVRKIVKTALNRFKNSKLTYAFESTYQYLLGLSLKISLVLQKILDNKKDLDIGAEIFILEKAL LGIKIYNGGITIINKVDFRQVEVTSQGINNITDTRSSKALNRYQSOLISYANGSHRDKSDEDFKLLVKNTRMGNLVSNYARTANKVSIFILNSAEPV
g074	31242:31733	18,425	51	MSKSYEFLPRSFVTHENGNSIDLKAPEIIFNGNHKLSGDIMAQLHYHFKDKIPPDYFHVKS GHFGFESSFSANVVLIDTDKISRLEKVSILDGVDTGKFKVQTDGRTVSI LMGALLMNNLNTIIEEVDGDQDIPAKCGKRGKCEFSYHVELAKSNSETWDDID
g079	33842:34639 (compl.)	29,550	720	MSEQVVGRIFYESAPVEDAIRSFKVLVPLNGVVKMEAILQDADAENRNRVYYPKSVINNSLQHPFVQEKLATNSLFGEMNHPDAKGLERQLKIDMHNVCHLIK NPRWDDENPNLLIGDVETGNAVGKDLAGLIVCNMPLSPFSMRGAGDVVNRGLAYVKDPMRLITWDNVHYPSHKVAYARAMNEDVNEVPITAKMLAEYVASNSKNTQMMLLEDVMSLGGAM SFSVDKGVLTIEKKSGLAGFALVEANLQAEYQDAIILS IMR
g081	35434:36822 (compl.)	49,684	181	MTAKAKVYDLNDKLI TELSTIDTSSKLLAGENNDLKAINIDIVKASLNYISAGSKADYNFAMEGILVLSNLYSDTFYSINQDMNDILNVLYSDRSLSAFVNSISSMSS ISSCCNKALGSCTSFANSINPANYSDISSYIKSIAQFLCPEASLLGDLGAAKSILNAVNTLRPAELLNAIKSDLASLMRDVSAASKYATSALGSLVNTIMS NLPPSANS LLGFIEKVQLYASAFIDLAQELWDAIESIMRMFGLTADDLFGIGILSDLGTSIKNAINNNSSNSNGIQIVPGIKVSPVDTIRCSQLAKIMQOQNSYRANNRYDYLNNS SIVSGITDSSIKLSFSNYSNYQTNASNANQSTKVYTSMGYLGK TALNSSMAAMVRNASSTRTINSLNGVNTDADYSAKSSIGNIANNLITVNVEMQKQTTMREGLKNIM TESNYNIPSTKFGV
g082	36917:37492	22,878	197	MGFNHSNDINVFNLEESSKRIPALYHISFKDNLEGEWKPRPPDGFQIKAKEKNKEKEVNITEPKTPRISCSTDLKKCFYAIYPNISFFFEKIKFPYMDFTVYSPRLNGEEK IWTPEYLTDDKLVHDAHVNEYCI LEKAFMQKVGHLRFVNPLKSNPPPEIYHYFPNNESEYERFLAYKVEIKVDSIK EK
g083	37498:45156	271,439	555	MAYLKGLATSLAKRVIDPVGNIKFGDELNTYKEIKDKVFTSREILRETQSNIKTTIQSFEFGFSKVGRLNLRQVRSRGKFGLTQDERDDIDMKAMGFDFDMSLGDSDLNALEQ DGENENSSGSSKASKNIIIPMMGKSSRGRNGLDAGSALISERLAYITGIQAQIGQTQITLLQGLGSLAGIGNFLSEQTVSHYQGTLEYQKLMSDNLSLSREVNLDLMKEKLM AEKKAEEKMATGSSGIGWEGFDPTRIAKDINEGPLGMLFGEMSPLRGFFQQVSVDPDPLGTIMALGLAHGAGSLFGGAIKAKGAFKTAGVRLQSTLEKWSATYDSDSFTGMFKS WIGDTRLAKRESATVAKLGAYDKGAAQFDGTRKAIVDVIMPYLAKILKAISGNSTHEVYDYNKGAFITSKAGAAEHVKNVKELRGEFLGGISNELTGLRGDPSVSKSDFG KIRASMIRASASSLSHEIKLDQVKMRDARLQETDYKYIKSLDEKSKASLYHSFQSHKYDVAALKDYQNSMYDENKGVETSAILEGVNPRILGYDAVEKANRKL DENERK KLKDRIKDSVKKLNGIKDRRRGGGNPPSGGAGDSGNSGGSDGDDAAGDSWFKRWLKGKFSGGKFSLTSTFTBGLVGTLTIGMKNLSRGFNRHILFPMKALLGKDADTRT IARTPPIAAVGTKFDKTVLFPPLKVVLLGGNEKQAESMIFGAIQKSFDRNIVFPFKRLLGKTDVVRTILKTSFIKSLGIRIDKSVLTPFKTMLLGGNKVKAEKVTFVTALK DTFKKNVLPMLKALLGKDTDLKTLTKTPFTKAIQGRFEQSILNPLKGLLLGKHSTTNGGSDVPTKTGFFKALGLSFDNRNIGRPIKTMFLFGQGNRNSFFKNIGETLSPFFN KLLFGSEKAEKAGFMANLKD FGSNVFKGFSDFKIKPIKDAFSDFFGKTMQDFIGTLRDSAKDFLGLSLNINIGKGFKEGAKGFFKTVFGDETVKILRDNIVTPLNDVTKKLS SIGMIAKLLLRVPANFIKGITNTLKLNRVKS GKNYSDEEKARLLDMEKKGSLFNFLDVGNESNKKISSLKNKVTGFLNRGKATDGSANANVETLDTNSANAKGTES ISDRARNLFRNKGKASSVDGSNNPNTDADGNIISASSTNGKKGTKTGFIDGLNRVSGKNSNSVDENGNPVEGGKSRFGI IERLPSRDKVGSVLFGRGKQNSEDSVQGSTKT KAGTLTGIPVGSAGLGGIVGHLNAPKLASRVDTISKMSELTSSASSAATAKSNMDILNPLKHHLSNNTQKLTKIVNAVAKGKTVSKISGEDKESVSFFRNPLKWAMRTSMKV LGFPLSMMKGMF SVGSKMVKGLMEVPKKIIGITVDLGTKLASMTLNVAASLTKGIFKAVDFVAQIGKTIETIGKAVNSIVRSITSLIATSIDI AAKAASLARGVNLATKA IGRGLLPAFEGLTAMVGRLGKGLFELGTLVIDGFVKS IGMTKLAGRMASGLMGKVFGRDRKLSASFSSITNFKDLILATKKRPMHVYIEDGKVATYEQKPPSTFFKDKVE DRSLNLSKRDKAKEKDNITSIFGKIANALGSLFSGLLAKGLGALTGILKGGGIFGKILLAITVMSKGLMSLIGTIKINVATLVFGRKAGDLIDNLPGRRRGRGLRAGKAGN IADRYYTAAKAEKQAIKDRASKAADKLANARNAVAEKQASAKTSGNKIAEKLANARTVAEGGSLRSRSCAGAAEAGANAASKGGGMFSRRAINGADFFYKAGERLATGGGIGC AVKGFGRLLKGGWKGFGISAIGGLADTFFDKGTAGNDVMTGASYGAWGATIGSVIPGVAVGAIIGGLVAVKGAIPHLIETFKKPIIGMTFTFAEIPDRISQYAEQL PNKISGLVKGIPEMINGWFLGDAEETPKIDEKTQVVIQKPKGIMGSLFSSIGASVMSVAKALPRIAATLVEGLVKGIILTTVSLMGNVNSAMNFGVGLGVDIKMMAARAI V AIKNKTFLLSDEEAQYKELDKQAADKAAIVSSSNTMEFVNLTKSIAFSVADKIPTGIESSVMQEKLYEKSMKESNGDKAKAKDAFIKSYTSGGKSAKEGEDAWAY QER AQNIRNGTDDGVNTPPEYKGFQGSISEAITKAANEEGVPVGLMMAKIESNPNKAMSPNGTTFKGLYQMGPDAWKENSPNVNGDIFDPYLSVAAAARFIKKNINYMKS KGVPI DNTSVYLAHQQGGGIVKIYNAAKNGTPLDQTTIRNMTNRRPQDGNKSATS DPNFELRWAVAVKNKGGGSDGLTDSLPAFNANSGITQGTASAI SGINNATTTAT GSSGFDP SALLMGMSKGLNPNLVSTVEVETGNRPNVGSNPNRPTAQVAMKSSMNA MPTTGSNVTAGAATATASP SVTSQEVVKAQAVNNAIVTADTAKYSQIKDNINNA EANSVLAKLLEKTNEILTGIKSDTGVAENTKHDEVTESIRYNSQSSNK TASRSQDERVGSNRRFFENGLDLYPSEGAKRIAQGGAS
g085	45613:46983	50,534	114	MEIINTAMNALNTATSFVSGVSNSTGKSFVAGDNPYSNLYNDSAGFIFGKPLAFNPKADPNQRVQRTLNRNDTIVNFVPGVPYTNQSMVDQANTILTAHREKLDIILLNSP VPDERAIANLCQDTQDALLSQQCDLRYYTFKQDVAGFLRALQIMVNRVGTSVFGWNQVSGFVSDLVASNVS EDDMLRGRFKVWVDKATSISESVDNSFTQSVLESTQKSISS

				AVKQARFLGGAANIAGMNAESVVKSSDANLEAVGNMNELASRTLKAGSSFPQVQVFDSSKFSR SYELYFRFMSPYGDDR SVFYHVLFPFLAMLTAALPRQDNISGYTSPFIMQ VDAPGYFSCPMGVVTALSF IKGDEKLFNRRGLPLIIEGTLISIQDLYSNLSLPLSFDQFATNLGTS AFLSNMGGMSLYSVMDV SARNYVQNWSKDTLTKFIQPYNWMNEESL KVRRYFGLY
g088	48997:49254	9,958	72	MESTERLSYLDNRQLNESVKRIVTTAVENYVKKVQQSNADSNVNTDRAIEALKNQLFLKFTGIDYELLFQSVYLDRKLNNGGFTTYE
g089	49250:50341	40,170	119	MSKPCSLNLCYGNTEKEDVNSLKENLPMNLLNLLGLLNCANNNTLKNAPGNWGLNI FAMQKLVDDINDVNNGGGRYDISDYNNIPDNIYNNLINDYNDLLCRQGGGLIDI DDKYNGNGDGSYVRPLFPDGLTDADKDLIIDHLNMMNNNGNDNINNGENNLVIVPPDYTDGSGTSGSSTNDDIFGESGNPIYNIIPNYEFNDGIDINKSVFKCFEPVFKLDNYI YSNSEKLFITPTK VANIKRVHNEILLPLYNFYGNTTAPSCQIKVFFGIGSRSEINSVAAGSSFSRHSRGEAVDFTMVGIDYKFLSDLKSGSLNLFNGILIPSNGIHMTLP YTYEGYTIQNTVINSKSKANSLSIEYL
g094	53058:53420 (compl.)	12,141	400	MALNPKLNDLKFPEGVSYNIVKLSALVANAADTQVTLITTRTPVLAKSSTASTQITIVYVDAGAVVQKTRAPNSEGKFTTDFAAAAGSGGGVTTEAMNAAIKAATDPLAAR TAALEGAGG
g099	57455:61633 (compl.)	157,842	976	MAINIDRVRPNKIYKVEKLLVNI SKRKKPLCNILFNDNSVTSIERDFEMIRVIPNSTKKIFYPPIKINPNTMTTRKDSV GANDYYLELKNRILF SR LKTIDLIESPKFSN KPVFADLSKYRESYNOYCANKKKTYVTGYMMFKNI IK SFISQLPKKVDYQFLDLDEE P DFFINTFLYGLKKEPKVINDLFKGIDL VFFHQIGNLVYRPGDLNDINKK TQIIN LINKVIKLGKAI AQQEVNDPDLLEVVTXIDPKDQELYDQSLQAYQHDDEEDSESTDNPEPSKRFTATDDSTDESEEEKTNDNSTDNDSVEDDKYITGSSTDEAPDLEIVD GALKDGDIDLDTVLNTSTV NAYTEK VRENEKFLN NN LK LQEAALKEFEELANKLAESNTLEDIRVNDSDIINPKVKSFR TSSITHSYKKGFKDIANI IKCLNNDPEHPV VVTKLEVTNVS NPLTKANEYHI EYMDKTFKRHQFRVNVPI LSHDGFMLINGKNKFI AKQSIVLVPVKEANDRVOITTYKKTFLYRKGDKINGQIDRVIRILVKGK EIASITK KHGNSYESNLEFNISIPYNYLSRKLFAISLEGYDIRLNQKEIRETLANKKITIDYDKYIPIGFKGKTVILEE IATRNI VGYETGK KYEALSDNLTKFLDILT GSKDPEIQD ALKSTKPSLSL SYTNMKIVSTSV ALGVLVAAYRGLLPALDLYDIPYRIEKRVAKTDSEVLLTFSDFYVVIDTKFDPAKELFANGLLFLNANEYKLEEATRMSPIFLDYFET HSGSRNTAKALVNFENSMIDPI TLEILKELKMPNFTTELLYGNLLGDYQRDRKNDMYHFRIRDSEVIAVAFYNSLVD AFN NYKRSTGSGVPTISVKRDDVIRRIQAMPN VEEYSTLNPIQIEIAKSKVTFKGPSGLNSED SYIAENRAYSKTMLGLYGIFSPISAQIGVNRSMAFNPKIDTIRGYIREKDIKDFADNLFSAAELVNGFTPKHADAPRAIM ATTQNKHIINTKVOHP TLVGTGVDKALAHIQEQFAWKAVKDGKITSIDEEKNLVFI EYIDGTTSAIDISMKAAKNSGGGFYIQNKLDLVNGIKVGSFKAGKV IAYDKNFF KETLDG SICFAGGR LSKVAIMALPETFEDSAVVTDTLSKLSSEVINKR DVI LKANR I IKMAQVGEIEVNDALVIFEVGNDEALAALEKLDKQTQASISELARSSAK AKYAGRIVDMKLYNTPIEKMHPTLRKVQSYIEKYDEKSSSLKNISPEFIQTPSTEMESNKLMGNEVQGVVIQYLIQHL DPLDVGDKITFTSLKTIVANTVDTEKAPY SELHP EEEVSALLSPLSII SRMTMDFYLQCGANKVLI ELKRQVLDILEE
g101	63492:64532	38,694	128	MAKKPGFLDKLKKGKTASKEEETLKGKAAVEDDEEEEEEEVAPPKSSKKPVLKGGKPKVVEEEDDEEEEDDAEEETEDDESEESDDEEEEEEEVAPKSSKSL PVKKGKAPVVEDDEEEEDSEEEEDDEEEVAPKSSK K VAPKKAAPAKKA AKKSKASDDEEEDDEESNGGKPLFGGKAKKAATLRTPKEGATMPRENIISMIAEKLD CTKAEAKTALVGIIEVLKEVGEQYSFRFMDSNFRRRVNVTRDYKAGGLTERVSSDVAAMTTRVESHVSLKGEVITYDRRTIRGVGEGDDFREGNMVGK FVAGTWDEEGNFT PSEKVKKAKK
g128	79877:81196	51,650	249	MAKVKRILTTEELVNRNLEQFTKSDSTETYPAFQESTPVFVTYYNK SAYASADIN FENFNQAVGVESPNLFY EIASLPLYEMTTADFSTEETDSGRKSNITGTCTIIPGTVT PSVEDQFLFDYHSQKYL FVITSIEPDNFNKYYRVSQYLSPNLDEIKAQVYKEMVVDY NLIGRTDSALIERKYS DYLNII RAVFDEFRKSYRYSYENQSCFANKRTEV IDQFINYFAANN DILPAFEDYRNTLTIHNKIIDDVSEQQYKLSPPFDLEKRRDPVSVS V FVEPVDRTPYKFAEFYKTKYNFTRYV KYTDEEVEIPSAEELAQMTEEEQLAV MDSIVKGYNLPELLEPINSNDPYNLQFIKKYIATFSKLRKTDQYKSQLIECLDMVEDELEEYFC DALYIKGASDSDKKEYDFNYIIGMIVL FALKDLYS FITNS
g129	81243:81872	22,677	158	MKDLTNELLTGGNLLAESVDQVETNVAIQEELQMANPTALLAESMEDMLFSE DARIDGNNIEMPSDAGIDSLQFFNQDNVQSVADSTLTGLDPRQNDFGNAAFASNPK GKDMVQLRSVGDNVALGESGIPSLFTEGGDEECDTETKDIANGVDDQTSIGDAGDSLES LKLFKEDDDNDSESDVEEDDEDEDDEYEDLGL
g130	81947:82549	22,697	51	MSNLDPVVYVTTI STARQIPFINRRGPITSKISMKSTIDTLKAMGFVKIIESEVNLKNGKQYADDASKFVPADNQSESNVDVNI VDAIKPNDEEIFSGSTEGLEAKKDDLS HSQQEEETPESEEEVDDTQEKPAEESTEEEDDDFSLESYETYEKWNATKLYIISDTIIPRLEEEDVPDPIPTSKKDCLAFLKTLLE
g131	82729:83088	12,634	128	MATVNDLLAQVOTNLKAQGV DSTKKECGVALEAVLNAIQTV ALDQGSIRTAIGTFKRKECAARQAHPKIGDKVDVPAKITVSFKSYYNEVVGEEKKAAKKAAPAKAAAPAAK KVVKKGLK
g133	83554:84015	17,241	92	MTLIEQLNDKKEIISGFPGIGKSELFKIYGEARVSDSDSSKFPKDSFPANYVAHIKDLIVNSNKKYILVSSHKVV RDALVAEGIPFLVYPNENCKDEYINRYIKRGSPEAF VSLLSNNFN NFVQECKTTNSPIATHIELSDGQYLSDVLTIG
g134	84113:84796	25,523	372	MSALFLSTLLQEDSADGGAKSSADL DLTGDKGDLK DQKDVKS DSDSDDEDNTNDDNDVSDN SDDTETDSSSPNDSSDDTEATDDNDSEEDDEVLNQR ELEEQRKKL NRYNQLKSLVVYTYRNL DGIYDKLLKLDLPGGLPNTIRMFKDKIAYNRESL DKL LLDVPIFNNTVSELTAIHNVM S DLR AVSTNLKVLVNSSVLNHTKDNRAILSII SRDA ENRR
g135	84901:86454	57,483	2760	MSVDYGFANLRSKTKHNSKATAHLKTLVEDYKGEVDILSEG YVDVVRQVAFQEYERLSEGTSEETGEQLAILSENTRMVMLQESMIGGVNPIAALSPLMRI GYPKMAV REGLPTEPVEQPKFKVTTKRPYVLDTATGEKHWLPGAFO TKKELFGLPR LQONDI TVPAGGLIDYDMLAPISK NALLGDEIDPRFTVVEVTD SADVAISVSFELDTNLNAIV GSFV DSTNAAVQLGKVDRSKGLLTLVTPGNAIKVKVIQGYVSSEMNAATQVGFIDGVDITIGTQPIESPINIQQMTDVMAMYNVDSTLTHMETMSTALAQSTDLEGV QFIYDCFRNTRKI QETFDAMPSPNFLAGPVEWRKNIHRNFRDLVMSQMTEFNIYTGHAVVFCPLDAQVAVDVRWMTYNTTEQPNDAVDFRVGTYSGTTTYVVLQSPNFV QGHYEVYIPSEADFKSLVYYPYAFNVIRGAASPNTPLPAIQMIKRYR FQKFNMMVARLNI VNNNGVGEV
g146	93545:94273	28,165	79	MEILDFVRESYSTLDGALEDFLSSKADLFSSAPRND FLEGLMSRIDS AKLSEDLSDDDKDELQLESALYHKIISHFRKDYSINEE EIANIFENAQITRFYIQKIYDLFYTGK REILINFLANFI FDNKGF AKQYKAQVSKKDY EYSLRKLDSL SKPEYCI I I KAVEIVNDILEDTESDFK E VLLGSDLEVEDVETVYSLFDSSDNFSR FIDTLIGSINQ DRISLEVKDKLGNMLKNAK
g148	95191:95886	26,368	140	MASKKVEKYIVKGTSLVCKQDVKI IIPKTFNFDNNLSEFKDSQIITQGLFKMSSPDGKQDFEVKLP LSVKLN FYDMTSDSNFYFPYKENDVIEDMVFI PNVKQASAFNLN

				LTAGKINTNSFTDVTVDVFKKSMAYNKVINVPSELVECMIAEMARYKSDTTIPFR IAYGTIKGVSESDYTMFNKYKDIARMTSVFGAVSFEDIKKSMSQSAVFMTR TNQDQTIT PTESVLKY
g149	96008:97777	65,490	831	MADYLHPSVSSR IVDNSAVYATAAGNTVLYAAIHSAGRDNAVEFVTTTDEFLEPKFGNPNLSK YGQTSYNILNLWQSGGTAYVLR VMPDDAK FANSLISIKTTAAADPAKAT VLVTAQAQTNTASKNAMK TILSGGTAGETPLCFIVPKGRGENYNYGFRLLSLR SDYDNTYNFR TYNLSVTVK DSTGADVVEGPIVVSFDPEAKDKSR OSIYIANINKYS QYVEIVDNR SAFETISEFVVGDEADPQKVDIIFGQER AVTPAETIHANVVWK SSSVTDTPSYDATAANFNNIQYLTEGSEGTWGTGNEESALLVK GYSGVLAPEILDKQQ YEIDVLLDGNNEVAVKNAMSDLCSEQR GDCIAILDSCFQGDAAQTIDYR TGNISMSTYFTAIFGQHMNVYDEYNGETITVTSTYFLASMI PSNDDQNGIQWTFVGPGRGVIS GFTDINFYPNEPWK EKLILAQVNYIERDPKKISFATQLTSQTSR SALSYINNVR VLLRIRREVEKMMADYR QEFQDNTTYDSMSYSLNNYLQQWVANRACSSISGTVYASDY DKQOSIARVKVELVFTGVIERIAIDLNVNK
g150	97791:98540	28,108	2625	MANSDILTYDSGLTGHDTSATGDNFFTGAIN TGQISHDPYVTGF AFWVWVSVPTWLSNPEEFKALSQKNFLAFQGLSNLEMETEGVRGGFTANETHYAK OTGAKPTEFTLRY QIHSGSPLDR FYNEWYSGIRDPHTGIATYPKQFGLYHSSNHTGVGLVYVTRPDANNFPAKNI EFATLITHMQPKRINMEHFNFPEQGTNDFATGEMPFSGVMHFGQQVAQYA ESYISDK VYNFITENNFSNINTYGS
g160	104515:107148 (compl.)	98,235	363	MAKQKEPDI LSGAVSKTAVNKISERFERSNNRIEK IQNNLDKEDLDLNAK INKLNADIDK LATNVARR LLASANNTE SAAPPVDSMFYLANPPK DKEERKKRKANGKKIK Q ELEEKKKEITTSQQSLIRNVVESNKSRFWEYLSTYNLIIIRIIPKMLLAK TVANTIISPDDFTK CALNIMIDSENVEEDHAQIKDR VESLITYYDINTKLQNDVIAYLR DG KLIYLVLSMNEEK QMLSESASELDKVDYSYK TLNLSEGLNSADKSQLLENASFSK QGSIFLEEFK DAFGIKKBATGVKVGAAALDKFSEALKDFTFIIGDSSHLSEYSDVL LSEDMNISGFFGSAVSPFNATGEGNTQPGSNRKLADPEKEK ILKTKLGGNEKAIKVRISPGNIVDLTFEDNILGYLYLDIVEVDPDGTMPSPDKVNGNEGYTFMPSQVSGN GNVLQNMVYSGKDIPI DNGGGHSSGAKNVENPNQGR LDVADDAR LQFLAAAFANR L SDESNIK LIKK SATIKQAIYNTLSIR KLTRKDKVRVIYLPKEEVVMINRGHSIFDN ILFFAKIYITTLTLLMQNVLRGAPKRAVYVEVGLDNNPNANATQQAIRDVKSKEISSITNMDMQSIIINYVGEFQDYIIPVVDGEKPIFETIDALDAKSLDDDFLNLWLSNNI FSGMGIPSAYLTEVENVDFAK T L S M Q N S R F I R D I I G D Q V L S K G Y T E L V R K I Y N L N F K S N E V D D K S D P A K D E I L T K N A K S Y T D T Q S L A K A A I K Y F D I N N I S V K F P S P A S L N M N N L S E Q I S N V N N F V T T L T E N L T F D D T I P Q D D Q E K L K K K L I M R F T K K N L P N V D W D E L D S I M D E I V R E Y T G E K V E K S I S T S E E N E E G S E D V G G G F
g161	107154:108185 (compl.)	38,885	72	MIDIVNESDLR L F I S K Y L K E G T F G T F S K L F N L E D K D N Y K G I F S L L T R K E I S L S D V E N L P I V N D Q Y V G E F L A N Y E K N Y F L P S N G S T T D F V A A D W L I P S G S P V R D L L V D Y N E K E N S G L L D K V I Y K V P Y I S Y K A T V G G L Y S N L S D F K H S F D E K A L N D I L T G M S V T D I S T L Q F S P D S T N C T D Q M K E F V N G I S S Q M N S L N K L S N D K L F N Y I V A L P D S N S A F G V S T A I F N N V F N Y K T S S K V I S F L V T N E I F H I N S L G E A P D E T Y M E A K Y S D V S S I K N Y P V F D D M F A T M K F Y I Q E Y L L A A T S E N N S L I N K L N A D L T L I A E Y E K F F S D I R K V P D L T M S A T N K Y L L V N T D
g162	108185:109204 (compl.)	38,088	107	MNEGTISNSDIQEWVNTLDGTDTRVQKVLGAIQLLDGNNLQQLVYFLINGVFDISDLYTMNINNTTIRNK LVIYDLNDNK TAITQYKTKIKSFASYTNLINEMNGKTADRFA T A S L S P F R I A V S K V I D P L G I T D G S V S S V T V L M S A Y T Q I I N D S S L V A L Q N S I I S Q A A A I Q T K Y G Y R R L N N L V T N F F K K I S D Y Y F I F T T D N T L S L K I N Y T D L E F Y I A K G G F K Q A K Y T S F I E N F M S A L N G M V L Y Y D A D E T K D G S Y A V N Q L K N N V S F S S S F T D T I L S S S L F E S O F N Y A T F V R I N L Y S T V Y G S A F F N D T A T V K E E F D S I I S L F D N M Y T Y C T A I R T A C Q G V G V
g163	109217:109738 (compl.)	20,229	225	MIHIKELQDK F S S L D V N S K N Y V L Y E R L L L L D I S D Y K P L M Q E I L D G R L D I E D T R Q L L N N V I P S S N K T S T N R N L S E N I V D T M K S N L I A R L T G E N T S D I I K E N Q L F F N N L D D Y R Y A I A K E Y E S D L L A E T L T D D E I N S I L V V V Q E T Y Y Q L N L E D I D G Q N K F T Q I F M S N K M S Y K P I S K L
g168	112269:112745	17,193	894	MKNTIVDELQNTSVTTNDLKNNASGDGAEFADATNAALTDPDLDIHADGIPLEEQEASLAQLAEFYKSMEGEGVLTBAARNIVRLNRQTKLNLNTR SALVMAARKGDPL FAKYAKFNKMRK A I R A Q I V Q K Y G A K A T S Y A R K L M T Q S V D N T G G K A K
g174	118419:119195	29,598	57	MFNNPPIIKIVGNSETLQALREKLATFTPEQIEETFRGGYSNDFFNRIINEARNKDKKPSVFEDILTDLIIKATKK KAETPKPTEEVEEVR IDLLNMSSEDIQNLKSKED LENILKSLDAYHHR IQSSRDMVWNTIMLREQEAHNSLTEIKSNLKDFIQIVNSLATDVK IQLPEIDKAK SVVWTWISDR LLVIEQKTKQVLNGAKTVAKAIQNGMICDNPT FFSEQVKVLEDKMLRIDQAKQLGKQLNELLSAVGY
g176	120095:120997	33,503	84	MAQQQKGGGNNRRQNSTYSSANFSSYDEQVSSLSFSYANEGSFIEVAPIIPTMIGKTPKQGEK VYDYDNRVSTFISPQHAVLLRRALNNLLELIEADEEIPAK TEVVMPTG QGIR RITIIAPGRVSLKIKGKK T S F P D N F M L K L E T K K S D D E T M L A Y H I L Q N S T V S Y I S G E E T S E E V V E H D M I M L M E F C E M A I E L S L G K V R H T A S M T R N L V N A N Q S S S S N G S K K G S A F N S M D D D D G E D D E D Y D N E T E E E T S R K S K K S S S G K S A K Q S P A K K T R K S I S S E M E D D D G D E D T D Y D D D I P M
g178	122027:123442	53,465	240	MKLDIYVDVK FIEVNGIK EVTSAYVFAADGAPDPDGI FSYDI FGRVGNDRKNFNGYIDLKRFHFPLIYNVLMQFR NLPLVLAGEK YCIVSPK TGGIK IVPPEEGAETG IDFFVDNWSK IKNVNDSTSREK BALFETIPEELIFVEKWPVLPAYYR DVNFHSHKESKGVDMDEFNTFYIKLINTVQSNLI FTGAYATQNAQTI IVDIHYI SKKISGK GGIIRKAIMGK TVDYAITNVISAPRFNAEFYK DQMI P Y N Y I G V P L Y N V C A L F F P L I V N Q L E N M F Y D I A Q S T E F F L T Y D E K T G E A K Y A E T G A M V T E K L N S E S L R K L V A T F I K D K T K S I R T A R F N L G E E H M V I S S L E R M L G R P A T V T D I L F N A A A D V F N K H V L S T R F P V T G S E S L I I N R I K I L T T E D T V D L S N I S D R L T G S L K E Y Y K A Y P L I P S A N D G T I R K D K V K W I D T V I P N N S F L K G M G G D F D G K F
g196	135742:138291	95,261	76	MANNDLQEMLRFRYPLNSTSMQPIITGQEGAVSMCKANLVTLQELQREADSAYTKFDISVK EARSSVGNNTINPNIPASSVTAQYLMGR IDMGVSVAVR LONEINKVLNPNL PDYVKMAYTPSVVDFGGVQAFNITYTKLKDIEKRAAIKEDNKPEIKEATVSSANTEYSLENKTFLKSMYDISIQMILPDNTQNDITQYITDILYSFNADNMVMPYISVKM KMPATLYDSIRKYFEGIKWYITVKLYTKSDTSQDQFAIPLSIKDFEELVAIDPFTSPGNDDSNRPEGLPIYTMK M D F V S A K G N A L N S I V K A R V F S N C K L I D V I S A L C S Q V K Q E Y B S T G N V D S S K E I K Y T I S P P D N T Q T Y E Q I I I K P G S F T A L H A L Q R E Y G V Y K F G I O V L F D G N H N I K D T N G V K P I S Y L T I T D K G G T A P A T N N I T D A L F E V V D P K Y V K N M I H E T G Y L I N N D T S S V T V R T F E P Y H L V K T N S D R L S R G D S V R M S T S O N S A S T V C D N P N D F S T Y N Q R F Y Y S K F D N P Y A L T Q L Q D N I R E R N Q M L T I E A R D I D I F T F N A N L N Y Q V K F Y S R D D E T G S G A Y R L K K Y A A Y F G F T R S G V K S R V E T S A V F E F G N I P E L R E N G V T A P R K T Y A E K V G S T N K S A S N S T L K G A T S S G T G G G F T S N S S N T P S V P F K T N F R N K D Y L G N T I P E T I P A S Y K M S Q O T T F E D V Y V T Q T G T D L Y K G N G L A N N F A Y F I N A Q R F A S E I L D P L F N S E G K S K L V D F Y K Y E T D G S Y G A S A H L V A A D V S F N K S G D E L I K S F I K I S N M G L D F D Q L I L Q D G D S S W K Y I H I G K N M N E T N R K Q V I L S P T S V N N D Y K V I N K A L T E E D V P K L K I T E F R K Y F
g197	138340:139185	32,333	754	MDFIYSDKLFSLMEEESSQFKKVILAEELFNSAIIKLEADENSPEGENSEAAESFKKIKIKFLKRIETVR KYLIKLRDFAKKIFAK ILGKFERKSKGDKVK LKGFANFEK

	(compl.)			FVEESDKFANGVTNIIISFVAQETSPEAAKASLSYNGNADGMKDI PAQDKQLVTEIFGNLSRFTDLTENEYTAQEVEAFKQKLMQIMGQFNSTNEQLLK TISHLEELKAKIE QLFKLYKMNGKQISLDMLTRRTRKALFNNSVFLGEIAKFMNNWLKIHSDYSSKLVSGKE
g198	139190:139978 (compl.)	29,967	460	MNLDYTFNEIDNLEHHECSIPLKVAILTETTYLQDMQALTNSVDPKLIARLEGFLNDNSINKIFFSHNELVNRYSQSHEYDLASTLKS LNGR IECLSMGER VLHSLFDV NISDFDNFSFTILSVNNHISKLCDQLDNANMYLNGNATDSNNTIFLEQIRTNLQALRQCSWIIDKFFPMLIKELKEIQHQFVDNNIQSPNIRLLDQMERLVSVVIALVNVK AIRSPFFR VIGGSANSKSMVRMSDKGLGNSKMYSDQGF
g199	140087:140863 (compl.)	27,259	7977	MAMFLESANTQKVEAPSLFEALAEITDVISDSAILNESIMVADYKLEKQLSALTEEGGDAAKEDKAGFLGNVKEKVKAMAQRAYEAIMKVYNAVSAFISKQVVRKAAALG AKTLELPESFKNAPKLVSAIEKAAANALGKDEFDAAYQEGLAAYRAAAEKGEVVKFLATDVSKVETMAKAANDRVKVAESLKAASAGGGDAAVV'TLNKAQRQASKLASL SSKFVHALFKATHGARGVAADKAAEAKEADKK
g203	141850:142593 (compl.)	27,984	158	MITAKNRKKAekli IDTISAMDPGSMNSRYKELFKSMSDKDFVEYFKMKKSEDNQFVEMDLYGKNDITLDNIEKAAKHINVPLEEYVYIRHKTTDGKVIPTPYKVPVW VHMKRMQQLLSKKVKSNDISSGNVRSRLTGS LNQDNKSGRFTDYDTQALLSVTLDDPSKEDANGYSPIIKEVLGARGDNIENKFLQDISLPGNASIEEVD TAKSGQAVQ TLDVFLMAAGLKS DLVSHSLINK
g204	142599:144113 (compl.)	56,354	431	MLTFAEQDELKESLEKFFLEEGICNFNTLEKNILGKTLLEDVYFAKPELLKMEKAITDDIRTKISKEKNVKKLHSETRELGNLIAKYFGFAGCNINLGVLD SGFIAILVYD ALTKGKTKELPMALMAFMPNAFTFKSMSQVNR FIPFDNSVIENTGKQIKYKNPMVAGILDITITPLFADALTSGEILSVILHEIGHNFYRGSYSAKLVSACQSAFDFPVV LLAKLLNLLRNNMPDEMEYKVFGGINKVMMAFINGIGNFRKTSVNTAKQFEAIMGISIPLSLINIAQQITSILGFVSLPSKIMSMIAMDGSAEEKFDDFAAMHGYPGEL ASGIAKLGMYGVRDKRAFVDVQDKLNKNPLNSDDYLLIAAFTSNYILEGFVELINIADPHPSNKARPKHIIIEGLKR LKSKTKDPVI IAKIDRDIKEIANNKFSIFPQDKKD IDSMVNDNIYSYLFK TANPGEIAKELITDIWTLPLNFKPGLKQLGSVTTDAMGTNTEK
g205	144126:145283 (compl.)	44,455	186	MSTFILKNFEK KINDAINDKVKFANLKHINIYMDKNHEVIFAQGPMTRLLFNTRDQDVILEFVGVTEAEVKA VIKQVPSIKASWQILNPNFYILSCYMRDLDKQKRQADL ELILMYSMLK IYSGLEPKYFKHGVNQQIMAF TLNLDLSDKFYKAAKNNYGVIKDVVFNHSHQYIEKLLREGNDEMLNDYLTQIFSR LNK IIMNIAKEEYNNRAKKNYLNVTVK YDESGNLLDYENSTGVIQNLAEQVTYFYMSSNVDLISVQVAAKNSIPSSVYQTLASIRKNESPTTVKEFMTHLVGVYDADPALIGR VCTTDFAITALKQLSVSNSSNQ NLIALKDELDRLLSIYSSKYAMTQRLATKMSYRNALYSYFVFLILINHCK
g206	145361:146332 (compl.)	36,761	427	MLKHVKKSQRRSAFRTLAEQRAAKRESEKILKEEVTQRFSPFKSSRSKQSKIAEKRNKHIQALKEVADIVRNACPIDGIKETPIDRVQAINCVDNFRSVVVS NPF ETLREQSSHRYRVEIGGANKSTPADLINKVAAMRVDDPHFNAGEGQEVKFKAINVVDVTARGGKADEKIYIMNEGIAASLVGK MSETIRDRVVKSYANEVQRIQNSVA FNEEDVMEGLNPIAAKNIKERIRKENNAHNITREIFSTVKIMNEHNSGPEPDLNETVLNMTILETLMNFMGLVESNVEPALKMRNARKEYMERSSTELI
g207	146390:148978 (compl.)	98,038	180	MATSTTTSSITESEYLFQEKSLR LAQKFYSINDISLLK TGMFGYTTALNAHAMRDSAFHRDVLNFNEFFLINANLNGTLYNWA TLNYDISLAVPSKMPIAIKLPVNEIEK IAT DSSSGNATKFTTISKDIKFDVSGYSFLLPYDLVINFYRSTSSKLAVNAYYDFESYSYKPI SIQTPYLK SFVSTEGGVDYITISAEIYQLTKQDFIFTISSNDILDAGIEVN YGSNIASFQASYNANNTTSSYSDLELIFNEIDQPTTSQCYTYFVGDDTLR IYFSSSAGFRPDYNSKVKISVYTTKAEEGNFSYSGSINIKDTRLNITVKVSATTGKST GGEAIKSFQETK LALIQQLRTRDNYTTTYDLETYFDKFKRENLSNSSEYQIVKLRDDIFRRQFTLYVLGKNSDGVITPTNTVDLEFSLSELENMYSIKPGTLVIYDRVTSK YRLLSDDEYDPVYLSNNDNFYIYAI PFLMNDIFKEFPKTNVYHTNYDK SITTEYSYNIRSPYQAIINTIDIARNPLTDINDFTVSCNVMTDSADPANLKVRLIYDGTILKG YFDLTRREDKTTVFSAKVNTEDTFTSDGYI I KDFI INPDNFTTYEDFPIETTYTIRLAVLINDTIDAAKTIDVYER MADLNNAFICDVGIDKISFADDLSDIMYHETII DTSSTKISISKVPVIGALFYMNQFQNNELMNDFVDNLTAIRSVTAQLENNSTIDTKFFNSYGVSRVFAASGTIDLR IKMSIALTTTATEAIDKEIKAEVVSFIEACNNQSDKR FSISNMITHLETNFSTIRFIKFFSVNGANIQNI EKTA YTSAKTENLPDYDVPYELTVTKVLPVSINTTQDFDYGIDITYI
g208	148983:149990 (compl.)	37,686	80	MSYEEVK SIDVAVERTDREFGFFLDRPTYPDATNLNYVFFTSPLSLTSGNFGSGVAGINTEVLLKNNSNFLKLPYGD SIYSTNIVNMLAGRSGIFMPLLTNRATSFP ASPVVLDSDQYAE TWNKYKILMGTTAKDSRIGGEFEMAFLEDQNTIVLKSMLKWESEYIEGAF LGDIMS AVAASNMGSSQTSYFDYMSSMYHFVTRPDGKTLIYWAKYGTIY PGNIPYDVFQSQDQGSII SVQC SFKFSYKEEMNIALNDFNLLGSSDSK SQIMDFTGKYFYENSVNLAASDVNPSIQKEQVNENGMPIYSLRMGETKDPFYPNGVNNWS
g209	150145:151266	41,209	855	MFEDNQTSNLLNEESVVEAILGTISK VGNFAIDK FSSSKGFDDLKAKENSLTLYALVSSNPADYK LISGGLEAKIATLIEVIAKNEISKDVASAKEFIQKSLTNTTMD SLINLGK TQANKALGLSEDEVKTPRRKVSADMNSP IVNRKIDVKTSELAGRDKPFI SPGR LFTFDIRATNARGDVVSLTYSIFVKVNLVIDSDMLLVAMSDAKESSTLYQYI KHRASGGGFGSFFKDVILNLKTLDRKVARATSNSLEDR ILNDLIRNKGLVAPNTTMVSDYLP LELKFMVVIDKSDADR LGSEFNSTLTSGSNLKAFFGNMMLSLCVIDTAR DKATMFSDSKPTEMMVSLRDVSQDKKIAEVFSKIYSR
g210	151284:152372	40,737	1132	MSNVSNFINEKTKAFYQNDLLSEQAYNR LNTTDKNSAKLQAVL TEYVNMKEVFKDFFGGKVDSSSEKELKALSNAKAPLQOIKSAGDIRKQVQDAANVRELLRFAISTNNSY TGSASKSDLYSFINVANNITDLFDKMAASANDFKKAMIR EASVASDKRFVNTNLYMSLAYLITVTGSVYSSSLEADFDYTSAPKIVKNLYFSYKSDAVPEQVRYIGFAADQF RKGSKKILQGMLEEFFAEERIALNENIFDVI VGLVAKSDDLSDLFLMPVYLLRWVVSFKYLVGSIK SVSDNIAESIALQKQOLVTKDQFDDYKQRSNKLAFSMNQ TAM KADSAVELEASKDKKTLNLR SQGLVL
g229	159873:160268 (compl.)	14,521	182	MGNNELVQCQLLEGVWEAKGNFKTNSTFTFRKAMHSLGISNFPDSEK KELFKMHTDKLOGMLDAADKSSLEGLVSKLEKWKI AAGGLKRIANTSGDAADIKNYNQAIIT ITTQLNNEFLGKAKAKLAKN
g231	161796:163067 (compl.)	47,442	340	MFDINDPEK TILGITEEDITKSFLLNVFAVDAATNKPVMDSRTRIKIEESLVPADSILRKFKNWKDIVGTETNAGRIVNFVIGSAGTINGIENSHLATILEYINKPTTAG VIGGLNRVISDAYDKYILSDVVKDFIDRMQWLGYTSAIYSIPSLDIKTIHPSEKVKKRDQMLKDNREVIESNDVMAFGKLENNEMIDFAKDLLEIGATGKMIYDSGFNGS WNNNFKVNISIFRGIAPKSDDPSTFSIVTSNLMDGVAKEDIAPHADLAVLGAAGRAVDTQGGYLLTKLFNAAFSGI IADKPDSDCGTTKFINVELTGDNKR EYRYIYDKNE LVCLTPALLEK YVGKTVKMRTPLLCQNKHYCHHCLGDMFYRLGVKNIGLHVSRSARIMNLSMKSFHDLVSQGEFDFIDHIEDINSK
g233	163355:164038 (compl.)	25,672	1147	MFEIDSQRVMTHDEKLLIQDAVRANKARENFI VIRDVPTNNILQVKNLNVQVGR EFTLRKLFKLPYSSENETLLHNRVYVCLFGIGSGGAPVSDPFPNITPTSADTQLNTAV PFR IANTANPMPSEDANFYDTDKVSGSDTQYMKKIFTKVELITDITVKDDYVVKATLTINEKDAR GSR ISELGLPSAVVPSDNTYNDIKIFSRITFETEPLSAETGKGLIVEY YVYA

g234	164092:166683 (compl.)	102,150	120	MKINDNTLEYIENLVQYIIKMVFK DELSADSQETENS MKASSEYMAAYLKMDNHTEAER QOIINDYVELNPPYK NLFDTYQIPYYTS RAKDLDI IYTDSDMETTYPYLSQY GPIYK KCLTYFYATNYAK SMETEENYR NFLCLMSINIMALINYIDNWLKYPPYDIDMMNESMIDDFILSGFISYFKSVPLQYKQIAK ININSLIR KNKGTDKVIINILDIFNFDN INIFKYLRVRETRKETDGTTNIINNPR FISHNIR IDSLSKA IKAGDYNDNTIADVTLGD DTWR TNTEI ARYNFVYQSKYFSIESGFELAK EGLNTVY LLNLLR KIRESY SSKEFMTFTASKISLDPITLEDTLVLLQILTLDYQGVDTIEYDENLIVSIYQFKDR SGTITYPVDIVNTPVD PNLYPVKDMTKETSFDEK TELEVFDY NIGIHS NIKDAL HTE TDYKR YKQLK Q IYYARFTQK LNYSLFNKA ETTYDYILSRN IDIYFN INAIR NIEDEPERQIEIER EITYYTDVVK SYLNDMDLFFG STSDV LISYLRDVI EVFK SFTV TLR ELSFVMVLKEELDGR IFDGVES MSDS IIPIEDHVTGATWDRYVFLGNFHY YDKVHLHDITSIFTYLGKQVTTDWIYFKDK QVFNADPHYK ELHLRDSFR ISVLP LRDY IPK GDYFYDFIYSNK WFNL FDRVNIHDDLVR YGTFT MKDSIK FNMYDQ IKVANSYFNQF DTVFVRDYVPVINTTYHNND STDTLE IDDNFPV PNSTYHI YDKNAK LRDFO TYS SFMDSVNK INFNDVFKKK LANYSI KEYNVLKE VIKSTK GNFYIDAL PDFGSEK AVPITSLPVK SKMLMQDNF VITV TGED
g236	167036:167455 (compl.)	15,304	50	MDSILNNFVSSYKNIILAICIAIVVVVIGLFLNGIADNAKASTEWEKGKYTAQK LVSDQLEAAL KAEQDKNKKLEAANA IQQTE IANLK AE DNKLN NDIKS NV EV IIKEVPAD KV ISKES S KALV KIAD NFD T VD NLE DL
g237	167557:168264	28,371	184	MSSKQFV F PLDK KKFSS KPL MYAYIEENYS NMLSE EMSPARLY FN LK YNK THGK S VI SGK P T WNEV TERYER F ANE E ER K EYR QMP F SR M IK K Y G K DHIL DEPE QOK VMLQ NRSIS KEY K WCDGSIS V VNSQ Y EYDFLNF V SVY RFDK N CFTE P PTI YY K DGSNMNFYLPDFY I PSL N LI IEV K S NGHYQ NR D S Y KE K LK A EAT IK EGF N F I QL N DR NYTP FNIF FKQ V LD N
g239	168981:169373	15,170	123	MPFFLKENLEDEE TEV FT EE EVDTIFGNFI AYA ERK KG YST KE ID ILAN RA KLNS EE VKR DL LSR IDA AKKA ANEAL QKA ESGS DKY REL RLQ I QV L GEL NS KVKS F KIT KYDE PE K KDDR D VID L DHN
g240	169388:169852	17,295	49	MDNVEVEGLDFVSATAESTKIETTLTLNAFGTTQNYTELQALAKRIQSLILLEPGTNPAN IE MGV GIR NYVTEFMDDITLTRLSR VTLDQV NK FI PT Q LV KD IT YGR SG GV SS FVFC Y VY IN MTS Q EF Q K NI LAI S FN SS SPAY TT TI ESE IFI
g244	173415:174284	32,789	69	MATLDAMEDSALEK EFT VS FL NN PM GDK SSV MG AAM LSK LHK DYQ TLL AMK RFK VG YFN IK KDY IKIK VP SE SET LYD VV IK FG KT NG HL PS LN PK IE FF S NA PS FT TY AH AFDAY G CL I PEL KK L F DK VFQD L PIN R ND LI TF Y ETIT F AL F Y IK EK G YLN T NAI AR DS AV TT TL QKL IA IV V PTA DE KLA EH KK AS DA AV AT R KQ ER AK KKA E K ED KK S ES AD FK NYIN R VP EAP K R AG R K PG S KN V NNT V MN Q VK ST VK SK VD N K VS
g270	191070:192023 (compl.)	35,757	1004	MSFKFR SL KD TYI GE IV ANR M L T Q V T G AM T K AED Y RL T K DQ L N L ISS D IK S KI SV SG K Q V IL DD LES G H ID CL Y VP G L N L T G WA IN G K D Y T G I AN L F G K I R DKY S FSS YQPR EV FG V L MY AH V QR K YA EN Y N L V Y ST PL L TNA I VYTR M VM R VID S MY AT AS IP VQ SA IS YS IAK F FLS Y V VER DHP L N K ETR N A I AM Q VAK Q SS I AD S AI E FG T RF N D EN L V S L P R LF S G L EA Y T V VK DL D V NT F IR K M I T Y CEK SL FM IE LL S GL Y GT ICS AY SS N L IK DF IF ES VS G KE GL IAN L F DL TR N
g271	192029:193801 (compl.)	67,577	590	MAYS N L ET SL GV GN FR EL S GL L EP LS K VQ P INS V T ED LS D V IF NR FK Q L GE AL RYCF Y AT GD IA IR DKR Y LET FL SH FT NP V DR S YQ F ANK T A H AN AN WD ANG V P IT Y PA STR VE YN V D CL GL N F AA I TH FV FD GK LL PK SE YDL V NTAY GV K CY IK A TH IND S V V G IK I H R LY NT SK DY VT K IT ST DL T G G CT LY VP AN LGT F Y HH K Y IN V V AK R T ID GL T AY V K F R Y G ID F ITE INT T G D TAR I IF ST L PQ VAD Q L F IM NG VY W Y EA ST NG D T S NE VE IK E KL ED N S Y R P V IG S IE D F VY L NG Y HL T AG K H Y TL Y G G NE F S G Y IK LL LANK ID DA A EW IR I V K NE C VM DD AD CI I IR Q DT M Q E Q GL I T TNT DI A I PL M TR L GS S F L N Q K Y LS N D W F K PK H R Q VM EL D V S K VD TL KD F EY RL R I V NT I D IS V IK F V VD NY SE FD IA EV GL D N IM NTI Y A DE P N V T PD NQ GV EE Y ST VF EFF SP Y TA A VE QLR V IS N Y Y QR I NNR S DL V L D AN D S PT E Y P ID IT EQ ML ED N V LL D AN MK Q S I F V L D AN Q Y F ING F P T S Q E
g272	193826:195130 (compl.)	50,670	107	MSDI Q F T DN L V T K Y EM T V N AL Y SS L I GV RE Y FI S LP P T Y L K Q V F IG NS IA AV TE Q PL DD L KN MK PL PR MS IK M NY EN Q T G D S FL AD P FL M AN Y TR FR AN RE MY T KI V Y D E E T K Y IS A ITS RAR H N F D II S SN K T EM S IN IK G Y LS N K L G T NR YR Y L N D Y S L EV PL PASS I G Y A K H K G Y N L K D K D S L N L FA EW I H K I S D G K V T Y K TH S T G K Q FF YR Y G T N L IR I P N F EV PD A Q K DNK AV S ES MLR FT VE VE Y NN H NY F IA E SY N LAN E P ID DE FL IT E GG M N D V M H WT L N W GI P H Q IGN K S I L G L EV TE AN SP IE E ID F GE Q L AR AP NY Y I D Y L R S Q D N F DEL K AT N L I Y K L R DR K EL VE DE D Y Q IN W K T K V I T IL SP Y AN D Y KF FM Y A P IS EV N D Y DE Y L Q N R AD V NN V IVES
g275	197299:197973	25,470	539	MAMF L EST V TE QQ PE LT Y V ST LS IME EL ND IT I EL SEN F ES FF Q E YS ILK EE S DEV KAE KK ES FK D K VG K TI T N IL DR IR V AV V K V Q FL LY I SK A IG S K R F V RA P Q ELV R FM A SH S IS Y V K NF AD DS K L D E IK DL EAD KA A E K AK AD T MT MP A E K FK G FL N PA K KS PE E IES G IK S N T OLS PEG AK VL Q AC V TF S LA V V HA IT AK T K GL D KE PS D K E
g280	200731:201084 (compl.)	13,507	280	MED F IK SL FM VR T AV H ML H L SS Q SY SE H K A L E T LY N EL G DD IDI I E CH Q G D I G Q L I V P SS IK E V D LAG N AL TF I S D FL S Y IK K ER S F P V S K GV N S L ID N L E TR L Q RA V R L K FL K
g281	201097:201483 (compl.)	14,163	448	MLG D N V L VE ELL Q E D FR PL EAL K L V FT GK S D T V I ST L K K D IAN AD SK SR I TE I I S D I D Y L N ATE S R IK G D I GR K F V DL F IS N FM G GL V MG ASH FI Q EN K AK GL L TF K S Q L E AV K S IA V K K K AS L K
g285	203743:204606 (compl.)	31,614	58	M N L S K QYQ E IS Y E GR K R G V V V R ED PK GES R IG V V I PS L MP F ED PN K PA KK DR VE T T GT SG R AN PE IT S M INT SL QR S N Y I W R P T V SS GS NR VP Y L GAN V V FE D GP K K GY L P Y R P TL NG ET VS Q D T IN M AS D S F N PS N K P N V H IE S F AD GT TI Y Y EN GG T RE Y G IK T S D G F E FT I AG N AG K S I A TP PK G FL L N IDE L N Q Y V Q M KE D N H V LS D KE Q F I EV K T T GG H VS R L D D KE K I T I Q T SG G N ST V MD D G AG T V TT Q NS SG AK I V H Q GS N VM IN
g294	209703:213608 (compl.)	144,703	233	MS Q T R ES AR S R N R AL ID G L NY Q D G T LS SG K P T SS SLK T QS FF Q I FK G K AV K FG G NY K L TFL G NS IK N SS S Y I IL G N L Y K D I Y ID G L G I AG S N K V K I SL K V V NG Y V K Y T M EN N E NY Q N SL V V Y NG IE Y L AL AL SN V IA AP VI D AV MS ED T LE T S Y ID PT Q FT TI K P W AV Y T D G S K I Y ES FN L PL NG W IDA E AV LT D G T V LL T G GF Y N P V L V D E Y S T S Q T K V V T K SY IF N P ADR IFR ET S N L P Y P V N H R L CA IK NN CA L Q L G G A Y S D Q ANT N Q W K TV G D IL LY T S D N V SG G Q W V L G Q WS NY L FR SH V T Y D D V ND Y L Y V Y IP FG T SP D G AD V A IN H F G R K I Y GS D T NN V L G A VE V LA S I P TL S MP S SS G SS G NY PN WR N T L HY Y NG K I Y L P L Q D N AS N S F S A Y G K W N Y GL V D V A T N T W K W INT I AA G D GS Q V VR IK DK LY TT G W V S QL KE GT T AT V P L Q F E S D V N AD ML CF DL L FG ST MY K S PL N L T N AS Y DI KN IL S A G SS C MA AL D GR IA I F G D R G NS F SG V PE Y LN T K L V N L NR PS V GF I Y P D N D T S D T F SG I V VER E IS D N Y SS F N Q P FK Y KG D F Y L ST G K S N D FG T P D G SK L IK Y SV Y GT ET IL N VD K LT K GL K GA AT AV S KS GR V F IL G G D K E G V L SN S FF T N L I T Y S D T W AA IT SLP IS V A Y GRA K VT PD N K V LY I G GV K S D GT FN R N V Y AY D LT ST W S IL T TL P Y D F I AG D I T ENT F NG N TY Y F Y GL HP D T TT NN F S I I S NT GT V TT KL SD L ST A IS NS Y IV G AS Y LN K D V IR L V D NT GR HW D Y DL V NS K IT V SL TR R NN F Y P Q NG I V K L D SD K CT T V D K Y GR Y I Y ITH L Y QR EF S E R ID H P D R D K I

				SGFKTDVGTAYNDVVGWESLVDIYTAPMVICNEYDNTSIFFTFTVQSRAPNNTTTPQNLIGIYNRVTEKFTTFPSSQLSSVYGATCSDRDGNPIIVGGVKNKTLVQFTENG ISYIDSRNVNNVFMFVYKFDMATKTIISQIATMPYQYGNAAFYDKDTRKLYMFGGYTAQKVYPWDIASNSLSTQRCMVYDVTNTFTSLIASLNPYGIWGNCFRKKDKKII IFGGYIGLVDDKVDLPSLRQDLGNPNKLIYEYSISSNTYKDVSVVEYSFIQNFRLTNGKNEFCVHFAERIPNCKEKEILFYNLQNTSKDSSSGFIHTGYTPKFLGGQLNIL NLDTKEFRVVDTANMVPPIASAALWSPYLSASKSLLHYGVDGKFFYEMNKYLAGSASFVNSPTKIGIVRNYL
g295	213718:217833 (compl.)	151,924	156	MSLSRKSTKFSRQIISGNLAKDKKGNITGELDIDPKEFASVLLSGSNAEALKNSTFKFEVQGVISDVNDILILAETSKDFTFSGKINFNVDYDASATFDFAYNKSNDY IKEGSFFSVAFKITLVDVTKGDKSYIGLKFDAALGSYILEAVSDSSQVVTLLKSTEYESYSVSTFLSNMFRSIPDIPKIGMAMRNVSGTVNSKNNKFYLYPGCAGENGLATG LTYPNEIGFSSETGCNNVFEYNPETNSWR EMAPFPEALYMSSTVEYKPGKNMIINYGVR SIGIPVVGGSYFYIHDIENTYERIPVSHQFFSTAACIVNDTIYVFCGR ETAY VNSVPTTMVPPFFGLYKYNLLTGEETYVANK YTAVERNENVANPTAFDQDSFNSSVYLPENLGIIFTAGIFRNNHKGTVLYKFDVATEEITKFMSPDFFLVSPKIRSPINDNK LIMFTGVSGETTIPTVTGTGNGAYYTSAFRDIIEIDFNKDYRYREKLPYPGFVCGNGIGIFKDKRVILAGGCVNYGEGESSFFSDPDNKISTGLMHSNHAMIVDFNKSVIA TNPVLSRLIPLNDNCTIPNTPVANWYNTQSASTSSNKVYTTILEKKN TAYLFDTVKKTWTVSPVPAVVNYRLNMKGAAIDFDEANTIYSA TDPSSEFSLQKINNFNSTITGY SLVLLGSTAGTKDAADSSDRTLNKSTLFYFSPDGSITWAGKPYFKDNFYTSNDNRRILREYGNFYFNSTQNVFYMIGGKDEMSLNKSSGYGVYSVTNSVTR YHDLIES GSSIPLEKYAKYGSALVSI SDTECIVFGGTTFDENGDKVYSSETFKLSLVNK GDYLTPIYTQLQSIPIEFESSEGW MR ATK INTNEIVLYAGR AKNYSGTLVFK YIIDTNK EMVTQTE CFALADLVKTPK GEFYIPFGMVK GSNGSISVTPPFKINAAYKPKIK FTSFGISPVFTLGNR TETNDIFITASTK VLDSNDNIHYFGGVI SGGAKSGESVFNH WKAKVTENS DGLFTLGTWQEDTIPSSLCQNSNGVAKTLPDGR IVLIGGLWDLSSASTDLQTNPIKMYR STLTIFDPVTNTYTEKQPTY INNPLSVNVK FINGTMSL TNNP DEILIFGGYK YWSTVTDYHANTCFNK VFKYSISANTLTYLGK LSITDNPMAISPFR NYAAVADNGTLFCFSSDNLSPDAMGNTGKILNNQVTVFDTLNTTTDEADANGLFK TIDLTKESVIPAALESGR VRPTYINLTGK HSNGNFYIFPVWETK STITAGIFNPYNK KYNTGIPWQLISVPIEANKMIAPVTTALSTSTGIFSSDSSAYPEMFEKLSGK IIVLPGMFGSRLNKSVYINGLGMFVVIS
g297	220803:221486 (compl.)	25,675	117	MATDIFQEV TIPDSMTKAVFDGNGFCIKEGYVGTYNIHSTTREYIGSNREYVPGVTVAAAYAR DLPPFAEYKQKVAAGVPOAILR AVDYKGWQIVIDYRGYTIYNTTDKTP KVVTSGFDIPTGFTLLAPTSEFDEWDGAK WVVENKKLOASIDAANK TQKTLMSEALEHINIYQDMLDFEMADDIEAVEAKLAEWKRYIAVYKISTVDPSTIVWPEKPVSM FGDN
g298	221502:223613 (compl.)	74,343	589	MANLSRIAFLVR TIPNLNSEEITGIYPGELFLVGDKK AIFYK QDGTVIS SPSEAASTFLTIEGNLEEISDKAK ARENLGKSAALKESGTDKGLYSVGDNFNSFLEESTIIS ALDFNTYKFKSNEVIVLAK AAATNIPVGLDLGNNVINMATGTSSK EGFIAPVSDLANEKASYVLRCSYGTSSR TFAVESIYTDKHKPNLNDLNVYSKETAK DTFFQISNL FSEIVDDDDKEIARLNLGLGTASVLNK GTAK GNVIVVGDGGLGINDNKVAALKSVIPAFYQLTDDSPDKPEVDKDYGVLSFPISK TLTGVLVSSDGLVSLGIDGADGVELS KVYLANNLPTLNELKALGTENNLSIDKDKAK ARNLGLGDSATK NVGTETGTVAAGD DSR ITGAAQDK NLKDLDKSIAR TNLELGSAAQKTAQTNITD TTSGSLMQV GAF GLGGSGVPPFSGEGKDLIAFLKNAPSGMYR FGDSSSYTNGIPSGQYSDILLMDHGGSNSKSAIAVDFSGQVFAICLRADGTSSGWVTIFSDTDFSNLLLP IGVPIPYPLATP PK NFLKANGSAFDQVAFPKLLNYYP SGVLPDLR GEFIRGWDDGRNVDTGRN ILTNQGD AIRNITGKLNVLSSSDEPFSTADLFWANEGALDIIR TGFSHSSGGGGGAGGGV SGFTFDASK SGVPTANENRPRNIAFNIIIRAK
g300	224277:226469 (compl.)	77,968	142	MANSRIVFILKRIPSMADPSINDVYPGEHFVTEDSGQLYKDLQGIIVTVGGGKAAVPASWVYTPSAKNFTFNFEGSNLEALEPFSSSIRFLTNANTVIGAGQNTITINIT KNGSTLATR DFPIVDIIGR ELRKASIPGNAICDILFDG TKFRVYNTYGFSS TDGYSYFFDNNVVDGGSFVINSNDKK GFVVLEDGTIVVGDGDKL TNVTDL FQLK GNST FDGNVNTGKLTGNNRVYSSK GYDGILLSDIPTLDWSKLTN ILPQDKVPAVTDNSNKIATTKVWADYIAYLGLTGGGAGGNGAVIDTKIQSTTATEGQTSFDFEYNADSLIV LSTGHVLVGGEDYATDGT TIVFTDPCDADEELMFINIGKLDQEKYLTADQINAIITGSNFATVTVYVQEQIDELIGGSSDALD TLK ELADALNNDPDFATTITNQLAGKADK QHSHTWADITSGLP TATESK SGLIQLLTVSEAIK GTDTTKAMHAAGTRELIRGVGFNGFNGVANYATSTVLTKNSIGRWRNFNQP GI VVTLV YTELDATGWSIVPIYNYST GPIITKCPTEHGFPTNSGIQITITLNAYSSVFCVFE GNSTLNGESNMMFGEALAYGAGLKKVLGVNGYQVLP SGIIIQW GKLSINIANLIQNTYTDGRK TSSITNGTY PIPFPTSC LFAFGGSNDTANTFVTTVY PANSSGFAYKYESTGGPASAGSNTLSFIAIGY
g326	241918:243252	50,944	56	MPEMNPDTIRNQVIEIAKGEAVITVNPVPCSLDTFLSAMTSPEVVINTFLPSDLYDKDLSNEEFKEIAFSVNTVLSYTYGATNVVPPSFDNCPPLARQNL LLQNFIEA GGTK ETLDVLGVK GFEKAFSMLLLPLGIAQCANMTDVRKSEI ALNKINDGEDRIKNNINVIMGNIPLETSLDDIEALAREGNIDVNLFLARQKPEFNTVEV IHRFKEFLDP SRVMIQVDFTEDAFRTLLDEGYFS DVRDLFTVANPMIMELADLREKYWDKFSWERALVKYTFSEIEILNRI ISEGKMNAVASV FQVMSDEMILANDEITFWRPEFNFDYW FKGENDNKELRAEYAKHTPTNETVAVANVEVIRRHKKQELLMLAAITVSEMPIAVELTDES GS TVESRVSNPQNDKPV DQRIKHKLLIPEVVDGKLQYSEEKDLSNFYEG
g333	249517:250233 (compl.)	27,166	75	MLNLNLLVNHIKARVGFTHRKLELSEEAIACLQ QETLPTISTYMPFCMSSCNLGTDKVAQNMTY WIEAK SDEFDILDVTLVMPVVR GMNSSAYMLTPIGADYQTIIGAI TTSKFASTLNSAVLNPETWQFLPPNQLRVYSTYSLTELALVLK TTHRKDFSTLPV GASEMVKL AYYDVMMDV GFGRKYFSQLNTAVAQINLNMDELSSIPDKRDDLIERLR KEQLKYSPTKKIYVA

Table S2. HHpred-Homology detection & structure and function prediction by HMM-HMM comparison of ϕ R1-37 ORFs. Included are hits with a probability >50%*. Predicted functional categories for the hits are highlighted by different colors: **DNA replication and repair**, **Nucleotide metabolism**, **ATPases**, **Toxin-antidote system**, **RNA-polymerase and RNA interactions**, **Virion structural proteins**, and **Lysis proteins**

Gp (size aa)	Region of similarity (size of hit, aa)	Function	Organism	PDB ^s code	Probability (%)	E-value
Gp017(110)	12-50(54)	Putative regulator of transfer genes ARTA	<i>Staphylococcus aureus</i>	3gxq	72.13	2.2
Gp024(67)	21-43(113)	Fibrin, gpwac E; bacteriophage T4 structural protein bacteriophage assembly attachment protein	T4	1aa0	59.60	3.7
Gp031(180)	84-138(470)	IMP-GMP specific 5'-nucleotidase	<i>Legionella pneumophila</i>	2bde	69.69	4.6
Gp039(85)	58-85(199)	Uncharacterized protein; zonular occludens toxin	<i>Neisseria meningitidis</i>	2r2a	75.70	0,89
Gp045(368)	217-261(106)	TM1634; membrane protein	<i>Thermotoga maritima</i>	2vkj	58.49	1.8
Gp046(202)	17-133(167)	Peptidoglycan-recognition protein-LC isoform LCA; protein-peptidoglycan complex	<i>Drosophila melanogaster</i>	2f2l	91.2	0.27
Gp049(261)	15-91(146)	Lysozyme zinc amidase and an inhibitor of T7 RNA polymerase	T7	1lba	87.91	0.46
Gp051(117)	48-260(208)	Phosphate starvation-inducible protein; PHOH2, ATPase, PFAM: PF02562	<i>Corynebacterium glutamicum</i>	3b85	99.88	1.1e-21
Gp052(147)	5-92(204)	3'-5' exonuclease ERI1; nuclease domain	<i>Homo sapiens</i>	1w0h	59.91	2.4
Gp056(180)	10-87(1095)	RNA-dependent RNA polymerase	Simian Rotavirus	2r7r	65.48	0.75
Gp060(145)	1-175(240)	Molybdate/tungstate ABC transporter	<i>Archaeoglobus fulgidus</i>	2onk	100.00	4.1e-33
Gp061(929)	1-142(134)	AP4A hydrolase; nudix protein, diadenosine polyphosphate	<i>Aquifex aeolicus</i>	3i7u	99.91	3.3e-23
Gp066(118)	1-909(1005)	DNA double-strand break repair RAD50 ATPase	<i>Methanocaldococcus jannaschii</i>	3auy	99.98	1e-29
Gp068(274)	5-68(492)	Insulinase family protease	<i>Yersinia pestis</i>	3go9	64.23	7.9
Gp069(217)	2-274 (326)	Ribose-phosphate pyrophosphokinase 1	<i>Homo sapiens</i>	3s5j	100.00	5.1e-69
Gp070(584)	59-168(174)	HNH homing endonuclease; HNH catalytic motif-helix-turn-helix DNA binding domain	Bacillus phage SPO1	1u3e	97.91	4e-07
Gp071(58)	8-584(484)	Nicotinamide phosphoribosyltransferase	<i>Homo sapiens</i>	3dhf	100.00	3.2e-113
Gp074(164)	21-55(56)	SVTR protein, transcription regulator; homodimer, ribbon-helix-helix, transcription repress	<i>Sulfolobus islandicus</i> rod-shaped virus	2kel	83.10	0.64
Gp076(121)	98-143(106)	Ferredoxin VI	<i>Rhodobacter capsulatus</i>	1uwm	75.93	1.3
Gp077(131)	42-75(109)	Replicase polyprotein 1AB; RNA replicationATP-binding	Human coronavirus 229E	2j97	81.04	1.8
Gp081(463)	3-110(161)	Lincosamide antibiotic adenylyltransferase LinA	<i>Staphylococcus haemolyticus</i>	4e8j	99.68	6.4e-17
Gp083(2553)	181-253(340)	Diaphanous protein homolog 1; helix bundle, protein binding	<i>Mus musculus</i>	1v9d	67.80	28.39
Gp084(119)	2141-2310(268)	Lytic transglycosylase; lysozyme	<i>Pseudomonas</i> phage ϕ KZ144	3bkh	98.75	2.4e-08
Gp086(178)	2-119(213)	Muramoyl-pentapeptide carboxypeptidase	<i>Streptomyces albus</i>	1lbu	99.91	7.3e-24
Gp089(364)	42-176(120)	Crossover junction endodeoxyribonuclease RUSA	<i>Escherichia coli</i>	2h8e	99.97	1.1e-31
Gp092(231)	267-335(213)	Muramoyl-pentapeptide carboxypeptidase	<i>Streptomyces albus</i>	1lbu	85.86	0.72
Gp093(165)	28-165(166)	Calcium binding protein	<i>Streptomyces coelicolor</i>	3akb	56.59	36
Gp096(690)	88-164(123)	Anti-sigma-factor antagonist (STAS)	<i>Moorella thermoacetica</i>	3zxn	63.79	25
Gp097(219)	146-410 (596)	Telomerase reverse transcriptase	<i>Tribolium castaneum</i>	3kyl	99.83	1.2e-20
Gp099(1393)	19-219 (277)	Putative ATP-dependent CLP protease proteolytic S	<i>Homo sapiens</i>	1t6g	100.00	9.2e-38
Gp100(358)	731-1392(1342)	DNA-directed RNA polymerase subunit beta	<i>Escherichia coli</i>	3lu0	100.00	1.4e-132
Gp101(347)	81-351 (203)	Non-canonical purine NTP pyrophosphatase	<i>Coxiella burnetii</i>	3tqu	100.00	2.8e-56
Gp102(603)	214- 327(115)	Polyadenylate-binding protein 1; RRM fold translation initiation, RNA-binding	<i>Homo sapiens</i>	4f25	99.71	1.7e-17
Gp103(171)	90-561(1524)	DNA-directed RNA polymerase beta' chain	<i>Thermus thermophilus</i>	2a6h	100.00	2.5e-92
Gp106(30)	9-145(193)	dCMP deaminase	Enterobacteria phage T4	1vq2	100.00	6e-37
Gp107(175)	8-29(145)	BH1478 protein; unknown function	<i>Bacillus halodurans</i>	2qup	92.12	0.13
Gp111(144)	21-152(174)	HNH homing endonuclease	Bacillus phage SPO1	1u3e	99.93	1.8e-25
Gp114(69)	1-96(102)	Aprataxin (resolves abortive DNA ligation intermediates during DNA repair)	<i>Homo sapiens</i>	3kt9	82.11	9
Gp117(51)	1 -66(122)	Hypothetical protein	<i>Vibrio cholerae</i>	1u9d	83.38	1
Gp120(124)	12-43	Fusion protein	Human respiratory syncytial virus	1g2c	80.89	0.54
Gp129(210)	34-107(294)	Putative diguanylate cyclase / phosphodiesterase	<i>Thiobacillus denitrificans</i>	2r6o	76.93	2.9
	162-207(386)	RAN GTPase-activating protein 1	<i>Schizosaccharomyces pombe</i>	2ca6	91.73	0.076

Gp130(201)	13-47 (73)	RSGI RUH-014, rhomboid family protein	<i>Arabidopsis thaliana</i>	livg5	60.8	5
Gp131(120)	2-100 (94)	DNA binding protein HU	<i>Anabaena sp</i>	1p71	99.9	1.4e-26
Gp133(154)	6-153(175)	Gluconate kinase	<i>Escherichia coli</i>	1knq	99.52	5.2e-13
Gp143(210)	54-170 (592)	Terminase, DNA packaging protein Gp17	T4	3cpe	98.4	2.5e-7
Gp145(469)	1-374 (592)	Terminase, DNA packaging protein Gp17	T4	3cpe	100	1.9e-36
Gp147(268)	1-78 (463)	Maltose ABC transporter periplasmic protein	<i>E.coli</i>	2xz3	59.9	40
Gp149(590)	1-458 (510)	Tail sheath protein GP18	T4	3foa	100	3e-49
Gp150(250)	39-86 (234)	5'-methylthioadenosine nucleosidase	<i>E.coli</i>	3o4v	63.3	4.4
Gp151(585)	272-376(105)	Homing endonuclease	Bacillus phage 0305phi8-36	3rp3	98.91	5.8e-9
	271-377 (136)	DNA mismatch endonuclease	<i>E.coli</i>	1cw0	98.79	1.5e-8
Gp154(240)	91-173 (238)	Zink protease, cell division protein FTSH homolog	<i>Aquifex aeolicus</i>	2di4	87.4	1.4
Gp157(89)	4 -64(333)	MutL DNA mismatch repair protein	<i>Escherichia coli</i>	1b63	65.63	7.7
Gp163(174)	3-59(155)	Calcineurin subunit B type 1	<i>Homo sapiens</i>	3l18	69.85	5.2
Gp164(242)	7-123 (786)	POL II, DNA polymerase II	<i>E.coli</i>	3k59	92.0	0.14
	7-125(919)	DNA polymerase delta catalytic subunit	<i>Saccharomyces cerevisiae</i>	3iay	74.2	2.5
Gp167(420)	4-342 (919)	DNA polymerase delta catalytic subunit	<i>Saccharomyces cerevisiae</i>	3iay	100	1.1e-40
Gp168(159)	114-137 (63)	Complexin-1; snare proteins	<i>Homo sapiens</i>	3rk3	76.7	2.2
Gp170(576)	15-563 (919)	DNA polymerase delta catalytic subunit	<i>Saccharomyces cerevisiae</i>	3iay	100	3.7e-63
Gp171(561)	265-367 (155)	DNA mismatch endonuclease	<i>E.coli</i>	1cw0	99.6	7.1e-16
Gp172(61)	20-61(126)	DNA repair protein RadC	<i>Chlorobium tepidum</i>	2qlc	82.11	2.2
Gp174(259)	4-52 (195)	Protoglobin	<i>Methanosarcina acetivorans</i>	2veb	74.5	1.7
Gp176(301)	233-261 (386)	RAN GTPase-activating protein 1	<i>Schizosaccharomyces pombe</i>	2ca6	81.4	0.31
Gp177(292)	47-125(247)	Beta-phosphoglucomutase	<i>Bacteroides vulgatus</i>	3dv9	76.52	3.7
Gp178(472)	10-414 (1524)	DNA-directed RNA polymerase subunit beta'	<i>Thermus thermophilus</i>	2a6h	100	5e-71
Gp179(310)	211-309 (155)	DNA mismatch endonuclease	<i>E.coli</i>	1cw0	99.3	4.5e-12
Gp187(203)	11-102 (195)	Hypothetical protein	<i>Aquifex aeolicus</i>	1xm7	99.9	7.6e-25
Gp188(28)	2-18(58)	Hypothetical protein	<i>Schizosaccharomyces pombe</i>	2elk	69.65	1.3
Gp189(59)	11-29(119)	CREB-binding protein	<i>Homo sapiens</i>	3p1f	81.72	0.59
	27-56(141)	Ephrin-B3	<i>Mus musculus</i>	3d12	62.10	3.1
Gp190(394)	160-232 (169)	Putative uncharacterized protein; YSCO, type III secretion apparatus	<i>Chlamydia trachomatis</i>	3k29	58.7	81
Gp191(440)	237-393(295)	Putative sensory BOX/ggdef family protein	<i>Vibrio parahaemolyticus</i>	3lid	73.5	15
Gp193(325)	100-171(178)	Diguanylate cyclase/phosphodiesterase	<i>Caldicellulosiruptor saccharolyticus</i>	3mtk	82.29	1.6
Gp197(282)	183-218 (57)	Septum site-determining protein diviva	<i>Bacillus subtilis</i>	2wuj	69.45	3.9
	55-222 (470)	DNA gyrase subunit A	<i>Mycobacterium tuberculosis</i>	3ilw	57.2	1.1e+2
Gp198(263)	144-189 (78)	NK-lysin; saposin fold, antibacterial peptide	<i>Sus scrofa</i>	1nkl	55.6	0.45
Gp203(248)	23-240 (1524)	DNA-directed RNA polymerase subunit beta'	<i>Thermus thermophilus</i>	2a6h	93.3	0.099
Gp204(505)	164-200(107)	Probable protease HTPX homolog	<i>Vibrio parahaemolyticus</i>	3cqb	96.35	0.00085
	146-344(482)	Nuclear membrane zinc metalloprotease ZMPSTE24 (FACE1)	<i>Homo sapiens</i>	4aw6	96.35	0.0051
Gp211(214)	78-97(167)	Peptidyl-Lys metalloendopeptidase	<i>Grifola frondosa</i>	1g12	83.16	0,57
Gp213(229)	2-173 (154)	Ribonuclease H1, HS-RNase HC, RNase H1	<i>Homo sapiens</i>	2gkb	100	4.7e-33
Gp214 (106)	16-101(124)	Hypoxia sensor histidine kinase response regulator Dost	<i>Mycobacterium tuberculosis</i>	3zxx	82.67	4.5
Gp218 (133)	69-124(133)	MCTP2 protein	<i>Homo sapiens</i>	2ep6	72.13	6.2
Gp219(119)	1-102 (755)	Biodegradative arginine decarboxylase	<i>E.coli</i>	2vyc	94.3	0.049
	1-100 (153)	Two-component response regulator, LUXR family	<i>Aurantimonas sp.</i>	2cz5	92.8	0.59
Gp229 (132)	15-59(59)	PRE-mRNA processing protein PRP40	<i>Saccharomyces cerevisiae</i>	2b7e	92.38	0,029
Gp230(469)	162-275(155)	DNA mismatch endonuclease	<i>E.coli</i>	1cw0	99.5	1.3e-13
Gp231(424)	5-423 (1524)	DNA-directed RNA polymerase subunit beta'	<i>Thermus thermophilus</i>	2a6h	100	3.5e-94
Gp232 (31)	13-25(274)	Urease Accessory Protein UreF	<i>Helicobacter pylori</i>	3cxn	86.94	0,2
Gp236 (140)	83-130(97)	NF-YC; histone-like PAIR	<i>Homo sapiens</i>	1n1j	82.70	0.23
Gp237(236)	132-221 (138)	Endonuclease, endodeoxyribonuclease I	T7	1m0d	97.2	0.00031
Gp239(131)	16-83 (135)	Transcriptional regulator CUER	<i>E.coli</i>	1q06	57.6	35
Gp243(574)	9-574 (543)	CPN60(groel); chaperonin, chaperone, groel, HSP60	<i>Thermus thermophilus</i>	1we3	100	1.1e-92

Gp244(290)	30-90(179)	Ubiquitin-conjugating enzyme E2 H; UBE2H	<i>Homo sapiens</i>	61.69	17
Gp247(353)	9-307(407)	DNA primase	<i>Aquifex aeolicus</i>	2au3	99.52
Gp250(407)	5-376(2050)	Protein RECArecombinase A	<i>E.coli</i>	3cmu	100
Gp258 (124)	14-69(106)	UvrB interaction	<i>Geobacillus stearothermophilus</i>	3fpn	70.59
Gp261(658)	5-633(1342)	DNA-directed RNA polymerase subunit beta	<i>E.coli</i>	3lu0	100
Gp267(697)	11-697(671)	DNA-ligase	<i>E.coli</i>	2owo	100
Gp269(110)	2-108(168)	Single-stranded DNA-binding protein	<i>Mycobacterium leprae</i>	3afp	99.98
Gp274(452)	2-452(1407)	DNA-directed RNA polymerase subunit beta'	<i>E.coli</i>	3lu0	100
Gp278(295)	13-76(144)	Ubiquitin carboxyl-terminal hydrolase 8	<i>Homo sapiens</i>	2a9u	54.00
Gp280(118)	1-115((179)	oxidative stress DPS DNA binding	<i>Lactococcus lactis</i>	1zuj	99.61
Gp282(225)	11-76(127)	Hypothetical protein	<i>Shewanella frigidimarina</i>	2ook	52.90
Gp284(98)	9-82(97)	Putative RNA-binding protein	<i>Lactobacillus plantarum</i>	3kwr	99.75
Gp285(288)	15-288(491)	Type VI secretion system component	<i>E. coli</i>	2p5z	99.87
	(211)	GPV baseplate assembly protein V	Enterobacteria phage P2	3qr8	99.58
Gp286(477)	68-477(574)	Exodeoxyribonuclease V	<i>Deinococcus radiodurans</i>	3e1s	100
Gp287(225)	53-223(174)	HNH homing endonuclease	Bacillus phage SPO1	1u3e	99.97
Gp289(251)	28-232(206)	Endo-type membrane-bound lytic murein transglycosylase	<i>E.coli</i>	3t21	99.95
Gp294(1302)	665-1163(568)	Kelch-like protein 12	<i>Homo sapiens</i>	2vpj	98.92
Gp295(1372)	206-538(624)	Kelch-like ECH-associated protein 1	<i>Homo sapiens</i>	1zgj	98.96
Gp296(928)	482-847(568)	Kelch-like protein 12	<i>Homo sapiens</i>	2vpj	98.58
Gp297(228)	90-222(102)	Uncharacterized protein	<i>Chromobacterium violaceum</i>	2kz6	99.9
Gp298(704)	547-704(312)	Bacteriophage T4 short tail fibre	Bacteriophage T4	1ocy	99.97
Gp299(173)	74-146(102)	Uncharacterized protein	<i>Chromobacterium violaceum</i>	2kz6	99.6
Gp300(731)	251-497(312)	Bacteriophage T4 short tail fibre	Bacteriophage T4	1h6w	94.89
Gp301(210)	94-140(151)	Uncharacterized protein	<i>Streptococcus mutans</i>	2qkp	79.44
	94-136(118)	Sensor protein DCUS	<i>E.coli</i>	2w0n	61.12
Gp302(109)	4-107(94)	Protein LUXT	<i>Vibrio parahaemolyticus</i>	3b4s	73.9
Gp303(180)	74-128(163)	ATP-dependent RNA helicase DBPA	<i>Bacillus subtilis</i>	2hvj	74.99
Gp304(430)	2-197(331)	Uncharacterized protein	<i>Cytophaga hutchinsonii</i>	2r44	99.58
	14-275(447)	Replication-associated recombination protein A	<i>E.coli</i>	3pvs	99.55
Gp305(611)	470-576(218)	Conserved hypothetical protein BATB	<i>Bacteroides thetaiotaomicron</i>	3ibs	98.86
Gp308(130)	1-103(104)	Antidote protein of plasmid maintenance	<i>Coxiella burnetii</i>	3trb	99.79
	1-99(104)	Antidote protein of plasmid maintenance	<i>Nostoc punctiforme</i>	3cec	99.72
Gp309(104)	1-103(104)	Antidote protein of plasmid maintenance	<i>Coxiella burnetii</i>	3trb	99.75
	1-99(104)	Antidote protein of plasmid maintenance	<i>Nostoc punctiforme</i>	3cec	99.7
Gp310(87)	2-29(110)	60S acidic ribosomal protein P11-P1 (L12P)	<i>Triticum aestivum</i>	3iz5	74.58
Gp311(441)	12-405(391)	DNA gyrase BGYRB	<i>E. coli</i>	1ei1	100
Gp313(248)	132-242(136)	VSR endonuclease	<i>E. coli</i>	1vsr	99.7
Gp314(357)	2-357(420)	DNA gyrase subunit B	<i>E. coli</i>	3nuh	100
Gp326(445)	175-309(208)	Response regulatory protein	<i>Ps. fluorescens</i>	1yio	52.42
Gp327(70)	1-70(496)	DNA topoisomerase 4 subunit A	<i>Streptococcus pneumoniae</i>	2nov	100
Gp328(311)	9-121(141)	TC3 transposase	<i>Caenorhabditis elegans</i>	1u78	99.03
	127-304(199)	homing endonuclease I-DMOI	<i>Desulfurococcus mobilis</i>	2vs7	99.93
Gp329(749)	1-677(716)	Topoisomerase IV subunit A	<i>E. coli</i>	1zvu	100
Gp331(238)	89-232(206)	Endo-type membrane-bound lytic murein transglycos	<i>E. coli</i>	3t21	99.85
Gp334(413)	10-389 (379)	Exonuclease	<i>Thermotoga maritima</i>	3tho	100
Gp335(251)	37-245(444)	DNAB replication FORK helicase	<i>Thermus aquaticus</i>	2q6t	96.49
Gp336(282)	132-252(136)	VSR endonuclease	<i>E. coli</i>	1vsr	99.7
Gp337(357)	1-268(444)	DNAB-like replicative helicase	Bacillus phage SPP1	3bgw	100
Gp341(87)	43-79(442)	Adenylosuccinate synthetase	<i>Plasmodium falciparum</i>	1p9b	64.3
Gp352(97)	17-74(109)	Arginyl-tRNA synthetase	<i>Klebsiella pneumoniae</i>	3gdz	60.0
Gp360(82)	12-51(85)	Splicing factor 3 subunit 1	<i>Homo sapiens</i>	2dt7	81.5

Gp363(84)

12-80(157)

Alanyl-tRNA synthetase

Pyrococcus horikoshii

1v4p

81.90

0,37

*Probability is the most sensitive criterion; a value >95% indicates almost certain homology; value >50% indicates a possible hit.

§ PDB = Protein Data Bank

Table S3. PHIRE repeat search (length 20, mismatches 4). Alignment of the AU-rich repeat hits. shown are the repeats with 5 nt of flanking sequences.

Promoter motif logo		
Promoter #	Sequence	Before gene
x	UCUUAAGGAUAUAUAAUUAAAUGAAGAAG	overlapping g001
1	UUCUAAGGAUAUAUAAUUAAUUAAGAAUCU	g011
2	UUUUAUUGAUUAUAUUAUUAUAAUGGAGGAA	g029
3	UUUCAUACAUAUAUUAUUACUAAGAUAAAU	g060
4	UUUUAUUGAUUAUAUUAUUUUUAUUGUAGUUU	g063
5	UUUUAGUAAUAUAUUAUAAUAAUGAUAAAC	g064
6	UUUUGUUCAUUAUAUUAUUCUAAUGAUACAU	g073
7	CAUUAUUGAUUAUAUUAUUUUAGUGUUAGUA	g073
8	UAUCAUUGAUUAUAUUAUUUUAAUGUGAAAU	g093
9	CAUCAAGGAUAUAUUAUCUUUGUGAUCAA	g091
10	UUUUAUAAUAUAUUAUUAAAUGAUUCUCA	g096
11	AAUUGCUGAUUAUAUUAUUUUAGUGAGUAGG	g101
12	ACUUACGAGUAUAUUAUGAAAAGAAUUUCG	tRNA-Arg g106
13	UUUUAUUCAUAUAUAAUUAUAAUAGAAUAU	g107
14	AAUAAGCAUAUAUUAUUACCAUGAAGUAG	g112
15	CUUUAAGAAUAUAUUAUUUUAGUGAUUUUG	g118
16	UUUUAUUGAUUAUAUUAUUAAAUUGAUUGUU	g122
17	UUUUAUUGAUUAUAUAAUCAUAUUGAAUAAU	g132
18	UUUCAUGCAUAUAUUAUAAAGAAAGAUACA	g136
19	UAUCAAAGAUUAUAUAAUUUUUAUAGAUAAU	g137
x	AUUCAGUGAUUAUAUAAUUAUUUGAAGGUU	Within g138
20	AAUUAAGAUUAUAUUAUUUAGGUGAUAGAA	g144
21	UUAUAGAAAUAUAUUAUUAAAUGAAAUAU	g170
22	AUUCAUUGAUUAUAUAAUUUUUAUUGUACUU	g173
23	AUAUGGUCAUUAUAUUAUUUAUUGACGAAU	g176
24	AAUAGUGAUUAUAUUAUUUAAAUGUACUU	g179
25	UUUUAAAGAUUAUAUAAUAAUAAUGAUAAU	g182
26	UCUUAUUGAUUAUAUUAUUUUAAUGAAUAGU	g188

27	UUUGAUGAAUAUAUAAUUAAAUGAUCUUU	g194
x	GAAAUUCAUUUAUUAUAAUGAUGUUUUAGAU	Within g194
28	UUUCGUUAAUAUAUUUAAAGUGAAAAGUA	g211
29	AUUGAAAAUAUAUUUAUUAACUUGAAUGAC	g245
30	UAAUAAUGAUUAUUAUUUUUAUUGAAGUAA	g248
31	UAUCAAGUAUAUAUUAUUUJAGUGAAAA GA	g251
32	UUUCAUUAAUAUAUUUAAUUGAUGAUAUCA	g252
33	UUUAAACAUAUAUUUAAAUUGUGUAUAGU	g261
34	UUUUUUCAUAUAUUUAAAUGAUAAAU	g262
35	UUUUAAAAUAUAUAAUUAAAUAGAAUAAU	g264
36	AUAUUUAUAUAUUUUUAUAAGAAUUAUU	g275
37	UUUUACAUAUAUUUAAUUGUAAUCU	g283
38	AGUUUUUGAUUAUUUAUUUUGAUAGCU	g286
39	UUUUUAGGAUAUAUAAUUAUUGGAAUACC	g303
40	AUUUAUUGAUUAUUAUUUUAGUGAAUGAG	g304
41	AAAUAAUGAUUAUUAUUUUUAUGAAUAAA	g306
42	UUUUAAAUAUAUAUUUAAUAAUGAAUCCA	g310
43	UUUUAAUGAUUAUUAUUAUCAUGAUUUUC	g316
44	AAAUAAAAUAUAUUUAAAUAAGAAUACU	g325
45	UUUGAAAUAGAUUAUUUUUAGUGAGAGGA	g326
46	UUUUUUGUAUAUAUAAUUAUAAGGAUUAUU	g327
47	UUUUAAAUAUAUAUUUAAUUAUGAUAAUC	g330
48	UUUCAUUAAUAUAUUUUACAUGUAAUUG	g283
49	UUUUAAUGAUUAUAUUUAGUUUGAUAAUA	g334
50	CAUUGUAGAUUAUUAUUAUAAUUAAGAGGGG	g339
51	AAAUAAAGAUAUAUUAUAAUUUUGAUUAUCU	rev g059
52	UUUCAUUGUAUAUAUAAUUAUAGUGUGAAUA	rev g152
53	UUUUAAAAUAUAUUUAAUAAUGAUAAUG	rev g159
54	UUAAAAUAAUAUAAAAUUUAUUGAAUAAGG	rev g166
55	UUUUAAUAAUAUAUUUCAUAAUAAUUUUU	g169
56	CUUCAAGAUUAUAUUUUUAAUGCGGUCU	g232
57	AUAAAAUAUAUUUAUAAUUAUGAAGAGUUU	g274
58	UAAAUAUAUAUUAAUGAAAGCAUUAUU	g282
59	CUAAUAUAUAUCUUAAAAGGUUAUAUUAA	g333

Table S4. ϕ R1-37 σ^{70} promoters identified by the BROM search tool from intergenic regions without a phage-specific promoter and the g233 antisense promoter.

Region between genes	The intergenic sequence upstream the ORF starting at the ATG codon. The -35 and -10 sequences of promoters are in bold upper case. Ribosomal binding sites underlined.
g081-082	tggtatgacttctctattcgattagtataactaa TTGATG ttccatcaacattcatt TAGTAACG tttatattaaataaaaaacg <u>GATC</u> aaatt ATG
g082-081	aattgatccgtttttaattataataaacgttactaa ATGAAT ggtgatggaacatc AAATTAGTAT actaatcgaat <u>AGGA</u> aaagtcatacc ATG
g094-095	tatctcagcctaaactatagactattaagtgt TTTATA Aaatcataact TTTTATTAG aaaaataataaacacggtattagtactaaa <u>AGGT</u> taaaataaaa ATG
g095-094	accgtgttatttttctaataaaaagtattatga TTTATA Aaacactaata GTCTATAGT ttaggctagagatattctatatagaaaat <u>TAAGGA</u> ataaaa ATG
g099-100	aaatttaaaatgataaataaacacataattttac TTGACT taagatagttattaa TCTATAAT Aatattaataacgaattacaaccaaggc <u>GAG</u> aaacgaagat ATG
g100-099	agtgtaatagggaat TTTATT tcaatgtaaa TCTTACTAT ttttagaaaagtgtaatttttagtattttcaaacatagttttagaattacatacataactatattggt <u>AGGT</u> aat ATG
g116-117	aaataccaaatgttaaattggtgga ATGCCT atccgagataactac AGATAGAAT agttaaaaggcggttagaaaaattcactgcatccccgttcctaatttat <u>TAAGGA</u> acgaaa ATG
g117-116	ggtgcacgcttctactacctagaaaaataaacactacatccccgtaacgtggg TTGTTA atgcccttatctag AGTTATCTA aataaggcacccttcagcctga <u>AAGGA</u> attc ATG
g167-168	atctaactagatttatctctataaaaactagtt TTGAAA taaacactaaaatagaaag TGATCAATT ttctatttcgtaaaccctta <u>TAAGGA</u> tttc ATG
g168-167	atgaaatcctataagggttaacgaaa TAGAAA attgatcact TTCTATTTT agtgtatttcaaaactagtttatagagataa <u>ATC</u> tagttaa ATG
g208-209	aattcaaftaatattaaaagaataaatctaattatagaacaagtg TTTAGT tctatagaaac TAATAAGCT tatttgttatag <u>AAGGA</u> atctctat ATG
g209-208	aataagcttattag TTTCTA tagaactaaacttg TTCTATAAT tagatttattcttttaattataaattgaaatttataatatttacttatatggaacaaaagtttagattctaattag <u>GGT</u> aaatc ATG
g303-302	atatcctaaaaaga TTGTTT gtttacacaaag TTTTAAAA Ttatatgatccacattaaggttttgaacaagaatttagaatctttatgtttgtttaaaatttttag <u>AAGGA</u> ttgctat ATG
g286-285	taatttcttcagacaaaatagctatcaataataatataatca ATAACT agtatataaaa TATTATACT tacaacaaataattgtaaaacta <u>TAAGGAG</u> ttttat ATG
g236-237	ttctctcacctttaaggatagattcgttatttta TTGAAA aactatataga TGTTAACAT gaatatagttaacttatagaaaat <u>AGG</u> catctata ATG
g237-236	ttataagatgcctatattttctataagttaaactatattcatg TTAACA tctatagtttt CAATAAAAT aacgaatcatataaccttaa <u>AGGTGA</u> agaga ATG
antisense g233	ttcattcttattagcctgagttgatgcttgtaataagaacagattgatctacaacgagaccgattaaaataatatact TTGAAG attatgcgtatacg TAGTATTCT acaatcagacctttaccagtttctgctgatagaggtctg gta ← g232 start aat ← g233 end

Table S5. Rho-independent transcriptional terminators of ϕ R1-37.

No	Range	Dir	Hairpin sequence ^a	Downstream 15 bp ^b
1	6581:6608	for	GAAAAGGTTCTCAAAAGAGAACCTACTC	TTTTTTTATTGATAT
2	14453:14484	for	CTCCCCATACCCTCTTTTCGAAGGTATGGGGAG	TCTTTGTTATGCTTC
3	14453:14484	rev	CTCCCCATACCCTTCGAAAGAGGGTATGGGGAG	TTTTATTTCTTTTTTTT
4	17055:17082	rev	CCCCTTATAGGCTTGAGCCTATAAGGGG	TTTTATTTTTTTAAGG
5	21973:22000	for	CCCTAATATACGAGGCTGTATATTAGGG	GTTTTATTTTCAAAG
6	24980:25009	for	CCCCATATCCTTCCAAAGAGGGATATGGGG	TTATAAAACCATTTT
7	25644:25674	for	TTTCCTCTGTATCAAAACGATACAGAGGAAA	TTATAAATTCTTTTTT
8	30034:30063	for	CCCTTATACCTTCTTCGAGGGTATAAGGG	GTTTTACTTTTTTTGTA
9	30796:30818	for	CCGTATCTTGATAAAAAGATACGG	ATTTTTTAATCACTTT
10	32414:32443	for	ACCCTATACCTTCTTCGAGGGTATAGGGT	TCATTCTTTTTTTTTT
11	32926:32955	for	CCCTTATACCTTCTTTGGAGGGTATAAGGG	AAGTTTTTTTAGAAAA
12	33804:33831	rev	CCCTAATCTAAAATGAATTAGATTAGGG	AGTTTTTTATCAAAAA
13	45186:45211	for	CCCCTGTAGGTTATTTCTACAGGGG	ATTCTTTTTTTATGAT
14	51008:51035	for	CCCCTATAGTCTATCAAGGCTATAGGGG	AATTTTTAACTTTCTC
15	53016:53044	for	CCCTATATCCCCTTTTGAGGGATATAGGG	TATTTTTTTTTTTAGC
16	53016:53044	rev	CCCTATATCCCCTCAAAAGGGGATATAGGG	GTTTTATTACATCAA
17	54072:54099	for	CCCATATCTCTCTATTGAGGGATATGGG	TAAATTTATAAATTT
18	56250:56263	for	GCGGGGAAACCCGC	TATTTTTTTTATCCAT
19	57404:57432	for	CCCTATATCCCCTCTTTTCGGGGATATAGGG	GTTTTATTTACTTTTTT
20	57404:57432	rev	CCCTATATCCCCGAAAGAGGGATATAGGG	AGTTTTTATTACACTA
21	64568:64595	for	CCCATAACGTTTAAAAGCGTTATGGGG	ATCCTTCAATATTAT
22	67551:67574	for	GGGATAATAACTTTTCATTATCCC	TTCTTTTTCATAAAGT
23	68134:68161	for	CCCATATCCTTCGAAAGAGGGATATGGG	AGTTTTATTTCTTTTTT
24	69994:70021	for	CCCATATCCTTCCAAAGAGGGATATGGG	AGTTTTTATTACAAAA
25	70443:70458	for	GCCTACTTAGGTAGGC	TTTTATTTTTATATAA
26	71235:71265	for	GAAGTGGGATAAATGGATTTTATCCCCTCC	TATTTAGTATTTTTTA
27	72748:72775	for	CCCATATCCTCTATTAAGGGGATATGGG	GTTTTCTTTCTTTCT
28	74131:74144	for	GACCTTATAAGGTC	TTTTTTTTTTTTTCATG
29	79770:79801	for	TACCCTATACGTACTTAGTTGCGTATAGGGTA	GAATATTTTTTTTTTTA
30	81890:81919	for	GCAGTAAGAATCTCTTATGTTCTTACTGC	ATTTTTATTTAATGGA
31	83119:83148	for	CCCCATATCTCCTGAAAAGGTGATATGGGG	TTTACTTTGCACTTA
32	83514:83535	for	AAGACCTCCATAGCGGGGTCTT	TTTTTTTTTTGAGGTT
33	86477:86504	for	CCCATATCCTCTATAAAGGGGATATGGG	GTTTTCTTTACAATT
34	88717:88738	for	GGAGAAACTTTAAAGTTTCTCC	TTTATATCTTTATTT
35	91965:91978	for	CAGGGTAAGCCCTG	TTTTTTTTTTTCATTA
36	95128:95158	for	TCCCCTATAGTCTGAACAAGGCTATAGGGGA	TTAACTCTTATTTTTT
37	98568:98593	for	CCCATATCCCTCCAAAGGGATATGGG	TTTCTTTTATACTATT
38	100420:100444	rev	CCCATATCTCTTTAAGAGATATGGG	TTATTCTTATTTTTTT
39	101607:101632	rev	CCATATCCTTCTTTTGAGGGATATGG	ATTATTTTTTAATTAT
40	104480:104505	rev	CCCCATACCTGTAAAAAGGTATGGGG	ATTTTTTATTCTTACT
41	109811:109842	rev	AACCCTATACCTCATAAAAAGGGGTATAGGGGG	ATTCTTTTTTTTTTTA
42	112757:112785	for	CCCTATATCCCCTTTTCGAGGGATATAGGG	AGTTTTTATTACTCTT
43	121025:121052	for	CCCCCTATAGGTAAGGCCTATAGGGGG	TTTATACTTTTTTTAT
44	121974:121999	for	AAATGGTACTCCTAAGAGTACCATTT	AATTATTTTTTTTTTC
45	125016:125043	for	CCTATATCCTTCTTTTGAGGGATATAGG	GTTTAATTTTTGTTTT
46	129212:129228	for	GGGGATGAAGTATCCCC	TTTTATTTTTAATACT
47	131858:131885	for	CCCCTATCCTTCTTTTGAGGGATAGGGG	TTATTAACTTTTTTAT
48	132926:132952	for	CATAGGGAGTTTAAAAACTCCCTATAG	TTTAAACAATTTTTTG
49	134552:134580	for	GACACAATACCTTGAACGGGTATTGTGTC	TATTTATTTTTTTATT
50	135679:135705	for	CCTTAACGGCTTATACGGTCGTTAAGG	ATTATATTCTTTTTTT

51	138301:138329	for	CCCTATATCCCCGAAAGAGGGATATAGGG	AGTTTTATTACTCTT
52	138301:138329	rev	CCCTATATCCCTCTTTTCGGGGATATAGGG	GTTTATTTAAAAGTA
53	140035:140063	rev	CCCCATATCATCTTAAAGGTGATATGGGG	TTTTCTTTCTATTTT
54	141817:141840	rev	CCCTATTCATTTTCGGTGAATAGGG	TAATTTCAACACATA
55	145324:145351	rev	CCCCATATAGGTTTCATTCCTATATGGGG	ACTCTTTAACAACCTC
56	152382:152409	for	CCCATATCCCTCTATCAAGGGGATATGGG	GTATTTACTCTTTTT
57	153158:153188	for	TATTCATACCCCTTAATCGGGGTATGAGTA	TTTGATTATTTTTGT
58	154249:154276	for	CCTATATCCCTCTTTTTCGGGGATATAGG	GTTTATACTACTTAT
59	154648:154675	for	CCCATATCCCTCTAAAGAGGGATATGGG	TTTCTAATAAGGAAA
60	158561:158578	for	GGGAGTAATTTTACTCCC	TTTATCTTATATATT
61	159512:159540	rev	CCCTATACCTTCATAAAGAGGGTATAGGG	AGTTTTATTACAACCT
62	159512:159540	for	CCCTATACCCCTCTTTATGAAGGTATAGGG	AATTATTTTTTTTTTT
63	161725:161765	rev	CCTCTCATGAACATAAAAATTAGTTAAATTGTTTCATGAGAGG	TTTTTATGCAAATTT
64	163099:163125	rev	CGCCTATGATCTTCACGGTCATAGGCG	TATTTATTTCTTTTT
65	171623:171655	for	GTCGCTAATATCTAATTTAAAGGTATTAGCGGC	TTTTTATTTCCATAG
66	179829:179856	for	CCCCTATAAGGTTAATTCCTTATAGGGG	ATCTTTTTTTTACGT
67	180392:180413	for	CCCGCCTATAGATTTAGGCGGG	ATTTTTTTTTTATCGT
68	181368:181382	for	GCCCCTGTATGGGGC	TTTATTTTTATGGTG
69	184003:184037	for	TTGAGAAACTCTAATAATCATTAGAGTTTCTCAA	AGTTAATCTTTTTTT
70	186090:186117	for	CGGGTAATAGTGTAAGCTATTACCCG	TATCTATCTTATTTT
71	186452:186472	for	CCCCTAAGAAATTCCTTAGGGG	TTTTATAAATCTTTT
72	187091:187115	for	CCCATATCCTCAAAGGGATATGGG	TTTCTTAATTTCTTT
73	191025:191053	for	CCCTATATCCCTTTTTCGGGGATATAGGG	AGTTTCTTTTTTTTTT
74	191025:191053	rev	CCCTATATCCCTCAAAGGGGATATAGGG	GTTTATTTACTTTTT
75	195161:195186	rev	CCCCTATAGGCTCAGTCCTATAGGGG	ATTCTTTTACTTTTT
76	197983:198011	for	CCCTATACCCCTCTTTATGAAGGTATAGGG	TTTATAATTTGAATT
77	198857:198885	for	CCCTATACCTTCATAAAGAGGGTATAGGG	AGTTATTTTTTTTTTT
78	198857:198885	rev	CCCTATACCCCTCTTTATGAAGGTATAGGG	AATTTTATTACACGA
79	200697:200721	rev	CCCTATACTATTGAATAGTATAGGG	TATTTTTAACTTTTT
80	203703:203728	for	TCTAGGGGAAAGGATTTATCCCCTAGA	TTTTTTTTTTTATTAG
81	203703:203728	rev	TCTAGGGGATAATCCTTTCCCCTAGA	TTGGTAAAATTAACCT
82	206958:206987	for	GGTACAATAAGGACTCTCCCTTATTGTACC	ATTTTTTTTTTTTATT
83	206958:206987	rev	GGTACAATAAGGGAGAGTTCCTTATTGTACC	AGAAACTGCGTCGTT
84	208214:208241	rev	CCCCATATCCCTTAAAGTGGATATGGGG	TGTTTTATTTTCTAA
85	209119:209140	rev	GCCCTATGGAAACCCATAGGGC	TTTTCTTTTTTAAGG
86	213671:213700	rev	CCCCTAATAAGCATCATGTCTTATTAGGGG	ACTTTTATTTACGAA
87	217885:217912	rev	GACCTTATAACCAGAAATGGTATAAGGTC	AATATTTTTTAACAA
88	220767:220794	rev	CCCCTATAACTTATGAAAGTTATAGGGG	TTCTTTAACAATTTCT
89	223706:223737	rev	CTCCCCTATAATCTCTTAAAGGCTATAGGGGAG	ATATTTTTCAATAAAA
90	226528:226554	rev	CCCTAATGATCTTTCTGGTCATTAGGG	ATTTCTTAACAAAAA
91	229650:229678	for	CGGGTAATGGGGTCTAATCCCATTACCCG	ATATTTATTTTTTACA
92	231559:231574	for	GAGTCCTTCGGGACTC	TTTTTTTTTATTTAA
93	233194:233222	for	TCCCCTAATAGTTAAAGCTATTAGGGGA	TAATCTTTTTTTTTTT
94	237107:237126	for	GAAGGTGGATTGCCACCTTC	TTTTTTTTTAAGGTT
95	237492:237520	for	CCCATATCCCTCTATAAAGGGGATATGGG	GTTAATATCTTTTTTT
96	243269:243286	for	GATCCTCTTCGGAGGATC	TTTTTTTTTTTTTATA
97	247206:247233	for	GGGTATAAGGATTAATTTCTTTATACCC	TATATTTTTTTATTT
98	247691:247710	for	GGGGGTGTTAATACACCCCC	TTACACATTTACAC
99	249466:249493	for	CCCTATACCTTCTTCGGAGGGTATAGGG	AGTTTATTTTTTTTTT
100	254530:254559	for	CTCCATATCCCTCTACAAAGGGGATATGGGG	TTTATTTTTTATTGTG

^aThe complementary bases forming the hairpin stem are highlighted. T represents dU.

^bTs within 15 downstream nucleotides highlighted.