

MGV-GENOME-0312075\_r;5592;6677  
OHGK01000149.1\_r;5669;7036  
NL\_vir006651\_f;37477;38559  
uvig\_425355\_r;11757;13058  
uvig\_425355\_f;39603;40904  
MGV-GENOME-0272770\_r;499;1920  
MGV-GENOME-0297536\_f;14441;15565  
MGV-GENOME-0310879\_f;320;1423  
MGV-GENOME-0305083\_f;11909;13054  
MGV-GENOME-0212520\_f;17251;18312  
Han\_2018\_ERR1398212\_NODE\_689\_r;4598;5668  
MGV-GENOME-0244886\_f;3097;4260  
MGV-GENOME-0353320\_f;45034;46209  
uvig\_205988\_r;1876;3051  
MGV-GENOME-0371861\_r;58522;59658  
MGV-GENOME-0371861\_f;4661;5845  
SRS415869\_NODE\_11\_r;74172;75341  
Manrique\_Person1\_NODE\_54\_r;53028;54197  
MGV-GENOME-0374715\_f;81873;83135  
MGV-GENOME-0321287\_f;36741;37898  
MGV-GENOME-0289925\_r;1259;2290  
OLKK01000549.1\_f;26052;27305  
uvig\_384903\_f;24457;25497  
ERS698742\_NODE\_4\_f;359;1546  
MGV-GENOME-0279285\_f;36563;37810  
MGV-GENOME-0347346\_f;54338;55741

```
..MSFVRSFNSNSNNVNRNVNTDGSNNNNAYNGNGVVRPAL.....VEHRDQVDRR.....SGGRKQRPHTQRNEYVPVQRRNAEDKHMTPTPRSPTIPRGPGRGYRGVWREVY.....  
.....MRREGYI.....I.....  
.....MRREGYI.....I.....  
.....MKRAK.....  
.....MDPRNQ.....  
.....MKRKG.....  
.....NKKRIG.....  
.....MISFNGVS.....  
.....MKRVR.....  
.....MKRVR.....  
.....QMTLEEFLLFQNFQAQTILNTNVSNT.....RTYAYETVSEHF  
.....MNSKERHEIRYQRRVAA.....RQAKRIA  
.....MSRRKGGRYERRKTR.....REENLR  
.....MTSEERREARYKRRRAR.....RQARLQA  
.....MNEDKILQMFFD.....IGQWT.....KAIEKGV  
.....MTSEERKEARYQRRKAS.....RQRRREE  
.....MNSEQRRRAARRKRRREEK.....RAKAKAE  
.....MKRLS.....  
.....CYTADTCFNQYSFEDCG.....LYVGDGTGKIFISQAKKI.....  
.....MGVVKT.....EYGL.....  
..MG.NRAVNKYLDDIRRTLLAWYRIGREPVFHACYAMRLGLPVKLLIQGCRTRYHFDNRL..LPESGGNLI MTSEERKEARFQRRKAK.....REAKRER
```

	1	10	20	30	40	50	60	70								
MGV-GENOME-0312075_r;5592;6677	..MQKTKVNF	..DTVYEFET	..LYNAYRASRR	GKRW	..KNTVAKVEMNALEAT	AVLQEE	..LSTGTYR	..PGGYREFY	..VE	..P	..KKRL					
OHGK01000149.1_r;5669;7036	..VEQDRKQNF	..DEVCDGF	..NLYKAYRASRR	GKRW	..KNTVAKIELNALEAV	AYLQNE	..LSEGTYK	..PGDYREFY	..VF	..P	..KKRL					
NL_vir006651_f;37477;38559	.....MY	..ERYGFD	..NLHKAFKLARR	GKRW	..KPATARFEVNLL	ENLRLSRE	..LQDKTYE	..LSEYHTFK	..VY	..P	..KKRD					
uvig_425355_r;11757;13058	.....EE	..IIEYSNMSE	AFDSVLR	GTGR	..KRSRGREFLLAHREKI	IAELTAS	..IADGSFR	..LGGYHERE	..IE	..Y	..GKKRI					
uvig_425355_f;39603;40904	.....EE	..IIEYSNMSE	AFDSVLR	GTGR	..KRSRGREFLLAHREKI	IAELTAS	..IADGSFR	..LGGYHERE	..IE	..Y	..GKKRI					
MGV-GENOME-0272770_r;499;1920	.....MGEQSN	..LSIGHSPGDEP	GKTKT	..VNAES	..SNAWYVNMNNGN	VNTNNK	.....	.....	.....	..TNAG	..RVPVSATDKPIYDIPLSSIVH					
MGV-GENOME-0297536_f;14441;15565	.....MQGAEF	..EQVYDFG	..NLYAGFLKARR	GKRRH	..KPSVAKFEANLLEAL	CLLSEM	..LTKTKTYR	..PSDYFVFK	..VY	..P	..KKRI					
MGV-GENOME-0310879_f;320;1423	.....DLY	..QKLISDEN	..LRLAILTVNA	THKWH	..PHH	..RPNKT	..VLRVE	..ADIDGYVEK	..REII	..INGYD	..AAPPRIARRWDKSA	..GKWRD				
MGV-GENOME-0305083_f;11909;13054	.....SLL	..ERYISWEN	..LIDAYHEAAS	EKWY	..RNDVTAFAANLEEN	..LISIQND	..LIWHTYK	..VGRYRQFY	..VH	..P	..KKRI					
MGV-GENOME-0212520_f;17251;18312	.....NIY	..EKITDLN	..NIETAIYRASK	GKGN	..RKSVEKILDSPTY	YAMQVQQA	..LINTKYV	..PNKYVEMK	..IRD	..GAN	..KKRI					
Han_2018_ERR1398212_NODE_689_r;4598;5668	.....YLM	..EKLCTRE	..NALLAIEAVNE	PRKK	..NKTAQWVESTKEAR	ADELCELL	..LRDF	..HPKKPRTFP	..RY	..DSTA	..GKWRD					
MGV-GENOME-0244886_f;3097;4260	.....EQLYIPEEQ	IKD	.....IYNASK	GKSK	..KEQAQIVKANVEHYR	KELDKR	..LKNNTFA	..PKRRHKT	..KI	..QENSC	..KTKRK					
MGV-GENOME-0353320_f;45034;46209	.....IY	..QQIISDEN	..LRLAIQDVNR	GHRRN	..GDY	..SLNKKVMEIE	..EHI	..DEYVVK	..LKKFIED	..LVTGDEHMHKPLQRRK	..WDRNADSGKG	..KWRD				
uvig_205988_r;1876;3051	.....VY	..KEIISDEN	..LRLAIREVNA	GHRRN	..GNH	..SLNKKVIEIE	..NNM	..DEYVEK	..LRAFIQGLVDG	..DEHMHPPPLKRRR	..WDRNADSGKG	..KWRD				
MGV-GENOME-0371861_r;58522;59658	TSRIDTDALI	..RKLVRFND	OTEALRAQER	STLYE	..TFH	..IPKKS	.....	.....	.....	.....	..GGLRR					
MGV-GENOME-0371861_f;4661;5845	..YSESFGRY	..EDVFSYE	HLYOAGKNCK	GVMW	.....	..KNS	..TQSYMSRIT	..TNTASTHDA	..LLRREFR	.....	..SRGFHDFDLIE	..RGKLRH				
SRS415869_NODE_11_r;74172;75341	..RAATVGG	..LDHVF	GYDMMYKAGKKCCN	GVRW	.....	..KNS	..TQRFEMHLFSG	TARRRRL	..LLEKWI	.....	..PGAYVHFTISE	..RGKTRP				
Manrique_Person1_NODE_54_r;53028;54197	..RNAAIGTL	..EEAFS	YRAMFYFGKKCCN	GVRW	.....	..KNS	..TQRYMKDYL	RNAVLSRRD	..LEGRDI	.....	..CRGFIRFDLW	..RGKLRH				
MGV-GENOME-0374715_f;81873;83135	..LKDIRKQDL	..IRLTD	EHTRMAMAYAMRR	GKYE	..ISPHTAQIPKDN	.....	..KAS	..VQRYEMNLLRN	..INNVTKA	..LES	..GENV	..SGGFIVFWLCE	..RGKLRH			
MGV-GENOME-0321287_f;36741;37898	..RLKEYDDF	..DRVKDAN	..NLIAPFKSKS	GVDW	.....	..KAS	..VQRYEMNLLRN	..INNVTKA	..LES	..GENV	..SGGFIVFWLCE	..RGKLRH				
MGV-GENOME-0289925_r;1259;2290	.....MENKF	..TDICTF	EVLYKAYLAAR	GKRS	.....	..RAATAHYEVHLL	ENLVNLY	..ITTKTYR	.....	..PGVFRVFFY	..VY	..P	..KKRL			
OLKK01000549.1_f;26052;27305	..RVKTC	..TL	..ETVADLN	SLCKASQQAAR	GVMW	.....	..KAS	..TQRYMKDYL	RNAVLSRRD	..LEGRDI	.....	..CRGFIRFDLW	..RGKLRH			
uvig_384903_f;24457;25497	.....NLY	..EQIISLD	..NLHLADEKARK	GKLR	.....	..SYGVK	..KHDRNREAN	..LALHES	..LKNKTFV	.....	..NSKYEVFI	..IRD	..PKERL			
ERS698742_NODE_4_f;359;1546	.....KNVY	..HLIYEC	SNLIRAYKAAQ	GKGE	.....	..RTEIS	..KFENILEN	..LDSLYWD	..LKNETYT	.....	..PGEYRIKVI	..YE	..PKERV			
MGV-GENOME-0279285_f;36563;37810	.....MTY	..QELCS	FGTLLWTAYHRARR	CKRG	.....	..KKS	..TAPFEYSAIE	..ELLILSKS	..LQGT	..HQ	.....	..PDPLDAFY	..YE	..P	..KKRL	
MGV-GENOME-0347346_f;54338;55741	..VLEE	..HG	..DY	..YKVIS	RNAISKSAIEAAK	..GVS	Y	.....	..KAS	..VKRYMLRRLTN	..VAATNKK	..TYCED	..I	.....	..HKGFICFGLNE	..RGKLRH

MGV-GENOME-0312075\_r;5592;6677  
OHGK01000149.1\_r;5669;7036  
NL\_vir006651\_f;37477;38559  
uvig\_425355\_r;11757;13058  
uvig\_425355\_f;39603;40904  
MGV-GENOME-0272770\_r;499;1920  
MGV-GENOME-0297536\_f;14441;15565  
MGV-GENOME-0310879\_f;320;1423  
MGV-GENOME-0305083\_f;11909;13054  
MGV-GENOME-0212520\_f;17251;18312  
Han\_2018\_ERR1398212\_NODE\_689\_r;4598;5668  
MGV-GENOME-0244886\_f;3097;4260  
MGV-GENOME-0353320\_f;45034;46209  
uvig\_205988\_r;1876;3051  
MGV-GENOME-0371861\_r;58522;59658  
MGV-GENOME-0371861\_f;4661;5845  
SRS415869\_NODE\_11\_r;74172;75341  
Manrique\_Person1\_NODE\_54\_r;53028;54197  
MGV-GENOME-0374715\_f;81873;83135  
MGV-GENOME-0321287\_f;36741;37898  
MGV-GENOME-0289925\_r;1259;2290  
OLXK01000549.1\_f;26052;27305  
uvig\_384903\_r;24457;25497  
ERS698742\_NODE\_4\_f;359;1546  
MGV-GENOME-0279285\_f;36563;37810  
MGV-GENOME-0347346\_f;54338;55741

```

      80      90      100      110      120      130
...QTNSF.KDKIVQHAFCDFILYDV...TRPFILDNYGQIKGKTHFGLNRRDFFREYY...
...QTNSF.KDKIVQHAFCDFILYDV...SRPFILDNYGQIKGKTHFGLNRRSGFMREYY...
...VMSNSF.RDKVVQHAFCDFILYDV...RKNFLYDNYASQVKGKTDFFLNRRDGFHKKFY...
...LQILSM.KDRIAQVFAIMNVVDRHL...QKRYIRTGASIKRRGTHDLNMCIRTDLQK...
...LQILSM.KDRIAQVFAIMNVVDRHL...QKRYIRTGASIKRRGTHDLNMCIRTDLQK...
AFDVCCKNKRNTDDCIEFSFEYDIDLAVAVDAIRYGRYEPDYSKCFIRKKPVLRE...
...VMTNAP.KDKVVQHAFCDFILYDV...SKAFIRDNYASQSGRGTHDGLYLRLEEFMRSYFTRKAN
...ISEPRLWPQDQYHHAIVQLLEPVL...MRGMDKFCGSIKGRGIHYGVKAIKKWMRT...
...VMALGF.RDKVVQHAIVQLLEPVL...DNGMIYHSGYGRVGRGTTIRADRLQ...YW...
...YKPRFYQDQYHHAIVQLLEPVL...MRGMDKFCGSIKGRGIMRGMRIKKLVLQ...
...INEPALWPQDQYHHAIVQLLEPVL...MRGMDKFCGSIKGRGIMRGMRIKKLVLQ...
...IVKPOQMYEQMAHHSVMRVFVPIA...MRGMYHYVYGSIPGRGVHRGKRTVERWIRE...
...INEPLLWPQDQYHHAIVQLLEPVL...KRSMDKYCIAVPGRGNSYGVKVLKKWMKN...
...INEPLLWPQDQYHHAIVQLLEPVL...KRSMDKYCIAVPGRGNSYGVKVLKKWMKN...
...IDAP...KPELMNAL...RNLKTIFFEDFHALYHTSAFAYVKNRCTVDAVK...
...IRSVHI.SERVVQRCLCDNILEPVL...SHSFVFDNAASLKGKGVDFAMDRDLDRHLHRY...
...IDAPRI.QDRQVHKVYTKKVLPLLY...REPMIYNNGASLEKGKFEFSKRMLEDLHWHF...
...IDAPRI.TDRQIHKTECNMILPLLY...TPHMIYDNGASRRGMGLHFAYHRLLEQQDLHWHF...
...VYVNEP.MDRVVGLGIANDLLEFELMPEMLHSSCKSYQTGLIGCGK...VVTIEVSHRMT...
...IKSVHI.WERVVQRCLCDNILEPVL...QTGLIYDNGASMEKGKIHFAFNRLDAHLHRY...
...VQAPAF.VDKVVQHAVVDNILEPVL...TNSFILDNYASQKNKGKLFHGLDRLKGLFTDYW...
...ISAVHF.PERVVQRCLCDNILEPVL...VPTLIAANSANIKGRGTDYALKLKRHLADHW...
...IYRLPYPPDRILVHAIMNILEPVL...VSLFTEDTYSICKDRGIIKKAADVKKALKE...
...IMIAFPYPPDRILVHAIMNILEPVL...TNFFIANITYACIKGRGIHKCMEDVHTALII...
...IQAPT.PDKVVQHALTDYIVYDEL...ARSFTLNTYAAQYKGKTHYGLEMLKRRMRTYFLRRKGA
...MSVHF.SERVVQRCLCDNILEPVL...TRSLIHDNGASQKKGKTSFAMKRLVTHLRRHY...

```

MGV-GENOME-0312075\_r;5592;6677  
OHGK01000149.1\_r;5669;7036  
NL\_vir006651\_f;37477;38559  
uvig\_425355\_r;11757;13058  
uvig\_425355\_f;39603;40904  
MGV-GENOME-0272770\_r;499;1920  
MGV-GENOME-0297536\_f;14441;15565  
MGV-GENOME-0310879\_f;320;1423  
MGV-GENOME-0305083\_f;11909;13054  
MGV-GENOME-0212520\_f;17251;18312  
Han\_2018\_ERR1398212\_NODE\_689\_r;4598;5668  
MGV-GENOME-0244886\_f;3097;4260  
MGV-GENOME-0353320\_f;45034;46209  
uvig\_205988\_r;1876;3051  
MGV-GENOME-0371861\_r;58522;59658  
MGV-GENOME-0371861\_f;4661;5845  
SRS415869\_NODE\_11\_r;74172;75341  
Manrique\_Person1\_NODE\_54\_r;53028;54197  
MGV-GENOME-0374715\_f;81873;83135  
MGV-GENOME-0321287\_f;36741;37898  
MGV-GENOME-0289925\_r;1259;2290  
OLXK01000549.1\_f;26052;27305  
uvig\_384903\_r;24457;25497  
ERS698742\_NODE\_4\_f;359;1546  
MGV-GENOME-0279285\_f;36563;37810  
MGV-GENOME-0347346\_f;54338;55741

```

      140      150      160      170      180
...RRHGS...ADGMVLKADVHHYFASIRHDIKQDVRELLDER.SLALTDLIIIDST...PG...
...RRHGS...AYGMVLKADVHHYFSSIRHDIKQDVRELLHDER.SLALSDAIIIDST...PG...
...RQHGL...EGMVLKCDIRKIFYISIPHEYIKRILEPYVPED.VRWLLWYIIIDST...A...
...DPEGTL...YAYKFDIRRFYDNARQDFVMWCFFRVFKDER.LLVLLERFVKLL...
...DPEGTL...YAYKFDIRRFYDNARQDFVMWCFFRVFKDER.LLVLLERFVKLL...
...YTKDA...YIFKGDFFKSFFMSMSKSLWE...MIDFIRDN...KGDDIECLLYILRTVIFHQPYKCYRKSPHLHWDLPDKSLF
AERERRAAGL.PPPGPEEVRRHYS DGMVLKCDIRKIFYISIQHEPIKQMTRKYIKDPD.ILWLVDLIVDST...E...
...DPKGT...YAAEIDIRHHFYDSLTIETVMARLRLRVKDRR.MLDVCCERLM...IPPS
CTLVDRKPKK.W...YLLKIDVSKIFYRVDHRLVGLILRRKFPNEDGYLWMETIINC...HTPFGLPPEGKSADE...
...DRKHGT...YCLKIDVSKIFYPSIDKQILKNKFRRLVKDKD.TLNLIDLIIIDSS...
...DQKGT...YAAEIDIRKIFYPSISPREVMRFLRRKIKDEA.FLGLIWRII...
...DSRNC...YIYKIDIRHHFESVPHRRLLKKALKRKIRDER.LLKKLFIIIDSH...
...DPVGT...YCAECDIRHHCFVEVDPPYVIAHALKRLFKDER.TLWLCDAIM...
...DVEGT...YCCEDIRHHCFEELDPPYVIEALKRVFKDTE.TLWLCDAIM...
...RHQKNNS.K...WFGKIDLDHDFFGSTTLDYVIMKFMVFPFSE...IVKFPNGEA...EL...RKALD
...RKFGV...EGVESGVLVTGDFSDFFNSAPHSIIYREAERRIHDD.VRRITACQFMEDF...G...
...RRYG...RDGNVILDFKQFFPSVSHHEEIKFRHEKLLLNPD.IRKIGDDVVNTV...SG...
...RRYG...RAGAVLLIDLKFFFSAPHATIIYQRHQRLILDPS.LRGLADSLVASS...PCPTP...
...ETGNSGC...LGWKSIDLSKYFDVSVPIRIDE.AFDKVEAKHGQSALIDVLRKYY...HSDLYFDEDN...R
...RRNGF...SNDGFIYLVDFSKIFYDNILHEPVYQDLQKNFTDER.IINLAAOLII...ADRK
...NKNHT...ADGVIYLLCDVRHHFFASIDHDLKEKLLK...DL.EPALYDILCVYV...DC...
...RRHGT...REGVILLLGDFSDYFARIAHQPVKDDQVASALLDP.VVLEHRLIDA...QG...
...DPEHT.T...YCLKMDIKKFYPSIDHDLKTTILRRKKIKDKD.LLWLVDVILDS...
...DRKGT.R...YCLKMDIKKFYPSIDHDLKTTILRRKKIKDKD.LLWLVDVILDS...
DEVARKAAGL.PHRPMEWDYAEGMVVKIDIRHHFFQSIDHQLKKAALPRFPDPD.IRALMWRVILDAV...
...KHHG...TEGVLLFDENKNYFNIDHDLAKQIIRRAFPDDK.IVWLTNRILDSYEHYLLKMAIKKGENPDT...VEHK

```

A



369

MGV-GENOME-0312075\_r;5592;6677  
OHGK01000149\_1\_r;5669;7036  
NL\_vir006651\_f;37477;38559  
uvig\_425355\_r;11757;13058  
uvig\_425355\_f;39603;40904  
MGV-GENOME-0272770\_r;499;1920  
MGV-GENOME-0297536\_f;14441;15565  
MGV-GENOME-0310879\_f;320;1423  
MGV-GENOME-0305083\_f;11909;13054  
MGV-GENOME-0212520\_f;17251;18312  
Han\_2018\_ERR1398212\_NODE\_689\_r;4598;5668  
MGV-GENOME-0244886\_f;3097;4260  
MGV-GENOME-0353320\_f;45034;46209  
uvig\_205988\_r;1876;3051  
MGV-GENOME-0371861\_r;58522;59658  
MGV-GENOME-0371861\_f;4661;5845  
SRS415869\_NODE\_11\_r;74172;75341  
Manrique\_Person1\_NODE\_54\_r;53028;54197  
MGV-GENOME-0374715\_f;81873;83135  
MGV-GENOME-0321287\_f;36741;37898  
MGV-GENOME-0289925\_r;1259;2290  
OLXK01000549\_1\_f;26052;27305  
uvig\_384903\_f;24457;25497  
ERS698742\_NODE\_4\_f;359;1546  
MGV-GENOME-0279285\_f;36563;37810  
MGV-GENOME-0347346\_f;54338;55741

.....NQRPAR.....  
.....VESADRGESQVR.....  
.....GVVV.....SIRE..LVNLPIVVK.DFETGI....KTEQGEDRCIVAIEVNGEAKKFFTNSEEMKNILAQVKEMPDGFPFETTIIKTETFGKGRTKYV.FT  
.....GVVV.....SIRE..LVNLPIVVK.DFETGI....KTEQGEDRCIVAIEVNGEAKKFFTNSEEMKNILAQVKEMPDGFPFETTIIKTETFGKGRTKYV.FT  
.....VFVLKNEYNFKKQ.....LKKQIRKGNAKKYLTPEIG..  
.....RE.....HTRKEQARWNM.....CTEPLKSTA..  
.....IELPQEVITQWELNQF.....ESKKSRKH.....NKT..  
.....RN.....ASRKERILC.....SST..  
.....RKENERN.....GMVRSRKHC.....RRETKRAGYNKQPAPCI..  
.....RR.....ECRRLQRLY.....PPYRAA..  
.....RK.....ECRRLQALY.....PPYQAA..  
.....LIK.....DDLRL.....LARAC..  
.....VS.....EERQ.....SELKTAASA..  
.....IE.....NFRE.....RGKKDEICC.....A..  
.....GLK.....C.....SRAYNTLVASL..  
LGYVVRITKPDGCIVRGRGRNVKA.NRDKTDRD....IPGYLTV.....GCMRNALLT..  
.....VPA.....GEWR.....  
.....AIQ.....TWRRHGWEHAPTLAT...EGATHDG.TGNRRG..D.....QRLQTAA..  
.....FNNCLN.....DGSKCFDY.....KKQVYDTLRF..  
.....IPAARPGPDERR....TKRYGTVTFQSGKRQPGEAGGKQQ..A.....RAAG..  
.....IE.....NWQR...EEVPLYG.....HQVHQAGQ.....