

Poxviruses capture host genes by LINE-1 retrotransposition

Sarah M. Fixsen¹, Kelsey R. Cone¹, Stephen A. Goldstein¹, Thomas A. Sasani¹, Aaron R. Quinlan¹, Stefan Rothenburg², & Nels C. Elde^{1*}

¹ Department of Human Genetics, University of Utah, Salt Lake City, USA

² Department of Microbiology, University of California, Davis, USA

*Correspondence to: nelde@genetics.utah.edu

Abstract

Horizontal gene transfer (HGT) provides a major source of genetic variation. Many viruses, including poxviruses, encode genes with crucial functions directly gained by gene transfer from hosts. The mechanism of transfer to poxvirus genomes is unknown. Using genome analysis and experimental screens of infected cells, we discovered a central role for Long Interspersed Nuclear Element-1 (LINE-1) retrotransposition in HGT to virus genomes. The process recapitulates processed pseudogene generation, but with host messenger RNA directed into virus genomes. Intriguingly, hallmark features of retrotransposition appear to favor virus adaption through rapid duplication of captured host genes on arrival. Our study reveals a previously unrecognized conduit of genetic traffic with fundamental implications for the evolution of many virus classes and their hosts.

Summary

Active selfish genetic elements in infected cells aid virus adaptation by catalyzing transfer of host genes to virus genomes.

Introduction

Vertical inheritance of genetic material from parent to offspring can be augmented by horizontal gene transfer (HGT) (Keeling and Palmer, 2008). HGT can shortcut gradual mutational processes by facilitating adaptive leaps through single mutational events where genes can be transferred across kingdoms of life (Keeling and Palmer, 2008). HGT plays a central role in the diversification of many infectious microbes. This process is particularly well-described in bacteria, where HGT contributes to the emerging crisis of multi-antibiotic resistant pathogens (Wijayawardena et al., 2013). Many classes of viruses also acquire host genes by HGT (Caprari et al., 2015). A conspicuous example is Rous Sarcoma virus, in which recognition of the horizontal transfer of the host *c-Src* gene into the virus genome led to the discovery of oncogenes (Swanstrom et al., 1983). It has become increasingly clear that acquisition of host genes is a common mechanism by which viruses of many classes gain adaptive advantages to propagate and evolve. Despite the prevalence of HGT in diverse viruses and its importance in virus evolution, the mechanisms by which genes are transferred from host to virus genomes are not understood.

To investigate mechanisms of transfer we examined poxviruses, which encode a plethora of host genes gained by HGT. Based on phylogenetic analysis, more than 25% of poxvirus genes were acquired by HGT from hosts (Hughes and Friedman, 2005), although similarities not evident in sequence comparisons but revealed by shared protein structures suggest that a higher proportion of genes were acquired by horizontal transfer (Bahar et al., 2011). Many captured genes adapt to act as inhibitors of host immune defenses to favor virus replication (Elde and Malik, 2009). Some act as host range factors because their deletion leads to restriction of infectible cell types or species, implying that new acquisitions may aid in host jumps (Bratke et al., 2013; Haller et al., 2014). While the most notorious poxvirus, variola virus which caused smallpox, has been eradicated, other extant poxvirus strains are capable of infecting humans and pose the risk of new pandemics (Sklenovska and Van Ranst, 2018).

Results

We hypothesized that examination of genes recently acquired by HGT might provide clues for revealing mechanisms of gene transfer. Virus homologs of the host Golgi Anti-Apoptotic Protein (GAAP; also called transmembrane Bax inhibitor motif protein family 4 or TMBIM4), found in several orthopoxviruses, are ~75% identical to human GAAP (Gubser et al., 2007; Saraiva et al., 2013), suggesting a recent acquisition and/or highly conserved function favoring virus replication. Strikingly, we observed nearly identical 21 base pair sequences flanking the open reading frame of virus GAAP (vGAAP), indicative of a target site duplication (TSD), consistent with gene capture resulting from LINE-1 mediated retrotransposition of host mRNA (Figs. 1A and S1). The vGAAP gene is situated near one of two identical inverted terminal repeat (ITR) regions of the virus genome, and on opposite ends of the genome in various viruses, suggesting that it was originally transferred into the ITR. A single copy of the putative 21bp TSD at an analogous position near the other ITR in the cowpox genome suggests a pre-integration sequence or pseudo empty site (Fig. 1A). Together with a published account of a non-autonomous Short Interspersed Nuclear Element (SINE) in the genome of taterapox virus (Piskurek and Okada, 2007), these observations suggest that host genes can be retrotransposed into poxvirus genomes by LINE-1 activity in a process akin to the generation of processed pseudogenes (Esnault et al., 2000).

Phylogenetic analysis revealed a single origin for the horizontal transfer of GAAP into poxvirus genomes. In addition, at least two other classes of DNA viruses, herpesviruses and iridoviruses, independently acquired GAAP by HGT (Fig. 1B) (Bellehumeur et al., 2016; de Groof et al., 2015). Another study revealed that the S1 gene of Squirrel Monkey cytomegalovirus encodes a protein more than 95% identical to squirrel monkey Signaling lymphocytic activation molecule family 6 (SLAMF6), followed by portions of the 3' untranslated region (UTR) of the host mRNA (Perez-Carmona et al., 2015). These findings suggest that many DNA viruses acquire host genes by retrotransposition. However, the majority of host to virus gene transfers lack features of retrotransposition, which can quickly degrade through mutation, motivating our development of an experimental strategy for capturing HGT events in real time.

FIGURE 1

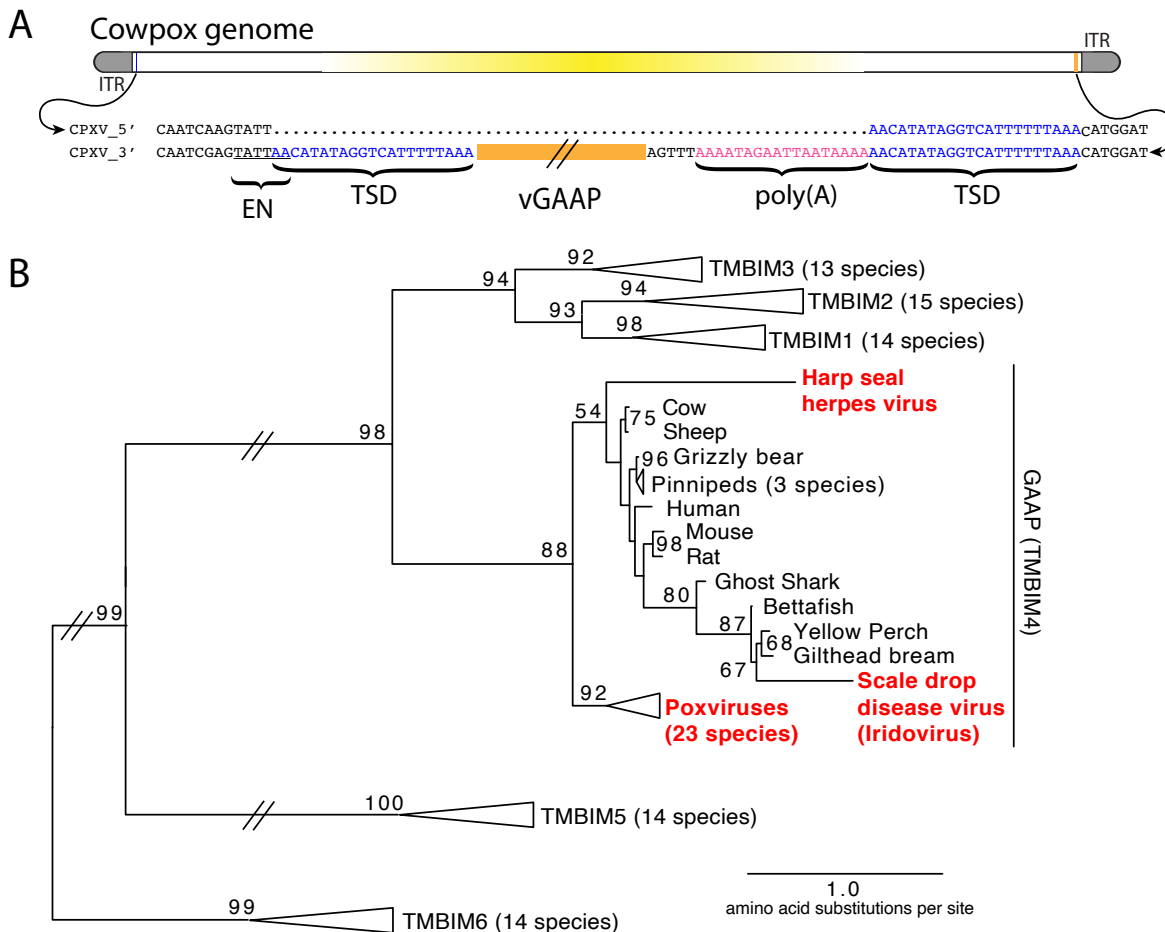


Fig. 1. Retrotransposon-mediated transfer of GAAP into poxvirus genomes. (A) The poxvirus gene vGAAP (orange) is encoded near the inverted terminal repeat (ITR; gray) of cowpox genomes. Signatures of retrotransposition include TSDs (blue), a LINE-1-like endonuclease site (EN; underlined), and a partially degraded poly(A) tail (pink). Some cowpox genomes (CPXV_5'; top line) contain a pseudo-empty site, with a single copy of the TSD sequence. **(B)** Phylogeny of TMBIM proteins and virus-encoded TMBIM4s (red). Bootstrap values >50 are indicated.

We developed a selection scheme using vaccinia virus (Copenhagen strain), which encodes nearly 200 genes in a 191,737 bp genome (See methods). We focused on K3L, a gene that was presumably acquired from a host by an ancient poxvirus based on sequence identity with human eukaryotic initiation factor 2 α (eIF2 α ; 27.6% identical) (Beattie et al., 1991; Elde et al., 2009; Elde and Malik, 2009). K3L blocks the antiviral Protein Kinase R (PKR) pathway from arresting protein translation during viral infections by phosphorylating eIF2 α (Fig. S2) (Beattie et al., 1991). Using a mutant virus deleted for K3L and another PKR inhibitor, E3L (Brennan et al.,

2014), we infected a rabbit kidney (RK13) cell line expressing complementing chromosome-integrated copies of K3L fused to the gene encoding mCherry fluorescent protein (RK13-K3L) (Fig. 2A and see Methods for details). Viruses grown in RK13-K3L cells were transferred to a non-complementing baby hamster kidney (BHK) cell line to select for viruses that “reclaimed” the mCherry-K3L gene and/or somehow adapted in the absence of K3L to block host cell PKR and replicate.

We screened ~500 million viruses over many experiments and recovered ten virus isolates from mCherry positive infected cell clones. In the restrictive BHK cells, viruses encoding K3L replicate 100-fold better than the parent virus deleted for K3L (Fig. 2B), providing strong selective pressure to screen for adapted viruses. After plaque purification of mCherry isolates, we infected BHK cells with each recovered virus isolate and observed titers similar to viruses encoding K3L (Fig. 2C). By sequencing the genome of each recovered virus isolate, we found that the mCherry-K3L fusion gene was integrated into the virus genome at a unique position in every isolate (Figs. 2D and S3), providing the opportunity to compare each independent integration event.

All integrations were consistent with a mechanism of LINE-1 retrotransposition-mediated HGT. The sequence of each transferred gene revealed hallmark features of mRNA gene capture, including 5’ and 3’ UTRs, precise intron splicing, and poly(A) tails (Fig. 3). For seven of ten isolates, TSDs flanked mCherry-K3L insertions and contained cut site sequences matching the consensus (TTTT/AA) of LINE-1 endonuclease (Fig. 3C). Although there were no TSDs flanking mCherry-K3L in isolates 5, 9, and 10, these integrations might result from a endonuclease-independent LINE-1 mechanism at DNA breaks (Morrish et al., 2002). Because our experiments lack any exogenous reverse transcriptase, we tested for endogenous activity in the RK13 cells and detected robust LINE-1 activity (Fig. S4). These results suggest that LINE-1 mediated retrotransposition may be a common mechanism of HGT into poxvirus genomes.

FIGURE 2

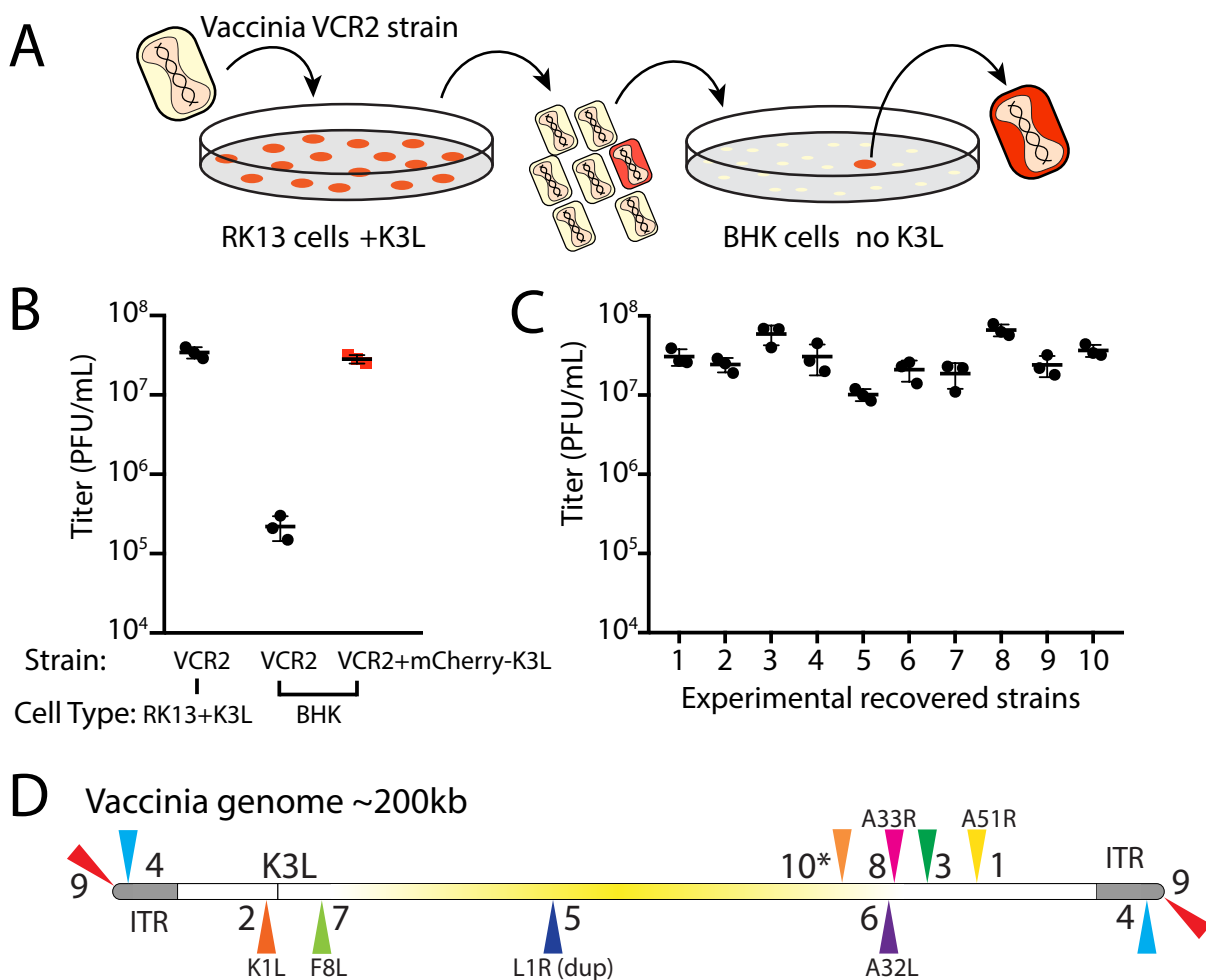


Fig. 2. Experimental capture of K3L by horizontal gene transfer. (A) Viruses lacking K3L (VCR2) replicated in RK13 cells expressing mCherry-K3L and were transferred to BHK cells lacking K3L to select for capture of mCherry-K3L. (B) Replication of VCR2 in RK13-K3L cells and BHK cells compared to VCR2+mCherry-K3L (see Methods). (C) Replication of recovered isolates in BHK cells. Three biological replicates of each strain/isolate are shown in (B) and (C). (D) Vaccinia genome illustrating K3L integrations in recovered isolates indicated by colored triangles. Endogenous K3L location is shown. The central region of the genome is highlighted in yellow. Triangles above and below the genome denote positive and negative sense orientation respectively. Virus genes interrupted by K3L integrations are indicated. The asterisk denotes a genomic rearrangement in isolate 10 (Fig S3).

Poxvirus genomes are arranged with the essential genes, shared by all poxviruses, near the center of the genome. Clade- or species-specific genes, which more commonly exhibit evidence of HGT, are enriched near the ends of the genome (Lefkowitz et al., 2006; McLysaght et al., 2003). This pattern may reflect that gene transfer into the central region of genome is more likely to

interrupt genes vital for virus replication (Lefkowitz et al., 2006). Consistent with this, all but one recovered isolate integrated K3L outside a roughly defined central region of the genome containing 90 core chordopoxvirus genes (Lefkowitz et al., 2006) (Fig. 2C). In the case of isolate 5, the K3L integration involved a duplication of two essential genes, G9R and L1R (Fig. S3). Because gene duplications can appear frequently in poxvirus genomes (Hughes and Friedman, 2005), this activity might enable the persistence of horizontally transferred genes by preserving existing ones, consistent with a process described in a contemporaneous study (Rahman *et al.*, 2020).

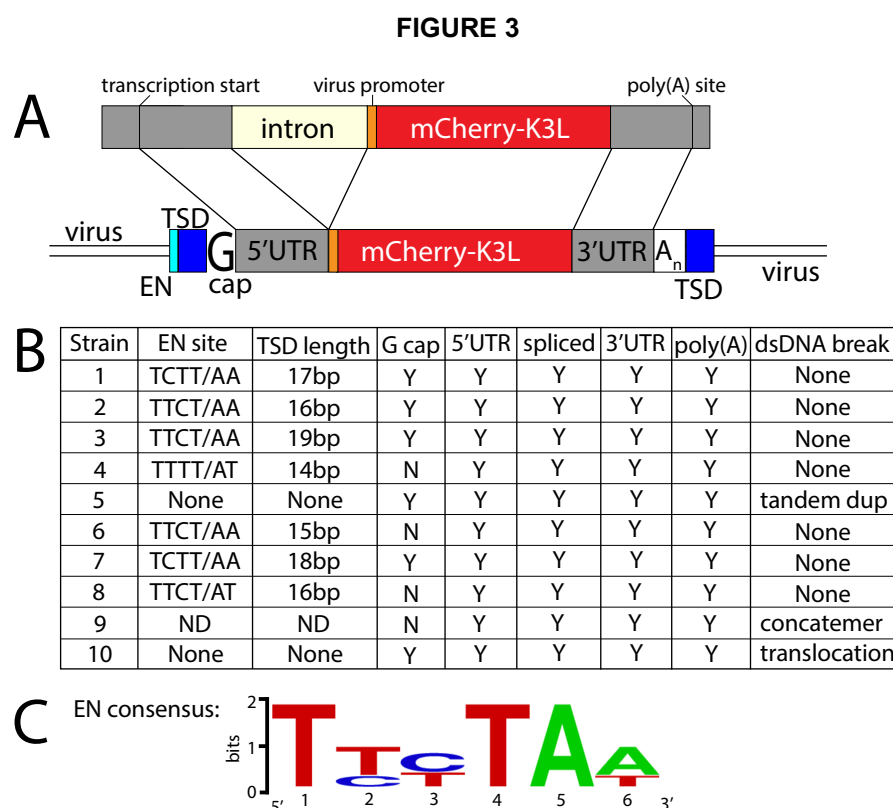


Fig. 3. All virus isolates exhibit signatures of LINE-1 mediated retrotransposition. (A) Schematic of construct integrated in RK13-K3L cells (top) compared to K3L integrations in recovered viruses (bottom). Shared features include spliced introns, 5' and 3' untranslated regions (UTRs), and poly(A) tails (A_n). Seven isolates have target site duplications (TSDs, blue) with LINE-1 endonuclease cut sites (EN, cyan). Six isolates encode a guanine (G cap) adjacent to the 5'UTR, indicative of 7-methylguanylate mRNA capping. (B) Table highlights details of each isolate. (C) Logos plot of endonuclease cut site consensus sequence.

Previous studies found that duplication of poxvirus genes, followed by rapid gene copy amplification, can boost virus replication (Brennan et al., 2014; Elde et al., 2012). In our

analysis, we discovered two isolates where K3L had undergone polymorphic copy number increases (isolates 4 and 5; Fig. 4A). Homologous recombination, including unequal crossovers leading to duplications, requires as little as 16bp of homology in poxviruses (Yao and Evans, 2001), suggesting that LINE-1 induced TSDs might facilitate gene amplification. Genes transferred to within the ITR are duplicated on the other end by replication-based mechanisms (McFadden and Dales, 1979), as we observed in isolates 4 and 9 (Fig. 2C). Genes integrated into the near the termini of the genome by HGT, like isolate 4, may be especially prone to duplication due to stretches of short, high copy tandem repeats found in and near the inverted terminal repeat regions (Fig. 4A). HGT therefore likely occurs evenly across poxvirus genomes, but with an enrichment of transferred genes persisting near the ends, where insults to essential genes are less probable and duplication-based adaptation is favored (Fig. S5). Consistent with this idea, gene families found only in specific lineages of poxviruses, indicating more recent acquisition, are more likely to be present in multiple copies than more host genes acquired in (Hughes and Friedman, 2005).

While our experimental system revealed LINE-1 driven mechanisms of HGT, the scheme involved selection for viruses regaining an existing virus gene. To better model virus adaptation following acquisition of a host gene, we engineered viruses with the human eIF2 α gene in place of K3L (Fig. S6; see Methods). We performed four serial infections of viruses expressing eIF2 α or an eIF2 α variant lacking the PKR phosphorylation site (S51A) in RK13 cells. While the recombinant viruses replicate weakly, we observed a nearly 10-fold increase in virus replication for both eIF2 α -encoding viruses after passaging (Fig. 4B). By sequencing the eIF2 α region in evolved virus strains, we found independent gene copy increases of eIF2 α in both viruses (Fig. 4C). These results support a critical role for gene copy increases following HGT in poxvirus evolution, as also observed for K3L and other genes (Brennan et al., 2014; Elde et al., 2012).

Discussion

Horizontal gene transfer is an important process facilitating virus evolution (Caprari et al., 2015; Deeg et al., 2018; Hughes and Friedman, 2005; Koonin and Yutin, 2019; McLysaght et al., 2003). Our discovery of LINE-1 retrotransposition-mediated HGT from host to poxvirus

genomes has several notable implications. In addition to self-propagation in “host” genomes, retrotransposition to virus genomes could greatly extend the range of LINE-1 or related elements that subsequently mobilize to a newly infected virus host (Piskurek and Okada, 2007). The extensive host range of some poxviruses, for example between rodent and primate species, suggests the possibility of cross host lineage transfer of LINE-1s, non-autonomous Alus, and host transcripts between mammals. A process involving the frequent appearance of DNA transposons in some insect viruses suggests a parallel mechanism for the spread of transposable elements between species (Gilbert et al., 2016). In cases of retrotransposition of host transcripts, the resulting processed pseudogene products are enriched for highly expressed genes, including ribosomal genes (Zhang et al., 2004). The over representation of certain transcripts might provide clues for understanding and predicting how viruses might adapt with repurposed genes mimicking host functions (Elde and Malik, 2009). Currently recognized cases of HGT involving oncogenes and regulators of translation, like eIF2 α , provide cases of transcripts exhibiting high gene expression and cellular functions easily repurposed to promote virus replication.

Repetitive target site duplications generated by LINE-1 activity may also enhance poxvirus adaptation by facilitating rapid gene copy number amplification of horizontally transferred genes on arrival. Increases in gene copy number augment the probability of beneficial mutations appearing, allowing for swift diversification of newly acquired genes (Elde et al., 2012). Thus, form follows function as the repetitive termini of poxvirus genomes further enable the emergence and persistence of acquired genes predominately at the ends of the genome (Fig. 4D). As beneficial virus gene variants endure, they end up closer to the center of the genome as subsequently captured genes appear near the ends.

Similar mechanisms of virus adaptation involving HGT may be shared by other diverse classes of nucleocytoplasmic DNA viruses (NCLDV), in addition to poxviruses (Deeg et al., 2018; Koonin and Yutin, 2019). Many giant viruses, which can exceed a megabase in genome size and infect diverse host species, encode a variety of genes acquired by HGT, including ribosomal genes. Given a primary role for LINE-1 elements in facilitating gene transfer, we speculate that other classes of transposable elements may spur NCDLV evolution by mobilizing host genes into virus genomes across diverse ecosystems.

FIGURE 4

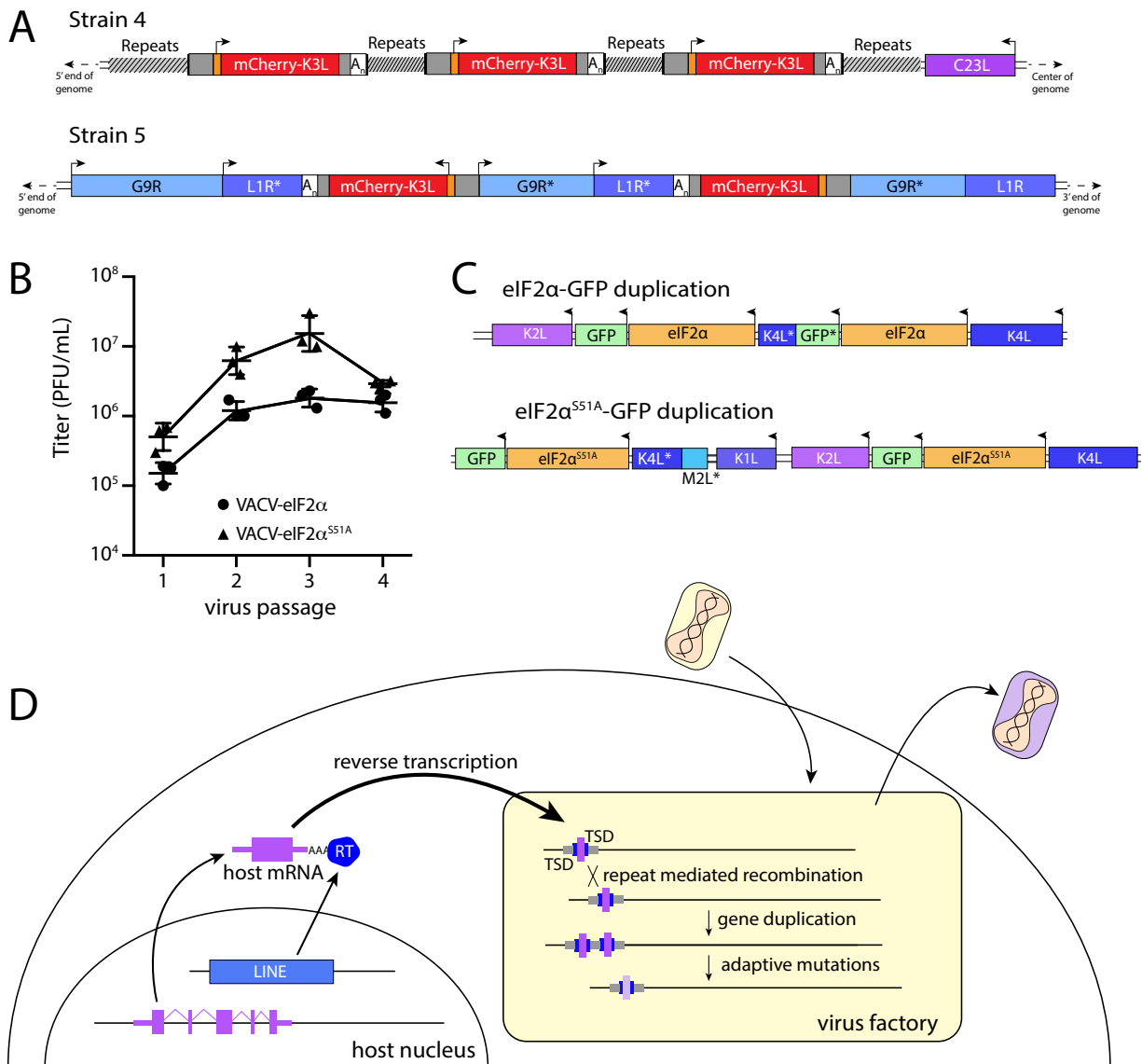


Fig. 4: Homology driven gene duplication of K3L and eIF2 α . (A) Schematics of mCherry-K3L duplications in isolates 4 and 5. (B) Virus titers after serial infections of eIF2 α viruses in HeLa cells (see Methods). (C) Schematics of eIF2 α duplications following virus passaging. Asterisks denote gene truncations (D) Model diagram highlighting how host genes (purple) are reverse transposed by LINE-1 reverse transcriptase (RT, blue), into virus genomes (yellow box). Following horizontal transfer, gene duplication is facilitated by unequal crossover recombination between TSDs (blue) or flanking repeat sequence (gray).

Materials and methods

Phylogenetic analysis and sequence comparison of vGAAP

TMBIM sequences for various species were downloaded from NCBI (Supplemental data) and aligned using the COBALT Multiple Alignment Tool. Amino acid Phylip alignments (see supplemental materials) were analyzed by PhyML (<http://www.atgc-montpellier.fr/phyml/>) for tree building with 100 bootstraps (Guindon et al., 2010).

Cell culture

BHK and HeLa cells were cultured in Dulbecco minimum essential medium (DMEM; HyClone) supplemented with 10% FBS, 2 mM L-glutamine and 100 µg/mL of pen/strep. RK13 cells were cultured in MEM-alpha supplemented with 10% fetal bovine serum (FBS; Gibco), 2 mM L-glutamine (HyClone) and 100 µg/mL of penicillin-streptomycin (pen/strep; HyClone). RK13 cells expressing SLP-mCherry-K3L (RK13-K3L) were generated by transfecting 3x10⁵ RK13-E3K3 cells (Brennan et al., 2014) with 2.5ug piggybac vector (see supplemental sequences) and 0.5 ug transposase vector (gift from Ed Grow, University of Utah) with 9uL Fugene-HD. Cells with the piggybac construct integrated were selected by the presence of 1-2 ug/mL puromycin, and populations were monitored for mCherry expression. After two weeks of selection, remaining polyclonal cells were propagated in MEM-alpha media supplemented with 10% FBS, 2 mM L-glutamine and 100 µg/mL of pen/strep. All cultures were maintained and infections performed at 37°C in a humidified 5% CO₂ incubator.

Virus strains

VC-2 refers to the Vaccinia virus-Copenhagen strain (Goebel et al., 1990) and VCR2 is a genetically modified isolate of VC-2 deleted for E3L and K3L (Brennan et al., 2014). To generate VCR2+mCherry-K3L, we cloned pBlue_165_mCherry-K3L (see supplemental sequences) into the MCS of pBluescriptIIS (-) (Addgene). This plasmid was transfected into BHK cells, which were then infected with VCR2 virus. Recombinant viruses were plaque purified in BHK cells and checked for purity by PCR amplifying across the mCherry-K3L insertion site. The amplicon was gel purified and sequenced.

To make eIF2 α viruses, sequences flanking K3L from VACV were amplified from VC-2 viral DNA: 680bp of 5' homologous sequence was amplified with the primers K2Lflank_F (5'-CTTCTTATC GATTTTTTATACCGAACATAAAAATAAGGTTAATTA) and K2Lflank_R (5'-CTTCTTCATATGG TGATTGTATTCCTTGCAATTTAG), and 1024bp of 3' homologous sequence (including the native K3L promoter) was amplified with the primers K4Lflank_F (5'-CGTCGTGCGGCCGCCTTGTTAACGGGCTCGTAAATT) and K4Lflank_R (5'-CGA GCGGAGCTCGTACGATACATAGATATTACAAATATCCTAG). A VACV synthetic early/late promoter (SLP)(Chakrabarti et al., 1997) was created by annealing primers SLP_F (5' TCGACAATTGGATCAGCTTTTTTTTTTTTTTTTTTTTGGCATATAAATA AGAAGCTTCCCGGGTCTAGAC) and SLP_R (5'-AGCTCAGATCTGGGCCCTTCG AAGAATAAATATACGGTTTTTTTTTTTTTTTTTTTCGACTAGGTTAAC). EGFP was amplified from pN1-EGFP (Addgene) using primers EGFP_F (5'-GGAGGACTCGAGATGGTGAGCAAGGGCGA) and EGFP_R (5'-GGAGGTATCGA TTTACTTGTACAGCTCGTCCATGC). Each PCR product was digested with restriction enzymes (New England Biolabs), gel purified (Zymo Research), and sequentially cloned into pB.2 as follows: 5' flank with ClaI and NdeI, 3' flank with NotI and SacI, SLP with SalI and XhoI, and EGFP with XhoI and ClaI. The resulting plasmid contained EGFP following SLP, between the two K3L flanking sequences (pB.2-EGFP). The K3L open reading frame was amplified from VC-2 viral DNA using primers K3L_F (5'-GTTGTAG GATCCATGCTTGCATTTTGTATTTCGTTGC) and K3L_R (5'-GTTCTTGTCTGACTT ATTGATGTCTACACATCCTTTTG). The eIF2 α gene was amplified from human cDNA using primers eIF2 α _F (5'-GATGTAGGATCCATGCCGGGTCTAAGTTGTAGAT) and eIF2 α _R (5'-CTACTTGTCTGACTTAATCTTCAGCTTTGGCTTCCAT). The resulting PCR products were cloned into pB.2-EGFP with BamHI and SalI, placing K3L or eIF2 α immediately following the native K3L promoter, and upstream of SLP-EGFP to create pB.2-K3L and pB.2-eIF2 α , respectively. Site-directed mutagenesis was performed on pB.2-eIF2 α or pB.2-eIF2 α - Δ C using primers eIF2 α _S51A_F (5'-GATAC GCCTTCTGGCTAATTCATAAGAAGAATCATGCCTTC) and eIF2 α _S51A_R (5'-GAAGGCATGATTCTTCTTAGTGAATTAGCCAGAAGGCGTATC) to generate pB.2-eIF2 α -S51A.

Recombinant eIF2 α -encoding viruses were constructed by replacing the K3L gene using homologous recombination. RK13-E3K3 cells were infected with VCR2 (MOI = 1.0) and transfected at 1-hour post-infection with pB.2-EGFP, pB.2-K3L, pB.2-eIF2 α , or pB.2-eIF2 α -S51A plasmids by use of FuGENE6 (Promega) according to the manufacturer's protocol. Infected cells were collected at 48 hours post infection, and viruses were released by one freeze-thaw cycle followed by sonication. Resulting viruses were plaque purified in RK13-E3K3 cells four times, selecting for recombinants expressing EGFP. Final virus clones were verified by PCR and sequencing of viral DNA across the K2L-K4L region of the genome.

Experimental screen for horizontal gene transfer events

Confluent 150mm dishes of RK13-K3L cells were infected with VCR2 at a MOI of 0.1. Cell-associated virus was collected after 48 hours of infection, and released from the cell by one freeze-thaw cycle followed by sonication as described (Isaacs, 2004). Confluent 150mm dishes of BHK cells were then infected with $\sim 10^6$ PFUs of this virus stock for 7 days, during which they were monitored for mCherry expression. 500 plates were screened, which accounts for an estimated 500 million viruses. For mCherry positive clones observed by fluorescence microscopy, cell-associated virus was collected and released by one freeze-thaw cycle followed by sonication. mCherry-expressing virus was then plaque purified in BHK cells as described (Isaacs, 2004).

Analysis of horizontal gene transfer events by inverse PCR (iPCR)

Virus DNA was extracted from mCherry-expressing clones as previously described from infected BHK cells (Esposito et al., 1981). Purified DNA was digested with XbaI and SpeI or BglII (New England Biolabs (NEB)), diluted, and ligated (Quick ligase; NEB) to circularize linear fragments. Primers pointing away from each other in both mCherry (F- CGTGGAAACAGTACGAACGCG and R- CCATGTTATCCTCCTCGCCC) and K3L (F- GAGCATAATCCTTCTCGTATACTC and R- GAATATAGGGATAAACTGGTAGGG) were used to amplify regions flanking mCherry-K3L insertions. Resulting PCR bands were gel purified and Sanger sequenced with the iPCR primers. From this sequencing, the general location of the mCherry-K3L insert could be inferred. Primers flanking this region were then used to amplify the putative insertion, along with flanking sequences. These amplicons were then

gel purified, Sanger sequenced, and compared to the parent (VCR2) genome to characterize gene integrations. Because isolates 7 and 9 incorporated mCherry-K3L within the repetitive ITR, each end was PCR amplified separately using unique flanking primers and internal primers (mCherry-R and K3L-F, above). However, only the 5' end of the insertion of isolate 9 was amplified, despite multiple attempts. Thus, we cannot be sure whether this isolate includes a TSD or was cut at a LINE-1 endonuclease site.

Long read genome sequencing of virus isolates

Virus particles from plaques expressing mCherry-K3L were isolated from BHK cells and virus cores were purified by ultracentrifugation through a 36% sucrose cushion at 60k rcf for 1 hr. Virus DNA was extracted as previously described (Esposito et al., 1981). The SQK-LSK108 library kit (Oxford Nanopore Technologies) was used to prep isolate 1 DNA, which was sequenced on a FLO-MIN106 cell. A SQK-RBK001 kit was used to prep DNA from isolates 2-5, which were multiplexed on a FLO-MIN107 cell. Isolates 6-10 were also multiplexed using the SQK-RBK004 library prep kit, and sequenced on FLO-MIN107. Reads were base-called with the Oxford Nanopore Albacore program and aligned to a reference genome that included the VCR2 genome on one contig and the cellularly-expressed mCherry-K3L construct on a separate contig using default NGMLR parameters (Sedlazeck et al., 2018). Integrative Genomics Viewer (IGV) software (Broad Institute) was used to visualize insertions (Robinson et al., 2011). Sequencing data is deposited in the NCBI SRA database, accession number: PRJNA614958.

Measuring titers of virus strains and isolates

We plated 5×10^6 cells in 100mm dishes and infected them 16 hours later at an MOI of 0.1. Three plates (biological replicates) were infected with each isolate or strain. After 48 hours, media was aspirated, and cell-associated virus stock was collected by one freeze-thaw and sonication cycle. Then, 6-well plates were seeded with 5×10^5 BHK cells per well, and, 16 hours later, infected with 10-fold serial dilutions of virus stock (2 wells per dilution) in 200uL media. After 2 hours at 37°C, 2mL of media was added to each well. Cells were fixed and stained with 20% Methanol + 0.2% crystal violet 48 hours post infection. Wells with 10-100 plaques were counted and averaged to calculate virus titer. Each of the biological replicates is shown in Figure 2 and Figure 4 along with an average of the three and calculated standard deviations.

Alu assay of RK13 cells

5×10^5 HA-HeLa or RK13 cells were plated per 100mm dish (3x each cell type). 16 hours later, each plate was transfected with 5ug Alu-Neo (Dewannieux et al., 2003). After 2 days, all cells were treated with 2ug/mL G418. Media was replaced daily for 7 days, after which time it was removed, and cells were fixed and stained with 20% Methanol + 0.2% crystal violet.

Detection of duplication events

Single Nanopore reads with multiple tandem copies of the mCherry-K3L fusion gene were analyzed using IGV (Robinson et al., 2011). Duplications were confirmed by PCR, using the inverse PCR primers in mCherry (F- CGTGGAACAGTACGAACGCG and R- CCATGTTATCCTCCTCGCCC). PCR amplicons were then gel purified and Sanger sequenced to determine break points. However, due to the repetitive nature of the sequence flanking the isolate 7 insertion, the exact breakpoint was not able to be determined.

Serial passage of VACV-eIF2 α virus strains

For each passage, 150-mm dishes were seeded with an aliquot from the same stock RK13- cells (5×10^6 cells/dish). For P1, dishes were infected with eIF2 α -GFP or eIF2 α^{S51A} -GFP virus (MOI = 1.0) for 2 hours in 5mL and then supplemented with 15mL medium. After 48 hours, cells were washed, pelleted, and resuspended in 1mL of medium. Virus was released by one freeze-thaw cycle followed by sonication. 900 μ L of virus was then used to infect a new dish of cells for P2, and the process was repeated for subsequent passages. Viral titers were determined using the remaining 100 μ L of reserved virus stocks from each passage by 48-hour plaque assay in RK13-E3K3 cells.

Analysis of viral-encoded eIF2 α genes

RK13-E3K3 cells were infected with P4 eIF2 α -GFP or eIF2 α^{S51A} -GFP viruses (MOI = 0.1) for 24 hours. Virus-infected cells were collected, and total viral DNA extracted as previously described (Esposito et al., 1981). The region between K2L and K4L containing the different viral-encoded eIF2 α genes was amplified by PCR with primers K2L_seq_F and K4L_seq_R (5'-GGCATTGGTAAATCCTTGCAGA and 5'-CACCTTTTAGTAGGACTAGTATCGTACAA, respectively). SNV detection was performed by sequencing across eIF2 α using primers eIF2 α _F

and eIF2 α _R for eIF2 α -GFP and eIF2 α ^{S51A}-GFP PCR products, CNV analysis was performed by PCR using primers eIF2 α _rep_F (5'-CCTCCTATGGAAGCCAAAGCTGAAGATGAA) and eIF2 α _rep_R (5'-CCTCCTATCTACAACCTTAGACC CGGCAT) for eIF2 α -GFP and eIF2 α ^{S51A}-GFP viral DNA. Any CNV PCR products formed were sequenced using the same primers for breakpoint detection.

References

- Bahar, M.W., Graham, S.C., Chen, R.A., Cooray, S., Smith, G.L., Stuart, D.I., and Grimes, J.M. (2011). How vaccinia virus has evolved to subvert the host immune response. *J Struct Biol* 175, 127-134.
- Beattie, E., Tartaglia, J., and Paoletti, E. (1991). Vaccinia virus-encoded eIF-2 alpha homolog abrogates the antiviral effect of interferon. *Virology* 183, 419-422.
- Bellehumeur, C., Nielsen, O., Measures, L., Harwood, L., Goldstein, T., Boyle, B., and Gagnon, C.A. (2016). Herpesviruses Including Novel Gammaherpesviruses Are Widespread among Phocid Seal Species in Canada. *J Wildl Dis* 52, 70-81.
- Bratke, K.A., McLysaght, A., and Rothenburg, S. (2013). A survey of host range genes in poxvirus genomes. *Infect Genet Evol* 14, 406-425.
- Brennan, G., Kitzman, J.O., Rothenburg, S., Shendure, J., and Geballe, A.P. (2014). Adaptive gene amplification as an intermediate step in the expansion of virus host range. *PLoS Pathog* 10, e1004002.
- Caprari, S., Metzler, S., Lengauer, T., and Kalinina, O.V. (2015). Sequence and Structure Analysis of Distantly-Related Viruses Reveals Extensive Gene Transfer between Viruses and Hosts and among Viruses. *Viruses* 7, 5388-5409.
- Chakrabarti, S., Sisler, J., and Moss, B. (1997). Compact, synthetic, vaccinia virus early/late promoter for protein expression. *Biotechniques* 23, 1094-1097.
- de Groof, A., Guelen, L., Deijis, M., van der Wal, Y., Miyata, M., Ng, K.S., van Grinsven, L., Simmelink, B., Biermann, Y., Grisez, L., *et al.* (2015). A Novel Virus Causes Scale Drop Disease in *Lates calcarifer*. *PLoS Pathog* 11, e1005074.

- Deeg, C.M., Chow, C.T., and Suttle, C.A. (2018). The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. *Elife* 7.
- Dewannieux, M., Esnault, C., and Heidmann, T. (2003). LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* 35, 41-48.
- Elde, N.C., Child, S.J., Eickbush, M.T., Kitzman, J.O., Rogers, K.S., Shendure, J., Geballe, A.P., and Malik, H.S. (2012). Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* 150, 831-841.
- Elde, N.C., Child, S.J., Geballe, A.P., and Malik, H.S. (2009). Protein kinase R reveals an evolutionary model for defeating viral mimicry. *Nature* 457, 485-489.
- Elde, N.C., and Malik, H.S. (2009). The evolutionary conundrum of pathogen mimicry. *Nat Rev Microbiol* 7, 787-797.
- Esnault, C., Maestre, J., and Heidmann, T. (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 24, 363-367.
- Esposito, J., Condit, R., and Obijeski, J. (1981). The preparation of orthopoxvirus DNA. *J Virol Methods* 2, 175-179.
- Gilbert, C., Peccoud, J., Chateigner, A., Moumen, B., Cordaux, R., and Herniou, E.A. (2016). Continuous Influx of Genetic Material from Host to Virus Populations. *PLoS Genet* 12, e1005838.
- Goebel, S.J., Johnson, G.P., Perkus, M.E., Davis, S.W., Winslow, J.P., and Paoletti, E. (1990). The complete DNA sequence of vaccinia virus. *Virology* 179, 247-266, 517-263.
- Gubser, C., Bergamaschi, D., Hollinshead, M., Lu, X., van Kuppeveld, F.J., and Smith, G.L. (2007). A new inhibitor of apoptosis from vaccinia virus and eukaryotes. *PLoS Pathog* 3, e17.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* 59, 307-321.
- Haller, S.L., Peng, C., McFadden, G., and Rothenburg, S. (2014). Poxviruses and the evolution of host range and virulence. *Infect Genet Evol* 21, 15-40.
- Hughes, A.L., and Friedman, R. (2005). Poxvirus genome evolution by gene gain and loss. *Mol Phylogenet Evol* 35, 186-195.
- Isaacs, S.N., ed. (2004). *Vaccinia Virus and Poxvirology* (Humana Press).

- Keeling, P.J., and Palmer, J.D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* 9, 605-618.
- Koonin, E.V., and Yutin, N. (2019). Evolution of the Large Nucleocytoplasmic DNA Viruses of Eukaryotes and Convergent Origins of Viral Gigantism. *Adv Virus Res* 103, 167-202.
- Lefkowitz, E.J., Wang, C., and Upton, C. (2006). Poxviruses: past, present and future. *Virus Res* 117, 105-118.
- McFadden, G., and Dales, S. (1979). Biogenesis of poxviruses: mirror-image deletions in vaccinia virus DNA. *Cell* 18, 101-108.
- McLysaght, A., Baldi, P.F., and Gaut, B.S. (2003). Extensive gene gain associated with adaptive evolution of poxviruses. *Proc Natl Acad Sci U S A* 100, 15655-15660.
- Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., and Moran, J.V. (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet* 31, 159-165.
- Perez-Carmona, N., Farre, D., Martinez-Vicente, P., Terhorst, C., Engel, P., and Angulo, A. (2015). Signaling Lymphocytic Activation Molecule Family Receptor Homologs in New World Monkey Cytomegaloviruses. *J Virol* 89, 11323-11336.
- Piskurek, O., and Okada, N. (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proc Natl Acad Sci U S A* 104, 12046-12051.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nature Biotechnology* 29, 24-26.
- Saraiva, N., Prole, D.L., Carrara, G., Maluquer de Motes, C., Johnson, B.F., Byrne, B., Taylor, C.W., and Smith, G.L. (2013). Human and viral Golgi anti-apoptotic proteins (GAAPs) oligomerize via different mechanisms and monomeric GAAP inhibits apoptosis and modulates calcium. *J Biol Chem* 288, 13057-13067.
- Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., and Schatz, M.C. (2018). Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods* 15, 461-468.
- Sklenovska, N., and Van Ranst, M. (2018). Emergence of Monkeypox as the Most Important Orthopoxvirus Infection in Humans. *Front Public Health* 6, 241.
- Swanstrom, R., Parker, R.C., Varmus, H.E., and Bishop, J.M. (1983). Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. *Proc Natl Acad Sci U S A* 80, 2519-2523.

Wijayawardena, B.K., Minchella, D.J., and DeWoody, J.A. (2013). Hosts, parasites, and horizontal gene transfer. *Trends Parasitol* 29, 329-338.

Yao, X.D., and Evans, D.H. (2001). Effects of DNA structure and homology length on vaccinia virus recombination. *J Virol* 75, 6923-6932.

Zhang, Z., Carriero, N., and Gerstein, M. (2004). Comparative analysis of processed pseudogenes in the mouse and human genomes. *Trends Genet* 20, 62-67.

Acknowledgments We thank D. Hancks, E. Choung, and E. Grow for reagents and advice.

Funding This work was supported by NIH grants R01GM114514 and R35GM134936 (N.C.E.), T32GM007464 (S.M.F. and T.A.S), and T32AI055434 (K.S.R.). N.C.E. is a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease and H.A. and Edna Benning Presidential Endowed Chair.

Author contributions N.C.E. and S.R. conceptualized the project. N.C.E., S.M.F. and S.R. contributed to experimental design. S.M.F. and K.R.C. performed experiments. S.M.F, K.R.C., and T.A.S. analyzed and validated data. S.M.F., S.A.G. and N.C.E. visualized data. N.C.E. supervised. S.M.F wrote the original draft and N.C.E., S.M.F., and S.A.G. reviewed and edited the manuscript.

Competing interests The authors declare no competing interests.

Data and materials availability Bioproject PRJNA614958

Supplemental Materials

Figures S1-S6

Table S1-S2

Supplementary sequences

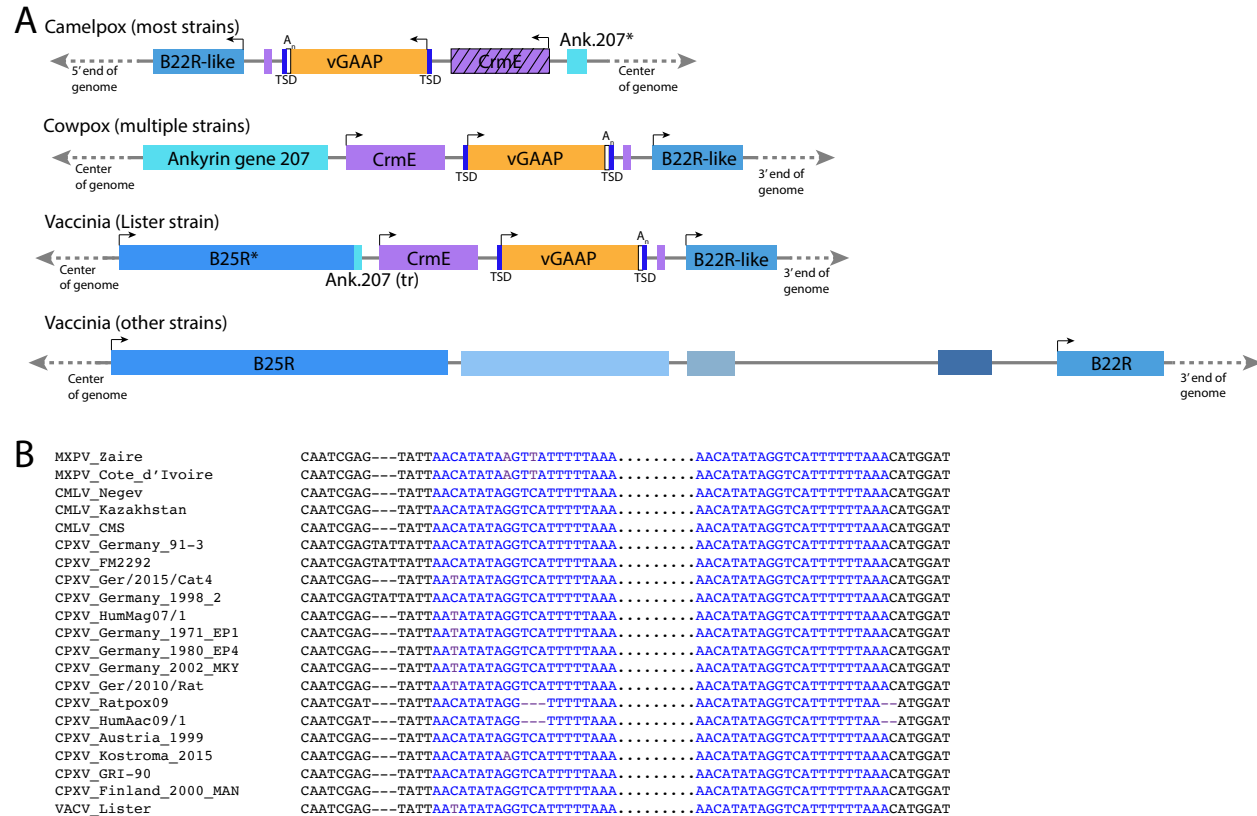


Fig. S1.

(A) v-GAAP (orange bar) is encoded by many orthopox species, including most camelpox strains, many cowpox strains, and at least one vaccinia strain. In camelpox viruses, v-GAAP is encoded on the left end of the genome; in all other viruses, v-GAAP is on the right end. In all viruses, v-GAAP is flanked by CrmE and vaccinia-B22R. In vaccinia strains that do not encode v-GAAP, CrmE is also missing, suggesting recombination of the region between B22R and B25R, which is truncated in the Lister strain. (B) Alignment of target site duplications (blue) and flanking sequence in 21 v-GAAP-encoding poxvirus strains. MPXV, monkeypox; CMLV, camelpox; CPXV, cowpox; VACV, vaccinia. Asterisks denote gene truncations.

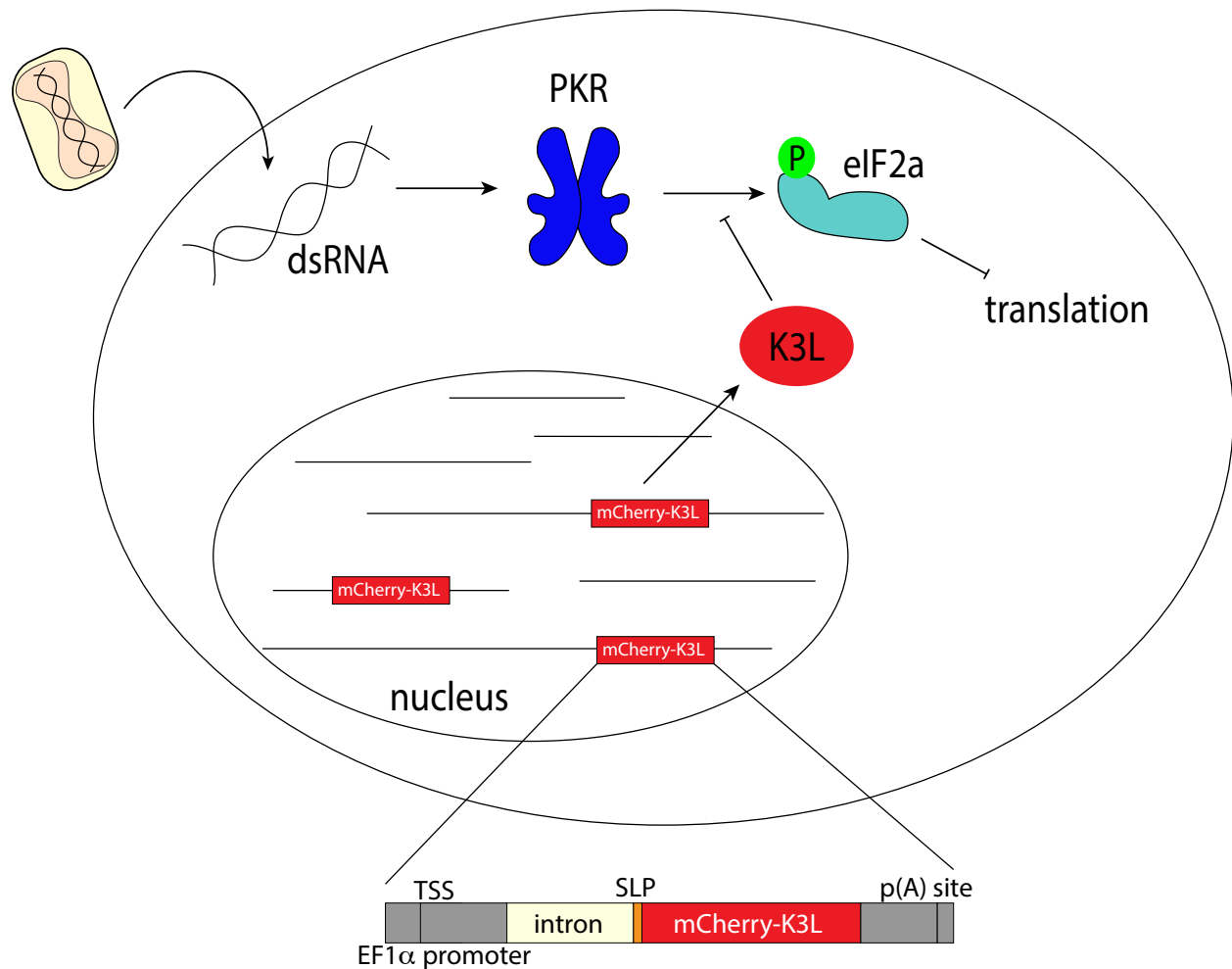


Fig. S2.

Protein Kinase R (PKR) is activated by binding dsRNA which is made in the cytoplasm during the course of a viral infection. When activated, PKR binds and phosphorylates the eukaryotic initiation factor, eIF2 α , leading to a block in translation and preventing viral replication. K3L acts as a pseudo-substrate of PKR, preventing activated PKR from binding eIF2 α , which is then free to initiate translation. In the complementing RK13-K3L cells, an mCherry-K3L fusion gene was integrated into cellular chromosomes. The cellular mCherry-K3L expression construct is shown below. The fusion gene is driven by a strong EF1 α promoter, but the construct also includes a poxvirus-specific promoter, the SLP, so that mCherry-K3L could be immediately expressed if transferred.

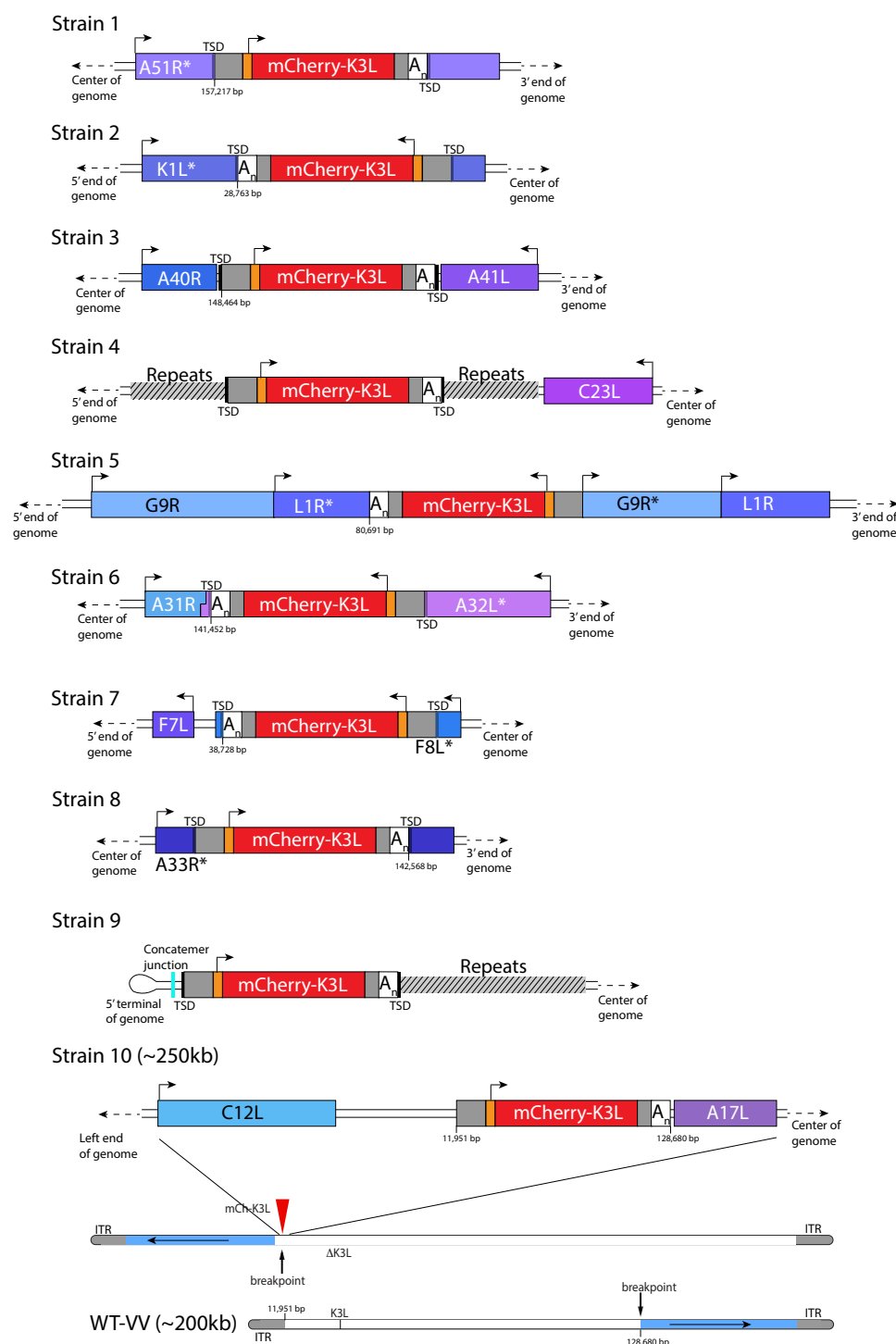


Fig. S3.

A detailed view of the region K3L inserted in each experimentally captured virus isolate. Arrows above cartoons indicate reading frame orientation. Flanking and/or interrupted (*) viral genes (blue/purple boxes), 3'/5' untranslated regions (UTRs; gray), and poly(A) tails (A_n ; white) are shown. Genomic rearrangements included in isolate 10 are shown below details. Some isolate 4 viruses include a second insertion that precisely matches that of isolate 2. We believe that this is due to recombination between the two strains during purification.

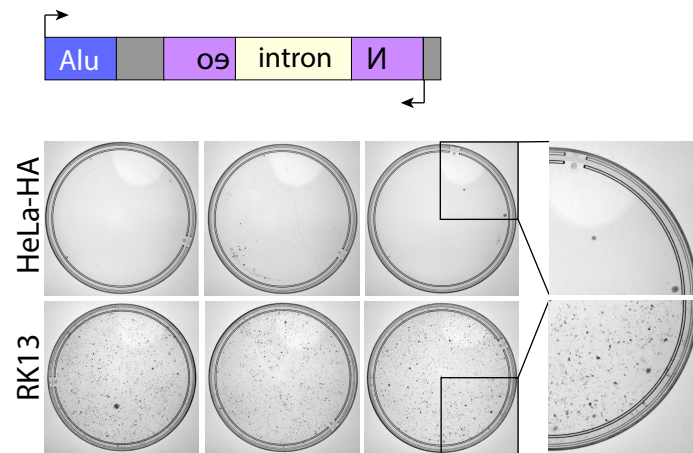


Fig. S4.

An Alu-Neo reporter (cartoon above) was used to interrogate the endogenous LINE-1 activity of the RK13-K3L cells compared to HeLa-HA cells. Below, representative plates from several experiments showing Neo-resistant colonies, with zoomed-in views to the right.

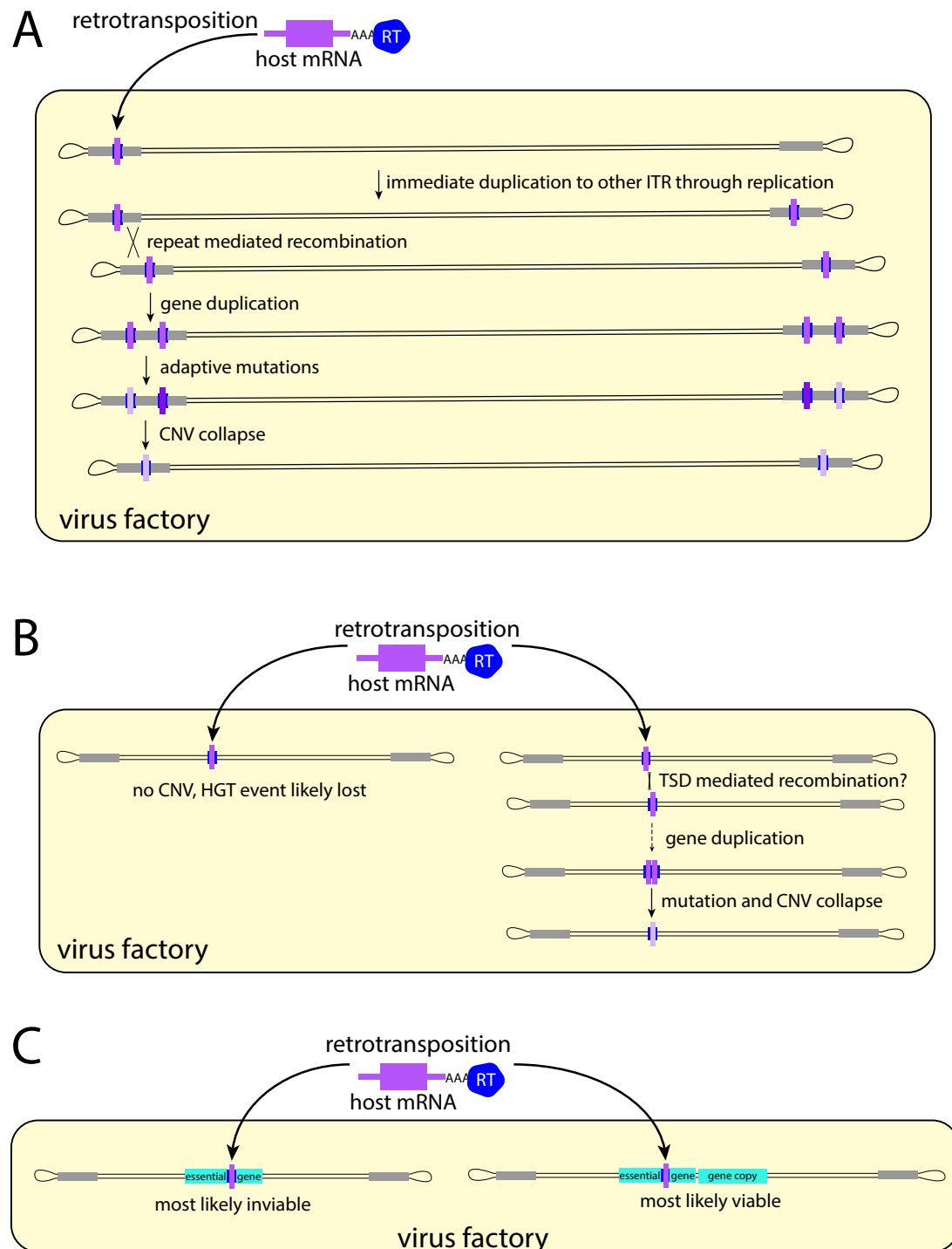


Fig. S5.

Models of evolution of newly acquired virus genes. (A) Genes that land in the repetitive region of the ITR are likely to be duplicated by both replication mechanisms (to the other ITR) and recombination mechanisms (tandem duplications). Duplications provide more chances for advantageous mutations to arise, increasing the likelihood of fixation. (B) Genes acquired outside the ITR may still be duplicated via TSD-mediated recombination. (C) Acquired genes that interrupt essential genes are unlikely to be maintained.

WT Vaccinia



VCR2



eIF2 α -GFP



eIF2 α^{S51A} -GFP



Fig. S6.

Structure of K3L region in eIF2 α vaccinia strains engineered for experimental evolution compared to wild type (WT) vaccinia and VCR2.

Organism	protein	Accession number
Cowpox strain Ger14Cat1	vGAAP/TMBIM4	SNB58186.1
Cowpox strain Ger10Rac	vGAAP/TMBIM4	SNB48400.1
Cowpox strain Ger02MKY	vGAAP/TMBIM4	ADZ30397.1
Cowpox strain Ratpox09	vGAAP/TMBIM4	CRL86727.1
Cowpox strain Ger17vFME	vGAAP/TMBIM4	SPN68348.1
Vaccinia strain EvansCDS	vGAAP/TMBIM4	AAV98625.1
Cowpox strain Finland_2000_MAN	vGAAP/TMBIM4	ADZ29327.1
Cowpox strain Kost15Hs	vGAAP/TMBIM4	AQQ13071.1
Vaccinia strain Lister	vGAAP/TMBIM4	Q49P94.1
Cowpox strain Aus99cat	vGAAP/TMBIM4	ADZ24213.1
Vaccinia strain USSR_CDS	vGAAP/TMBIM4	AAW21699.1
Camelpox strain CMS	vGAAP/TMBIM4	AAG37461.1
Camelpox strain Negev16	vGAAP/TMBIM4	QCW07309.1
Camelpox strain M96_Kazakstan	vGAAP/TMBIM4	NP_570396.1
Cowpox strain GRI-90	vGAAP/TMBIM4	CAD90752.1
Cowpox strain FM2292	vGAAP/TMBIM4	CRL87022.1
Cowpox strain Ger07Vole	vGAAP/TMBIM4	SNB49536.1
Cowpox strain Ger91-3	vGAAP/TMBIM4	ABD97561.1
Monkeypox strain USA03_044	vGAAP/TMBIM4	AAV96807.1
Monkeypox strain Cote d'Ivoire 1971	vGAAP/TMBIM4	AKG51345.1
Monkeypox strain DRC_07	vGAAP/TMBIM4	AGR36454.1
Monkeypox strain Zaire96	vGAAP/TMBIM4	NP_536611.1
Monkeypox strain Congo03	vGAAP/TMBIM4	AAV97191.1
Scale Drop Disease Iridovirus	TMBIM4	YP_009163856.1
Harpseal herpes virus	TMBIM4	AJG42933.1
Homo sapiens	TMBIM4	NP_057140.2
Mus musculus	TMBIM4	NP_080893.1
Rattus Norvigicus	TMBIM4	NP_954547.1
Bos taurus	TMBIM4	NP_001014914.1
Ovis aries	TMBIM4	XP_014950252.2
Grizzly bear	TMBIM4	XP_026355879.1
CA Sea lion	TMBIM4	XP_027451274.1
Stellar sea lion	TMBIM4	XP_027966737.1

Weddell seal	TMBIM4	XP_006731256.1
Northern fur seal	TMBIM4	XP_025738732.1
Ghost shark	TMBIM4	NP_001279731.1
Bettafish	TMBIM4	XP_029017613.1
Yellow perch	TMBIM4	XP_028426613.1
Gilt head Bream	TMBIM4	XP_030296356.1
Homo sapiens	TMBIM3	NP_000828.1
Mus musculus	TMBIM3	NP_075657.1
Rattus Norvigicus	TMBIM3	NP_695220.4
Bos taurus	TMBIM3	NP_001032682.1
Ovis aries	TMBIM3	NP_001265493.1
Grizzly bear	TMBIM3	XP_026344576.1
CA Sea lion	TMBIM3	XP_027465497.1
Weddell seal	TMBIM3	XP_006745469.1
Northern fur seal	TMBIM3	XP025727080.1
Whale shark	TMBIM3	XP020385145.1
Bettafish	TMBIM3	XP_029023514.1
Yellow perch	TMBIM3	XP_028453020.1
Gilt head Bream	TMBIM3	XP_030268118.8
Homo sapiens	TMBIM2	NP_036438.2
Mus musculus	TMBIM2	NP_082500.2
Rattus Norvigicus	TMBIM2	NP_653357.1
Bos taurus	TMBIM2	NP_001068886.1
Ovis aries	TMBIM2	XP_0040066421.1
Grizzly bear	TMBIM2	XP_026357704.1
CA Sea lion	TMBIM2	XP_027450780.1
Weddell seal	TMBIM2	XP_006729299.1
Northern fur seal	TMBIM2	XP_025739188.1
Whale shark	TMBIM2	XP_020373190.1
Bettafish	TMBIM2	XP_029006890.1
Yellow perch	TMBIM2	XP_028437735.1
Atlantic sea bass	TMBIM2	XP_018521446.1
Gilt head bream	TMBIM2	XP_030278166.1
Danio rerio	TMBIM2	XP_005163554.1
Homo sapiens	TMBIM1	NP_001308356.1

Mus musculus	TMBIM1	NP_081430.3
Rattus Norvigicus	TMBIM1	NP_001007714.1
Bos taurus	TMBIM1	NP_991367.1
Ovis aries	TMBIM1	XP_004004971.1
CA Sea lion	TMBIM1	XP_027443873.1
Weddell seal	TMBIM1	XP_006750605.1
Northern fur seal	TMBIM1	XP_025747486.1
Ghost shark	TMBIM1	XP_007882730.1
Bettafish	TMBIM1	XP_028986890.1
Yellow perch	TMBIM1	XP_028448525.1
Atlantic sea bass	TMBIM1	XP_018532965
Gilt head bream	TMBIM1	XP_030257449.1
Danio rerio	TMBIM1	NP_001005992.2
Homo sapiens	TMBIM5	NP_055209.2
Mus musculus	TMBIM5	NP_001186051.1
Rattus Norvigicus	TMBIM5	NP_001005908.1
Bos taurus	TMBIM5	NP_001029224.1
Ovis aries	TMBIM5	XP_004021578.1
Grizzly bear	TMBIM5	XP_026360365.1
CA Sea lion	TMBIM5	XP_027452399.1
Monk seal	TMBIM5	XP_021555821.1
Northern fur seal	TMBIM5	XP_025736444.1
Ghost shark	TMBIM5	AFK11640.1
Bettafish	TMBIM5	XP_028990283.1
Yellow perch	TMBIM5	XP_028423643.1
Atlantic sea bass	TMBIM5	XP_18554801.1
Danio rerio	TMBIM5	NP_956885.1
Homo sapiens	TMBIM6	NP_003208.2
Mus musculus	TMBIM6	NP_001164505.1
Rattus Norvigicus	TMBIM6	NP_062254.2
Bos taurus	TMBIM6	NP_001069882.1
Ovis aries	TMBIM6	XP_004006423.2
Grizzly bear	TMBIM6	XP_026357713.1
CA Sea lion	TMBIM6	XP_027449007.1
Weddell seal	TMBIM6	XP_006729302.1
Northern fur seal	TMBIM6	XP_025738444.1

Ghost shark	TMBIM6	XP_007910207.1
Bettafish	TMBIM6	XP_029013526.1
Yellow perch	TMBIM6	XP_028439299.1
Atlantic sea bass	TMBIM6	XP_018531358.1
Gilt head bream	TMBIM6	XP_030277962.1

Table S1.

Accession numbers for sequences used in phylogenetic analysis (Fig. 1B)

109	436	
VACLister	-----	
CML_Ngv16	-----	
CMLV_Orig	-----	
CMLV_CMS	-----	
CPX_Ger07	-----	
CPXFM2292	-----	
CPX_Gri90	-----	
CPX_Aus99	-----	
CPX_Kos15	-----	
CPX_Fin00	-----	
CPX_Ger17	-----	
CPX_Rat09	-----	
CPX_Ger91	-----	
CPX_Ger02	-----	
CPX_Ger14	-----	
CPX_Ger10	-----	
CPX_Ger80	-----	
CPX_Ger98	-----	
MPX_USA03	-----	
MPX_Ivr71	-----	
MPX_Zre96	-----	
MPX_Con03	-----	
MPX_DRC07	-----	
SDDIridoV	-----	
HSherpesV	-----	
Human_LG4	-----	
Mouse_LG4	-----	
NRat_LFG4	-----	
Cow_LFG4	-----	
Sheep_LG4	-----	
GrizzlyL4	-----	
CAlionL4	-----	
StClionL4	-----	
WeddellL4	-----	
Nfurseal4	-----	
GShark_L4	-----	
betta_LG4	-----	
YlwPrchL4	-----	
GHbreamL4	-----	
Human_LG1	MSHEKS-FLVSGDNYPPNPGYPGGPQPP--MPPYAQPPYPGAPYPQPPFQ	
Mouse_LG1	MSHEKS-FLVSGDSYPPQNI---VGQPAP--MPPYVQAPYPGAPYPQAPFQ	
NRat_LFG1	MSHEKS-FLVSGDSYPPNPGYPVGPQAP--MPPYVQPPYPGAPYPQAAFQ	
Cow_LFG1	MSHEKS-FLVSGDSYPPNPGYPGGPQPS--MAPY-----PGAPYPQAPFQ	
Nfseal_L1	MPHEKS-FLVSGDSYPPNPGYPGGPQPS--MPPY-----PVAPYPQAPFQ	
WeddellL1	MPHEKS-FLVSGDSYPPNPGYPGGPQPS--MPPY-----PGAPYPQAPFQ	
CAlionL1	MPHEKS-FLVSGDSYPPNPGYPGGPQPS--MPPY-----PVAPYPQAPFQ	
Sheep_LG1	MSHEKS-FLVSGDSYPPNPGYPEGPQPS--MAPY-----PGAPYPQAAFQ	
GrizzlyL1	MPHEKS-FLVSGDSYPPNPGYPGGPQPS--MPPY-----PGAPYPQAPFQ	
WhSharkL1	MSQEKT-YLVSGD---PQQPPY---PRNEMSYGNDHYPPQP-PPYGQPVFP	
Betta_LG1	MSQDKSGYPVMGETNPLHSNMY-GPPQPFGFMPPPNYSQAPGGFYPPA---	
GHbreamL1	MSQDKSGYPIMGETNPLHNNVY-GPPQPFGFMPPPNYSQAPGGFYPPA---	
YlwPrchL1	MSQEKSGYPVMGENNPLHNNVY-GPPQPFGFMPPPNYSQAPGGFYPPA---	
Human_LG2	MTQGKLSVANKAPGTEGQQ-QVH---GEKK-----EAPAVPSAPPSYE	
Mouse_LG2	MTQGKLSVANKAPGTEGQQHQAN---GEKK-----DAPAVPSAPPSYE	

13

16

AAFLALTTYTLOSKRDFSKLGAGLFAALWILILSGLLGI----F-VONETVKLVLSAFGAL

18

19

TPPPPKKCPAVPCACFPFCCSLTNN-----
V-----MCGFVLFDTQLIEKAE---N--GDKDYVWHCVDLFLDFITIF
I-----MCGFVLFDTQLIEKAE---N--GDKDYVWHCVDLFLDFITIF
I-----MCGFVLFDTQLIEKAE---N--GDKDYVWHCVDLFLDFITIF
I-----MCGFVLFDTQLIEKAE---N--GDKDYVWHCVDLFLDFITIF

LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
-----SMTRTH-----
-----SMTRTH-----

LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----

LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----
LHLLQLLEVSNNK-----

MRILHILKSIKNQ-----
MDFLRILGESKSESISITV
LHLLRFLEAVNKK-----
LHLLKFLEAVNKK-----
LHLLKFLEAVNKK-----
LHLLRVLEAANKK-----
LHLLRVLEAVNKK-----
LHLLRVLEAVNKK-----
LHLLRFLEAVQKK-----
LHLLRFLEAVNKK-----
LHLLRFLEAVNKK-----
LHLLRFLEAVNKK-----
LHLLRFLEAVNKK-----
LHLLRFLEAVNKK-----
LHLLRFLEAVNKK-----
LHILRVLDMSMKKH-----
LHILRLDMSMKKH-----
LHILRLDMSMKKH-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----
LYILTIIGRAKE-----

LYILAIVGRSRE-----
LYILAIVGRSRE-----
LYILTIVGRSRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFFLQLFGTNRE-----
TFMLQIFGSNRE-----
SFLQLTGQGRD-----
SFFLQIFGTKQE-----
SFFLQIFGTKQE-----

```

SFFLQIFGTKRE-----
LFFLQLFGS-RE-----
TFVLQLMGDRN-----
TFVLQLVGSRD-----
TFVLQLVGNRD-----
TFVLQLVGSRD-----
TFVLQLVGRQD-----
TFVLQLVGDRN-----
TFVLQLVGDRN-----
TFMLQLVGDRN-----
INLLQLFGSRE-----
LFLQLIIGASTK-----
LFLQLIIGASTK-----
IFLLQLIGAATK-----
LFLQLIIGASTK-----
IFLLQLIIGYAER-----
MRVATMLATGG-NRKK---
MRVATMLATGS-NRKK---
MRVATMLATGS-NRKK---
MRVATILATGG-NRKK---
MRVASILATGG-NRKK---
MRVASILATGS-NRKK---
MRVASILATGG-NRKK---
MRVASILATGS-NRKK---
MRVASILATGS-NRKK---
IRVAAMLGNGGGSRRK---
MRLVMILANGGGSRRK---
MRLVMILANGGGGKRR---
MRLVMILANGGGSRRK---
MRLVMILSGAG-SRRK---
RKLMMILAMNEKDKKKEKK
RKLMLILAFNEKDKKKEKK
RKLMLILAFNEKDKKKEKK
RKLMMILAMNEKDKKKEKK-
RKLMMILAMNEKDKKKEKK-
RKLMMILAMNEKDKKKEKK
RKLMMILAMNEKDKKKEKK
RKLMMILAMNEKDKKKEKK
RKLMMILAMNEKDKKKEKK
RKLMMILAMNEKDKKKEKK
-----
RKLMMILALNDKDKKKEKK
RKLMMILAMNDKDKKKEKK
RKLMMILALNDKDKKKEKK
RKLMMILAMNDKDKKKEKK

```

Table S2.

Phylog alignment for sequences used in phylogenetic analysis (Fig. 1B)

Supplementary sequences:

piggybac_SLP-mCherry-K3L:

actcttcttttcaatattattgaagcatttatcagggtattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaatagggttccgcgcacatttccccga
aaagtgccacctaattgttaagcgttaataattttgttaaaattcgcgttaaaattttgttaaatcagctcatttttaaccaataggccgaaatcgccaaatcccttataatcaaa
agaatagaccgagatagggttgagtgtgttcagtttggacaagagtcactattaaagaacgtggactccaacgtcaaaaggcgaaaaaccgtctatcaggggcgatg
gcccactacgtgaaccatcacccctaatacagtttttgggtcagagtgccgttaaagcactaaatcggaaccctaaagggaagccccgatttagagcttgacggggaag
ccggcgaaacgtggcgagaaagggaagggaagaaagcgaaggagcggcgctagggcgctggaagtgtagcgggtcacgctgcgcgttaaccaccacacccgccg
cgcttaatgcgcgtacagggcgctcccatcgcattcaggctgcgcaactgttgggaaggcgatcggtgcgggctcttctgctattacgccagctggcgaaagg
gggatgtgtgcaaggcgattaagtgggtaacgccagggtttccagtcacgacgtgttaaacgacggccagtgagcgcgctctgctcattcacgttttgaacccgt
ggaggacggcgagactcgggtgcaaatgtgtttacagcgtgatggagcagatgaagatgctgcacacgctgcagaacacgcagctagattaaccctagaaagataa
tcattgtgacgtacgttaagataatcatgcgtaaaatgacgcatgtgtttatcggctgtatcagaggtttatttataattgaatagatattaagttttattatttacact
acatactaataataaattcaacaacaattttattgtttattttattataaaaaaactcaaaatttcttataaagtaacaaaactttatagggacagcccccc
cccaagccccaggatgtaattacgtccctccccgctagggggcagcagcgagccgcccgggctccgctccggcgctcccccccatcccagccgagccg
gcagcgtgcggggacagcccgggcacggggaagggtggcagggatgcttctctgaacgcttctcgtgctctttagcctgcagacacctggggggatagggg
aaaaGGCCTCCACGGCCAGACTAGAACTTTTAAAGAAAAGGGGGGATTGGGGGGGTACAGTGCAGGGG
AAAGAATAGTAGACATAATAGCAACAGACATACAACTAAAGAATTACAAAAACAAATTACAAAAAT
TCAAAATTTTATCGATCACGAGACTAGCCTCGAGGGATCacccgggtcccgtcagtgggcagagcgacatcgccacagtcccc
gagaagtggggggagggtgcgcaattgaacgggtgcctagagaagggtggcggggtaaaactgggaaagtgatgtcgtgactggtccgcttttcccagggt
gggggagaacctatataagtgcaagtagtcgccgtgaacgttcttttgcacaggggttggccgcagaacacagctgaagcttcgaggggctcgcattctctcttcacgc
gcccgcgcctacctagggcgcccatccacgcgggtgagtcgcttccgctccgctggtggtgctcctgaactgcgtccgctcgttaggtaagttaaagct
caggtcgagaccgggcttctgctggcgctcccttgagcctacactagactcagccggtctccacgcttgcctgacctgcttgcactctacgttctgttctgttct
gttctgcgcgttacagatcaagctgtgaccggcgccGAATTCCTTCAGGGTGAGTTTGGGGACCCTTGATTGTTCTTTCTTT
TTCGCTATTGTAAATTCATGTTATATGGAGGGGGCAAAGTTTTCAGGGTGTTGTTTAGAATGGGAAG
ATGTCCCTTGTATCACCATGGACCCTCATGATAATTTTGTTCCTTCACTTTCTACTCTGTTGACAACCA
TTGTCTCTCTTATTTTCTTTTCATTTTCTGTAACCTTTTCGTTAAACTTTAGCTTGCATTTTGTAAACGAAT
TTTTAAACTACTTTTGTATTGTGTGAGATTGTAAGTAACTTCTCTAATCACTTTTTCCTAAGGCAATC
AGGGTATATTATATTGTACTTCAGCAGACTTTTAGAGAACTTGTATAATTAATGATAAGGTAGA
ATATTTCTGCATATAAATTCTGGCTGGCGTGGAATATTCTTATTGGTAGAAACAACTACATCCTGGTC
ATCATCCTGCCTTTCTCTTTATGGTTACAATGATATACACTGTTTGGAGATGAGGATAAAATACTCTGAG
TCCAAACCGGGCCCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTTCTACAGCTCCTGGGCAACGGT
ACCGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGGGAATATAAATGGTGAGC
AAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCT
CCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGAC
CGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCTCAGTTCA
TGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCGACTACTTGAAGCTGTCTTCCCC
GAGGGCTTCAAGTGGGAGCGCGTGATGAACCTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACT
CCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGC
CCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCGAGGACGGCG
CCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAA
GACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGAC
ATCACCTCCCAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCA
CCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTGTATTTCGTTGCCCA
ATGCGGGTGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATATATTTATCTTTTTGACT
ATCCTCACTTTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGATATGTTGAATATAGG
GATAAACTGGTAGGGAAAAGTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAAGGATATATAG
ATGTCAATTACAAAAGGATGTGTAGACATCAATAATAGCGGCCGCTAaactgtttattgcagcttataatggttacaaataa
agcaatagcatcacaatttcacaataaagcattttttactgcattctagttgtgtttgtccaaactcatcaatgtatcttatcatgtctGGATCCggctccgcgcg
ggttttggcgctccgcggcgccccctcctcacggcgagcgtgccacgtcagacgaaggcgagcgcgctcctgaccttccgccggagcgtcaggaca
gcggccgctgctcataagactcggccttagaaccacagatcagcagaaggacattttaggacgggacttggtgacttagggcactggttttcttcagagagcgg
aacaggcgaggaagtagtcccttctcggcgattctcggcagggatcctcgtggggcggtgaacgccgatgattatataaggacgcgccgggtgtggcacagctagt
tccgtcgagccgggatttgggtcgcgggttctgttggatcgtgtgctgctgacttggtgagtagcgggctgctgggtggccgggcttctgtggccgccgggccc
gctcggtgggacggaagcgtgtggagagaccgccaaggcgtgtgtgctgctgagcgaaggttgcctgaactgggggttgggggagcgcagcaaaatggc
ggctgttcccgaagcttgaatggaagacgttgtgaggcgggcgtgtgagtcgttgaacaaggttggggggcatggtggcgccgaagaacccaaggtcttgaggcctt
cgctaatgcgggaagccttattcgggtgagatgggctggggcaccatcgtgggacacctgacgtgaagtttgcactgactgagagaactcggttctgtctgttgcggg
ggcggcagttatggcggtgcccgttgggcagtgacccgtaccttgggagcgcgcgccctcgtcgtgtgacgtcacccgttctgttggcttataatgcagggtggg

ccacctgccgtaggtgtgcgtaggcttttctcgtcgcaggacgcagggttcgggcttaggtaggctctcctgaatcgacaggcgccggacctctggtgagggga
gggataagttaggcgtcagtttcttggtcggtttatgtacatctcttaagtagctgaagctccggtttgaactatgcgtcggggtggcgagtgtgtttggaagtttt
taggcaccttttgaatgtaatcatttgggtcaatatgtaatttcagtgtagactagtaaatgtccgtaaatctggccgttttggctttttgtagacgtagcgatAG
CGCTTCTCATTGGGCATTCCAGCCTACCCAGCTCGGAGTTAGTTACTCCGTAAGTGTGGCCGGAACAG
AGTTCGTCCATCTAAAAAGGGAGGGGACCGGCGAACACATCTAGAGCCACCATGGCCACCGAGTACA
AGCCACAGGTGCGCCTCGCCACCCGCGACGACGTCCCCGGGGCGTACGCACCCTCGCCGCCGCGTTC
GCCGACTACCCCGCCACGCGCCACACCGTGCAYCCGGACCGCCACATCGAGCGGGTACCCGAGCTGC
AAGAACTCTTCTCACGCGCGTCCGGCTCGACATCGGCAAGGTGTGGGTCTCGCGGACGACGGCGCCGCG
GTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCYCGCGCA
TGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCTGGCGCCGCACCG
GCCCAAGGAGCCCGGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCCGACCACAGGGCAAGGGTCTGG
GCAGCGCCGTCGTGCTCCCCGAGTGGAGGCGGCCGAGCGCGCYGGGGTGGCCGCTTCCTGGAGAC
CTCCGCGCCCCGCAACCTCCCTTCTACGAGCGGCTCGGCTTACCGTCAACGCCGACGTCGAGGTGCC
CGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCGTCGACAATCAACCTCTGGATTACA
AAATTTGTGAAAGATTGACTGGTATTCTTAACATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTT
AATGCCTTTGTATCATGCGTTAACTAACTAaactgtttattgcagctataatggttacaataaagcaatagcatcacaatttcacaataaa
gcatttttctactgcattctagtgtgtgtgttgcacaaactcatcaatgtatcttcatgtctggaattgactcaaatgatgtcaattagtctatcagaagctatctggtctccctcc
gggggacaagacatccctgtttaataatfaaacagcagtggtccaaactgggtcttataccctgtctgtgcaaccaggttcaggggttctgtctctcacaggaacgaa
gtccctaaagaacagtggtgagccaggtttagccccgaattgactggttcttcttttagggcccaattggtatggtcttttccccgtatccccccaggtgtctgcaggctca
aagagcagcgagaagcgttcagaggaagcgtccgtccaccttccccgtgcccgggtgtccccgcacgtgcccggctggggtatgcggggggagcgccgg
accggagcggaagccccggcggtctgctgctgcccctagcgggggagggagcttaattacatccctgggggctttgggggggggctgtccctgatatacacaag
aaaatatataataaagtattcacgtaagtagaacaatgaataaacaataataattatcgtatgagtaaatctaaagtcacgtaaaagataatcatgctcatttgactacg
cggtcgttatagttcaaatcagtgacacttaccgcatgacaagcagcctcagggagctcaagcggcgactgagatgtcctaatgcacagcgacggattcgcgct
atttagaagagagagcaatattcaagaatgcatgctcaatttaccgagactatcttctagggtaattctagctgcatcaggatcatactgcgggtcttttccggtc
gtcatcgcccaagctggcgctatctggcgatcgggaggaagaagcccgcttctccgcgaggttgaagcggcatggaagagtttccgaggatgactgctgct
gcattgacgttgagcgaaaacgcagctttaccatgatgattcggaaggtgtggccatgcacgccttaacgggtgaactgtcgttcaggccacctgggataccagttcgt
cgcggtctttccggacacagttccggatggtcagcccgaagcgcatcagaacccgaacaataaccggcgacagccggaactcgctgcccgtgtgcagattaatgac
agcggtgcggcgctgggataattacgacgagggaggggtatctggtggtgcccagaaatggacatggatacccgtagttaccggcgggcgcgcttggcg
taatcatggtcatagctgttctgtgtgaaattgttatccgctcacaattccacacaacatacagccggaagcataaagttaaagcctggggtgctaatgagtgcgta
actcacattaattgcgttcgctcactgcccgtttccagtcgggaacacctgtcgtgccagctgcattaatgaatggccaacgcgcggggagagggcggttgcgtattgg
gcgtcttccgcttctcgtcactgactgctgctgctggctgctggcgagcggtatcagctcactcaaaaggcggaataacggttatccacagaatcagggg
ataacgcaggaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaggccgcgttgcgtggcttttccataggtctccggccccctgacgagcat
cacaataatcgacgctcaagtcagaggtggcgaaacccgacaggaactataagataccaggcggttccccctggaagctccctgctgctctctctgttccgacctgc
cgcttaccggatactgtccgccttctcccttccgggaagcgtggcgcttctcatagctcagctgtaggtatctcagttcgggtgtaggtcgttcgctccaaagctggcgctg
tgcacgaacccccctgacggcgacctgcgccttatccggttaactatcgtctttagtccaacccggtaagacacgacttatcggcactggcagcagccactggtaac
aggattagcagagcgaggtatgtaggggtgctacagagtcttgaagtgggtggcctaactacggctacactagaaggacagattttggtatctgcgctcgtgaagcca
gttaccttggaaaaagagttgtagctcttgatecggcaaaacaccacgctgtagcggtgtgtttttgttgaagcagcagattacgcgcaaaaaaaggatctc
aagaagatccttctgatttctacgggtctgacgctcagtggaacgaaactacgcttaagggttttggcatgagattcaaaaaggatcttcacctagatccttttaa
ttaaaatgaagtttaaatcaatctaaagtatatagtaaaactgtgctgacagttacaaatgcttaacagtgaggcacctatctcagcgtatgtctatttcttcacata
gttgctgactcccgtcgtgtagataactacgatacgggaggggttaccatctggccccagtgctgaatgataccgcgagaccacgctcaccggctccagattatca
gcaataaaccagccagccggaaggccgagcgcaaggtgctcgtcaacttctccctcactcagctcttaattgttggcggaagctagagtaagtagttgcc
agttaatagtttgcgaacgttggcattgctacagggatcgtggtgtcagcgtcgtgttggtagtggttcattcagctccgggtcccaacgatcaaggcgagttacatg
atcccccatgttgcgaataaagggttagctccttggctcctccgacgtgtgtcagaagtaagttggcgaggttatcactcatggtatggcagcactgcataattctctt
actgtcatgccatccgtaagatgctttctgtgactggtgagctactcaaccaagtcactctgagaatagtgtatgcggcgaccgagttgctcttccccggcgtaacacggga
taataccgcgccacatagcagaactttaaaagtgtcatcatttggaacacgttcttggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgtatgaacca
ctcgtgacccaactgatctcagcatcttttaccagcgtttctgggtgagcaaaaacaggaaggcaaatgccgcaaaaaagggaataaggcgacacggaaa
tgttgaatactcat

pBlue_165_mCherry-K3L:

GAGCTCtgcaacctagataaaaaaactattgtttatccactttctcgtatgttcgaatgagattataatcgttattatagggatgtggaagaattggaacacacg
cgtgctatataatgaagagataaaatacactccagtcgaagtatttcttttataaaaaatccatatataatttatattctgtaacatgttatccctttcaattaacaagttagttata
aaaaataaagaagcgaaatcaatgattaatagatgttaagaactataattacagtgattataagttagttggttagttaaagaagagataaacggttactaataattg
ttagttattgtctatatgatattacaacctattttgttcttatgttacattaataaaattttatgtgacacccattcatctggagaataactcttgataccattagttatcatata
ccatgattctggccaactctgcaaacacagcacGTGCAGCCCTGCAGATCGATTTAAAAATTGAAATTTTATTTTTTTTTTTT
GGAATATAAATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGTTCA
AGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCC
CTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCTGGGAC

ATCCTGTCCCTCAGTTCATGTACGGCTCCAAGGCTACGTGAAGCACCCCGCCGACATCCCCGACTAC
TTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGT
GACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCA
ACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATG
TACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACT
ACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGT
CAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGAACAGTACGAACGCGCC
GAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATT
TTGTTATTCGTTGCCCAATGCGGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTAT
AcATTTATCTTTTTGACTATCCTCACTcTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAG
ATATGTTGAATATAGGGATAAACTGGTAGGGAAAACTGTAAAAGTTAAAGTGATTAGAGTTGATTATA
CAAAAGGATATATAGATGTCAATTACAAAAGGATGTGTAGACATCAATAAataggcgttaaacaattggaATCAGC
TAAAAATTGAAATTTTATTTTTTTTTTTTGGAAATATAAatgattgaacaagatggattgcacgcagggttccggccgcttgggtggag
aggctattcggctatgactgggcacaacagacaatcggtgctctgatccgctgttccgctgtcagcgaggggcggcggcttcttgcagaccgacctgtc
cggtgccctgaatgaactgcaggacgaggcagcgcgctatcggtggtccacgacggcggttcttgcgcagctgtctgcagcttgcactgaagcgggaagg
actggtgctattggcggaagtgcggggcaggatctctgtcatctcacttgccttgcgagaaagtatccatcatgctgatgcagcgcgctgcatacgctt
atccggctacctgccattcgaccaccaagcgaaacatcgcatcgagcgagcagctactcgatggaagccggtcttgcgacgagatgatcgacgaagagcatc
aggggctcgcgccagccgaactgttcgcaggctcaaggcgcgcatgcccagcgcgaggatctcgtcgtgacctatggcgatgcttgccttgcgaatatcatggtg
gaaaatggcgcttttctggattcatcgacttggcggtggtggtggtggcgaccgctatcaggacatagcggttggctaccggtgatattgctgaagagcttggcgga
atgggtgaccgcttctcgtgctttacgggtatcgccgctcccgattcgagcgcatcgcttctatcgcttcttgcagaggttcttgaGCGGGACTCTGGGG
TTCGCGAAATGATCGACCAAGCGACGCCAACCTGCCATCACGAGATTTTCGATTCCACCGCCGCCTTC
TATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCT
CATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAG
CATCACAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTATCGATaacagcatcttccactttg
atagcgacacgtatgtcgaggtagcctcatccccagggttatatacctgatgaatcgacacgcgactgtatgctcttcttctcagaaatgcacaacacaaagtctt
tgatatgctttctatattctaccagctgtaggtatgtaaccaccgtgaatccaaactttgtgtgtattattatcaactccaatgaacaaatccacccttgtgtg
gtaaaagatgagagtataaaggagctgctgccgtatcgttcttagttgctgtaaacagatgatgtttaaacaattaatagaattaccaaatgactaattaccag
cctgaagtctgatctattgaagaactctctaccaatctcaattgattcagtgcttccactccatctggatattcaaatcctgcatttctggtctgggactccatccactga
ttccttcagttcgtGCTAGCCATATGGGTACC

Isolate sequences:

Isolate 1:

caagtcctacacgcaaatggaatatccattagatttctagatttactaggataagagaggataaaacgtggaaagaatctactcatctaaacgatttagtaaacttgactaa
atcttaatagttacatacaaaactgaaaataaaataacaccatttagtgggtgctgccatggatgggtgtattgtatactgtctaaacgcgttagtaaaacatggcgaggaaat
aaatcatataaaaaatgattcatgattaaacatgttgtgaaagagttgtgaaaaagttaagaacgttcacattggcgacaatctaaaaacaatacagtgattgcagatttg
ccatataatgataatgctgtatcgatgtatgaattcactgtataaaaaagaatgtatcaagaataatccagatttgctaatttgataaagatagatgacgatgacaagactccta
ctggtgtatataattatttaaacctaaagatgttattctgttatcatatctataggaaaggataaagatgtctgtgaactattaatctcatcagacatctgtgtgcatgctgga
gttaaatcatataaagtagccattcttccatggatgttccctttttacaaaggaaatgcacattgattattctctgttggatttctctatcatgatgcagcacctctctaaagaag
tgtaaccgatagcttttgcgaacgggttgcggcagaacacagctgaagcttcgaggggtcgcacatctctctcagcgcccgccctacctgaggccgccatcc
acgcgggttgagtcgcttctgccgctcccgctgtggtgctcctgaactgcgtccgctctagCTCCTGGGCAACGGTACCGGATCCCTG
CAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGGAAATATAAATGGTGAGCAAGGGCGAGGAG
GATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCA
CGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAG
GTGACCAAGGGTGGCCCCCTGCCCTTCGCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAG
GCCTACGTGAAGCACCCCGCGACATCCCCGACTACTTGAAGCTGTCTTCCCCGAGGGCTTCAAGTG
GGAGCGCTGATGAACCTTCGAGGACGGCGCGTGGTGACCGTGACCCAGGACTCCTCCCTCGAGGAC
GGCGAGTTCATCTACAAGGTGAAGCTGCGCGGACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAA
GAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAG
ATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGG
CCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAAC
GAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACG
AGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTGTATTTCGTTGCCCAATGCGGGcGATGTAA
TAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATAcATTTATCTTTTTGACTATCCTCACTcTGAAG
CTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGATATGTTGAATATAGGGATAAACTGGTAGGG
AAAAGTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAAGGATATATAGATGTCAATTACAAAAG
GATGTGTAGACATCAATAATAgGCGGCCGCCTAaactgtttattgcagcttataatggttacaaataaagcaatagcatcacaaaAAAA

AAAgaagtgaaccgataataatgttattatctagacacc
agcgtctacatgacgagctccgagttcaattggtcaagttttacataagataaagtcgactattgttctattatattatggttggtgatgctgtgatgacgatagct
gataatagaactcacgcaattattagcaaaatattagacaatactacgattaacgatgagtgatgctgttatttgaaccacagattaggattcttgatagagatgaga
tgctcaatgatcatcgtgtgatgaacagacattgtattatgatgaatttacctgatgtaggcgaatttgatctagtattgtgggaaatatgaacctgacatgattaagatt
gctcttccggtggttgtaatttaataagaatacagactacattccgggagacgaggatagctactacgtttacggtatagcctctagataa

Isolate 2*:

atgattcgttattgatactatcgttaacgtgtcagtgctacacatcgtctgcaaggatttccaatgccggtatagtagcgataaaaaatttcaaatgataatattgtctct
caccgtttggttattcgtttctatgtttatgtcgtattgcctgcatcaggaataactagaataattattgaagactatggattgagaaaaagagatctgggtccagcattta
cagaattaatacaggattagctaaagctgaaaacatcctaatactacgactgatctaacttatacaagtttcgtagataactgtgtgcattaaaccgtgtattatcaacaat
atcatagattcggcctatatagattaaacttttagacgagatgcggttaataaaatttattctatagtagaacgtagatccggtatgtctaattgtagtagattctaatatgctcgac
aataatactctatgggcaatcattaatactatattttaaggtacatggcaatatccgtttgatctactaaacacgcaatgctagtttactaataagtagcgtgacgaaaac
gggtcccatgatgaacgtagttactaaattgcaaggaaatacaatcacaatcgtacgaagaatatgatgtgctgcctccgtataaggatgctaattatgtagtaccct
ggcaataggtgataatatgaccatttcacagattctattacggctgcaaaattagactattgtgtcgtttcaattagggaataaagtgtacaactttaaaccccaattttctat
cgaaaaataagaggatattaaagtcgtagaccgaatgatggctcctagtagtttaaccagataatgcgtcgtttaacatatgactagggaccattatattataaaatg
ttcagaatgcaaaagatagatgtcgacgaacaaggactgtagcagaggcatctactatcatggttagctacggcgagatcatctctgaaaaactggaatttaatacaccat
ttgtgtcatcattagacatgataattactggatttatattgtttatgggtaaggtagaatctcctaattattgtttatggatagcggggaaggaaactattttattatattgatgggt
acgtgaaatctgaattttcttaataatatttttttaaaatgtgtatatgtgttttgcgtagccatgtatctactaatcagatctattagagataatttaattctggtgcaatatg
acaaaattataaaaaatgaaaaataacactaattagcgtctcgtttcagacatggatctgtcacgaattaacttggaaagtctaaagcagctgaaaagctttctctctagtaa
agatacatttaagcggatgtccatggacatagtccttattatgcaatagctgataaactgcgtcgtctagtagtgcgttgaacgctggagcattgaaaaacttctta
gagaatgaatttccattacatcaggcagccacattagaagatacacaataagcttttgcgaacgggttggccgagaaacacagctgaagcttcgaggggctcgcactct
cttccacgcgccgccctacctgagggccatccacgcccgttgagtcgcttgcgcctcccgctgtggtgcctcctgaactgcgtccgcccgtctagCTC
CTGGGCAACGGTACCGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGAATA
TAAATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTTCATGCGCTTCAAGGTGC
ACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGA
GGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGT
CCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAG
CTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAGGACGGCGGCGTGGTGACCGT
GACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCC
CCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCC
GAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGAC
GCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCCTGCAGCTGCCCGGCGCCTACAACGTCAACA
TCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGG
CCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTGTTA
TTCGTTGCCCAATGCGGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATAcATTTA
TCTTTTTGACTATCCTCACTcTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGATATGT
TGAATATAGGGATAAACTGGTAGGGAAAACTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAA
GGATATATAGATGTCAATTACAAAAGGATGTGTAGACATCAATAATAgGCGGCCGCCTAaactgtttattgcag
cttataatggttacaataaagcaatgcatcacaAAAAAaAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAagaagatacaaaatagtaaagatttgcattcagtggaatggatgattcacaatttgatgacaaggaaacaccgattgtattatgcggttgatagtgtaacatgc
aaacggtagaactgtttgttaagaaaattggagactgatgttctatgggaaactggatggaaaactcattttatcatgccgtatgcttaattgatgaattgtatcatcac
tttcttcagaataaccatctatttgcctgtgctattctccttattgtattacacactataaaaaatggacacgtggatgatgattcttctcgtcactatagacgtcgac
aaacaccaataatcccttcttcttccggacattaaattggctatagataaaagacattgagatgttacaggtctgttcaatacagacattatctactctgttaattcg
gaaatgtactattggatgatccgaaataactaatgattatgaaaaatggttaatacaagctgactcctatacaaaagatctcgatctgcaagaataataaattg
gatgaataattagcaaaaacaaggaaactcagactcatgtacgtcaattgtgtaagaaaaactaa

Isolate 3:

atgaacaaacctaaagacagattatgctggttatgcttgcgtgtaatatgcggtctaattgtcggaaattttttacagcgacactattaaaaagttgtagaacgtaaattagttcat
acaccatcaatagataaaacgataaaagatgcataattagagaagattgtcctactgactggataagctataataataaatgtatccatttatctactgatcgaacacctgg
gaggaaggacgtaatacatgcaaaagcttaaatccaaatcggatctaattagatagagactccaaacgagtttaagtttttaagaagccttagacgaggttattgggtag
gagaatccgaatatataaacagacaacccatataattttatagctaaaaatgccaggaagaattgaaatatttttagcacaacgaatactccaaactgcattcgtgtt
acatatataacaattacactacattttatcataccactactcgtttagatgttttagaaaaataaatatcgcGcttttgcgaacgggtttccgccgaacacagctgaa
gttcgaggggctcgcactctctccttcacgcgccgccgcccactcagggcgccatccacgcccgttgagtcgctgttgcgcttgcgcccctccgctgtgtgtcctctgaa
ctgcgtccgcttagCTCTTGGCAACGGTACCGGATCCCTGCAGAAAGCTTCTAAAAATTGAAATTTTATTTTT
TTTTTTTTTGAATATAAATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTTCATG
CGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGG
GCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCC

TGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCC
CGACTACTTGAAGCTGTCCTTCCCCGAGGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCG
GCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGC
GGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGGCCTCCTCCGA
GCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGGCTGAAGCTGAAGGACGGC
GGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCT
ACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGAACAGTACGA
ACGCGCCGAGGGCCGCGCACTCCACCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGC
TTGCATTTTGTATTTCGTTGCCCAATGCGGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATG
CTCTATAcATTTATCTTTTTGACTATCCTCACTcTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATAT
GGATAGATATGTTGAATATAGGGATAAACTGGTAGGGAAAAGTGTAAAAGTTAAAGTGATTAGAGTT
GATTATACAAAAGGATATATAGATGTCAATTACAAAAGGATGTGTAGACATCAATAATAgGCGGCCGC
CTAaactgtttattgcagcttataatggttacaataaagcaatagcatcacaAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAagaaaaataataatcgccgtaccgttctgtttataaaataacaattaacaattatcaatttttcttaataattttacgtggtgacct
tctgtgtgtaaaataatctttagtgttggaatggaatgctgttaattgttccgacatcatgtatatttgacgtatgcagtcacatggttacgcaatagtcagactgtatgtct
atcatgcttctcatcagaagagggaacagtttaagctcctgtgtttaaactattgcatgatttcatgaaatccttgtttatccacttcacatttaataaatgtccacta
tacattctctgttaatttactagatcgatggtcatggtataggttcgtagtccatggatccaaactagcaaaactcgcgtatagcgtatcgcgattagtgatcac
caacagatgaaaataagaaacagtttaatgatcaacagaaatatttaactcctcgttgatagatagtcgccatatttggatttcggttcacacgtgtgtgtgagg
ggtcgtctagcgttctctacgttaaaactcgaatccatatttcttattgtcagaatcgataccgatttatcatcatcactgtttgaaaactaaatggtatcacatcaaaat
aataaataaacgagtagat

Isolate 4*:

gagaaagagaaagagatagttagcttagatattttcttagtacaagtgcaatgttttaaatatatggacaagaatttctgtgtataaaaactgtgtgaaatttgtacaaag
aaaaaatgtgagcagtatccctacatggattttactagatcatattatatacaaaaaatattatagatctacgtttattatattgatttaacgtgtaaaattataaactattttatg
atatacaattgtctggaacatagatggcataggggatgttgataagctcgacgagtagatgtgttgacgttattgtttaagaaatagttgatgcacagaagagaataa
aaaatatttttagtgagaccatcgaaagagagaaagagataaaaactttttacgactccatcagaagagggttaataattttgtgagaccatcgaaagagagaaagagaataaaa
atatttttagactccattgaagagagaaagagaaaatgagaatgagaataaaaatatttttagtgacaccatcagaagagggttaataattttgtgagaccatcgaaagagaga
aagagaataaaaatatttttagactccattgaagagagaaagagaaaatgagaatgagaataaaaatatttttagtgacaccatcagaagagggttaataatttttagagacc
atcaagagagagaaagagaataaaaatattttgtaaaacttttttagagaccatcaagagagaaagaGAATAAAAAATATTTTTGTAAAACCTTT
TTTTGTGAGACCATCAAAGAGAGAAAGAGAATAAAAAATATTTTTGTAAAACCTTTTTTATGAGACCAT
CAAAGAGAGAAAGAGAATAAAAAATATTTTTGTAAAACCTTTTTTATGAGACCATCAAAGAGAGAAAG
AGAATAAAAAATATTTTTGTAAAACCTTTTTTATGAGACCATCAAAGAGAGAAAGAGAATAAAAAATATT
TTTGTAAAACCTTTTTTATGAGACCATCAAAGAGAGAAAGAGAATAAAAAATATTTTTGTAAAACCTTTTT
TTATGAGACCATCAGAAAGAGGTTTAATATTTTTGTGATACCCTGAAAGGAAATAGGAATAGGAATAG
GAATAGTGTCTAATCGTATCACACTATTGAGACAGAAAAAGAAGAAGTAACGAGAGGTAACCTTTTTTG
TGAATGTAGTTAAGAACATTTTTGTTTTGCAAACCGGAATATAGTGTCCGGTACACTTTTAGACCATCG
AAGAGAGAAAGAGAATAAAAAATATTTTTATGACTCCATTGAAGAGAGAAAGAGAAAATGAGAATGAG
AATAAAAAATATTTTTAGTctttttcgcaacgggtttgccgcgaacacagctgaagcttcgaggggctcgcatctctcttcacgcgcccgcgcctac
ctgagggccgcatccacgccggttgagtcgcttgcgcctccgcctgtgtgctcctgaactgcgtccgccgtctagCTCCTGGGCAACGGTAC
CGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGGAAATATAAATGGTGAGCAA
GGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCC
GTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCG
CCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGT
ACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG
GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTC
CCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCG
TAATGCAGAGAAGACCATGGGCTGGGAGTGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCT
GAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGAC
CACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCA
CCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCGCACTCCACCGG
CGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTTGTATTTCGTTGCCCAATGC
GGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATAcATTTATCTTTTTGACTATCC
TCACTcTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGATATGTTGAATATAGGGATA
AACTGGTAGGGAAAAGTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAAGGATATATAGATGTC
AATTACAAAAGGATGTGTAGACATCAATAATAgGCGGCCCGCTAaactgtttattgcagcttataatggttacaataaagcaat
agcatcacaatttcacaataaagcattttttcactgcaAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAtatttttagtgacaccatcagaagagggttaataattttgtgagaccatcgaaagagagaaagagaataaaaatattttgtaaaacttttttagagaccatcaa

ataatcgatcatctcttctgctatctctgtattgtgtgcgcgtaaatgatgcaaaatgattcacatattggtacactagcatctttactacataatgttgcattctattatacagattt
attaacgattgttgacctctacagttctattactctattaaaggctgaaccgatccactgatggcatatgtttctatcgaacgtatccccctgacaccagtcgaataaatcaac
atcgcatcttccagtgctgtgaacgtctggccaacacgattctaatactgcacctcttcatacttctgatattcttccatcttttccaataatgagtacgattaaaagtgcga
cagcaattgggtgcggttcattatagcattgtcttaacaatgccgaaagaccaccgggtcctgtgtgactaatctaaactctggatattcttcttacttcttcttcttctt
gctagtgtgctctatagcacatattctaaaagcttagcgggagctatcacgtcatgtttatccacgtttaataacatctcatcagtgaggactcccgaggcgatcccg
ttagggagctcaacacttactccgccaccatattatctcattgaaagtattatctaaaacgccataaagatgttgatcttaaaggattgaactctatccgaaaacacatt
cctagaatgttctcgtcattatccattacgattctagtttcaaaaacattgactctcttttgaaatccctcgtagtgttgtagagacgagatagctatttgaagaaaactttttagt
tcttgagaacattcagtcatagaatattccctggaaaatgcatcagttaccaggagcttctcataataattgtcatctttaaacaataagccagatgctgatgctgactaata
cacttgataaagcccaacacattctaaatgaatgacgggtctaccacaacattttcttactactgtgaaaataaacacataagatttctttgatctatacttagacaatag
tagtatctgtcctaataaggcatcagttcttattacaatcgacacttactacatgataactagaaggtttactagatttttccagatcaggttctatctatgatggcatcattgt
gaaaactacacaacacgattttacttggacacaggaagattaaacacaatgtttcggctcctcgtagaacactgatgactgagaggtataatggcccaaatgtttac
agatccgcccaaggcggcaaaaatacattaaactcatccgtcgagctacattatagacacttctcactgaactctgaaaaatagccacaatttggcgagtttatcatt
tttatacggatgctcattttaaattttgtaaatttttaaagttaaattggctgcagaacagcgtcttcaattttgacatagtttcaaatgtatagtcgaactgtattaaagag
atatacttataattctgaatacagagtcgcaaaagctaaacaattgtgctattgtccggcatcgaaaaaggaatcagtgattaatggtatcacaattgttgcgagtcgaaatata
gaaataatggacaagagcagctattaaaatattggacaatcttcgatgtcattcggctcatgtatgtaacgccacagatttctggagactatataattcgttaaacggttta
ctcatactaccgcatctttaaatacatgcaagcccactattctagccacg

Isolate 6:

atgcctcttcttctcctactaacgacctagtttccatattttgatttattatcaaattaatttagtaactgtaataataattatgaattgttccaagaaaaacaattttcaaga
gaaaatttataaaaaatgccgttagaattgttttaacgggaggtatcgtgctgaaaaactatcttattactatctctgtttctacactagttaaaaaataaacatattctt
gtttacaccgtttataaccagattatgatggatacttggccaatcataaatttctgttagtagtcaggaaatcttagaataaatctgatacgaactaaaagtaacataga
aaaatgtattgtctgcacaaaaatcataaaaaatcagcacatttttacttattttgatgatgtaggcgataaactatcaaaatgcaatactctaatagaattctaaacttggga
aggcattaaacacgtctattatttactatgccaaacttatagacacgtaccaatattaggacgggctaacattacgcattttttagtatttcaacttccatctcagacgcggaa
aatatgtctacgatcgtgcctgtaaagggggaaacgaaggatatttaaacatgttgaatatgatacagacagctagatccaataatcagttggctattattatcgaagactc
cgtattttgtgaagggtgaattacgtatattgacggataccgccgataaggacgtcatagaacaaaaggttaaacatcgatatttttagtaaatcaatttgcacatgaaaagaa
tctaaacgctatattagaaagttaaaaaaaccttttgcgaacgggttggccgcagaacacagctgaagcttcgaggggctcgcactctcttctcagcgcggccgcct
acctgagggccgcatccacgcgggttagtgcgttctgcgcctccgcctgtgtgtgcctcctgaactcgcgtccgcgtctagCTCCTGGGCAACGGTA
CCGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGGGAATATAAATGGTGAGCA
AGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGGCTC
CGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACC
GCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCAT
GTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCTTCCCCG
AGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTC
CTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCC
CCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGC
CCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAG
ACCACCTACAAGGCCAAGAAGCCCCTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACAT
CACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACC
GGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCAATTTTGTTATTCGTTGCCCAAT
GCGGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATAcATTTATCTTTTTGACTAT
CCTCACTcTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGATATGTTGAATATAGGGAT
AAACTGGTAGGGAAAACTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAAGGATATATAGATG
TCAATTACAAAAGGATGTGTAGACATCAATAATAgCGCGCCGCCTAaactgtttattgcagcttataatggttacaataaagc
aatagcatcacaAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAagaaggtaaaaaaca
attgtgcaatagcgaatcatcgtcgtcaaaaaatgtatcatcataattttaaacttttgaacaatagtcctatttagagaaagtctatcgatagacgatccaatttgaatt
actccactactgtcactccatgatataatccatgtaaaatgacatcatcttttagatcattattgttaccttcccaatacaacaaatcatgatataacctctccag
acaagtatttaaacacggtagaatgctttggcttataaaatacaaatgacattcccttatgtttaatcttaacttttctttagttattgaatcgttacaattataaatgatgttttcc
aaaaacctaaagtgtatttaaaatagatgccat

Isolate 7:

atggagggatctaaacgcaaacacgatagtcggcgaccacaacaagaacaggagcagcctcgtccacgtacaccgccatcatatgaagaaattgcaaaataggacac
tcatttaacgtgaaaaatgattacgaatgaagaaatgtgtcttaagaatgattatccacgaGctttttcgcaacgggttggccgcagaacacagctgaagcttcgaggggt
cgcatctctcttcacgcgcccgcgacctacgtgagggcgccatccacgcgggttagtgcgttcgctcgcctccgctgtgtgcctctccgctcgttcgctcgttcgctcgt
ctagCTCCTGGGCAACGGTACCGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTG
GAATATAAATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAA
GGTGACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCC
TACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACAT

CCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACG
ACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAA
CATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAG
GGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTG
TTATTCGTTGCCCAATGCGGGTGTATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATATA
TTTATCTTTTTGACTATCCTCACTTTGAAGCTATCTTGGCAGAGAGTGTTAAGATGCATATGGATAGAT
ATGTTGAATATAGGGATAAACTGGTAGGGAAAAGTGTAAAAGTTAAAGTGATTAGAGTTGATTATACA
AAAGGATATATcGATGTCAATTACAAAAGGATGTGTAGACATCAATAATAgGCGGCCCGCTAaactgtttatt
gcagcttataatggttacaataaagcaatagcatcacaAAAAA.....aaaa
atcaatcgcttataagaccatcagaagagggttaattttgtgagaccatcgaaggagaagagataaaactttttacgactccatcagaagagggttaattttgt
gagaccatcgaaggagaagagataaaactttttacgactccatcagaagagggttaattttgtgagaccatcgaaggagaagagataaaactttttacgactcca
tcagaagagggttaattttgtgagaaaggagaagagataaaactttttacgactccatcagaagagggttaattttgtgagaccatcgaaggagaagagataa
aactttttacgactccatcagaagagggttaattttgtgagaccatcgaaggagaagagataaaactttttacgactccatcagaagagggttaattttgtgaga
ccatcgaaggagaagagataaaactttttacgactccatcagaagagggttaattttgt

Isolate 10:

atggatatctttaagaactaatcgtaaaacacctgatgaaatgtttgatttccagtttctattttatctactttatctataatcatggagcagctggttctacagctgaa
caactatcaaatatagagaatgatgagaatacaccgatgacaataatgacatggacgtagatattccgtattgtgcgacactagctaccgcaataaaatatacg
gtagcgatagatcaggttcacgacctctctacaaaaataaaagacgatttcaactgtaaactttaataatgctaacaaacaaggaactaatcaacgaatgggtt
aagacaatgacaatggttaaattatcttattgactagtcctcattacataaactcgtatgacagttgttagcgccgtcattttaagcaatgtggaatatccatttct
aaacatctacatatacagacaagttttatattctaagaatagttaccagttgttatgatgatgtgggtaccgagaataactgcaatatgtacatattaatgaattatcggg
ggattctctattatcgatattccatagcagggaactctagtaggtgattatactaccggacgacatagaaggatatataacatagaaaaataaacaagatgaaaattt
aaaaatggtgtgtattctactaaaagtatacgtgtatagcgaaggttaaggtggaatgacagaaaccgtataatctggtaccgattttagaaaatttaggactta
ctaataattcggatattatgcagattttagcaagatgtgtaataaactatcactgtagaaaaattctacatagcagctttatagatgttaatgaggagtatacagaagcatcg
gccgttacaggagattttacgattaaactttcgtggtatctgtacgaaggtctacataaaccatcattcatgtacatgattaaagacaccacaggacgtatacttttatagg
gaaatactgctatccgcaataaataaacaataagacttttatcagctttatctatgtctaaatattacaataagtaatagtataaaactaaagctgataacttaaaaaataata
atatcatttacaattaatagataaaactaaaaataaacaatacgtttattataagtaatacacaatgatgatatacggattaatagcgtgtcttatctgtgacttcatccatcgct
agtccaccactacggaagataaactcgttaaatgtgtagaggattattgttctgttttagatgacaaaaagactatacggtaacttctcagttcaataactacactatcg
atacaaaagactggactatccacacctgatggtttggatataccattgactaatataactattggtcaggtttactataggtcgtgactgttcaaatcagagtctgaggat
tttcaaaagaaaatgatttctaggtgtttctatagaatgtaagaagtcgtcgacattacttactttttgaccgtgcgtaaaatgactcgagttaataaatttcagatatg
gcttattatcgaggagactgtttaaagccgtttatgtaacaatgacttataaaatactaaaactggagagactgattacacgtacctctaatgggggttgcctgcatacta
tcgtaatggggctgatggttattattgattatattctgttttgcgaacgggttgcggccagaacacagctgaagcttcgaggggtcgcactctctctcactcgcgc
ccgcccctactcgaggccgcatccacgcccgttgagtcgcttgcgcccctccgctgtggtgctcctgaactgcgtccgctctagCTCCTGGGC
AACGGTACCGGATCCCTGCAGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTTTGAATATAAATG
GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGG
AGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCAC
CCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTC
AGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCC
TTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCA
GGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCG
ACGGCCCCGTAAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGA
CGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAG
GTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTT
GGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCAC
TCCACCGGCGGCATGGACGAGCTGTACAAGTCTAGATCTGCCACCATGCTTGCATTTTGTATTTCGTTG
CCCAATGCGGGcGATGTAATAAAGGGCAGAGTATACGAGAAGGATTATGCTCTATAcATTTATCTTTTT
GACTATCCTCACTcTGAAGCTATCTTGGCAGAGTAGTGTAAAGATGCATATGGATAGATATGTTGAATAT
AGGATAAACTGGTAGGGAAAAGTGTAAAAGTTAAAGTGATTAGAGTTGATTATACAAAAGGATATA
TAGATGTCAATTACAAAAGGATGTGTAGACATCAATAATAgGCGGCCGCCTAaactgtttattgcagcttataatggtta
caataaagcaatagcatcacaAAATTCACAAATAAAGCATTTTTTTCACTGCAAAAAAAAAAAAAAAAAAAAAA
AAAaagagtaacagctgccccattcttaataatcgtcagattttaactgttaaatgttggtatataacatctacctattttcccgagtataaggtttgtgcaggtatact
gttcaggaatggttacatttatacttctctatagtcctgttcttgatgttcatcacatagcaagaacagaataaacaataatgtaagaataatattaaatctgtgaatt
cgtaaatacattgattgccataataattacagcagctacaatacacacaatagacattcccacaggttgccattacccacgatacatttgagtactaagcaataggaata
actaagctagtaagaggcaatagaaaagatgagataaatcatcaatatagattagaggggctatatagagccaagacgaacaaaatcaaacaggatgaacgttct
aacatcatttttgaagattcccaataatcattctctccataatcgttttgcatacactccatcttaggcataaacgattgctgctgttctctgtataaataatctttatca
agcactccagcaccgcagagaagtcgaagcatattgtaatatcttaataactcat

*Note that some isolate 4 viruses carry a second insertion that precisely matches that of the isolate 2 insert. We believe that this is a result of recombination between two separate strains during plaque purification of both.

Nanopore reads showing K3L-CNV:

Isolate 4:

TGTTGATAATGGGAAGATGGGATATATGATGTTGCATGAACGTTGATAGTTCACTGGTCATCAATATG
CCGACATTCAATGCCAGAATCATGGTTAATTTTCGCCAATACATACCACGATACTTCCAGCGTGAGAA
GTCGCGTAAACAATATCCTCATCTCCTCTGCGGACGCTATTAATTGTGATAGTTGAATACCAGCCAGCG
CTTGTAATTTTCGGCGTCGTTAGCGCAGTCACTGGTTTGAATGATAGCAGCAACGTCAACATATTTTCATC
ATTTTATGAGGATTCGTGCTATGCCAGCCAGCGTGTGTTTATAACAATATTTTCGTAGCCCATATTATTA
ATTCAGTAATATAAACTATCAATATTGTGTTACTGTAAAATGTCATGCGGTCAAATGTTATTAGTCGTA
CACGCATGCCAATATAATATAGTCGTAATTATGCGCAACACGGAAGATATCACCATTTGTCATTGTGCG
CTTCATCATATCTTGAGCTTAGAACATCATCAATCCATTTTCATGCCGGTGACACGGTATGCACAAGCGA
TTGATTGCACGCAATACAACAATGAAGTATGAATACGGGGTACAAACAACCGAAATTTCAACATTACT
ATTTATGATATACACGTTATGTCGCCATTGCTTCATACCATGGATTGTGAATCAATCGTAAACAATTTT
CTGCTTATTTTTTATTACTTTTTTATTCGTTTGCCCATATAGGAACAACATATCGTACAGCATAGTCGGCGA
ATGGCATCGATAACGACGAGTCATTTATGCAGGACGCGCAAGATATTCGGTACAGCATCATACGCAAC
AAAGCGTACGGGCGAAATAAAAGACATTTTCAACCGGAACCTTACACGCTACCAAACAAGTGACCA
ATCAATCTAGCGGTAAAAGTCCGCCATTCACGTATGACAGCCAAAGGCGCCGCCCATTTTGGGTACGC
TTTCAAACAGCCCACATAGTACAGACAAGTTTTTATATTTTCAATGAGTTCGGAAGGCCTCGTGGG
TCATGCAGTTGAGTATTACGACTGAGTCAGTGACACTTTTATTATCGATACCATGAACCTCAGTATACG
ATTATTTATGTAGCTGCGATCAAGATGAAAAATTTGCGTTCGGTTCGGTATGTTAGACTTCGTTAAAAGC
TTGGAAAGTGAATGACAGAATCGCATGGGCGTCGCATTTTGAATTTAGACTTTTCAATACATTCG
GACATTGGCGCATTTGCTAAAGATGCGTAATGTCACTGCAGTAGGCCTTACATACGACGTTACAGATG
TTACGAGGAGTGACATCGGCCGCCACAGGAGAGTTTACGATTAATCTCTTTTCGATGGCACTACGTACA
TGATTAAGACACCACAGGACGCAAGGTTTTATAGAAATACGTTATCCGCACAAATATGTTTCAATAA
AATATTTTCATCACGTTTGTCTAAATATTACAAATAAAGTGGGTGAGTACGCAACTAAAAAGCCGACAAT
GGTTTCAAAAATAATATCATTTTTACTACCACAGGAACTGTACTAAACAACGTTATTAGTAATATCA
AAATGATGACATACGATTATCAGCGCAGACTTTGAAGATGTAACAAGTCAATAGCGTAGAGGCATAGT
TTCGAAAGTTTGAGATTTATCAAAAGAATATACGGTAATCAGTTCACAATCACAATTGTGCGATAAATA
AAGACTGCCATTCATCACCCGACGGGATGAGTGTATCGACTGGTCAGATAGCTGGTCACGCATTATA
GGCCATCATTTCGAGTTTCCAAAGAAATGACCAGGCATAGCGTAAGAAGTCGCACGACAGGACTCGAG
TATAAATTATCGCGCATAGCCAAGTTGTTACTTATCGTATGCTCCATTTGCCGACATTAGCATATGCC
CTACTCACACAAAAAGAACATTTTTTATAAACATGATCTGAGAGGCTTAGAGCAATCCAGCATGTTTA
ATATGCAACATTTTCGCTGTTTCATAAAATGAAGGTATAGTCATTAATATGACGCTTAGTTTACATCGGC
AATCATACCATTTGCCCCGGCCATACTATTTGGACAGCATTAGCGTTTCGTGGCAAAGAAAACATGCATG
GCGACATTGGACTGGTTATTATCAATAAAGAAGAGTGTAGTGTCAAAAGTAAAGTCAACAAAAGTAA
CGGTATCGTTGGCCGCTGAGCTTGACTGACAGACATACGAGGGTGTAGAAATGTCCATTAAAATATGT
GACATCAACGGAACATAATATATAAATGCACATTGCCGTTGGCATAATTGCCGATTTCACAGAGTTAA
AGTTATTCAATATGAACCTTACGCGACCATACTGTCGATGCATACACTGTTATTACCACGAAGAATCAT
GCGACTCTCTACATCTCATCGATGATAACATGTACAATATTTCTCGAATAAGTCTTTTAAATAATATGA
AAAACATACATGCAGATGCAAGATGTTTAAATGATGATACGATTTTCATCTTCAGCGAGAAGATGTGCTG
CCAGAATACTTGTGATAACCGTTTAAATAATATCAATTTAGTCACGATAATACAACATTTGTTAAGTTCA
AACTTTCCACAGCCAGACTGCTTGTAGAATGCCGTGACATAAACAGTCCAGTAGAAGGTACATTCCCA
CACACTGTTAATTCGCTGGTCAATTATGATTTGACCATTTTAAAACCTTTCAAGGAGTATCAGCCAC
TCATTATCATGATGAAGTTGCGAATTATATGCTTTCTCTCGATAGTATAGATAGCTTAGCTAATGTATA
ACATAATTACCATAGCAGTTTACCGATTTAAAACGATAAACTATTGATTACGCACTGTATAAGTTGC
CGGGAAAAGGAGCAGACCCCAATATGTAGATTATAGAGGGCTTCTTCTGCCACGCCTCGCATCTATAT
GTTTATTCTCATGGAAAAGATGGTTCGAGTGCCATTTGTTGTAAGAGTTAAAACAGAGTCTGAAAGT
ACCTCATAACTACGCACTCCCTCAATATTGCTCTCACTTCTCATCTCACAACCCACAACAAATGGAA
ATAAAAGCCCTCATTGGGCATAAACGCCTGGGTTGCCTTCTCGCTGCTCAGACCCCTCGCCAGCCGGC
GACCCCTACGTGTGCGATGTCTGACTGGTCACACCTCGTGTTTCATAGAGACCCTGCAGCTCCTTCAGC
CTCATTACAAACGAGCTTGGCGGTGAAGCCGAGTACTGCGGACAATTGCGTTATTAAATGTTTCATCA
GAATGAAACCGCCGAAATATTGCATCATTGACTTTGGTTCATCAATATCGTTCCCGGCCAACGATGGGT
ATAGTCGACATACATGCCGAGGGTAATCCATAGTCATGCCACTGGGTTGACTTTTTTGTCAATTCATGAG

ACGATACGCTTTCGTGAAATAGAAGGACAGCGTTGGCGATTATTGCGGAAAGCACCAGCTTCATGCGT
CGATGAAGATGCAAACCTGAGCATTGAAACGTCACGGGGTCTCATCGAGCAGACAGCATCTATAAGTC
GGTCGCACATGTCTCCACACGCGATACAGACAAATTCGACTGCCTGCGAGTCGAGAGAGTCAGGAAA
CAACGAAGCGATAACTCGCCAGCGATGCCAGTGTTACAGACATGCAAAATTCAGACAGTACTCCACG
ATGTCCGGTACATTTCAATTTCAACGAATGATTATGATGTCCGTTTCAGAAAGATGCGAGGCGCTTAGAA
AATTCCTCATATACGAGGAAACGTTAATCTTCTTACGTCATGGTTTCGTGCGATGCAATGCGGTACAGCC
ACATCTGGCAAGGTGCGGTCTCCACTGGCTTGGTCATTTCAAGAGATGAATAACAAATCCAGTTACGT
TGATAAACTACTGATGCGACAGGGTCTGTTTCAGATGGACGACAACAATAATTTTGTAAATTAACGTA
GGTTATCGCGGAAATCAGATACGCGGCTAAAATCACACAATTATGCTATTGTGCGAGATTTACTGCAAG
GTCACTGTTATGAAACCATACTACGGCTTGATATACCACAAGTACATGGACGATGACATGACGACAAT
AAAATCATTATAGTTACGCGGTGCGATGTCAATAATATACAATTTCAAAGCGGTTGGTTATGGAGCATC
AATGCCGATAATAACAGCGTCATTACAGTAGCAGTTGCCAAAGAGAAAAATAAAATCTATGGAGAAGC
CGATTTTATTTACAGTCATCCAGAGACAGACGAGCCTGGGTCGATGCGCATGCTGCCTATATCTCGA
AATCATTCCGGTATTGTACGCCCAGATGATTTTTGCAAAAGATATCGCCGGTCATAACTGTGCTATAG
AGCAGACATTAATAGAATGAAGAATGCCTGGCTATTGGTCAGGCGTTCCATGCTTGATATATGCTTAA
ACGAAAGAAACGCCACAGACTGAGGATACGCAAGCATTACATCAAAATGGTACAGTAATCCACCCAAA
GTCCATCACAGAAAACCGTACACGAATATTACAGACGATATTCTGATTTTCGAAATTTCCACATTCAT
TGCCATCACGCGGAGCGAGCTGGTCGTCTCAGCCAGCTTCATGATACAGTTGAACGATGATAAAAGAT
ATCGGCATCATAAACACGGGAGTTTGTATATGCAAAATCCGCATAAACATTAATAAAATATATTGTA
CCTGTTTCTTTCTTTTCGTACATCTTCTACTGCGAAAAGTAAATGATGCAAAGGAGATGGGCGATGAGA
TTGAATAAAACCGCATTGCGTTATACCTAACGTAGCAGGACGATGTACGCTTTCAATGTTTGTCAACAT
AAGGGTTTACGTTTCGATGTGGACAAAGACCATGTTTAAAGCGTTGTTTATTGCAATAAACAAAAATCCTC
CATTTACCGCCATTGTTTAAATGCCACGGTTTGGTCCATTTCGTACAGCATTGGCACAATAACAACAGCA
ACCCGATAGCGATGCAGGTGCCACTACTCGAGCCCATTGTCATCAATGAGCTGTCAAAGGTTTCTTTTCG
ATACACGACGTTCTTTAAACATCGATCTGTTGCTGCAAGTTATTTTACCGTCGGGAAAGCAAATTGAT
TCATTTAATTGGCTCTTTTATTGGTCATTGCAAAACACATGTTACGTTCTCGTGCTCATTGTTGTGTTA
CATAGAAATTGCCGCGTGAGCAAACAGATAACGTAAGTCAGCACTACCTTCAACAGCATGTCAAACAT
GTTCTTCGCATGCTCTTTTGGCCGAGTGACAGCGGCCACCGTGCCAATTGTCATGTAAGTTACAGCTCT
CTGCTCGGCATCAAAGGAAGAAGGACACCCTTTGCCCATCATAAAACATATCGCGATACACTCAATA
TACCACGGCATTCGCCAGAGAAGACGGCCACAATACGCTCAAAATTAAATATACATATGAATATCAA
CCTTCCGATGATCGCTCGACTTTCTGAACAGAAACATCGTTATTAAGCAAGTGATGTTGTGACGACAC
ACACAGACTGTATTACCCATGATTAAATGTTACAAGATGCTCAGAGTTCTCGTCGATTACTGCAGCCG
AAGTCATCGTTTACAGTGATTATCATTTTTATAAAATTGCTTAAATGAGATTATAAAAGTGCAATTATTG
AAAATATCATTTCTCGACAAGATAGATTCTGCCTGTCGCGACATAAGTCGCTCAAAGTTATATTACCCCG
TATGCTTGCTTCAAAGTTGTCTTGAAATATGGGGACAGAGTCAATATTACTTCTTCTCATACGACGAT
CCGGTCAGTGTCATCTGTGCGATGGAAAGCGTAAACCCAAAATGACAAAGTTACCAATAAATTGGGAA
AGTTCACATCTGGCCGTATCATGCGCACACACCAAAGAAAGGCAGCTGTATTTGAAAGAAATCATTGC
GTCATTATTATGACAGAAGAAAATGAATTGTTGGAGAAGTAAAAGACAATATTACAATTATGAATAT
ACTGTATATTGCTCTCTCATGCATATCAAAGACAAACATCCGATGGAAAGTGGAAGACAATTTAACAA
ACGTTACGGGTTTCAGTCTGGGGAACATGCTTCCAGACAGCAGGTTAAATGGTCGCCAAAGCCGCTGTA
ACACTGCAATAATAACCGTTGTCGCCGTATGGAAGACGAGTCAAACATCTTCAATAAAGTCAGGCCCTC
CGTGAGCTCTCCCCGATATTGCCCGCATGCTTTCAACAAAATGTGAAGGATTACGGAGAACAGCCCG
GCTCGTGTTTCGGAATATCAAACCTTCTTCGGCAGAATCACAGAAATGTTTCCACCATAACGATGCGGAAA
TTAACCAACGCCGACGACACTGGCGCTCGCGACCACAATGCCGGAATTATTGCTTTGCCAAAAGATC
TCAAGAGCCAAGATGGATATTCTCAGGATTTTTCACAGGACCCCGGAATTTTCGAGTCAAAGGCAACAAA
GATTCTTCTCCAAATAATCTAGTAAAGACAGAACATGTACAGCAGGAGACTGCCTATCATTACATACTAT
GCGGGTCAATAATACTACATTTGAAAACCGATACAATCTCTTACTTTTCTATGTCGAGGTTATTATGCT
CGGTATTAGCGATGACCCCGCTATAGTAAAGTTTTCACCATAACAATAACAATAAATTTCGTAAAAG
CAGAAAATATATTCTGGTTCTGTACGGCAGCGAAGCAGAATATGTTAACAATAAAGGGCCGGCGAAA
GAATATACGTCTCGGTGATGGATGTTTCGCAATAGCCTCGCAAATCATCTCATTCTTCTTTTCGCATGG
AAGAAGAAAACAAACATCATGAAGGGGAACGATGTTACATCAAAAGTCACAGTCACTTTATCATGAT
AAGATTTGCCGACCGCATAGAGTTCTGAGTCAGTTCAGATCTCTTTATAGTGGGAACAGAAGATGATC
AACATCAAAATGAGATTCTCATCTGATACCATCATCGGTGAGGTTTCAGCGAACGATCACCAATGT
CCAGACACCATTGAGAATCCGCCGATTGAAACACAGTGAATGGCACAGTCTGGGTACCTCGCTTTCTC
TCGCGATCACAGAAAGAAGTTTCGCTGTCAAAGACTGTGAGGTGCTCCACGACACGACAAGACAGC
GACACCATTGAGTTCAACATCATACAGTGAAGATCGGTTCAACATCGACAAAATGCATCAGAATCT

CAGAAGTTTGAAGTAGACAGTAATCACCTGACCTGAGTCGGCGATGGATTCTACGTCGACGATCGGGC
 ATTCGAAGGCAACCGACTCGTCGATCAACAAAGTGTGTTTCGACGATAATTTGTTTCATTCGAAATTGGA
 TAATCGAACGATCAGGCACACCACGAATTGCAAAATTATATTTGCTTCAAACAAAATGTTCTAATTGC
 CATCTTCACAAAAGTTATTTCTGTTACTTCTTTTCTGTTCAATAGAAGATTGTCAATTATTCCTGTTCCCC
 TATTTCTTTGCTTCTCGGCCAAAATATTAAACCTTCTTTTCGATGGTTTCATAAAAAGCTTCAAAATA
 GTTTTATTCTTCTTTGATAACTGATAGGGTTTTATTGTGTATGATTTTGATCCATAGGGAAGTTTGTAAC
 ATTTTAAAGTCCAGGGAAGTTCTTGACAGTTTTATGTCTCTGATGAACCATAGGGTCTTTATTTTATCTT
 CATAGGCCTATGTGACAAAATATTTATTCTGAGTCCTCATTGTAACGTTTATTTGAATTTCTGATTGCCG
 TGAGGTCTGCTATTGATAAACCATTGTCATCCCAGGGACTGACCATAATTTATCCTGACCCCTTATGTTT
 TCCCGTCATTTGTTTCATATCAACATGTCTATTTCATGGCCTGACAAACAAGATTAAAGTGAGACAGTCAT
 AGCATAATACTTCGCCTTTATCACATCACCGCACAAAGGACGAATGTCGAATGCAGCATGGCGGCAGA
 TCTGACTGCAGTTTCGTTTCATGCCAGCGTACTGTACGTTTCATCATGGTAGTCTCGCCGCGGGGGTGGGT
 GACGTCGACGCCGGGCAATGGCTTCGGGTTGTGCAGACCAGCGTCGCAGTGGCCGCCGCCCTGTTTTTC
 AAAGTTGACTGCCCTTTCAGGGCTGACCGAGCCTCCAGTTTCATTGTTTTCCCGCGAGTTTGTAAGTCGT
 CGGAGGAGAAAACAGCAAAGCCTGTAGTAATCCGTCCTGCAGGAGTCTCGGGTTCGCGCCACGCGCC
 CTCATTCTTCTCGGTAGTTCGGGATGTGCGCAGCAAGCCACGTAGTTCGAGCCGATTACGAACAT
 GAAAGGTGTTGGCCAGTTTTCGCGCAATTTGCCAGGTCTCGCCCGCTCGCCTGGGCGCATTTTGTATGA
 TGGCCATGTTTTCTCGCCCCGTACCACTTTCATATTCAAAAAAAAATAAAATTAATTTTATAGAGTTTT
 TCTGTAGGGGATTTCGAATAAGCAGTTCAGGAGATTTATGATGAGCCAGAGGTACCACAGGGGAGGT
 GGCAGAAGTTAGTAAGGGCGGGCAGAGATGTGAGCTCGAGTTGCCGTGTTCTGGAAAGTGCAGTAAC
 CCGTTGCGAAAAGACCAATATTCTCATTATTTTCTTTCAATGGAACATAAAAAATATTATCTCTTCGATA
 AAGCGTACCGGACATCTGCTGCTTCAAACAAAATGCCTTTCATCCACATCACAAAGTTGCCTTCGTCAT
 TTCTTTCTCTTTTCTCTGTTTCACAGGTGATGAATATTTTCATCAATTATTCCTTGCCCTTATTCCTGAGCC
 TCTCAATAAAACATAAACCTCTTCGTGAGTCCAGGTTTTTATAACCAGCCTTATTAAAGATGTTTCAGGG
 AAGTCTATGTTTTATGTTGTCTCATTAGGTTCTTGATATCCATAGGTTTTTATTGTCCATAGGGTTTTGAC
 AGGGAAGATGGAGGAACATGAATTCTGATTTATGTCAATTAATAATTTAACAGCGATTTAACTGACAT
 GGTTCACCATCGATCCTTTTGTAAATCGACACTATATCCTTTGTAGCATCAACTTCAATCACTTTACTTTG
 GCTGTTTTCTCATAACATAACCCACCATATGCATCTAACATTTTCAGTGCAGACAGTCGTTATCATAGCG
 TATTGGCTGGCCCTTCGTACTGCCCTTGATACCGCATTGCCAAATTTAACAACAAAATGCAAAGTTTGC
 TCGTACAGTCAAACACGGTGTGATGTAGCATGGCGCCCGACGATGTAGTTCCGCCGTGGAGACCAATC
 CGATGCCGATAGATGCCGGGCATAAGCCGCACGCTTCAGCACCGCTCGCAGGCCGCGCTTTCTCCT
 CGCTCAGGTGTTCCGTTCTCGCCTCATTATTTTCGACAGCGTAGTCAGTTGTGCGAAAAGTCACCTGTAG
 ATGACCGTCTGCCCGCAGGGAGAGCAAAACAATGTCTTCGAGTCACGCGCCCTGCCTCGGGAAGCGA
 CAGCAATGGGGGATGTGCGCGGGCAAGCCATCTGGAGTCGCAATACTTGTGAGGGGGACACAGGCTA
 GCAGCAGGGGGCATCAGCTTGCAGCCAGCCTGCAGGGGGCGGCCCTCGCCCTCGCCTCGATCTCGACC
 ACCATTTGTGACCTGAGCATGAACCTCCCGGATGATGCACATGTTATACCTCCCCGCCCGCTCAAATT
 ACTATATTTCAAAAAAAAATAAAATTTGGTCTAGAAAGTCAGCAAGACGGGACGCAGCTCAGA
 GGCATGGGAGGCAACGCGAGCCAACCGGTGGATGGCTCAGTAGGGCGGCGGGCGCGAGCCCCCTCGAA
 GCTTCAGTTGGAAGAGTGGCAACTCCGTTGCGAAAAGTTGACTGCCTTTCTTCTCTCATCGGAGTCAT
 AAAAATATTTTTGTTCTTCTTTTTCGATGTTCAAAAATGGACATCGTTTGCAAAGTGGGGTCAAGCCTA
 ACAGATGATCATGGACACTACTCCTGTCTCACTTTCAGTATGTTGGAAGTTATTTTCTTATTGAAT
 AGGAAAGTTTATTCTGATCCTGATAGAAAGTGCCATTTTATTCTGATGAACCTGATAGGAAGCTTATT
 AAATATTTTTTAGTTAGTTTTTGATTATGTGGATGATAGATAGGGTTTTATTAAATATTTTTTCTTCTTT
 CTGATCCAGGTTGTATTTTGGATGAACGTATTGAACCTCTCTGATGGTGTACCAAATATTTTGCAG
 ATGAATTTAGGACAAGTTATGCTGTTTGGGTCCATAGTCAGTTTATGATTGTTATTATCCTTAGACTGA
 TCTGAACCTCAGGGATCCTTGCCCATCAACCACATGTATTCAATTCGCCAAGATATGAGATAGTCGTCGA
 GGCAGAGCGGTCTCTGCCTGCCCTTGTGAATACCGCATTGGGCGCATAGTCAAGCATGGTGACAGAT
 GCCGGTGGAGTGAAGCTCATTTTTCGTTGTTTCTCGTTGCGGGAGATACCACGCGGCCGGGTGCCGCAC
 TGTTTTCTTGCTCCGCAGGTGGTTCGACCAGCGTCGGGCGGCCCGCTCCTCAGCTGGTCGCTGATCTC
 GCCCTTTCAGGGCGCCGTCTCCGGGGTAGATCTCGGAGAGTTTCGCTGTAGTTGCCGCGTAGTTTCATCT
 GTATCCTCGTCTCGCATGAGTCATGCTCACTCGAAGTTGCACGCGGGTGCTGCCGGCCTTTGAGCCG
 CACATGACCGAGGGGACATAGTCACCTCGTCGGCCTGCGCTGGGGCGGCTCGCTCTTTCGATTTGATG
 AACTTCTGATGATGGCCATACCTTCGGTCTGCTCACCCGCCTATATTCAAAAAAAGGGCGCTTGGGAT
 TCAGGTCACGTCTGACGATGATTTAGAACAACGGGAGAGCAGGTTATCTTCTCGAGTTTTCTCAGTTGT
 TCGGCGGTAAATCTGTTGGTGAAAAGACTTTGTGTCATGCTTTTGATTATCTTCATTTTCTTCTTCAT
 CGGAGTCGCCTTTTATTCTTCTCGCCGATTTGTCTAAAAGCAGTACCGACGATCTAGTTCTCGGTTTCAT

AAAATTCAAGACCATCTCACGTCTGCGCTTCTTTCTTTTCAATCCGTGATTATGATCATTTCATTCTAG
TTCCTGTTCTATTTTCTTCGTCAAATATTAAGTTTTTATAAAAAATATTTTATTCTTCCTGATGAAACC
ATGTGTTTTTGCAATATTTTTATCTTCTTCTTCCTTAGATGAGTTCATAGTGGCCACAAACATTTT
TGCCCTTTTGATGACTGATAGTGCCTTACAAAATATTTATTCTTCTTTTCTTTCTTGATGAGTCTTCTGA
TAGGGTTTTATAGGGTTTTATTCTGATCTCATAGGGGTTTATTAATAATATATTCTTTTCTGATGAGCTTA
TGTAACATATGGATTCTTTTCATTAGTAAAGTTTTATTAATAATATTTTATTCTTCTTTTGATGTTTCATAA
AGTCTTTATTTCTTGATGTCTCATAAAATATTAGTCCTTTCTGATGGAGAAACACTAAAGTTTTATTTTA
TTCTGCCACTGAGTCATAAAATATTTTATTTTGGATCTCATTGTGGGCTAGTCCCTTCTCAAAACATTGGTA
ATATTTATTTTCATTCTTTTGACATTTTCTTTTCTTTCCATAGAGTTAAATATTTCTTTTGTTCTCTTGATG
GTCTCATTGTTAGTCCCTTGATATTTAAGTTTATCTTCTGATCCCCACTGTAATATTTTATTTCTTCTGAT
GCATCAACCATTTCAACATCCAACAGCTCCACGATCAACAGCCTCGGCTATCTATATCACTGCACGTCA
ATGCCGGTTAAACATATAACGTGATCGCATATGGCGGCGTAAATGATTTCAGTAAACCACGGGGAGA
CGCTCACGGCCTTTTCTTGGCGCTGTGTCAATGTTTACACAGCATCAAATTTGTCCGGCGATACTAAA
ACGAGCCGACTCGCGCCATGGAAAACAGCCCAGACTGGTCCGTCGATGGTTCATGCAAATATTGAAAA
GTTTCTTCTCGATATTTCCCGTCAGAGGCGAACCTTCTTTTCGATGTCTGTAAGCCTATTTTTCGATGG
TCTCGCCAGTCTTTTCTGGGATAAGTCTCGAAGATGGTTACAGTTGAGGCCAACCTCTCGGATAAAA
GCATTTCTCGATGTGAGCCCTCACGTAAATATTGAAATTTTCGATGGAGTCGTAAGCCTTATCTTCTC
GAGGAAGTTCTCTTCAACCTTTCGATGGAGTCGTTTTATTCTTTTCGATGGTTAAATTCGATGTCTG
GCTGATGTCTCGTCAAAAATATTAACCTCTCGATGGAGTCGTAAGTTTTTATCTCTTTCTTTTCGATAAA
AATATACATCCCTTTCGAGTGTTTTATCTCTTCTTTTCGATGGTCTCGGTCAAAAATATAACTCTTTTCGAGT
CGTAATATTCTCTCTCGATGGTCTCATATTAACCTTCTTTTCGATGTCTGCTTATCTTCTTTCTTCATTT
GTCCCTCACATATCCTGTTGATAGTCAGTTTTATTTTTTGATATTAATAGTCCTGAGTCAGCTTTATC
CACAGCTGAACCTTGATGATTGTCTTTGTCTCACAGTTAGTCCTGAGTCAGGACCATTAAATCCTTTCTT
TGATGTCAAAGTGATCTTTTTCTGATATCCAACCTTCTGAGTCGTGTTTTATCTCTTTTCTTTCTGTCTC
ATTAGTCCTTGATGAAAGTCAGATCTTTTCTGATTGAATCTCATTGATCAAAATGACCTCACACAGCAT
CCCTTTCGAAACCATCAGGGTCGTGTTTTATTTTCTCCTGATTGTCCAAAATAGTTAGTGATCTTGATGT
TCATTGTATAGCTGATGGTCTCACCTCTGATGTTTTATCTTTTTTTCTTCTCACAGTCCTTTTTCTGATG
AGTCAGCTTATCTTTCTCATCAGGAAGACAGTCCTCTTTTCTGATGTTCCACAGAACAGTCCTGATGTG
GGCTTTTTCTGCCTGATCCCTGATGGAGCTTATTCTCAAAGTCTCAACCTCTCATCAGCCAACTCTTGTC
AGCCTATGCCCTTTCTCTCGATGTCTCACCTCTTTTCTGATGGTCGTTTTATCTTCTCTCACTGATGTC
AAGCCTGATCTTCATTAATAAATACCTTCTTTTTCGATGGAGTCGTGCCGTCTTCTTTCTTTTCGATGGTTC
AATAAAATAGATGGAAATGTGCCATCTTCTTTTCGATGAAACTTCATTGTACATAGTCCTTTTTTCGATCT
TATATGATGATATTTTTTTAGACACTAACAAATATTTAACATAATATTAATATACTAAAATTTATGTAT
TATTAATTTATTTAACTAAAGTTATAAATTATATACATACTCTTTATAATTAATTTGACTTTACTATTTA
TTTAGTGTCTAGAAAAATGTGCGGACCAACGACCGTAGGAACTCTAGAGGGTAAGAAAAATCAAT
CGCTTTATAAGACCATCAGAAAGAGGTTTAATATTTTGTGAGACCATCGAAGGAGAAAGAGATAAAA
CTTTTTACGACTCCACTCAGAAAGAGGTTTAATATTTTGTGAGACCATCGAAGGAGAAAGAGATAAAA
CTTTTTACGACTCCATCAGAAAGAGGTTTATTGTTTTTACGAGACCATCGAAGGAGAAAGAGATAAAA
TTTTTACGACTCCATCAGAAAGAGGTTTAATATTTTGTGAGAAAGGAGAAAGAGATAAAAACCTTTTT
ACGACTCCACTGGGGAAAGAGGTTTCTAATATTTTGTGAGACCATCGAAGGAGCAAAGAGATAAACTTT
TTTTATGACTCCATCAGAAAGAGGTTTATGATTTTACGAGACCATGAAGGAGAAAGAGGATAAACTTT
TTACGACTCCATCAGAAAGAGGTTTAAATATTTTGTGAGACCATCATTGAGAAAGAGATAAACTTTTT
ACGACTCCATCAGAAAGAGGTTTAAATATTTTATAGAGACTACTCTAGCTGAGAAAGAGATAAACTTT
TTTGACTCCACCCACCAGAAAGAGGTTTACTATTTTGTGAGACTACTCGAAGGAGACTCCATCAGAAAA
GAGGTTGCAATTATTTTGTGAGAAAGGAGAAAGATAAACTTTTTTACGACTCCATCAGAAAGAGGCTT
GAATATTTTGTGAGAGAGTGAGAAAGAGATAAACTTTTTTACGACTCCATCAGAGAGGCTTCTAATATT
TTTGTGAGACCATCGAAGGAGAGAAAGAGATAAACTTTTTTACGAATTCATCAGAAAGAGGCTTAAATTA
TTTTTGTGAGACCATCTAGAAGGAGAAAGAGATAAACTTTTACACTCCATCAGAAAGAGGAATAATATT
TTCAGGAGACCATCATCGAAGAGAGAAAGAGATAAACTTTTAAATGACTCCACTCAGAAAGAGGTTT
AATATTTTATAGAGACCATCGAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGTTT
GCATTTTGTGAGACCATCGAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGTTT
AATATTTTATAGAGACCATCAGAAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGTTT
ATTATTTTATAGAGACTATCAAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGTTTAA
TATTTTGTGAGAGACCATCGAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGGT
TTAATATTTTATAGAGACCATCGAAGAGAGAAAGAGATAAACTTTTACGACTCCATCAGAAAGAGGTT
TTAATATTTTGTGAGACCATCGAAGAGAGAAAGAGATAAACTTTTTTACTGACTCCATCAGAAAGA

GGTCAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTTACGACTCCCATCAGA
AAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTACGACTCCCATCAGA
AAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTACGACTCCCATCAGAA
AGAGGTTTAATATTTCTGTGGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTTACGACTCCACCCA
TCAGAAAGAGGGCGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTACGACTC
CATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTTACGACTCC
ATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTTACGACTCCA
TCAGAAAGAGGTTTAATATTTTTGTGAAAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTACGACCC
ATCAGAAAGAGGTTTAATACTTTTTGTGAGACCATCGAAGAGAGAGAGAAAAGAGATAAACTTTTTTACGACTC
ATCAGAAAGAGGCTTAATATTTTTGTGGGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTACGACT
CCATCAGAAAGAGGTTTAATATTGCAGGGAGACCATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTACGA
CTCCATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTAATGAC
TCCATCAGAAAGAGGTTTAATATTTTTGTGAATTATCGAAGAGAGAGAAAAGAGATAAAACTTTTTTACGACT
CCATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTTACGACT
CCATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAACTTTTTTACGACTCC
ATCAGAAAGAGGTTTAATATTTCTGTGAGACCATCGAAGAGAGAGAAAAGAGATAAGTAGTTAGTTTCAG
ATATTTTCTGTACAAAAGTCAATAATGTTTTAAATAATAAGGGACAAGAATTTGTCTGTATAAAAA
CTTGTGTGAATTTTGTACCAAAAGAAAAATGTGAGGTATCCCTCATATGGATTTTACAGATCATTTAT
ATAACAAAAATATTATACGATCTACGTTTTATTATATGATTTTAAACGTAAATTATAAACATTATTTTGG
CGACATACAATTAATTGTAAACCTAGATAAGCATAGGGGATGTTGATAAGCTCGACGAGTATATGCTG
TTGGACGTTGTTTAAGAAATAGTTGATGCATCAGAAAGAGAATAAAAAATATTTTAGTGAGACCAATTG
AAGAGAGAGAGAGATAAAACTTTTTTACGACTCCCATCAGAAAGAGGGGTTTAATATTTTTGTGAGACCAT
CGAAGAGAGAGAGAGAGAGATAAAATATTTTATGACTCCATTGAAGAGAGAGAGAGAGAGAGAGAGAGAG
GAGAATAAAAAATATTTTGAAGGACACCACTGTGAAAGAGGTTTAATATTTTTGTGAGCAGACCATCAG
AAG
ATAAAAATATTTTAGTGACATTCATCAGAAAGAGGTTTAATATTTTTTATGAGACCACTAAAGAGAGAGAG
AG
TAAAACTTTTTTGTGAGACCATCGAAG
GGACCATCAAAG
AG
TGTAACCTTTTTTATGAGACCATCAAAG
AGACCATCAAAG
AAAG
AATTCAGGCGAG
AAGAAG
ATATAGTGTCCGGTACACTTTTAGACCATCGAAG
GAGAGTAG
CACAGCTGAAGCTTCGAGGCTCGCATCTCTCTCACGCGCCCGCCCTACCTGAGGCGCCCATCCACGCC
GGCTGAGCTGCTGCTCCTGCCGCTCCCGCTGTGGTGCTCTCTGAAGTGCAGCTCCGCGCTAGCTC
CCGGCAACGGTACCGGACTCGCAGAGCTTCTAAATTGAAATTTTTATTTTTTTTTTTTGGAATATA
AATGCAGGAGCAAGGGCGAGAGGATAACATGGCCATCATCAAGGAGTTTCAGCGCGCTTCAAGGTCCA
CATGGAGGGGCTCCGTGAACGGCCAATGAGTTTCGAGATCGAGGGCGAGAGGGCGAGGGCCGCCCCCTA
CGAGGGGGCACCCAGACCCCTGCCAAGCTGAAGGTGACCAAGCGGGGTGGCCCCCTGCTCTTCGCCTGGG
GACATCCTGTCCCTCAGTTCATGTACGCTCCAAGGCTCACGTGAAGCACCCCAAGGCCGACATCTCCG
GACTACTTGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAAGTTCGAGGACGGCGG
CGTGAGACCGTGACCCAGGACTCCTCCTGCAGGATTTTAAGGAGTTCATCTGGCTAAGGTGAAGCTGC
GCGGCACCAACTTCCCCTCCCGACGGCCCCGTAATGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
CGGAGCGGATGTACTTCGAGGACGGCGCCCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCCGGAAG
GATATGGCCACTACGACGCTGGAGAGGTCAAGACCACCTACAAGGCTGGGCAGTCTCCGTGCAGCTGC
CCGGCGCTCAACGTCAACATCAAGTTGGACATCACCTCCCCACAACGAGGACTGGCTACCATCGTGGA
ACAGTACAACGCGCCGAGGGCCGCCACCCACCGCGGCATGGGATGAGCTGTGCCAGTCTAGATTTCG
CCACCATGTTTGCATTTTGTATTTCGTTGCCAATBCGGGTGATGTAATAAAGGCAGAGTATACGAGA
AGGATTATGCTCATATATTTATCTTTTTTACTATCCTCACTTTGAAGCTATCTTACAGAGAGTGCCTTA
AGATGCATGGATAGATATGTTGAATATAGATAAACTGGTAGGAAAACCCGTAAAAGTTAAAGTGATT
AGAGTTGATTATACAAAAGGATATATAGATGTCAATTACAAAGGATGTAGACATCAATAATAGCGGCC
GCCCAACTCTGTTTATTGACGCTTATAATGGTTGGGCAACAAAAGCAATAGCATCACAAATTACTGAA

TAAAGTGATTTTTTCACTGGGTAACATTTTTGAGTGATTAATCAGATCAGAAAGAGGGTTTTAATATTT
CAGGGAGATCCATCGAATAGAGAAAGAGAATAAAAAATATTTTTGTAACTTTTTATATATCCACTAA
AGAGAGAAAAGAGAATAAAAAATATTTTTGTAACTTTTTATGTGAGACCATCAAGAGAGAAAGAGAAT
AGCAATATTTTTGTAACTTTTAGAGACCATCAAAGAGAGGAAAGAGAATAAAAAATATTTTTGTAACTT
TTATGAGACCATCGAAGAGAGAAAGAGAATTGTAACATTTTAGAACTTTTTATGAGACTGACCAAAGA
GAAAGAGAATAAAAAATATTTTTAGAACTTTTTATGAGATCCCGACCGAAGAGAGAAAGAGAATAAAAA
TATTTCCCACTTTTTTATGTGAGACCATCAGAAAGAGTTTTAAATATTTTGTGATACTCCTGGAAAGG
AAATAGGAATAGGAATAGGAATAGTGTGCATAATCGTATCACACTATTGAGACAGAAAAAGAAGAAGT
AACGAGAGGTAACTTTTGTGAATGTAGTTAAGAACATTTTTGTTTTGCAAACCGGAACATAGTGTCCG
GTACACTTTTAGACCATCGAAGAGAGAAAGAGAATAAAAAATATTTTATGACTCATTGAAGAGAGAGA
GAAATGAGAATGAGAATAAAAAATATTTTAGTCTTTTTTCGCAACGGGTTTGCTGCCAGAACACAGCTGA
AGCTTCGAGGCTCGCATCCTCCTTCACGCGCGCCCGCCGCTACCCGAGGCCGCCATCCACGCCGGTT
GAGTCGCGTTCTGCGCCTCCCGCTGTGGTGCTCCTGAAGTGTGTCGCGCTAGCCTGGGCAACGG
TACGATCCCTGCGCAGAAGCTTCAAAAATTGAAATTTTTATTTTTTTTTTGAATATAAATGGTGAGC
AAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTCCACATGGAGGGCT
CCGAACGGCCACGAGTTTCGAGATCGGAGGGCGAGGGCGAGGGCCGCCCTACGGAGGGGCAACCCAGA
CGGGCCAAGCTGAAGGTGACTGGCGCGGCCCTGCTCCTTCGCTCCGGGGCCATCCTGTCCCCTC
AGTTCATGCAGTACGGCTCCAAGGCTCACGTGACCCCGCGACATCCCTCCTCTCGACTACTTGAAGCC
GTCCTTCCCCGAGGGCTTCAACAAGTGGGAGGCAAAATGAACTTCTGAGGACGGCGGGCGTGGTGACC
GTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATTTCATAAGGTGAAAGCTGCGCGGCACCAAC
TTCCCTCCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGCA
CCCCGAGGACGGCGCCCTGAAGGGCGAGAGATCAAGCAGAGGCTGAGCTGAAGGACGGCGGCCAAT
ACGACGCTGAGGCCAAGACCACCTACAAGGCCATAGCCCGCAGCTGCCCGCGCCTACAACGTCAAC
ATCAAGTTGACATCACCTCCACAACGATGACCACAGCCTTCACGTGGAACAGTACGAACGCGCCGAG
GGCCGCCGCCACAATCTACCCGGCGGCATGGACGAGCTGTACAAGCTCCATCTGCCACCATGCTTGCA
TTTTGTTACCTGTTGCCCAATGCGGGTGATGCAATAAGGCAGAGTATACGAGAAGGATTATGCTCATA
TATTTTATCTTTTGAATATCCTCACTTTGAAGCTATCTTGGCAGAGAGTGTAAAGATGCAGGCGGATAT
GTTGAATATAGGGATAAACCAGGAAAACCTGTAAAAGTTAAAGTGGATAATTATAATAAAGGATATAT
ATAGATGTCAATTCACAAAGGATGTAGACATCAATAGGCGGCCGCCAAACTTGTATTATGCAGCTTA
TAATATTACAAATAAAGAATAGTGATCACAATTTTCATTGAATAAAGCATTTTTTCATCAGCAATATTT
TAGTGACACCATCAGAAAGAGGTTTAATATTTTTAGGAGACCGATCGAAGAGAGAGAAAGAGAGAAT
AAAAATATTTTTGTAAACCTTTTTATGAGATCCATCAAAGAGAGAAAGAGAATAAAAAATATTTTAGAACT
TATGAGACCATCAAGAGAGAGAAAGAGAATAAGTAAATATTTGTAACTTTTTATGAGACCATCAAAGAGAGAAAGAGAA
CAAAAATATTTTTGTAAACCTTTATGAGACCATCAAAGAGAGAGAAAGAGAATAAAAAATATTTTTGTAA
ACTTTTTTTATGAGACCATCAAAGAGAGAGAAAGAGAATAAAAAATATTTTTGTAACTATGAGAGACCAT
CAGAAAGAGGTTTATTATTTTGTGATAGAAAGGAAATAGAATAGAGGGAGTGTGCATAATCGTATCACA
CTATTGACAGAAAAGAAGAAGTAACGAGAGGTAACCTGTGAAGTTACAACATTTTTGTTTGCAAACCG
AATATAGTATGTCCGTAATACACCATCGAAGAGAGAGAAAGAGAATAAACAACAAAAATATCGACGAAGA
GAGAAAGAGAAAAATGAGAATGAGAATAAAAAATATTTTAGTTCTTTTCGCAACGGGTTTCGGCGCGAA
ACACGCTGAGGGGGCTGCAGAGATTCTCTTCACGGCGCCGCCATCCTTGAGGGCCGTCGTACT
CCTCCCGCCAGGGTGTCTGTTTTCTGAATCAGGCCAGTCCGCCAGCCGCCGGGGCGGTACCGGAT
CCCTGGTAGGTTTTTAAAAATTGAAATTTGTTTTTGAACAGTAAATGGTGAGGCGGGCGAGGGACATG
GCCATCATCATGAGTCTGCTGCGTTCACATGATCGTGACCCAGGCCACGAGATCGAGGAGAGG
GCAGAATTCTGCGCACCCGACCGCGTTGAAGGTGACCACCAAGTGCCTCTGCCTGGGGACATCCT
GTCCTGCGCTGGCGTACGGCCACGGTCACGTGAAGCAATCTCCCGCCGACATCCTCCTTCGAAGGAC
TGCTCACGTTTCGTCCTTCGAGGGCGAGCGCGCGCTATGAACTTTCAGGCGTGAAACCTCCCGCGAG
TTCCTACAGCCGCCAGCGGCACCAACTCCGCAGAAGGGAGTGCATACCGTGCGAAGGACCATATG
CTGGCGCCTACAACCGTCAATAACATCAAGTTGGACATCGCCTCCACAACGAGGACTACACCATCGT
GGAACAGTAATGAACGCCATTGAGGGCCGCACTCCACCGCGGCATGGACGAGCTGTACAAGTCTA
GATCTGCCACCATGTTTGTGCATTTTGTGTCTAAGGTTCACTGTGGCAGATGTGAATAGCAGAGTATA
CGAGGAAGGATTATGCTCTATATTTATCTTTTGAATATCCTCACTTTGAAGCTATCTTGGCAGAGAGG
CAGATGTATATGGATAGATATGCTGAATATAGGGATAAAGTGGTAGGAAAACCTGTAAAAGCAGTGAT
TGAGTTGATTTATAAAGGATTGGTATAGATAAAGTCACTCATAAAGGGATGTAGATCAATAATAGG
GTGGTCAGCCTAACTTGTATTATTTGTAGCTTATATTTGTTACAAATAAAGCTAAATAGTGACTGCAAT
TTCACAAATAAAGTATTTTTTCACTGCAAAAAATATTTTAGTGACAATTCCATCAGAAAGAGGTTTAAT

ATTTTGTAGACCACTGAAGAGAGAGAGAATAAAAAAATATTTTGTAAAACCTTTTATGAGACCATCAAA
GAGAGAAAGAGAATAAAAAATATTTCTGTAAAACCTTTTATGAGACCATCAAAGAAAGAGAAAGAGAAT
AAAAATTATTTTGTAAAACCTTTTATGAGACCATCAAAGAGAGAAAGAGAATAAAAAATATTTTGT
AACTTTTATGAGACCATCAAAGAGAGAAAGAGAATAAAAAATATTTTGTAAAACCTTTTATGAGACCAT
CAAAGAGAGAAAGAGAATAAAAAATATTTTGTAAAACCTTTTATGAGACCATCAAAGAGAGAA
GAGAATAAAAAATATTTTGTAAAACCTTTTATGAGACCATCAGAAAGAGGTTTAATATTTTGTGATACCCTG
AAAGGAAATAGGAATAGAAGAGGAATAGTGTCTAATCGTATCACACTATTGAGACAGAAAAAGAAG
AAGTAACGAGAGGTAACCTTCTTTTGTGAATGTAGTTAAGAACATTTTATTTTGTCAAACCGGAATATAG
TGTCGGCAATAATTTTAAATTGTGGTGTCTGAATCGTTTCGATTAAACCTACCCATCCAATTTTCAAGATGA
ATAGAAGCTATCGATTTCAGCATCATCGCTTTGAGTTTGTGAATCGATGAGTGAAGTATCATCGGTTG
CACAGATGCCGATCATACTTGAATCCATCCTTGACCTCAAGTTCAGATTTTCGATTCTTGCACATGTT
CCGATACGAACGCTAACTCTAGATTCTTTATCATCATTATTTTGTATCGACGATCGTTGGAACCGAC
GATATCTTCGTAACCTCACTTCTCTTATATTAGACCGGAGTACTGATGGGTCTTGATGTCGCTGTCTTCTC
TTCGCTATGATCTGATGTCGATAGACACCCTCACAGTCTTGATCATAAAGCCAGAGCTTCTTCACGAGT
GGATCGCGGAGAGTCTTACCTGTCTGTACGCTGGACAATCTAGCATTCACTGTGTTTCCATCGAAGC
GGATTCTGAGGATGTGATTTAATCTGAGGACGGTAATCCAAAGTTCATTCAAGGACCTCCAATTGATGA
TGGAGTACAAGTGGTAGGATCAGATCTCGACTGGATGTGGAATCACTTCTGATTTCCAATCTCGAT
CTGGATCTGACTCGGACTCTGTAATTTGCTACGGATTAAATCTTATCATTGGTCTGGTGTGGTCTTGCT
TTGGTGGACTTTGATAATAACATCGATTCTTATAAGGATGTTTGTCTTCTTCAAACACGAGGAGGA
GGATGAGGATTGCTGAAGACTGGCAGCAATGATGCATGCCAGGACGATATATTGTTTCATAATTTCGTT
ATTGATTGAGTACTGTTCTTTATGATTCTACTTTTACTCATGCACAATTAGAATATATTTTCTAATTTTA
CGAGAAATTAATTTATTGTATTTATTATTTATGGGTGAACTTAGCCTATAAAAGCGGTGGTTTGGAAAT
TAGTGATCAGTTTATGTATATCGCAACTACCGGCATAGTATTATCGACATCGAGAACATTGTACCTAC
ATGATAAAGAGATTGTATCAGTTTCGTAGTCTTGAGTATTGGTATTACTATATAGTATATAGATGTCTGA
TAGATATATAGTCTCCGAATGCGGCATGATACCGTCATCATTCTTTGCTTCGCTAACTGTTTGGAGGA
AGAATCTTTGTTGCATTTAACTCGAAATTCAGAGTGCACACCTTTCTCTCTGTAAAGAATCCTGAAGT
CGCTACCTTATTAAGAACGAGAAGTATCCATCACGAAATACGGGATTATGCCTTATGGATTTCATAGTA
ATAGTTAGTTCCGACGTTGAGATGGATTCAATTGAGACCGGTAGTGGTCGTCGAGGCTAATGACGTC
GTTGACGGGACATGATTAATTTCCACATCGATATAGTTAAAGGTATTTCTGGGTATGGGTTTCGATTTA
TCTGCGGAAGAGACGGTGTGAGAATATGTTCCGAGACCACACGGAGAACAGATGACGTCTCCGGATA
CTCCGCATCCTATTCCACATTTTGTGTTGGGAAACACATGCCTTGCATCCGGATGAGATCCTTTTGAGAA
CATTGGGCCAATATCTGGGAGAGCAGACCAGATTCTATTGTGAGTGTTACACGATCGTCTTCCGTTACA
ACTTAGACAAGCGGGTAAATGATTATTGGTGAGATGTGAAGTACTCCGAACCACACGGCGTACATTGT
GTGTTAGCGTCTTGCTAATGCGCATAATCTGAAGCGCGTATGTTCCCGGACACAAATTATGGCGTTTGT
ATTGCTTGGTTTCTTTACACTTTCCATCGGATGGCATGCGGTGCTATATCTCTTCCGTTTATTATTAAT
ATACATGAGAGAAACAATATATACGAGTATAATACGGACTTCATAATTTAATAATGTAGTAATCAAGT
TGTCGTTCCCTGTTTCTTACTTCTCCAATCATATAGATATTTTCTTTCTATCATGGATAATATTTGTAATG
GTTCTTCCGGTACAACATACTGTTTAGATGATATTGCGCATAATTTCCGAGGCAAATACGATAGTTCAG
ATTGACCGATGGTAGACTCCTAATTTATTGAGTGTCTTGTGCGACGAGTTTCATTTCTTACGCTCCATCG
ATAGATGGCACTGTTCTATGAGATCGTCGTACATGGGAAATGAAATGTGTTTCGTCGGAATGTATGATT
CCAAGATAGCTGTGATACCGTATACAGGTCGGTATGCTGAAAGATTGGAATCTCTTTGAGGCGACTTA
TGTCACGATGATGGAATCTATCTTACTGAATGATATATTTTTCATATAAATACACTTTTATAGTCCGTTTA
AACAGAATTTACTATGTAGTTCCGCGAATGACTCGTCCCTTAATAGGCAGTAGGCTAGTATCTTTTACG
TAGTAATCGTCGTAGGGAGAGAATTCTGACATCTTGTAGAACAACGATTTAATCATAGGTAGAGATAC
TTTCAGTCTGTGGTGGATGATAAACATTACAAACATCCGCTTGTATATGATGTTTCTGTTTCAAACA
CCAAGTCAATACCGTCTTAGTCGGAAGGTTGATGTCGTATCCGATGTATGAGGCAAACATTGAGTTG
TTACACTTGAATAATATGGCATATAGTATTCGTCTTTCTGAATGTGCAACCTATCTAGTAGATACCGTAG
TATATTGAGAGTGTATCCTTGATTATGTTTTATGAATAGATAAAGTAGATGGCTGTCCTTCTTCTTTT
GTTTCGCCAATTCGAGTAACATTGGCGAGAATATGACCTGTTGCACAATCGTTCCATGATGGGTGTACA
ATCAAGATTATTACGTATCCTCGAGATAAAGAGCATAACACCACACGAGACTATGTTTGGTATACTGT
TGAAGGTAAGTGTGTAACCGCGTTAATGTTTGTCTCATAACGGCTATCGCGTAGATGAATTGCCGGCT
CGCATCTTAGTGACTTAACTTGAATAATTGCTTTGGTAGAACGTGGATATGTGTTTACAGTAGCAATG
AAGAGAAGAAAGCGTTCATCCTCGTCGGCGCAATTAGGCTTCCATCCTTTGTACAGAACGCAGTAGTT
TAAGCCCTACTGAATTTATATCTAAGATAACACAGCAATAGATCGGATGATTTTACTAAAGTCATCAA
TGGTGTCCGAAAGTTAGTATATCAAAGATCTTGTTATCGATTGATAGTGAATGAATCAGATAGTGGTG
TCCTCCTTTTATCCTTGTCTATCAAAGCATGCAGTGCCGAGGTAACAACAATATCTTAATACAGATGG

ATTTCAACCGTGTATTCATCGTATAGCAATGTAATGGAGAGTTACCCTGTTTATTCAGATCGCAAGTGT
TTAATAACTAGCTTAACAGATGAGACGATGTATCCACATCAAAGAACGTAAAAATACATATGACAAAC
ATTGCTGACAGAAACGTGACCTTCATTCTTACCTCGTCGCCATAAATACGTTAGGTATGTACCATATA
CTGTGCGCAACGATGCTCACAATCTCGTCCATCTCGCCCATCTCTTATTATTTCCATCATTACTTTTTT
CATAATTAGAGATGTACGAAAGAAAAAGAAAAAGAAAAACAGAACAATATATTTTTTAGTAA
TAATGTTTATGCGAGACATATAAAAAATAAACTCCGTGTTTATGATAAAGCTGGTAAATGTTTTATCCAT
CTTTGGACGGAATCAGATTTTGTAAATATGTCATGGAAACAAATGAAACAGGACATTATCAATTCCATG
ATAAACAATTATTTAATGGAGTAATAGCATTCCATGGGTAATTTGAAATCAAGTTTCGTCTGTATTAAT
GTTGTCCACTATGGAGTCGATCCTCCTAATTGTTCTTACAGTTTCTGCAATGACGTTAGTTTTTTGTAC
CATTTGATGTGCGATTCTTTGCGTATTCTAGTCTGTGGCGTTTGCTTCTGTTTAAATAATATATCAAACA
TGGAGACGCCTGATATGTAGGCATTCTTCATTCTATTAATGTTCACTTATAGCGCTTAGTTTCTTATGA
TTAATCGGTGATATCAATACTTACTTTAGAAGGAAAAATCATCATCTAGATTAAGGCGCATCTGATACA
GGCAATAATGTCCAGGATATAGATATGCATATCTCATTAAATGCGTCAATCATAGTCTCAGAGGGAT
GGCAGCTAAGTAATAAATCAAACCTATCCCTCGTTTTGTTTTCTTCTTGGTAACTGCTTTTCTCTGGATG
GCCGTATTGATTATCGAGCGTGATGTTGTAACACTCTGCCCATATTCCAATAACCGCTTTGCAAATTGT
ATATTATGACATCGACCGTGTAATATAGTAGAGTTTTATTCTCATTATCGATCATATCTATATCATCCA
TGTACTTGCTTTAGTATATCAAATACATCTATTAGTATGGTTTCATAACAGTGATACCCGCAATTATTA
AATCTCGATATCAGACCGTACATACATACAGGGTCATTGTTTAGATATGTGATTTAAGGCAGCCGCGT
GTCCATATTTTCCACGATAAACCCATGACGTTTTAGCATCGGACGAGACATATTAACAAAGTAGCGTC
GTGCAGAGGGATAGTTGTTGTCCGTTATCTAACATGCACGAACTAATTCATATCGCCGTAATGTAAG
TAGTTTAAACATCAACATGGCTTGGTACGATGGATTTCATCCTGTTCAACCCTTTAGAATGTTATCGATGA
TGTAAGTGGTTATATTCTTGAATCGTACGGAAGTAATACTACGCATTACGTCGACAAGAGCATGACGTC
TCTCAACAAGAAGATTAACGATTTGGCGTTACATTATATGGGGTTACTGCAACGTTTAAAGCTTAGATA
ATACGCCTCTTAATATAGGCTGACGTCGTATACCTTACACGTGTCCACATCCTTTATTAATAATAATTT
AACAATCTCTATATCTATGGTTGAGAAAGACCAGTAGTATTGGATGGGTAAAGATCCTCCTTCGTCTCG
CCATATGGATGGAACATTGCGCTATGGATCAAACATTTAATTACATCCCTGGATAGAGATTGAGATT
CTCTATGAGACGATATATAGTAATGAAGAGAGTTCTTACAATATTATCACTGTTGTATATACAGGTACG
AAAGATGCAACCGGTGCTGTAAACATCTCTGATTAAATAGCCATAAGTTAGTTTCTGGTCTCGGATTAG
GCGTCGTTACGTATATATCCACCCAATCCGATGAATTCATTGATTGTATAATTGTAATTGGACGGACGT
ATCCGTTTATCCACAATTAGGTATTTGAAGACGTAAGTCGATTATCCGAAGACAGATCGAAATCATTT
ATATTCGACTTGAGTTCGTTAGAGGAATTGCAATAGCTGGATATCGCAGGATGCACAAATCTGAGATT
TTGTATTTACTTCCATGTTTACTGTATGCTCCTAGCGGAGTTAACCTTCGTTGTTTCTACAAAGTCTCGA
CTCCGCGAGAGAGTAACAGTTCGAACAATCTTAATGTCTGTTATCGCATTTATTGGAGACGTAACAATA
AAGTGCATTGTTTCTCGTTTCATCTATATATGTTTTGATAAGTTGTGACACGTTTCAATCTCTAGTTTATT
TTTTTTGTACGTCACATCTTCATCCAGTAGACGACATAGAATAGTGTACCTTCTACCACAATAATCCAT
AGCTATTCTTGGTGCTAATATTCCTATTTTACGAAAAATGATGAAGGCAATCTGGCCTCATAAGATGAT
GATAAAAAAGTGTAGTGAGAGAGCATGAAGGAGATTTAGTATTTAGCAGTGCGGATGTTGATCCAAG
AGGGTGAGATAGTCGTTCTCGTCTGTGAATCTTCTCGCAGCATAAGTAGTATGCTCGATATATTTATC
GTTGGAAGACTCTTCCAGAGACGATAGCTGATTGAGTACAAAGTCCAATGATTGCATCGAAGTTCTCG
GCGGTTTTTCACTGGAGTCATTTCTGATGAAACATTTAATGATCTCACTTGCCTGTGATATTGTCCAC
GGAAGTGAACCTTTCCAACCTACCACCAAAGAGCTCCGTTGCATCAGTTCTGAAAGAGATGAGAAGCC
TGTAAGAGAGACCCTGCGTTTTCTCTATGGGTCCATCTATGAGAAACCCACAGGATGTATTCAGTCAGA
CAATGTTCTCCGACGTCGGCCACGGTATTCAGGGAGTCCCTTAGTATGTGGCAATGACAGGGTCTGAA
TGTTACCAAGGAAAGGCCATTGTAAAGGTAGACCTAGCCGTTTATGCTAATAGAGGCTTTAATTTTCATT
TTTGAATGGGTTGTGGATGAGGAATGAGAGTGATATCATATTGAGAGCGTAGTTATGTAGAGGTGTA
TTTCCTATATTATTTACTTTTCGGTTTATTTTACCAACTCTTTTAAATAAATTTCTTTTACAGATGCATCTTA
GTTGAATGATCGCGTTTTTTCATAAGTGGACATATAGATGCAGAAGTAATGAAGAAAAAGTATTACTCT
ATCATCTACATAATTAGGGTCTGCTCCTTTTTTAACTTATAGCATGTAGTAGTAGTTTACTAGTTCA
AATCAAGTCTAGAATATATAGTGGATTAATATATTTTTATATTAGCTAAAAAGCTTTATCTATACTCAT
CAGAAAGCATATCATTCTCAACTTCATCATGAGTTAAATATTTGTGTAATGGATGTCATGAACATTAA
CGTATTCATGACATACCTTTAATAGGTTTTTAAACAGTGATGATTCAAATCCTTCCATTTCATTAGATAA
CAGTGTAAACGGAGTCAAATTCTCTTACTAGTTTGTATATATCATAAGCATCTCTACAAACAGTCTAAAA
CAACAATAGAGAAGACGGACAGATTCTTTTAAACGCATAAATGACACATGTTATCAAATGATATTTCGCT
GGATGAATTATTATTAACGTAAGTTATGATAAATGATCTCTAACGACATCTCTCGCTAGAGATAAAAT
CTAGTATCAAGAACATTAAACATCTTTGCATCATACTCGCATAGCATAGTTTTTCATAATTATCACAAT
ATTTAAAAAGACTTATTCGGAAAAAGTATTTTACATGTATCATCGATGGAGATCCATATGAGGAGTCAC

TAGTTCCTTCAGTAGTAATAACAGTGCTATCATTGATAGTATAATTATATAAGGCGTAGAAGGTTTCAT
ATGTTGTTGTAATTGAGTAACCGTTGGTAGTTCCTGTGGAATCAATAATTATACTAACAGCAAATAGTAT
ATAATATATAAATATGTTCCGTTGATATCATTATTTAATGAACCCATTCTAACACCCTGTTATATCGT
CCAATTAATGTAGCCAAACAATCTACCATGTTCTCGATTGACTACTTGTACGGTAGCGACGCTGACTA
TCTTTTATTGTCTCTGCTCCAATTGAATGTCGCGATACAACGCAGTTTCTTCTTAAGTGCATGTTTCATA
TCACCACGAACATAAGTCGGCAAGATATATAGCCAGAGATACTTGTAAATCATGATTGCTTGGTCATAA
ACAAGCCCGTCAATAATTGTATACATATTCAGTTTAGAGCAAAATAATTAAGCACAATAGCGCTTAAT
CTCAAAATATGTTATGTTATTTTTTCATATTAACATACTGGTTAAAAATCCTCTAAAGGCTGATCTTCT
ATCTATAAATCAAGACTACCATTTAGACAGTGGTTTCATGTTTATAAAATGTTCTTTTGTGTGAATAAG
GAATATGTTTAAATCAATAATCAACCATCGACCCCATACGATAGTATGCAGGTTAAACCCCATAGAG
AGGTACGTGTAATCAGTCTCTCCAGTTTATGATTTTTTATAAGTCATTACTACATAAACGGCTTTTAAA
CAGTCTCCTCGATAATAAGCCATATCTGAAATTCTATTAAATACTCGAGTCATTTTACGCACGGTCAAA
AAAGTAAGTAATGCCGACGACTTCTACATCTCTATAGAACACCTTAGAATACTCATTTTCTTTGGAAAA
TATCCTCAGACTCTGATTTAACAATGCACGACCTATAGTAAACCGCAGACCAAGTTATATTAGTCAA
TGGTATATCCAACCATCAGGTGTGGATAGTCCAGTCTTTGGTATCGATAGTGTAGTTATTGAACTGAGA
AGTTACCGTATAAATTTTTGGTCATCTCTAAACAAGGAACTAATAATTCTACACTATTGAACGATTATA
TCTTCCGTAATGGGTGGACTAGCGATGATGAAGTCACGAATATAAGACACGCTATTAATCCGTATATC
ATCATTTTGATATTACTTATAATAACGATTTGTTAATTTTAGTTTATACTATTAAATTGTAAATGATATT
ATTATTTTTTAAGTATTATCAGTTTAGTTTATACTATTACTATTTCTGTAATATTTAGACATAGATAAAC
GTGATAAAGTCTATTTGTTTATATTTATTGCGGATAGCAGTATTTCCATAAACCGCAGCAAGGAGATGTC
CTAGGGTTTACTATACATGTGAATGATTGACGAGGTAGATCTGTAGCCGATAGGTTTCATCGAAAAGA
CTGTAATACTCCTGTAACGGGCTGATGCTTCCTAGGCACCTTTTATTAATGACGAGAACGTCGGGAATTT
TTCATAAAGATAGTTTCTTATTACATCTCATAAATATCCGGAATTATTGCAGTCCCGAATTTCTTTAAA
ATCGAGAAATTCAGATTATACGGTCTATTATTTCCAGACTTTTAACTTTGTATTATATTAAGTCTATA
CTTTTGGTAGATAACATACCATTAAATCATTTTAATTTTTTCATCCATGCTTTTATGTTATACACCTCTATC
GTCCAGTATAATCACCATTACTAGAGTTTCTCGTATGGAATATCGATAATAGAGACCCTCCGAATAATT
CATTAGCGGCTATTGCAAGTTATTCTCGGTACCACCATCATATCAACACTGCACCATATTCTTAGAAAT
ATAAACTTGTCTGTATATGTAAGATGTGTTTAGAAAAGCCGATATTTACATTGCTTTGTAAGTAAATGG
ACGGCGCTAACAACCTGTCATATTTATATTAATGGATAGCGGATTAGTCAATAAGGAATTACTCTTACC
ATTTGTCACGTCTTAACCCCATTCGCTGATTAGTTCCTTTGTTTGGTTAGCATTATTAAGCTTACAGCT
TGAAAATCGTCTTATTTTTGTAGGAAGGAGGCGTGGAATCCCGATACTATCGCTACCGTATATTTTAT
TTGCGGTAGCGCTAGTGTCTGCACAATACGAATATCTACGTCCATGTCATTATTGCTCATCGGGTGTAT
TCTCATTCATATTCTCTATATATTTTGATAGTTGTTTCAGCTGTAGAACCAGCTGCCCATGATTTAGAACA
GACAGTAGATAAAATAGAAAACCGAGAAACAAAACATTTTCATCAGGGTGTTTTACGATTAGTTCTTT
AAAGATATCCATGGTATGTGGACCAACAATAACGATAACGATATATATCATAAATAGCCAATGTTAA
ATTTTCAGTTTATGTTTGGTACCCCGTATTCATACATTTCAATAAATTGGTATTGCGTACACAATCAATC
ATATTACATACCATTAAATAATGCAAGCATAAAAAATCGTTAGTAGATGTTTCAACATAGGTTCCGTAAG
CAAAGAATATAAGAATGAAGCGGTAATGATAAAATCAATCGTTATCTAAAATGATCATACTCATTTAT
TTTATTCTATTATATTAACACATACATTTTTAAATAACAGCAACACATTCAATATTGTATTGTTATTTTA
TATATTATTTACACAATTAACAATATATTACTAGTTTATATTACTGAATTAATAATATAAAATTTCCAT
CTTGTATATAACACACACTGGAGAACAAAGCATAAACACAGAATCCATCAAAAATGTCGATGAAATA
TTCGATGTTGTTGTTGCGCGTTATGATAATCAGATCATTTCGCCGATAGTGGTAACGCTATCGAAACGAC
ATTGTTGCAAAATTGAAAACGCTACAAACAGATATTCCAGCTATCAGATTATGCGGTCCAGAGGGAGA
TGGATATTGTTTACACGGTGACTGTATCCACTAGAGATATCGACGGTATGTATTGTAGATGCTCTCATG
GTTATACAGGCATTAGATGTCAGCATGTAGTATTAGTAGACTATCAACGTTTCAGAAACCTAAACACTA
CAACGTCATATATCCCATCTCCCATTTATGCTTCATTAGGCATTATTATTGAAGCGGCGCACGAAAACGC
GAAAGCGTTTCACGATAAATGCGAAACTAGGGAAACACGATAGAACTGGAAAGGCTTCACGTAACGT
AACGAAGTACAACAG

Isolate 5

ACGGTACTTCGTTTCAGTTACGGCGTATTGCTCTTGTCCAGGGTTTGTGGCCTTGTTCCTGCATTTATCGT
GAAACGCTTTTCGCGTTTTTTCGTGCGCCACTACCCCAAGATGTTCAATACTTCATCTGAAGCAAAAGATC
TTGTGTCTAAACAATGAATACGCTTTATACGATTTTCATCTTGAATAGATCCATAAACGGGGGCTCCATG
AATATCGTGTGTTGAAGTAACGTCGTCGGATACATAAGAATACTTTGTTCAATTACGGCTTTTGGATTTG
CTCCAATATCATCCATTCTTCATCACTCCGTCGAAATCAGCATTTTTTGAGAGTTGGCAATTCGGGAGAT
ATTTTGATAGTATCTCCTTCGGTAGCTCCTGATAGATGAAACGATGACGTTTAACCATCTATGTAGAGA
CGGCTGTCTTCCAAAATAATGCCTTATATTCTTGAACGGCTTCTTCACTGCCCAATCACCAGGCAATAA

ATGTATTTTATTTTAATAAACTTTCTTGGCATATTCTTGTTAATTGGTTTAAATCGTTTATTAAGTAGAA
TTTACGAGTTTGGAGACGCTAATAGTTGTTTAACTTATCCTGTAAAGGCATTAACAAATATCTTTTCTGT
AGTGTATTTCTGTATATGCGGGCATTCTACCTCGTAACGGTGATAGATGTACTGGGACCAATTACAG
ATCTAGCGGTCTGATCTTTTCGTGCGAGGGCGACAATATAACTTCTAATCTTATTATTTTTCGCGGATGTA
ATATATGATAAATTGATACTGGAAGTGTTGTAGAAATAATTTTAAATATCATCGTATTCTGTACTGCCTT
CTGGATAACCTGTTCATCCAGCATTCAAGTTACAATTCCTAACGATCATACCTAATAAAGTAAGTTAAT
TCATTGGTCTCTTTGGGTACATACTATCTATCAAAAATAATAGCCGGTCTAATAATCAGCGGAGGGAT
GGAAAGTAGTCGCAACATAAAAATAAGTTGGCTGGATATTGATGGGTATAATAATGGCTAAAACCTTTTT
CATGAATAGAAATTAACCTTTTTGATAGATGAAGGAAAGTGGAAACGTTAATATCATCCAACCTTGTTGAC
GAAACAAACCTTTCTTGAAAAGTAATTTTTGATACGGTTGCATACATTCGCTGTTCCAACATGAGCTTT
TCTTGATAATATTTATCCTTTAATGCTCCCAGAGCGTGTCCCATTCAACTCTTTAGGTTAATATCGTCC
AGACCGGTTCTCGTGAACACCAATAATCCGCAGTGAATACATATATGATTCCGTAAACGAATAATTTT
TGAAATAAATTCAGGCTTAACCTATGAGGTTTTATAAATACTTACTTACCCAGTGACCAAGAAACATTC
AATTCGTTTTCCACAAGTCTTGTAATGCCCCATCCATAGCACCTAGTCTACGCTCTTTAACGGTACCGA
TATCGTCGTCATTTTTTAACACTGACTAATGATAATATCTGTATTAAATCTCTTTTGATCATATAGACTAT
ATACGTAACCTAGAGATTACAACCATTTTTATCAAGTCAGTTTCTTTTAAAGAACCAAGTATACAATCA
AATTTCCCTTTTTATTACAACCTATAAAATAAATAGTTATATTTACACACTTTAAATTTTTTATCATGACAA
GGACGAACAAATTTATGCATTCTGTGATGCTAACAAAGACGATATACGATGTAAATGTATTTATCCTG
ACCAAAAGCATAGTACGGATAGGAATAGATACAAGATACCCTATTATTATTGTTGGTACGAATAACCA
TGTAACGAAGCGATGCGTTGTTACCAGCCTCTTTAAAAAATAAACAATAATGCAATGTGTATCGGAT
TGTACCATTTTATTGGGAAACGTTTCCATTACAGATAGTAAATTAGATATGTAAATAATGTTTGTGAT
TCCCAAACGGTAGCTACCGAGAATATAGCTGTCCGCTATTTAAATCAGGAAATTAGATACCCTAATTA
TGATTGATGGCTTCCGATTGGATTACTAGTTAGCTATTTTAAATATTACTTTTCTAAGCAAGATATAGAA
TTATAAGATACCAGATATATATTGTGATTGTATAATGTTCTTATCTCATCTCTACTAATTGATTAATCAG
CGACTGAAATAGCAGATCTATCGGCTATCTCTACTCCAGTTACCATGTTGTACTTGGAAAAATCTAACA
ATTTTTAATGGTATATTAGGGCGATAGAATCTTGACAACCTATTTCCGCTCTCCAACATTTTGGGAAACG
AAAGTCTTTTTTACCATCGCCTGTTTGTAGTACACTATTAACCTATATTAGTTTCTGTAAGTATCTAGTCT
AACAAGCTGAAATAGAAGTCGAGATATTACCACATCTATGTATAGAAAATTTACAGAACTTGTTATG
ACCTTTTTGAAGATACGATGTTATAATGATTGGTTCTCCATAGTCATCATCAACCCGGGTATAATTTAT
GGATGATTTCGTTAAATACCTGAAACACAGTTTCCATAGAATCAATTTTATGTCTAACAGTTGTCAAAAA
GTAACATAATCATTATACTCTATATCGGTAGTATATCTCAGTAGTACGTTTCTATTTATTACGCGTACG
GATCTCTTGCTATTTCTGTTTTAGACATAGAATTTTGCTAGATATTTTACGTTGTATTGGTTCTTGACT
AACTTTTCAGTTGATGTTGTTGAATATTTAAGAAACGAAATATAGGTGTTGTAGAAATAGTACCTTTGC
TTTAGTAGTAGGAAATGTTTTATTGCAGTACACGGTCTCAGCATAAAGTACATGTAGAAAATAGTCA
TATTCTGATTAGGATAATCAAAGTTAACTACTTTGTTACGGACGATCTTATTAAGGTAGTACATCT
TTTCATAATTTACAGCGTCTGATTGGTAACTCGAGGTCCAGTCTCATGTTCTCTTTCGGTATAAATACT
TAATAATCTCATTAGCTGAATATGAAGGAGCAAGGTTGTAACATTTTATTACAGTGTGGGATATAAA
GTCCTTGATCCAGTGATCTGGAACGGGCATCTCCATTTGGAGCTAGATGCCACGGGTTTAAAAACT
AATCATGACATTTTGTAGAGCGTAATTGCTTAGTAAATCCGCTTAACACTAGGTTTCAATTCCTCCTCGT
TTGGATCTTATCAGAAATTAATAATCTTAGAAGGATGCAGTTGTTTTGATGGATCGTAGATATTCCTC
ATCAACGAACCGAGTCACTAGAGTCACATACGCAATCCATTTAGGAATAGGATCATGATGGCGTCCG
TCAATTAGCATCCATTTGATAATCATTCTAAATTGCCAAATGATCTCTCAAAATAACGTATATGTGTA
CCAGGAGCCGATCTATGCATACACTACGGTGGCACCATCTAATATACCGTGTGCTGTAACTTACTGA
AAAAAATAATTTCTCCCTGAGCCAATAGTTTTAACGCTGTCTTGATACGGCAGTTTTGCGACCTCATT
GCACTTTCTGGTTCGTAATCTAACTCATTATCGAATTTCACTCAAAATACATAAACGGTTTATCTAACG
ACACAACATCCATTTTTAAGTATTATATTCTAAAAATTTAATCAATGTTTATTTTTAGTTTTGAATAAA
AATATAATATTATGAGTCGATGTAACACTTTCTACACACCGATTGATACATATCATTACCTCTATTATT
TCTATCTCGTTTTCTCACCCAATGTTTAGAAAAAGGCCTCCTTAAAGCATTTTATACACACAGCAGTT
AGTTTTACCACCATTTTACAGATAATGGAATAAGATTCAAAATATTATTAAACGGTTTACGTTGAAATGTC
CCCATCGAGTGCGGCTACTAATTTTCTTCTGTTGCTGCCATACGCTCACAGAATTCACCAATGTC
TGGAAGAACTGTCCTTCATCGATACCTATCACGGAGAAATCTGTAATTGATTCCAGGGCATCCACATA
GTTTAGTTGCTTCCAATGCTTCAAAAATTATTCTTATCATGCGTCCATAGTCCGTTCCATTATCTATTAT
CGTTAGAATATTTATAGTCACGCATTTATATTGAGCTATTTGATAACGTCTAACTCGTCTATTAATTCTG
TACTTTTACCTGAAAACATGGGGCCGATTATCAACTTCAGATAATGTCCGCCGTTTATGATGACAATAA
AGAATTAATTATTGTTCACTTTATTCGGCATTTAATATATCCATCACGTTAGAAAAATGCGATATCATGA
CGAGGATCTATGTATCTAATGGGATCTATTGCGGTGGTAGCTAGGGGATTCTTTTTGAATCGCATCAAA

CTAATCACAAGTCGAACAAATATCCTTTATTAAGTTGACCCTTCCATCTGTAACAATAGGGACCAGCGT
TAACTGGTTTTAAATCTTGAGAGTCTGTGAATTTTGTGGTGTCTGTATTCTCTGAAGAGATTCATAAC
AATGACCCACGGCTCTAATTTGTTTTGATTGGATCAATAATAGTGTAGAAAGTCTAGATATTGAGTGAT
TGCAATATATCAGATAATGAGAGATTCATCATCTTGACTAGCCAAATAGCAAATGAATCATCCTCTGC
GAGAACATCGTTAAGGATACTGGTTGTGATCCATTTATTGATCATAAAAGCTTTGCACAATCTTTATAC
TATCGGTTTACTATTTATTGATAACGCAGATGTTTGAGTTGTCATCCATGGTAATCCATAGATCATTAA
TTTATCGTCTTCACGCTAGATTAGCACGTCGTAATCTGTCAATAGAGGATCGAGGTATTTTTCGTTTTG
ACGCGAAGAACATATTTAATTCAAGATCTAAAAATACATATATTAGGAATGAATACAACAGGAATACT
TCAAATAAAACTATTAATCTGTGTTTATAAACGCAATAAAGGAATGTTTAAACGTGGGCTCTATAAAC
TTTTAAGAGTTAAAGTATACATTAGGAACATTCTCCATTTATAGTAATCAATCCTTTGTCCGGAATATC
TGTTAGAGAATATTTCTTTAACACATTCCAATAGTCCCAGAAAATCATTAACTAATAGGTTCTTGAA
GAGGCTAGAACTATACGATTCTAGTTCCTTATTAGTACTAAGTTTCTCTAGTTTATTTTTATTGGTGAAT
CCGTAAATGGCATTCAATCTCATGGATGTGGAAGGAGGAATACACATTCCAAGGACATTATGAATTCC
GCATATTTGGAATATCAAAGCCTTGTTGTCTCCGATAGGTGTATACTCAGTCGATGCGGATTCCATATT
TTTCTTTATAAATATTAATCTTTTACGAGTTTCGAAAATGCACAAGATGAGTGATGATCCTTGACGAAA
TAACCATGTTCTTATTCTTTCTCCCTACCACCTTTTTAGGAATGACGGCAGTCATATCGTATTATTAATA
TAAGGGGCATAGCTTCATTGTTAGGATCCAAAATGGTTTTCAATGGCTTCAGCAGATTTGGTTTATCAT
CTCCCTAGACCCAAATTGTTTTCTTATATTATACTGTTGTAGTAAACATTGCTCCTGTGATATTTGCCTG
GAGTATGAATCAACTCCAACCTCATATTTTGACATCGTTCTAGGTACTATATTAATAAGAAAGCATAG
ATCTTGGAATTTGGATTTTGCGCCGGCAATAACCATTGTAAATCATCATACTCAGATATGCTTCCTGC
AAAAATATGGTATCTTCTTCGATGAGGTTTTCTAGCAGGCGCTCATTTAGAAGTTTTTGTGATAAATGA
ATACCCATACCGATGTTTACAAACGATAATATAGACAAAAATCCAATAAACGGTGATAGAAATATAC
CAGGAAGAAATGAAAAGATTGTGGCCAAAATAGATTCAACAACGATAGACCAAGGCCTAAACCAAAG
ACTACAGCCTAATCAGCCACCGAAACAAGATAATAAATGCAGAAGTGGAGATTTTATCAATATTAGTT
GTGTGCCTACGAAAAGGAATATTGCAATGACGGATATCTATCTCCTGCCTATTATATGTTAAACAGG
TGGATGATGAAGAAATGAGTTGCTGGTCAGAACTATCGTCGTTGTGAGATCCAGAAGGCGGTAGGATT
TCCTCTATTAAGGCGGCTAAACGTATTTCTGTGGATCGATGCTATATTTTAAACAGTTTAAAAACA
GTAAAGTTGTGAAATTAACCCACTGGGATTAATGTTTAAATGATACTGTTATTTTCAAACGTAGTTA
TTTTATATTCAATGTATAAACGTGGCATATATTCTAACAATTTGTTTTGATCTGGTTTCTATTCCCAGAA
CAGACATTATTTTTCTGTTAATCAATTAATGTTTAAACATTTGTGCAGACATATTGGTAGTTCTATCTAT
TTACGGCAACCGGCTCTATAGAACAATCTACCACAGTCGTGTTACTTAAATTTTCATGCAGCCATGAGA
CAATAGCCGTGAGGATATGTTACTCAATTACTTTTTCGGTGGTTGATAAAAAATCACATATCGCTATTG
ACCAAGCAGAACGATGGATATTCTCAGAGTAAAGAAAAGAAGTATGCTACAGGAGCACCAGTAAATG
AGTTGTTGAACCTGGTACACTGGTATATGTGCCCAAGAAGATTATTACTTTATGAGCATGTCACCTACC
GATGTGTCAATTAGCAATAATGTCAGGTATTATTTTCCACAGATGGAATAGTGTTAGAAATAGAAGAC
TTTAATATCAAGCATTTGTTATGGCAGGTGAATGTTTGTTAGAAGTCAGTCTAGTACTATTATGATATA
AAGTAATAAAAAATAGTTAATGTGATGACTAGCGCCACCAACGCCAACAACATTTGATAATTTCTACC
TAGACGTACCGTAAAAATATATAAATTACTATAACAAATAATGATATATCAATAAACACCTAATTAA
TGGTCGAGTATAGCGAGGACATTGATGCTCTAGACCGTGTATAACAAAATCTACAAATTTTTCGTCCG
CTATATTTTGTTCCTATATCGTCGATGATCAGCGACCAACTTCCATGTTAATCTATTAATAAACCA
TCAGCGTATATTTTTCGTTTGCATATCCGTGGTAGCAATAACCATCGGAGAAGTTCTAAAGAATGTGT
CATGTAAGTAGTCAATGGACGTTTCCTTATTGGCTAAATAAGTTTGATTTTATCATTGGTGGATGTGA
ACAGCATAACGCTTGTAGATACATAAGCACTGCCAATATTATAACACCGATAACAATCATATAAAAACT
GAACTCCTGTACCAGCAACTTGTCTAGGTGCTATTTAGGTAGTGGCCTTGGTAGTCCAATTGCATCAAC
GCCTTAATGGCACAATTTCCCTTGTAGATCCTGTTAATAAGTTCCAAATTTGTTGGAGATCCTGGGGC
TCCTTAACATTCATCTATGATTACGTTTGTATCTTTAATTTGTTATCGACGACCATACTTCAAATTTACA
GGTTGTTTTACATAATTTTCAAAATCTCTGGCAACAGTGTTTACACTCGTCTGAATGTTTAACTTGCAG
CAGTAAACATAGCTGGCGTATGCTTTGTTCCGGTGTAAATCCACTACTATGTTTCTGTAGCGGCTGATA
ACCACAGCATCAACTGAGCATCCGCGTCTGAAACACATATTTTTAACTGTGAGTTACATCCATGGTTT
TGGCCGGATATAAAATTTCCGATTTCTATATCACATTTGTTTTGAACATAGCATTCACTTCTTGTCTAA
TTTAGACGAGATACGTTTCTAGGTGATTACCGTCGTCTGGCAATAGCGCTGCGGCTTTTCCATTT
TAAATAGCTACAATTAGTATCCATGTGCAAGAGATAATAAACTGATCAAATGCAATTTTAGGTGCAAC
GAGTGTTTAATATTATGTTGAACACTTTTTGCTCTATTGACGGGAACAGTAGAAAATCTATCACTATT
GCTTAATCCACATGACAATCTTAATGAGAGTTTATCCATCTGTAAGTTGTTACAGCTAGTATTACATCG
TACAATATTGCAAAGTCTAAATTATTATAATTACGTGTTAGCAAGAAATTAACATTGGCATTGCAACCT
CTGGATCCCAACATTCTCGAGGTTCCGCATTTAATGACTCTTCTAACTTATCTCTAGTGGGATAACTA

CATCTCCTTATATTTCTGTTTAAAGTCCGCAATTTATTTGTCTTAGAATATAATCGATCATCTCTTTGCT
ATCTTCTGTATTGTGTGCGTAAATGATGCAGAAATGATTACATATTGGTACACTAGCATCTTTACTAC
ATAATGTTTGCATCTTATTATACAGATTTATTAACGATTGTTGACCCTCTACAGTTCTATTACTCTATTA
AAATGAACCGATCCACTGATGGCATATGTTTCTATCGAACGTATCCCCTGACACCAGTCGAATAGAGA
AATCAGCATCGCATTTCCAGTGTCTGTAACGTCTGGCCAACGTTCTAATACTGCACCCTCTTCATACTT
GTCTGATATCTTTCCATCCGCTTTTCCAATAATGAGTACGATTAAAGTGGAGCAGCAGTGGGTGCGGT
TCCACGACAACCGCGATTGTCTTAACAATGCCGAAAGACCACCGGGTCCTGTTGACTAATCTAAACTC
ACGGATATCTTTTTCTTTACTAGCTTTTCACCGAGCAGGTTTGGCGCAGAACGCAACTGGAAGCTTCGAG
GGGCTCGCATCTCTCTTCACGCGCCCCGCCCTACCTGAGGCCGCCATCCGCGCCGGTTGAGTCGCG
TTCTGCCGCTCCGCTGTGGTGCCTCTGAACTACGTCCGCGCTCGCTCCTGGGCAGCAGTGGATCCCT
GCGGAAGCTTCTAAAAATTGAAATTTTATTTTTTTTTTGGAAATATAAATAGTGAGCAAGAACGAGGAGG
ACAACATAGCCATCATCAAAGGAGTTTCAAGCTGCGCTTCAAAGGTCACATGGAGGGCTCCGTGAACA
GCCACGAGTTTCGATCGAGGGCGGGGCGAGGCCCGCCCCCTACGAAGGCACCAGACCGCCGGCTGAA
AGGTGACCAAGGGTGGCCTGCCCTTCGCTTGGAAATCCCTCAGTTCATGTATACGGCTCAAAGCTA
CGTGAAGCACCCGCCGACATCCCCGACCGCGGGGCGCAATGTCTTCCCAGGGAACCTCAGAAGTGCG
AGACGCGTGATGAACCTTCGAGGACGGCGGCGTGGTGACCGTGACAGGACTCCTCCCGGACGGCAGT
TCATCTGAGTGAAGCTGCGCGCACCCAACTCCCTCCGACGGCCCCGTAATGCGAAAGACCATGGG
CTGGAGGCCTCCTCCGAGCGGATGTACGGAGACGGCGCCCTGAGGGCGAGATCAAAGCAGAGGCTGA
AGCTGAAGGACGGCGGCCGCACTACGGCCGCTGAGGTCAAGACCACCTACAAGGCCAGAAGCCCGTG
CAGCTGCCACGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACCTGCCACC
ATCGTGGAACAGTACGAAGACGCGCCGAGGGGCCGCACTCCACCGGCAGCATGGACGAGCTGGCCT
GGAAGTCTAGATCTGCCACCATGCTTGTGTTTGTATTCTGTTGCCCAATGCGGGTGATGTAATAAAGAGG
GCAGTATACGAAGGATTATGCTCTATATATTTATCTTTGACTATCCTCACTTTATGAACTATCTTGGC
AGAGAGTGTTGAATGCATATGGATAGATATGTTGAATATGAGGGATAAACTGGTAGGGAAAAGTGT
AAAAGTTAAAAGTGATTAGAGTTGATTATACAAAAGGATATATAGATGATCAATTACAAGGATGTGG
AGCATCAATAATAGGCGACCGCTAAACTTGTGTTTATGTCAGCTTATAATGGTTACAAATAAAGCAAT
AGCATCTAAATTTACAAAATAAAGCATTTTTTCACTTTTGCAAAAATAAATTCCAAATTTGTTGGAGATC
TGGGGCTCCGTAAACATTCATCTATGATTACGTTTTGTATCTTTTAATTTGTTATCGACTGACCACTTGC
TAAATTACAAGTTGTTTTACATAATTTCAAATCTCTAACAACAGTGTTACTCGTCTGAATGTTTAAC
GCAGCCAGTAAACATAGCTGGCACGTGTCTTTGTTCAAGTTAATCACTATATGTTTCTGTAGCGGCCTG
ATAACGGCATCCAAGTGAAGCATCCGCGTCCGACAGACACATATTTTTTAACAGTGAGGTTACATCCAT
GGTTTGTGCGATATAAAAAATTTCCGATTTCTATATCACATTTTGTGTTGACACTAGCATTCTTCTGTTCT
AATTTAGACGAGATACGTTTCGCTGAGTGTATTACCGTCGCTGTATGCTTGCTGCGGCACCATTTAAA
TAGCTACAATTATCCATATTACCAAGAGGATAATAAACTGATCAAATGCAATTTAGGTGAGTGAGTGT
TTAATATTATGTTGAACTACTTTTGTCTATTGACGGGAACAGTAGAAGAATCTATCACTATTGCTTAA
TCCGCGACAATCTTAATGAGGAAGTTTATCCATCTGTAAGTTGTTTCGCGCTAGTATTACATCGTACAA
TATTACCCAAGTCTAAATTATTATAATTACGTGTTAGCAAAATTAACATTGGCATTGCAACACTCTGGA
TCCCAGCATTCTCGAGGTTCCATATTTAATAGCTCTTCTAACTTATCTCTATTAAGGATAACCCATCTCA
TATATTTCTGTTTAAAGTCCGACAGCTGTTGTCTTAGAATATAATCGATCATCTCTTTGCTATCTTCTTA
TTGTGTGCGCGTAAATGATGCGAAATGATTACATATTGGTACACTAGCATCTTTACTTACATAATGTT
TGCATCTTATTATACAGATTTATTAACGATTGTTGACCCCTACAGTTCTATTACTTATTAAAGGCTGA
ACCGATCCGCTGATGGCATGTTTCTATCGAACGTATCCCCTGACGCAGTCGAATAAATCAACATCGCA
TTTTCCAGTGTCGTGAACGTCTGGCCAACACGATTCTAATGCTGCACCTCTTCTATTACTTATCACTGAT
ATCTTTCCATCCTTTTCCAATAATGAGTACGACAAAGTGCAATTGTAGTGGGTCGCAGTTCCATTATGCG
ATTGTGTAACAATGCCGAAAGAACACCACCGGGTCCTGTGTTGACTAATCTAAACTCTGGATATCTTTTTC
TTTACCTAGCTTTTTTCGCTTGAGTTTGGCGCCAGAACACAGCTGAAGCTTCGAGGAGACTCGCATCTC
TCCTTCACGCGCCCCGCCCTACCTGAGACCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCC
CGCTGTGGTGCCTCCTGAACTGCGTCCGCGCTAGTCTCCTGGGCAACGGTGCGGATCCCTGCAGAA
GCTAAAAATTGAAATTTTATTTTTTTTTTTTGGAAATATAAATGGTGAGCAAGGGCGAGGAGGACAACAT
GGCCATCATCCAAGGAGTTTATGCGCTTCAAAGGTCCACATGGAGGAGCTCCGTGAACGGCCACGAGT
TCAGTCGAGGGCGAGGGCGAGGGGCCGCCCTACGGGGAGCGAACCAGCAAGCTAGAAGGTGATAAG
AGGTGGCCTGCCCTTCATACAGGACATCCTGTCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTAAAC
GCACCCGCCGACATCTGACTGCAAGCTGTCTTCCCCGAGGGAGCTTCAGAAGTGGGAGCGCGTGATG
AACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCGGGACGGCGAGTTCATCTG
AGTGAAGCTGCGCGGCGCAACTTCCCCTCCGACGGCCCCAGTAATGCAGAAGAAGACCACATGGGGC
TGGGAGGCCTCCTCGAGCGGATGTACCCGAGGACGGCGCCCCCTGAAGGGCGAGATCAAGCAGAGGC

TGAAGCTGAAGGACGGCGGCCACTTTCCGACGCTGAGGTCAAGACCACCTGAGGGACAAGAAGCCCG
TGCAGCTGCCACGGCGCCTACAATCAACATCAGGTTGGACATCACCTCCCACAACGAGGACTACACC
AATCGTGAACAGTACGAACGCGCGGGGCCCACTCCACCGGCAACCAGACCGAGCTGTACAAGTC
TAGATCTACCACCATGCTTACATTTTGTATTTCGTTACCCAATGCGGTGATGTAATAAAGGAGCAGGAA
ATTACGAAGGATTATGCTCTATATATTTATCTTTGACTATCCTCACTGAAGCTAATATTGGCAGAGTG
TTCAAAGATGCATATGGATAGATATGTTGAATATAAGGGATAAACTAGTGAGGAAAACCTGTAAAAGTT
AAAGTGATTAGAGTTGATTATACAAGGATATATAGATGTCAATTACAAAGGATGTGTGAACATCAATA
ATGACGGCCACCCTAACTTAATTTACTTGGCTTATAATGGTTACAAATAAACAATAACATCTAA
ATTCTAAATCGTATTTTACAATAAATTCCAAGTTGTTGGAGATCCTGGGGCTCCGTAACGTTTCATCTA
TGATTACGTTTTGTATCTTTAATTTGTTGTATCGACGACCGCGCTAGAGATTACAAGTTTGTTTTACAGG
TACTAAAATCTCTAACAACAGTGTTTACTCGTCTGAATGTTTAAACACAGCAGTAAACATAGCTGGCAC
GTATGCTTTTTGTTCCGGTGTTAATCCGCTATATGTTTTCTGTAGCGGCTGATAACACAGCATCAGCGC
TGAGCATCCGCGTCCGCAGGCACATATTTTAAACAGTGAGGTTACATCCATGGTTTTGTGCATATAAAG
AATTTCCGATTTCTATATCACATTTGTTTGACACTAACATTCGGCTTCTTGTCTAATTTAGACGAGATA
CGCGTTCGCTGGAGTGTATTCACCGTCGCTGTGTCTTGTGTCGGCCTCCATTTAAATAGCTACAATTG
GTATCCATATTGCAAGAGAGATAATAAACTGAGTAAAAATGCAATTTTAGGTCGAACGAAGAGTGTTA
ATATTATGTTGAACACTACTTTTGCTCTATTGACAGGAACAGTGAAAATCTATCACTATTGCTTAATCCAC
ATGACAATCTTAATGAGGAAGTTTATCCATCTGTAAGTTGTTACGCTAGTATTACATCGTACAACAGC
CCGCAAGTCCTAAAATTATTATAATTACGTGTTAGCAAGAAGTGCTGTGGCATTCCAGACACTCTGGA
TCCCAACATTCTCGAGGTTCCGCATATTTATTGACTCTTCTAATAACTTATCTCTAGTGGGATAACTAC
ATCTATATATTTCTGTTTAAAGTCCGCAGACTGTTGTCTTAGAATATAATCAATCATCTCTTTTGCTATC
TTCTGTATTGTGTGCGCGTAAATGATGCAAAAAATGATTCTATATTGGTACACTAGCATCTTTACTACA
TAATGTTTGCATCTTATTATACAGATTTATTAACGATTGTTGACCTCTACAGTTCTATTACTCCCTATTA
AAGGCTGAACCGATCACTGATGGCATATGTTTCTATCAGACGTATCCCCTGACACCAGTCGAATAAAT
CAACATCCTCATTTTTCAGTGTCTTGAACGTCTATTAACACGATTCTAATACTGCTTTTCTTCTACTT
ATCTGATATCTTTTATCCTTTTTTCAATAATGAGTACGATTAAAGTGCGACAGCAATTGGGGTGCAGTT
CCATTATACGATTGTCTTAACAATGCCGAAAGACCACCGGGTCTGTGTTGACTAATCTAAACTCTGGAT
ATCTTTTCTTTACTTTATCCTCTTTATCTTTTGCTAGTATTGGTCCTATATGCACATATTCTAAAAGCTTA
GCGGGGAACATATCGTCATGCATTTTATCCACGTTTAAATAACATCTCATCAGTGGGGTACTCCCGAGGCG
GATCCCGTTTAGGAGCTCAGCACTACTCCGCCACCTATTTATCTCATTGAAATTATTAATCTAAAAAC
GCCATAAAGATGTTGATCTTAAAGGATTGAACTCTATCCGAAAACAACATTCTAGAATGTTATCGTCA
TTATCATTACGATTCTAGTTTCAAAAACGATGACTCTCTTTTGAATCCTCGTAGTTTGTGAGACAGAA
TAGCTATTTTGAAGAAAACCTTTTGTAGTTCTTGAGAACATTCAGTCATAGAATATTCCTGGAAAATGCA
TCAGTATTACCAGAGGTCTTCATAATAATATTGTCATCTTTAAACATAATAGCCAGATGCTGATGCTGA
CTAATACATGATAAAGCCCAACAACCATTTCTAAATGAATGACAGTTTCTACCACAACATTTTTCTTCAT
GCGAAAATTAAACACATAAAGATTTCTTTTGATCTATACTTAGACAAATAGTAGTATCTGTCCTAATGA
CATCAATTCATATTACAATCGACACTTACACATGATACTAGAAAAGTTTACTAGGTATTTTTCAGATC
AGGTTCTATATCTATGATGGCATCATTGTAGAAAACACAAACACGATTTTACCCCTTGGACACAGGA
AGATTAACACAATGTTTCGGCTCCTCGGTAGAACACTGATGCATTTGAGAGAGGGTATAATAGACT
TTAAATGTTTACAAGATCCGCCAAGGCGGCAAAATACATTAACCTCATCCGTCGAGGTCAATTTATAGAC
ACTTCTTCACTGAACCTCTGAAAAATATGCACCTGGTAGCGCAGTTTATCGATTTTATACGGATGCTCA
TTTTAAATTTTTTGTAATTTATTTAAGTTAAATGGCTGCAGAACAGCGTCGTTCTACAATTTTGACATA
GTTTCAAATGCCATAATTGCAATCTGTATTGAATATATCTATTAATTCTGAATACATAAGTCCAAAGCT
AAACAATTGTGCTGTGTGCGCATCGAAAAGTCGGTGATTAATGGTATCTACAATTGTTGCGAGTCAAA
TATAGAAATAATGGACAAGAGCAGCTATTAATAATATTGGACAATCTTCGATGTCATTCGGCTCTATG
TATGTAACACCACAGATTTCTGGGGCTATATAATTCGTTAAAAACAGTTTACACTCATACTACCACATTC
TTTATACATAACAAGCCCACTATTCTAACACGCTAAACACTTTATTAACCTGATTTTTATCTAACAAGT
TGTGTATGCGGCAGAAATGGTAGAGTATCTGAAGACCAACTAGATTCATCAATAAATCAATGTCTCA
AGAGCTAGCAAGAATTGTGTGGAATGAAATATGCTCTCATTAACTGGTACAATATAGGATTTTGCC
AATGATCATCAGTGAGCCTATTATAGTAGCTGTGGATTTCTGGTTAAAACCAACTCTGATTATTCTGCA
GAAGTGGAAGGCTAATGGAACCTACCAGTTAAACCTGATATAGTGAATACCATGACTGTAGCCAGAA
AGGTATTGATACTAGCAACAATATGAAACAGAATATATAAACCGGCAGAAAATAGAAGGAAATTGAA
AAGGTAGAAAAATATTTACCAGAAGTTATATCTACAATTGCCAATAGTAATATAATAAAAAATAAAAA
TCTATCTTTCAGGCCAATATCAACGATAAACAGATCATGGAATGCTCTGAATGCTAGACACGAGTGAG
AAATACTCTGGGATACTCGAAACTGATGGAACGTGTGACTAGTCCATTGACGGGAAATAATACAATTAC
AACATTTATACCAATTTCTGCGTCCGATATGCAAAAGTTTACCATTTTAGAATATCTTTACATTATGAG

GAAATTGATGGCAAACAACGTTAAGAAAGAACGAGAGGAAAAACAACGGAGGGTAGTTATGCATAT
 TAACTCACCCCTTTAAGGTAATCAATTTGCCAAATGTTAAAATTAGAATCATATATCTTTATCCATCTAT
 TATCGTCTGTGAAGTTTATCAGAAAATATCTTTACCTAATATTCGTTTTAGACGATGTTTCAACACCGG
 CATTTTATAACAGTCAGCTACCAATTTAAAACAATACATTCTATTATGACCAAATCCATACGGAATAC
 CCAATAAAATTGGCTTGTATCAGCTGCTATTTTCATAACTTGAATATTCTCTAACATGTATGCTTTTGCA
 CTAAGATATCCTTCAAAAACCTATCCAATGGGACATTCAAACCTCGCACTATGGGTAGATTCAAGCTACCA
 AGTGTTCCGATAATAACGTACTATAATAAATACTTATGTTTTTCTATCGATGGATTCAAATGACTTGT
 TAATGACATGGTACTCAATTAATAAAATAATAGAACCTCTAGGCGCATATGTCTTGATAAATTAACCG
 GATCCATATTAGTTTATATCCAATAGAGGTTGCACTGTTAAAGTTACGTTGAGGTTCTATTTGTGTGAA
 TATAATTTTAACCTACAACATTCGTTCTTTACACTGACAGTACATCCTTTAATGATTTTTCGTATAATCA
 ATTTATATCGTTTCGTGAGAAATATCTTTACCGCATGTAGAGGCACACGGTTGAACACCATTTTAATCTG
 GTAACCTATATTTATTATATTATCACTTTTAGTATTGGTTGATGACAGACTATCGTTGAAGTAAAAGA
 ATTATTTATACTCAGTAGTAAATTATCATATAGATGCGAATACTATCATTATTTACAGTTTTCGTTGGG
 CTTAATCCGAAGTCGAGGATATAGATTTTCAACATCGATATCTGATTTAGTATTTTCGAAATCTATATA
 AGATGATAAAGTTTCTTCTGAAACATCAATCTCCTTGCGTGGTTCTAGCGCGTACATTAATTCAACAGG
 GTAACGGAACGGTCTCTAAATAGATGATTTAAAAGAATTGTTCCATTTTCATCTAATGCGGAAATATA
 AATTCTTGAAACAGTGCATTGACAAGGAGGAACATACAGTAGACAT