Coding and non-coding elements comprise a regulatory network controlling transcription in Kaposi's sarcoma-associated herpesvirus

David W Morgens[1*], Leah Gulyas[1], Alejandro Rivera-Madera[2], Annabelle S. Souza[2], Britt A Glaunsinger[1,2,3*]

[1]Department of Plant and Microbial Biology, UC Berkeley, Berkeley, CA, USA; [2]Department of Molecular and Cell Biology, UC Berkeley, CA, USA; [3]Howard Hughes Medical Institute, UC Berkeley, CA, USA

*Corresponding authors: glaunsinger@berkeley.edu, dmorgens@berkeley.edu

Abstract

Gene regulation in eukaryotes relies on many mechanisms for optimal expression, including both protein transcription factors and DNA regulatory elements. CRISPR-based screens of both protein coding genes and non-coding regions have allowed identification of these transcriptional networks in human cells. Double-stranded DNA viruses also invoke human-like regulation to control transcription of viral genes that are required at different stages of the viral lifecycle. Here, we applied CRISPR-based tools to dissect regulation of a viral gene at high resolution in the oncogenic human herpesvirus Kaposi's sarcoma-associated herpesvirus (KSHV), whose compact, densely encoded genome provides unique challenges and opportunities for studying transcriptional networks. Through a combination of CRISPR-interference (CRISPRi) and Cas9 nuclease screening, we mapped a novel regulatory network comprised of coding and noncoding elements that influence expression of the essential KSHV protein ORF68 at early and late stages of the viral lifecycle. ORF68 encodes an essential protein involved in packaging the replicated viral DNA into nascent capsids. Although ORF68 expression initiates early in the viral lifecycle, we found that it is primarily required at later times. This work demonstrates the ability to exhaustively identify features controlling a given locus, capturing a complete viral regulatory circuit that functions within the human nucleus to control transcription.

Introduction

Complex transcriptional control is a conserved feature of eukaryotes, as well as the viruses that infect them. Underlying the lifecycle of all DNA viruses is a highly regulated cascade of viral gene transcription. In human herpesviruses, many of these genes are transcribed in a host-like

manner; nuclear dsDNA viral genomes can be chromatinized by human histones[1], decorated with enhancer marks[2], driven by human transcription factor binding[3], and even form transcription-associated domains with human CTCF and cohesion[4,5]. Exhaustively identifying and characterizing these regulatory features can both help us understand the biology of human pathogens such as Kaposi sarcoma-associated herpesvirus (KHSV) – a major cause of cancer in AIDS and other immunocompromised patients – as well as use these as models for transcriptional regulation in human cells.

Numerous examples of non-coding regulatory sequences have been found in DNA viruses. Enhancer sequences include the first described enhancer on SV40[6], the EIIA enhancer on adenovirus [7], and the major immediate enhancer element in the betaherpesvirus human cytomegalovirus[8,9]. Other viral regulatory elements act through expression of non-coding elements. In the gammaherpesvirus MHV68 there are tRNA-like elements that control latency and egress[10,11]. And from KSHV itself, many different functional elements beyond coding mRNAs are transcribed: including miRNAs that regulate cancer phenotypes[12,13], origin RNAs that regulate viral DNA replication[14,15], circular RNAs[16,17], and long ncRNAs with various functions[18–22]. While these individual events have been explored to different degrees, systematic searches for regulatory sequences have been limited by traditional methods that rely on deletions to perturb noncoding elements, which in the densely encoded KSHV genome may have unintended effects on surrounding elements.

For the human genome, the discovery of functional regulatory sequences has been greatly accelerated by the use of CRISPR interference or CRISPRi [23–27]. Notably, by recruiting repressive chromatin regulators to DNA, CRISPRi can repress gene expression not only through proximal promoter elements but also via more distal, enhancer elements. While this tool has been applied widely on the host, it can also effectively repress transcription from the KSHV genome[28], and thus has the potential to provide deep insight into the components and structure of viral transcriptional networks.

Here we combine CRISPRi with a library of guide RNAs densely tiling the KSHV genome, allowing a thorough interrogation of potential regulatory activity controlling the expression of a single viral gene, ORF68. We selected ORF68 as a proof of principle for this study as it plays an essential role late in the viral life cycle during packaging of new viral DNA[29,30]. Although its transcriptional regulation is largely undefined, ORF68 is expressed much earlier than would be required for its role in packaging[31,32]. By complementing the CRISPRi approach with a Cas9 nuclease screen, we identified functional promoters that control expression through their

associated coding regions, as well as non-coding regulatory elements. Notably, targeting these non-coding regions allowed us to specifically inhibit ORF68 early in the viral life cycle, revealing that its expression appears dispensable early but is required late in the KHSV lifecycle. These findings illustrate the potential of this approach to map viral gene regulatory networks on a genome wide scale.

Results

CRISPRi tiling identifies regulatory regions across the viral genome

To identify a regulatory network in the KSHV genome, we tested the effect of each viral locus on a protein reporter of ORF68, an early gene involved in packaging[29,30]. Cas9-KRAB (CRISPRi) was delivered lentivirally to iSLK cells latently infected with a copy of the BAC16 KSHV that incorporated a HaloTag fused to the N-terminus of ORF68 at the endogenous locus (HaloTag-ORF68)[33]. We then delivered a library of guide RNAs densely tiling the KSHV genome with an average of one guide every 8 base pairs[33]. After four days, the virus was reactivated to the lytic cycle by addition of doxycycline and sodium butyrate as previously described, and cells were treated with a fluorescent Halo ligand to monitor ORF68 protein levels. 24 hours post-reactivation, cells were fixed and sorted for high and low ORF68 protein expression (**Figure 1a**). By sequencing the sgRNA locus from both populations, we calculated an average guide enrichment from two replicates. A negative value signifies that the guide was enriched in the low ORF68 signal population, indicating that CRISPRi recruitment to that locus inhibits the expression of ORF68. Similarly, a positive value signifies that the guide was enriched in the high ORF68 signal population, indicating that CRISPRi recruitment to that locus increases expression of ORF68 (**Figure S1a,b**).

To identify regions of the viral genome that affect ORF68 expression, we performed a sliding window analysis and identified 8 loci with target guide RNA scores that differed significantly from negative controls (**Figure 1b; Supplementary Data 1, 2**). Our strongest signal corresponded to a peak poised immediately upstream of HaloTag-ORF68 itself, confirming successful transcriptional inhibition of the locus by CRISPRi. We observed that the center of most other peaks also corresponded with known transcriptional start sites (TSSs); since we expect CRISPRi to work primarily by impeding transcription, these peaks are named by their nearest TSS for simplicity (**Figure S1c-h**). For example, we find one peak near the TSS of ALT, a lncRNA of unknown function which runs antisense to many genes expressed during latency[21],

that we will refer to as TSSalt (**Figure S1e**). The exceptions include one peak that roughly maps to the ORF50 coding locus (**Figure S1c**) – likely targeting the exogenous copy of ORF50 used to reactivate the virus – as well as an additional peak which does not correspond to a previously described TSS (**Figure S1h**) [34] but is located near the shared polyA sites of ORF75, ORF74, K14, K15, and the ALT lncRNA[21,35]. We will refer to these as ORF50 and polyA75 respectively.

We next validated these regions' effect on ORF68 by delivering a pool of three guides per locus and measuring the percentage of cells expressing HaloTag-ORF68 at two timepoints post-reactivation (**Figure 1C**). These validated our screen results at the protein level, but we did note that some of the effects on ORF68 levels were lost at later timepoints, such as when targeting the TSSalt. These may reflect changing regulatory events as the viral life cycle progresses and provide an opportunity to study the effects of suppressing ORF68 expression early in the viral life cycle without effecting its known function in later steps. While guides targeting the EF1a promoter displayed the expected effect as well, we also noted a strong toxicity of these guides, most likely through their association with the BAC-encoded EF1a-EGFP-HygroR selection locus or the possible inhibition of the host EF1a site; we thus excluded this peak from further analysis.

Despite the correspondence of most peaks to TSSs, clearer identification of important underlying regulatory features was made impossible by the large observed footprint of CRISPR, likely due to the spread of KRAB-induced heterochromatinization[36], possibly exasperated in the viral lytic environment or in the presence of sodium butyrate. Illustrating this, the TSS68 peak is comprised of guides targeting not only the entirety of the ORF68 coding region, but also many surrounding genes (**Figure 1D**). This equates to approximately a 2-5 kb window of CRISPRi repression from a single guide, which is mirrored at other loci (**Figure S1D**). This prompted us to further interrogate the global transcriptional effects of CRISPRi at each regulatory region on the viral genome.

CRISPRi recruitment to the viral genome inhibits many genes locally

To query overall changes in transcription during CRISPRi at our target loci, we performed polyA+ RNA-seq at 24 hours post-reactivation on the previously validated guide pools (**Figure S2a**). Examining viral RNA expression relative to our negative controls, we observed that our regulatory loci act on the mRNA levels of ORF68, causing changes consistent with those observed at the protein level (**Figure 2a, Figure S2b,c, Supplementary Data 3**). We observed that CRISPRi targeting ORF50, TSS75, or TSS57 caused a dramatic downregulation of nearly

all viral genes, aligning with the critical roles of ORF50, ORF75, and ORF57 protein in the viral lifecycle. ORF50 is a master regulator of KSHV lytic reactivation[37], ORF75 is required to prevent innate immune suppression of viral gene expression [38,39], and ORF57 has many reported viral functions including the export of viral mRNA from the nucleus[40].

In contrast, guides targeting TSS68 and TSSalt have a limited global effect, only strongly inhibiting a small number of genes each. We find the strongest downregulated genes are within the local region of the guides (**Figure 2b,c**), suggesting again that CRISPRi inhibits transcription in region of approximately ~2.5 kb around the targeting site. This appears true even when the effect of CRISPRi increases ORF68 mRNA expression, with targeting the TSS72 region causing a global increase in viral gene expression (**Figure 2a**) despite still locally decreasing transcription (**Figure 2d**). Many of these locally repressed regions include ORF73 or LANA, whose expression is required for latency maintenance[3,41]; to test whether this effect could explain our results we grew our knockdown lines in the absence of hygromycin selection and observed a more rapid loss of latency when targeting TSS75 or polyA75 (**Figure S2d**). Thus, RNA-seq corroborates our screening data, and reveals both local and global effects of CRISPRi at our target loci.

The regulatory regions we identify affect both protein and RNA levels; to test whether this results from regulation of transcription itself we used metabolic labeling of nascent RNA. Incubating the cells with ethynyl uridine (EU) for two hours at 24 hours post-reactivation, we measured ongoing transcription of ORF68 and ORF75 (**Figure 2e**). As expected, targeting TSS68 or TSS75 had a dramatic effect on transcription of their respective RNAs, and across all loci we observe a reduction or increase in transcription consistent with total RNA levels. The exception is TSS57 where we see no significant effect on ORF68 transcription at 24 hours, which may indicate the small reduction in ORF68 total RNA and protein levels are acting either post-transcriptionally at the RNA level or represents accumulated transcriptional regulation occurring before the 24-hour timepoint measured here. Overall, the loci identified are regulating transcription of viral genes.

Viral knockouts identify associated coding features

While CRISPRi at a given locus may effectively suppress multiple viral genes, not every gene will be responsible for the observed effect on ORF68 expression. For example, while guides targeting the TSS68 repress expression of ORFs 65, 66, 67, 68, and 69 (**Figure 2b**), downregulation of the ORF68 reporter is most likely due to direct repression of the ORF68

promoter. Furthermore, CRISPRi alone is unable to distinguish how regulatory regions may act to control ORF68 expression, whether by preventing protein-coding transcription of a regulator or by impeding a non-coding mechanism. To overcome these limitations, we directly assessed the role of coding loci underlying each regulatory locus by performing a CRISPR nuclease tiling screen in the same HaloTag-ORF68 reporter line.

In previous screens, we have reported a strong background effect, where targeting any locus on the viral genome with Cas9 nuclease resulted in a decrease in reporter expression[33]. Here, we observed unexpectedly that this background effect was variable across the viral genome, with targeting to the region upstream of the ORF68 coding region having a stronger effect on reporter expression than targeting downstream (**Figure S3a, Supplementary Data 4,5**). The reason for this difference is unknown (possibly a local effect of DNA damage on the viral genome), but to adjust for the differences in local background we used a boundary method to identify coding regions: for each candidate coding exon, we compared the enrichment of the coding region to the immediate adjacent non-coding region, and considered the coding region a hit if the boundaries were both significantly increased or both significantly decreased.

This yielded an exhaustive list of viral coding regions that control ORF68 expression at 24h post-reactivation: ORF50, HygroR, ORF68, and ORF75 (**Figure 3a-c; Figure S3b**). Disruption of these coding sequences had similar effects on ORF68 expression as CRISPRi repression at their corresponding loci (EF1a for HygroR). Loss of functional protein is thus likely the mechanistic explanation for these regions' effect on ORF68. However, for the other CRISPRi loci, the nuclease screen identified no corresponding effect of disrupting coding regions. For example, despite RNA-seq data showing that targeting TSS57 and TSS72 does result in inhibition of ORF57 and ORF72 mRNA expression respectively (**Figure 2a**) – along with other local genes – we observe no effect on ORF68 of targeting these coding regions (**Figure S3c,d**). ORF72 (also known as vCyclin) encodes a homolog of human protein cyclin D, plays a role promoting oncogenesis during latency[42,43], and has no known role in the viral lytic cycle. The TSSalt region has no clear association with the promoter of a coding region, and indeed we find no evidence of a nearby coding element responsible for its effect on ORF68 expression. PolyA75 is adjacent to the ORF75 coding locus, but no other coding elements are identified that could explain its activity. These regions with no associated coding region thus likely act through noncoding elements that are unable to be efficiently disrupted by Cas9.

To test whether the identified coding regions have the expected effect on the mRNA levels of ORF68, we next cloned pooled guides for nuclease-targeting at the coding regions of ORF50

and ORF75, including ORF57 to validate the negative result. ORF50 and ORF75 pools had the expected effect of decreasing ORF68 protein levels (**Figure 3d**) as well as corresponding depletion of ORF68 RNA (**Figure 3e**). We noted that targeting ORF50 also reduced RNA levels of ORF75. This allows us to hypothesize a regulatory model of protein coding elements where ORF50 regulates ORF75 which, in turn, regulates ORF68. To test this model directly, we transfected a reporter plasmid containing 230 bp of upstream of the ORF68 start codon driving a HaloTag reporter along with an expression plasmid of either ORF50 or ORF75 into HEK293T cells. We saw a significant increase in expression from the ORF68 promoter only with ORF75, consistent with our model (**Figure 3f**). We also observed that targeting either ORF50 or ORF75 reduced RNA levels of ORF57 (**Figure 2e**). Interestingly, ORF57 pools reciprocally decreased ORF75 RNA levels, yet ORF57 protein disruption had little or no effect on ORF68 mRNA levels – as predicted by the nuclease screen – suggesting an unclear regulatory relationship.

Mapping regulatory targets of noncoding elements

Since we cannot identify coding elements to directly explain all CRISPRi peaks, these seemingly noncoding loci could instead act in a distal manner, indirectly impacting expression of one of the three viral proteins that affect reporter expression. We therefore returned to our RNA-seq data and measured the correlation between each sample (**Figure 4a**). We hypothesized that despite local effects of CRISPRi, the regulatory regions would correlate most strongly with their regulatory target. TSSalt and polyA75 most strongly correlated with TSS68 and TSS75 respectively, suggesting that these regions promote ORF68 and ORF75 expression. Conversely, TSS72 had a strong anticorrelation with polyA75 and TSS75. Given that recruitment of CRISPRi to TSS72 causes an increase in viral gene expression, this suggests that TSS72 acts to repress expression of ORF75. TSS57 had only weaker correlation with both ORF50 and the TSSalt, preventing any firm conclusion. We can thus use these correlative regulatory interactions along with data from **Figure 3** to create a model regulatory network consisting of both coding and non-coding elements controlling ORF68 expression (**Figure 4b**). Of note, as we performed these experiments while overexpressing ORF50 from an exogenous promoter, we were likely unable to detect any regulation of the viral copy of ORF50.

Finally, we evaluated how the components of this ORF68 regulatory network impacted virion production in KSHV infected cells using a supernatant transfer assay. BAC16 derived KSHV expresses GFP, which enables quantification of infected recipient HEK293T cells by flow cytometry. As expected, targeting any of the coding elements via their promoters (TSS57, TSS68, TSS75) caused a severe loss in infectious virion production (**Figure 4c**). Targeting

polyA75 also reduced virion production, albeit more modestly, consistent with the regulatory network. In contrast, guides targeting TSS72 or TSSalt did not negatively impact virion production (**Figure 4c**). This was expected for TSS72, which was predicted by the model to reduce ORF68 expression. It is notable that targeting TSSalt specifically disrupts ORF68 expression at the early 24h but not the late 48h late time point (see **Figure 1c**). Thus, the observation that targeting the TSSalt locus did not impair virion production suggests that while ORF68 expression is essential late in infection, it may be dispensable at early time points.

Discussion

Here, we use CRISPRi and a tiling library to identify regulatory events across the KSHV genome that control the expression of the early gene ORF68. Our analyses indicated that CRISPRi on the viral genome effectively suppresses gene expression in a ~2-5 kb local region during lytic reactivation and in the presence of sodium butyrate. In the densely encoded herpesviral genome, this footprint covers multiple elements, any of which could be responsible for the observed regulatory activity. We thus used a complementary Cas9 nuclease screen to disrupt the coding elements in these regions and found that for many of these regions their activity must be explained by non-coding elements that cannot be disrupted by Cas9. Using RNA expression patterns to predict functional targets of these elements, we present a complete model for viral regulation of ORF68 and demonstrate that these regulatory elements play key roles in completion of the KSHV lifecycle. By perturbing specific regulatory elements that selectively disrupt ORF68 expression early in the lytic cycle, we conclude that ORF68 expression is primarily required at late times, in accordance with its established role in viral genome packaging.

Our functional model does not make predictions for the mechanistic nature of these regulatory events. For example, the coding elements may or may not directly regulate transcription on the viral genome. Indeed, while ORF50 has been previously characterized as a viral transcription factor[37,44], ORF75 more likely acts indirectly by preventing inhibition of viral transcription by host factors[45]. Similarly, the non-coding elements could involve ncRNAs, microRNAs, or enhancers, as CRISPRi is expected to be able to inhibit each of these; though as we see a direct effect on transcription (**Figure 2e)**, we favor an enhancer model. Given the density of the viral genome and the width of CRISPRi footprint, it is possible a peak corresponds to multiple regulatory events or that two nearby peaks both inhibit the same regulatory locus. The TSSalt here could be an example of the former, as there are nearby K12 TSS and miRNA loci (**Figure S1e**) while the polyA75 peak could be an example of the latter as it behaves very similarly to the

nearby TSS75 peak. Though with the viral density and propensity for multiple functions encoded within the same locus, these issues may be intrinsic to the study of gene regulation on the viral genome, but it is possible that other functional perturbations such as using base editors or dCas9 could allow us to pinpoint the underlying regulatory sequence.

Our model is unable to place ORF57 protein and the TSS57 into the network. Our observation that Cas9-based disruption of ORF57 did not affect ORF68 mRNA or protein suggests that ORF57 protein does not regulate ORF68. This is in agreement with previous reports using an ORF57 deletion virus, which also indicated little regulatory effect on ORF68[46]. ORF57 protein does have a broad effect on viral transcription, but this is likely through direct effects on DNA replication components, upon which the majority (but not ORF68[31]) of gene expression is partially dependent. It is possible that the TSS57 promoter contributes to ORF68 regulation independent from the ORF57 protein, as many human promoters act as enhancers for distal genes[47]. However, we have not yet been able to test this hypothesis because we cannot disrupt the promoter and its hypothetical distal activity using CRISPRi without also disrupting the protein levels of ORF57.

Our approach demonstrates the power of combining CRISPR screening tools for the discovery of viral gene regulatory networks. Given that KSHV is a double stranded DNA virus, it incorporates the same spectrum of regulatory mechanisms as the host – transcription factors, non-coding RNAs, enhancers, and DNA structural elements. By studying these networks on the viral genome, we can thus learn both how these regulatory mechanisms function and, ultimately, how the virus controls gene expression under diverse cellular conditions and in various cell types.

Figure legends

Figure 1. CRISPRi screen identifies novel viral regulatory regions. A) Schematic of screen. The viral genome encodes a constitutive fluorescent marker (green) and a HaloTag-ORF68 fusion (blue). B) Summary of results from the CRISPRi screen. X-axis identifies the genome coordinate on the BAC16 KSHV genome. y-axis represents the log-transformed p-value of each locus relative to negative control distribution. Red dotted lines indicate the location of transcriptional start sites. Blue dotted line identifies the two peaks that do not associate with a TSS. C) Validation of pooled guides targeting each peak. Three guides were used to target each locus identified on the x-axis. Y-axis displays the relative percent of cell expressing the HaloTag-ORF68 for two time points. Error bars represent standard error of four replicates from independent reactivations. D) Enrichment of individual guides at the ORF68 locus. Each dot represents a single guide, with the target location displayed on the x-axis and the average enrichment from two replicates on the y-axis. Arrows represent coding regions of ORF68 (in blue) and surrounding genes in grey.

Figure S1. Supplementary screen data. A) Enrichment of individual guides at the ORF68 locus. Each dot represents a single guide, with the target location displayed on the x-axis and the average enrichment from two replicates on the y-axis. B) Reproducibility of guide enrichments from two replicates. C-H) Smoothed enrichment of guides at indicated locus with annotated transcription start sites[34]. For each guide, the median enrichment of a 100 bp window centered at the target locus was calculated along with an interquartile range (IQR) to represent the range of values. Median value is shown as a point and IQR is shown as shaded region. Regions significant ($p<10^{-11}$) in sliding window analysis are shown in blue. NA indicates unannotated TSSs.

Figure 2. Local effects caused by CRISPRi. A) Heat map in genome order of changes to viral gene expression at 24 hours relative to vSAFE negative controls, with ORF68 pulled out and presented at the top. Average of three replicates. B-D) Change in RNA-level of each viral gene relative to vSAFE negative controls in genome-order. Genes whose start codons are within 2.5 kb of at least one guide in targeting pool are highlighted. E) Nascent RNA expression at 24 hours post-reactivation from two-hour treatment with EU. Values are presented relative to parental cells, and error bars are standard error from three independent replicates.

Figure S2. *Supplementary for RNA-seq*. A) Schematic showing set-up of RNA-seq experiment on CRISPRi cells infected with a three-guide pool, reactivated, and polyA+ RNA-seq at 24 hours

post-reactivation. B) Heatmap indicating viral gene expression relative to the matched vSAFE replicate of each individual replicate. Replicate from three independent reactivations. C) RT-qPCR after knockdown of individual elements. Error bars represent standard error of four independent reactivations. D) Effect on latency measured by loss of virally encoded EGFP expression over 10 days. Values are presented relative to parental cells, and error bars are standard error from three parallel replicates. P-values are calculated by t-test.

Figure 3. Knockout screen maps associated coding regions. A-C) Median smooth enrichments from Cas9 nuclease screen at associated coding locus. Dotted lines indicate exon boundaries. For each guide, the median enrichment of 500 bp window centered at the target locus was calculated along with an IQR. Median value is shown as a point and IQR is shown as shaded region. Regions were considered significant if the guides on each side of the boundary were significantly different and in consistent directions. D) Percent of cells expressing HaloTag-ORF68 24 hours post-reactivation for the indicates pool of coding region-targeting guides. Values are averages of four independent replicates and error bars represent standard error. E) RT-qPCR of indicated viral gene at 24 hours post-reactivation. Relative to 18S and vSAFE. Error bars show standard error from four technical replicates. F) Average fluorescence from HEK293T cells transfected with an ORF68-promoter-driven HaloTag and a plasmid expressing the indicated viral protein. Error bars are standard error from seven independent replicates.

Figure S3. Coding screen supplement. A) Enrichment of individual guides. Each dot represents a single guide, with the target location displayed on the x-axis and the average enrichment from two replicates on the y-axis. Red dotted lines represent the median guide enrichment for the two regions indicated. B-D) Median smoothed data from coding screen at indicated locus. For each guide, the median enrichment of a 500 bp window centered at the target locus was calculated along with an IQR. Median value is shown as a point and IQR is shown as shaded region. B) For the EGFP-HygroR locus, location of functional units shown in color. C,D) Exon boundaries are marked by dotted line.

Figure 4. Mapping regulatory network by effect on viral transcription. A) Co-correlation matrix of RNA-seq data. For each indicates sgRNA pool, Pearson correlation was calculated between each pool. Data from Figure 2A. B) Model of regulatory events controlling transcription of ORF68 locus. C) Supernatant transfer after knockdowns. Error bars represent standard error from six independent reactivations.

Figure S4. Supplementary mapping data. A) Individual replicate correlation among RNA-seq.

Methods

Plasmid and oligos

pMD2.G (Addgene plasmid # 12259), pMDLg/pRRE (Addgene plasmid # 12251) and pRSV-Rev (Addgene plasmid # 12253) were gifts from Didier Trono. pMCB320 was a gift from Michael Bassik (Addgene plasmid # 89359). lentiCas9-Blast was a gift from Feng Zhang (Addgene plasmid # 52962). Lenti-dCas9-KRAB-blast was a gift from Gary Hon (Addgene plasmid # 89567). Sequences used are listed in **Supplementary Data 6**.

Cell culture and plasmids

iSLK cells and Lenti-X 293T (Takara) were maintained in DMEM (Gibco +glutamine, +glucose, -pyruvate) with pen-strep (Gibco; 100 I.U./mL) and 1X Glutamax (Gibco) along with 10% FBS (Peak Serum). iSLK cells were maintained in 1 ug/mL puromycin, 50 ug/mL G418, and 125 ug/mL hygromycin B (Gibco). Cas9+ and CRISPRi+ cells were maintained in 10 ug/mL blasticidin. 0.05% Trypsin (Gibco) was used to passage cells. All cells were maintained at 37C and 5% CO2 in a humidity-controlled incubator. Lenti-X 293T cells were obtained from the UCB Cell Culture Facility.

Generation of CRISPRi Halo-ORF68 iSLK line

iSLK line latently infected with a copy of BAC16 containing a HaloTag-ORF68 fusion as the endogenous locus were lentivirally infected with dCas9-KRAB (CRISPRi). Briefly, Lenti-X cells were transfected with third-generation lentiviral mix (pMDLg/pRRE, pRSV-REV, pMD2.G) and dCas9-KRAB (blastR) with polyethylimine (Polysciences). Supernatant was harvested at 48 and 72 hours and 0.45 um filtered before applying to iSLK cells for 48 hours. This process was then repeated to ensure high CRISPRi expression.

CRISPRi screening and analysis

Library of guides tiling the KSHV BAC16 genome was delivered lentivirally to the CRISPRi+ iSLK cells above. Four days afterwards cells were reactivated with 5 ug/mL doxycycline, 1 mM sodium butyrate, 10 nM JF646 Haloligand (Promega). 24 hours post reactivation, cells were fixed in 4% PFA, and sorted for high and low ORF68 expression using a BD Aria II. Cells were then unfixed overnight in 150 mM NaCL and 60C and 50 ug/mL protease K (Promega). DNA was then extracted using a single column of QIAamp DNA Blood Mini Kit (Qiagen), following the manufacturers protocol and adjusting initial reaction volume. sgRNA locus was then amplified

and library adaptors ligated as previously described[48]. Libraries were sequenced on an Illumina NovaSeq 6000.

Counts for individual guides were converted to enrichment scores by calculating the log2 ratio of counts between high and low populations relative to the median negative control. Enrichment values from two replicates were averaged. To calculate significance of a given window, a 500 bp sliding window was used, comparing the enrichment of each guide and a 500 bp neighborhood to the enrichment scores of all negative controls using a Mann-Whitney test. An arbitrary p-value cutoff was used to identify peaks.

Individual guide delivery and characterization

For each pool of sgRNAs, independent lentiviruses were produced as above, then pooled and applied to CRISPRi-positive HaloTag-ORF68 iSLK cells for 48 hours. For protein analysis, cells were reactivated in the presence of 10 nM JF646 Haloligand (Promega) with doxycycline and sodium butyrate as above. Cells were then analyzed for fluorescence at 24 and 48 hours post-reactivation from four independent reactivations on a BD Accuri C6 plus.

For latency analysis, cell lines were maintained in triplicate with blasticidin, puromycin, and G418 but in the absence of hygromycin. Cells were split every 48 hours and GFP levels were measured by a BD Accuri C6 plus. The percent of GFP positive cells on day 11 was normalized to the percent positive on day 1, and this ratio was again normalized to the loss of GFP observed in parental cells without a guide RNA.

For RT-qPCR analysis, RNA was extracted at 24 hours post-reactivation using a RNeasy Plus Micro kit (Qiagen), treated with DNAase I (Lucigen), and reverse transcribed using AMV RT (Promega) and random 9-mers in the presence of RNasin (Promega). qPCR was then performed on a Quantstudio 3 using the indicated targets and with iTaq Universal SYBR Green (BioRad). RQ values were calculated using a standard curve. Results are from four independent reactivations.

For nascent RNA analysis, the manufacturer's protocol was followed from the Click-iT™ Nascent RNA Capture Kit (Thermo). Briefly, cells were treated 24 hours post-reactivation with 200 uM 5-ethynyl uridine (EU) for two hours. RNA was then extracted using a RNeasy Plus Micro kit (Qiagen) and biotinylated via a click reaction. The reaction was then purified by ethanol precipitation and biotinylated RNA was enriched using Dynabeads MyOne Streptavidin T1 beads (Thermo). Reverse transcription was performed on bead with the SuperScript VILO cDNA

Synthesis Kit (Thermo), and qPCR was then performed on a Quantstudio 3 with iTaq Universal SYBR Green (Biorad). RQ values were calculated using a standard curve. Results are from three independent reactivations.

RNA-seq and analysis

RNA samples from above were sent for library preparation and sequencing at the QB3-Berkeley Genomics core labs (RRID:SCR_022170). Total RNA quality as well as poly-dT enriched mRNA quality were assessed on an Agilent 2100 Bioanalyzer. Libraries were prepared using the KAPA RNA Hyper Prep kit (Roche KK8581). Truncated universal stub adapters were ligated to cDNA fragments, which were then extended via PCR using unique dual indexing primers into full length Illumina adapters. Library quality was checked on an AATI (now Agilent) Fragment Analyzer. Library molarity was measured via quantitative PCR with the KAPA Library Quantification Kit (Roche KK4824) on a BioRad CFX Connect thermal cycler. Libraries were then pooled by molarity and sequenced on an Illumina NovaSeq 6000 S4 flowcell for 2 x 150 cycles, targeting at least 25M reads per sample. Fastq files were generated and demultiplexed using Illumina bcl_convert and default settings, on a server running CentOS Linux 7.

Sequencing quality was assessed with MultiQC and reads were preprocessed with HTStream version 1.3.0 including deduplication. Genome indices were prepared using STAR 2.7.1a. The human GRCh38.p13 genome assembly was indexed with Gencode v43 annotations. Due to overlapping transcripts on the KSHV-BAC16 genome, individual exon coordinates were assigned to the corresponding parent transcript. Preprocessed reads were then aligned using STAR and counts files were generated for transcripts. Any transcript with no reads in all replicates and conditions was eliminated from further analysis. Reads from *E.coli* genes on the BAC were also removed. Raw viral counts were normalized to total input reads for each sample and subsequently normalized to within replicate vSafe condition values. Correlations and heatmaps were generated using the matplotlib, pandas, and seaborn packages in Spyder 5.3.3.

CRISPR nuclease screen and analysis

Cas9 screen was performed as the CRISPRi screen above using a Cas9+ iSLK cell line infected with a copy of BAC16 containing a HaloTag-ORF68 fusion. The library was then amplified using staggered primers[49] and a modified amplification protocol previously described[50]. qPCR was used to determine cycle number at ¼ CT. All PCR product was run over a single Minelute column (Qiagen) for each sample. Size selection was performed using a gel extraction (Thermo). Library was sequenced on Illumina NextSeq 2000 with a 150bp single-

read using Illumina sequencing primers. Adaptors were removed using cutadapt[51], and reads were aligned to library using bowtie[52].

Log2 values were calculated as above, and enrichment values were averaged from two replicates. For each exon boundary, a window of 500 bp on one side was tested against a 500 bp on the other to calculate p-values using MW test. Median values were used to determine the sign of the shift. Exons with p values < 0.001 and consistent signs were considered hits.

RT-qPCR

RNA extractions were performed using RNeasy Micro Plus Kit (Qiagen). Samples were treated with DNase I (Lucigen). Reverse transcription was performed using AMV RT (Promega) in the presence of RNasin (Promega). qPCR was performed with iTaq Universal SYBR Green (BioRad) on a Quantstudio 3. Relative quantities were calculated using a standard curve.

Transfection of ORF68 reporter

HEK293T cells were transfected using PolyJet In Vitro DNA Transfection Reagent (SignaGen) with a plasmid containing 240 basepairs upstream of ORF68's start codon[29] driving a HaloTag. Cells were cotransfected with either a plasmid expressing ORF75 or ORF50. 24 hours later, cells were treated with JF646 Halo Ligand (Promega) for 24 hours before analysis on a BD Accuri C6 plus. Cells were then gated for 646 signal, and the average fluorescence for the 646 positive cells was calculated. Data are from seven independent replicates.

Supernatant transfer

Cells were reactivated with 5 ug/mL doxycycline and 1 mM sodium butyrate. 72-hour post reactivation, supernatant was filtered through a 0.45 um PES filter and applied to naïve HEK293T cells for 24 hours. HEK cells were then counted on a Accuri C6 Plus (BD) and percent GFP positive was used to calculate infection.

Supplementary Data

Supplementary Data 1. CRISPRi count files. Raw counts for sequencing from CRISPRi screens.

Supplementary Data 2. CRISPRi screen results. Processed values for CRISPRi screen.

Supplementary Data 3. RNA-seq counts. Viral counts for RNA-seq.

Supplementary Data 4. Cas9 count files. Raw counts for sequencing from Cas9 screens.

Supplementary Data 5. Cas9 screen results. Processed values for Cas9 screens.

Supplementary Data 6. Sequences. Sequences used.

Acknowledgments

The authors would like to thank the members of the Glaunsinger lab for their feedback and support, and Dr. C. Kimberly Tsui for the stagger primer sequences. Cell lines were obtained from the UCB Cell Culture Facility which is supported by The University of California Berkeley (SCR_017924). RNA-seq was performed by the Berkeley functional genomics core (QB3 Genomics, UC Berkeley, Berkeley, CA, RRID:SCR_022170). Flow cytometry and FACS were conducted at the CRL Flow Cytometry Facility. We thank Hector Nolla and Alma Valeros of the UC Berkeley Cancer Research Laboratory Flow Cytometry Facility for training and expertise. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. B.A.G. is an investigator of the Howard Hughes Medical Institute and D.W.M. is a Howard Hughes Medical Institute Awardee of the Life Sciences Research Foundation. This research was also funded by NIH grant AI122528 to B.A.G.

1. Kutluay SB, Triezenberg SJ. Role of chromatin during herpesvirus infections. Biochim Biophys Acta. 2009;1790: 456–66. doi:10.1016/j.bbagen.2009.03.019

2. Toth Z, Maglinte DT, Lee SH, Lee HR, Wong LY, Brulois KF, et al. Epigenetic analysis of KSHV latent and lytic genomes. PLoS Pathog. 2010;6: 1–17. doi:10.1371/journal.ppat.1001013

3. Qi Y, Zheng G, Di C, Zhang J, Wang X, Hong Y, et al. Latency-associated nuclear antigen inhibits lytic replication of Kaposi's sarcoma-associated herpesvirus by regulating let-7a/RBPJ signaling. Virology. 2019;531: 69–78. doi:10.1016/J.VIROL.2019.02.019

4. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. EMBO J. 2008;27: 654. doi:10.1038/EMBOJ.2008.1

5. Campbell M, Watanabe T, Nakano K, Davis RR, Lyu Y, Tepper CG, et al. KSHV episomes reveal dynamic chromatin loop formation with domain-specific gene regulation. Nat Commun 2017 91. 2018;9: 1–14. doi:10.1038/s41467-017-02089-9

6. Banerji J, Rusconi S, Schaffner W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell. 1981;27: 299–308. doi:10.1016/0092-8674(81)90413-X

7. Loeken MR, Brady J. The Adenovirus EIIA Enhancer. J Biol Chem. 1989;264: 6572–6579. doi:10.1016/s0021-9258(18)83386-5

8. Stinski MF, Isomura H. Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. Med Microbiol Immunol. 2008;197: 223–231. doi:10.1007/S00430-007-0069-7/FIGURES/2

9. Dooley AL, O'Connor CM. Regulation of the MIE Locus During HCMV Latency and Reactivation. 2020;9. doi:10.3390/PATHOGENS9110869

10. Feldman ER, Kara M, Oko LM, Grau KR, Krueger BJ, Zhang J, et al. A Gammaherpesvirus Noncoding RNA Is Essential for Hematogenous Dissemination and Establishment of Peripheral Latency. mSphere. 2016;1. doi:10.1128/MSPHERE.00105-15

11. Hoffman BA, Wang Y, Feldman ER, Tibbetts SA. Epstein-Barr virus EBER1 and murine gammaherpesvirus TMER4 share conserved in vivo function to promote B cell egress and dissemination. Proc Natl Acad Sci U S A. 2019;116: 25392–25394.

doi:10.1073/pnas.1915752116

12. Hu M, Wang C, Li W, Lu W, Bai Z, Qin D, et al. A KSHV microRNA Directly Targets G Protein-Coupled Receptor Kinase 2 to Promote the Migration and Invasion of Endothelial Cells by Inducing CXCR2 and Activating AKT Signaling. PLOS Pathog. 2015;11: e1005171. doi:10.1371/JOURNAL.PPAT.1005171

13. Gay LA, Stribling D, Turner PC, Renne R. Kaposi's Sarcoma-Associated Herpesvirus MicroRNA Mutants Modulate Cancer Hallmark Phenotypic Differences in Human Endothelial Cells. J Virol. 2021;95. doi:10.1128/JVI.02022-20

14. Wang Y, Tang Q, Maul GG, Yuan Y. Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: dual role of replication and transcription activator. J Virol. 2006;80: 12171–12186. doi:10.1128/JVI.00990-06

15. Wang Y, Li H, Chan MY, Zhu FX, Lukac DM, Yuan Y. Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: cis-acting requirements for replication and ori-Lyt-associated RNA transcription. J Virol. 2004;78: 8615–8629. doi:10.1128/JVI.78.16.8615-8629.2004

16. Tagawa T, Gao S, Koparde VN, Gonzalez M, Spouge JL, Serquiña AP, et al. Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA. Proc Natl Acad Sci U S A. 2018;115: 12805–12810. doi:10.1073/pnas.1816183115

17. Tagawa T, Oh D, Santos J, Dremel S, Mahesh G, Uldrick TS, et al. Characterizing Expression and Regulation of Gamma-Herpesviral Circular RNAs. Front Microbiol. 2021;12. doi:10.3389/FMICB.2021.670542/FULL

18. Rossetto CC, Pari GS. Kaposi's sarcoma-associated herpesvirus noncoding polyadenylated nuclear RNA interacts with virus- and host cell-encoded proteins and suppresses expression of genes involved in immune modulation. J Virol. 2011;85: 13290–7. doi:10.1128/JVI.05886-11

19. Campbell M, Izumiya Y. PAN RNA: transcriptional exhaust from a viral engine. J Biomed Sci 2020 271. 2020;27: 1–10. doi:10.1186/S12929-020-00637-Y

20. Sun R, Lin SF, Gradoville L, Miller G. Polyadenylylated nuclear RNA encoded by Kaposi sarcoma-associated herpesvirus. Proc Natl Acad Sci U S A. 1996;93: 11883–11888. doi:10.1073/PNAS.93.21.11883

21. Schifano JM, Corcoran K, Kelkar H, Dittmer DP. Expression of the Antisense-to-Latency Transcript Long Noncoding RNA in Kaposi's Sarcoma-Associated Herpesvirus. J Virol. 2017;91. doi:10.1128/JVI.01698-16

22. Chandriani S, Xu Y, Ganem D. The Lytic Transcriptome of Kaposi's Sarcoma-Associated Herpesvirus Reveals Extensive Transcription of Noncoding Regions, Including Regions Antisense to Important Genes. J Virol. 2010;84: 7934. doi:10.1128/JVI.00645-10

23. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol 2015 335. 2015;33: 510–517. doi:10.1038/nbt.3199

24. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods 2015 1212. 2015;12: 1143–1149. doi:10.1038/nmeth.3630

25. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science. 2017;355. doi:10.1126/science.aah7111

26. Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, et al. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science. 2016;354: 769. doi:10.1126/SCIENCE.AAG2445

27. Tycko J, Wainberg M, Marinov GK, Ursu O, Hess GT, Ego BK, et al. Mitigation of off-target toxicity in CRISPR-Cas9 screens for essential non-coding elements. Nat Commun. 2019;10. doi:10.1038/s41467-019-11955-7

28. Brackett K, Mungale A, Lopez-Isidro M, Proctor DA, Najarro G, Arias C. Crispr interference efficiently silences latent and lytic viral genes in kaposi's sarcoma-associated herpesvirus-infected cells. Viruses. 2021;13. doi:10.3390/v13050783

29. Gardner MR, Glaunsinger BA. Kaposi's Sarcoma-Associated Herpesvirus ORF68 Is a DNA Binding Protein Required for Viral Genome Cleavage and Packaging. J Virol. 2018;92. doi:10.1128/JVI.00840-18

30. Didychuk AL, Gates SN, Gardner MR, Strong LM, Martin A, Glaunsinger BA. A pentameric protein ring with novel architecture is required for herpesviral packaging. Elife.

2021;10: 1–31. doi:10.7554/ELIFE.62261

31. Gabaev I, Williamson JC, Crozier TWM, Schulz TF, Lehner PJ. Quantitative Proteomics Analysis of Lytic KSHV Infection in Human Endothelial Cells Reveals Targets of Viral Immune Modulation. Cell Rep. 2020;33. doi:10.1016/j.celrep.2020.108249

32. Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, et al. KSHV 2.0: A Comprehensive Annotation of the Kaposi's Sarcoma-Associated Herpesvirus Genome Using Next-Generation Sequencing Reveals Novel Genomic and Functional Features. Dittmer DP, editor. PLoS Pathog. 2014;10: e1003847. doi:10.1371/journal.ppat.1003847

33. Morgens DW, Nandakumar D, Didychuk Id AL, Yang KJ, Glaunsingerid BA. A Two-tiered functional screen identifies herpesviral transcriptional modifiers and their essential domains. Gewurz BE, editor. PLOS Pathog. 2022;18: e1010236. doi:10.1371/JOURNAL.PPAT.1010236

34. Ye X, Zhaoid Y, Karijolich J. The landscape of transcription initiation across latent and lytic KSHV genomes. PLoS Pathog. 2019;15: e1007852. doi:10.1371/journal.ppat.1007852

35. Gregory Bruce A, Barcy S, Dimaio T, Gan E, Jacques Garrigues H, Lagunoff M, et al. Quantitative Analysis of the KSHV Transcriptome Following Primary Infection of Blood and Lymphatic Endothelial Cells. Pathog (Basel, Switzerland). 2017;6. doi:10.3390/PATHOGENS6010011

36. Lensch S, Herschl MH, Ludwig CH, Sinha J, Hinks MM, Mukund A, et al. Dynamic spreading of chromatin-mediated gene silencing and reactivation between neighboring genes in single cells. Elife. 2022;11. doi:10.7554/ELIFE.75115

37. Guito J, Lukac DM. KSHV Rta promoter specification and viral reactivation. Front Microbiol. 2012;3: 30. doi:10.3389/FMICB.2012.00030/BIBTEX

38. Komander D, Rape M. The ubiquitin code. Annu Rev Biochem. 2012;81: 203–29. doi:10.1146/annurev-biochem-060310-170328

39. McCollum CO, Didychuk AL, Liu D, Murray-Nerger LA, Cristea IM, Glaunsinger BA. The viral packaging motor potentiates Kaposi's sarcoma-associated herpesvirus gene expression late in infection. PLOS Pathog. 2023;19: e1011163. doi:10.1371/JOURNAL.PPAT.1011163

40. Majerciak V, Zheng ZM. KSHV ORF57, a Protein of Many Faces. Viruses. 2015;7: 604. doi:10.3390/V7020604

41. Uppal T, Banerjee S, Sun Z, Verma SC, Robertson ES. KSHV LANA—The Master Regulator of KSHV Latency. Viruses. 2014;6: 4961. doi:10.3390/V6124961

42. Jones T, Da Silva SR, Bedolla R, Ye F, Zhou F, Gao SJ. Viral Cyclin promotes KSHV-induced cellular transformation and tumorigenesis by overriding contact inhibition. http://dx.doi.org/104161/cc27758. 2014;13: 845–858. doi:10.4161/CC.27758

43. Kumar Singh R, Pei Y, Bose D, Lamplugh ZL, Sun K, Yuan Y, et al. Kshv-encoded vcyclin can modulate hif1a levels to promote dna replication in hypoxia. Elife. 2021;10. doi:10.7554/ELIFE.57436

44. T K-N, Y X, CC R, K C, I P, GS P. Overexpression of the kaposi's sarcoma-associated herpesvirus transactivator K-Rta can complement a K-bZIP deletion BACmid and yields an enhanced growth phenotype. J Virol. 2007;81: 13519–13532. doi:10.1128/JVI.00832-07

45. Full F, Jungnickl D, Reuter N, Bogner E, Brulois K, Scholz B, et al. Kaposi's Sarcoma Associated Herpesvirus Tegument Protein ORF75 Is Essential for Viral Lytic Replication and Plays a Critical Role in the Antagonization of ND10-Instituted Intrinsic Immunity. PLOS Pathog. 2014;10: e1003863. doi:10.1371/JOURNAL.PPAT.1003863

46. Verma D, Li D-J, Krueger B, Renne R, Swaminathan S. Identification of the Physiological Gene Targets of the Essential Lytic Replicative Kaposi's Sarcoma-Associated Herpesvirus ORF57 Protein. J Virol. 2015;89: 1688–1702. doi:10.1128/jvi.02663-14

47. Chen PB, Fiaux PC, Zhang K, Li B, Kubo N, Jiang S, et al. Systematic discovery and functional dissection of enhancers needed for cancer cell fitness and proliferation. Cell Rep. 2022;41: 111630. doi:10.1016/J.CELREP.2022.111630

48. Deans RM, Morgens DW, Ökesli A, Pillay S, Horlbeck MA, Kampmann M, et al. Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. Nat Chem Biol. 2016;12: 361–366. doi:10.1038/nchembio.2050

49. Tsui CK, Twells N, Doan E, Brooks J, Kulepa A, Webster B, et al. Combinatorial CRISPR screens and lectin microarrays identify novel glycosylation regulators. bioRxiv. 2023; 2023.10.23.563662. doi:10.1101/2023.10.23.563662

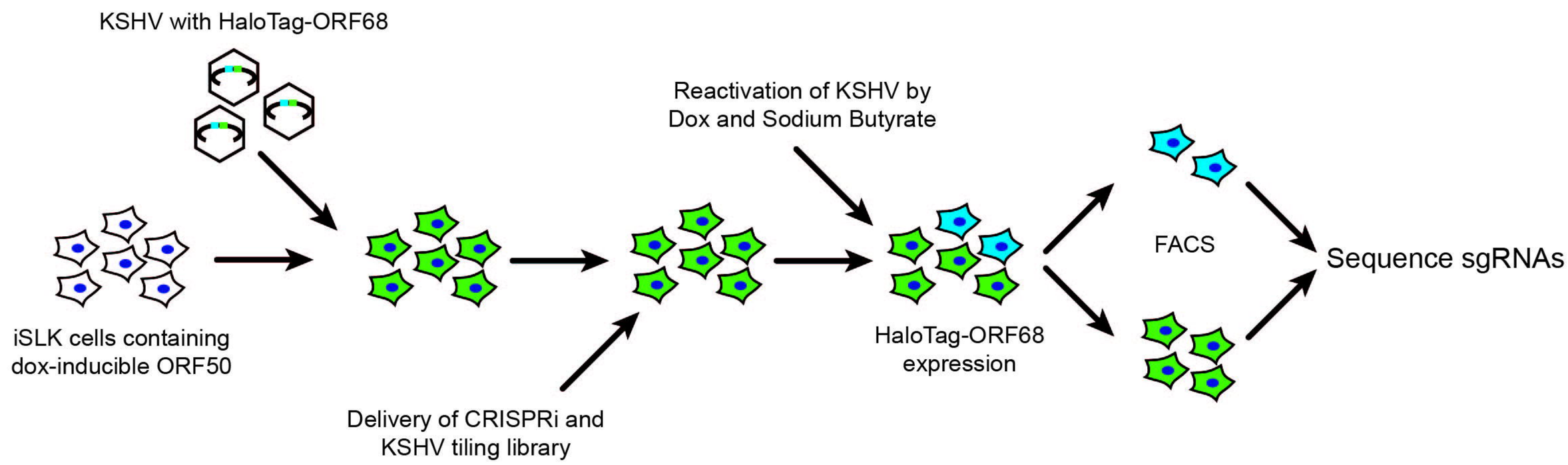
50. Morgens DW, Chan C, Kane AJ, Weir NR, Li A, Dubreuil MM, et al. Retro-2 protects cells from ricin toxicity by inhibiting ASNA1-mediated ER targeting and insertion of tail-anchored proteins. Elife. 2019;8. doi:10.7554/eLife.48434

51. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17: 10. doi:10.14806/ej.17.1.200

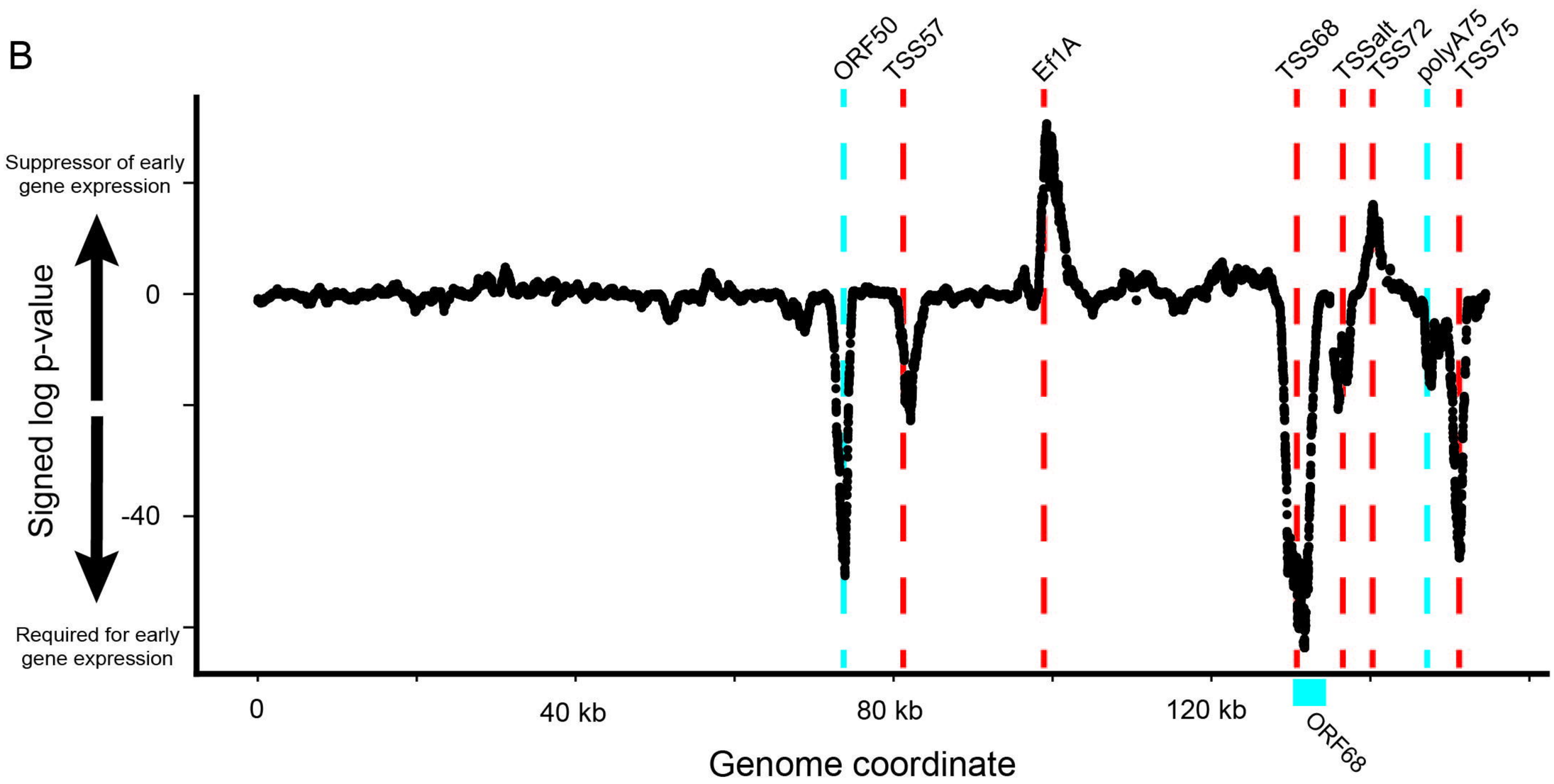
52. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10: R25. doi:10.1186/gb-2009-10-3-r25

Figure 1

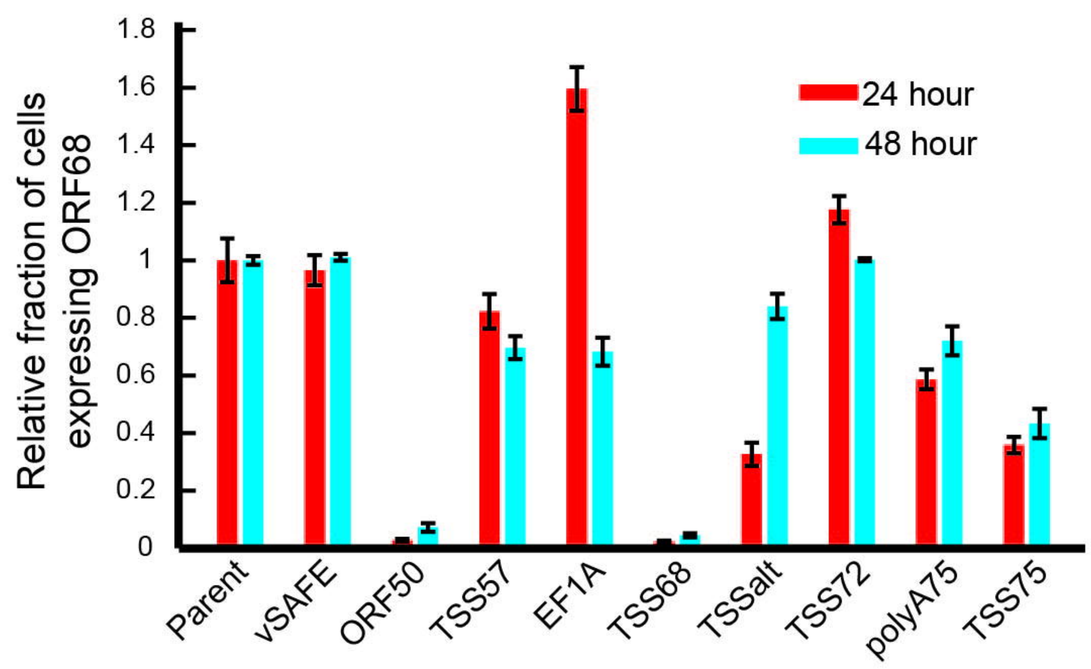
A



B



C



D

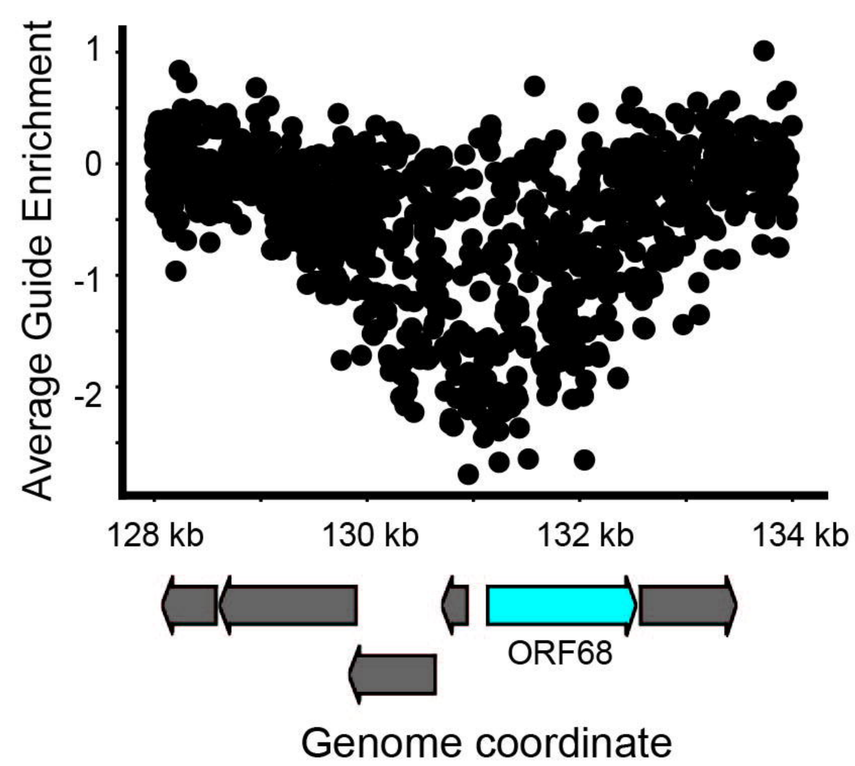
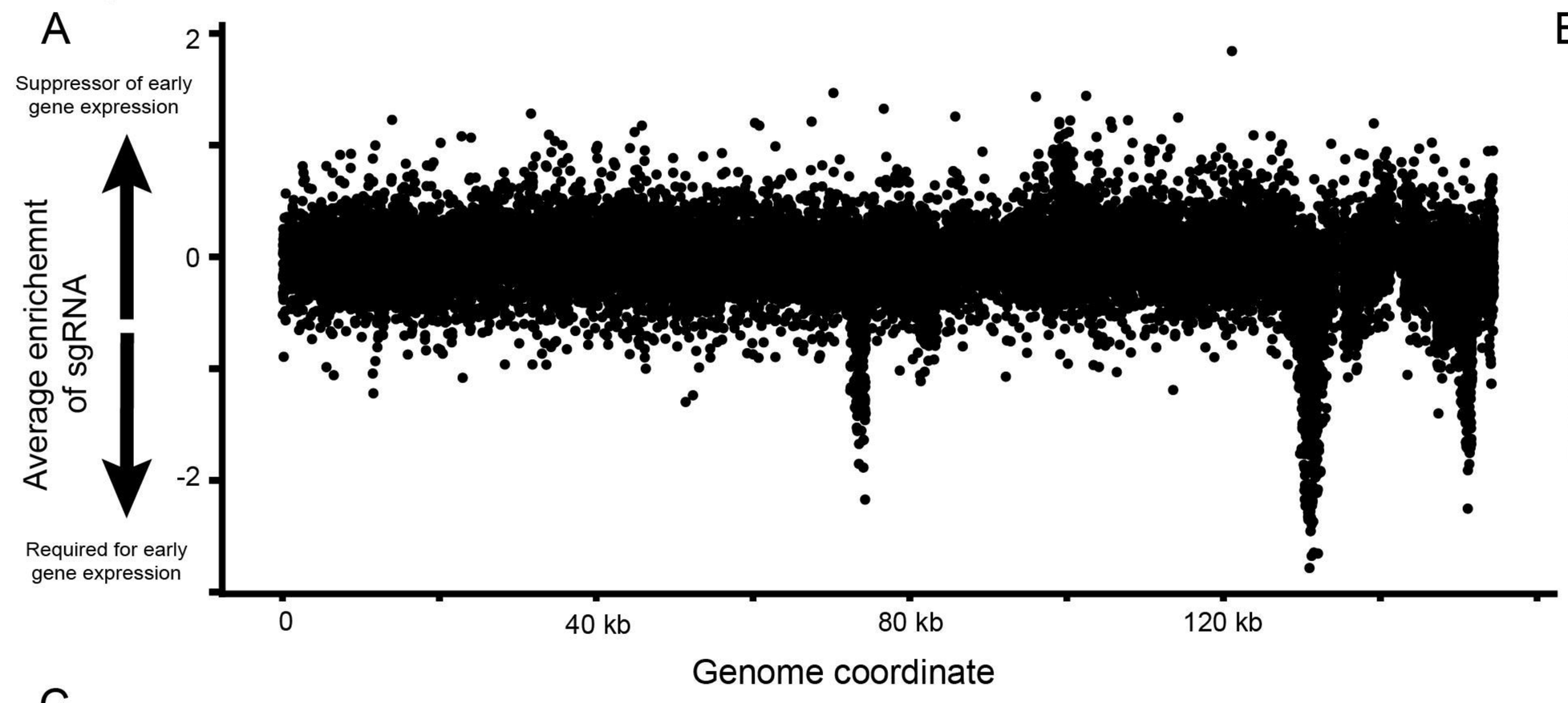
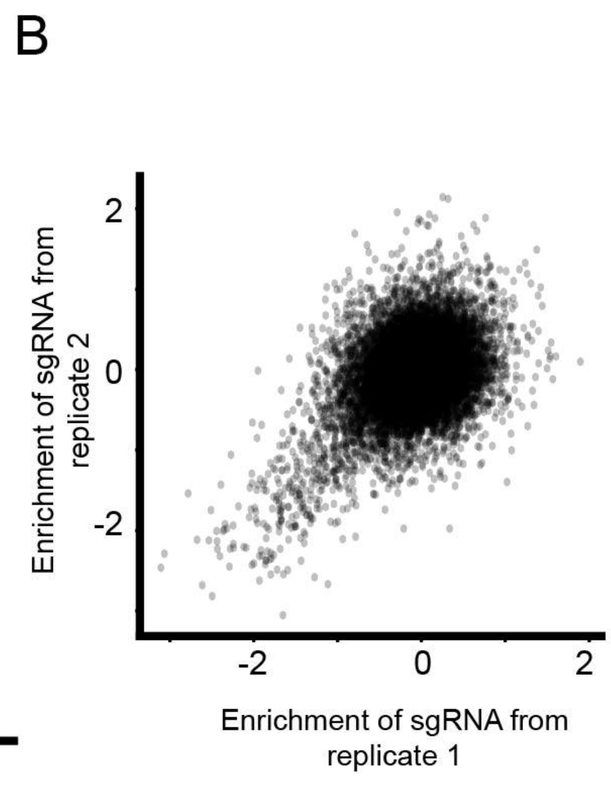
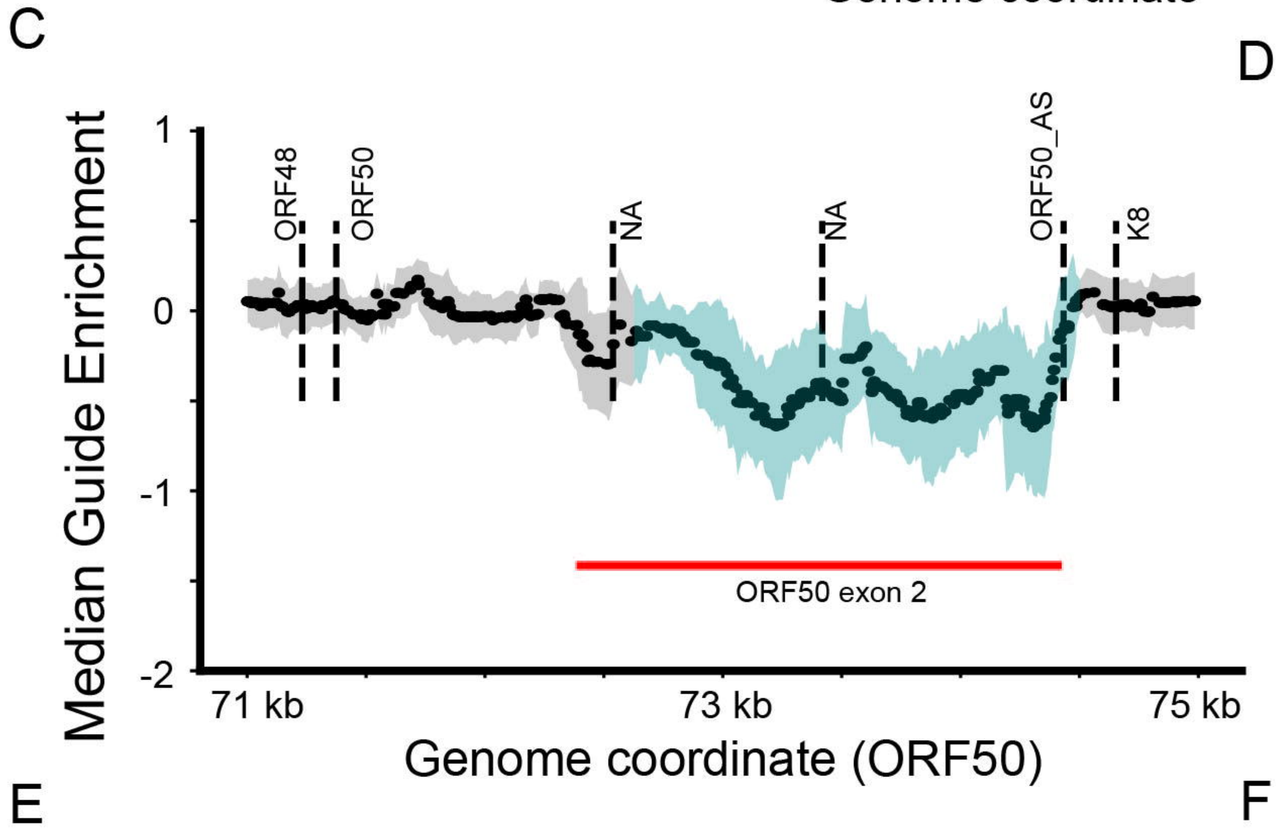
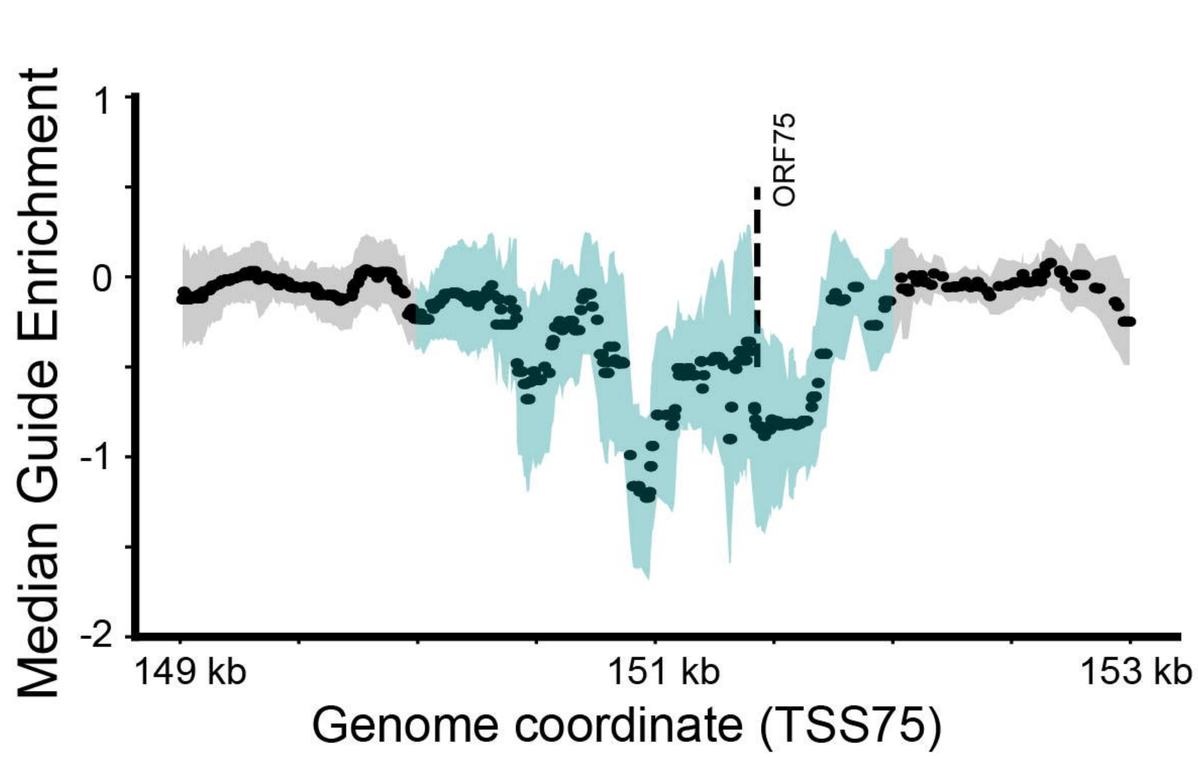


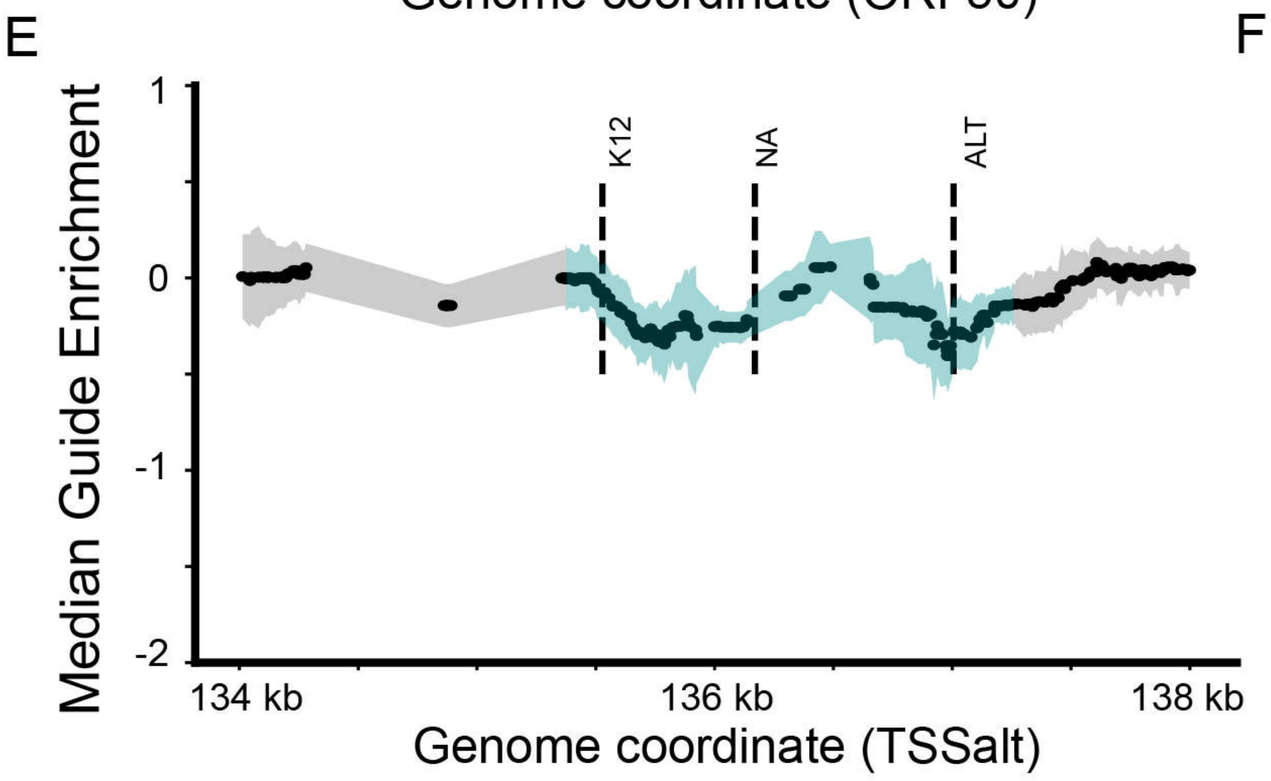
Figure S1

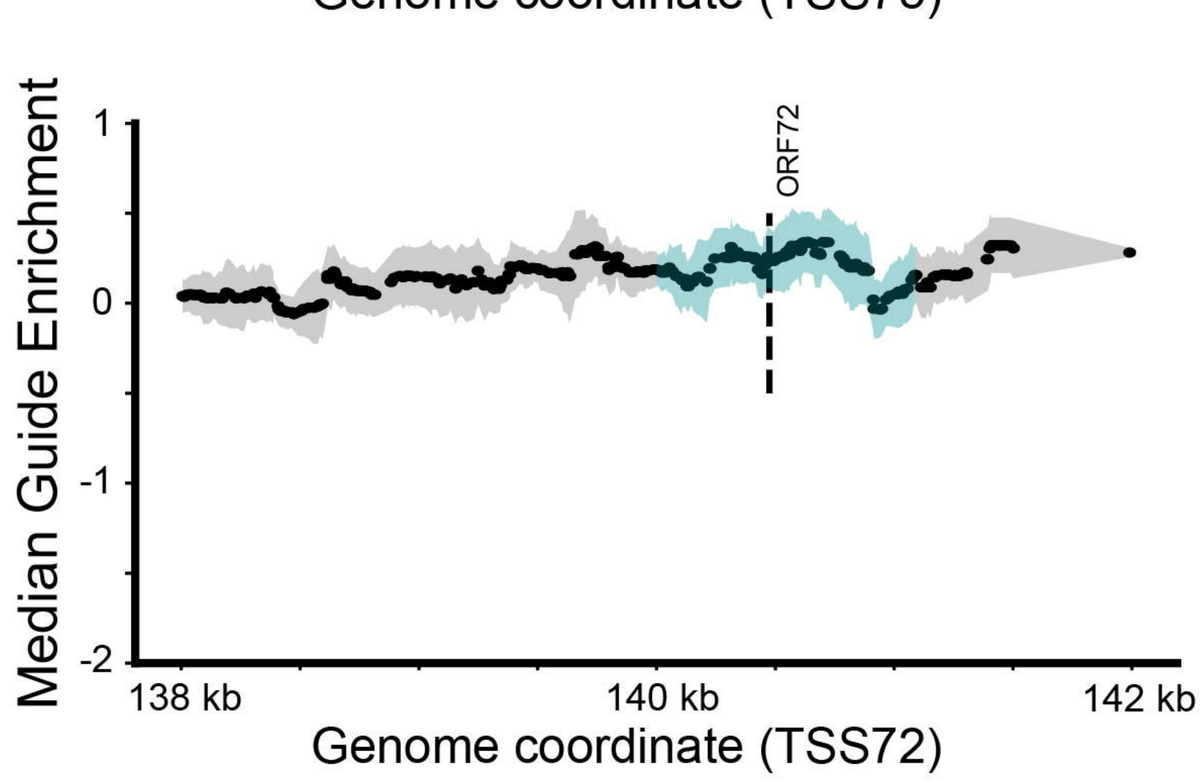

A
Suppressor of early gene expression
Required for early gene expression
Average enrichemnt of sgRNA
2
0
-2
0
40 kb
80 kb
120 kb
Genome coordinate

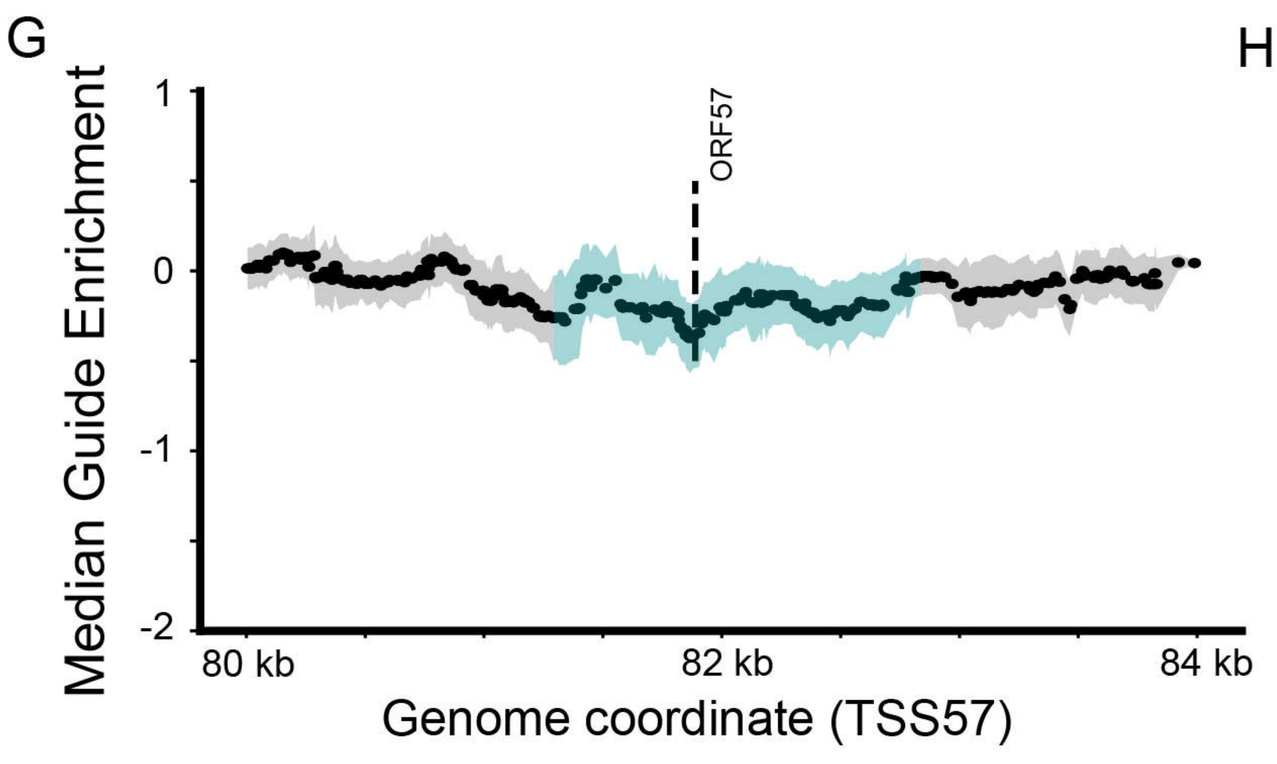

B
Enrichment of sgRNA from replicate 2
Enrichment of sgRNA from replicate 1
2
0
-2
-2
0
2


C
Median Guide Enrichment
ORF48
ORF50
NA
NA
ORF50_AS
K8
ORF50 exon 2
1
0
-1
-2
71 kb
73 kb
75 kb
Genome coordinate (ORF50)


D
Median Guide Enrichment
ORF75
1
0
-1
-2
149 kb
151 kb
153 kb
Genome coordinate (TSS75)


E
Median Guide Enrichment
K12
NA
ALT
1
0
-1
-2
134 kb
136 kb
138 kb
Genome coordinate (TSSalt)


F
Median Guide Enrichment
ORF72
1
0
-1
-2
138 kb
140 kb
142 kb
Genome coordinate (TSS72)


G
Median Guide Enrichment
ORF57
1
0
-1
-2
80 kb
82 kb
84 kb
Genome coordinate (TSS57)

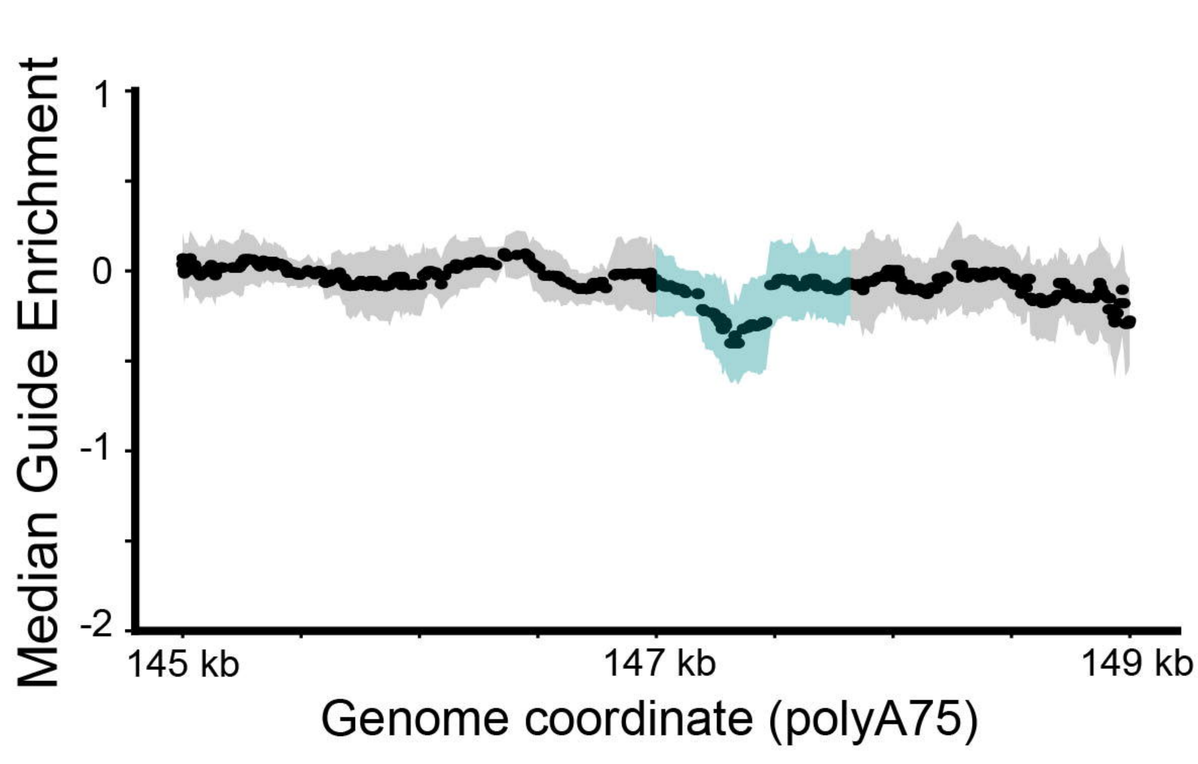

H
Median Guide Enrichment
1
0
-1
-2
145 kb
147 kb
149 kb
Genome coordinate (polyA75)

Figure 2

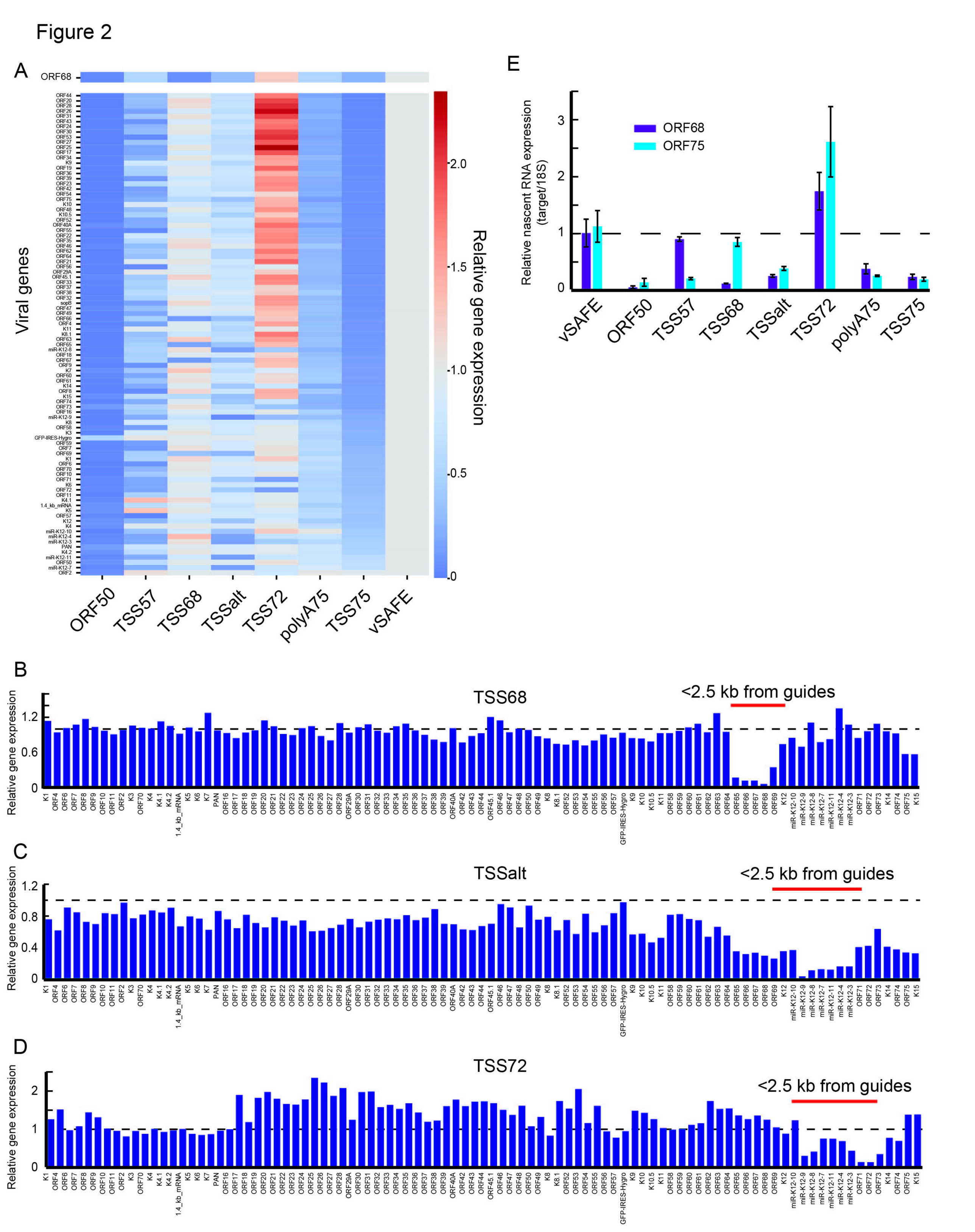
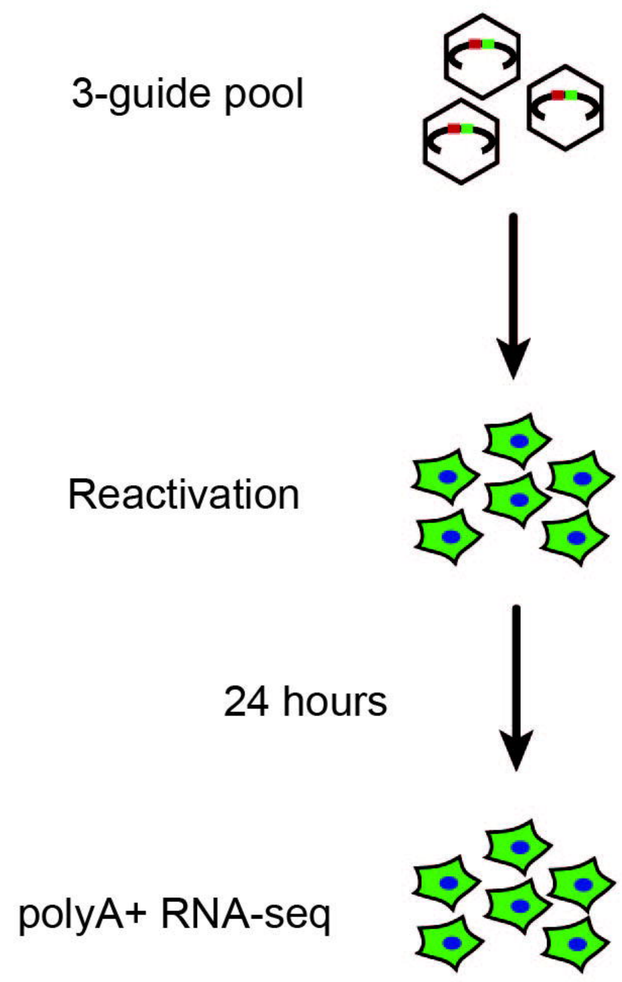
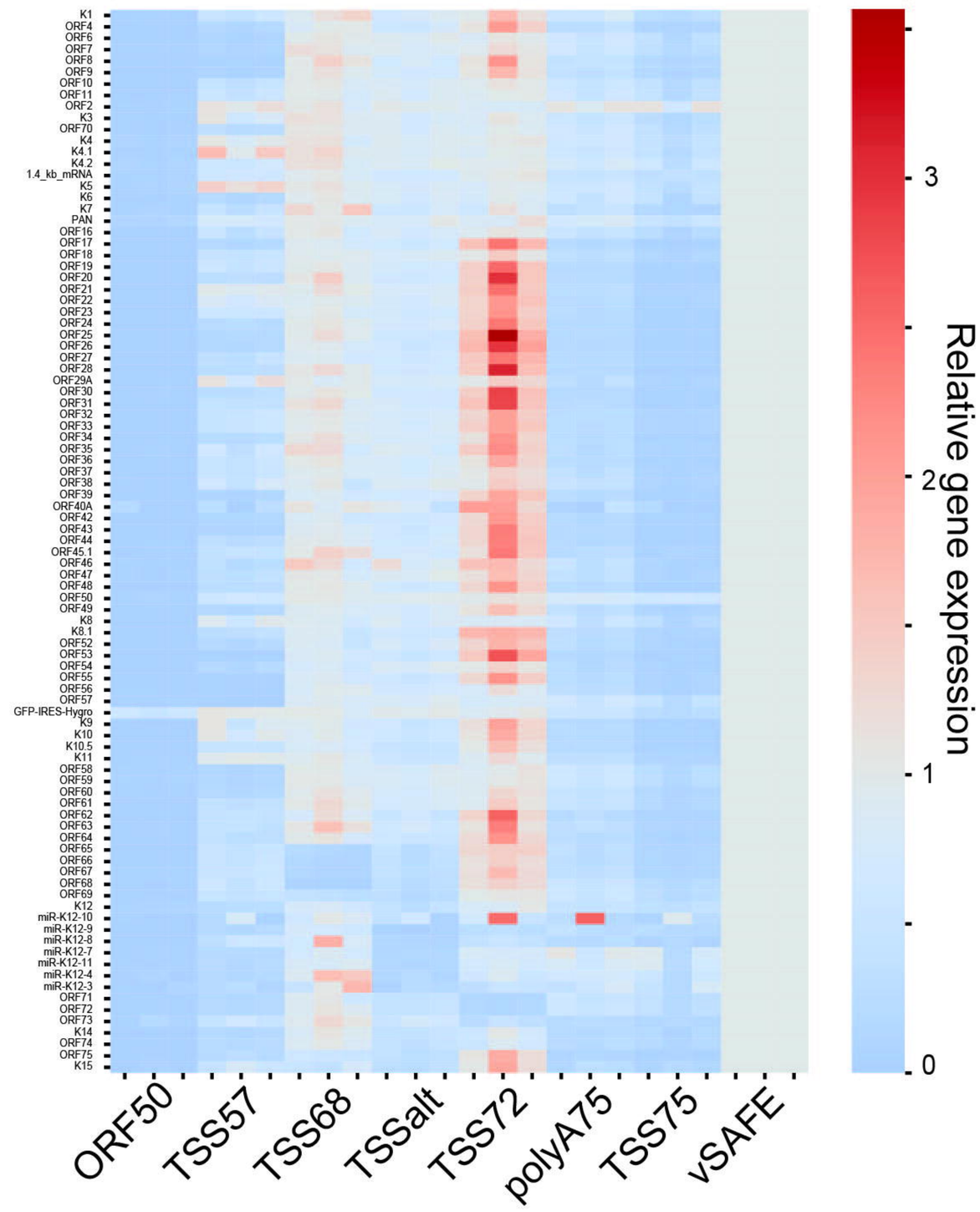

A
ORF68
Viral genes
Relative gene expression
2.0
1.5
1.0
0.5
0
ORF50
TSS57
TSS68
TSSalt
TSS72
polyA75
TSS75
vSAFE
E
Relative nascent RNA expression (target/18S)
ORF68
ORF75
vSAFE
ORF50
TSS57
TSS68
TSSalt
TSS72
polyA75
TSS75
B
TSS68
<2.5 kb from guides
Relative gene expression
C
TSSalt
<2.5 kb from guides
Relative gene expression
D
TSS72
<2.5 kb from guides
Relative gene expression

Figure S2

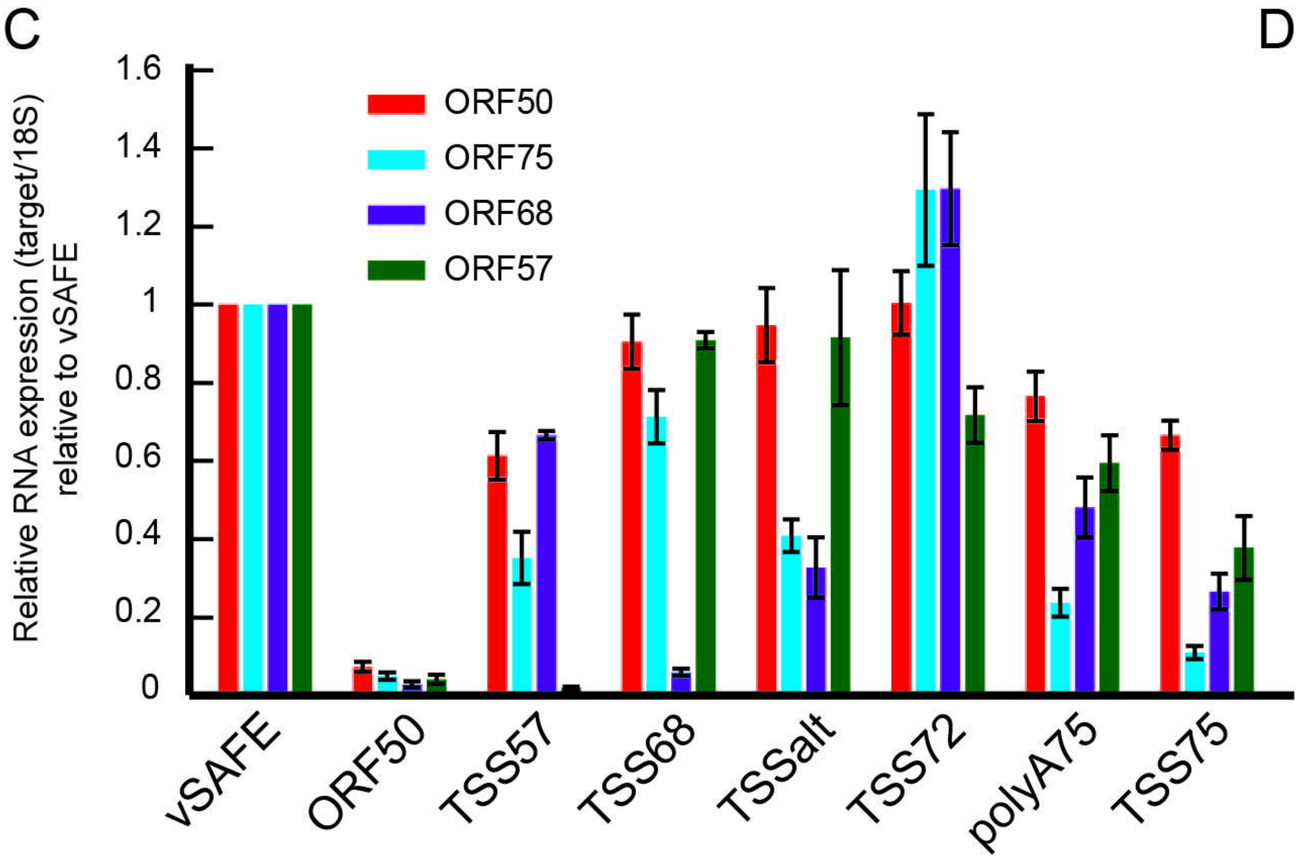
A



B



C



D

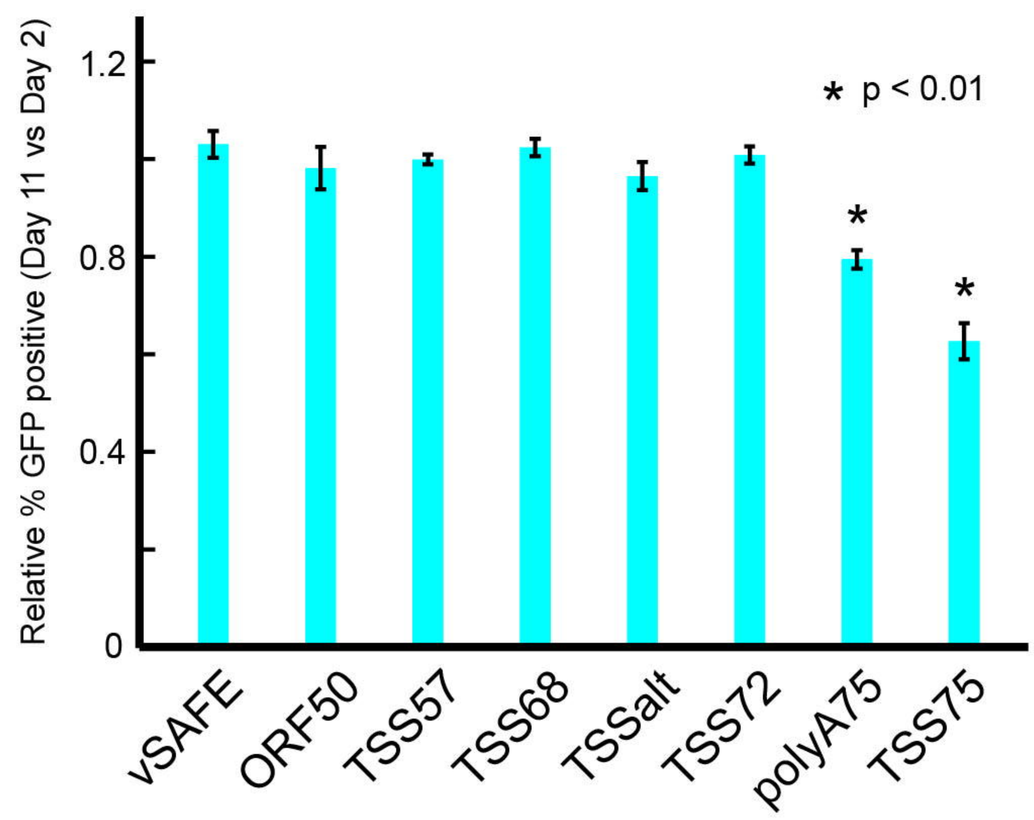
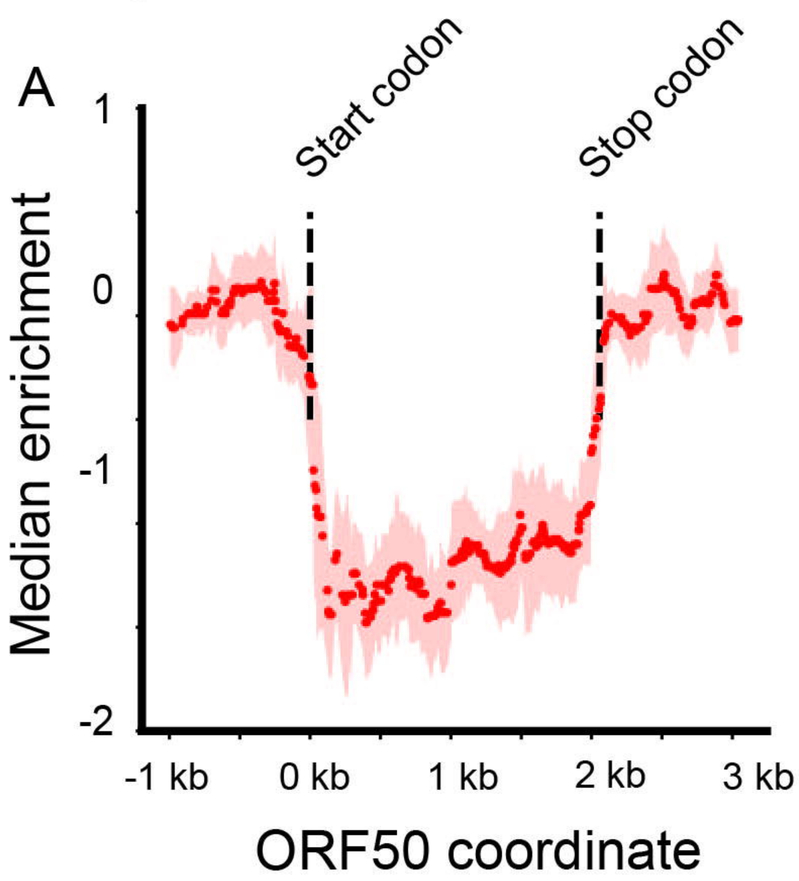
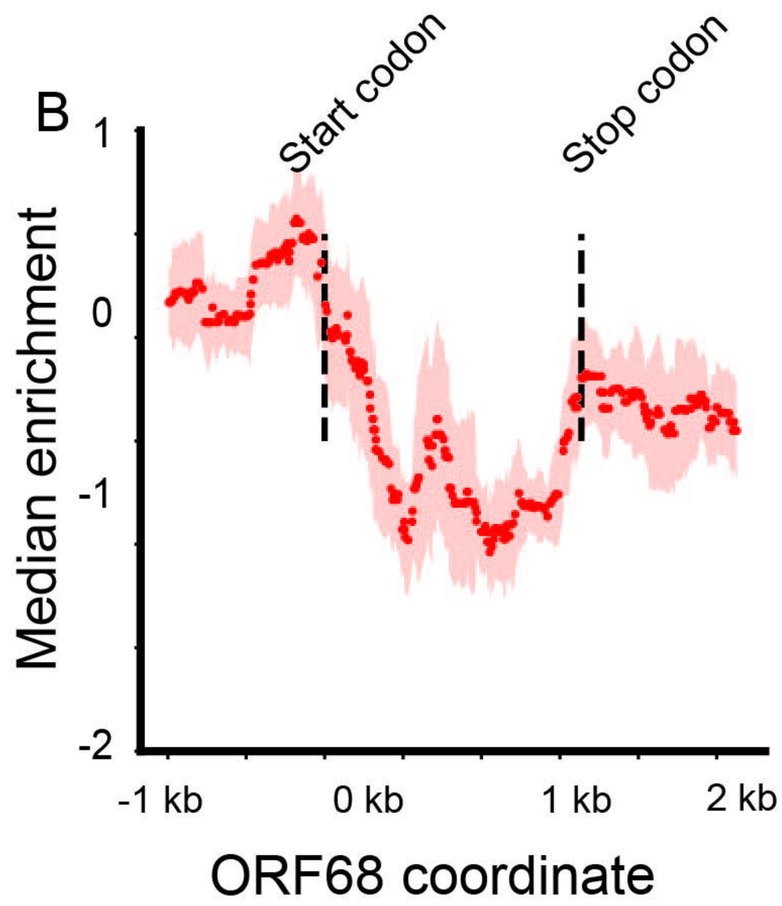
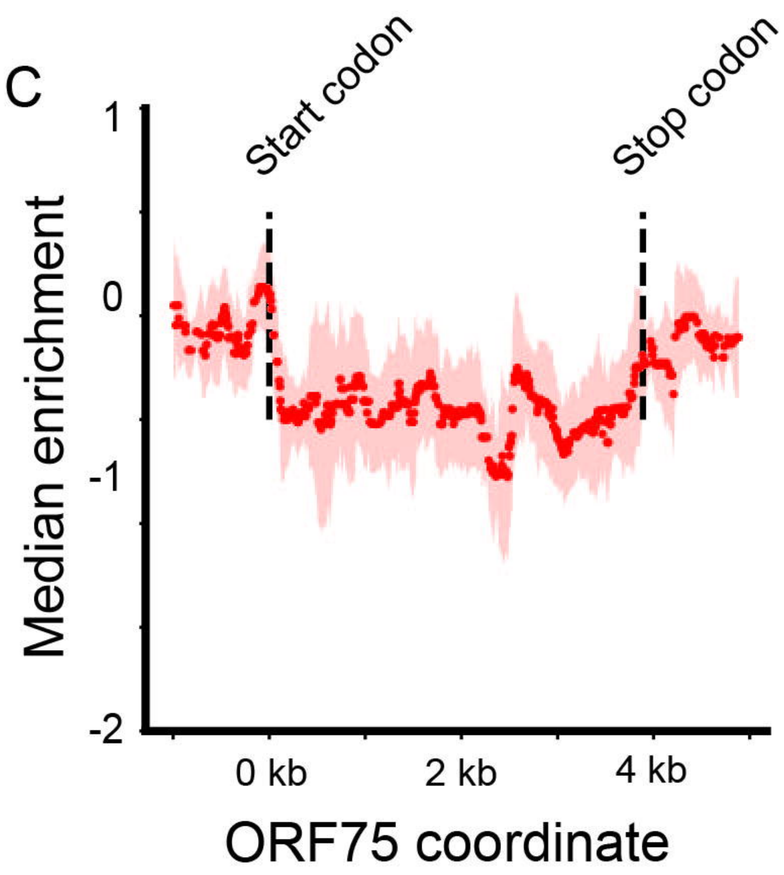
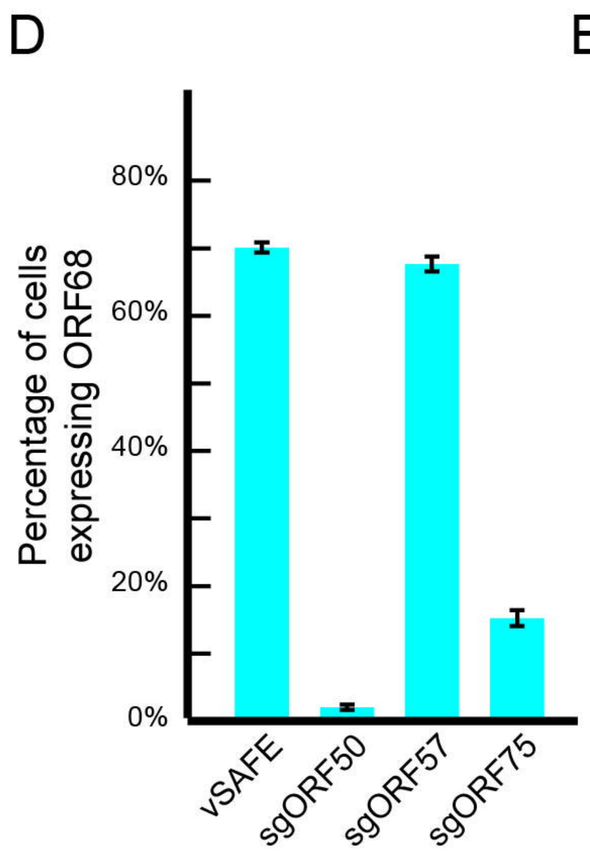


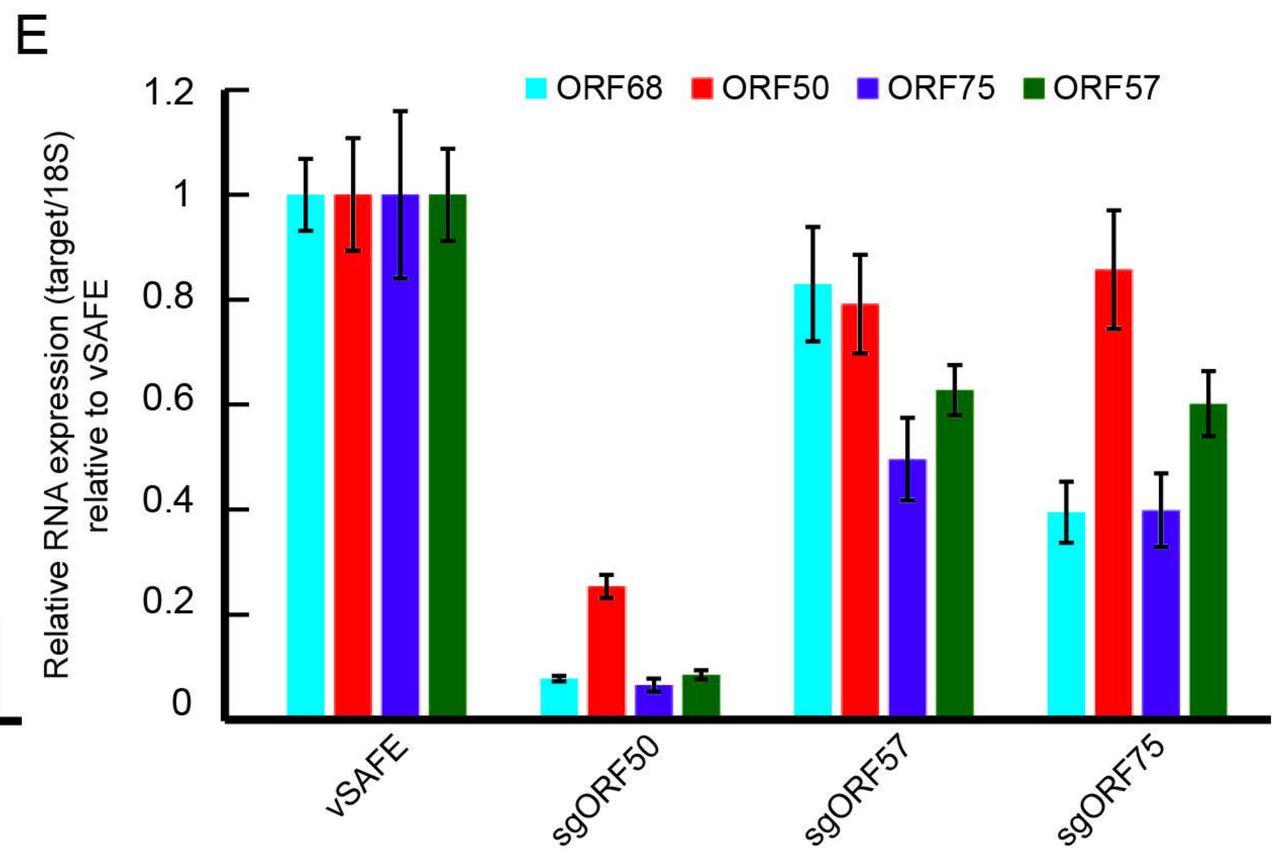
Figure 3


A
Start codon
Stop codon
Median enrichment
1
0
-1
-2
-1 kb
0 kb
1 kb
2 kb
3 kb
ORF50 coordinate


B
Start codon
Stop codon
Median enrichment
1
0
-1
-2
-1 kb
0 kb
1 kb
2 kb
ORF68 coordinate


C
Start codon
Stop codon
Median enrichment
1
0
-1
-2
0 kb
2 kb
4 kb
ORF75 coordinate


D
Percentage of cells expressing ORF68
80%
60%
40%
20%
0%
vSAFE
sgORF50
sgORF57
sgORF75


E
ORF68
ORF50
ORF75
ORF57
Relative RNA expression (target/18S) relative to vSAFE
1.2
1
0.8
0.6
0.4
0.2
0
vSAFE
sgORF50
sgORF57
sgORF75

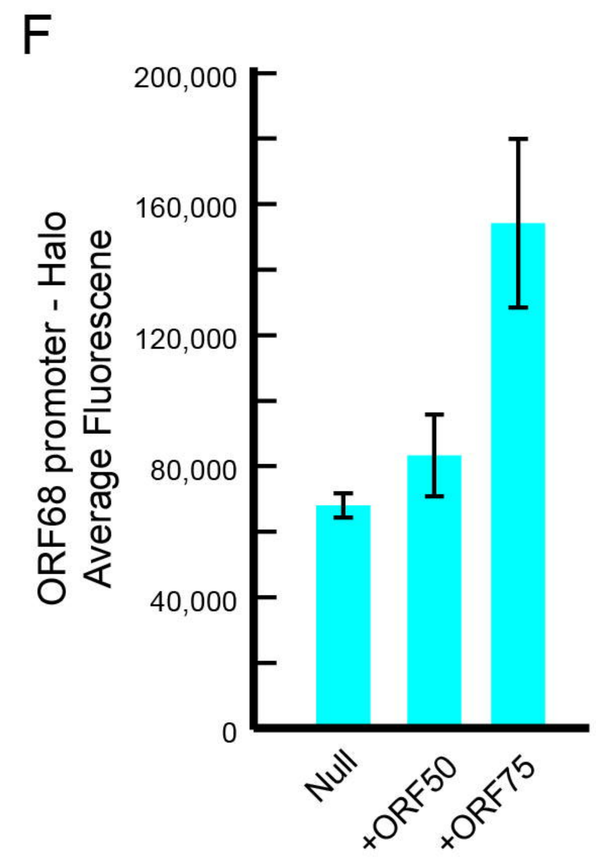
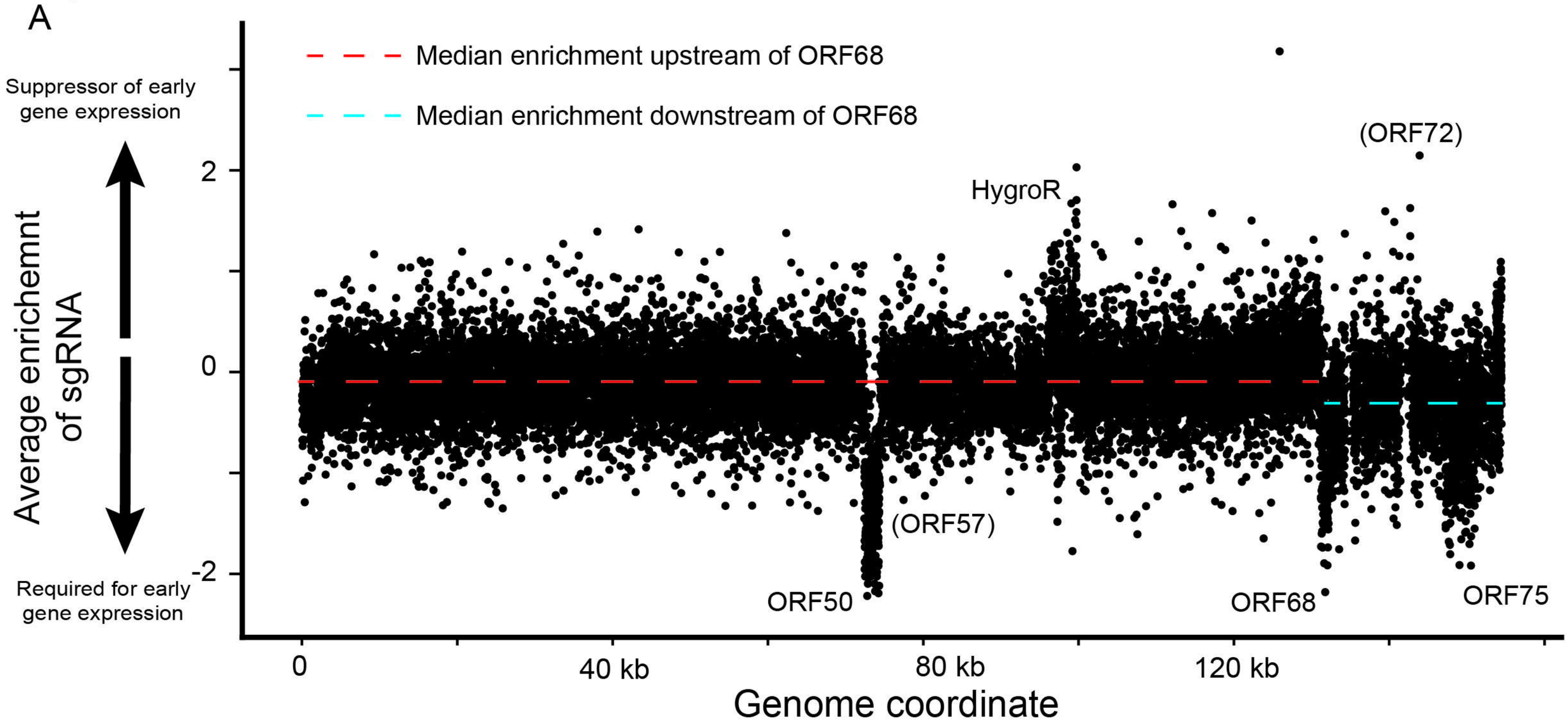
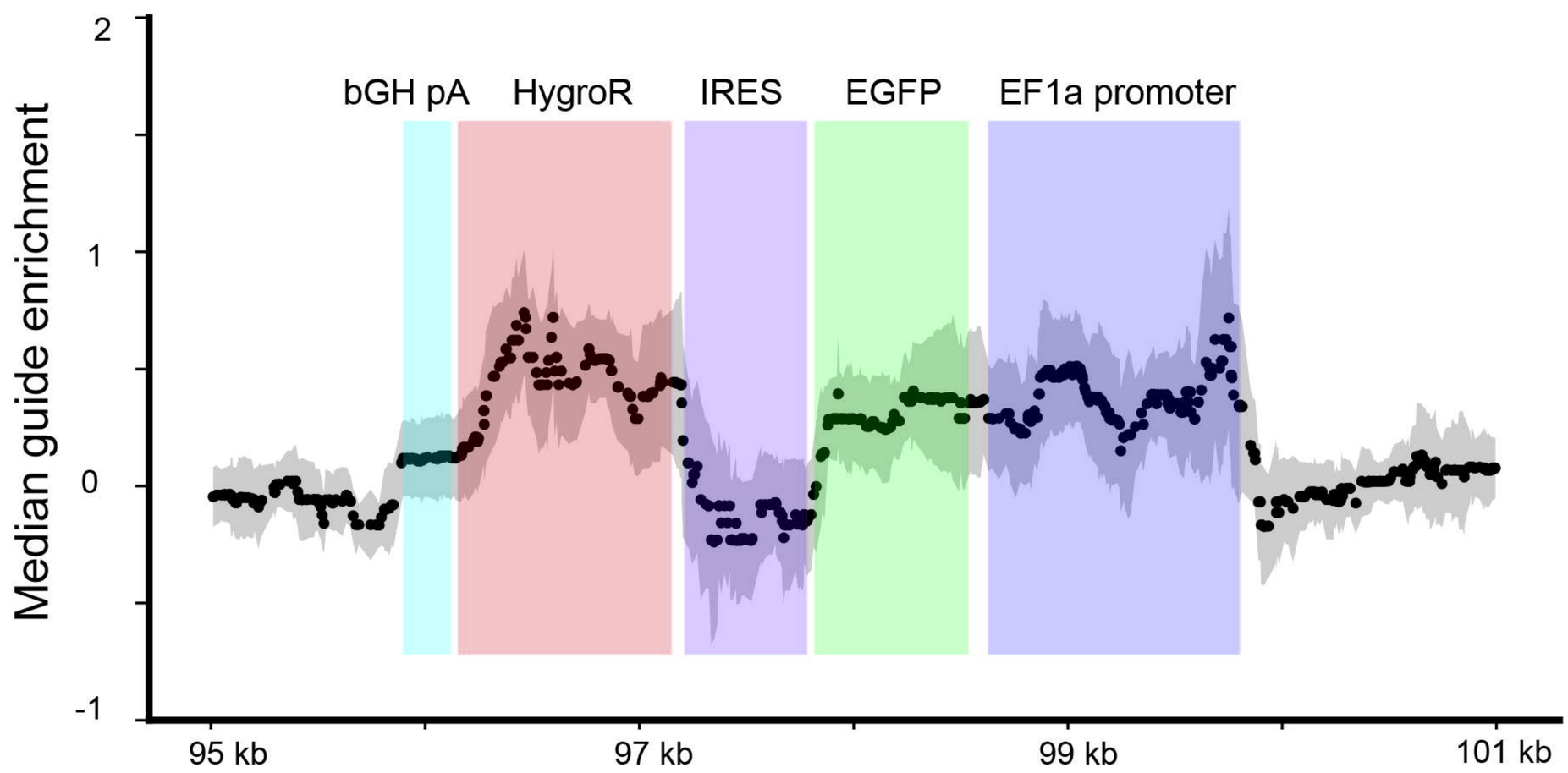

F
ORF68 promoter - Halo Average Fluorescene
200,000
160,000
120,000
80,000
40,000
0
Null
+ORF50
+ORF75

Figure S3

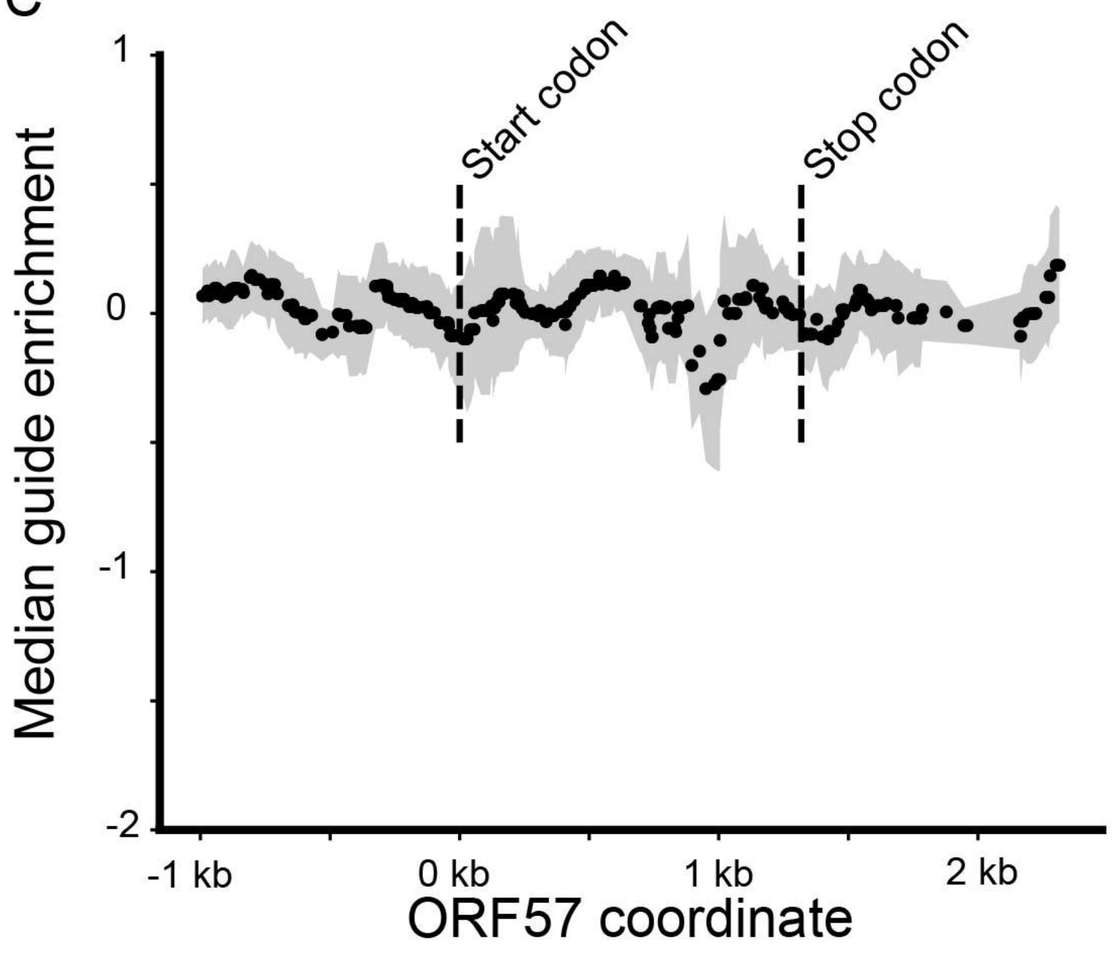
A



B



C



D

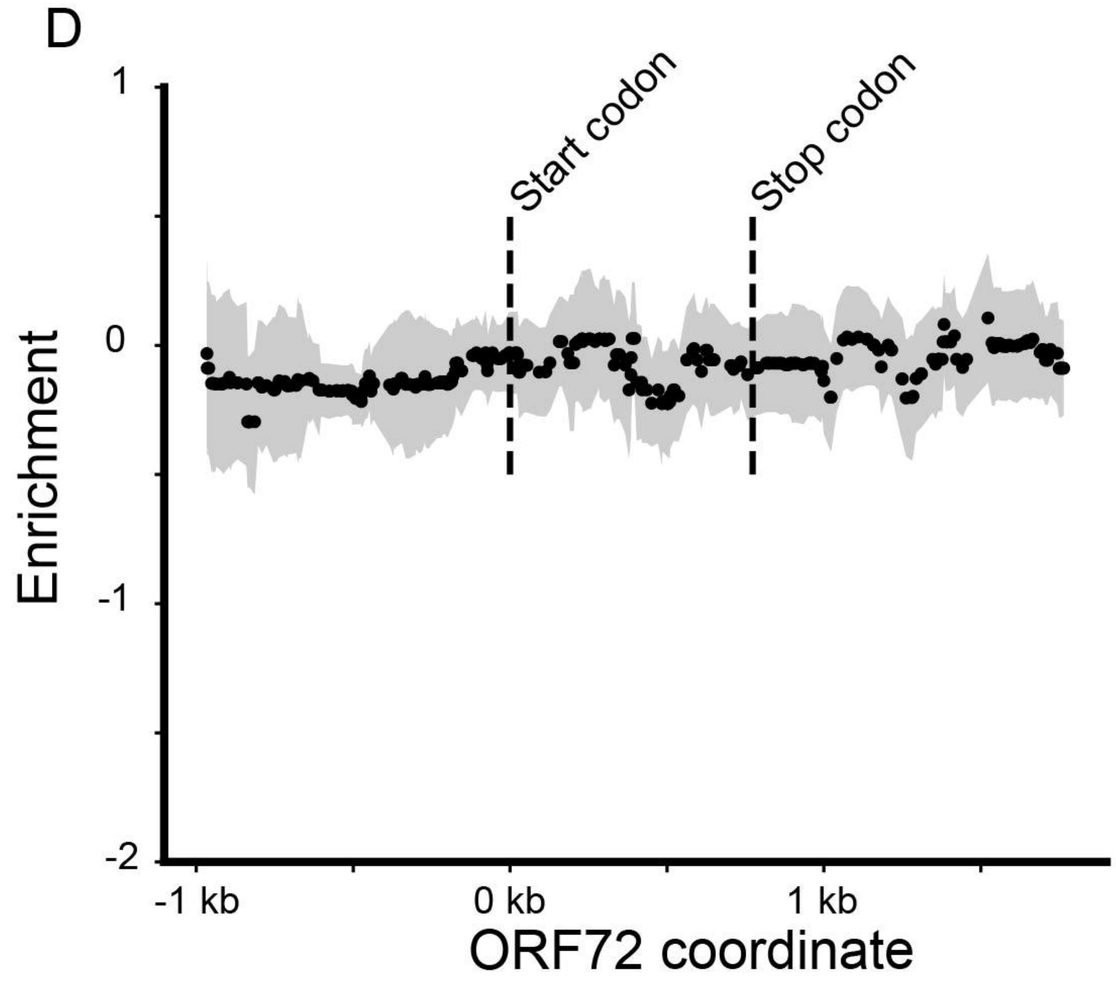
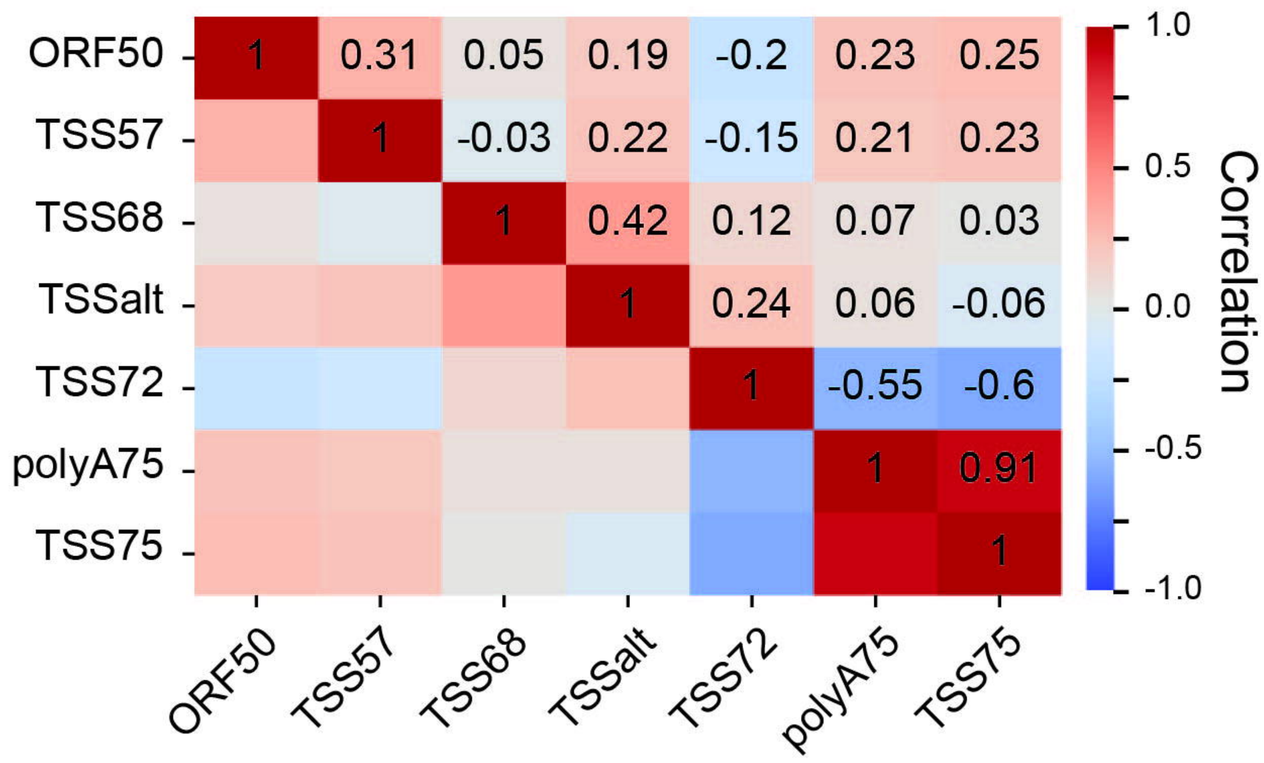
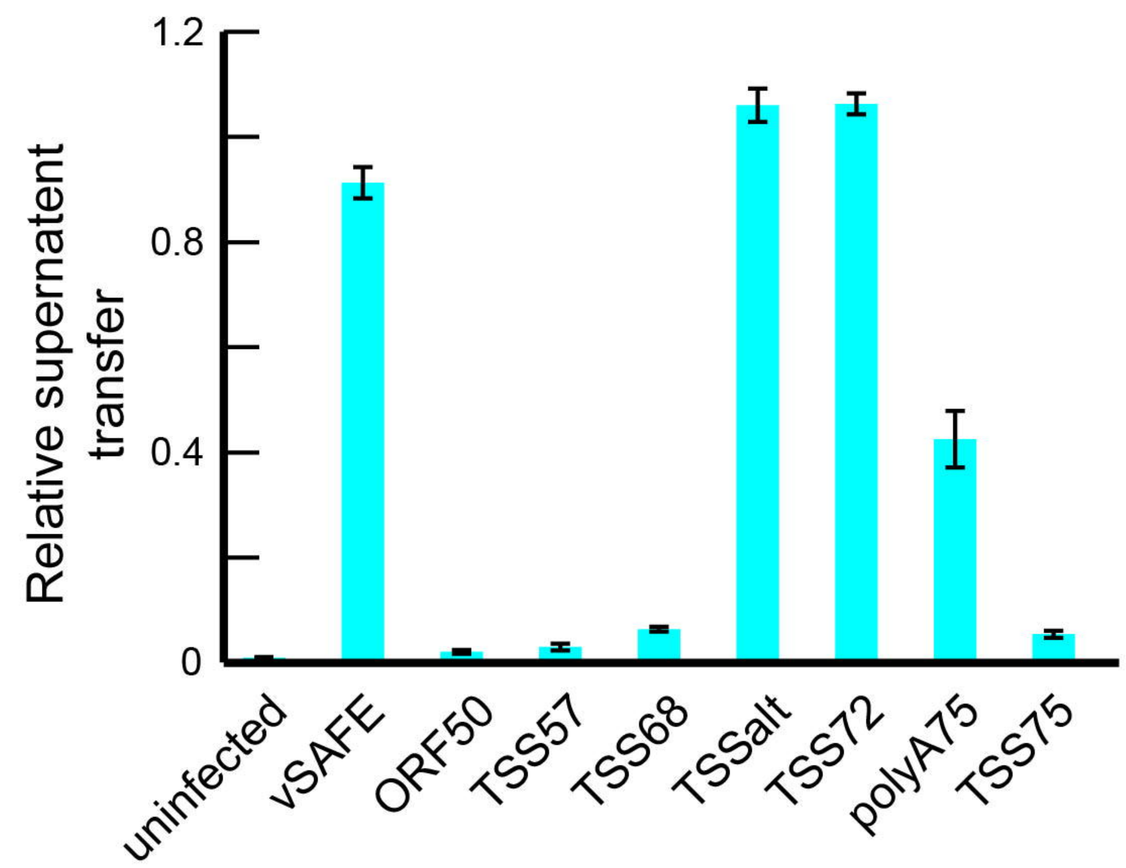


Figure 4

A


ORF50
TSS57
TSS68
TSSalt
TSS72
polyA75
TSS75
1
0.31
0.05
0.19
-0.2
0.23
0.25
1
-0.03
0.22
-0.15
0.21
0.23
1
0.42
0.12
0.07
0.03
1
0.24
0.06
-0.06
1
-0.55
-0.6
1
0.91
1
1.0
0.5
0.0
-0.5
-1.0
Correlation

C


Relative supernatent transfer
1.2
0.8
0.4
0
uninfected
vSAFE
ORF50
TSS57
TSS68
TSSalt
TSS72
polyA75
TSS75

B

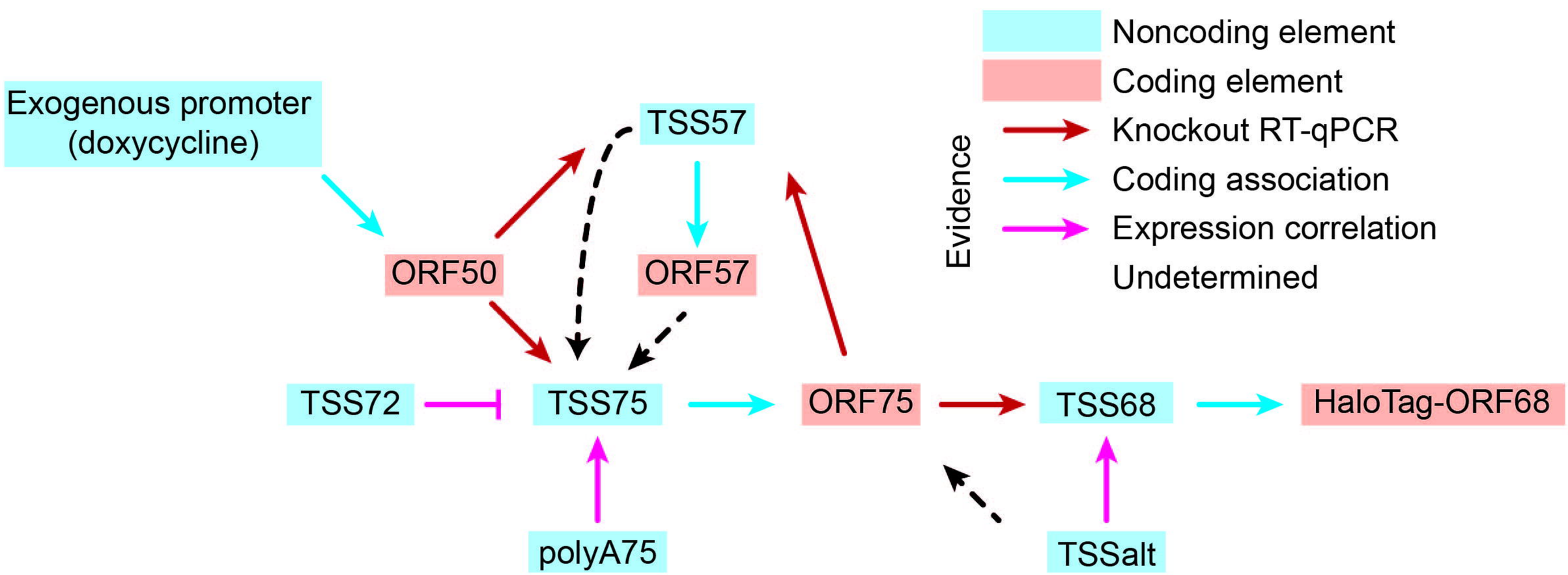
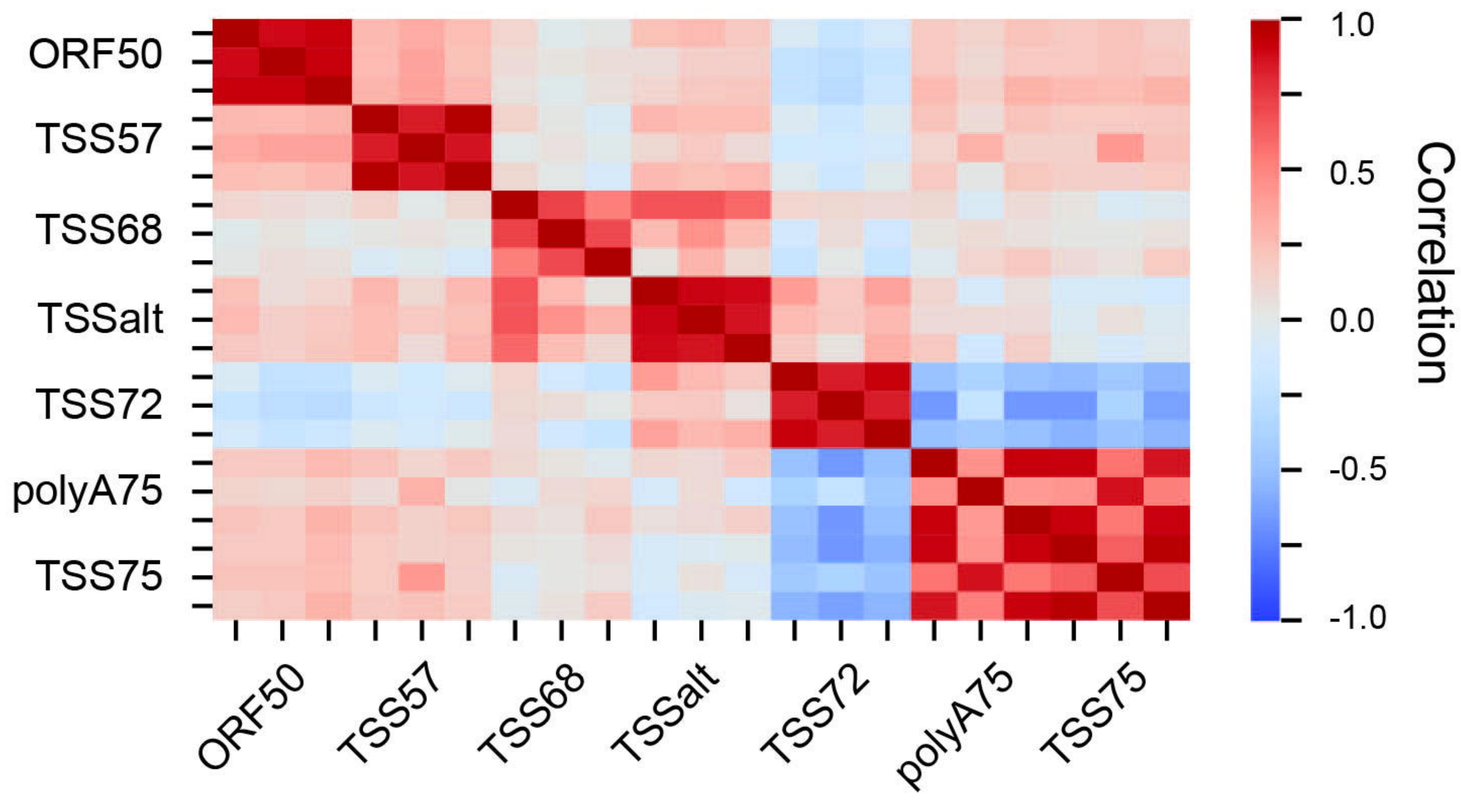

Exogenous promoter (doxycycline)
TSS57
ORF50
ORF57
TSS72
TSS75
ORF75
TSS68
HaloTag-ORF68
polyA75
TSSalt
Evidence
Noncoding element
Coding element
Knockout RT-qPCR
Coding association
Expression correlation
Undetermined

Figure S4

A


ORF50
TSS57
TSS68
TSSalt
TSS72
polyA75
TSS75
1.0
0.5
0.0
-0.5
-1.0
Correlation