# The polypyrimidine tract binding protein, PTBP1, regulates selenium homeostasis via the Selenoprotein P 3' untranslated region.

Sumangala P. Shetty, Nora T. Kiledjian and Paul R. Copeland*

Department of Biochemistry and Molecular Biology

Rutgers, Robert Wood Johnson Medical School

Piscataway, NJ

*To whom correspondence should be addressed:

Paul R. Copeland

Biochemistry and Molecular Biology

Rutgers-Robert Wood Johnson Medical School

675 Hoes Ln

Piscataway, NJ 08854

ph: (732) 235-4670

Email: paul.copeland@rutgers.edu

Keywords: selenocysteine, selenoprotein, selenium, SELENOP, 3' UTR, PTBP1, RNA binding, SECIS

Abstract

Selenoproteins contain the 21st amino acid, selenocysteine (Sec), which is incorporated at select UGA codons when the encoding mRNA contains a specialized hairpin sequence in its 3' UTR. This hairpin, the so-called Sec insertion sequence (SECIS) element, is found in all selenoprotein mRNAs, but the sequence surrounding these elements is widely variable and in many cases of considerable length. In order to determine the function of one such SECIS context, we chose to focus on the plasma selenoprotein, SELENOP, that is required to maintain selenium homeostasis. It is unique in that its mRNA contains two SECIS elements that lie in the context of a highly conserved 843-nucleotide 3' UTR. Prior work has attempted to examine the functions of the SECIS context but none were identified. Here we have used CRISPR/Cas9 genome editing to delete the region between the two SECIS elements. We found that this sequence is required to mediate an increase in SELENOP synthesis under conditions of peroxide stress. Using RNA affinity chromatography, we have identified PTBP1 as the major RNA binding protein that specifically interacts with this region.

## INTRODUCTION

The role of untranslated regions in post transcriptional gene regulation is well established. A pointed example of this is the requirement for a specific 3' UTR sequence that directs selenocysteine (Sec) incorporation. Sec is incorporated into specific sites within 25 human proteins at specific UGA codons, which would otherwise signal translation termination. Sec incorporation requires a stem-loop structure in the 3‘ UTR’s of mammalian selenoprotein mRNA known as SECIS element. All but one of the human selenoproteins contain a single Sec codon in their coding region, and a single SECIS element in their 3‘UTRs. The only exception to this is selenoprotein P (SELENOP), which in humans contains 10 Sec codons in the coding region and 2 SECIS elements in the 3’UTR. While the SECIS element is known to be necessary and sufficient for Sec incorporation, albeit with varying efficiency (Latrèche et al. 2009), the role of surrounding sequences in the extremely diverse array of selenoprotein 3' UTRs has only recently been the subject of investigation, for example in the case where sequence upstream of the selenoprotein S SECIS element regulates incorporation (Cockman et al. 2019). In this case, SELENOS mRNA is known to respond to cytokines (Curran et al. 2005), but a connection to modulation of Sec incorporation by regulatory sequences in the 3' UTR has not been made. The potential for regulation is substantial considering that the size of selenoprotein mRNA 3' UTRs range from ~300 to ~5000 nt (Gonzalez-Flores et al. 2013). In the case of SELENOP, the two SECIS elements sit in the context of 800 nt of well-conserved sequence. Surprisingly, systematic deletion analysis of non-SECIS sequences in the SELENOP 3' UTR did not reveal any role in regulating Sec incorporation in vitro or transfected cells (S. P. Shetty et al. 2018). We

expect that the conserved sequences in the SELENOP 3' UTR will impact three different aspects: 1) Sec incorporation efficiency and processivity, 2) general translation efficiency, and 3) mRNA stability. There is a hint that the sequence upstream of the first SECIS element may play a role in the first of these as there is a polymorphism 14 nt downstream of the stop codon that correlates with increased premature translation termination at the 2nd UGA codon (Méplan et al. 2007).

As a hepatokine regulator of whole body selenium homeostasis, SELENOP levels are regulated by a multitude of physiological conditions including insulin levels, exercise and, of course, selenium availability (reviewed in Saito 2020). To date, none of the studies have examined the specific role of 3' UTR sequences in mediating regulation. Since SELENOP supports general selenoprotein synthesis by delivering selenium to the extrahepatic tissues, it is a key player in maintaining cellular redox balance through the myriad of selenoenzymes that resolve oxidative stress. In addition, SELENOP itself has a thioredoxin motif that may directly be involved in maintaining redox balance. As such, SELENOP plays a central role in maintaining whole body redox balance.

Here we have used an oxidative stress model in human hepatoma cells to analyze SELENOP mRNA and protein expression in the context of a targeted deletion in the endogenous gene. Focusing on the highly conserved U-rich sequence between the two SECIS elements, we found that this region is required for regulating SELENOP translation during oxidative stress in human liver cells in culture. Using RNA affinity chromatography, we identified the polypyrimidine tract-binding protein (PTBP1) as one of the key RNA binding proteins that interacts specifically with this region.

RESULTS

**Targeting the SELENOP 3' UTR.**

The 821 nucleotide (nt) SELENOP 3'UTR is highly conserved among mammals (76% identity) and can be divided into 5 regions (Figure 1A): 1) the sequence upstream of SECIS-1 (254 nt); 2) SECIS-1 (83 nt); 3) the sequence between the two SECIS elements, hereafter the "interSECIS," (358 nt); 4) SECIS-2 (81 nt); 5) the sequence downstream of SECIS-2 (63 nt). One notable feature of the entire SELENOP 3'UTR is that the AU content is quite high at 71%. Since our prior work showed that the first SECIS is sufficient to promote wild-type levels of SELENOP in vitro and transfected cells, we chose to focus on the sequence downstream of SECIS-1, which has the highest AU content and alterations there should not impact the function of SECIS-1. In order to analyze the function of the interSECIS region, we generated a genomic deletion in HepG2 cells using CRISPR/Cas9 (Figure 1A "ΔinterSECIS"). Besides avoiding the artifacts associated with transient transfections, the creation of a genomic mutation was necessary for two reasons: 1) mammalian SELENOP cannot be expressed from transfected cDNA for unknown reasons, and 2) the expression of zebrafish SELENOP, which can be expressed in mammalian cells, is not affected by systematic deletion mutations in the 3' UTR (S. P. Shetty and Copeland 2018). We used a set of single guide RNAs (sgRNAs) targeting the region 37 nts downstream of SECIS-1 and 45nts upstream of SECIS-2. The sgRNAs were electroporated along with a GFP vector into HepG2 cells and 48 hours later the top 3% of fluorescent cells were sorted into a 96-well plate in a single-cell format. Upon confluence, the clonal populations were analyzed for genomic deletion by PCR of genomic DNA. As shown in Figure 2B,

genomic PCR verified the deletion, and sequencing of PCR fragments confirmed that the expected 282 nt fragment was eliminated. We also confirmed the genomic deletion by RT-PCR of RNA derived from the deletion clones (Figure 1B).

**Deletions in the SELENOP 3'UTR influences SELENOP expression**

In addition to screening clones by PCR, we performed a functional screen of GFP positive clones by metabolic labeling with $^{75}$Se-selenite. We expected to observe altered SELENOP expression in mutant clones, which would be observable as changes in $^{75}$Se incorporation into SELENOP, which is a 42 kDa secreted protein with multiple glycosylation sites, thus altering the apparent molecular weight. In addition, a fraction of the secreted protein is truncated as a result of translation termination at the 2nd UGA codon depending on selenium availability and cell type (S. P. Shetty, Shah, and Copeland 2014). As such, the expected products are observable as full length and truncated bands at 68 and 52 kDa respectively. Clonal GFP-positive cells were incubated with 100 nM $^{75}$Se-selenite for 24 hours. The media and cell lysates were analyzed by SDS PAGE followed by phosphorimager analysis. Figure 2A shows the two predicted bands for the untransfected control (Figure 2A, lane 1). However, under normal conditions, HepG2 cells produce mainly full length SELENOP (S. Shetty, Marsicano, and Copeland 2018). It is important to consider that the conditions used for electroporation and single cell cloning do not include supplemental selenium as is present during maintenance, and this is likely why we observe that the truncated form appears with the same intensity as full length. Considering that there are 10 Sec residues in full-length and only 1 in the truncated form, the efficiency of producing full length SELENOP is only about 10%. In the case of four clones (Figure 2A, lanes 2,3,5

and 6), we observed virtually no full length SELENOP production, suggesting a defect in processive Sec incorporation. To determine if the expression of any other selenoproteins was affected by the interSECIS deletion, we analyzed the cytoplasmic lysates of the metabolically labeled cells. In this case there were no significant differences in the expression of detectable intracellular selenoproteins (Figure 2B).

This analysis of clonal populations lacking the interSECIS region of SELENOP shown in Figure 2A indicates that this sequence may play a key role in promoting synthesis of full-length SELENOP. However, after establishing the mutant lines over several passages, this phenotype was lost. Figure 2C shows immunoblot and $^{75}$Se-selenite analysis of conditioned media derived from wild-type and ΔinterSECIS mutant lines after the cells had been passaged to the maintenance stage (~6 passages). Note, in order to improve electrophoretic resolution, the media was treated with PNGAseF, which removes both N= and O-linked glycosylation. After such treatment, the full length and truncated versions of SELENOP migrate at 42 and 30 kDa respectively, and we observed slightly more total SELENOP synthesis (both full length and truncated) in the ΔinterSECIS mutant. This increase correlated well with a 1.3 fold increase in SELENOP mRNA levels as determined by qRT-PCR (Figure 3C, right panel). Together these results establish that the SELENOP interSECIS region may have an impact on SELENOP mRNA stability and appears to be required to maintain wild-type levels of full length SELENOP expression under stressful conditions like single cell culture.

**Role of pre-SECIS-1 and InterSECIS during oxidative stress**

Considering the role of selenoproteins in responding to oxidative stress, we determined whether the intserSECIS region is required for SELENOP expression under stress conditions. We chose peroxide stress as a blunt but effective tool for inducing ROS in mammalian cell culture, and it is a well-established method to induce stress in HepG2 cells (Tani et al. 2019). We chose to use a broad range of concentrations for an extended time period to mimic a chronic stress that might be similar to that observed during the CRISPR clonal selection process. To assess the effect of peroxide treatment on SELENOP expression, we treated wild type and ΔinterSECIS mutant HepG2 cells with a range of hydrogen peroxide concentrations (5 - 1000 uM) in the presence of $^{75}$Se-selenite. Figure 3A and the graph in Figure 3B shows that peroxide treatment caused an ~1.5 fold increase in SELENOP levels in wild-type cells with maximum in the 50-250 uM range. As expected, the ΔinterSECIS mutant showed higher SELENOP expression than wild-type in the absence of treatment, and a slight further increase in response to peroxide. This result was also evident in the metabolic labeling of SELENOP with $^{75}$Se-selenite (Figure 3A, bottom panel). Interestingly, the ratio of full length to truncated SELENOP did not significantly change as a result of peroxide treatment, so peroxide treatment was unable to replicate the condition that caused predominantly truncated $^{75}$Se-selenite-labeled protein expression observed in Figure 2A. In addition, the peroxide treatment did not affect intracellular selenoprotein synthesis as shown by examining the $^{75}$Se-selenite-labeled cell lysates (Figure S1). These results indicate that the interSECIS sequence is required to moderate SELENOP expression so that levels can increase in response to stress. In order to determine the

contribution of RNA levels to the change in SELENOP expression, we performed qRT-PCR. Figure 3C shows that ΔinterSECIS SELENOP mRNA levels were generally higher than wild-type and increased to more than 2-fold over the wild-type level in response to peroxide. Interestingly, the maximal SELENOP protein expression occurred at just 5 uM peroxide while the maximal RNA levels occurred at 50 uM. This suggests that the interSECIS sequence may modulate mRNA stability as well as modulating the response to stress and that SELENOP production is maximal under limiting oxidative stress.

**Identification of interSECIS RNA binding proteins**

In order to determine if specific mRNA binding proteins may be involved in mediating the function of the interSECIS region, we performed RNA affinity chromatography based on the MS2/MS2 coat protein interaction (Selby and Peterlin 1990). As an initial unbiased approach we appended 3 copies of the MS2 sequence tag onto the 3' end of the full length SELENOP 3' UTR. As a control RNA we used a segment of the SECISBP2L coding region that had a similar GC content and the same length. These RNAs were synthesized in vitro and attached to glutathione agarose magnetic beads that had been pre-bound with GST-tagged MS2 coat protein, which binds with high affinity to the MS2 RNA sequence tag. For these experiments the rat SELENOP 3' UTR was used because they were constructed prior to the CRISPR/Cas9 analysis. As such, we began by analyzing extract from rat liver and the rat hepatoma cell line McArdle 7777. Figure 4A shows the proteins that eluted from control RNA as compared to those eluting from the SELENOP 3' UTR. We observed a series of bands at ~65 kDa that

specifically eluted from SELENOP 3' UTR but not the control RNA. To identify these proteins, the bands were excised from the gel and subjected to LC MS/MS. The predominant peptides identified corresponded to the polypyrimidine tract binding protein 1 (PTBP1). It is likely that the other bands correspond to some of the known PTBP1 isoforms (there are 7 transcript variants listed in the NCBI RefSeq database). To verify the identity of these bands, we repeated the RNA affinity chromatography with HepG2 lysate. Figure 4B shows that we obtained a similar result with three bands in the ~65 kDa range eluting specifically from the SELENOP 3' UTR. Immunoblot analysis using anti-PTBP1 monoclonal antibody confirmed that all three bands are recognized by the antibody (Figure 4C). In an effort to uncover additional RNA binding proteins, we repeated the RNA affinity chromatography and submitted the entire eluate for LC MS/MS analysis. Figure 4D shows the list of top hits that showed significant selectivity for the SELENOP 3' UTR. In addition to the predominant number of peptides corresponding to PTBP1, several other proteins were recovered, including SECISBP2 and SECISBP2L, serving as internal controls for the experiment. These results confirm the initial identification of PTBP1 as a SELENOP 3' UTR binding protein, and expand the list of candidates that may play roles in fine tuning SELENOP expression either through modulating translation or mRNA stability.

METHODS

**CRISPR/Cas9 genome editing**

Two guide RNAs against the interSECIS-region of SELENOP 3'UTR (Figure S3) were synthesized and cloned separately into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Ran et al. 2013). This vector was a gift from Feng Zhang (Addgene plasmid # 62988 ; http://n2t.net/addgene:62988 ; RRID:Addgene_62988). HepG2 cells were grown to 75% confluence in 10%FBS. 2.5 µg each of plasmid DNAs harboring the SELENOP targeting sgRNAs were electroporated using the Amaxa 4D Nucleofector protocol for HepG2 cells. of pmaxGFP plasmid (0.15 µg) was co-transfected to allowed calculation of electroporation efficiency. After 48 hrs GFP positive (top 3%) cells were sorted as single cells into a 96 well plate containing 50% HepG2 conditioned media. The single cell clones were grown for 3 weeks and then split and screened for deletions by $^{75}$Se-selenite labeling. The genome editing was confirmed using genomic DNA PCR with Herculase II enzyme

**$^{75}$Se Labeling and oxidative stress .**

Cells were grown to 80% confluence and then the complete EMEM media was replaced with serum free media which was supplemented with 100 nM of $^{75}$Se-selenite (specific activity 6.29µCi/µl; Research Reactor Center, University of Missouri, Columbia). After 12 hours, the media was collected and centrifuged at 2500g for five minutes at 4°C. The top 80% of the centrifuged media was transferred into a new tube. 1.5% of the total centrifuged media was used for analysis by 12% SDS-PAGE. The adhered cells were then gently washed with cold PBS and lysed in 1% NP-40 buffer (50mM Tris-HCl pH 8.0, 150 mM sodium chloride, 1% NP-40, Roche Complete protease inhibitor). The

lysate was then cleared by centrifugation at 17,000 x g for 10 min at 4°C. For oxidative stress, $^{75}$Se-selenite was added 24 hours prior to lysis and the cells were treated with the indicated range hydrogen peroxide starting six hours prior to lysis. The conditioned media and cell lysate was analyzed as described above.

**Quantitative RT-PCR**

Total RNA was purified using the RNAeasy kit (Qiagen) and reconstituted using nuclease free water and quantitated using a Nanodrop 2000c spectrophoptometer. 200 ng of total RNA was converted to cDNA using the reverse transcriptase (RT) Superscriptase III (Applied Biosystems). 20 µl reactions were set up according to the manufacturer's protocol using oligo dT primer. Gene specific forward and reverse primers were designed to amplify 57-71 base long reference and target amplicons (Table 1). Actin was used as a reference gene. qRT-PCR was performed using an ABI Step One Plus Real Time PCR System (Applied Biosystems) and PCR Master Mix (ABI) with Power SYBR Green (Invitrogen) and ROX reference Dye (Invitrogen). All reactions were performed in triplicate. The total reaction volume was 20 µl with 5 µl of 1:5 diluted RT reaction. Working concentration of the primers in the reaction was 0.25 µM. Thermal cycling conditions were 95°C for 10 min followed by 40 cycles of 95°C for 15 sec, 60°C for 1min. Melt curve analysis was performed for each sample to ensure a single amplification product. Samples were analyzed in triplicate for both the reference gene and the target gene. Quantitation was performed using the comparative ΔΔCt method. We used actin as the normalizer and the calibrator sample used was the endogenous human SELENOP from untreated samples. Primers in the acceptable efficiency range (90-110%) were determined using the standard curve method.

**Immunoblot analysis**

For SELENOP immunoblots, 15 µl of the lysate (10% of the total cell lysate which is ~30 ug) as described above was resolved by 12% SDS-PAGE blotted to nitrocellulose membrane (Amersham Biosciences), blocked in 5% nonfat dried milk for 1 hour at room temperature, and probed using a monoclonal horseradish peroxidase-conjugated (HRP) anti-SELENOP antibody (Thermo-Fisher 37A1) at a 1:1,000 dilution overnight, and was developed using the SuperSignal West Femto kit (Pierce) according to the manufacturer's protocol.

**Affinity Chromatography.**

The entire rat SELENOP 3'UTR was cloned into the pcDNA3.1 vector using the TOPO -TA cloning kit (Invitrogen). 3 copies of the viral RNA MS2 sequence separated by random sequence was synthesized commercially from IDT technologies (Table 1). These synthetic constructs were ligated downstream of the rat SELENOP 3' UTR sequence. As a control, we also ligated viral MS2 sequence to the 3'end of a non specific sequence corresponding to a fragment of the coding region for SECISBP2L, which has a similar GC content. sequence of similar length as SELENOP 3'UTR. The SELENOP 3'UTR and the control plasmid were linearized using Bsb1 and in vitro transcribed using Ribomax kit (Promega). The RNA was then purified using p30 size exclusion columns and quantified. The GST-MS2 protein was derived from a clone obtained from Rachel Green (HHMI) and was purified as described earlier (Youngman and Green 2005).

Bead preparation: 10ul of Magnetic Glutathione beads (Pierce) were incubated with 100 µg of purified GST-MS2 protein for 2 hours at 4˚C in buffer A (20 mM Tris -OAc, pH 7.5,

100 mM KOAc, 2 mM DTT, 2.5 mM $Mg(OAc)_2$, 0.25 mM spermidine, 0.4 mM GTP). The protein-bound beads were then washed with buffer A three times (10 minutes each) and then incubated with in vitro transcribed 10 µg of MS2 tagged RNA in buffer A containing RNAsin for an hour and again washed with buffer A. The beads were then incubated with precleared cell lysate for 2 hours at 4˚C and then washed buffer A three times (10 minutes each) and eluted with 1M NaCl for 20 minutes. The elution was sent for LC MS/MS analysis at the Rutgers Biological Mass Spectrometry Facility.

## DISCUSSION

All selenoprotein 3'UTR's contain an approximately 100 nt SECIS element required for recoding a UGA to allow Sec incorporation. However, most selenoprotein 3'UTR's contain long sequences adjacent to the SECIS element with no known function. It is therefore likely that selenoprotein 3' UTRs contain regulatory motifs in their 3' UTRs. Such motifs have been revealed over the last several decades as key regulators of translation, mRNA localization and mRNA decay (reviewed in Mayr 2019). Naturally, RNA binding proteins are key players in most 3' UTR-mediated regulatory processes, many serving as a platform for highly complex ribonucleoprotein complexes. One of the many pleiotropic RNA binding proteins that falls in this category is polypyrimidine tract binding protein (PTBP1), which has demonstrated roles in regulating pre-mRNA splicing, translation and mRNA stability (reviewed in Romanelli, Diani, and Lievens 2013). Here we have found that PTBP1 plays a role in regulating SELENOP synthesis through an interaction with its 3' UTR.

The human SELENOP mRNA contains multiple U-rich sequences, many of which are canonical PTBP1 binding sites (Figure S3). The fact that these sites cluster in the region between the two SECIS elements is consistent with our experimental determination that PTBP1 is only recovered from RNA affinity chromatography when the interSECIS region is present. In the absence of any stressors, we have found that loss of the PTBP1 binding site is correlated with a small but consistent increase in SELENOP mRNA, suggesting a role in mRNA decay. The addition of hydrogen peroxide as an inducer of oxidative stress revealed that SELENOP production increases in response and that the PTBP1 binding sites are required for this response. While the data suggest a potentially

complex mechanism, overall it is clear that the PTBP1 RNP, likely consisting of multiple proteins, plays a role in regulating SELENOP production.

Extensive analysis of SELENOP production in vitro and in transfected cells has failed to reveal a role for the sequences surrounding the SECIS elements (S. P. Shetty et al. 2018; S. P. Shetty, Shah, and Copeland 2014). Considering that the results obtained here indicate a regulatory role for the interSECIS region, it is likely that the functional RNP complex may form co-transcriptionally or coincident with RNA splicing. Considering the high concentration of PTBP1 in the nucleus and its role in splicing, it is logical that it binds to the mature SELENOP mRNA and is exported in a conformation that is able to respond to cellular conditions. In the case of oxidative stress, it is possible that the complex is removed, allowing increased SELENOP expression. The idea that an active SELNEOP mRNP may form during pre-mRNA processing may explain why it is not possible to express detectable protein from transfected cDNA (S. P. Shetty and Copeland 2018).

The response of selenoprotein expression to oxidative stress is surprisingly complex. While one might expect robust induction of selenoprotein expression under stress, only two selenoproteins have been reported to respond in the NRF2 dependent transcriptional response to reactive oxygen species (GPX2 and thioredoxin reductase). In a recent study in mammalian cells, treatment with hydrogen peroxide had a minimal positive impact on the selenoproteins analyzed but it increased SECIS-dependent expression of a stably transfected reporter construct by more than 5-fold (Touat-Hamici et al. 2014). Considering the apparent complexity of post-transcriptionally regulating selenoprotein expression to meet the needs of a variety of stress conditions it seems

likely that the 3' UTR sequences in each selenoprotein may be required for unique aspects of how selenoprotein expression responds to cellular signalling and stress.

ACKNOWLEDGEMENTS

This work was supported by funding from the National Institutes for Health R01GM077073 (PRC). We acknowledge technical support from Joanna Duong.

REFERENCES

Cockman, Eric M., Vivek Narayan, Belinda Willard, Sumangala P. Shetty, Paul R. Copeland, and Donna M. Driscoll. 2019. “Identification of the Selenoprotein S Positive UGA Recoding (SPUR) Element and Its Position-Dependent Activity.” *RNA Biology* 16 (12): 1682–96.

Curran, Joanne E., Jeremy B. M. Jowett, Kate S. Elliott, Yuan Gao, Kristi Gluschenko, Jianmin Wang, Dalia M. Abel Azim, et al. 2005. “Genetic Variation in Selenoprotein S Influences Inflammatory Response.” *Nature Genetics* 37 (11): 1234–41.

Gonzalez-Flores, Jonathan N., Sumangala P. Shetty, Aditi Dubey, and Paul R. Copeland. 2013. “The Molecular Biology of Selenocysteine.” *Biomolecular Concepts* 4 (4): 349–65.

Latrèche, Lynda, Olivier Jean-Jean, Donna M. Driscoll, and Laurent Chavatte. 2009. “Novel Structural Determinants in Human SECIS Elements Modulate the Translational Recoding of UGA as Selenocysteine.” *Nucleic Acids Research* 37 (17): 5868–80.

Mayr, Christine. 2019. “What Are 3’ UTRs Doing?” *Cold Spring Harbor Perspectives in*

*Biology* 11 (10). https://doi.org/10.1101/cshperspect.a034728.

Méplan, Catherine, Lynne K. Crosley, Fergus Nicol, Geoffrey J. Beckett, Alexander F. Howie, Kristina E. Hill, Graham Horgan, John C. Mathers, John R. Arthur, and John E. Hesketh. 2007. “Genetic Polymorphisms in the Human Selenoprotein P Gene Determine the Response of Selenoprotein Markers to Selenium Supplementation in a Gender-Specific Manner (the SELGEN Study).” *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 21 (12): 3063–74.

Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. 2013. “Genome Engineering Using the CRISPR-Cas9 System.” *Nature Protocols* 8 (11): 2281–2308.

Romanelli, Maria Grazia, Erica Diani, and Patricia Marie-Jeanne Lievens. 2013. “New Insights into Functional Roles of the Polypyrimidine Tract-Binding Protein.” *International Journal of Molecular Sciences* 14 (11): 22906–32.

Saito, Yoshiro. 2020. “Selenoprotein P as an in Vivo Redox Regulator: Disorders Related to Its Deficiency and Excess.” *Journal of Clinical Biochemistry and Nutrition* 66 (1): 1–7.

Selby, M. J., and B. M. Peterlin. 1990. “Trans-Activation by HIV-1 Tat via a Heterologous RNA Binding Protein.” *Cell* 62 (4): 769–76.

Shetty, Sumangala, John R. Marsicano, and Paul R. Copeland. 2018. “Uptake and Utilization of Selenium from Selenoprotein P.” *Biological Trace Element Research* 181 (1): 54–61.

Shetty, Sumangala P., and Paul R. Copeland. 2018. “The Selenium Transport Protein,

Selenoprotein P, Requires Coding Sequence Determinants to Promote Efficient Selenocysteine Incorporation." *Journal of Molecular Biology* 430 (24): 5217–32.

Shetty, Sumangala P., Ravi Shah, and Paul R. Copeland. 2014. "Regulation of Selenocysteine Incorporation into the Selenium Transport Protein, Selenoprotein P." *The Journal of Biological Chemistry* 289 (36): 25317–26.

Shetty, Sumangala P., Ryan Sturts, Michael Vetick, and Paul R. Copeland. 2018. "Processive Incorporation of Multiple Selenocysteine Residues Is Driven by a Novel Feature of the Selenocysteine Insertion Sequence." *The Journal of Biological Chemistry* 293 (50): 19377–86.

Tani, Hidenori, Ayaka Numajiri, Motohide Aoki, Tomonari Umemura, and Tetsuya Nakazato. 2019. "Short-Lived Long Noncoding RNAs as Surrogate Indicators for Chemical Stress in HepG2 Cells and Their Degradation by Nuclear RNases." *Scientific Reports* 9 (1): 1–10.

Touat-Hamici, Zahia, Yona Legrain, Anne-Laure Bulteau, and Laurent Chavatte. 2014. "Selective up-Regulation of Human Selenoproteins in Response to Oxidative Stress." *The Journal of Biological Chemistry* 289 (21): 14750–61.

Youngman, Elaine M., and Rachel Green. 2005. "Affinity Purification of in Vivo-Assembled Ribosomes for in Vitro Biochemical Analysis." *Methods* 36 (3): 305–12.

FIGURE LEGENDS

**Figure 1.**

Genome editing of the SELENOP 3' UTR. A) Diagram and multiple sequence alignment of the mammalian SELENOP 3' UTR. The alignment of sequences from the species indicated was performed using the MUSCLE algorithm, identical positions are black. B) PCR Analysis of genomic DNA (left panel) and cDNA (right panel) from CRISPR/Cas9 edited HepG2 cells.

**Figure 2.**

Edited HepG2 cells show initial defects in SELENOP synthesis that are lost after passaging. A) Single-cell derived clonal populations of HepG2 cells that had been electroporated with a plasmid expressing sgRNAs, Cas9 and GFP were incubated with 100 nM $^{75}$Se-selenite. 30 µl of conditioned media from the no DNA control and 6 clones was analyzed by SDS-PAGE and phosphorimage analysis (top) and Coomassie stain of the gel prior to drying and exposure to phosphor screen (bottom). B) Analysis of cell lysates by phosphorimage analysis (top) and Coomassie (bottom) from the cultures described in A). C). Conditioned media from wild-type (wt) and ΔinterSEICS (Δ IS) genome edited HepG2 cells after at least 6 passages was analyzed by anti-SELENOP immunoblot (left) and phosphorimage analysis (middle). Total RNA was isolated from these cells and analyzed by qRT-PCR normalized to actin and the wt set to 1.

**Figure 3.**

Hydrogen peroxide treatment reveals regulatory role for the interSECIS sequence. A) Wild-type (wt) or ΔinterSECIS genome edited HepG2 cells were treated with the hydrogen peroxide concentrations indicated and 100 nM $^{75}$Se-selenite for 6 hours. 30 µl

ofConditioned media was analyzed by SDS PAGE followed by immunoblot (top) or phosphorimaging (bottom). B) Quantitation of the band corresponding to full length (FL) SELENOP from immunoblot (top) and phosphorimage (bottom). C) Total RNA was isolated from these cells and analyzed by qRT-PCR normalized to actin and the wt set to 1.

**Figure 4.**

RNA affinity chromatography identifies PTBP1 as a SELENOP 3' UTR binding protein. A) Immobilized in vitro transcribed SELENOP 3' UTR (SP 3' UTR) or control RNA (Cntrl RNA) were attached to glutathione beads via an MS2/MS2 coat protein tag and either rat liver extract (left) or rat hepatoma (McArdle) cell lysate (right) was incubated with the beads. After washing, proteins were eluted in 1M NaCl and analyzed by SDS-PAGE and Coomassie staining. B) The entire eluate from the RNA affinity column incubated with McArdle cell lysate was analyzed by LC MS/MS. The most enriched proteins (ratio comparing SELENOP 3' UTR to control above 4) are shown. The two known SECIS binding proteins are highlighted. C) Same experiment as in A) except human hepatoma (HepG2) cell lysate was used. D) Anti-PTBP1 immunoblot analysis of eluate from panel C).

**Figure S1.**

Intracellular selenoprotein expression is not affected by hydrogen peroxide treatment of HepG2 cells. A) Wild-type (wt) or ΔinterSECIS genome edited HepG2 cells were treated with the hydrogen peroxide concentrations indicated and 100 nM $^{75}$Se-selenite for 6 hours. Cytoplasmic lysates were analyzed by SDS PAGE followed by Coomassie stain

(top) or phosphorimaging (bottom). B) Quantitation of total $^{75}$Se-selenite (FL) SELENOP from phosphorimage (bottom).

**Figure S2.**

Schematic of RNA affinity chromatography. Affinity matrix was assembled as shown and RNA protein complexes were eluted with 1M NaCl. Analysis of eluted RNA by gel electrophoresis is shown (two bands typically result from in vitro transcription indicating multiple structural forms).

Table 1 Oligonucleotide sequences used in this study

| | |
|---|---|
| SELENOP qPCR Fwd | CTAGGAGCTGATGCTGCCATT |
| SELENOP qPCR Rev | GGTGATTGCAGACCCTGTTTTT |
| Actin qPCR Fwd | GCGCGGCTACAGCTTCA |
| Actin qPCR Rev | CTTAATGTCACGCACGATTTCC |
| Genomic PCR Primer Fwd | ATATTTAAAATAGGACATACTCCCC |
| Genomic PCR Primer Rev | CAGCTTTAAGGTTTTTATTGAATTTATTTG |
| MS2 sequence | CTCGAGACTAGTAGATCTTTTTTTACTAGTAGAT CTTTTTTTGATGAGGATTACCCATCACTAGTAGA TCTTTTTTTGATGAGGATTACCCATCACTAGTAG ATCTTTTTTTACTAGTAGATCTTTTTTTGATGAG GATTACCCATCACTAGTAGATCTTTTTTTACTAG TAGATCTTTTTTTCTCGAG |

Figure 1, Shetty et al

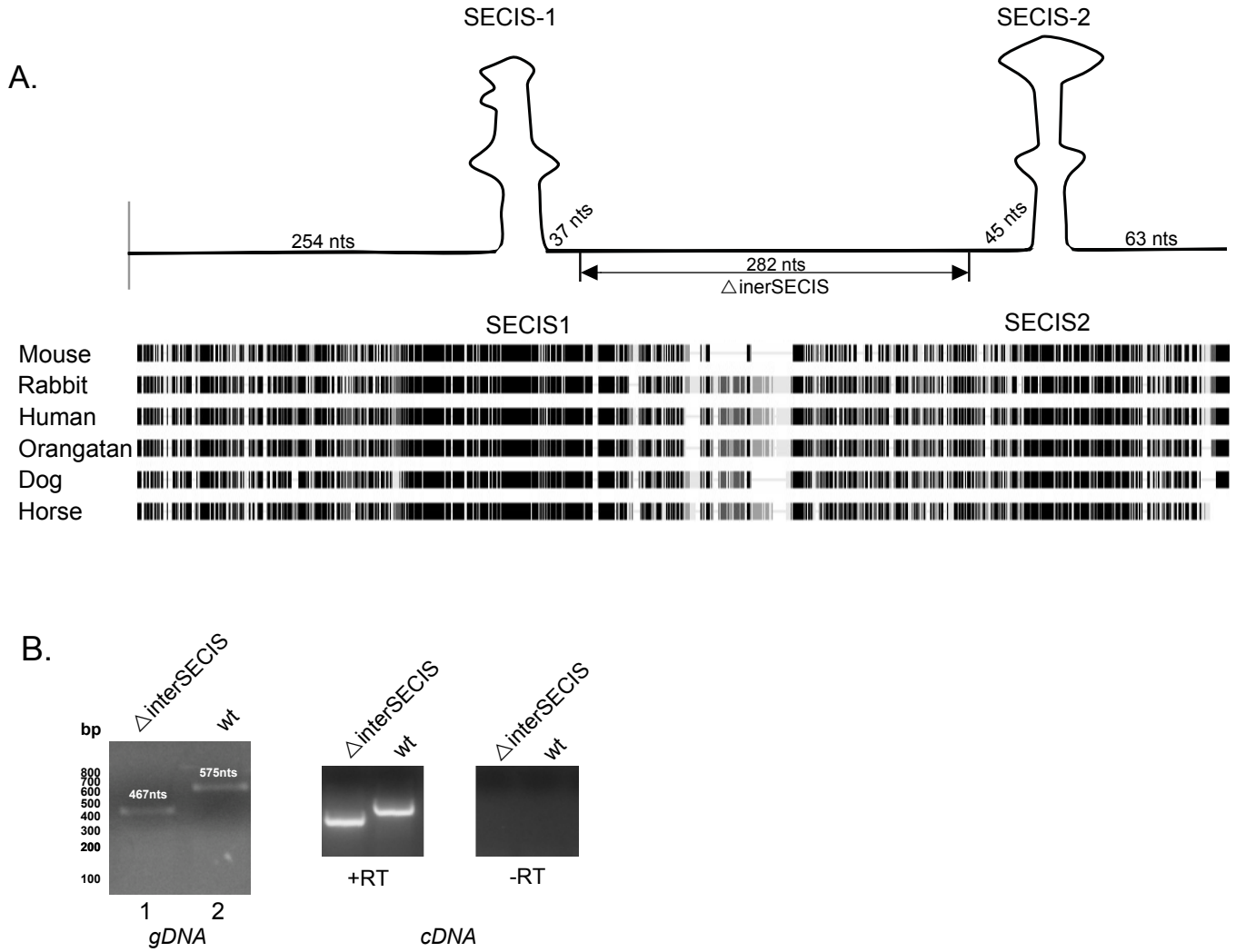


Figure 2, Shetty et al

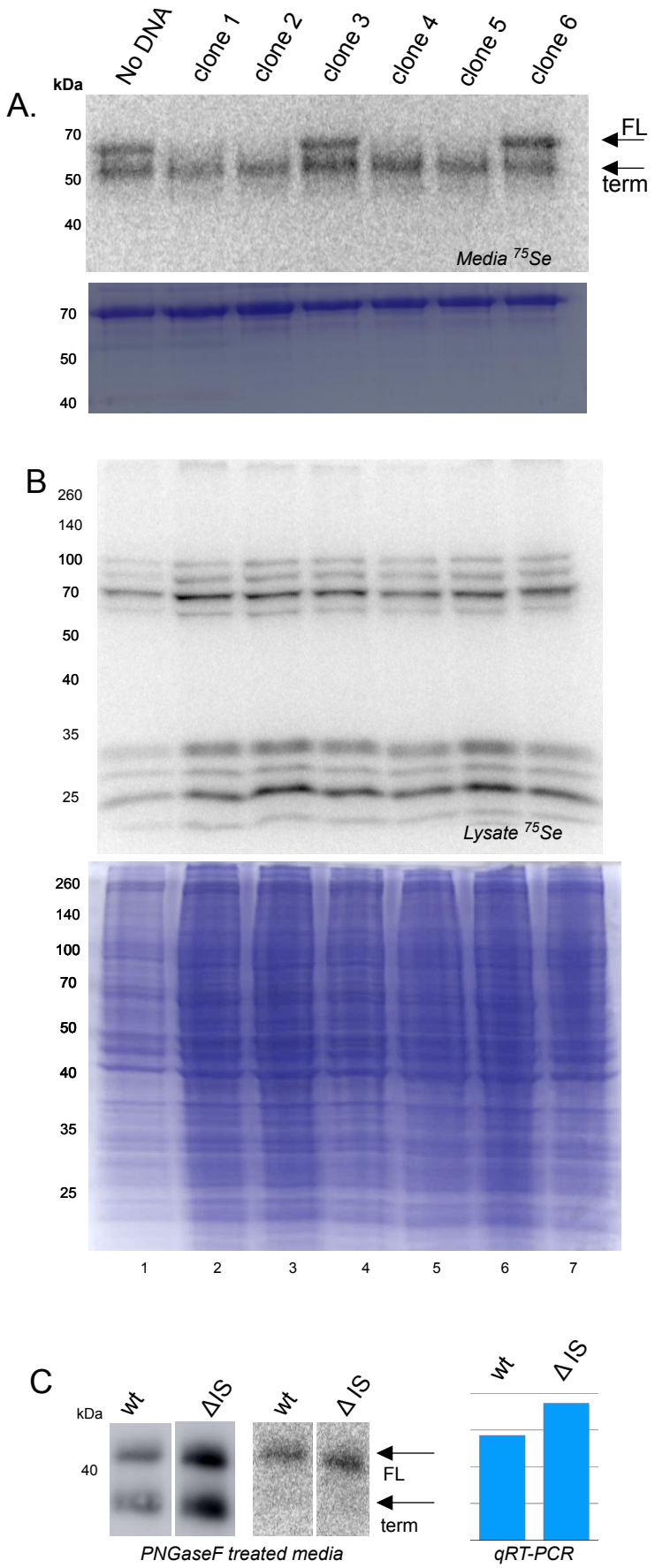
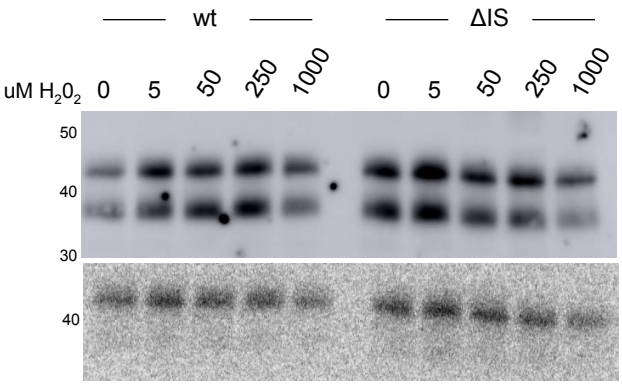
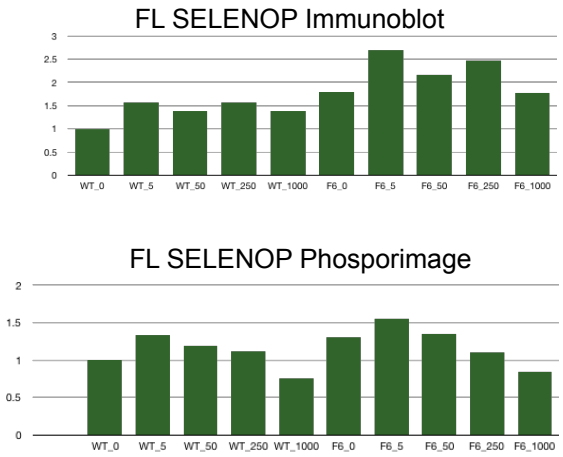


Figure 3, Shetty et al

A.




B.




C.

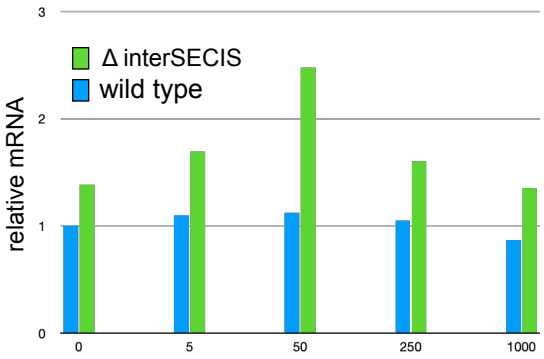
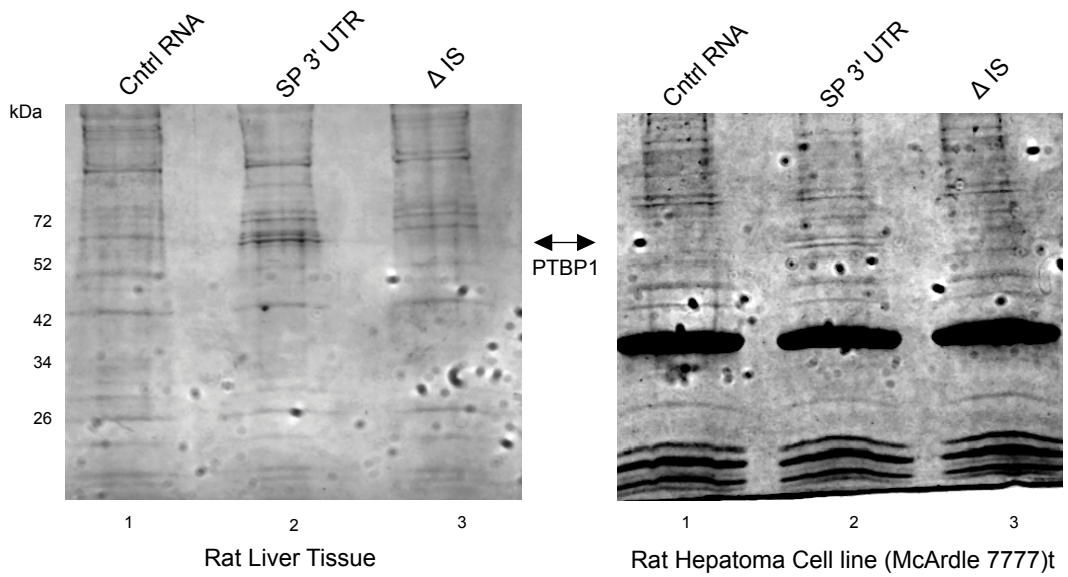
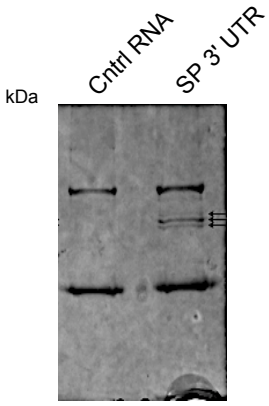


Figure 4, Shetty et al

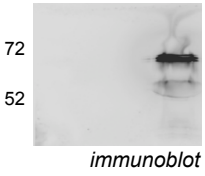
A




B




C

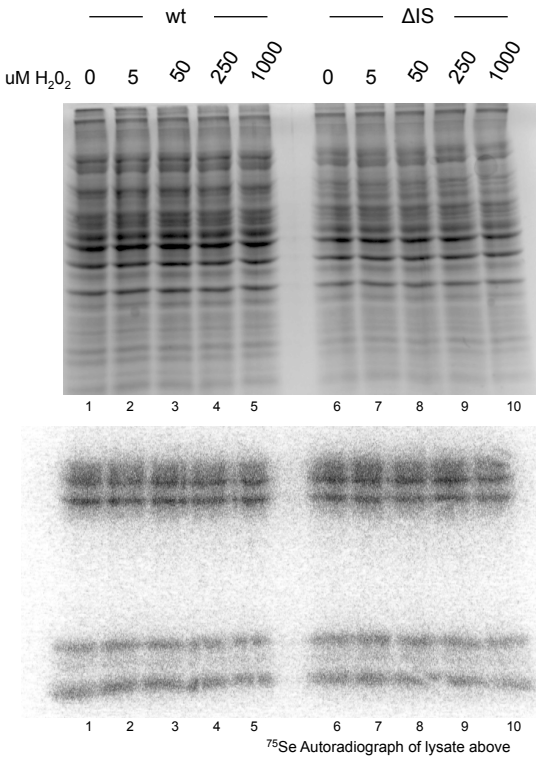



D.

| UniProt ID | Description | Control RNA | SP 3' UTR | Ratio |
|---|---|---|---|---|
| ENSRNOG00000020710 | Raver-1, Ribonucleoprotein PTB-binding 1 | 1 | 96 | 96 |
| ENSRNOG00000010827 | Polypyrimidine tract-binding protein 2 | 1 | 80 | 80 |
| ENSRNOG00000016334 | Polypyrimidine tract-binding protein 3 | 1 | 73 | 73 |
| ENSRNOG00000002408 | RNA-binding protein 47 | 1 | 55 | 55 |
| ENSRNOG00000001069 | ELAV-like protein 1 | 1 | 54 | 54 |
| ENSRNOG00000010448 | Polypyrimidine tract-binding protein 1 isoform b | 19 | 782 | 41 |
| ENSRNOG00000019875 | Matrin-3 | 5 | 162 | 32 |
| ENSRNOG00000033195 | APOBEC1 complementation factor | 1 | 28 | 28 |
| ENSRNOG00000002292 | Heterogeneous nuclear ribonucleoprotein C | 1 | 22 | 22 |
| ENSRNOG00000018295 | uncharacterized protein LOC306007 | 1 | 15 | 15 |
| ENSRNOG00000006929 | Protein Hnrpll | 1 | 11 | 11 |
| ENSRNOG00000013773 | selenocysteine insertion sequence-binding protein 2 | 3 | 32 | 11 |
| ENSRNOG00000008629 | SECIS binding protein 2-like | 3 | 20 | 7 |
| ENSRNOG00000003645 | heterogeneous nuclear ribonucleoprotein A/B | 4 | 22 | 6 |
| ENSRNOG00000020235 | heterogeneous nuclear ribonucleoprotein L isoform b | 21 | 112 | 5 |
| ENSRNOG00000019113 | Heterogeneous nuclear ribonucleoprotein K | 7 | 37 | 5 |

Figure S1, Shetty et al

A.




B.

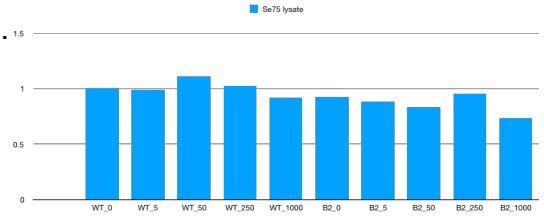
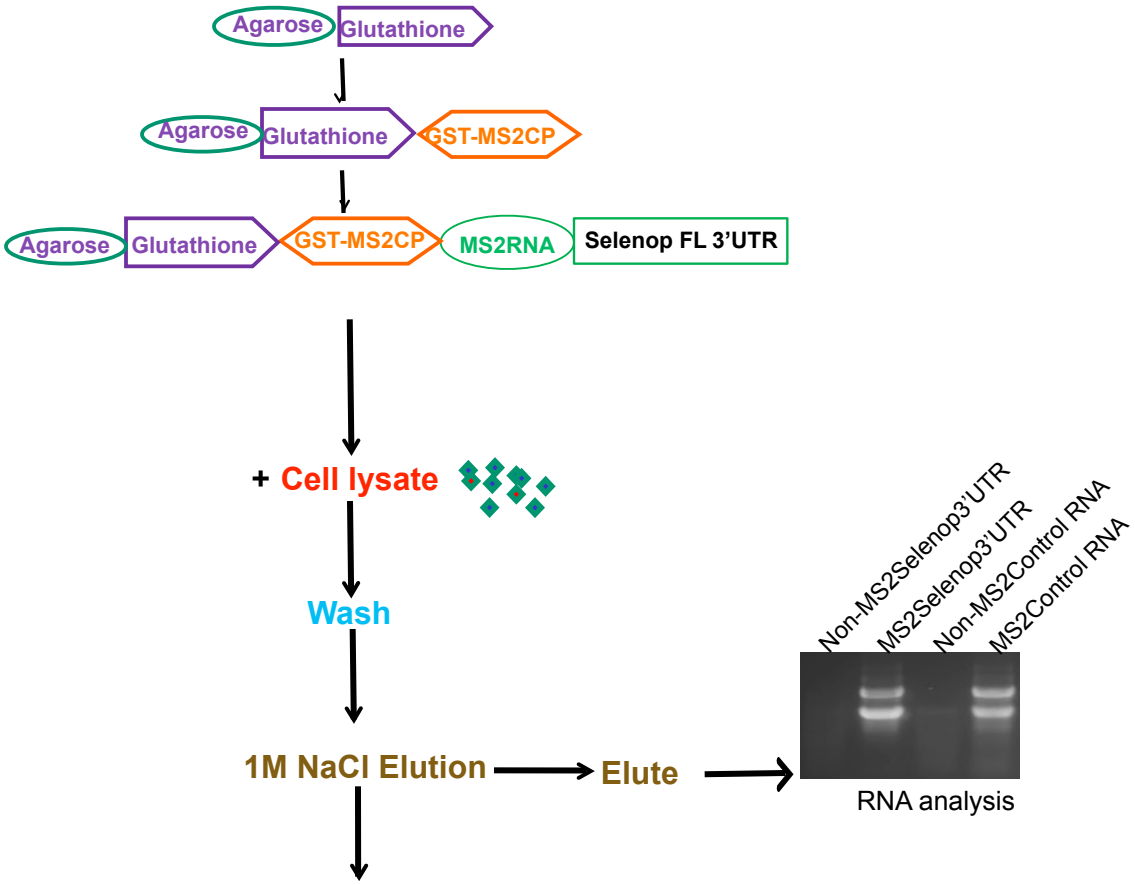



Figure S2, Shetty et al