A broad response to intracellular long-chain polyphosphate in human cells

Emma Bondy-Chorney[1,3], Iryna Abramchuk[2,3,4], Rawan Nasser[1,3,4], Charlotte Holinier[1,3], Alix Denoncourt[1,3], Kanchi Baijal[1,3], Liam McCarthy[1,3], Mireille Khacho[2,3], Mathieu Lavallée-Adam[2,3], Michael Downey[1,3]

1. Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada, K1H 8M5

2. Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada, K1H 8M5

3. Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, Ontario, Canada, K1H 8M5

4. Equal contribution from these authors

* Michael Downey and Emma Bondy-Chorney

Email: mdowne2@uottawa.ca and ebond062@uottawa.ca

ORCID iDs: 0000-0002-5234-6666 (M.D), 0000-0003-1957-0209 (E.B.-C.)

Classification

Biological Sciences. Cell Biology.

Keywords

Polyphosphate, lysine polyphosphorylation, ERK1/2, EGR1, mTOR, Ppk1

Author Contributions

E.B.-C. and M.D. devised experiments and wrote the paper. E.B.-C., I.A, R.N., C.H., A.D., K.B., L.M., M.K, M.L.-A., and M.D. carried out experiments and/or analysis. All authors reviewed the manuscript.

This PDF file includes:

Main Text

Figures 1 to 7

Supplemental Figure 1 to 2

Supplemental Tables 1 to 4

Abstract

Polyphosphates (PolyP) are composed of long chains of inorganic phosphates linked together by phosphoanhydride bonds. They are found in all kingdoms of life, playing roles in cell growth, infection, and blood coagulation. A resurgence in interest in polyP has shown links to diverse aspects of human disease. However, unlike in bacteria and lower eukaryotes, the mammalian enzymes responsible for polyP metabolism are not known. Many studies have resorted to adding polyP to cell culture media, but it is not clear if externally applied polyP enters the cell to impact signaling events or whether their effect is mediated exclusively by extracellular receptors. For the first time, we use RNA-seq and mass spectrometry to define a broad impact of polyP produced inside of mammalian cells via ectopic expression of the *E. coli* polyP synthetase Ppk1. RNA-seq demonstrates that Ppk1 expression impacts expression of over 350 genes enriched for processes related to transcription and cell motility. Analysis of proteins via label-free mass spectrometry identified over 100 changes with functional enrichment in cell migration. Follow up work suggests a role for internally-synthesized polyP in promoting activation of mTOR and ERK1/2-EGR1 signaling pathways implicated in cell growth and stress. Finally, fractionation analysis shows that polyP accumulated in multiple cellular compartments and was associated with the relocalization several nuclear/cytoskeleton proteins, including chromatin bound proteins DEK, TAF10, GTF2I and translation initiation factor eIF5b. Our work is the first to demonstrate that internally produced polyP can activate diverse signaling pathways in human cells.

Significance Statement

For many years following its discovery in 1890, polyphosphates (polyP) were dismissed as evolutionary fossils. Best understood for its role in bacteria and yeast, our understanding of polyP in mammals remains rudimentary because the enzymes that synthesize and degrade polyP in mammalian systems are currently unknown. In our work, we carried out large-scale transcriptome and proteome approaches on human cells designed to accumulate internally produced polyP via ectopic expression of a bacterial polyP synthetase. Our work is the first to systematically assess the impact of increased intracellular polyP.

Introduction

Polyphosphates (polyP) are molecules comprised of inorganic phosphates joined by high energy phosphoanhydride bonds in chains of 3-1000 units in length. Although polyP itself appears to be present across all kingdoms of life, its mechanisms of synthesis appear poorly conserved. In bacteria, polyP is synthesized by the Ppk1 enzyme that uses ATP as a substrate (Ahn and Kornberg, 1990). In the enterobacterium *E. coli* in particular, polyP levels are low during logarithmic growth in rich media but increase dramatically to approximately 20 mM (measured in terms of individual phosphate units) in response to a variety of stresses such as amino acid starvation (Rao et al., 1998). These chains range from 100-1000 units (Kornberg et al., 1999; Kumble et al., 1996). PolyP accumulation is countered by Ppx, an exopolyphosphatase that degrades polyP beginning at the end of the chain (Akiyama et al., 1993; Rashid et al., 2000). In yeast, polyP is synthesized by the VTC complex (Hothorn et al., 2009). The VTC complex is situated within the vacuolar membrane and, like Ppk1, synthesizes polyP using ATP as a substrate. PolyP is translocated into the vacuole as it is synthesized (Gerasimaitė et al., 2014), where it accumulates to remarkably high levels (>200 mM total cellular concentration, comprising up to 10% of the dry weight of the cell) (Auesukaree et al., 2004; Kornberg et al., 1999). Lower levels of polyP can be found in other subcellular compartments including the cytoplasm, nucleus, plasma membrane and mitochondria (Gerasimaitė and Mayer, 2016; Lichko et al., 2006). How polyP localizes to these areas of the cell is not understood.

In stark contrast to bacteria and yeast, the enzymes that synthesize polyphosphate in mammalian cells have remained elusive (Desfougères et al., 2020). There are no clear Ppk1 or VTC homologs in mammalian cells. On the other hand, Prune, previously characterized as an interactor of GSK-3β and a regulator of Wnt signaling can degrade polyP *in vitro* (Tammenkoski et al., 2008), but it is not clear if this enzyme has any role in polyP metabolism *in vivo.* The same is true for alkaline phosphatase and three members of the Nudix family of hydrolases, DIPP1-DIPP3, which are homologs of yeast Ddp1 (Lonetti et al., 2011). Whole-cell levels of polyP (25-120 μM – Pi residues) in human cells and tissues (Kumble and Kornberg, 1995) are considerably lower than those measured in bacteria or yeast. Higher local concentrations may exist in certain subcellular compartments. For example, Jimenez-Nunez *et al.* reported high local concentrations of polyP in the nucleus of myeloma cells (Jimenez-Nuñez et al., 2012). PolyP also accumulates to levels of 130 mM in the dense granules of platelets (Ruiz et al., 2004), a lysosome-related organelle similar in many ways to the yeast vacuole. Release of polyP from these structures is thought to play an important role in the activation of the blood coagulation cascade (Travers et al., 2015). The accumulation of polyP in these structures may suggest elements of conservation with yeast in terms of polyP synthesis or storage.

With no enzymes responsible for *in vivo* polyP synthesis and degradation identified, studying polyP in mammalian systems presents a unique challenge. Several groups have expressed yeast Ppx1 exopolyphosphatase enzyme in mammalian cells in an attempt to deplete endogenous polyP. Wang *et al.* first

used this approach to study the role of polyP role in mammalian target of rapamycin (mTor) activation in MCF-7 mammary cancer cells (Wang et al., 2003). Subsequently others have used ectopic expression of Ppx1 to study the role of PolyP in mitochondria (Abramov et al., 2007; Seidlmayer et al., 2012), wound healing (Simbulan-Rosenthal et al., 2015), the DNA damage response (Bru et al., 2017), and glycolysis (Nakamura et al., 2018).

The impact of increased polyP has been studied by adding polyP chains of various lengths directly to culture media (Dinarvand et al., 2014; Hassanian et al., 2015; Holmström et al., 2013; Xie et al., 2019). For example, previous work from the Rezaie group showed that polyP chains of 70 units interact with the RAGE and P2Y1 receptors on the plasma membrane and activates the mTORC1 and TORC2 pathways in endothelial cells (Dinarvand et al., 2014; Hassanian et al., 2015). In these and other studies it is not clear whether polyP is entering the cell and/or whether this is required for the observed effects. Indeed, the requirement of external receptors for activation of these pathways suggests that polyP may be acting extracellularly (Dinarvand et al., 2014; Hassanian et al., 2015). In this study, we investigate the impact of internally-synthesized polyP in mammalian cells produced via ectopic expression of the *E. coli ppk1+* gene. We report that ectopic expression of the Ppk1 protein results in polyP accumulation throughout the cell, including the nucleus. Using transcriptomic and proteomic analyses, we demonstrate a broad impact of polyP on diverse pathways, including activation of the ERK1/2-EGR1 signaling axis. PolyP accumulation also results in redistribution of chromatin bound proteins DEK, TAF10 and GTF2I, and

translation initiation factor eIF5b from the nucleus/cytoskeleton to the cytoplasm. Our work will serve as a novel resource to interrogate polyP biology in higher eukaryotes.

RESULTS:

Internal synthesis of polyP in mammalian cells:

To determine the impact of polyP accumulation in mammalian cells, we transiently transfected either an empty vector (control) or a vector expressing the *E. coli ppk1+* gene under the CMV promoter **(Fig. 1A)**. The *ppk1+* gene encodes Ppk1, the enzyme responsible for synthesizing polyP chains in *E. coli* (Ahn and Kornberg, 1990; Kornberg et al., 1956). Ppk1 expression in mammalian cells has been used previously as a gene reporter system in magnetic resonance imaging techniques (Ki et al., 2007). We also previously used this experimental set-up to identify the first human targets of polyphosphorylation, a recently described post-translational modification wherein polyP chains are covalently attached to lysine residues (Bentley-DeSousa et al., 2018). However, we did not directly examine polyP production or downstream impacts of its accumulation. Ppk1 expression resulted in accumulation of long-chain polyP (>130 residues) in HEK293T cells **(Fig. 1B, lanes 6 and 7)**. The size distribution of polyP produced is similar to the polyP made in *E. coli* under starvation (stress), and contrasts with the shorter and more heterogeneous chains synthesized by the VTC complex in yeast **(Fig. 1B, lanes 8 to 11)**. As we demonstrated previously, Ppk1 expression caused a dramatic decrease in electrophoretic mobility for polyphosphorylation targets Nucleolin and MESD when analyzed on NuPAGE gels **(Fig. 1C)**. This change in electrophoretic mobility, visible on NuPAGE, but not SDS-PAGE gels, is a

hallmark of lysine polyphosphorylation (Azevedo et al., 2015; Bentley-DeSousa et al., 2018).

Next, we investigated the intracellular localization of the recombinant Ppk1 protein. As there are no commercially available antibodies for *E. coli* Ppk1 protein, we expressed an HA-tagged Ppk1 and carried out cell fractionation experiments. Commonly used cell fractionation markers showed expected separation for control and HA-Ppk1 samples **(Fig. 1D)**. We found that some HA-Ppk1 protein was localized in each of these compartments, with the majority observed in the cytoplasmic and nuclear/cytoskeletal fractions **(Fig. 1D).** Surprisingly, polyP concentration did not correlate with the level of HA-Ppk1 expression. While polyP was observed in all fractions, the highest concentrations were in the cytoplasm and the lowest in the nuclear/cytoskeletal fractions **(Fig. 1E, lanes 10 - 13)**. Nevertheless, these results demonstrate that mammalian cells are capable of supporting the ectopic expression of Ppk1 and the accumulation of polyP *in vivo.*

No significant impact of polyP on cell-cycle, apoptosis or autophagy markers:

With a system in place to successfully induce accumulation of polyP inside mammalian cells, we set out to determine the impact of elevated and internally-synthesized polyP on common cell signaling pathways. PolyP is important for cell cycle regulation in yeast and bacteria (Bru et al., 2016; Racki et al., 2017). Coordination of the cell cycle is a highly complex process involving the activities

and levels of numerous proteins, including cyclins (Ingham and Schwartz, 2017; Schafer, 1998). Altered levels of cyclins serve as an indicator of cell-cycle changes (Malumbres and Barbacid, 2005; Yam et al., 2002). However, western blotting analysis revealed no apparent changes in cyclins protein levels when Ppk1 was expressed as compared to control, suggesting no gross disruption of normal cell cycle activities **(Supplemental Fig. 1A and B)**. Next, we examined common markers of apoptosis (programmed cell death), which is regulated by a balance between pro- and anti-apoptotic pathways, including transcriptional regulation (Kumar and Cakouros, 2004; Portt et al., 2011). Indeed the external application of polyP has been reported to induce apoptosis in select human plasma cells (Hernandez-Ruiz et al., 2006) or under specific conditions, such as in combination with chemotherapeutic agents (Xie et al., 2019). We used RT-qPCR to determine the levels of *p53*, *BAX*, and *CASPASE-9*, three mRNAs commonly involved in the regulation of apoptotic pathways (Cummings, 1996; Liu and Lobie, 2007; Raman et al., 2000). Analysis of this data suggested no consistent change of these three apoptosis-related mRNAs (**Supplemental Fig. 1C)**. Lastly, we tested if we could detect changes in autophagy, a highly-conserved intracellular degradation and recycling system (Shibutani et al., 2015), in Ppk1-expressing cells. Measuring the conversion of LC3-I to LC3-II by immunoblotting is a commonly used method for monitoring autophagy (Mizushima et al., 2010). We saw no significant change in the LC3-I/LC3-II ratio between control and Ppk1-expressing cells **(Supplemental Fig. 1D)**. Taken together, these data suggest that the expression of Ppk1 in HEK293T cells does

not significantly disrupt cell cycle progression and does not induce apoptosis, or autophagy. To take a broad, unbiased look at the impact of Ppk1 expression, we used RNA-sequencing (RNA-seq) and label-free mass spectrometry to uncover changes in gene and protein expression.

Hundreds of differentially expressed genes in cells making excess polyP:

To investigate the impact of polyP accumulation on the transcriptome, we carried out a large-scale RNA-seq analysis in HEK293T cells **(Fig. 2A)**. NuPAGE analysis of known polyphosphorylated proteins DEK and MESD were used to confirm polyP production in Ppk1 samples used for RNA-seq **(Fig. 2B)**. We found 313 downregulated and 47 upregulated genes (FDR-adjusted p-value<0.05, $\log_2$ -fold change threshold > 1.25 and < 0.8) in Ppk1-expressing cells **(Fig. 2C and Supplemental Table 2)**. Among the top protein-coding genes that showed significant downregulation in Ppk1-expressing cells were three members of the Inhibitor of DNA-binding/differentiation (ID) family (*ID1*, *ID2*, *ID3*), *CCR4*, *DHRS2*, and *SFRP4*. Included among the genes that showed significant upregulation included *MYH15*, *VGF*, *ETV5*, and multiple members within the same protein family, such as the Early Growth Response (EGR) transcription factor family (*EGR1*, *EGR3*). Notably, top-ranked down and upregulated genes also included pseudogenes and lncRNAs **(Supplemental Table 2)**. To validate hits from our RNA-seq we carried out RT-qPCR on 12 candidate targets that were predicted to have differential expression based on the RNA-seq data **(Fig. 2D - F and Supplemental Table 2 and 3)**. All results shown are normalized to controls

within paired biological replicates. In total, 9 out of the 12 genes tested showed statistically significant differences in the direction that was predicted by the RNA-seq data including *HIST1H1D*, *ETV5*, *ID3*, and *CCR4*, resulting in a validation rate of 75% **(Fig. 2D - F)**. Additionally, we confirmed that 2 genes that we worked with previously, *NCL* and *NOP56* (Bentley-DeSousa et al., 2018), did not show any significant differences in differential expression, as predicted by the RNA-seq analysis **(Fig. G)**. We conclude that our dataset is of high quality and that polyP production in mammalian cells stimulates significant reprograming of gene expression. GO-term enrichment analysis performed on all significantly differentially expressed genes revealed 11 enriched terms **(Fig. 2H and Supplementary Table 2**; FDR-adjusted p-value < 0.05**)**. The top 4 most significantly enriched GO-terms were associated with DNA, specifically transcription regulation and DNA binding. Altogether, these results show significant changes in the transcriptome in cells accumulating internally-synthesized polyP.

Global changes in protein levels in polyP producing cells:

We next investigated if changes in gene expression were accompanied by changes in protein levels. We used label-free mass spectrometry to measure global protein levels in five biological replicates of HEK293T cells transiently transfected with either Ppk1 or control plasmids **(Fig. 3A)**. NuPAGE analysis was again used to test for polyphosphorylation of DEK and MESD to confirm production of polyP in samples processed for mass spectrometry **(Fig. 3B)**. Our

analysis of the mass spectrometry data revealed statistically significant changes in the levels of 14 proteins in cells with Ppk1 expression as compared to control **(Fig. 3C; FDR<15%)**. Importantly, we detected our recombinant Ppk1 protein as present only in Ppk1 conditions **(Supplemental Table 4)**. Amongst the proteins identified was CUL1, two Minichromosome Maintenance Complex Component (MCM) proteins MCM3 and MCM4, and RRBP1. We further characterized an additional group we called “singletons”, composed of proteins that show changes in their levels but could not be included in our statistical analysis as they were not detected in all five biological replicates in both conditions **(Fig. 3D, details defining singletons in materials and methods)**. This group included Dehydrogenase/Reductase 2 (DHRS2), a protein encoded by the *DHRS2* gene, which was also one of the hits we uncovered in our RNA-seq to be decreased with Ppk1 expression **(Supplemental Table 2)** and discussed further below. Further comparative analysis revealed 95.6% of the proteins detected in our mass spectrometry did indeed have their corresponding RNAs detected in our RNA-seq. Interestingly, DHRS2 and DHCR24 were the only hits from the mass spectrometry that also showed significant changes in the RNA-seq. This data suggests that the proteomic changes we observe with Ppk1 expression may be occurring at a post-transcriptional level. GO-term enrichment analysis of the significantly differentially expressed and singletons proteins revealed only one significantly enriched term, positive regulation of epithelial cell migration (adjusted p-value = 0.028).

Extracting the protein-protein interactions between the significantly differentially expressed proteins and singletons using STRING (Szklarczyk et al., 2019) revealed several subsets of conserved interaction networks **(Supplemental Fig. 2)**. We further analyzed this STRING network by identifying GO-terms that were associated with the individual nodes with this network. From this, we saw several dense subnetworks where the majority of proteins within the nodes were associated with a specific GO-term, including RNA processing, DNA replication, and transcription elongation from the RNA polymerase II promoter. In these cases, very few proteins associated with these terms were identified outside of the isolated nodes. Finally, we also saw that a number of the proteins within our overall STRING-derived network were associated with the GO-terms nucleic acid metabolic processes and cellular nitrogen compound metabolic processes **(Supplemental Fig. 2)**. Altogether, these results confirm that internal polyP production is associated with changes at the proteome level.

Internal synthesis of polyP leads to downregulation of the dehydrogenase/reductase DHRS2:

Taken together, our RNA-seq and mass spectrometry data reveals global changes in both the transcription and proteome of Ppk1-expressing cells. To extend and further validate these analyses, we followed up on several of these targets including one that came out as a top hit in both our RNA-seq and mass spectrometry experiments, *DHRS2*. *DHRS2* is a gene that encodes for the DHRS2 protein, also known as Hep27 (Donadel et al., 1991; Gabrielli et al.,

1995). DHRS2 is a member of the short-chain dehydrogenases/reductases family and has been reported to catalyze the NADPH-dependent reduction of specific dicarbonyl compounds (Shafqat et al., 2006). Data from our RNA-seq revealed that *DHRS2* mRNA levels were decreased by 3.44 fold (deseq. adjusted p-value = 2.90 x 10-14) in Ppk1-expressing cells **(Supplemental Table 2)**. Validation by RT-qPCR supported this expression level change, as *DHRS2* was downregulated by 3.47 fold (p-value = 0.013) with Ppk1 expression **(Fig. 4A)**. To investigate whether the decrease in the *DHRS2* mRNA levels was also reflected in the protein levels, we examined DHRS2 protein changes by western blot **(Fig. 4B)**. In confirmation of our mass spectrometry data **(Fig. 3D and Supplemental Table 4)** we observed a significant decrease in DHRS2 protein, as compared to normalized control **(Fig. 4B and C)**. DHRS2 is reported to stabilize p53 levels via Ser15 phosphorylation in esophageal squamous-cell carcinoma cells (Zhou et al., 2018). However, we observed no changes in p53 Ser15 phosphorylation in Ppk1-expressing cells, as compared to controls **(Fig. 4B and D)**. This difference in our results and previously published data may stem from the use of different cell lines, as detailed in the discussion. It would be interesting to test the impact of polyP-induced DHRS2 downregulation on additional pathways regulated by DHRS2, such as dicarbonyl reduction (Shafqat et al., 2006).

PolyP synthesis activates ERK1/2 signaling:

Analysis of our RNA-seq data revealed that upon Ppk1 expression, *EGR1* levels were increased by 2.28 fold (deseq. adjusted p-value = 2.10-24) **(Supplemental Table 2)**. EGR1 is a transcription factor that plays roles in many cellular pathways including cell growth, differentiation, and survival (Adamson and Mercola, 2002; Milbrandt, 1987; Sukhatme et al., 1988; Veyrac et al., 2014). Analysis by RT-qPCR and western blotting demonstrated that both EGR1 mRNA and protein levels were increased with Ppk1 expression **(Fig. 5A - C)**. *EGR1* mRNA transcription is regulated by activation of the ERK1/2 pathway (Roskoski Jr, 2012). Indeed, we found that Ppk1 expression induced ERK1/2 phosphorylation on tyrosine residues Y202 and Y204, without impacting overall ERK1/2 levels **(Fig 5. D and E)**. These residues are phosphorylated by the MEK1/2 kinases in the signaling cascade that depends on upstream Raf and Ras signaling. Interestingly, Hassanian et al. previously reported that incubation of endothelial cells with external polyP induced the phosphorylation of ERK1/2 (Hassanian et al., 2015). They also found that this same treatment resulted in activation of the mTORC1 pathway as measured by phosphorylation of p70S6K (Hassanian et al., 2015). Likewise, we found that p70S6K was phosphorylated in cells expressing Ppk1 **(Fig. 5 F and G)**. For the first time, our work demonstrates that internal production of polyP can activate conserved pathways that impinge on cell growth and development.

Relocalization of nuclear proteins in cells accumulating polyP:

Changes in protein localization can dictate a protein's function and/or activity independently of absolute protein levels (Bauer et al., 2015). Changes in protein localization are often dictated by post-translational modifications. Indeed, work in yeast has demonstrated that polyphosphorylation of RNA binding protein Nsr1, and DNA topoisomerase Top1 modulates the distribution of both proteins between the nucleoplasm and the nucleolus (Azevedo et al., 2015). To examine the impact of polyP accumulation in HEK293T cells on protein localization, we tested whether Ppk1 expression altered the localization of the known polyphosphorylated proteins NOP56, eIF5b, GTF2I, and DEK (Azevedo et al., 2018; Bentley-DeSousa et al., 2018). Upon HA-Ppk1 expression, NOP56 did not show any localization changes from control to HA-Ppk1 conditions **(Fig. 6A)**. Interestingly, eIF5b and GTF2I both showed decreases in protein amount present in the nuclear/cytoskeletal fraction when Ppk1 was expressed **(Fig. 6A and B)**. Fractionation markers for the representative biological replicate in **Fig. 6B** are shown in **Fig. 1D**. DEK, in agreement with previous studies (Hu et al., 2007; Kappes et al., 2001), was localized to the nuclear/cytoskeletal fraction in control conditions. In contrast, in HA-Ppk1-expressing cells, DEK showed a dramatic redistribution to other cellular compartments **(Fig. 6B)**. Based on these findings, we expanded our results to include a screen of 11 additional proteins, using available antibodies targeting both polyphosphorylated and non-polyphosphorylated targets **(Fig. 6 and Materials and Methods)**. Most proteins showed identical fractionation patterns in control and Ppk1-expressing cells. This

result suggests that re-distribution of proteins in cells making excess polyP is specific to individual proteins rather than a general phenomenon. However, we did find a consistent decrease in the nuclear/cytoskeletal fraction of one non-polyphosphorylation target, TAF10, following Ppk1 expression **(Fig. 6A)**. PolyP concentrations in all compartments were high enough to confer polyphosphorylation of DEK **(Fig. 6C)**. Taken together, these results show that internal accumulation of polyP alters protein localization of both targets and non-targets of polyphosphorylation in mammalian cells.

Discussion

PolyP is a molecule that plays many roles across all kingdoms of life (Kulaev and Vagabov, 1983). PolyP research in mammalian cells is hampered by a lack of knowledge concerning the enzymes involved in polyP metabolism (Desfougères et al., 2020). Many studies have relied on the addition of polyP to cell culture media in order to stimulate signaling from extracellular receptors. Here we demonstrate for the first time that internal accumulation of polyP results in changes at both the transcriptome and proteome levels. Proteomic changes include alterations in protein abundance, but also activation of conserved signaling cascades and mislocalization of a subset of nuclear proteins **(Fig. 7)**. Our work provides a novel resource for studying the impact of mammalian polyP in more depth.

We found that ectopic expression of *ppk1*+ results in the production of polyP throughout the cell. Interestingly, the pattern of localization of polyP does not mirror the localization of HA-Ppk1. For example, while most HA-Ppk1 localized to the nucleus/cytoskeleton, this compartment actually contained the lowest amount of polyP. This may reflect either a reduced ability to make polyP or variations in polyphosphatase activity within the nucleus relative to other cellular compartments. It is also possible that polyP synthesized in one cellular compartment could be transported to other organelles. The transport of polyP within cells is poorly understood and represents an important area for future research.

PolyP chains synthesized by Ppk1 in mammalian cells are of the long chain type (>130 phosphate moieties), with the longest chains found in the nuclear/cytoskeleton compartment. These are similar to the polyP synthesized by *E. coli* under stress conditions and qualitatively different from the more heterogeneous chains synthesized by the vacuolar bound VTC complex in budding yeast. A wide range of sizes of polyP chains have been reported in mammalian tissues and cells, ranging from 2 - >5000 Ps in length (Gabel and Thomas, 1971; Kumble and Kornberg, 1995; Pisoni and Lindley, 1992). Short chains (60 – 100 Ps) have been reported in human platelets (Müller et al., 2009; Ruiz et al., 2004). The longest chains have been reported in specific cell and tissue types, such as the lysosomes of human fibroblasts (100-600 Ps) (Pisoni and Lindley, 1992), the rodent brain, and in cell lines such as HEK293s where chains were found to predominantly be 800 residues long (Kumble and Kornberg, 1995). This suggests that the chains made via expression of Ppk1 are reflective of the sizes found in some mammalian cell types.

Notably, the amounts of polyP synthesized in cells transfected with ectopic *ppk1+* are much greater than polyP levels from untransfected cells. While this remains an important caveat of our study, it is possible that untransfected cells contain localized pools of highly concentrated polyP that could impact pathways similar to those uncovered here. The Ppk1 system also provides a method to model the impact of high concentrations of intracellular polyP delivered via nanoparticles, which is being investigated as a method to deliver polyP intracellularly (Fernandes-Cunha et al., 2018). Long-chain polyP produced

by Ppk1 may mimic the impact of bacterial long-chain polyP that could be released during infection and taken up by host cells.

We investigated the impact of polyP accumulation on both the human transcriptome and proteome with RNA-seq and mass spectrometry **(Fig. 7)**. This work is the first example of wide-spread changes to the cellular landscape via increased internal polyP synthesis. A previous study examined the impact of external polyP application to SaOS-2 cells using microarrays and found 23 upregulated and 11 downregulated genes that showed expression changes (p-value = 0.05, F.C$\geq$1.5) after a 3 hour incubation with polyP (65 Ps) (Lui et al., 2016). The genes from this study with greatest expression changes (*IL11* and *SLC30A1*) did not show a significant change with Ppk1 expression. However, in agreement with our work, the authors did find significant upregulation of *EGR1* mRNA and an increase in phosphorylation of ERK1/2. Finally, several genes (*ID3*, *ID1*, *MGAM*), were upregulated in the microarray experiment but downregulated in our RNA-seq. This suggests that the mode of exposure (internal vs. external) of polyP is a critical factor in the control of polyP-regulated pathways. Other experimental factors between the two studies could also contribute to the observed differences such as the length of polyP (65 Ps vs. 130 Ps), or exposure time to polyP (3 vs. 48 hours).

Our mass spectrometry analysis revealed that internal polyP accumulation resulted in 14 proteins with significant changes in levels as compared to the controls. Moreover, we uncovered >100 proteins, termed singletons, that showed

changes in their levels with polyP, but could not be included in our statistical analysis because they were only detected in 3 or 4 of our biological replicates, instead of all 5 in both conditions. One such example is DHRS2, which showed drastic changes in mRNA levels and was completely undetected in all five Ppk1 biological replicates. DHRS2 protein downregulation via shRNA was previously shown to decrease p53 phosphorylation at Ser15 (Zhou et al., 2018). We observed no significant change in phosphorylated p53 (Ser15) levels accompanying polyP-induced DHRS2 downregulation. This difference could be due to several reasons. First, our study was carried out in HEK293T cells, whereas the silencing of DHRS2 and decrease of p53 (Ser15) phosphorylation was in KYSE510 and KYSE30 cells, both from esophageal squamous cell carcinoma. Additionally, it is possible that signaling cascades impacted by polyP do not reflect “canonical” pathways. It is intriguing to consider that polyP’s effect on some pathways, such as those regulated through DHRS2, may only be apparent under conditions such as stress or infection.

Of our top candidates (FDR<15% and singletons detected in 5 out of 5 replicates in one condition), 3 proteins (RPPB1, MCM3 and DHX9) have PASK motifs, which may be hotspots for lysine polyphosphorylation (Bentley-DeSousa et al., 2018). It is intriguing to speculate that stability of these proteins may be altered by polyphosphorylation. PolyP may also act non-covalently to modulate stability via protein folding (Xie and Jakob, 2019). Finally, polyP may impact protein abundance indirectly by regulating translation or turnover via the ubiquitin proteasome.

Previous work found that external application of polyP activates the ERK1/2 signaling pathway and TOR signaling pathways by interacting with the RAGE and P2Y1 receptors on the exterior of the cell (Hassanian et al., 2015). Here we show that these same pathways can be activated by internal production of polyP **(Fig. 7)**. While we cannot completely exclude the possibility that polyP is being exported to interact with external receptors, we favor a model where polyP functions by promoting activation of upstream regulators. For example, Wang *et al,* previously showed that polyP directly interacts with mTOR to increase its activity (Wang et al., 2003). PolyP may have a similar impact on the RAS, RAF and MEK kinases that sit upstream of the ERK1/2-EGR1 signaling cascade investigated here. Indeed, previous work by the Saiardi lab identified H-RAS as a candidate polyP interacting protein (Azevedo et al., 2018). It is also possible that polyP impacts this pathway indirectly. Given our findings with the ERK1/2-EGR1 pathway, it is perhaps surprising that we did not observe any major changes in cell cycle, autophagy, or apoptosis markers. An effect on these processes may become apparent under specific conditions.

We found that polyP accumulation coincided with the relocalization of several known chromatin bound proteins, GTF2I (Makeyev et al., 2012), DEK (Alexiadis et al., 2000; Waldmann et al., 2002), and TAF10 (Soutoglou et al., 2005) from the nuclear/cytoskeleton fraction. This group includes factors previously described as targets of lysine polyphosphorylation (eIF5b, GTF2I and DEK) **(Fig. 7)**. Interestingly, Ppk1 expression resulted in a decrease in nuclear/cytoskeleton eIF5b without a corresponding increase in other fractions.

PolyP may promote the degradation of eIF5b in this fraction. For DEK, we observed a corresponding increase in protein accumulation in the cytoplasm and the mitochondrial/organelle fractions. Since polyP produced by Ppk1 is sufficient to cause polyphosphorylation of DEK in all fractions, it is tempting to speculate that polyphosphorylation might regulate DEK localization directly, perhaps by preventing nuclear import. Indeed, DEK's predicted nuclear localization signal is immediately adjacent to one of its three poly-acidic serine and lysine rich (PASK) regions that are common sites for polyphosphorylation (Bentley-DeSousa et al., 2018). Our work provides a unique resource to explore polyphosphorylation and other aspects of polyP biology in greater detail.

Materials and Methods:

Cell lines & maintenance*:* The HEK293T cell lines used in this study were obtained from the American Type Culture Collection (ATCC CRL-3216). Cells were grown at 37 °C and 5% CO2 to ~70% confluency in DMEM (Wisent, 319-015-CL) supplemented with 10% Fetal Bovine Serum (Wisent, 080-150) and 1% Penicillin-Streptomycin-Amphotericin B solution (Wisent, 450-201-EL). For the RNA-seq, ERK1/2, and p70SK6 experiments cells were grown with the addition of 1 mM sodium pyruvate (Sigma, S8636) in the media. All cell lines used were regularly tested for mycoplasma contamination using a PCR Mycoplasma Detection Kit (Applied Biological Materials, G238).

Yeast and bacterial strains: Wild-type and *vtc4*Δ::NatMX *S. cerevisiae* strains containing empty vectors allowing growth in SC-Uracil minimal media are of the BY4741 S288C background with the genotype *MAT*a *leu2*Δ*0 ura3*Δ*0 met15*Δ*0 his3-1*. Strains were grown in SC-URA media at 30 °C to maintain the plasmids. For bacterial polyP extraction, WT MG1655 *E. coli* of genotype *F- lambda- ilvG- rfb-50 rph-1* was grown as described below for 'polyphosphate analysis'.

Plasmids: Yeast strains contain the plasmid p416 (ATCC) to allow growth in synthetic media lacking uracil. Mammalian *ppk1+* expression vector contains the *E. coli ppk1+* sequence cloned into a version of pcDNA3.1(-) containing a bleomycin resistance marker. This plasmid has been described previously (Bentley-DeSousa et al., 2018) and is available from Addgene (Plasmid #108850). Mammalian HA*ppk1+* expression vector contains the *E. coli ppk1+*

sequence including a HA-tag on the N-terminal, cloned into a version of pcDNA3.1(-) containing a bleomycin resistance marker. This plasmid will be available on Addgene.

Polyphosphate analysis:

The polyP isolation protocol was adapted from (Bru et al., 2017). HEK293T cell pellets were resuspended in 400 µL of LETS buffer (100 mM LiCl, 10 mM EDTA, 10 mM Tris-HCl pH 7.4, 0.2% SDS) kept at 4 °C. Samples were mixed with 600 µL of phenol pH=8 and 150 µL of water in a screw cap tube, and vortexed for 20 seconds. They were then heated for 5 minutes at 65 °C then cooled down on ice. 600 µL of chloroform were added followed by 20 seconds of vortexing. Samples were then centrifuged at room-temperature for 2 minutes at 13 000 x g and the top layer/water phase was transferred to a new screw cap tube containing 600 µL of chloroform. Samples were vortexed and centrifuged again for 2 minutes at 13 000 x g, and the top layer was transferred to a clean 1.5 mL tube. 2 µL of 10 mg/mL RNAse A and 2 µL of 10 mg/mL DNAse I were added followed by 1 hour incubation at 37 °C. Samples were transferred to pre-cold 1.5 mL tubes containing 1 mL of 100% ethanol and 40 µL of 3 M sodium acetate pH=5.3. polyP was left to precipitate at -20 °C for at least 3 hours. Samples were centrifuged for 20 min at 13 000 x g at 4 °C and the supernatant was discarded. PolyP pellets were washed with 500 µL of 70% ethanol and centrifuged for an additional 5 minutes. The supernatant was discarded and the pellets air dried for a few minutes before being resuspended in 20 µL of water.

Yeast polyP preps were from cells diluted at OD600 = 0.2 in YPD and grown at 30 °C until the OD reached 0.8-1. The equivalent of 6 OD600 units were used in the prep and the final polyP pellet was dissolved in 30 µL of water. For *E. coli* polyP preps, overnight cultures were diluted to OD600 = 0.1 in 20 mL of LB media in the absence of antibiotics and grown at 37 °C. Cells were grown to exponential phase (OD600 = 0.4-06) prior to harvesting (control) or switching to MOPs minimal media to induce starvation. OD600 units were harvested for control samples and these cells were washed with cold water, pelleted and frozen on dry ice. For starvation samples, 6 OD equivalent of cells were washed once with 1 X PBS [100 mL of 10 X 1L Stock (54.36 g NaH_2PO_4 (Sigma-Aldrich S9638-1KG), 160 g NaCl (Fisher Scientific, AM9759), 4 g KCl (Sigma-Aldrich 746436-500G), 9.6 g KH_2PO_4 (Sigma-Aldrich P0662-500G), fill to 1 L ddH2O), and 900 mL ddH2O], resuspended in 10 mL of 1 X MOPS (Sigma-Aldrich M1254-1KG) media and incubated at 37 °C for and additional 2 hours. Starved cells were harvested similar to control cells by centrifugation, washing with cold water, and freezing on dry ice. Cells were then processed as described for HEK293T cells.

Extracted polyP with loading dye (10 mM Tris-HCl pH 7, 1 mM EDTA, 30% glycerol, bromophenol blue) was run on a 15.8% acrylamide TBE-Urea gel (5.8 M Urea - 5.25 g Urea, 7.9 mL 30% acrylamide, 3 mL 5 X TBE, 150 µL 10% APS, 15 µL TEMED) at 100 V for 1 hour and 45 minutes in 1 X TBE buffer.

Negative DAPI staining protocol was adapted from (Smith and Morrissey, 2007). Briefly, gel was incubated for 30 minutes in fixative solution (25% methanol, 5% glycerol, 50 mM Tris base - pH 10.5) containing 2 µg/mL of DAPI at room temperature. Following this incubation, the gel was destained for 1 hour in fixative solution with a replacement of fresh fixative solution after 30 minutes. Gel was exposed on a 365 nm UV transilluminator to photobleach DAPI bound to PolyP. This step required adjustable times (average several minutes) based on the amount of polyP. Finally, the gel was imaged using a Chemi Doc at 312 nm.

Transfection and lysate harvest:

HEK293T cells were grown to ~70% confluency and transiently transfected with 2-28 µg of either control plasmid (pcDNA3.1-) or *ppk1*+ expression plasmids (pcDNA3.1-*ppk1*+), (pcDNA3.1-HA*ppk1*+) using Lipofectamine 2000 (Thermo Fisher Scientific, 11668019) or Lipofectamine LTX Reagent (Thermo Fisher Scientific, 15338100). Transfected cells were grown for 48 hours in OptiMeM (Fisher Scientific, 31985-070) and then harvested.

Protein:

To isolate total protein, cells were washed with ice-cold 1 X PBS, scraped, and lysed in ~200 µL (3.5 cm plate) or ~1 mL (10 cm plate) of ice-cold radioimmunoprecipitation assay buffer (RIPA) [50 mL ice cold lysis buffer [10 mM Tris-HCl pH 7.4, Hydrochloric acid (Fisher Scientific, A144 212), 100 mM NaCl (Fisher Scientific, AM9759), 1 mM EDTA (Sigma-Aldrich, 03690-100ml), 1%

IGEPAL (Sigma-Aldrich, I8896-50ML), 0.5% Sodium Deoxycholate (Sigma-Aldrich, D6750-25G), 0.1% SDS (Thermo Fisher Scientific, 15525-017)], 500 µL Sodium Fluoride (0.5 M, Sigma-Aldrich, 201154-5G), 500 µL Glycerol-2-Phosphate (1 M), 500 µL Nicotinamide (1 M Sigma-Aldrich, N3376-100G), 500 µL PMSF (1 mg/mL Sigma-Aldrich, P7626-1G), 500 µL Sodium Butyrate (1 M Sigma-Aldrich, 303410-100G), and 1 cOmplete protease inhibitor tablet (Roche, 04693159001)]. Following a 15 minute incubation on ice, the cells were centrifuged for 15 minutes at 13,500 RPM. Supernatant was collected and either used immediately for analysis with immunoblotting or stored at -20°C.

To isolate the cell population for cell fractionation experiments, cells were washed with ice-cold 1 X PBS, scraped, and centrifuged for 5 minutes at 4°C at 350 RCF. Cell pellet was then resuspended in 1 X PBS and cell fractionation was carried out using the Cell Fractionation Kit (Cell Signaling Technology, 9038) according to manufacturer's protocol.

RNA:

To isolate total RNA, cells were washed with ice-cold 1 X PBS and scraped in 1 mL of ice-cold TRIzol (Thermo Fisher Scientific, 15596026). Cells were then briefly vortexed and incubated at room temperature for 5 minutes. Following incubation, 200 µL of chloroform (Sigma-Aldrich, 472476-500ml) was added and cells were vortexed for 30 seconds and incubated for 3 minutes at room temperature. Cells were then centrifuged for 15 minutes at 4°C at 12,000 RCF. The aqueous phase was transferred to a new tube and 500 µL of ice-cold

isopropanol (Sigma-Aldrich, 278475-1L) and 2 µL of Glycoblue (Thermo Fisher Scientific, AM9515) was added. Mixture was incubated for 10 minutes at room temperature and then centrifuged for 15 minutes at 4°C at 12,000 RCF. Isopropanol was carefully removed and pellet was resuspended in 500 µL of ice-cold ethanol (Commercial Alcohols, P016EAAN). Following a 10 minute centrifugation at 4°C at 7, 600 RCF the pellet was air-dried for 5 minutes at room temperature. Dried pellet was resuspended in 25 µL of RNase-free water. All RNA work was carried out with filter pipette tips.

Western blotting:

Following protein harvest and preparation (cell fractionation) samples were analyzed by western blotting using SDS-PAGE and NuPAGE. 5 X Laemmli Sample Buffer was added to the protein samples prior to boiling at 95°C for 2 minutes and briefly centrifuged. Fraction samples used for cell fractionation experiments were boiled for 5 minutes and centrifuged for 3 minutes at 15,000 x RCF, as recommended according to manufacturer's protocol. All antibodies and dilutions used in this study are described in **Supplemental Table 1**. Please note that SDS-PAGE gels do not resolve polyP-induced protein shifts (Azevedo et al., 2015; Bentley-DeSousa et al., 2018).

SDS-PAGE:

Samples were separated using 10-15% SDS-PAGE gels or 4–20% Criterion TGX Stain-Free Protein Gel (Bio-Rad, 5678095). Gels were run at 100 V using 1 X

SDS-PAGE Running Buffer [100 mL of 10 X 1 L Stock [30.2 g Tris Base (Fisher Scientific, BP152-5), 188 g Glycine (Sigma-Aldrich, G7126-500G), 10 g SDS (Thermo Fisher Scientific, 15525-017)] and 900 mL ddH2O] and transferred to Polyvinylidene fluoride membranes (Bio-Rad, 162-0177) at 85 V for 1 hour in 1 X SDS-PAGE Transfer Buffer [100 mL of 10 X 1L Stock [30.275 g Tris Base (Fisher Scientific, BP152-5), 166.175 Glycine (Sigma-Aldrich, G7126-500G), 200 mL Methanol (Fisher Scientific, A412P-4)] and 700 mL ddH2O.

NuPAGE:

Samples were separated using 4-12% Bis-Tris NuPAGE gels (Thermo Fisher Scientific, NP0336BOX) at 200 V using 1 X NuPAGE Running Buffer [50 mL of 20 X 1 L Stock [209.2 g MOPS (Sigma-Aldrich, M1254-1KG), 121.1 g Bis-Tris (Sigma-Aldrich, B9754-1KG), 20 g SDS (Thermo Fisher Scientific, 15525-017), 12 g EDTA (Sigma-Aldrich, 03690-100ml)] and 950 mL ddH2O]. The gels were then transferred to Polyvinylidene fluoride membranes at 85V for 1 hour using 1 X NuPAGE Transfer Buffer [50 mL of 20 X 1 L Stock [81.6 g Bicine (Sigma-Aldrich, B3876-250G), 104.8 g Bis-Tris (Sigma-Aldrich, B9754-1KG), 6 gEDTA (Sigma-Aldrich, 03690-100ml)], 200 mL Methanol (Fisher Scientific, A412P-4) and 750 mL ddH2O].

Immunoblotting:

In general, membranes were blocked for 30 minutes with either 5% milk or BSA (VWR, 97061-416) diluted in 1 X TBST [100 mL of 10 X 1L Stock (90 g Tris Base

(Fisher Scientific, BP152-5), 88 g NaCl (Fisher Scientific, AM9759), 2 g KCl (Sigma-Aldrich 746436-500G), adjust pH with HCl to 7.5, fill to 1 L ddH2O), 10 mL of 10% Tween 20 (Fisher Scientific BP337500) and 850 mL ddH2O] under gentle agitation at room temperature. Primary antibodies were diluted either with 5% milk or BSA diluted in 1 X TBST and incubated with membrane at 4°C overnight or 1 hour at room temperature. Following use, primary antibodies were stored at 4°C with 1% sodium azide. After primary antibody incubation, membranes were washed with 1 X TBST for 10 minutes at room temperature 3 times followed by a 30 minute incubation with horseradish peroxidase-conjugated secondary antibodies. Following 3 more 10 minute washes with 1 X TBST at room temperature, membranes were incubated with Chemiluminescence HRP Substrate, Luminata Forte (Fisher Scientific, WBLUF0500) and exposed to autoradiography film (Harvard Apparatus Canada, DV-E3018). Of note, Vimentin blots were incubated with fluorescent IR Dye 800CW secondary antibodies for 1 hour, washed in 1 X TBST 3 times (10 minutes), and visualized using a LI-COR Odyssey imaging system.

Protein screen:

Proteins that were screened (11 proteins) for fractionation experiments were Hsp90, Nucleolin, CDC5L, Cyclin A, Cyclin B1, H3, KAT2A/Gcn5, Lamin A/C, p53, RALY, TAF10. All antibodies and dilutions are listed in **Supplemental Table 1**.

RT-qPCR setup and analysis:

cDNA synthesis was carried out with isolated total RNA using 5 X All-In-One RT MasterMix kit (Applied Biological Materials, G490) according to manufacturer's protocol. Reverse transcription conditions: 25°C for 10 minutes, 42°C for 50 minutes, and 85°C for 5 minutes. Newly synthesized cDNA was used immediately or stored at -20 °C. Primer design for qPCR primers was completed using NCBI Primer-BLAST web-based program with standard parameters (product size between 80-200 bps, T_m=~60°C (±3°C), GC content ≥50%). qPCR reactions were carried out using SYBR Green Supermix (Bio-Rad, 1708880) according to manufacturer's protocol with the following conditions: 95°C 2 minutes, (95°C 30 sec, 60°C 30 sec, 72°C 45 sec)×40 cycles, 72°C 10 minutes). Standard curves were completed for every qPCR experiment and to confirm accurate primer pair efficiency for each primer pair used in this study **(Supplemental Table 3)**. All qPCR reactions were completed in 3 technical replicates and included ≥3 biological replicates and normalized to a control gene. The changes in mRNA expression is represented as fold change and was calculated using the ΔΔCt method (Schmittgen and Livak, 2008).

RNA-sequencing (RNA-seq):

Preparation of RNA-seq samples:

RNA-seq was carried out on HEK293T cells transiently transfected in either CTRL or Ppk1-expressing conditions. Three biological replicates were used for

this experiment. Following a 48 hour transfection, RNA for RNA-seq was extracted using an miRNeasy Mini Kit (Qiagen, 217004). Protein was also isolated from each sample (detailed in “Transfection and lysate harvest”) and analyzed on NuPAGE gels to confirm detection of shifts on target proteins DEK and MESD **(Fig. 2)**. RNA pellets were stored at -80 °C and shipped on dry ice to Genome Quebec Innovation Centre for RNA sequencing.

The RNA-seq was carried out at Genome Quebec as follows. Total RNA was quantified with a NanoDrop Spectrophotometer ND-1000 (NanoDrop Technologies, Inc.). RNA integrity was assessed on a 2100 Bioanalyzer (Agilent Technologies). To generate libraries from 250 ng of total RNA, the following steps were taken. First, mRNA enrichment was performed using the NEBNext Poly(A) Magnetic Isolation Module (New England BioLabs followed by cDNA synthesis using the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). Next, library preparation was completed using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). The adapters and PCR primers used were purchased from New England BioLabs. Quantification of libraries was carried out using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems) with the average size fragment determined with a LabChip GX (PerkinElmer) instrument. Libraries were first normalized, then denatured in 0.05 N NaOH followed by dilution to 200 pM and finally neutralized using HT1 buffer. Clustering was carried out on an Illumina cBot. The flowcell was run on a HiSeq

4000 for 2x100 cycles in paired-end mode following the manufacturer's instructions. Used as a control was a phiX library mixed with libraries at 1% level. The Illumina control software used was HCS HD 3.4.0.38 and RTA v. 2.7.7 was the program used for real-time analysis. Finally, the bcl2fastq v2.20 program was used to demultiplex samples and generate fastq reads.

Analysis of RNA-seq was completed by Canadian Center for Computational Genomics (C3G) as follows. Paired-ends sequencing reads were first clipped for adapter sequence and then trimmed for minimum quality (Q30) in 3'. Following this, sequencing reads were filtered for minimum length of 32 bp using Trimmomatic (Bolger et al., 2014). The surviving read pairs were then aligned to the Genome Reference Consortium Human Build 38 using STAR (Dobin et al., 2013). This alignment used the two-passes method. Gene quantification (gene-level count-based) against Ensembl annotations was carried out using HT-seq count (Anders et al., 2015) (mode used: intersection-nonempty). Following gene quantification, exploratory analysis was then conducted with various functions and packages from R and the Bioconductor project (Huber et al., 2015). To complete differential expression both edgeR (Robinson et al., 2010) and DESeq (Anders and Huber, 2010) were used. Significance was determined by DESeq adjusted p-values < 0.05. Gene Ontology terms were tested for enrichment with the GOseq (Young et al., 2010) R package (FDR threshold of 0.05). As expected, *Ecppk1* sequences were detected in Ppk1 expressing samples, but not in the CTRL samples. Each processing step

described above was accomplished through the GenPipes framework (Bourgey et al., 2019).

Liquid Chromatography – Mass Spectrometry:

Preparation of mass spectrometry samples:

Mass spectrometry was carried out on HEK293T cells transiently transfected in either CTRL or *ppk1*+ plasmid expressing conditions (five biological replicates for each condition). Following a 48 hour transfection (detailed in "Transfection and lysate harvest"), proteins were harvested, pelleted, and frozen at -80 °C. A small amount of protein was used immediately for analysis on NuPAGE gels to confirm detection of shifts on target proteins DEK and MESD. Protein pellets were shipped overnight on dry ice to the UC Davis Genome Center Proteomics Core Facility. Extracts were prepared via sonication in 5% SDS, 50mM TEAB, supplemented with protease and phosphatase inhibitor tablets (Roche). Extracts were clarified via centrifugation at 15,000xg for 10min @ 4 °C prior to trypsin digestion using S-Trap Mini spin columns (PROTIFI) as per manufacturer's instructions.

Liquid chromatography and mass spectrometry analysis:

LC separation was carried out on a Proxeon Easy1200 HPLC (Thermo Fisher Scientific) with a Proxeon nanospray source (operating in positive ionization mode). After LC separation, digested peptides were then reconstituted in 2% acetonitrile/0.1% trifluoroacetic acid. 1 µg of each sample was then loaded

onto a 100 µm×25 mm Magic C18, 100 Å 5 µm reverse phase trap wherein samples were desalted online, and directly eluted onto an analytical column (maintained at 40º C) for on-line separation. The analytical column used was a 75µm×150 mmMagic C18 200 Å 3µm reverse phase column. Peptides were then eluted into the mass spectrometer using a 2-buffer gradient (Buffer A: 0.1% formic acid in water and Buffer B: 0.1% formic acid in 100% acetonitrile) at a flow rate of 300 nL/min.

A 144 minute gradient was run as follows: 2 to 25% Buffer B (90 minutes), 25% to 40% Buffer B (26 minutes), 40 to 100% Buffer B (4 minutes), held at 100% Buffer B (3 minutes). Finally, ending with 2% Buffer B (21 minutes). The analytical column used tapers off into the electrospray emitter. The peptides were then ionized (spray voltage 2.0 kV). An Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) was used to collect the mass spectra. Mass spectra was collected in a data-dependent mode with the following: MS survey scan (m/z range 350–1,600), 70,000 resolution, target of 1 × 106 ions or a maximum injection time of 30 ms, 15 MS/MS scans (where the top 15 ions in the MS spectra were subjected to high-energy collisional dissociation, and normalised fragmentation energy of 27%. The MS/MS spectra were obtained at a resolution of 17,500 and a target of 5 × 104 ions or a maximum injection time of 50 ms. The isolation mass window used for precursor ion selection was 1.6 m/z, with accepted charge states 2-4, and a normalized collision energy of 27% used for fragmentation. Finally, precursor ion dynamic exclusion was set to 20 seconds.

Protein identification:

LC-MS/MS RAW data files were converted to mzML files using MSConvert from ProteoWizard (Adusumilli and Mallick, 2017). mzML files were then searched using X! Tandem (Bjornson et al., 2008) against the UniProt human protein sequence database (downloaded March 4, 2019, from uniprot.org;) (Consortium, 2015) using a target-decoy strategy. The database contained 73,948 protein sequences amended with 49 potential contaminants from the cRAP database of common laboratory contaminants (www.thegpm.org/crap) and an equal number of reversed decoy sequences. The database search was performed using trypsin as the digestive enzyme and allowing for one missed cleavage, as well as carbamidomethylation of Cys as a fixed modification. Up to five variable modifications were allowed per peptide, including oxidation of Met and Trp, as well as n-terminal acetylation. Instrument parameters and match tolerances were set to Orbitrap defaults. A protein identification required at least one non-redundant peptide (unique) identification. Protein identifications were grouped and clustered using Scaffold (v4.1, Proteome Software). Protein and peptide identification confidence threshold was set to 1% FDR and were matched between runs and second peptides enabled.

Protein differential expression analysis:

Protein and protein cluster spectral counts from Scaffold were normalized based on the total number of spectral counts in each replicate. Protein differential expression was statistically assessed using a two-tailed, two-sample Student's *t*-

test assuming unequal variance, which was performed on normalized spectral counts of proteins or protein clusters detected in all 5 replicates in each condition. This criterium was selected to maximize the accuracy of the *t*-test statistical assessment. An FDR was estimated at different *p*-value thresholds **(Supplementary Table 4)**. On the other hand, singleton proteins are defined as proteins that were identified in at least 3 replicates in one condition and 0 in another (All or None (5 replicates/5 replicates vs. 0 replicates/5 replicates) – Mostly All or None (3 or 4)/5 vs. 0/5) or proteins that were identified in all 5 replicates in one condition but in 2 or less in the other condition (All or Mostly None (5/5 vs. (1 or 2)/5) **(Supplementary Table 4)**.

Gene Ontology enrichment analysis:

Differentially expressed proteins (FDR <15%) and singleton proteins were investigated for Gene Ontology enrichments using Ontologizer (Bauer et al., 2008). All quantified proteins were used as background.

Protein-protein interaction network construction with STRING:

The set of differentially expressed proteins (FDR <15%) and singleton proteins was used to query their protein-protein interactions from the STRING database (Szklarczyk et al., 2019). Protein-protein interactions derived from experiments and databases with a medium confidence level were used to build the network. All proteins without any protein-protein interactions with a protein in the query set

were not included in the network. Gene Ontology term overlay on the network was performed using STRING.

Data availability:

RNA-seq data has been deposited into the GEO database under the identifier GSE149011.

Mass spectrometry data have been deposited to the ProteomeXchange Consortium (Deutsch et al., 2016) via the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifiers PXD018896 and 10.6019/PXD018896. Reviewers can access the data on PRIDE using the following credentials: username: reviewer22839@ebi.ac.uk and password: GcJJmmdb.

Acknowledgments:

We thank the Downey lab members for critical reading of the manuscript. We thank T. Shiba (Regentiss, Japan) for the generous gift of polyP standards. This study was funded by a Canadian Institutes of Health Research Project Grant to MD (PJT-148722). E. B.-C. was supported in part by a postdoctoral fellowship from The Natural Sciences and Engineering Research Council of Canada (NSERC). I.A. was supported in part by a NSERC Undergraduate Student Research Award. M.L.A. holds a NSERC Discovery Grant.

The Canadian Center for Computational Genomics (C3G) is a Genomics Technology Platform (GTP) supported by the Canadian Government through Genome Canada.

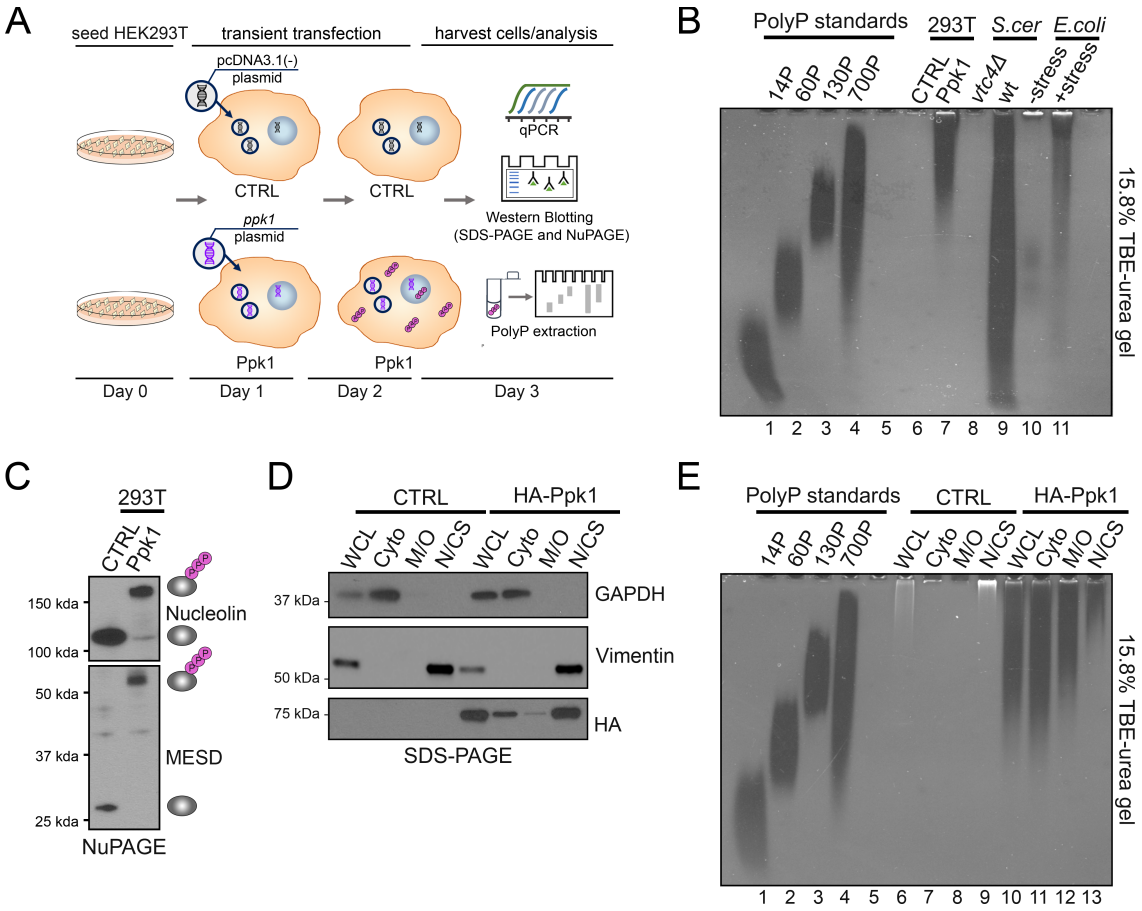


Figure 1. Characteristics of cells that accumulate internal polyP. *(A)* HEK293T cells were seeded on day zero and on day 1 were transfected with either empty vector (pcDNA3.1(-)) or *ppk1*+ expression plasmid (pcDNA3.1(-)*ppk1*+) to generate control or Ppk1 conditions, respectively. After 48 hours cell lysates were harvested and analyzed with follow up experiments including, but not limited to qPCR, western blotting (SDS-PAGE and NuPAGE), and polyP extraction. *(B)* PolyP extractions analyzed on a 15.8% TBE-urea gel stained with DAPI from control and Ppk1-expressing HEK293T cells, wt and *vtc4Δ* yeast, and *E. coli* in unstressed or stressed conditions (amino acid starvation via growth in MOPS minimal media). PolyP standards [14P, 60P, 130P (Regentiss, Japan)

and 700P (Kerafast)] are presented for comparison. Image shows results from one biological replicate, which is representative of N=3. *(C)* NuPAGE analysis of known polyphosphorylated proteins. Total protein was extracted using RIPA buffer prior to NuPAGE analysis via western blotting (same blot) with indicated antibodies. *(D)* 48 hours after transfection HEK293T cells were fractionated using Cell Signaling Technologies Cell Fractionation Kit according to manufactures protocol. SDS-PAGE analysis (4–20% Criterion TGX Stain-Free Protein Gel) of whole cell lysate (WCL) and cell fractions from control and Ppk1 conditions with antibodies known to be localized to specific fractions: Cytoplasm (Cyto) – GAPDH and Nuclear/Cytoskeleton (N/CS) – Vimentin. Image is representative of N≥3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(E)* PolyP extractions analyzed on a 15.8% TBE-urea gel from cell fractionations of control and Ppk1 conditions HEK293T cells. PolyP was extracted directly from cell fraction aliquots either immediately following fractionation or from fractions stored at -80 °C.

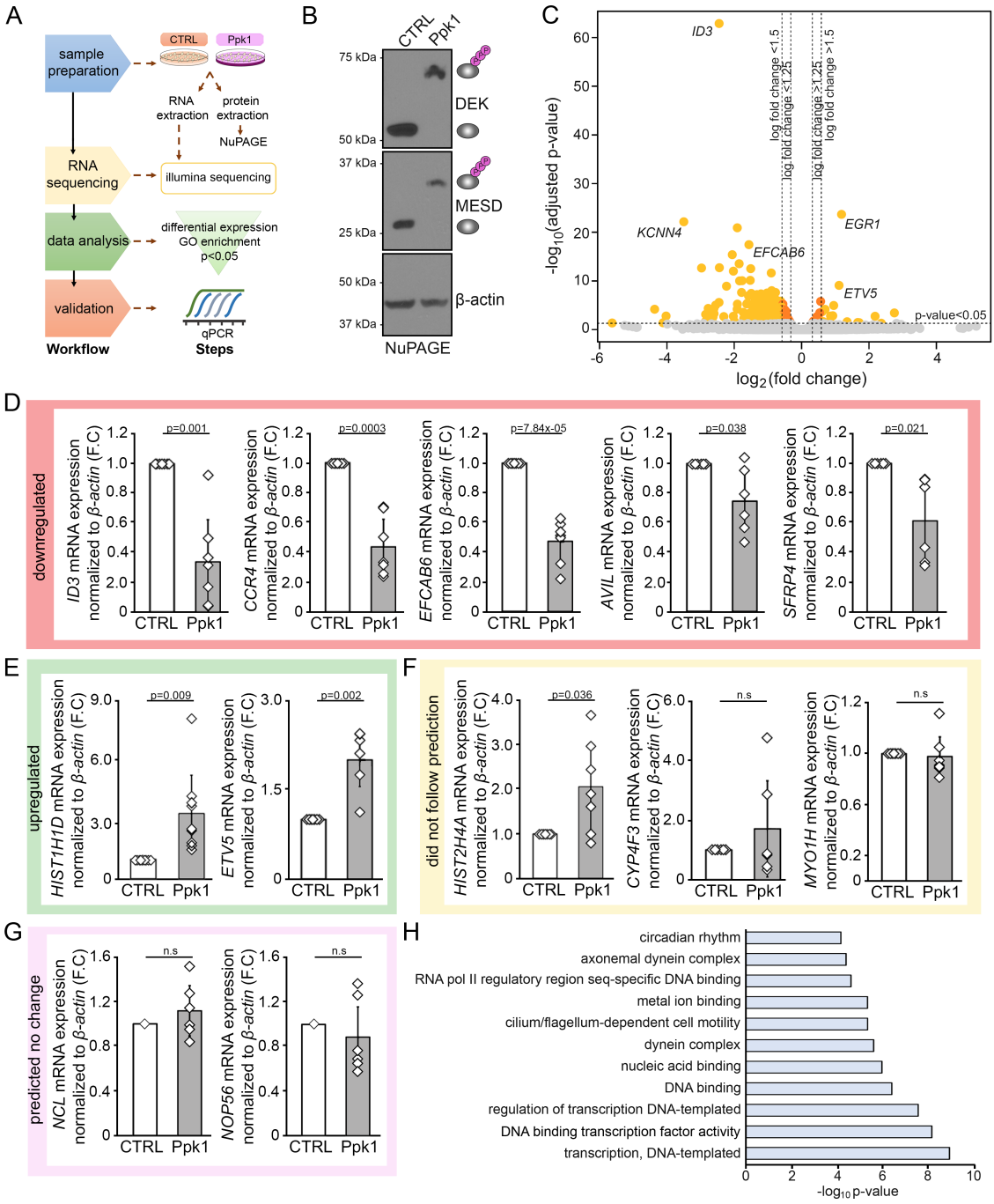


Figure 2. RNA-seq analysis of cells that accumulate polyP. *(A)* Workflow of RNA-seq analysis. 48 hours after transfection with either control or *ppk1+* expression plasmid, RNA and protein were harvested from cells. RNA was immediately shipped to Genome Quebec for RNA-seq. RNA-seq data was

analyzed and a small subset of genes were validated by RT-qPCR (details in materials and methods). *(B)* NuPAGE analysis showing polyphosphorylation induced shifts of DEK and MESD in lysates harvested from cells sent for RNA-seq. β-actin, a non-polyphosphorylated protein, was used as a loading control. Image is representative of N=3. Representative image shown here is from the same biological sample from different blots analyzed in parallel (MESD and β-actin were analyzed on one blot separate from DEK). *(C)* Volcano plot shows $\log_2$(fold change) and $\log_{10}$(FDR-adjusted p-value) for all genes obtained from the RNA-seq experiment. Genes were classified as significant if their FDR-adjusted p-value<0.05. Statistical analysis of RNA-seq data is included in detail in the materials and methods. *(D,E,F,G)* RT-qPCR validation of top hits from RNA-seq differential expression analysis. Changes in mRNA levels are represented by fold change (F.C). *(D,E)* Hits that were found to show expression changes in the same direction as predicted by the RNA-seq data are shown inside green (up) or pink (down) boxes. *(F)* Hits that did not show expression changes that followed predicted patterns from RNA-seq data are inside a yellow box. *(G)* Graphs inside the light pink box are hits that were predicted and found by RT-qPCR not to show significant expression changes from control to Ppk1 conditions. Primers used for RNA-seq validation are included in **Supplemental Table 3**. Graphs represent N≥4 for each mRNA. P-values are shown as numerical values or if p≥0.05 as non-significant (n.s). Statistical tests performed were one-sample t-Test (unequal variances) where error bars represent standard

deviation. *(H)* GO-term analysis of differentially expressed genes (FDR-adjusted *p*-value<0.05). GO-term analysis is described in materials and methods.

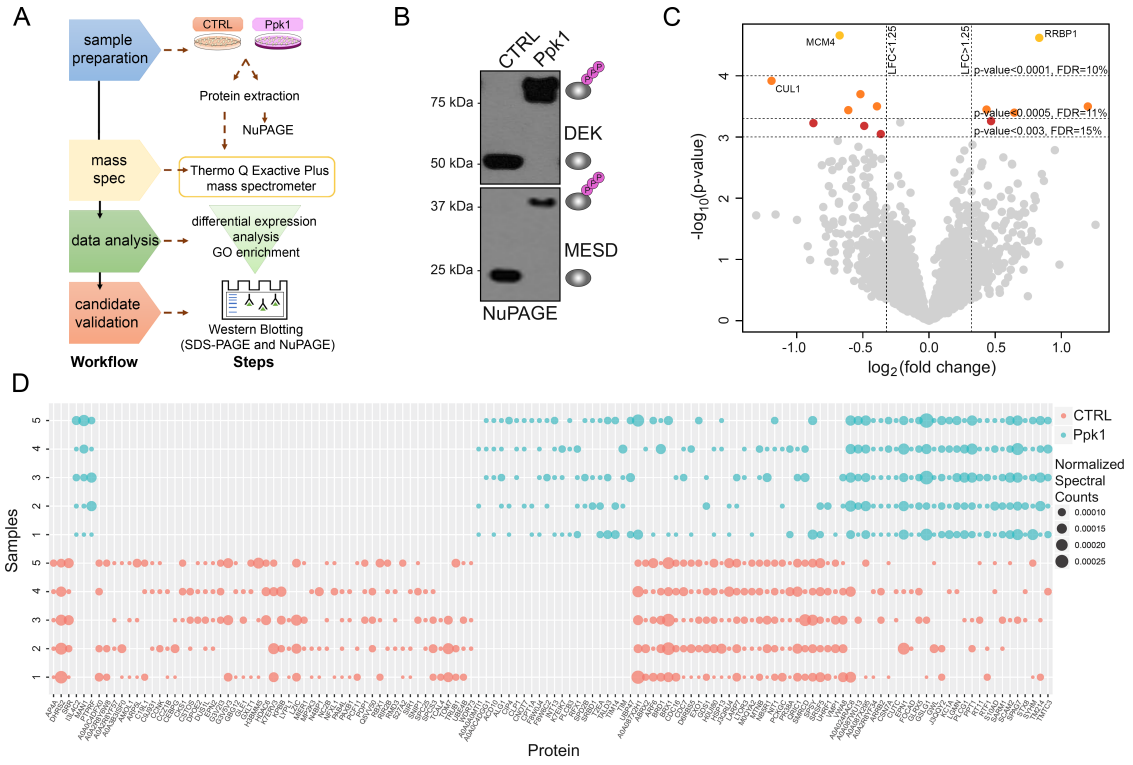


Figure 3. Global changes in protein levels in cells that accumulate polyP. *(A)* Workflow of mass spectrometry analysis. 48 hours after transfection with either control or *ppk1*+ expression plasmid, protein was harvested from cells and a small portion of sample was used to confirm polyphosphorylation of targets by NuPAGE in all samples. The remaining portion of each sample was immediate shipped to UC Davis for mass spectrometry (details included in materials and methods). *(B)* NuPAGE analysis showing polyphosphorylation induced shifts of DEK and MESD in lysates harvested from cells sent for MS. Image is representative of N=5. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(C)* Volcano plot shows $\log_2$(fold change) and $-\log_{10}$(p-value) for all proteins identified in all 5 replicates of each condition. Statistical analysis of mass spectrometry data is included in detail

in the materials and methods. *(D)* “Singletons” from mass spectrometry data. Categories shown in both control (pink) and Ppk1 (blue) conditions include “All or None” (proteins identified in 5/5 in one condition and 0/5 in the other), 2) “Mostly All or None” (proteins identified in (3 or 4)/5 in one condition and 0/5 in the other), and 3) “All or Mostly None” (proteins identified in 5/5 in one condition and (1 or 2)/5 in the other). Size of bubbles reflect normalized spectral counts, a proxy of the relative abundance of the protein within each biological replicate (samples 1-5), for each condition.

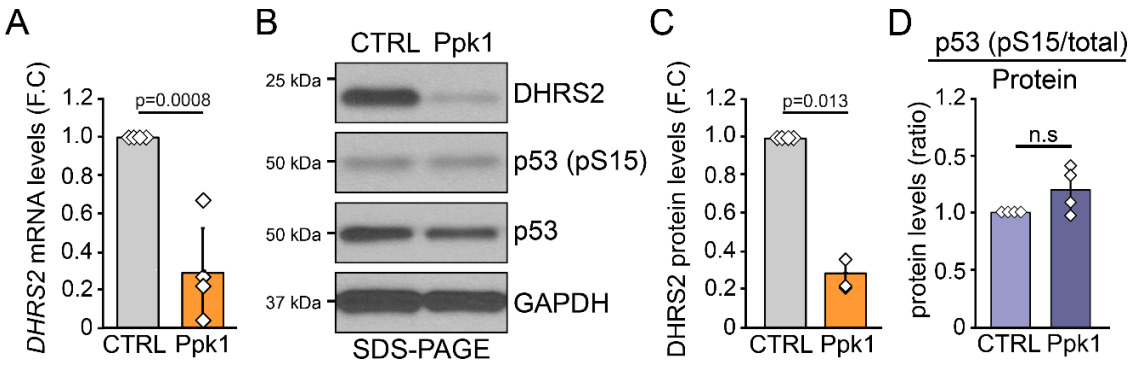


Figure 4. Accumulation of polyP leads to upregulation of DHRS2. RNA and protein were harvested from HEK293T cells following a 48-hour transient transfection of either a control or *ppk1+* expression plasmid. *(A)* RT-qPCR analysis of *DHRS2* mRNA in control and Ppk1 conditions. Changes in mRNA levels are represented by fold change (F.C). *(B)* Western blot analysis with SDS-PAGE (4–20% Criterion TGX Stain-Free Protein Gel) of proteins in control and Ppk1 conditions using antibodies against DHRS2, p53, and phosphorylated p53 (Ser-15). GAPDH was used as a loading control. Image is representative of N$\geq$3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(C,D)* Semi-quantitative analysis of protein levels shown in panel B. P-values are shown as numerical values or if p$\geq$0.05 as non-significant (n.s). Statistical tests performed were one-sample t-Test (unequal variances) where error bars represent standard deviation. All experiments shown here are representative of N$\geq$3.

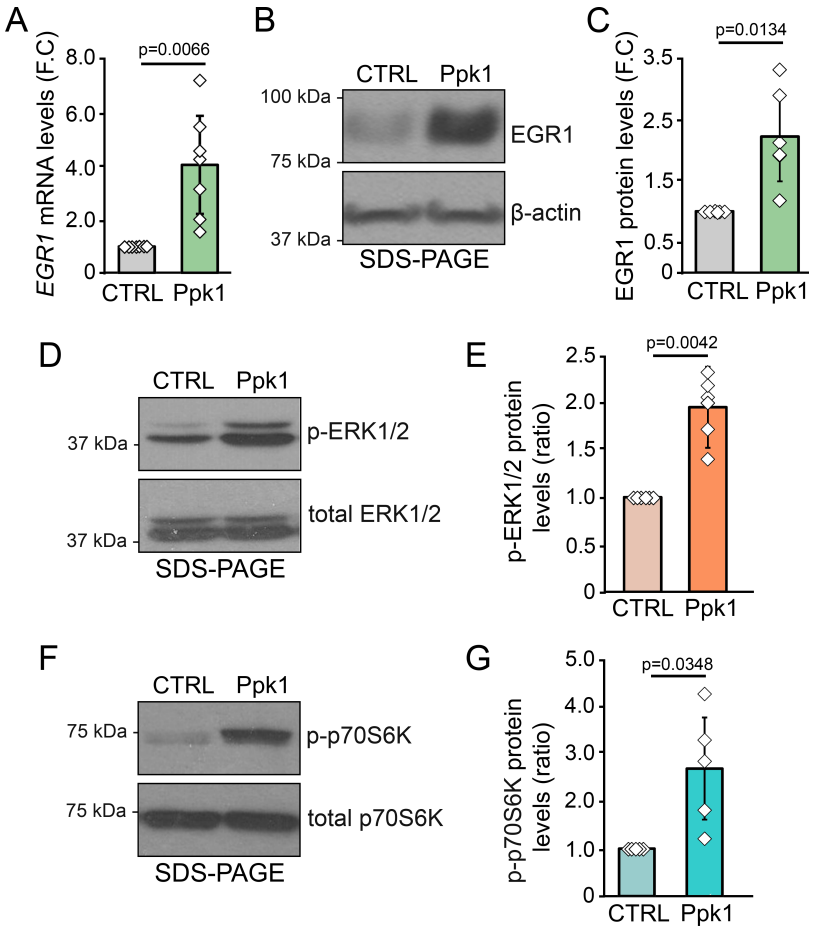


Figure 5. Activation of ERK1/2 signaling by production of polyP. RNA and protein were harvested from HEK293T cells following a 48-hour transient transfection of either a control or *ppk1+* expression plasmid. *(A)* RT-qPCR of *EGR1* mRNA levels in control and Ppk1 conditions. Changes in mRNA levels are represented by fold change (F.C). *(B)* western blotting analysis by SDS-PAGE (4–20% Criterion TGX Stain-Free Protein Gel) in control and Ppk1 conditions cells using the EGR1 antibody with β-actin as a loading control. Image is representative of N≥3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(C)* Semi-quantitative analysis of protein levels shown in panel B. *(D)* Western blotting analysis by SDS-PAGE in control and

Ppk1 conditions cells using antibodies against total ERK1/2 and phosphorylated ERK1/2. Image is representative of N≥3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(E)* Semi-quantitative analysis of protein levels shown in panel D. *(F)* Western blotting analysis by SDS-PAGE in control and Ppk1 conditions cells using antibodies against total p70S6K and phosphorylated p70SK6. Image is representative of N≥3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. *(G)* Semi-quantitative analysis of protein levels shown in panel G. P-values are shown as numerical values or if p≥0.05 as non-significant (n.s). Statistical tests performed were one-sample t-Test (unequal variances), where error bars represent standard deviation.

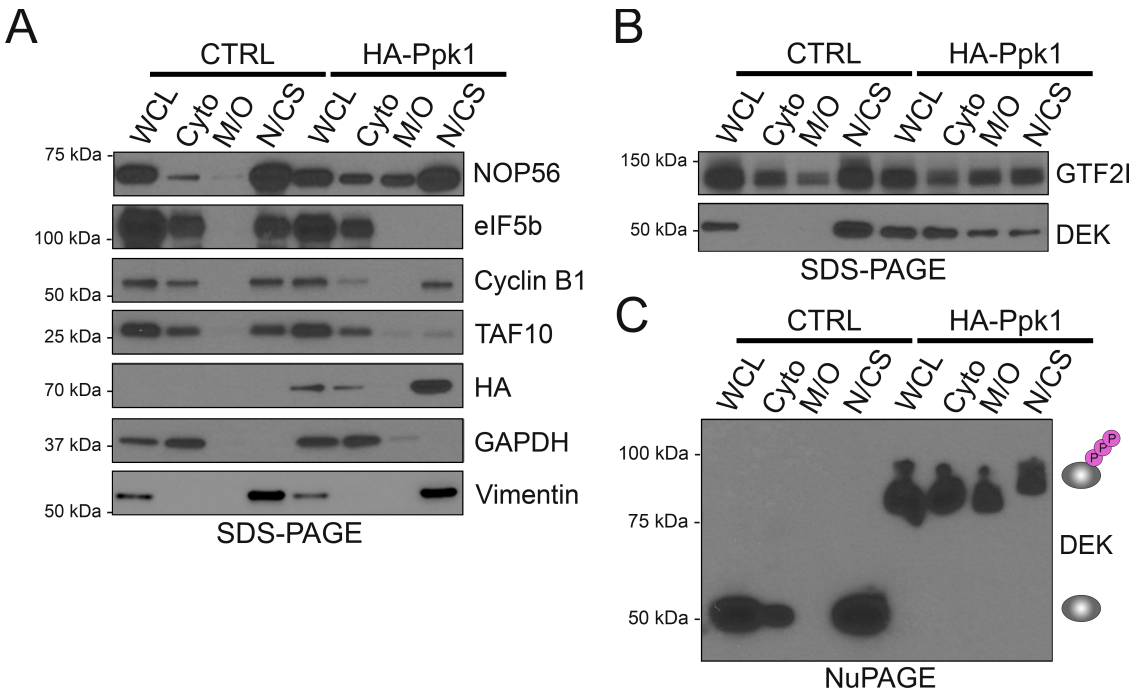


Figure 6. Changes in protein localization in both polyphosphorylated and non-polyphosphorylated proteins with polyP accumulation. HEK293T cells were transiently transfected for 48 hours with either a control or *ppk1+* expression plasmid. *(A, B)* Fractionation of cells was performed using Cell Signaling Technologies Cell Fractionation Kit according to manufactures protocol. SDS-PAGE analysis (4–20% Criterion TGX Stain-Free Protein Gel) of whole cell lysate (WCL) and cell fractions from control and Ppk1 conditions with antibodies against known polyphosphorylated proteins NOP56, eIF5b, GTF2I, and DEK and non-polyphosphorylated proteins, Cyclin B1 and TAF10 were also assessed using antibodies against indicated proteins. Image is representative of N$\geq$3. Representative image shown here is from the same biological sample from different blots analyzed in parallel. Amount of cell material used to generate each fraction was consistent between all biological replicates tested. Fractionation

markers for panel *B* are shown in **Fig. 1D**. *(C)* NuPAGE analysis of cell fractionations from HEK293T control and Ppk1 conditions of the polyphosphorylated protein target DEK. Image is representative of N=3.

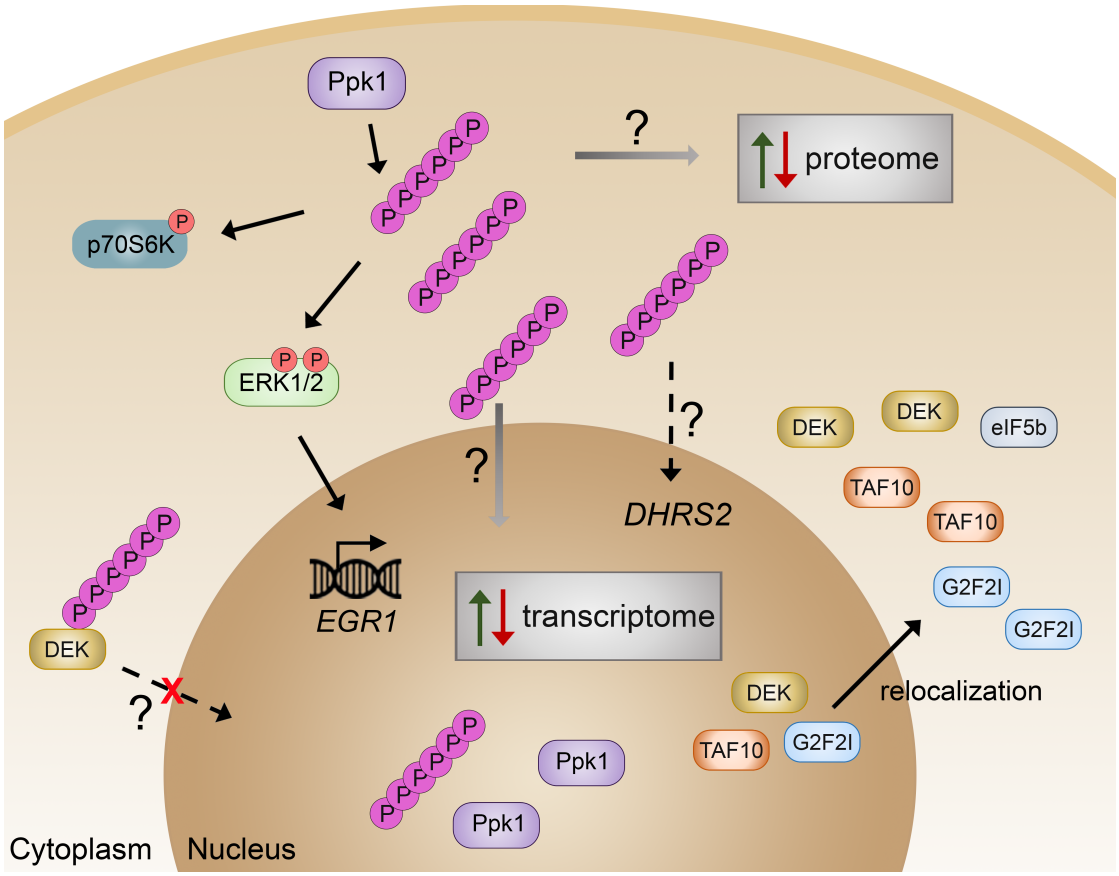
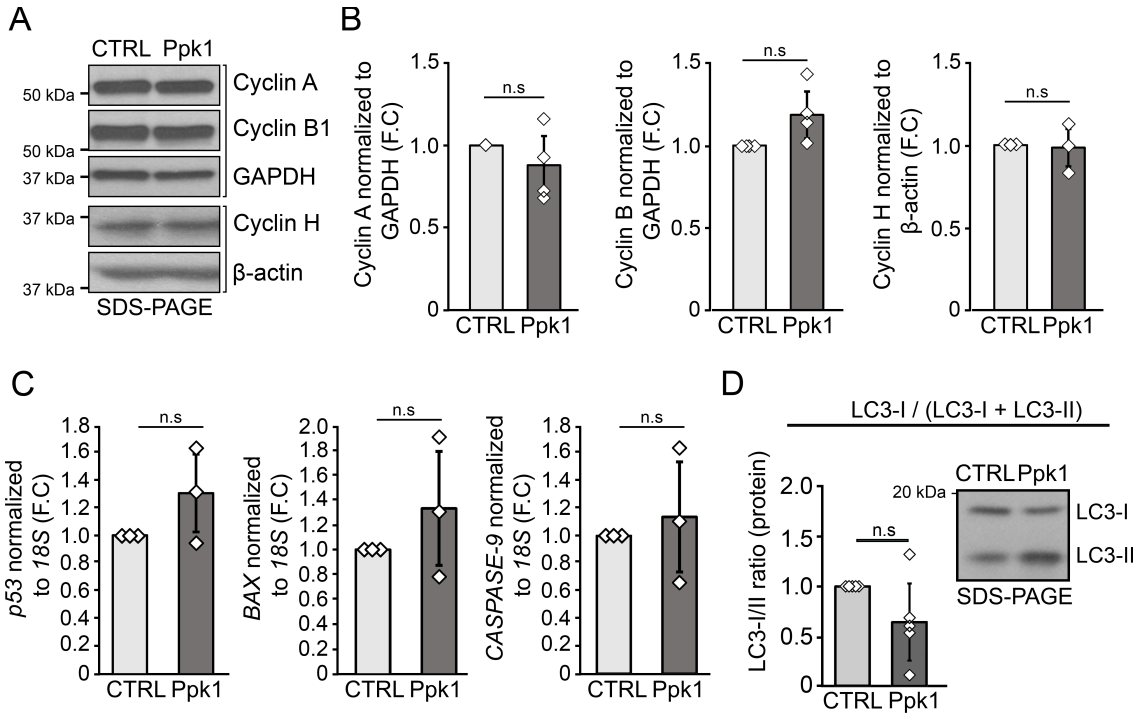
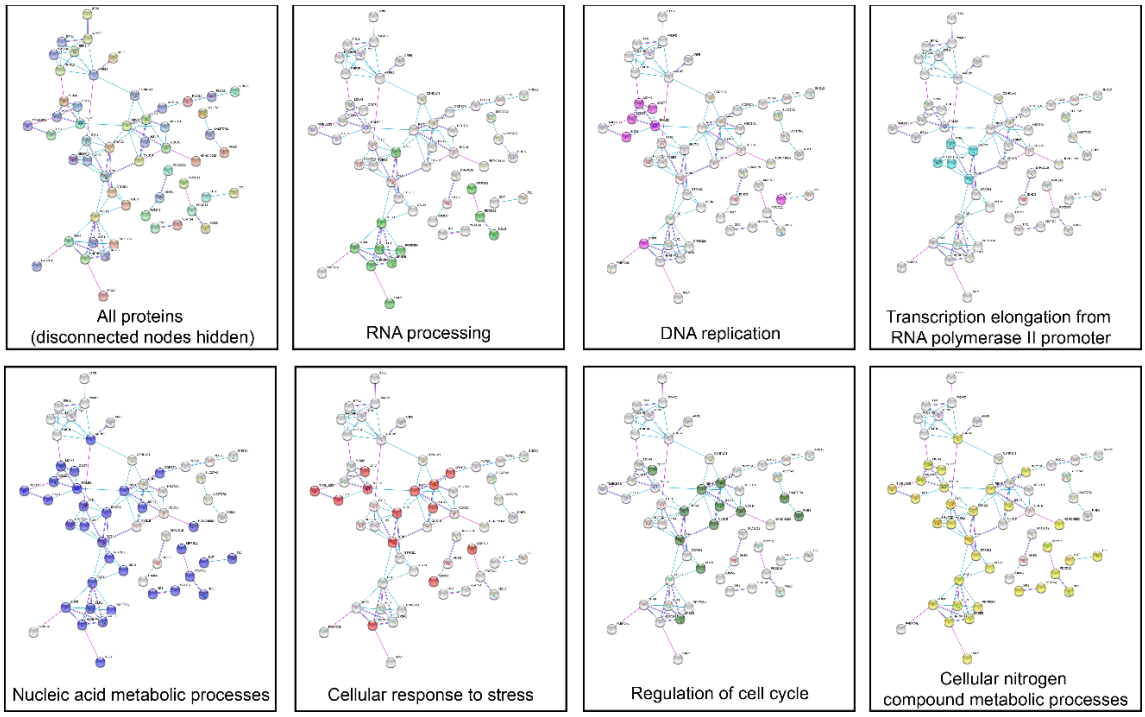


Figure 7. Model. PolyP produced via Ppk1 accumulates inside the cell and results in changes to the transcriptome and the proteome, and relocation of several nuclear proteins. The question marks represent mechanisms still to be explored.



Supplementary Figure 1. No significant impact of polyP on cell cycle, apoptosis, or autophagy. RNA and protein were harvested from HEK293T cells following a 48-hour transient transfection of either a control or *ppk1+* expression plasmid. *(A)* Western blotting analysis by SDS-PAGE in control and Ppk1 conditions cells using antibodies against Cyclin A, Cyclin B1, GAPDH, Cyclin H, and β-actin. Image is representative of N≥3. Brackets indicate the separation of biological replicates shown. For each bracketed image, the same biological sample was used from different blots analyzed in parallel. GAPDH or β-actin were used as loading controls. *(B)* Semi-quantitative analysis of protein levels shown in panel A. *(C)* RT-qPCR of *p53, BAX, and CASPASE-9* mRNA levels in control and Ppk1 conditions. Changes in mRNA levels are represented by fold change (F.C). *(D)* Western blotting analysis by SDS-PAGE and semi-quantitative analysis of protein levels of LC3 in control and Ppk1 conditions cells. P-values are shown as

numerical values or if $p \geq 0.05$ as non-significant (n.s). Statistical tests performed were one-sample t-Test (unequal variances) where error bars represent standard deviation. Image shows results from one biological replicate which is representative of $N \geq 3$.



Supplementary Figure 2. STRING. Proteins classified as significantly differentially expressed and singletons based on the mass spectrometry experiment **(Supplemental Table 4 and materials and methods)** were analyzed using the web-based STRING program v11.0 (Szklarczyk et al., 2019). Network was generated using known interactions (curated databases and experimentally determined) with disconnected nodes hidden from network. GO-terms associated with proteins shown here belong to the “biological processes” sub-ontology. Nodes in color with the terms ‘RNA processing’, ‘DNA replication’, and ‘transcription elongation from RNA polymerase II promoter’ were nodes forming isolated clusters.

Supplementary Table 1. Antibodies used in this study.

Antibody target	Dilution	Source	Identifier
Primary antibodies			
Beta-Actin HRP-conjugate	1:40,000	Santa Cruz	Sc-47778
CDC5L	1:1,000	BD Transduction Laboratories	612362
Cyclin A	1:1,000	Santa Cruz	SC-596
Cyclin B1	1:1,000	Santa Cruz	SC-752
Cyclin H	1:1,000	Santa Cruz	SC-855
DEK	1:1,000	Santa Cruz	136222
DHRS2	1:1,000	Cedarlane	GTX123431-25UL
EGR1	1:500	Cell Signalling Technologies	44D5
eIF5b	1:200	Santa Cruz	393564
ERK1/2	1:1,000	Cell Signaling Technologies	4695S
GAPDH HRP	1:40,000	Abcam	Ab85760
H3	1:10,000	Abcam	Ab1791
HA	1:500	Santa Cruz Biotechnology	sc-57592
Hsp90	1:1,000	Santa Cruz Biotechnology	Sc-13119
KAT2A/Gcn5	1:1,000	Gift from Dr. Laszlo Tora	2GC-2C11
Lamin A/C	1:1,000	Abcam	Ab108922
LC3	1:1,000	Novus	NB100-2220
MESD	1:1,000	Cell Signalling Technologies	2763
Nucleolin	1:5,000	Abcam	Ab22758
p-p70S6K	1:1,000	Cell Signaling Technologies	9234S
p53	1:1,000	Cell Signaling Technologies	2524
p70S6K	1:1,000	Cell Signalling Technologies	2708S
p-ERK1/2	1:1,000	Cell Signalling Technologies	9101S
p-p53 (Ser15)	1:1,000	Cell Signalling Technologies	9284
RALY	1:1,000	EMD millipore	AB12226E104
TAF10	1:1,000	Gift from Dr. Laszlo Tora	6TA 2B11
TFII-I	1:1,000	Santa Cruz	sc-46670

Vimentin	1:1,000	BD Pharmingen	550513
Secondary antibodies			
Goat anti-Mouse HRP	1:10,000	Biorad	172-1011
Goat anti-Rabbit HRP	1:10,000	Biorad	170-6515
IR Dye 800CW Goat anti-Mouse IgG	1:5,000	LI-COR	925-32210

Supplementary Table 2. Downey_RNA-seq_Analysis_Processed_data_2020.

GEO submission identifier: GSE149011

Supplementary Table 3. RT-qPCR primers used in this study. Primer design is detailed in materials and methods.

mRNA	Primer (5' → 3')	Primer pair efficiency (%)
β-actin - Forward	AAGAGAGGCATCCTCACCCTGAA	108.54
β-actin - Reverse	ACCGCTCATTGCCAATGGTGAT	
AVIL - Forward	TCAACGATGGTGCTGAGTCG	110.17
AVIL - Reverse	TCTGGCTTGGTGTGTAGCAG	
BAX - Forward	TGGCAGCTGACATGTTTTCTGAC	103.31
BAX - Reverse	TCACCCAACCACCCTGGTCTT	
CASPASE9 - Forward	TGTGGTGGTCATCCTCTCTCA	108.54
CASPASE9 - Reverse	GTCACTGGGGGTAGGCAAACT	
CCR4 - Forward	AATCTCGGGGAATTGCTTGTT	104.21
CCR4 - Reverse	ATTTCGGGTTCCATCTGGGTC	
CYP4F3 - Forward	CCTGACAGAAGGATGCCACA	110.81
CYP4F3 - Reverse	CCAATTCCGTTTCGGGGGTT	
DHRS2 - Reverse	TCAGCAAAGTGTTTCATGGGA	114.81
DHRS2 - Forward	CACGATTCCTGCACAGTCCT	
EFCAB6 - Forward	AGGAGTGTCAGCAGCTCATT	122.20
EFCAB6 - Reverse	CCGGCTTCTTTCATCTTGTGT	
ETV5 - Forward	ACCCACTTTGAAGACAGCCC	84.71
ETV5 - Reverse	TTGCCTGAATGGGAACTGCT	
EGR1 - Forward	TCCCCGACTACCTGTTTCCA	108.93
EGR1 - Reverse	GACAGAGGGGTTAGCGAAGG	
HIST1H1D - Forward	GCTCCACTTGCTCCTACCAT	97.13
HIST1H1D - Reverse	TGCCACTGCCTTGGTGATAA	
HIST2H4A - Forward	AGCTGTCTATCGGGCTCCA	118.64
HIST2H4A - Reverse	GGCTTGGTGATGCCCTGAAT	
ID3 - Forward	AGCGCGTCATCGACTACATT	93.34
ID3 - Reverse	TGACAAGTTCCGGAGTGAGC	
MYO1H - Forward	TTAACAGTGGTGTCAGTCCG	84.97
MYO1H - Reverse	TCCCCACTTCCCCTGTAGTT	
NOP56 - Forward	TTCGTCTGGTGGCCTTTTGT	97.02
NOP56 - Reverse	AACTCCTCCTGTATTGCGGC	
Nucleolin - Forward	AATGGCTCCTCCTCCAAAGG	105.95

Nucleolin - Reverse	GGGGAAACGACCACCTTCTT	
p53 - Forward	GCAGTCAGATCCTAGCGTCG	118.15
p53 - Reverse	TTTTCAGGAAGTAGTTTCCATAGGT	
SFRP4 - Forward	GTAAACGCCTAAGCCCCGAT	88.16
SFRP4 - Reverse	CGTTGTGACCTCATTGCAGC	

Supplementary Table 4. Differential Protein Expression Summary Downey Ppk1 Mass Spectrometry. Normalized spectral counts and statistical analysis of differential expression analyses derived from mass spectrometry data.

PRIDE identifier PXD018896 and 10.6019/PXD018896

References:

Abramov, A. Y., C. Fraley, C. T. Diao, R. Winkfein, M. A. Colicos, M. R. Duchen, R. J. French, and E. Pavlov, 2007, Targeted polyphosphatase expression alters mitochondrial metabolism and inhibits calcium-dependent cell death: Proceedings of the National Academy of Sciences, v. 104, p. 18091-18096.

Adamson, E. D., and D. Mercola, 2002, Egr1 transcription factor: multiple roles in prostate tumor cell growth and survival: Tumor biology, v. 23, p. 93-102.

Adusumilli, R., and P. Mallick, 2017, Data conversion with ProteoWizard msConvert, Proteomics, Springer, p. 339-368.

Ahn, K., and A. Kornberg, 1990, Polyphosphate kinase from Escherichia coli. Purification and demonstration of a phosphoenzyme intermediate: Journal of Biological Chemistry, v. 265, p. 11734-11739.

Akiyama, M., E. Crooke, and A. Kornberg, 1993, An exopolyphosphatase of Escherichia coli. The enzyme and its ppx gene in a polyphosphate operon: Journal of Biological Chemistry, v. 268, p. 633-639.

Alexiadis, V., T. Waldmann, J. Andersen, M. Mann, R. Knippers, and C. Gruss, 2000, The protein encoded by the proto-oncogene DEK changes the topology of chromatin and reduces the efficiency of DNA replication in a chromatin-specific manner: Genes & development, v. 14, p. 1308-1312.

Anders, S., and W. Huber, 2010, Differential expression analysis for sequence count data: Nature Precedings, p. 1-1.

Anders, S., P. T. Pyl, and W. Huber, 2015, HTSeq—a Python framework to work with high-throughput sequencing data: Bioinformatics, v. 31, p. 166-169.

Auesukaree, C., T. Homma, H. Tochio, M. Shirakawa, Y. Kaneko, and S. Harashima, 2004, Intracellular phosphate serves as a signal for the regulation of the PHO pathway in Saccharomyces cerevisiae: Journal of Biological Chemistry, v. 279, p. 17289-17294.

Azevedo, C., T. Livermore, and A. Saiardi, 2015, Protein polyphosphorylation of lysine residues by inorganic polyphosphate: Molecular cell, v. 58, p. 71-82.

Azevedo, C., J. Singh, N. Steck, A. Hofer, F. A. Ruiz, T. Singh, H. J. Jessen, and A. Saiardi, 2018, Screening a protein array with synthetic biotinylated inorganic polyphosphate to define the human polyP-ome: ACS chemical biology, v. 13, p. 1958-1963.

Bauer, N. C., P. W. Doetsch, and A. H. Corbett, 2015, Mechanisms regulating protein localization: Traffic, v. 16, p. 1039-1061.

Bauer, S., S. Grossmann, M. Vingron, and P. N. Robinson, 2008, Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration: Bioinformatics, v. 24, p. 1650-1651.

Bentley-DeSousa, A., C. Holinier, H. Moteshareie, Y.-C. Tseng, S. Kajjo, C. Nwosu, G. F. Amodeo, E. Bondy-Chorney, Y. Sai, and A. Rudner, 2018, A screen for candidate targets of lysine polyphosphorylation uncovers a conserved network implicated in ribosome biogenesis: Cell reports, v. 22, p. 3427-3439.

Bjornson, R. D., N. J. Carriero, C. Colangelo, M. Shifman, K.-H. Cheung, P. L. Miller, and K. Williams, 2008, X!! Tandem, an improved method for running X! tandem in parallel on collections of commodity computers: The Journal of Proteome Research, v. 7, p. 293-299.

Bolger, A. M., M. Lohse, and B. Usadel, 2014, Trimmomatic: a flexible trimmer for Illumina sequence data: Bioinformatics, v. 30, p. 2114-2120.

Bourgey, M., R. Dali, R. Eveleigh, K. C. Chen, L. Letourneau, J. Fillon, M. Michaud, M. Caron, J. Sandoval, and F. Lefebvre, 2019, GenPipes: an open-source framework for distributed and scalable genomic analyses: GigaScience, v. 8, p. giz037.

Bru, S., J. M. Martínez-Laínez, S. Hernández-Ortega, E. Quandt, J. Torres-Torronteras, R. Martí, D. Canadell, J. Ariño, S. Sharma, and J. Jiménez, 2016, Polyphosphate is involved in cell cycle progression and genomic stability in Saccharomyces cerevisiae: Molecular microbiology, v. 101, p. 367-380.

Bru, S., B. Samper-Martín, E. Quandt, S. Hernández-Ortega, J. M. Martínez-Laínez, E. Garí, M. Rafel, J. Torres-Torronteras, R. Martí, and M. P. Ribeiro, 2017, Polyphosphate is a key factor for cell survival after DNA damage in eukaryotic cells: DNA repair, v. 57, p. 171-178.

Consortium, U., 2015, UniProt: a hub for protein information: Nucleic acids research, v. 43, p. D204-D212.

Cummings, M. C., 1996, Increased p53 mRNA expression in liver and kidney apoptosis: Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, v. 1315, p. 100-104.

Desfougères, Y., A. Saiardi, and C. Azevedo, 2020, Inorganic polyphosphate in mammals: where's Wally?: Biochemical Society Transactions, v. 48, p. 95-101.

Deutsch, E. W., A. Csordas, Z. Sun, A. Jarnuczak, Y. Perez-Riverol, T. Ternent, D. S. Campbell, M. Bernal-Llinares, S. Okuda, and S. Kawano, 2016, The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition: Nucleic acids research, p. gkw936.

Dinarvand, P., S. M. Hassanian, S. H. Qureshi, C. Manithody, J. C. Eissenberg, L. Yang, and A. R. Rezaie, 2014, Polyphosphate amplifies proinflammatory responses of nuclear proteins through interaction with receptor for advanced glycation end products and P2Y1 purinergic receptor: Blood, The Journal of the American Society of Hematology, v. 123, p. 935-945.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras, 2013, STAR: ultrafast universal RNA-seq aligner: Bioinformatics, v. 29, p. 15-21.

Donadel, G., C. Garzelli, R. Frank, and F. Gabrielli, 1991, Identification of a novel nuclear protein synthesized in growth-arrested human hepatoblastoma HepG2 cells: European journal of biochemistry, v. 195, p. 723-729.

Fernandes-Cunha, G. M., C. J. McKinlay, J. R. Vargas, H. J. Jessen, R. M. Waymouth, and P. A. Wender, 2018, Delivery of inorganic polyphosphate into cells using amphipathic oligocarbonate transporters: ACS central science, v. 4, p. 1394-1402.

Gabel, N.-W., and V. Thomas, 1971, Evidence for the occurrence and distribution of inorganic polyphosphates in vertebrate tissues: Journal of neurochemistry, v. 18, p. 1229-1242.

Gabrielli, F., G. Donadel, G. Bensi, A. Heguy, and M. Melli, 1995, A nuclear protein, synthesized in growth-arrested human hepatoblastoma cells, is a novel member of the short-chain alcohol dehydrogenase family: European journal of biochemistry, v. 232, p. 473-477.

Gerasimaitė, R., and A. Mayer, 2016, Enzymes of yeast polyphosphate metabolism: structure, enzymology and biological roles: Biochemical Society Transactions, v. 44, p. 234-239.

Gerasimaitė, R., S. Sharma, Y. Desfougères, A. Schmidt, and A. Mayer, 2014, Coupled synthesis and translocation restrains polyphosphate to acidocalcisome-like vacuoles and prevents its toxicity: J Cell Sci, v. 127, p. 5093-5104.

Hassanian, S. M., P. Dinarvand, S. A. Smith, and A. R. Rezaie, 2015, Inorganic polyphosphate elicits pro-inflammatory responses through activation of the mammalian target of rapamycin complexes 1 and 2 in vascular endothelial cells: Journal of Thrombosis and Haemostasis, v. 13, p. 860-871.

Hernandez-Ruiz, L., I. González-García, C. Castro, J. A. Brieva, and F. A. Ruiz, 2006, Inorganic polyphosphate and specific induction of apoptosis in human plasma cells: Haematologica, v. 91, p. 1180-1186.

Holmström, K. M., N. Marina, A. Y. Baev, N. W. Wood, A. V. Gourine, and A. Y. Abramov, 2013, Signalling properties of inorganic polyphosphate in the mammalian brain: Nature communications, v. 4, p. 1-8.

Hothorn, M., H. Neumann, E. D. Lenherr, M. Wehner, V. Rybin, P. O. Hassa, A. Uttenweiler, M. Reinhardt, A. Schmidt, and J. Seiler, 2009, Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase: Science, v. 324, p. 513-516.

Hu, H.-g., I. Scholten, C. Gruss, and R. Knippers, 2007, The distribution of the DEK protein in mammalian chromatin: Biochemical and biophysical research communications, v. 358, p. 1008-1014.

Huber, W., V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, and T. Girke, 2015, Orchestrating high-throughput genomic analysis with Bioconductor: Nature methods, v. 12, p. 115.

Ingham, M., and G. K. Schwartz, 2017, Cell-cycle therapeutics come of age: Journal of Clinical Oncology, v. 35, p. 2949.

Jimenez-Nuñez, M. D., D. Moreno-Sanchez, L. Hernandez-Ruiz, A. Benítez-Rondán, A. Ramos-Amaya, B. Rodríguez-Bayona, F. Medina, J. A. Brieva, and F. A. Ruiz, 2012, Myeloma cells contain high levels of inorganic polyphosphate which is associated with nucleolar transcription: Haematologica, v. 97, p. 1264-1271.

Kappes, F., K. Burger, M. Baack, F. O. Fackelmayer, and C. Gruss, 2001, Subcellular localization of the human proto-oncogene protein DEK: Journal of Biological Chemistry, v. 276, p. 26317-26323.

Ki, S., F. Sugihara, K. Kasahara, H. Tochio, M. Shirakawa, and T. Kokubo, 2007, Magnetic resonance-based visualization of gene expression in mammalian cells using a bacterial polyphosphate kinase reporter gene: Biotechniques, v. 42, p. 209-215.

Kornberg, A., S. Kornberg, and E. S. Simms, 1956, Metaphosphate synthesis by an enzyme from Escherichia coli: Biochimica et Biophysica Acta, v. 20, p. 215-227.

Kornberg, A., N. N. Rao, and D. Ault-Riche, 1999, Inorganic polyphosphate: a molecule of many functions: Annual review of biochemistry, v. 68, p. 89-125.

Kulaev, I. S., and V. M. Vagabov, 1983, Polyphosphate metabolism in micro-organisms, Advances in microbial physiology, v. 24, Elsevier, p. 83-171.

Kumar, S., and D. Cakouros, 2004, Transcriptional control of the core cell-death machinery: Trends in biochemical sciences, v. 29, p. 193-199.

Kumble, K. D., K. Ahn, and A. Kornberg, 1996, Phosphohistidyl active sites in polyphosphate kinase of Escherichia coli: Proceedings of the National Academy of Sciences, v. 93, p. 14391-14395.

Kumble, K. D., and A. Kornberg, 1995, Inorganic polyphosphate in mammalian cells and tissues: Journal of Biological Chemistry, v. 270, p. 5818-5822.

Lichko, L., T. Kulakovskaya, N. Pestov, and I. Kulaev, 2006, Inorganic polyphosphates and exopolyphosphatases in cell compartments of the yeast Saccharomyces cerevisiae under inactivation of PPX1 and PPN1 genes: Bioscience reports, v. 26, p. 45-54.

Liu, D., and P. Lobie, 2007, Transcriptional activation of p53 by Pitx1: Cell Death & Differentiation, v. 14, p. 1893-1907.

Lonetti, A., Z. Szijgyarto, D. Bosch, O. Loss, C. Azevedo, and A. Saiardi, 2011, Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases: Journal of Biological Chemistry, v. 286, p. 31966-31974.

Lui, E. L.-H., C. K.-L. Ao, L. Li, M.-L. Khong, and J. A. Tanner, 2016, Inorganic polyphosphate triggers upregulation of interleukin 11 in human osteoblast-like SaOS-2 cells: Biochemical and biophysical research communications, v. 479, p. 766-771.

Makeyev, A. V., B. Enkhmandakh, S.-H. Hong, P. Joshi, D.-G. Shin, and D. Bayarsaihan, 2012, Diversity and complexity in chromatin recognition by TFII-I transcription factors in pluripotent embryonic stem cells and embryonic tissues: PloS one, v. 7.

Malumbres, M., and M. Barbacid, 2005, Mammalian cyclin-dependent kinases: Trends in biochemical sciences, v. 30, p. 630-641.

Milbrandt, J., 1987, A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor: Science, v. 238, p. 797-799.

Mizushima, N., T. Yoshimori, and B. Levine, 2010, Methods in mammalian autophagy research: Cell, v. 140, p. 313-326.

Müller, F., N. J. Mutch, W. A. Schenk, S. A. Smith, L. Esterl, H. M. Spronk, S. Schmidbauer, W. A. Gahl, J. H. Morrissey, and T. Renné, 2009, Platelet polyphosphates are proinflammatory and procoagulant mediators in vivo: Cell, v. 139, p. 1143-1156.

Nakamura, A., N. Kawano, K. Motomura, A. Kuroda, K. Sekiguchi, M. Miyado, W. Kang, Y. Miyamoto, M. Hanai, and M. Iwai, 2018, Degradation of phosphate polymer polyP enhances lactic fermentation in mice: Genes to Cells, v. 23, p. 904-914.

Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, and M. Eisenacher, 2019, The PRIDE database and related tools and resources in 2019: improving support for quantification data: Nucleic acids research, v. 47, p. D442-D450.

Pisoni, R. L., and E. R. Lindley, 1992, Incorporation of [32P] orthophosphate into long chains of inorganic polyphosphate within lysosomes of human fibroblasts: Journal of Biological Chemistry, v. 267, p. 3626-3631.

Portt, L., G. Norman, C. Clapp, M. Greenwood, and M. T. Greenwood, 2011, Anti-apoptosis and cell survival: a review: Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, v. 1813, p. 238-259.

Racki, L. R., E. I. Tocheva, M. G. Dieterle, M. C. Sullivan, G. J. Jensen, and D. K. Newman, 2017, Polyphosphate granule biogenesis is temporally and functionally tied to cell cycle exit during starvation in Pseudomonas aeruginosa: Proceedings of the National Academy of Sciences, v. 114, p. E2440-E2449.

Raman, V., S. A. Martensen, D. Reisman, E. Evron, W. F. Odenwald, E. Jaffee, J. Marks, and S. Sukumar, 2000, Compromised HOXA5 function can limit p53 expression in human breast tumours: Nature, v. 405, p. 974.

Rao, N. N., S. Liu, and A. Kornberg, 1998, Inorganic polyphosphate in Escherichia coli: the phosphate regulon and the stringent response: Journal of bacteriology, v. 180, p. 2186-2193.

Rashid, M. H., K. Rumbaugh, L. Passador, D. G. Davies, A. N. Hamood, B. H. Iglewski, and A. Kornberg, 2000, Polyphosphate kinase is essential for biofilm development, quorum sensing, and virulence of Pseudomonas aeruginosa: Proceedings of the National Academy of Sciences, v. 97, p. 9636-9641.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth, 2010, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data: Bioinformatics, v. 26, p. 139-140.

Roskoski Jr, R., 2012, ERK1/2 MAP kinases: structure, function, and regulation: Pharmacological research, v. 66, p. 105-143.

Ruiz, F. A., C. R. Lea, E. Oldfield, and R. Docampo, 2004, Human platelet dense granules contain polyphosphate and are similar to acidocalcisomes of bacteria and unicellular eukaryotes: Journal of Biological Chemistry, v. 279, p. 44250-44257.

Schafer, K., 1998, The cell cycle: a review: Veterinary pathology, v. 35, p. 461-478.

Schmittgen, T. D., and K. J. Livak, 2008, Analyzing real-time PCR data by the comparative C T method: Nature protocols, v. 3, p. 1101.

Seidlmayer, L. K., M. R. Gomez-Garcia, L. A. Blatter, E. Pavlov, and E. N. Dedkova, 2012, Inorganic polyphosphate is a potent activator of the mitochondrial permeability transition pore in cardiac myocytes: Journal of General Physiology, v. 139, p. 321-331.

Shafqat, N., J. Shafqat, G. Eissner, H.-U. Marschall, K. Tryggvason, U. Eriksson, F. Gabrielli, H. Lardy, H. Jörnvall, and U. Oppermann, 2006, Hep27, a member of the short-chain dehydrogenase/reductase family, is an NADPH-dependent dicarbonyl reductase expressed in vascular endothelial tissue: Cellular and Molecular Life Sciences CMLS, v. 63, p. 1205-1213.

Shibutani, S. T., T. Saitoh, H. Nowag, C. Münz, and T. Yoshimori, 2015, Autophagy and autophagy-related proteins in the immune system: Nature immunology, v. 16, p. 1014.

Simbulan-Rosenthal, C. M., A. Gaur, V. A. Sanabria, L. J. Dussan, R. Saxena, J. Schmidt, T. Kitani, Y. Chen, S. Rahim, and A. Uren, 2015, Inorganic polyphosphates are important for cell survival and motility of human skin keratinocytes: Experimental dermatology, v. 24, p. 636.

Smith, S. A., and J. H. Morrissey, 2007, Sensitive fluorescence detection of polyphosphate in polyacrylamide gels using 4′, 6-diamidino-2-phenylindol: Electrophoresis, v. 28, p. 3461-3465.

Soutoglou, E., M. A. Demény, E. Scheer, G. Fienga, P. Sassone-Corsi, and L. Tora, 2005, The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners: Molecular and cellular biology, v. 25, p. 4092-4104.

Sukhatme, V. P., X. Cao, L. C. Chang, C.-H. Tsai-Morris, D. Stamenkovich, P. C. Ferreira, D. R. Cohen, S. A. Edwards, T. B. Shows, and T. Curran, 1988, A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization: Cell, v. 53, p. 37-43.

Szklarczyk, D., A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, and P. Bork, 2019, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets: Nucleic acids research, v. 47, p. D607-D613.

Tammenkoski, M., K. Koivula, E. Cusanelli, M. Zollo, C. Steegborn, A. A. Baykov, and R. Lahti, 2008, Human metastasis regulator protein H-prune is a short-chain exopolyphosphatase: Biochemistry, v. 47, p. 9707-9713.

Travers, R. J., S. A. Smith, and J. H. Morrissey, 2015, Polyphosphate, platelets, and coagulation: International journal of laboratory hematology, v. 37, p. 31-35.

Veyrac, A., A. Besnard, J. Caboche, S. Davis, and S. Laroche, 2014, The transcription factor Zif268/Egr1, brain plasticity, and memory, Progress in molecular biology and translational science, v. 122, Elsevier, p. 89-129.

Waldmann, T., C. Eckerich, M. Baack, and C. Gruss, 2002, The ubiquitous chromatin protein DEK alters the structure of DNA by introducing positive supercoils: Journal of Biological Chemistry, v. 277, p. 24988-24994.

Wang, L., C. D. Fraley, J. Faridi, A. Kornberg, and R. A. Roth, 2003, Inorganic polyphosphate stimulates mammalian TOR, a kinase involved in the proliferation of mammary cancer cells: Proceedings of the National Academy of Sciences, v. 100, p. 11249-11254.

Xie, L., and U. Jakob, 2019, Inorganic polyphosphate, a multifunctional polyanionic protein scaffold: Journal of Biological Chemistry, v. 294, p. 2180-2190.

Xie, L., A. Rajpurkar, N. Taube, E. Quales, J. Erba, A. S. Rai, B. Sliwinski, M. Markowitz, U. Jakob, and D. Knoefler, 2019, Accumulation of Nucleolar Inorganic Polyphosphate Is a Cellular Response to Cisplatin-Induced Apoptosis: Frontiers in Oncology, v. 9, p. 1410.

Yam, C., T. Fung, and R. Poon, 2002, Cyclin A in cell cycle control and cancer: Cellular and Molecular Life Sciences CMLS, v. 59, p. 1317-1326.

Young, M. D., M. J. Wakefield, G. K. Smyth, and A. Oshlack, 2010, Gene ontology analysis for RNA-seq: accounting for selection bias: Genome biology, v. 11, p. R14.

Zhou, Y., L. Wang, X. Ban, T. Zeng, Y. Zhu, M. Li, X.-Y. Guan, and Y. Li, 2018, DHRS2 inhibits cell growth and motility in esophageal squamous cell carcinoma: Oncogene, v. 37, p. 1086-1094.