Investigating polyphosphate biology: From post-translational modification to rare disease

Amanda Bentley-DeSousa

A thesis submitted to the
Faulty of Graduate and Postdoctoral Studies
in partial fulfillment of requirements for the
Doctorate in Philosophy degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

Authorizations

Manuscript #1

Bentley-DeSousa, A., Holinier, C., Moteshareie, H., Tseng, Y-C., Kajjo, S., Nwosu, C., Amodeo, G.F., Bondy-Chorney, E., Sai, Y., Rudner A., Golshani, A., Davey, N.E., and Downey, M. (2018). A Screen for Candidate Targets of Lysine Polyphosphorylation Uncovers a Conserved Network Implicated in Ribosome Biogenesis. *Cell Reports* 22, 3427-3439.

Manuscript #2

McCarthy, L.*, **Bentley-DeSousa, A.***, Denoncourt, A., Tseng, Y-C., Gabriel, M., and Downey M. (2020). Proteins required for vacuolar function are targets of lysine polyphosphorylation in yeast. *FEBS Letters* 594, 21-30.
*Authors contributed equally

Manuscript #3

Bentley-DeSousa, A., and Downey, M. Vtc5 is localized to the vacuole membrane by the conserved AP-3 complex to regulate polyphosphate synthesis in budding yeast.

Manuscript submitted to BioRxiv and *in review* at mBio

Manuscript #4 (Review – Appendix C)

Bentley-DeSousa, A., and Downey, M. (2019). From underlying chemistry to therapeutic potential: open questions in the new field of lysine polyphosphorylation. *Current Genetics* 65, 57-64.

Abstract

The first report of polyphosphates (polyP) was in 1890 by L. Liberman and since then, polyP's role in biology has been explored. PolyPs are chains of phosphoanhydride-linked inorganic phosphates ranging from 3-1000s of units in length. These chains are implicated in many cellular pathways including blood clotting, bacterial virulence, and neuroproteotoxic disease. Given the diversity of polyP, they make an excellent candidate in the development of novel therapeutics. In yeast, polyP is synthesized by the vacuolar transporter chaperone (VTC) complex as a translocation event into the vacuole lumen. In 2015, polyP chains were found to act as a post-translational modification termed polyphosphorylation on yeast proteins (Nsr1 and Top1). This modification occurs non-enzymatically on lysine residues within poly-acidic, serine, and lysine (PASK) motifs and can only be detected via electrophoretic mobility shift on NuPAGE gels. We have since expanded the pool of yeast polyphosphorylated substrates to 25, with an enrichment of proteins with roles related to RNA biology. Additionally, we were the first group to demonstrate polyphosphorylation of 6 human proteins by expressing *E. coli* PPK1 in HEK293T cells. We next focused on elaborating how polyP is being regulated via the VTC complex by assessing which protein trafficking pathways are critical for VTC localization at the vacuole membrane. We found the adaptor protein 3 (AP-3) complex is responsible for localizing Vtc5 subunit to the vacuole membrane and in AP-3 mutants, Vtc5 becomes mislocalized to the vacuole lumen and degraded. Vtc5 degradation, upon AP-3 mutation, is mediated by the endosomal sorting complex required for transport (ESCRT) complex. The loss of polyP in AP-3 mutants is imparted by Vtc5 mislocalization. In humans, mutations in AP-3 cause a rare genetic disorder termed Hermansky-Pudlak Syndrome (HPS) which has a wide range of symptoms. These include defects in polyP accumulation in platelets, likely related to a loss of polyP. We expect that our work using yeast will provide a framework for understanding fundamental aspects of polyP biology related to HPS and other health conditions.

Contents

List of Tables

*Note: due to the size of these tables, they will be included in Appendices A and B.

List of Figures

List of Abbreviations

AP-3: adaptor protein 3 complex

CHX: cycloheximide

Cont.: control

DAPI: 4',6-diamidino-2-phenylindole

DAVID: database for annotation, visualization, and integrated discovery

DEX: dextrose

DMSO: dimethyl sulfoxide

DTT: dithiothreitol

ESCRT: endosomal sorting complex required for transport

FDR: false discovery rate

GAL: galactose

GFP: green fluorescent protein

GST: glutathione S-transferase

HA: hemagglutinin

HIR: DNA replication-independent histone chaperone complex

HPS: Hermansky-Pudlak Syndrome

HOPS: homotypic fusion and protein sorting complex

IB: immunoblot

ILV: intralumenal vesicle

MMS: methyl methanesulfonate

MVB: multi-vesicular body

ORF: open reading frame

PASK: poly-acidic, serine and lysine

PBS: phosphate-buffered saline

PolyP: polyphosphate

PPK: polyphosphate kinase

PTM: post-translational modification

RAP/RAPA: rapamycin

RIPA: radioimmunoprecipitation assay

RAFF: raffinose

SC: synthetic complete

SCFSaf1: Skp1-Cul1-F-Box SCF associated factor 1 ubiquitin ligase complex

SDS-PAGE: sodium dodecyl slfate polyacrylamide gel electrophoresis

SWI/SNF: SWItch/Sucrose Non-Fermentable complex

SUMO: small ubiquitin-like modifier

TAP: tandem affinity purification

TCA: trichloroacetic acid

TEF: translational elongation factor

TEMED: tetramethylethylenediamine

TGN: trans-golgi network

TOR: target of rapamycin

VTC: vacuolar transporter chaperone

WT: wild-type

WCE: whole cell extract

YPD: yeast extract peptone dextrose

Acknowledgements

When thinking about this section, I realized I'm fortunate to have had so many people contribute to who I am today and who I will be tomorrow. There are so many people who have not only supported in my endeavours but have been role models that I've looked up to my whole life. Whether it was athletics or academics, I couldn't thank everyone enough.

In terms of academia, I would like to specifically acknowledge and thank those who have supported me and taken a chance on me. All of those involved have truly helped me develop my passion for science and research. Specific thanks to my supervisor (Michael Downey), my thesis committee members, and those who are on my defence examining committee. I'd also like to thank all of the support staff at uOttawa who work behind the scenes and go unnoticed at times. Finally, thank you to all of my past and current friends and labmates. I am the luckiest person to be surrounded by people who not only share a similar passion as I, but who make every day just a little bit better.

Lastly, I'd like to thank my family and friends for all of the support they've given me my whole life. Even though my interests are certainly different than most of you, you've all been there no matter what.

Thank you to everyone who has been a part of my life. I'm on to new journeys post-grad and I can't be more excited.

Chapter 1: Introduction

1.1: Polyphosphate: an ancient molecule with wide applications

Polyphosphate (polyP) chains are comprised of units of inorganic orthophosphates linked via phosphoanhydride bonds. These chains are heterogeneous in nature, ranging from 3-1000s of phosphate units in length (**Figure 1**) (Rao et al., 2009). The first description of polyP stems from observations in 1890 by L. Liberman, however, despite over a century of research, the field of polyP research still contains many open questions (Kulaev et al., 2004).

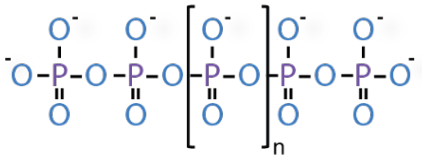


Fig. 1: PolyP is Comprised of Phosphate Units Linked via Phosphoanhydride Bonds. Chains can range from 3-1000s of units in length and are implicated in many cellular processes.

These chains are present in all prokaryotic and eukaryotic organisms studied to date and are thought to play critical roles in various aspects of cell biology (Rao et al., 2009). Given the high-energy nature of phosphoanhydride bonds, these chains were originally predicted to serve as an energy reservoir, but a plethora of research since has demonstrated that polyP acts in a variety of ways. At a foundational biochemical level, these chains have been implicated in maintaining the homeostasis of many cellular aspects such as: ion balance (Nguyen et al., 2019; Ryazanova et al., 2016; Trilisenko et al., 2019), phosphate homeostasis (Desfougeres et al., 2016; Eskes et al., 2018; Huang et al., 2002; Ogawa et al., 2000), and pH level maintenance (Eskes et al., 2018). More intriguingly, polyP chains have been associated with multiple aspects of human health, such as: bacterial virulence (Kim et al., 2002; Peng et al., 2016; Rao et al., 2009; Rashid and Kornberg, 2000; Rashid et al., 2000a; Rashid et al., 2000b; Shiba et al., 1997), blood coagulation (Morrissey, 2012; Morrissey et al., 2012; Morrissey and Smith, 2015), and neuroproteotoxic disease (Cremers et al., 2016). Interestingly, the heterogeneity of these

chains can confer different functions, particularly in the context of blood coagulation. In this instance, shorter chains (~60-100 phosphate units long) act predominantly to amplify blood coagulation by acting on factor XI, whereas longer chains (100s of phosphate units long) act largely on factor XII and fibrin (Morrissey et al., 2012). Given the diverse applications of these heterogeneous polyP chains, researchers emphasize the potential use of manipulating these chains for therapeutic purposes in a wide range of pathologies (Bentley-DeSousa and Downey, 2019).

1.2: PolyP chains are biochemically versatile molecules

Although polyP as a polymer is very simple in nature, these chains possess interesting biochemical features which add to the complexity and diversity of this molecule. Hydrolysis of polyP chains release the same amount of energy as the cleavage of phosphate from ATP (Kulaev et al., 2004). The hydrolysis of these chains occurs more rapidly with increases in temperature, a reduction in pH, and in combination with complex cations (Kulaev et al., 2004). These polyanionic chains can form complexes with cations, mostly Mg^{2+}, Ca^{2+}, and Mn^{2+}, often within acidocalcisomes, like the yeast vacuole, where polyP is typically stored at the highest concentrations (Kulaev et al., 2004). Additionally, polyP can also be found in complexes with specific amino acids (arginine and lysine), spermidine, RNA, DNA, and proteins (Kulaev et al., 2004). Perhaps the most intriguing biochemical quality of polyP is its association with proteins. There has been an abundance of reports demonstrating that polyP interacts with proteins through electrostatic interactions and can act as a protein chaperone (Azevedo et al., 2018; Gray et al., 2014; Mutch et al., 2010b). The most interesting examples of these phenotypes come from the blood coagulation cascade (Mutch et al., 2010b) and the bacteria heat shock response (Gray et al., 2014). Overall, this versatile molecule has very interesting biochemical

properties that affect how it functions *in vivo*. Some of these examples have been extensively studied, while others give rise to even more questions.

1.3: Methods to assess polyP qualitatively and quantitatively

Given that polyP has been extensively studied for the past half a century, many methods have been developed to detect and assess polyP. For the purpose of this thesis, the most common methods used in cell biology labs will be discussed, however, there exist other methods that are more commonly used in other fields (**Figure 2**) (Christ et al., 2020). Qualitative methods assess polyP by extracting the compound and analyzing it on a gel or by fluorescence microscopy. Typically, polyP is extracted using a phenol-chloroform or an acid-based protocol which includes DNAse and RNAse treatments (Bru et al., 2016a). The extracted polyP is then electrophoresed on polyacrylamide (Christ et al., 2020) or TBE-UREA gels (Smith and Morrissey, 2007). Following this, gels are stained with either DAPI or Toluidine Blue for visualization (Christ et al., 2020). Additionally, there are microscopy methods that take advantage of polyP's ability to bind multiple compounds such as: DAPI (Christ et al., 2020), polyP binding domains (mostly derived from *S. cerevisiae* Ppx1) coupled as a probe (Saito et al., 2005; Werner et al., 2007), and/or fluorescent probes (Angelova et al., 2014). When polyP interacts with DAPI, the emission wavelength of DAPI shifts from 475 nm (when bound to DNA) to 525 nm with an excitation at 360 nm. In order to determine the concentration of polyP in a quantitative fashion, the most common method involves the degradation of extracted polyP chains followed by methods to quantify phosphate release. Extracted polyP is degraded either enzymatically using *Saccharomyces cerevisiae* (*S. cerevisiae*) exopolyphosphatase Ppx1 or chemically by acid treatment *in vitro.* Concentration is then measured by release of phosphate units with Malachite Green, or other similar kits, by measuring absorbance at 650 nm (Bru et al., 2016a; Lonetti et al., 2011). There are pros and cons to each of the methods discussed and

choice of methodology largely depends on which information is to be gathered. Ultimately, qualitative methods can provide information about the length and localization of polyP whereas quantitative methods provide information about concentrations.

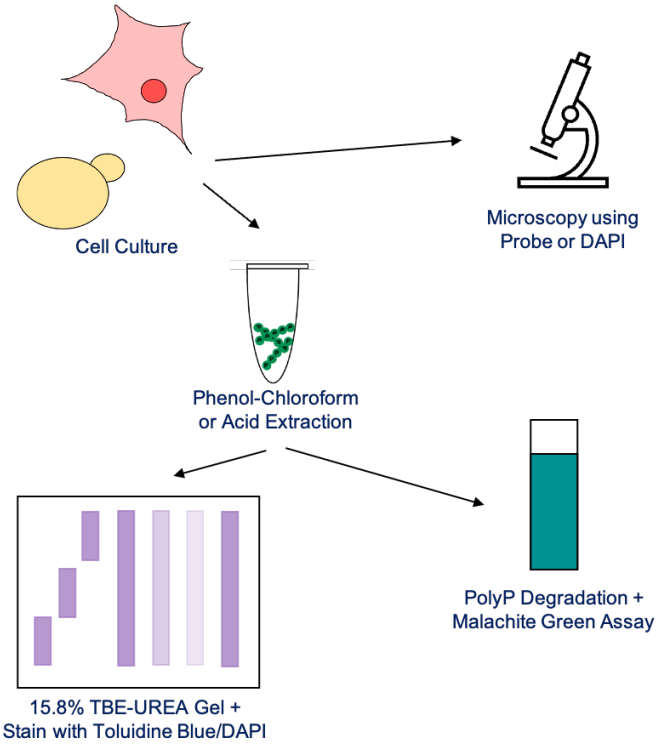


Fig 2: PolyP can be Assessed using Qualitative and Quantitative Methods. The two major means of assessing polyP involve either A) polyP extractions, or B) *in vivo* analysis using microscopy.

1.4: PolyP chains are synthesized by polyP synthetases and degraded by polyphosphatases

The metabolism of polyP chains is well characterized in multiple lower-order organisms. PolyP chains are synthesized by enzymes called polyP synthetases, which typically catalyze the creation of phosphoanhydride bonds by removing the terminal, gamma phosphate from ATP (or in some cases GTP) to create a growing polyP chain (**Figure 3**) (Hothorn et al., 2009; Kornberg, 1999). In prokaryotes, polyP is synthesized in the cytoplasm, whereas in lower-order eukaryotes the synthesis of these chains typically occurs within acidocalcisomes/vacuoles (Denoncourt and Downey, 2021).

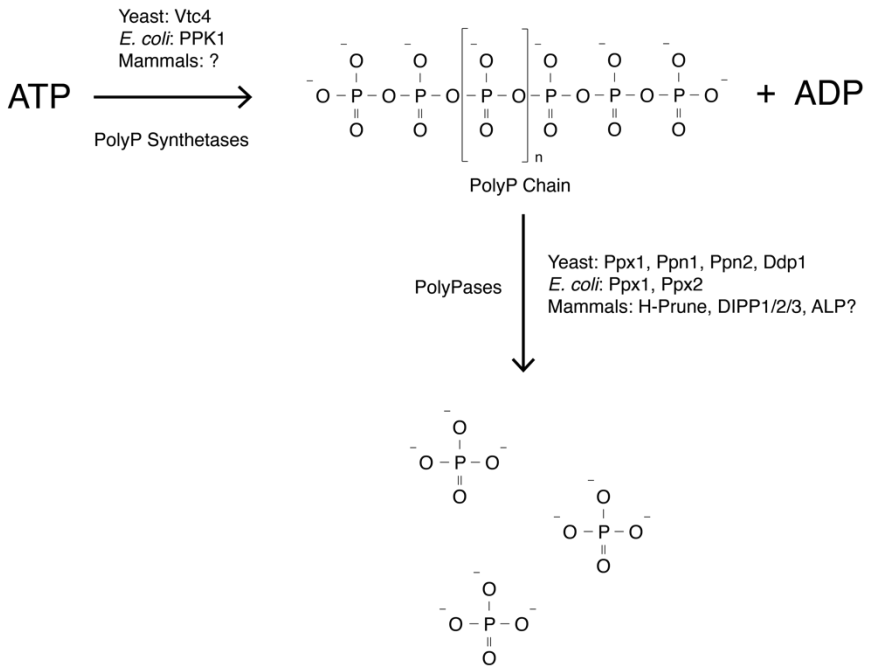


Fig. 3: PolyP is Synthesized by PolyP Synthetases and Degraded by Polyphosphatases. PolyP synthetases typically take the terminal gamma phosphate from ATP and catalyze the creation of phosphoanhydride bonds, elongating polyP chains. PolyP chains can then be degraded by exo- and endo-polyphosphatases. The metabolism of polyP is well known in lower-order organisms, such as yeast and *E. coli*, however, these enzymes are largely uncharacterized in mammals.

PolyP is degraded by common phosphatases, such as alkaline phosphatase (Lorenz and Schroder, 2001), and/or polyphosphatases (Denoncourt and Downey, 2021; Lonetti et al., 2011). Polyphosphatase enzymes possess either exo- or endo-polyphosphatase activity, and in some cases both activities can be performed. Exopolyphosphatases remove terminal phosphates from polyP chains to leave a polyP chain (n – 1) and a phosphate unit (Pi). Endopolyphosphatases cleave internal phosphoanhydride bonds within chains to create two shorter polyP chains (Bentley-DeSousa and Downey, 2019). In lower-order eukaryotes, polyphosphatases are often stored in different subcellular compartments, contain different affinities for polyP, and require specific cations for function (Denoncourt and Downey, 2021). Overall, the basic metabolism of these chains *in vivo* is well established in lower-order

organisms. Unfortunately, there are no proposed mammalian polyP synthetases. Nonetheless, short chain polyphosphatases have been proposed, however, their activity *in vivo* has not been characterized (Lonetti et al., 2011; Tammenkoski et al., 2008) (see section 1.6).

1.5: Saccharomyces cerevisiae is an ideal model system to study polyP biology

One model organism often used to study the foundations of polyP biology is the budding yeast *S. cerevisiae*. Apart from its general benefits as a model organism, such as low cost, ease of genetic manipulation, high-throughput analysis, and conservation of key pathways, many advancements in polyP biology have come from studies in *S. cerevisiae*. Firstly, polyP is stored in very high concentrations within yeast vacuoles at over 200 mM (Auesukaree et al., 2004; Kornberg, 1999). Secondly, the enzymes involved in the metabolism of polyP (polyP synthetase and polyphosphatases) are well characterized (**Figure 3**). Finally, high-throughput screens have been conducted in order to determine which genes impact polyP production on a genome-wide scale (Freimoser et al., 2006). To date, the only polyP synthetase in yeast is the vacuolar transporter chaperone (VTC) complex (**Figure 4**). This complex consists of 3 necessary subunits: Vtc1, Vtc2 or Vtc3, and Vtc4 (the catalytic subunit) (Hothorn et al., 2009). The VTC complex is embedded in the vacuole membrane where it synthesizes growing polyP chains as a translocation event, ultimately storing the chains within the vacuole lumen (Gerasimaite et al., 2014). Apart from its role in synthesizing polyP, the VTC complex has also been implicated in vacuole fusion, microautophagy, and in regulating V-ATPase subunits (Muller et al., 2002; Muller et al., 2003; Uttenweiler et al., 2007). Therefore, it plays other roles at the vacuole and is required for proper vacuole function in general. Interestingly, the VTC complex has also been found in other subcellular compartments, such as the plasma membrane, nucleus, and endoplasmic reticulum (Hothorn et al., 2009), however, its functions in these areas are unclear. Logically, one would hypothesize VTC is synthesizing polyP in these subcellular compartments

but testing this poses challenges. Additionally, it is possible VTC is imparting other cellular functions unrelated to polyP production at these locations given that it is responsible for maintenance of general vacuole functions.

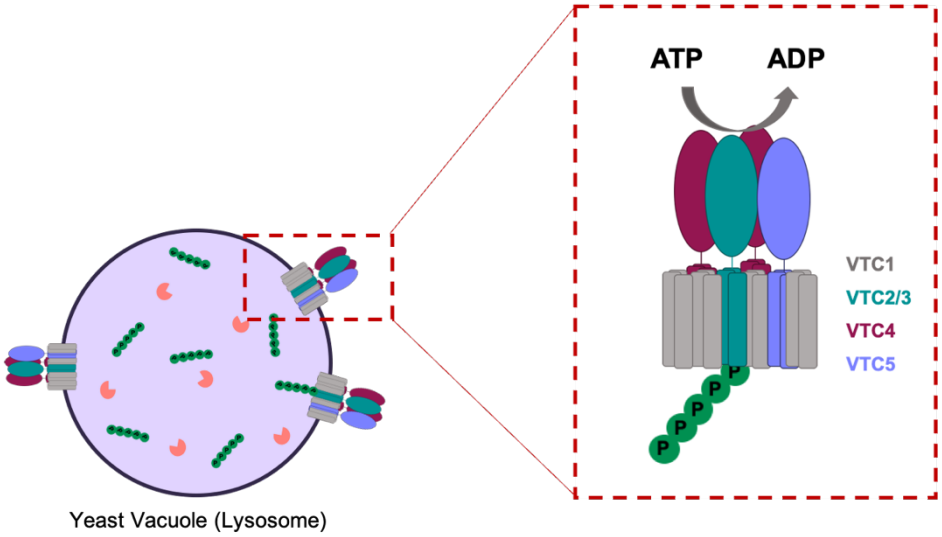


Fig. 4: PolyP is Synthesized by the VTC Complex in Yeast. The VTC complex is comprised of 3 necessary subunits: Vtc1, Vtc2 or Vtc3, and Vtc4 (the catalytic subunit). Vtc5 is a novel VTC subunit that acts to enhance the rate of polyP production by the VTC complex. This complex is predominantly localized at the vacuole membrane, where it synthesizes polyP chains as a translocation event into the vacuole lumen.

One subunit, Vtc2, is primarily found at the plasma membrane and nucleus but relocalizes to the vacuole membrane under phosphate starvation (Hothorn et al., 2009). The implications of this relocalization is not understood (see section 1.8). Recently, a novel VTC subunit, termed Vtc5, was found to act as a regulatory subunit where it enhances the rate of polyP production by the VTC complex through its interaction with the complex at the vacuole membrane (Desfougeres et al., 2016). In terms of polyP degradation, there are multiple polyphosphatases that have been characterized in yeast, each with different biochemical requirements and subcellular localizations. Notably, those that are responsible for polyP degradation are: Ppn1, Ppn2, Ppx1, and Ddp1 (Andreeva et al., 2019; Andreeva et al., 2015; Gerasimaite and Mayer, 2017; Lichko et al., 2006; Lonetti et al., 2011). Overall, yeast is a great

resource to study polyP biology given our extensive knowledge of basic polyP biology in this organism. Importantly, our knowledge of how polyP is regulated *in vivo* largely comes from studies done in yeast, which will be elaborated on in section 1.8.

1.6: PolyP is implicated in many aspects of human health

Research over the past century has demonstrated that polyP is very versatile and plays many critical roles in areas related to human health and medicine.

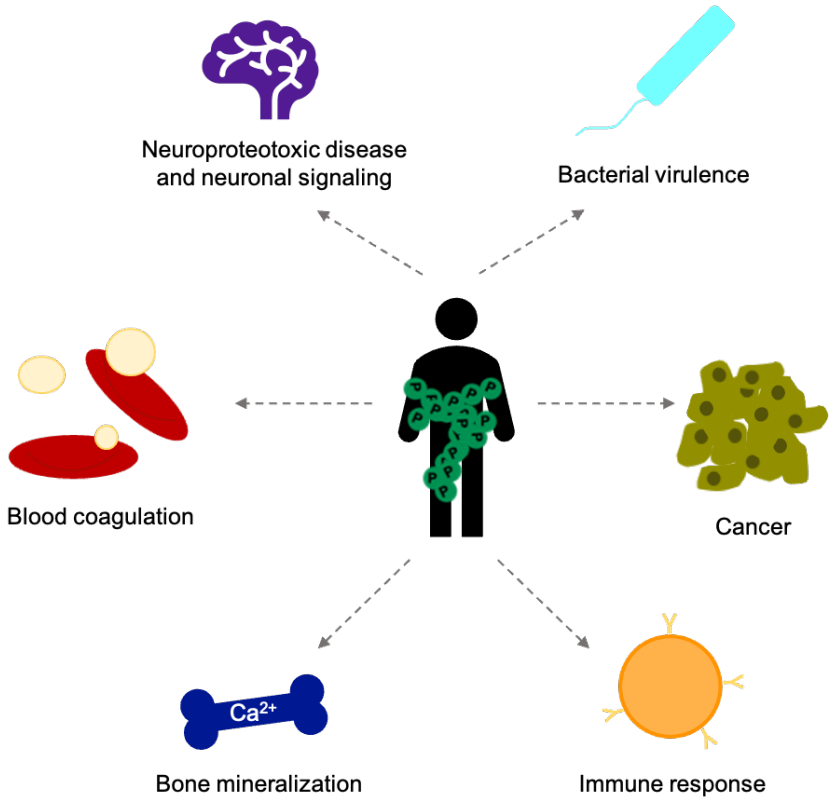


Fig. 5: PolyP is Implicated in Many Aspects of Human Health. A few examples where polyP has been implicated in human health are: neuroproteotoxic disease and neuronal signaling (Cremers et al., 2016; Holmstrom et al., 2013; Stotz et al., 2014), bacterial virulence (Kim et al., 2002; Peng et al., 2016; Rashid and Kornberg, 2000; Rashid et al., 2000a; Rashid et al., 2000b; Shiba et al., 1997), cancer (Arelaki et al., 2018; Boyineni et al., 2020; Jimenez-Nunez et al., 2012), immune response (Bae et al., 2012; Hernandez-Ruiz et al., 2006; Wat et al., 2014), bone mineralization (Hacchou et al., 2007; Hoac et al., 2013; Usui et al., 2010), and blood coagulation (Moreno and Docampo, 2013; Morrissey et al., 2012; Ruiz et al., 2004).

Three fascinating examples of polyP's function in maintaining human health are: blood coagulation, neuroproteotoxic disease, and bacterial virulence. As previously mentioned in 1.1, polyP plays critical roles at multiple parts of the blood coagulation cascade. Since polyP is found in platelet dense granules at the highest concentration in mammals, its role in promoting blood coagulation has been extensively studied (Ruiz et al., 2004). PolyP appears to promote coagulation at multiple different parts of the blood coagulation cascade, such as factor XII (Puy et al., 2013), factor XI/V (Choi et al., 2011, 2015), and thrombin (Choi et al., 2011; Mutch et al., 2010b; Smith et al., 2006). PolyP binds thrombin itself, perhaps acting as a chaperone (Mutch et al., 2010b). For a more extensive list of polyP's functions in the blood coagulation cascade see (Bentley-DeSousa and Downey, 2019) (included in **Appendix C**). PolyP is also associated with the reduction of toxic intermediates of aggregate-prone proteins such as β-amyloid, tau, and α-synuclein (Cremers et al., 2016). Here polyP is capable of accelerating the rate of mature fibril formation, which reduces the amount of toxic aggregate intermediates that are formed (Cremers et al., 2016). This aspect of polyP is be important in the development of neurodegenerative diseases associated with the accumulation of protein aggregates, such as Parkinson's disease, and could be considered as a therapeutic intervention (Ross and Poirier, 2004). Finally, it is possible the interplay between mammalian polyP and the production of polyP by resident lower-order organisms in the body is essential in maintaining proper health. In terms of the microbiome, there are interesting links between polyP production and the development of disease. In fact, the drug mesalamine that is currently used to treat Ulcerative Colitis and Inflammatory Bowel Disease is capable of reducing the amount of polyP produced by bacteria commonly found in the digestive tract through inhibition of polyphosphate kinase (PPK) enzymes (Dahl et al., 2017). Additionally, polyP accumulation in bacteria can increase their virulence, primarily by altering transcriptional profiles to enhance the expression of virulence factors and by increasing the ability for bacteria to form biofilms which may result in persistent

infection (Peng et al., 2016; Rashid and Kornberg, 2000; Rashid et al., 2000a; Rashid et al., 2000b; Shiba et al., 1997). This is extremely relevant for patients who suffer from chronic illness, such as Cystic Fibrosis, where infections with bacteria such as *Pseudomonas aeruginosa* become life-threatening (Fraley et al., 2007; Rashid et al., 2000b). Overall, it is obvious that polyP is an attractive molecule for the development of therapeutics, but further research into the endogenous regulation of polyP in mammals is essential.

One current caveat that remains in studying polyP biology is that we do not know which enzymes manipulate polyP homeostasis in mammals (**Figure 3**). BLAST alignment searches of both DNA and protein sequences do not reveal good candidates for mammalian polyP synthetases. There are a few proposed mammalian polyphosphatases, (h-Prune, DIPP1, DIPP2, and DIPP3) (Lonetti et al., 2011; Tammenkoski et al., 2008), which appear to function on incredibly short chains *in vitro*, however, their function *in vivo* is yet to be elucidated. Although not specifically classified as a polyphosphatase, alkaline phosphatase (ALP) has been shown to have exopolyphosphatase activity (Lorenz and Schroder, 2001) and appears to impact polyP levels in the process of bone mineralization (Mikami et al., 2016). To date, there are no known polyphosphatases that function *in vivo* or on chains larger than ~30-units (Lonetti et al., 2011; Tammenkoski et al., 2008). Given these challenges, many researchers typically study polyP in mammalian cells by using *in vitro* methods or by adding synthetic polyP to growth media. The latter comes with its own caveats, given that this results in the external application of the polymer and it is not possible to determine how this polyP is entering the cell and if it truly recapitulates endogenous *in vivo* polyP. Other research groups make use of the exopolyphosphatase Ppx1 from *S. cerevisiae* as a means to sequester and degrade polyP in mammalian systems to reduce the amount of available polyP (Labberton et al., 2016; Wang et al., 2020). Given our lack of a known polyP synthetase, a system wherein polyP levels can be manipulated intracellularly is required to increase our ability to study polyP. Ultimately, the

largest remaining open question in the field is: What is/are the mammalian polyP synthetase(s)? Additionally, do the previously proposed polyphosphatases act *in vivo*?

1.7: Polyphosphorylation as a new function for polyP at the protein level

In 2015, a novel function for polyP chains was uncovered wherein these chains were found to act as a non-enzymatic covalent post-translational modification (PTM) on proteins (Azevedo et al., 2015). This particular modification occurs on lysine residues within poly-acidic, serine, and lysine (PASK) rich motifs (Azevedo et al., 2015). Biochemically, the phosphate from polyP chains attach to the epsilon ε-NH_3^+ group on lysine residue side chains via phosphoramidate bond formation. The only way to detect polyphosphorylation to date is via electrophoretic mobility shift of substrates through western blotting using NuPAGE gels (**Figure 6**) (Azevedo et al., 2015). Originally, two yeast polyphosphorylated targets (Nsr1 and Top1) were characterized. These targets demonstrate delayed migration through Bis-Tris NuPAGE gels in wild-type yeast strains when compared to *vtc4*Δ mutant yeast strains, which lack the ability to synthesize polyP (Azevedo et al., 2015). The authors demonstrated that polyphosphorylation is capable of disrupting Nsr1-Top1 protein interaction, altering their subcellular localization, and negatively impacting Top1 topoisomerase activity *in vitro* (Azevedo et al., 2015). They additionally demonstrated polyphosphorylation is regulated by amount of polyP and that Top1 is capable of being polyphosphorylated *in vitro* by polyP extracted from the nucleus (Azevedo et al., 2015).

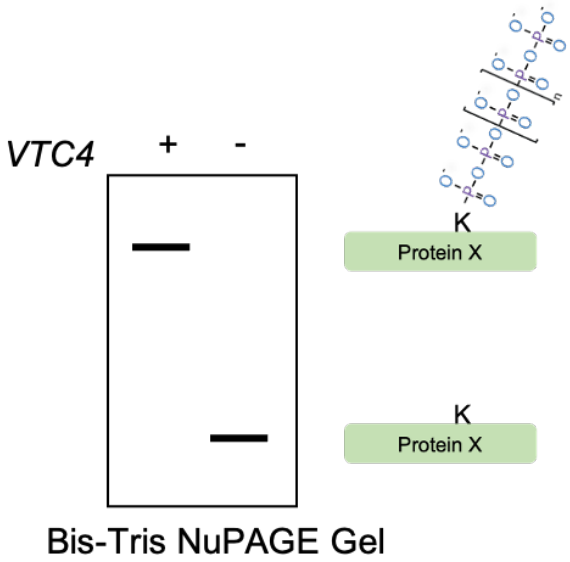


Fig. 6: Polyphosphorylation is Detected via Electrophoretic Mobility Shift on NuPAGE Gels. Protein extracts from wild-type (*VTC4*) and mutant (*vtc4*Δ) yeast strains are separated on 4-12% Bis-Tris NuPAGE gels. An electrophoretic mobility shift is indicative of protein polyphosphorylation, wherein the polyphosphorylated protein (in wild-type strains) will be at a higher molecular weight than the unmodified protein (in *vtc4*Δ mutant strains).

1.8: Intracellular regulation of polyP through localization and polyphosphatases

Since Nsr1 and Top1 are nuclear proteins, one of the areas that remains to be explored is how polyP is regulated intracellularly given that it is synthesized and constrained within the vacuole in yeast (Gerasimaite et al., 2014). PolyP is found in many subcellular compartments such as the nucleus, plasma membrane, mitochondria, and cytoplasm (Denoncourt and Downey, 2021; Gerasimaite and Mayer, 2016; Kornberg, 1999). However, these polyP concentrations are much lower, and the amounts of polyP detected in these compartments vary depending on the methods used in the study. Given this variability, it is difficult to deduce the exact amounts of polyP in these subcellular organelles. The VTC complex is localized to other subcellular compartments, however, if it produces polyP at these locations is unknown as previously mentioned. It is possible that polyP synthesized within the vacuole is being

transported to other areas in the cell, similar to the transport of lipids intracellularly. We know that in yeast cytoplasmic accumulation of polyP was shown to be toxic to the cell, resulting in decreased cell growth and abnormal cell morphology (Gerasimaite et al., 2014). In this case, the authors expressed the polyP synthetase from *E. coli* (*Ec*PPK1) in yeast, which produced high concentrations of long chain polyP in the cytoplasm. They demonstrated that synthesis of polyP must occur as a translocation event that requires the electrochemical gradient produced at the vacuole membrane (Gerasimaite et al., 2014). Whether these same toxicities exist when polyP is produced at high concentrations in organelles other than the vacuole remains to be explored. Not surprisingly, multiple polyphosphatases (Ppn1, Ppn2, Ppx1, and Ddp1; **Figure 3**) in yeast are localized throughout the cell as a mechanism to regulate polyP concentrations both within and outside of the vacuole. Ppn1 and Ppn2 are both polyphosphatases that are localized to the vacuole lumen and differ in their activities. Ppn1 possesses both exo- and endo-polyphosphatase activity whereas Ppn2 only possesses endopolyphosphatase activity (Andreeva et al., 2019; Andreeva et al., 2015; Gerasimaite and Mayer, 2017). Ppx1 is an exopolyphosphatase localized in the cytoplasm (Lichko et al., 2006; Lichko et al., 2003) and is critical for reducing polyP-induced cytotoxicity (Gerasimaite et al., 2014). Finally, Ddp1 is an endopolyphosphatase and its localization has not been extensively studied though it is proposed to be localized to the cytoplasm and the nucleus (Huh et al., 2003; Lonetti et al., 2011; Trilisenko et al., 2015). Biochemically, all of these enzymes are known to function in the presence of Zn^{2+}, Co^{2+}, and Mg^{2+} however their activities vary depending on the type and concentration of divalent cation (Andreeva et al., 2019). Interestingly, Ddp1 is also responsible for inositol pyrophosphate metabolism, acting to remove phosphates from IP7 and IP6 (Azevedo et al., 2015; Lonetti et al., 2011; Saiardi, 2012a, b). Given these roles, it is possible Ddp1 acts as a mediator to connect both inositol phosphate and polyP metabolism – two pathways that are intimately linked (Lonetti et al., 2011). Since the vacuole contains such a high concentration of

polyP in yeast (Auesukaree et al., 2004; Kornberg, 1999), regulation of polyP at the vacuole itself likely contributes to the largest changes in cellular polyP levels. The biggest known regulator of polyP production in yeast is the phosphate pathway. Intracellular overproduction of polyP through the VTC complex is known to activate the PHO phosphate starvation response pathway (Auesukaree et al., 2004; Desfougeres et al., 2016). Despite the plethora of polyP knowledge in yeast, there remains a gap in knowledge concerning how the VTC complex itself is regulated, which in turn, would affect polyP and phosphate regulation.

1.9: Trafficking pathways regulate protein localization and function

One way to regulate the metabolism of compounds within a cell is through regulation of the proteins/ complexes responsible for the compound's metabolism. Since the VTC complex localizes mostly to the vacuole (Desfougeres et al., 2016; Hothorn et al., 2009), altering VTC localization and subunit levels is likely to impact polyP production. In yeast, there are two pathways responsible for transporting proteins to the vacuole from the trans-golgi network (TGN) (**Figure 7**) (Feyder et al., 2015). The first pathway, named the CPY or indirect pathway, transports proteins from the TGN to the endosomal system. The endosome then fuses with the vacuole to deliver proteins. This pathway is well known for its delivery of soluble proteases, such as CPY (*Prc1*). The second pathway, named the AP-3 or direct pathway, directly transports proteins from the TGN to the vacuole through the cytoplasm (Feyder et al., 2015).

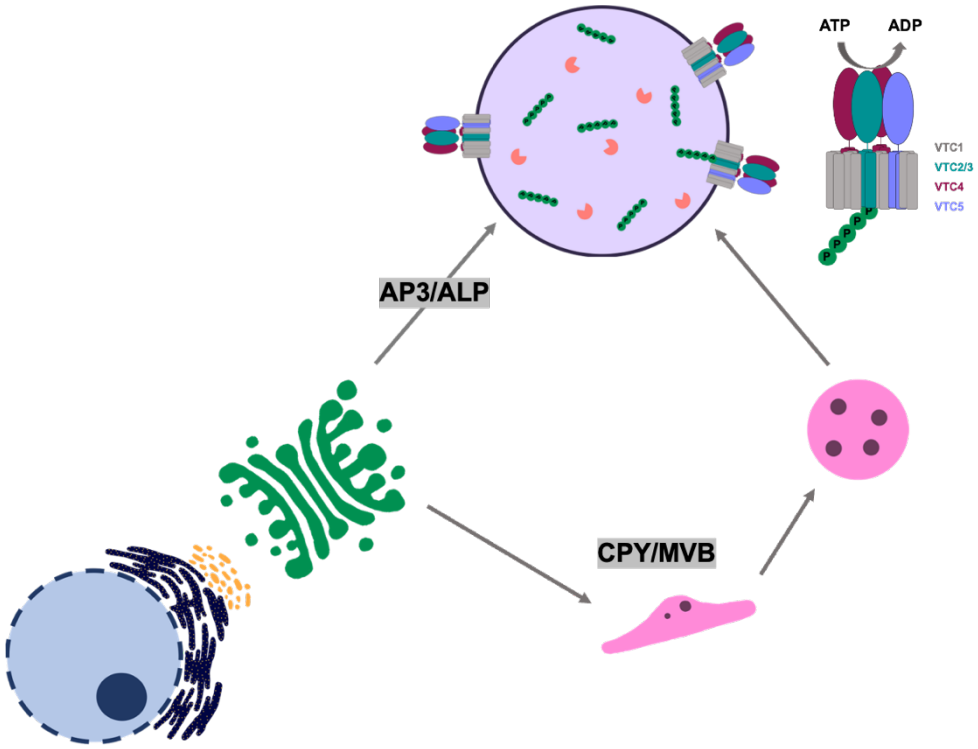


Fig. 7: Yeast Proteins are Localized from the TGN to the Vacuole Primarily Via the CPY and AP-3 Pathways. The CPY pathway results in cargo delivery to the vacuole indirectly, using the endosomal system. The AP-3 pathway allows for cargo delivery directly to the vacuole.

The AP-3 complex is a eukaryotic conserved heterotetrameric complex that includes two large subunit (*Apl5*/AP3D1 and *Apl6*/AP3B1), one medium subunit (*Apm3*/AP3M1), and one small subunit (*Aps3*/AP3S1) (Odorizzi et al., 1998). There already exist links between the AP-3 complex and polyP accumulation. In yeast, the AP-3 complex was found to be important for maintaining polyP levels in a genome-wide large-scale screen (Freimoser et al., 2006). Surprisingly, there are also links between this complex, a rare disorder, and polyP accumulation in mammals. Mutations in AP3δ1 and AP3β1 cause the development of a rare genetic disorder called Hermansky-Pudlak Syndrome (HPS) (Ammann et al., 2016; El-Chemaly and Young, 2016; Huizing et al., 2020; Mohammed et al., 2018; Wenham et al., 2010). This syndrome is characterized by a loss of the ability to properly synthesize lysosome-related organelles in specific cell types. Patients with HPS commonly have bleeding diathesis, albinism, and visual impairment (El-Chemaly and Young, 2016; Huizing et al., 2020). Depending on the HPS

subtype, patients may have additional symptoms specific to the gene/protein that is mutated. For example, patients with mutations in AP3δ1 often also display immunodeficiency and neurological dysfunction (Ammann et al., 2016; Mohammed et al., 2018). Biochemically, HPS patients are incapable of accumulating polyP in platelet dense granules, where polyP is found at the highest abundance in mammals (Muller et al., 2009). It is hypothesized that this characteristic is at least in part responsible for the bleeding diathesis HPS patients face.

1.10: Many pathways can be activated to degrade proteins

Another means to regulate polyP metabolism is through the degradation of the VTC complex. Depending on which cell signals are activated, there are multiple pathways the cell can use to degrade proteins such as: autophagy, ubiquitin-proteasome degradation, endosomal sorting complex required for transport (ESCRT) complex, and general proteases (**Figure 8**).

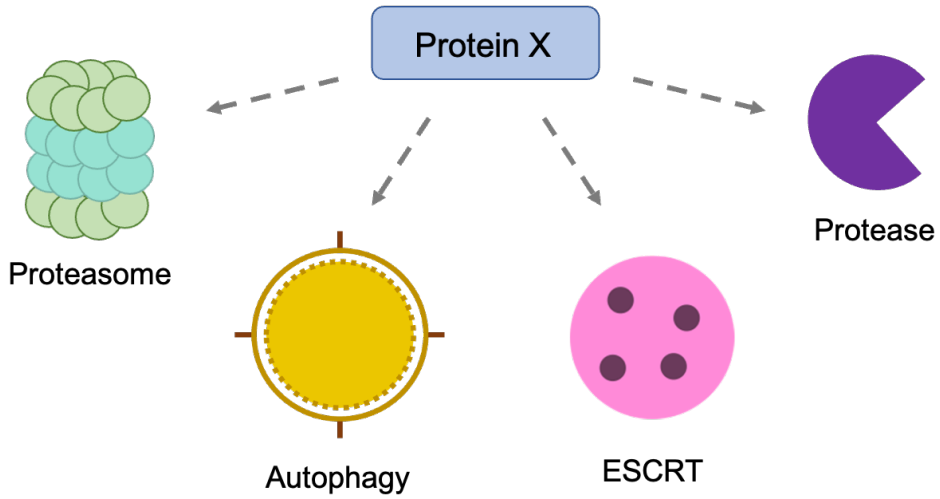


Fig. 8: Proteins Destined for Degradation can be Targeted by Many Pathways. A few examples of pathways are ubiquitin-dependent proteasome-mediated degradation, autophagy, ESCRT complex, and general protease activity.

Although there exist other pathways that degrade proteins endogenously, these four categories will be discussed as primary means for protein degradation in response to abnormal cellular events. Autophagy refers to the delivery of cytoplasmic materials to the

vacuole/lysosome for degradation and recycling of simple biomolecules, such as amino acids, within the cell. There are variations of the autophagic pathway, but the three main categories of are: macroautophagy, microautophagy, and chaperone-mediated autophagy (Glick et al., 2010). Macroautophagy involves the non-specific degradation of cytoplasmic material via the vacuole/ lysosome. It involves the creation of autophagosomes containing cytoplasmic material that end up fusing with vacuoles/lysosomes (Reggiori and Klionsky, 2013). Variations of microautophagy exist which consist of the targeted degradation of cargo via the invagination of membranes, as opposed to the direct fusion of membranes seen in macroautophagy (Reggiori and Klionsky, 2013). This pathway performs normal physiological roles, such as transporting proteins required for metabolism, organelle remodeling, and general quality control, however it can become upregulated under stress conditions or with drug treatment, for example with the use of the TOR pathway inhibitor rapamycin (Kamada et al., 2004; Sarkar et al., 2008). There are two major types of microautophagy, fission-type and fusion-type, that differ based on the specific proteins and machinery used (Reggiori and Klionsky, 2013; Schuck, 2020). Finally, chaperone-mediated autophagy requires chaperone proteins but results in the delivery of cytosolic proteins to the vacuole/lysosome by transporting them directly across the membrane (Dice, 2007; Tekirdag and Cuervo, 2018). This pathway contributes to quality control and is highly selective in nature. Although it is well characterized in higher-order organisms, this pathway is not characterized in yeast (Kirchner et al., 2019).

Both the ubiquitin-proteasome degradation and ESCRT-mediated degradation pathways are selective pathways that recognize ubiquitylated proteins and target them for degradation. Ubiquitylation of proteins occurs on lysine residues and involves a cascade of reactions with four basic steps (Ravid and Hochstrasser, 2008). Briefly, ubiquitin is activated by an E1 ubiquitin activating enzyme. Next, ubiquitin becomes conjugated to an E2 ubiquitin conjugating enzyme. Finally, E3 ubiquitin ligases recruit E2 conjugating enzymes to targets and catalyze the transfer

of the ubiquitin moiety to lysine residues of proteins (Ravid and Hochstrasser, 2008). Interestingly, ubiquitin can also form chains of different linkages (polyubiquitylation) which results in different outcomes. In the ubiquitin-proteasome pathway, K48-linked ubiquitylated proteins become localized to the proteasome, a large complex responsible for protein degradation (Swatek and Komander, 2016). On the other hand, K63-linked chains often play a variety of roles in cellular signaling pathways (Sun and Chen, 2004; Swatek and Komander, 2016). Other types of chain links exist alongside mixed chain linkages, all acting to increase the complexity and ultimately the function of proteins (Swatek and Komander, 2016). Both monoubiquitylation and chain-specific polyubiquitylation can affect many functions such as: protein activity, subcellular localization, and protein-protein interactions, among other functions (Sun and Chen, 2004). After proteins have been targeted, ubiquitin moieties become removed by proteins termed deubiquitinases (Komander et al., 2009).

The ESCRT pathway often recognizes mono- or poly-ubiquitylated transmembrane proteins for degradation. This pathway can be activated endogenously as a means to recycle proteins or in response to a cellular stressor. The ESCRT complex begins as a complex of two subunits (ESCRT-0) and matures through multiple variations (ESCRT-I, ESCRT-II, and ESCRT-III) to ultimately delivery cargoes through the creation of intralumenal vesicles (ILVs) into the multi-vesicular bodies (MVBs) of the endosomal system for delivery to the vacuole lumen (Henne et al., 2011; Schmidt and Teis, 2012a). Prior to the final delivery of ubiquitylated substrates, the deubiquitinase Doa4 (in concert with Bro1) (Luhtala and Odorizzi, 2004), removes ubiquitin groups in order to maintain cellular ubiquitin homeostasis. Alternatively, ESCRT may function in ubiquitin-independent fashions. The ESCRT pathway is also known to be involved in type 1 fission microautophagy by allowing for the invagination of the endosomal membranes (Schuck, 2020).

Finally, proteases are enzymes that cleave peptide bonds in proteins, allowing for proteins to become degraded into to amino acid components for recycling. Since the vacuole is a major site of protein degradation, there exist 7 proteases within this subcellular compartment (Hecht et al., 2014). The main protease Pep4 undergoes auto-activation where it is then responsible for the activation of other proteases within the vacuole. Pep4 is an aspartyl endoprotease and becomes activated upon delivery to the vacuole where it then activates Prb1, another major vacuolar protease (Hecht et al., 2014). Ultimately, this activation cascade results in the downstream activation of other proteases (Hecht et al., 2014). Overall, cells have multiple pathways it can utilize to degrade proteins and portions of organelles either in response to damage, a stimulus, or endogenously as a means to maintain proper cellular function. Although multiple aspects and pathways can function in collaboration with one another, which pathway is chosen depends on the substrate and how it is destined for degradation.

1.11: Open questions in the field of polyP biology

Although there are many aspects of polyP biology that require further investigation, two interesting areas that have yet to be explored in depth are the polyphosphorylation of proteins and the regulation of the VTC complex. As previously mentioned, polyphosphorylation has only been demonstrated on two yeast substrates and there are other modified proteins in other organisms is unknown. This poses many questions regarding the overall function of polyphosphorylation which primarily requires an expansion of known polyphosphorylated substrates. Additionally, how the VTC complex is regulated remains largely uncharacterized. Since the VTC complex is the only known polyP synthetase in yeast, understanding how it is regulated will not only expand our knowledge of polyP metabolism in yeast, but it may also provide insight into how the elusive polyP synthetase in mammals may be regulated. Ultimately, regulation of polyP synthetases and polyphosphatases will impact polyP accumulation and subsequently, may impact polyphosphorylation of substrates due to fluctuating levels of polyP.

1.12: Description of rationales and hypotheses

In this thesis, I present two manuscripts, which elaborate on the polyphosphorylation of proteins in both yeast and mammals. Finally, I present a third manuscript that assesses the regulation of polyP through the VTC complex with an emphasis on the Vtc5 subunit. Overall, these works contribute to the field of polyP biology by elaborating on novel roles for polyP and by exploring the regulation of the VTC complex.

1.12.1: A Candidate Screen for Novel Polyphosphorylated Proteins

1.12.1.1: Rationale

The original article describing protein polyphosphorylation in yeast only characterized two substrates, Nsr1 and Top1. In order to enhance our understanding of this PTM, it is necessary to expand the pool of polyphosphorylated substrates. Therefore, we aimed to perform a prioritized screen for polyphosphorylation of PASK-like proteins in yeast. Additionally, one question that remained was whether human proteins are capable of being polyphosphorylated. To this end, we aimed to create a novel system to study mammalian polyP biology, all while assessing the polyphosphorylation status of PASK-like proteins in humans.

1.12.1.2: Hypotheses

I hypothesized there are more polyphosphorylated yeast substrates apart from the currently known substrates Nsr1 and Top1. Additionally, I hypothesized mammalian proteins are polyphosphorylated given their homology to known yeast substrates.

1.12.2 A Follow Up Yeast Screen to Complete a PASK Protein-Wide Set of Substrates

1.12.2.1 Rationale

Previous efforts by our group demonstrated that proteins enriched in nuclear proteins with functions related to ribosome biogenesis are polyphosphorylated. These proteins were selected for screening based on the number and size of PASK motifs. Apart from this, there remained ~350 yeast proteins containing PASK-like sequences that remained to be screened. Thus, we chose to test whether these proteins were polyphosphorylated, using similar screening methods as in 1.12.1.

1.12.2.2 Hypotheses

I hypothesized other, previously un-screened, PASK are polyphosphorylated. Given the predicted localization and functions of the remaining ~350 proteins, I hypothesized substrates

that localize to other subcellular organelles are polyphosphorylated and this will provide insight into polyP's function in other cellular compartments.

1.12.3 Assessing the Role of Protein Trafficking Pathways in VTC Complex Regulation

1.12.3.1 Rationale

One open question in the field of yeast polyP biology is: how is the VTC complex regulated? A novel VTC subunit, termed Vtc5, was found to act as an accessory subunit to increase the rate of polyP production by the VTC complex. In order to answer this question, it is important to explore how the VTC complex is regulated. We have done so by examining how it is localized to the vacuole membrane and how it becomes degraded within the cell.

1.12.3.2 Hypotheses

I hypothesized Vtc5 is localized to the vacuole membrane by the AP-3 pathway in yeast, given the links between the AP-3 complex and polyP accumulation in yeast and mammals. Upon AP-3 mutation, Vtc5 becomes mislocalized within the cell and subsequently transported and/or degraded by one, or many, of the degradation pathways available in the cell.

2.0: Manuscript #1:

A Screen for Candidate Targets of Lysine Polyphosphorylation Uncovers a Conserved Network Implicated in Ribosome Biogenesis

Publication information:
Bentley-DeSousa, A., Holinier, C., Moteshareie, H., Tseng, Y-C., Kajjo, S., Nwosu, C., Amodeo, G.F., Bondy-Chorney, E., Sai, Y., Rudner A., Golshani, A., Davey, N.E., and Downey, M. (2018). A Screen for Candidate Targets of Lysine Polyphosphorylation Uncovers a Conserved Network Implicated in Ribosome Biogenesis. *Cell Reports* 22, 3427-3439.

Author's Contribution:
ABD and MD devised experiments and wrote the paper. ABD, HM, YCT, YS, CH, CN, AR, GFA, SK, EBC, NED, and MD carried out experiments and/or analysis. AG supervised HM. All authors reviewed the manuscript.

Contributions by figure: Figure 9 (ABD), Figure 10 (ABD, CH, MD), Figure 11 (ABD, MD), Figure 12 (ABD, MD), Figure 13 (ABD, MD), Figure 14 (ABD), Figure 15 (ABD, SK, YCT), Figure 16 (ABD, MD, YCT), Figure 17 (CN, MD, YCT), Figure 18 (HM, MD), Figure 19 (CH), Figure 20 (CH), Figure 21 (ABD), Table 1 (NED), Table 2 (NED).
MD contributed data analysis and/or preparation to all figures.

A Screen for Candidate Targets of Lysine Polyphosphorylation Uncovers a Conserved Network Implicated in Ribosome Biogenesis

Amanda Bentley-DeSousa,[1,4] Charlotte Holinier,[1,4,7] Houman Moteshareie,[3,4,7] Yi-Chieh Tseng,[1,4] Sam Kajjo,[2,4] Christine Nwosu,[1,4] Giuseppe Federico Amodeo,[6] Emma Bondy-Chorney,[1,4] Yuka Sai,[1,4] Adam Rudner,[2,4] Ashkan Golshani,[3,4] Norman E. Davey,[5] and Michael Downey[1,4,8,*]

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

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

[3]Department of Biology and Institute of Biochemistry, Carleton University, Ottawa, Ontario K1S 5B6, Canada

[4]Ottawa Institute of Systems Biology, Ottawa, Ontario K1H 8M5, Canada

[5]Conway Institute of Biomolecular & Biomedical Research & UCD School of Medicine & Medical Science, University College Dublin, Belfield, Dublin 4, Ireland

[6]Department of Basic Sciences, New York University, College of Dentistry, 345 East 24th Street, New York, NY 10010, USA

[7]These authors contributed equally

[8]Lead contact

*Correspondence: mdowne2@uottawa.ca

2.1: Abstract

Polyphosphates (polyP) are chains of inorganic phosphates found in all cells. Previous work has implicated these chains in diverse functions, but the mechanism of action is unclear. A recent study reports that polyP can be non-enzymatically and covalently attached to lysine residues on yeast proteins Nsr1 and Top1. One question emerging from this work is whether so-called ‘‘polyphosphorylation’’ is unique to these proteins or instead functions as a global regulator akin to other lysine post-translational modifications. Here, we present the results of a screen for polyphosphorylated proteins in yeast. We uncovered 15 targets including a conserved network of proteins functioning in ribosome biogenesis. Multiple genes contribute to polyphosphorylation of targets by regulating polyP synthesis, and disruption of this synthesis results in translation defects as measured by polysome profiling. Finally, we identify 6 human proteins that can be modified by polyP, highlighting the therapeutic potential of manipulating polyphosphorylation in vivo.

2.2: Introduction

Polyphosphates (polyP) are chains of inorganic phosphates, ranging from three to thousands of moieties in length. PolyP is considered to be an 'ancient' molecule and is found in diverse cell types from bacteria to human cells (Kornberg et al., 1999; Moreno and Docampo, 2013; Rao et al., 2009). All known synthesis of polyP chains occurs through the sequential generation of phosphoanhydride bonds by polyP polymerases, which add gamma-phosphates from ATP or other nucleoside triphosphates to growing chains. In bacteria, this is accomplished via the action of polyP kinases PPK1 or PPK2 (Akiyama et al., 1993; Tzeng and Kornberg, 2000; Zhang et al., 2002). In the budding yeast *S. cerevisiae*, polyP synthesis occurs via the action of Vtc4, which complexes to Vtc1 and either Vtc2 or Vtc3, to form the V-ATPase transporter chaperone (VTC) complex (Hothorn et al., 2009; Morrissey et al., 2012). Most VTC complexes are vacuolar bound, and the synthesis of polyP is coupled to its translocation into the vacuole (Gerasimaite et al., 2014; Hothorn et al., 2009). Although the vast majority of polyP is constrained in the vacuole, significant levels have been detected in other cellular compartments. However, the estimated size of these non-vacuolar polyP pools varies considerably from study to study (Azevedo et al., 2015; Gerasimaite et al., 2014; Lichko et al., 2006; Rao et al., 2009; Urech et al., 1978).

In yeast, the hydrolysis of polyP chains by exopolyphosphatase Ppx1, and endopolyphosphatases Ppn1 and Ppn2, results in the generation of free inorganic phosphates or chains of smaller length. Ppx1 is thought to be mostly cytoplasmic (Wurst et al., 1995) whereas Ppn1 and Ppn2 localize predominately to the vacuole (Gerasimaite and Mayer, 2017; Kumble and Kornberg, 1996). In both bacteria and yeast, polyP has been proposed to function as an energy store, a phosphate reservoir, and a metal chelator (Moreno and Docampo, 2013). Phenotypically, bacteria lacking polyP show defects in biofilm formation, motility, and other

processes important for infection (Rashid et al., 2000a; Rashid et al., 2000b; Shi et al., 2004). In yeast, deletion of *VTC4* results in microautophagy defects (Uttenweiler et al., 2007) and reverses stress hypersensitivity of cells mutated for the cyclin-encoding gene PHO80 (Huang et al., 2002).

In contrast to lower eukaryotes, mammalian polyP synthesis and degradation machineries are uncharacterized (Hooley et al., 2008). One exception is h-prune, originally characterized as a regulator of cell migration, which also acts as a short-chain exopolyphosphatase *in vitro* (Carotenuto et al., 2014; Tammenkoski et al., 2008). Whether h-prune acts on polyP chains *in vivo* is unknown. Hence, many studies using cell lines have relied on the ectopic expression of yeast Ppx1 or the addition of synthetic polyP to culture media to alter polyP levels in vivo (Rao et al., 2009). In mammals, polyP chains have been implicated as regulators of the blood coagulation cascade (Smith et al., 2010; Smith et al., 2006; Yeon et al., 2017), neuroprotection (Cremers et al., 2016), and cell growth (Jimenez-Nunez et al., 2012; Rao et al., 2009; Shiba et al., 2003; Wang et al., 2003). Given its diverse roles across multiple types of cells, modulation of polyP levels is suggested to have therapeutic potential (Labberton et al., 2016; Moreno and Docampo, 2013). However, the mechanisms by which polyP chains exert their molecular functions are poorly understood.

Azevedo *et al.* reported that polyP chains can be covalently attached to lysine residues on yeast proteins Nsr1 and Top1 as a non-enzymatic post-translational modification (PTM) (Azevedo et al., 2015). Polyphosphorylation of these proteins confers a dramatic decrease in their mobility on NuPAGE gels (Azevedo et al., 2015). The resulting electrophoretic ‘‘shift’’ is observed for Nsr1 and Top1 purified from wild-type cells but is absent when these proteins are purified from *vtc4*Δ mutants, which lack polyP. Modified lysine residues for both Nsr1 and Top1 proteins were mapped to poly-acidic, serine, and lysine-rich (PASK) motifs of several dozen amino acids in length (Azevedo et al., 2015). Polyphosphorylation impacts binding of Nsr1 to Top1, the *in vitro*

topoisomerase activity of Top1, and the subcellular localization of both proteins (Azevedo et al., 2015). Whether additional targets of lysine polyphosphorylation exist and whether this new mechanism of protein modification is conserved in other systems are unknown.

Here, we describe a screen for lysine polyphosphorylation targets in yeast. We identified 15 additional targets, including an evolutionarily conserved network involved in ribosome biogenesis. Based on these results, we uncovered an unreported role for the Vtc4 polyP polymerase in promoting ribosome function. We also report that lysine polyphosphorylation is controlled by proteins involved in vacuolar biology and that bypassing this control mechanism by synthesizing polyP directly in the cytoplasm allows for target modification but renders cells sensitive to drugs that impact protein translation and cell growth. Finally, we show that human cells can be engineered to produce polyP chains that modify human PASK-containing proteins. Our work recasts polyphosphorylation of lysine as a global PTM with the potential to regulate diverse cellular processes.

2.3 Results

Additional Targets of Lysine Polyphosphorylation in Yeast

To search for polyphosphorylated proteins in yeast, we generated a list of proteins that contain PASK-like sequences similar to those found in the two known targets, Nsr1 and Top1. PASK motifs in these proteins contain stretches of glutamic acid (E) and aspartic acid (D) interspersed with serine (S) and lysine (K) residues (Azevedo et al., 2015). Notably, the serine residues of Nsr1 and Top1 PASK sequences are not required for their modification with polyP (Azevedo et al., 2015). Nevertheless, we chose to retain serine as a PASK feature for our analyses, since it may be important for regulation under specific circumstances. We identified 427 proteins (out of ~6,000 total proteins in yeast) with one or more 20-amino-acid stretches made up of at least 75% D/E/S with at least one K (**Figure 9A; Table 1, Appendix A**). Gene ontology (GO)-term analysis of this set of proteins using DAVID (Sherman et al., 2007) indicated enrichment in functions related to ribosome biogenesis and rRNA processing (**Figure 9B**). Consistent with this, we also observed enrichment for nuclear and nucleolar subcellular localizations (**Figure 9C**). We chose 90 of these PASK-containing proteins for analysis based on the length and number of PASK motifs (**Table 1, Appendix A**), with priority given to proteins with long or multiple PASK motifs. Our goal was to use NuPAGE analysis to measure the electrophoretic shift of each candidate in the presence and absence of VTC4, as was previously done for Nsr1 and Top1(Azevedo et al., 2015). To do this, we took advantage of the yeast GFP-tagged collection wherein each open reading frame is expressed from its genomic locus as a fusion with GFP (Huh et al., 2003).

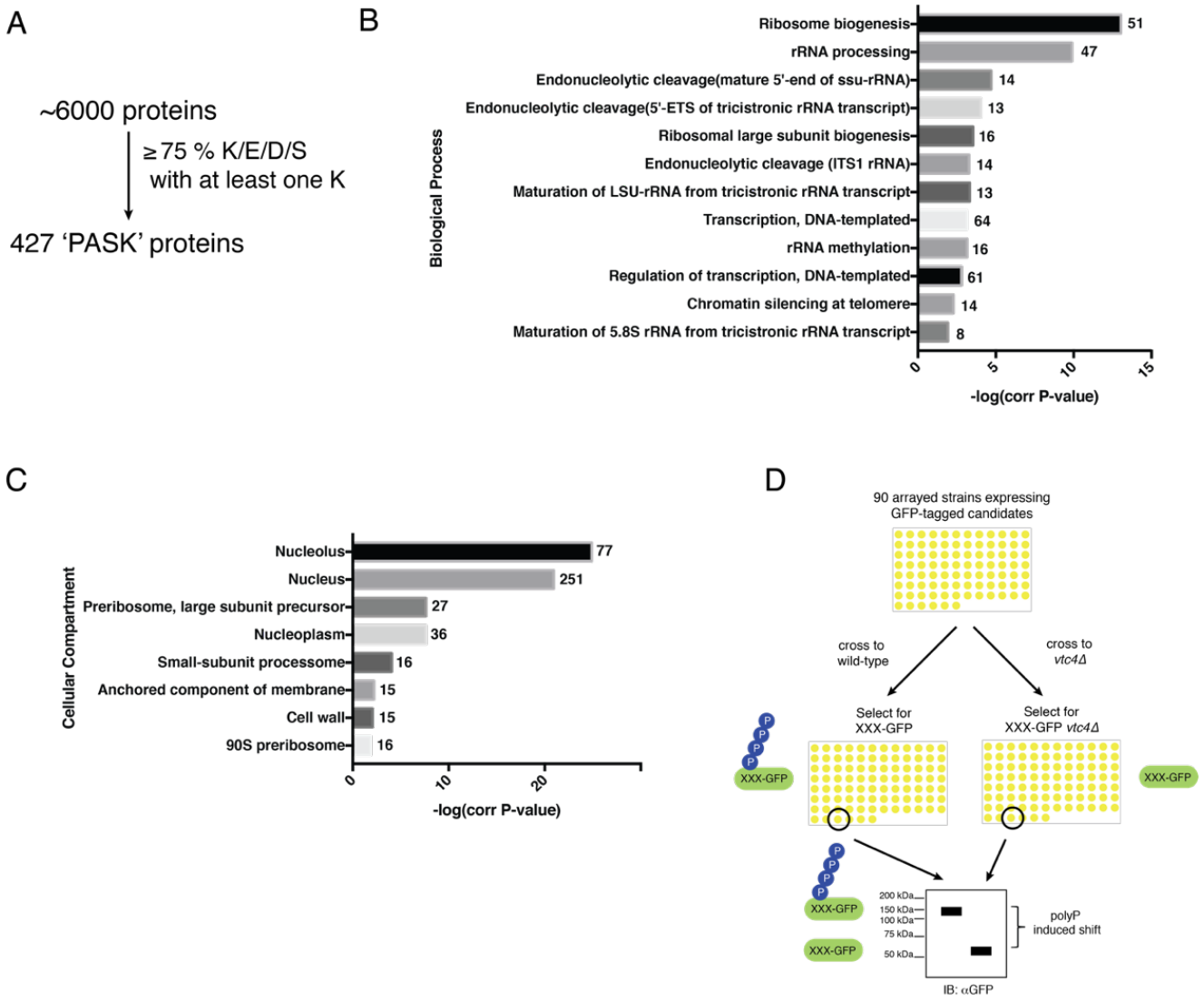


Fig. 9: Analysis of PASK-Containing Proteins in Yeast. A) Parameters used to identify PASK-like motifs in *S. cerevisiae* proteins. B) PASK-containing proteins function in ribosome biology. The number of proteins in each category is shown. C) PASK-containing proteins are enriched for nucleolar localization. The number of proteins in each category is shown. Bioinformatics analyses were carried out using DAVID with default parameters. D) A screen for polyP targets. See Experimental Procedures for details. Also see **Table 1 (Appendix A)**.

Starting with strains pulled from this set, we used high-throughput yeast mating and selection techniques to generate isogenic *VTC4* and *vtc4*Δ strains expressing our 90 prospective targets as GFP fusions (**Figure 9D**, also see Experimental Procedures). For each GFP-candidate fusion, paired protein extracts from wild-type and *vtc4*Δ mutants were analyzed using NuPAGE

followed by western blotting and detection of tagged candidates with an antibody against GFP (**Figure 9D**). Of 90 candidates, 15 fusion proteins displayed *VTC4*-dependent electrophoretic shifts characteristic of lysine polyphosphorylation (**Figures 10A, 9B**, and **11A–11C**). Not all GFP-fusions displayed altered mobility in *vtc4*Δ strains compared to wild-type controls, demonstrating that the GFP tag was not being polyphosphorylated (**Figure 11C**).

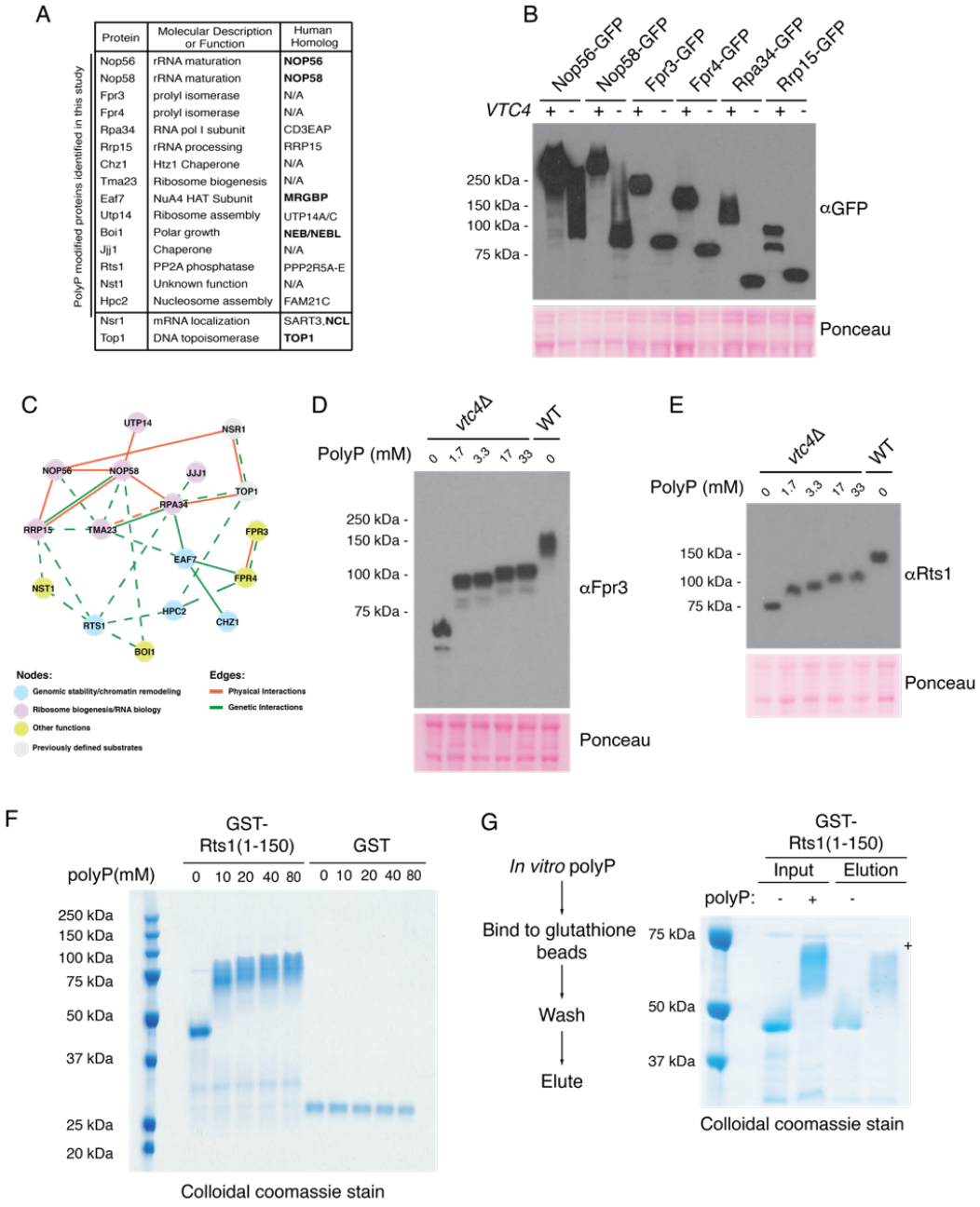


Fig. 10: Identification of Polyphosphorylated Proteins. A) 15 additional polyphosphorylated proteins were uncovered. Indicated in this table are known polyphosphorylated proteins in

yeast. For human homologs, bolded text indicates conservation of PASK-like motifs. See Experimental Procedures. B) Polyphosphorylation causes electrophoretic shifts on NuPAGE gels. Extracts from *VTC4* or *vtc4*Δ strains expressing the indicated GFP-fusion proteins were run on NuPAGE gels and transferred to PVDF membrane. Proteins were detected using an antibody against the GFP epitope. C) PolyP targets are physically and genetically connected. Interactions were determined using the Genemania tool (http://genemania.org/) (Zuberi et al., 2013). D-E) Modification of targets with synthetic polyP. Extracts from the indicated strains were incubated with increasing concentrations of synthetic polyP with a modal length of 75 residues. Fpr3 (D) or Rts1 (E) were detected with antibodies against native proteins following NuPAGE analysis and western blotting. F) *In vitro* polyphosphorylation of a bacterial-purified fragment of Rts1 fused to GST. G) GST-Rts1(1-150), GST, or buffer was incubated with biotinylated polyP. Biotinylated polyP was detected with Strep-horseradish peroxidase (HRP) after NuPAGE and western blotting. Also see **Figure 11**.

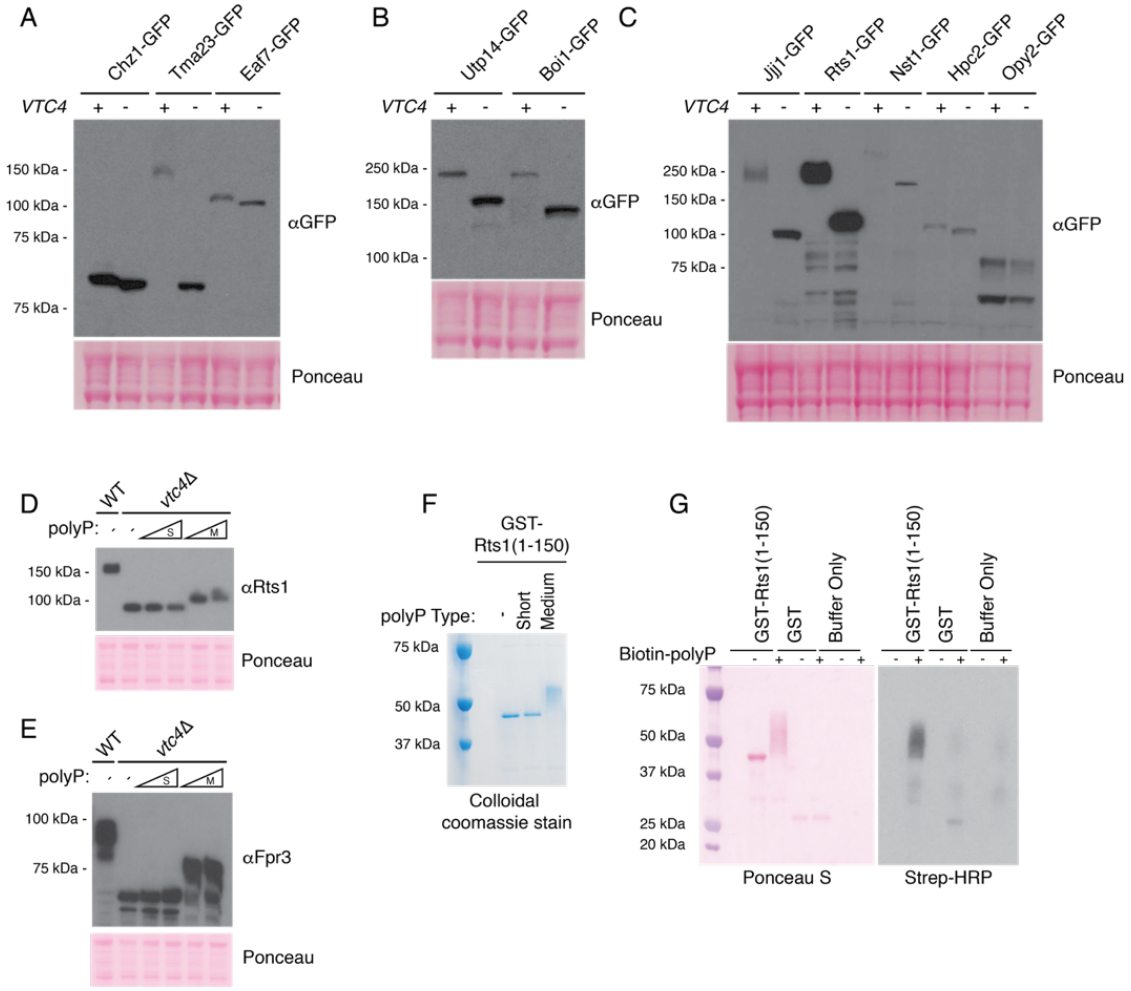


Fig. 11: Identification of Polyphosphorylated Candidates and *In Vitro* Analysis. A-C) Identification of polyphosphorylated candidates. Proteins from the indicated strains were prepared using TCA-style lysis and separated using NuPAGE prior to western blotting and

detection with αGFP antibody and Ponceau stain (lower). Opy2-GFP was found to be unaffected by *vtc4*Δ in our screen and is included here as a negative control. D-E) Short (14 units) or medium (60 units) chain polyphosphate were used in *in vitro* assays at a concentration of 0.415 and 0.83 mM. F) Polyphosphorylated GST-Rts1(1-150) was captured with glutathione beads, and washed prior to analysis by NuPAGE and colloidal blue staining.

Many of the polyphosphorylated targets that we identified have functions related to ribosome biogenesis (**Figure 10A**). For example, Nop56, Nop58, and Utp14 are members of the small subunit processome (SSU) and are required for 18S rRNA processing and modification (Gautier et al., 1997; Lafontaine and Tollervey, 2000; Oruganti et al., 2007; Zhu et al., 2016). Rrp15 is a constituent of pre-60S ribosomal particles and is required for the processing of 27S rRNA (De Marchis et al., 2005). Rpa34 functions upstream of ribosomal assembly as a regulatory subunit of RNA pol I, which is required for transcription of 37S pre-rRNA from rDNA repeats (Werner et al., 2009). Other targets uncovered in our screen do not function directly in ribosome biogenesis but are part of a larger network connected by both genetic and physical interactions that also includes the previously identified targets Nsr1 and Top1 (**Figure 10C**). For example, Rts1 is a regulatory subunit of the PP2A subcomplex that has been linked to the control of cell size (Artiles et al., 2009). Targets Fpr3 and Fpr4 are prolyl isomerases and FK506 binding proteins involved in nucleosome assembly and transcriptional regulation (Benton et al., 1994; Park et al., 2014). Other chromatin regulators recovered in our screen include Eaf7, a member of the NuA4 histone-acetyltransferase complex (Mitchell et al., 2013; Rossetto et al., 2014); Chz1, a chaperone for the Htz1 histone variant (Luk et al., 2007); and Hpc2, a member of the HIR complex that contributes to nucleosome remodeling by recruiting the SWI/SNF complex to DNA (Balaji et al., 2009). Many of these targets have one or more human homologs (**Figure 10A**), which will be discussed later. Our ability to identify 15 targets from a prioritized pool of just 90 proteins (**Table 1, Appendix A**) suggests that polyphosphorylation is likely to be a widespread

PTM.

Molecular Characterization of Polyphosphorylation

In support of non-enzymatic polyphosphorylation being responsible for the shifts observed in our experiments, we were able to largely restore shifts to targets Rts1 and Fpr3 with the addition of synthetic polyP to boiled *vtc4*Δ extracts (**Figures 10D and 10E**). In these experiments, Rts1 and Fpr3 are detected with antibodies directed against native proteins. Our inability to completely restore shifts may be due to differences in chain length between polyP attached in vivo versus the 75-unit synthetic polyP used in our experiments. Indeed, a similar experiment using synthetic polyP of different lengths showed that electrophoretic shift is chain-length dependent for these targets (**Figures 11D and 11E**). To gain further support for polyP being the agent responsible for electrophoretic shifts, we carried out in vitro polyphosphorylation assays with bacterially purified GST-Rts1(amino acids 1–150), which contains the 2 predicted PASK sequences for Rts1. GST-Rts1(1–150), but not GST alone, shifted on NuPAGE gels following incubation with a range of synthetic polyP concentrations (**Figure 10F**). In support of the idea that polyP was directly responsible for the change in electrophoretic mobility, synthetic biotinylated polyP was observed to co-migrate with the purified GST-Rts1(1–150) (**Figure 10G**). Finally, the GST-Rts1(1–150) shift was retained when excess polyP was washed away following capture of the fusion protein on glutathione beads (**Figure 11F**). We conclude that polyP can directly modify Rts1 and suggest that this is also true for our other targets.

Polyphosphorylation of Targets Requires PASK Motifs

In order to test the contribution of the PASK motifs to target polyphosphorylation, we focused on 5 targets with C-terminal PASK motifs (Nop56, Nop58, Tma23, Rpa34, and Chz1) that could be easily deleted using homologous recombination in yeast. Deletion of PASK motifs from Nop56,

Nop58, Tma23, and Rpa34 resulted in collapse of the *VTC4*-dependent shift observed with NuPAGE analysis (**Figures 12A–12C**). These data are consistent with the PASK motif being the site of modification for these four proteins. For Rpa34, we found that mutation of PASK lysine residues to arginines largely prevented electrophoretic mobility shifts observed in wild-type strains (**Figures 12D and 12E**) while leaving Rts1 polyphosphorylation unaffected (**Figure 13A**).

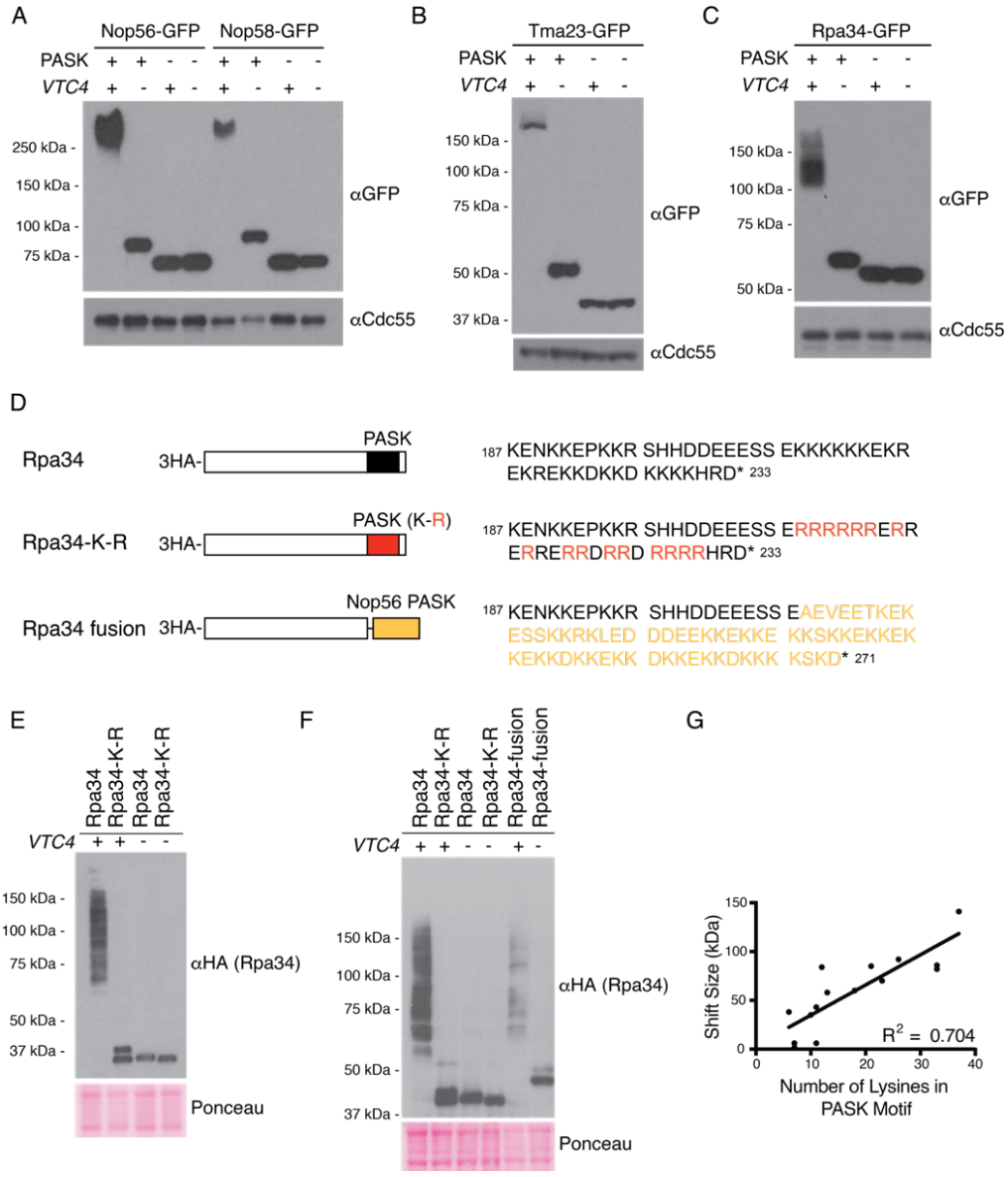


Fig. 12: Molecular Characterization of Polyphosphorylation. A-C) PASK motifs are required for target modification. GFP fusions with Nop56 and Nop58 (A), Tma23 (B), or Rpa34 (C) were detected with an antibody against GFP following NuPAGE and western blotting. D-F) Rpa34

constructs shown in (D) were analyzed via NuPAGE (E and F). G) Shift size in NuPAGE analysis correlates with lysine content of target PASK motifs. Also see **Figures 13-15**.

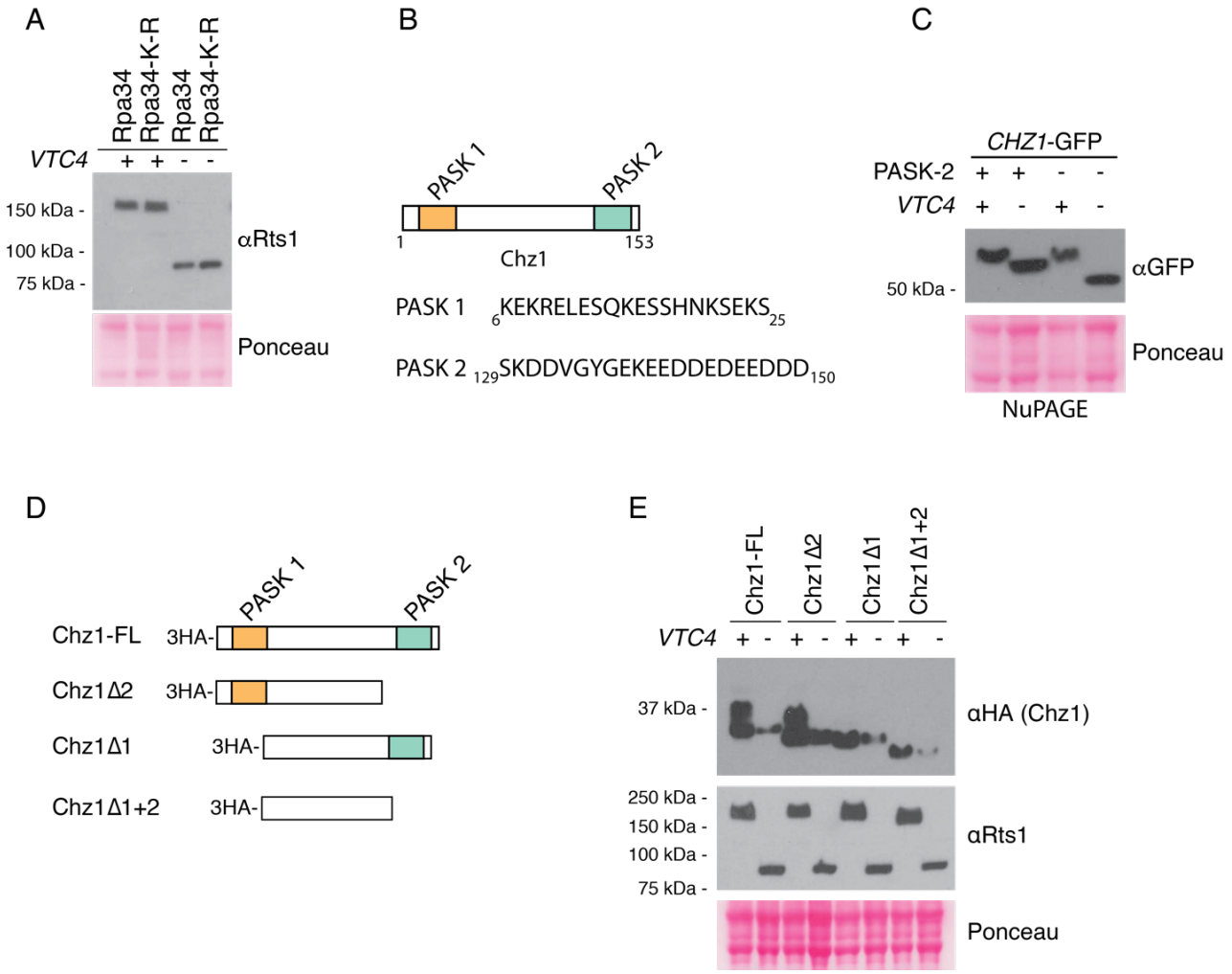


Fig. 13: Characterization of Rpa34 and Chz1 PASK Motifs. A) Rts1 shift is not affected in Rpa34 PASK mutants. Strains were analyzed via NuPAGE analysis. The same samples were used in Figure 11D. B) Schematic representation showing the positions of the two PASK motifs in the Chz1 protein. C) The C-terminal PASK motif of Chz1 is not the region of polyP attachment. Total proteins from the indicated strains were extracting using a 'polyP' lysis protocol and separated using 4-12% NuPAGE. D) Chz1 deletion constructs expressed under the GAL promoter. E) The indicated strains expressing the Chz1 deletion constructs described in (D) were analyzed via NuPAGE and probed with the indicated antibodies following western blotting.

Moreover, a synthetic construct wherein Rpa34 lacking its PASK motif was fused to the PASK from Nop56 retained wild-type levels of polyphosphorylation, confirming that PASK motifs are portable (**Figure 12F**). In contrast to other deletion mutants analyzed, the net change in electrophoretic mobility for Chz1-GFP (wild-type versus *vtc4*Δ) was not significantly impacted in the absence of its C-terminal PASK when analyzed via NuPAGE (**Figures 13B and 13C**). Notably, Chz1 has a second PASK motif in its N terminus that could serve as the site of modification (**Figure 13B**). To test this idea, we expressed a series of 3HA-tagged Chz1 constructs from the GAL promoter in wild-type and *vtc4*Δ cells (**Figure 13D**). Consistent with our observations with Chz1-GFP, NuPAGE analysis revealed electrophoretic shifts for 3HA-Chz1 in wild-type cells, and this shift was maintained in constructs lacking the C-terminal PASK (**Figure 13E**). In contrast 3HA-Chz1 lacking the N-terminal PASK motif or both PASK motifs migrated identically in extracts prepared from both wild-type and *vtc4*Δ strains (**Figure 13E**). Altogether, these experiments indicate that Chz1 is polyphosphorylated in its N terminus.

In our screen and follow-up experiments, we observed that the size of the polyP-induced shift varied dramatically depending on the target in question. The apparent molecular weight of Nop58-GFP from wild-type cells was almost 150 kDa greater than that of Nop58-GFP from *vtc4*Δ cells (**Figure 10B**). In contrast, the shifts measured for proteins Chz1 and Hpc2 were consistently smaller (**Figures 11A and 11C**). We analyzed target PASK motifs and found that the shift size for each target correlated strongly with the number of lysine residues present within these regions (**Figure 12G**). In contrast, we observed no correlation with the serine or acidic amino acid content (**Figures 14A and 14B**).

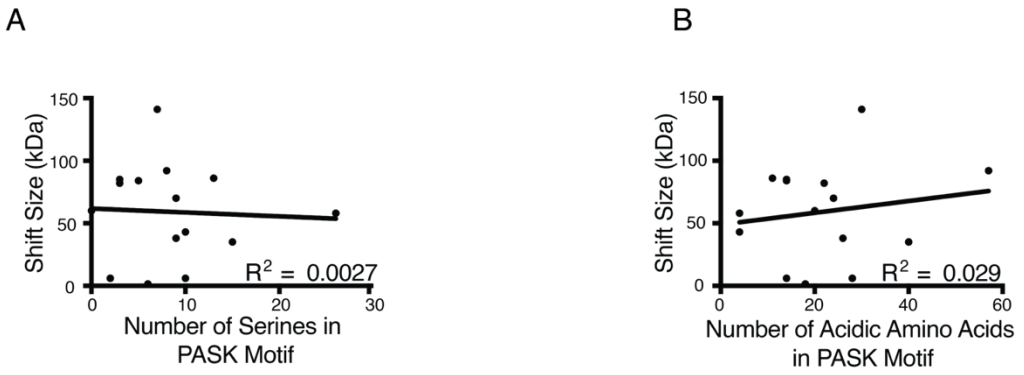


Fig. 14: PASK Serine and Acidic Amino Acid Content Do Not Impact Shift Size. A-B) PASK serine or acidic amino acid content (aspartic or glutamic acid) does not correlated with shift size.

Notably, an electrophoretic shift associated with canonical serine/threonine phosphorylation was not at all impacted by *vtc4*Δ mutation (**Figure 15A**). These data are consistent with lysine, rather than serine, being the site of polyP attachment. Presumably, the size of the shift is partially dependent on the number of occupied sites, with greater occupancy yielding a greater change in apparent molecular weight. Throughout this study, we observed that in contrast to NuPAGE gels, polyP-induced shifts are not resolved on SDS-PAGE gels (**Figure 15B**). According to the manufacturer, NuPAGE gels, like other Bis-Tris systems, are polymerized in the absence of tetramethylethylenediamine (TEMED). To test whether the absence of TEMED was critical for resolving polyP-induced shifts, we made our own Bis-Tris gels with or without TEMED (see Experimental Procedures). We found that addition of TEMED to Bis-Tris gels largely collapsed the shift associated with polyphosphorylated GST-Rts1(1–150) (**Figure 15C**). While the mechanism behind this collapse is unclear, investigation of interaction between TEMED and polyP may provide insight into the stability of polyphosphorylated lysine.

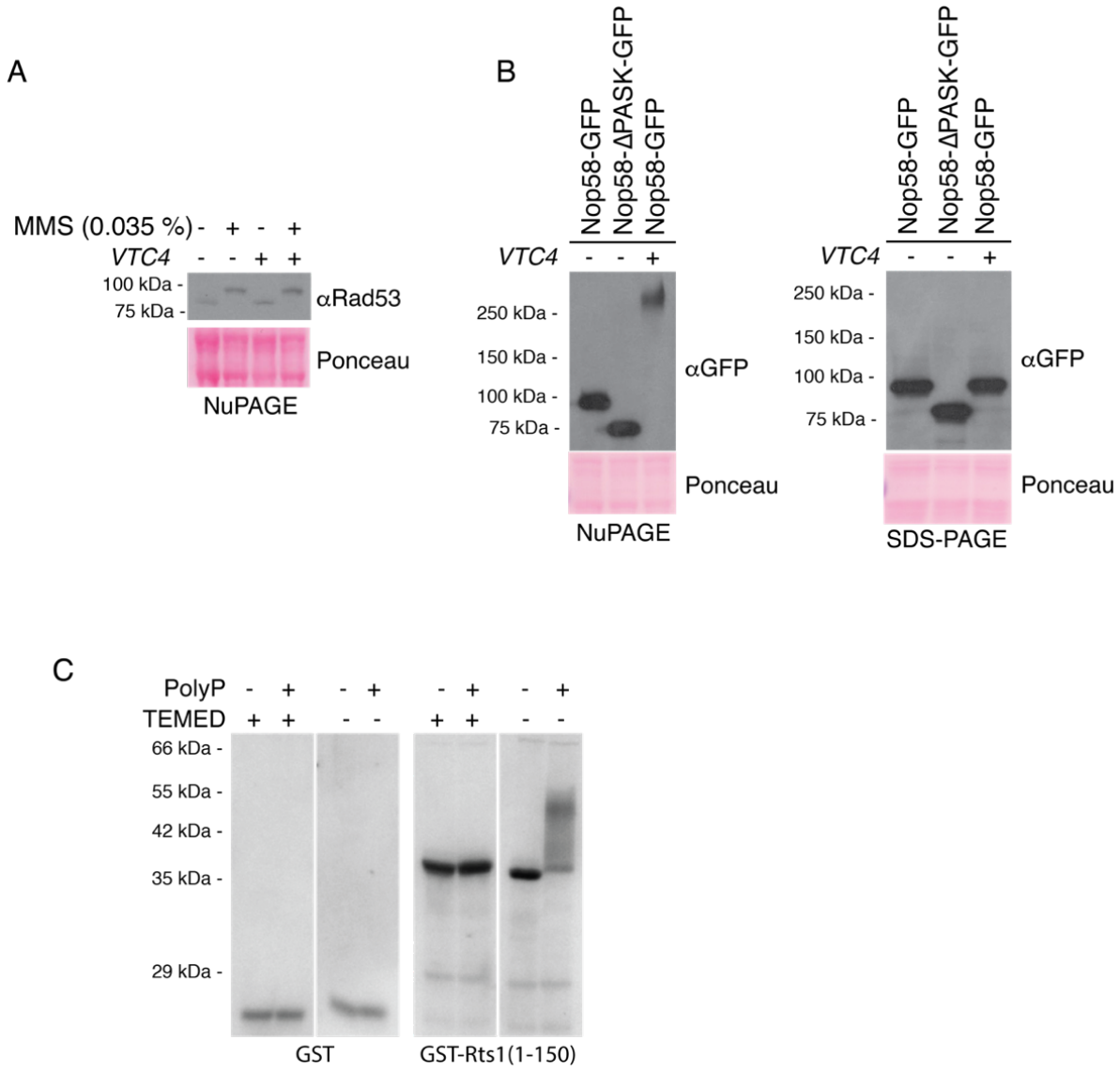


Fig. 15: Analysis of polyphosphate induced electrophoretic shifts. A) Vtc4 does not impact electrophoretic shift of Rad53 following DNA damage. Extracts from the indicated strains were run on Bis-Tris NuPAGE gels prior to transfer to PVDF membrane and analysis with an antibody against Rad53. B) Polyphosphate-induced shifts are not resolved on SDS-PAGE gels. Extracts from the indicated strains were prepared using TCA lysis and analyzed vis NuPAGE or SDS-PAGE and transferred to PVDF membrane prior to detection of Nop58-GFP constructs with an antibody against GFP. The ΔPASK strain is shown to demonstrate that conditions used for SDS-PAGE can resolve differences in molecular weight in a manner similar to NuPAGE. The only difference is in the ability to detect polyphosphorylation. C) GST or GST-Rts1(1-150) was analyzed following *in vitro* polyphosphorylation (10 mM) on homemade Bis-Tris gels with or without TEMED.

Regulation of Polyphosphorylation by polyP Synthesis

PTMs play important roles in cell signaling because they are often regulated at the step of their attachment to protein targets by specific enzymes. In contrast, lysine polyphosphorylation is non-enzymatic (Azevedo et al., 2015). Therefore, we focused on the initial synthesis of polyP chains and the targeted action of polyphosphatase enzymes as mechanisms of regulation. First, Freimöser *et al.* identified 254 genes besides *VTC4* that contribute to the synthesis or maintenance of polyP pools in yeast (Freimoser et al., 2006). We analyzed target polyphosphorylation in 4 of these mutants that contain polyP levels 10-fold lower than wild-type cells. Deletion of *VTC1*, *VMA22*, and *VAM3* resulted in a loss of Rts1 and Nsr1 polyphosphorylation comparable to that observed in *vtc4*Δ cells (**Figures 16A and 17A**).

A subtle but reproducible defect in the polyphosphorylation of both proteins was observed in *yol019W*Δ cells (**Figures 16A and 17A**). All of these genes are linked to vacuolar biology. *VTC1* encodes a member of the VTC complex required for polyP synthesis (Hothorn et al., 2009). *VMA22* and *VAM3* encode proteins involved in vacuolar protein maturation or localization (Graham et al., 1998; Srivastava and Jones, 1998). Finally, although its molecular function is unknown, *YOL019W* encodes a predicted transmembrane protein that localizes to the vacuole and outer cell membrane (Cherry et al., 2012). These data indicate a particularly important role for the vacuole in regulating polyphosphorylation through polyP synthesis and/or storage. Phenotypes shared among these and other mutants with low polyP levels may stem in part from changes to lysine polyphosphorylation. Although polyP is found throughout the cell, including the nucleus (Azevedo et al., 2015), the majority is retained in the vacuole following its synthesis by the VTC complex (Gerasimaite et al., 2014). How polyP is transported to other compartments is unknown. To determine whether the mechanism of polyP synthesis impacts target

modification, we took advantage of a previously described system that allows for *E. coli* PPK1 (*Ec*PPK1) expression in yeast (**Figure 16B**) (Gerasimaite et al., 2014).

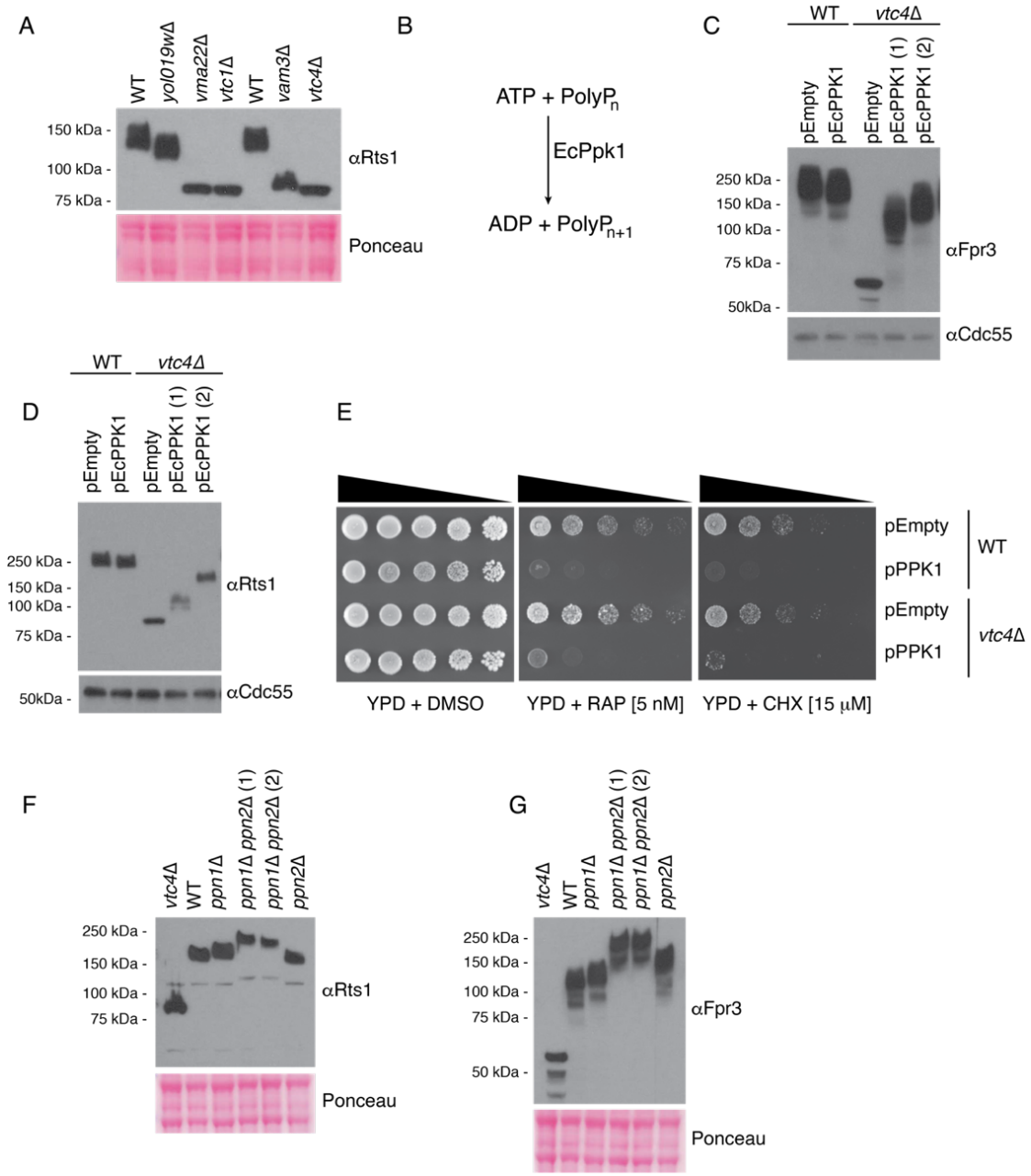


Fig. 16: Regulation of Polyphosphorylation. A) Polyphosphorylation of Rts1 is impacted by multiple genes. B) *Ec*PPK1 functions analogously to Vtc4 but synthesizes polyP independently of the vacuole. C-D) Expression of *Ec*PPK1 rescues polyphosphorylation defects of Fpr3 (C) and Rts1 (D) in *vtc4*Δ mutants. Two isolates of the *Ec*PPK1 expression in *vtc4*Δ mutants are shown. E) *Ec*PPK1 causes hypersensitivity to rapamycin and cycloheximide. Indicated strains were spotted in a 5-fold dilution series. F-G) Increased shift of targets Rts1 (F) and Fpr3 (G) in *ppn1*Δ *ppn2*Δ double mutants. Also see **Figure 17**.

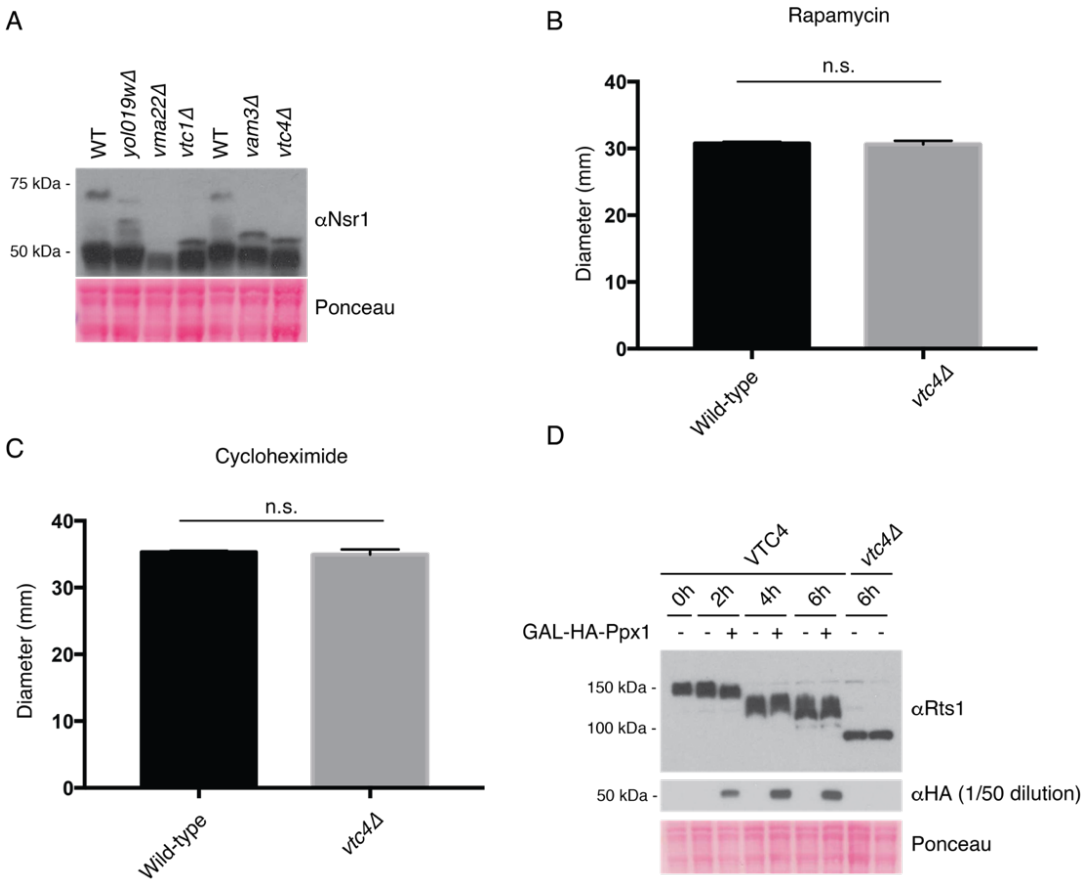


Fig. 17: Analyzing the Regulation of Polyphosphorylation. A) Proteins from the indicated strains were resolved via NuPAGE and probed with an antibody against Nsr1 following western blotting. B-C) *vtc4*Δ mutants are not sensitive to rapamycin or cycloheximide. The indicated strains were spread on YPD plates and filter disks containing 4.16 μg rapamycin or 2.5 μg cycloheximide. The diameter of the area of growth inhibition was measured after 22 hours of incubation at 30°C. No significant difference was found via student T-test (GraphPad Prism 7.0). Error bars represent the standard error. N=4 for each. D) Endogenous HA-Ppx1 was overexpressed via the GAL promoter and Rts1 polyphosphorylation was measured via NuPAGE analysis. HA-Ppx1 was detected by running a diluted version of the same samples on a different gel.

In contrast to yeast *Vtc4*, *Ec*PPK1 synthesizes polyP independently of vacuolar translocation and storage (Gerasimaite et al., 2014). *Ec*PPK1 expression in a *vtc4*Δ background was able to restore lysine polyphosphorylation of multiple targets (**Figures 16C and 16D**). These data suggest that vacuolar-coupled synthesis of polyP is not necessary for target modification. The presence of polyP, not its mechanism of synthesis, is the critical factor in target modification.

Expression of *Ec*PPK1 in yeast was previously found to result in slow growth that could be mitigated by the expression of the cytoplasmic Ppx1 polyphosphatase (Gerasimaite et al., 2014). In an extension of this observation, we found that *Ec*PPK1-expressing cells were sensitive to cycloheximide, which directly inhibits protein translation, and rapamycin, an inhibitor of the TORC1 nutrient-sensing kinase (**Figure 16E**). These phenotypes held true in both wild-type and *vtc4*Δ backgrounds. In contrast, *vtc4*Δ alone did not impact sensitivity to these drugs in either spot tests (**Figure 16E**) or halo assays (**Figures 17B and 17C**). Thus, while vacuole-coupled synthesis of polyP is not required for target modification, constraining polyP in the vacuole is still important for resistance to drugs that impact important aspects of cell growth. In the Discussion, we elaborate on a previous model suggesting that *Ec*PPK1 toxicity in yeast stems from unwanted accumulation of polyP in the cytoplasm (Gerasimaite et al., 2014).

Polyphosphatase-Dependent Regulation of Polyphosphorylation

Azevedo *et al.* demonstrated that the exopolyphosphatase Ppx1 and endopolyphosphatase Ddp1 are able to degrade polyP chains attached to target proteins in vitro (Azevedo et al., 2015). However, Ppx1 is largely cytoplasmic (Wurst et al., 1995), while most of our targets are nuclear. In order to test the function of Ppx1 more rigorously, we compared target polyphosphorylation in strains where the only copy of *PPX1* was placed under an inducible GAL1/10 promoter. In this experimental setup, Ppx1 is not detectable when cells are grown in raffinose but is overexpressed in the presence of galactose. While target polyphosphorylation decreased with culture time in some experiments (see Discussion), no difference was observed in the presence or absence of *Ppx1* overexpression (**Figure 17D**). These data suggest that Ppx1 may not directly reverse polyphosphorylation on target proteins. Similarly, we did not detect consistent changes in target polyphosphorylation in single deletions of polyphosphatases *PPX1* or *DDP1*, or vacuolar endopolyphosphatases *PPN1* or *PPN2*. In contrast, we observed a

consistent increase in electrophoretic shifts for Rts1 and Fpr3 when NuPAGE was used to analyze protein extracts from *ppn1*Δ *ppn2*Δ double mutants (**Figures 16F and 16G**). Interestingly, *ppn1*Δ *ppn2*Δ mutants were previously shown to accumulate longer polyP chains in all phases of growth, without an overall change to polyP levels (Gerasimaite and Mayer, 2017). We favor a model wherein overall polyphosphorylation of targets under normal growth conditions is dictated by polyP availability rather than removal of chains by polyphosphatases. However, since Ppx1 and Ddp1 can act on polyphosphorylated targets in vitro (Azevedo et al., 2015), we cannot completely rule out that these enzymes have a greater impact in vivo under certain conditions.

PolyP Regulates Ribosome Function

The 17 known polyphosphorylated proteins (our 15 targets plus Nsr1 and Top1) show preferential localization to the nucleolus, and a subset of these form a conserved network of interacting proteins implicated in ribosome biogenesis (**Figure 10C**). Since the regulation of ribosome biogenesis is tightly coupled to a cell's capacity to support translation and cell growth, we predicted that *vtc4*Δ cells, in which polyP is absent, would also show defects in ribosome function. To test this, we performed polysome profiling to examine ribosome assembly and translation elongation along mRNA. Here, *vtc4*Δ mutants showed a strong defect in 80S monosome and polysome assembly compared to wild-type, with a concomitant increase in free 40S and 60S subunits (**Figure 18A**).

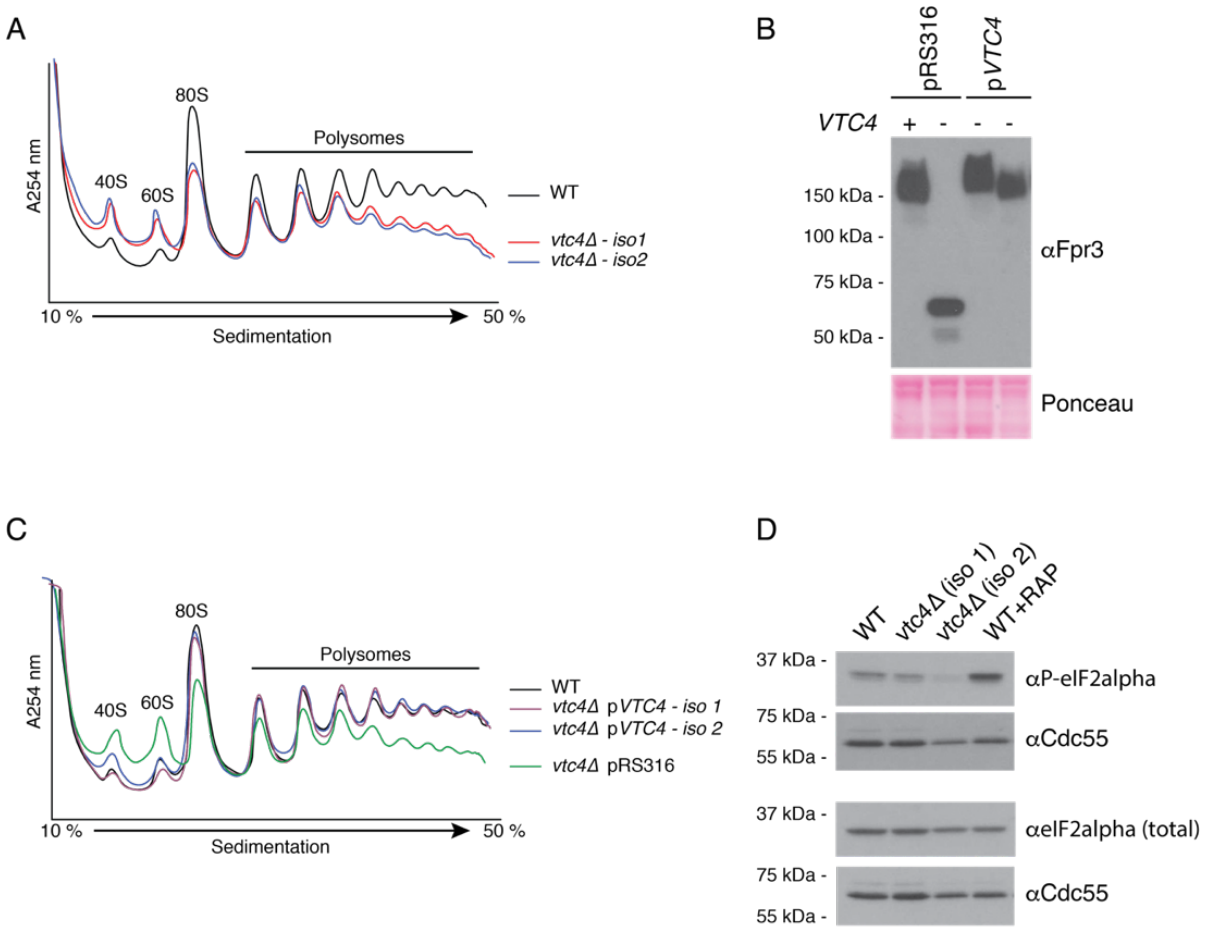


Fig. 18: Vtc4 Impacts Ribosome Function. A) Polysome profiles of cycloheximide-treated wild-type and *vtc4*Δ mutant strains (see Experimental Procedures for details). B-C) p*VTC4* complements polyphosphorylation (B) and polysome defects (C) in a *vtc4*Δ mutant strain. Two isolates of the *VTC4* rescue are shown. D) eIF2alpha is not hyperphosphorylated in *vtc4*Δ mutants.

To confirm that polysome defects observed in *vtc4*Δ cells were due to loss of the *VTC4* gene, we re-introduced the open reading frame (ORF) on a low copy plasmid under its own promoter. Expression of this construct in *vtc4*Δ mutant strains rescued target polyphosphorylation (**Figure 18B**) and reversed defects in polysome assembly (**Figure 18C**). The polysome profiles of *vtc4*Δ mutant cells are somewhat complex but are most similar to mutants having defects in translation initiation, consistent with a role for polyP in ribosome function (Li et al., 2009). Notably, cells lacking *VTC4* do not show elevated phosphorylation of eIF2a (**Figure 18D**), an event associated

with inhibition of translation following diverse cellular stresses (Dever, 2002). Therefore, it is unlikely that the polysome profiles stem from a general stress response triggered by the absence of Vtc4 or polyP. Since we have not yet been able to identify a PASK mutant that shares this phenotype, we favor the hypothesis that this phenotype represents the combined effects of polyphosphorylation changes to multiple targets.

Polyphosphorylation of Human Homologs

The core network of yeast polyphosphorylation targets implicated in ribosome biogenesis is conserved in human cells (**Figure 19A**). This led us to consider whether polyphosphorylation of lysine also occurs in higher eukaryotes. In order to understand the potential scope of lysine polyphosphorylation in mammalian cells, we screened the human proteome for proteins containing PASK motifs (75% S/D/E with at least one K). 1,134 proteins fit these criteria, and this set was enriched for nuclear/nucleolar factors with diverse functions (**Figures 19B and 19C**). Critically, included within this list (**Table 2, Appendix A**) were 6 homologs of polyphosphorylated yeast proteins (bolded in **Figure 10A**), including nucleolin, a functional homolog of Nsr1 that functions in various aspects of ribosome biogenesis. Nucleolin was previously reported to migrate at an apparent molecular weight of ~100 kDa on SDS-PAGE (Ginisty et al., 1999), and we found that it migrated at a similar position on NuPAGE gels (**Figure 19D**, lane 1). PolyP concentration in most mammalian tissues and cell lines is significantly lower than in yeast (~50 mM versus > 100 mM) (Kumble and Kornberg, 1995).

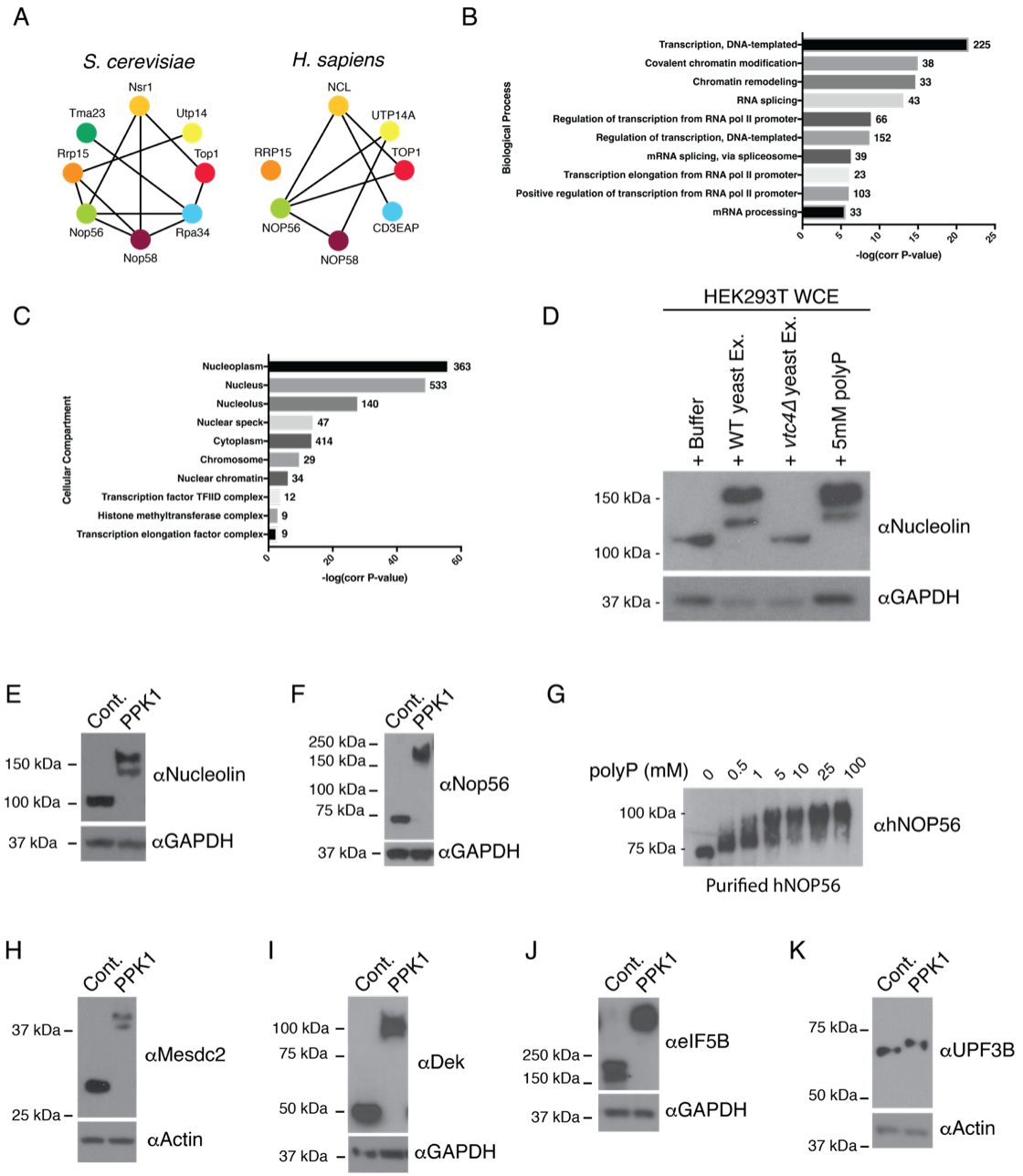


Fig. 19: Polyphosphorylation of Human Proteins. A) A network of physically interacting polyphosphorylated yeast targets is conserved in human cells. Network connections determined using Genemania. B-C) Analysis of PASK-containing proteins in human cells using DAVID with default settings. The number of proteins in each category is shown for biological process (C) and cellular component (D). D) Nucleolin can be polyphosphorylated *in vitro*. Extracts from HEK293T cells were incubated with wild-type or *vtc4*Δ mutant yeast extracts or with synthetic polyP prior to NuPAGE analysis. E-F) Expression of *Ec*PPK1 in HEK293T cells confers polyphosphorylation of human targets nucleolin (E) and hNOP56 (F). G) *In vitro* polyphosphorylation of bacterially purified hNOP56. H-K) Polyphosphorylation of human proteins Mesd (H), DEK (I), eIF5b (J), and UPF3B (K) following *Ec*PPK1 expression. Also see **Table 2, Appendix A** and **Figure 20**.

We reasoned that the polyP concentration in HEK293T cells used for our experiment may be too low to cause the dramatic changes in electrophoretic mobility observed for yeast targets. In support of this interpretation, mixture of HEK293T cell lysate with extracts from wild-type, but not *vtc4*Δ yeast, conferred a dramatic electrophoretic shift to the entire population of nucleolin protein (**Figure 19D**, lanes 2 and 3). A similar result was obtained by adding synthetic polyP to the extract (**Figure 19D**, lane 4).

The enzymes that synthesize polyP in human cells are currently unknown. Therefore, to test whether human cell lines could be engineered to produce polyP in quantities sufficient to confer lysine polyphosphorylation visible by NuPAGE analysis, we transfected cells with a plasmid expressing *Ec*PPK1. Nucleolin recovered from these cells displayed the electrophoretic shifts characteristic of lysine polyphosphorylation on NuPAGE gels (**Figure 19E**). This was also true for hNOP56, the homolog of yeast target Nop56 (**Figure 19F**). *In vitro*, synthetic polyP shifted hNOP56 purified from bacteria in a concentration-dependent manner (**Figure 19G**). To further demonstrate the utility of **Table 2 (Appendix A)**, we cross-referenced this list with commercially available antibodies and selected additional candidates for testing with NuPAGE analysis. We found evidence for *Ec*PPK1-induced polyphosphorylation of Mesd, DEK, eIF5b, and UPF3B (**Figures 19H–19K**). In contrast, the electrophoretic mobility of PASK-containing proteins Hsp90alpha and HIP were not impacted by *Ec*PPK1 expression (**Figures 20A and 20B**). Therefore, as we saw in yeast, the presence of a PASK motif alone is not enough to predict quantitative polyphophorylation. Critically, the observation that human cells can support the production of polyP in quantities allowing for NuPAGE detection of protein modification provides an avenue to study lysine polyphosphorylation in higher eukaryotes.

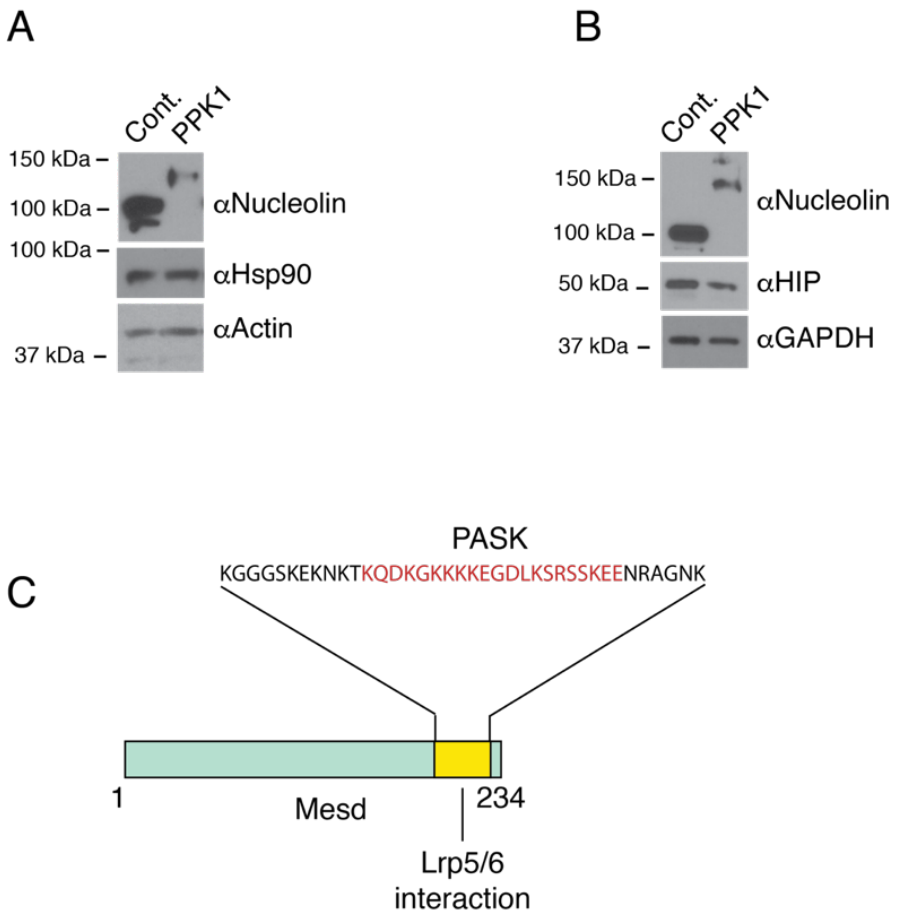


Fig. 20: Hsp90 and HIP are not polyphosphorylated. A-B) Hsp90 and HIP are not polyphosphorylated following *Ec*PPK1 expression. Proteins from the indicated HEK293T lysates were probed with the indicated antibodies following NuPAGE and western blotting. Nucleolin is used to show that the transfection with *Ec*PPK1 was successful. Nucleolin blot from membrane processed in parallel. C) Position of the PASK motif within a fragment of Mesd shown to mediate Wnt signaling (Lin et al., 2013).

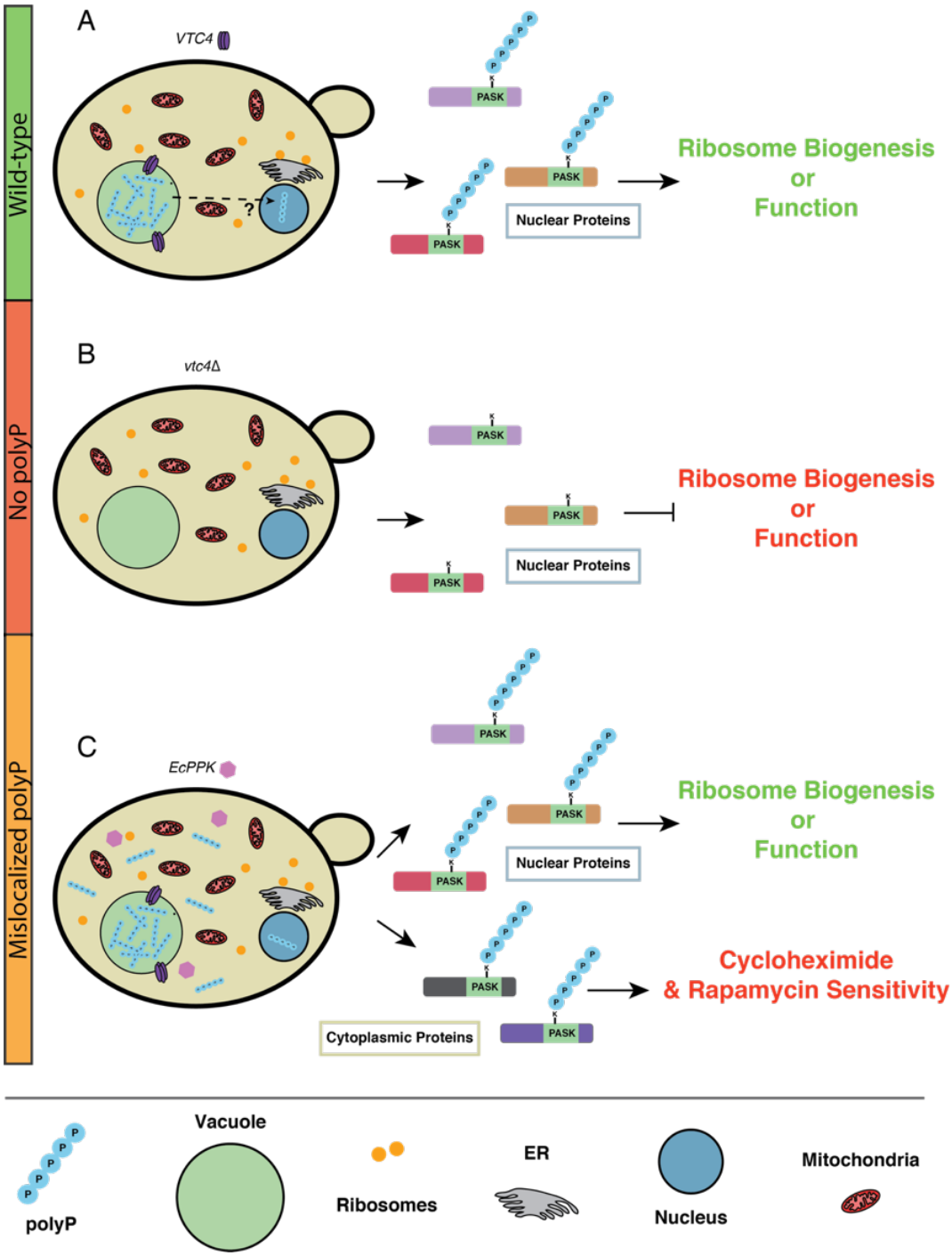


Fig. 21: An Updated Model for polyP Function in Yeast. A) In wild-type cells, polyP is synthesized by Vtc4 and is stored in the vacuole. It is transported to other areas of the cell by unknown mechanisms. In the nucleolus, polyphosphorylation of one or more ribosome biogenesis proteins is important for ribosome function. B) In the absence of *VTC4*, key substrates are no longer polyphosphorylated and ribosome biogenesis and/or translation are defective. C) Expression of *Ec*PPK1 permits extra polyP accumulation in the cytoplasm. Targets can still be polyphosphorylated, but the extra polyP disrupts important cellular functions.

2.4 Discussion

Polyphosphorylation as a Global Regulator

Polyphosphorylation of lysine was originally identified as a novel PTM on yeast proteins Nsr1 and Top1 (Azevedo et al., 2015). In this work, we have exploited tools unique to the budding yeast model to identify 15 previously unreported targets, including a conserved network of proteins implicated in ribosome biogenesis. Our work provides evidence of polyphosphorylation as a common modifier. Critically, our 15 targets were found in a set of only 90 prioritized candidates. There are an additional 337 proteins in yeast that contain PASK motifs, and these function in diverse pathways (**Table 1, Appendix A**). As such, lysine polyphosphorylation is likely to have a broad impact on cell function. To facilitate the use of our work as a resource, we have fully annotated **Tables 1 and 2 (Appendix A)** to indicate PASK motifs in both yeast and human proteins that overlap with unique functional, structural, or regulatory features. We centered our search for targets around the PASK motif. However, we cannot discount the possibility that sequences lacking PASK characteristics are also polyphosphorylated. As such, it will be critical to develop other methods to identify polyphosphorylated proteins. Notably, Azevedo et al. were unable to detect polyP linkage to Nsr1 using mass spectrometry (Azevedo et al., 2015). While we are revisiting this approach for other targets, the unbiased identification of new targets may be possible with other high-throughput methods. The use of protein chips (Zhu et al., 2001)—in which thousands of test proteins are immobilized on glass slides that can be analyzed using a micro-array scanner—is an attractive option. Our work will serve as a means to benchmark new methods for identifying targets of polyphosphorylation in yeast and mammalian systems.

A Role for polyP in Ribosome Function

Based on our recovery of a conserved network of targets that function in ribosome biogenesis, we predicted that cells that cannot make polyP would have defects in ribosome function. Indeed, cells lacking *VTC4* accumulate free 40S and 60S subunits and have a corresponding decrease in joined 80S monosomes and polysomes (**Figure 18**). These data are consistent with polyP being critical for a biogenesis step that impacts late-stage assembly, although a more direct role for polyP in the regulation of translation is also possible. As yet, we have been unable to directly link polyphosphorylation of any one target to defects in ribosome biogenesis or translation. It is possible that modulating the polyphosphorylation of multiple targets is required to confer observable phenotypes (**Figure 21**). Alternatively, modification of single targets with polyP chains could function redundantly with other pathways.

Molecular Function of Polyphosphorylation

We envision that polyphosphorylation of targets that we uncovered could regulate protein function through diverse molecular mechanisms. First, the addition of polyP chains would alter the charge of targets, which could have dramatic consequences for their interactions with other proteins and nucleic acids. Second, polyphosphorylation could also compete directly with other global lysine modifications such as acetylation, ubiquitylation, or SUMOylation. Additionally, polyphosphorylation could cross-talk with non-lysine modifications. For example, many serine residues within PASK domains are known sites of (mono) phosphorylation (**Table 1, Apendix A**). Finally, the phosphoanhydride bonds formed between inorganic phosphate residues in polyP chains are of very high energy (Rao et al., 2009). Polyphosphorylation of nucleolar proteins could have the secondary effect of concentrating this energy source in the nucleolus. In this scenario, polyphosphatase-mediated hydrolysis of polyP attached to protein targets could be coupled to energy-dependent events driving ribosome biogenesis.

Regulation of Polyphosphorylation

For all of our polyphosphorylated targets, the vast majority of the protein population appears to exist in the modified state as judged by NuPAGE analysis. This is in contrast to other lysine modifications such as acetylation and SUMOylation, which often occur on only a small fraction of total target protein (Sarangi and Zhao, 2015; Weinert et al., 2014). We suggest that the near complete modification of targets by polyphosphorylation results from the very high concentration of polyP (>100 mM) in budding yeast (Auesukaree et al., 2004). Yet, because polyP is added non-enzymatically (Azevedo et al., 2015), we cannot discount the possibility that some polyphosphorylation is occurring in solution following cell lysis. *In vivo,* target modification may be sub-stoichiometric due to local variation in polyP concentrations, competition with other PTMs or occlusion of PASK motifs by protein-protein or protein-nucleic acid interactions.

We found that deletion of genes required for the maintenance of polyP levels impacted lysine polyphosphorylation of multiple targets (**Figures 16A and 17A**). This highlights that polyP synthesis is likely to be a critical regulator of target modification. In some experiments, we noticed that target polyphosphorylation diminished when cultures were grown for long periods of time (**Figure 17D**). However, this was somewhat inconsistent, with other trials showing persistent target polyphosphorylation even after many days at saturation. While we cannot currently explain this observation, it is possible that subtle variations in growth conditions can trigger pathways that prevent complete degradation of polyP chains. In an elegant study, Gerasimaité et al. showed that polyP synthesis in the cytoplasm causes slow growth and aberrant cellular morphology. The authors proposed a model wherein the obligate coupling of polyP synthesis to vacuolar storage prevents this toxicity in wild-type cells (Gerasimaite et al., 2014). We used this same system to demonstrate that cytoplasmic polyP synthesis renders cells sensitive to the translation inhibitor cycloheximide and the TOR inhibitor rapamycin (**Figure 16E**). We suggest that high concentrations of cytoplasmic polyP could drive unwanted

polyphosphorylation of PASK-containing proteins to interfere with important cell functions (**Figure 21**).

Conservation of polyP and Lysine Polyphosphorylation

Many of the targets we identified are conserved in human cells. While the human polyP polymerase is currently unknown, our finding that 6 human proteins (nucleolin, hNOP56, Mesd, DEK, eIF5b, and UPF3B) can undergo polyphosphorylation suggests that regulation of proteins by this PTM may also be conserved. Of these targets, Mesd, a molecular chaperone involved in Wnt/b-catenin signaling, is particularly notable for the location of its 22-amino-acid PASK motif. This PASK is situated within a 38-amino-acid C-terminal and vertebrate-specific region of Mesd (**Figure 20C**). Peptides from this region inhibit Wnt ligand binding to LRP5 and dampen downstream transcription of Wnt/b-catenin-regulated genes (Lin et al., 2013). Notably, expression of the *Ec*PPK1 polyP kinase was required to observe quantitative polyphosphorylation of Mesd and other human proteins by NuPAGE analysis in HEK293T cells (**Figures 19E–19K**). The lack of quantitative polyphosphorylation in cells transfected with control plasmids likely stems from the lower concentration of polyP relative to that seen in yeast (Kumble and Kornberg, 1995). As such, polyphosphorylation could be more selective in higher eukaryotes. PolyP has previously been associated with cell growth. Decreasing polyP concentration through the ectopic expression of yeast exopolyphosphatase Ppx1 inhibits proliferation of the MCF7 breast carcinoma cell line (Wang et al., 2003). Mechanistically, polyP activates the nutrient and energy sensing kinase mTOR by promoting its (canonical) phosphorylation (Wang et al., 2003). Our work raises the possibility that signaling through mTOR and other pathways governing cell growth and division could be mediated in part by lysine polyphosphorylation. Intriguingly, we also found that alternative isoforms of many human proteins include changes to predicted PASK sequences (**Table 2, Appendix A**). Differences in

lysine polyphosphorylation could underlie variations in function observed for such isoforms. Our study provides a foundation for investigating this potential mode of regulation in higher eukaryotes.

2.5 Experimental Procedures

Yeast Strains and Plasmids

S. cerevisiae strains were generated and grown using standard techniques (Goldstein and McCusker, 1999; Longtine et al., 1998). Synthetic DNA constructs and strain genotypes are listed in **Tables 4 and 5 (Appendix B)**.

Statistical Analyses

For halo assays (see Supplemental Experimental Procedures), n = 4 with error bars representing SE. No significant difference was found via Student's t test. Description of corrected p values for bioinformatics is contained in the Supplemental Experimental Procedures.

Western-Blotting Screen for Polyphosphorylated Proteins

From the 427 proteins with PASK domains, 90 candidates were prioritized largely based on known expression level as well as number and length of PASK sequences present in each protein. Generation of matched *ppn1*Δ and *ppn1*Δ *vtc4*Δ GFP-tagged strains was achieved using Synthetic Genetic Array Technology (Tong et al., 2001). Details are contained in the Supplemental Experimental Procedures. Protein extracts were prepared using a trichloroacetic acid (TCA)-lysis method (Rossl et al., 2016) and analyzed on Bis-Tris NuPAGE gels (Thermo Fisher Scientific NP0336BOX). *VTC4*-dependent electrophoretic shifts of potential targets were considered a ''hit.'' See **Table 3 (Appendix B)** for antibody details.

In Vitro Polyphosphorylation Assays

For analysis of Rts1 and Fpr3, synthetic polyP (Kerafast EUI005) was added to TCA lysis extracts for 1 hr at room temperature. For the chain-length experiment in **Figure 11**, polyP standards from Regenetiss, Japan were used. Samples were analyzed by NuPAGE.

Transfections and Protein Preps from EcPPK1-Expressing HEK293T Cells

HEK293T cells cultured as described in Supplemental Experimental Procedures were grown to ~70% confluency and transfected with 2–28 mg of DNA using Lipofectamine 2000 (Invitrogen 11668-019). To isolate protein, cells were washed with ice-cold 1X PBS, scraped, and lysed in 0.2–1 mL of RIPA Lysis buffer (see supplemental buffer recipes). Cells were incubated for 15 minutes on ice and centrifuged for 15 min at 13,500 rpm. The supernatant was collected prior to addition of Laemmli sample buffer and analysis using NuPAGE.

2.6 Acknowledgements

We thank Dr. A. Mayer for yeast *Ec*PPK1 expression constructs and Drs. C. Azevedo and A. Saiardi for sharing strains and for valuable advice. We thank Dr. T. Shiba for providing polyP standards. This work was supported by CIHR grant number PJT-148722 to M.D. and an NSERC grant to A.G. A.R. is supported by NSERC grant number 355719.

2.7 Supplemental Experimental Procedures

Yeast Strains

ORF disruptions were confirmed using PCR to test for (1) the correct positioning of knock-out cassettes, and (2) for the absence of the wild-type ORF. Epitope-tagged strains were confirmed via immunoblotting and/or PCR to check for correct location of the tagging constructs prior to use. Our screen was carried out in synthetic complete media. All other experiments were in YPD media unless indicated otherwise. For the MMS experiment, cells were diluted to OD600 = 0.2 and grown to OD600 ~ 0.5. Afterwards, MMS was added to cultures (0.035%) for 3 hours.

Plasmids

The VTC4 open reading frame was amplified from yeast gDNA prepared using the YeaStar Genome DNA kit (Zymo Research D2002) with primers CGAGGTCGACGGTATCGATAACATATTAAAGCTTTCAAAA and GTGGATCCCCCGGGCTGCAGTTTCATTCTATCAAGTTAAC. These primers amplify a cassette that contain 400 bps of the promoter region and 200 bps of the terminator sequence and contain homology for assembly into EcoRI (NEB R3101T) and HindIII (NEB R3104S) digested pRS316 to yield (pRS316-VTC4) using the NEB Gibson Assembly Cloning Kit (NEB E5510S). For EcPPK1 expression in human cells, E. coli PPK1 was amplified from a yeast construct obtained from Andreas Mayer using primers TCCACCACACTGGACTAGTGATGGGTCAGGAAAAGCTATAC and TGATCAGCGGTTTAAACTTATTATTCAGGTTGTTCGAGTG. The cassette was introduced into BamHI (NEB R3136S) and HindIII (NEB R3104S) cut pcDNA3.1(-) using Gibson assembly. Plasmid sequences were confirmed by Sanger sequencing (McGill University and Génome Québec Innovation Centre). VTC4 and EcPPK1 expression constructs generated for this work will be made available through Addgene (Addgene.org). Other plasmids encoded PASK mutants under the GAL promoter.

Bioinformatics

In silico PASK screens were carried out using custom scripts to search Uniprot for yeast or human sequences with a window of 20 amino acids to identify regions containing 75% S/D/E/K with at least one K. Windows were extended until the percentage of these residues dropped below 75%. GO-enrichments for PASK datasets were carried out using DAVID version 6.8 with default settings (Sherman et al., 2007). P-values shown are corrected using the Benjamini-Hochberg method. Functional and structural annotations of PASK motifs in **Tables 1 and 2 (Appendix A)** were done using a variant of SLiMSearch (Krystkowiak and Davey, 2017). Human homologs of yeast targets were identified using Yeastmine (Balakrishnan et al., 2012). Nucleolin is considered a functional homolog of Nsr1 based on (Edwards et al., 2000) and included here based on discussion as such in Azevedo et al. 2015.

Details of Western-blotting screen:

Strains from the GFP-tagged yeast collection were mated to YMD1178 or YMD1179. Diploid selection was carried out on –HIS + cloneNAT [100 mg/L; Werner 5.005.00]. Following sporulation, sequential selection of haploid strains containing desired markers was carried out by pinning of strains on selection media. For *ppn1*Δ strains, final selection was on canavanine [50 mg/L; Sigma C9758], thialysine [50 mg/L; Sigma a2636], -LEU, -HIS, and cloneNAT [100 mg/L; Werner 5.005.00]. Final selection for *ppn1*Δ *vtc4*Δ double-mutant haploid strains was on canavanine [50 mg/L; Sigma C9758], thialysine [50 mg/L; Sigma a2636], -LEU, -HIS, cloneNAT [100 mg/L; Werner 5.005.00], and G418 [200 mg/L; Wisent 400-130-IG]. Control and *vtc4*Δ strains expressing the same GFP-tagged target were prepared as described below and separated on 4-12% Bis-Tris NuPAGE gels (ThermoFisher NP0336BOX). In the beginning of our screen we used ‘PolyP WCE preparations (see below) in order to extract proteins for 4-12% Bis-Tris NuPAGE analysis (ThermoFisher NP0336BOX). After screening ~ 1/3 of our candidates, we found that shifts could be preserved using standard TCA lysis, and switched to

this method of screening for the remainder of the screen. TCA lysis had the benefit of allowing full extraction of low abundance proteins. 15 of 17 initial targets had reproducible VTC4-dependent shifts upon reanalysis. Strains used for our screen and some (but not all) of our follow-up work were deleted for PPN1. We reasoned that more polyP (or greater chain length) in *ppn1*Δ mutants might result in more polyphosphorylation of PASK motifs, making new targets easier to identify. In subsequent experiments, we found that PPN1 had only a minor impact on target modification. As such, we have transitioned away from working in the *ppn1*Δ background.

TCA protein Preps

Yeast cell cultures were diluted in YPD or SC selective media from an overnight culture to OD600 = 0.8 and grown for 1 hour at 30°C. Cells were pelleted at 3,000 RPM (4°C) and washed with cold water. Cells were resuspended in 300 μL of 20 % TCA (Sigma T6399) and 100 μL of glass microbeads were added. Cells were lysed by bead beating samples twice for 1.5 minutes, with incubation on ice between bursts. The supernatant (S1) was collected in an Eppendorf tube. Samples were vortexed for 10 seconds in 300 μL of 5% TCA (Sigma T6399) and this supernatant (S2) was pooled with S1. The pooled supernatant was pelleted at 10,000 RPM (4°C) and subsequently removed. The pellet was resuspended in 100 μL SDS-PAGE sample buffer (see Supplemental Buffer recipes) and boiled at 100°C. Samples were pelleted at 10,000 RPM (RT) and the supernatant was retrieved. Samples were ran on a 4-12% Bis-Tris NuPAGE gel (ThermoFisher NP0336BOX) or an 8% SDS-PAGE gel and subsequently transferred to a PVDF membrane (Biorad 162-0177). Samples loaded on a 4-12% Bis-Tris NuPAGE gel (ThermoFisher NP0336BOX) were run at 200V for approximately 1 hour. For NuPAGE analysis, shifts were preserved in different sample buffers (NuPAGE (ThermoFisher NP0007/NP0009), BOLT (ThermoFisher B0007/B0009), Laemmli, SDS-PAGE and in both the presence and absence of the NuPAGE anti-oxidant in the running buffer (ThermoFisher NP0005). Samples loaded on 8% SDS-PAGE gels were run at 120V for approximately 2 hours.

PolyP WCE Preparations

The protocol for "polyP" style lysis was adapted from Azevedo et al. (2015). Yeast cell cultures were diluted in YPD or synthetic complete (SC) selective media from an overnight culture to OD600 = 0.8 and grown for 1 hour at 30°C. Cells were pelleted at 3,000 RPM (4°C) and washed with cold water. Cells were resuspended in 300 μL of ice-cold polyP lysis buffer (see Supplemental Buffer recipes). Approximately 100 μL of glass microbeads were added. Cells were lysed by bead beating samples twice for 1.5 minutes (BioSpec Mini Bead Beater), with ice incubation in between bursts. Cellular remnants were pelleted at 10,000 RPM (4°C) and the supernatant was kept. A 1:1 ratio of supernatant to BOLT LDS loading buffer (ThermoFisher B0007, B0009) was boiled at 100°C. Samples were loaded on a 4-12% Bis-Tris NuPAGE gel (ThermoFisher NP0336BOX) and subsequently transferred to a PVDF membrane (Biorad 162-0177) for immunoblotting.

Immunoblotting

Antibody sources and dilutions used are described in **Table 3 (Appendix B)**. In general, membranes were blocked for 30 minutes with 5 % milk in 1X TBST, prior to incubation with primary antibody overnight at 4°C or for 1 hour at RT. Following 3 X 5 minute washes with 1X TBST, blots were incubated with secondary antibodies conjugated to HRP. After an additional 3 X 5 minutes of washing with 1X TBST, blots were incubated with Millipore Immobilon ECL reagent (Sigma WBKLS0500) and exposed to autoradiography film from Progene (Ultident 39-20810). Equal transfer and loading across membranes was verified by staining with Ponceau S (Sigma P-3504). For blots with biotin, nitrocellulose membrane (Bio-Rad 1620112) was used prior to blocking with 5 % milk, 5 x 10 minute washes and incubation with Strep-HRP (Abcam Ab7403).

TEMED Collapse of PolyP-Induced Shifts

For this assay ~375 ng of stock Rts1 in 1X PBS was incubated with 10 mM sodium phosphate

(Fisher BP329-500 and BP332-1) or polyP (Kerafast EUI005) in a 10 μL reaction for min at room temperature. For home-made Bis-Tris gels, the following recipe was used: 10% 30:2.5% w/v Acrylamide:Bis (Fisher BP170, BioRad 161-0201), 0.33M Bis-Tris [pH 6.5-6.8 (Sigma B9754) with HCl (Fisher A144-212)]. Gel mixes were degassed before addition of 40 μL of 10% w/v ammonium persulphate (fresh, Sigma A3678) per 10 mL. Where used, TEMED (Fisher BP150) was added at a concentration of 5 μL per 10 mLs. Gels were polymerized for 1 hour with an overlay of 0.1 % SDS (Fisher BP166). Gels were run in 1X MOPS running buffer (20X = 1M MOPS (Sigma M1254), 1M Tris base (Fisher BP152-5), 20mM EDTA (Sigma ED2SS), 2% SDS (Invitrogen 15525-017)). Sodium bisulfite (Fisher S-654) was added to running buffer at a concentration of 5mM. Gels were run at 145mA and 200V for 1.5 hours. Gels were stained using the Invitrogen Colloidal Blue Staining Kit (Invitrogen, LC6025) as per manufacturer instructions for Bis-Tris gels.

PASK Motif Mutagenesis Experiments

Plasmids encoding GAL-3HA-RPA34, GAL-3HA-RPA34allK-R, GAL-3HA-RPA34ΔPASKNOP56, GAL-3HA-CHZ1, GAL-3HA-CHZ1ΔN, GAL-3HA-CHZ1ΔC, and GAL-3HACHZ1Δ(N+C) were purchased from ATUM (See **Table 4 in Appendix B**). For galactose induction experiments cells were grown overnight in SC-URA dropout media with 2% raffinose. Cultures were diluted to OD600 ~ 0.4, grown for an addition two hours in SC-URA + 2% raffinose, prior to addition of 2% galactose for 2 hours. Cells were pelleted at 3,000 RPM (4°C) and washed with cold water. TCA protein lysis was performed and samples were run on 4-12% Bis-Tris NuPAGE Gels (ThermoFisher NP0336BOX) as previously described. Immunoblotting was performed as described above.

Anti-Sui2 Antibody Generation

Anti-Sui2 rabbit polyclonal antibody sera was generated by injecting 1mg of GST-Sui2 (aa2-304) into rabbits every 4 weeks for 12 weeks (approval and handling via uOttawa animal

facility), rabbit serum was harvested, clarified by centrifugation. GST-Sui2 was expressed from pAR1395, in which Sui2 (aa2-304) was cloned as a BamH1/EcoR1 fragment into pGEX-6P-1.

Spot-tests & Halo Assays

Freshly-grown yeast cells were taken directly from agar plates, resuspended in water and serially diluted 5-fold. Each dilution was spotted on the indicated plates and spots were allowed to dry at room temperature. Plates were imaged after the indicated time at 30 °C. Drug concentrations used are indicated in each figure legend. For halo assays, 1mL of diluted cultures of equal density (OD600 = 0.5-0.7) were spread over the surface of the plate. Excess culture was removed and plates were dried at room temperature. Sterile filter disks containing the indicated amounts of cycloheximide (Sigma 01810) or rapamycin (Sigma R8781) were placed in the centre of the plate. After ~22 hours the diameter of the area of growth inhibition was measured from 2 directions and averaged.

Shifts of bacterially purified Rts1

Preparation of GST-Rts1(1-150) and GST controls were described previously (Kennedy et al., 2016). For GST-Rts1(1-150) polyphosphorylation shift assays, 2 μL of Rts1 (1/10 stock at 3.2 mgs/mL) was incubated in 20 μL reactions with the indicated concentration of polyP (Kerafast EUI005) and 10 μL polyP lysis buffer (see Supplemental Buffer Recipes) for 1 hour at room temperature prior to addition of NuPAGE LDS sample (as previously described). Samples were separated on 4-12 % Bis-Tris NuPAGE gels (ThermoFisher NP0336BOX) prior to staining with the Colloidal Blue Staining Kit (Invitrogen LC6025) as per manufacturer instructions for BisTris gels. Where indicated glutathione sepharose beads (Thermo Fisher 16101) were incubated with reactions described above for 40 minutes prior to washing 3 x 500 μL of polyP lysis buffer and elution with NuPAGE sample buffer. Biotinylated polyP (Kerafast EUI007) was used in reactions with GST-Rts1(1-150) or GST denatured in SDS-PAGE sample buffer at the indicated concentrations.

Polysome Profiling

Strains were grown either in YPD or Synthetic Complete (SC) media. SC media lacking uracil (-URA) was used to maintain the plasmids. Total RNA was isolated from yeast cells at mid-log phase according to the protocol of Esposito, et al. (Esposito et al., 2010)(Esposito et al., 2010) and then loaded on a 10–50% sucrose gradient made by an automated gradient maker (Biocomp). Samples were then centrifuged at 40,000 rpm for 2 hours at 4C° in a Beckman Optima LE-80K ultracentrifuge to separate the molecules through the sucrose gradient. After centrifugation, samples were analyzed on Biocomp gradient station, the absorbance from the gradient was recorded at A260 using a flow cell coupled with a spectrophotometer (Bio-Rad Econo UV monitor) where untranslated mRNAs (top fractions) were separated from polysome-associated mRNAs (bottom fractions) as in Faye, et al. (Faye et al., 2014). For all experiments cycloheximide was added to a concentration of 100 μg/mL at the time of harvest (OD600 =0.6-0.8) and included in the wash solution (100 μg/mL), lysis buffer (80 μg/mL), and sucrose gradient buffer (100 μg/mL).

Mammalian Cell Culture

HEK293T cell lines were obtained from ATCC (ATCC CRL-3216). Cells were grown to ~80% confluency in DMEM (Wisent Multicell 319-015-CL) supplemented with 10% Fetal Bovine Serum (Wisent Multicell 080-150), 1mM sodium pyruvate (Wisent Multicell 600-110-EL), and Penicillin-Streptomycin-Amphotericin B solution (Wisent Multicell 450-115-EL). Final concentrations of Penicillin, Streptomycin, and Amphotericin B were 50,000 IU/500 mL, 50 mgs/500 mL and 125 μg/500 mL, respectively. Cell cultures in our lab are routinely monitored for mycoplasma contamination using the e-Myco VALiD Mycoplasma PCR Detection Kit (Cat. 25239). Cells were grown in 100mm x 20mm polystyrene dishes (VWR CABD353003) at 37°C and 5% CO2.

In vitro polyphosphorylation assay (HEK293T)

Cells were washed twice with 1X PBS (Wisent 311-010-CL), scraped and centrifuged for 5 minutes at 5,000 rpm. PBS was removed and the pellet was stored at -80 °C until use. The HEK293T cell pellet was thawed and 1 mL of mammalian polyP lysis buffer (see Supplemental Buffer recipes) was added. The mixture was vortexed, kept on ice for ~15 minutes and revortexed. The sample was passed through a 21-gauge needle ~5 times and centrifuged for 15 minutes at 13,500 RPM. The supernatant was collected in a new microcentrifuge tube and the protein extract was kept on ice. For each experiment, 4 reactions were assembled. In tube #1, 5 μL of HEK293T WCE was added to 45 μl of mammalian polyP lysis buffer. In tube #2, 5 μL of HEK293T WCE was added to 45 μl of control yeast lysate (prepared with polyP WCE protocol). In tube #3, 5 μL of HEK293T WCE was added to 45 μL of vtc4Δ yeast lysate (prepared with polyP WCE protocol). In tube #4, 5 μl of HEK293T WCE was added to 45 μL of synthetic biotinylated polyP (Kerafast EUI007, final concentration = 0.5mM). After incubation with endover-end rotation for 1 hour at 4°C, BOLT loading buffer (see Supplemental Buffer recipes) was added to each sample in a 1:1 ratio. Samples were boiled for 5 minutes prior to separation via 4-12% Bis-Tris NuPAGE (ThermoFisher NP0336BOX). Antibody sources and concentrations can be found in the Antibodies table.

hNOP56 shift assay

A master mix was prepared with 40 μl of polyP Lysis Buffer without inhibitors, 5 μl of purified hNOP56 (OriGene TP761154) and 5μl of Laemmli Sample Buffer. This master mix was boiled before being used in 7 samples containing synthetic polyP with concentrations ranging from 0 mM to 100 mM. The samples were incubated at room temperature for 1 hour. Laemmli sample buffer was added to each sample in a 1:1 ratio and the samples were boiled before analysis using 4-12% Bis-Tris NuPAGE gels (ThemoFisher NP0336BOX), as described above.

SUPPLEMENTAL BUFFER RECIPES

PolyP Lysis Buffer (10 mL)

9.7 mL ice cold lysis buffer [150mM NaCl (Sigma S9888), 50mM Tris-HCl pH 8.90 (Tris base Fisher BP152-5, Hydrochloric acid Fisher A144-212), 7mM EDTA (Sigma ED2SS)], 20 μL 1M Imidazole (Sigma i0250), 10 μL 1M Sodium Fluoride (Sigma 201154), 11.5 μL 1M Sodium Molybdate (Sigma 243655), 50 μL 200mM Sodium Orthovanadate (NEB P0758S), 50 μL fresh 1M DTT (BioBasic DB0058), 10 μL 5mM Cantharidin (Sigma C7632), 100 μL 2.5mM (-)-pbromotetramisole Oxalate (Sigma 190047), 100 μL Sigma Protease Inhibitor Cocktail (Sigma P8215).

Mammalian PolyP Lysis Buffer

9.7 mL ice cold lysis buffer [150mM NaCl (Sigma S9888), 50mM Tris-HCl pH 8.90 (Tris base Fisher BP152-5, Hydrochloric acid Fisher A144-212), 7mM EDTA (Sigma ED2SS)], 20 μL 1M Imidazole (Sigma i0250), 10 μL 1M Sodium Fluoride (Sigma 201154), 11.5 μL 1M Sodium Molybdate (Sigma 243655), 50 μL 200mM Sodium Orthovanadate (NEB P0758S), 50 μL fresh 1M DTT (BioBasic DB0058), 10 μL 5mM Cantharidin (Sigma C7632), 100 μL 2.5mM (-)-pbromotetramisole Oxalate (Sigma 190047), 100 μL Sigma Protease Inhibitor Cocktail (Sigma P8215), 100μL 100 mgs/mL PMSF (Sigma P7626), 1 cOmplete protease inhibitor tablet (without EDTA; Roche 4693132001).

RIPA Lysis Buffer

50 mL ice cold lysis buffer [10mM Tris-HCl pH 7.4 (Tris base Fisher BP152-5, Hydrochloric acid Fisher A144-212), 100mM NaCl (Invitrogen AM9759), 1mM EDTA (Sigma 03690), 1% IGEPAL CA-630 (Sigma I8896), 0.5% Sodium Deoxycholate (Sigma D6750), 0.1% SDS (Invitrogen 15525-017)], 500 μL Sodium Fluoride (0.5M, Sigma 201154), 500 μL Glycerol-2-Phosphate (1M, Sigma G9422), 500 μL Nicotinamide (1M Sigma N3376), 500 μL PMSF (1 mg/mL Sigma P7626), 500 μL Sodium Butyrate (1M Sigma 303410), and 1 cOmplete protease

inhibitor tablet (Roche 04693116001).

SDS-PAGE Running Buffer (1X)

100 mL of 10X 1 litre Stock [30.2 g Tris Base (Fisher BP152-5), 188 g Glycine (Fisher BP381-5), 10 g SDS (Fisher BP166)] and 900 mL ddH2O.

NuPAGE Running Buffer (1X)

50 mL of 20X 1 litre Stock [209.2 g MOPS (Sigma M1254), 121.1 g Bis-Tris (Sigma B9754), 20 g SDS (Fisher BP166), 12 g EDTA (Sigma ED2SS)] and 950 mL ddH2O.

SDS-PAGE Transfer Buffer (1X)

100 mL of 10X 1L Stock [30.275 g Tris Base (Fisher BP152-5), 166.175 Glycine (BioBasic GB0235], 200 mL Methanol (Fisher A412P-4) and 700 mL ddH2O.

NuPAGE Transfer Buffer (1X)

50 mL of 20X 1L Stock [81.6 g Bicine (Sigma B3876), 104.8 g Bis-Tris (Sigma B9754), 6 g EDTA Sigma ED2SS)], 200 mL Methanol (Fisher A412P-4) and 750 mL ddH2O.

SDS-PAGE Sample Buffer: (Yeast, 3X)

800 μL of stock [160 mM Tris-HCl pH 6.8 (Tris base Fisher BP152-5, Hydrochloric acid Fisher A144-212), 6% SDS w/v (Fisher BP166), 30% Glycerol (Fisher BP229-4), 0.004 % Bromophenol Blue (Fisher BP115-25)] supplemented with 100 μL 1M DTT (BioBasic DB0058), and 100 μL 1.5M Tris-HCl (pH 8.8) [Tris base (Fisher BP152-5), Hydrochloric acid (Fisher A144-212)].

Laemmli Sample Buffer: (Mammalian, 5X)

10% SDS (Fisher BP166), 0.5M DTT (BioBasic DB0058), 25% B-mercaptoethanol (Sigma M6250), 50% Glycerol (Fisher BP229-4), 4-Bromophenol Blue (Fisher BP115-25), 250mM Tris-HCl pH 6.8 [Tris base (Fisher BP152-5), Hydrochloric acid (Fisher A144-212)].

3.0: Manuscript #2

Proteins required for vacuolar function are targets of lysine polyphosphorylation in yeast

Publication Information:
McCarthy, L.*, **Bentley-DeSousa, A.***, Denoncourt, A., Tseng, Y-C., Gabriel, M., and Downey M. (2020). Proteins required for vacuolar function are targets of lysine polyphosphorylation in yeast. *FEBS Letters* 594, 21-30.
*Authors contributed equally

Author's Contribution:
LM, ABD, AD, YCT and MG conceived and carried out experiments and interpreted data. MD conceived experiments, interpreted data, and wrote the paper with LM and ABD.

Contributions by figure: Figure 22 (YCT), Figure 23 (AD), Figure 24 (AD), Figure 25 (MD), Figure 26 (LM), Figure 27 (ABD, LM), Figure 28 (ABD), Figure 29 (MD).
MD contributed data analysis and/or preparation to all figures.

Proteins required for vacuolar function are targets of lysine polyphosphorylation in yeast

Liam McCarthy[1,2], Amanda Bentley-DeSousa[1,2], Alix Denoncourt[1,2], Yi-Chief Tseng[1,2], Matthew Gabriel[1,2], and Michael Downey[1,2]

[1]Department of Cellular and Molecular Medicine, University of Ottawa, Canada

[2]Ottawa Institute of Systems Biology, University of Ottawa, Canada

Liam McCarthy and Amanda Bentley-DeSousa contributed equally to this work

3.1 Abstract

Polyphosphates (polyP) are long chains of inorganic phosphates that can be attached to lysine residues of target proteins as a nonenzymatic post-translational modification. This modification, termed polyphosphorylation, may be particularly prevalent in bacterial and fungal species that synthesize large quantities of polyP. In this study, we evaluated the polyphosphorylation status of over 200 candidate targets in *Saccharomyces cerevisiae*. We report eight new polyphosphorylated proteins that interact genetically and physically with previous targets implicated in ribosome biogenesis. The expanded target network includes vacuolar proteins Prb1 and Apl5, whose modification with polyP suggests a model for feedback regulation of polyP synthesis, while raising questions regarding the location of polyphosphorylation in vivo.

3.2 Introduction

Polyphosphate (polyP) is a polymer of inorganic phosphate moieties joined together in linear chains ranging from 3 to 1000 s of residues in length. Evidence suggests that these chains exist in all cells, although their length and concentration vary widely. PolyP is particularly abundant in the budding yeast Saccharomyces cerevisiae, comprising upward of 10% of the dry weight of the cell and reaching internal concentrations of over 200 mM in terms of individual phosphate units (Auesukaree et al., 2004; Kornberg et al., 1999). Most of the polyP in yeast is stored in the vacuole following synthesis by the vacuolar membrane-bound VTC complex (Gerasimaite et al., 2014). Overall chain length is controlled by exo- and endopolyphosphatases Ddp1, Ppn1, Ppn2, and Ppx1 (Gerasimaite and Mayer, 2017). In addition to its role in ion homeostasis and phosphate metabolism (Eskes et al., 2018; Rosenfeld et al., 2010), Azevedo et al. (Azevedo et al., 2015), showed that polyP chains can be noncovalently added to lysine residues as a post-translational modification (PTM). They characterized this modification on proteins Nsr1 and Top1, providing evidence that polyP chains can regulate protein interactions and topoisomerase activity of purified Top1 in vitro, as well as the localization of Nrs1 and Top1 *in vivo*.

Polyphosphorylation is nonenzymatic and occurs in polyacidic serine and lysine rich (PASK) motifs (Azevedo et al., 2015; Azevedo and Saiardi, 2016). While the sequence requirements for polyP addition are unclear, lysine residues appear to be the modified site (Azevedo et al., 2015; Azevedo and Saiardi, 2016; Bentley-DeSousa et al., 2018). Consistent with this, mutation of lysine residues within predicted PASK motifs to arginine prevents polyphosphorylation of Nsr1, Top1, and Rpa34 (Azevedo et al., 2015; Bentley-DeSousa et al., 2018).

Polyphosphorylation causes dramatic electrophoretic shifts when protein extracts are run on denaturing Bis-Tris NuPAGE gels, but this same shift does not occur with traditional SDS/PAGE gels (Azevedo et al., 2015; Bentley-DeSousa et al., 2018). The underlying reasons for this

difference are not completely clear, although it could be related to the use of TEMED in the polymerization of SDS/ PAGE gels (Bentley-DeSousa et al., 2018).

Recently, we reported the identification of an additional 15 polyphosphorylation targets in yeast, including a conserved network implicated in ribosome biogenesis (Bentley-DeSousa et al., 2018). Many of the PASK motifs in these proteins are contained within intrinsically disordered regions that appear to be under evolutionary selection (Zarin et al., 2019). Based on enrichment in ribosome-related functions, we uncovered a novel function for VTC4, encoding the catalytic subunit of the VTC complex (Hothorn et al., 2009), in polysome assembly (Bentley-DeSousa et al., 2018). We also demonstrated that six human proteins, including homologs of our novel yeast targets, could be polyphosphorylated following the ectopic expression of *Escherichia coli* PPK, the bacterial polyP polymerase, in HEK293T cells (Bentley-DeSousa et al., 2018). Subsequent studies using protein microarrays identified an additional eight human proteins as targets, in addition to five proteins that appear to bind polyP chains noncovalently (Azevedo et al., 2018). Altogether, work from our group and others suggests that lysine polyphosphorylation has the capacity to be a global protein modifier, akin to other lysine-based modifications such as acetylation or ubiquitylation (Azevedo and Saiardi, 2016; Bentley-DeSousa and Downey, 2019).

Our previous strategy to identify targets of lysine polyphosphorylation in yeast took advantage of a set of strains wherein each open reading frame is expressed as a fusion with the GFP epitope (Huh et al., 2003). Candidate fusions from this set were crossed into a wild-type or *vtc4*Δ mutant background prior to analysis by NuPAGE and immunoblotting with an antibody against GFP. In this assay, a *VTC4*-dependent electrophoretic shift signals a candidate polyphosphorylated protein. In our previous work (Bentley-DeSousa et al., 2018), we prioritized 90 targets from 427 proteins containing putative PASK motifs, defined by a sliding window of 20 amino acids with 75% D, E, S, or K amino acids with at least 1 K. This set of 90 proteins was enriched for multiple or long PASK motifs containing many lysine residues. In an attempt to further define the

landscape of polyphosphorylation in *S. cerevisiae*, we have now screened the remainder of the candidates present in this GFP collection (See Methods section) using a similar strategy.

3.3 Methods

Yeast Strains

All strains are in the S288C BY474 background and were generated using standard methods. Strains were analyzed via PCR in a manner similar to that described by the yeast deletion consortium (http://www-sequence.stanford.edu/ group/yeast_deletion_project/protocols.html) to confirm the presence of the knockout cassette and the absence of the wild-type open reading frame for gene deletions and the correct position of the tag-marker cassette for epitope-tagged strains. This analysis was also done for strains streaked or generated from strains taken from the GFP-tagged set. Here, a common forward primer was used to check the position of GFP-HIS3MX cassette with a gene-specific reverse primer outside of the area of homology used for recombination at the 30 end of the gene (typically 300–500 bp downstream of the stop codon). The common forward primer P1 (also known in our lab as KandCV) has the sequence 50-TTCGCCTCGACATCATCTGCCC-30 and primes in the common TEF terminator of the 'MX' style cassettes. Thus, this primer is useful in confirming cassette positions associated with other yeast collections including the TAP-tagged and deletion libraries. A sample PCR analysis is shown in **Figure 22**.

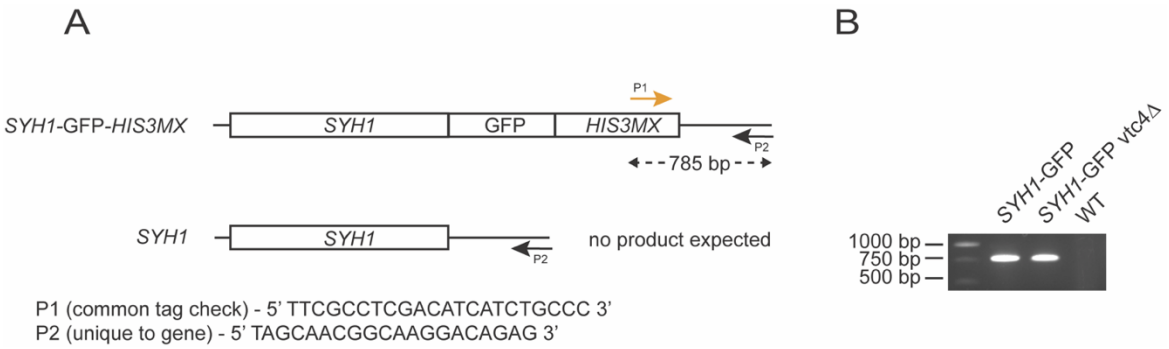


Fig. 22: Strategy for TAG checking. A) Diagram of *SYH1*-GFP tagged locus (top) and WT locus (bottom). Primers used for confirmation are indicated. B) gDNA prepared from the indicated strains was used in PCR reactions with the primers indicated in A. Strains used in the analysis are YMD2258, YMD2259, and YMD963. Relates to Materials and Methods.

Strain confirmations were always done on colony-purified isolates, whose genotypes are described in **Table 7 (Appendix B)**. For the GAL shut-off experiment used to examine Prb1 processing, wild-type GALpr-PRB1 and GALpr-PRB1 *vtc4*Δ mutant strains were grown in synthetic complete (SC) media with 2% raffinose overnight. Cells were diluted to OD600 = 0.3 in SC media with 2% raffinose and grown until OD600 = 0.6. Cultures were then supplemented with 2% galactose for 30 min to induce Prb1 expression. Subsequently, glucose was added to a final concentration of 2% and time points were taken every 15 min. For all other experiments, cells were grown in YPD media supplemented with 0.005% adenine and tryptophan. All strains were grown at 30 °C.

Screening Protocol

Candidate target strains were recovered from the yeast GFP-tagged collection and the *vtc4*Δ and *ppn1*Δ mutations were crossed into the strains as previously described (Bentley-DeSousa et al., 2018). Paired protein extracts were generated and analyzed using NuPAGE analysis following the preparation of protein extracts via TCA lysis (see below). Rts1 was used as a control polyphosphorylated protein for each sample analyzed. Candidates were confirmed in a secondary screen before streaking strains for single colonies and confirming the correct position of the GFP tag. As described throughout the results section, select hits were reconfirmed by re-tagging candidates in wild-type and *vtc4*Δ strains and/or with the use of antibodies directed against native targets. A list of screened targets and their predicted PASK motifs is included in **Table 6 (Appendix B)**. We have now analyzed 321 of 427 PASK containing proteins in yeast.

Protein Analysis

Protein extraction was done using a TCA lysis method that has been described previously (Bentley-DeSousa et al., 2018). A similar protocol is reiterated here. Briefly, 3–6 OD600 equivalents of yeast cells were lysed in 300 μL of 20% TCA with 100 μL of acid- washed glass

beads. Supernatant was recovered and beads were washed with 300 μL of 5% TCA and combined with the first supernatant. Samples were centrifuged at 16 000 g for 4 min at 4 °C. Pellets were resuspended in 100 μL SDS/PAGE sample buffer (0.8 mL SDS/PAGE sample buffer, 100 μL 1.5 M Tris-HCl pH 8.8, 100 μL 1 M DTT). Samples were boiled for 5 min and centrifuged again at 16 000 g for 4 min. Supernatant was removed for immediate analysis or stored at -80 °C until use. Typically, 10 μL of extract (~ 75 μg) was loaded for each lane, but this varied depending on the expression of the protein under analysis. We have observed that protein samples tend to lose their polyphosphorylation after multiple freeze thaw cycles. We recommend same-day NuPAGE analysis of extracted proteins.

Western Blotting

Buffers used for NuPAGE and SDS/PAGE analysis were described previously (Bentley-DeSousa et al., 2018). SDS/PAGE gels used for this work were 10% final concentration made with 30% acrylamide/bis solution at a ratio of 37.5:1 (BioRad 1610158, Hercules, CA, USA). Please note that SDS/PAGE analysis does not resolve polyP-induced shifts. PVDF membrane (Biorad 162-0177) was used for Western blotting of separated protein samples. Primary and secondary antibodies used for immunoblotting, including dilution information for each antibody, are described in **Table 8 (Appendix B)**. Washed blots were exposed to ECL from Millipore (Burlington, MA, USA: Sigma WBKLS0500 or WBLUF0500). Ponceau S stain (Sigma P-3504, St. Louis, MO, USA) was used to verify equal loading across the blots and to monitor the quality of transfer for each blot.

Homologs and Network Analysis

Homologs were determined via the YeastMINE tool (Balakrishnan et al., 2012). Genetic networks were generated using GeneMANIA (GeneMANIA.org; June 12, 2019) (Warde-Farley et al., 2010). Gene attributes were set to zero, only genetic and physical interactions were

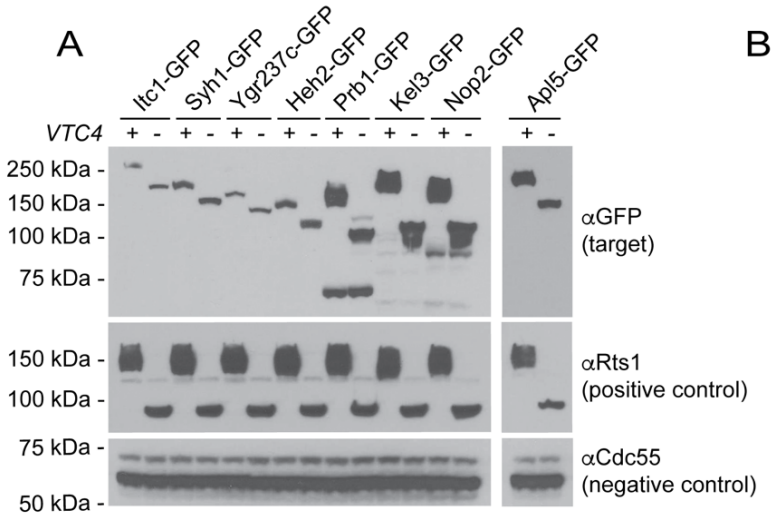
computed and shown, all other settings were left as default. FDR displayed was as calculated for nucleolar localization by the GeneMANIA tool.

In Vitro Polyphosphorylation Assays

In vitro polyphosphorylation assays were carried out by mixing the indicated concentrations of polyP with TCA-style protein preps, followed by incubation at room temperature for 20min before re-boiling of reactions and NuPAGE analysis, as described above. PolyP chains used for this analysis are an average of 75 units in length and were purchased from Kerafast (Medium chain, EUI005, Boston, MA, USA). For supporting *in vitro* analyses, short (14 units), medium (60 units), and long (130 units) chains were gifted by T. Shiba (Regentiss, Kunitachi City, Japan). Extra-long polyP chains (700 units) were from Kerafast (EUI002).

3.4 Results and Discussion

Using our previously defined screening approach, we have now uncovered an additional eight candidate targets. As observed in our previous work, the shifts conferred by polyP (i.e., in the presence of a functional *VTC4* gene) are dramatic and quantitative, with the whole population of target proteins being affected (**Figure 23**).



B

Protein	Description or Function	Homolog
Itc1	Chromatin remodeller	N/A
Syh1	Unknown	GIGYF1/2
Ygr237c	Unknown	N/A
Heh2	Inner nuclear membrane	LEMD2/3
Prb1	Vacuolar protease	PCSK9
Kel3	Unknown	KLHDC2
Nop2	rRNA methyltransferase	**NOP2**
Apl5	Cargo transport to vacuole	**AP3δ1**

Fig. 23: Identification of new polyphosphorylated targets in yeast. A) NuPAGE analysis of polyphosphorylated proteins. Paired GFP-tagged proteins were extracted from wild-type or *vtc4*Δ mutant cells using TCA lysis prior to NuPAGE analysis, western blotting to PVDF membrane and detection using the indicated antibodies. Rts1 serves as a positive control. Cdc55 serves as a negative control. Images representative of $N = 4$ or greater for each target. B) Table of new polyphosphorylated targets uncovered in this study. Human homologs were determined using the YeastMINE tool. Bold indicates the presence of a PASK sequence in human homolog (as defined by (Bentley-DeSousa et al., 2018)).

We attribute this 'all or nothing' pattern to the high concentration of polyP in cells and in extracts following cell lysis, which could be quantitatively modifying proteins in vitro. In vivo polyphosphorylation of any given target may be sub-stoichiometric. As we previously found

forother targets (Bentley-DeSousa et al., 2018), polyP-induced shifts collapsed when samples were analyzed via SDS/PAGE instead of NuPAGE (**Figure 24**).

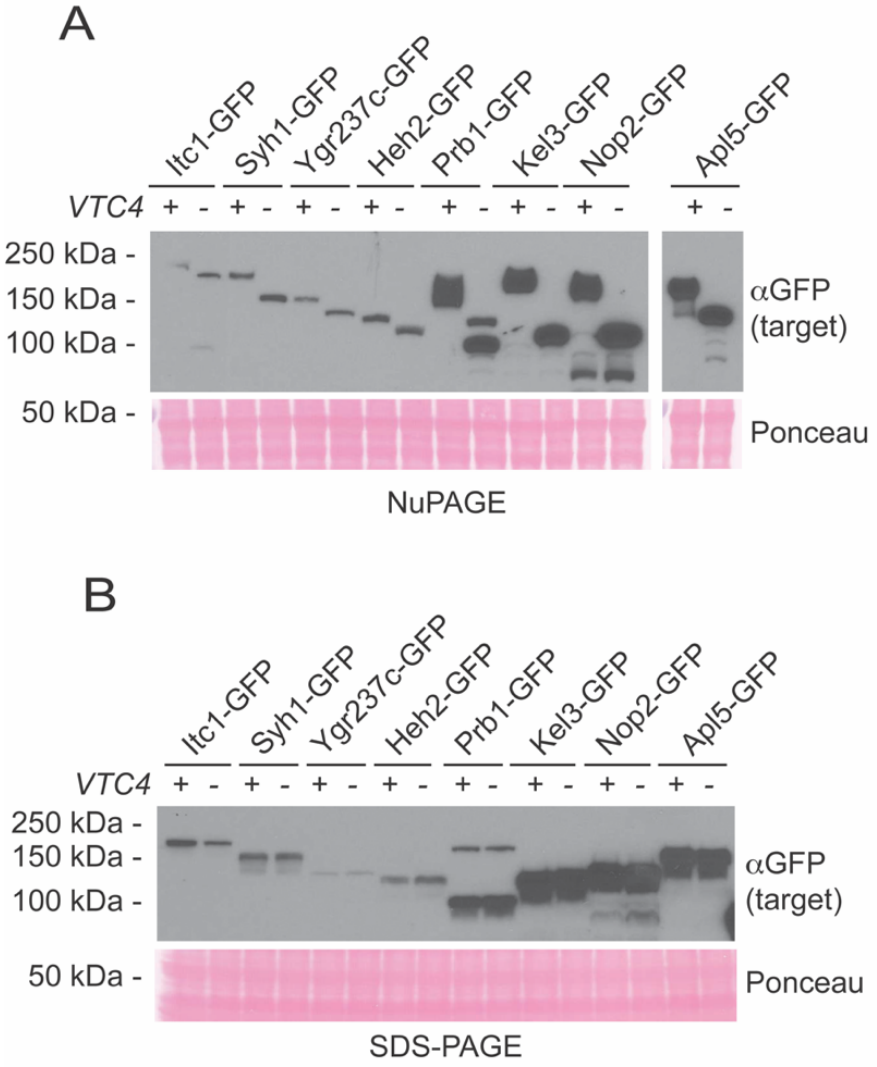


Fig. 24: Collapse of polyphosphate-induced shifts on SDS-PAGE gels. A) Analysis of protein samples from the indicated strains by NuPAGE was followed by immunoblotting with anti-GFP antibody. Representative of at least N=4 for each target. B) As in A but with separation on SDS-PAGE gels instead of NuPAGE. Representative of N=2. Relates to **Figure 23**.

New polyphosphorylated targets uncovered in our screen include homologs of proteins implicated in human disease (**Figure 23**, see below). Despite the identification of proteins predicted to function in diverse processes, new targets display extensive physical and genetic interactions with previously identified polyphosphorylated proteins (**Figure 25**). Heh2 is an inner

nuclear membrane protein involved in ensuring the quality control of nuclear pore complexes (Webster et al., 2014). Ygr237c and Kel3 are uncharacterized proteins with diverse physical and genetic interactions and putative cytoplasmic localization (Cherry et al., 2012).

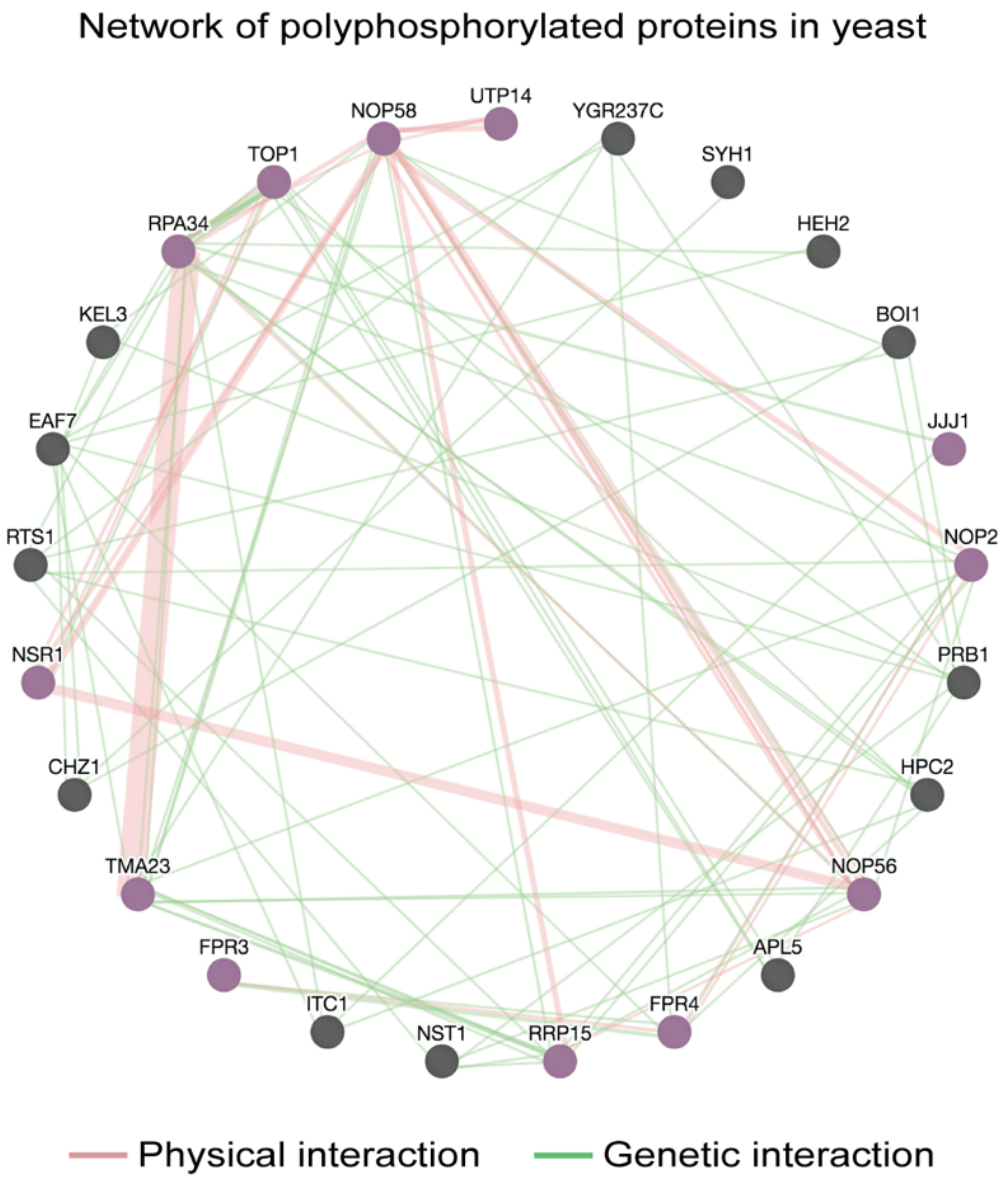


Fig. 25: Polyphosphorylated targets are connected by genetic and physical interactions. Genetic and physical associations between new and previously identified targets were computed using the GeneMANIA tool (see Methods section). Each node (circle) represents an individual polyphosphorylated protein (n = 22). Physical and genetic interactions are indicated with green and pink lines, respectively. Purple nodes indicate localization to the nucleolus (n = 12; FDR = 1.85×10^{-7}).

Itc1 is a member of the Itc1-Isw2 chromatin remodeling complex involved in meiotic gene repression (Sugiyama and Nikawa, 2001). Nop2 is an rRNA methyltransferase required for 27S rRNA processing and 60S ribosome biogenesis (Hong et al., 1997; Sharma et al., 2013). While largely uncharacterized, Syh1 co-purifies with ribosomes (Fleischer et al., 2006), suggesting that it too functions in translation or ribosome assembly. SYH1 also displays genetic interaction with EBP2, encoding a protein involved in 60S biogenesis (Okano et al., 2015). The identification of Nop2 and Syh1 is consistent with our previous observation that *vtc4*Δ cells show defects in polysome formation and the previous identification of negative genetic interactions between *vtc4*Δ mutants and deletions of genes impacting ribosome biogenesis (RPP1A, MAK11, and RPC25) and translation (ELP3, ELP6, and MNR1) (Chatr-Aryamontri et al., 2017; Costanzo et al., 2016). These data are also consistent with findings in other eukaryotes. In particular, polyP localizes to the nucleolus, the site of ribosome biogenesis, and interacts with nucleolar proteins in trypanosomes (Negreiros et al., 2018). Nucleolar polyP is also thought to regulate RNA polymerase I in myeloma cells (Jimenez-Nunez et al., 2012). Our expanded target list also included a number of proteins associated with vacuolar biology. Prb1 encodes the yeast vacuolar proteinase B (Moehle et al., 1987). Apl5 is involved in transporting protein cargo from the ER-Golgi system to the vacuole (Angers and Merz, 2009). Since the vast majority of polyP localizes to the vacuole in *S. cerevisiae*, these proteins are ideal candidate effectors of polyP-mediated signaling. Therefore, we carried out further experiments to validate the polyphosphorylation of these new targets and to determine amino acid sequences required for the modification.

We first sought to confirm that the electrophoretic shift detected for Prb1-GFP in the presence of *VTC4* was due to polyP. Since polyphosphorylation occurs nonenzymatically even in harsh conditions (Azevedo et al., 2015), we added synthetic polyP of 75 units in length to *vtc4*Δ whole-cell protein extracts generated via a trichloroacetic acid (TCA) lysis protocol. PolyP

addition caused a concentration-dependent electrophoretic shift (**Figure 26A**). Addition of 5 mM polyP conferred an electrophoretic shift to Prb1-GFP in *vtc4*Δ extracts that was equal to the shift observed in the wild-type control. In contrast, the same concentration of sodium phosphate had no impact on the mobility of Prb1- GFP, confirming that polyP is required for the modification (**Figure 26A**). PolyP-induced shifts also scaled with chain length, with longer chains producing a more dramatic change in electrophoretic mobility (**Figure 27A**).

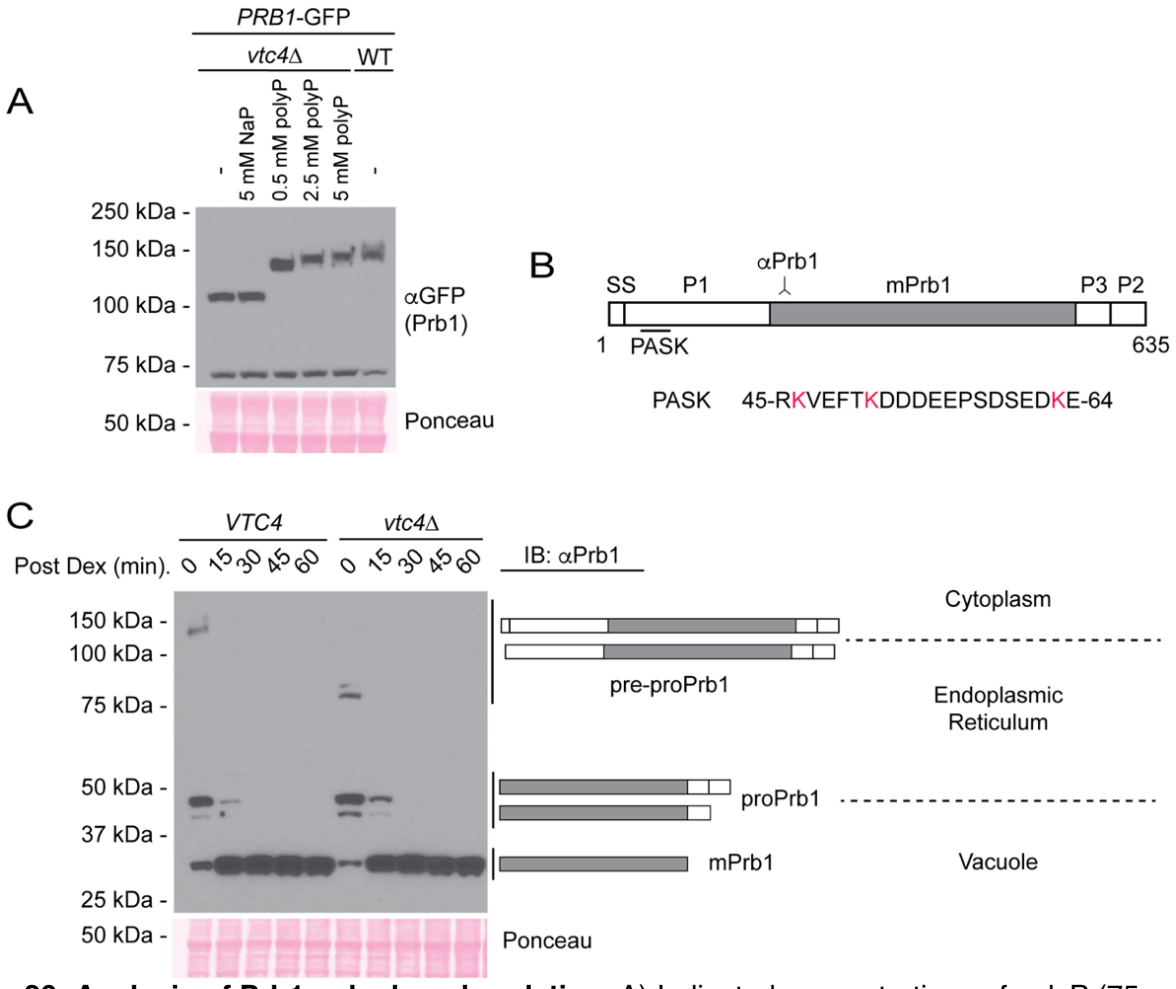


Fig. 26: Analysis of Prb1 polyphosphorylation. A) Indicated concentrations of polyP (75 units) or sodium phosphate control were added to *vtc4*Δ mutant cell extracts expressing GFP-tagged Prb1 and polyP-induced shifts were analyzed by western blotting with an antibody against GFP after NuPAGE analysis. The band at the bottom is likely a breakdown product containing GFP. Representative of $N = 2$. B) Schematic of inactive pre-pro form of Prb1. The region recognized by the antibody is shown. Red lysines indicate candidate sites for polyP attachment. C) Prb1 was expressed under the GAL promoter (see Methods section) for 2 h prior to the addition of glucose to stop transcription. Processing and polyphosphorylation of Prb1 at

the indicated time points was monitored via NuPAGE analysis. The proposed schematic for Prb1 processing indicated on the right was adapted from that proposed in (Mark et al., 2014). Representative of $N = 2$.

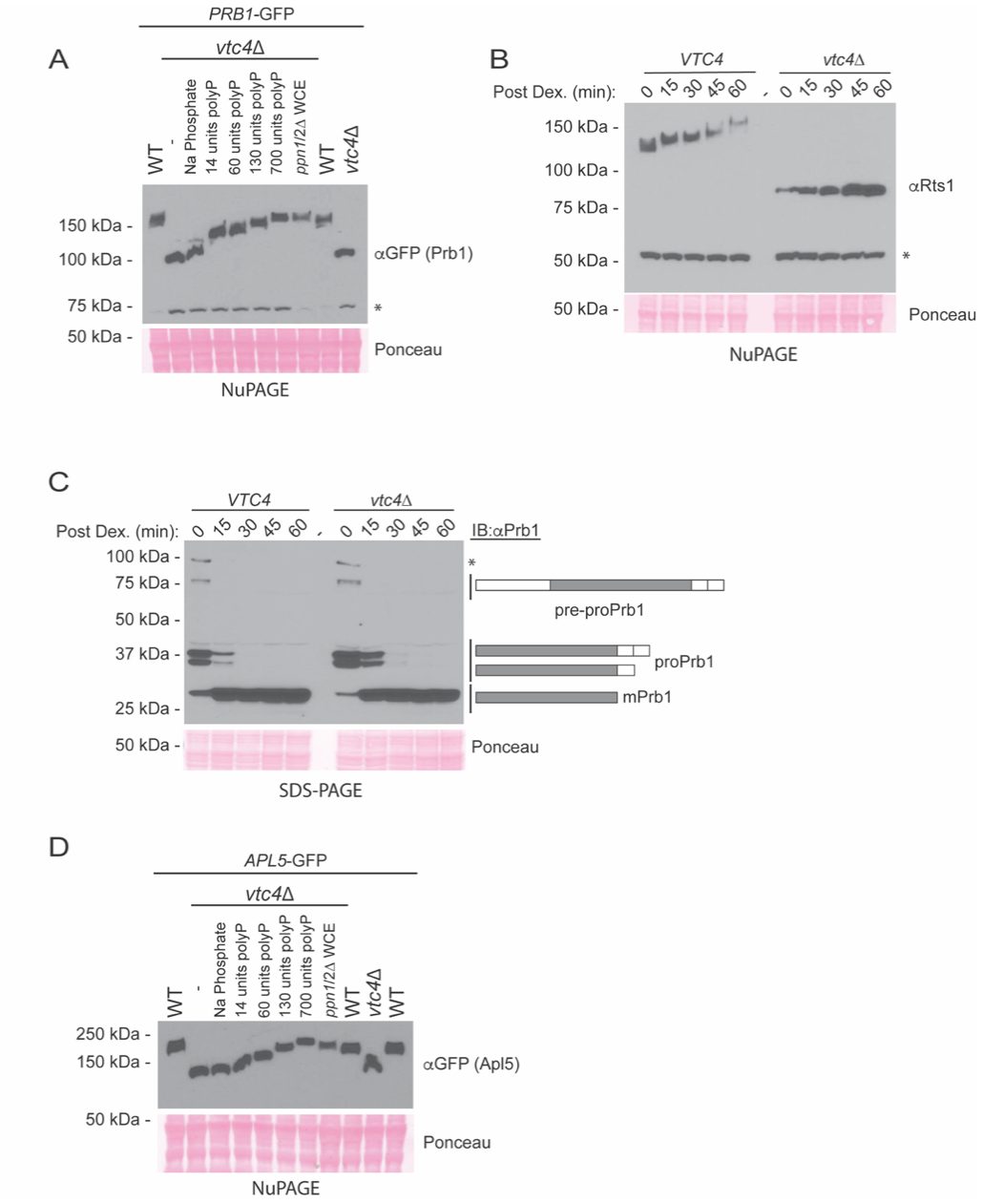


Fig. 27: Analysis of Prb1 and Apl5 polyphosphorylation. A) Chain length dependence of polyP-induced shifts. 10 mM chains of the indicated lengths or extracts from *ppn1*Δ *ppn2*Δ deletion mutants were mixed with extracts from *vtc4*Δ cells expressing Prb1-GFP and resolved via NuPAGE. The band marked with an asterisk at the bottom is likely a breakdown product containing GFP. Representative of N=2. B) Polyphosphorylation of Rts1 was monitored during the experiment presented in **Figure 26C** using an antibody against Rts1. The band marked with an asterisk at the bottom is likely a breakdown product containing GFP. Representative of N=2. C) Processing of pre-pro Prb1 was monitored for the experiment presented in **Figure 26C** using

SDS-PAGE and immunoblotting with an antibody against Prb1. Note: SDS-PAGE does not resolve shifts due to lysine polyphosphorylation. Asterisk may represent a previously uncharacterized form of Prb1. Representative of N=2. D) 10 mM chains of the indicated lengths or extracts from *ppn1*Δ *ppn2*Δ deletion mutants were mixed with extracts from *vtc4*Δ cells expressing Apl5-GFP and resolved via NuPAGE. Representative of N=4. Relates to **Figures 26 and 28**.

The Prb1 protease is a heavily processed protein that undergoes a series of glycosylations and programed cleavages to convert the zymogen precursor (pre-pro) form to the fully active enzyme (mPrb1) (Moehle et al., 1987). Processing occurs in step with its transit through the ER-Golgi system to the vacuole (**Figure 26B,C**). Since the C-terminal GFP tag of Prb1-GFP is predicted to be removed during its processing, **Figures 23A and 26A** only confirmed polyphosphorylation of pre-proPrb1. In order to study polyphosphorylation of Prb1 in the context of its processing from zymogen to active enzyme, we used a previously described galactose over-expression system (Mark et al., 2014). Induction of PRB1 transcription with the addition of galactose, followed by addition of glucose to turn off that transcription, allowed for the production of a population of native Prb1 whose processing could be monitored over time. Following analysis by NuPAGE and western blotting, we detected Prb1 with a previously published antibody that recognizes the first 14 amino acids of the fully processed and active (mPrb1) protein (**Figure 26B**) (Moehle et al., 1987). We expect this antibody to detect all forms of Prb1. Only pre-proPrb1, but not any of the processed forms, exhibited an electrophoretic shift in *VTC4* strains relative to *vtc4*Δ mutants (**Figure 26C**). As a control, the previously identified target Rts1 (Bentley-DeSousa et al., 2018) showed polyphosphorylation throughout the time course in VTC4 cells (**Figure 27B**). These data suggest that the P1 segment of Prb1, which is removed by the Pep4 protease in the first cleavage event within the endoplasmic reticulum (Moehle et al., 1987), is the modified species. Indeed, the predicted PASK motif is located in the P1 fragment, as indicated in **Figure 26B**. Importantly, analysis of these same samples via

traditional SDS/PAGE, which does resolve polyP-induced shifts (Bentley-DeSousa et al., 2018), showed identical migration of all species, consistent with the electrophoretic shift being due to polyphosphorylation (**Figure 27C**). In this analysis, we also observed slightly delayed processing of Prb1 in *vtc4*Δ mutants, although it is unclear if this is biologically meaningful (**Figure 27C**).

Polyphosphorylation of the P1 fragment may be required to ensure early processing steps occur in a manner consistent with the production of an active mPrb1 protein. In this capacity, polyP may function redundantly or in coordination with other PTMs, including glycosylations. Alternatively, polyP may function in a quality control capacity to regulate the function or degradation of Prb1 that is improperly processed by the ER-Golgi system. A similar role has been suggested for the SCFSaf1 ubiquitin ligase, which ubiquitylates Prb1 but does not impact the bulk turnover of the protein population (Mark et al., 2015; Mark et al., 2014). The hypothesis that polyP modulates protein stability is consistent with its previously described role as a molecular chaperone in other biological systems (Gray et al., 2014).

Prb1 is an interesting target in part because it emphasizes the importance of determining where within the cell polyphosphorylation occurs. While most polyP is localized to the yeast vacuole, it has also been detected in other compartments including the mitochondria, cytoplasm, and nucleus (Azevedo et al., 2015; Breus et al., 2012; Pestov et al., 2004; Saito et al., 2005). Importantly, the bioavailability of these pools of polyP to act in polyphosphorylation reactions *in vivo* is not clear. To our knowledge, polyP localization to the ER lumen has not been well-documented. But, a fraction of the polyP-synthesizing VTC complex was proposed to reside at the ER membrane (Gerasimaite and Mayer, 2016).

Apl5, the second vacuole-related protein uncovered in our screen, functions as a member of the conserved AP3 complex. AP3 selectively transports protein cargo from the Golgi to lysosome-related organelles such as the yeast vacuole (Cowles et al., 1997; Odorizzi et al., 1998). As with

Prb1-GFP, Apl5-GFP polyphosphorylation was rescued by the addition of increasing concentrations of polyP to *vtc4*Δ extracts, and polyP-induced shifts scaled with chain length (**Figure 28A, Figure 27D**). Other subunits of AP3 were not polyphosphorylated (**Figure 28B**), suggesting that Apl5 may be a focal point for AP3 regulation by polyP. As predicted, the PASK motif of Apl5 is required for the electrophoretic shift observed in *VTC4* strains (**Figure 28C**). This PASK motif is located in the 'ear domain' of the protein (**Figure 28D**), previously demonstrated to be required for interaction with AP3 regulators such as Vps41 (Angers and Merz, 2009; Rehling et al., 1999).

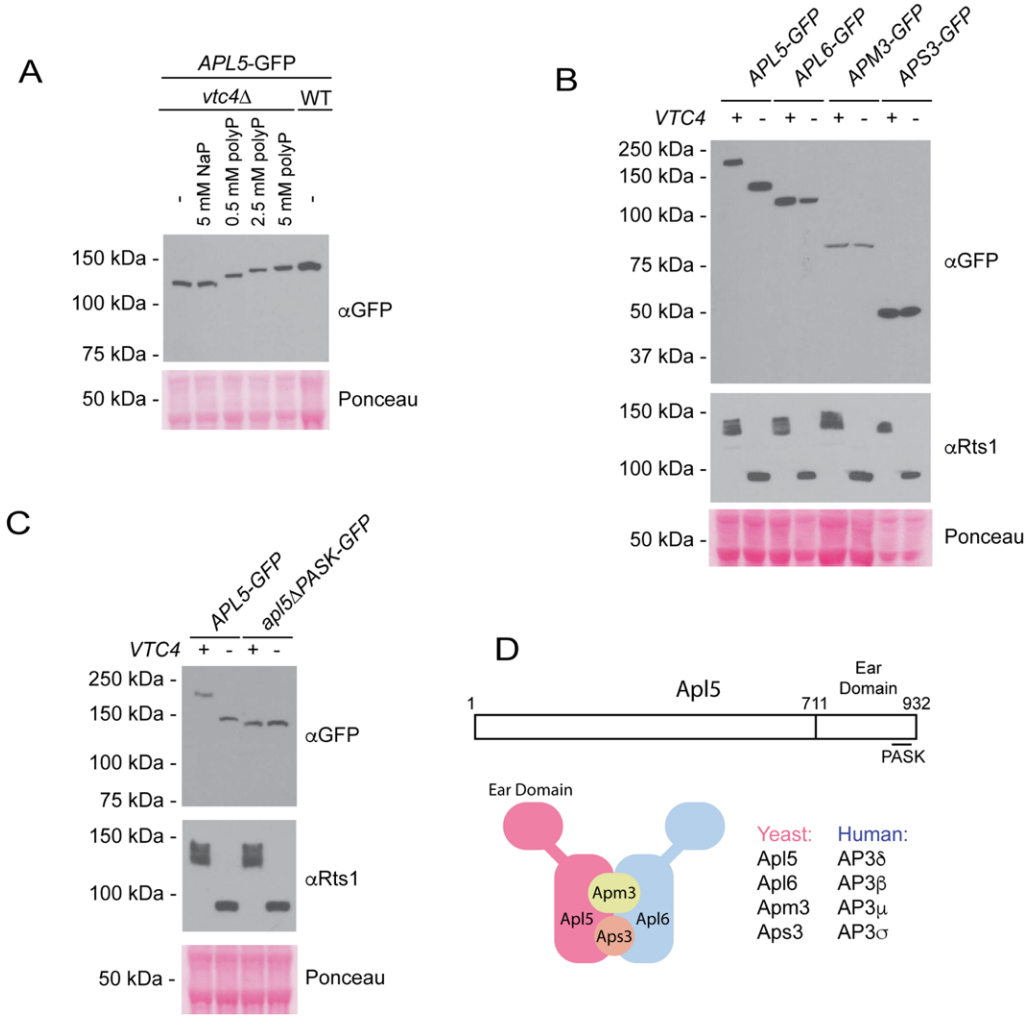


Fig 28: Analysis of Apl5 polyphosphorylation. A) Indicated concentrations of polyP (75 units) or sodium phosphate control were added to *vtc4*Δ mutant cell extracts expressing GFP-tagged Apl5 and polyP-induced shifts were analyzed by western blotting with an antibody against GFP

after NuPAGE analysis. Representative of *N* = 3. B) The indicated components of the AP3 complex were expressed as GFP fusions in WT and *vtc4*Δ strains and analyzed by western blotting following NuPAGE analysis of protein extracts. Representative of *N* = 3. C) The PASK motif of Apl5 is required for its polyphosphorylation. Protein extracts from the indicated strains were analyzed as described in B. Representative of *N* = 3. D) Representation of the conserved Apl5 protein and the AP3 complex. The schematic of the AP3 complex is adapted from that proposed previously (Cowles et al., 1997).

This interaction is thought to occur when Vps41 is integrated into the HOPS complex and promotes AP3 docking to the vacuole membrane, which allows proper delivery of vacuolar cargo (Angers and Merz, 2009). The other large subunit of the AP3 complex, Apl6, also has a PASK motif (Bentley-DeSousa et al., 2018), although we did not detect reproducible *Vtc4*-dependent electrophoretic shifts of this protein on NuPAGE gels (**Figure 28B**). Notably, Apl6 only has two lysines in its PASK motif, compared to 13 for Apl5 (**Figure 29A**).

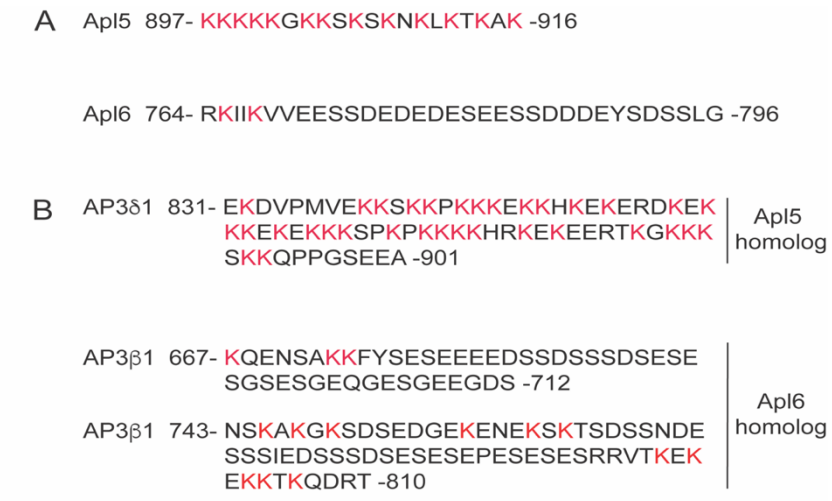


Fig. 29: PASK sequences within the AP3 complex. A) PASK sequences of Apl5 and Apl6. Numbers indicate amino acid positions. B) PASK sequences of AP3β1 and AP3δ1. Numbers indicate amino acid positions. Relates to **Figure 28**.

Since we previously found that electrophoretic shifts due to polyphosphorylation scale in proportion to the number of lysines in the PASK motif (Bentley-DeSousa et al., 2018), it is possible that Apl6 polyphosphorylation is not easily detected via this method.

AP3 mutant strains were previously described as having low levels of polyP (Freimoser et al., 2006). AP3 could impact polyP synthesis by transporting proteins involved in polyP synthesis or storage to the vacuole (Besteiro et al., 2008; Freimoser et al., 2006; Huang et al., 2011). In the context of this model, polyphosphorylation of AP3 may serve as means of feedback to regulate AP3 function. Importantly, the impact of AP3 on polyP metabolism could be mediated by multiple cargos, as over 200 genes are known to impact polyP synthesis in S. cerevisiae (Freimoser et al., 2006). In humans, mutations in the AP3b and AP3d subunits of the AP3 complex, homologs of Apl6 and Apl5, respectively (**Figure 28D**) (Cowles et al., 1997), give rise to Hermansky–Pudlak syndrome (HPS) subtypes 2 and 10 (Ammann et al., 2016; Mohammed et al., 2018; Wenham et al., 2010). HPS is a pleiotropic disease characterized by defects in the transport of proteins to lysosome-related organelles (Huizing et al., 2000; Huizing et al., 2002). Notably, a hallmark of HPS is bleeding diathesis and a failure to accumulate or retain polyP in the dense granules of platelets (Huizing et al., 2000; Huizing et al., 2002). PolyP has a well-characterized role in the blood coagulation cascade and works by stimulating activation of several clotting factors (Travers et al., 2015). AP3b1, AP3b2, and AP3d also have PASK motifs as defined by our criteria (**Figure 29B**), and Azevedo et al. (Azevedo et al., 2018) provided evidence that AP3b1 can be polyphosphorylated. Direct regulation by polyP may be a conserved feature of AP3 biology. As with Prb1, where within the transit pathway to the vacuole AP3 subunits interact with polyP is a critical question that remains to be answered. Altogether, our work expands the scope of polyphosphorylation in yeast and provides new avenues for exploring the functional consequences of this intriguing PTM. Expanding the catalog of polyphosphorylated proteins in other eukaryotes will be a critical step in understanding the molecular and cellular functions of polyP in these systems.

3.5 Acknowledgements

We thank the members of the Downey lab for critical reading of the manuscript and reviewers for insightful comments. We thank Dr. Elizabeth Jones' lab for the anti-Prb1 antibody and T. Shiba for polyP standards. ABD was supported by an Ontario Graduate Scholarship. This work was funded by a Canadian Institutes of Health Research Project Grant to MD (PJT- 148722). We also acknowledge funding from a Rare Disease Foundation Microgrant (grant number 1937).

4.0: Manuscript #3

Vtc5 is localized to the vacuole membrane by the conserved AP-3 complex to regulate polyphosphate synthesis in budding yeast

Publication Information:
Bentley-DeSousa, A., and Downey, M. (2020). Vtc5 is localized to the vacuole membrane by the conserved AP-3 complex to regulate polyphosphate synthesis in budding yeast. Submitted to *BioRxiv* and *in review* at mBio.

Author's Contribution:
ABD and MD conceived and carried out experiments.

Contributions by figure: Figure 30 (ABD), Figure 31 (ABD), Figure 32 (ABD), Figure 33 (ABD), Figure 34 (ABD), Figure 35 (ABD), Figure 36 (ABD), Figure 37 (ABD), Figure 38 (ABD), Figure 39 (ABD).
MD contributed data analysis and/or preparation to all figures.

Vtc5 is localized to the vacuole membrane by the conserved AP-3 complex to regulate polyphosphate synthesis in budding yeast

Amanda Bentley-DeSousa,[1,2], and Michael Downey[1,2,3,*]

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

[2]Ottawa Institute of Systems Biology, Ottawa, Ontario K1H 8M5, Canada

[3]Lead contact

*Correspondence: mdowne2@uottawa.ca

4.1: Abstract

Polyphosphates (polyP) are energy-rich polymers of inorganic phosphates assembled into chains ranging from 3-1000s of residues in length. They are thought to exist in all cells on earth and play roles in an eclectic mix of functions ranging from phosphate homeostasis to cell signaling, infection control, and blood clotting. In the budding yeast *Saccharomyces cerevisiae*, polyP chains are synthesized by the vacuole-bound VTC complex, which synthesizes polyP while simultaneously translocating it into the vacuole lumen where it is stored at high concentrations. VTC's activity is promoted by an accessory subunit called Vtc5. In this work, we find that the conserved AP-3 complex is required for proper Vtc5 localization to the vacuole membrane. In human cells, previous work has demonstrated that mutation of AP-3 subunits gives rise to Hermansky-Pudlak Syndrome, a rare disease with molecular phenotypes that include decreased polyP accumulation in platelet dense granules. In yeast AP-3 mutants, we find that Vtc5 is rerouted to the vacuole lumen by the ESCRT complex, where it is degraded by the vacuolar protease Pep4. Cells lacking functional AP-3 have decreased levels of polyP, demonstrating that membrane localization of Vtc5 is required for its VTC stimulatory activity *in vivo.* Our work provides insight into the molecular trafficking of a critical regulator of polyP metabolism in yeast. We speculate that AP-3 may also be responsible for the delivery of polyP regulatory proteins to platelet dense granules in higher eukaryotes.

4.2: Introduction

Polyphosphates (polyP) are chains of inorganic phosphates found in all cell types studied to date. PolyP chains are variable in size, ranging from 3-1000s of units in length, and are linked together via high-energy phosphoanhydride bonds (Rao et al., 2009). While once dismissed as 'molecular fossils', recent work suggests that polyP plays critical roles in diverse processes across both prokaryotic and eukaryotic organisms including bacterial virulence, infection control, blood coagulation, protein folding, and diverse aspects of cell signaling (Bondy-Chorney et al., 2020; Dahl et al., 2017; Gray et al., 2014; Roewe et al., 2020; Travers et al., 2015; Xie and Jakob, 2019). As such, polyP chains have gained significant interest as a potential target for therapeutics in a wide variety of pathologies. In contrast to bacteria and fungi, the enzymes that synthesize polyP chains in higher eukaryotic cells are largely unknown. There are no clear homologs of either prokaryotic or fungal polyP synthetases in mammals. Recent work by the Abramov group suggests that the mammalian mitochondrial F_0F_1 ATPase has polyP synthesis capabilities (Baev et al., 2020), but the overall contribution of this enzyme to total cellular pools of polyP remains to be tested.

One model organism used to study polyP at a foundational level is the budding yeast *S. cerevisiae*. Here, polyP is present in high concentrations (>200 mM), and the enzymes responsible for its metabolism have been identified (Auesukaree et al., 2004; Gerasimaite and Mayer, 2016; Kornberg et al., 1999). In yeast, polyP is synthesized by the vacuole transporter chaperone (VTC) complex. The minimal (core) VTC complex is composed of 3 subunits: Vtc1, Vtc2 or Vtc3, and Vtc4 (the catalytic enzyme) (Gerasimaite and Mayer, 2016; Hothorn et al., 2009). The Vtc3 subunit mostly localizes to the vacuole membrane alongside Vtc1 and Vtc4 (Hothorn et al., 2009). On the other hand, the Vtc2 subunit is found at the nucleus and cell periphery, and relocalizes to the vacuole membrane under phosphate-limited conditions (Hothorn et al., 2009). PolyP synthesis by VTC requires simultaneous translocation of polyP

into the vacuole lumen (Gerasimaite et al., 2014), where it makes up over 10 % of the dry weight of the cell (Kornberg et al., 1999). PolyP is also found in lower concentrations in the cytoplasm, plasma membrane, mitochondria, and nucleus, although observed concentrations vary by method of analysis (Denoncourt and Downey, 2021). In addition to being the only known polyP synthetase in yeast, the VTC complex has been demonstrated or suggested to play a role in a myriad of cellular activities such as the stability of vacuolar V-ATPase subunits, in vacuole fusion, and in microautophagy (Muller et al., 2002; Muller et al., 2003; Uttenweiler et al., 2007). PolyP itself has also been implicated in the regulation of pH balance (Eskes et al., 2018), ion homeostasis (Nguyen et al., 2019; Ryazanova et al., 2016; Trilisenko et al., 2019), and phosphate metabolism (Desfougeres et al., 2016; Eskes et al., 2018; Huang et al., 2002; Ogawa et al., 2000).

Recent work has identified several exciting aspects of VTC regulation. First, activity of the complex is promoted by binding of inositol pyrophosphate InsP7 to the SPX domains of VTC proteins (Gerasimaite et al., 2017; Wild et al., 2016). As such, loss of Kcs1, which catalyzes the formation of InsP7, drastically reduces polyP levels (Azevedo et al., 2015; Lonetti et al., 2011). Although canonical (serine/threonine) phosphorylation and lysine ubiquitylation sites have been identified on multiple VTC subunits (Albuquerque et al., 2008; Beltrao et al., 2012; Holt et al., 2009), their functions are currently unknown. Recently, a new VTC regulatory subunit termed Vtc5 was identified (Desfougeres et al., 2016). Vtc5 localizes exclusively to the vacuole membrane and interacts with the VTC complex to increase the rate of polyP production (Desfougeres et al., 2016). PolyP levels in *vtc5*Δ mutants are reduced to 20% of those of wild-type cells (Desfougeres et al., 2016). The mechanism by which Vtc5 exerts its positive effects on VTC is unknown, although it appears to function independently of effects imparted by inositol pyrophosphates (Desfougeres et al., 2016). Given the importance of Vtc5 as a regulator of the

core VTC complex, we sought to identify the pathways that are responsible for localizing Vtc5 to the vacuole membrane.

Membrane proteins are synthesized by ribosomes at the endoplasmic reticulum and transported to the trans-golgi network (TGN) prior to being sorted to their final destination (Feyder et al., 2015). Vacuolar proteins are sorted by two well established protein transport pathways, the CPY (indirect / Carboxypeptidase Y) or AP-3 (direct / Adaptor Protein Complex 3) transport pathways. The CPY pathway transports cargoes to the vacuole in an indirect fashion using the endosomal system as an intermediate path. In this pathway, cargoes localize to endosomes prior to fusion with the vacuole membrane for delivery (**Fig. 30A**)(Feyder et al., 2015). In contrast, AP-3 cargoes are selected at the TGN, where the complex buds from the TGN with its cargoes to create AP-3 coated vesicles. These vesicles are transported directly to the vacuole through the cytoplasm, and AP-3 docks at the vacuole membrane to release its cargoes (**Fig. 30A**)(Feyder et al., 2015). There are links between the AP-3 complex and polyP storage in higher eukaryotes. In humans, mutations in AP-3 are associated with a rare disorder called Hermansky-Pudlak Syndrome (Ammann et al., 2016; El-Chemaly and Young, 2016; Huizing et al., 2020). Hermansky-Pudlak Syndrome patients lack the ability to synthesize lysosome-related organelles in diverse cell types throughout the body, which results in broad phenotypes that include albinism, visual impairment, and bleeding problems (Huizing et al., 2020). Notably, this includes the inability to generate dense granules in platelets, where polyP is usually stored in its highest concentrations (130 mM) (Ruiz et al., 2004). PolyP promotes blood clotting by acting on proteins involved in the coagulation cascade including factor XII (Puy et al., 2013), factor XI/V (Choi et al., 2011, 2015), and thrombin (Smith et al., 2006). PolyP addition re-establishes clotting stimulatory activity of platelets derived from Hermansky-Pudlak patients (Muller et al., 2009).

In this study, we describe a role for the yeast AP-3 complex in localizing Vtc5 to the vacuole membrane. In AP-3 mutants, Vtc5 is rerouted to the vacuole lumen by the Endosomal Sorting Complex Required for Transport (ESCRT) complex where it is degraded by the vacuolar protease Pep4. Mislocalization of Vtc5 in AP-3 mutants is accompanied by a decrease in VTC protein levels and decreased levels of polyP. Overall, our study explains the polyP accumulation defects in AP-3 mutants and provides novel insights into the regulation of the VTC polyP synthetase in budding yeast.

4.3: Results

Vtc5 localization to the vacuole membrane is disrupted in AP-3 mutants

The CPY and AP-3 pathways comprise the two major routes of protein transport to the vacuole from the TGN (**Fig. 30A**). To test which of these pathways is responsible for Vtc5 localization to the vacuole membrane, we used live-cell fluorescence microscopy to analyze GFP-Vtc5 localization in AP-3 and CPY pathway mutants. Notably, all N-terminal GFP-Vtc fusions used in this work are functional, with the GFP tag facing the cytoplasmic side of the vacuole membrane to ensure the tag is not degraded within the vacuole lumen(Desfougeres et al., 2016; Hothorn et al., 2009). These GFP fusions are expressed from constitutive promoters integrated at endogenous *VTC* loci (See Experimental Procedures). In wild-type cells, GFP-Vtc5 (green) localized exclusively to the vacuole membrane as demonstrated by colocalization with FM4-64 (magenta), a dye which labels the vacuole membrane (**Fig. 30B**). CPY (*pep12*Δ) mutants showed GFP-Vtc5 localization similar to that of wild-type cells (**Fig. 30B**). Conversely, GFP-Vtc5 was mislocalized to the vacuole lumen in AP-3 mutant (*apl5*Δ) cells (**Fig 30B**). AP-3 is highly conserved from yeast to humans. It consists of a heterotetramer of four protein subunits: two large subunits (Apl5/AP3β1and Apl6/AP3δ1), one medium subunit (Apm3/AP3μ1), and one small subunit (Aps3/AP3σ1) (**Fig. 31A**)(Odorizzi et al., 1998). Deletion of genes encoding any one of the four AP-3 subunits results in defects in AP-3 cargo transport (Cowles et al., 1997). Therefore, to corroborate our results with *apl5*Δ, we examined GFP-Vtc5 localization in *aps3*Δ cells and again observed its mislocalization to the vacuole lumen in this mutant (**Fig. 30C**). Overall, these data suggest that the primary pathway responsible for localizing GFP-Vtc5 to the vacuole membrane is the AP-3 transport pathway, with little if any compensation by the CPY pathway.

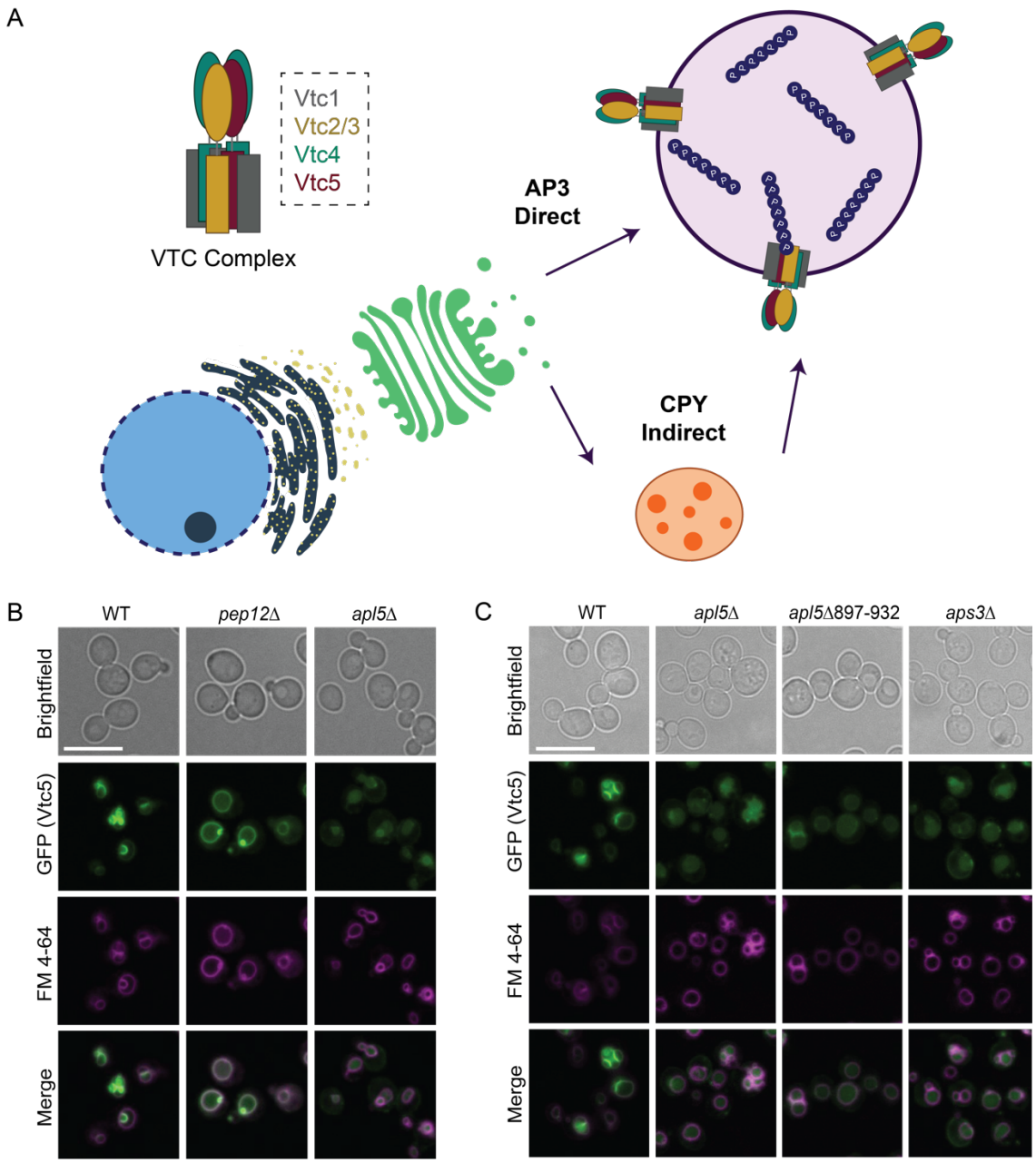


Fig. 30: GFP-Vtc5 is localized to the vacuole membrane via the conserved AP-3 pathway. A) Simplified schematic of protein transport to the vacuole in *S. cerevisiae*. B) The indicated strains were grown in YPD prior to incubation with FM 4-64, which accumulates in the vacuole membrane, for 2 hours. Cells were then washed with fresh YPD for 30 minutes, transferred to synthetic media, and imaged. Live-cell fluorescence microscopy was performed using a Leica DMI 6000B microscope at 63X with oil immersion. For AP-3 mutants, green (GFP-Vtc5) images were taken at longer exposure times to account for differences in signal intensity. Images were processed in FIJI. Images were adjusted by altering min/max thresholds for visualization purposes. Scale bar represents 10 μm. C) The indicated strains were processed as in (B).

Intriguingly, polyP chains can be covalently attached to protein targets as a post-translational modification termed polyphosphorylation (Azevedo et al., 2015; Azevedo and Saiardi, 2016). Polyphosphorylation is the non-enzymatic addition of polyP chains onto lysine residues, principally within poly-acidic, serine, and lysine (PASK) rich clusters (Azevedo et al., 2015; Bentley-DeSousa and Downey, 2019). We previously reported that Apl5 is polyphosphorylated in its C-terminus PASK cluster (amino acids 897-932)(McCarthy et al., 2020). Therefore, to test if polyphosphorylation impacts Apl5's role in localizing GFP-Vtc5 to the vacuole membrane, we first deleted the PASK cluster in its entirety. Deletion of Apl5's PASK cluster (*apl5*Δ897-932) resulted in mislocalization of GFP-Vtc5 to the vacuole lumen, similar to what we observed in AP-3 null mutants (**Fig. 30C**). To define the contribution of polyphosphorylation more specifically, we generated a strain wherein Apl5 is expressed with 13 lysine to arginine (K-R) amino acid substitutions in its PASK cluster. Polyphosphorylation results in an electrophoretic shift of target proteins separated on Bis-Tris NuPAGE gels (Azevedo et al., 2015; Bentley-DeSousa et al., 2018), and this is currently the only method described for evaluation of this new modification. As demonstrated by NuPAGE analysis, the resulting mutant ($Apl5_{PASK}$K-R) was unable to undergo polyphosphorylation. (**Fig. 31B**). However, this did not impact GFP-Vtc5 localization (**Fig. 31C**), nor did it mimic other known *apl5*Δ phenotypes such as defects in the maturation of AP-3 target vacuolar alkaline phosphatase Pho8 (Stepp et al., 1997) (**Fig. 31D**), or enhanced sensitivity to nickel chloride (Ruotolo et al., 2008) or rapamycin (Kapitzky et al., 2010) (**Fig. 32A-B**). We note that the PASK cluster lies within Apl5's Vps41 binding domain. Vps41 is a member of the Homotypic Fusion and Protein Sorting (HOPS) complex required for vesicle docking at the vacuole and delivery of AP-3 cargoes (Angers and Merz, 2009; Rehling et al., 1999). As such, the role of the Apl5 PASK in GFP-Vtc5 delivery may stem from disruption of this interaction rather than a defect in polyphosphorylation. Interestingly, the Apl5 C-terminal PASK deletion mutant (*apl5*Δ897-932) showed defects in maturation of Pho8, but it did not

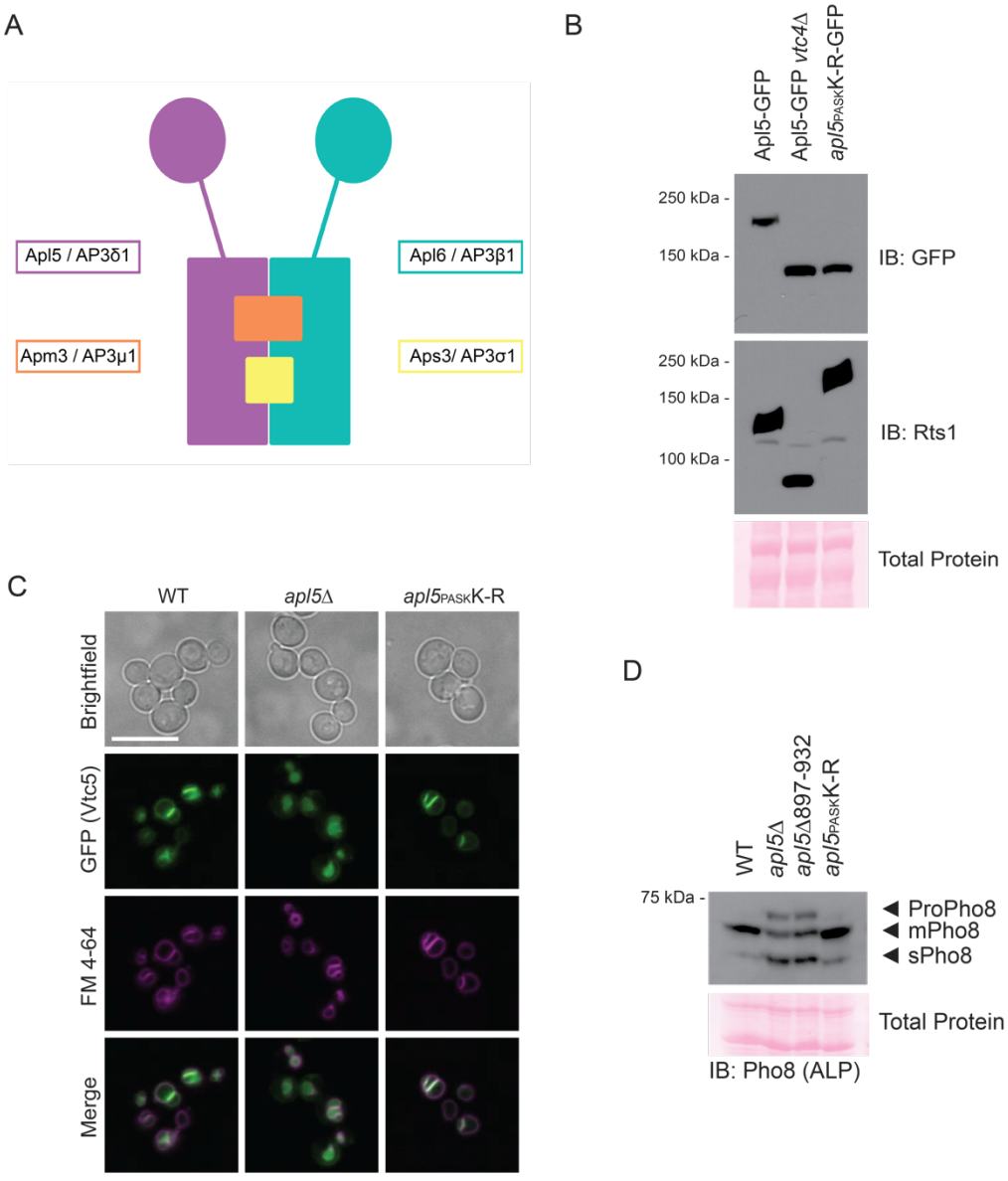


Fig. 31: Polyphosphorylation of Apl5 does not impact protein transport. A) The AP-3 complex is a conserved heterotetramer which includes two large subunits (Apl5/AP3β1 and Apl6/AP3δ1), one medium subunit (Apm3/AP3μ1), and one small subunit (Aps3/AP3σ1). B) Mutating lysines to arginines (K-R) within Apl5's PASK domain causes a collapse in the electrophoretic shift on NuPAGE, indicating a loss of polyphosphorylation. Proteins from the indicated strains were extracted using a TCA protein extraction protocol, electrophoresed on a 4-12% NuPAGE gel, and transferred to a PVDF membrane. The membrane was developed with autoradiography film after immunoblotting with an anti-GFP antibody to detect C-terminal Apl5-GFP fusion alleles. Anti-Rts1 serves as a positive control for polyphoshorylation. Ponceau S stain is used as a loading control. C) GFP-Vtc5 is localized to the vacuole membrane in Apl5$_{PASK}$K-R mutants. Cells were grown in YPD prior to incubation with FM 4-64, which accumulates in the vacuole membrane, for 2 hours. Cells were then washed with fresh YPD for

30 minutes, transferred to synthetic media, and imaged. Live-cell fluorescence microscopy was performed using a Leica DMI 6000B microscope at 63X with oil immersion. In AP-3 null mutant (*apl5*Δ), green (GFP-Vtc5) images were taken at longer exposure times to account for differences in signal intensity. Images were processed in FIJI. Images were adjusted by altering min/max thresholds for visualization purposes. Scale bar represents 10 μm. D) Apl5$_{PASK}$K-R does not impact maturation of AP-3 target Pho8. Proteins were extracted from the indicated strains using a TCA protein extraction protocol, electrophoresed on a 12% SDS-PAGE gel, and transferred to a PVDF membrane. The membrane was developed with autoradiography film after immunoblotting with an anti-Pho8 antibody. Ponceau S stain is used as a loading control.

result in sensitivity to nickel chloride or rapamycin (**Figs. 31C, 32A-B**). This separation of function mutant may serve as a useful tool to dissect AP-3's role underlying these distinct phenotypes.

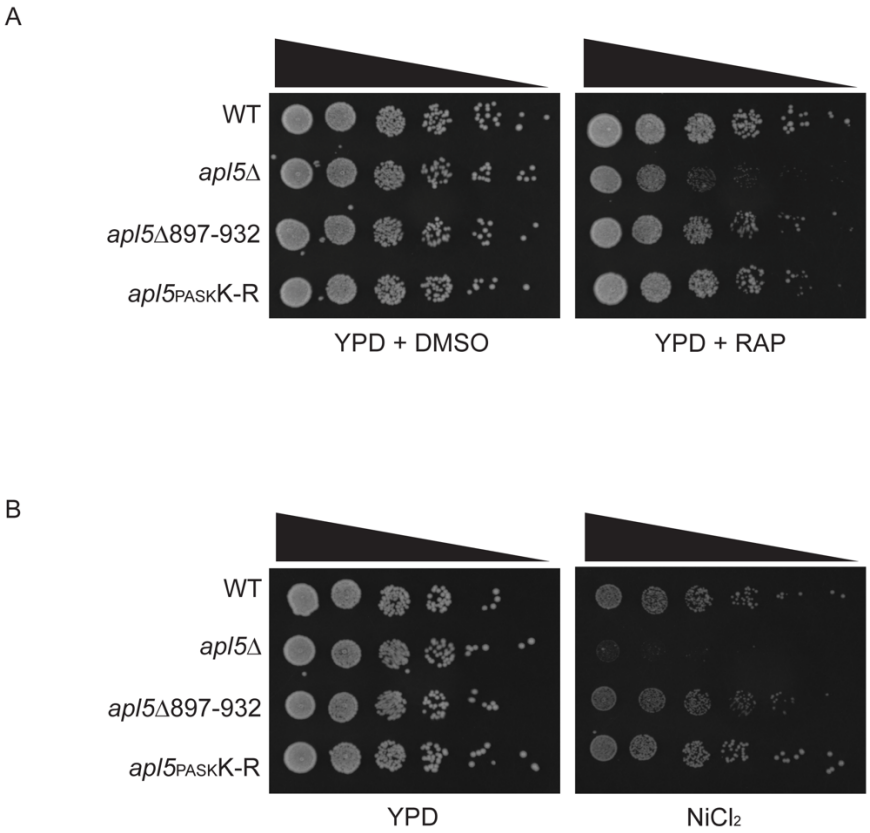


Fig. 32: Deletion of the Apl5 PASK motif does not impact sensitivity to rapamycin or nickel chloride. A) Functional AP-3 is required for growth in the presence of 0.003 μM rapamycin. Cells were grown in YPD to log phase and diluted to 0.1 OD_{600} prior to 5-fold serial dilutions in H_20. 4 μL of each dilution was spotted on the indicated media prior to incubation at

30 °C for 2-4 days. Images were taken on a Bio-Rad Chemi-Doc. DMSO is used as a control for rapamycin treatment. B) Functional AP-3 is required for growth in the presence of 0.75 mM $NiCl_2$. The same methodology was used as described for Figure 32A with YPD plates used as a control.

We next sought to test whether vacuolar localization of additional VTC subunits was also impacted by *APL5* deletion. We observed no difference in the localization of GFP-Vtc3 in wild-type versus *apl5*Δ mutants (**Fig. 32A**). GFP-Vtc4 showed an intermediate phenotype, with increased localization to the cytoplasm in *apl5*Δ, with some protein also remaining at the vacuole membrane (**Fig. 32B**). These data suggest that although the AP-3 pathway is responsible for the proper vacuolar localization of GFP-Vtc5, other pathways may contribute to the localization of core VTC subunits. Given the importance of Vtc5 as a regulator of VTC.

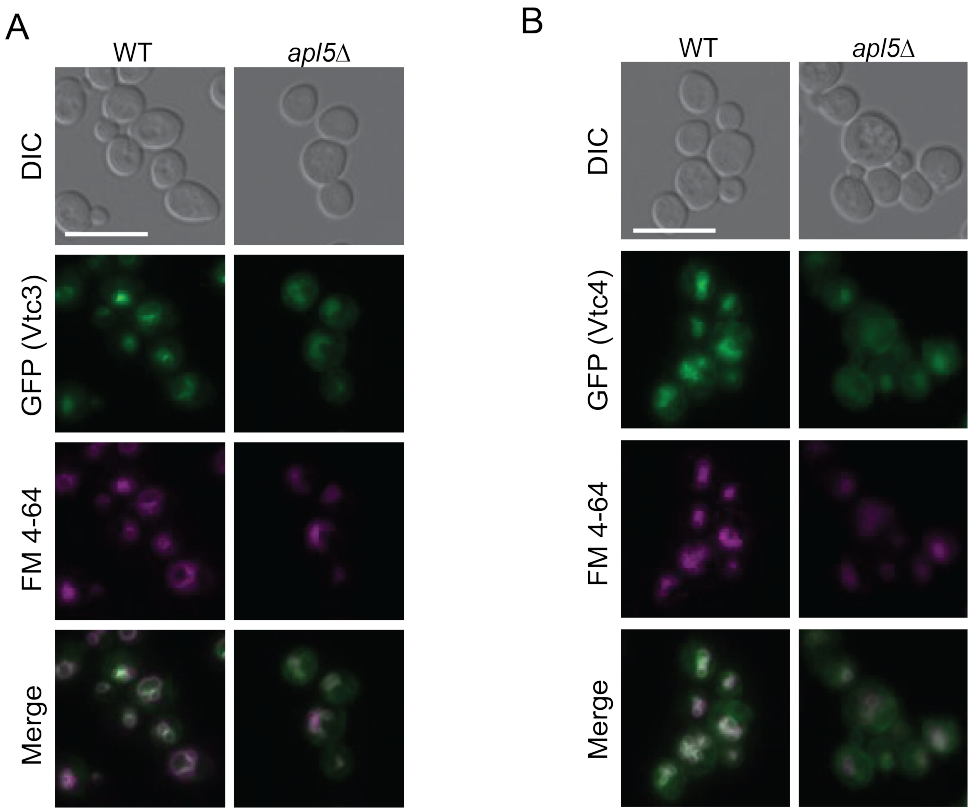


Fig. 33: Partial disruption of GFP-Vtc3 and GFP-Vtc4 localization in AP-3 mutants
A) The indicated strains were grown in YPD media prior to incubation with FM 4-64, which accumulates in the vacuole membrane, for 2 hours. Cells were then washed with fresh YPD for 30 minutes, transferred to synthetic media, and imaged. Live-cell fluorescence microscopy was

performed using a Zeiss AxioObserver 7 at 63X with oil immersion. Images were processed in FIJI. B) The indicated strains were processed as in (A). Scale bar represents 10 μm.

GFP-Vtc subunits are degraded in cells lacking functional AP-3

We next used western blotting to gain insight into the fate of mislocalized GFP-Vtc5. Relative to wild-type controls, *apl5*Δ and *aps3*Δ cells showed a striking accumulation of free GFP, which is known to be resistant to degradation (**Fig. 34A-B**) (Chalfie et al., 1994). This was not the result of increased protein expression, as full length GFP-Vtc5 was reduced in these mutants (**Fig. 34A & 34C**). This same pattern has been observed previously for other proteins mislocalized to the vacuole lumen and is attributed to cargo degradation (Li et al., 2015; Pokrzywa et al., 2009). We also tested if GFP-Vtc4 and GFP-Vtc3 levels were similarly affected. Both GFP-Vtc4 and GFP-Vtc3 also accumulated free GFP at the expense of decreased full-length fusions, although the effect was not as dramatic as that observed for GFP-Vtc5 (**Fig. 34D-34I**). Previous work from the Mayer group showed that *vtc5*Δ cells have decreased levels of core VTC subunits at the vacuolar membrane (Desfougeres et al., 2016). Degradation of GFP-Vtc3 and GFP-Vtc4, and partial mislocalization of GFP-Vtc4 in *apl5*Δ cells, are consistent with a model wherein mislocalized Vtc5 is largely non-functional in the absence of AP-3. However, we cannot exclude that these molecular phenotypes stem in part from defects in the transport of other AP-3 cargos (see Discussion).

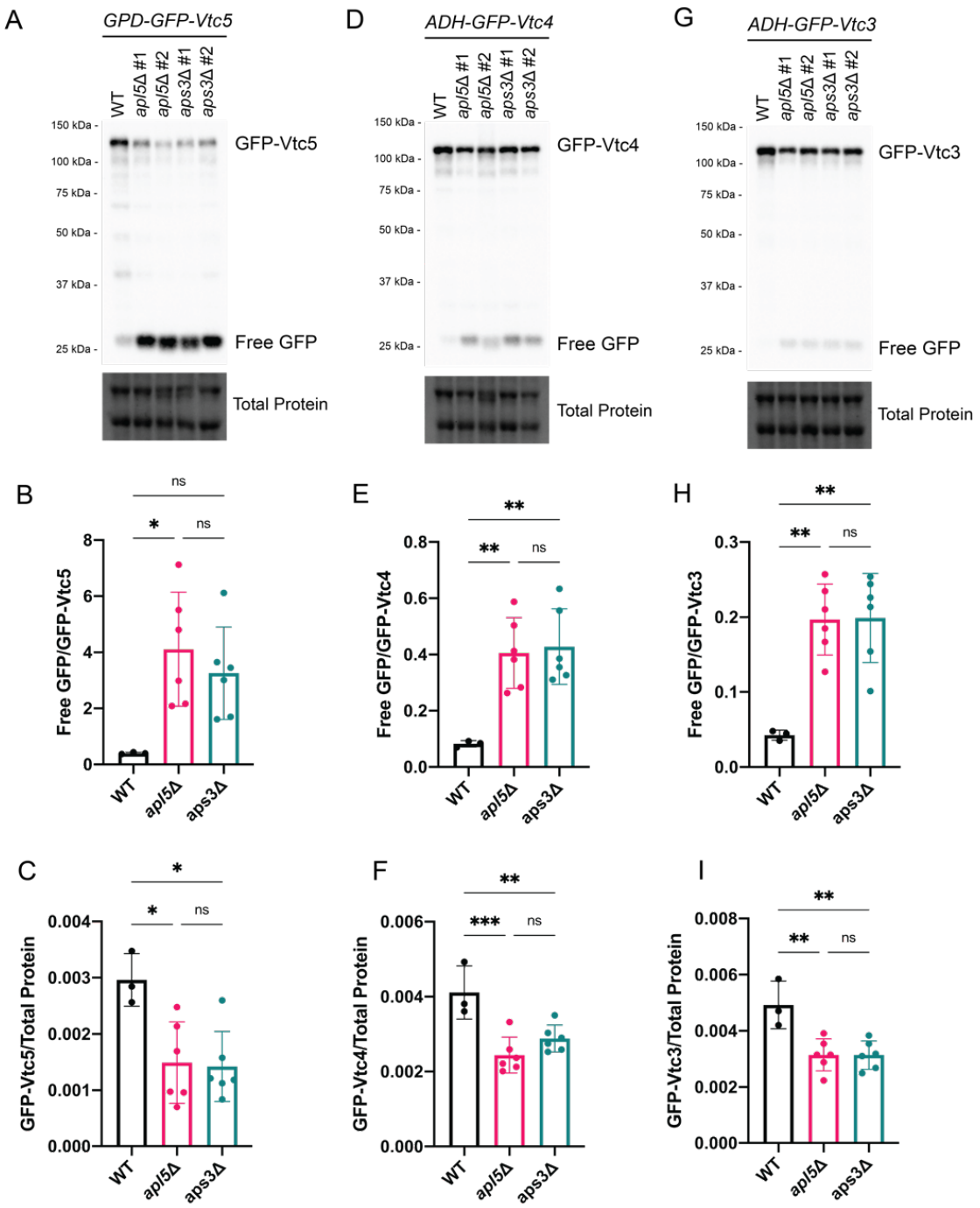


Fig. 34: AP-3 mutation causes degradation of GFP-Vtc proteins. A) Upon AP-3 mutation (*apl5*Δ and *aps3*Δ), full-length GFP-Vtc5 protein levels are reduced with a concomitant increase in free GFP. Proteins were extracted from the indicated strains using a TCA protein extraction protocol, separated on a 10% BioRad TGX Stain-Free™ FastCast™ acrylamide gel, and transferred to a nitrocellulose membrane. The membrane was imaged on a BioRad ChemiDoc after immunoblotting with anti-GFP. Total protein was imaged as a loading control. B-C) Quantification of free GFP/GFP-Vtc5 ratios and GFP-Vtc5 full length levels. Quantifications were done using BioRad ImageLab software and graphs were created using Prism GraphPad

Software. One-way ANOVAs were performed with Tukey posthoc tests. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Error bars represent standard deviation of the mean. n = 3 for WT and n = 6 for AP-3 mutants. D)-F) In AP-3 mutants, GFP-Vtc4-expressing strains accumulate free GFP at the expense of full-length protein. Methodology was as in Figure 34A-C. G)-I) In AP-3 mutants, GFP-Vtc3-expressing strains accumulate free GFP at the expense of full-length protein. The same methodology was used as in Figure 34A-C**.**

Mislocalized Vtc5 is rerouted to the vacuole lumen by the ESCRT pathway

We next investigated which pathways are responsible for localizing GFP-Vtc5 to the vacuole lumen in the absence of functional AP-3, with a focus on the autophagy and ESCRT pathways. Autophagy is a process whereby cytoplasmic material becomes sequestered in vesicles that fuse to the vacuole to deliver their contents for degradation (Reggiori and Klionsky, 2013). The ESCRT pathway is responsible for detecting ubiquitylated transmembrane proteins to sort them into multi-vesicular bodies (MVBs) in the endocytic pathway for delivery to the vacuole (Henne et al., 2011). Disruption of ESCRT (*vps27*Δ), but not autophagy (*atg8*Δ) in AP-3 mutants resulted in a loss of GFP signal from the vacuole lumen (**Fig. 35A**). In contrast, neither pathway impacted GFP-Vtc5 localization in wild-type cells (**Fig. 36A**). When ESCRT complex subunits are mutated, the complex becomes defective for proper MVB biogenesis and fusion at the vacuole membrane (Henne et al., 2011). Indeed, in *apl5*Δ *vps27*Δ double mutants, GFP-Vtc5 appeared to accumulate in FM4-64 labelled MVBs around the vacuole membrane (**Fig. 35A**). Deletion of *VPS27* also resulted in reversal of free GFP accumulation in western blots (**Fig. 35B**). This same molecular phenotype was observed across multiple ESCRT complex mutants (ESCRT-0, -I, -II, and -III, **Fig. 36B)**.

The ESCRT pathway recognizes cargoes that have been ubiquitylated by E3 ubiquitin ligases (Henne et al., 2011; Schmidt and Teis, 2012b). These cargoes are then transported to the vacuole via MVBs that fuse with the vacuole membrane to deliver its contents. As one of the last steps in this process, the deubiquitinase Doa4 removes ubiquitin moieties from ESCRT-targeted

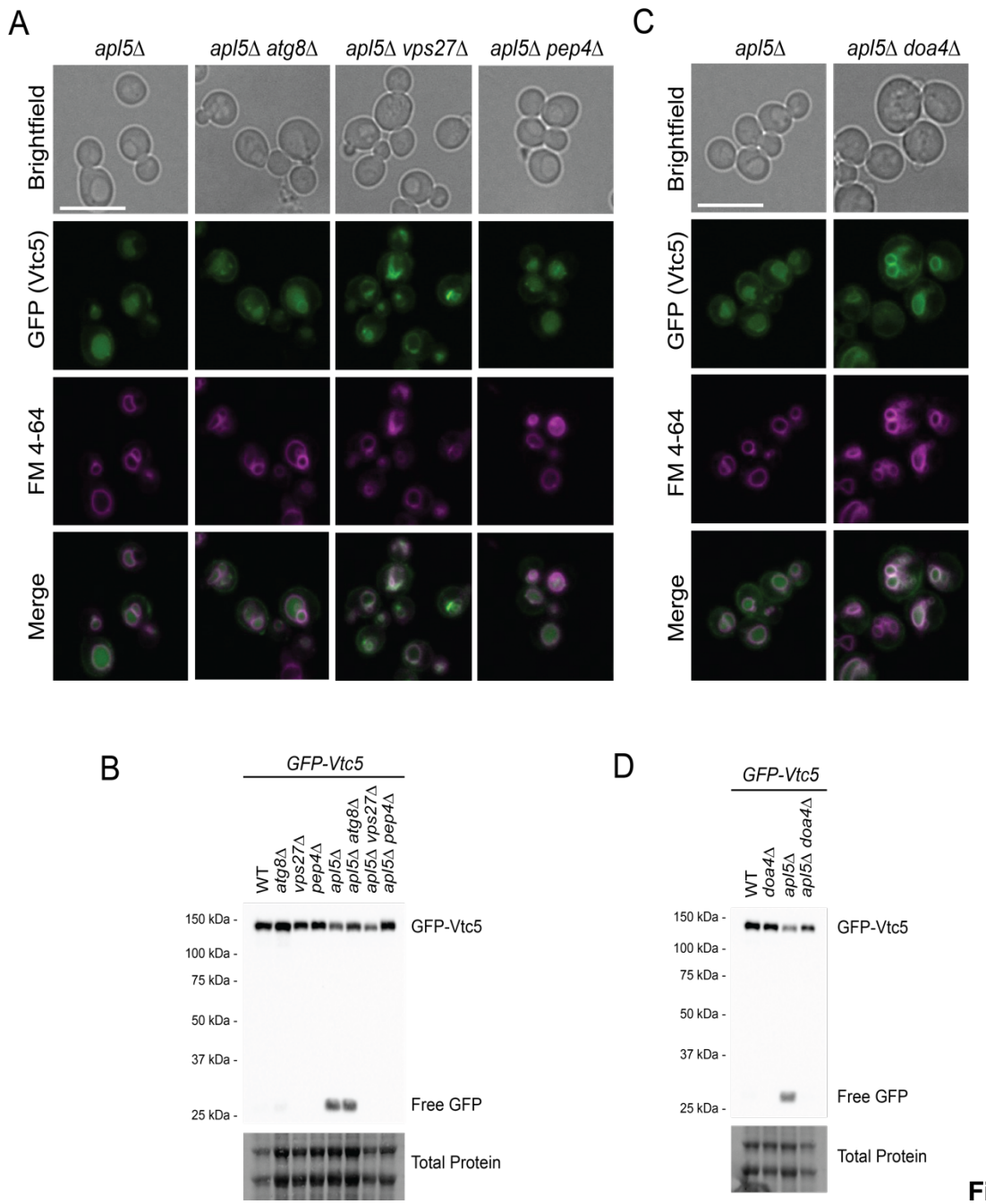


Fig. 35:

Degradation of mislocalized GFP-Vtc5 depends on ESCRT and the Pep4 protease. A) The mislocalization of GFP-Vtc5 upon AP-3 mutation is mediated by the ESCRT pathway. Cells were grown in YPD prior to incubation with FM4-64, which accumulates in the vacuole membrane, for 2 hours. Cells were then washed with fresh YPD for 30 minutes, transferred to synthetic media, and imaged. Live-cell fluorescence microscopy was performed using a Leica DMI 6000B at 63X with oil immersion. Images were processed in FIJI. Images were adjusted by altering min/max thresholds for visualization purposes. Scale bar represents 10 μm. B) Free GFP accumulation in GFP-Vtc5, mediated by AP-3 mutation, is reversed by ESCRT (*vps27*Δ) and *pep4*Δ mutation. Proteins from the indicated strains were extracted using a TCA protein

extraction protocol, separated on a 10% BioRad TGX Stain-Free™ FastCast™ acrylamide gel, and transferred to a nitrocellulose membrane. The membrane was imaged for total protein and probed using an anti-GFP antibody to detect GFP-Vtc5 protein using a BioRad ChemiDoc. C) GFP-Vtc5 degradation and free GFP accumulation is reversed by *DOA4* deletion. The same methodology was used as in Figure 34B. D) Deletion of *DOA4* rescues localization of GFP-Vtc5 in *apl5*Δ mutants. The same methodology was used as in Figure 35A.

proteins in order to recycle ubiquitin levels within the cell upon cargo delivery to the vacuole (Henne et al., 2011; Schmidt and Teis, 2012b). Mutation of *DOA4* alone has no impact on GFP-Vtc5 processing or localization (**Figs. 35C and 36C**). In AP-3 mutants, however, *doa4*Δ prevented the accumulation free GFP, mirroring what was seen for ESCRT mutants (**Fig. 35C**). However, in contrast to the ESCRT mutants, *doa4*Δ also restored GFP-Vtc5 localization to the vacuole membrane (**Fig. 35D**). Altogether, we conclude that Doa4 is a key player in localizing Vtc5 to the vacuole lumen in the absence of AP-3. Finally, when proteins accumulate in the lumen, they become accessible to vacuole proteases (Hecht et al., 2014). Deletion of the major vacuole protease Pep4 largely rescued both full-length GFP-Vtc5 and reversed the accumulation of free GFP in AP-3 mutants (**Fig. 35A-B**), suggesting that Pep4 is a major protease responsible for Vtc5 degradation.

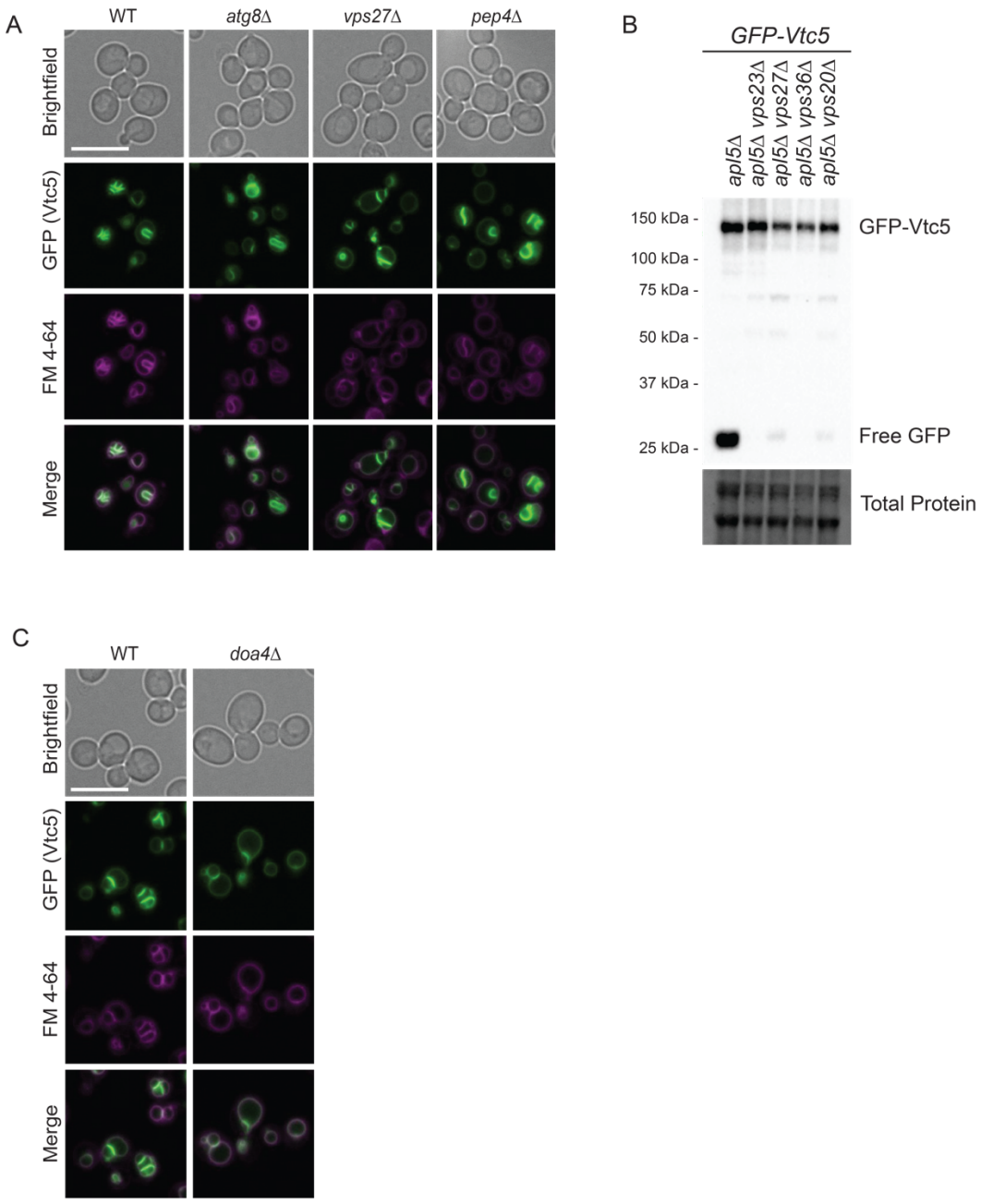


Fig. 36: ESCRT complex is responsible for GFP-Vtc5 degradation only when AP-3 is non-functional. A) In wild-type cells, GFP-Vtc5 remains localized to the vacuole membrane in autophagy- (*atg8*Δ), ESCRT- (*vps27*Δ), and protease-deficient (*pep4*Δ) cells. The indicated strains were grown in YPD prior to incubation with FM 4-64, which accumulates in the vacuole membrane, for 2 hours. Cells were then washed with fresh YPD for 30 minutes, transferred to synthetic media, and imaged. Live-cell fluorescence microscopy was performed using a Leica DMI 6000B microscope at 63X with oil immersion. Images were processed in FIJI. Images were adjusted by altering min/max thresholds for visualization purposes. Scale bar represents 10 μm. B) Mutations in ESCRT-0 (*vps27*Δ), ESCRT-I (*vps23*Δ), ESCRT-II (*vps36*Δ), and ESCRT-III

(*vps20*Δ) reverse free GFP accumulation in AP-3 mutants expressing GFP-Vtc5. Proteins were extracted using a TCA protein extraction protocol, electrophoresed on a 10% BioRad TGX Stain-Free™ FastCast™ acrylamide gel, and transferred to a nitrocellulose membrane. The membrane was imaged using a BioRad ChemiDoc after immunoblotting with an anti-GFP antibody to detect GFP-Vtc5. Total protein was imaged as a loading control. C) Deletion of *DOA4* does not impact wild-type GFP-Vtc5 localization. The same methodology was used as in Figure 35A.

The AP-3 complex is required for maintenance of wild-type polyP levels

Finally, we tested the contribution of AP-3 to polyP homeostasis. Deletion of AP-3 subunits resulted in a clear decrease in polyP levels (**Figure 37A**). This finding is consistent with previous data wherein AP-3 mutants were found to be important for polyP accumulation in a large scale screen (Freimoser et al., 2006). We observed a similar result in our strains used for prior analyses where GFP-Vtc5 is expressed under a constitutive *GPD1* promoter (**Figure 37B**). Together with our previous results, these data suggest that correct localization of Vtc5 to the vacuole membrane by the AP-3 complex is important for its function. Notably, AP-3 mutants, which have clear defects in GFP-Vtc5 localization to the vacuole membrane (**Figure 30**), still have more polyP than cells lacking Vtc5 altogether (*vtc5*Δ, **Figure 37C**). This observation could be explained by residual localization of Vtc5 to the vacuole membrane in AP-3 mutants. Interestingly, deletion of *DOA4*, which rescues GFP-Vtc5 protein levels and localization to the vacuole membrane, was unable to reverse the decreased polyP levels in *apl5*Δ mutants (**Figure 37D**). Thus, proper function of GFP-Vtc5 requires localization to the vacuolar membrane, specifically through the AP-3 pathway.

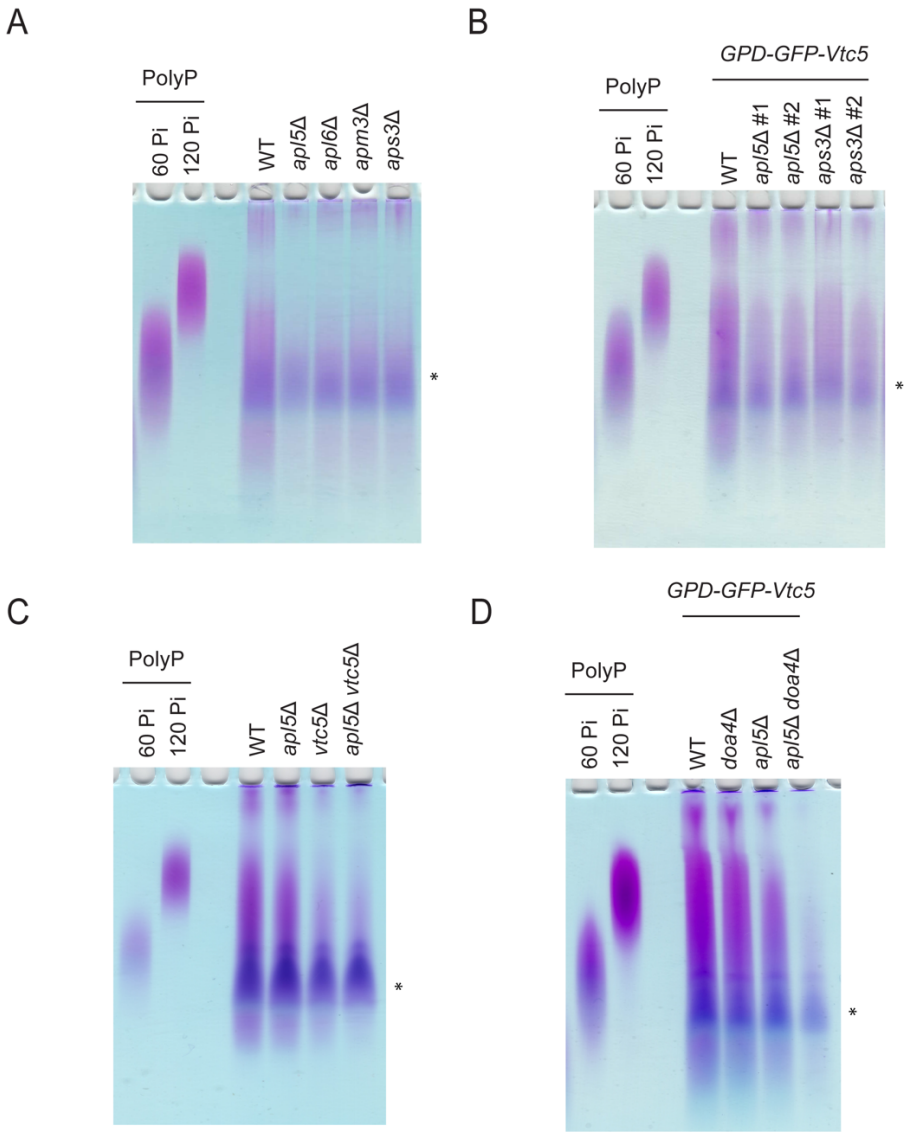


Fig. 37: Functional AP-3 is required for the maintenance of polyP levels. A) AP-3 subunit mutant causes a reduction in polyP levels. There is a loss of polyP when mutating each of the four AP-3 subunits (*apl5*Δ, *apl6*Δ, *apm3*Δ, and *aps3*Δ). PolyP was extracted as described in Experimental Procedures. After extraction, samples were mixed with polyP sample buffer and separated on a 15.8% TBE-Urea Acrylamide gel. The gel was incubated in fixing solution with toluidine blue for 15 minutes prior to destaining and imaging. PolyP standards (14, 60, and 120 Pi units) are included as standards. B) AP-3 mutants (*apl5*Δ and *aps3*Δ) cause a reduction in polyP levels when GFP-Vtc5 is expressed under a *GPD1* promoter. The same methodology was used as in Figure 36A. C) Mutating both AP-3 (*apl5*Δ) and *vtc5*Δ results in a largely epistatic decrease in polyP levels. The same methodology was used as in Figure 36A. D) Although *doa4*Δ is capable of rescuing GFP-Vtc5 localization and protein degradation when AP-3 is mutated (*apl5*Δ), it is incapable of restoring polyP levels. The same methodology was used as in Figure 36A.

4.4: Discussion

AP-3 regulation of Vtc5 localization

Interest in polyP research has experienced a resurgence in recent years on the heels of exciting connections between polyP and diverse aspects of cell signaling and protein homeostasis. With the key players involved in mammalian polyP metabolism largely uncharacterized, model systems have proved essential to our understanding of polyP dynamics and function. In *S. cerevisiae*, polyP is synthesized by the vacuole-bound VTC complex and stored at high concentrations in the vacuole (Gerasimaite et al., 2014). The activity of the core VTC complex (consisting of Vtc1, Vtc2 or Vtc3, and Vtc4) is increased dramatically by the Vtc5 subunit (Desfougeres et al., 2016). This is the only protein known to act directly on the VTC complex to stimulate polyP production. Our study supports a model where Vtc5 localization to the vacuole membrane depends on the evolutionarily conserved AP-3 complex. This finding provides insight into the regulation of the VTC complex and identifies Vtc5 as a cargo whose mislocalization underlies polyP accumulation defects observed in AP-3 mutants.

In wild-type cells (**Fig. 38)**, we propose that Vtc5 is sorted into AP-3 coated vesicles at the TGN and transported directly to the vacuole. The AP-3 complex selects protein cargoes by tyrosine-based motifs (YXXØ where X represents any amino acid and Ø is a bulky hydrophobic amino acid) and/or dileucine-based motifs ([D/E]XXXL[L/I] where X represents any amino acid) (Park and Guo, 2014). Notably, Vtc5 has 6 tyrosine-based motifs and 2 dileucine-based motifs which together could be used as a targeting signal. Alternatively, Vtc5 could be transported in conjunction with other AP-3 cargoes such as Vam3, which is also required for wild-type levels of polyP accumulation (Freimoser et al., 2006). Regardless, we propose that AP-3 coated vesicles may bind to Vps41 of the HOPS complex to facilitate docking and release of cargo, including Vtc5, into the vacuole membrane. In the absence of AP-3, Vtc5 is mislocalized to the vacuole lumen. Mislocalized Vtc5 (e.g. in *apl5Δ*), is recognized and sorted by the ESCRT pathway into

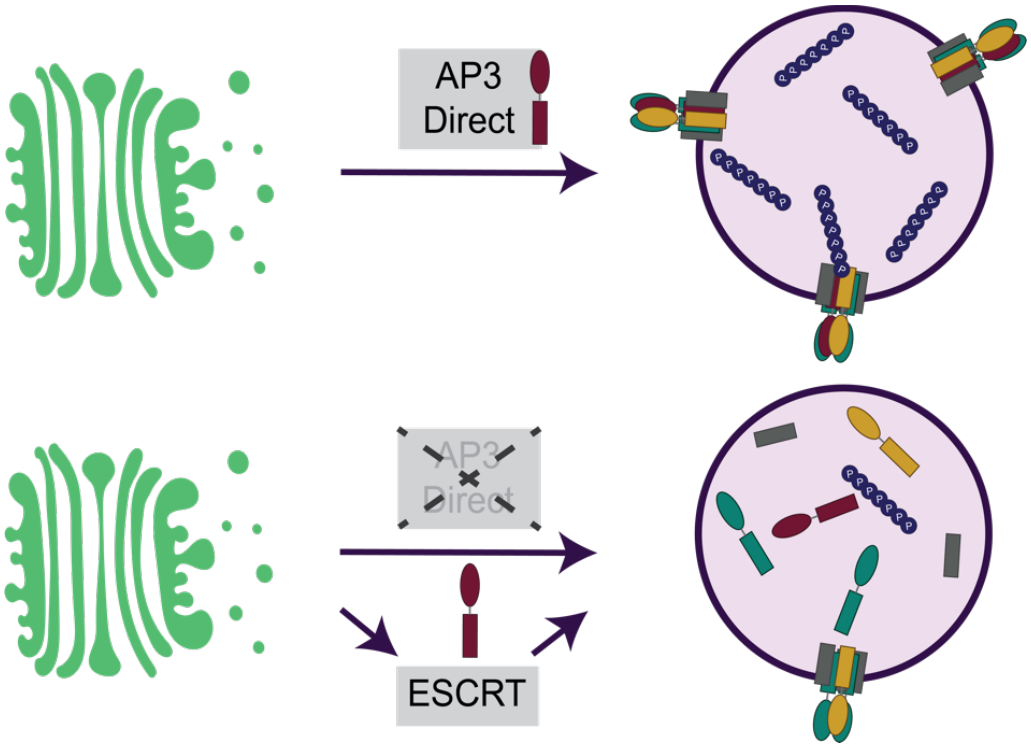


Fig. 38: GFP-Vtc5 is localized to the vacuole membrane via the AP-3 complex. In the absence of AP-3, GFP-Vtc5 is transported to the vacuole lumen via the ESCRT pathway where it becomes degraded in manner dependent on vacuole protease Pep4. See text for details.

the vacuole lumen via the endosomal system where it is eventually degraded in a manner dependent on vacuole protease Pep4. Since Pep4 is also required for the activation of vacuole proteases Prb1 and Prc1 (Distel et al., 1983; Mechler et al., 1982), it is possible that these also these play a role in Vtc5 degradation in the vacuole lumen.

We found that degradation of Vtc5 in AP-3 mutants requires the Doa4 deubiquitylase. The involvement of Doa4 suggests that mislocalized Vtc5 is subject to regulation by ubiquitylation. In support of this assertion, several large-scale studies have identified ubiquitylation sites in the middle and C-terminal regions of the protein (Oughtred et al., 2019). We speculate that these sites may be targets of the Rsp5 E3 ubiquitin ligase that has been implicated in the ESCRT-dependent delivery of proteins to the vacuole lumen (McNatt et al., 2007; Stringer and Piper, 2011; Zhu et al., 2017). Notably, Vtc5 is also rich in phosphoserines targeted by both the Cdk1

(cell cycle) and Nkk1(nitrogen metabolism) (Oughtred et al., 2019). These modifications could modulate Vtc5 localization in wild-type or AP-3 mutant cells or may instead be involved in regulating its activity towards the core VTC complex.

In contrast to GFP-Vtc5, we found that localization of GFP-Vtc3 to the vacuole membrane was not impacted by disruption of AP-3 and GFP-Vtc4 was only partially affected. However, deletion of AP-3 subunits still reduced GFP-Vtc3/4 protein levels and the increased appearance of free GFP in western blots, albeit to a lesser degree than that observed for GFP-Vtc5. We suggest that transport of core VTC subunits occurs through multiple transport routes, although these may function redundantly with AP-3. This type of dual regulation has been documented for other AP-3 cargos such as Sna4 (Pokrzywa et al., 2009) and Ypq1 (Llinares et al., 2015). It is also possible that changes observed in core subunits stem from mislocalized Vtc5. There is strong evidence that maximal polyP production requires VTC localization to the vacuole. However, polyP has also been detected in variable amounts at the plasma membrane, cytoplasm, mitochondria, and nucleus(Denoncourt and Downey, 2021). Whether vacuolar polyP is somehow transported to these compartments or if it is made by VTC residing locally remains an open question. In the latter case, it is possible that Vtc5 association (or lack thereof) with core VTC in these areas also determines subcellular concentrations of polyP outside of the vacuole. In future work it will be intriguing to identify pathways that regulate localization of GFP-Vtc2, which is found at the plasma membrane or peripheral endoplasmic reticulum in phosphate replete conditions, becoming enriched at the vacuole only under conditions of phosphate starvation (Gerasimaite and Mayer, 2016; Hothorn et al., 2009; Uttenweiler et al., 2007).

Our work shows that AP-3 mutant cells have reduced levels of polyP. This was true in both a wild-type background and under conditions of Vtc5 overexpression. These findings are consistent with work from Feimoser *et al.* identifying AP-3-encoding genes in a genome-wide screen for deletion mutations with defects in polyP accumulation (Freimoser et al., 2006).

Interestingly, deletion of *DOA4* in an AP-3 mutant background rescued both GFP-Vtc5's localization to the vacuole membrane and restored wild-type levels of full-length GFP-Vtc5 protein without restoring polyP levels. In fact, polyP levels in *apl5*Δ *doa4*Δ double mutants were lower than in either single mutant. While this may seem counterintuitive, we suggest that Vtc5 is not correctly positioned within the vacuole membrane under these circumstances and unable to stimulate VTC activity.

The reduction in polyP levels seen in AP-3 is not as dramatic as that observed in cells lacking Vtc5 altogether, suggesting either residual localization of Vtc5 to the vacuole membrane in AP-3 mutants, or that mislocalized Vtc5 remains competent to stimulate core VTC activity to some degree. Notably, Freimoser *et al*. identified over 200 additional genes that impact polyP metabolism (Freimoser et al., 2006), and some of these could function as direct regulators of Vtc5 or AP-3. We recently described the Apl5 subunit of AP-3 as a target of lysine polyphoshorylation (McCarthy et al., 2020). However, our analysis of mutant Apl5 that cannot be polyphosphorylated suggests that this modification does not impact GFP-Vtc5 delivery to the vacuole or other AP-3 related phenotypes that we tested. Polyphosphorylation of AP-3 may become important under select stress conditions that remain to be identified. Since polyP has been described as having chaperone activity (Gray et al., 2014), another possibility is that polyphosphorylation promotes degradation or refolding of a small fraction of AP-3 that is itself mislocalized to the polyP-rich vacuole lumen.

Conservation of the AP-3 complex

The AP-3 complex is highly conserved in mammalian cells in terms of subunit organization and function. For example, the human homolog of Ypq1, PQLC2, is transported to the vacuole membrane by the AP-3 complex when it is expressed in yeast (Llinares et al., 2015). Notably, mutations in human AP-3 give rise to Hermansky-Pudlak Syndrome. Relevant here is the observation that molecular phenotypes of Hermansky-Pudlak Syndrome patients include a failure to accumulate polyP in platelet dense granules, a type of lysosome-related organelle conceptually similar to the yeast vacuole. Based on the conserved role of AP-3 subunits in polyP metabolism, we suggest that identification and characterization of AP-3 cargoes in human cells that accumulate high levels of polyP may provide unique insights into the human polyP synthetases, the identity of which remains a critical open question in the field.

4.5: Methods

Yeast strains & handling:

Yeast strains were constructed using standard techniques via transformation of PCR products containing selectable markers. For gene deletions, PCR analyses were used to confirm the position of gene deletion and the absence of wild-type gene copy. PCR was also used to confirm the correct genomic location of cassettes used for epitope tagging. Genotypes for all strains used in this study are listed in **Table 9 (Appendix B)**. For all experiments performed strains were grown in YPD (2% glucose supplemented with 0.005% Adenine and 0.005% Tryptophan).

Apl5 K-R Mutagenesis:

A construct containing the lysine (K) to arginine (R) mutations within Apl5 PASK motif was created custom by GenScript and used for the experiments listed in **Figure 31 and Figure 32**. The construct was integrated with a GFP tag or flag tag at the endogenous *APL5* locus using transformation of overlapping PCR products.

Electrophoresis & Immunoblotting:

Methods for protein extraction were described previously (Bentley-DeSousa et al., 2018) and are summarized here using similar wording for clarity. A BioSpec bead-beater was used to lyse cell pellets corresponding to 3-6 OD_{600} units in the presence of 100 □L of acid washed beads and 300 □L of 20 % TCA (Sigma-Aldrich T6399). 2 x 3-minute pulses were used. The supernatant was recovered, and cells were washed with 300 μL of 5 % TCA (Sigma-Aldrich T6399). The supernatant from the second wash was combined with the first and this mixture was clarified by centrifugation at 4 °C at 16,000 g for 4 minutes. The supernatant was removed, and the pellet was resuspended in SDS-PAGE sample buffer (see buffer recipes section) supplemented with 1/10 volume of 1.5M Tris-HCl pH 8.8 (Tris Base Fisher BP152-5, Hydrochloric Acid Fisher A144-212) and 1/10 volume 1 M DTT (Bio Basic DB0058). Samples

were boiled for 5 minutes before an additional centrifugation at 4 °C at 16,000 g for 4 minutes. The supernatant was recovered and stored or used immediately for SDS-PAGE or NuPAGE analysis (Thermo Fisher NP0336). Where indicated, BioRad TGX Stain-Free™ FastCast™ 10% acrylamide (BioRad 1610183) was used for quantification using total protein in place of a loading control. When BioRad TGX acrylamide was used, nitrocellulose membrane (BioRad 162-0112) was used, and exposures were obtained using a Bio-Rad Chemi-Doc system. When non-TGX 12% SDS-PAGE and NuPAGE gels (Thermo Fisher NP0336) were used (as in Supplemental Fig. 1), PVDF membrane (BioRad 162-0177) was used and exposures were obtained using autoradiography film (DiaMed DIAFILM810). In all cases, Chemiluminescence Luminata Forte ECL (Fisher Scientific WBLUF0500) was used. All antibodies used for immunoblotting are described in **Table 10 (Appendix B).**

*Spot tests***:**

Cells were diluted to 0.1 OD_{600}, grown for 4 hours, then diluted to 0.1 OD_{600} prior to being serially diluted 5-fold in H_20. 4 μL of each dilution was spotted on the indicated media prior to incubation at 30 °C for 2-3 days. Plates contained the following chemicals/drugs: 0.028% DMSO (VWR CA97061-250), 0.75 mM $NiCl_2$ (Fisher N54-250), and 0.003 μM Rapamycin (Sigma-Aldrich R0395-1MG). Images were taken using a Bio-Rad Chemi-Doc.

*Microscopy***:**

Live-cell fluorescence imaging was conducted using a Leica DMI 6000B with the Volocity® imaging program. Briefly, cells were diluted to OD_{600} = 0.2 from overnight cultures in YPD. After 3 hours of growth at 30 °C, FM 4-64 (Thermo Fisher T13320; 1.64 mM stock in DMSO) was added to a final concentration of 1.64 μM for an additional 2 hours. Cells were then washed out in YPD for 30 minutes at room temperature. In instances where there was a difference in protein level expression, exposure times were taken at longer intervals to account for this discrepancy

(**Figure 30**). For microscopy in **Figure 33**, images were taken on a Zeiss AxioObserver 7 with a Hamamatsu ORCA-Flash LT camera using Zeiss Zen 3.0 Pro Software. All images were taken with oil immersion at 63X. All images were then analyzed in FIJI. Backgrounds were subtracted with a rolling ball radius of 50 pixels and images taken with FM4-64 dye were converted from red to magenta.

Polyphosphate Extractions

Polyphosphate was extracted from yeast pellets containing 8-12 OD_{600} units using an adapted protocol from Bru *et al* 2017 (Bondy-Chorney et al., 2020; Bru et al., 2017). Cells were resuspended in 400 μL of cold LETS buffer (see buffer recipes section). Subsequently, 600 μL of neutral phenol pH 8 (Sigma-Aldrich P4557) and 150 μL of mH_2O were added. Samples were vortexed for 20 seconds and heated for 5 minutes at 65°C followed by a 1-minute incubation on ice. 600 μL of chloroform (Sigma-Aldrich 472476) was added, samples were vortexed for 20 seconds and spun down at room temperature for 2 minutes at 13,000 rcf. The top layer was then transferred to a new tube containing 600 μL of chloroform (Sigma-Aldrich 472476), vortexed for 20 seconds and spun down at room temperature for 2 minutes at 13,000 rcf. The top layer was transferred to a new tube and 2 uL of RNAse A (10 mg/mL; Thermo Fisher R1253) and 2 μL of DNAse I (10 mg/mL; Thermo Fisher AM2222) were added followed by a 1-hour incubation at 37°C. The mixture was transferred to a pre-chilled tube containing 1 mL 100% ethanol (Commercial Alcohols P006EAAN) and 40 μL of 3M sodium acetate pH 5.3 (Sigma-Aldrich S7899). Samples were left at -20°C overnight and then centrifuged for 20 minutes at 13,000 rcf at 4°C. The pellet was washed in 500 μL of cold 70% ethanol (Commercial Alcohols P006EAAN), centrifuged for 5 minutes at 13,000 rcf at 4°C then the supernatant was discarded, and the pellet was dried. The pellet was resuspended in 20-30 μL mH_2O. Samples were mixed 1:1 with polyP loading dye (see buffer recipes section) and

electrophoresed on a 15.8% TBE-Urea Acrylamide Gel at 100V for 1 hour and 45 minutes in 1X TBE buffer (see buffer recipes section). The gel was incubated in fixing solution (see buffer recipes section) with toluidine blue for 15 minutes and then destained in destaining solution. PolyP standards (a gift from T. Shiba) were used to assess polyP chain length.

*Statistical analyses***:**

For statistical analyses performed on western blots in **Figure 34**, a one-way ANOVA was performed with Tukey posthoc tests at 95% confidence intervals. Error bars, p-values and number of biological replicates (n) are defined in the relevant figure legends.

Buffer recipes:

The following buffer recipes are taken from our previous study (Bentley-DeSousa *et al.*, 2018) and are listed here again verbatim for convenience.

SDS-PAGE Running Buffer (1X working):

100 mL of 10X 1 litre Stock [30.2 g Tris Base (Fisher BP152-5), 188 g Glycine (Fisher BP381-5), 10 g SDS (Fisher BP166)] and 900 mL ddH_2O. * Please note SDS-PAGE gels don't resolve polyP shifts*

NuPAGE Running Buffer (1X working):

50 mL of 20X 1 litre Stock [209.2 g MOPS (Sigma M1254), 121.1 g Bis-Tris (Sigma B9754), 20 g SDS (Fisher BP166), 12 g EDTA (Sigma ED2SS)], 5 mL of 1M Sodium Bisulfate (Fisher S654-500) and 950 mL ddH_2O.

SDS-PAGE Transfer Buffer (1X working):

100 mL of 10X 1L Stock [30.275 g Tris Base (Fisher BP152-5), 166.175 Glycine (BioBasic GB0235], 200 mL Methanol (Fisher A412P-4) and 700 mL ddH_2O.

NuPAGE Transfer Buffer (1X working):

50 mL of 20X 1L Stock [81.6 g Bicine (Sigma B3876), 104.8 g Bis-Tris (Sigma B9754*)*, 6 g EDTA Sigma ED2SS)], 200 mL Methanol (Fisher A412P-4) and 750 mL ddH_2O.

SDS-PAGE/NuPAGE Sample Buffer (3X stock):

800 μL of stock [160 mM Tris-HCl pH 6.8 (Tris base Fisher BP152-5, Hydrochloric acid Fisher A144-212), 6% SDS w/v (Fisher BP166), 30% Glycerol (Fisher BP229-4), 0.004 % Bromophenol Blue (Fisher BP115-25)]. For TCA preps, 3X is supplemented with 100 μL 1M DTT (BioBasic DB0058), and 100 μL 1.5M Tris-HCl (pH 8.8) [Tris base (Fisher BP152-5), Hydrochloric acid (Fisher A144-212)].

The following buffer recipes are used in Bentley-DeSousa *et al.,* 2020.

TBE (1X working):

200 mL of a 5X Stock [67.5 g Tris Base (Fisher BP152-5), 34.37 g Boric Acid (Fisher BP168-1), and 25 mL 0.5M EDTA (Sigma-Aldrich 03690)]

LETS Buffer (1X working):

100 mM LiCl (Fisher L120-500), 10 mM EDTA (Sigma-Aldrich 03690), 10 mM Tris-HCl pH 7.4 [(Tris Base Fisher BP152-5, Hydrochloric Acid Fisher A144-212)], 20% SDS (Thermo Fisher AM9820)

PolyP Loading Dye (6X):

10 mM Tris-HCl pH 7 [(Tris Base Fisher BP152-5, Hydrochloric acid Fisher A144-212)], 1 mM EDTA (Sigma-Aldrich 03690), 30% Glycerol (Fisher BP229-4), and Bromophenol Blue (Fisher BP115-25).

Toluidine Blue Fixing Solution:

25% Methanol (Fisher A412P), 5% Glycerol (Fisher BP229-4), 0.05% Toluidine Blue (Sigma-Aldrich T3260)

Destaining Solution:

25% Methanol (Fisher A412P), 5% Glycerol (Fisher BP229-4)

4.6: Acknowledgements

We thank A. Mayer for yeast strains, T. Shiba (Regentiss, Japan) for polyP standards and members of the Downey lab for critical reading of the manuscript. We also acknowledge the Cell Biology and Image Acquisition Core funded by the University of Ottawa, Ottawa, Canada and the Canada Foundation for Innovation for microscopy training and expertise. Work for this project was funded by a Canadian Institutes of Health Research (CIHR) Project Grant (PJT-148722), and by an Early Researcher Award from the Ontario Ministry of Innovation and Research to MD. ABD was funded in part by a graduate scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Chapter 5: Discussion

5.1: Polyphosphorylation: a novel PTM in yeast and human cells

To date, polyphosphorylation of both yeast and mammalian proteins has been studied (Azevedo et al., 2015; Azevedo et al., 2018; Bentley-DeSousa et al., 2018; McCarthy et al., 2020). Polyphosphorylation was originally discovered by the Saiardi lab on two yeast substrates, Nsr1 and Top1 (Azevedo et al., 2015). The original aim in manuscripts 1 (Chapter 2) and 2 (Chapter 3) were to expand our knowledge of substrate polyphosphorylation. Our lab has now increased the set of known yeast polyphosphorylated substrates to 25 (Bentley-DeSousa et al., 2018; McCarthy et al., 2020). We also discovered 6 polyphosphorylated human proteins (Bentley-DeSousa et al., 2018), a number which was later expanded by the Saiardi lab to a total of 14 (Azevedo et al., 2018). We approached the question of which other proteins are polyphosphorylated by using screen-based methods in both yeast and human cells with the help of bioinformatics datasets using a PASK motif description of a minimum of 20 amino acid stretch containing >75% acidic and serine residues, with at least one lysine (Bentley-DeSousa et al., 2018). Overall, both our labs have been instrumental in expanding the field of polyP biology by exploring protein polyphosphorylation.

5.2: Polyphosphorylation modifies 25 yeast substrates known to date

In order to expand the dataset of known polyphosphorylated substrates in yeast, we performed two high-throughput screens to test for protein polyphosphorylation of PASK-like proteins using NuPAGE analysis (Bentley-DeSousa et al., 2018; McCarthy et al., 2020). To date, compiling all data published on yeast polyphosphorylation, there are 25 polyphosphorylated yeast substrates (Azevedo et al., 2015; Bentley-DeSousa et al., 2018; McCarthy et al., 2020). These substrates are enriched in nuclear localization and the majority function in ribosome biogenesis and RNA biology. Additionally, we found that Vtc4 is required

for proper ribosome biogenesis when comparing polysome profiles of wild-type cells to *vtc4*Δ mutant cells (Bentley-DeSousa et al., 2018). One difficulty in studying non-enzymatic PTMs is that defining a precise phenotype for the PTM relies on creating specific mutant proteins that can no longer be modified. In the case of polyphosphorylation, mutating lysines to arginines within PASK motifs causes a loss of polyphosphorylation (Azevedo et al., 2015; Bentley-DeSousa et al., 2018). Therefore, in order to determine specifically if polyphosphorylation impacts ribosome biogenesis, one would have to systematically mutate endogenous PASK lysines to arginines for all of these substrates. It is possible that polyphosphorylation of some or all proteins involved in RNA biology is necessary for ribosome biogenesis. Additionally, it is possible polyP may be acting in a non-covalent manner to impart its effects on polysome assembly, however the creation of PASK mutants is critical for this distinguishment. Another functional enrichment for polyphosphorylated proteins is for roles in chromatin maintenance, which can affect replication and transcription (Azevedo et al., 2015; Bentley-DeSousa and Downey, 2019). One way to assess this would be to perform ChIP-seq and RNA-seq on wild-type and *vtc4*Δ cells. For example, Chz1 is a polyphosphorylated substrate (Bentley-DeSousa et al., 2018) that is a histone chaperone for H2A.Z/H2A-H2B dimers, acting to increase promoter accessibility in repressed regions (Luk et al., 2007). It is possible polyphosphorylation of Chz1 impacts its histone chaperone capability, altering transcription. Additionally, Eaf7 is a polyphosphorylated substrate (Bentley-DeSousa et al., 2018) of the NuA4 acetyltransferase complex (Mitchell et al., 2013; Rossetto et al., 2014). It would be interesting to assess whether polyphosphorylation of Eaf7 impacts substrate acetylation. In order to specifically link any changes to modification of these substrates specifically, similar mutational analysis would be required as previously discussed. Finally, after assessing the original 15 substrates (Bentley-DeSousa et al., 2018), we were incapable of demonstrating any changes in protein localization

related to polyphosphorylation similar to the changes observed originally for Nsr1 and Top1 (data not shown). We hypothesize that changes in protein localization related to polyphosphorylation may be substrate- or condition-specific (also see section 5.4 and 5.5).

After completing our screens, we performed bioinformatics analysis in an attempt to further understand what a PASK motif is. Despite expanding the original dataset, we were incapable of obtaining a more detailed description of what constitutes a PASK motif. There are two interesting approaches to address this: 1) perform a non-biased genome-wide screen for protein polyphosphorylation, and 2) create synthetic PASK motif constructs. For the former approach, one would use the same approach outlined in our manuscripts but done genome-wide without enrichment for PASK motifs. Alternatively, combinations of synthetic PASK-like constructs with an epitope tag (such as GFP) could be expressed in wild-type and *vtc4*Δ strains to define a minimal PASK motif. It could also be informative to perform bioinformatics analyses on PASK-like proteins that were found not to be modified. Ultimately, the exact function of a PASK motif remains unknown and will be critical to determine to advance the field. Overall, our expansion of known yeast polyphosphorylated substrates adds to the complexity of how polyP functions intracellularly and may impact protein function and regulation at the PTM level.

5.3: Towards developing experimental methods to detect polyphosphorylation

One limitation in the field of polyphosphorylation is that the only way to detect this PTM is via electrophoretic mobility shift on NuPAGE gels. This becomes an issue for a few reasons: 1) our current lack of knowledge of PASK motifs (as previously discussed), 2) a lack of understanding of mammalian polyP biology (see section 5.4), and 3) electrophoretic mobility shifts may only resolve highly modified proteins, akin to hyper-phosphorylation shifts. Both the Saiardi lab and our lab have proposed developing mass spectrometry methods to detect protein polyphosphorylation on a large-scale basis (Bentley-DeSousa and Downey, 2019). Previous

groups have been capable of detecting lysine phosphorylation via mass spectrometry (Hardman et al., 2019; Hardman et al., 2017). This could be non-canonical phosphorylation however it is also possible that these detected fragments could be remnants of polyphosphorylation. In this case, the polyP chains could be breaking down in the acidic conditions used during sample preparation. Given the heterogeneity of polyP chains, it becomes difficult to know how to analyze any mass spectrometry data. It would be interesting to go through these published datasets and assess whether any proteins with phosphorylated lysine residues are polyphosphorylated using our established methods (Bentley-DeSousa et al., 2018). Overall, the field of polyphosphorylation would benefit greatly from a development like this as it would allow the possibility for unbiased large-scale analysis. One could then interrogate whether specific cellular stressors or drug treatments alter the polyphosphorylation of substrates on a high throughput basis.

5.4: EcPPK1: a novel system to study human polyP biology and polyphosphorylation

Polyphosphorylation of human proteins was demonstrated for the first time by our lab (Bentley-DeSousa et al., 2018). Since then, the Saiardi lab have developed an assay to assess human protein polyphosphorylation on a larger scale (Azevedo et al., 2018). Combining these datasets, there are 14 polyphosphorylated substrates in humans. In our case, we took advantage of *E. coli*'s polyP synthetase (*Ec*PPK1) to produce polyP intracellularly within HEK293T cells to increase the amount of polyP from endogenous levels. We screened for human polyphosphorylated substrates based on bioinformatics analysis, conservation to known yeast substrates, and antibody availability (Bentley-DeSousa et al., 2018). The Saiardi lab then published 8 polyphosphorylated proteins by taking advantage of a proteome chip and incubating it with 8-unit biotinylated polyP (Azevedo et al., 2018). To confirm results from their screen, they incubated HeLa extracts with 100-unit polyP then performed NuPAGE analysis and native

PAGE analysis to assess protein polyphosphorylation and electrostatic interactions respectively. Interestingly, there were many differences in our datasets. None of the proteins we uncovered to be modified were positive hits in their screens and we obtained different results for HPS90B1 polyphosphorylation. A simple explanation for the first difference is that the proteins we found to be polyphosphorylated cannot bind polyP as small as 8-units in length and therefore would not have been a hit in their screen. Additionally, we found HPS90B1 was not polyphosphorylated in HEK293Ts using our experimental setup (Bentley-DeSousa et al., 2018), while they found HPS90B1 to be polyphosphorylated. Overall, this suggests there may be variability in human protein polyphosphorylation which may depend on antibodies, isoforms, cell lines, and/or polyP chain length. This also may suggest there are protein-specific requirements for polyphosphorylation *in vivo* that may not be available *in vitro*. For example, there may be factors that promote or inhibit protein polyphosphorylation that are inactive *in vitro*. Finally, it is possible that different cell lines have variance in protein polyphosphorylation. In fact, we have observed differences in protein polyphosphorylation between HEK293Ts, HeLas, and MDA-MB-231s using *Ec*PPK1 expression (data not shown). This could be explained by differences in *Ec*PPK1 expression and/or polyP production as our lab has shown different cell lines produce different amounts of polyP using our system (Bondy-Chorney et al., 2020). More interestingly, these observations could have fascinating biological relevance. Despite this, the concept of cell-type specific protein polyphosphorylation would be an interesting project to pursue in the future.

5.5: Polyphosphorylation may be implicated in human disease

Previous research has demonstrated that polyP is associated with tumourigenesis and cancer (Arelaki et al., 2018; Boyineni et al., 2020; Jimenez-Nunez et al., 2012; Wang et al., 2003). All of the known human polyphosphorylated substrates either have prognostic or mutational associations with cancers (Azevedo et al., 2018; Bentley-DeSousa et al., 2018;

Uhlen et al., 2010), therefore it is possible protein polyphosphorylation may contribute to its development. One example that warrants further investigation is the proto-oncogene DEK. Our lab found that upon *Ec*PPK1 transfection, DEK is polyphosphorylated (Bentley-DeSousa et al., 2018) and its localization is drastically altered. DEK is normally localized to the nucleus and cytoskeletal fractions, however, upon *Ec*PPK1 expression, DEK becomes largely localized to the cytoplasm and mitochondrial fractions (Bondy-Chorney et al., 2020). Therefore, it is possible that modification of this protein alters its functions *in vivo* and may contribute to its role in cancer biology. Alternatively, polyphosphorylation could impact protein function in other examples like the blood coagulation cascade. Given that polyP accumulates in platelet dense granules (Ruiz et al., 2004), it is logical to hypothesize that proteins within this cell signalling pathway may be polyphosphorylated. One interesting candidate is thrombin since lysine residues are critical for polyP binding (Mutch et al., 2010b). Investigating whether polyP acts as a PTM in examples where protein function is essential for human health is an avenue that remains to be explored.

5.6: The regulation of polyP as seen through the trafficking of VTC

Overall, there remain a lot of open questions in the novel field of polyphosphorylation. In both yeast and mammals, polyphosphorylation could be regulated endogenously by specific cellular stressors. Alternatively, protein polyphosphorylation of some substrates may be a means to transport polyP intracellularly. Overall, all of these hypotheses require prior knowledge of how polyP is regulated in general *in vivo*. One of the areas that remains to be explored in the field is how the VTC complex is regulated in yeast. Since the AP-3 complex was already shown to be required for proper polyP maintenance in yeast and humans (Freimoser et al., 2006; Muller et al., 2009), and Apl5, an AP-3 subunit, is polyphosphorylated (McCarthy et al., 2020), we decided to address this question from a protein trafficking standpoint. We found that the VTC subunit Vtc5 becomes localized to the vacuole membrane via the AP-3 pathway (Bentley-

DeSousa and Downey, *in review*). Upon AP-3 mutation, Vtc5 is mislocalized to the vacuole lumen and VTC subunits are degraded, resulting in the accumulation of free GFP moieties. The mislocalization of Vtc5 is largely imparted by the ESCRT complex and vacuolar protease Pep4. We hypothesize this mislocalization underlies the loss of polyP in AP-3 mutants and in patients with HPS. This research provides insight to how the VTC complex is being regulated, however, this work raises many interesting questions which will be elaborated on in subsequent sections.

5.7: Vtc5 as a means to regulate VTC complex function

Although these data elaborate on previous work demonstrating Vtc5 is critical for polyP production (Bentley-DeSousa and Downey, *in review*), how Vtc5 actually stimulates polyP production still remains unknown. The Mayer lab proposed that Vtc5 may act to stabilize the VTC complex at the vacuole membrane for optimal polyP production (Desfougeres et al., 2016). This hypothesis is in line with our data demonstrating the degradation of VTC subunits upon AP-3 mutation, however, the specific details of how it may be stabilizing the complex remain to be explored. One could perform experiments to assess subunit stability using cycloheximide chase assays. Alternatively, the interaction of subunits could be assessed using Bimolecular Fluorescence Complementation (Sung and Huh, 2007) and co-immunoprecipitation assays. Finally, one of the most interesting open questions is whether Vtc5 is being localized via the AP-3 pathway in a direct or indirect manner. We demonstrated this pathway is critical for proper Vtc5 localization to the vacuole membrane, however, it is possible that there is a secondary, intermediate, protein mediating this localization. For example, the AP-3 complex could transport a cargo that is in return responsible for Vtc5 localization. In this case, AP-3 would not be transporting Vtc5 directly as a cargo. In order to investigate this, assessing Vtc5 localization after mutating its putative AP-3 binding motifs will be informative. Another interesting result from our work was that Apl5 subunit polyphosphorylation did not impact AP-3's function in Pho8 and

Vtc5 transport. Ultimately, this suggests Apl5 polyphosphorylation may be tightly regulated *in vivo*, and its role as a PTM may be critical under cellular changes where polyP concentrations need to be altered wherein Apl5 polyphosphorylation may act as a sensor to either inhibit or enhance polyP accumulation. There is evidence in *Candida humicola* that polyP accumulates more when grown in acidic media (pH 5.5) (McGrath and Quinn, 2000). Here, one could test if Apl5 polyphosphorylation impacts Vtc5 localization when changing the acidity of growth media or in other different conditions that may cause alterations in the regulation of polyP biology.

5.8: In AP-3 mutants, Vtc5 is localized to the vacuole lumen in an ESCRT-dependent manner

As previously mentioned, we found that upon AP-3 mutation, Vtc5 becomes rerouted to the vacuole lumen in an ESCRT-dependent manner, where it becomes degraded in a Pep4-dependent manner (Bentley-DeSousa and Downey, *in review*). Traditionally, proteins following a similar localization phenotype and degradation have been found to be ubiquitylated and targeted by ESCRT (Henne et al., 2011; Li et al., 2015; Schmidt and Teis, 2012a). Since we demonstrated the deubiquitinase Doa4 is important in this phenotype, we hypothesize ubiquitylation of Vtc5 may be a critical event upon AP-3 mutation. It is possible that upon AP-3 mutation Vtc5 becomes rerouted to the plasma membrane originally, where it then becomes endocytosed and targeted by ESCRT. This localization to the plasma membrane is characterized for other AP-3 cargoes (Badolato and Parolini, 2007; Yang et al., 2000). Using a method that allows time-dependent tracking (Casler and Glick, 2020) of Vtc5 localization may provide interesting information regarding the localization of Vtc5 between its original mislocalization in AP-3 mutants to its final destination in the vacuole lumen. Finally, it is possible that ESCRT-mediated degradation of Vtc5 may also involve microautophagic pathways (Schuck, 2020). We determined that macroautophagy did not impact Vtc5 localization by mutating *Atg8*, however, both microautophagy and ESCRT may work in concert.

5.9: Subcellular localization of other VTC subunits may use other pathways

Subcellular localization of other VTC subunits may use other trafficking pathways. The majority of our work focused on the regulation of the Vtc5 subunit, however, we also assessed whether Vtc3 and Vtc4 localization were altered in AP-3 mutants and found that Vtc4 contains a partial disruption of localization (Bentley-DeSousa and Downey, *in review*). We also found degradation of both Vtc3 and Vtc4 in AP-3 mutants. These data may suggest a role for Vtc5 in maintaining VTC stability and/or localization at the vacuole membrane. Since Vtc3 and some Vtc4 remained localized at the vacuole membrane, it is possible other VTC subunits are localized in an AP-3-independent manner. In this case, these subunits may be localized via the CPY pathway which would create another interesting layer of VTC regulation, by having two separate protein trafficking pathways responsible for proper localization that can be regulated independently. Finally, as the VTC complex is localized to other subcellular compartments (Hothorn et al., 2009), but Vtc5 is entirely vacuolar (Desfougeres et al., 2016), other trafficking pathways must be responsible for its localization elsewhere at the cell. As previously mentioned, the Vtc2 subunit is predominantly localized to the plasma membrane and nucleus, however upon phosphate starvation it relocalizes to the vacuole membrane (Hothorn et al., 2009). Understanding how other VTC subunits are localized will provide us more information about its trafficking and may help our understanding how polyP is localized to other subcellular compartments. Performing high-throughput microscopy screens for VTC subunit localization using the yeast collection deletion set (Giaever and Nislow, 2014) would be an interesting approach. Alternatively, adding specific localization signals to VTC subunits (for example, a nuclear or mitochondrial localization signal) will help our understanding of VTC's function outside the vacuole.

5.10: The AP-3 approach: a means to uncover the elusive mammalian polyP synthetase?

As previously mentioned, there are no proposed candidate polyP synthetases in mammals to date, despite many efforts by polyP biologists. In order to understand mammalian polyP biology and to enhance our ability to manipulate it *in vivo*, it is imperative to uncover the mammalian polyP synthetase. A potential way to approach answering this question is through evolutionary enzymology. In this case, analyzing the folding and 3D structure of known polyP synthetases may contain similarities to some mammalian proteins. Alternatively, we propose interrogating the AP-3 interactome in mammals. In particular, assessing AP-3 binding proteins in platelets, where dense granules are known to store polyP at high concentrations (Ruiz et al., 2004), may provide a list of candidate polyP synthetases. Given the eukaryotic conservation of the AP-3 transport pathway (Odorizzi et al., 1998), it is possible the elusive polyP synthetase is being regulated in a similar fashion as we demonstrated for Vtc5. Interestingly, PQLC2 is a mammalian AP-3 cargo and it is capable of being expressed in yeast and localized to the vacuole membrane (Llinares et al., 2015). This further exemplifies the extent of conservation of this pathway and its function. One way to address this would be to perform BioID (Sears et al., 2019) to get an understanding of which proteins AP-3 may be interacting with. Alternatively, mass spectrometry may be performed on platelets in wild-type and AP-3 mutant cells. Once a list of candidate polyP synthetases is obtained, we can approach validation in two complementary ways: 1) expression in yeast and search for polyP production and localization (similar to PQLC2) and, 2) knockdown or knockout of candidates in human cells, followed by analysis of polyP concentrations and/or protein polyphosphorylation. Further testing would be required but this creative way of answering a question that has persisted for over half a century and could provide us with evidence towards our understanding of mammalian polyP biology.

Chapter 6: References

Akiyama, M., Crooke, E., and Kornberg, A. (1993). An exopolyphosphatase of Escherichia coli. The enzyme and its ppx gene in a polyphosphate operon. J Biol Chem *268*, 633-639.

Albuquerque, C.P., Smolka, M.B., Payne, S.H., Bafna, V., Eng, J., and Zhou, H. (2008). A multidimensional chromatography technology for in-depth phosphoproteome analysis. Mol Cell Proteomics *7*, 1389-1396.

Ammann, S., Schulz, A., Krageloh-Mann, I., Dieckmann, N.M., Niethammer, K., Fuchs, S., Eckl, K.M., Plank, R., Werner, R., Altmuller, J.*, et al.* (2016). Mutations in AP3D1 associated with immunodeficiency and seizures define a new type of Hermansky-Pudlak syndrome. Blood *127*, 997-1006.

Andreeva, N., Ledova, L., Ryazanova, L., Tomashevsky, A., Kulakovskaya, T., and Eldarov, M. (2019). Ppn2 endopolyphosphatase overexpressed in Saccharomyces cerevisiae: Comparison with Ppn1, Ppx1, and Ddp1 polyphosphatases. Biochimie *163*, 101-107.

Andreeva, N., Trilisenko, L., Eldarov, M., and Kulakovskaya, T. (2015). Polyphosphatase PPN1 of Saccharomyces cerevisiae: switching of exopolyphosphatase and endopolyphosphatase activities. PLoS One *10*, e0119594.

Angelova, P.R., Agrawalla, B.K., Elustondo, P.A., Gordon, J., Shiba, T., Abramov, A.Y., Chang, Y., and Pavlov, E. (2014). In Situ Investigation of Mammalian Inorganic Polyphosphate Localization Using Novel Selective Fluorescent Probes JC-D7 and JC-D8. ACS Chem Biol *9*, 2101-2110.

Angers, C.G., and Merz, A.J. (2009). HOPS interacts with Apl5 at the vacuole membrane and is required for consumption of AP-3 transport vesicles. Mol Biol Cell *20*, 4563-4574.

Arelaki, S., Arampatzioglou, A., Kambas, K., Sivridis, E., Giatromanolaki, A., and Ritis, K. (2018). Mast cells co-expressing CD68 and inorganic polyphosphate are linked with colorectal cancer. PLoS One *13*, e0193089.

Artiles, K., Anastasia, S., McCusker, D., and Kellogg, D.R. (2009). The Rts1 regulatory subunit of protein phosphatase 2A is required for control of G1 cyclin transcription and nutrient modulation of cell size. PLoS Genet *5*, e1000727.

Auesukaree, C., Homma, T., Tochio, H., Shirakawa, M., Kaneko, Y., and Harashima, S. (2004). Intracellular phosphate serves as a signal for the regulation of the PHO pathway in Saccharomyces cerevisiae. J Biol Chem *279*, 17289-17294.

Azevedo, C., Livermore, T., and Saiardi, A. (2015). Protein polyphosphorylation of lysine residues by inorganic polyphosphate. Mol Cell *58*, 71-82.

Azevedo, C., and Saiardi, A. (2015). Why always lysine? The ongoing tale of one of the most modified amino acids. Adv Biol Regul.

Azevedo, C., and Saiardi, A. (2016). The new world of inorganic polyphosphates. Biochem Soc Trans *44*, 13-17.

Azevedo, C., Singh, J., Steck, N., Hofer, A., Ruiz, F.A., Singh, T., Jessen, H.J., and Saiardi, A. (2018). Screening a Protein Array with Synthetic Biotinylated Inorganic Polyphosphate To Define the Human PolyP-ome. ACS Chem Biol *13*, 1958-1963.

Badolato, R., and Parolini, S. (2007). Novel insights from adaptor protein 3 complex deficiency. J Allergy Clin Immunol *120*, 735-741.

Bae, J.S., Lee, W., and Rezaie, A.R. (2012). Polyphosphate elicits pro-inflammatory responses that are counteracted by activated protein C in both cellular and animal models. J Thromb Haemost *10*, 1145-1151.

Baev, A.Y., Angelova, P.R., and Abramov, A.Y. (2020). Inorganic polyphosphate is produced and hydrolyzed in F0F1-ATP synthase of mammalian mitochondria. Biochem J *477*, 1515-1524.

Balaji, S., LIyer, L.M., and Aravind, L. (2009). HPC2 and ubinuclein define a novel family of histone chaperones conserved throughout eukaryotes. Mol Biosyst *5*, 269-275.

Balakrishnan, R., Park, J., Karra, K., Hitz, B.C., Binkley, G., Hong, E.L., Sullivan, J., Micklem, G., and Cherry, J.M. (2012). YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database (Oxford) *2012*, bar062.

Bashatwah, R.M., Khanfar, M.A., and Bardaweel, S.K. (2018). Discovery of potent polyphosphate kinase 1 (PPK1) inhibitors using structure-based exploration of PPK1 pharmacophoric space coupled with docking analyses. J Mol Recognit.

Beltrao, P., Albanese, V., Kenner, L.R., Swaney, D.L., Burlingame, A., Villen, J., Lim, W.A., Fraser, J.S., Frydman, J., and Krogan, N.J. (2012). Systematic functional prioritization of protein posttranslational modifications. Cell *150*, 413-425.

Bentley-DeSousa, A., and Downey, M. (2019). From underlying chemistry to therapeutic potential: open questions in the new field of lysine polyphosphorylation. Curr Genet *65*, 57-64.

Bentley-DeSousa, A., Holinier, C., Moteshareie, H., Tseng, Y.C., Kajjo, S., Nwosu, C., Amodeo, G.F., Bondy-Chorney, E., Sai, Y., Rudner, A.*, et al.* (2018). A Screen for Candidate Targets of Lysine Polyphosphorylation Uncovers a Conserved Network Implicated in Ribosome Biogenesis. Cell Rep *22*, 3427-3439.

Benton, B.M., Zhang, J.H., and Thorner, J. (1994). A novel FK506- and rapamycin-binding protein (FPR3 gene product) in the yeast Saccharomyces cerevisiae is a proline rotamase localized to the nucleolus. J Cell Biol *127*, 623-639.

Besteiro, S., Tonn, D., Tetley, L., Coombs, G.H., and Mottram, J.C. (2008). The AP3 adaptor is involved in the transport of membrane proteins to acidocalcisomes of Leishmania. J Cell Sci *121*, 561-570.

Bondy-Chorney, E., Abramchuk, I., Nasser, R., Holinier, C., Denoncourt, A., Baijal, K., McCarthy, L., Khacho, M., Lavallee-Adam, M., and Downey, M. (2020). A Broad Response to Intracellular Long-Chain Polyphosphate in Human Cells. Cell Rep *33*, 108318.

Boyineni, J., Sredni, S.T., Margaryan, N.V., Demirkhanyan, L., Tye, M., Robert, J., Gonzalez-Nilo, F., Hendrix, M., Pavlov, E., Soares, M.B.*, et al.* (2020). Inorganic polyphosphate as an energy source in tumourigenesis. Oncotarget *11*, 4613-4624.

Breus, N.A., Ryazanova, L.P., Dmitriev, V.V., Kulakovskaya, T.V., and Kulaev, I.S. (2012). Accumulation of phosphate and polyphosphate by Cryptococcus humicola and Saccharomyces cerevisiae in the absence of nitrogen. FEMS Yeast Res *12*, 617-624.

Bru, S., Jimenez, J., Canadell, D., Arino, J., and Clotet, J. (2016a). Improvement of biochemical methods of polyP quantification. Microb Cell *4*, 6-15.

Bru, S., Martinez-Lainez, J.M., Hernandez-Ortega, S., Quandt, E., Torres-Torronteras, J., Marti, R., Canadell, D., Arino, J., Sharma, S., Jimenez, J.*, et al.* (2016b). Polyphosphate is involved in cell cycle progression and genomic stability in Saccharomyces cerevisiae. Mol Microbiol *101*, 367-380.

Bru, S., Samper-Martin, B., Quandt, E., Hernandez-Ortega, S., Martinez-Lainez, J.M., Gari, E., Rafel, M., Torres-Torronteras, J., Marti, R., Ribeiro, M.P.C.*, et al.* (2017). Polyphosphate is a key factor for cell survival after DNA damage in eukaryotic cells. DNA Repair (Amst) *57*, 171-178.

Cai, L., Sutter, B.M., Li, B., and Tu, B.P. (2011). Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. Mol Cell *42*, 426-437.

Carotenuto, M., De Antonellis, P., Liguori, L., Benvenuto, G., Magliulo, D., Alonzi, A., Turino, C., Attanasio, C., Damiani, V., Bello, A.M.*, et al.* (2014). H-Prune through GSK-3beta interaction sustains canonical WNT/beta-catenin signaling enhancing cancer progression in NSCLC. Oncotarget *5*, 5736-5749.

Casler, J.C., and Glick, B.S. (2020). A microscopy-based kinetic analysis of yeast vacuolar protein sorting. Elife *9*, e56844.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chatr-Aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N., O'Donnell, L., Oster, S., Theesfeld, C., Sellam, A.*, et al.* (2017). The BioGRID interaction database: 2017 update. Nucleic Acids Res *45*, D369-D379.

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R.*, et al.* (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res *40*, D700-705.

Choi, S.H., Smith, S.A., and Morrissey, J.H. (2011). Polyphosphate is a cofactor for the activation of factor XI by thrombin. Blood *118*, 6963-6970.

Choi, S.H., Smith, S.A., and Morrissey, J.H. (2015). Polyphosphate accelerates factor V activation by factor XIa. Thromb Haemost *113*, 599-604.

Christ, J.J., Willbold, S., and Blank, L.M. (2020). Methods for the Analysis of Polyphosphate in the Life Sciences. Anal Chem *92*, 4167-4176.

Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D.*, et al.* (2016). A global genetic interaction network maps a wiring diagram of cellular function. Science *353*, aaf1420.

Cowles, C.R., Odorizzi, G., Payne, G.S., and Emr, S.D. (1997). The AP-3 adaptor complex is essential for cargo-selective transport to the yeast vacuole. Cell *91*, 109-118.

Cremers, C.M., Knoefler, D., Gates, S., Martin, N., Dahl, J.U., Lempart, J., Xie, L., Chapman, M.R., Galvan, V., Southworth, D.R.*, et al.* (2016). Polyphosphate: A Conserved Modifier of Amyloidogenic Processes. Mol Cell *63*, 768-780.

Dahl, J.U., Gray, M.J., Bazopoulou, D., Beaufay, F., Lempart, J., Koenigsknecht, M.J., Wang, Y., Baker, J.R., Hasler, W.L., Young, V.B.*, et al.* (2017). The anti-inflammatory drug mesalamine targets bacterial polyphosphate accumulation. Nat Microbiol *2*, 16267.

De Marchis, M.L., Giogi, A., Schinina, M.E., Bozzone, I., and Fatica, A. (2005). Rrp15p, a novel component of pre-ribosomal particles required for 60S ribosome subunit maturation. RNA *11*, 495-502.

Denoncourt, A., and Downey, M. (2021). Model systems for studying polyphosphate biology: a focus on microorganisms. Curr Genet.

Desfougeres, Y., Gerasimaite, R.U., Jessen, H.J., and Mayer, A. (2016). Vtc5, a Novel Subunit of the Vacuolar Transporter Chaperone Complex, Regulates Polyphosphate Synthesis and Phosphate Homeostasis in Yeast. J Biol Chem *291*, 22262-22275.

Dever, T.E. (2002). Gene-specific regulation by general translation factors. Cell *108*, 545-556.

Dice, J.F. (2007). Chaperone-Mediated Autophagy. Autophagy *3*, 295-299.

Distel, B., Al, E.J., Tabak, H.F., and Jones, E.W. (1983). Synthesis and maturation of the yeast vacuolar enzymes carboxypeptidase Y and aminopeptidase I. Biochim Biophys Acta *741*, 128-135.

Donovan, A.J., Kalkowski, J., Szymusiak, M., Wang, C., Smith, S.A., Klie, R.F., Morrissey, J.H., and Liu, Y. (2016). Artificial Dense Granules: A Procoagulant Liposomal Formulation Modeled after Platelet Polyphosphate Storage Pools. Biomacromolecules *17*, 2572-2581.

Edwards, T.K., Saleem, A., Shaman, J.A., Dennis, T., Gerigk, C., Oliveros, E., Gartenberg, M.R., and Rubin, E.H. (2000). Role for nucleolin/Nsr1 in the cellular localization of topoisomerase I. J Biol Chem *275*, 36181-36188.

El-Chemaly, S., and Young, L.R. (2016). Hermansky-Pudlak Syndrome. Clin Chest Med *37*, 505-511.

Eskes, E., Deprez, M.A., Wilms, T., and Winderickx, J. (2018). pH homeostasis in yeast; the phosphate perspective. Curr Genet *64*, 155-161.

Esposito, A.M., Mateyak, M., He, D., Lewis, M., Sasikumar, A.N., Hutton, J., Copeland, P.R., and Kinzy, T.G. (2010). Eukaryotic polyribosome profile analysis. J Vis Exp.

Faye, M.D., Graber, T.E., and Holcik, M. (2014). Assessment of selective mRNA translation in mammalian cells by polysome profiling. J Vis Exp.

Feyder, S., De Craene, J.O., Bar, S., Bertazzi, D.L., and Friant, S. (2015). Membrane trafficking in the yeast Saccharomyces cerevisiae model. Int J Mol Sci *16*, 1509-1525.

Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L., and Link, A.J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. Genes Dev *20*, 1294-1307.

Fraley, C.D., Rashid, M.H., Lee, S.S., Gottschalk, R., Harrison, J., Wood, P.J., Brown, M.R., and Kornberg, A. (2007). A polyphosphate kinase 1 (ppk1) mutant of Pseudomonas aeruginosa exhibits multiple ultrastructural and functional defects. Proc Natl Acad Sci U S A *104*, 3526-3531.

Freimoser, F.M., Hurlimann, H.C., Jakob, C.A., Werner, T.P., and Amrhein, N. (2006). Systematic screening of polyphosphate (poly P) levels in yeast mutant cells reveals strong interdependence with primary metabolism. Genome Biol *7*, R109.

Gautier, T., Berges, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol Cell Biol *17*, 7088-7089.

Gerasimaite, R., and Mayer, A. (2016). Enzymes of yeast polyphosphate metabolism: structure, enzymology and biological roles. Biochem Soc Trans *44*, 234-239.

Gerasimaite, R., and Mayer, A. (2017). Ppn2, a novel Zn(2+)-dependent polyphosphatase in the acidocalcisome-like yeast vacuole. J Cell Sci *130*, 1625-1636.

Gerasimaite, R., Pavlovic, I., Capolicchio, S., Hofer, A., Schmidt, A., Jessen, H.J., and Mayer, A. (2017). Inositol Pyrophosphate Specificity of the SPX-Dependent Polyphosphate Polymerase VTC. ACS Chem Biol *12*, 648-653.

Gerasimaite, R., Sharma, S., Desfougeres, Y., Schmidt, A., and Mayer, A. (2014). Coupled synthesis and translocation restrains polyphosphate to acidocalcisome-like vacuoles and prevents its toxicity. J Cell Sci *127*, 5093-5104.

Giaever, G., and Nislow, C. (2014). The Yeast Deletion Collection: A Decade of Functional Genomics. Genetics *197*, 451-465.

Ginisty, H., Sicard, H., Roger, B., and Bouvet, P. (1999). Structure and functions of nucleolin. J Cell Sci *112*, 761-772.

Glick, D., Barth, S., and Macleod, K.F. (2010). Autophagy: cellular and molecular mechanisms. J Pathol *221*, 3-12.

Goldstein, A.L., and McCusker, J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast *15*, 1541-1553.

Graham, L.A., Hill, K.J., and Stevens, T.H. (1998). Assembly of the yeast vacuolar H+-ATPase occuts in the endoplasmic reticulum and requires a Vma12p/Vma22p assembly complex. J Cell Biol *142*, 39-49.

Gray, M.J., Wholey, W.Y., Wagner, N.O., Cremers, C.M., Mueller-Schickert, A., Hock, N.T., Krieger, A.G., Smith, E.M., Bender, R.A., Bardwell, J.C.*, et al.* (2014). Polyphosphate is a primordial chaperone. Mol Cell *53*, 689-699.

Hacchou, Y., Uematsu, T., Ueda, O., Usui, Y., Uematsu, S., Takahashi, M., Uchihashi, T., Kawazoe, Y., Shiba, T., Kurihara, S.*, et al.* (2007). Inorganic polyphosphate: a possible stimulant of bone formation. J Dent Res *86*, 893-897.

Hardman, G., Perkins, S., Brownridge, P.J., Clarke, C.J., Byrne, D.P., Campbell, A.E., Kalyuzhnyy, A., Myall, A., Eyers, P.A., Jones, A.R.*, et al.* (2019). Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation. EMBO J *38*, e100847.

Hardman, G., Perkins, S., Ruan, Z., Kannan, N., Brownridge, P.J., Byrne, D.P., Eyes, P.A., Jones, A.R., and Eyers, C.E. (2017). Extensive non-canonical phosphorylation in human cells revealed using strong-anion exchange-mediated phosphoproteomics. BioRxiv.

Hecht, K.A., O'Donnell, A.F., and Brodsky, J.L. (2014). The proteolytic landscape of the yeast vacuole. Cell Logist *4*, e28023.

Henne, W.M., Buchkovich, N.J., and Emr, S.D. (2011). The ESCRT pathway. Dev Cell *21*, 77-91.

Hernandez-Ruiz, L., Gonzalez-Garcia, I., Castro, C., Brieva, J.A., and Ruiz, F.A. (2006). Inorganic polyphosphate and specific induction of apoptosis in human plasma cells. Haematologica *91*, 1180-1186.

Hoac, B., Kiffer-Moreira, T., Millan, J.L., and McKee, M.D. (2013). Polyphosphates inhibit extracellular matrix mineralization in MC3T3-E1 osteoblast cultures. Bone *53*, 478-486.

Holmstrom, K.M., Marina, N., Baev, A.Y., Wood, N.W., Gourine, A.V., and Abramov, A.Y. (2013). Signalling properties of inorganic polyphosphate in the mammalian brain. Nat Commun *4*, 1362.

Holt, L.J., Tuch, B.B., Villen, J., Johnson, A.D., Gygi, S.P., and Morgan, D.O. (2009). Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. Science *325*, 1682-1686.

Hong, B., Brockenbrough, J.S., Wu, P.Y., and Aris, J.P. (1997). Nop2p is required for pre-rRNA processing and 60S ribosome subunit synthesis in yeast. Mol Cell Biol *17*, 378-388.

Hooley, P., Whitehead, M.P., and Brown, M.R. (2008). Eukaryote polyphosphate kinases: is the 'Kornberg' complex ubiquitous? Trends Biochem Sci *33*, 577-582.

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

Huang, D., Moffat, J., and Andrews, B. (2002). Dissection of a complex phenotype by functional genomics reveals roles for the yeast cyclin-dependent protein kinase Pho85 in stress adaptation and cell integrity. Mol Cell Biol *22*, 5076-5088.

Huang, G., Fang, J., Sant'Anna, C., Li, Z.H., Wellems, D.L., Rohloff, P., and Docampo, R. (2011). Adaptor protein-3 (AP-3) complex mediates the biogenesis of acidocalcisomes and is essential for growth and virulence of Trypanosoma brucei. J Biol Chem *286*, 36619-36630.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686-691.

Huizing, M., Anikster, Y., and Gahl, W.A. (2000). Hermansky-Pudlak syndrome and related disorders of organelle formation. Traffic *1*, 823-835.

Huizing, M., Boissy, R.E., and Gahl, W.A. (2002). Hermansky-Pudlak syndrome: vesicle formation from yeast to man. Pigment Cell Res *15*, 405-419.

Huizing, M., Malicdan, M.C.V., Wang, J.A., Pri-Chen, H., Hess, R.A., Fischer, R., O'Brien, K.J., Merideth, M.A., Gahl, W.A., and Gochuico, B.R. (2020). Hermansky-Pudlak syndrome: Mutation update. Hum Mutat *41*, 543-580.

Jimenez, J., Bru, S., Ribeiro, M.P., and Clotet, J. (2017). Polyphosphate: popping up from oblivion. Curr Genet *63*, 15-18.

Jimenez-Nunez, M.D., Moreno-Sanchez, D., Hernandez-Ruiz, L., Benitez-Rondan, A., Ramos-Amaya, A., Rodriguez-Bayona, B., Medina, F., Brieva, J.A., and Ruiz, F.A. (2012). Myeloma cells contain high levels of inorganic polyphosphate which is associated with nucleolar transcription. Haematologica *97*, 1264-1271.

Kamada, Y., Sekito, T., and Ohsumi, Y. (2004). Autophagy in yeast: a TOR-mediated response to nutrient starvation. Curr Top Microbiol Immunol *279*, 73-84.

Kapitzky, L., Beltrao, P., Berens, T.J., Gassner, N., Zhou, C., Wuster, A., Wu, J., Babu, M.M., Elledge, S.J., Toczyski, D.*, et al.* (2010). Cross-species chemogenomic profiling reveals evolutionarily conserved drug mode of action. Mol Syst Biol *6*, 451.

Kim, K.S., Rao, N.N., Fraley, C.D., and Kornberg, A. (2002). Inorganic polyphosphate is essential for long-term survival and virulence factors in Shigella and Salmonella spp. Proc Natl Acad Sci U S A *99*, 7675-7680.

Kirchner, P., Bourdenx, M., Madrigal-Matute, J., Tiano, S., Diaz, A., Bartholdy, B.A., Will, B., and Cuervo, A.M. (2019). Proteome-wide analysis of chaperone-mediated autophagy targeting motifs. PLoS Biol *17*, e3000301.

Komander, D., Clague, M.J., and Urbe, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat Rev Mol Cell Biol *10*, 550-563.

Kornberg, A. (1999). Inorganic polyphosphate: a molecule of many functions. Prog Mol Subcell Biol *23*, 1-18.

Kornberg, A., Rao, N.N., and Ault-Riche, D. (1999). Inorganic polyphosphate: a molecule of many functions. Annu Rev Biochem *68*, 89-125.

Krystkowiak, I., and Davey, N.E. (2017). SLiMSearch: A framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. Nucleic Acids Res *45*, W464-W469.

Kulaev, I., Vagabov, V.M., and Kulakovskaia, T.V. (2004). The Biochemistry of Inorganic Polyphosphates, 2 edn (Chichester, West Sussex, PO19 8SQ, England: John Wiley & Sons, Ltd).

Kumble, K.D., and Kornberg, A. (1995). Inorganic polyphosphate in mammalian cells and tissues. J Biol Chem *270*, 5818-5822.

Kumble, K.D., and Kornberg, A. (1996). Endopolyphosphatases for long chain inorganic polyphosphate in yeast and mammals. J Biol Chem *271*, 27146-27151.

Labberton, K., Kenne, E., Long, A.T., Nickel, K.F., Di Gennaro, A., Rigg, R.A., Hernandez, J.S., Butler, L., Maas, C., Stavrou, E.X.*, et al.* (2016). Neutralizing blood-borne polyphosphate in vivo provides safe thromboprotection. Nat Commun *7*, 12616.

Lafontaine, D.L., and Tollervey, D. (2000). Synthesis and assembly of the box C+D small nucleolar RNPs. Mol Cell Biol *20*, 2650-2659.

Li, M., Rong, Y., Chuang, Y.S., Peng, D., and Emr, S.D. (2015). Ubiquitin-dependent lysosomal membrane protein sorting and degradation. Mol Cell *57*, 467-478.

Li, Z., Lee, I., Moradi, E., Hung, N.J., Johnson, A.W., and Marcotte, E.M. (2009). Rational extension of the ribosome biogenesis pathway using network-guided genetics. PLoS Biol *7*, e1000213.

Lichko, L., Kulakovskaya, T., Pestov, N., and Kulaev, I. (2006). Inorganic polyphosphates and exopolyphosphatases in cell compartments of the yeast Saccharomyces cerevisiae under inactivation of PPX1 and PPN1 genes. Biosci Rep *26*, 45-54.

Lichko, L.P., Pestov, N.A., Kulakovskaya, T.V., and Kulaev, I.S. (2003). Effect of PPX1 inactivation on the exopolyphosphatase spectra in cytosol and mitochondria of the yeast Saccharomyces cerevisiae. Biochemistry (Mosc) *68*, 740-746.

Lin, C., Liu, W., Zhang, W., Londono-Joshi, A.I., Buchsbaum, D.J., Bu, G., and Li, Y. (2013). The C-terminal region Mesd peptide mimics full-length Mesd and acts as an inhibitor of Wnt/B-catenin signaling in cancer cells. PLoS One *8*, e58102.

Llinares, E., Barry, A.O., and Andre, B. (2015). The AP-3 adaptor complex mediates sorting of yeast and mammalian PQ-loop-family basic amino acid transporters to the vacuolar/lysosomal membrane. Sci Rep *5*, 16665.

Lonetti, A., Szijgyarto, Z., Bosch, D., Loss, O., Azevedo, C., and Saiardi, A. (2011). Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases. J Biol Chem *286*, 31966-31974.

Longtine, M.S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.

Lorenz, B., and Schroder, H.C. (2001). Mammalian intestinal alkaline phosphatase acts as highly active exopolyphosphatase. Biochim Biophys Acta *1547*, 254-261.

Luhtala, N., and Odorizzi, G. (2004). Bro1 coordinates deubiquitination in the multivesicular body pathway by recruiting Doa4 to endosomes. J Cell Biol *166*, 717-729.

Luk, E., Vu, N.D., Patteson, K., Mizuguchi, G., Wu, W.H., Ranjan, A., Backus, J., Sen, S., Lewis, M., Bai, Y., *et al.* (2007). Chz1, a nuclear chaperone for histone H2AZ. Mol Cell *25*, 357-368.

Mark, K.G., Meza-Gutierrez, F., Johnson, J.R., Newton, B.W., Krogan, N., and Toczyski, D.P. (2015). Prb1 protease activity is required for its recognition by the F-box protein Saf1. Biochemistry *54*, 4423-4426.

Mark, K.G., Simonetta, M., Maiolica, A., Seller, C.A., and Toczyski, D.P. (2014). Ubiquitin ligase trapping identifies an SCF(Saf1) pathway targeting unprocessed vacuolar/lysosomal proteins. Mol Cell *53*, 148-161.

McCarthy, L., Bentley-DeSousa, A., Denoncourt, A., Tseng, Y.C., Gabriel, M., and Downey, M. (2020). Proteins required for vacuolar function are targets of lysine polyphosphorylation in yeast. FEBS Lett *594*, 21-30.

McGrath, J.W., and Quinn, J.P. (2000). Intracellular Accumulation of Polyphosphate by the Yeast Candida humicola G-1 in Response to Acid pH. Appl Environ Microbiol *66*, 4068-4073.

McNatt, M.W., McKittrick, I., West, M., and Odorizzi, G. (2007). Direct binding to Rsp5 mediates ubiquitin-independent sorting of Sna3 via the multivesicular body pathway. Mol Biol Cell *18*, 697-706.

Mechler, B., Muller, M., Muller, H., Meussdoerffer, F., and Wolf, D.H. (1982). In vivo biosynthesis of the vacuolar proteinases A and B in the yeast Saccharomyces cerevisiae. J Biol Chem *257*, 11203-11206.

Mikami, Y., Tsuda, H., Akiyama, Y., Honda, M., Shimizu, N., Suzuki, N., and Komiyama, K. (2016). Alkaline phosphatase determines polyphosphate-induced mineralization in a cell-type independent manner. J Bone Miner Metab *34*, 627-637.

Mitchell, J.L., Lionikiene, A.S., Georgiev, G., Klemmer, A., Brain, C., Kim, P.Y., and Mutch, N.J. (2016). Polyphosphate colocalizes with factor XII on platelet-bound fibrin and augments its plasminogen activator activity. Blood *128*, 2834-2845.

Mitchell, L., Huard, S., Cotrut, M., Pourhanifeh-Lemeri, R., Steunou, A.L., Hamza, A., Lambert, J.P., Zhou, H., Ning, Z., Base, A., *et al.* (2013). mChIP-KAT-MS, a method to map protein interactions and acetylation sites for lysine acetyltransferases. Proc Natl Acad Sci U S A *110*.

Moehle, C.M., Aynardi, M.W., Kolodny, M.R., Park, F.J., and Jones, E.W. (1987). Protease B of Saccharomyces cerevisiae: isolation and regulation of the PRB1 structural gene. Genetics *115*, 255-263.

Mohammed, M., Al-Hashmi, N., Al-Rashdi, S., Al-Sukaiti, N., Al-Adawi, K., Al-Riyami, M., and Al-Maawali, A. (2018). Biallelic mutations in AP3D1 cause Hermansky-Pudlak syndrome type 10 associated with immunodeficiency and seizure disorder. Eur J Med Genet.

Moreno, S.N., and Docampo, R. (2013). Polyphosphate and its diverse functions in host cells and pathogens. PLoS Pathog *9*, e1003230.

Morrissey, J.H. (2012). Polyphosphate: a link between platelets, coagulation and inflammation. Int J Hematol *95*, 346-352.

Morrissey, J.H., Choi, S.H., and Smith, S.A. (2012). Polyphosphate: an ancient molecule that links platelets, coagulation, and inflammation. Blood *119*, 5972-5979.

Morrissey, J.H., and Smith, S.A. (2015). Polyphosphate as modulator of hemostasis, thrombosis, and inflammation. J Thromb Haemost *13 Suppl 1*, S92-97.

Muhl, L., Galuska, S.P., Oorni, K., Hernandez-Ruiz, L., Andrei-Selmer, L.C., Geyer, R., Preissner, K.T., Ruiz, F.A., Kovanen, P.T., and Kanse, S.M. (2009). High negative charge-to-size ratio in polyphosphates and heparin regulates factor VII-activating protease. FEBS J *276*, 4828-4839.

Muller, F., Mutch, N.J., Schenk, W.A., Smith, S.A., Esterl, L., Spronk, H.M., Schmidbauer, S., Gahl, W.A., Morrissey, J.H., and Renne, T. (2009). Platelet polyphosphates are proinflammatory and procoagulant mediators in vivo. Cell *139*, 1143-1156.

Muller, O., Bayer, M.J., Peters, C., Andersen, J.S., Mann, M., and Mayer, A. (2002). The Vtc proteins in vacuole fusion: coupling NSF activity to V(0) trans-complex formation. EMBO J *21*, 259-269.

Muller, O., Neumann, H., Bayer, M.J., and Mayer, A. (2003). Role of the Vtc proteins in V-ATPase stability and membrane trafficking. J Cell Sci *116*, 1107-1115.

Mutch, N.J., Engel, R., Uitte de Willige, S., Philippou, H., and Ariens, R.A. (2010a). Polyphosphate modifies the fibrin network and down-regulates fibrinolysis by attenuating binding of tPA and plasminogen to fibrin. Blood *115*, 3980-3988.

Mutch, N.J., Myles, T., Leung, L.L., and Morrissey, J.H. (2010b). Polyphosphate binds with high affinity to exosite II of thrombin. J Thromb Haemost *8*, 548-555.

Negreiros, R.S., Lander, N., Huang, G., Cordeiro, C.D., Smith, S.A., Morrissey, J.H., and Docampo, R. (2018). Inorganic polyphosphate interacts with nucleolar and glycosomal proteins in trypanosomatids. Mol Microbiol *110*, 973-994.

Nguyen, T.Q., Dziuba, N., and Lindahl, P.A. (2019). Isolated Saccharomyces cerevisiae vacuoles contain low-molecular-mass transition-metal polyphosphate complexes. Metallomics *11*, 1298-1309.

Odorizzi, G., Cowles, C.R., and Emr, S.D. (1998). The AP-3 complex: a coat of many colours. Trends Cell Biol *8*, 282-288.

Ogawa, N., DeRisi, J., and Brown, P.O. (2000). New components of a system for phosphate accumulation and polyphosphate metabolism in Saccharomyces cerevisiae revealed by genomic expression analysis. Mol Biol Cell *11*, 4309-4321.

Okano, A., Wan, K., Kanda, K., Yabuki, Y., Funato, K., and Mizuta, K. (2015). SMY2 and SYH1 suppress defects in ribosome biogenesis caused by ebp2 mutations. Biosci Biotechnol Biochem *79*, 1481-1483.

Oruganti, S., Zhang, Y., Li, H., Robinson, H., Terns, M.P., Terns, R.M., Yang, W., and Li, H. (2007). Alternative conformations of the archaeal Nop56/58-fibrillarin complex imply flexibility in box C/D RNPs. J Mol Biol *371*, 1141-1150.

Oughtred, R., Stark, C., Breitkreutz, B.J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R.*, et al.* (2019). The BioGRID interaction database: 2019 update. Nucleic Acids Res *47*, D529-D541.

Park, S.K., Xiao, H., and Lei, M. (2014). Nuclear FKBPs, Fpr3 and Fpr4 affect genome-wide transcription. Mol Genet Genomics *289*, 125-136.

Park, S.Y., and Guo, X. (2014). Adaptor protein complexes and intracellular transport. Biosci Rep *34*.

Peng, L., Jiang, Q., Pan, J.Y., Deng, C., Yu, J.Y., Wu, X.M., Huang, S.H., and Deng, X.Y. (2016). Involvement of polyphosphate kinase in virulence and stress tolerance of uropathogenic Proteus mirabilis. Med Microbiol Immunol *205*, 97-109.

Pestov, N.A., Kulakovskaya, T.V., and Kulaev, I.S. (2004). Inorganic polyphosphate in mitochondria of Saccharomyces cerevisiae at phosphate limitation and phosphate excess. FEMS Yeast Res *4*, 643-648.

Pokrzywa, W., Guerriat, B., Dodzian, J., and Morsomme, P. (2009). Dual sorting of the Saccharomyces cerevisiae vacuolar protein Sna4p. Eukaryot Cell *8*, 278-286.

Puy, C., Tucker, E.I., Ivanov, I., Gailani, D., Smith, S., Morrissey, J.H., Gruber, A., and McCarty, O.J. (2016). Platelet-derived short-chain polyphosphates enhance the inactivation of tissue factor pathway inhibitor by activated coagulation factor XI. PLoS One *11*, e0165172.

Puy, C., Tucker, E.I., Wong, Z.C., Gailani, D., Smith, S.A., Choi, S.H., Morrissey, J.H., Gruber, A., and McCarty, O.J. (2013). Factor XII promotes blood coagulation independent of factor XI in the presence of long-chain polyphosphates. J Thromb Haemost *11*, 1341-1352.

Rao, N.N., Gomez-Garcia, M.R., and Kornberg, A. (2009). Inorganic polyphosphate: essential for growth and survival. Annu Rev Biochem *78*, 605-647.

Rashid, M.H., and Kornberg, A. (2000). Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *97*, 4885-4890.

Rashid, M.H., Rao, N.N., and Kornberg, A. (2000a). Inorganic polyphosphate is required for motility of bacterial pathogens. J Bacteriol *182*, 225-227.

Rashid, M.H., Rumbaugh, K., Passador, L., Davies, D.G., Hamood, A.N., Iglewski, B.H., and Kornberg, A. (2000b). Polyphosphate kinase is essential for biofilm development, quorum sensing, and virulence of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *97*, 9636-9641.

Ravid, T., and Hochstrasser, M. (2008). Diversity of degradation signals in the ubiquitin-proteasome system. Nat Rev Mol Cell Biol *9*, 679-689.

Reggiori, F., and Klionsky, D.J. (2013). Autophagic processes in yeast: mechanism, machinery and regulation. Genetics *194*, 341-361.

Rehling, P., Darsow, T., Katzmann, D.J., and Emr, S.D. (1999). Formation of AP-3 transport intermediates requires Vps41 function. Nat Cell Biol *1*, 346-353.

Roewe, J., Stavrides, G., Strueve, M., Sharma, A., Marini, F., Mann, A., Smith, S.A., Kaya, Z., Strobl, B., Mueller, M.*, et al.* (2020). Bacterial polyphosphates interfere with the innate host defense to infection. Nat Commun *11*, 4035.

Rosenfeld, L., Reddi, A.R., Leung, E., Aranda, K., Jensen, L.T., and Culotta, V.C. (2010). The effect of phosphate accumulation on metal ion homeostasis in Saccharomyces cerevisiae. J Biol Inorg Chem *15*, 1051-1062.

Ross, C.A., and Poirier, M.A. (2004). Protein aggregation and neurodegenerative disease. Nat Med *10(Suppl)*, S10-S17.

Rossetto, D., Cramet, M., Wang, A.Y., Steunou, A.L., Lacoste, N., Schulze, J.M., Cote, V., Monnet-Saksouk, J., Piquet, S., Nourani, A.*, et al.* (2014). Eaf 5/7/3 form a functionally independent NuA4 submodule linked to RNA polymerase II-coupled nucleosome recycling. EMBO J *33*, 1397-1415.

Rossl, A., Bentley-DeSousa, A., Tseng, Y.C., Nwosu, C., and Downey, M. (2016). Nicotinamide Suppresses the DNA Damage Sensitivity of Saccharomyces cerevisiae Independently of Sirtuin Deacetylases. Genetics *204*, 569-579.

Ruiz, F.A., Lea, C.R., Oldfield, E., and Docampo, R. (2004). Human platelet dense granules contain polyphosphate and are similar to acidocalcisomes of bacteria and unicellular eukaryotes. J Biol Chem *279*, 44250-44257.

Ruotolo, R., Marchini, G., and Ottonello, S. (2008). Membrane transporters and protein traffic networks differentially affecting metal tolerance: a genomic phenotyping study in yeast. Genome Biol *9*, R67.

Ryazanova, L., Zvonarev, A., Rusakova, T., Dmitriev, V., and Kulakovskaya, T. (2016). Manganese tolerance in yeasts involves polyphosphate, magnesium, and vacuolar alterations. Folia Microbiol (Praha) *61*, 311-317.

Saiardi, A. (2012a). Cell signalling by inositol pyrophosphates. Subcell Biochem *59*, 413-443.

Saiardi, A. (2012b). How inositol pyrophosphates control cellular phosphate homeostasis? Adv Biol Regul *52*, 351-359.

Saito, K., Ohtomo, R., Kuga-Uetake, Y., Aono, T., and Saito, M. (2005). Direct labeling of polyphosphate at the ultrastructural level in Saccharomyces cerevisiae by using the affinity of the polyphosphate binding domain of Escherichia coli exopolyphosphatase. Appl Environ Microbiol *71*, 5692-5701.

Sarangi, P., and Zhao, X. (2015). SUMO-mediated regulation of DNA damage repair and responses. Trends Biochem Sci *40*, 233-242.

Sarkar, S., Ravikumar, B., Floto, R.A., and Rubinsztein, D.C. (2008). Rapamycin and mTOR-independent autophagy inducers ameliorate toxicity of polyglutamine-expanded huntingtin and related proteinopathies. Cell Death Differ *16*, 46-56.

Schmidt, O., and Teis, D. (2012a). The ESCRT machinery. Curr Biol *22*, R116-120.

Schmidt, O., and Teis, D. (2012b). The ESCRT machinery. Curr Biol *22*, R116-120.

Schuck, S. (2020). Microautophagy-distinct molecular mechanisms handle cargoes of many sizes. J Cell Sci *133*, jcs246322.

Sears, R.M., May, D.G., and Roux, K.J. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. Methods Mol Biol *2012*, 299-313.

Sharma, S., Yang, J., Watzinger, P., Kotter, P., and Entian, K.D. (2013). Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res *41*, 9062-9076.
Sherman, B.T., Huang, W., Tan, Q., Guo, Y., Bour, S., Liu, D., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). DAVID Knowledgebase: A gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics *8*, 426.

Shi, X., Rao, N.N., and Kornberg, A. (2004). Inorganic polyphosphate in Bacillus cereus: motility, biofilm formation, and sporulation. Proc Natl Acad Sci U S A *101*, 17061-17065.

Shiba, T., Nishimura, D., Kawazoe, Y., Onodera, Y., Tsutsumi, K., Nakamura, R., and Ohshiro, M. (2003). Modulation of mitogenic activity of fibroblast growth factors by inorganic polyphosphate. J Biol Chem *278*, 26788-26792.

Shiba, T., Tsutsumi, K., Yano, H., Ihara, Y., Kameda, A., Tanaka, K., Takahashi, H., Munekata, M., Rao, N.N., and Kornberg, A. (1997). Inorganic polyphosphate and the induction of rpoS expression. Proc Natl Acad Sci U S A *94*, 11210-11215.

Smith, S.A., Choi, S.H., Collins, J.N., Travers, R.J., Cooley, B.C., and Morrissey, J.H. (2012). Inhibition of polyphosphate as a novel strategy for preventing thrombosis and inflammation. Blood *120*, 5103-5110.

Smith, S.A., Choi, S.H., Davis-Harrison, R., Huyck, J., Boettcher, J., Rienstra, C.M., and Morrissey, J.H. (2010). Polyphosphate exerts differential effects on blood clotting, depending on polymer size. Blood *116*, 4353-4359.

Smith, S.A., and Morrissey, J.H. (2007). Sensitive fluorescence detection of polyphosphate in polyacrylamide gels using 4',6-diamidino-2-phenylindol. Electrophoresis *28*, 3461-3465.
Smith, S.A., and Morrissey, J.H. (2008). Polyphosphate as a general procoagulant agent. J Thromb Haemost *6*, 1750-1756.

Smith, S.A., Mutch, N.J., Baskar, D., Rohloff, P., Docampo, R., and Morrissey, J.H. (2006). Polyphosphate modulates blood coagulation and fibrinolysis. Proc Natl Acad Sci U S A *103*, 903-908.

Srivastava, A., and Jones, E.W. (1998). Pth1/Vam3p is the syntaxin homolog at the vacuolar membrane of Saccharomyces cerevisiae required for the delivery of vacuolar hydrolases. Genetics *148*, 85-98.

Stepp, J.D., Huang, K., and Lemmon, S.K. (1997). The yeast adaptor protein complex, AP-3, is essential for the efficient delivery of alkaline phosphatase by the alternate pathway to the vacuole. J Cell Biol *139*, 1761-1774.

Stotz, S.C., Scott, L.O., Drummond-Main, C., Avchalumov, Y., Girotto, F., Davidsen, J., Gomez-Garcia, M.R., Rho, J.M., Pavlov, E.V., and Colicos, M.A. (2014). Inorganic polyphosphate regulates neuronal excitability through modulation of voltage-gated channels. Mol Brain *7*, 42.

Stringer, D.K., and Piper, R.C. (2011). A single ubiquitin is sufficient for cargo protein entry into MVBs in the absence of ESCRT ubiquitination. J Cell Biol *192*, 229-242.

Sugiyama, M., and Nikawa, J. (2001). The Saccharomyces cerevisiae Isw2p-Itc1p complex represses INO1 expression and maintains cell morphology. J Bacteriol *183*, 4985-4993.
Sun, L., and Chen, Z.J. (2004). The novel functions of ubiquitination in signaling. Curr Opin Cell Biol *16*, 119-126.

Sung, M.K., and Huh, W.K. (2007). Bimolecular fluorescence complementation analysis system for in vivo detection of protein-protein interaction in Saccharomyces cerevisiae. Yeast *24*, 767-775.

Swatek, K.N., and Komander, D. (2016). Ubiquitin modifications. Cell Res *26*, 399-422.

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

Tekirdag, K., and Cuervo, A.M. (2018). Chaperone-mediated autophagy and endosomal microautophagy: Jointed by a chaperone. J Biol Chem *293*, 5414-5424.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Travers, R.J., Smith, S.A., and Morrissey, J.H. (2015). Polyphosphate, platelets, and coagulation. Int J Lab Hematol *37 Suppl 1*, 31-35.

Trilisenko, L., Andreeva, N., Eldarov, M., Dumina, M.V., and Kulakovskaia, T.V. (2015). Polyphosphates and Polyphosphatase Activity in the Yeast Saccharomyces cerevisiae during Overexpression of the DDP1 Gene. Biochemistry (Mosc) *80*, 1312-1317.

Trilisenko, L., Zvonarev, A., Valiakhmetov, A., Penin, A.A., Eliseeva, I.A., Ostroumov, V., Kulakovskiy, I.V., and Kulakovskaya, T. (2019). The Reduced Level of Inorganic Polyphosphate Mobilizes Antioxidant and Manganese-Resistance Systems in Saccharomyces cerevisiae. Cells *8*.

Tsutsumi, K., Matsuya, Y., Sugahara, T., Tamura, M., Sawada, S., Fukura, S., Nakano, H., and Date, H. (2017). Inorganic polyphosphate enhances radio-sensitivity in a human non-small cell lung cancer cell line, H1299. Tumour Biol *39*, 1010428317705033.

Tzeng, C.M., and Kornberg, A. (2000). The multiple activities of polyphosphate kinase of Escherichia coli and their subunit structure determined by radiation target analysis. J Biol Chem *275*, 3977-3983.

Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S.*, et al.* (2010). Towards a knowledge-based Human Protein Atlas. Nat Biotechnol *28*, 1248-1250.

Urech, K., M, D., T, B., A, W., and J, S. (1978). Localization of polyphosphate in vacuoles of Saccharomyces cerevisiae. Arch Microbiol *116*, 275-278.

Usui, Y., Uematsu, T., Uchihashi, T., Takahashi, M., Takahashi, M., Ishizuka, M., Doto, R., Hanaka, H., Komazaki, Y., Osawa, M.*, et al.* (2010). Inorganic polyphosphate induces osteoblastic differentiation. J Dent Res *89*, 504-509.

Uttenweiler, A., Schwarz, H., Neumann, H., and Mayer, A. (2007). The vacuolar transporter chaperone (VTC) complex is required for microautophagy. Mol Biol Cell *18*, 166-175.

Wang, J., Wang, J., Shin, B.S., Kim, J.R., Dever, T.E., Puglisi, J.D., and Fernandez, I.S. (2020). Structural basis for the transition from translation initiation to elongation by an 80S-eIF5B complex. Nat Commun *11*, 5003.

Wang, L., Fraley, C.D., Faridi, J., Kornberg, A., and Roth, R.A. (2003). Inorganic polyphosphate stimulates mammalian TOR, a kinase involved in the proliferation of mammary cancer cells. Proc Natl Acad Sci U S A *100*, 11249-11254.

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T.*, et al.* (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res *38*, W214-W220.

Wat, J.M., Foley, J.H., Krisinger, M.J., Ocariza, L.M., Lei, V., Wasney, G.A., Lameignere, E., Strynadka, N.C., Smith, S.A., Morrissey, J.H.*, et al.* (2014). Polyphosphate suppresses complement via the terminal pathway. Blood *123*, 768-776.

Webster, B.M., Colombi, P., Jager, J., and Lusk, C.P. (2014). Surveillance of nuclear pore complex assembly by ESCRT-III/Vps4. Cell *159*, 388-401.

Weinert, B.T., Iesmantavicius, V., Moustafa, T., Scholz, C., Wagner, S.A., Magnes, C., Zechner, R., and Choudhary, C. (2014). Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae. Mol Syst Biol *10*, 716.

Wenham, M., Grieve, S., Cummins, M., Jones, M.L., Booth, S., Kilner, R., Ancliff, P.J., Griffiths, G.M., and Mumford, A.D. (2010). Two patients with Hermansky Pudlak syndrome type 2 and novel mutations in AP3B1. Haematologica *95*, 333-337.

Werner, M., Thuriaux, P., and Soutourina, J. (2009). Structure-function analysis of RNA polymerases I and III. Curr Opin Struct Biol *19*, 740-745.

Werner, T.P., Amrhein, N., and Freimoser, F.M. (2007). Specific localization of inorganic polyphosphate (poly P) in fungal cell walls by selective extraction and immunohistochemistry. Fungal Genet Biol *44*, 845-852.

Whyte, C.S., Chernysh, I.N., Domingues, M.M., Connell, S., Weisel, J.W., Ariens, R.A., and Mutch, N.J. (2016). Polyphosphate delays fibrin polymerisation and alters the mechanical properties of the fibrin network. Thromb Haemost *116*, 897-903.

Wild, R., Gerasimaite, R., Jung, J.Y., Truffault, V., Pavlovic, I., Schmidt, A., Saiardi, A., Jessen, H.J., Poirier, Y., Hothorn, M.*, et al.* (2016). Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. Science *352*, 986-990.

Wurst, H., Shiba, T., and Kornberg, A. (1995). The gene for a major exopolyphosphatase of Saccharomyces cerevisiae. J Bacteriol *177*, 898-906.

Xie, L., and Jakob, U. (2019). Inorganic polyphosphate, a multifunctional polyanionic protein scaffold. J Biol Chem *294*, 2180-2190.

Yang, W., Li, C., Ward, D.M., Kaplan, J., and Mansour, S.L. (2000). Defective organellar membrane protein trafficking in Ap3b1-deficient cells. J Cell Sci *113*, 4077-4086.

Yeon, J.H., Mazinani, N., Schlappi, T.S., Chan, K.Y., Baylis, J.R., Smith, S.A., Donovan, A.J., Kudela, D., Stucky, G.D., Liu, Y.*, et al.* (2017). Localization of Short-Chain Polyphosphate Enhances its Ability to Clot Flowing Blood Plasma. Sci Rep *7*, 42119.

Zarin, T., Strome, B., Nguyen Ba, A.N., Alberti, S., Forman-Kay, J.D., and Moses, A.M. (2019). Proteome-wide signatures of function in highly diverged intrinsically disordered regions. eLife *8*, e46883.

Zhang, H., Ishige, K., and Kornberg, A. (2002). A polyphosphate kinase (PPK2) widely conserved in bacteria. Proc Natl Acad Sci U S A *99*, 16678-16683.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T.*, et al.* (2001). Global analysis of protein activities using proteome chips. Science *293*, 2101-2105.

Zhu, J., Liu, X., Anjos, M., Correll, C.C., and Johnson, A.W. (2016). Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. Mol Cell Biol *36*, 2101-2105.

Zhu, L., Jorgensen, J.R., Li, M., Chuang, Y.S., and Emr, S.D. (2017). ESCRTs function directly on the lysosome membrane to downregulate ubiquitinated lysosomal membrane proteins. Elife *6*.
Zuberi, K., Franz, M., Rodriguez, H., Montojo, J., Lopes, C.T., Bader, G.D., and Morris, Q. (2013). GeneMANIA prediction server 2013 update. Nucleic Acids Res *41*, W115-W122.

Chapter 7: Appendices

7.1: Appendix A

Table 1: PASK-like proteins in *Saccharomyces cerevisiae*

*Note – this table has been amended for ease of inclusion in this thesis. Only the first 6 columns are included. For the full table, see (Bentley-DeSousa et al., 2018).

ProteinAcc	ProteinName	GeneName	Hit	SeqStart	SeqStop
P38904	Protein SPP41	**SPP41**	AAPNDKLSSKKSSKKKKKDKSK	203	224
Q01477	Ubiquitin carboxyl-terminal hydrolase 3	**UBP3**	ADDAEKTTESDAKKSSKSKS	531	550
Q03281	Inner nuclear membrane protein HEH2	**HEH2**	ADRKKTLKSKKLESSSSESK	72	91
P40316	Securin	**PDS1**	ADSGKNEESSDDDEGNEDSESK	177	198
Q05468	Uncharacterized protein YDR333C	**YDR333C**	AEAAKDKGSDDDDDDEEFDK	111	130
Q04660	Ribosome biogenesis protein ERB1	**ERB1**	AEASESDEDDDEYESAVEEKESSSDKEAQDDSDDDSDAELNK	41	82
P40161	Histone chaperone RTT106	**RTT106**	AEDGEEAEEDDEEDDEEEDKKGQSA	378	402
P27476	Nuclear localization sequence-binding protein	**NSR1**	AEEEEDEESSNKKQKNEETEE	146	166
P05453	Eukaryotic peptide chain release factor GTP-binding subunit	**SUP35**	AESDKKEEEKSAETKEPTKE	166	185
Q12460	Nucleolar protein 56	**NOP56**	AEVEETKEKESSKKRKLEDDDEEKKEKKEKKSKKEKKEKK EKKDKKEKKDKKEKKDKKKKSKD	442	504
P38174	Methionine aminopeptidase 2	**MAP2**	AKADESDPVESKKKKNKKKKKKKSN	30	54
P34218	Histone acetyltransferase SAS3	**SAS3**	AKEDEDEDEDFTLDDDIEDE	754	773
P32642	Monothiol glutaredoxin-4	**GRX4**	AKGPKSTSDEESSGSSDDEEDE	119	140
P36100	Transcription initiation factor IIE subunit alpha	**TFA1**	AKKQAKLNKEEEEEEEEEEDEEEEEEEEMEDVMDDNDE	365	402
P38779	Proteasome-interacting protein CIC1	**CIC1**	AKKRSSSELEKESSESEAVKK	353	373
P23201	Protein SPA2	**SPA2**	AKLASLTSEKEKEKKEEKEEKEKEKN	312	337
P32521	Actin cytoskeleton-regulatory complex protein PAN1	**PAN1**	AKQESDDEDEDDEEKRLQEE	1131	1150
Q06810	Protein OPY2	**OPY2**	AKQSDSQLSSLSSSSSSSSS	64	84
Q12355	Cell wall mannoprotein PST1	**PST1**	ASASSVSSSGASSSSSKSSK	398	417
P38079	Protein YRO2	**YRO2**	ASFKSSDSDSSIKEKLKLKKK	292	312
P22146	1,3-beta-glucanosyltransferase GAS1	**GAS1**	ASGSSSKSNSGSSGSSSSSSSSSASSSSSSKKNA	496	529
O13547	Covalently-linked cell wall protein 14	**CCW14**	ASSSSASSSTKASSSSASSS	113	132
O13547	Covalently-linked cell wall protein 14	**CCW14**	ASSSSKASSSTKASSSSASSS	101	121
Q12117	Protein MRH1	**MRH1**	ATPNLSKDKKKKSKKSKKSKKSKKSEE	294	320
P12954	ATP-dependent DNA helicase SRS2	**SRS2**	CVFNDDKKDESEEDEEEDQE	672	691
Q05022	rRNA biogenesis protein RRP5	**RRP5**	DAADEKKSKVEDAEYESSDDEDEKLDKSNE	171	200
Q12499	Nucleolar protein 58	**NOP58**	DADTAKAASDSESDSDDEEEEKKEKKEKKRKRDDDEDSKDSKKA KKEKKDKKEKKEKKEKKEKKEKKEKKEKKSKKEKKEKK	430	511
P16140	V-type proton ATPase subunit B	**VMA2**	DDADEDEEDPDTRSSGKKKD	490	509
P32432	Transcription factor SFP1	**SFP1**	DDDDDDDDDDTENGSSSNGKS	546	566
Q08438	Phosphopantothenoylcysteine decarboxylase subunit VHS3	**VHS3**	DDDDDDDEDDEDEDEDDEGKKKEDKGGLQRS	644	674
P38959	Vacuolar protein sorting-associated protein 41	**VPS41**	DDDDDDDEDEDDEDEPPLLKY	83	103
Q12206	Transcriptional modulator WTM2	**WTM2**	DDDDDDDNDDDDEEGNGKTKS	87	107
P40572	Uncharacterized protein YIR016W	**YIR016W**	DDDDDDEDDDDSSMSRDLSKALDMSSSSSSSPR	90	122
Q05543	Regulator of Ty1 transposition protein 103	**RTT103**	DDDDDDGDNDDDDDDDDDDKNYDDRSNDSN	269	298
P51862	RHO1 GDP-GTP exchange protein 2	**ROM2**	DDDDEEKPSWTSSVSKEDWE	632	651
P40157	Vacuolar import and degradation protein 27	**VID27**	DDDDELDSSSDDFQDAKDTS	188	207
Q07897	Protein CMS1	**CMS1**	DDDDVDDIDGKKEERNSEDDS	43	63
P38439	RNA polymerase-associated protein LEO1	**LEO1**	DDDDVKKSETEKSDSDSDEDDEGENINH	71	98
Q12515	Protein PAR32	**PAR32**	DDDEDKFNLKDDNGKEKKKKKKKKSG	260	285
P17890	DNA-directed RNA polymerase III subunit RPC7	**RPC31**	DDDFEEDEDEEDDDDYNAEKY	210	230
P07250	Inositol polyphosphate multikinase	**ARG82**	DDDNDDDDDDDAEGSSEGPKDKKT	291	314

P54784	Origin recognition complex subunit 1	**ORC1**	DDEDKDLIESNKAKDDNDDDDD	745	766
P38691	Serine/threonine-protein kinase KSP1	**KSP1**	DDENDCDSSDCDDDEDTKVD	63	82
P36053	Transcription elongation factor 1	**ELF1**	DDGDEGSDSDYESDSEQDAK	89	108
P32567	Phosphatidic acid phosphohydrolase 1	**PAH1**	DDKDLDSRVSDEFDDDEFDEDE	838	859
P38439	RNA polymerase-associated protein LEO1	**LEO1**	DDLFGDDNDDDDDDDVKKSE	60	79
P42846	Protein KRI1	**KRI1**	DDQTEEGEVEKEQKEEDEEE	399	418
P36158	Uncharacterized protein YKR078W	**YKR078W**	DDSERSVKSVSVSISDDEDSK	27	47
P39523	Uncharacterized protein YMR124W	**YM8564.06**	DDSSSVSFSDFEKELRKKSSK	861	881
P24276	Protein SSD1	**SSD1**	DDVWESKKEKEEKKRRKDAS	411	430
Q12161	RING finger protein PSH1	**PSH1**	DEDDEEEDEEEEEEMDSDLKDFIEDDEDDEDEDGSRRNLV	259	298
P34248	Probable intramembrane protease YKL100C	**YKL450**	DEDKKKKKKSKRFFDMMDEK	137	156
Q99299	Altered inheritance of mitochondria protein 44	**AIM44**	DEEDEEDEEDDEDDEEEEEKRKEGEGRNLAKEVDEL	672	707
P53927	Ribosome biogenesis protein 15	**NOP15**	DEEDEKDEDEIEGLAASDDE	40	59
P38911	FK506-binding nuclear protein	**FPR3**	DEEDNDGEEEQEEEEEEEQKEEVKPE	223	248
P40340	Tat-binding homolog 7	**YTA7**	DEEEEEADEFEDEYLDEDSKDN	134	155
P36044	Protein MNN4	**MNN4**	DEEKKKQEEEEKKKNEEEEKKKQEEGHSN	1150	1178
Q04373	Pumilio homology domain family member 6	**PUF6**	DELDISDDSEEHENENEEKE	72	91
Q03880	V-type ATPase assembly factor PKR1	**PKR1**	DELDKDANKKDDSAIKEDSEDK	79	100
Q06214	WD repeat-containing protein JIP5	**JIP5**	DELDKDLKEDHQEEKESNSKS	445	465
Q06205	FK506-binding protein 4	**FPR4**	DELVKKDEKKKNNKKDSKRK	225	244
P32583	Suppressor protein SRP40	**SRP40**	DEVPKLSVKEKEIEEKSSSSSSSSSSSSSSSSSSSSSSSSSGESS SSSSSSSSSSSSDSSDSSDSESSSSSSSSSSSSSSSSSDSESSSESD SSSSGSSSSSSSSSDESSSESESEDETKKRARESDNEDAKETKK AKTEPESSSSSESSSSGSSSSSESESGSESDSDSSSSSSSSSDSES DSESDSQSSSSSSSSDSSSDSDSSSSDSSSDSDSSSSSSSSSSDS DSDSDSSSDSDSSGSSDSSSSSDSSSDESTSSDSSDSDSDSDSGSS SELETKEATADESKAEETPASSNESTPSASSSSSANKLNIPAGTDEI KEGQRKHFSRVDRSKINFEAWELTDNTYKGAAGTWGEKANEKLG RVRGKDFTKNKNKMKRGSYRGGSITLESGSYKFQD	9	406
P40991	25S rRNA (cytosine(2870)-C(5))-methyltransferase	**NOP2**	DEVSDRKKKKSKPFKKSRKEEEEV	43	66
P38041	Protein BOB1	**BOI1**	DGSSEEKSKKKLFSSTKESF	707	726
P15442	Serine/threonine-protein kinase GCN2	**GCN2**	DHDTSSTSSSESQDDTDKESKS	740	761
P53072	tRNA acetyltransferase TAN1	**TAN1**	DIKEGEDESENDEKKDLSIED	64	84
Q12425	Iron-sulfur assembly protein 2	**ISA2**	DIKNDVKDKEFSDDLDDDDSKD	105	126
P34250	Eisosome protein SEG2	**SEG2**	DKDENVDDDENVDDDDDDDDDDDDEYHDSYDV	619	650
P53261	Pescadillo homolog	**NOP7**	DKDVNKSKNKKRKVDEEEEEKK	532	553
Q04712	RNA polymerase I-specific transcription initiation factor RRN11	**RRN11**	DKEDSDEERKFWKKYEKPEKS	117	137
P40328	Probable 26S protease subunit YTA6	**YTA6**	DKESSASSSLDSRKEDILKS	426	445
P53335	Protein PXR1	**PXR1**	DKKEHKKHKKEEKRLKKEKR	193	212
Q02455	Protein MLP1	**MLP1**	DKLESKEKKSKQSLQKIESE	558	577
P53067	Importin subunit beta-5	**KAP114**	DKLKEYIDDDVDEEADDDSDD	934	954
P47035	Nucleolar protein NET1	**NET1**	DKNNSKEKEDSKSKQVSQKK	907	926
P35177	Transcriptional activator SPT7	**SPT7**	DLEKEIEDMEKDKDYELDEEEE	556	577
Q05777	Cell wall protein YLR194C	**YLR194C**	DNTASSSKSSSGSSSKPESS	170	189
P40052	Uncharacterized protein YER079W	**YER079W**	DPKNKKKNKKKKNDKDDKYK	118	137
P53336	Putative methyltransferase YGR283C	**YGR283C**	DQSKDKKHEKKSRSKETISD	81	100
Q04958	Lysophospholipase NTE1	**NTE1**	DRDEEDEEEDDDIDNYDTKS	281	300
P25588	Mediator of replication checkpoint protein 1	**MRC1**	DSDSDEYGDDDMDSIKLSKDES	408	429
P38852	Protein LIN1	**LIN1**	DSDSDSSVEDSTDNENSGKE	47	66
Q12334	Protein SCM3	**SCM3**	DSEASDSEVDADNDTEEEKDEK	182	203
Q03834	DNA mismatch repair protein MSH6	**MSH6**	DSESDEDEYLPDKNDGDEDDD	165	185
P23643	Vacuolar protein sorting-associated protein 3	**VPS3**	DSESSPHERENDKKKTEDDS	24	43
P40856	SIT4-associating protein SAP185	**SAP185**	DSNSTDESLESNSFKSEDEEEE	42	63
P53874	Ubiquitin carboxyl-terminal hydrolase 10	**UBP10**	DSSAMEKELPEEEENDSSSK	242	261

P32504	Centromere DNA-binding protein complex CBF3 subunit A	**CBF2**	DSSDESSTEDSNVFKKDGES	681	700
P32583	Suppressor protein SRP40	**SRP40**	DSSDSDSDSDSGSSSELETKEATADESKAEET	267	298
P43124	Non-structural maintenance of chromosome element 4	**NSE4**	DSSGETRKQKKSRSDEKSSSSKDGD	19	43
Q02208	Topoisomerase 1-associated factor 2	**TOF2**	DSSSSDDSDDDSDSRNVQVKK	707	727
P28319	Cell wall protein CWP1	**CWP1**	DSSSSSSASASSASASSSTK	151	170
P53599	MAP kinase kinase kinase SSK2	**SSK2**	DSTEKPSSSKMRQSSSSSSS	14	33
Q06132	Suppressor of glycerol defect protein 1	**SGD1**	DVIKKFEDKEFDSESDTDEEDD	724	745
Q03016	GLC7-interacting protein 3	**GIP3**	DVLGSSRDDSKESVDSSEDSK	1013	1033
P53734	ATP-dependent RNA helicase DBP6	**DBP6**	DYGASEEDSSEVEEEESKPS	69	88
P40352	DNA repair and recombination protein RAD26	**RAD26**	DYQMSGKESEDDEEEENDDK	202	221
P05318	60S acidic ribosomal protein P1-alpha	**RPP1A**	EAGEAEAEKEEEEAKEESDDD	79	99
P53950	Vacuolar segregation protein 7	**VAC7**	EAKHENDDDDDDDGDDGDDE	1079	1098
P38708	Putative proline--tRNA ligase YHR020W	**YHR020W**	ECEEDIKESSAKKDDGEEFEEDDKA	624	648
P14682	Ubiquitin-conjugating enzyme E2-34 kDa	**CDC34**	EDDDDDDDDSIDNDSVMDRK	254	273
P36024	Phosphopantothenoylcysteine decarboxylase subunit SIS2	**SIS2**	EDDDEEEDTEDKNENNNDDDDDDDDDDDDDD DDDDDDDDDDEDEDEAETPGIIDKHQ	506	562
P53874	Ubiquitin carboxyl-terminal hydrolase 10	**UBP10**	EDDEDEDDMDYDSSAMEKEL	231	250
P46949	Protein FYV8	**FYV8**	EDDGEVESSGALEKKEKSEEK	273	293
Q03388	Vacuolar protein sorting-associated protein 72	**VPS72**	EDEDDEDFEMMAKDDDDEGEEKEDET	69	94
P53207	U1 small nuclear ribonucleoprotein component SNU71	**SNU71**	EDEDEEDDTEDDLALEKRKEE	343	363
Q06523	Uncharacterized protein YPR148C	**YPR148C**	EDEEDEEDEEDDEDLSNLIKV	197	217
P38996	Nuclear polyadenylated RNA-binding protein 3	**NAB3**	EDEEDKKDKTKDKEVELRRE	154	173
P33441	THO complex subunit MFT1	**MFT1**	EDEENEEEDDMEVDVEDIKED	305	325
P47035	Nucleolar protein NET1	**NET1**	EDESSSSSSSSDEETSTSRKAR	990	1011
P32447	Histone chaperone ASF1	**ASF1**	EDGEGEAEEAAEEEEEEEEK	195	214
P40477	Nucleoporin NUP159	**NUP159**	EDHQSSEEDASEKDSRQSSE	1037	1056
P53743	Pre-rRNA-processing protein ESF2	**ESF2**	EDIESENESDIEEEQKQEEKED	43	64
Q05022	rRNA biogenesis protein RRP5	**RRP5**	EDLDEDMDSDTDAADEKKSK	160	179
P38277	Uncharacterized protein YBR138C	**YBR1014**	EDNKIDLKKEISKKKDSFSKE	438	458
P41811	Coatomer subunit beta'	**SEC27**	EDSEFKESNSEAVEAEKKEEE	844	864
P36775	Lon protease homolog, mitochondrial	**PIM1**	EDSPTSSADSKPKESVSSEEK	844	864
P40362	U3 small nucleolar RNA-associated protein 18	**UTP18**	EDSQDKNEDTMDVDDEDDSSSDDYSEDSEEAA	93	124
Q08979	Kelch repeat-containing protein 3	**KEL3**	EEDDESDQDLDEILSSFSKK	42	61
P08638	Regulatory protein LEU3	**LEU3**	EEDEDEEGEEEEEEEELSSKVPENMDSQ	682	709
P12954	ATP-dependent DNA helicase SRS2	**SRS2**	EEDEEEDQENSKKDASPKKT	683	702
P53741	UBP3-associated protein BRE5	**BRE5**	EEDKIRHESGVEKEKEKEKS	168	187
P04786	DNA topoisomerase 1	**TOP1**	EEEDKKAKEEEEEYKWWEKENEDDT	118	142
P46949	Protein FYV8	**FYV8**	EEEDNEDNESIEDKNEDNES	397	416
Q07915	Ribosome biogenesis protein RLP24	**RLP24**	EEEEDMEIDSDEEEEQLEKQK	163	184
P32784	Glycerol-3-phosphate O-acyltransferase 1	**SCT1**	EEEEEEEEEEEEEEEEEGKEGDA	736	759
P43597	Uncharacterized protein YFR016C	**YFR016C**	EEEEEEEEENSTFSKVKKEN	469	488
P32447	Histone chaperone ASF1	**ASF1**	EEEEEEEEKTEDNETNLEEEEEDIE	206	230
Q06631	Protein BFR2	**BFR2**	EEEEKEEDALSFRTDSEDEE	106	125
Q03338	U4/U6 small nuclear ribonucleoprotein PRP3	**PRP3**	EEEELKRKLKQEEDEKDKRK	113	132
P38911	FK506-binding nuclear protein	**FPR3**	EEEEQKEEVKPEPKKSKKEKKRKHEEKEEEKKA KKVKKVEFKKDLEEG	237	284
P38080	Serine/threonine-protein kinase AKL1	**AKL1**	EEEKESQSDQRKSTLSEDKSS	415	435
P36046	Mitochondrial intermembrane space import and assembly protein 40	**MIA40**	EEESSKKTVSDSENSAKQSESSDEEKEELRKQEEKQ	236	271
Q06839	PX domain-containing protein YPR097W	**YPR097W**	EEGEGEESDFDEYKDASDSK	444	463
P23255	Transcription initiation factor TFIID subunit 2	**TAF2**	EEGKNSKSKDAQDNDEEEEE	296	315
P00127	Cytochrome b-c1 complex subunit 6	**QCR6**	EEKAAEGEEKEEENGDEDEDEDEDEDDDDDDDE DEEEEEEVTDQLEDLREHFK	34	86

P53911	Chromatin modification-related protein EAF7	**EAF7**	EEKEEIEEDNTKDEEQKEKKEEIQE	258	282
P47136	Bud site selection protein 4	**BUD4**	EEKEVETKDENIETEKDESE	556	575
Q06218	ATP-dependent RNA helicase DBP9	**DBP9**	EEKSLEGEPENDKKPSKKKK	353	372
P29056	Transcription factor IIIB 70 kDa subunit	**BRF1**	EEKVKKVKTKTSEEKKENES	401	420
P20484	Protein MAK11	**MAK11**	EELFEKVEENEKQEDDDDKED	323	343
Q04213	ISWI one complex protein 4	**IOC4**	EEMDREKPSFSEDVKEEESK	453	472
P25582	27S pre-rRNA (guanosine(2922)-2'-O)-methyltransferase	**SPB1**	EEMHQKQDEADSSDESSSDDSD	599	620
P40085	Protein TAPT1 homolog	**YER140W**	EEQDESSAAKSEEEHEDDYE	51	70
P53104	Serine/threonine-protein kinase ATG1	**ATG1**	EEQDSDDAEEEDETLKKYKED	643	663
P39730	Eukaryotic translation initiation factor 5B	**FUN12**	EEQEKLEREEEERLEKEEEE	184	203
P38996	Nuclear polyadenylated RNA-binding protein 3	**NAB3**	EEQHELEDVNDEEEEDKEEK	57	76
Q06525	Pre-mRNA-splicing factor URN1	**URN1**	EESDSEDNSEDDSEVLEPTK	274	293
P38700	Adaptin medium chain homolog APM2	**APM2**	EESSSSSGSDSDSEYSNTNK	156	175
P40517	Ran-specific GTPase-activating protein 2	**YRB2**	EETSDKSEAPNKNDEEKKEE	45	64
P53911	Chromatin modification-related protein EAF7	**EAF7**	EEVDKEKNENEEGDDEREKS	289	308
P39985	DNA polymerase V	**POL5**	EGEEEFEEIKEENDASEDESK	703	723
P39985	DNA polymerase V	**POL5**	EGLSDDGGDDEDEESMDDEK	786	805
Q12532	Translation-associated element 2	**TAE2**	EGNTTDSSSDSDSEDMESSKE	470	490
P38789	Ribosome biogenesis protein SSF1	**SSF1**	EGQGKDDAMSDDESSSSDSE	416	435
P53842	Putative uncharacterized protein YNL266W	**N0800**	EGSLKSEEMLEDDEDKEFSS	106	125
Q03661	Silent chromatin protein ESC1	**ESC1**	EHELDDKELIEDIESSDSES	391	410
P53119	Probable E3 ubiquitin-protein ligase HUL5	**HUL5**	EHLEEESDEDMEKEIDLDKE	457	476
P53317	Nucleolar protein 19	**NOP19**	EHSKSSVSNSYSASDEEDED	152	171
P42846	Protein KRI1	**KRI1**	EIKSAFSDEENEESSGDEDD	171	190
P07276	DNA repair protein RAD2	**RAD2**	EILERESEKESSNDENKDDD	642	661
Q04500	U3 small nucleolar RNA-associated protein 14	**UTP14**	EKAELKQKKKKKGKSNDDED	699	718
P38881	Nucleus-vacuole junction protein 1	**NVJ1**	EKDCSSSSEVESQSKCRKES	231	250
O13555	Putative uncharacterized protein JIP3	**JIP3**	EKDEEDEESEEAEEELLVLESDEKL	2	26
P40480	Protein HOS4	**HOS4**	EKDIEYSDSDKDTDDVGSDD	198	217
P53903	Processing of GAS1 and ALP protein 2	**PGA2**	EKDKRDKEEKRSKDLIDKPDD	55	75
Q02792	5'-3' exoribonuclease 2	**RAT1**	EKDNEEEEIAKDSKKVKTEK	521	540
P33311	ATP-dependent permease MDL2, mitochondrial	**MDL2**	EKEDLNESKEHDDQKKDDNDD	705	725
P38911	FK506-binding nuclear protein	**FPR3**	EKEEEKKAKKVKKVEFKKDLEE	262	283
P53935	Stress response protein NST1	**NST1**	EKEEKKQKKREKEKEKKRLQ	640	659
P38712	ATP-dependent rRNA helicase RRP3	**RRP3**	EKEGGSESDSEEDATAEKKK	38	57
Q04632	Conserved oligomeric Golgi complex subunit 8	**COG8**	EKEKEEEVNTKDNKAEKEEEE	450	470
P53741	UBP3-associated protein BRE5	**BRE5**	EKEKEKEKSPEISKPKAKKET	179	199
P53874	Ubiquitin carboxyl-terminal hydrolase 10	**UBP10**	EKEKRKELSTSISSDDEIEDDEDEDDMDYDSSA	213	245
Q12460	Nucleolar protein 56	**NOP56**	EKKDKKEKKDKKEKKDKKKKSKD	482	504
Q06640	F-box protein YDR306C	**YDR306C**	EKKDKKGRKQSPSSSSTSSSK	71	91
P38789	Ribosome biogenesis protein SSF1	**SSF1**	EKKDNKEREKETEEEDVEMEE	290	310
P33322	H/ACA ribonucleoprotein complex subunit 4	**CBF5**	EKKEKKEKKRKSEDGDSEEKKSKKSKK	457	483
P32499	Nucleoporin NUP2	**NUP2**	EKKNDENSTSNSKPEKSSDS	326	345
Q06205	FK506-binding protein 4	**FPR4**	EKKQTTKKDKKAEKVKDSEESK	257	278
Q08421	Enhancer of translation termination 1	**ETT1**	EKKRKVESVEKKSDEPSRES	14	33
P43564	Uncharacterized membrane protein YFL034W	**YFL034W**	EKNNDDDDDDSYDEYEDDVE	249	268
P53301	Probable glycosidase CRH1	**CRH1**	EKQSSSSSKKTVASSSTSES	384	403
P53838	Boron transporter 1	**BOR1**	ELDRESSVSESRSDEESHEK	30	49
P35732	RNA polymerase II degradation factor 1	**DEF1**	ELKKEISDIKKDDQKSEASEEK	267	288
P34218	Histone acetyltransferase SAS3	**SAS3**	ENDEEEDTYEEDSDDDEDGKRK	778	799
P31385	Transcriptional regulatory protein DEP1	**DEP1**	ENEDESKEEEKSQELEEAIDSKEKSTDA	109	136
P35177	Transcriptional activator SPT7	**SPT7**	ENKEAGENNEEEEDDDDEDEDEDM	692	715
P40578	Protein MGA2	**MGA2**	ENSENSEEEDYEEEEEFLKK	1001	1020
P33441	THO complex subunit MFT1	**MFT1**	ENVDEKEESDHEYDDQEDEE	289	308
Q05515	Survival factor 1	**SVF1**	EQALQGEDEKEDEKEDEEEEE	368	388

Q12176	Ribosome biogenesis protein MAK21	**MAK21**	EQDNSDKKRSFAESSEEDESSEEEKEEEENKEVSAKR	958	994
P36121	DNA-directed RNA polymerase III subunit RPC5	**RPC37**	EQEEKSEEVKAEDDTGEEEEDD	47	68
Q12329	Heat shock protein 42	**HSP42**	EQEGEKGEEKDKKDKSEAPKEE	167	188
P42846	Protein KRI1	**KRI1**	EQKEEDEEEGPKRKKSKKEEKL	410	431
P53107	Ran-specific GTPase-activating protein 30	**YRB30**	ERKEEEEEENKLEDKDSSIT	361	380
P35178	Ribosomal RNA-processing protein 1	**RRP1**	ERLLKDGDEDDEDEENKEEE	182	201
Q04119	Endopolyphosphatase	**PPN1**	ERVDEKNVENEKKKKEKKDKKKKKP	448	472
P47018	Maintenance of telomere capping protein 1	**MTC1**	ESEDEDEDDEIIDPSEWVKE	435	454
Q04373	Pumilio homology domain family member 6	**PUF6**	ESELSKKEDAVSSSSDDDDLDDLSTSDSEAEEEADELD	38	75
P36134	Uncharacterized protein YKR041W	**YKR041W**	ESERDRLKRSHESESSSESE	113	132
Q12416	G1-specific transcriptional repressor WHI5	**WHI5**	ESESEDDENEEENENQKKYD	112	131
P25386	Intracellular protein transport protein USO1	**USO1**	ESKAKVEEGLKKLEEESSKEK	1477	1497
Q02455	Protein MLP1	**MLP1**	ESKKIKTEDEEEKETDKVNDE	1852	1872
P32380	Spindle pole body component 110	**SPC110**	ESLRKDIEEYKESAKDSEDK	554	573
P25588	Mediator of replication checkpoint protein 1	**MRC1**	ESSGSENDNRRESDSEKEDDE	567	587
Q04177	U3 small nucleolar RNA-associated protein 5	**UTP5**	ESYGSEEEEEGDSDNEEEQK	595	614
P36076	Coenzyme A biosynthesis protein 3	**CAB3**	ETASDESNDEEDEEDEEDVKTE	549	570
P36123	SIT4-associating protein SAP190	**SAP190**	ETNDDDYDSDDGKSKSAESE	849	868
P25357	Probable DNA-binding protein SNT1	**SNT1**	EVDESKEESTNTIEKEEKSE	753	772
P11745	Ran GTPase-activating protein 1	**RNA1**	EVDSEDEEGEDEEDEDEDEKLEEIETERLEKEL	357	389
P53333	Pre-mRNA-splicing factor CWC22	**CWC22**	EVEDAEEEEKKLREEEELEK	486	505
P22936	DNA-(apurinic or apyrimidinic site) lyase 1	**APN1**	EWLESKSESELLEDKEYKEK	291	310
Q06810	Protein OPY2	**OPY2**	FDDEDDDDEDDEDDGGMRKDES	146	167
Q01389	Serine/threonine-protein kinase BCK1/SLK1/SSP31	**BCK1**	FDSDDSDDDSSSSDDIIWSKKKTA	966	989
P36097	TEL2-interacting protein 1	**TTI1**	FDSKLREPDSDDDEEEREEE	808	827
P23615	Transcription elongation factor SPT6	**SPT6**	FDSSEEDEDIDEDEDEARKV	37	56
Q03280	E3 ubiquitin-protein ligase TOM1	**TOM1**	FKIEDHDEVEDEVDESDKEE	1875	1894
P32497	Eukaryotic translation initiation factor 3 subunit C	**NIP1**	FNESESESEADVDSDDSDAK	49	68
P40477	Nucleoporin NUP159	**NUP159**	FPSSGDQSEDESKSDVDSSS	667	686
P48561	Poly(A) RNA polymerase protein 1	**TRF5**	FSDSSEDETEQIKEDDDERSS	110	130
Q03388	Vacuolar protein sorting-associated protein 72	**VPS72**	FSSSESEDSSNENDEDAEEKEI	115	136
P38911	FK506-binding nuclear protein	**FPR3**	GDFDEDEIDEESSEEEEEEKTQKKKKSKGKKAESESED DEEDDDEDDEFQESVLLTLS	69	126
P22517	Calcium/calmodulin-dependent protein kinase II	**CMK2**	GEEEDDSKKTLHDDRESKSED	427	447
P25618	Protein CWH43	**CWH43**	GEEKSSELQKSGEKKVEKEK	234	253
P38764	26S proteasome regulatory subunit RPN1	**RPN1**	GESLEGEEIKSEEKKGKSSDKD	693	714
P02400	60S acidic ribosomal protein P2-beta	**RPP2B**	GGDAAEEEKEEEAKEESDDD	84	103
P47068	Myosin tail region-interacting protein MTI1	**BBC1**	GHEDREGDNDEEKEEEDSEE	393	412
Q04217	Probable ATP-dependent RNA helicase DHR1	**ECM16**	GKEDEMDFDTTSEDDDEEED	261	280
P47018	Maintenance of telomere capping protein 1	**MTC1**	GKEKEPQQQEKEEEEEEEEEEEEEETP	115	143
P53313	Protein SDA1	**SDA1**	GKESDSDLELSDDDDEKEVKDEQEDADID	600	628
P25037	Ubiquitin carboxyl-terminal hydrolase 1	**UBP1**	GKHKKKLNRKSSSKEDEEKS	173	192
P38798	Nonsense-mediated mRNA decay protein 2	**NMD2**	GKHQAKQDESEDEDDEDDDEDDDDDDDDD DDDGEEGDEDDDEDDDDEDDDDEEEEDSDS DLEYGGDLDADRDIEMKRMYEEYERK	877	961
Q12136	Something about silencing protein 10	**SAS10**	GKQKQEIDEEERESDEEEEEED	345	366
P40019	Histone H2A.Z-specific chaperone CHZ1	**CHZ1**	GSKDDVGYGEKEEDDEDEEDDD	129	150
P36024	Phosphopantothenoylcysteine decarboxylase subunit SIS2	**SIS2**	GYPKNNEEEDDDEDEEEDDDEEEDTEDKNE	490	519
Q07896	Nucleolar complex-associated protein 3	**NOC3**	HKAQEKPKDDDEEDEDSNDSSEDDEG	89	114
Q12176	Ribosome biogenesis protein MAK21	**MAK21**	HKLEDVDDDDIEEESTSSKESK	113	134
P40986	Cell division control protein 1	**CDC1**	HSKKSDDKAHYKSRSKKKSKS	14	34
Q01919	Serine/threonine-protein kinase KIN4	**KIN4**	HSSNKSDASSDKDSKKIYEKK	648	668

Q08208	Nucleolar protein 12	**NOP12**	IAETEESGNEKEEESSKKSD	178	197
P53847	Protein transport protein DSL1	**DSL1**	IDDEEEKTNQEKEKEPEEEE	433	452
P40157	Vacuolar import and degradation protein 27	**VID27**	IDDRSNEERDKESSESENDSEDEDDENDHSKR	371	402
P32892	ATP-dependent RNA helicase DRS1	**DRS1**	IDSKQEEETEKEKVEKENDSDDEEL	126	150
P42846	Protein KRI1	**KRI1**	IEASKSYVPSEDEEEDEEEEEEEDDY	43	68
P39730	Eukaryotic translation initiation factor 5B	**FUN12**	IEEKKDGKPILKSKKEKEKEKKEKEKQKKKEQA	78	110
P32833	Origin recognition complex subunit 2	**ORC2**	IKKDEKDTISSKKKRKLDKD	92	111
P39730	Eukaryotic translation initiation factor 5B	**FUN12**	IKSDSKKDSEVVPDDELKESED	297	318
P35724	Manganese resistance protein MNR2	**MNR2**	ILKMHQSDDSSDSDSSDSDSDS	741	762
Q08054	Chitin synthase 3 complex protein CSI2	**CSI2**	INDDESETSSHSDDSDISEK	188	207
P38703	Sphingosine N-acyltransferase LAG1	**LAG1**	IQKDERSDSDSDESAENEESKEKCE	387	411
P27637	Bud site selection protein 14	**BUD14**	ISEDEDEEENYSDDDDFDRK	211	230
Q08952	Oxidation resistance protein 1	**OXR1**	ISEKEESEQEGKEGKEEGDKEE	177	198
Q12211	tRNA pseudouridine synthase 1	**PUS1**	KADFDDEKDKKKDNDKHIDK	32	51
P38911	FK506-binding nuclear protein	**FPR3**	KAESESEDDEEDDDEDDEFQESVLL	99	123
P53946	Actin-related protein 5	**ARP5**	KAKEEKERVAKEEEEKKLKE	430	449
Q03640	Tricalbin-3	**TCB3**	KAKGKKEDENEDTTEKEEDE	634	653
P23615	Transcription elongation factor SPT6	**SPT6**	KAPSEEEEGEDVFDSSEEDED	25	45
Q06132	Suppressor of glycerol defect protein 1	**SGD1**	KDDEASEKAFSSDDDLSASD	254	273
P32773	Transcription initiation factor IIA large subunit	**TOA1**	KDDEEKEEDVEKTRKEKEQIE	178	198
P23255	Transcription initiation factor TFIID subunit 2	**TAF2**	KDEDKDSNLKNDEEGKNSKSKD	284	305
P38129	Transcription initiation factor TFIID subunit 5	**TAF5**	KDEEKKESELKVDGEKKDSNLSSP	393	416
P07213	Mitochondrial import receptor subunit TOM70	**TOM70**	KDEKKDTKDSQKETEGAKKS	46	65
P17442	Phosphate system positive regulatory protein PHO81	**PHO81**	KDISETNDEEDDDEISEDHDD	770	790
P36168	EST/SMG-like protein 1	**ESL2**	KDKKERKKSSNNDSSVTESS	609	628
P43610	Uncharacterized ATP-dependent helicase IRC5	**IRC5**	KDNSNSDDEEHSSKKRKTKKKS	164	185
P17883	Superkiller protein 3	**SKI3**	KDRVEDEVKDQDEEAKEDEEED	376	397
P43586	60S ribosomal subunit assembly/export protein LOC1	**LOC1**	KEALKQDKLEEKKDEIKKKSS	147	167
P33322	H/ACA ribonucleoprotein complex subunit 4	**CBF5**	KEDSLIKEVETEKEEVKEDDSKKEKKEKKDKKE KKEKKEKKDKKEKKEKKEKKRKSEDGDSEEKK SKKSKK	413	483
Q06839	PX domain-containing protein YPR097W	**YPR097W**	KEEDNIDEDEYEEEGEGEESD	432	452
Q12532	Translation-associated element 2	**TAE2**	KEEEEKEEEQQQDEDDSNEV	744	763
P53935	Stress response protein NST1	**NST1**	KEEEKRKREEEKERLKKELEE	663	683
Q12072	ISWI one complex protein 2	**IOC2**	KEEENKKKDLESEWFEKLDE	382	401
P39520	Protein IFH1	**IFH1**	KEEKEKEKEEQQKQQESNKKE	218	238
P34222	Peptidyl-tRNA hydrolase 2	**PTH2**	KEGKLHNDTDEEESESEDESDEDEDIESTS	51	80
P32909	Protein SMY2	**SMY2**	KEKEALEAEKQKKSEKTKKD	419	438
P09620	Pheromone-processing carboxypeptidase KEX1	**KEX1**	KEKEGNKDKDGDDDNDNDDDDEDDH	538	562
P25386	Intracellular protein transport protein USO1	**USO1**	KEKEVSELEDKLKASEDKNSK	1270	1290
P38079	Protein YRO2	**YRO2**	KEKLKLKKKHKKDKKKAKKAKK	304	325
P47006	DNA-directed RNA polymerase I subunit RPA34	**RPA34**	KEKREKREKKDKKDKKKKHRD	213	233
P40019	Histone H2A.Z-specific chaperone CHZ1	**CHZ1**	KEKRELESQKESSHNKSEKS	6	25
Q03703	Uncharacterized protein YML037C	**YML037C**	KELPKEKKKSDGDKTKKNKSK	193	213
P25386	Intracellular protein transport protein USO1	**USO1**	KENGGSNKSDKESDSDKDTD	460	479
P47006	DNA-directed RNA polymerase I subunit RPA34	**RPA34**	KENKKEPKKRSHHDDEEESSEKKKK KKEKREKREKKDKKDKKKKHRD	187	233
P39744	Nucleolar complex protein 2	**NOC2**	KENKKLKKKTTKEQSDEDSSSSEEEEDMG	83	111
Q07915	Ribosome biogenesis protein RLP24	**RLP24**	KEQERAESVSEQEESEEEEED	147	167
P31374	Serine/threonine-protein kinase PSK1	**PSK1**	KESELSLMSSASSSASSSKKSSSR	893	916
P36775	Lon protease homolog, mitochondrial	**PIM1**	KESVSSEEKAENNAKSSSEK	856	875
P06104	Ubiquitin-conjugating enzyme E2 2	**RAD6**	KETVEKSWEDDMDDMDDDDDDDDDDDDEAD	142	172

P15108	ATP-dependent molecular chaperone HSC82	**HSC82**	KEVEKEVPIPEEEKKDEEKKDEDDKKP KLEEVDEEEEEKKPKTKKVKEEV	211	260
P02829	ATP-dependent molecular chaperone HSP82	**HSP82**	KEVEKEVPIPEEEKKDEEKKDEEKKDED DKKPKLEEVDEEEEKKPKTKKVKEEVQEIEEL	211	270
P38280	Spore-specific protein YSW1	**YSW1**	KGLKKEYEKSFEEYSDDSDDE	198	218
P38041	Protein BOB1	**BOI1**	KHKSKHKHKNSSSKDGSSEEKSKKKL	693	718
P50089	Uncharacterized protein YGR237C	**G8581**	KHKVVDSDDEESDSDESVTS	569	588
P36049	rRNA-processing protein EBP2	**EBP2**	KHMSGDEDESGDDREEEEEEEEEEG	101	126
P39929	Vacuolar-sorting protein SNF7	**SNF7**	KIKQSENSVKDGEEEEDEEDEDEKA	206	230
P35187	ATP-dependent helicase SGS1	**SGS1**	KIKREKSSVSQKDEEDDFDDD	619	639
Q00723	Pre-mRNA-splicing factor 38	**PRP38**	KIYNSDEESSSSESESNGDSEDDND	213	237
P41920	Ran-specific GTPase-activating protein 1	**YRB1**	KKAEKPETKKDEEDTKEETKKE	32	53
P50105	Transcription initiation factor TFIID subunit 4	**TAF4**	KKDDKKKNKGTSKADSKDGK	107	126
Q06205	FK506-binding protein 4	**FPR4**	KKDSKRKHEEDEEESAKPAEKK	238	259
P31376	SWR1-complex protein 3	**SWC3**	KKEEEDVKEKGVKSEDTQKKED	398	419
P54858	Protein DOS2	**DOS2**	KKEKETEEKEVEWDDEEEEEDDDKVEAVAD	240	269
Q12499	Nucleolar protein 58	**NOP58**	KKEKKEKKEKKEKKEKKEKKEKKSKKEKKEKK	480	511
P12754	Translation initiation factor eIF-2B subunit delta	**GCD2**	KKGSEGESKDASNEEDSNSK	579	598
P20447	ATP-dependent RNA helicase DBP3	**DBP3**	KKHKKEKKGEKEVEVPEKESEKKPE	43	67
P53115	Putative DNA helicase INO80	**INO80**	KKKIEKEAMEQAKKEEEEKESK	554	575
P53743	Pre-rRNA-processing protein ESF2	**ESF2**	KKKISSKDDIFSKKVEDIESE	28	48
P40187	GSY2-interacting protein PIG2	**PIG2**	KKKKFKLSDDDSDIENDNDSDD	285	306
Q08951	AP-3 complex subunit delta	**APL5**	KKKKKGKKSKSKNKLKTKAK	897	916
P36133	Putative uncharacterized protein YKR040C	**YKR040C**	KKKKKKKKEKSLHTEREKKKKKF	146	167
P12945	N-terminal acetyltransferase A complex subunit NAT1	**NAT1**	KKKLDDLASLKKEVESDKSE	515	534
O94742	26S proteasome complex subunit SEM1	**SEM1**	KKKNEEINKKSLEEDDEFED	18	37
P33399	La protein homolog	**LHP1**	KKKNGKEESKEDSSAIADDDEE	251	272
P35724	Manganese resistance protein MNR2	**MNR2**	KKKRGSDDSSNKNSKSTSSDS	215	235
P53863	J protein JJJ1	**JJJ1**	KKKRKKKKKAKVDTETEESES	458	478
Q08965	Ribosome biogenesis protein BMS1	**BMS1**	KKLAKMEEEKSQRDKEKKKE	1140	1159
Q08438	Phosphopantothenoylcysteine decarboxylase subunit VHS3	**VHS3**	KKNDTGGKDEDNDDDDDDDDDDDDDD DDDDDDDDDDDDDDDDDDDDDDDDD DDDDEDDEDEDEDDEGKKKEDKGGLQRS	594	674
Q04373	Pumilio homology domain family member 6	**PUF6**	KKPRISIDSSDEESELSKKED	26	46
P38845	Cruciform DNA-recognizing protein 1	**CRP1**	KKQEKGSKEVKRSETSKEKK	401	420
P36008	Elongation factor 1-gamma 2	**TEF4**	KKQKAEKPKAEKSKAEKKKDE	218	238
Q03707	Inner nuclear membrane protein SRC1	**SRC1**	KKRKDPDSDDWSESNSKENK	174	193
P38798	Nonsense-mediated mRNA decay protein 2	**NMD2**	KKSEEPSSSKETYEELSKPKK	1015	1035
P38198	Protein STU1	**STU1**	KKSEGDEESDDAVDENDVKK	1010	1029
P42846	Protein KRI1	**KRI1**	KKSKKEEKLQKKKEKRKVNE	423	442
P41818	Protein GLC8	**GLC8**	KKSREALKDEDEDEDDSTTKE	208	228
P32565	26S proteasome regulatory subunit RPN2	**RPN2**	KKTKKEKGPNEEEKKKEHEEKEKERE	815	840
Q02203	Uncharacterized protein YKR005C	**YK104**	KKVSSENCSKDTDDKSGSKKE	241	261
P15108	ATP-dependent molecular chaperone HSC82	**HSC82**	KLEEVDEEEEEKKPKTKKVKEE	238	259
P02829	ATP-dependent molecular chaperone HSP82	**HSP82**	KLEEVDEEEEKKPKTKKVKEE	243	263
P47036	Putative uncharacterized protein APQ13	**APQ13**	KLFPSSDESPESSESDSDLESE	53	74
Q02875	SMY2 homolog 2	**SYH1**	KLKKEKKLKQKQKKEEEKLKKKKKEEGKLEKEKQKEL	444	480
P36044	Protein MNN4	**MNN4**	KLLEERKRREKKKKEEEEKKKKEEEEKKKK EEEEKKKKEEEEKKKKEEEEKKKKEEEEKK KQEEEEKKKKEEEEKKKQEEGEKMKNED EENKKNEDEEKKKNEEEEKKKQEEKNKK NEDEEKKKQEEEEKKKNEEEEKKKQEEGHSN	1032	1178
Q04660	Ribosome biogenesis protein ERB1	**ERB1**	KMSKKRAASEESDVEEDEDK	12	31
P40517	Ran-specific GTPase-activating protein 2	**YRB2**	KNDEEKKEEGKKDQEPSHKK	56	75
P38241	Pre-mRNA-splicing factor SLT11	**ECM2**	KNDKKKTSKKVHKDRSKKSK	337	356

P36044	Protein MNN4	**MNN4**	KNEDEENKKNEDEEKKKNEEEEKKKQEEK NKKNEDEEKKKQEEEEKKKNEEEEKKKQEEGHSN	1116	1178
P35177	Transcriptional activator SPT7	**SPT7**	KNEEQDMVEESSKTEDSSKD	655	674
P27692	Transcription elongation factor SPT5	**SPT5**	KNSDGDTKDEGDNKDEDDDEDDDDDDDD EDDDDEAPTKRR	134	173
Q00816	Resistance to glucose repression protein 1	**REG1**	KNTSDDDTSSQSSSSSHSDDEE	535	556
Q04213	ISWI one complex protein 4	**IOC4**	KPDEEEYVEEEEEENEPEKK	178	197
P46947	Pre-mRNA-splicing factor CWC26	**BUD13**	KPKNKTKKKKKESKSDANSDK	15	35
P47068	Myosin tail region-interacting protein MTI1	**BBC1**	KPSEEPKKKQHKEKEEEEPE	449	468
P31376	SWR1-complex protein 3	**SWC3**	KRGRDDDDDDDDEESDDAYDE	30	50
P25357	Probable DNA-binding protein SNT1	**SNT1**	KRKERSNDEEVEVDESKEES	742	761
P36534	54S ribosomal protein L40, mitochondrial	**MRPL40**	KSDLKEVDEKTSFKSEKEWK	43	62
Q05672	RNA-binding suppressor of PAS kinase protein 1	**RBS1**	KSEDDVKSTNSSGSSDSDNE	197	216
P53286	Protein BTN2	**BTN2**	KSEQEKAAKEDEERQKKEKE	282	301
P47035	Nucleolar protein NET1	**NET1**	KSESDSDDSGDSSDDGKSFISAKSAS	1141	1166
P53883	Nucleolar protein 13	**NOP13**	KSIEEYKEDAEKKKSGASEKD	75	95
P12683	3-hydroxy-3-methylglutaryl-coenzyme A reductase 1	**HMG1**	KSLSSAQSSSSGPSSSSEEDDS	567	588
P38904	Protein SPP41	**SPP41**	KSPASESTSKKKKKKKTVKES	392	412
P38845	Cruciform DNA-recognizing protein 1	**CRP1**	KSPAVSEEKEKKKKQEKGSKE	389	409
Q12495	Chromatin assembly factor 1 subunit p90	**RLF2**	KSRSSSPCSKRELSSSKKEE	120	139
Q04049	DNA polymerase eta	**RAD30**	KSSAGKEDEEKTTSSKADEK	527	546
P38764	26S proteasome regulatory subunit RPN1	**RPN1**	KSSDKDATTDGKNDDEEEEKE	709	729
P53148	Spindle pole body component SPC105	**SPC105**	KSSDYSVERSNDKSDLSKSE	253	272
P36775	Lon protease homolog, mitochondrial	**PIM1**	KSSSEKTKDNNSEKTSDDIE	870	889
P32623	Probable glycosidase CRH2	**UTR2**	KSSSSTATSSSKTSSDHSSS	358	377
P39520	Protein IFH1	**IFH1**	KSVKLSPKKENEEEQKEEKEKEKEEQ	203	228
P38144	ISWI chromatin-remodeling complex ATPase ISW1	**ISW1**	KTEHEEDAELLKEEDSDDDES	160	180
Q04225	Ribosome assembly protein RRB1	**RRB1**	KTLLKDDNEGEDDEEDDEDD	177	196
Q12532	Translation-associated element 2	**TAE2**	KTSGNEDNGDDDEEEEEEEEEEEEEEE EEEEEEEEKEEEEKEEEQQQDEDDSNEV NGLEKGGDSNDSTK	708	777
P04786	DNA topoisomerase 1	**TOP1**	KTTKKEEQENEKKKREEEEEEDKKAKEEEEEYKW	100	133
P32499	Nucleoporin NUP2	**NUP2**	KVEDVQKSSDSSSEDEVKVE	153	172
Q05515	Survival factor 1	**SVF1**	KVKKEVIPESDEEESSADEDD	196	216
Q03525	Protein TMA23	**TMA23**	KVKRKKLKKDKKTSNDSESKKKKKKKSKK ESKKGKKSKHSSDEGDKSKHKKSKKSKK HKKEESSARRDRKEHI	139	211
P20447	ATP-dependent RNA helicase DBP3	**DBP3**	KVVDEEVIEKKKSKKHKKDKKDKKEKK DKKHKKHKKEKKGEKEVEVPEKESEKKPEPTSA	12	71
P36046	Mitochondrial intermembrane space import and assembly protein 40	**MIA40**	LAEEDNKSSEDKDTDESKVS	144	163
Q03707	Inner nuclear membrane protein SRC1	**SRC1**	LDDVSNDDDDDDDDDDDNDKKDDPL	111	135
P54858	Protein DOS2	**DOS2**	LDKESKRKEILSKKEKETEEKE	228	249
Q04195	E3 SUMO-protein ligase SIZ1	**SIZ1**	LEDDDDSDSDSNDGSRSPEK	444	463
P25386	Intracellular protein transport protein USO1	**USO1**	LEEESSKEKAELEKSKEMMKK	1489	1509
P32892	ATP-dependent RNA helicase DRS1	**DRS1**	LEKGGKDDEIDEEDDSEEAK	193	212
O13570	Putative uncharacterized protein YPR142C	**P9659.16B**	LGKSSSSSSSSSSSSASLCS	39	58
P25441	DNA-directed RNA polymerase III subunit RPC4	**RPC53**	LIEDEDDGESEKSSDVDMDDEE	169	190
P38903	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit delta isoform	**RTS1**	LIKKTTGSSSSSSSKKKDKEKEKEKSSTTSSTSKKP	9	44
P09064	Eukaryotic translation initiation factor 2 subunit beta	**SUI3**	LKKKKKKSKSVSADAEAEKE	48	67

P42073	RNA end formation protein 2	**REF2**	LKKKLASSKAKIKDSEKEKEKEKDKSKVKMKTK	196	228
Q04500	U3 small nucleolar RNA-associated protein 14	**UTP14**	LKLQDTDISSESSSSEESESESEDDEEEEDPFDEISEDEEDIELN	154	198
P53911	Chromatin modification-related protein EAF7	**EAF7**	LKRSKEVKFEDEEKEEIEED	247	266
P20459	Eukaryotic translation initiation factor 2 subunit alpha	**SUI2**	LLESKELDNRSDSEDDEDESDDE	282	304
P36133	Putative uncharacterized protein YKR040C	**YKR040C**	LLKRVSLKSSEEKKKKKKKKKEKSL	132	156
Q06680	Condensin complex subunit 3	**YCG1**	LPDEKSDAMSIDEEDKDSES	1011	1030
P21339	Morphogenesis-related protein MSB1	**MSB1**	LRKKSGSNSSSSSFEEKSKD	580	599
P53269	Putative uncharacterized protein YGR115C	**G6166**	LRKPDSSSSSSSEKKSSNLD	82	101
P53115	Putative DNA helicase INO80	**INO80**	LSDIINEKDADEDEEDDEDE	132	151
P53191	Phosphatidylinositol 3-phosphate-binding protein 2	**PIB2**	LSKENKAKESSSSSTSSVSSSS	229	250
P46949	Protein FYV8	**FYV8**	LTSDEEEEEKGMSSDSDEGS	439	458
Q08492	Bud site selection protein 21	**BUD21**	LVISKKDKEVHSSSDEESDDDD	28	49
P53269	Putative uncharacterized protein YGR115C	**G6166**	LVSSSSSENSSSSMKSSNSSK	36	56
P53885	Signal transduction protein MDG1	**MDG1**	LYELSAEDLEKEEEEEDEDK	229	248
P27466	Calcium/calmodulin-dependent protein kinase I	**CMK1**	MDDKVSEKESSPKQTEEDSE	1	20
P53313	Protein SDA1	**SDA1**	MEDSDDEKDNAKGKESDSDLE	588	608
Q06511	Ribosomal RNA-processing protein 15	**RRP15**	MEGDEAEEDEQNSSSDESSK	43	62
P06774	Transcriptional activator HAP2	**HAP2**	MKSKKSGASDDPDDSHEDKK	223	242
Q12094	Ribosome biogenesis protein TSR3	**TSR3**	MKSREAQSEESEDEESGSKE	279	298
P10964	DNA-directed RNA polymerase I subunit RPA190	**RPA190**	MREAEKSSDEEGIDSDKESDSDSEDEDVDM	1395	1424
P39985	DNA polymerase V	**POL5**	NDASEDESKTGSESESESESDSDDADEKDEEDEANEDILN	715	754
P34250	Eisosome protein SEG2	**SEG2**	NDDDDDKDDDVNDDDKDENVDDDEN	605	629
P32361	Serine/threonine-protein kinase/ endoribonuclease IRE1	**IRE1**	NDGSLKSEKDNDDADEDDEKS	615	635
P11075	Protein transport protein SEC7	**SEC7**	NDKRHDDEGEDGDEDEDEDEDEDEDN	84	109
P43611	Outer spore wall protein 7	**OSW7**	NEDEDEDEDEDDVGIEDENKE	183	203
Q04632	Conserved oligomeric Golgi complex subunit 8	**COG8**	NEEKEIEEKEKETKKEEYQK	375	394
P23255	Transcription initiation factor TFIID subunit 2	**TAF2**	NGEESEKEKEDTPEHDEEEE	257	276
P40498	U3 small nucleolar RNA-associated protein 25	**UTP25**	NHEDDNDDDSSGDEKDIDSEDE	137	158
P53172	Protein SDS23	**SDS23**	NKSDTESLHKSISKSSSSSS	40	59
Q12176	Ribosome biogenesis protein MAK21	**MAK21**	NLSSDQEPESESAEKEKKEEKD	175	196
Q08281	Restriction of telomere capping protein 1	**RTC1**	NNDDGDDDDDDDDDDKIIES	751	770
P32892	ATP-dependent RNA helicase DRS1	**DRS1**	NNGDENQSEEEEEEEEKEEEEEEEEEQEEM	162	191
P09620	Pheromone-processing carboxypeptidase KEX1	**KEX1**	NSEGDDDDDDDDDEDDNNEKQS	563	584
P53234	Uncharacterized protein YGR053C	**YGR053C**	NSKEDEKKKNDGDGKRSKKK	218	237
P08964	Myosin-1	**MYO1**	NSKEEEIKSFNDKLSSSEED	1025	1044
Q03661	Silent chromatin protein ESC1	**ESC1**	PDAEDESSSNIDSDKDEDLE	175	194
Q12276	HMG2-induced ER-remodeling protein 1	**HER1**	PDASKISDSTDEPDKSKDEKKE	817	838
Q08438	Phosphopantothenoylcysteine decarboxylase subunit VHS3	**VHS3**	PDVSAGKEEEEDEDNDEEDD	573	592
Q12263	Serine/threonine-protein kinase GIN4	**GIN4**	PESGSSSHTEKEEENEEKEEKK	945	966
Q03690	Clustered mitochondria protein 1	**CLU1**	PIKERSKQEEKDEKSDPEEKK	152	172
P38788	Ribosome-associated complex subunit SSZ1	**SSZ1**	PKEENAEEDDESEWSDDEPE	466	485
P32794	ATPase family gene 2 protein	**AFG2**	PKSSSSGSKKKSSASSNSAD	3	22
P38764	26S proteasome regulatory subunit RPN1	**RPN1**	PNKKDKKKEEEEQLSEEDAK	25	44
P34250	Eisosome protein SEG2	**SEG2**	PNNDKDDDDDKDNDNDDDDDKDDDV	591	615
P15705	Heat shock protein STI1	**STI1**	PQKEESKESEPMEVDEDDSK	240	259
Q08965	Ribosome biogenesis protein BMS1	**BMS1**	PSEQAEDNSDKESEDEDENED	670	690
Q04779	Transcriptional regulatory protein RCO1	**RCO1**	PSHSNSSSPSSSSLSSSSSKEKKR	12	35
P38201	Uncharacterized protein YBL029W	**YBL0422**	PSSFSSSSSSSESTVSSSRK	227	246
P25361	Uncharacterized protein YCR043C	**YCR43C**	PSSKKSKFDWSKKKDDFKME	97	116
P38250	Increased sodium tolerance protein 2	**IST2**	PSSLSSASSPSLSSSSSSK	671	690

P12683	3-hydroxy-3-methylglutaryl-coenzyme A reductase 1	**HMG1**	PSSSSEEDDSRDIESLDKKI	579	598
Q06523	Uncharacterized protein YPR148C	**YPR148C**	PSTEDEKPKSEGAEEESKKE	319	338
P39113	Regulatory protein CAT8	**CAT8**	PSTSTKLKEKDEDKKLDDDS	328	347
P38222	Ribosomal lysine N-methyltransferase 3	**RKM3**	QAKSDNLSDDIESEEEEEEEE	407	427
P23615	Transcription elongation factor SPT6	**SPT6**	QDSTSKLEDFFSEDEEEEES	144	163
P46947	Pre-mRNA-splicing factor CWC26	**BUD13**	QEDSKIDDSSFSRSKYEDEK	132	151
P53927	Ribosome biogenesis protein 15	**NOP15**	QEKEELALETSSSSSDEEDEKDEDEIEG	25	52
P53935	Stress response protein NST1	**NST1**	QELEEEKRKKREKEEKKQKK	629	648
P53125	Imitation switch two complex protein 1	**ITC1**	QKESDAKNETNSESDSKSDSDSEERD	631	656
P38996	Nuclear polyadenylated RNA-binding protein 3	**NAB3**	QKGGNDDDDDDNEEEEEEEEDDDDDDDDDDDEEEEEEEEEEGNDNSSVGSDSAAEDGEDEEDKKDKTKDKEVELRRETLEKEQKD	96	181
Q06344	Pre-rRNA-processing protein ESF1	**ESF1**	QKKNKNKTSKKKKTDDSDSD	207	226
Q01389	Serine/threonine-protein kinase BCK1/SLK1/SSP31	**BCK1**	QKSASKLKSSRSSSESIKSK	225	244
Q08193	1,3-beta-glucanosyltransferase GAS5	**GAS5**	QQSCDAKDDDDEEDDDTSSSSSSSSSSSSASSSSESSSSTSKASSSSPSASETSLLKSAAS	388	449
Q12132	Nutrient and stress factor 1	**USV1**	QRKPSSSSLSSSSSASSSSS	107	126
P36026	Ubiquitin carboxyl-terminal hydrolase 11	**UBP11**	RDSSQDSSSSLSKVEKPKEEE	246	266
P38266	Altered inheritance of mitochondria protein 3	**AIM3**	RERKSGKKNSSDEEYDSEDE	48	67
P28496	Sphingosine N-acyltransferase LAC1	**LAC1**	RGILKDDRSDSESDEESDESS	385	405
P46682	AP-3 complex subunit beta	**APL6**	RKIIKVVEESSDEDEDESEESSDDDEYSDSSLG	764	796
Q05024	Protein TRI1	**TRI1**	RKRKGKSESKSKRKKKKNDS	94	113
P07866	Guanine nucleotide exchange factor LTE1	**LTE1**	RKSSKEEEEEFLNSDSSKFD	423	442
P09232	Cerevisin	**PRB1**	RKVEFTKDDDEEPSDSEDKE	45	64
P36049	rRNA-processing protein EBP2	**EBP2**	RLDLEKLAKSDSESEDDSESE	127	147
Q12017	Phosducin-like protein 2	**PLP2**	RLEDKDLSDLEELEDDEDED	55	74
P40362	U3 small nucleolar RNA-associated protein 18	**UTP18**	RPKWVDDESDSELDDEEDDEEE	174	195
P32900	Protein SKG6	**SKG6**	RRDDSDSSSSSASSTKNSKS	31	50
Q08723	26S proteasome regulatory subunit RPN8	**RPN8**	RVKDKQSKVSDDSESESGDKE	305	325
P54784	Origin recognition complex subunit 1	**ORC1**	SADELEEEEDEEEDEDEEEKEARHT	278	302
Q12208	U6 snRNA phosphodiesterase	**USB1**	SADYSSSDGSDTESESSNKSE	5	25
P17260	Protein KRE1	**KRE1**	SASSSVSSSVSSSGSSSSVK	152	171
P43587	Type 1 phosphatases regulator YPI1	**YPI1**	SDDDGSSSSGSSSSESENEKDLDFNE	88	113
Q12348	COP9 signalosome complex subunit 10	**RRI2**	SDDEGMESIEMEEETDDEDK	14	33
Q05881	Uncharacterized protein YLR287C	**YLR287C**	SDDIFLDFEDSESESDSQKEE	209	229
P40362	U3 small nucleolar RNA-associated protein 18	**UTP18**	SDDYSEDSEEAAWIDSDDEK	113	132
P47069	Spindle pole body assembly component MPS3	**MPS3**	SDEDYTDEADKSFIEDSDSDS	107	127
P32497	Eukaryotic translation initiation factor 3 subunit C	**NIP1**	SDEESDEEDGKKVVKSAKEKL	99	119
P38931	Mediator of RNA polymerase II transcription subunit 13	**SSN2**	SDKSDSMVDKELFGEDEDED	472	491
P53863	J protein JJJ1	**JJJ1**	SDLEKFDSADESVKEKEDID	386	405
P25386	Intracellular protein transport protein USO1	**USO1**	SDLEQTKEEIISKSDSSKDE	1032	1051
P39520	Protein IFH1	**IFH1**	SDSDSDSETSSDDENIDFVK	150	169
P38741	Meiotic activator RIM4	**RIM4**	SDSEEKLTSDGIYDDEDKDSE	320	340
Q01448	Histone promoter control protein 2	**HPC2**	SDSNISSSKKPTSEKAKKSSS	312	332
P40498	U3 small nucleolar RNA-associated protein 25	**UTP25**	SDVDSGDDFDIEDEEGKKEK	59	78
Q04217	Probable ATP-dependent RNA helicase DHR1	**ECM16**	SEDDDEEEDQEEEDKMHPSEN	272	292
P36049	rRNA-processing protein EBP2	**EBP2**	SEDDSESENDSEEDEDVVAKEESEEKEEQEEEQDVPLSDVE	140	180
Q12176	Ribosome biogenesis protein MAK21	**MAK21**	SEEEKEEEENKEVSAKRAKKK	978	998
P25582	27S pre-rRNA (guanosine(2922)-2'-O)-methyltransferase	**SPB1**	SEEFDSDYDSEEEKNQTKKEK	630	650
P36531	54S ribosomal protein L36, mitochondrial	**MRPL36**	SEENKKKTQIKKEEKEDVSEK	126	146
P38925	Manganese transporter SMF1	**SMF1**	SEEQDKVKKSKSTEEIMEEK	304	323

P47018	Maintenance of telomere capping protein 1	**MTC1**	SEERNKSYDQKKQKESEDEDEDDEI	421	445
P47101	UPF0508 protein YJR030C	**J1575L**	SEIDDSESIISEDGKDKRDKD	464	484
P39969	Protein BOI2	**BOI2**	SEKNSSPIVDKKSSKKSRSK	678	697
P41895	Transcription initiation factor IIF subunit alpha	**TFG1**	SESDIENKENESPVKKEEDSD	560	580
P10962	Protein MAK16	**MAK16**	SESESDSESESDSDSDEENKNSAKRRKKGTSAK	249	281
P40552	Cell wall protein TIR3	**TIR3**	SESSSAASSSASEAAKSSSS	176	195
P53616	Probable secreted beta-glucosidase SUN4	**SUN4**	SESSVTKSSSKVASSSESTE	84	103
P47074	Spindle assembly checkpoint component MAD3	**MAD3**	SFSQSKNSEIISDDDKSSSS	488	507
Q12373	HAT1-interacting factor 1	**HIF1**	SGDSEDSGEGSEEEEENVEKEEER	141	164
P32567	Phosphatidic acid phosphohydrolase 1	**PAH1**	SGSDTEDETSFSKEQSSKSEK	311	331
Q12196	Serine/threonine-protein kinase RIO1	**RIO1**	SGSEEDDEEEGEYYDDDEPK	417	436
P53246	Late endosome and vacuole interface protein 11	**ENV11**	SHEEEDEEEDDTEDIELIEKD	33	53
Q03177	WD repeat-containing protein YMR102C	**YM6228.01C**	SHQESFDGESSSGESKSKSK	33	52
P43579	Ino eighty subunit 1	**IES1**	SIGGDKSEEDGEGEDDKSEE	487	506
P53094	Negative regulator of sporulation MDS3	**MDS3**	SISSSSTNSISSSDLEEKEE	1261	1280
P25386	Intracellular protein transport protein USO1	**USO1**	SKAKEKSESELSRLKKTSSEE	1323	1343
Q07800	Phosphatase PSR1	**PSR1**	SKDDLYEEKYEVDEDEEIDDED	110	131
P53904	Tubulin-specific chaperone B	**ALF1**	SKDEESAEVEDVHDDVESDDE	233	253
P43560	Uncharacterized protein YFL042C	**YFL042C**	SKEEEEQDDKEVKNKLKEVD	590	609
P35177	Transcriptional activator SPT7	**SPT7**	SKFAEDEDYDDEDENYDEDS	230	249
P11745	Ran GTPase-activating protein 1	**RNA1**	SKFDDLEVDDFEEVDSEDEE	345	364
P32944	Mitosis inhibitor protein kinase SWE1	**SWE1**	SKFLGKIEEEEEEEEEGKDEES	81	102
Q04500	U3 small nucleolar RNA-associated protein 14	**UTP14**	SKGDTNSDEDDIDSESFEDEE	56	76
P38904	Protein SPP41	**SPP41**	SKKKKKDKSKNRESSKDKSSKKSKSSS HSKKHAKDRNKEKQSKPTN	215	260
P21372	Pre-mRNA-processing ATP-dependent RNA helicase PRP5	**PRP5**	SKKKSKRSKVKKKISFDFSDDDDSEI	90	115
Q03210	Probable RNA exonuclease NGL3	**NGL3**	SKKRNGEESDQDDEECDEKS	331	350
P25386	Intracellular protein transport protein USO1	**USO1**	SKLKDLGVEISSDEEDDEEDDEEDEEEGQ	1760	1788
P25617	Uncharacterized protein YCR016W	**YCR16W**	SKPGTKKKEKLSKDEKNSKK	74	93
P32657	Chromo domain-containing protein 1	**CHD1**	SKPKSRTKSKSKSKPKSQSEK	97	117
P27476	Nuclear localization sequence-binding protein	**NSR1**	SKQAKEEKAKAVSSSSSESSSSSSSSESESES ESESESSSSSSSSDSESSSSSSDSESEAETKKE ESKDSSSSSSDSSSDEEEEEEKEETKKEESKE SSSSDSSSSSSSDSESEKEESNDKKRKSEDAE EEEDEESSNKKQKNEETEEPATIFVGRLSW SIDDEWLKKEFEHIGGVIGARVIYERGTDR SRGYG	17	212
P25386	Intracellular protein transport protein USO1	**USO1**	SKSDSSKDEYESQISLLKEK	1043	1062
P12753	DNA repair protein RAD50	**RAD50**	SKSKLDELEVDSTKLKDEKE	767	786
P47035	Nucleolar protein NET1	**NET1**	SKSNEKSQTKAPSSSDDESSSDSDSNSS SDSVSDSSSDSKSESDSDDSGDSSD DGKSFISAKSASAALGKKKKPSG	1102	1177
P46997	J protein JJJ2	**JJJ2**	SKSSSPDEEKKNNKEPKRES	258	277
P22035	Myb-like DNA-binding protein BAS1	**BAS1**	SKTKKKEKRKSESSQHHSSSS	665	685
P32266	Dynamin-like GTPase MGM1, mitochondrial	**MGM1**	SLDDDESKRQGDPKDDDDEDDDDEDDENDSVDT	147	179
Q01448	Histone promoter control protein 2	**HPC2**	SLEKEKDEEGDVIELDDDED	513	532
O13547	Covalently-linked cell wall protein 14	**CCW14**	SLGDSSSSASSSASSSSKASSS	89	110
P09064	Eukaryotic translation initiation factor 2 subunit beta	**SUI3**	SLKKKKKKTKDSSVDAFEKE	80	99
Q04429	Protein HLR1	**HLR1**	SLSSSILDSEKQTDDEEEES	336	355
P46678	Transcription factor TFIIIB component B''	**BDP1**	SLTKEEQEEEEEKRKEERDK	326	345

P53335	Protein PXR1	**PXR1**	SNAKKRKREGDDSEDEDDDDKEDKD SDKKKHKKHKKHKKDKKKDKKDKKEHK KHKKEEKRLKKEKRAEKTKETKKTSKLK SSESASNIPD	147	236
Q05610	Donuts protein 1	**DON1**	SNEDSILKGKEKGKEEEKEK	220	239
P32497	Eukaryotic translation initiation factor 3 subunit C	**NIP1**	SNKFLKSSNYDSSDEESDEED	87	107
Q06648	GTPase-interacting component 2	**GIC2**	SNSNSAKSKESSSSSSSANK	86	105
P38883	Pre-rRNA-processing protein RIX1	**RIX1**	SQTKEEEKPMEDSEDEEQEE	725	744
Q12052	Trimethylguanosine synthase	**TGS1**	SRRESSEKEELSSENEELSK	274	293
P39712	Flocculation protein FLO9	**FLO9**	SSASSSSSISSESPKSTYSSSS	970	991
P50091	Peroxisomal membrane protein PEX21	**PEX21**	SSASSSSTLSSSVESKLSESK	215	235
Q03525	Protein TMA23	**TMA23**	SSDEGDKSKHKKSKKSKKHKKEESSARRDRKEH	178	210
P32623	Probable glycosidase CRH2	**UTR2**	SSDHSSSTKKSSKTSSTASSSSSSSSSSSSSSSTATKN	371	408
P38190	Putative uncharacterized protein YBL053W	**YBL0514**	SSDICSSSMSSSRVKSSSSSSSSL	70	93
P53049	Oligomycin resistance ATP-dependent permease YOR1	**YOR1**	SSDISTDVDKDTSSSWDDKS	30	49
Q04660	Ribosome biogenesis protein ERB1	**ERB1**	SSDKEAQDDSDDDSDAELNK	63	82
P47030	Protein TAX4	**TAX4**	SSDSDTDEDSNVMEIKDKKKKS	349	370
P39924	Hexose transporter HXT13	**HXT13**	SSEESKTEKKDWKKFLKFSK	542	561
O13563	26S proteasome regulatory subunit RPN13	**RPN13**	SSESDEEESNDEKQKAQDVD	132	151
P47160	Epsin-3	**ENT3**	SSESSKPSKELIQEDEKKADEEEDDDDEFSEFQS	244	277
P43597	Uncharacterized protein YFR016C	**YFR016C**	SSESVKKDTEEDAEVENSEK	806	825
Q12421	Autophagy-related protein 31	**ATG31**	SSHGDDKSNDEEEELSVDSD	128	147
P53120	Uncharacterized membrane protein YGL140C	**YGL140C**	SSMSEEKYKKIEKISKSAKE	822	841
P53908	Uncharacterized membrane protein YNL143C	**N1206**	SSNSSKKRPKDSSLLSEKKKKKKKKKKDVLSY	32	63
P32378	4-hydroxybenzoate polyprenyltransferase, mitochondrial	**COQ2**	SSRKRYTSSSSSSSPSSKES	29	49
Q06511	Ribosomal RNA-processing protein 15	**RRP15**	SSSDESSKIIDNEQSDAEEDDDEE EEDDDFPRKKKSKNSKHDDGS	55	99
Q12161	RING finger protein PSH1	**PSH1**	SSSDSEDRSLSYSGSSDVKD	357	376
P10863	Cold shock-induced protein TIR1	**TIR1**	SSSEAKSSSAAPSSSEAKSSS	144	164
P10863	Cold shock-induced protein TIR1	**TIR1**	SSSEAKSSSAAPSSSEAKSSS	156	176
P10863	Cold shock-induced protein TIR1	**TIR1**	SSSEAKSSSAAPSSSEAKSSS	168	188
P38931	Mediator of RNA polymerase II transcription subunit 13	**SSN2**	SSSEEEEDEEENGSSDEDLKSLNV	654	677
Q01389	Serine/threonine-protein kinase BCK1/SLK1/SSP31	**BCK1**	SSSESIKSKLKNSKSQEDIS	236	255
Q08193	1,3-beta-glucanosyltransferase GAS5	**GAS5**	SSSESSSSTSKASSSSPSASETSL	420	443
Q99248	Uncharacterized protein YOR019W	**OR26.09**	SSSGDVFESDDRNDKKKKKKKKKKSL	699	724
Q12138	Putative uncharacterized protein YLR125W	**L3101**	SSSGEKENSGLCSEESSEED	66	85
P53035	Regulatory protein MIG2	**MIG2**	SSSGSNEEEHESLDHESSKS	335	354
Q08446	Protein SGT1	**SGT1**	SSSKKKIDWSKLDIDEEADEE	315	335
P33322	H/ACA ribonucleoprotein complex subunit 4	**CBF5**	SSSQETKETEEEPKKAKEDS	397	416
Q03361	Uncharacterized protein JIP4	**JIP4**	SSSRSKSKSRSSSKSRIRDKSK	685	706
O13547	Covalently-linked cell wall protein 14	**CCW14**	SSSSAAPSSSKASSTESSSSSSSTK	136	161
P27476	Nuclear localization sequence-binding protein	**NSR1**	SSSSDSSSSSSSDSESEKEESNDKKRKSEDA EEEEDEESSNKKQKNEETEEPATIFVG	116	173
P10863	Cold shock-induced protein TIR1	**TIR1**	SSSSEAKSSSAAPSSSEAKSSS	143	164
P38827	Histone-lysine N-methyltransferase, H3 lysine-4 specific	**SET1**	SSSSEEEEEEAPDKKFKSESEPTT	682	705
P32768	Flocculation protein FLO1	**FLO1**	SSSSFISSESSKSPTYSSSS	1185	1204
P53269	Putative uncharacterized protein YGR115C	**G6166**	SSSSSEESKTSSPSSSSLGA	204	223
P53882	Topoisomerase I damage affected protein 7	**TDA7**	SSSSSKVPSNRPSSSSSDD	305	324
P38894	Flocculation protein FLO5	**FLO5**	SSSSSSISSESPKSPTNSSSS	733	753

P38903	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit delta isoform	**RTS1**	SSSSSSRSSSVSRSGSSSTKK	99	119
P32623	Probable glycosidase CRH2	**UTR2**	SSSSSSSSSSSSSTATKNGDK	391	411
Q05854	Uncharacterized transcriptional regulatory protein YLR278C	**L8003.10**	SSSSSSSSSSTLSFPSSQEK	876	895
P39009	DNA damage response protein kinase DUN1	**DUN1**	SSSSTDIENDDEKVSSESRS	139	158
Q03655	Probable 1,3-beta-glucanosyltransferase GAS3	**GAS3**	SSSTGSSSSTGSSSASSSSKSK	476	497
P39521	Pre-rRNA-processing protein FHL1	**FHL1**	SSTSSESESESDSGSEVDEKNNKNEK	848	873
P06786	DNA topoisomerase 2	**TOP2**	SSTSSSSIFDIKKEDKDEGE	1266	1285
P0CE41	Heme-responsive zinc finger transcription factor HAP1	**HAP1**	SSVEDEDVKVKQESSDELKKDD	1077	1098
P38996	Nuclear polyadenylated RNA-binding protein 3	**NAB3**	SSVGSDSAAEDGEDEEDKKDK	142	162
P53874	Ubiquitin carboxyl-terminal hydrolase 10	**UBP10**	STEPPSSSSEDDKVGKEEEEE	122	142
Q12465	Bud site selection protein RAX2	**RAX2**	STSSQPPSSSASSESKSKSKKKKI	1134	1157
Q12339	rRNA-processing protein UTP23	**UTP23**	SVKKKKKVNSPSDEVKDKED	206	225
Q06205	FK506-binding protein 4	**FPR4**	SVKSKKGKAVEQSESEDSEDE	92	112
Q01329	Pre-tRNA-processing protein PTA1	**PTA1**	SVPSSSSSKRKSDDDDDGND	489	508
Q04693	Pre-mRNA-splicing factor RSE1	**RSE1**	SVSLNNEEEEEEEEDDDDEKEEEEINSSG	784	812
Q05942	Ribosome assembly protein 3	**RSA3**	TADVSSSDSSSSDPSSESEKEEIQ	25	48
P53261	Pescadillo homolog	**NOP7**	TDETEKEEEQEKKQEKEQEKE	305	325
P53156	Uncharacterized protein YGL081W	**YGL081W**	TDTDTTEEKEEEEEKEEGDDEE	242	263
O13547	Covalently-linked cell wall protein 14	**CCW14**	TESSSSSSSSTKAPSSEESSST	150	171
P53269	Putative uncharacterized protein YGR115C	**G6166**	TFLASSSSSSISSSSSEESK	193	212
P38129	Transcription initiation factor TFIID subunit 5	**TAF5**	TGDDDKDKIKDKIAKDEEKKESELK	379	403
P32336	Protein NUD1	**NUD1**	TKEEDEEDTSNSFEFSSSSS	251	270
P25635	Periodic tryptophan protein 2	**PWP2**	TKRPSDDDDNESEDDDKQEE	221	240
P53313	Protein SDA1	**SDA1**	TMDSDKEYDVDMEDSDDEKD	577	596
P43597	Uncharacterized protein YFR016C	**YFR016C**	TNDISVEVEKEEEEEEEEEE	458	477
P38894	Flocculation protein FLO5	**FLO5**	TSSSTSGSSESKTSSASSSSSSSSISSESPKSP	715	747
P32773	Transcription initiation factor IIA large subunit	**TOA1**	TTVENIDDESEKKDDEEKEED	166	186
P38351	RNA polymerase II-associated protein 1	**PAF1**	TVDEEDDEDEEQPEDVKKESE	360	380
Q00416	Helicase SEN1	**SEN1**	VDEGEEADKAVKKKKKEKKKEKKKSKADDKKKNNKKAESPS	1913	1953
P48362	Protein HGH1	**HGH1**	VEEMPSKNAEEEEEEESEEEEDDDEEDEIVEV	362	393
Q03973	High mobility group protein 1	**HMO1**	VEKRPHDDDGSSEKKKKKKKKDKKKDKSNSSI	215	246
P41895	Transcription initiation factor IIF subunit alpha	**TFG1**	VKKEEDSDTLSKSKRSSPKK	573	592
P38934	Nuclear segregation protein BFR1	**BFR1**	VKKELEEKRLKEQEESEKDKE	449	469
P47083	U3 small nucleolar RNA-associated protein MPP10	**MPP10**	VKKHSDVKDPKEDEELDEEE	245	264
Q06632	Protein CFT1	**CFT1**	VNKNDDDEEEEDEDENIDDSE	147	167
P38827	Histone-lysine N-methyltransferase, H3 lysine-4 specific	**SET1**	VSKEHDEEDENMTSSSSEEEEEEA	669	692
P36152	Fe-S cluster assembly protein DRE2	**DRE2**	VVDDLIEDSDDDDFSSDSSK	198	217
Q08965	Ribosome biogenesis protein BMS1	**BMS1**	WRGEDDDSKDESDIEEDVDDD	567	587
P34756	1-phosphatidylinositol 3-phosphate 5-kinase FAB1	**FAB1**	YDSSRKDSEDKKSLHDEKAK	1500	1519
P48234	Ribosome biogenesis protein ENP2	**ENP2**	YDYEDEESDEEESDDETNQKS	543	563
P53107	Ran-specific GTPase-activating protein 30	**YRB30**	YEAIEEDEEEDEEEDEEEGKDGEERKEEEEEENKLEDKDSSITL	338	381
P53974	Actin-regulating kinase 1	**ARK1**	YEDSSSSSDESYSGDVDELKK	455	475
Q06631	Protein BFR2	**BFR2**	YEEVSENEDEEEEEEEEEEKEEDALSFR	91	118
P53064	RNA polymerase-associated protein RTF1	**RTF1**	YSDNEDEDDEEDYREEDYKDDE	180	201
P40048	Tyrosine-protein phosphatase 3	**PTP3**	YSKYKSMLSLESDSDSESDD	499	518

Table 2: PASK-like proteins in *Homo sapiens*
*Note – this table has been amended for ease of inclusion in this thesis. Only the first 6 columns are included. For the full table, see (Bentley-DeSousa et al., 2018).

ProteinAcc	ProteinName	GeneName	Hit	SeqStart	SeqStop
Q13442	28 kDa heat- and acid-stable phosphoprotein	**PDAP1**	AAGDPKKEKKSLDSDESEDEEDDY	47	70
Q12830	Nucleosome-remodeling factor subunit BPTF	**BPTF**	AASQKRKREEEKDSSSKSKKKK	2832	2853
Q8IYB5	Stromal membrane-associated protein 1	**SMAP1**	AAVDKNKLEKEKEKKKEEKK	157	176
Q8NFD5	AT-rich interactive domain-containing protein 1B	**ARID1B**	ADDSGKEEEDAECIDDDEEDEEDEEEDSEKTESDEKSSIALT	1712	1753
Q9P2R6	Arginine-glutamic acid dipeptide repeats protein	**RERE**	ADKDKDKDKEKDRDRDRDRE	3	22
Q8NFG4	Folliculin	**FLCN**	ADLEEESESWDNSEAEEEEK	290	309
Q01970	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-3	**PLCB3**	ADREDEEEDEEEEEQTDPKK	553	572
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	**ACIN1**	ADRNLKTEEEEEEEEEEEEDDEEEEGDDEG	263	292
Q13428	Treacle protein	**TCOF1**	ADSSSEDTSSSSDETDVEGK	196	215
Q9UKJ3	G patch domain-containing protein 8	**GPATCH8**	ADTEEKSSKAESGEKSKKRKK	693	713
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	ADTGSKSISEKGSEEDEEEK	2225	2244
Q5C9Z4	Nucleolar MIF4G domain-containing protein 1	**NOM1**	AEDEEKSENSSEDGDITDKS	311	330
A6NFI3	Zinc finger protein 316	**ZNF316**	AEDGSECDPDQEEEEEEEEK	18	37
Q70Z53	Protein FRA10AC1	**FRA10AC1**	AEEASKKKDKGHSSSKKSEDSL	253	274
Q76N89	E3 ubiquitin-protein ligase HECW1	**HECW1**	AEEEDGAEEESTLKDSSEKD	579	598
P07196	Neurofilament light polypeptide	**NEFL**	AEEEEKDKEEAEEEEAAEEEEA	476	497
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	AEEEKEEAEEALKEKEKPKEEE	379	400
Q9HC10	Otoferlin	**OTOF**	AEEEKEKKKKKKGTAEEPEEEE	1299	1320
Q96JQ2	Calmin	**CLMN**	AEEKEQKQESSKIPESSSDK	469	488
Q8N2M8	CLK4-associating serine/arginine rich protein	**CLASRP**	AEKPEEEESAAEEESNSDEDE	179	199
P46100	Transcriptional regulator ATRX	**ATRX**	AEKSTGKGDSCDSSEDKKSK	1064	1083
Q12872	Splicing factor, suppressor of white-apricot homolog	**SFSWAP**	AENKSDEKKKSGVSSDNEDDDDEEDG	269	294
Q8N1F8	Serine/threonine-protein kinase 11-interacting protein	**STK11IP**	AEPQEEEEEKEGKEEKEEGE	503	522
Q96SW2	Protein cereblon	**CRBN**	AESEEEDEMEVEDQDSKEAKK	23	43
Q9P0K7	Ankycorbin	**RAI14**	AESSSKLEEDKDKKINEMSKE	870	890
Q7Z401	C-myc promoter-binding protein	**DENND4A**	AETEQQQKEEEEEDEDDSKS	1175	1194
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog	**UTP18**	AETTKRKTSSDDESEEDEDD	197	216
Q8IZU1	Protein FAM9A	**FAM9A**	AEVIVVEDEEEEEKEEEEEKEEEEEGEEEGG	214	245
Q9NQ66	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1	**PLCB1**	AGEADTESDDDDDDDDCKKSS	504	524
Q9ULT8	E3 ubiquitin-protein ligase HECTD1	**HECTD1**	AGPSSDDENEEESKPEKEDE	628	647
O60885	Bromodomain-containing protein 4	**BRD4**	AGSSKMKGFSSSESESSSESSSSDSEDSETEMA	690	722
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	AKAPPKKAKSSDSDSDSSSEDE	157	178
Q8N9E0	Protein FAM133A	**FAM133A**	AKEKVKKKKKKQHKKHSKKKKKKSGSSHKS	218	247
Q5T6F2	Ubiquitin-associated protein 2	**UBAP2**	AKENSENKENREKKSEKESS	110	129
Q14BN4	Sarcolemmal membrane-associated protein	**SLMAP**	AKESDFSDTLSPSKEKSSDD	442	461
P54105	Methylosome subunit pICln	**CLNS1A**	AKFEEESKEPVADEEEEDSDDD	84	105
Q13428	Treacle protein	**TCOF1**	AKKASTKDSESPSQKKKKKKKK	1461	1482
Q14093	Cylicin-2	**CYLC2**	AKKDAKKDAKKNAKKDEKKD	323	342
Q13061	Triadin	**TRDN**	AKKDEKKEDSKKTKKPAEVE	371	390
B2RPK0	Putative high mobility group protein B1-like 1	**HMGB1P1**	AKKGVVKAEKSKKKKEEEEDEEDEEDEEEEDEEDEEDDDDE	171	211
P09429	High mobility group protein B1	**HMGB1**	AKKGVVKAEKSKKKKEEEEDEEDEEDEEEEEDEEDEDEEEDDDDE	171	215
Q5BKY9	Protein FAM133B	**FAM133B**	AKKKKSSEEREKATEKTKKKKK	206	227
P19338	Nucleolin	**NCL**	AKNGKNAKKEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMKA	134	177

Q99453	Paired mesoderm homeobox protein 2B	**PHOX2B**	AKNGSSGKKSDSSRDDESKE	167	186
Q13428	Treacle protein	**TCOF1**	AKRAKKPEEESESSEEGSESEEEA	260	283
O15347	High mobility group protein B3	**HMGB3**	AKVARKKVEEEDEEEEEEEEEEEEEDE	173	200
O15047	Histone-lysine N-methyltransferase SETD1A	**SETD1A**	ALDSEGEEASQESSSEKDEEDDEEDEEDEDREEAVDT	960	996
Q9NQ66	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1	**PLCB1**	ALEKSAKKDSKKKSEPSSPD	965	984
O15014	Zinc finger protein 609	**ZNF609**	ALTPAKDKKKKDKKKKESSKE	720	740
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	ANGKAASSSSSSSSSSSDDSEEEKA	210	235
P05386	60S acidic ribosomal protein P1	**RPLP1**	APAEEKKVEAKKEESEESDDD	87	107
Q9NW13	RNA-binding protein 28	**RBM28**	APAKSSDHSEEDSDLEESDS	276	295
Q86UX7	Fermitin family homolog 3	**FERMT3**	APEKKEKKKKEKEPEEELYD	144	163
Q96NK8	Neurogenic differentiation factor 6	**NEUROD6**	APGEETEKEEEEEDREEEDE	51	70
Q9UK55	Protein Z-dependent protease inhibitor	**SERPINA10**	APKEEEEDEQEASEEKASEEEK	44	65
Q13435	Splicing factor 3B subunit 2	**SF3B2**	APKKKGFEEEHKDSDDDSSDDE	418	439
Q15059	Bromodomain-containing protein 3	**BRD3**	APVNKPKKKKEKKEKEKKKKDKEKEKEKHKVKAEEEKKAKV	483	523
Q9NRE2	Teashirt homolog 2	**TSHZ2**	AQEEQLKEEEEIKEEEEEEDS	16	36
Q9BW71	HIRA-interacting protein 3	**HIRIP3**	AQKKEQREEEVEEEEKEEDEEK	233	254
P31273	Homeobox protein Hox-C8	**HOXC8**	ARDEEKVEEEGNEEEEKEEEEKEENKD	216	242
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	ASAKKGKAESSNSSSSDDSSEEEEEKLKGK	507	536
Q92609	TBC1 domain family member 5	**TBC1D5**	ASASSSNPSSSPDDDSSKDS	766	785
Q96ST2	Protein IWS1 homolog	**IWS1**	ASDSEEEAGKELSDKKNEEKD	437	457
Q96ST2	Protein IWS1 homolog	**IWS1**	ASDSESEELHRQKDSDSESEE	79	99
Q96T23	Remodeling and spacing factor 1	**RSF1**	ASEEEEEKESEEAILADDDE	873	892
Q6ZV73	FYVE, RhoGEF and PH domain-containing protein 6	**FGD6**	ASEELLEKSSYPSSEEKSSEKS	514	535
Q9NWH9	SAFB-like transcription modulator	**SLTM**	ASGQAKSSSKESKDSKTSSKDDKGS	347	371
O95218	Zinc finger Ran-binding domain-containing protein 2	**ZRANB2**	ASILKEVEDKESEGEEEDEDED	142	163
O75592	E3 ubiquitin-protein ligase MYCBP2	**MYCBP2**	ASKNSVQSGESDSDEEEESKE	133	153
P30414	NK-tumor recognition protein	**NKTR**	ASKSSSHSRSRSKSRSSSKS	538	557
O15054	Lysine-specific demethylase 6B	**KDM6B**	ASSFQESLQEEKESEDEESEE	1285	1305
Q96S99	Pleckstrin homology domain-containing family F member 1	**PLEKHF1**	ASSGDDDDSDEDKEGSRDGD	241	260
Q92620	Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16	**DHX38**	ASSKEEKDWKKEKSRDRDYD	146	165
Q5BKY9	Protein FAM133B	**FAM133B**	ATEKTKKKKKHKKHSKKKKKK	218	238
Q68CQ4	Digestive organ expansion factor homolog	**DIEXF**	ATLKNVSEEEEEDEEEEEEDS	77	98
Q5T4S7	E3 ubiquitin-protein ligase UBR4	**UBR4**	ATTQSKSSTKKSKKEEKEKEKD	3359	3380
Q12888	Tumor suppressor p53-binding protein 1	**TP53BP1**	AVEEEKGEELEQKEKEKEED	130	149
O75592	E3 ubiquitin-protein ligase MYCBP2	**MYCBP2**	AVLKKKECEKENKKSKKEKKKKEKAE	3153	3178
O75808	Calpain-15	**CAPN15**	CEDKDEEEKEEQEEEEGAAE	123	142
Q9NQX3	Gephyrin	**GPHN**	CEEEEEEKKDSGVASTEDSSSS	212	233
Q9BZI1	Iroquois-class homeodomain protein IRX-2	**IRX2**	CESGSECKDKYDDLEDDEDDDEEGE	250	274
O95425	Supervillin	**SVIL**	CESKAEEEEGEGEGEEKEED	423	442
Q5VYM1	Uncharacterized protein C9orf131	**C9orf131**	CKQKSEVEEEGEEEEEGEDE	92	111
Q96F63	Coiled-coil domain-containing protein 97	**CCDC97**	CLEEEEEEEDSDEEDQRSGKDSEA	247	270
C9JE40	Protein PAT1 homolog 2	**PATL2**	CQLEKEEENEGEEEEEEEDEED	24	45
O94759	Transient receptor potential cation channel subfamily M member 2	**TRPM2**	CSKILKELSKEEEDTDSSEE	655	674
O95153	Peripheral-type benzodiazepine receptor-associated protein 1	**BZRAP1**	CSKKLFSIPEEEEEEEEDEEEEKSG	1325	1349
Q9UMN6	Histone-lysine N-methyltransferase 2B	**KMT2B**	CWKKQEQKLDDEEEEKKEEEEKDKEGEEKEERA	353	385
Q9H1E5	Thioredoxin-related transmembrane protein 4	**TMX4**	DAEEEKDDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEEDNLA	243	286
Q9UL36	Zinc finger protein 236	**ZNF236**	DAEQEKEQESPEKLDKKEKK	441	460
Q9P2B7	UPF0501 protein KIAA1430	**KIAA1430**	DAGSDSHLSDSSPSSKSSKK	190	209
P11362	Fibroblast growth factor receptor 1	**FGFR1**	DALPSSEDDDDDDDSSSEEKETDNTK	120	145

P49750	YLP motif-containing protein 1	**YLPM1**	DANIEEQKEEKKDAEEEESE	1822	1841
P27824	Calnexin	**CANX**	DAPQPDVKEEEEEKEEEKDK	518	537
A6NGG8	Uncharacterized protein C2orf71	**C2orf71**	DASSLSDSEDSSPEEEEEDKMSSMSL	477	502
Q96MY1	Nucleolar protein 4-like	**NOL4L**	DDDDDDHDDHEDNDKMNDSE	161	180
Q6PL18	ATPase family AAA domain-containing protein 2	**ATAD2**	DDDDDEDDEDEEDGEEENQKRY	273	294
O75822	Eukaryotic translation initiation factor 3 subunit J	**EIF3J**	DDDDDEKKEEAEVKPEVKISEKKKI	53	77
P06748	Nucleophosmin	**NPM1**	DDDDDFDDEEAEEKAPVKKS	176	195
Q96AJ1	Clusterin-associated protein 1	**CLUAP1**	DDDDEDDDLEDESISLSPTK	376	395
Q86VP6	Cullin-associated NEDD8-dissociated protein 1	**CAND1**	DDDDQGSDDEYSDDDDMSWK	329	348
Q7Z3E5	LisH domain-containing protein ARMC9	**ARMC9**	DDEDEDDEEDHDIMEADLDKDE	584	605
O75167	Phosphatase and actin regulator 2	**PHACTR2**	DDEDEDEDEDGSGESALASK	453	472
Q15542	Transcription initiation factor TFIID subunit 5	**TAF5**	DDEDEEGENEEGKPKKKKPKKDSI	387	410
P35659	Protein DEK	**DEK**	DDEDKESEEEPPKKTAKREK	245	264
O15047	Histone-lysine N-methyltransferase SETD1A	**SETD1A**	DDEEDEEDEDREEAVDTTKKE	980	1000
P19338	Nucleolin	**NCL**	DDEEEEEEEEEEPVKEAPGK	260	279
Q6ZWH5	Serine/threonine-protein kinase Nek10	**NEK10**	DDELDISDNSSSSSSSPLKES	903	923
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	**MPHOSPH10**	DDELDSNKEDDEIAEEEAEE	284	303
Q96MI9	Cytosolic carboxypeptidase 4	**AGBL1**	DDEVDKDSDTEDGKVEDDDLE	262	282
Q9Y6M7	Sodium bicarbonate cotransporter 3	**SLC4A7**	DDKKKKKEKEEAERMLQDDDD	1147	1166
Q13017	Rho GTPase-activating protein 5	**ARHGAP5**	DDKKMKKKTHKVKEDKKQKKK	1223	1243
A1L162	Glutamate-rich protein 2	**ERICH2**	DDKLSESAEDDGEDDTNDEDDDEDSN	33	58
Q76FK4	Nucleolar protein 8	**NOL8**	DDKPKESKAKRKKKREEAEK	962	981
Q9HAU5	Regulator of nonsense transcripts 2	**UPF2**	DDKRKKEDKERKKKDEEKVKAEEESKKKEEEEK KKHQEEERKKQEEQAKRQQEEEA	62	117
Q92541	RNA polymerase-associated protein RTF1 homolog	**RTF1**	DDKSSEKSDRSSRTSSSDEEEEKEEIPPKSQP	321	352
Q6U841	Sodium-driven chloride bicarbonate exchanger	**SLC4A10**	DDLMPESKKKKLEDAEKEEE	1036	1055
P21817	Ryanodine receptor 1	**RYR1**	DDLSKAGEQEEEEEEVEEKK	3675	3694
Q9H8M2	Bromodomain-containing protein 9	**BRD9**	DDRSDHERERHKEKKKKKKKKSEKEKHLDDEER	53	85
P08238	Heat shock protein HSP 90-beta	**HSP90AB1**	DDSGKDKKKKTKKIKEKYID	259	278
Q58FF8	Putative heat shock protein HSP 90-beta 2	**HSP90AB2P**	DDSGKDKKKKTKKIKEKYID	181	200
O00193	Small acidic protein	**SMAP**	DDSPDPESPDDSESDSESEKEESAEELQA	125	153
Q9Y6R4	Mitogen-activated protein kinase kinase kinase 4	**MAP3K4**	DDTEGELKELESSTDESEEE	445	464
Q5VWT5	Uncharacterized protein C1orf168	**C1orf168**	DDVDLSEKESKDEDKLKMWK	588	607
Q13435	Splicing factor 3B subunit 2	**SF3B2**	DDVKKEKEKEPEKLDKLENS	397	416
P46100	Transcriptional regulator ATRX	**ATRX**	DDYTKKKKKGKKGKKDSSSSGS	1925	1946
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	DEDDDKLSESKSDGRERSKKSSVSDA	1364	1389
P04198	N-myc proto-oncogene protein	**MYCN**	DEDDEEEDEEEEIDVVTVEK	265	284
O94818	Nucleolar protein 4	**NOL4**	DEDEDDHEDHDDSEKVNETD	382	401
Q3BBV2	Putative neuroblastoma breakpoint family member 8	**NBPF8**	DEDEDEDVQVEEDEKVLESS	132	151
Q5RGN0	Putative neuroblastoma breakpoint family member 24	**NBPF24**	DEDEDEDVQVEEDEKVLESS	92	111
Q86T75	Neuroblastoma breakpoint family member 11	**NBPF11**	DEDEDEDVQVEEDEKVLESS	167	186
P26358	DNA (cytosine-5)-methyltransferase 1	**DNMT1**	DEDEDGDEKDEKKHRSQPKD	273	292
Q13562	Neurogenic differentiation factor 1	**NEUROD1**	DEDEDLEEEEEEEEEDDDQKPKRRG	61	85
P20042	Eukaryotic translation initiation factor 2 subunit 2	**EIF2S2**	DEDEILEKDEALEDEDNKKDD	135	155
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein	**EIF3CL**	DEDKKAAEKKREDKAKKKHD	264	283
Q99613	Eukaryotic translation initiation factor 3 subunit C	**EIF3C**	DEDKKAAEKKREDKAKKKHD	264	283
Q6ZMS4	Zinc finger protein 852	**ZNF852**	DEDKKKSTKDRYEEYKEVEE	119	138
Q12955	Ankyrin-3	**ANK3**	DEDKPPSKSSSSEKTPDKTD	3476	3495

O75037	Kinesin-like protein KIF21B	**KIF21B**	DEDSGSEESLVDSDSDPEEKEV	609	630
Q8IZQ1	WD repeat and FYVE domain-containing protein 3	**WDFY3**	DEDSSDSEADEQSISQDPKD	3280	3299
P46100	Transcriptional regulator ATRX	**ATRX**	DEEAKNQVNSESDSDSEESKK	1313	1333
Q9NW13	RNA-binding protein 28	**RBM28**	DEEDGVFDDEDEEEENIESK	241	260
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	DEEDQNEEKGEAGKEDKDEK	133	152
Q9P212	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase epsilon-1	**PLCE1**	DEEDSPSEGNSSRKSSLKDKS	784	804
Q8NFD5	AT-rich interactive domain-containing protein 1B	**ARID1B**	DEEEDSEKTESDEKSSIALT	1734	1753
P13667	Protein disulfide-isomerase A4	**PDIA4**	DEEEEEEEDDDEEEDDLEVKEENGVL	40	65
O00294	Tubby-related protein 1	**TULP1**	DEEEEEEEDEEDEEEEAEEKKEKILLPPKKPL	116	147
Q8N2W9	E3 SUMO-protein ligase PIAS4	**PIAS4**	DEEEEEEEEEDEDEEGPRPKR	478	498
Q8WYN3	Cysteine/serine-rich nuclear protein 3	**CSRNP3**	DEEEEEEEEEEEEEDDDDDKGDGFVEGLG	370	398
Q5JTC6	APC membrane recruitment protein 1	**AMER1**	DEEEEEEEEVELEEEEEEVKEEEEDDDLEYLWET	379	412
Q9Y5B6	PAX3- and PAX7-binding protein 1	**PAXBP1**	DEEEENEEVFKVKKSSYSKK	111	130
Q6ZV73	FYVE, RhoGEF and PH domain-containing protein 6	**FGD6**	DEEEIINSSDEDDVSSESSKGE	833	854
O00541	Pescadillo homolog	**PES1**	DEEGENEEEEEDAEAGSEKEEE	472	493
Q9ULD4	Bromodomain and PHD finger-containing protein 3	**BRPF3**	DEEGLKEGDGEEEEEEEVEEEE	409	430
Q8N4S0	Coiled-coil domain-containing protein 82	**CCDC82**	DEELDSDESFENDEELDSNK	52	71
P35663	Cylicin-1	**CYLC1**	DEESTDADSEPKGDSKKGKKDEKKGKKDSKKDDKK KDAKKNAESTEMESDLELKKDKKHSKEKKGSKKDIK KDARKDTESTD	440	521
Q96T23	Remodeling and spacing factor 1	**RSF1**	DEETKKEEETPKQEEQKESEK	232	252
O00213	Amyloid beta A4 precursor protein-binding family B member 1	**APBB1**	DEGEEKAAGEAEEEEEDDDDEEEEEDLSSP	147	176
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	DEGSEKADASSGSDSTSSSSEDSDSSNED	2869	2897
Q15061	WD repeat-containing protein 43	**WDR43**	DEIADKDSEDNWDEDEEESESEKDEDVEEE DEDAEGKDEENGEDRDTASEKEL	610	662
O75351	Vacuolar protein sorting-associated protein 4B	**VPS4B**	DEKGNDSDGEGESDDPEKKK	96	115
O43719	HIV Tat-specific factor 1	**HTATSF1**	DEKLFEDDDSNEKLFDEEEDSSEKL	693	717
Q9H6T3	RNA polymerase II-associated protein 3	**RPAP3**	DELDKDDSTHESLSQESESEED	103	124
Q5SVQ8	Zinc finger and BTB domain-containing protein 41	**ZBTB41**	DELEEEMSDEYSDIEEQSEKD	307	327
O60763	General vesicular transport factor p115	**USO1**	DELESGDQEDEDDESEDPGKD	938	958
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	DERKDKKEERDDETDEPKPK	809	828
Q9BXY0	Protein MAK16 homolog	**MAK16**	DESDISDFEDMDKLDASSDED	227	247
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	DESENDDEDTALSEASEKDE	1115	1134
P09874	Poly [ADP-ribose] polymerase 1	**PARP1**	DEVDGVDEVAKKKSKKEKDKDSKLEK	211	236
P31629	Transcription factor HIVEP2	**HIVEP2**	DFHSSKELSSSTEESKDPSSEKSQ	2421	2444
A6PVY3	Protein FAM177B	**FAM177B**	DGDIMEEYSTEEEEEEEKEE	30	49
O60832	H/ACA ribonucleoprotein complex subunit 4	**DKC1**	DGDSDTTKKKKKKKKAKEVE	491	510
Q92620	Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16	**DHX38**	DGEDKKKSKVSSYKDWEESKDDQKDAEEEG	67	96
O94913	Pre-mRNA cleavage complex 2 protein Pcf11	**PCF11**	DGKDDDVKEKRKTAEKKDKDE	413	433
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	DGKEKDKKDKDIDRYKERDK	1013	1032
Q9HCE9	Anoctamin-8	**ANO8**	DGPPGGKEEDEDDEEEEDEEEEEDEEEGEEGG	575	606
Q96GA3	Protein LTV1 homolog	**LTV1**	DIQKSENEDDSEWEDVDDEK	178	197
P21815	Bone sialoprotein 2	**IBSP**	DITNKATKEKESDEEEEEEEE	138	158
Q9H501	ESF1 homolog	**ESF1**	DKDALEEDSESVSEIGSDEESE	235	256
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	DKDEKKEDGKQAENDSSNDDK	304	324
Q96BU1	S100P-binding protein	**S100PBP**	DKDETDSSKDTEKLSSLGEE	161	180

A8MW92	PHD finger protein 20-like protein 1	**PHF20L1**	DKDHYRPKQKKKKKKKKSK	560	579
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	DKDRDKKKSHSIKPESWSKS	1517	1536
Q9HAU5	Regulator of nonsense transcripts 2	**UPF2**	DKDSKDSMTEGENLEEDEEEEE	1016	1037
Q8N7H5	RNA polymerase II-associated factor 1 homolog	**PAF1**	DKEEIFGSDADSEDDADSDDED	449	470
Q5SRN2	Uncharacterized protein C6orf10	**C6orf10**	DKEKERDAEKDPNKKEKGDK	500	519
Q15042	Rab3 GTPase-activating protein catalytic subunit	**RAB3GAP1**	DKEKGEVGKSWDSWSDSEEE	565	584
Q03468	DNA excision repair protein ERCC-6	**ERCC6**	DKEKRLKLEDDSEESDAEFDE	475	495
Q15858	Sodium channel protein type 9 subunit alpha	**SCN9A**	DKEKYEQDRTEKEDKGKDSKESKK	1965	1988
O43719	HIV Tat-specific factor 1	**HTATSF1**	DKELEENDSENSEFEDDGSEK	589	609
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	DKERKEKTKPERYKEKSSDKDKSEKSILEK	1016	1045
Q8IZD6	Solute carrier family 22 member 15	**SLC22A15**	DKESSLGSESEEEEEFYDADEE	521	542
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	DKGEKKPDSNEKGERKKEKKEK	320	341
Q9UGP8	Translocation protein SEC63 homolog	**SEC63**	DKGSDSEEEETNRDSQSEKDD	571	591
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	DKISKEKEKIFKEDKEKLKKEKVY	785	808
O43719	HIV Tat-specific factor 1	**HTATSF1**	DKKAEEGDADEKLFEESDDKEDEDAD	660	685
O60885	Bromodomain-containing protein 4	**BRD4**	DKKEKKKEKHKRKEEVEENKKSKAKEPP	542	569
Q8IYB3	Serine/arginine repetitive matrix protein 1	**SRRM1**	DKKHKKDKKHKKHKKHKKEKA	816	836
Q9NU22	Midasin	**MDN1**	DKLDERLWGDDDEEEDEEEED	4772	4792
Q76M96	Coiled-coil domain-containing protein 80	**CCDC80**	DKLLKSEKQMKKSEKKSKQEKEKSKKKKGGKTE	560	592
Q96JM4	Leucine-rich repeat and IQ domain-containing protein 1	**LRRIQ1**	DKLSISSLEKEDIESDAKSE	17	36
Q9Y5T5	Ubiquitin carboxyl-terminal hydrolase 16	**USP16**	DKNLKKTVEDEDQDSEEEKD	401	420
Q9UPR3	Protein SMG5	**SMG5**	DKPSEPASEEGSESEGSESS	613	632
Q66PJ3	ADP-ribosylation factor-like protein 6-interacting protein 4	**ARL6IP4**	DKRRKKKKKRKKLKKKGKEK	293	312
P83916	Chromobox protein homolog 1	**CBX1**	DKSEGGKRKADSDSEDKGEESK	78	99
Q8NI27	THO complex subunit 2	**THOC2**	DKSRERSREREKKDEKDRKE	1461	1480
Q9UKL3	CASP8-associated protein 2	**CASP8AP2**	DKSSRSSKTEKKDKVMSTSS	1002	1021
P20810	Calpastatin	**CAST**	DKVKEKAKEEDREKLGEKEE	456	475
Q13206	Probable ATP-dependent RNA helicase DDX10	**DDX10**	DKYRSSEDSDSEDMENKISD	799	818
Q9Y5B0	RNA polymerase II subunit A C-terminal domain phosphatase	**CTDP1**	DLDFDLSSDSESSSESEGTKSSSSASDGESEGK	449	481
Q96T23	Remodeling and spacing factor 1	**RSF1**	DLDSDSNLDEEESEDEFKISD	1093	1113
Q5C9Z4	Nucleolar MIF4G domain-containing protein 1	**NOM1**	DLESDSQDESEEEEEGDVEKEKKAQE	250	275
Q96MI9	Cytosolic carboxypeptidase 4	**AGBL1**	DLPEEDFEDDGDDEVDKDSD	251	270
Q16659	Mitogen-activated protein kinase 6	**MAPK6**	DLSNWKEQSKEKSDKKGKSK	475	494
Q9P2R6	Arginine-glutamic acid dipeptide repeats protein	**RERE**	DLSSASEDDFDSEDSEQELK	484	503
Q14093	Cylicin-2	**CYLC2**	DNKKDKKDSNKGKDSATESE	188	207
Q96QF7	Acidic repeat-containing protein	**ACRC**	DNSDDSSDDNSDDSDVPDDKSDDSDVPDDSSDDSDVPD DSSDDSEAPDDSSDDSEAPDDSSDDSEAPDDSSDDSEAP DDSSDDSEASDDSSDDSEASDDSSDDSEAPDDKSDDSDV PEDKSDDSDVPDDNSDDLEVPVPAEDLCNEGQI	199	347
Q96T23	Remodeling and spacing factor 1	**RSF1**	DNSSRESPSLEDEETKKEEE	221	240
Q96JN2	Coiled-coil domain-containing protein 136	**CCDC136**	DPEEAKSTEDQEENEEDKEEEEKEEDSEEEEDDADSSLESPEENNP	1068	1113
O00555	Voltage-dependent P/Q-type calcium channel subunit alpha-1A	**CACNA1A**	DPLPKKEEEKKEEEEDDRGED	1193	1213
Q06323	Proteasome activator complex subunit 1	**PSME1**	DPVKEKEKEERKKQQEKEDKDEKKKGEDEDKGP	67	99
Q13316	Dentin matrix acidic phosphoprotein 1	**DMP1**	DQEDSDSSEEDSSHTLSHSKSESREE	377	402
Q96JM2	Zinc finger protein 462	**ZNF462**	DQLEQMKEKMESSSSDDEDKEEEM	2366	2389
Q8WYA6	Beta-catenin-like protein 1	**CTNNBL1**	DRDGEEEEEEEEPLDESSVKK	64	84
P19447	TFIIH basal transcription factor complex helicase XPB subunit	**ERCC3**	DRDKKKSRKRHYEDEEDDEED	8	28

Q8WXX7	Autism susceptibility gene 2 protein	**AUTS2**	DRDVDKRDSSVSKDDKERES	827	846
Q9NQ38	Serine protease inhibitor Kazal-type 5	**SPINK5**	DREANERKKKDEEKSSSKPS	890	909
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	DREKDKEKDRERE REKEHEKD	308	328
O95232	Luc7-like protein 3	**LUC7L3**	DREQDRKSKEKEKRGSDDKKSS	369	390
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	DRKSEDDKEEEERKRQEEIE	694	713
Q86X95	Corepressor interacting with RBPJ 1	**CIR1**	DRLEKKKKKKDRKKKKFQKS	229	248
Q12955	Ankyrin-3	**ANK3**	DRLSEEEKKMQSELSDEEES	4012	4031
Q12872	Splicing factor, suppressor of white-apricot homolog	**SFSWAP**	DRPSSKSKDPPREEEKEKKKKK	739	760
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	**ACIN1**	DRRRKERAKSKEKKSEKKEK	1182	1201
Q8TA86	Retinitis pigmentosa 9 protein	**RP9**	DRSSSSSSEGKEKHKKKKKKEKHKKRKKEKKKKKKRKHKSSKSNEGSDSE	172	221
O60229	Kalirin	**KALRN**	DSADEKSKKGWGEDEPDEES	1816	1835
Q8N556	Actin filament-associated protein 1	**AFAP1**	DSDAMSSSYESYDEEEEDGK	114	133
Q8TED9	Actin filament-associated protein 1-like 1	**AFAP1L1**	DSDAMSSSYESYDEEEEEGKS	181	201
Q5M9Q1	NKAP-like protein	**NKAPL**	DSDDDKKRVKAKKKKKKKKHKTKKKKNKKTKKE SSDSSCKDSEEDLSEATWME	215	267
Q13435	Splicing factor 3B subunit 2	**SF3B2**	DSDDDSSDDEQEKKPEAPKLSKKKL	430	454
O43719	HIV Tat-specific factor 1	**HTATSF1**	DSDEKEEEEDTYEKVFDDESDEKEDEEYADEKGLEA	623	658
Q9UBN4	Short transient receptor potential channel 4	**TRPC4**	DSDEKSDSEGNSKDKKKNFSL	776	796
Q01538	Myelin transcription factor 1	**MYT1**	DSDGSEDTEVKDASVSDESE	90	109
Q15653	NF-kappa-B inhibitor beta	**NFKBIB**	DSDLEKEEEESEEDWKLQLE	182	201
Q5M9Q1	NKAP-like protein	**NKAPL**	DSDSNSESDTNSDSDDDKKRVK	203	224
Q9Y2K6	Ubiquitin carboxyl-terminal hydrolase 20	**USP20**	DSDSSDTDEKREGDRSPSEDE	262	282
Q68DQ2	Very large A-kinase anchor protein	**CRYBG3**	DSDSSLEKNSSASEDSSFLK	1011	1030
P83916	Chromobox protein homolog 1	**CBX1**	DSEDKGEESKPKKKKEESEKPRGF	90	113
Q9NZ63	Uncharacterized protein C9orf78	**C9orf78**	DSESEEDEQDSEEVRLKLEE	14	33
Q14093	Cylicin-2	**CYLC2**	DSESELKQGKKDSKKGKDIEK	129	149
Q9NU22	Midasin	**MDN1**	DSGNSNKDKSQQDKKEEKEE	4817	4836
Q9P0W5	Schwannomin-interacting protein 1	**SCHIP1**	DSGSDKDSDADDSKTETSLD	307	326
A3KN83	Protein strawberry notch homolog 1	**SBNO1**	DSGSESDASDNEESDYESSKN	732	752
P30414	NK-tumor recognition protein	**NKTR**	DSHPSSDKEEGEATSDSESE	902	921
Q8NFI4	Putative protein FAM10A5	**ST13P5**	DSKKVEEDLKADEPSTEESD	61	80
Q8TF50	Zinc finger protein 526	**ZNF526**	DSLEKEERNGLEEEEDDEEDEEDDEEMEDEEAM	221	254
Q9UGP8	Translocation protein SEC63 homolog	**SEC63**	DSQSEKDDGSDRDSDREQDEK	584	604
O95402	Mediator of RNA polymerase II transcription subunit 26	**MED26**	DSSKADSDAASSGGSDSKKKK	375	395
Q96CJ1	ELL-associated factor 2	**EAF2**	DSSSDSEDEDCKSSTSDTGN	192	211
Q9Y4B4	Helicase ARIP4	**RAD54L2**	DSSSGSEDEKSSRDEVIELSS	186	206
Q12830	Nucleosome-remodeling factor subunit BPTF	**BPTF**	DSSSKSKKKKMISTTSKETKKD	2844	2865
O95096	Homeobox protein Nkx-2.2	**NKX2-2**	DSSSKSPEPSADESPDNDKE	98	117
Q5TAX3	Terminal uridylyltransferase 4	**ZCCHC11**	DSSSLSTSKSSEIEPKLDKK	796	815
Q96KG7	Multiple epidermal growth factor-like domains protein 10	**MEGF10**	DSSSSPKQEDSGGSSSNSSSSSE	1118	1140
Q9UL68	Myelin transcription factor 1-like protein	**MYT1L**	DSSSVDECDDSDGTEDMDEKEEDEGEEYSEDNDE	83	116
P35659	Protein DEK	**DEK**	DSSTTKKNQNSSKKESESEDSSDDEP	286	311
Q3ZCN5	Otogelin-like protein	**OTOGL**	DSSVKASKYDDSEEFKHSSS	1328	1347
Q5VTL8	Pre-mRNA-splicing factor 38B	**PRPF38B**	DSVEKSKKREHSPSKEKSRK	480	499
Q9NY46	Sodium channel protein type 3 subunit alpha	**SCN3A**	DSVTKPDKEKFEKDKPEKESK	1971	1991
O60840	Voltage-dependent L-type calcium channel subunit alpha-1F	**CACNA1F**	DTEEEEEEGQEGVEEEDEKD	1638	1657
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	DTEGSEEEDDDDKDQDESDSDTEGEKTSMK	642	671
P30419	Glycylpeptide N-tetradecanoyltransferase 1	**NMT1**	DTGAKKKKKKQKKKKEKGSE	51	70

Q13206	Probable ATP-dependent RNA helicase DDX10	**DDX10**	DTGNEEQEEEEDDEEEMEEK	576	595
P15884	Transcription factor 4	**TCF4**	DTKSSEDKKLDDDKKDIKSITS	525	546
O15047	Histone-lysine N-methyltransferase SETD1A	**SETD1A**	DTTKKETEVSDGEDEESDSSSK	995	1016
Q6DN14	Multiple C2 and transmembrane domain-containing protein 1	**MCTP1**	DTVVEDMLEDEEEEDDKDDKDSEKKG	851	876
Q9HCG8	Pre-mRNA-splicing factor CWC22 homolog	**CWC22**	DVEQNKSSPSSSSSASSSSESDSSDSDSDSSDSSSESSSEESDSSSISSHSSASANDVRKKGHGKT	660	725
P07900	Heat shock protein HSP 90-alpha	**HSP90AA1**	DVGSDEEEEKKDGDKKKKKKIKEKYIDQEE	260	289
Q9UGP8	Translocation protein SEC63 homolog	**SEC63**	EAAVKEDEEEVSDKGSDSEEEE	559	580
Q9UNL4	Inhibitor of growth protein 4	**ING4**	EADLKEKQIESSDYDSSSSK	108	127
Q99592	Zinc finger and BTB domain-containing protein 18	**ZBTB18**	EADSTKKEEDASSCSDKVES	124	143
P78362	SRSF protein kinase 2	**SRPK2**	EAEDQEEKEDAEKENIEKDEDD	344	365
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	EAEEEKAKSGDETNKEEDEDDEEAEEEEEEEEEEEDEDDDDNNEEEEFECYPP	519	571
P07197	Neurofilament medium polypeptide	**NEFM**	EAEGEEAEAKEEKKVEEKSEE	574	594
Q5C9Z4	Nucleolar MIF4G domain-containing protein 1	**NOM1**	EAQSEDDDEDTEEEQGEEKEK	277	297
Q9BTK6	PAXIP1-associated glutamate-rich protein 1	**PAGR1**	EAQSEEERSDEEPEAKEEEEEK	140	161
Q14596	Next to BRCA1 gene 1 protein	**NBR1**	EAVMEEEEDEEDEEEEDELKDEVQS	699	723
Q96PY6	Serine/threonine-protein kinase Nek1	**NEK1**	EDAKEHEKEKSVSSDRKKWE	742	761
Q9H2P0	Activity-dependent neuroprotector homeobox protein	**ADNP**	EDASESEEKLDQKEDGSKYE	918	937
Q9H160	Inhibitor of growth protein 2	**ING2**	EDCDDQPPKEKKSKSAKKKK	171	190
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**	EDDDDSDYGSSKKKNKKMVKKSKPE	139	163
P51946	Cyclin-H	**CCNH**	EDDDYVSKKSKHEEEEWTDDD	298	318
P29692	Elongation factor 1-delta	**EEF1D**	EDDEDDDIDLFGSDNEEEDKE	150	170
Q9UER7	Death domain-associated protein 6	**DAXX**	EDDEEEDEEEEAAAGKDGDKS	475	495
Q7Z4S6	Kinesin-like protein KIF21A	**KIF21A**	EDDIDGGESSDESDSESDEKAN	621	642
Q13439	Golgin subfamily A member 4	**GOLGA4**	EDDKSKMEKKESNLETELKS	1503	1522
Q14515	SPARC-like protein 1	**SPARCL1**	EDDSHHKAEKSSVLKSKEES	61	80
P29375	Lysine-specific demethylase 5A	**KDM5A**	EDDSMEEKPLKVKGKDSSEKK	1485	1505
Q5H9L4	Transcription initiation factor TFIID subunit 7-like	**TAF7L**	EDEDDEDEDEDEDEDEDKEEEEEDCSEEYLERQLQAEFIESG	336	379
Q9BYV6	Tripartite motif-containing protein 55	**TRIM55**	EDEDEEEEEGGEGEKEGEGE	329	348
Q86TY3	Uncharacterized protein C14orf37	**C14orf37**	EDEEDEDEEEEDEEEDEEDKDADSLDEGLDGDTELPG	619	655
Q5M9Q1	NKAP-like protein	**NKAPL**	EDEEEVTHQKSSSSDSNSEE	155	174
Q8N5F7	NF-kappa-B-activating protein	**NKAP**	EDEEPKKSTTSASTSEEEKKKKSSRSKERSKKR	164	196
O15047	Histone-lysine N-methyltransferase SETD1A	**SETD1A**	EDEESDSSSKCSLYADSDGE	1007	1026
Q5TCS8	Adenylate kinase 9	**AK9**	EDEFPKDEEEMSGEEDEEQE	1237	1256
Q13490	Baculoviral IAP repeat-containing protein 2	**BIRC2**	EDEKREEEKEKQAEEMASDD	438	457
O95760	Interleukin-33	**IL33**	EDESYEIYVEDLKKDEKKDK	139	158
Q9BXW9	Fanconi anemia group D2 protein	**FANCD2**	EDGEEDEVSAGEKEQDSDES	1427	1446
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	EDGKEKGDKKEGKDVKVKEDEKERED	212	237
Q14692	Ribosome biogenesis protein BMS1 homolog	**BMS1**	EDIEKEVKEEIDPDEEESAKKK	803	824
O76041	Nebulette	**NEBL**	EDIKDETEEEKIGEEENEED	7	26
Q7Z401	C-myc promoter-binding protein	**DENND4A**	EDIQEEKDKKGSDCSSLSESES	924	945
Q99613	Eukaryotic translation initiation factor 3 subunit C	**EIF3C**	EDKAKKKHDRKSKRLDEEEED	275	295
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein	**EIF3CL**	EDKAKKKHDRKSKRLDEEEEED	275	296
P0C860	Putative male-specific lethal-3 protein-like 2	**MSL3P1**	EDKDEKISEECDIEEKTEVKEE	62	83
P20042	Eukaryotic translation initiation factor 2 subunit 2	**EIF2S2**	EDKDLEADEEDTRKKDASDD	50	69

Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	EDKEDKKGKKTDICEDEDEGD	1319	1339
Q9NVM9	Protein asunder homolog	**ASUN**	EDKEDKSEKAVKDYEQEKSW	578	597
O95674	Phosphatidate cytidylyltransferase 2	**CDS2**	EDKESESEAKVDGETASDSES	17	37
Q6Q759	Sperm-associated antigen 17	**SPAG17**	EDKPKLEKDKGKAKSPKEKK	152	171
Q05682	Caldesmon	**CALD1**	EDKPTFKKEEIKDEKIKKDKE	453	473
Q9Y2W2	WW domain-binding protein 11	**WBP11**	EDMDQDKHDDSTDDSDTDKSD	260	280
Q15047	Histone-lysine N-methyltransferase SETDB1	**SETDB1**	EDPEESNDDSSDDNFCKDED	896	915
O00567	Nucleolar protein 56	**NOP56**	EDPSISFSKPKKKKSFSKEE	497	516
Q9BU76	Multiple myeloma tumor-associated protein 2	**MMTAG2**	EDQTESSCESHRKSKKEKKKKKKRK	172	196
O15259	Nephrocystin-1	**NPHP1**	EDSEDSGGEEEDAEEEEEEKEENESHKWST	124	153
P55042	GTP-binding protein RAD	**RRAD**	EDSEDSLSSGGSDSDESVYK	74	93
Q14686	Nuclear receptor coactivator 6	**NCOA6**	EDSEMDFDSGLEDDDTKSDS	22	41
Q8ND61	Uncharacterized protein C3orf20	**C3orf20**	EDSESVKKAESEDIQGSSSS	841	860
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	EDSLATSLSSSASSSSEESKESSR	1533	1556
Q9Y6V0	Protein piccolo	**PCLO**	EDSSEEEELREEEELLKEQEK	1702	1722
P08238	Heat shock protein HSP 90-beta	**HSP90AB1**	EDVGSDEEDDSGKDKKKKTKKIKEKY	251	276
Q58FF8	Putative heat shock protein HSP 90-beta 2	**HSP90AB2P**	EDVGSDEEDDSGKDKKKKTKKIKEKY	173	198
Q14568	Putative heat shock protein HSP 90-alpha A2	**HSP90AA2**	EDVGSDEEEEKKDGDKKKKKTKEKYIDQEE	259	288
Q5VTH9	WD repeat-containing protein 78	**WDR78**	EDVLESAKHEEVEEESKKEEEEEI	438	461
Q9UQD0	Sodium channel protein type 8 subunit alpha	**SCN8A**	EDVSSESDPEGSKDKLDDTSSSEGS	1111	1135
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein	**EIF3CL**	EDWDTGSTSSDSDSEEEEGK	229	248
Q99613	Eukaryotic translation initiation factor 3 subunit C	**EIF3C**	EDWDTGSTSSDSDSEEEEGK	229	248
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	EEAEDEEDDRDEEEMTKRDDK	876	896
P07196	Neurofilament light polypeptide	**NEFL**	EEAKDEPPSEGEAEEEEKDKEE	464	485
Q02952	A-kinase anchor protein 12	**AKAP12**	EECKEEGEEKQEKEPSKSAES	263	283
P06748	Nucleophosmin	**NPM1**	EEDAESEDEEEEDVKLLSIS	120	139
P19338	Nucleolin	**NCL**	EEDDEDEDEDEDEIEPAAMK	157	176
Q96M86	Dynein heavy chain domain-containing protein 1	**DNHD1**	EEDDESEESNEAEDQTKEQK	3603	3622
Q13144	Translation initiation factor eIF-2B subunit epsilon	**EIF2B5**	EEDEDDGEFSDDSGADQEKDK	457	477
Q8TDN2	Potassium voltage-gated channel subfamily V member 2	**KCNV2**	EEDEDGEEEDQWKDDLAEED	56	75
Q9P1Z0	Zinc finger and BTB domain-containing protein 4	**ZBTB4**	EEDEEEEDEEEEEEDEEESKAGGEDQLWR	637	665
P45379	Troponin T, cardiac muscle	**TNNT2**	EEDEEEEEAKEAEDGPMEESK	54	74
Q86WV1	Src kinase-associated phosphoprotein 1	**SKAP1**	EEDEEEEEKEETYDDIDGFDS	220	240
Q8IWZ3	Ankyrin repeat and KH domain-containing protein 1	**ANKHD1**	EEDEENKPKENSELPEDEDEEE	1472	1493
Q15911	Zinc finger homeobox protein 3	**ZFHX3**	EEDKEEESDLEDKQSPTGSDS	1490	1510
Q9UIS9	Methyl-CpG-binding domain protein 1	**MBD1**	EEDKEENKDDSASKLAPEEE	545	564
Q9H4A5	Golgi phosphoprotein 3-like	**GOLPH3L**	EEDSNWEKSPDNEDSGDSKD	24	43
Q8N9Q2	Protein SREK1IP1	**SREK1IP1**	EEDTSKQKKQKYQKKEKKKEKKSKSKKGKH	101	130
Q96HJ3	Coiled-coil domain-containing protein 34	**CCDC34**	EEDVDDEEDVDEDAHDSEAK	87	106
Q9UPZ3	Hermansky-Pudlak syndrome 5 protein	**HPS5**	EEDVSSDTCPKEEDTEEEKE	580	599
P07196	Neurofilament light polypeptide	**NEFL**	EEEAAEEEEAAKEESEEAKEEEEG	488	511
Q9NW82	WD repeat-containing protein 70	**WDR70**	EEEAEEEEEEEEEEENPVHK	150	169
Q8TAQ2	SWI/SNF complex subunit SMARCC2	**SMARCC2**	EEEAKEKTSEAPKKDEEKGKE	783	803
Q24JP5	Transmembrane protein 132A	**TMEM132A**	EEEARKEETEAREEEEEEEEE	815	835
P55209	Nucleosome assembly protein 1-like 1	**NAP1L1**	EEECEWKPDEEDEISEELKEK	129	149
Q6AI12	Ankyrin repeat domain-containing protein 40	**ANKRD40**	EEEDDDDDDDDNLPQLKKESE	96	116
Q9UD71	Protein phosphatase 1 regulatory subunit 1B	**PPP1R1B**	EEEDDEEEEEEEDSQAEVLKV	124	144

Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	**ACIN1**	EEEDDEEEEGDDEGQKSREA	279	298
Q8NEJ9	Neuroguidin	**NGDN**	EEEDEAEDDQSEASGKKSVK	146	165
Q9H501	ESF1 homolog	**ESF1**	EEEDEDEEEDEDEDSEDDDKSDSGPDLARGKGN	276	308
Q9UQL6	Histone deacetylase 5	**HDAC5**	EEEDEEDDGEEEEDCIQVKDEE	587	608
O75419	Cell division control protein 45 homolog	**CDC45**	EEEDEEHSGNDSDGSEPSEK	137	156
Q9Y3S2	Zinc finger protein 330	**ZNF330**	EEEDEYEAEDDEEEEDEGRKDSDTESSDLF	270	299
O14776	Transcription elongation regulator 1	**TCERG1**	EEEDPKEEPIKEIKEEPKEEE	490	510
Q9Y5B9	FACT complex subunit SPT16	**SUPT16H**	EEEDSDEDYSSEAEESDYSKESLGSEEESGKDWDELEEEARK	958	999
A6NKF2	AT-rich interactive domain-containing protein 3C	**ARID3C**	EEEEDAEEDEEKREEAGAEEE	51	71
Q9ULV3	Cip1-interacting zinc finger protein	**CIZ1**	EEEEDDEDEEEIEVEEELCKQ	745	765
Q9UHR5	SAP30-binding protein	**SAP30BP**	EEEEDENSRQSEDDDSETEKPE	62	83
Q9NQC3	Reticulon-4	**RTN4**	EEEEEDEDEDLEELEVLERK	39	58
P46100	Transcriptional regulator ATRX	**ATRX**	EEEEEDENDDSKSPGKGRKK	1460	1479
Q6ZS11	Ras and Rab interactor-like protein	**RINL**	EEEEEDLEGKEEGREDDPEEE	233	253
Q06481	Amyloid-like protein 2	**APLP2**	EEEEEEDEEEDYDVYKSEFP	222	241
P21817	Ryanodine receptor 1	**RYR1**	EEEEEEDEEEEGEEEDEEEKEEDEEETAQEK EDEEKEEEEAAEGEKEEGLEEGLLQMK	1873	1930
Q08495	Dematin	**DMTN**	EEEEEEEDDDSGEEMKALRE	216	235
Q9ULW6	Nucleosome assembly protein 1-like 2	**NAP1L2**	EEEEEEEEEDDIEATGEENKEEEDPK	213	238
Q9Y5Q8	General transcription factor 3C polypeptide 5	**GTF3C5**	EEEEEEEEEDFKPSDGSENEME	491	512
Q15911	Zinc finger homeobox protein 3	**ZFHX3**	EEEEEEEEEEEEEEEDEGCKGLFPSE	473	498
Q4KMQ1	Taperin	**TPRN**	EEEEEEEEEEEEEGSGSEEKPFAL	609	632
Q96MU7	YTH domain-containing protein 1	**YTHDC1**	EEEEEEEEEEEYEQDERDQKEEGND	239	263
Q96JN2	Coiled-coil domain-containing protein 136	**CCDC136**	EEEEEEEEEEVEEEEEQVQKG	17	37
A6NFI3	Zinc finger protein 316	**ZNF316**	EEEEEEEEKGEEVQEVEEEEEEIV	29	52
Q9HAW4	Claspin	**CLSPN**	EEEEEEEEMTDESEEDGEEKVEKEEKEEELEEEEEKEE EEEEEGNQETAEFLLSSEEIETKDEKEMDKEN	630	699
Q8N7H5	RNA polymerase II-associated factor 1 homolog	**PAF1**	EEEEEEEMETEEKEAGGSDEE	377	397
P30414	NK-tumor recognition protein	**NKTR**	EEEEEEIDDKQVTQESKEKK	1024	1043
O60721	Sodium/potassium/calcium exchanger 1	**SLC24A1**	EEEEEEQEEEEEEEEEEEKGNEEPLSLDW	872	901
Q8IX03	Protein KIBRA	**WWC1**	EEEEEEVEEEEGEEDVFTEK	855	874
Q8TC90	Coiled-coil domain-containing glutamate-rich protein 1	**CCER1**	EEEEEKKNDDEEEYDQEVCD	277	296
Q14160	Protein scribble homolog	**SCRIB**	EEEEENRAEEEEASTEEEDKE	675	695
Q969Q1	E3 ubiquitin-protein ligase TRIM63	**TRIM63**	EEEEFIEEEDQEEEESTEGKEE	329	350
P78332	RNA-binding protein 6	**RBM6**	EEEEIKEKKPTSQGKSSSKKE	826	846
Q9NQ38	Serine protease inhibitor Kazal-type 5	**SPINK5**	EEEEKKKKEGKSRNKRQSKS	410	429
Q9P2H5	Ubiquitin carboxyl-terminal hydrolase 35	**USP35**	EEEEKVEKETEKEAEQEKEEDSLG	714	737
Q9Y6F6	Protein MRVI1	**MRVI1**	EEEEQKSESPEEPEEVEETEEEEKG	777	801
Q9NZ53	Podocalyxin-like protein 2	**PODXL2**	EEEEREKEEVEKQEEEEEEELL	169	190
Q9NVM9	Protein asunder homolog	**ASUN**	EEEERKKRGRKREDKEDKSEK	566	586
Q70CQ2	Ubiquitin carboxyl-terminal hydrolase 34	**USP34**	EEEGEEEEEEEDILSLAEEK	2489	2508
Q9P2H5	Ubiquitin carboxyl-terminal hydrolase 35	**USP35**	EEEGKEERTEKEEVGEEEES	680	699
Q9NQW8	Cyclic nucleotide-gated cation channel beta-3	**CNGB3**	EEEGKENEDKQKENEDKQKE	710	729
Q9NPG3	Ubinuclein-1	**UBN1**	EEEKDKEQRDRICSDEEEDEEK	480	501
P23327	Sarcoplasmic reticulum histidine-rich calcium-binding protein	**HRC**	EEEKEKEEDPGSHEEDDESSEQGEKG	483	508
Q5JQC4	Cancer/testis antigen 47A	**CT47A12**	EEEKEQEKEKDAENKVKNSK	267	286
P0C2W7	Cancer/testis antigen 47B	**CT47B1**	EEEKEQEKEKDVENKVKNSK	278	297
Q4G1C9	GLIPR1-like protein 2	**GLIPR1L2**	EEEKKEKEEMEMEIMEMEEEKEEREEEEEETQKEKMEEEEK	304	344
Q6ZU64	Coiled-coil domain-containing protein 108	**CCDC108**	EEELGKEEIEEKEEERDEKEEKVSW	1787	1811
Q7Z2Y5	Nik-related protein kinase	**NRK**	EEELRQVDKDKEDESSDNDE	739	758
A1A519	Protein FAM170A	**FAM170A**	EEENGNEKEEEEKPEAKEEE	281	300
Q8TBY9	WD repeat-containing protein 66	**WDR66**	EEERKTGEEEGEEEGKEDKK	52	71
Q70CQ2	Ubiquitin carboxyl-terminal hydrolase 34	**USP34**	EEESSKSSDPFSLWSTDEKEK	136	156

O60237	Protein phosphatase 1 regulatory subunit 12B	**PPP1R12B**	EEETPKSQEMEEENKESSSSSSEEEEGEDEASESETEKEADKKPEA	341	386
Q9NUY8	TBC1 domain family member 23	**TBC1D23**	EEEYDTDEIDSSSMSDDDRKE	557	577
Q12830	Nucleosome-remodeling factor subunit BPTF	**BPTF**	EEFEDQSLEKDSDDKTPDDD	595	614
Q9Y4C8	Probable RNA-binding protein 19	**RBM19**	EEGADNSSAKMEEEEEEEEEEESL	703	727
Q8N7H5	RNA polymerase II-associated factor 1 homolog	**PAF1**	EEGDRDEASDKSGSGEDESSEDEA	421	444
O60237	Protein phosphatase 1 regulatory subunit 12B	**PPP1R12B**	EEGEDEASESETEKEADKKPE	365	385
O14974	Protein phosphatase 1 regulatory subunit 12A	**PPP1R12A**	EEGKKDESSCSSEEDEEDDSESEAETDKTKPLAS	346	379
P35269	General transcription factor IIF subunit 1	**GTF2F1**	EEGPKGVDEQSDSSEESEEEK	295	315
Q2LD37	Uncharacterized protein KIAA1109	**KIAA1109**	EEGRRDDSLSSTSEDSEKDEKDEDHE	3714	3739
Q9ULH1	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 1	**ASAP1**	EEIDESDDDLDDKPSPIKKE	712	731
O75152	Zinc finger CCCH domain-containing protein 11A	**ZC3H11A**	EEIKSKKMKEKSKKQGEGSS	218	237
Q3YEC7	Rab-like protein 6	**RABL6**	EEKAAKKKSKHKKSKDKEEGKEERR	665	689
P15923	Transcription factor E2-alpha	**TCF3**	EEKEDEENTSAADHSEEEKKE	500	520
Q8IZU1	Protein FAM9A	**FAM9A**	EEKEEEEEEGEEEGGGEEGEE	231	251
P27824	Calnexin	**CANX**	EEKEEEKDKGDEEEEGEEKLEEKQKSDAEEDGGT	529	562
P14625	Endoplasmin	**HSP90B1**	EEKEESDDEAAVEEEEEEKKPKTKKVEKT	301	329
Q13061	Triadin	**TRDN**	EEKEGEKKKAEKKVTSETKKKEKEDIKKKSEKETAID	310	346
Q9BWT7	Caspase recruitment domain-containing protein 10	**CARD10**	EEKEKEKEKEKEPDNVDLVSE	262	282
Q13111	Chromatin assembly factor 1 subunit A	**CHAF1A**	EEKEKLKEEAKRAKEEAKKKKEEEKELKEKERREKREKDEKEKAEKQR	355	402
Q8N954	G patch domain-containing protein 11	**GPATCH11**	EEKEQDEDEYKSEDLSVLEK	199	218
P78559	Microtubule-associated protein 1A	**MAP1A**	EEKKDKEKKEIKKERKELKKDEGRKEEKKDAKKEEKRKDTKPELKKISKPDLK	413	465
Q58FF7	Putative heat shock protein HSP 90-beta-3	**HSP90AB3P**	EEKPKIKDVGSDEEDDSKEY	224	243
P49756	RNA-binding protein 25	**RBM25**	EEKQEKEEKREEPMEEEEEPE	588	608
O15318	DNA-directed RNA polymerase III subunit RPC7	**POLR3G**	EELEKRGDGEKSDEENEEKE	146	165
Q03164	Histone-lysine N-methyltransferase 2A	**KMT2A**	EELSKDRDADKSVEKDKSRE	856	875
A6NGS2	Glutamate-rich protein 4	**ERICH4**	EELVTILEEEEESSKEEEED	84	103
Q7L014	Probable ATP-dependent RNA helicase DDX46	**DDX46**	EEMKQGKKWSLEDDDDDEDD	190	209
O00472	RNA polymerase II elongation factor ELL2	**ELL2**	EENHSMSHKKSKKKSKKHKEKD	452	473
Q9P1Z9	Coiled-coil domain-containing protein 180	**CCDC180**	EENVKGQGEKKEESEEEDEKEEEEEEEKLEEEKEEKE AQEEQESLSVGEEEDKEEGLEEI	893	952
Q58FF3	Putative endoplasmin-like protein	**HSP90B2P**	EEPDEEPEETAEDKEQDKDKE	356	376
Q5SWA1	Protein phosphatase 1 regulatory subunit 15B	**PPP1R15B**	EEPSDSEKDLSGKSDLENSS	505	524
P10645	Chromogranin-A	**CHGA**	EEPSSKDVMEKREDSKEAEKS	122	142
Q9NRZ9	Lymphoid-specific helicase	**HELLS**	EEQKKKEKLERKKESLKVKK	92	111
Q9NW75	G patch domain-containing protein 2	**GPATCH2**	EEQKVSDELMSESDSSSLSS	246	265
P78316	Nucleolar protein 14	**NOP14**	EEQSKEASDPESNEEEGDSS	346	365
Q9BVP2	Guanine nucleotide-binding protein-like 3	**GNL3**	EEREDDKDSDQETVDEEVDE	482	501
Q6ZQQ6	WD repeat-containing protein 87	**WDR87**	EERKEEEEGEEKQVEKEEEEKKKKKKEKKKEEVQEKEEVFE EKEEIMSEEETESLSDEEEEEESCSLEEEVDREKEILKKEKQFKL	2267	2352
Q96M86	Dynein heavy chain domain-containing protein 1	**DNHD1**	EERKNEQEKEQEENEEKEEEK	3624	3644
O43719	HIV Tat-specific factor 1	**HTATSF1**	EESDDKEDEDADGKEVEDADEKLFEDDDSNEKL	674	706
Q70Z53	Protein FRA10AC1	**FRA10AC1**	EESSHKKSRLSSAEEASKKKDK	241	262
Q8IZU1	Protein FAM9A	**FAM9A**	EETEEEEEEEEEEEEEQIKAF	259	280
Q5JRA6	Melanoma inhibitory activity protein 3	**MIA3**	EETRDTMDLESSSSEEEKEDDDDA	395	418
Q92665	28S ribosomal protein S31, mitochondrial	**MRPS31**	EETSKETSESQDSEKENTKKD	78	98
Q00722	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-2	**PLCB2**	EEVEEEEEEESGNLDEEEIKK	508	528
Q8IZU0	Protein FAM9B	**FAM9B**	EEYITDEQKEEEEEEGEEEE	110	129
O75155	Cullin-associated NEDD8-dissociated protein 2	**CAND2**	EFSEQESEDEYSDDDDMSWK	331	350
Q12789	General transcription factor 3C polypeptide 1	**GTF3C1**	EGEDTFLSESDSEEERSSSK	469	488
O95218	Zinc finger Ran-binding domain-containing protein 2	**ZRANB2**	EGEEEDEDEDLSKYKLDEDEDEDDADLSKY	154	183
P20962	Parathymosin	**PTMS**	EGEEEDEEEEEEDDEGPALK	61	80
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	EGEPKEEEEEEAKEEKQERKK	637	657

Q5W0A0	Glutamate-rich protein 6B	**ERICH6B**	EGESLEDKEYLEEEEDLEEEE	52	72
A6NDY0	Embryonic polyadenylate-binding protein 2	**PABPN1L**	EGGEGKEEKEEEEDAEEDQD	43	62
Q6IPX3	Transcription elongation factor A protein-like 6	**TCEAL6**	EGKPEDEVEPDDEGKSDEEEK	15	35
Q969E4	Transcription elongation factor A protein-like 3	**TCEAL3**	EGKPEDEVEPDDEGKSDEEEK	15	35
Q15413	Ryanodine receptor 3	**RYR3**	EGKVESEKADMEDGEKEDKDKEEEQ	4337	4361
Q5MJ09	Sperm protein associated with the nucleus on the X chromosome N3	**SPANXN3**	EGPSKEDKDLDSSEGSSQEDED	106	127
Q9H1I8	Activating signal cointegrator 1 complex subunit 2	**ASCC2**	EGQEEDDDDEEDDADEEAPK	655	674
Q4UJ75	Ankyrin repeat domain-containing protein 20A4	**ANKRD20A4**	EHKKKILKKEKSDVGSSDES	253	272
Q5SQ80	Ankyrin repeat domain-containing protein 20A2	**ANKRD20A2**	EHKKKILKKEKSDVGSSDES	253	272
Q5TYW2	Ankyrin repeat domain-containing protein 20A1	**ANKRD20A1**	EHKKKILKKEKSDVGSSDES	253	272
Q5VUR7	Ankyrin repeat domain-containing protein 20A3	**ANKRD20A3**	EHKKKILKKEKSDVGSSDES	253	272
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	**MPHOSPH10**	EIEADDKEDLEDLEEEEVSD	121	140
Q9Y6V0	Protein piccolo	**PCLO**	EIEQQQRKSSSKKSKKDKDE	1725	1744
P55209	Nucleosome assembly protein 1-like 1	**NAP1L1**	EISEELKEKAKIEDEKKDEEKEDPKG	141	166
Q9H4A5	Golgi phosphoprotein 3-like	**GOLPH3L**	EISKNSEKKMESEEDSNWEKS	12	32
Q9BXY0	Protein MAK16 homolog	**MAK16**	EKDDDDDDEEDVGKREFVED	204	223
Q9UKE5	TRAF2 and NCK-interacting protein kinase	**TNIK**	EKDETEYEYSGSEEEEEENDS	315	335
Q5T8P6	RNA-binding protein 26	**RBM26**	EKDIKKEEITKEEEREKKFS	102	121
Q13547	Histone deacetylase 1	**HDAC1**	EKDPEEKKEVTEEEKTKEEKPE	450	471
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	EKDRSKEIDEKRKKDKKSRT	344	363
O14617	AP-3 complex subunit delta-1	**AP3D1**	EKDVPMVEKKSKKPKKKEKK	831	850
Q9HCF6	Transient receptor potential cation channel subfamily M member 3	**TRPM3**	EKEAEEPEKPTKEKEEEDME	834	853
Q9Y618	Nuclear receptor corepressor 2	**NCOR2**	EKEAEKEEEKPEVENDKEDL	526	545
P25054	Adenomatous polyposis coli protein	**APC**	EKEAEKTIDSEKDLLDDSDDDD	1550	1571
Q9H1E5	Thioredoxin-related transmembrane protein 4	**TMX4**	EKEDLGDEDEAEEEEEEDNL	266	285
Q9UL68	Myelin transcription factor 1-like protein	**MYT1L**	EKEEDEGEEYSEDNDEPGDEDEEDEEGDREEEEEIEEED EDDDEDGEDVEDEEEEEEEEEEEEEEEEEEENEDHQMNCHNT	101	178
O60663	LIM homeobox transcription factor 1-beta	**LMX1B**	EKEKDLLSSVSPDESDSVKSEDEDGD	170	195
O94880	PHD finger protein 14	**PHF14**	EKEKEKEREKEKEKATVSEN	124	143
Q99460	26S proteasome non-ATPase regulatory subunit 1	**PSMD1**	EKEKKEEEKMEVDEAEKKEEKEKKKEPEPNFQ	845	876
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	EKEKKIKHEHKSEKDKLDLSE	842	862
Q99442	Translocation protein SEC62	**SEC62**	EKEKKKDGEKEESKKEETPGT	138	158
P18583	Protein SON	**SON**	EKEKKRDSSLRSRSKRSKSSE	1813	1833
Q13061	Triadin	**TRDN**	EKEKPEKKATHKEKIEKKEK	176	195
O14974	Protein phosphatase 1 regulatory subunit 12A	**PPP1R12A**	EKEKQDKEKQEEKKESETSREDEYKQKYS	731	759
P11055	Myosin-3	**MYH3**	EKEKSEFKLEIDDLSSSMES	1219	1238
Q58FF8	Putative heat shock protein HSP 90-beta 2	**HSP90AB2P**	EKEREKEISDGKAEEEKGEKEEENKDDEEKPKIEDV	140	175
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	EKEREKEKREREKREKEKEKHK	720	741
Q9Y520	Protein PRRC2C	**PRRC2C**	EKERKQEKEKELERQKEKEKE	556	576
Q8IYW2	Protein CFAP46	**CFAP46**	EKERSKEKENERSKEKDKEKGKEEKVKEPKQSQS	1374	1407
Q15459	Splicing factor 3A subunit 1	**SF3A1**	EKFGESEEVEMEVESDEEDDK	315	335
P46100	Transcriptional regulator ATRX	**ATRX**	EKFLKKDQSDETSEDDKKQSKK	966	987
Q92736	Ryanodine receptor 2	**RYR2**	EKFQEQKAKEEEKEEKEETKSE	4417	4438
Q9BZI7	Regulator of nonsense transcripts 3B	**UPF3B**	EKGDEKELDKREKAKKLDKE	288	307
Q8TEW0	Partitioning defective 3 homolog	**PARD3**	EKGDKTDRKKDKTGKEKKKD	983	1002
P46821	Microtubule-associated protein 1B	**MAP1B**	EKGEAEQSEEEADEEDKAED	1009	1028

Q6H9L7	Isthmin-2	**ISM2**	EKGEEKEEDEDYPSEDIEGED	264	284
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	EKGKNKEKDRELDKKEKSRDKESI	947	970
O14511	Pro-neuregulin-2, membrane-bound isoform	**NRG2**	EKGRCSSYSDSSSSSSERSSSSSSSSSESGSSSRSSSNNSSISR	15	58
O94913	Pre-mRNA cleavage complex 2 protein Pcf11	**PCF11**	EKITKKELDQLDSKSKSKSKS	350	370
Q9Y2J2	Band 4.1-like protein 3	**EPB41L3**	EKKAEEERDEEEDKRRKGEE	471	490
Q9Y2X3	Nucleolar protein 58	**NOP58**	EKKAKKAKIKVKVEEEEEEK	456	475
P12036	Neurofilament heavy polypeptide	**NEFH**	EKKDTKEEKAKKPEEKPKTE	959	978
O94880	PHD finger protein 14	**PHF14**	EKKEEEENGERPRKKKEKEKEKEKEKEKEREKEKEKATVSENV	100	144
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	EKKHAAEDKAKSKHKEKSDKE	1243	1263
Q8TA86	Retinitis pigmentosa 9 protein	**RP9**	EKKKKKKRKHKSSKSNEGSDSE	200	221
P26038	Moesin	**MSN**	EKKKREMAEKEKEKIEREKEE	326	346
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	EKKKVKDKDRDRDRDKDRDKKKSHS	1503	1527
P29973	cGMP-gated cation channel alpha-1	**CNGA1**	EKKSKSDDKNENKNDPEKKKKKKDKEKKKKE EKSKDKKEEEKKEVVVIDPSGNTYY	111	166
Q9UQR1	Zinc finger protein 148	**ZNF148**	EKKSSGMDKESALDKSDLKKDK	321	342
Q8IYB3	Serine/arginine repetitive matrix protein 1	**SRRM1**	EKLASMKKQDEDKDKRDKEEKESSREKRE	139	167
Q9UIF9	Bromodomain adjacent to zinc finger domain protein 2A	**BAZ2A**	EKLKEKVKREKKEKVKMKEKEE	766	787
Q5VZ18	SH2 domain-containing adapter protein E	**SHE**	EKNGKSNYPSSSSSSSSSSS	154	174
Q8IXW5	Putative RNA polymerase II subunit B1 CTD phosphatase RPAP2	**RPAP2**	EKQYESSSSSTHSDSSSDNE	201	220
Q8WVS4	WD repeat-containing protein 60	**WDR60**	EKREKYSKEKSNSFSDKGEE	235	254
Q8IWZ3	Ankyrin repeat and KH domain-containing protein 1	**ANKHD1**	EKRKEKRKKKKEEQKRKQEEDEENK	1454	1478
Q9ULU4	Protein kinase C-binding protein 1	**ZMYND8**	EKSDSSDSEYISDDEQKSKNEPEDTEDKEG	593	622
Q9NPI7	Lysine-rich coiled-coil protein 1	**KRCC1**	EKSEEERSKHKRKKSCEEID	169	188
P18583	Protein SON	**SON**	EKSKKNKNRDKGEKEKKRDSS	1801	1821
Q5VUA4	Zinc finger protein 318	**ZNF318**	EKSKSPEKVSSFSNSSSNKESK	1006	1027
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	EKSPMKKKQKKKENKENKEK	1384	1403
Q96MT7	Cilia- and flagella-associated protein 44	**CFAP44**	EKSVTSKSDGKKSLRSSKSES	12	32
Q9NWH9	SAFB-like transcription modulator	**SLTM**	EKVKGDPSKKEMKKENDEKSSS	458	479
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	ELDSGGEEDEKESLSASSSSS	1028	1048
Q08AE8	Protein spire homolog 1	**SPIRE1**	ELDSSESEEETLHKSTSSSS	461	480
Q9H116	GDNF-inducible zinc finger protein 1	**GZF1**	ELEELSKKAGPEEEEEEEEEDEEGEKKKSN	287	316
Q7Z6E9	E3 ubiquitin-protein ligase RBBP6	**RBBP6**	ELEKSQKHKHKKKKSKKNKDKEKEKEKDDQKVKSVT	1756	1791
Q13111	Chromatin assembly factor 1 subunit A	**CHAF1A**	ELKEKERREKREKDEKEKAEK	380	400
O75971	snRNA-activating protein complex subunit 5	**SNAPC5**	ELSTKSHVTEEEEEEEEEESDS	77	98
Q8NEW7	Transmembrane inner ear expressed protein	**TMIE**	ELTEVPGEDKKKKKKKKKDS	114	133
Q96MC2	Dynein regulatory complex protein 1	**DRC1**	EMEGEKEESLVEGEKEEEEE	593	612
Q96MT7	Cilia- and flagella-associated protein 44	**CFAP44**	EMGEDGEKEFQEEEEEKEEEEEEEEP	713	738
Q8N5I9	Uncharacterized protein C12orf45	**C12orf45**	EMNQSDSKEVDSSEESSQDSSE	119	140
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	ENDEEKEKEAKKTEEEVPEEE	286	306
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	ENDSSNDDKTKKSEGDGDKEEKKEDSEKEAKKSSKKRNRKHS	316	357
Q92922	SWI/SNF complex subunit SMARCC1	**SMARCC1**	ENEKNSEKEQDSEVSEDTKSEEKETEENKE	811	840
Q92954	Proteoglycan 4	**PRG4**	ENQESSSSSSSSSSSTIRK	162	181
O75592	E3 ubiquitin-protein ligase MYCBP2	**MYCBP2**	EPEEEEDEENKTSKENSEQEKD	3569	3590
P0CB47	Putative upstream-binding factor 1-like protein 1	**UBTFL1**	EPEENRKKDREKEESSNSSD	363	382
Q15653	NF-kappa-B inhibitor beta	**NFKBIB**	EPEGEDEKSGPCSSSSDSDS	300	319
P47710	Alpha-S1-casein	**CSN1S1**	EPEKMESSISSSSEEMSLSK	79	98
A1YPR0	Zinc finger and BTB domain-containing protein 7C	**ZBTB7C**	EPGGDGGEEDDKEDDDDDEDDDDEEDEEEEE EEEEDDDDDTEDFADQENLPDP	128	180
O15050	TPR and ankyrin repeat-containing protein 1	**TRANK1**	EPGKESPGGEEEEEEEDEEEEDSIE	1078	1102
Q13569	G/T mismatch-specific thymine DNA glycosylase	**TDG**	EPKKPVESKKSGKSAKSKEK	75	94

Q8TBF4	Zinc finger CCHC-type and RNA-binding motif-containing protein 1	**ZCRB1**	EPPKKKEKKKKKKAPEPEEE	129	148
Q99081	Transcription factor 12	**TCF12**	EPPSSDDMKSDDESSQKDIK	531	550
Q13435	Splicing factor 3B subunit 2	**SF3B2**	EPSDEESSEEEEEEESDEDKPDETGFIT	712	739
O95714	E3 ubiquitin-protein ligase HERC2	**HERC2**	EPSGTKKEDLNDKEKKDEEE	67	86
Q8IW19	Aprataxin and PNK-like factor	**APLF**	EPTDEDSDWEPGKEDEEKED	477	496
Q9ULW6	Nucleosome assembly protein 1-like 2	**NAP1L2**	EPTEEECEYKSDSEDCDDEE	160	179
A1L162	Glutamate-rich protein 2	**ERICH2**	EQDGENSDEDSSGESKGESDEE	122	143
Q15751	Probable E3 ubiquitin-protein ligase HERC1	**HERC1**	EQEDGEEEEKKVDSSGETEKKD	1868	1889
P41236	Protein phosphatase inhibitor 2	**PPP1R2**	EQESSGEEDSDLSPEEREKK	118	137
Q6NXS1	Protein phosphatase inhibitor 2-like protein 3	**PPP1R2P3**	EQESSGEEDSDLSPEEREKK	118	137
Q96PQ5	Putative protein phosphatase inhibitor 2-like protein 1	**PPP1R2P1**	EQESSGEEDSDLSPEEREKK	118	137
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein	**EIF3CL**	EQSADEDAEKNEEDSEGSSDEDEDEDGVS	164	192
Q99613	Eukaryotic translation initiation factor 3 subunit C	**EIF3C**	EQSADEDAEKNEEDSEGSSDEDEDEDGVS	164	192
O15042	U2 snRNP-associated SURP motif-containing protein	**U2SURP**	ERDKKDKEKLESRSKDKKEKDECTP	896	920
Q9UPQ9	Trinucleotide repeat-containing gene 6B protein	**TNRC6B**	EREEQLMEDKKRKKEDKKKKE	7	27
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	EREKEHEKDRDKEKEKEQDKEKEREKDRSKEIDEKRKKDKKSRTP	320	364
Q5VTR2	E3 ubiquitin-protein ligase BRE1A	**RNF20**	EREKEKERE REKQKLKESEKE	583	603
P49756	RNA-binding protein 25	**RBM25**	ERERE REKDKKRDREEDEED	420	439
Q8TBK6	Zinc finger CCHC domain-containing protein 10	**ZCCHC10**	ERKAKKKRSKSVTSSSSSSSDSSASDSSSESEETSTS SSSEDSDTDESSSSSSSSASSTTSSSSSDSDSDSSSS SSSSTSTDSSSDDEPPKKKKKK	97	192
Q8TF01	Arginine/serine-rich protein PNISR	**PNISR**	ERKKERSRSIDKDRKKKDKE	664	683
Q8IUX7	Adipocyte enhancer-binding protein 1	**AEBP1**	ERQTDEEKEELKKPKKEDSS	329	348
Q8WUA2	Peptidyl-prolyl cis-trans isomerase-like 4	**PPIL4**	ERSKKRDRSRSPKKSKDKEKSK	469	490
Q9UPT8	Zinc finger CCCH domain-containing protein 4	**ZC3H4**	ERSRKEKGEKHHSDSDEEKS	80	99
O15234	Protein CASC3	**CASC3**	ESAEDSEGEEGEYSEEENSK	93	112
P24588	A-kinase anchor protein 5	**AKAP5**	ESDFKENGITEEKSKSEESK	314	333
Q9P243	Zinc finger protein ZFAT	**ZFAT**	ESDLELEKKCKEDDREKASK	152	171
Q96GN5	Cell division cycle-associated 7-like protein	**CDCA7L**	ESDLSDDGKASLVSEEEEDEEEDKA	104	128
Q13367	AP-3 complex subunit beta-2	**AP3B2**	ESDNEDQDEDEEKGRGSESE	717	736
O95149	Snurportin-1	**SNUPN**	ESEEENKKDDEEMDIDTVKK	74	93
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	**MPHOSPH10**	ESEEQEREEDGSEIEADDKED	109	129
Q14093	Cylicin-2	**CYLC2**	ESEGEKGGTEKDSKKGKKDSKK	205	226
Q5H9L2	Transcription elongation factor A protein-like 5	**TCEAL5**	ESEGKPEDEGSTEDEGKSDEEEKPD	19	43
Q9UHQ7	WW domain-binding protein 5	**WBP5**	ESEPKHEEEPKPEEKPEEEEK	14	34
P35659	Protein DEK	**DEK**	ESESEDSSDDEPLIKKLKKP	300	319
Q8N9E0	Protein FAM133A	**FAM133A**	ESESESKESVKKKKKSKDETEKEKDVRSLSKK RKKSYPDDKPLSSESSSESDYEEDVQAKKKRR	148	211
Q56P03	E2F-associated phosphoprotein	**EAPP**	ESESSSEDEFEKEMEAELNS	52	71
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	ESGTEDKKGSDSKESEGKTE	2736	2755
Q0VAK6	Leiomodin-3	**LMOD3**	ESKGSSNIQETDEEDEEEEDDDDDDEGEDDGEESEETN	133	170
Q86UW6	NEDD4-binding protein 2	**N4BP2**	ESKIMEKRPEEESEDSKMDS	113	132
Q99549	M-phase phosphoprotein 8	**MPHOSPH8**	ESKKPKKDEVKETKELKKVKK	212	232
Q96JM4	Leucine-rich repeat and IQ domain-containing protein 1	**LRRIQ1**	ESKTSRKSLLKSEKEKKISEE	1525	1545
Q6ZQQ6	WD repeat-containing protein 87	**WDR87**	ESLLDELEKQESLSSEEEEE	2231	2250
Q8N841	Tubulin polyglutamylase TTLL6	**TTLL6**	ESQEGENSEEKGDSSKEDPKE	8	28
Q5T8A7	Protein phosphatase 1 regulatory subunit 26	**PPP1R26**	ESSEDKSSSLDSDEDLDTAIKD	678	699

Q13316	Dentin matrix acidic phosphoprotein 1	**DMP1**	ESSESSEGSKVSSEEQANED	46	65
Q5VT06	Centrosome-associated protein 350	**CEP350**	ESSKKSEIKEIEYTKLKKSK	2251	2270
Q16594	Transcription initiation factor TFIID subunit 9	**TAF9**	ESSNALKRKREDDDDDDDDDDD	239	260
Q8IUR6	CREB3 regulatory factor	**CREBRF**	ESSSLSASTSVSDSSQKKEE	315	334
O60885	Bromodomain-containing protein 4	**BRD4**	ESSSSDSEDSETEMAPKSKKKG	708	729
Q9UQ35	Serine/arginine repetitive matrix protein 2	**SRRM2**	ESSSSRSPTPKQKKKKKKKD	176	195
P43243	Matrin-3	**MATR3**	ESSTEGKEQEEKSGEDGEKD	619	638
P35663	Cylicin-1	**CYLC1**	ESTEMESDLELKKDKKHSKEKK	482	503
Q9NW13	RNA-binding protein 28	**RBM28**	ESVKKKGREEEDMEEEENDDDDDDDDEEDGVFDDEDEEEENIESKV	216	261
Q8IVG5	Sterile alpha motif domain-containing protein 9-like	**SAMD9L**	ETDIEKDKSKFLEFKKSKEE	666	685
Q5VZP5	Inactive dual specificity phosphatase 27	**DUSP27**	ETESKSSSYKFSKSQSEEQD	970	989
Q8NDB2	B-cell scaffold protein with ankyrin repeats	**BANK1**	ETGDEPKGEKEKKEEEKEQEEEEDP	553	577
Q8NI27	THO complex subunit 2	**THOC2**	ETKERTPKSDKEKEKFKKEEK	1316	1336
Q9UIF9	Bromodomain adjacent to zinc finger domain protein 2A	**BAZ2A**	ETKSLKQKEAKKKSKAEKEK	741	760
Q96EV2	RNA-binding protein 33	**RBM33**	ETLELQKDIKEESDEEEEDDEESG	193	216
P55198	Protein AF-17	**MLLT6**	ETSESSRESKGKKSSSHSLS	290	309
Q9ULG3	Uncharacterized protein KIAA1257	**KIAA1257**	ETSSKNSEEYEKSLKMDDSS	277	296
Q01804	OTU domain-containing protein 4	**OTUD4**	EVADEDNSEISDSEDDSCKSK	192	212
Q8TC90	Coiled-coil domain-containing glutamate-rich protein 1	**CCER1**	EVCDAKEASEEEEEVEDEEEE	293	313
Q92793	CREB-binding protein	**CREBBP**	EVDEKKPEVKVEVKEEEESSS	1047	1067
Q15784	Neurogenic differentiation factor 2	**NEUROD2**	EVKEEGELGGEEEEEEEEEE	72	91
P46821	Microtubule-associated protein 1B	**MAP1B**	EVKKEEKKEVKKEEKEPKKEIKKLPKDAKKSST	712	744
B2RXH8	Heterogeneous nuclear ribonucleoprotein C-like 2	**HNRNPCL2**	EVKNAKSEEEQSSSSMKKDE	214	233
O60812	Heterogeneous nuclear ribonucleoprotein C-like 1	**HNRNPCL1**	EVKNAKSEEEQSSSSMKKDE	214	233
B7ZW38	Heterogeneous nuclear ribonucleoprotein C-like 3	**HNRNPCL3**	EVKNAKSEEEQTSSSSKKDK	214	233
P0DMR1	Heterogeneous nuclear ribonucleoprotein C-like 4	**HNRNPCL4**	EVKNAKSEEEQTSSSSKKDK	214	233
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	EVKNEEEDQKEDEEDQNEEK	122	141
A8MV65	Transcription cofactor vestigial-like protein 3	**VGLL3**	EVTLPSKQEEEDEEEEEEEKD	56	76
Q9NYI0	PH and SEC7 domain-containing protein 3	**PSD3**	EWAVDDEEKKKSPSESTEEK	737	756
Q15291	Retinoblastoma-binding protein 5	**RBBP5**	EYEERESEFDIEDEDKSEPE	344	363
Q92771	Putative ATP-dependent RNA helicase DDX12	**DDX12P**	EYESDEEKKVASGVDEDEDD	220	239
Q96FC9	Probable ATP-dependent RNA helicase DDX11	**DDX11**	EYESDEEKKVASRVDEDEDD	201	220
Q8WWF3	Serine-rich single-pass membrane protein 1	**SSMEM1**	EYGSEESNSEASSWKESESE	149	168
O94874	E3 UFM1-protein ligase 1	**UFL1**	EYKIKKVKKKGRKDDDSDDES	442	462
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	EYSTEKHKKHKKEKKKVKDKD	1491	1511
Q92541	RNA polymerase-associated protein RTF1 homolog	**RTF1**	FEIKKKLKTAKKKEKKEKKKK	222	242
Q9NVU7	Protein SDA1 homolog	**SDAD1**	FFLGKDEDEKQDSDSESEDD	220	239
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	FGKRSDKFCSSESESESSES	469	488
P25205	DNA replication licensing factor MCM3	**MCM3**	FKKVLEKEKKRKKRSEDESE	654	673
P0CJ78	Zinc finger protein 865	**ZNF865**	FKPKAEVPSSSSSSSSSSSSSSSSSSSSSSSQAKKPDPP	85	125
P16383	GC-rich sequence DNA-binding factor 2	**GCFC2**	FKSVEEFMDSSVEDSKKESSSDKKV	525	549
Q8IWT3	Cullin-9	**CUL9**	FLEQEDEEEKRLEEEEEEEEEEAEKEL	1659	1686
Q5TD94	Radial spoke head protein 4 homolog A	**RSPH4A**	FNSIQKNEEEEEEEDEEKDDSD	560	581
Q5TAX3	Terminal uridylyltransferase 4	**ZCCHC11**	FRLKKKDSEEEKEGNEEEKDS	1316	1336
Q8IV63	Inactive serine/threonine-protein kinase VRK3	**VRK3**	FSDGDSSESEDTLSSSERSK	77	96
Q96GW7	Brevican core protein	**BCAN**	FSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALW	445	480

Q15911	Zinc finger homeobox protein 3	**ZFHX3**	FSEKVEPAEEEAEEEEEEE	452	471
Q99250	Sodium channel protein type 2 subunit alpha	**SCN2A**	FSESSSVASKLSSKSEKELK	483	502
Q9NV56	MRG/MORF4L-binding protein	**MRGBP**	FSSSGSLGKASEKSSKDKEK	144	163
Q12802	A-kinase anchor protein 13	**AKAP13**	FSYIKNKMSSSKKSKEKEKEKDKIKEKEKDSKDKEKDKKTVNGH	1749	1792
Q14568	Putative heat shock protein HSP 90-alpha A2	**HSP90AA2**	FVEKECDKEVSDDETEEKEDKEEEKEKEEKESKDKPEIE DVGSDEEEEKKDGDKKKKKTKEKYIDQEELNKTKP	221	294
P07900	Heat shock protein HSP 90-alpha	**HSP90AA1**	FVEKERDKEVSDDEAEEKEDKEEEKEKEEKESEDKPEI EDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPI	221	296
Q8WU90	Zinc finger CCCH domain-containing protein 15	**ZC3H15**	FVLKKDKKKEEKEDEISLED	215	234
P20020	Plasma membrane calcium-transporting ATPase 1	**ATP2B1**	GAGGEEEEKKDEKKKEKKNKK	292	312
Q13316	Dentin matrix acidic phosphoprotein 1	**DMP1**	GDDKDDDEDDSGDDTFGDDDS	99	119
Q7L0X2	Glutamate-rich protein 6	**ERICH6**	GDPGKKDQKESEEELEEEEEEEEVEEEEEEVE EEEEEVEEEEEEVVEEELVGEE	11	64
Q13428	Treacle protein	**TCOF1**	GDQSNPKSKKEKKKSDKRKKDKEKKEKKKKA KKASTKDSESPSQKKKKKKKKTAEQT	1431	1487
P37275	Zinc finger E-box-binding homeobox 1	**ZEB1**	GDSDERESLTREEDEDSEKEEEEEDKEMEELQEEKECEKP	1020	1059
Q03001	Dystonin	**DST**	GDSLKRLEEEEKSLEEEKKE	1422	1441
Q9H7S9	Zinc finger protein 703	**ZNF703**	GDSRKDSGSSSVSSTSSSSSSS	164	185
Q9H9B1	Histone-lysine N-methyltransferase EHMT1	**EHMT1**	GEEEEGGDESDLSSESSIKKK	405	425
Q9P2H5	Ubiquitin carboxyl-terminal hydrolase 35	**USP35**	GEGEREKEEEVEEEEEKVEKE	702	722
Q05DH4	Protein FAM160A1	**FAM160A1**	GEKEKEGKKELEDEEDDFDSF	794	814
O00203	AP-3 complex subunit beta-1	**AP3B1**	GEKENEKSKTSDSSNDESSSIEDSSSDSESE SEPESESESRRVTKEKEKKTKQDRT	755	810
Q8WYB5	Histone acetyltransferase KAT6B	**KAT6B**	GEKPEDDLIKPEEEEEEEEEEEEEEEEEEGEEEEGGG	1340	1376
Q15746	Myosin light chain kinase, smooth muscle	**MYLK**	GEKPEEPKDEVEVSDDDEKE	1425	1444
Q7Z7B0	Filamin-A-interacting protein 1	**FILIP1**	GEKSLSEDAKKKKKSNRKEDD	28	48
Q96I76	G patch domain-containing protein 3	**GPATCH3**	GEPEEEVGKEEEEESHSDEDDD	274	295
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	GERKKEKKEKTEKKFDHSKKSEDT	332	355
A1L162	Glutamate-rich protein 2	**ERICH2**	GESKGESDEELSDESSDEGEDGS	134	156
Q08945	FACT complex subunit SSRP1	**SSRP1**	GESSKRDKSKKKKKVKVKMEKKST	619	642
O15066	Kinesin-like protein KIF3B	**KIF3B**	GGEEEEEEGEEGEEEGDDKDD	392	412
P01880	Ig delta chain C region	**IGHD**	GGEEKKKEKEKEEQEERETK	138	157
Q6SPF0	Atherin	**SAMD1**	GGEERVLEKEEEEDDDEDEDEEDDVSEGSE	320	349
Q15910	Histone-lysine N-methyltransferase EZH2	**EZH2**	GGENNDKEEEEKKDETSSSSE	389	409
Q8IXZ2	Zinc finger CCCH domain-containing protein 3	**ZC3H3**	GGSASPSSSKASSSSSSSS	862	881
Q8N7H5	RNA polymerase II-associated factor 1 homolog	**PAF1**	GGSDEEQEKGSSSEKEGSEDE	392	412
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	GIISEDFSEKKDDKKGKEKS	1093	1112
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	GKAAAKASESSSSEESSDDDDEEDQKKQ	117	144
Q9H2P0	Activity-dependent neuroprotector homeobox protein	**ADNP**	GKEDDSSSDSFENLEEESNES	869	889
Q13185	Chromobox protein homolog 3	**CBX3**	GKEKDGTKRKSLSDSESDDSKSKKKRDA	83	110
P78559	Microtubule-associated protein 1A	**MAP1A**	GKKHLKEKISKLEEKKDKEKKE	401	422
P51825	AF4/FMR2 family member 1	**AFF1**	GKKHSSEKRSSDSSSKLAKK	794	813
Q15424	Scaffold attachment factor B1	**SAFB**	GKKTSDKRDSDGKKEKSSNSD	488	508
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	GKNKDRDKEREKDREKDKEKD	296	316
P60006	Anaphase-promoting complex subunit 15	**ANAPC15**	GKPASEHYDDEEEEDDEDDEDSEEDSEDDEDMQDMDEM	55	92
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	GKRKKQDSSDEDDDDDEAPK	200	219
Q96M83	Coiled-coil domain-containing protein 7	**CCDC7**	GKSEDSEKKMSPEKEFKIKED	353	373
Q76FK4	Nucleolar protein 8	**NOL8**	GKTSGKLFDSSDDDESDSEDDS	828	849
Q92688	Acidic leucine-rich nuclear phosphoprotein 32 family member B	**ANP32B**	GLDEEDEDEDEDEEEEEGGKGEKRK	217	241
P23327	Sarcoplasmic reticulum histidine-rich calcium-binding protein	**HRC**	GLSLNQEEEEEEDKEEEEEEEDEER	527	551

Q6IN85	Serine/threonine-protein phosphatase 4 regulatory subunit 3A	**SMEK1**	GLVDYPDDDEDDDEDEDKED	802	821
Q9NZW4	Dentin sialophosphoprotein	**DSPP**	GNDDNDKSDSGKGKSDSSDSDSSDSSNSSDSSDSSDS DSSDSNSSSDSDSSDSDSSDSSDSDSSDSSNSSDSSDSS DSSDSSDSSDSSDSKSDSSKSESDSSDSDSKSDSSDSNS SDSSDNSDSSDSSNSSNSSDSSDSSDSSDSSSSSDSSNS SDSSDSSDSSNSSESSDSSDSSDSDSSDSSDSSNSNSSD SDSSNSSDSSDSSNSSDSSDSSDSSNSSDSSDSSDSSNS SDSSDSSDSSDSSDSSNSSDSNDSSNSSDSSDSSNSSDS SNSSDSSDSSDSSDSDSSNSSDSSNSSDSSDSSNSSDSS DSSDSSDGSDSDSSNRSDSSNSSDSSDSSDSSNSSDSS DSSDSNESSNSSDSSDSSNSSDSDSSDSSNSSDSSDSSN SSDSSESSNSSDNSNSSDSSNSSDSSDSSDSSNSSDSSN SSDSSNSSDSSDSNSSDSSDSSNSSDSSDSSDSSDSSDS SDSSNSSDSSDSSDSSDSSNSSDSSNSSDSSNSSDSSDS SDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSDSSDS SDSSDSSESSDSSDSSNSSDSSDSSDSSDSSDSSDSSDSS DSSDSSNSSDSSDSSDSSDSSDSSNSSDSSDSSESSDSS DSSDSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSS DSSDSSDSSDSSDSSDSSDSSDSSDSSDSSDSSDSSDSN ESSDSSDSSDSSDSSNSSDSSDSSDSSDSTSDSNDESDS QSKSGNGNNNGSDSDSDSEGSDSNHSTSDD	533	1301
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	GPKKKQKKKAKNKEDSEKDEK	636	656
Q8WYB5	Histone acetyltransferase KAT6B	**KAT6B**	GQLLELSKESSEEEEEEEDEEEEEEEEEEEDEEEE EEEEEEEEEENIQSSPPRLTKP	1059	1116
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	GRDKKEKVFEKHKEKKDKES	1194	1213
P35269	General transcription factor IIF subunit 1	**GTF2F1**	GRKKKKKKGSDDEAFEDSDD	244	263
Q9NQV6	PR domain zinc finger protein 10	**PRDM10**	GRLEAKEEEDEDEDEDTEEDEEEDGEDTD	128	156
Q9H0M4	Zinc finger CW-type PWWP domain protein 1	**ZCWPW1**	GSCLEKKEKEEELEKEEGEK	450	469
Q8WVC0	RNA polymerase-associated protein LEO1	**LEO1**	GSDADSEAERKDSDSGSDSDSD	9	30
O95232	Luc7-like protein 3	**LUC7L3**	GSDDKKSSVKSGSREKQSED	383	402
Q6PKG0	La-related protein 1	**LARP1**	GSDSKESPKTKSDESGEEKN	214	233
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	GSEEDEEEKLEDDDKSEESSQPEAG	2236	2260
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	GSEEDSSSSEDSDDSSSEVKRKKHKDEDWQMSGS	121	154
Q8NFJ8	Class E basic helix-loop-helix protein 22	**BHLHE22**	GSGSGSGGSSSSSSSSSKKSKE	218	239
Q03188	Centromere protein C	**CENPC**	GSKKSSTRKDKEESKKKRFSSESKNK	482	507
P49754	Vacuolar protein sorting-associated protein 41 homolog	**VPS41**	GSLEESTDESEEEESEEEPKLK	10	31
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	GSRDSSSDSDSSSSEEEEEKTSKSAVKKKPQKV	465	497
Q8NEE8	Tetratricopeptide repeat protein 16	**TTC16**	GSSEGEAEAPEEEEEKEKEKKEEKKSELI	567	595
Q92833	Protein Jumonji	**JARID2**	GSSQDEEEVEEEDDETEDVK	177	196
P25054	Adenomatous polyposis coli protein	**APC**	GSSSSILSASSESSEKAKSEDEKH	2568	2591
O43172	U4/U6 small nuclear ribonucleoprotein Prp4	**PRPF4**	GTDALKKTKKDDEKSKKSKEE	141	161
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	GTDSDIPSSKDSEDSNEDEEEDDEEEDEEDDEDDE SDDSQSESDSNSESDTEGSEEEDDDDKDQDESDSD TEGEKTSMKLNKTTSSVKSPSMSLTG	593	688
Q15811	Intersectin-1	**ITSN1**	HEEEKLKREESVKKKDGEEK	677	696
Q5JSZ5	Protein PRRC2B	**PRRC2B**	HEEVDYSEKLKFSDDEEEEE	376	395
P55197	Protein AF-10	**MLLT10**	HEKEKKKYKEKDKHKQKHKK	228	247
Q8IUR6	CREB3 regulatory factor	**CREBRF**	HELSENEEEEEEEEDYEDDKDDDISDTFSE	377	406
Q16534	Hepatic leukemia factor	**HLF**	HHEDAFSKDKDKEKKLDDES	35	54
Q13200	26S proteasome non-ATPase regulatory subunit 2	**PSMD2**	HICSEHFDSKEKEEDKDKKEKKDKDKKEAP	614	643

O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	HIEVKENPEEEEEEEEEEEEDEESEEEEEEEGES EGSEGDEEDEKVSDEKDSGKTLDKKPSKEMSS DSEYDSDDDRTKEERAYDKAKR	524	611
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	HIPEEESSSSSSDEDEDDRKQ	148	168
Q15311	RalA-binding protein 1	**RALBP1**	HKEEKHKEEKHKEKKSKDLT	136	155
Q13061	Triadin	**TRDN**	HKEKEKGKEKVREKEKPEKK	164	183
Q9HCK1	DBF4-type zinc finger-containing protein 2	**ZDBF2**	HLDDAFSEEEEEDEDKVEDED	81	101
Q8IXF0	Neuronal PAS domain-containing protein 3	**NPAS3**	HLPEKTSESSETSDSESDSKD	450	470
Q8WVC0	RNA polymerase-associated protein LEO1	**LEO1**	HSDDEKWGREDKSDQSDDEK	139	158
Q6VN20	Ran-binding protein 10	**RANBP10**	HSKYPAPSSSSSSSSSSSSS	402	422
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	**ACIN1**	HSPRKSSSISEEKGDSDDEK	201	220
Q70EL1	Inactive ubiquitin carboxyl-terminal hydrolase 54	**USP54**	HSRDWEIESTSSESKSSSSSK	551	571
Q14185	Dedicator of cytokinesis protein 1	**DOCK1**	HSRSQDKLDKDDLEKEKKDKKKEKR	1678	1702
Q9Y2W1	Thyroid hormone receptor-associated protein 3	**THRAP3**	HSRVESSKRKSAKEKKSSSKDS	166	187
Q9UGU5	HMG domain-containing protein 4	**HMGXB4**	HTDGHSEKKKKKEEKDKERE	382	401
Q9NP64	Nucleolar protein of 40 kDa	**ZCCHC17**	HTSKDSKAAKKKKKKKKHKKK	218	238
Q9UJX5	Anaphase-promoting complex subunit 4	**ANAPC4**	IDDEWELDESSDEEEEASNK	748	767
Q9NS56	E3 ubiquitin-protein ligase Topors	**TOPORS**	IDSDSDKDSEVKEDTECDNS	910	929
Q5BKZ1	DBIRD complex subunit ZNF326	**ZNF326**	IEGDEEDEEKIDEPIEEEEDEDEEEEAEEV	483	512
Q96JP5	E3 ubiquitin-protein ligase ZFP91	**ZFP91**	IEKVTTDKDPKEEKEEEDDS	116	135
O14974	Protein phosphatase 1 regulatory subunit 12A	**PPP1R12A**	IESLEQEKVDEEEEGKKDESS	334	354
Q9BRR8	G patch domain-containing protein 1	**GPATCH1**	IFASSSDEKSSSSEDEQGDSEDDQ	757	780
Q9Y5B9	FACT complex subunit SPT16	**SUPT16H**	IFLKNEDEEEEEEEKDEAED	430	449
Q9Y570	Protein phosphatase methylesterase 1	**PPME1**	IIEEEEEDEEGSESISKRKKEDDME	254	278
Q96JM7	Lethal(3)malignant brain tumor-like protein 3	**L3MBTL3**	IKDKDQKEERDVEEDNEEED	147	166
Q12802	A-kinase anchor protein 13	**AKAP13**	IKEKEKDSKDKEKDKKTVNG	1772	1791
P23759	Paired box protein Pax-7	**PAX7**	IKFGKKEEEDEADKKEDDGEKK	162	183
O60832	H/ACA ribonucleoprotein complex subunit 4	**DKC1**	IKKEKKKSKKDKKAKAGLES	466	485
P46100	Transcriptional regulator ATRX	**ATRX**	IKVQEDSSSENKSNSEEEEEEKEEEEEEEEEEE EEEEDENDDSKSPGKGRKKIRKILKDDKL	1428	1489
Q14940	Sodium/hydrogen exchanger 5	**SLC9A5**	ILTVESEEEEEESDSSETEKEDDEG	697	721
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	IPKSKVKKEEKKRHKSSSSSSSSSSDSDSSSDSQS SSDSSDSESATEEKSKKRKKKHRKNSRKHKKEK KKRKKSKKSASSESEAENLEAQPQS	178	270
Q9Y2K5	R3H domain-containing protein 2	**R3HDM2**	ISCPSDKEEEKSTKDVSEKEDKDKNKEKI	105	133
O43159	Ribosomal RNA-processing protein 8	**RRP8**	ISDSEEEEEERKKKCPKKAS	61	80
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	ISNKRKKGSSSEEDEREESD	281	300
P42568	Protein AF-9	**MLLT3**	ISSKSDSEQPSPASSSSSSSS	370	391
Q9NUW8	Tyrosyl-DNA phosphodiesterase 1	**TDP1**	ISSSDESEEEKPKPDKPSTSS	12	32
Q01082	Spectrin beta chain, non-erythrocytic 1	**SPTBN1**	ITSESSPGKREKDKEKDKEK	2336	2355
P07199	Major centromere autoantigen B	**CENPB**	ITTSLKSEGEEEEEEEEEEEEEGE	397	421
Q14093	Cylicin-2	**CYLC2**	KADEKKDEDGKKDANKGDESKD	238	259
Q99442	Translocation protein SEC62	**SEC62**	KADLKKDEKSETKKQQKSDSEEKSDSEKKEDEEGKVGP	290	327
Q04726	Transducin-like enhancer protein 3	**TLE3**	KAEEKDSLSRYDSDGDKSDD	228	247
Q9HC44	Vasculin-like protein 1	**GPBP1L1**	KAEFEDSDTETSSSETSDDD	452	471
Q13428	Treacle protein	**TCOF1**	KAEKPMDNSESSEESSDSADSEEA	600	623
Q14692	Ribosome biogenesis protein BMS1 homolog	**BMS1**	KAIFGDEDESGDSDDEEDDE	430	449
Q13206	Probable ATP-dependent RNA helicase DDX10	**DDX10**	KAKDEEEAFLDWSDDDDDDDD	768	788
P55265	Double-stranded RNA-specific adenosine deaminase	**ADAR**	KAKDSGKSEESSHYSTEKESEK	574	595
Q9GZP8	Immortalization up-regulated protein	**IMUP**	KAKKPKVKKKEKGKKEKGKKKE	82	103
P50502	Hsc70-interacting protein	**ST13**	KAKSEENTKEEKPDSKKVEED	48	68
Q8NFI4	Putative protein FAM10A5	**ST13P5**	KAKSEENTKEEKPDSKKVEED	48	68
Q8IZP2	Putative protein FAM10A4	**ST13P4**	KAKSEENTKEEKPDSKVEED	45	64

O15117	FYN-binding protein	**FYB**	KAKTEEKDLKKLKKQEKEEKD	674	694
P35659	Protein DEK	**DEK**	KCPEILSDESSSDEDEKKNKEESSDDEDKESEEEPPKKTAKREKPK	221	266
Q14093	Cylicin-2	**CYLC2**	KDAKKVAKKDTEKESADSKKD	298	318
Q15424	Scaffold attachment factor B1	**SAFB**	KDDAKKGDDGSGEKSKDQDD	522	541
O94874	E3 UFM1-protein ligase 1	**UFL1**	KDDDSDDESQSSHTGKKKPE	454	473
P27797	Calreticulin	**CALR**	KDEDEEDEEDKEEDEEEDVPGQAKDEL	391	417
Q9UPW6	DNA-binding protein SATB2	**SATB2**	KDEELLTESEENDSEEGSEE	690	709
P23470	Receptor-type tyrosine-protein phosphatase gamma	**PTPRG**	KDEKSESEDGEREHEEDGEKDSEKKEKSGVTH	582	613
P52951	Homeobox protein GBX-2	**GBX2**	KDESKVEDDPKGKEESFSLESD	183	204
P78312	Protein FAM193A	**FAM193A**	KDESPQISSTSSSSSEADDEE	340	360
P11387	DNA topoisomerase 1	**TOP1**	KDKDKKVPEPDNKKKKPKKEEE	179	200
Q15911	Zinc finger homeobox protein 3	**ZFHX3**	KDKDSEKVKEKEKAHKGKGE	3233	3252
Q2WGJ9	Fer-1-like protein 6	**FER1L6**	KDKDSKSSKGKDKADKTEDGKSQQ	427	450
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	KDKEFDKCFKEKKDTKEKHKD	1048	1068
Q86TJ2	Transcriptional adapter 2-beta	**TADA2B**	KDKKEKEKALKRKITKEEKE	235	254
P35663	Cylicin-1	**CYLC1**	KDKKHSKEKKGSKKDIKKDARKDTES	494	519
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	KDKKPRKKAMESSSNSDSDS	230	249
Q15075	Early endosome antigen 1	**EEA1**	KDKLSKVSDSLKNSKSEFEKE	859	879
P27797	Calreticulin	**CALR**	KDKQDEEQRLKEEEEDKKRKEEEEAEDKEDDED KDEDEEDEEDKEEDEEEDVPGQAKDEL	358	417
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	**EFTUD2**	KDLDEMDDDDDDDDVGDHDDD	31	51
Q9H9J4	Ubiquitin carboxyl-terminal hydrolase 42	**USP42**	KDPLEEPKAKKHKKSKKKKKSKDKH	1184	1208
Q8TF01	Arginine/serine-rich protein PNISR	**PNISR**	KDRKKKDKEREREQDKRKEK	675	694
Q92766	Ras-responsive element-binding protein 1	**RREB1**	KDSDKEERGEEDSENESTHS	1573	1592
Q14966	Zinc finger protein 638	**ZNF638**	KDSEPERKRKKTEDSSSGKS	1891	1910
Q8NFN8	Probable G-protein coupled receptor 156	**GPR156**	KDSSFSPGKEEKISDSKDFSD	461	481
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**	KDVKTKKDDSHSAEDSEDEKEDHKN	64	88
Q86YA3	Protein ZGRF1	**ZGRF1**	KDYFEKQVEEKQKKKSEKEKSKDKSHS	2078	2104
Q15878	Voltage-dependent R-type calcium channel subunit alpha-1E	**CACNA1E**	KEAEIREDEEEVEKKKQKKEK	1100	1120
O75264	Small integral membrane protein 24	**SMIM24**	KEAKEKEEKRKKEKKTAKEGES	87	108
Q5JRC9	Protein FAM47A	**FAM47A**	KECSSELKYSMELDEKDEDK	648	667
O60216	Double-strand-break repair protein rad21 homolog	**RAD21**	KEDDEEEEDEDASGGDQDQEE	533	553
Q6ZU65	Ubinuclein-2	**UBN2**	KEDDIEMKKRKRKEEGEKEKK	272	292
Q5T1B0	Axonemal dynein light chain domain-containing protein 1	**AXDND1**	KEDEEESKEDRKLQEENKER	848	867
P21817	Ryanodine receptor 1	**RYR1**	KEDEEKEEEEAAEGEKEEGLEE	1903	1924
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	KEDEKEREDGKEDEGGNEEE	229	248
Q9BZF1	Oxysterol-binding protein-related protein 8	**OSBPL8**	KEDISQNKDESSLSMSKSKSESKL	78	101
Q5VTL8	Pre-mRNA-splicing factor 38B	**PRPF38B**	KEDKDDRRHRDDKRDSKKEKK	404	424
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	KEDKDEKGEEDGKEDKNGNEK	146	166
Q58FG1	Putative heat shock protein HSP 90-alpha A4	**HSP90AA4P**	KEDLELPEDEEEKKKQEEGK	254	273
Q6ZNL6	FYVE, RhoGEF and PH domain-containing protein 5	**FGD5**	KEDLTSDEEQRSSEEEDSAS	859	878
Q9Y2D9	Zinc finger protein 652	**ZNF652**	KEDRENSDDTEEEEEEVSYK	94	113
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	KEDSEKDEKRDEERQKSKRG	648	667
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	KEDSEKEAKKSSKKRNRKHS	338	357

Q9UKN7	Unconventional myosin-XV	**MYO15A**	KEEDEEKKAKKGKKGKKAPE	3	22
Q5HYW3	Retrotransposon gag domain-containing protein 4	**RGAG4**	KEEEEEIRNKNEEEGESKDEEDEDEDGG	404	431
Q9Y4W2	Ribosomal biogenesis protein LAS1L	**LAS1L**	KEEEKEEKEVLPDQVEEEEE	565	584
Q58FF8	Putative heat shock protein HSP 90-beta 2	**HSP90AB2P**	KEEENKDDEEKPKIEDVGSDEEDDSGKDKKKKTKKIKEKYID	159	200
Q15506	Sperm surface protein Sp17	**SPA17**	KEEETSVTILDSSEEDKEKEE	95	115
Q9BTT0	Acidic leucine-rich nuclear phosphoprotein 32 family member E	**ANP32E**	KEEIQDEEDDDDYVEEGEEEEEEEEG	223	248
Q92736	Ryanodine receptor 2	**RYR2**	KEEKEETKSEPEKAEGEDGEKEEKAKEDKGKQK	4429	4461
Q96JP5	E3 ubiquitin-protein ligase ZFP91	**ZFP91**	KEEKEKKEIKVEVEVEVKEEE	254	274
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	KEEKKGRRMKSKPKDKDSKK	562	581
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	KEEKQERKKVEQEEAEEEDE	649	668
Q9NRG0	Chromatin accessibility complex protein 1	**CHRAC1**	KEEKREEDEENDNDNESDHDE	108	128
P29973	cGMP-gated cation channel alpha-1	**CNGA1**	KEEKSKDKKEEEKKEVVVIDPS	140	161
Q9NWK9	Box C/D snoRNA protein 1	**ZNHIT6**	KEEKVKEEVMDWSEVKEEKD	138	157
Q9NVU0	DNA-directed RNA polymerase III subunit RPC5	**POLR3E**	KEEPVSEEGEEDEEQEAEEE	498	517
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	KEERKKKSKDRPSKLEKKND	762	781
Q5T1B0	Axonemal dynein light chain domain-containing protein 1	**AXDND1**	KEFIEPEIDESFKEDEEESKED	836	857
P49321	Nuclear autoantigenic sperm protein	**NASP**	KEGEETEGSEEDDKENDKTEE	472	492
Q9NPG3	Ubinuclein-1	**UBN1**	KEGGEKIKKKKKDDTYDKEKKSKKSKFSKA	194	223
P08238	Heat shock protein HSP 90-beta	**HSP90AB1**	KEGLELPEDEEEKKKMEESK	538	557
Q58FF7	Putative heat shock protein HSP 90-beta-3	**HSP90AB3P**	KEGLELPEDEEEKKKMEESK	411	430
Q58FF7	Putative heat shock protein HSP 90-beta-3	**HSP90AB3P**	KEGLELPEDEEEKKKMEESKEK	411	432
P07900	Heat shock protein HSP 90-alpha	**HSP90AA1**	KEGLELPEDEEEKKKQEEKK	546	565
Q58FG0	Putative heat shock protein HSP 90-alpha A5	**HSP90AA5P**	KEGLELPEDEEEKKKQEEKK	231	250
O94880	PHD finger protein 14	**PHF14**	KEGSDGDNEDDEDEGSGSDEDE	229	250
Q0VFZ6	Coiled-coil domain-containing protein 173	**CCDC173**	KEHEEEEERRKKYEEKDAEE	223	242
P11387	DNA topoisomerase 1	**TOP1**	KEHKKEKDREKSKHSNSEHKDSEKKHKEKEKTKHK DGSSEKHKDKHKDRDKEKRKEEKVR	36	95
Q86U86	Protein polybromo-1	**PBRM1**	KEIEEDKLKREEEKREAEKSEDSSG	912	936
Q14683	Structural maintenance of chromosomes protein 1A	**SMC1A**	KEIEKDKKRMDKVEDELKEKKKEL	247	270
Q96B01	RAD51-associated protein 1	**RAD51AP1**	KEIKKKEVKVKSPVEKKEKKSKSKC	230	254
P46100	Transcriptional regulator ATRX	**ATRX**	KEIQSGSSSSDAEESSEDNKKKKQ	1147	1170
Q96DR7	Rho guanine nucleotide exchange factor 26	**ARHGEF26**	KEKALDIDSDEESEPKEQKSDEKI	384	407
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	KEKCEGKEEEEETDGSGKESK	2184	2204
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	KEKDKVKEKEKDKETGRETK	554	573
Q9UQ13	Leucine-rich repeat protein SHOC-2	**SHOC2**	KEKDSKEKDPKVPSAKEREKE	7	27
Q13061	Triadin	**TRDN**	KEKEDIKKKSEKETAIDVEKKE	330	351
Q9NRL2	Bromodomain adjacent to zinc finger domain protein 1A	**BAZ1A**	KEKEDKEKKREELKKIVEEER	340	360
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	KEKEDNSSEEEEEIEPFPEE	289	308
Q659C4	La-related protein 1B	**LARP1B**	KEKEEKVEKRSNSDSKENRE	30	49
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	KEKEHNKEPDSSVSKEVDDKD	459	479
Q8NCQ7	Protein PROCA1	**PROCA1**	KEKEKDKEEMDEKAKLKKKAKKGQ	240	263
O14776	Transcription elongation regulator 1	**TCERG1**	KEKEKLEEKIKEPIKEPSEE	464	483
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	KEKEKQSDSKGKDQERSRSKEKSKQ	498	522
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1	**SREK1**	KEKERVKEKDREKERERERE KE	269	288
Q86VM9	Zinc finger CCCH domain-containing protein 18	**ZC3H18**	KEKGKSKKEDGVKEEKRKRDSS	754	775
Q5W041	Armadillo repeat-containing protein 3	**ARMC3**	KEKGWRKSKGKKEEEKVKEEEE	677	698
Q8NI27	THO complex subunit 2	**THOC2**	KEKIEKKEKRDSSGGKEEKKH	1565	1585

Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	KEKKDKESTEKYKDRKDRAS	1206	1225
Q9Y6R1	Electrogenic sodium bicarbonate cotransporter 1	**SLC4A4**	KEKKEKERISENYSDKSDIE	52	71
Q92541	RNA polymerase-associated protein RTF1 homolog	**RTF1**	KEKKEKKKKQEEEQEKKKLTQ	234	254
Q86VM9	Zinc finger CCCH domain-containing protein 18	**ZC3H18**	KEKKKEDDDGEIDDGEIDDDDLEEGE	184	209
Q8IYB5	Stromal membrane-associated protein 1	**SMAP1**	KEKKKEEKKREKEPEKPAKP	168	187
Q9NQ38	Serine protease inhibitor Kazal-type 5	**SPINK5**	KEKLEREAAEKKKKEDEDRS	811	830
Q03701	CCAAT/enhancer-binding protein zeta	**CEBPZ**	KEKQKRDADEESIEDVDDEE	824	843
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	KEKQMSSRKDKEGDKERKKSKDKKEKPKSGDAKSSSKSKRSQ	1401	1442
Q5SNT6	WASH complex subunit FAM21B	**FAM21B**	KEKSKKKVEAKSIFDDDMDD	1190	1209
Q5SRD0	Putative WASH complex subunit FAM21	**FAM21D**	KEKSKKKVEAKSIFDDDMDD	245	264
Q641Q2	WASH complex subunit FAM21A	**FAM21A**	KEKSKKKVEAKSIFDDDMDD	1278	1297
Q9Y4E1	WASH complex subunit FAM21C	**FAM21C**	KEKSKKKVEAKSIFDDDMDD	1255	1274
Q13523	Serine/threonine-protein kinase PRP4 homolog	**PRPF4B**	KEKTTRSKSKERKKSKSPSK	209	228
P46821	Microtubule-associated protein 1B	**MAP1B**	KEKTVKKETKVKPEDKKEEKEK	634	655
A6NHT5	Homeobox protein HMX3	**HMX3**	KELDSKSPDEIILEESDSEESKKEGE	165	190
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2	**GCC2**	KELEEKIENLEKECKEKEEK	947	966
Q9NZB2	Constitutive coactivator of PPAR-gamma-like protein 1	**FAM120A**	KELKSKSGESKSSAMSSDGS	1029	1048
Q9UBW5	Bridging integrator 2	**BIN2**	KELLEEEEIEKEGSEASSSEEDEP	313	336
Q9NXE8	Pre-mRNA-splicing factor CWC25 homolog	**CWC25**	KELLQMSLEKKEKKKKKEKKKK	164	185
Q76L83	Putative Polycomb group protein ASXL2	**ASXL2**	KELSEGSEESSDGQSDSQSSE	92	112
Q9NPJ6	Mediator of RNA polymerase II transcription subunit 4	**MED4**	KENEDDVEIMSTDSSSSSSESD	249	270
Q9NQW8	Cyclic nucleotide-gated cation channel beta-3	**CNGB3**	KENEDKQKENEDKGKENEDKDK	721	742
O14974	Protein phosphatase 1 regulatory subunit 12A	**PPP1R12A**	KENEREGEKREEEKEGEDKS	816	835
Q0JRZ9	FCH domain only protein 2	**FCHO2**	KENHFYSSSDSDSEDEEPKK	335	354
Q9H307	Pinin	**PNN**	KENVSALDMEKESEEKEEKESE	446	467
P42568	Protein AF-9	**MLLT3**	KEPSRDHNKSSKESSKKPKE	219	238
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	KEQEPQKSKGKKKKEKKKQD	33	52
Q9Y5Q8	General transcription factor 3C polypeptide 5	**GTF3C5**	KEQLTYESGEDEEDEEEEEEEEEDFK	477	502
P35498	Sodium channel protein type 1 subunit alpha	**SCN1A**	KEQSGGEEKDEDEFQKSESEDS	507	528
Q9NSV4	Protein diaphanous homolog 3	**DIAPH3**	KERREEEDIEEKKSIKKKIKE	691	711
Q05519	Serine/arginine-rich splicing factor 11	**SRSF11**	KERSRDERERSTSKKKKSKDKEKDRE	384	409
Q14093	Cylicin-2	**CYLC2**	KESADSKKDAKKNAKKDAKKD	310	330
Q9UQ13	Leucine-rich repeat protein SHOC-2	**SHOC2**	KESKEKEPKTKGKDAKDGKKDSSA	36	59
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	KESKEKPEKRSQIKEKDIEK	816	835
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	KESLCMENSSNSSSDEDEEE	849	868
P43243	Matrin-3	**MATR3**	KESPSDKKSKTDGSQKTESS	602	621
Q99569	Plakophilin-4	**PKP4**	KESPSKDSEPSCWGKKKKKKK	774	794
O43719	HIV Tat-specific factor 1	**HTATSF1**	KESSPEKEAEEGCPEKESEE	450	469
Q9ULL1	Pleckstrin homology domain-containing family G member 1	**PLEKHG1**	KETDGDEDDYVEIKSEEDESE	1271	1291
Q6EMB2	Tubulin polyglutamylase TTLL5	**TTLL5**	KETEQMAEKKSKKKVEEEEED	761	781
P46939	Utrophin	**UTRN**	KFVEKHLESSDEDEKMDEES	1787	1806
Q9BQG0	Myb-binding protein 1A	**MYBBP1A**	KGAEDKSEEGEDNRSSESEEESEGEESEEEERDGDVD	717	753
Q9NS69	Mitochondrial import receptor subunit TOM22 homolog	**TOMM22**	KGDAEKPEEELEEDDDEELDE	22	42
Q14093	Cylicin-2	**CYLC2**	KGDESKDAKKDAKEIKKGKKDKKKPSSTDSDSKDDVKK ESKKDATKDAKKVAKKDTEKESADSKKDAKKN	253	322
P23760	Paired box protein Pax-3	**PAX3**	KGEEEEADLERKEAEESEKK	164	183
O94826	Mitochondrial import receptor subunit TOM70	**TOMM70A**	KGEKSDEDKDKEGEALEVKE	275	294
Q5SSJ5	Heterochromatin protein 1-binding protein 3	**HP1BP3**	KGEPENEEKEENKSSEETKKDEKDQSKEKEKKVKKTI	97	133
Q9UKL3	CASP8-associated protein 2	**CASP8AP2**	KGEPKTESKSSKFKSNSDSD	293	312

P07197	Neurofilament medium polypeptide	**NEFM**	KGEQKEEEEKEVKEAPKEEK	698	717
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	KGGDSDYDDDDDDDSDDQGDEDDEDEEDK EDKKGKKTDICEDEDEGDQA	1293	1341
P50607	Tubby protein homolog	**TUB**	KGISSSMSFDEDEEDEEENSSSSSQ	188	212
Q6Q759	Sperm-associated antigen 17	**SPAG17**	KGKEKPKESLKEEEHPKEEEKKEEEVEPE	1182	1210
Q9BQI5	SH3-containing GRB2-like protein 3-interacting protein 1	**SGIP1**	KGKHFYSSSESEEEEESHKK	97	116
O95400	CD2 antigen cytoplasmic tail-binding protein 2	**CD2BP2**	KGKHSLDSDEEEDDDDGGSSK	42	62
Q6PD62	RNA polymerase-associated protein CTR9 homolog	**CTR9**	KGKIKSKAIISSSDDSSDEDK	1005	1025
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	KGKKQSFDDNDSEELEDKDSKSKKT	102	126
Q8IVL1	Neuron navigator 2	**NAV2**	KGNEKEKEKQQREKDKEKSKD	506	526
Q9UKN5	PR domain zinc finger protein 4	**PRDM4**	KGPTSSSSAPEEEEEDDSEEED	753	774
Q9NU22	Midasin	**MDN1**	KGQEKDKEDPDSKSDIKGED	4705	4724
P38432	Coilin	**COIL**	KGSVSVCSKESPSSSSESES	233	252
Q14331	Protein FRG1	**FRG1**	KGTKTKSKKKKSKDKKRKREEDEET	15	39
P55198	Protein AF-17	**MLLT6**	KGVSSFTSASSSSSSSSSS	319	338
P11387	DNA topoisomerase 1	**TOP1**	KHKDGSSEKHKDKHKDRDKEK	68	88
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	KHKEKSDKEHSKERKSSRSAD	1255	1275
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	KHKEREKEKHKKEIEGEKEK	656	675
Q9BRR8	G patch domain-containing protein 1	**GPATCH1**	KHKGKQKNKKPEKSSSSESSDSSDSQSDEETAD	881	913
Q13523	Serine/threonine-protein kinase PRP4 homolog	**PRPF4B**	KHRSKHKKHKHSSEEDKDKK	49	68
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	KHSGDDSFDEGSVSESESESES	355	376
Q5T5U3	Rho GTPase-activating protein 21	**ARHGAP21**	KIADSQKSSEDSGSRKDSSSE	906	926
Q5VV16	Forkhead box protein D4-like 5	**FOXD4L5**	KIDVLGEEEDEDEVEDEEEE	26	45
Q6VB85	Forkhead box protein D4-like 2	**FOXD4L2**	KIDVLGEEEDEDEVEDEEEE	26	45
Q8WXT5	Forkhead box protein D4-like 4	**FOXD4L4**	KIDVLGEEEDEDEVEDEEEE	26	45
Q9NU39	Forkhead box protein D4-like 1	**FOXD4L1**	KIDVLGEEEDEDEVEDEEEE	26	45
Q14151	Scaffold attachment factor B2	**SAFB2**	KIEKKEEKKPEDIKKEEKDQDELKP	528	552
Q14151	Scaffold attachment factor B2	**SAFB2**	KIEKTVIKKEEKIEKKEEKK	517	536
Q9ULD9	Zinc finger protein 608	**ZNF608**	KIKKESEEDAEKKDKAEQLDSKKVD	1029	1053
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	KIKKESEKSFREEKIKDLKEE	753	773
Q13144	Translation initiation factor eIF-2B subunit epsilon	**EIF2B5**	KINMEEESESESEQSMDSEE	518	537
Q9BZI7	Regulator of nonsense transcripts 3B	**UPF3B**	KKAESTESIGSSEKTEKKEE	405	424
Q9H1X3	DnaJ homolog subfamily C member 25	**DNAJC25**	KKAKEKGKNKKSKEEIRDEEE	199	219
P55081	Microfibrillar-associated protein 1	**MFAP1**	KKAKEQEAEPEEQEEDSSSD	64	83
Q8N5F7	NF-kappa-B-activating protein	**NKAP**	KKAKKKEKKKKHRSKKYKKKRSKKSRKESSDSSSKESQEEFLENP	233	277
Q86XH1	IQ and AAA domain-containing protein 1	**IQCA1**	KKAKKQKKGTKEKNKEEDEK	364	383
Q6V0I7	Protocadherin Fat 4	**FAT4**	KKAKNPKEEKKPKEKKKKGSE	4532	4552
Q13547	Histone deacetylase 1	**HDAC1**	KKAKRVKTEDEKEKDPEEKKE	438	458
Q96M83	Coiled-coil domain-containing protein 7	**CCDC7**	KKAVKTVKKKDKGKSEDSEKK	341	361
Q58FG0	Putative heat shock protein HSP 90-alpha A5	**HSP90AA5P**	KKDADKKKKKSKEKYIDQEL	95	114
Q13129	Zinc finger protein Rlf	**RLF**	KKDECSSETDLESSCEETESK	1247	1267
O75475	PC4 and SFRS1-interacting protein	**PSIP1**	KKDEEGQKEEDKPRKEPDKKE	232	252
Q8TAQ2	SWI/SNF complex subunit SMARCC2	**SMARCC2**	KKDEEKGKEGDSEKESEKSDGDPIVD	795	820
P0CB48	Putative upstream-binding factor 1-like protein 6	**UBTFL6**	KKDGEEEESSNSLDCSSGED	376	395
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	KKDKDKREKEKVKDKGREDKMK	627	648
Q6PKG0	La-related protein 1	**LARP1**	KKDMKEQEKGEGSDSKESPK	203	222
Q14093	Cylicin-2	**CYLC2**	KKDSKKGKKDAEKGKDSATESEDEKG	158	183
P35663	Cylicin-1	**CYLC1**	KKDSKTDNKKSVKNDEESTD	426	445
P46821	Microtubule-associated protein 1B	**MAP1B**	KKEDKTPIKKEEKPKKEEVKKE	662	683
Q8IUZ0	Leucine-rich repeat-containing protein 49	**LRRC49**	KKEEEKKRESHKQSLLKEKK	319	338
Q8IWE2	Protein NOXP20	**FAM114A1**	KKEEKEEKSQDPQEDKKEEKKTKTIEEV	416	443
Q5T4S7	E3 ubiquitin-protein ligase UBR4	**UBR4**	KKEEKEKEKDGETSGSQEDQ	3371	3390
Q9H158	Protocadherin alpha-C1	**PCDHAC1**	KKEETKKKKKKKKGNKTQEKKEKG	931	954
Q9UN72	Protocadherin alpha-7	**PCDHA7**	KKEETKKKKKKKKGNKTQEKKEKG	905	928

Q9UN73	Protocadherin alpha-6	**PCDHA6**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9UN74	Protocadherin alpha-4	**PCDHA4**	KKEETKKKKKKKGNKTQEKKEKG	915	938
Q9UN75	Protocadherin alpha-12	**PCDHA12**	KKEETKKKKKKKGNKTQEKKEKG	909	932
Q9Y5H5	Protocadherin alpha-9	**PCDHA9**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9Y5H6	Protocadherin alpha-8	**PCDHA8**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9Y5H7	Protocadherin alpha-5	**PCDHA5**	KKEETKKKKKKKGNKTQEKKEKG	904	927
Q9Y5H8	Protocadherin alpha-3	**PCDHA3**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9Y5H9	Protocadherin alpha-2	**PCDHA2**	KKEETKKKKKKKGNKTQEKKEKG	916	939
Q9Y5I0	Protocadherin alpha-13	**PCDHA13**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9Y5I1	Protocadherin alpha-11	**PCDHA11**	KKEETKKKKKKKGNKTQEKKEKG	917	940
Q9Y5I2	Protocadherin alpha-10	**PCDHA10**	KKEETKKKKKKKGNKTQEKKEKG	916	939
Q9Y5I3	Protocadherin alpha-1	**PCDHA1**	KKEETKKKKKKKGNKTQEKKEKG	918	941
Q9Y5I4	Protocadherin alpha-C2	**PCDHAC2**	KKEETKKKKKKKGNKTQEKKEKG	975	998
Q9BU76	Multiple myeloma tumor-associated protein 2	**MMTAG2**	KKEKKKKKKRKHKKEKKKKDKEHRRP	186	211
Q14839	Chromodomain-helicase-DNA-binding protein 4	**CHD4**	KKEKKSKSKRKEEEEEEDDDDDSKEPKSSAQLLEDWGMEDID	123	164
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	KKEKPKSGDAKSSSKSKRSQ	1423	1442
Q9NWH9	SAFB-like transcription modulator	**SLTM**	KKENDEKSSSRSSGDKKNTSD	470	490
P46821	Microtubule-associated protein 1B	**MAP1B**	KKETPPKEVKKEVKKEEKKE	701	720
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	KKEVKSSKEKPEREKTPSEDK	718	738
Q96ST2	Protein IWS1 homolog	**IWS1**	KKFHSSDSEEEEHKKQKMDSDEDEKEGEEEKVAK	358	391
Q99543	DnaJ homolog subfamily C member 2	**DNAJC2**	KKFKEEEKAKKEAEKKAKAE	281	300
Q9H8M9	Protein eva-1 homolog A	**EVA1A**	KKFLQDRESSSDSSDSEDGSED	69	90
Q5T200	Zinc finger CCCH domain-containing protein 13	**ZC3H13**	KKGQKKKSIEKKRKKSKGDSD	995	1015
P25789	Proteasome subunit alpha type-4	**PSMA4**	KKHEEEEAKAEREKKEKEQKEKDK	238	261
P30622	CAP-Gly domain-containing linker protein 1	**CLIP1**	KKHEEEKKELERKLSDLEKK	995	1014
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	KKIKNKLDKEKDKDEKYSPK	639	658
P11387	DNA topoisomerase 1	**TOP1**	KKIKTEDTKKEKKRKLEEEED	150	170
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	KKKDEDEDDNDDGCLKEPKSS	114	134
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	KKKDKKKKKGEKEEKEKEKKKGPSKATVKA	314	343
P21817	Ryanodine receptor 1	**RYR1**	KKKEEKPEEERSAEESKPRS	2074	2093
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	KKKEGSNDELSENEEDLEEKSESEGSDYS	66	94
A2RUB6	Coiled-coil domain-containing protein 66	**CCDC66**	KKKEKEVSEKWNDPWKKSESDK	278	299
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	KKKEKKKQDFDEDDILKELEE	44	64
Q8N9Q2	Protein SREK1IP1	**SREK1IP1**	KKKEKKSKSKKGKHHKKEKKKRKKEKHSSTPNSSEFS	117	153
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	KKKEKQRDREREKDKNKDKSKEKDKVKEKEKDKETGRE	534	571
Q8N9Q2	Protein SREK1IP1	**SREK1IP1**	KKKEKSKEKIKLKKKRKRSYSSSSTEEDTSKQKK	76	109
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	KKKERESAATSSSSSSPSSSS	1265	1285
Q6Q759	Sperm-associated antigen 17	**SPAG17**	KKKGKDNAEKEDSRSLKKKS	968	987
Q58FG0	Putative heat shock protein HSP 90-alpha A5	**HSP90AA5P**	KKKKESNDKPEIEDVGSDEEEEKKDADKKKKKSKEKYIDQEL	73	114
Q9ULG1	DNA helicase INO80	**INO80**	KKKKFKEEKKLKAKLKKVKKK	208	228
Q86X95	Corepressor interacting with RBPJ 1	**CIR1**	KKKKFQKSRSKHKKHKSSSSSSSSSSSSSSSTETSESSSE SESNNKEKKIQRKKRKKNK	241	298
O14617	AP-3 complex subunit delta-1	**AP3D1**	KKKKHRKEKEERTKGKKKSKK	873	893
P18583	Protein SON	**SON**	KKKKKEKEKKYKRQPEESESKTKSHDDG	119	146
Q9NXE8	Pre-mRNA-splicing factor CWC25 homolog	**CWC25**	KKKKKEKKKKHKKHKHRSSSSDRSSSEDEHSAG	176	208
A0JLT2	Mediator of RNA polymerase II transcription subunit 19	**MED19**	KKKKKKEEDPDRKRKKKEKKKKKNRHSPD	200	228
O95391	Pre-mRNA-splicing factor SLU7	**SLU7**	KKKKKKKKKHRKSSSDSDDEEKKHEKLKKALNAEEA	499	534
Q9H9J4	Ubiquitin carboxyl-terminal hydrolase 42	**USP42**	KKKKKKKKRHSRKSEDFVKDSE	1234	1255
A8MW92	PHD finger protein 20-like protein 1	**PHF20L1**	KKKKKKKSKQHDYSDYEDSSLE	571	592
Q5FYB0	Arylsulfatase J	**ARSJ**	KKKKPSKNQAEKKQKKSKKKKKKQ	560	583
Q8N5F7	NF-kappa-B-activating protein	**NKAP**	KKKKSSRSKERSKKRRKKKSSKRKHKKYSED SDSDSDSETDSSDEDNKRRAKKAKKKEKK KKHRSKKYKKKRSKKSRKESSDSSSKESQEEFLENPWKDR	182	281

A6NE01	Protein FAM186A	**FAM186A**	KKKRKEMKSFSEDKSKSPTE	500	519
Q9H8M2	Bromodomain-containing protein 9	**BRD9**	KKKSEKEKHLDDEERRKRKEEKKRK	71	95
Q05519	Serine/arginine-rich splicing factor 11	**SRSF11**	KKKSKDKEKDRERKSESDKDVKQVT	398	422
P35251	Replication factor C subunit 1	**RFC1**	KKKTKSSKPSKPEKDKEPRK	1121	1140
P49756	RNA-binding protein 25	**RBM25**	KKLEEEKGKKEKERQEIEKE	289	308
Q9P2I0	Cleavage and polyadenylation specificity factor subunit 2	**CPSF2**	KKLEQSKEADIDSSDESDIEED	407	428
Q14141	Septin-6		KKLHQDEKKKLEDKKKSLDDE	372	392
Q9NQ66	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1	**PLCB1**	KKLKEICEKEKKELKKKMDKK	1055	1075
Q15544	Transcription initiation factor TFIID subunit 11	**TAF11**	KKLKIDTKEKKEKKQKVDEDE	82	102
Q15858	Sodium channel protein type 9 subunit alpha	**SCN9A**	KKLSSGEEKGDAEKLSKSESEDSI	486	509
P11234	Ras-related protein Ral-B	**RALB**	KKMSENKDKNGKKSSKNKKS	179	198
O14617	AP-3 complex subunit delta-1	**AP3D1**	KKPKKKEKKHKEKERDKEKKKEKEKKKSPKPKKK KHRKEKEERTKGKKKSKKQPPGSEEA	842	901
Q15311	RalA-binding protein 1	**RALBP1**	KKPSFSKKKEKDFKIKEKPKEEKHKEEKHKEEKHKEKKSKDLTA	113	156
P30414	NK-tumor recognition protein	**NKTR**	KKPTHSEGSDSSSNSSSSSESSSESELEHE	190	219
O15117	FYN-binding protein	**FYB**	KKREKEEKKRLELEKKEQKEKEKKEQEIKKKFK	473	505
O14713	Integrin beta-1-binding protein 1	**ITGB1BP1**	KKRHSSSSSQSSEISTKSKS	6	25
Q99549	M-phase phosphoprotein 8	**MPHOSPH8**	KKRISEAKEELKESKKPKKDE	200	220
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	KKRKKSKKSASSESEAENLE	246	265
Q2KHR3	Glutamine and serine-rich protein 1	**QSER1**	KKRKKWKEEFSSSQSDSSPE	1499	1518
Q96GQ7	Probable ATP-dependent RNA helicase DDX27	**DDX27**	KKRKTEDKEAKSGKLEKEKE	126	145
Q04724	Transducin-like enhancer protein 1	**TLE1**	KKRKVDDKDSSHYDSDGDKSDD	225	246
P46087	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	**NOP2**	KKRPAPGSDEEEEEEDSEED	104	123
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	KKRQIDSSEEDDDEEDYDNDK	209	229
P25205	DNA replication licensing factor MCM3	**MCM3**	KKRSEDESETEDEEEKSQEDQEQKRK	665	690
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	KKRSKAKKDPKEPKEPKEKKE	644	664
Q8IYB4	PEX5-related protein	**PEX5L**	KKSEPSSKTSSLKKKADGSD	126	145
Q9UIF9	Bromodomain adjacent to zinc finger domain protein 2A	**BAZ2A**	KKSKAEKEKGKTKQEKLKEK	752	771
Q9UKJ3	G patch domain-containing protein 8	**GPATCH8**	KKSSKHKRKHKADTEEKSSK	682	701
Q8N9E0	Protein FAM133A	**FAM133A**	KKSYPDDKPLSSESSSESDYEEDV	181	204
Q9Y6F6	Protein MRVI1	**MRVI1**	KKTKELQDLKEEEEEQKSES	766	785
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	KKTKILSDSEDSESEEQDRED	1814	1834
P35251	Replication factor C subunit 1	**RFC1**	KKTKSSPAKKESVSPEDSEKK	355	375
Q9H3N1	Thioredoxin-related transmembrane protein 1	**TMX1**	KKVEEEQEADEEDVSEEEAESKEG	233	256
Q8WWQ0	PH-interacting protein	**PHIP**	KKVPKNKTKKAESSSDEEEESEKQK	867	891
Q8TDI8	Transmembrane channel-like protein 1	**TMC1**	KKVQIKVEEKEDETEESSSEEEEEVEDKL	4	32
Q07866	Kinesin light chain 1	**KLC1**	KKYDDDISPSEDKDTDSTKE	155	174
Q9BXY0	Protein MAK16 homolog	**MAK16**	KLDASSDEDQDGKSSSEEEEEKALS	239	263
Q96JN2	Coiled-coil domain-containing protein 136	**CCDC136**	KLDGLAKEEEKKEEMEEEKK	1029	1048
Q9HCK8	Chromodomain-helicase-DNA-binding protein 8	**CHD8**	KLEDEDDSDSELDLSKLSPSSSSSSSSSSSSSTDE SEDEKEEKLTDQSRSKLYDEESLLSL	2062	2123
P11387	DNA topoisomerase 1	**TOP1**	KLEEEEDGKLKKPKNKDKDKK	164	184
Q9NQ38	Serine protease inhibitor Kazal-type 5	**SPINK5**	KLEEEEKKNDKEEKGKVEAEK	537	557
Q9BT43	DNA-directed RNA polymerase III subunit RPC7-like	**POLR3GL**	KLETLEKKEEEVTSEEDEEKEEEEEKEEEEEEEYDEEEHEEETDYIMSY	150	198
Q9NR09	Baculoviral IAP repeat-containing protein 6	**BIRC6**	KLISEQKDDKEKKNHEEKEK	3882	3901
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	KLLTRKADSSSSEEESSSSEEEKTKK	415	440
Q9UPV0	Centrosomal protein of 164 kDa	**CEP164**	KLSTSGAIKKKKKKKEKKDKKD	102	123
Q15061	WD repeat-containing protein 43	**WDR43**	KLVYEEESSEEESDDEIADKDSEDN	596	620
Q9NU22	Midasin	**MDN1**	KMHDGELEEQEEDDEKSDSE	4736	4755
Q05682	Caldesmon	**CALD1**	KNDWRDAEENKKEDKEKEEEEEEKPK	174	199

Q9NQ66	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1	**PLCB1**	KNKKKSHKSSEGSGKKKLSE	465	484
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	KNKVKKRAKSKSRSKSKEKSKSKERDSKH	449	477
A6NCI8	Uncharacterized protein C2orf78	**C2orf78**	KNSSKKSEESKQSGKKVKVEEK	578	599
Q2M1Z3	Rho GTPase-activating protein 31	**ARHGAP31**	KPAKDDSPSSLESSKEEKPK	1245	1264
Q8IZP2	Putative protein FAM10A4	**ST13P4**	KPDSKVEEDLKADEPSSEESD	56	76
O95467	Neuroendocrine secretory protein 55	**GNAS**	KPEDKDPRDPEESKEPKEEK	195	214
Q16186	Proteasomal ubiquitin receptor ADRM1	**ADRM1**	KPEQKEGDTKDKKDEEEDMS	386	405
Q02952	A-kinase anchor protein 12	**AKAP12**	KPKEDEVEASEKKKEQEPEK	315	334
P46821	Microtubule-associated protein 1B	**MAP1B**	KPKKEEVKKEVKKEIKKEEKKEPKKEVKKETP	674	705
Q99856	AT-rich interactive domain-containing protein 3A	**ARID3A**	KPKWEEEEMEEDLGEDEEEEEEDYEDEEEEEDEEGLG	124	160
Q8WY36	HMG box transcription factor BBX	**BBX**	KPPKKKVKSREKKMSKEKSSD	525	545
P26358	DNA (cytosine-5)-methyltransferase 1	**DNMT1**	KPQEESERAKSDESIKEEDKD	179	199
Q7Z6E9	E3 ubiquitin-protein ligase RBBP6	**RBBP6**	KPVKEEKVKKDYSKDVKSEK	1103	1122
Q8IWY8	Zinc finger and SCAN domain-containing protein 29	**ZSCAN29**	KQAEEADEATEEDSDDDEED	524	543
Q03164	Histone-lysine N-methyltransferase 2A	**KMT2A**	KQAKAVKKKEKKSKTSEKKDSKESSV	1208	1233
O95751	Protein LDOC1	**LDOC1**	KQCFGWDDDEDDDDEEEEDD	126	145
Q14696	LDLR chaperone MESD	**MESDC2**	KQDKGKKKKEGDLKSRSSKEE	204	224
O94888	UBX domain-containing protein 7	**UBXN7**	KQDSRSDEESESELFSGSEE	309	328
Q9HD67	Unconventional myosin-X	**MYO10**	KQEEEEKKKREEEERERERE	823	842
O00203	AP-3 complex subunit beta-1	**AP3B1**	KQENSAKKFYSESEEEEDSSDSSSDSESESGSESGEQGESGEEGDS	667	712
Q9GZP8	Immortalization up-regulated protein	**IMUP**	KQGHHSSSDSSSSSSDSDTD	46	65
Q8TDR0	TRAF3-interacting protein 1	**TRAF3IP1**	KQKEELKEDRKPREKDKDKEK	183	203
Q9UQ90	Paraplegin	**SPG7**	KQKNKEKDKSKGKAPEEDEEE	107	127
Q01484	Ankyrin-2	**ANK2**	KQKRDYKKEPKQEESSSSSD	2606	2625
Q2TBE0	CWF19-like protein 2	**CWF19L2**	KQKYEKNNESSDSSSSSEDEW	99	119
Q8TA86	Retinitis pigmentosa 9 protein	**RP9**	KQLLEDSTSDEDRSSSSSSE	161	180
Q969S3	Zinc finger protein 622	**ZNF622**	KQQEEDSEEEEEDLDGDDWED	192	212
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	KQSDNSEYSKSEKGKNKEKD	936	955
Q58FG0	Putative heat shock protein HSP 90-alpha A5	**HSP90AA5P**	KQVSDAEAEKKEDKRKKKKES	58	78
Q9NRL2	Bromodomain adjacent to zinc finger domain protein 1A	**BAZ1A**	KREKADALEAKKKEKEDKEKK	328	348
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	KREKSDKDKSKVSTSGSSVDSSKKTSESKN	1354	1383
Q08170	Serine/arginine-rich splicing factor 4	**SRSF4**	KRGSKRDSKAGSSKKKKKED	379	398
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	KRKDDSKDDDETEEDNNQDE	850	869
Q6ZT98	Tubulin polyglutamylase TTLL7	**TTLL7**	KRPKYCSSDSSYDSSSSSSESDENEKEEY	542	570
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	KRRKSGDDKDKKEDRDERKKEDKRKDDSKDDDETEEDNNQDE	828	869
O75400	Pre-mRNA-processing factor 40 homolog A	**PRPF40A**	KRSRSRSGSDSDDDDSHSKKK	824	844
Q5VTL8	Pre-mRNA-splicing factor 38B	**PRPF38B**	KRSRSRSKEKSSKHKNESKEKS	446	467
Q5JRM2	Uncharacterized protein CXorf66	**CXorf66**	KRSSKSSCSKKLSKSSHLEK	162	181
Q15652	Probable JmjC domain-containing histone demethylation protein 2C	**JMJD1C**	KRTYESGSESGDSDESESKSE	1619	1639
Q9UHL9	General transcription factor II-I repeat domain-containing protein 1	**GTF2IRD1**	KRVSEGNSVSSSSSSSSSSS	903	923
Q8IWA0	WD repeat-containing protein 75	**WDR75**	KSAKEIPEDVDMEEEKESEDSDEEND	762	787
Q92736	Ryanodine receptor 2	**RYR2**	KSCHDEEDDDGEEEVKSFEEKE	3697	3718
Q15413	Ryanodine receptor 3	**RYR3**	KSCQSGEDEEEDEDKEKTFEEKEMEK	3586	3611
P15884	Transcription factor 4	**TCF4**	KSDDEGDENLQDTKSSEDKK	514	533
P07197	Neurofilament medium polypeptide	**NEFM**	KSDQAEEGGSEKEGSSEKEE	544	563
Q8NHV1	GTPase IMAP family member 7	**GIMAP7**	KSEEEKEKEIKLLKLKYDEK	241	260
Q70CQ2	Ubiquitin carboxyl-terminal hydrolase 34	**USP34**	KSERKEGFKEVSDHSKDSES	2126	2145
Q96AJ1	Clusterin-associated protein 1	**CLUAP1**	KSGSNDDSDIDIQEDDESDSE	307	327

Q9ULG1	DNA helicase INO80	**INO80**	KSILLSDESSEADSQSEDDDEEEL	127	150
Q8IU89	Ceramide synthase 3	**CERS3**	KSIQDVRSDDEDYEEEEEEEEEEA	333	356
Q15287	RNA-binding protein with serine-rich domain 1	**RNPS1**	KSKDRSKDKGATKESSEKDR	38	57
Q9BZ95	Histone-lysine N-methyltransferase NSD3	**WHSC1L1**	KSKHESRKEKRKKSNKHDSS	187	206
Q9ULD4	Bromodomain and PHD finger-containing protein 3	**BRPF3**	KSKKPSSKGKKKESCSKHAS	89	108
Q6ZUT1	Uncharacterized protein C11orf57	**C11orf57**	KSKKSHKKKQKKRSHKKQKKSKKEA	152	176
O60524	Nuclear export mediator factor NEMF	**NEMF**	KSKMKKMKEKYKDQDEEDRE	873	892
Q9NXR8	Inhibitor of growth protein 3	**ING3**	KSKNNNKSSSQQSSSSSSSS	301	321
Q86Y13	E3 ubiquitin-protein ligase DZIP3	**DZIP3**	KSKQRKKKKTKNKKNKDSKED	656	676
P67936	Tropomyosin alpha-4 chain	**TPM4**	KSLEAASEKYSEKEDKYEEE	169	188
Q12789	General transcription factor 3C polypeptide 1	**GTF3C1**	KSLGKDGSLEDDEDEEDDLDE	1604	1624
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	KSNKKWDGSEEDEDNSKKIKE	156	176
Q9Y4D8	Probable E3 ubiquitin-protein ligase HECTD4	**HECTD4**	KSPEKSKPDEKDPEKSPTKK	3380	3399
Q9NPI1	Bromodomain-containing protein 7	**BRD7**	KSPSKENKKKDKDMLEDKFKS	288	308
Q9UKJ3	G patch domain-containing protein 8	**GPATCH8**	KSPSQYSEEEEEEDSGSEHS	831	850
P11171	Protein 4.1	**EPB41**	KSQVSEEEGKEVESDKEKGE	91	110
Q9UHB7	AF4/FMR2 family member 4	**AFF4**	KSREIIETDTSSSDSDESES	639	658
Q6KC79	Nipped-B-like protein	**NIPBL**	KSRKRVDSDSDSDSEDDINS	2502	2521
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	KSRSLKHSSKDIKKKDENKSDDKDGKEVDSSHEK	1021	1054
O60763	General vesicular transport factor p115	**USO1**	KSSEEDKKEEEVKKTLEQHD	630	649
O75330	Hyaluronan mediated motility receptor	**HMMR**	KSSESKESQKNDKDLKILEKE	70	90
Q13439	Golgin subfamily A member 4	**GOLGA4**	KSSHEKSNKSLEDKSLEFKK	1189	1208
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	KSSKKSSNKKPDSEASALKKK	693	713
Q5M9Q1	NKAP-like protein	**NKAPL**	KSSKRKHRKYSDSDSNSESD	192	211
P17480	Nucleolar transcription factor 1	**UBTF**	KSSRTTLQSKSESEEDDEEDEDDEDEDEEEE DDENGDSSEDGGDSSESSSEDESEDGDENE EDDEDEDDDEDDDEDEDNESEGSSSSSSSSGDSSDSDSN	665	764
Q9NYV4	Cyclin-dependent kinase 12	**CDK12**	KSSSKESRSSKLHKEKTRKE	168	187
Q6DN12	Multiple C2 and transmembrane domain-containing protein 2	**MCTP2**	KSSSSSLKQSEEELDWSQEE	90	109
Q6ZUT1	Uncharacterized protein C11orf57	**C11orf57**	KSSTHESRKHKKSKKSHKKK	141	160
Q5T8A7	Protein phosphatase 1 regulatory subunit 26	**PPP1R26**	KTDEARRLDEKESSEDKSSS	667	686
Q9UEW8	STE20/SPS1-related proline-alanine-rich protein kinase	**STK39**	KTEDGDWEWSDDEMDEKSEE	376	395
P12036	Neurofilament heavy polypeptide	**NEFH**	KTEEKKDSKKEEAPKKEAPK	858	877
Q8TEW0	Partitioning defective 3 homolog	**PARD3**	KTGKEKKKDRDKEKDKMKAKK	994	1014
Q5M9Q1	NKAP-like protein	**NKAPL**	KTKKESSDSSCKDSEEDLSEAT	243	264
P49792	E3 SUMO-protein ligase RanBP2	**RANBP2**	KTPEKAKEKKKPEDSPSDDD	2612	2631
P21127	Cyclin-dependent kinase 11B	**CDK11B**	KTSSAESSSAESGSGSEEEEEEEEEEEEEEGSTSEES EEEEEEEEEEEEETGSNSEEASEQSAEEVSEEEMSEDEERENENHL	288	369
Q9UQ88	Cyclin-dependent kinase 11A	**CDK11A**	KTSSAESSSAESGSGSEEEEEEEEEEEEEEGSTSEES EEEEEEEEEEEEETGSNSEEASEQSAEEVSEEEMSEDEERENENHL	276	357
Q9Y2X3	Nucleolar protein 58	**NOP58**	KVEEEEEEKVAEEEETSVKKKKKR	467	490
Q9P2E9	Ribosome-binding protein 1	**RRBP1**	KVEKKKKEKTVEKKGKTKKKEEKP	56	79
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	KVKEEIKSDSSPLPSEKSDEDDDKLSESKSDGRE	1346	1379
Q8TEW8	Partitioning defective 3 homolog B	**PARD3B**	KVKKTKEKEKKKEKGKLKVKEKKRKEENED	833	862
P46060	Ran GTPase-activating protein 1	**RANGAP1**	KVLASLSDDEDEEEEEGEEEEEEAEEEEEEDEEEE EEEEEEEEEPQQRGQGEKSA	352	408
P11388	DNA topoisomerase 2-alpha	**TOP2A**	KVPDEEENEESDNEKETEKSDS	1096	1117
P46100	Transcriptional regulator ATRX	**ATRX**	KVSSEDSEDSDFQESGVSEE	1373	1392
O60885	Bromodomain-containing protein 4	**BRD4**	KVVAPPSSSDSSSDSSSDSDSS	478	499
P55010	Eukaryotic translation initiation factor 5	**EIF5**	KWLKEAEEESSGGEEEDEDE	380	399
Q14432	cGMP-inhibited 3',5'-cyclic phosphodiesterase A	**PDE3A**	KWVEDSDESGDTDDPEEEEEE	1025	1045
Q99442	Translocation protein SEC62	**SEC62**	KYDKDIKKEKDKGKAESGKEEDKKSKKENIKDEKTKKEKE KKKDGEKEESKKEETPGTPKKKETKKKFKLEP	101	172

Q9NZ81	Proline-rich protein 13	**PRR13**	KYHKHGKHSSSSSSSSSSDSD	128	148
Q02880	DNA topoisomerase 2-beta	**TOP2B**	KYTFDFSEEEDDDADDDDDD	1369	1388
Q9HD67	Unconventional myosin-X	**MYO10**	LAEKREQEEKKKQEEEEKKK	812	831
Q8WY36	HMG box transcription factor BBX	**BBX**	LAKKLLDFSEEEEEEDEEED	32	51
Q9HAZ2	PR domain zinc finger protein 16	**PRDM16**	LASEKQEDVEEEDDDDLEEDDEDSL	1110	1134
O75554	WW domain-binding protein 4	**WBP4**	LDESKSSDSHSDSDGEQEAEE	217	237
Q96MC2	Dynein regulatory complex protein 1	**DRC1**	LDGKKESEEDQSKSYKQKEES	59	79
P48380	Transcription factor RFX3	**RFX3**	LDKDEGSEVESEMDEELDDSSE	668	689
Q8IWI9	MAX gene-associated protein	**MGA**	LDLEEDDEDDNEKTDDSIDE	2260	2279
Q9ULU8	Calcium-dependent secretion activator 1	**CADPS**	LDPSSSEEESDEIVEEESGKE	2	22
Q9UBV2	Protein sel-1 homolog 1	**SEL1L**	LDSEESELESSIQEEEDSLKS	61	81
Q32MZ4	Leucine-rich repeat flightless-interacting protein 1	**LRRFIP1**	LDSSQKKTKNKKKKNKKKKS	562	581
Q8NAV1	Pre-mRNA-splicing factor 38A	**PRPF38A**	LEEDMDDVESSEEEEEEDEKLERV	184	207
Q5QJE6	Deoxynucleotidyltransferase terminal-interacting protein 2	**DNTTIP2**	LEEEDKASEVAIEEEKEEEEDEKSEEDSSDHDENEDEFSDEEDFLNSTK	505	553
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	LEEKSESEGSDYSPNKKKKKK	82	102
P18858	DNA ligase 1	**LIG1**	LEEQSEDEDREAKRKKEEEEEE	137	158
Q92793	CREB-binding protein	**CREBBP**	LEESIKELEQEEEERKKEES	1549	1568
Q8IY81	pre-rRNA processing protein FTSJ3	**FTSJ3**	LEGEEKDGISDSDSSTSSEEEESWE	607	631
Q9NTJ3	Structural maintenance of chromosomes protein 4	**SMC4**	LEKEKEKEEKKLKEVMDSLK	449	468
Q58FF7	Putative heat shock protein HSP 90-beta-3	**HSP90AB3P**	LEKEQDKEISDDEAEEEKGEKEEEDKDDEEK PKIKDVGSDEEDDSKEYGEFYKS	196	249
P08238	Heat shock protein HSP 90-beta	**HSP90AB1**	LEKEREKEISDDEAEEEKGEKEEEDKDDEEKPKIEDVGS DEEDDSGKDKKKKTKKIKEKYIDQEEL	217	282
Q14165	Malectin	**MLEC**	LEKKEEEEEEEEYDEGSNLKK	227	247
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	LEKKSKLEKNIKDDKSTKEK	720	739
Q12872	Splicing factor, suppressor of white-apricot homolog	**SFSWAP**	LEKNRVKLDDDSDDDEESKE	593	612
Q9H3M9	Putative ataxin-3-like protein	**ATXN3L**	LEKVSEESDESGTSDQDEED	209	228
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	**MPHOSPH10**	LENIEKEEERKDDNDEEEED	214	233
Q9BXY0	Protein MAK16 homolog	**MAK16**	LEQQEAESDSSDTEEKDDDDDDEEDV	190	215
Q9P266	Junctional protein associated with coronary artery disease	**KIAA1462**	LESFNKELQEEEESSSSSSSSSSSSEESEAEPQ	835	867
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	LESSSSDDEDGPAEENDEEKEKEAKK	272	297
Q5SSJ5	Heterochromatin protein 1-binding protein 3	**HP1BP3**	LFPKKEPDDSRDEDEDEDESSEEDSEDEEPPPK	422	454
Q6ZU64	Coiled-coil domain-containing protein 108	**CCDC108**	LGKKEEGEEEKGEEEEEELEEEEEEEEETEEEELGK EEIEEKEEERDEKEEKVSWAG	1757	1813
Q13610	Periodic tryptophan protein 1 homolog	**PWP1**	LGSKLSKKKKKKGKKSSSAE	234	253
Q8TAD8	Smad nuclear-interacting protein 1	**SNIP1**	LHESSDTSEIDRKDDEDEEEEEEVSDS	370	396
Q9NUL5	UPF0515 protein C19orf66	**C19orf66**	LILEDLKEEEEEEEEVEDEE	267	286
Q8N5C6	S1 RNA-binding domain-containing protein 1	**SRBD1**	LKDEFSSFSELSSASEEDDKEDSA	17	40
P51608	Methyl-CpG-binding protein 2	**MECP2**	LKDKPLKFKKVKKDKKEEKE	21	40
Q8N9T8	Protein KRI1 homolog	**KRI1**	LKDRYGDRDSSSDSSSESDSSDER	32	55
P08575	Receptor-type tyrosine-protein phosphatase C	**PTPRC**	LKHELEMSKESEHDSDESSDDDSDSEEPSKYI	985	1016
Q86X95	Corepressor interacting with RBPJ 1	**CIR1**	LKHESSREDSKWSHSDSDKKS	339	359
Q76M96	Coiled-coil domain-containing protein 80	**CCDC80**	LKKAKESKKHEKLEKPEKEKKKKMK	531	555
Q96HJ3	Coiled-coil domain-containing protein 34	**CCDC34**	LKKKNAEECERKKKEKEKEK	240	259
Q6P4R8	Nuclear factor related to kappa-B-binding protein	**NFRKB**	LKKKVKEKEEKKKKKIKTIKSE	308	329
A6NLX3	Speedy protein E4	**SPDYE4**	LKRKSEWSDESEEELEEELE	51	70
Q9Y4C8	Probable RNA-binding protein 19	**RBM19**	LKSKMVKAGSSSSSEEEESEDE	211	232

Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	LKVEEQKEEQKELEKSEKEEDEDDDRKSEDDKEEEERKRQEEIER	670	714
Q14324	Myosin-binding protein C, fast-type	**MYBPC2**	LLKKREVVEEEKKKKKKDDDD	187	207
Q92614	Unconventional myosin-XVIIIa	**MYO18A**	LMKKDKDKDGGRKEKKEKKEKKER	4	27
Q8N9T8	Protein KRI1 homolog	**KRI1**	LNKRYEEEEEEEEDEEEMEEEE	245	266
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	LPEKLFEEKEKVKEKEKKKDKKEKKKKKEKEKEKKEKER EKEKREREKREKEKEKHKHEKIKVEPVAL	685	752
Q8IYW2	Protein CFAP46	**CFAP46**	LPKKEKENERSKEKEKERSKEKENE	1360	1384
Q8IXQ4	GPALPP motifs-containing protein 1	**GPALPP1**	LPPNYKSSSSDSSDSDEDSSS	35	55
O95405	Zinc finger FYVE domain-containing protein 9	**ZFYVE9**	LPSKEDSVTEEKEIEESKSE	495	514
Q8N9Q2	Protein SREK1IP1	**SREK1IP1**	LQEKRINEEEEKKKEKSKEK	65	84
Q86YN6	Peroxisome proliferator-activated receptor gamma coactivator 1-beta	**PPARGC1B**	LQQQEEEDEEEEEEEEEEKEEEEEWGRK	426	454
Q92794	Histone acetyltransferase KAT6A	**KAT6A**	LRCQSSSKRKSKDEEEDEESDD	1094	1115
O60231	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX16	**DHX16**	LRKKREEEEEEEASEKGKKK	128	147
Q8N954	G patch domain-containing protein 11	**GPATCH11**	LRLEEETEEDEEEKEQDEDE	188	207
O95260	Arginyl-tRNA--protein transferase 1	**ATE1**	LSDDIKESLESEGKNSKKEE	143	162
Q8NFU7	Methylcytosine dioxygenase TET1	**TET1**	LSEAPSENSSPSKSEKDEESE	900	920
Q5RHP9	Glutamate-rich protein 3	**ERICH3**	LSEKSGKHVSAEEKEKDKSK	684	703
Q08945	FACT complex subunit SSRP1	**SSRP1**	LSESFKSKEFVSSDESSSGE	656	675
Q5BKY9	Protein FAM133B	**FAM133B**	LSGSESSSKKRQRKKKEKKKS	81	101
Q13316	Dentin matrix acidic phosphoprotein 1	**DMP1**	LSHSKSESREEQADSESSES	392	411
Q99460	26S proteasome non-ATPase regulatory subunit 1	**PSMD1**	LSITAKAKKKEKEKEKKEEEK	833	853
Q96KB5	Lymphokine-activated killer T-cell-originated protein kinase	**PBK**	LSNDDDDEDKTFDESDFDDE	250	269
Q9NNZ6	Protamine-3	**PRM3**	LSSAGEEEEEEEEEGEEEEKEELP	46	69
Q8N5Y2	Male-specific lethal 3 homolog	**MSL3**	LSSSSDCSENKDEEISEESDIEEKTE	129	154
Q13233	Mitogen-activated protein kinase kinase kinase 1	**MAP3K1**	LSSSSTSTSSSENSIKDEEE	421	440
P19021	Peptidyl-glycine alpha-amidating monooxygenase	**PAM**	LSTEGSDQEKEDDGSESEEE	941	960
Q9HD90	Neurogenic differentiation factor 4	**NEUROD4**	LTEEHDSIEEEEEEEEDGEK	47	66
Q05DH4	Protein FAM160A1	**FAM160A1**	LTKEEEGKEESKGEKEKEGKKE	782	803
P58658	Protein eva-1 homolog C	**EVA1C**	LVPGSDKVEEDSEDEEEEED	361	380
Q96A61	Tripartite motif-containing protein 52	**TRIM52**	LWSKEDEEDQNEEEDEWEEEEDEEA	47	71
P11388	DNA topoisomerase 2-alpha	**TOP2A**	MDLDSDEDFSDFDEKTDDED	1328	1347
Q7Z4N2	Transient receptor potential cation channel subfamily M member 1	**TRPM1**	MEDDEPPAKGKKKKKKKKEEE	575	595
Q8IYE0	Coiled-coil domain-containing protein 146	**CCDC146**	MEDSSTDTEKEEEEEKDEKDQEPI	1	24
Q12906	Interleukin enhancer-binding factor 3	**ILF3**	MEEDGEEKSPSKKKKKIQKKEEKAE	374	398
O95391	Pre-mRNA-splicing factor SLU7	**SLU7**	MELHQEKLKEEKKKKKKKKKK	487	507
O75381	Peroxisomal membrane protein PEX14	**PEX14**	MEVQGEEEKREDKEDEEDEEDDDVS	311	335
Q9BY89	Uncharacterized protein KIAA1671	**KIAA1671**	MFKDSTEEKSPRKEESDEEE	1686	1705
A6NE52	WD repeat-containing protein KIAA1875	**KIAA1875**	MGEKPGEEGEEDKKEEEEEKEDEELD	1089	1114
Q8N128	Protein FAM177A1	**FAM177A1**	MKKEEEEEEEENRMSEEAEK	142	161
Q13127	RE1-silencing transcription factor	**REST**	MKKSTKKKTLKNKSSKKSSK	575	594
Q9BXN1	Asporin	**ASPN**	MMLKDMEDTDDDDDDDDDDDDDEDNSL	29	56
Q2Y0W8	Electroneutral sodium bicarbonate exchanger 1	**SLC4A8**	MPESKKKKLDDAKKKAKEEEE	1008	1028
P53999	Activated RNA polymerase II transcriptional coactivator p15	**SUB1**	MPKSKELVSSSSSGSDSDSE	1	20
Q9NZ53	Podocalyxin-like protein 2	**PODXL2**	MPPREEEEEEEEEEREKEE	158	177
Q9Y618	Nuclear receptor corepressor 2	**NCOR2**	MPRSSQEEKDEKEKEKEAEKEEEKPE	512	537
Q9BWU0	Kanadaptin	**SLC4A1AP**	MRMKDEPEVEEEEEEEEEEEKEKEEHEKKKLEDGSLS	639	675
Q53EV4	Leucine-rich repeat-containing protein 23	**LRRC23**	MSDEDDLEDSEPDQDDSEKEEDEKETEEGEDYRKEGEE	1	38
Q8NEM8	Cytosolic carboxypeptidase 3	**AGBL3**	MSEDSEKEDYSDRTISDEDESDEDM	1	25

Q96TC7	Regulator of microtubule dynamics protein 3	**RMDN3**	NAESDNERDSDKESEDGEDE	180	199
Q15910	Histone-lysine N-methyltransferase EZH2	**EZH2**	NDDDDDDDGDDPEEREEKQKD	182	202
Q9NSY1	BMP-2-inducible protein kinase	**BMP2K**	NDESESDFESDPPSPKSSEEEE	737	758
O14967	Calmegin	**CLGN**	NDGEMLEKEEESEPEEKSEEE	543	563
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	NEEEAGKEKEDLKEEEEGKEEDEIKEDDGKKEEP	245	278
Q8NHV1	GTPase IMAP family member 7	**GIMAP7**	NEEIKLVEEDKHKSEEEKEKE	229	249
P80303	Nucleobindin-2	**NUCB2**	NEEKRKEEESKFEEMKKKHE	196	215
Q5TID7	Coiled-coil domain-containing protein 181	**CCDC181**	NENKDTDSKKSEEYEDDFEKD	2	22
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	NGDKEEDDEGKKEDKKGKFK	1683	1702
P82970	High mobility group nucleosome-binding domain-containing protein 5	**HMGN5**	NGEDGKEKGEDEKEEEDRKE	185	204
O60583	Cyclin-T2	**CCNT2**	NHHSKMSKSSKSSGSSSSSSS	629	650
O43719	HIV Tat-specific factor 1	**HTATSF1**	NKESEEEVGPTKESEEDDSEKESDEDCSEKQSEDGSEREFEE	526	567
Q9HAU5	Regulator of nonsense transcripts 2	**UPF2**	NKESNKDDTKEAKESKENKE	492	511
Q9BZI1	Iroquois-class homeodomain protein IRX-2	**IRX2**	NKSEDEDEDEGDATRSKDES	184	203
Q03164	Histone-lysine N-methyltransferase 2A	**KMT2A**	NKSETKSGDKIKKKDSKSIEKK	200	221
P29973	cGMP-gated cation channel alpha-1	**CNGA1**	NNSSNKDQEPEEKKKKKKEKKSKSDDKNENKND	93	125
Q86XH1	IQ and AAA domain-containing protein 1	**IQCA1**	NQEEEEKNKKKKKKKEKQPKK	345	365
Q96M86	Dynein heavy chain domain-containing protein 1	**DNHD1**	NQKRESKTDMKEEDDESEES	3592	3611
Q5T1M5	FK506-binding protein 15	**FKBP15**	NQQEQEKEESSSEEEEKAEE	930	950
Q9Y2W1	Thyroid hormone receptor-associated protein 3	**THRAP3**	NSDKSSSDRSRRSSSSRSSS	145	164
Q86X95	Corepressor interacting with RBPJ 1	**CIR1**	NSDSEEKDKSKKRKLHEELSSS	304	325
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	NSEDEQSLKSEIEEEELKDE	1677	1696
O43364	Homeobox protein Hox-A2	**HOXA2**	NSEGKCKSLEDSEKVEEDEEEK	209	230
Q8NBR6	Protein FAM63B	**FAM63B**	NSERKRKEPREKDKEKEKEK	596	615
Q9UEE9	Craniofacial development protein 1	**CFDP1**	NSESEGSSSEEEDDAAEQEK	79	98
O00203	AP-3 complex subunit beta-1	**AP3B1**	NSKAKGKSDSEDGEKENEKSK	743	763
Q3B820	Protein FAM161A	**FAM161A**	NSQDSYKEKDEANEESEEEKS	635	655
Q4LE39	AT-rich interactive domain-containing protein 4B	**ARID4B**	NWKTELKEDSSSSEAEEEEEEEDDEKEKEDNSSEEEEEIEPFPEEREN	264	311
Q9UPY3	Endoribonuclease Dicer	**DICER1**	NYVSWSDSEDDDEDEEIEEKEK	408	429
Q9UQ35	Serine/arginine repetitive matrix protein 2	**SRRM2**	PAKRKRRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSPSPAKPGPQALPK	2600	2661
P21675	Transcription initiation factor TFIID subunit 1	**TAF1**	PDEKEEATSNSPSKESKKESS	514	534
Q8IZX4	Transcription initiation factor TFIID subunit 1-like	**TAF1L**	PDEKEEATSNSPSKESKKESS	533	553
Q9Y3C0	WASH complex subunit CCDC53	**CCDC53**	PDGESEKTVEESSDSESSFSD	174	194
Q8IWT3	Cullin-9	**CUL9**	PDPEEESKSEASFSEEETES	592	611
P50502	Hsc70-interacting protein	**ST13**	PDSKKVEEDLKADEPSSEESD	60	80
O60942	mRNA-capping enzyme	**RNGTT**	PDWCFEDDEDEDEDEDGKKESE	190	211
Q96NB3	Zinc finger protein 830	**ZNF830**	PDYEDEEEEEEEEEGDGERK	157	176
P46821	Microtubule-associated protein 1B	**MAP1B**	PEDKKEEKEKPKKEVAKKEDKT	646	667
Q9Y5X0	Sorting nexin-10	**SNX10**	PEEDEEGKKENDIDYDSESSSS	157	178
O60216	Double-strand-break repair protein rad21 homolog	**RAD21**	PELELLPEKEKEKEKEKEDDEEEEDEDASG	517	546
Q8IYB3	Serine/arginine repetitive matrix protein 1	**SRRM1**	PEPKEPSPEKNSKKEKEKEK	254	273
Q13367	AP-3 complex subunit beta-2	**AP3B2**	PERKGEASSSDEGSDSSSSSSESEM	751	775
Q8N554	Zinc finger protein 276	**ZNF276**	PERKVSGKKSESKEAKKSEE	387	406
Q8NG31	Protein CASC5	**CASC5**	PESTSESHSQSKSSSDECEE	621	640
P07197	Neurofilament medium polypeptide	**NEFM**	PEVKEEEGEKEEEEGQEEEEEEDEG	518	542
Q58F21	Bromodomain testis-specific protein	**BRDT**	PFRKLNKKKEKSKKEKKKEK	444	463
Q14839	Chromodomain-helicase-DNA-binding protein 4	**CHD4**	PGKKKKKKLGPKKEKKSKSK	112	131
O95835	Serine/threonine-protein kinase LATS1	**LATS1**	PKEDESEKSYENVDSGDKEKK	588	608
Q5VUA4	Zinc finger protein 318	**ZNF318**	PKEEKKEKKAKAVKEVKEDDK	1208	1228

P07197	Neurofilament medium polypeptide	**NEFM**	PKEEKVEKKEEKPKDVPEKKK	713	733
Q3YEC7	Rab-like protein 6	**RABL6**	PKESSEEGKEGKTPSKEKKKKKKKGKEEEEKAAKKKS KHKKSKDKEEGKEERRR	637	690
Q8N4C8	Misshapen-like kinase 1	**MINK1**	PKKAMDYSSSSEEVESSEDDEEEGE	817	841
P17096	High mobility group protein HMG-I/HMG-Y	**HMGA1**	PKKLEKEEEEGISQESSEEE	87	106
Q6UN15	Pre-mRNA 3'-end-processing factor FIP1	**FIP1L1**	PKPKVTETEDDSDSDSDDDEDD	74	95
O15085	Rho guanine nucleotide exchange factor 11	**ARHGEF11**	PKTKKSSNSKKEKDALEDKK	513	532
Q9NRR4	Ribonuclease 3	**DROSHA**	PKTKLDEDLESSSESECESDEDST	469	492
Q9Y330	Zinc finger and BTB domain-containing protein 12	**ZBTB12**	PLDEDLELKAEEEDEDEDED	185	204
O95983	Methyl-CpG-binding domain protein 3	**MBD3**	PLDKACAEDDDEEDEEEEEEE	261	281
Q92541	RNA polymerase-associated protein RTF1 homolog	**RTF1**	PLKTSEVYSDDEEEEEDDKSSEKSDRSSR	305	333
Q9NSY1	BMP-2-inducible protein kinase	**BMP2K**	PLLMDSEDEEEEEKHSSDSD	786	805
Q5H9B9	Putative BMP-2-inducible kinase-like protein	**BMP2KL**	PLLMDSEDEEEEEKRSSDSD	41	60
Q9UKE5	TRAF2 and NCK-interacting protein kinase	**TNIK**	PMKKVTDYSSSSEESESSEEEEEDGESET	822	850
Q8NI27	THO complex subunit 2	**THOC2**	PNEKDKEKNKSKSSGKEKGSDS	1523	1544
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	PPAKKAAESSSDSSDSDSSEDDEA	353	376
Q86UE4	Protein LYRIC	**MTDH**	PPEIDKKNEKSKKNKKKSKSD	152	172
P32745	Somatostatin receptor type 3	**SSTR3**	PPEKTEEEDEEEEDGEESRE	344	363
A0JLT2	Mediator of RNA polymerase II transcription subunit 19	**MED19**	PPETPSDSDHKKKKKKKEED	189	208
P25440	Bromodomain-containing protein 2	**BRD2**	PPGLAKSSSESSSEESSSESSSEEEEEEDEEDEEEEESESS DSEEERAHRLAELQEQ	470	526
P62851	40S ribosomal protein S25	**RPS25**	PPKDDKKKKDAGKSAKKDKD	2	21
Q8N488	RING1 and YY1-binding protein	**RYBP**	PPKKEKKEKVEKQDKEKPEKDKEIS	75	99
Q08211	ATP-dependent RNA helicase A	**DHX9**	PPPKDKKKKDKDDDGGEDDD	586	605
P19338	Nucleolin	**NCL**	PPPKEVEEDSEDEEMSEDEEDDSSGEEV	19	46
Q8IY57	YY1-associated factor 2	**YAF2**	PPTQSKKEKKDKVEKEKSEKE	70	90
Q5TCY1	Tau-tubulin kinase 1	**TTBK1**	PQANGKEEEEEEEEDEEEEEDEEEEEEEEEEEEE EEEEEEEEEAAAAVALGE	727	780
P51858	Hepatoma-derived growth factor	**HDGF**	PQEEEEEEDEEEEATKEDAE	211	230
Q96T23	Remodeling and spacing factor 1	**RSF1**	PQEPDFSEDQEEKKKDSKKSK	984	1004
O60885	Bromodomain-containing protein 4	**BRD4**	PQQNKPKKKEKDKKEKKKEK	531	550
Q14028	Cyclic nucleotide-gated cation channel beta-1	**CNGB1**	PRELSRIEEEKEDEEEEEEEEEEEEEEEVTEV	344	375
Q96JE7	Protein transport protein Sec16B	**SEC16B**	PRSISESSASSAKEDEKESSDE	864	885
Q92766	Ras-responsive element-binding protein 1	**RREB1**	PSDGESAAEKRSSEKSDDDKK	1474	1494
Q6H9L7	Isthmin-2	**ISM2**	PSEDIEGEDQEDKEEDEEEQ	276	295
Q9ULX3	RNA-binding protein NOB1	**NOB1**	PSEEEEEEENGFEDRKDDSDDD	183	204
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	PSKLEKKNDLKEDKISKEKEK	773	793
O75376	Nuclear receptor corepressor 1	**NCOR1**	PSQEEKVEEKEEDKAEKTEKKEEEKKDEEEKDEKE DSKENTKEKDKIDGTAEETEEREQA	508	567
P38432	Coilin	**COIL**	PSSSSESESCDESISDGPSK	244	263
Q7Z4V5	Hepatoma-derived growth factor-related protein 2	**HDGFRP2**	PSSSSSDSDSDEVDRISEWK	298	317
P26583	High mobility group protein B2	**HMGB2**	PTGSKKKNEPEDEEEEEEEEDEDEEEEDEDEE	178	209
Q8IUR6	CREB3 regulatory factor	**CREBRF**	PTKCSPEEDEEDEEDVDDED	349	368
P18583	Protein SON	**SON**	PTKKSKKHKKHKNKKKKKKKEKEKKYKR	104	131
P29144	Tripeptidyl-peptidase 2	**TPP2**	PTKTKNGSKDKEKDSEKEKD	997	1016
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	PVAKKAKKKASSSDSEDSSEEEEEV	73	97
Q13045	Protein flightless-1 homolog	**FLII**	PVEYEEEEKKEDKEEKAEGKE	943	963
Q13127	RE1-silencing transcription factor	**REST**	PVNDEESSTKKKKKVESKSK	533	552
P29084	Transcription initiation factor IIE subunit beta	**GTF2E2**	PVVEKRSASSESSSSSSKKKK	20	40
Q03169	Tumor necrosis factor alpha-induced protein 2	**TNFAIP2**	PYREEEEAAKKKKEKKKKSK	19	38
Q13428	Treacle protein	**TCOF1**	QAGKREEDSQSSSEESDSEEE	402	422
Q9NPG3	Ubinuclein-1	**UBN1**	QASESEDDFIKEKKKKSPKK	171	190

Q7Z4V5	Hepatoma-derived growth factor-related protein 2	**HDGFRP2**	QASVSPSEEENSESSSESEK	179	198
Q13428	Treacle protein	**TCOF1**	QATKAPESSDDSEDSSDSSS	1135	1154
Q9GZR7	ATP-dependent RNA helicase DDX24	**DDX24**	QAVSEEEEEEEGKSSSPKKK	79	98
O60315	Zinc finger E-box-binding homeobox 2	**ZEB2**	QDGDEEFEEEEEESENKSMD	1154	1173
Q8TF01	Arginine/serine-rich protein PNISR	**PNISR**	QDSKKSTTKDSKKHSGSDSS	738	757
P37275	Zinc finger E-box-binding homeobox 1	**ZEB1**	QEEKECEKPQGDEEEEEEEEE	1051	1071
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**	QEKDSGSDEDFLMEDDDDSD	126	145
Q4G0J3	La-related protein 7	**LARP7**	QEKVMEEESTEKKKEVEKKK	8	27
A6NCM1	Putative IQ and AAA domain-containing protein 1-like	**IQCA1P1**	QENRKKEQEKSKEKGKDEKEKKKGKEEKAKKGE	346	378
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	QEREETESKCDSEGEEDEED	550	569
Q63HN8	E3 ubiquitin-protein ligase RNF213	**RNF213**	QESKKKKRKKKKKGNKSASSE	86	106
Q8IVL6	Prolyl 3-hydroxylase 3	**LEPREL2**	QESQEEEEEEEEEMPSKDPS	692	711
O43670	BUB3-interacting and GLEBS motif-containing protein ZNF207	**ZNF207**	QESQKKKQQDDSDEYDDDDS	102	121
Q96BK5	PIN2/TERF1-interacting telomerase inhibitor 1	**PINX1**	QETTDSSDKKEKKSFSLEEKSK	97	118
P78332	RNA-binding protein 6	**RBM6**	QGKSSSKKEMSKRDGKEKKD	838	857
Q8NCQ7	Protein PROCA1	**PROCA1**	QGSKVIKKVKKKKEKEKDKEE	228	248
P49756	RNA-binding protein 25	**RBM25**	QIKQEPESEEEEEEKQEKEEK	576	596
P40818	Ubiquitin carboxyl-terminal hydrolase 8	**USP8**	QKAKEEMEKKESEQAKKEDKE	519	539
Q14160	Protein scribble homolog	**SCRIB**	QKEEEEEEEGSPQEEEVEEEEE	658	679
Q04637	Eukaryotic translation initiation factor 4 gamma 1	**EIF4G1**	QKEFEKDKDDDEVFEKKQKE	848	867
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	QKGDLGKDSEESEEDGEEEEESEEEEETSDLR	875	906
Q9NRF9	DNA polymerase epsilon subunit 3	**POLE3**	QKKKDKDKKTDSEEQDKSRDEDNDEDEERLEEEEQNEEEEVD	105	146
Q9H6L5	Protein FAM134B	**FAM134B**	QKKRERSEADKEKSHKDDSE	263	282
P06748	Nucleophosmin	**NPM1**	QKKVKLAADEDDDDDDEEDDDEDDDDDDFDDEEAEEKAPVKKSI	153	196
Q5VTR2	E3 ubiquitin-protein ligase BRE1A	**RNF20**	QKLKESEKERDSAKDKEKGK	595	614
P55196	Afadin	**MLLT4**	QKRLQESKQKDEDDEEEEDDD	1567	1587
O00401	Neural Wiskott-Aldrich syndrome protein	**WASL**	QKRSKAIHSSDEDEDEDDEED	476	496
O43283	Mitogen-activated protein kinase kinase kinase 13	**MAP3K13**	QKSGDDSSEEEEGEVDSEVE	809	828
Q92794	Histone acetyltransferase KAT6A	**KAT6A**	QKSREKIKDKEETELDSEEE	1357	1376
P31629	Transcription factor HIVEP2	**HIVEP2**	QKSSGPSESKESSDELDIDE	1554	1573
Q8N766	ER membrane protein complex subunit 1	**EMC1**	QKSSSSEDGSMGSFSEKSSSKDSL	343	366
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	QKSVSDPVEDKKEQESDEEEEEEEEDEPSG	2971	3000
P54257	Huntingtin-associated protein 1	**HAP1**	QLYSDSDEEDEDEEEEEEKEAEEEQEEEEAEEDLQC	254	290
O15042	U2 snRNP-associated SURP motif-containing protein	**U2SURP**	QNQEEESEDEEDTQSSKSEE	794	813
Q9UQ35	Serine/arginine repetitive matrix protein 2	**SRRM2**	QPAKERRSSSSSSSSSSSSSSSSSSSSSSSSSSGSSSSDSEGSSLPVQPE	2525	2573
Q96MG8	Protein-L-isoaspartate O-methyltransferase domain-containing protein 1	**PCMTD1**	QPLDSEEDEKMEEDNKEEEEKD	298	319
Q63HK5	Teashirt homolog 3	**TSHZ3**	QPSSEKNNGSSSSSSSSSS	140	159
Q5HYW3	Retrotransposon gag domain-containing protein 4	**RGAG4**	QQEMKEKEEEEMKKEEEMKKKEEKEEEEEEEEMKQKEE EEEIRNKNEEEGESKDEEDEDEDGGQKPE	370	435
Q9ULL8	Protein Shroom4	**SHROOM4**	QQQKQQEEEEEEEEEEEEEEEEEEEEAEEEEEELPP	1126	1161
Q8IVV2	Lipoxygenase homology domain-containing protein 1	**LOXHD1**	QRKKKKRKGSDEEDEGEEEESSSSEESSSEEEEMEEEEEEEEFGPGMQE	659	707
P38570	Integrin alpha-E	**ITGAE**	QRRALEKEEEEDKEEEEDEEEEEA	176	199
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	QRSSASESETEDSDDDKKPK	1096	1115

Q5VUA4	Zinc finger protein 318	**ZNF318**	QSELKRKLSEKPKEEKKEKK	1197	1216
Q15032	R3H domain-containing protein 1	**R3HDM1**	QSFEKEEKPSKDEAEKEKASDK	110	131
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B	**BAZ2B**	QSIEESEDDDSDSESEAQHKS	277	297
Q86XP1	Diacylglycerol kinase eta	**DGKH**	QSSDYDSTETDESKEEAKDD	652	671
Q8IYB3	Serine/arginine repetitive matrix protein 1	**SRRM1**	QSSSDSGSSSSSEDERPKRS	491	510
Q68DL7	Uncharacterized protein C18orf63	**C18orf63**	QSSSSKKQILDSDKSKLKKS	660	679
Q8WUB2	Protein FAM216A	**FAM216A**	QTVSSDDSESHMSEEKKEED	233	252
P51665	26S proteasome non-ATPase regulatory subunit 7	**PSMD7**	RDAEKKEGQEKEESKKDRKEDKEKDKDKEKSDVKKEEKKEKK	283	324
Q9NW82	WD repeat-containing protein 70	**WDR70**	RDCSKSSSRDTSSSESEQSSDSSDDELI	91	118
Q7L4S7	Protein ARMCX6	**ARMCX6**	RDDSEKLEEEGEEEWDDDQE	31	50
Q8N4S0	Coiled-coil domain-containing protein 82	**CCDC82**	RDFEDSEKESCPSSDEVDEEEEEDN	259	283
Q6PCB5	Round spermatid basic protein 1-like protein	**RSBN1L**	RDKIKDKIKERDKEKEREKKK	195	215
Q05519	Serine/arginine-rich splicing factor 11	**SRSF11**	RDYDEEEQGYDSEKEKKEEKK	423	443
P25788	Proteasome subunit alpha type-3	**PSMA3**	REEAEKYAKESLKEEDESDDD	233	253
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	REEEEEEKEEFEEERSREEKRS	462	483
Q6ZQQ6	WD repeat-containing protein 87	**WDR87**	REEEEEREEEEVREEEEERKEEEEGEEKQ	2251	2279
Q7Z6M4	Transcription termination factor 4, mitochondrial	**MTERF4**	REEEESESSTSDDKRASLDEDEDDDDEEDNDEDDNDED DDDEDDDEAEDNDEDEDDDEEE	322	381
Q13409	Cytoplasmic dynein 1 intermediate chain 2	**DYNC1I2**	REEKKRKEEERKKKETDQKKE	22	42
P35659	Protein DEK	**DEK**	REESEEEEDEDDEEEEEEEKEKSLIVEGKREKKKVERL	29	66
Q9BSC4	Nucleolar protein 10	**NOL10**	REKEEEEEPEGKPSDAESSESSDDEKAW	530	557
O15042	U2 snRNP-associated SURP motif-containing protein	**U2SURP**	REKEKELERERERDKKDKEK	885	904
Q8TC05	Nuclear protein MDM1	**MDM1**	REKSKADKMKEGSDSSVSSEK	489	509
Q9NRF9	DNA polymerase epsilon subunit 3	**POLE3**	REQKGKKEASEQKKKDKDKK	94	113
Q9C0B9	Zinc finger CCHC domain-containing protein 2	**ZCCHC2**	RESFESEEEKDRDTDSNSEDS	647	667
Q9BUQ8	Probable ATP-dependent RNA helicase DDX23	**DDX23**	RGKDFKSRKDRDSKKDEEDE	112	131
Q8N5B7	Ceramide synthase 5	**CERS5**	RGKVSKDDRSDVESSSEEED	341	360
Q05519	Serine/arginine-rich splicing factor 11	**SRSF11**	RHKKEKKKDKDKERSRDERE	373	392
O75400	Pre-mRNA-processing factor 40 homolog A	**PRPF40A**	RHKSDSPESDAEREKDKKEKD	880	900
Q15858	Sodium channel protein type 9 subunit alpha	**SCN9A**	RIAERKSKEPKEEKKDDDEE	26	45
Q92769	Histone deacetylase 2	**HDAC2**	RIEEDKKETEDKKTDVKEEDKSKDNS	446	471
A3KN83	Protein strawberry notch homolog 1	**SBNO1**	RKDHKKNKEKKKKKSIDPDS	780	799
Q8NDX6	Zinc finger protein 740	**ZNF740**	RKDSDKSRSRKDDDSLSEAS	56	75
Q03111	Protein ENL	**MLLT1**	RKDSESKSSSKELEREQAKSSKDTS	206	230
Q70Z53	Protein FRA10AC1	**FRA10AC1**	RKEIKSKKRKDKTKKDCEESS	224	244
Q96IZ7	Serine/Arginine-related protein 53	**RSRC1**	RKGRDKEKREKEKDKGKDKE	139	158
Q5BKY9	Protein FAM133B	**FAM133B**	RKKKEKKKSGRYSSSSSSSSDSSSSSSDSEDEDKKQGKRRKK KKNRSHKSSESSMSETESDSKDSLKKKKKSKDGTEKEKDIKG LSKKRKMYSEDKPLSSESLSESEYIEEV	93	204
Q5TCS8	Adenylate kinase 9	**AK9**	RKKKKEAQKDGKGEEEEEEEE	185	205
Q8N9E0	Protein FAM133A	**FAM133A**	RKKKRKKKSCRSSSSSSSSDSSSSSSDSEDEEKKQ GKRRKKKKNRSYKSSQSST	93	146
Q7Z388	Probable C-mannosyltransferase DPY19L4	**DPY19L4**	RKKPKSSENKESAKEEKISD	14	33
Q9P2B7	UPF0501 protein KIAA1430	**KIAA1430**	RKKYCKVSSSSSSLSSSS	163	182
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	RKNSRKHKKEKKKRKKSKKS	235	254
Q9NR30	Nucleolar RNA helicase 2	**DDX21**	RKQTEEKEKKEKPKSDKTEE	24	43
Q5T2S8	Armadillo repeat-containing protein 4	**ARMC4**	RKSAEKIEETVSDSSSESEEDEEP	413	436
Q5T200	Zinc finger CCCH domain-containing protein 13	**ZC3H13**	RKSSKSPKRKSSPKSSSASKKD	213	234
Q96N46	Tetratricopeptide repeat protein 14	**TTC14**	RKSTSSSSVSSADESVSSSSSSSSSG	473	498
P51825	AF4/FMR2 family member 1	**AFF1**	RLEKEIKSQSSSSSSSHKESSK	828	849
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	**ACIN1**	RLKQSADSSSSRSSSSSSSSS	569	589
Q96T23	Remodeling and spacing factor 1	**RSF1**	RLSSSESEESYLSKNSEDDE	1243	1262

Q6ZMV5	Putative SMEK homolog 3	**SMEK3P**	RLVDHPDDEEEKEEDEEEKEEDKEDETSPKKKP	796	828
Q9H307	Pinin	**PNN**	RNKTSKSRSRSSSSSSSSSS	567	587
O75807	Protein phosphatase 1 regulatory subunit 15A	**PPP1R15A**	RPGEDTEEEEDEDVDSEDKEDDSEA	435	459
P51531	Probable global transcription activator SNF2L2	**SMARCA2**	RQKIAKEEESEDESNEEEEEEDEEESESEAKSVKVKIK	1503	1540
Q96L91	E1A-binding protein p400	**EP400**	RQNDLDIEEEEEEEEEEEKSE	419	440
Q9UQ35	Serine/arginine repetitive matrix protein 2	**SRRM2**	RRERKKSSKKKKHRSESESKK	206	226
Q96T23	Remodeling and spacing factor 1	**RSF1**	RRKTPKKKYSDDDEEEESEE	1163	1182
O00267	Transcription elongation factor SPT5	**SUPT5H**	RSAAGSEKEEEPEDEEEEEEEEEYDEEEEEEDDDRPPKKP	31	70
Q5VUB5	Protein FAM171A1	**FAM171A1**	RSAHEEEEDDDDDDQGEDKKS	854	874
Q7Z4V5	Hepatoma-derived growth factor-related protein 2	**HDGFRP2**	RSASSSSSSSSSDSDVSVKKP	252	273
P78411	Iroquois-class homeodomain protein IRX-5	**IRX5**	RSEDEEEEENIDLEKNDEDE	184	203
Q8TC27	Disintegrin and metalloproteinase domain-containing protein 32	**ADAM32**	RSESSSQADTSKSKSEDSAE	749	768
Q9H4Z3	Phosphorylated CTD-interacting factor 1	**PCIF1**	RSHSSGSSSSSSSEAKDRDS	672	691
Q7Z6E9	E3 ubiquitin-protein ligase RBBP6	**RBBP6**	RSHSSSASSAESQDSKKKKKKKEKKKHKKHKKHKKHKK	1712	1749
Q9Y2F9	BTB/POZ domain-containing protein 3	**BTBD3**	RSKKSSKKANTSSSSSNSSK	25	44
Q14781	Chromobox protein homolog 2	**CBX2**	RSKLKEPDAPSKSKSSSSSSS	92	113
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	RSQLSSSSTSSTSDKDDDDDDSDDRDESENDDEDTA	1090	1125
Q7L014	Probable ATP-dependent RNA helicase DDX46	**DDX46**	RSRSKEKTDGGESSKEKKKDKDDKEDEKEKDAGN	117	150
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	RSSLPVTEDEKEEESSEEEDEDKR	145	168
A7MD48	Serine/arginine repetitive matrix protein 4	**SRRM4**	RSSSYSPSPVKKKKKKSSKK	122	141
Q8N9E0	Protein FAM133A	**FAM133A**	RSYKSSQSSTHESESESKES	137	156
Q7KZ85	Transcription elongation factor SPT6	**SUPT6H**	RVKKMSDDEDDDEEEYGKEE	120	139
Q8IU57	Interferon lambda receptor 1	**IFNLR1**	RWKKDLAEDEEEEDEEDTED	317	336
Q96M86	Dynein heavy chain domain-containing protein 1	**DNHD1**	SADLKTSSDSLYSEEEDEEEDSKDEF	599	624
Q86Y26	NUT family member 1	**NUTM1**	SADRAKGKEKKKKEAEEEDEE	996	1016
P20962	Parathymosin	**PTMS**	SAKDLKEKKEKVEEKASRKE	13	32
Q5CZC0	Fibrous sheath-interacting protein 2	**FSIP2**	SAKEKEEEEREKEKVREEIKSE	5743	5764
P23497	Nuclear autoantigen Sp-100	**SP100**	SAKKRVVKAEKSKKKKEEEEDEEDEQEEENEEDDDK	844	879
Q9UBN4	Short transient receptor potential channel 4	**TRPC4**	SANASKESSNSADSDEKSDSE	764	784
P05387	60S acidic ribosomal protein P2	**RPLP2**	SAPAAAEEKKDEKKEESEESDDDM	86	109
Q13428	Treacle protein	**TCOF1**	SAQVGKWEEDSESSSEESSDSSDGE	535	559
Q15386	Ubiquitin-protein ligase E3C	**UBE3C**	SASCHDSASDSEEESEEADK	358	377
O75400	Pre-mRNA-processing factor 40 homolog A	**PRPF40A**	SASEHSSSAESERSYKKSKK	852	871
Q99543	DnaJ homolog subfamily C member 2	**DNAJC2**	SASFQELEDKKELSEESEDEE	47	67
Q2V2M9	FH1/FH2 domain-containing protein 3	**FHOD3**	SASSVSSSSSTLEREEKEDK	626	645
P07814	Bifunctional glutamate/proline--tRNA ligase	**EPRS**	SATGAEDKDKKKKEKENKSEK	954	974
Q96PV7	Protein FAM193B	**FAM193B**	SCKSQSCGDDSHSSSSSSSSSSSSSSSCP	158	187
Q9NQ75	Cas scaffolding protein family member 4	**CASS4**	SCSTTSTDDSSSSSSEESAKE	412	432
Q96ST2	Protein IWS1 homolog	**IWS1**	SDADDSDSDAVSDKSGKREK	415	434
P35663	Cylicin-1	**CYLC1**	SDAESEDSKDAKKDSKKVKKNVKKDDKKKDVK KDTESTDAESGDSKDERKDTKKDKKKLKKDDK KKDTKKYPESTDTESGDAKDA	296	380
O95602	DNA-directed RNA polymerase I subunit RPA1	**POLR1A**	SDAKRKEKQEEEVDYESEEEEE	1413	1434
Q96QF7	Acidic repeat-containing protein	**ACRC**	SDDSEASDDSSDDSEAPDDKSDDSDVPEDKSDDSDVPDDNSDDLEVPV	289	336
O43719	HIV Tat-specific factor 1	**HTATSF1**	SDEDCSEKQSEDGSEREFEE	548	567
Q58FF6	Putative heat shock protein HSP 90-beta 4	**HSP90AB4P**	SDEEDDTDKNNKKKTKKIKEK	228	248
Q5SNT6	WASH complex subunit FAM21B	**FAM21B**	SDEEEDDDGCDLFADSEKEEED	162	183
Q641Q2	WASH complex subunit FAM21A	**FAM21A**	SDEEEDDDGCDLFADSEKEEED	250	271
Q9Y4E1	WASH complex subunit FAM21C	**FAM21C**	SDEEEDDDGCDLFADSEKEEED	250	271
Q6ZQQ6	WD repeat-containing protein 87	**WDR87**	SDEEEEEESCSLEEEVDREKEI	2322	2343
Q8WWQ0	PH-interacting protein	**PHIP**	SDEEEESEKQKQKQIKKEKKKV	881	902
Q8WVC0	RNA polymerase-associated protein LEO1	**LEO1**	SDEEHRHSDDEEEQDHKSES	254	273
O15318	DNA-directed RNA polymerase III subunit RPC7	**POLR3G**	SDEENEEKEGSKEKSKEGDDDDDDDAAEQEEYDEEEQEEENDYI	157	200

P26358	DNA (cytosine-5)-methyltransferase 1	**DNMT1**	SDEKDEDEKEEKRRKTTPKE	312	331
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	SDEKDSGKTLDKKPSKEMSSDSEYDSDDDR	570	599
Q5F1R6	DnaJ homolog subfamily C member 21	**DNAJC21**	SDENEMEEHELKDEEDGKDSDE	283	304
Q15397	Pumilio domain-containing protein KIAA0020	**KIAA0020**	SDESAAKKPKWDDFKKKKKE	100	119
Q13061	Triadin	**TRDN**	SDIISSEDEEDDDGDEDTDKGE	111	132
Q569K4	Zinc finger protein 385B	**ZNF385B**	SDKSEDKGKLKASSSSQPSSSESGS	220	244
Q9HAZ2	PR domain zinc finger protein 16	**PRDM16**	SDLDSDVDSDPDKDKGKGKS	623	642
P46100	Transcriptional regulator ATRX	**ATRX**	SDMMENSKEEGTSSSEKSKSS	54	74
Q5VZP5	Inactive dual specificity phosphatase 27	**DUSP27**	SDNKRSSLFKKKKVKEDEDD	861	880
Q9BQI7	PH and SEC7 domain-containing protein 2	**PSD2**	SDPSLKDGLSDSDSELSSSE	274	293
Q8IYB3	Serine/arginine repetitive matrix protein 1	**SRRM1**	SDQEGGGKKKKKKKDKKHKKDKKHKK	802	827
Q5SNT6	WASH complex subunit FAM21B	**FAM21B**	SDRGSIVDTEEEKEEEESDED	125	145
Q641Q2	WASH complex subunit FAM21A	**FAM21A**	SDRGSIVDTEEEKEEEESDED	213	233
Q9Y4E1	WASH complex subunit FAM21C	**FAM21C**	SDRGSIVDTEEEKEEEESDED	213	233
Q7Z4V5	Hepatoma-derived growth factor-related protein 2	**HDGFRP2**	SDRMESDSDSDKSSDNSGLK	137	156
P46100	Transcriptional regulator ATRX	**ATRX**	SDSDETSMSLSSDDYTKKKKK	1913	1933
Q99549	M-phase phosphoprotein 8	**MPHOSPH8**	SDSDQQSETKEDTSPKKKKKK	136	156
P01106	Myc proto-oncogene protein	**MYC**	SDSEEEQEDEEEIDVVSVEK	250	269
A1A519	Protein FAM170A	**FAM170A**	SDSEPSGEEKEHEERTESDS	183	202
Q96ST2	Protein IWS1 homolog	**IWS1**	SDSESEELPKPQVSDSESEE	209	228
Q96KR1	Zinc finger RNA-binding protein	**ZFR**	SDSLSEHEKNKNKEGDDKKE	759	778
Q9C0F0	Putative Polycomb group protein ASXL3	**ASXL3**	SDSPSGSESKNGEADSSDKE	194	213
Q76L83	Putative Polycomb group protein ASXL2	**ASXL2**	SDSQSSENSSSSSDGGSNKE	106	125
O43623	Zinc finger protein SNAI2	**SNAI2**	SDTSSKDHSGSESPISDEEE	92	111
P51531	Probable global transcription activator SNF2L2	**SMARCA2**	SDYEEEDEEEESSRQETEEKI	640	660
P46100	Transcriptional regulator ATRX	**ATRX**	SEDDKKQSKKGTEEKKKPSDFKKKV	978	1002
Q8NHP6	Motile sperm domain-containing protein 2	**MOSPD2**	SEDETSSKEDIESDGKETLE	261	280
Q9Y496	Kinesin-like protein KIF3A	**KIF3A**	SEEDDDEEGEVGEDGEKRKK	388	407
O14576	Cytoplasmic dynein 1 intermediate chain 1	**DYNC1I1**	SEEDEEDEEMVESKVGQDSE	179	198
Q9HCG8	Pre-mRNA-splicing factor CWC22 homolog	**CWC22**	SEEDEEEEEEEGEEDEEGQKVT	422	443
Q6RI45	Bromodomain and WD repeat-containing protein 3	**BRWD3**	SEEDENMGGEDKEKKETKEKS	1579	1599
O43719	HIV Tat-specific factor 1	**HTATSF1**	SEEDSPKKESKKKTLKNDCEE	494	514
Q14687	Genetic suppressor element 1	**GSE1**	SEEEEEEDDEDGEDEEEVPKRKWQ	1101	1124
P51532	Transcription activator BRG1	**SMARCA4**	SEEEEEGEEEGSESESRSVKVK	1575	1596
Q96PN7	Transcriptional-regulating factor 1	**TRERF1**	SEEEEELEEEEEEDPEEDRKSTKEEESEVPKSPEPP	955	990
Q8NE71	ATP-binding cassette sub-family F member 1	**ABCF1**	SEEEGEGEEEEEEGGESKADD	228	248
Q9Y6V0	Protein piccolo	**PCLO**	SEEEIKESQEERKDTFKKDS	1395	1414
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	SEEEREKSSCDETESDYEPK	166	185
Q2TAZ0	Autophagy-related protein 2 homolog A	**ATG2A**	SEEKEDEREEEGDGDTLDSDE	1341	1361
Q13316	Dentin matrix acidic phosphoprotein 1	**DMP1**	SEESHSEEDDSDSQDSSRSKEDSNSTESKSSSEEDGQLKNIE	444	485
Q01538	Myelin transcription factor 1	**MYT1**	SEESSKQKGILSHEEEDEEEEEEEEEEEDEEEEE EEEEEEEEEEEEEEEEEEEEEEEEEAAPDVIFQ EDTSHTSAQKAPELR	245	329
P49593	Protein phosphatase 1F	**PPM1F**	SEFRKLPREEEEEEEDDDEEEK	93	114
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	SEGKSDSCSSDSETEDALEK	689	708
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	SEGSGSSEEDEWRSDRKSES	834	853
Q9Y5B6	PAX3- and PAX7-binding protein 1	**PAXBP1**	SEHGEDEMDMESEKEEEKPK	180	199
O75554	WW domain-binding protein 4	**WBP4**	SEIKQKSLDKAKEEEKASKE	45	64
Q8N7H5	RNA polymerase II-associated factor 1 homolog	**PAF1**	SEKEGSEDEHSGSESEREEGD	404	424
Q9ULU8	Calcium-dependent secretion activator 1	**CADPS**	SEKEKEELERLQKEEEERKK	98	117
Q9Y2B4	TP53-target gene 5 protein	**TP53TG5**	SEKELKSKKLESTGDPKKKE	106	125
Q8NFW9	Rab effector MyRIP	**MYRIP**	SEKETSSGEDQESEPKTESE	595	614
Q9BVI0	PHD finger protein 20	**PHF20**	SEKGKVSEKSLPKNEKEDKE	203	222

Q8N3X1	Formin-binding protein 4	**FNBP4**	SEKIDENSDKEMEVEESPEK	492	511
Q92766	Ras-responsive element-binding protein 1	**RREB1**	SEKSDDDKKPKTDSPKSVASK	1486	1506
O43719	HIV Tat-specific factor 1	**HTATSF1**	SEKVLDEEGSEREFDEDSDEKEEEEDTYEK	607	636
O75475	PC4 and SFRS1-interacting protein	**PSIP1**	SESDIITEEDKSKKKGQEEK	206	225
Q7Z5K2	Wings apart-like protein homolog	**WAPAL**	SESEDDEDDDCQVERKTSKK	459	478
Q96GN5	Cell division cycle-associated 7-like protein	**CDCA7L**	SESEDDSRDESQESSDALLK	197	216
O00203	AP-3 complex subunit beta-1	**AP3B1**	SESEPESESESRRVTKEKEKK	784	804
Q5T3I0	G patch domain-containing protein 4	**GPATCH4**	SESKPPKKKKKKRRQKEEEE	197	216
Q9BXX2	Ankyrin repeat domain-containing protein 30B	**ANKRD30B**	SESKQKDDEENSWDSESPCE	563	582
Q9BXX3	Ankyrin repeat domain-containing protein 30A	**ANKRD30A**	SESKQKDYEESSWDSESLCE	792	811
Q86TI0	TBC1 domain family member 1	**TBC1D1**	SESSFKLLGSSEDLSSDSES	556	575
Q7Z4V5	Hepatoma-derived growth factor-related protein 2	**HDGFRP2**	SESSSESEKTSDQDFTPEKK	190	209
Q86X95	Corepressor interacting with RBPJ 1	**CIR1**	SESSSESESNNKEKKIQRKK	274	293
Q9ULU4	Protein kinase C-binding protein 1	**ZMYND8**	SETASASKEKETSAEKSKES	1098	1117
Q2PPJ7	Ral GTPase-activating protein subunit alpha-2	**RALGAPA2**	SETDSKEASSESSGHKRSSS	461	480
Q8N108	Mesoderm induction early response protein 1	**MIER1**	SEVEEESEEDEDYIPSEDWKKE	160	181
Q13061	Triadin	**TRDN**	SEVTESGKKKTEISEKESKEK	605	625
Q8WZ42	Titin	**TTN**	SEYKSELDFMSKEEKSRKKS	33143	33162
Q03111	Protein ENL	**MLLT1**	SFKSESAQSSPSNSSSSSDSSSDSDFEPS	366	394
Q9Y6R1	Electrogenic sodium bicarbonate cotransporter 1	**SLC4A4**	SFLDDVIPEKDKKKKEDEKKKKKKKGSLDSDNDDSDC	1000	1036
Q5TID7	Coiled-coil domain-containing protein 181	**CCDC181**	SFQESKLESQKDLEEEEDEE	103	122
Q7RTP6	Protein-methionine sulfoxide oxidase MICAL3	**MICAL3**	SFSEDSDLSSDDVLEKSSQKS	1792	1812
Q9UHI6	Probable ATP-dependent RNA helicase DDX20	**DDX20**	SGDSESDSDSYSSRTSSQSK	649	668
Q562E7	WD repeat-containing protein 81	**WDR81**	SGDSSQDLKQSEGSEEEEEEEDSC	1139	1162
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein	**EIF3CL**	SGESRKFLKKMDDEDEDSEDSEDDEDWDT	205	233
Q99613	Eukaryotic translation initiation factor 3 subunit C	**EIF3C**	SGESRKFLKKMDDEDEDSEDSEDDEDWDT	205	233
A8MTJ3	Guanine nucleotide-binding protein G(t) subunit alpha-3	**GNAT3**	SGISSESKESAKRSKELEKK	3	22
O14958	Calsequestrin-2	**CASQ2**	SGKINTEDDDEDDDDDDNSDEEDNDDSDDDDDE	367	399
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	SGKMKSSKSEGSSSSKLSSSM	1111	1131
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**	SGKNSQEDSEDSEDKDVKTKKDDSHS	50	75
Q8N9E0	Protein FAM133A	**FAM133A**	SGNESSSKKRERKKKRKKKS	82	101
Q13367	AP-3 complex subunit beta-2	**AP3B2**	SGPTESADSDPESESESDSKSSSESGSGESSSESDNEDQDE DEEKGRGSESEQSEEDGKRK	685	745
Q8TF01	Arginine/serine-rich protein PNISR	**PNISR**	SGSDSSGRSSSESPGSSKEKK	752	772
P22459	Potassium voltage-gated channel subfamily A member 4	**KCNA4**	SGSEEKILRELSEEEEDEEEEEEEEEGR	111	139
Q9BXB5	Oxysterol-binding protein-related protein 10	**OSBPL10**	SGSELVLSEDEKSDNEDKEE	368	387
O14646	Chromodomain-helicase-DNA-binding protein 1	**CHD1**	SGSPSQSGSDSESEEEREKSS	154	174
Q14201	Protein BTG3	**BTG3**	SGSSSSDEETSKEMEVKPSS	144	163
Q9NZN5	Rho guanine nucleotide exchange factor 12	**ARHGEF12**	SHDFDPTDSSSKKTKSSSEES	41	61
Q6ZUT1	Uncharacterized protein C11orf57	**C11orf57**	SHSDDSASSSSEESEERDTKKTKRKK	228	253
P55201	Peregrin	**BRPF1**	SHSEGEEDEDEEEDEGKGWSSEKVKK	460	485
Q8NEY1	Neuron navigator 1	**NAV1**	SHSSIGSSKDADAKKKKKKS	1193	1212
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	SHSSSSSSSASTSGKMKSSKSE	1099	1120
Q96N64	PWWP domain-containing protein 2A	**PWWP2A**	SIDSTDDLKSSNSECSSSES	584	603
P0C860	Putative male-specific lethal-3 protein-like 2	**MSL3P1**	SIEEKNENDENSLSSSSDSSEDKDEKISEECDIEEKTE	42	79

Q6DN12	Multiple C2 and transmembrane domain-containing protein 2	**MCTP2**	SIQDSQESTDIDDEEDEDDKESEKKG	731	756
Q5FWF5	N-acetyltransferase ESCO1	**ESCO1**	SIQEKSKENSSKVTKKSDDK	3	22
Q5VUA4	Zinc finger protein 318	**ZNF318**	SISKEEILESSKDKEDGKTE	1294	1313
Q86SE9	Polycomb group RING finger protein 5	**PCGF5**	SKADKPKVDEEGDENEDDKD	112	131
O15397	Importin-8	**IPO8**	SKAEKADMEENEEISSDEEE	888	907
Q2WGJ9	Fer-1-like protein 6	**FER1L6**	SKALKELKLPSKDKDSKSSK	416	435
Q9HAU5	Regulator of nonsense transcripts 2	**UPF2**	SKAPEDKKKRLEDDKRKKEDKE	50	71
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	**NUCKS1**	SKASKEKTPSPKEEDEEPES	195	214
P35663	Cylicin-1	**CYLC1**	SKDERKDTKKDKKKLKKDDKKKDTKKYPESTDTESGD	340	376
Q7Z5L7	Podocan	**PODN**	SKDRGRLGKEKEEEEEEEEEEET	589	612
Q7L4I2	Arginine/serine-rich coiled-coil protein 2	**RSRC2**	SKEGRRHESKDKSSKKHKSEE	67	87
Q13123	Protein Red	**IK**	SKEKEEEELMEKPQKETKKDED	180	201
O60870	DNA/RNA-binding protein KIN17	**KIN**	SKEKKKKKSALDEIMEIEEEKK	251	272
Q13427	Peptidyl-prolyl cis-trans isomerase G	**PPIG**	SKEKSKQLESKSNEHDHSKSKEKDR	516	540
Q13523	Serine/threonine-protein kinase PRP4 homolog	**PRPF4B**	SKERKSKRSETDKEKKPIKS	309	328
Q6UB98	Ankyrin repeat domain-containing protein 12	**ANKRD12**	SKERNFKEERDKIKKESEKS	742	761
Q9Y2D9	Zinc finger protein 652	**ZNF652**	SKETPVLKTSSEEEEEESEEE	141	161
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	SKGEEKGNENEDENKDSEKS	2841	2860
Q8TAM6	Ermin	**ERMN**	SKHDEEQKVWDEEIDDDDDD	173	192
Q8IX21	Protein FAM178A	**FAM178A**	SKKDKERSSSKECSGHSTES	499	518
P49711	Transcriptional repressor CTCF	**CTCF**	SKKEDSSDSENAEPDLDDNEDEEEP	604	628
Q92889	DNA repair endonuclease XPF	**ERCC4**	SKKEKISEKMEIKEGEETKKE	352	372
Q14093	Cylicin-2	**CYLC2**	SKKGKDIEKGKEEKLDAKKDSKKGKKDAEK	141	170
Q15021	Condensin complex subunit 1	**NCAPD2**	SKKKPKVVFSSDESSEEDLS	1357	1376
O75164	Lysine-specific demethylase 4A	**KDM4A**	SKKKSSSSLGSGSSRDSISSDSETSE	504	529
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like 1	**BOD1L1**	SKKSEDTQKVKDEKQAKEKE	349	368
Q9P0M6	Core histone macro-H2A.2	**H2AFY2**	SKKSKPKDSDKEGTSNSTSED	157	177
Q92772	Cyclin-dependent kinase-like 2	**CDKL2**	SKKSQNRKKEKEKDDSLVEE	314	333
Q5TCQ9	Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 3	**MAGI3**	SKMERMDSSLPEEEEDEDKE	227	246
P54577	Tyrosine--tRNA ligase, cytoplasmic	**YARS**	SKMSSSEEESKIDLLDRKED	221	240
Q6NUN7	Uncharacterized protein C11orf63	**C11orf63**	SKPFSELSDSDLEEKSSSLS	207	226
P12036	Neurofilament heavy polypeptide	**NEFH**	SKPKAEKAEKSSSTDQKDSK	993	1012
P55199	RNA polymerase II elongation factor ELL	**ELL**	SKPKKKSKKHKDKERAAEDK	446	465
P25440	Bromodomain-containing protein 2	**BRD2**	SKPKRKREKKEKKKKRKAEK	543	562
O95835	Serine/threonine-protein kinase LATS1	**LATS1**	SKPSKEDQPSLPKEDESEKS	577	596
Q9UGU5	HMG domain-containing protein 4	**HMGXB4**	SKPSKKTGEKSSGSSSHSESKKEH	124	147
Q14839	Chromodomain-helicase-DNA-binding protein 4	**CHD4**	SKRKRSSSEDDDLDVESDFDD	303	323
Q9NWB6	Arginine and glutamate-rich protein 1	**ARGLU1**	SKRSSLDEKQKREEEEKKAE	90	109
Q14008	Cytoskeleton-associated protein 5	**CKAP5**	SKSDKETEKGASRIDEKSSK	1807	1826
Q8NI27	THO complex subunit 2	**THOC2**	SKSKEREMDKKDLDKSRERS	1448	1467
Q9UHB7	AF4/FMR2 family member 4	**AFF4**	SKSMLKDDLKLSSSEDSDGE	376	395
P10451	Osteopontin	**SPP1**	SKSNESHDHMDDMDDEDDDD	76	95
O95218	Zinc finger Ran-binding domain-containing protein 2	**ZRANB2**	SKSRSSHSRSSSRSSSPSSS	204	223
Q03164	Histone-lysine N-methyltransferase 2A	**KMT2A**	SKSSSLKGEKTKVLSSKSSE	2298	2317
Q14781	Chromobox protein homolog 2	**CBX2**	SKSSSSSSSSTSSSSSDEEDDSDLDAKRGPRG	104	136
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	SKTKFLKIESSEEDSKSHDSD	1746	1766
Q9P2D1	Chromodomain-helicase-DNA-binding protein 7	**CHD7**	SLDGSQEEKKKKKRSKAKKD	633	652
Q9Y266	Nuclear migration protein nudC	**NUDC**	SLDSPGKQDTEEDEEEDEKDK	136	156
Q9H091	Zinc finger MYND domain-containing protein 15	**ZMYND15**	SLEDGEEGEEEEEEDEEEEKREDGG	108	132
Q9HCK1	DBF4-type zinc finger-containing protein 2	**ZDBF2**	SLEDKSSNSYSPEESSDSNDS	1353	1373

Q16352	Alpha-internexin	**INA**	SLKKEEEEEEASKVASKKTS	445	464
Q92620	Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16	**DHX38**	SLKRREREEKDDGEDKKKSK	56	75
Q96CJ1	ELL-associated factor 2	**EAF2**	SLMDQMSSCDSSSDSKSSSSSSSEDSSSDSEDEDCKSSTSDTGNC	168	212
P22670	MHC class II regulatory factor RFX1	**RFX1**	SLNPLDPDKDEEEEEEEESEDE	915	936
P55197	Protein AF-10	**MLLT10**	SLSDSSSHSQDKHHEKEKKK	215	234
P18846	Cyclic AMP-dependent transcription factor ATF-1	**ATF1**	SLSESEESQDSSDSIGSSQK	34	53
C9JSJ3	Basic helix-loop-helix and HMG box domain-containing protein 1	**BHMG1**	SLSGNSKAPSSSSSSSSSSSSEDSDSEPL	444	473
Q9H7U1	Serine-rich coiled-coil domain-containing protein 2	**CCSER2**	SLSSLSSSDKNDLSEDFSDD	398	417
O60315	Zinc finger E-box-binding homeobox 2	**ZEB2**	SMDDSSEDGKMETKSDHEED	1189	1208
P07197	Neurofilament medium polypeptide	**NEFM**	SMKEEKKEAAEEKEEEPEAEEEEV	483	506
Q9BZF1	Oxysterol-binding protein-related protein 8	**OSBPL8**	SMSKSKSESKLYNGSEKDSS	91	110
Q8NI27	THO complex subunit 2	**THOC2**	SNKAVKENDKEKGKEKEKEKKEKT	1262	1285
Q9ULK2	Ataxin-7-like protein 1	**ATXN7L1**	SNKKRKPQSSTSSSSSSSSS	641	661
Q8NEJ9	Neuroguidin	**NGDN**	SNMMSKLSSEDEEEDEAEDD	135	154
Q6PD62	RNA polymerase-associated protein CTR9 homolog	**CTR9**	SNSDSDEDEQRKKCASSESDSDEN	1039	1062
Q6NWY9	Pre-mRNA-processing factor 40 homolog B	**PRPF40B**	SPESETDPEEKAGKESDEKE	795	814
Q9Y3E1	Hepatoma-derived growth factor-related protein 3	**HDGFRP3**	SPGDEDDKDCKEEENKSSSEG	162	182
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	SPNKKKKKKLKDKKEKKAKRKKKDEDEDDNDDGC	94	127
Q5TF21	Protein SOGA3	**SOGA3**	SPRDSDAESDAGKKESDDDS	747	766
P35251	Replication factor C subunit 1	**RFC1**	SPSKKESESKKSRPTSKRDS	518	537
Q9Y388	RNA-binding motif protein, X-linked 2	**RBMX2**	SPSLSESSEDEKPTKKHKKDKKEKKKKKKEKEKADREVQAE	142	182
Q9BYP7	Serine/threonine-protein kinase WNK3	**WNK3**	SPSSPMSSDDESEIEDEDLK	1515	1534
Q13523	Serine/threonine-protein kinase PRP4 homolog	**PRPF4B**	SPTDDKVKIEDKSKSKDRKKS	257	277
Q15287	RNA-binding protein with serine-rich domain 1	**RNPS1**	SPTKRKDRSDEKSKDRSKDK	27	46
Q3KQU3	MAP7 domain-containing protein 1	**MAP7D1**	SPVQKKEKKDKERENEKEKS	410	429
Q2TBE0	CWF19-like protein 2	**CWF19L2**	SQEHSVKKKKKKDKHSKKAKKEKKKKSKKQKYEKNN ESSDSSSSSEDEWVEAV	71	123
Q8WWA1	Transmembrane protein 40	**TMEM40**	SQEQYERNKSSSSSSSSSSSSSSSSSSSESNDEDQQPR	41	80
P98174	FYVE, RhoGEF and PH domain-containing protein 1	**FGD1**	SQEVDSDLEEEDDEEEEEEKDREIP	332	356
Q5SW79	Centrosomal protein of 170 kDa	**CEP170**	SQKSSESELSKSASAKSIDSK	122	142
A2RTX5	Probable threonine--tRNA ligase 2, cytoplasmic	**TARSL2**	SQSQDKDMKKKKMKESEADSE	103	123
Q08170	Serine/arginine-rich splicing factor 4	**SRSF4**	SQSRSRSKKEKSRSPSKEKS	236	255
Q13428	Treacle protein	**TCOF1**	SQTKAGKPEEDSESSSEESSDSEEET	328	353
Q53F19	Uncharacterized protein C17orf85	**C17orf85**	SRDASEDKSAEKRKKDKQEDSSDDDEAEEGEVEDENSSDVE	189	229
P29536	Leiomodin-1	**LMOD1**	SRDKDKKREEMKEVAKKEDDEK	201	222
O60706	ATP-binding cassette sub-family C member 9	**ABCC9**	SREAKAQMEDEDEEEEEEEDEDDN	943	966
Q15751	Probable E3 ubiquitin-protein ligase HERC1	**HERC1**	SREENEMREEKESKEEEKGK	2263	2282
Q6ZMG9	Ceramide synthase 6	**CERS6**	SRGKVSKDDRSDIESSSDEEDSEP	331	354
Q5T1B0	Axonemal dynein light chain domain-containing protein 1	**AXDND1**	SRKESKEEKENQDEREVKEEEE	970	991
Q8N5F7	NF-kappa-B-activating protein	**NKAP**	SRKESSDSSSKESQEEFLEN	257	276
Q9NP64	Nucleolar protein of 40 kDa	**ZCCHC17**	SRKRKKEKKKKKHRDRKSSDSDSSDSESDTGKR	183	215
Q9NYF8	Bcl-2-associated transcription factor 1	**BCLAF1**	SRSSSSSASPSSPSSREEKESKKEREEEFK	749	778
Q05519	Serine/arginine-rich splicing factor 11	**SRSF11**	SRSTSKTRDKKKEDKEKKRSK	304	324
O15014	Zinc finger protein 609	**ZNF609**	SRSVAGSKKEKENSSSKSKKE	150	170
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	SSASDKEEEQESTEEEEEAEEEEEEEVPRS	1062	1091
Q96JM2	Zinc finger protein 462	**ZNF462**	SSDDEDKEEEMNSKAEDREL	2379	2398
Q96SI9	Spermatid perinuclear RNA-binding protein	**STRBP**	SSDEKSDNESKNETVSSNSS	465	484
Q08945	FACT complex subunit SSRP1	**SSRP1**	SSDESSSGENKSKKKRRRSEDSEEEELAS	667	695

Q9C0A6	SET domain-containing protein 5	**SETD5**	SSDHEEVDNPEEKPEEEKEE	476	495
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	SSDKDKSEKSILEKCQKDKEFDKC	1032	1055
Q15468	SCL-interrupting locus protein	**STIL**	SSDNSEDEEEPPDNADSKSE	1131	1150
O43818	U3 small nucleolar RNA-interacting protein 2	**RRP9**	SSDSESESLAPRKPEEEEEEE	50	70
Q9NZW4	Dentin sialophosphoprotein	**DSPP**	SSDSSDSTSDSNDESDSQSKSGNG	1255	1278
Q2VWA4	SKI family transcriptional corepressor 2	**SKOR2**	SSEDEDDEEEEQEVDVEGHK	731	750
P52655	Transcription initiation factor IIA subunit 1	**GTF2A1**	SSEEDEDEEEDYDDDEEEDKEKDGAEDGQVEEEPLNSEDDVSDEEG	280	325
Q13523	Serine/threonine-protein kinase PRP4 homolog	**PRPF4B**	SSEEDKDKKHKHKHKHKKHK	60	79
P35269	General transcription factor IIF subunit 1	**GTF2F1**	SSEESEEEKPPEEDKEEEEEKKAPTP	307	332
Q9BTC0	Death-inducer obliterator 1	**DIDO1**	SSEKVKGGDDHDDTSDSDSD	138	157
Q99250	Sodium channel protein type 2 subunit alpha	**SCN2A**	SSESDMEESKEKLNATSSSE	1123	1142
Q9NY46	Sodium channel protein type 3 subunit alpha	**SCN3A**	SSESELEESKEKLNATSSSE	1121	1140
P16383	GC-rich sequence DNA-binding factor 2	**GCFC2**	SSESKDDQGLSSDSSSSLGEKE	106	127
Q68DE3	Basic helix-loop-helix domain-containing protein KIAA2018	**KIAA2018**	SSESLIPESVSKSKSAEKSS	871	890
P98082	Disabled homolog 2	**DAB2**	SSFDSLKSPDQKKENSSSSS	309	328
Q9UBL0	cAMP-regulated phosphoprotein 21	**ARPP21**	SSFSSLQEEDKSRKDDSEREKEKDKNKDKTSEKPKI	95	130
P17948	Vascular endothelial growth factor receptor 1	**FLT1**	SSGFQEDKSLSDVEEEEDSD	968	987
Q8NI27	THO complex subunit 2	**THOC2**	SSGGKKESRHDKEKIEKKEK	1554	1573
Q8NI27	THO complex subunit 2	**THOC2**	SSGKEKGSDSFKSEKMDKISS	1535	1555
Q92541	RNA polymerase-associated protein RTF1 homolog	**RTF1**	SSGSSDKDSSAESSAPEEGE	135	154
Q5H9R7	Serine/threonine-protein phosphatase 6 regulatory subunit 3	**PPP6R3**	SSGSTDSEESTDSEEEDGAK	639	658
Q5UIP0	Telomere-associated protein RIF1	**RIF1**	SSIENSESDSSEAKEEGSRKK	1547	1567
O14976	Cyclin-G-associated kinase	**GAK**	SSKESESALMEDRDESEVSDE	811	831
Q9NQB0	Transcription factor 7-like 2	**TCF7L2**	SSKHQDSKKEEEKKKPHIKK	333	352
Q9BYW2	Histone-lysine N-methyltransferase SETD2	**SETD2**	SSKKEDSHIGKDEEIPDSSK	278	297
Q7L4I2	Arginine/serine-rich coiled-coil protein 2	**RSRC2**	SSKKHKSEEHNDKEHSSDKG	79	98
Q96JM3	Chromosome alignment-maintaining phosphoprotein 1	**CHAMP1**	SSKKLKKDNQESSDAELSSSE	615	635
Q56NI9	N-acetyltransferase ESCO2	**ESCO2**	SSKNKEKLIKDSSDDRVSSKE	278	298
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	SSKSVEVKEEDSRIKDQEDK	2836	2855
Q9NWH9	SAFB-like transcription modulator	**SLTM**	SSKTQASVKKEEKRSSEKSEKKESKDTKKIEGKDEKND	492	529
Q01954	Zinc finger protein basonuclin-1	**BNC1**	SSMKSESSSHSSWDSDGVSEE	862	882
Q9UQ84	Exonuclease 1	**EXO1**	SSNASKLSQCSSKDSDSEESD	686	706
Q96T23	Remodeling and spacing factor 1	**RSF1**	SSNDESEGSGSEKSSAASEEEEEKESEEAI	857	886
Q9UPT8	Zinc finger CCCH domain-containing protein 4	**ZC3H4**	SSSDDFSDFSDDSDFSPSEKG	131	151
Q58F21	Bromodomain testis-specific protein	**BRDT**	SSSDLSSSDSSDSESEMFPKF	653	673
Q9Y5J3	Hairy/enhancer-of-split related with YRPW motif protein 1	**HEY1**	SSSDSELDETIEVEKESADE	9	28
Q8WUF8	Protein FAM172A	**FAM172A**	SSSDSSDEPAEKRERKDKVSKE	215	236
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	SSSEESKESSRARESSSRSG	1546	1565
P46100	Transcriptional regulator ATRX	**ATRX**	SSSEKSKSSGSSRSKRKPSI	66	85
O60583	Cyclin-T2	**CCNT2**	SSSERHSSSDEGSGKSKHSS	530	549
Q6ZT98	Tubulin polyglutamylase TTLL7	**TTLL7**	SSSESDENEKEEYQNKKREK	558	577
O14647	Chromodomain-helicase-DNA-binding protein 2	**CHD2**	SSSESSESQSESESESAGSKSQPV	55	78
Q9NRH2	SNF-related serine/threonine-protein kinase	**SNRK**	SSSETSDDDSESRRRLDKDS	539	558
Q5VUA4	Zinc finger protein 318	**ZNF318**	SSSFGKFSWKKPEKEEEKSS	1270	1289
Q6AI08	HEAT repeat-containing protein 6	**HEATR6**	SSSFSSSSWKRVSSSESDFSD	383	403
Q9NP61	ADP-ribosylation factor GTPase-activating protein 3	**ARFGAP3**	SSSFSSWDDSSDSYWKKETSKD	365	386
P56915	Homeobox protein goosecoid	**GSC**	SSSKASPEKREEEGKSDLDSDS	236	257
Q70Z53	Protein FRA10AC1	**FRA10AC1**	SSSKKSEDSLLRNSDEEESASESEL	265	289
Q9BQI6	Ankyrin repeat domain-containing protein 32	**ANKRD32**	SSSLKKLKKKSEGELSCSKE	770	789
P30414	NK-tumor recognition protein	**NKTR**	SSSLSSHHSSKRDWSKSDKD	487	506

A6NMT0	Homeobox protein DBX1	**DBX1**	SSSPGKPSDFSDSEEEEEGEE	316	336
Q9Y6V0	Protein piccolo	**PCLO**	SSSSDEYKQEDSQGSGEEED	1466	1485
Q03164	Histone-lysine N-methyltransferase 2A	**KMT2A**	SSSSEMKQSSASDLVSKSSS	2283	2302
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	SSSSGSSTTSPSSSASDKEEE	1050	1070
Q8NI27	THO complex subunit 2	**THOC2**	SSSSIGSASKSDESSTEETDKS	1209	1230
P51826	AF4/FMR2 family member 3	**AFF3**	SSSSSDSDLESEQEEYPLSK	673	692
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	SSSSSKKSHDRERAKKEKAEKKEKGEDYKE	1353	1382
Q14C86	GTPase-activating protein and VPS9 domain-containing protein 1	**GAPVD1**	SSSSSPSKDSSRGETEERKDSDDEKSDRN	946	974
C9JSJ3	Basic helix-loop-helix and HMG box domain-containing protein 1	**BHMG1**	SSSSSSEDSDSEPLWKQRED	460	479
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	SSSSSSSSEESDSDEEEAQKRESTTH	2186	2211
O15417	Trinucleotide repeat-containing gene 18 protein	**TNRC18**	SSSSSSSSGSETEGEEEGDKN	2570	2590
O15265	Ataxin-7	**ATXN7**	SSSSSSSSSSSSHSMESFRKN	717	737
Q6ZN04	RNA-binding protein MEX3B	**MEX3B**	SSSSSSSSSSSSSSGLRRKGS	494	515
Q9UQ35	Serine/arginine repetitive matrix protein 2	**SRRM2**	SSSSSSSSSSSSSSPSPAKPG	2634	2655
Q66PJ3	ADP-ribosylation factor-like protein 6-interacting protein 4	**ARL6IP4**	SSSSSSSSSSSSSSSSDGRKKRGKYKDKRRKKKKKRKKLKKKGKEKAEA	267	315
P42568	Protein AF-9	**MLLT3**	SSSSSSSSSSSSSSTSFSKPHKLMKEHKEKPSKDSRE	176	213
Q12791	Calcium-activated potassium channel subunit alpha-1	**KCNMA1**	SSSSSSSSSSSSSSVHEPKMDAL	46	69
Q8TBK6	Zinc finger CCHC domain-containing protein 10	**ZCCHC10**	SSSSSSSTSTDSSSDDEPPKKKKKK	168	192
O00567	Nucleolar protein 56	**NOP56**	SSSTPEECEEMSEKPKKKKK	465	484
Q9UHB7	AF4/FMR2 family member 4	**AFF4**	SSSTSKQKKTEGKTSSSSKE	858	877
Q8N302	Angiogenic factor with G patch and FHA domains 1	**AGGF1**	SSTKQSKDKKLKKKRKDPDSS	264	284
Q8IYX8	Centrosomal protein CEP57L1	**CEP57L1**	SSVSNLKHSKEKKKSSKKTK	232	251
Q96QE3	ATPase family AAA domain-containing protein 5	**ATAD5**	SSVSSSSNAEESKTGDEESK	1593	1612
Q9HCK8	Chromodomain-helicase-DNA-binding protein 8	**CHD8**	STDESEDEKEEKLTDQSRSK	2094	2113
Q96T23	Remodeling and spacing factor 1	**RSF1**	STDEYSEADEEEEEEEGKPS	1277	1296
Q8IWB6	Inactive serine/threonine-protein kinase TEX14	**TEX14**	STEDEQEETSKESPKELKEKD	1153	1173
Q8WYB5	Histone acetyltransferase KAT6B	**KAT6B**	STEKEDSARLDDHEEEEEEDEE	1396	1417
Q6UB99	Ankyrin repeat domain-containing protein 11	**ANKRD11**	STEKKDKNDSEREPSKKIEKE	1424	1444
Q7Z5K2	Wings apart-like protein homolog	**WAPAL**	STEKTDGTEEKHKKEEEDEE	1083	1102
Q9NR64	Kelch-like protein 1	**KLHL1**	STFWKKPSSSSSSSSPSSSSSF	66	89
Q9P0P8	Uncharacterized protein C6orf203	**C6orf203**	STKSTKKSLQKVDEEDSDEES	90	110
Q96N64	PWWP domain-containing protein 2A	**PWWP2A**	STSSTSSSSKEEKKLSNSLK	615	634
A1Z1Q3	O-acetyl-ADP-ribose deacetylase MACROD2	**MACROD2**	SVDDNNEEEEDVEMKEDSDE	241	260
Q9H814	Phosphorylated adapter RNA export protein	**PHAX**	SVDSSEESFSDSDDDSCLWK	62	81
Q8IWE2	Protein NOXP20	**FAM114A1**	SVDVAKVSEEETKKEEKEEKS	404	424
Q86VM9	Zinc finger CCCH domain-containing protein 18	**ZC3H18**	SVGEKESLEAAKEKKKEDDD	173	192
O60870	DNA/RNA-binding protein KIN17	**KIN**	SVKRKESSQSSTQSKEKKKKKS	238	259
P51532	Transcription activator BRG1	**SMARCA4**	SVRQKIEKEDDSEGEESEEEEE	1559	1580
Q8WVS4	WD repeat-containing protein 60	**WDR60**	SVSKVRSEEKDEDSERGDED	163	182
Q9Y6V0	Protein piccolo	**PCLO**	SVSSLDEDSDSSPSHKKGESK	1664	1684
Q15648	Mediator of RNA polymerase II transcription subunit 1	**MED1**	SVSSSGSKSHHSHSSSSSSS	1088	1107
P11229	Muscarinic acetylcholine receptor M1	**CHRM1**	SWKEEEEEDEGSMESLTSSE	273	292
Q9Y6M7	Sodium bicarbonate cotransporter 3	**SLC4A7**	SWLDDLMPESKKKKEDDKKKKEKEEAER	1132	1159
Q99250	Sodium channel protein type 2 subunit alpha	**SCN2A**	SYDSVTKPEKEKFEKDKSEKEDKGKD	1974	1999
O75400	Pre-mRNA-processing factor 40 homolog A	**PRPF40A**	SYKKSKKHKKKSKKRRHKSDS	865	885
Q02880	DNA topoisomerase 2-beta	**TOP2B**	SYSQKSEDDSAKFDSNEEDS	1452	1471
P31268	Homeobox protein Hox-A7	**HOXA7**	TAAADKADEEDDDEEEEDEEE	210	230
O60841	Eukaryotic translation initiation factor 5B	**EIF5B**	TAAEDDNEGDKKKKDKKKKK	303	322

Q86X27	Ras-specific guanine nucleotide-releasing factor RalGPS2	**RALGPS2**	TASEKSSSSESLSDKGSELKKS	16	37
Q96B23	Uncharacterized protein C18orf25	**C18orf25**	TCDSSTSSSDDDEEVSGSSK	212	231
Q8N129	Protein canopy homolog 4	**CNPY4**	TDGEEKTEGEEEQEEEEEEEEEG	207	230
P46100	Transcriptional regulator ATRX	**ATRX**	TDGEKKSKKIRDKTSKKKDE	1040	1059
Q96T58	Msx2-interacting protein	**SPEN**	TDSEGKMDDKKEDHKEEEQE	1502	1521
P24534	Elongation factor 1-beta	**EEF1B2**	TDSKDDDDIDLFGSDDEEESEE	93	114
Q5TCS8	Adenylate kinase 9	**AK9**	TEAEVDEELEEEEEEEGEDK	902	921
O15355	Protein phosphatase 1G	**PPM1G**	TEDEDEKEKVADEDDVDNEE	122	141
P35498	Sodium channel protein type 1 subunit alpha	**SCN1A**	TEDFSSESDLEESKEKLNESSSSSEGST	1129	1156
Q8N5Y2	Male-specific lethal 3 homolog	**MSL3**	TEEKDENDENSLSSSSDCSE	118	137
P12036	Neurofilament heavy polypeptide	**NEFH**	TEEVTEEEEKEAKEEEGKEEE	467	487
Q9NRE2	Teashirt homolog 2	**TSHZ2**	TELKKESKKERPEETSKDEK	458	477
Q5BKY9	Protein FAM133B	**FAM133B**	TESDSKDSLKKKKKSKDGTEKEKDIKGLSKKRK	150	182
Q14093	Cylicin-2	**CYLC2**	TESEDEKGGAKKDNKKDKKDS	176	196
Q5T7N2	LINE-1 type transposase domain-containing protein 1	**L1TD1**	TFIDSVEDSESEEEEEGKSSE	468	488
P35663	Cylicin-1	**CYLC1**	TFSTDSESELESKESQKDEKKDKKDSKTDNKKSVKNDEESTD	404	445
Q12904	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1	**AIMP1**	TGDEKKAKEKIEKKGEKKEKK	117	137
Q76L83	Putative Polycomb group protein ASXL2	**ASXL2**	TGESSSSKEDDTDEESTGDE	1170	1189
P23470	Receptor-type tyrosine-protein phosphatase gamma	**PTPRG**	TKDGEGTEEGEKDEKSESED	571	590
Q14191	Werner syndrome ATP-dependent helicase	**WRN**	TKEEEEDDENEANEGEEDDDKD	505	526
Q05086	Ubiquitin-protein ligase E3A	**UBE3A**	TKEELKSLQAKDEDKDEDEKEK	174	195
Q96KN7	X-linked retinitis pigmentosa GTPase regulator-interacting protein 1	**RPGRIP1**	TKGKDTKDSSKISSEEEKAS	942	961
Q06481	Amyloid-like protein 2	**APLP2**	TKIIGSVSKEEEEEDEEEEEEEDEEEDYDV	206	235
P35663	Cylicin-1	**CYLC1**	TKKDTKKNAKKSSDAESEDSKD	284	305
Q7KZ85	Transcription elongation factor SPT6	**SUPT6H**	TKKFVEEEDDDEEEEEENLDD	27	47
Q9H270	Vacuolar protein sorting-associated protein 11 homolog	**VPS11**	TKLKDSSKLEEFIKKKSESE	462	481
Q9BQ67	Glutamate-rich WD repeat-containing protein 1	**GRWD1**	TKPPPSEGSDEEEEEEDEEDEEERK	114	138
Q14919	Dr1-associated corepressor	**DRAP1**	TKSKDKKLSGTDSEQEDESED	126	146
O95202	LETM1 and EF-hand domain-containing protein 1, mitochondrial	**LETM1**	TLEKEEKVEEKEKAKEKAEKE	713	733
P23921	Ribonucleoside-diphosphate reductase large subunit	**RRM1**	TLNKEKLKDKEKVSKEEEEKE	751	771
Q9NYB0	Telomeric repeat-binding factor 2-interacting protein 1	**TERF2IP**	TPEEDSETQPDEEEEEEEEK	286	305
Q4G1C9	GLIPR1-like protein 2	**GLIPR1L2**	TPEESEAGNEEEEKEEEKKEKEEME	290	314
P29374	AT-rich interactive domain-containing protein 4A	**ARID4A**	TPKQKEKKIKKQEDSDKDSDEEEEKSQEREETESKCDSEGEEDEEDMEP	524	572
P35269	General transcription factor IIF subunit 1	**GTF2F1**	TPQEKKRRKDSSEESDSSEESD	331	352
Q6WKZ4	Rab11 family-interacting protein 1	**RAB11FIP1**	TPSVDSDDESVVKDKKKKSK	197	216
Q13428	Treacle protein	**TCOF1**	TQRPAEDSSSSEESDSEEEK	686	705
Q13428	Treacle protein	**TCOF1**	TQVKAEKQEDSESSEEESDSEE	752	773
O14974	Protein phosphatase 1 regulatory subunit 12A	**PPP1R12A**	TREQENEEKEKEEKEKQDKEK	719	739
Q96LT9	RNA-binding protein 40	**RNPC3**	TSGSEKKKRSDDPVEDDKEKKE	112	133
Q12830	Nucleosome-remodeling factor subunit BPTF	**BPTF**	TSIEREEKEKVKKKEKKQEEEE	979	1000
Q9UPS6	Histone-lysine N-methyltransferase SETD1B	**SETD1B**	TSKAEATSSSESSESSEFESSSESSPSSSEDEEEVVA	1147	1183
Q9Y3E1	Hepatoma-derived growth factor-related protein 3	**HDGFRP3**	TSKKSSKQSRKSPGDEDDKD	151	170
Q14151	Scaffold attachment factor B2	**SAFB2**	TSRSKERSSKSQDRKSESKEK	588	608
Q5T5P2	Sickle tail protein homolog	**KIAA1217**	TSSSKDEEEEEEEGDKIMAE	1117	1136

P51826	AF4/FMR2 family member 3	**AFF3**	TSVPSSKGSSSSSSSGSSSSSSDSESSSGSDSETESS SSESEGSKPPHFSSPE	405	457
Q86UP3	Zinc finger homeobox protein 4	**ZFHX4**	TSVSLSHSSSESSKMSESKD	421	440
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	TTPTRKSSSSEDSSSDEEEE	258	277
O95196	Chondroitin sulfate proteoglycan 5	**CSPG5**	TTSFYDDLDEEEEEEEDDKD	268	287
O15355	Protein phosphatase 1G	**PPM1G**	VAKSKFFEDSEDESDEAEEEEEDSEECSEEEDGYSSEE AENEEDEDDTEEAEEDDEEEEEEMMVP	251	315
Q8TDZ2	Protein-methionine sulfoxide oxidase MICAL1	**MICAL1**	VAMEKEEKESPFSSEEEEED	863	882
Q53HC9	Protein TSSC1	**TSSC1**	VDDDDISDQEDHRSEEKSKE	314	333
O14513	Nck-associated protein 5	**NCKAP5**	VDEDSRSESSSTDEGKEKTK	164	183
Q92576	PHD finger protein 3	**PHF3**	VDKKPDKPKSEDYEKDKEREKSKH	1992	2015
Q9Y2W2	WW domain-binding protein 11	**WBP11**	VEEFSEDDDEDDSDDSEAEKQSQKQ	349	373
P14625	Endoplasmin	**HSP90B1**	VEEPMEEEEAAKEEKEESDDE	289	309
Q9NWH9	SAFB-like transcription modulator	**SLTM**	VEKEARESSKKAESGDKEKD	317	336
P12036	Neurofilament heavy polypeptide	**NEFH**	VEKPKESKVEAKKEEAEDKKK	888	908
Q14978	Nucleolar and coiled-body phosphoprotein 1	**NOLC1**	VEKQQPVESSEDSSDESDSSSEEEKKPP	313	340
Q5XX13	F-box/WD repeat-containing protein 10	**FBXW10**	VEKTKQKKNKEKEEEKEENS	692	711
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	**HNRNPC**	VEMKNDKSEEEQSSSSVKKDE	226	246
Q13428	Treacle protein	**TCOF1**	VGKQEEDSRSSSEESDSDRE	468	487
Q5VWK0	Neuroblastoma breakpoint family member 6	**NBPF6**	VHKLSPENDEDEDEDEDDKDEE	159	180
Q86XG9	Putative neuroblastoma breakpoint family member 5	**NBPF5P**	VHKLSPENDEDEDEDEDDKDEE	159	180
Q96M43	Neuroblastoma breakpoint family member 4	**NBPF4**	VHKLSPENDEDEDEDEDDKDEE	159	180
Q3L8U1	Chromodomain-helicase-DNA-binding protein 9	**CHD9**	VHKRGSESSSDSDSDSERSS	2123	2142
Q9UGU5	HMG domain-containing protein 4	**HMGXB4**	VIDDSYREIKKKKKSKKSKKKKDKEKHKEKRHSKSKRS	308	345
Q86XH1	IQ and AAA domain-containing protein 1	**IQCA1**	VKAGKKKDKKGKKGKKKEKK	463	482
P19338	Nucleolin	**NCL**	VKAKNVAEDEDEEEDDEDEDDDDDEDDEDDD DEDDEEEEEEEEEPVKEAPGKRKKEMAKQKAA	227	290
Q6KC79	Nipped-B-like protein	**NIPBL**	VKDKRKERKSSPSKENESSDSEEEVS	2473	2498
P23327	Sarcoplasmic reticulum histidine-rich calcium-binding protein	**HRC**	VKDRSHLRKDDSEEEKEKEED	471	491
Q9Y6X9	MORC family CW-type zinc finger protein 2	**MORC2**	VKEEKKDSNELSDSAGEEDS	766	785
Q9P2K5	Myelin expression factor 2	**MYEF2**	VKMENDESAKEEKSDLKEKS	52	71
P51826	AF4/FMR2 family member 3	**AFF3**	VPKSKEFIETESSSSSSSDSD	659	680
Q8NDV7	Trinucleotide repeat-containing gene 6A protein	**TNRC6A**	VQEEEQLMEEKKKKKDDKKKKE	21	42
Q9P287	BRCA2 and CDKN1A-interacting protein	**BCCIP**	VQRDEEEEKEVENEDEDDDDSDKEKDEEDEVIDEEVN	22	58
Q5VWG9	Transcription initiation factor TFIID subunit 3	**TAF3**	VRKEKEKHKDKKKDREKGKKDKDKREKEKVKDKG	609	642
Q15311	RalA-binding protein 1	**RALBP1**	VSDDEKDHGKKKGKFKKKEK	61	80
P26374	Rab proteins geranylgeranyltransferase component A 2	**CHML**	VSDKDGDKDESKSTVEDKADE	191	211
Q9BXP5	Serrate RNA effector molecule homolog	**SRRT**	VSESESESESGQAEEEKEEAEE	367	388
Q9H3P7	Golgi resident protein GCP60	**ACBD3**	VSESSDDDEEEEENIGCEEK	447	466
Q15059	Bromodomain-containing protein 3	**BRD3**	VSKGAESSRSSEESSSDSGSSDSEEERA	433	460
Q9ULD9	Zinc finger protein 608	**ZNF608**	VSLKDKKKKEKRKLKDKEGKE	806	826
Q5W0B1	RING finger protein 219	**RNF219**	VSNKDSSEDDISRSENEKKSE	437	457
Q86UV5	Ubiquitin carboxyl-terminal hydrolase 48	**USP48**	VSSSETEEDKEEAKPDGEKD	885	904
P51784	Ubiquitin carboxyl-terminal hydrolase 11	**USP11**	VTKPNSDDEDDGDEKEDDEEDKDDV	643	667
O43491	Band 4.1-like protein 2	**EPB41L2**	VTQKDSTSLSSESSSSSSESEEEDV	729	753
Q9UN42	Protein ATP1B4	**ATP1B4**	VTVVPKSEEEEEEEEKEEEEEEEKEEEEGQ	45	74
Q8NI27	THO complex subunit 2	**THOC2**	VVLSSEKMDEREKEKEKEEEK	316	336
Q8WZ42	Titin	**TTN**	VVSDEKQDESLKPSEEKEESSSESG	4243	4267
O14896	Interferon regulatory factor 6	**IRF6**	WDEKDNDVDEEDEEDELDQS	134	153
Q12873	Chromodomain-helicase-DNA-binding protein 3	**CHD3**	WEAKEEEEEYEEEGEEEGEKEEEDDH	430	455
Q92563	Testican-2	**SPOCK2**	WEDEEEKETEEAGEEAEEEE	392	411
Q9H792	Pseudopodium-enriched atypical kinase 1	**PEAK1**	WEDPDDPEKDEDDMEETEED	1405	1424

O75822	Eukaryotic translation initiation factor 3 subunit J	**EIF3J**	WEGEDEDEDVKDNWDDDDDEKKEEAE	39	64
Q8TDI0	Chromodomain-helicase-DNA-binding protein 5	**CHD5**	WEPKDDDDEEEEGGCEEEEDD	394	414
Q9P289	Serine/threonine-protein kinase 26	**STK26**	WKAEGHSDDESDSEGSDSES	294	313
Q9NSI6	Bromodomain and WD repeat-containing protein 1	**BRWD1**	WRSDRKSESYSESSSDSSSR	845	864
Q8IZL8	Proline-, glutamic acid- and leucine-rich protein 1	**PELP1**	YDKEEASDVEISLESDSDDS	768	787
O15117	FYN-binding protein	**FYB**	YEDIEASKEREKKREKEEKK	462	481
P78415	Iroquois-class homeodomain protein IRX-3	**IRX3**	YGSEREEEDEEEDEEDGKRE	206	225
Q8IX12	Cell division cycle and apoptosis regulator protein 1	**CCAR1**	YKSLLSLPEKEDKKEKDKKSKKDERKDKKEERDDETDEPKPK	787	828
Q58FF6	Putative heat shock protein HSP 90-beta 4	**HSP90AB4P**	YLEKEPEKEISDDEEEKGEKEEEDKDDKEKPKT EDVGSDEEDDTDKNNKKKTKK	191	244
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	YLFETDSEEEEEEELKKEDEE	1552	1572
O75264	Small integral membrane protein 24	**SMIM24**	YQDQSEDKREKKEAKEKEEK	76	95
P98175	RNA-binding protein 10	**RBM10**	YRTEQGEEEEEEEDEEEEK	107	126
Q9H5I5	Piezo-type mechanosensitive ion channel component 2	**PIEZO2**	YRWEPSDESSEKREEEEEEKEE	450	471
Q9ULU4	Protein kinase C-binding protein 1	**ZMYND8**	YTAVEHSDSEDSEKSDSSDSE	581	601
Q8TAA3	Proteasome subunit alpha type-7-like	**PSMA8**	YVTEIEKEKEEAEKKKSKKS	236	255

7.2: Appendix B

Table 3: Antibodies used in this work (Bentley-DeSousa et al., 2018)

Antibody Target	**Dilution**	**Source**	**Identifier**
GFP (JL-8) Monoclonal	1:2,000	Living Colors (Clontech)	632381
Fpr3 C-terminus	1:2,000	Dr. Jeremy Thorner	PMID: 7525596
Rts1	1:5,000	Dr. Adam Rudner	PMID: 26715668
Anti-Sui2 (eIF2)	1:40,000	Described here	N/A
Anti-phospho-eIF2alpha	1:2000	Cell Signaling	9721S
12CA5	1:1000	Covance	Ascites
Cdc55	1:5,000	Dr. Adam Rudner	PMID: 26715668
Nsr1	1:10,000	AbCam	Ab4642
Rad53	1:2,000	Santa Cruz	Sc-6749
Nucleolin	1:5,000	AbCam	Ab22758
hNOP56	1:1,000	Bethyl	A302-720A-T
MesD	1:1,000	Cell Signaling	2763
DEK	1:1,000	Santa Cruz	136222
eIF5B	1:200	Santa Cruz	393564
UPF3B	1:500	Santa Cruz	48800
Hip	1:1,000	Santa Cruz	136175
HSP90alpha	1:1,000	Santa Cruz	515081
Streptavidin (HRP)	1:5,000	Abcam	Ab7403
Beta-Actin HRP-conjugate	1:40,000	Santa Cruz	Sc-47778
GAPDH HRP-conjugate	1:40,000	Abcam	Ab85760
Goat anti-Mouse HRP Conjugate secondary	1:10,000	Biorad	172-1011
Goat anti-Rabbit HRP Conjugate secondary	1:10,000	Biorad	170-6515
Donkey anti-Goat HRP Conjugate secondary	1:10,000	Santa Cruz	sc-2020

Table 4: Synthetic DNA constructs (ATUM) (Bentley-DeSousa et al., 2018)

Plasmid	**Description**
GAL-3HA-RPA34	Wild-type Rpa34
GAL-3HA-RPA34allK-R	Rpa34 with PASK lysine (K) to arginine (R) mutations from amino acid 208 to 233
GAL-3HA-RPA34ΔPASKNOP56	Rpa34 PASK deletion (from amino acid 208 to 233); Wild-type Nop56 PASK replacement (from amino acid 442 to 504)
GAL-3HA-CHZ1	Wild-type Chz1
GAL-3HA-CHZ1ΔC	Chz1 with N-terminal PASK deletion (from amino acid 1 to 131)
GAL-3HA-CHZ1ΔN	Chz1 with C-terminal PASK deletion (Methionine + amino acid 26 to 153)
GAL-3HA-CHZ1Δ(N+C)	Chz1 with N-terminal and C-terminal PASK deletions (Methionine + amino acid 26 to 131)

Table 5: Strains used in this study (Bentley-DeSousa et al., 2018)

Strain number	Strain Genotype	Source
YMD963	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+	Boone Lab
YMD964	MATalpha leu2d0 ura3d0 his3-1 MET15 lys2d0	Boone Lab
YMD1031	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc4::KANMX	This study
YMD1178	MATalpha can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX	This study
YMD1179	MATalpha can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study
YMD1182	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RTS1GFP(HIS) ppn1::NATMX	This study
YMD1183	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RTS1GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1184	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ EAF7GFP(HIS) ppn1::NATMX	This study
YMD1185	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ EAF7GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1188	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ FPR3GFP(HIS) ppn1::NATMX	This study
YMD1189	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ FPR3GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1190	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ FPR4GFP(HIS) ppn1::NATMX	This study
YMD1191	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ FPR4GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1192	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NOP56GFP(HIS) ppn1::NATMX	This study
YMD1193	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NOP56GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1194	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NOP58GFP(HIS) ppn1::NATMX	This study
YMD1195	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NOP58GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1196	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RPA34GFP(HIS) ppn1::NATMX	This study
YMD1197	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RPA34GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1198	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ BOI1GFP(HIS) ppn1::NATMX	This study
YMD1199	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ BOI1GFP(HIS) ppn1::NATMX vtc4::KANMX	This study

YMD1200	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ JJJ1GFP(HIS) ppn1::NATMX	This study
YMD1201	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ JJJ1GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1202	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ UTP14GFP(HIS) ppn1::NATMX	This study
YMD1203	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ UTP14GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1204	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RRP15GFP(HIS) ppn1::NATMX	This study
YMD1205	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ RRP15GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1206	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NST1GFP(HIS) ppn1::NATMX	This study
YMD1207	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ NST1GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1208	MATa can1^::STE2pr-LEU2 lyp1^ ura 3^0 LYS2+ CHZ1GFP(HIS) ppn1::NATMX	This study
YMD1209	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ CHZ1GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1210	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ TMA23GFP(HIS) ppn1::NATMX	This study
YMD1211	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ TMA23GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1212	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ HPC2GFP(HIS) ppn1::NATMX	This study
YMD1213	MATa can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ HPC2GFP(HIS) ppn1::NATMX vtc4::KANMX	This study
YMD1270	MATalpha NOP56Δ442-504-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX	This study
YMD1272	MATalpha NOP56Δ442-504-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study
YMD1274	MATalpha RPA34Δ213-233-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX	This study
YMD1275	MATalpha RPA34Δ213-233-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study
YMD1315	MATalpha NOP58Δ429-511-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX	This study
YMD1317	MATalpha NOP58Δ429-511-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study
YMD1416	MATalpha TMA23Δ139-211-GFP(HIS) can1^::STE2prLEU2 lyp1^ ura3^0 LYS2+ ppn1::NATMX	This study
YMD1420	MATalpha TMA23Δ139-211-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura 3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study

YMD1424	MATalpha CHZ1Δ131-150-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura 3^0 LYS2+ ppn1::NATMX	This study

YMD1426	MATalpha CHZ1Δ131-150-GFP(HIS) can1^::STE2pr-LEU2 lyp1^ ura 3^0 LYS2+ ppn1::NATMX vtc4::KANMX	This study
YMD1574	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc4::NATMX	This study
YMD1576	MATalpha leu2d0 ura3d0 his3-1 MET15 lys2d0 vtc4::NATMX	This study
YMD1770	MATalpha leu2d0 ura3d0 his3-1 MET15+ lys2 vtc4::NATMX [pURA3]	This study
YMD1771	MATalpha leu2d0 ura3d0 his3-1 MET15+ lys2 vtc4::NATMX [pURA3]	This study
YMD1774	MATalpha leu2d0 ura3d0 his3-1 MET15+ lys2 vtc4::NATMX [pPPK1-URA3]	This study
YMD1775	MATalpha leu2d0 ura3d0 his3-1 MET15+ lys2 vtc4::NATMX [pPPK1-URA3]	This study
YMD1776	MATalpha leu2d0 ura3d0 his3-1 MET15+ lys2 vtc4::NATMX [pPPK1-URA3]	This study
YMD2000	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc4::NATMX [pRS316]	This study
YMD2001	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc4::NATMX [pRS316-VTC4]	This study
YMD2005	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc4::NATMX [pRS316-VTC4]	This study
YMD2013	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ pRS316	This study
N/A	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc1::KANMX	This study
N/A	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ vam3::KANMX	This study
N/A	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ vma22::KANMX	This study
N/A	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ yol019w::KANMX	This study
YMD2299	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34ΔPASK::NOP56]	This study
YMD2300	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34ΔPASK::NOP56]	This study
YMD2301	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34ΔPASK::NOP56] vtc4::NATMX	This study
YMD2302	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34ΔPASK::NOP56] vtc4::NATMX	This study
YMD2306	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1]	This study

YMD2307	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1]	This study
YMD2308	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔC]	This study
YMD2309	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔC]	This study

YMD2310	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔN]	This study
YMD2311	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔN]	This study
YMD2312	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1Δ(N+C)]	This study
YMD2313	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1Δ(N+C)]	This study
YMD2314	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1] vtc4::NATMX	This study
YMD2315	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1] vtc4::NATMX	This study
YMD2316	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+[pURA-GALHA-CHZ1ΔC] vtc4::NATMX	This study
YMD2317	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔC] vtc4::NATMX	This study
YMD2318	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔN] vtc4::NATMX	This study
YMD2319	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1ΔN] vtc4::NATMX	This study
YMD2320	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1Δ(N+C)] vtc4::NATMX	This study
YMD2321	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-CHZ1Δ(N+C)] vtc4::NATMX	This study
YMD2322	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34]	This study
YMD2323	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34]	This study
YMD2324	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34allK-R]	This study
YMD2325	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34allK-R]	This study
YMD2326	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34] vtc4::NATMX	This study
YMD2327	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34] vtc4::NATMX	This study

YMD2328	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34allK-R] vtc4::NATMX	This study
YMD2329	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ [pURA-GALHA-RPA34allK-R] vtc4::NATMX	This study
YMD2284	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppx1::KANMX	This study
YMD2285	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppx1::KANMX	This study
YMD2286	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ddp1::KANMX	This study
YMD2287	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ddp1::KANMX	This study
YMD2289	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn2::KANMX	This study
YMD2290	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn2::KANMX	This study
YMD2291	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn1::KANMX	This study
YMD2330	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn1::KANMX ppn2::NATMX	This study
YMD2331	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn1::KANMX ppn2::NATMX	This study
YMD2332	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn1::KANMX ppn2::NATMX	This study

Table 6: Candidates Screened (McCarthy et al., 2020)
*Note – this table has been amended for ease of inclusion in this thesis. ORF names denoted in red were not screened and information on why is included in (McCarthy et al., 2020)

ORF-GFP FUSIONS SCREENED IN THIS WORK	HITS	ORF-GFP FUSIONS SCREENED IN THIS WORK	HITS
YFR019W		YPL212C	
YHR023W		YDR195W	
YML059C		YBR138C	
YDR457W		YGL056C	
YLR430W		YMR215W	
YOR341W		YMR131C	
YBL034C		YKL081W	
YER164W		YPR169W	
YDR283C		YGR189C	
YML072C		YDL101C	
YAL024C		YJL013C	
YMR280C		YFR039C	
YGL197W		YML043C	
YPR189W		YOR229W	
YNR031C		YOR198C	
YMR190C		YOL041C	
YNL088W		YDL189W	
YLR278C		YDR473C	
YPL137C		YPR152C	
YOL138C		YNL206C	
YDR293C		YBR279W	
YOR227W		YDR055W	
YDR301W		YDL150W	
YLR084C		YPR148C	
YGR270W		YDR528W	
YJR092W		YBL091C	
YDR097C		YNL066W	
YGL140C		YNL175C	
YKR096W		YOR051C	
YGR233C		YGR142W	
YLR371W		YDL105W	
YDR507C		YDR309C	
YNL250W		YGR187C	
YMR012W		YOR057W	
YGL133W	YES	YKL114C	
YCR033W		YBR065C	
YML049C		YDR171W	
YOR188W		YHR052W	
YJR035W		YBL029W	
YML010W		YKR071C	
YDR325W		YDR173C	

YIL112W		YLR287C	
YBR245C		YOR261C	
YKL033W		YGR283C	
YBR059C		YML037C	
YJL098W		YOL007C	
YFL034W		YHR156C	
YHR082C		YPL157W	
YPR104C		YNL322C	
YCR057C		YOR006C	
YGR140W		YJR007W	
YER075C		YDL173W	
YDR028C		YDR054C	
YOR048C		YOR083W	
YMR124W		YGR239C	
YGL241W		YJL115W	
YIL075C		YDL051W	
YDR356W		YKR025W	
YDR495C		YCR016W	
YGL180W		YLR003C	
YLR451W		YOR281C	
YOR373W		YDR087C	
YBR237W		YLR132C	
YPL105C	YES	YGL237C	
YDR409W		YGR280C	
YER151C		YNL148C	
YDL040C		YGR075C	
YGL141W		YOR004W	
YGR071C		YER174C	
YOR233W		YDL139C	
YMR102C		YLR194C	
YLR095C		YDR174W	
YKR099W		YKL096W	
YDL122W		YMR233W	
YLR397C		YMR311C	
YNL258C		YLR221C	
YPL158C		YBL057C	
YPL074W		YER079W	
YAL043C		YOR078W	
YGR237C	YES	YBR122C	
YKR010C		YDR022C	
YHR149C		YFR001W	
YNL059C		YGR251W	
YHR197W		YPR067W	
YOR206W		YGL058W	
YAR014C		YLR421C	
YDR333C		YKL160W	
YDR458C	YES	YNL143C	

YLR002C		YCR043C	
YLR403W		YNL149C	
YDR452W		YMR123W	
YMR075W		YDR382W	
YNL176C		YDL081C	
YNL299W		YDR363W-A	
YDR398W		YGL137W	
YEL060C	YES	YCR017C	
YNR038W		YHR079C	
YNL020C		YBR086C	
YGL023C		YIR006C	
YDR419W		YDR080W	
YPL263C	YES	YPL195W	YES
YOL117W		YGR281W	
YBR060C		YPL270W	
YOR027W		YGR261C	
YGR013W		YBL011W	
YNL039W		YHL019C	
YNL061W	YES	YNL275W	
YKR078W		YMR307W	
YGR246C		YKL008C	
YER140W		YLL010C	
YGR278W		YBR127C	
YKL088W		YKL021C	
YGL244W		YHL003C	
YHR064C		YNR041C	
YBR030W		YHR195W	
YIL045W		YDR033W	
YFR019W		YNL173C	

Table 7: Yeast strains used in this work (McCarthy, L., *et al* (2020))

Strain Number	Background	Full Genotype	Source
YMD2256	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 ITC1-GFP-HIS3MX	GFP collection
YMD2257	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ ITC1-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2258	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 SYH1-GFP-HIS3MX	GFP collection
YMD2259	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ SYH1-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2260	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 YGR237C-GFP-HIS3MX	GFP collection
YMD2261	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ YGR237C-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2262	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 HEH2-GFP-HIS3MX	GFP collection
YMD2263	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ HEH-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2264	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 PRB1-GFP-HIS3MX	GFP collection
YMD2265	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ PRB1-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2266	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 KEL3-GFP-HIS3MX	GFP collection
YMD2267	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ KEL3-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2268	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 NOP2-GFP-HIS3MX	GFP collection
YMD2269	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ NOP2-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2272	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APL5-GFP-HIS3MX	GFP collection
YMD2273	S288C BY474	MATa can1::STE2pr-LEU2 lyp1 ura 30 LYS2+ APL5-GFP-HIS3MX ppn1::NATMX vtc4::KANMX	This study
YMD2380	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APL5-GFP-HIS3MX	This study
YMD2381	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APL6-GFP-HIS3MX	This study
YMD2382	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APS3-GFP-HIS3MX	This study
YMD2531	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APL5-GFP-HIS3MX vtc4::NATMX	This study
YMD2537	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APL6-GFP-HIS3MX vtc4::NATMX	This study

YMD2539	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d0 APS3-GFP-HIS3MX vtc4::NATMX	This study
YMD2548	S288C BY474	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APM3-GFP-HIS3MX	This study
YMD2551	S288C BY474	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APM3-GFP-HIS3MX vtc4::NATMX	This study
YMD2583	S288C BY474	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APL5PASKΔ896-932-GFP-HIS3MX	This study
YMD2644	S288C BY474	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ KANMX-GALpr-PRB1	This study
YMD2647	S288C BY474	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ KANMX-GALpr-PRB1 vtc4::NATMX	This study
YMD2663	S288C BY474	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APL5PASKΔ896-932-GFP-HIS3MX vtc4::NATMX	This study
YMD2330	S288C BY474	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ ppn1::KANMX ppn2::NATMX	Bentley-DeSousa et al., PMID: 29590613
YMD963	S288C BY474	MATa his3d1 leu2d0 met15d0 ura3d	Boone lab via Adam Rudner

Table 8: Antibodies used in this study (McCarthy et al., 2020)

Antibody Target	Host	Source	Dilution	Conditions	Notes
GFP	Mouse	Clontech JL-8 (#632381)	1/2000	5 % milk tbst	can re-use dilution
Prb1	Rabbit	Elizabeth Jones' Lab (PMID 2645294)	1/1000	5 % milk tbst	can re-use dilution
Rts1	Rabbit	Rudner Lab (uOttawa, PMID 26715668)	1/5000	5 % milk tbst	can re-use dilution
Cdc55	Rabbit	Rudner Lab (uOttawa, PMID 26715668)	1/5000	5 % milk tbst	can re-use dilution
Rabbit IgG	Goat	BioRad	1/10,000	5 % milk tbst	

Table 9: Yeast strains used in this study (Bentley-DeSousa, A., & Downey, M., *in review*)

STRAIN	GENOTYPE	SOURCE
YMD963	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+	Rudner Lab via Boone Lab
YMD2333	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apl5::KANMX	Collection Set
YMD2334	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apl6::KANMX	Collection Set
YMD2335	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ aps3::KANMX	Collection Set
YMD2336	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apm3::KANMX	Collection Set
YMD2380	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ APL5-GFP(HISMX)	Collection Set
YMD2531	MATa his3d1 leu2d0 met15d0 ura3d0 APL5-GFP-HIS3MX vtc4::NATMX	McCarthy et al
YMD2556	Matalpha leu2d0 ura3d0 h+B41is3-1 lys2d0 MET15+ natNT2-GPD1-GFP-Vtc3	Andreas Mayer
YMD2558	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-ADH1-GFP-Vtc4	Andreas Mayer
YMD2560	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5	Andreas Mayer
YMD2571	Matalpha leu2d0 ura3d0 his3-1 lys2d0 MET15+ natNT2-ADH-GFP-Vtc3 apl5::KANMX	This study
YMD2577	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-ADH-GFP-Vtc4 apl5::KANMX	This study
YMD2583	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APL5PASKΔ896-932-GFP-HIS3MX	McCarthy et al 2020
YMD2595	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX	This study
YMD2614	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 aps3::KANMX	This study
YMD2663	MATa leu2d0 ura3d0 his3-1 met15d0 LYS2+ APL5PASKΔ896-932-GFP-HIS3MX vtc4::NATMX	McCarthy et al 2020
YMD2739A	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+natNT2-GPD1-GFP-Vtc5 apl5PASKΔ-FLAG(HYGMX)	This study
YMD2746A	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apl5PASKK-R-GFP(HISMX)	This study
YMD2750A	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apl5::KANMX	This study
YMD2885	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 Apl5PASKK-R-FLAG(HYGMX)	This study
YMD2953	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 atg8::HYGMX	This study

YMD2961	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+natNT2-GPD1-GFP-Vtc5 apl5::KANMX atg8::HYGMX	This study
YMD2967	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 vps27::HYGMX	This study
YMD2976	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 vps36::HYGMX	This study
YMD2983	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX vps36::HYGMX	This study
YMD2992	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 vps20::HYGMX	This study
YMD2997	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 pep12::HYGMX	This study
YMD3002	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX vps23::HYGMX	This study
YMD3013	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+natNT2-GPD1-GFP-Vtc5 vps23::HYGMX	This study
YMD3018	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX vps27::HYGMX	This study
YMD3024	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX vps20::HYGMX	This study
YMD3204	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 pep4::HYGMX	This study
YMD3210	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 apl5::KANMX pep4::HYGMX	This study
YMD3371	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc5::KANMX	This study
YMD3382	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 doa4::ura3	This study
YMD3432	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ apl5::KANMX vtc5::HYGMX	This study
YMD3491	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ vtc5::HYGMX apl5::KANMX	This study
YMD3576	Mata leu2d0 ura3d0 his3-1 met15d0 LYS2+ natNT2-GPD1-GFP-Vtc5 doa4::ura3 apl5::KANMX	This study

Table 10: Antibodies used in this study (Bentley-DeSousa, A., & Downey, M., *in prep*)

ANTIBODY	**SOURCE**	**IDENTIFIER**
GFP (host: mouse, 1:2,000 dilution)	Clontech	632381
PHO8 (host: rabbit)	Abcam	b113688
RTS1 (host: rabbit, 1:5,000 dilution)	Dr. Adam Rudner	PMID: 26715668
Mouse IgG (1/10000 dilution)	Santa Cruz	sc-2025
Rabbit IgG (1/10000 dilution)	NEB	2729S

7.3: Appendix C

Bentley-DeSousa, A., & Downey, M. (2019). From underlying chemistry to therapeutic potential: open questions in the new field of lysine polyphosphorylation. *Curr Genet* 65, 57-64.

Abstract

Polyphosphorylation is a newly described non-enzymatic post-translational modification wherein long chains of inorganic phosphates are attached to lysine residues. The first targets of polyphosphorylation identified were S. cerevisiae proteins Nsr1 and Top1. Building on this theme, we recently exploited functional genomics tools in yeast to identify 15 new targets, including a conserved network of nucleolar proteins implicated in ribosome biogenesis. We also described the polyphosphorylation of six human proteins, suggesting that this unique post-translational modification could be conserved throughout eukaryotes. The study of polyphosphorylation seems poised to uncover novel modes of protein regulation in pathways spanning diverse biological processes. In this review, we establish a framework for future work by outlining critical questions related to the biochemistry of polyphosphorylation, its therapeutic potential, and everything in between.

Introduction

Polyphosphates (polyP) are long chains of inorganic phos- phates that are found in diverse cell types. These unique molecules, which can range in size from 3 to thousands of phosphate residues in length, have important roles in diverse processes including phosphate homeostasis, cell growth, infection control, and blood coagulation (Eskes et al., 2018; Kornberg, 1999; Morrissey et al., 2012). The budding yeast S. cerevisiae is an ideal system to study polyP biology as the regulators of its metabolism have been well characterized in this organism. Remarkably, polyP makes up more than 10% of the dry weight of yeast with estimated concen-

trations of over 200 mM in free phosphate units (Auesukaree et al., 2004; Kornberg, 1999). It is synthesized from ATP by the VTC complex consisting of a vacuolar-bound polyP synthetase, called Vtc4, along with Vtc1, Vtc5, and one of either Vtc2 or Vtc3 (Desfougeres et al., 2016; Gerasimaite et al., 2014; Hothorn et al., 2009). PolyP is degraded by exopolyphosphatase Ppx1, endopolyphosphatases Ppn2 and Ddp1, or dual exo/endopolyphosphatase Ppn1 (Andreeva et al., 2015; Gerasimaite and Mayer, 2017; Lichko et al., 2006; Lonetti et al., 2011). While these aspects of polyP metabolism are well defined, the molecular functions of polyP have remained more enigmatic.

Remarkably, in 2015 Azevedo et al. reported that polyP can be covalently attached to lysine residues as a post-translational modification (PTM). This PTM, termed 'polyphosphorylation', occurs non-enzymatically and confers a dramatic decrease in electrophoretic mobility when proteins are analyzed via Bis-Tris polyacrylamide gel electrophoresis (Azevedo et al., 2015). Azevedo et al. used these gels, sold commercially under the 'NuPAGE' brand name, to characterize polyphosphorylation of two yeast proteins: Nsr1 and Top1. These proteins are both modified on lysine residues in regions referred to as poly-acidic, serine and lysine rich (PASK) motifs (Azevedo et al., 2015). Evidence suggests that polyphosphorylation impacts the localization and physical interaction of Nsr1 and Top1, and the topoisomerase activity of Top1 (Azevedo et al., 2015).

We recently used functional genomics tools to test the polyphosphorylation status of an additional 90 PASK-containing proteins in yeast (Bentley-DeSousa et al., 2018). We identified 15 novel targets including a conserved network involved in ribosome biogenesis. Based on these targets we defined a new function for Vtc4 in ribosome biology. We also identified 6 human proteins that can undergo polyphosphorylation, suggesting conservation of this PTM (Bentley-DeSousa et al., 2018). While our work points to the exciting possibility that polyphosphorylation is a wide-spread PTM (Bentley-DeSousa et al., 2018), its function and

regulation remain unclear. In this mini review, we outline the critical open questions in the new field of lysine polyphosphorylation.

Open questions in the field of polyphosphorylation

What determines whether a protein is polyphosphorylated?

We have little understanding of why polyphosphorylation happens on some proteins and not others. The 15 targets uncovered in Bentley-DeSousa et al. (2018) were identified from a pool of > 400 shortlisted candidates with stretches of amino acids similar to the PASK motifs of the two founding targets, Nsr1 and Top1. Evidence supports that lysines are the site of polyphosphorylation in the PASK motifs (Azevedo et al., 2015). While the serine residues are dispensable for modification, there has been little effort to deter- mine the contribution of the acidic residues or the importance of lysine spacing (Azevedo et al., 2015). Expressing PASKs as fusions with non-targets confers polyphosphorylation of those proteins, suggesting that the requirements for modification are contained therein (Azevedo et al., 2015; Bentley-DeSousa et al., 2018). It is unknown if polyphosphorylation can occur in other sequence contexts. The field would benefit from the use of peptide arrays to systematically identify amino acid sequences that promote or inhibit polyphosphorylation. This experiment is a critical step in understanding the chemistry of how polyphosphorylation occurs.

Although polyphosphorylation bears hallmarks of a covalent modification (Azevedo et al., 2015), future work should continue to test this assertion, which is likely to remain somewhat controversial until explicitly demonstrated. Peptides bearing polyphosphorylated lysine residues have not been observed by mass spectrometry (MS). In contrast, MS has been used to identify monophosphorylated lysine residues proteome-wide (Hardman et al., 2017). Thus, minimally, the Nε–P phosphoramidate bond that is also characteristic of polyphosphorylation can be

preserved and detected by current MS methods. It is possible some monophosphorylated lysines mapped in this study correspond to the remnants of longer polyP chains.

How is polyphosphorylation regulated if chain addition is non-enzymatic?

Regulation by polyP synthesis

Our recent work shows that deletion mutants of the VTC complex or vacuolar regulators VAM3 or VMA22 show decreased polyphosphorylation of multiple targets in yeast (Bentley-DeSousa et al., 2018). These mutants have low levels of polyP compared to wild-type cells (Freimoser et al., 2006), highlighting polyP bioavailability as a key point of regulation for lysine polyphosphorylation (**Figure 39**). This connection is not entirely surprising and is reminiscent of the regulation of lysine acetylation by fluctuations in Acetyl-CoA (Cai et al., 2011; Weinert et al., 2014). Surprisingly, over 200 other non-essential gene deletion mutants also show measurable defects in polyP levels (Freimoser et al., 2006). Whether all of these mutants result in decreased target polyphosphorylation is unknown.

Most polyP is constrained in the vacuole following its synthesis by the VTC complex (Gerasimaite et al., 2014). The exact proportion of polyP in other organelles is unclear. Although most polyphosphorylated targets identified to date localize to the nucleus, estimates of the nuclear polyP levels vary widely (from very little to 18% of total) depending on the study in question (Azevedo et al., 2015; Saito et al., 2005). The Mayer lab demonstrated that accumulation of cytoplasmic polyP is toxic in yeast (Gerasimaite et al., 2014), suggesting that transit of polyP throughout the cell (from the vacuole to the nucleus, for example) might be highly regulated. As an alternative to transport, polyP could be synthesized locally. Interestingly, the Vtc2 variant of the VTC complex localizes around the nucleus in addition to the vacuole, but it is unclear if this complex impacts accumulation of nuclear polyP (Hothorn et al., 2009). PolyP synthesis also appears to be cell cycle regulated, and it would be interesting to determine the

impact of this regulation on unique pools of polyP (Bru et al., 2016b; Jimenez et al., 2017). The identification of mutants influencing polyP accumulation in specific organelles would provide a unique avenue to study regulation of polyphosphorylation in vivo.

Via NuPAGE analysis, the target protein populations exist almost exclusively in the modified (supershifted) state (Azevedo et al., 2015; Bentley-DeSousa et al., 2018). Yet, because polyphosphorylation can occur non-enzymatically even under harsh denaturing conditions (Azevedo et al., 2015; Bentley-DeSousa et al., 2018), the percentage modification occurring in vivo versus in vitro after cell lysis has occurred is unclear. This issue limits our understanding of how polyphosphorylation occurs and how it might be removed (see below). Careful competition experiments pairing in vivo orthophosphate labeling with lysis buffers containing very high concentrations of unlabeled polyP may provide insight into this critical question.

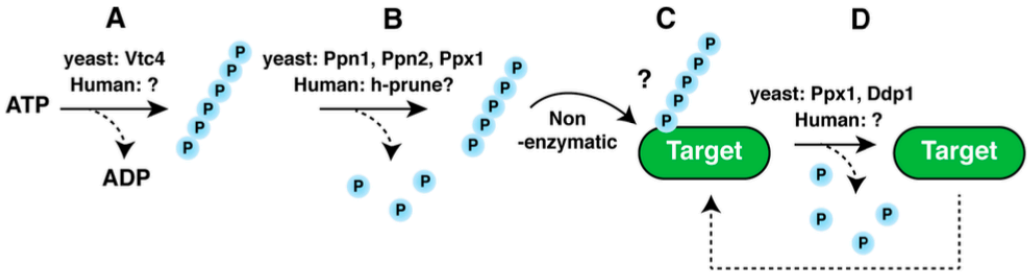


Fig. 39: Critical points of regulation for polyphosphorylation. A) Polyphosphate biosynthesis requires Vtc4 in yeast but the human polyP synthesis machinery is unknown. B) Exo and Endo polyphosphatases may trim chains prior to addition to protein targets. C) Little is known about the non-enzymatic mechanism of polyP attachment to lysine residues or the impact of amino acids surrounding the modified residues. D) Polyphosphatases can also trim polyP chains attached to protein targets in vitro, but whether this occurs in vivo is unclear.

Regulation by polyphosphatases

Polyphosphatases could either indirectly or directly regulate target modification. Indirect regulation refers to polyphosphatase regulation of the pool of polyP before it is attached to protein targets. We reported that simultaneous deletion of *PPN1* (*ppn1Δ*) and *PPN2* (*ppn2Δ*)

resulted in an increased electrophoretic shift for target proteins Fpr3 and Rts1 as judged by NuPAGE analysis (Bentley-DeSousa et al., 2018). Since *ppn1*Δ *ppn2*Δ double mutants show a striking increase in polyP chain length (Gerasimaite and Mayer, 2017), we suggest a model wherein the length of chains added is controlled primarily at the level of free polyP. In this model, the enhanced electrophoretic shifts observed for targets isolated from *ppn1*Δ *ppn2*Δ double mutants stem from an increase to the overall length of chains available for non-enzymatic addition to unmodified targets (**Figure 39**). It is unclear how the length of a polyP molecule impacts its ability to be added to target proteins.

Direct regulation refers to the action of polyphosphatases on polyP chains that are already attached to protein targets (**Figure 39**). In support of direct regulation, Azevedo et al. (2015) observed that purified exopolyphosphatases Ppx1 and endophosphatase Ddp1, can depolyphosphorylate Nsr1 in vitro. In contrast, our analysis of single *ppx1*Δ and *ddp1*Δ mutants and *PPX1*-overexpression strains showed wild-type polyphosphorylation of multiple targets (Bentley-DeSousa et al., 2018). These apparently contradictory data can be reconciled in a number of ways. First, Ppx1 and Ddp1 could function redundantly. In this case, analysis of targets extracted from *ppx1*Δ *ddp1*Δ double mutants may uncover changes to polyphosphorylated targets not observed in either single mutant. Second, depolyphosphorylation could occur locally and depolyphosphorylated proteins could exist for only a short period of time before they are repolyphosphorylated. The possibility of such direct regulation raises a number of important questions. For example, do polyP-binding proteins restrict polyphosphatase access to polyP chains? Even if chains can be completely removed, are there enzymes that can cleave the phosphoramidate (Nε–P) bond to restore lysine side chains to their unmodified form? Alternatively, do lysine residues on depolyphosphorylated proteins retain this bond and remain refractory to other modifications? Answers to these

questions will provide greater insight into the functional relevance of direct regulation by polyphosphatases.

What is the molecular function of polyphosphorylation?

Downstream functions of polyphosphorylation will depend on the roles played by individual protein targets. However, we anticipate that polyP chains could impact protein function by common mechanisms at the molecular level. Azevedo et al. (2015) provided evidence suggesting that polyP can impact localization, protein–protein interactions and enzymatic activity of protein targets Nsr1 and Top1. Our expanded dataset of 15 additional targets in yeast provides new opportunities to test the generality of these molecular phenotypes (Bentley-DeSousa et al., 2018).

How might polyP attachment mediate such changes? We and others have previously discussed several possibilities. These include (a) competition between polyP and other lysine-based PTMs (Azevedo et al., 2015; Azevedo and Saiardi, 2015; Bentley-DeSousa et al., 2018) (b) polyP occlusion of binding surfaces otherwise occupied by nucleic acids or other effector molecules (Azevedo et al., 2015; Bentley-DeSousa et al., 2018); or (c) the use of polyP phosphoanhydride bonds as a source of energy to drive molecular remodelling (Bentley-DeSousa et al., 2018). PolyP chains could have different functions depending on the length of chain. Short chains may be sufficient to block other PTMs, while longer chains could disrupt large binding surfaces. How longer chains interact with the protein may also be important for function: Do they flow outward from the modified residue or collapse back onto the protein surface? In addition to chain length, polyP function could be dictated by unique counterions or polyP binding proteins. It will be important to determine the full repertoire of such polyP ‘complexes’ that can be used in polyphosphorylation reactions.

How is polyphosphorylation (and polyphosphate) regulated in higher eukaryotes?

Concentration of polyP in higher eukaryotes

As discussed above, polyP is likely a critical regulator of polyphosphorylation. The concentrations of polyP in mammalian tissues and cells are much lower than the > 200 mM reported for yeast (Auesukaree et al., 2004). In 1995, Kumble and Kornberg measured appreciable concentrations of approximately 100 μM in mammalian heart and brain tissues (Kumble and Kornberg, 1995). The Ruiz lab also reported concentrations of ~ 65 μM in myeloma cells (Jimenez-Nunez et al., 2012). While these values are still low relative to yeast, localization of polyP to specific organelles could increase local concentrations. For instance, platelets are estimated to contain concentrations of up to 130 mM in substructures called dense granules (Ruiz et al., 2004), and this is important for functions related to blood coagulation (see below). In another example, polyP accumulates in nucleoli of myeloma cells (Jimenez et al., 2017). This finding is intriguing given our identification of a conserved network of nucleolar proteins as targets for polyphosphorylation (Bentley-DeSousa et al., 2018).

Regulation of polyP in higher eukaryotes

One of the more important open questions in the field of polyP biology is the identity of polyP regulators in higher eukaryotes. Both the polyP synthetase(s) and polyphosphatases are largely unknown in mammalian cells. Vtc4 and *E. coli* polyP synthesis enzymes have no convincing homolog in human cells. One exception may be the human protein h-prune (homolog of *Drosophila* prune) that belongs to a superfamily of phosphodiesterase enzymes and demonstrates exopolyphosphatase activity on short polyP chains (~ 3–4 phosphate units long) in vitro (Tammenkoski et al., 2008). Whether this in vitro activity translates to in vivo function relevant for polyphosphorylation is unclear. The discovery of enzymes responsible for polyP metabolism in mammals will be a significant step forward for the field.

Tools to study polyP in mammalian cells

Because these regulators remain uncharacterized, we and others have studied polyP biology in mammalian cell lines via ectopic expression of polyP metabolic enzymes from other species (Bentley-DeSousa et al., 2018; Wang et al., 2003). For example, we used ectopic expression of PPK1 from *E. coli* (*Ec*PPK1) in HEK293T cells to artificially increase the concentration of polyP (Bentley-DeSousa et al., 2018). We identified 6 human proteins that showed an electrophoretic shift on NuPAGE gels following *Ec*PPK1 transfection, consistent with an ability of these proteins to undergo polyphosphorylation (Bentley-DeSousa et al., 2018). We argue that the use of the *Ec*PPK1 could exaggerate polyphosphorylation that is already occurring at lower levels in control cells. Testing this idea will be facilitated by the identification of the human polyP synthetase(s).

Is there a role for polyphosphorylation in human disease?

As yet, there are no demonstrated links between mis-regulation of lysine polyphosphorylation and human disease. Considering the novelty of the PTM and the challenges described above, this is hardly surprising. Two points make us hopeful for an eventual medical application. First, over 100 human proteins contain putative PASK motifs and these proteins are together implicated in diverse pathologies (Bentley-DeSousa et al., 2018). Moreover, in many instances human PASK motifs are situated in regions that may be differentially expressed among predicted isoforms (Bentley-DeSousa et al., 2018). Second, polyP, independently of polyphosphorylation, has been well studied in the context of pathways disrupted in important medical conditions. For example, polyP is likely to play roles in cell growth and cancer (Arelaki et al., 2018; Jimenez-Nunez et al., 2012; Tsutsumi et al., 2017; Wang et al., 2003), bone formation (Hacchou et al., 2007; Hoac et al., 2013; Usui et al., 2010), bacterial virulence (Kim et al., 2002; Peng et al., 2016; Rashid and Kornberg, 2000; Rashid et al., 2000a; Rashid et al.,

2000b; Shiba et al., 1997), and immune function (Bae et al., 2012; Hernandez-Ruiz et al., 2006; Wat et al., 2014). Below we expand briefly on additional polyP connections to blood coagulation and neurological function, with an emphasis on where polyphosphorylation could play a role.

PolyP and the blood coagulation cascade

As detailed above, polyP chains are stored within the dense granules of platelets (Ruiz et al., 2004). They are released following activation of the coagulation cascade (Moreno and Docampo, 2013; Ruiz et al., 2004). PolyP chains induce clotting by affecting factor V (Choi et al., 2015; Smith and Morrissey, 2008; Smith et al., 2006), factor VII (Muhl et al., 2009), factor XII (Mitchell et al., 2016; Muller et al., 2009; Puy et al., 2013), fibrinogen/fibrin clots (Mutch et al., 2010a; Whyte et al., 2016), thrombin (Choi et al., 2011; Mutch et al., 2010b) and tissue factor pathway inhibitor (TFPI) (Puy et al., 2016; Smith et al., 2006). Notably, the critical amino acids for polyP binding to thrombin are lysine residues (Mutch et al., 2010b). Testing whether this interaction is impacted by covalent polyphosphorylation may shed light on the mechanism of activation for thrombin and other proteins. Finally, the impact of polyP on the coagulation cascade is chain length dependent: long chains (> 500 phosphates long) activate the contact pathway, whereas shorter chains (~ 100 phosphates long) play roles in factor V activation and the tissue factor pathway (Smith et al., 2010).

PolyP and neurological function

Astrocytes, non-neuronal glial cells that support neuron homeostasis, respond to exogenous polyP by releasing polyP of their own, which is taken up by neurons. Given this property, it was hypothesized that polyP may act as a gliotransmitter (Holmstrom et al., 2013). Follow-up work demonstrated that polyP is also released by neurons themselves in response to depolarization to regulate neuronal activity. This suggests a widespread role for polyP as a neuromodulator

(Stotz et al., 2014). Intriguingly, polyP was also observed to localize in discrete punctae along axon projections (Stotz et al., 2014). It is tempting to speculate that these concentrated regions of polyP may provide an environment conducive to polyphosphorylation.

PolyP chains have also been described as molecular chaperones (Cremers et al., 2016; Gray et al., 2014). This may be particularly relevant in neurodegenerative diseases such as Alzheimer's, Huntington's, and Parkinson's disease, which involve the accumulation of potentially toxic protein aggregates (Ross and Poirier, 2004). Elegant work by Cremers et al. suggests a model wherein polyP accelerates the formation of protein aggregates called fibrils while bypassing the accumulation of more toxic fibril intermediates (Cremers et al., 2016). Whether these effects involve lysine polyphosphorylation should be a topic of future investigation.

Is polyphosphorylation a viable therapeutic target?

Given the links to human disease described above, polyP chains are attractive targets for the development of novel therapeutics. Recent work has seen the development of strategies to increase and decrease polyP in vivo. While these strategies are designed to regulate polyP bioavailability, similar techniques could also be effective in modulating pathways uniquely impacted by lysine polyphosphorylation.

Increasing polyP levels

Therapies to increase polyP levels in blood could serve as novel treatment for bleeding disorders or as a front line first aid treatment. These may be particularly relevant for deep or internal injuries that are difficult to treat using other methods. With this goal in mind, several groups have made progress in technologies having potential to deliver polyP chains in vivo. A recent advancement is the development of artificial dense granules (Donovan et al., 2016). Artificial dense granules are designed to mimic mammalian dense granules and consist of a

lipid outer layer containing polyP within its core. These artificial dense granules can induce factor XII activation in vitro (Donovan et al., 2016). Ultimately, it will be important to test whether similar strategies can be used to increase polyphosphorylation of prospective mammalian targets.

Inhibiting polyP

Strategies to inhibit polyP function are also being sought by various groups. For example, (Labberton et al., 2016)describe expression of polyP binding domains to reduce polyP bioavailability or activity in vivo. Small molecules such as spermine have also been observed to inhibit polyP's binding to thrombin (Smith et al., 2012). In this context, small molecules could serve a similar purpose to polyP blocking proteins, although their effectiveness for broader applications is uncertain.

PolyP inhibition may also be a strategy to combat fungal or bacterial infections (Rao et al., 2009). Certain fungi and bacteria have increased virulence due to increased transcription of virulence-promoting genes or by developing biofilms and these traits correlate with polyP levels (Peng et al., 2016; Rashid and Kornberg, 2000; Shiba et al., 1997). Prokaryotic polyP kinase (most commonly called PPK) does not share significant homology with any human proteins, making it an attractive target for infection control. Notably, mesalamine, currently used in the treatment of inflammatory bowel disease and ulcerative colitis, was recently identified as a PPK inhibitor (Dahl et al., 2017). It is tempting to speculate that its success may be due to its ability to kill or inhibit the growth of gut bacteria via disruption of polyP synthesis. Recently, in silico drug screening has uncovered new candidate PPK1 inhibitors, although the impact on polyP synthesis is currently unclear (Bashatwah et al., 2018). Nevertheless, this strategy has promise for the modulation of polyP synthesis and, potentially, polyphosphorylation of target proteins in prokaryotic cells.

Outlook

The new field of polyphosphorylation is exciting, with each experiment raising new questions in addition to providing answers to previous ones. There exists significant potential for polyphoshorylation and polyphosphate biology in general to have an important impact in diverse areas of health research. But, along with the push for an application for polyphosphate biology, we cannot neglect the fundamental questions related to biochemistry and molecular function. Answers to these questions will provide a solid foundation for all subsequent discoveries.

Acknowledgements

We thank the members of the Downey lab for critical review of the manuscript. Polyphosphate work in the Downey lab is funded by the Canadian Institutes of Health Research (PJT-148722)