The gag-like gene *RTL8* antagonizes PEG10-mediated virus like particles in humans

Will Campodonico[1], Holly H. Black[1], Cristina I. Lau[1], and Alexandra M. Whiteley[1]

[1]Department of Biochemistry, University of Colorado, Boulder, CO, USA

*e-mail: alexandra.whiteley@colorado.edu

Abstract

Transposable elements can cause catastrophic genomic instability and are tightly regulated by the host through multiple restriction pathways. Domesticated elements are derived from these ancestors and have evolved adaptive roles but may retain pathological activity. For example, *PEG10* is required for placentation but also promotes cancer and neurodegeneration. While much of *PEG10* remains poorly understood, a unique feature is its ability to readily form virus-like particles (VLPs) which may contribute to its adaptive and pathological nature. However, comparable restriction networks that antagonize the harmful potential of domesticated genes like *PEG10* remain unknown. Here, we describe restriction of PEG10 VLP abundance via *UBQLN2* and the domesticated retrotransposon *RTL8*. The gag-like *RTL8* antagonizes PEG10 through competitive incorporation into VLPs reminiscent of transposable element inhibitors from diverse eukaryotes. These results represent the first known instance of a retroelement-derived restriction factor antagonizing another domesticated retrotransposon and have implications for our understanding of PEG10 biology.

Introduction

Transposable elements (TEs), which include DNA transposons and RNA retrotransposons, are virus-like genetic elements capable of replicating within the genome through a variety of complex mechanisms. Due to the catastrophic consequences of unchecked genomic integration, eukaryotes have developed extensive networks to restrict transposable element expression and activity[1]. These restriction mechanisms act at transcriptional, translational, and post-translational stages of the TE lifecycle, and are found in diverse eukaryotes from yeast to humans[1]. Inhibition or loss of these restriction factors results in elevated TE activity[2,3], which can result in genomic instability and human disease.

Rather than being restricted by host factors, expression of a subset of TE-derived genes is essential for organismal health[4]. These genes are referred to as 'domesticated'[4,5] and have a variety of adaptive roles, including the maintenance of genomic integrity[6,7], development[5], and viral restriction[8–10]. However, despite the selective pressure towards host fitness, recent studies have shown that a subset of these domesticated elements retain pathological effects and can directly contribute to human disease[6,11,12]. Therefore, it would be advantageous for the host to use the same, or adapted TE restriction factors to regulate their domesticated counterparts. However, no such restriction factors of domesticated elements have been identified.

The domesticated retrotransposon *Paternally Expressed Gene 10* (*PEG10*) embodies a balance between adaptation and pathology: *PEG10* is necessary for placentation in mice[13,14], but also drives cancer and neurological disease in humans[15–19]. The mechanisms of *PEG10*-mediated adaptive and pathological roles remain poorly understood, but may involve the protein's rare ability to form virus-like particles (VLPs) that are exported from the cell[20–22]. For the domesticated gene *Arc*, formation and release of a VLP is essential to its adaptive function in the organism[23,24]. Like *Arc*, *PEG10*-derived VLPs may contribute to either its adaptive or pathological roles.

Here, we describe a multifaceted regulatory network that antagonizes PEG10 VLP formation and release from human cells. It was recently discovered that the proteasome shuttle factor ubiquilin 2 (UBQLN2) selectively targets PEG10 for degradation[18], and here we show that this translates to decreased VLP abundance. We also describe a novel role for the human gene *RTL8*, a domesticated retrotransposon with structural similarity to the N-terminal lobe of PEG10 gag[25], in the restriction of *PEG10*-derived VLPs. RTL8 binds to the PEG10 N-terminal lobe, is found in PEG10 VLPs, and causes PEG10 to be retained in the cell. Together, these findings suggest a model in which RTL8 competitively incorporates into PEG10 VLPs, thereby inhibiting their assembly or release. RTL8's restriction of PEG10 appears to be species-specific, suggesting that the two domesticated retroelements co-evolved in conflict. These results illuminate a new role for *RTL8* in the inhibition of PEG10 activity and may have important

implications for reproductive and neurological health. These results also identify a novel conflict between two domesticated retrotransposons in the human genome and highlight the unique demands of balancing beneficial and detrimental roles of TEs and TE-derived genes.

Results

PEG10 spontaneously releases virus-like particles in human cell lines

TEs come in many forms, with diverse genetic structures and mechanisms of propagation[4,26,27]. *PEG10* evolved from the Ty3/*gypsy* family[25] of LTR retrotransposons that replicate through an RNA intermediate and encode a classical *gag-pol* genetic structure. Like its ancestors, *PEG10* retains an intact *gag* and partial *pol* open reading frame and utilizes a programmed -1 ribosomal frameshift to allow for translation of both gag and gag-pol polyproteins from a single transcript (Fig. 1a)[28,29]. The *gag* open reading frame codes for the structural capsid and nucleic acid-binding zinc finger (NC)[20,21,30]. The *pol* open reading frame codes for an aspartic protease (PR) but lacks the reverse transcriptase (RT), RNAse (RH), and integrase (IN) domains necessary for genomic reincorporation (Fig. 1a)[25].

Expression of PEG10 gag and gag-pol in cells results in the release of VLPs into cell culture medium[20–22]. The contribution of PEG10-derived VLPs to the adaptive and pathological roles of PEG10 remains unclear; however, they have been engineered for the production of RNA-containing particles that can deliver functional nucleic acid to human cells[21]. In both cases, the factors that regulate PEG10-derived VLP production and release are unknown. To better understand PEG10 VLP production in a native context, we quantified VLP abundance in conditioned medium from a range of cultured human cell lines. We tested the abundance of PEG10 in cell lysate (Fig. 1b,d) and conditioned medium (Fig. 1c,e) following enrichment of extracellular particles by ultracentrifugation. PEG10 abundance in lysate and the VLP fraction was measured by western blot using a polyclonal antibody generated against PEG10 gag, allowing us to separately quantify the abundance of gag and gag-pol proteins, as well as their respective contribution to VLPs, in each cell line.

All cell lines expressed detectable levels of PEG10 gag and gag-pol protein (Fig. 1b,d). The hepatocellular carcinoma HepG2 and the human trophoblast hTR-8 cells had a significantly higher ratio of gag-pol:gag protein than other lines tested (Extended Data Fig. 1a), which could be caused by unique regulation of *PEG10*'s frameshifting or differential stability of the two products in the cells.

In contrast to cell lysate, gag and gag-pol proteins were only detectable in the VLP fraction of some cell lines (Fig. 1c,e). HepG2 cells had the highest lysate abundance and VLP production of all cell lines tested (Fig. 1b-e). However, across all cell lines, lysate PEG10 only correlated weakly with VLP abundance (Extended Data Fig. 1b). For example, hTR-8 cells had significantly less PEG10 in lysate (Fig. 1b,d) but

nearly identical VLP abundance compared to HepG2 cells (Fig. 1c,e). The remaining cell lines (HEK293, A549, and U87) had comparable PEG10 abundance in cell lysate to hTR-8 cells, but negligible VLP abundance (Fig. 1c,e).

Elevated PEG10 protein expression has been implicated in the neurodegenerative disease Amyotrophic Lateral Sclerosis (ALS) and the neurodevelopmental disease Angelman syndrome[18,19,22]. In both conditions, primary cultured neurons are commonly used as a model of cellular dysfunction. We tested for PEG10 VLPs in conditioned media from human iPSC-derived neurons and observed a strong PEG10 gag band (Extended Data Fig. 2), though the presence of a cytoplasmic control band means that we cannot rule out the possibility of contamination with nonspecific cellular fragments.

UBQLN2 regulates PEG10 gag-pol abundance and VLP release in a human trophoblast cell line

Elevation of the ratio of gag-pol:gag protein levels in HepG2 and hTR-8 cells, paired with the enhanced VLP abundance in these cell lines, suggested that a component of the pol open reading frame may increase the efficiency of PEG10-derived VLP formation and/or release. Recent studies have shown that the proteasome shuttle factor ubiquilin 2 (UBQLN2) specifically regulates the abundance of PEG10 gag-pol, but not gag, by targeting the protein for proteasomal degradation[18,22]. To test whether UBQLN2 similarly restricts VLP abundance, we stably introduced shRNAs targeting *UBQLN2* to HepG2 and hTR-8 cells (Fig. 2a,b, Extended Data Fig. 3a,b). Consistent with previous work[18], knockdown of *UBQLN2* in hTR-8 cells led to a greater than two-fold increase in PEG10 gag-pol abundance in cell lysate, with no significant change to gag abundance (Fig. 2a,c-d). In hTR-8 cells, knockdown of *UBQLN2* also resulted in a two-fold increase in VLP abundance (Fig. 2e,f). Depletion of *UBQLN2* in HepG2 cells did not result in changes to gag-pol or gag in cell lysate, or PEG10 VLP abundance (Extended Data Fig. 3), which may be due to a large excess of PEG10, or additional regulatory factors masking a UBQLN2-dependent effect.

PEG10 is a regulatory intermediate between UBQLN2 and RTL8

UBQLN2 has also been reported to share a functional relationship with *RTL8*, a domesticated retrotransposon that resembles *PEG10*[18,31]. Like *PEG10*, *RTL8* is a member of the Mart (or Mar) family of domesticated retrotransposons derived from the Ty3/*gypsy* lineage[25]. However, *RTL8* consists of a partial gag gene with high sequence homology to the capsid^{NTD} of *PEG10* (Fig. 3a). While the disruption of *UBQLN2* in mouse tissues and human cells leads to accumulation of PEG10 protein, it simultaneously results in a profound and unexpected loss in RTL8 protein levels independent of transcription, translation, or degradation rates (Fig. 3b)[18]. Similarly, knockdown of *UBQLN2* in hTR-8 cells drove a 50% reduction in RTL8 abundance with no transcriptional changes (Fig. 3c,e, Extended Data Fig. 4a). This relationship was more pronounced upon genetic deletion of *UBQLN2* – in a HEK293 cell line lacking

UBQLN 1, *2*, and *4* (HEK293^{TKO})[32], we observed a robust increase in PEG10 gag-pol (Extended Data Fig. 4b) and near complete loss of RTL8 (Fig. 3d,f).

The existence of a regulatory relationship between *UBQLN2* and two highly similar genes of the same domesticated retrotransposon family led us to examine whether *PEG10* and *RTL8* share a direct relationship in addition to their regulation by *UBQLN2* (Fig. 3b). We introduced an shRNA against *PEG10* into *UBQLN*^{TKO} cells and saw a partial rescue of RTL8 (Fig. 3g, h), confirming that *UBQLN2*'s influence on RTL8 is at least partially mediated by PEG10.

RTL8 interacts with PEG10 capsid$^{NTD\text{-}lobe}$

After confirming that PEG10 is a regulatory intermediate between UBQLN2 and RTL8, we tested whether PEG10 and RTL8 interact to mediate this regulation. *RTL8* exists as three nearly identical genes (*RTL8a*, *b*, and *c*); we expressed HA-PEG10 and FLAG-RTL8c in HEK293 cells and performed crosslinking co-immunoprecipitation. Both proteins were readily visualized by western blot of cell lysate (Fig. 4a,b). HA-PEG10 was visible as four distinct bands representing gag-pol, gag, and two self-cleavage products generated via activity of the PEG10 protease domain[22]. The 37 kDa band represents a retrovirus-like capsid fragment, and the 22 kDa band reflects a protein fragment consisting of the N-terminal ~100 amino acids of the protein (referred to as 'capsid$^{NTD\text{-}fragment}$'). Upon immunoprecipitation of HA-PEG10, we saw robust co-immunoprecipitation of FLAG-RTL8c (Fig. 4a). When we performed the reciprocal experiment (IP-FLAG), we observed co-immunoprecipitation of HA-PEG10 gag-pol, gag, and capsid, but not the shorter proteolytically cleaved capsid$^{NTD\text{-}fragment}$ (Fig. 4b), indicating that while gag and capsid are sufficient to interact with RTL8c, the very N-terminal region of PEG10 is insufficient.

Retroelement gag proteins have two distinct lobes of the structural capsid (Fig. 1a): the N-terminal domain (capsid$^{NTD\text{-}lobe}$) and C-terminal domain (capsid$^{CTD\text{-}lobe}$), each of which contribute specific multimerization properties to the final capsid structure[33]. During retroelement capsid assembly, NTD lobes self-associate into penta- and hexameric cones, which are held together by homotypic CTD:CTD interactions on adjacent cones[33–35]. Based on sequence alignment and structural modeling, PEG10 capsid closely resembles that of the ancestral Ty3 retrotransposon and maintains the N-terminal cone structure as well as C-terminal dimerization domain (Extended Data Fig. 5a,b). Indeed, recent structural studies confirm that PEG10 capsid$^{CTD\text{-}lobe}$ closely resembles ancestral Ty3 capsid and is capable of dimerization[36].

To more precisely map the interaction between PEG10 and RTL8, we generated PEG10 capsid truncation constructs representing the N- and C-terminal lobes (HA-capsid$^{NTD\text{-}lobe}$ and HA-capsid$^{CTD\text{-}lobe}$ Fig. 4c), co-expressed these constructs in HEK293 cells with FLAG-RTL8c, and immunoprecipitated

FLAG. FLAG-RTL8c interacts with HA-PEG10 gag-pol, gag, and capsid$^{NTD\text{-}lobe}$, but not capsid$^{CTD\text{-}lobe}$ (Fig. 4d, Extended Data Fig. 5c), indicating that capsid$^{NTD\text{-}lobe}$ is both necessary and sufficient for interaction with RTL8. Co-immunoprecipitation of HA-gag and HA-capsid$^{NTD\text{-}lobe}$ with FLAG-RTL8c results in the appearance of a higher molecular weight band that stains positive for both FLAG and HA (Fig. 4d, Extended Data Fig. 5c, asterisks). This band is at the correct molecular weight to be the crosslinked dimer of PEG10 with RTL8, which furthers the hypothesis that these proteins interact in the cell.

It was surprising that FLAG-RTL8c was unable to co-immunoprecipitate the small fragment of PEG10 generated upon self-cleavage (capsid$^{NTD\text{-}fragment}$) (Fig. 4b) given that the slightly larger HA-capsid$^{NTD\text{-}lobe}$ construct was sufficient (Fig. 4d). We have previously predicted with high confidence that PEG10 self-cleavage within the NTD lobe occurs at amino acid 113-114 and cleavage after the CTD lobe occurs at amino acids 259-260[22]. We generated capsid truncation constructs that mimic these natural cleavage events (HA-capsid$^{NTD\text{-}fragment}$, amino acids 1-111, and HA-capsid$^{CTD\text{-}fragment}$, amino acids 112-259, Extended Data Fig. 5d) and again immunoprecipitated FLAG-RTL8c upon co-expression. Neither HA-capsid$^{NTD\text{-}fragment}$ nor HA-capsid$^{CTD\text{-}fragment}$ were capable of interaction with FLAG-RTL8c (Extended Data Fig. 5e), consistent with results in Fig. 4b and suggesting that the entirety of the capsid$^{NTD\text{-}lobe}$ is necessary for interaction between PEG10 and RTL8. Based on these data, as well as the structural homology of PEG10 capsid$^{NTD\text{-}lobe}$ and RTL8, we conclude that the PEG10:RTL8 interaction resembles the homotypic capsid NTD:NTD interactions typical of the retroelement capsid assembly process.

RTL8 incorporates into PEG10 VLPs and decreases the efficiency of release

Due to the structural similarity (Fig. 5a) and demonstrated interaction between RTL8 and PEG10 capsid (Fig. 4), we hypothesized that RTL8 can incorporate into PEG10 VLPs. FLAG-RTL8c was co-expressed in HEK293 cells with empty vector or with HA-PEG10 gag-pol, followed by analysis of the VLP fraction in ultracentrifuged media by western blot. FLAG-RTL8c was only found in the VLP fraction when HA-PEG10 was co-expressed (Fig. 5b), indicating that RTL8 is unable to independently form VLPs but can incorporate into VLPs produced by PEG10. To determine an approximate ratio of PEG10:RTL8 abundance in ultracentrifuged VLPs, we expressed HA-tagged PEG10 along with an HA-tagged RTL8c to compare the abundance of HA tags as a proxy for PEG10:RTL8 stoichiometry. The ratio of HA-RTL8c:HA-PEG10 is decreased in the VLP fraction relative to cell lysate, and HA-RTL8c exists at approximately 5% the abundance of HA-PEG10 in the VLP fraction (Extended Data Fig. 6), together indicating either that RTL8 incorporation into PEG10-derived VLPs is inefficient, or that only low levels of RTL8 incorporation are tolerated before VLP formation or release is impossible. Consistent with the latter possibility, co-expression of FLAG-RTL8c also decreased the overall abundance of HA-PEG10 VLPs in conditioned media (Fig. 5c,d, Extended Data Fig. 7). The inhibitory effect was more pronounced against forms of PEG10 lacking a functional pol region: gag alone, or a protease dead mutant (gag-

pol^{ASG}), had a larger magnitude of change when co-expressed with FLAG-RTL8c despite higher baseline abundance of VLPs (Fig. 5d, Extended Data Fig. 8).

If RTL8 inhibits VLP assembly or release from cells, co-expression of PEG10 with RTL8 should lead to accumulation of intracellular PEG10 due to the decreased export of PEG10 in VLPs. We examined retention of intracellular PEG10 with a flow cytometry-based fluorescent reporter of PEG10 abundance (Fig. 5e, Extended Data Fig. 9a). As proof of principle, we overexpressed PEG10-Dendra2 in HEK293^{WT} and HEK293^{TKO} cells and saw an increase in fluorescence in cells lacking *UBQLN1, 2,* and *4* (Extended Data Fig. 9b). Co-expression of the PEG10 reporter construct with FLAG-RTL8c also led to increased intracellular fluorescence (Fig. 5f), consistent with intracellular accumulation of PEG10 due to RTL8 restriction and furthering the hypothesis that RTL8 inhibits either PEG10 VLP assembly or release.

Three paralogs of *RTL8* (*RTL8a*, *b*, and *c*) are all found on the X-chromosome and share >92% identity (Extended Data Fig. 9c)[25]. We tested all three *RTL8* paralogs, without and with their 3' UTRs, for the ability to induce retention of intracellular PEG10-Dendra2 and saw no differences in the intracellular retention of PEG10 (Extended Data Fig. 9d). Mice also express three *RTL8* paralogs from the X-chromosome which share comparatively less identity (Extended Data Fig. 9c). Expression of mouse FLAG-RTL8b, which shares the highest homology to human RTL8, had no effect on the intracellular retention of human PEG10-Dendra2 by flow cytometry (Fig. 5g), highlighting the species specificity of the PEG10:RTL8 relationship.

RTL8 is a competitive inhibitor of PEG10 VLPs

Given that RTL8 closely resembles and interacts with PEG10 capsid$^{NTD\text{-}lobe}$, we speculated that RTL8 restricts VLP abundance by interfering with essential PEG10 NTD:NTD interactions necessary for proper VLP formation. By binding to the capsid$^{NTD\text{-}lobe}$ of PEG10, RTL8 competes with other PEG10 monomers for capsid$^{NTD\text{-}lobe}$ self-association, thereby interfering with the proper assembly of VLPs. In this model of interference, co-transfection of a truncated PEG10 capsid$^{NTD\text{-}lobe}$ with full-length PEG10 should similarly restrict VLP abundance. PEG10-Dendra2 was co-expressed with an unlabeled capsid$^{NTD\text{-}lobe}$ fragment, resulting in similar levels of intracellular PEG10-Dendra2 retention as compared to RTL8 co-expression (Fig. 5f). These data support the model that RTL8 antagonizes VLP release through its interactions with capsid^{NTD} and suggest that RTL8 specifically acts as a competitive inhibitor of PEG10 capsid assembly.

Discussion

Based on the data presented here, we propose a complex regulatory model of PEG10 VLP formation and release from human cells (Fig. 6). *PEG10* is tightly regulated at the transcriptional level, both in healthy and diseased tissues[37]. Once translated, PEG10 gag-pol protein is degraded by UBQLN2, which

limits the availability of PEG10 monomers for VLP assembly. In parallel, RTL8 inhibits VLP formation or release through competitive incorporation via RTL8:capsid$^{NTD\text{-}lobe}$ interactions. Low levels of RTL8 incorporation are compatible with PEG10-derived VLPs, as RTL8 was detectable in PEG10 VLP preparations despite an overall loss in VLP abundance. This relationship may explain the paradoxical gain and loss of PEG10 and RTL8 observed upon *UBQLN2* perturbation: accumulation of PEG10 upon loss of *UBQLN2* increases VLP abundance, which promotes maximal incorporation of RTL8 into VLPs and subsequent loss of RTL8 protein from cell lysate. Presumably, any higher levels of RTL8 incorporation would prevent the formation or release of VLPs from the cell, resulting in further intracellular accumulation of PEG10.

An outstanding question raised by our results is where RTL8 acts in the capsid assembly and release process to restrict PEG10 VLP abundance. One possibility is that upon RTL8 incorporation into early capsid multimers, the absence of an RTL8 capsid$^{CTD\text{-}lobe}$ prohibits the formation of higher-order assemblies generated through both homotypic CTD:CTD and heterotypic NTD:CTD interactions that are necessary for retroelement capsid assembly[38–41]. A second possibility is that RTL8 incorporation disrupts proper localization of the PEG10 capsid or protein:protein interactions necessary for trafficking and release of VLPs. In support of this hypothesis, recent studies have shown RTL8 influencing the localization and function of UBQLN2[31], and RTL8 could have a similar effect on PEG10. Detailed investigation of PEG10 capsid assembly and release is warranted to distinguish these possibilities.

The ability of *RTL8* to antagonize PEG10-derived VLPs may be particularly relevant when considering the utility of PEG10 in generating biocompatible mRNA delivery systems, such as SEND[21]. Generation of PEG10-derived VLPs for the purposes of nucleic acid delivery in human cells may be limited in cell lines expressing abundant *RTL8*. Further, RTL8 may contaminate PEG10-derived VLPs due to low levels of incorporation. Alternatively, production of human PEG10 VLPs in cell lines derived from other species of origin may not be similarly inhibited, given the inability of murine RTL8 to induce cellular retention of human PEG10.

The regulation of PEG10 by UBQLN2 and RTL8 highlights the theme of genetic conflict. During embryonic development, there is conflict between paternally expressed genes, which promote growth of the offspring even at the expense of the mother, and maternally expressed genes, which limit the activity of paternal genes to maintain maternal health[42,43]. *PEG10* is a paternally expressed gene which is essential for placental development, but is linked to cancer and neurological disease[15–17,19,22]. *UBQLN2* and *RTL8* are located on the X-chromosome and in mice, placental expression of genes on the X-chromosome occurs exclusively from the maternal allele[44]. Our results therefore raise the intriguing possibility that both *UBQLN2* and *RTL8* represent maternal restriction factors that limit unchecked

activity of PEG10 in reproduction, and perhaps in other tissues, as directed by the paternal allele. In this sense, the domestication of *PEG10* could be considered incomplete; PEG10 retains the ability to promote pathology in non-reproductive tissues, and specific restriction factors have evolved to limit its activity to the tissues and expression levels that render it essential for survival.

Due to the overwhelming danger of unchecked genomic integration, the eukaryotic genome is littered with restriction factors that regulate TEs[1]. Some of these restriction factors are themselves derived from TEs: recently, a domesticated human *env* gene was discovered to limit the infectious capacity of type D retroviruses[9]. A second subset of restriction factors are derived from retroelement *gag* genes and act as competitive inhibitors of capsid formation, release, or uncoating. The sheep gag-derived gene *enJS56A1* restricts infection by the enJSRV family of viruses by disrupting capsid assembly and release[8,45]. The mouse *Fv1* gene is a gag-like restriction factor that targets incoming murine leukemiavirus (MuLV) particles by coating the endocytosed virus and preventing disassembly of the capsid[46]. In a mechanism that closely mimics the PEG10:RTL8 relationship, the Ty1 retrotransposon in yeast is tightly regulated by the presence of a cryptic start site within the *gag* open reading frame, which results in the production of a truncated capsid^{CTD} fragment. The truncated fragment prevents homotypic CTD:CTD interactions of intact capsid through competitive incorporation, thereby limiting Ty1 re-integration[47]. In each system, the *gag*-like genes use an affinity for capsid to inhibit the formation, release, or uncoating of the infectious particle. *Gag*-like genes are ubiquitous in the human genome[48], but no such *gag*-like restriction mechanisms for retroviruses or domesticated retrotransposons have been described in humans. Based on our data, we posit that *RTL8* can be included in this list of 'gag decoys' which act at the step of TE capsid formation or release.

The ability of *RTL8* to antagonize *PEG10* raises the possibility that *RTL8* is a yet-uncharacterized susceptibility locus in neurological diseases. Population analyses and *in vitro* screens have failed to link *RTL8* mutation to human disease. However, *RTL8* is under positive selection[25], implying the existence of an undiscovered adaptive function. These failures are likely due, at least in part, to the existence of three highly similar and potentially redundant *RTL8* genes, and due to their presence on the X-chromosome, which is excluded from typical GWAS studies. RTL8 interacts with UBQLN2 and contributes to protein quality control (PQC)[31], and defects in PQC pathways are a hallmark of neurodegeneration[49–51]. In this context, *RTL8* loss or mutation may mimic *UBQLN2*-mediated disease[52], as loss-of-function mutations in *RTL8* would result in elevated PEG10 VLP release. Due to the myriad roles of PEG10 in development, diverse disease, and applications of targeted mRNA delivery, further study of *RTL8* and its ability to inhibit PEG10 activity may provide valuable insight to advance both medicine and biotechnology.

1 **Materials and Methods**

Target	Clonality	Species	Vendor	Catalog #	Dilution	Application
UBQLN1/2	Monoclonal M03	Mouse	Abnova	H00029978	1:1000	WB
PEG10	Polyclonal	Rabbit	Proteintech	14412-1-AP	1:1000	WB
FAM127B (RTL8)	Polyclonal	Rabbit	Proteintech	20282-1-AP	1:1000	WB
FLAG	Monoclonal M2	Mouse	Sigma	F3165	1:5000	WB, IP
FLAG	Monoclonal D6W5B	Rabbit	CST	14793	1:5000	WB
HA	Monoclonal HA-7	Mouse	Sigma	H3663	1:5000	WB, IP
HA	Monoclonal C29F4	Rabbit	CST	3724	1:5000	WB
GAPDH	Monoclonal 14C10	Rabbit	CST	2118	1:5000	WB
Tubulin	Monoclonal DM1A	Mouse	Novus	NB100-690	1:10000	WB
α-mouse IgG 680	Polyclonal	Goat	Licor	926-68070	1:20000	WB
α-mouse IgG 800	Polyclonal	Goat	Licor	926-32210	1:20000	WB
α-rabbit IgG 680	Polyclonal	Goat	Licor	926-68071	1:20000	WB
α-rabbit IgG 800	Polyclonal	Goat	Licor	926-32211	1:20000	WB

2

3 Constructs

Construct Name	Insert	Tag	Species	NCBI protein reference	Vector
gag-pol	PEG10 (AA1-708)	2x HA (N-	*Homo*	NP_055883.2	pcDNA3.1

		terminal)	*sapiens*		
gag-pol^{ASG}	PEG10 (AA1-708)* D370A	2x HA (N-terminal)	*Homo sapiens*	NP_055883.2	pcDNA3.1
gag	PEG10 (AA1-325)	2x HA (N-terminal)	*Homo sapiens*	NP_001035242.1	pcDNA3.1
capsid$^{NTD\text{-}lobe}$	PEG10 (AA1-160)	2x HA (N-terminal)	*Homo sapiens*		pcDNA3.1
capsid$^{CTD\text{-}lobe}$	PEG10 (AA161-259)	1x HA (C-terminal)	*Homo sapiens*		pcDNA3.1
capsid$^{NTD\text{-}fragment}$	PEG10(AA1-111)	2x HA (N-terminal)	*Homo sapiens*		pcDNA3.1
capsid$^{CTD\text{-}fragment}$	PEG10 (AA112-259)	2x HA (C-terminal)	*Homo sapiens*		pcDNA3.1
FLAG-*Hs*RTL8a	RTL8a (AA1-113)	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071640.1	pcDNA3.1
FLAG-*Hs*RTL8b	RTL8b (AA1-113)	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071641.1	pcDNA3.1
FLAG-*Hs*RTL8c	RTL8c (AA1-113)	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071639.1	pcDNA3.1
FLAG-*Hs*RTL8a-3'UTR	RTL8a (AA1-113), with 700bp 3'UTR	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071640.1	pcDNA3.1
FLAG-*Hs*RTL8b-3'UTR	RTL8b (AA1-113), with 859bp 3'UTR	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071641.1	pcDNA3.1
FLAG-*Hs*RTL8c-3'UTR	RTL8c (AA1-113), with 821bp 3'UTR	3x FLAG (N-terminal)	*Homo sapiens*	NP_001071639.1	pcDNA3.1
HA-*Hs*RTL8c	RTL8c (AA1-113)	2x HA (N-terminal)	*Homo sapiens*	NP_001071639.1	pcDNA3.1
HA-*Hs*RTL8c-3'UTR	RTL8c (AA1-113), with 821bp 3'UTR	2x HA (N-terminal)	*Homo sapiens*	NP_001071639.1	pcDNA3.1

*Mm*RTL8b	RTL8b (AA1-113)	3x FLAG (N-terminal)	*Mus musculus*	NP_001018073.1	pcDNA3.1
gag-pol flow reporter	PEG10 (AA1-708)	Dendra2 (C-terminal), IRES-CFP	*Homo sapiens*	NP_055883.2	pDendra2

Cloning

All cloning was performed by Gibson assembly (Gibson HiFi master mix, Invitrogen) and transformed into chemically competent DH5α *E. coli* cells (Invitrogen). Transformed *E. coli* were plated on either 50 mg/mL kanamycin (Teknova) or 100 mg/mL carbenicillin (Gold Biotechnology) LB agar (Teknova) plates overnight at 37°C. Single colonies were picked and grown overnight in 5 mL LB Broth (Alfa Aesar) with kanamycin or carbenicillin at 37°C with shaking at 220 rpm. Shaking cultures were mini-prepped (Zymo) and sent for Sanger Sequencing (Azenta). Sequence verified samples were then grown in 50 mL LB Broth overnight with appropriate antibiotic at 37°C with shaking at 220 rpm. 50 mL cultures were midi-prepped (Zymo) for transfection.

Cell lines

hTR-8/SVneo cells were purchased from ATCC (CRL-3271). A549 and U-87 MG cells were a gift from Dr. Roy Parker (Department of Biochemistry, CU Boulder). HepG2 cells were obtained from ATCC (HB-8065) via the CU Boulder Biochemistry Shared Cell Culture Facility. HEK293 WT and *UBQLN1, 2, 4* knockout (TKO) cells were a gift from Dr. Ramanujan Hegde (Medical Research Council Laboratory of Molecular Biology, Cambridge England) and are described in Itakura et al.[32]

Induced cortical neurons were differentiated and cultured as previously described[53]. Cells were cultured on Matrigel (Corning) coated plates. Following differentiation, cells were maintained with half-media changes as described. The removed media from these half-media changes was collected, flash frozen, and stored at -80°C. Once ~20mL media had been collected, samples were thawed and virus-like particles isolated as described below.

All cells were maintained at 37°C with 5% CO_2. HEK293 and A549 cells were maintained in DMEM (Invitrogen) supplemented with 1% penicillin/streptomycin (Invitrogen), 1% L-glutamine (R&D 2 Systems, Inc.), and 10% FBS (Millipore Sigma). hTR-8/SVneo cells were maintained in RPMI 1640 (Invitrogen) supplemented with penicillin/streptomycin, L-glutamine, and 10% FBS. HepG2 and U87 MG cells were maintained in MEM (Invitrogen) supplemented with penicillin/streptomycin, L-glutamine, and 10% FBS.

Transfection

Cells were grown to 70% confluency and transfected with Lipofectamine 2000 (ThermoFisher) according to manufacturer's instructions. For 6-well plates, 2.5μg plasmid DNA was transfected per well. For 12-well plates, 1μg plasmid DNA was transfected per well. For 96-well plates, 0.1μg plasmid DNA was transfected per well. For cotransfections, equal mass amounts of each plasmid were added to the total amount listed above. Transfection mixture was prepared at a ratio of 1μg DNA:2.5μL Lipofectamine 2000. Unless otherwise stated, cells and media were harvested 48h following transfection.

Generation of stable cell lines

MISSION lentiviral packaging plasmids and shRNA plasmids were purchased from Sigma-Aldrich. Generation of Lentiviral delivery vectors was performed according to manufacturer's recommendation. Briefly, ~1.7×10^6 HEK293T cells were plated on 10cm dishes. The following day, cells were transfected with three plasmids containing VSV-G, Δ8.9 packaging plasmid, and the target shRNA using Lipofectamine 2000 (Invitrogen). The following day, virus-containing media was collected, filtered through a 0.45μm filter, and stored at -80°C until infection.

Target cells were plated in 6-well plates for lentiviral infection. 24 hr after plating (at 30-50% confluency) cells were infected with virus-containing media. Polybrene (EMD Millipore) was added to virus-containing media to a final concentration of 8μg/mL. 1mL of this infection media was added per well, and the plate centrifuged at 1800RPM in an Allegra X-12R tabletop centrifuge for 45 min at room temperature. The cells were incubated at 37°C for 3 hr, then 2mL additional growth media added and the cells incubated at 37°C overnight. The following day, this spinfection was repeated exactly. Two days after the second infection, selection for infected cells was initiated through the introduction of 2mg/mL (hTR-8/SVneo) or 8mg/mL (HepG2) puromycin (Gibco). Once stably infected, cells were maintained in growth media supplemented with puromycin.

Virus-like particle isolation

Virus-like particles were harvested from cultured media as previously described[22]. Briefly, T75 flasks were plated at 70% confluency (for endogenous VLP production), or 6-well plates (for overexpression experiments) were transfected as described above. Cells were grown for 48 hr and conditioned media was collected. Media was first centrifuged at 4700 x g for 10 min to remove cellular debris. The extracellular fraction was collected by ultracentrifugation using a preparative ultracentrifuge (Beckman Coulter) at 134,000 x g (Beckman SW41Ti rotor) for 4 hr at 4°C over a 30% sucrose (MP Biomedicals)

cushion. Media and sucrose were aspirated and the VLP-containing pellet was resuspended in western blot lysis buffer.

Western blotting

Cells were harvested by trypsinization, pelleted, and washed in PBS. Cell pellets were resuspended in in 8M urea (Fisher Chemical) containing 75 mM NaCl (Honeywell), 50 mM HEPES (Millipore Sigma) pH 8.5, with 1x cOmplete Mini EDTA-free protease inhibitor (Roche) and incubated at room temperature for 15 min. Lysate was cleared by centrifugation at 21,300 x g and the supernatant collected for western blot. Total protein was quantified by BCA assay (Pierce).

For endogenous protein analysis, 15μg total protein was loaded into NuPage 4-12% Bis-Tris precast protein gels (Life Technologies) per sample. For overexpression experiments, 4μg total protein was loaded per sample. Protein concentrations for gel loading were normalized with excess urea lysis buffer along with 1x Laemmli sample buffer supplemented with βME (Sigma Aldrich). Proteins were separated by SDS-PAGE using 1x NuPage MES SDS running buffer (Life Technologies). For samples lysed in urea, wells of the protein gel were equilibrated in urea lysis buffer for 10 min, and urea washed out before loading protein samples.

Proteins were transferred to nitrocellulose membrane (Amersham) using the Invitrogen Mini Blot Module. Membranes were blocked with Intercept (PBS) blocking buffer (LICOR) for 30 min at room temperature. Membranes were incubated with primary antibody overnight at 4°C. Membranes were incubated with secondary antibodies for 30 min at room temperature. After both antibody incubations, membranes were washed three times for 5 min each in TBST. Membranes were imaged on a LICOR Odyssey CLx and analyzed with LICOR ImageStudio software.

Unless otherwise stated, western blots were analyzed as follows. For cell lysate, target protein signal was normalized to tubulin as a loading control. This ratio was then normalized to the mean of the ratio for all samples within the experiment to account for technical variability across replicates. For endogenous VLP production, target protein signal was normalized to the mean target signal for all samples within the experiment. For overexpression VLP production, the overexpressed target protein signal was normalized to the tubulin-normalized cell lysate abundance of that target to account for variability in transfection efficiency. This ratio was then normalized to the mean of the ratio for all samples within the experiment.

Crosslinking Co-Immunoprecipitation

Cells were harvested by pipetting in ice cold PBS and centrifuged at 300 x g for 3 min, then resuspended in 300μL 0.1% PFA in PBS and incubated for 7 min at room temperature to crosslink proteins. Cells were collected by centrifugation at 300 x g for 3 min and washed three times with ice cold PBS. Crosslinked cells were lysed in 150μL lysis buffer containing 1% Triton X-100 (Sigma), 100mM NaCl (Honeywell), 10mM Hepes pH 7.5 (Sigma), 10mM EGTA (Sigma), 10mM EDTA (Sigma), and 1x protease inhibitor (Roche), and incubated for 30 min on ice. Lysate was pre-cleared by centrifugation at 16,000 x g for 5 min at 4°C and the supernatant was collected for immunoprecipitation.

Unless otherwise stated, all incubations and washes were performed with end-over-end rotation at room temperature. All centrifugation to collect beads was performed 2,500 x g for 2 min. Primary antibodies (5μg) or isotype controls (5μL) were incubated with 40μL protein A (Invitrogen) or protein G (Pierce) bead slurry for 30 min to couple antibodies to beads. After coupling, beads were collected by centrifugation and the supernatant discarded. Beads were washed twice (1mL 0.2M sodium borate, pH 9 for 1 min). Beads were resuspended in 400μL 20mM dimethyl pimelimidate (DMP, Thermo) and incubated for exactly 30 min to crosslink antibodies. After crosslinking, beads were washed twice (1mL 0.2M ethanolamine (EMD Millipore) pH 8 for 5 min) and incubated in 1.4mL 0.2M ethanolamine for two hr. Beads were then washed twice (1mL PBS for 5 min).

50μL crosslinked cell lysate and 50μL Triton lysis buffer were added to antibody-conjugated beads and incubated overnight at 4°C with end-over-end rotation. After incubation, beads were collected by centrifugation and washed three times (500μL PBS, 0.1% Triton X-100 for 2 min). To elute proteins, beads were resuspended in elution buffer (20μL 5x Laemmli samle buffer and 30μL PBS) and incubated at 90°C for 10 min. Following elution, the samples were centrifuged to collect the beads and the supernatant removed for analysis by western blot as described above. For SDS-PAGE, 7.5μL protein lysate (5%) and 25μL (50%) immunoprecipitated sample were loaded.

Sequence alignment

Sequence alignment was performed on the EMBL-EBI MUSCLE online interface[54]. Alignments used the ClustalW algorithm using the default parameters.

Structure prediction

Structures of *Homo sapiens* PEG10 gag (AA 1-325) and RTL8 (AA 1-113) were modeled using the Phyre 2.0 web server[55] using the intensive modeling mode. Structures were visualized using UCSF Chimera[56].

Flow cytometry

HEK293 cells were transfected in a 96-well plate with plasmids encoding a PEG10-Dendra2 fusion protein followed by an IRES-CFP cassette under control of a CMV promoter. Cells were lifted by pipetting in FACS buffer (D-PBS, 2% FBS, 0.1% sodium azide) and analyzed on a FACSCelesta (BD Biosciences). 20,000 events were collected per sample on the cytometer. Flow cytometry analysis was performed using Flowjo software (Treestar). Single cells were first gated on FSC-A vs SSC-A, followed by gating of CFP+ cells. PEG10-Dendra2 signal was normalized to an IRES-CFP expressed from the same plasmid to account for transfection differences by quantifying the geometric mean of Dendra2 signal divided by the geometric mean of CFP signal for each sample.

Statistics

Error bars represent mean ± SEM for all Figures. All statistical analysis was performed using GraphPad Prism software. Standard one-way ANOVA was corrected for multiple comparisons by Dunnett's test as recommended. Standard two-way ANOVA was corrected for multiple comparisons by Šídák's test as recommended. For all figures, statistical tests are listed in the figure legend and *$p<0.05$, **$p<0.01$, ***$p<0.001$, and ****$p<0.0001$.

Author contributions:

WC, HB, and AMW conceived of the project and designed the experiments. WC, HB, and CL executed the experiments. WC analyzed the data, in discussion with AMW. WC and AMW wrote the manuscript with input from all authors.

References

1. Goodier, J. L. Restricting retrotransposons: a review. *Mob. DNA* **7**, 16 (2016).
2. Brown, K. D. & Robertson, K. D. DNMT1 knockout delivers a strong blow to genome stability and cell viability. *Nat. Genet.* **39**, 289–290 (2007).
3. Zamudio, N. *et al.* DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes Dev.* **29**, 1256–1270 (2015).
4. Bourque, G. *et al.* Ten things you should know about transposable elements. *Genome Biol.* **19**, 199 (2018).
5. Senft, A. D. & Macfarlan, T. S. Transposable elements shape the evolution of mammalian development. *Nat. Rev. Genet.* **22**, 691–711 (2021).
6. Lee, S.-H. *et al.* The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc. Natl. Acad. Sci.* **102**, 18075–18080 (2005).
7. Nakamura, T. M. & Cech, T. R. Reversing Time: Origin of Telomerase. *Cell* **92**, 587–590 (1998).
8. Murcia, P. R., Arnaud, F. & Palmarini, M. The Transdominant Endogenous Retrovirus enJS56A1 Associates with and Blocks Intracellular Trafficking of Jaagsiekte Sheep Retrovirus Gag. *J. Virol.* **81**, 1762–1772 (2007).
9. Frank, J. A. *et al.* Evolution and antiviral activity of a human protein of retroviral origin. *Science* **378**, 422–428 (2022).
10. Srinivasachar Badarinarayan, S. & Sauter, D. Switching Sides: How Endogenous Retroviruses Protect Us from Viral Infections. *J. Virol.* **95**, e02299-20.
11. Antony, J. M. *et al.* The Human Endogenous Retrovirus Envelope Glycoprotein, Syncytin-1, Regulates Neuroinflammation and Its Receptor Expression in Multiple Sclerosis: A Role for Endoplasmic Reticulum Chaperones in Astrocytes. *J. Immunol.* **179**, 1210–1224 (2007).
12. Sankowski, R. *et al.* Endogenous retroviruses are associated with hippocampus-based memory impairment. *Proc. Natl. Acad. Sci.* **116**, 25982–25990 (2019).
13. Shiura, H. *et al.* PEG10 viral aspartic protease domain is essential for the maintenance of fetal capillary structure in the mouse placenta. *Development* **148**, dev199564 (2021).

14. Ono, R. *et al.* Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat. Genet.* **38**, 101–106 (2006).
15. Akamatsu, S. *et al.* The Placental Gene PEG10 Promotes Progression of Neuroendocrine Prostate Cancer. *Cell Rep.* **12**, 922–936 (2015).
16. Deng, X. *et al.* PEG10 plays a crucial role in human lung cancer proliferation, progression, prognosis and metastasis. *Oncol. Rep.* **32**, 2159–2167 (2014).
17. Li, X. *et al.* PEG10 promotes human breast cancer cell proliferation, migration and invasion. *Int. J. Oncol.* **48**, 1933–1942 (2016).
18. Whiteley, A. M. *et al.* Global proteomics of Ubqln2-based murine models of ALS. *J. Biol. Chem.* **296**, 100153 (2021).
19. Pandya, N. J. *et al.* Secreted retrovirus-like GAG-domain-containing protein PEG10 is regulated by UBE3A and is involved in Angelman syndrome pathophysiology. *Cell Rep. Med.* **2**, 100360 (2021).
20. Abed, M. *et al.* The Gag protein PEG10 binds to RNA and regulates trophoblast stem cell lineage specification. *PLOS ONE* **14**, e0214110 (2019).
21. Segel, M. *et al.* Mammalian retrovirus-like protein PEG10 packages its own mRNA and can be pseudotyped for mRNA delivery. *Science* **373**, 882–889 (2021).
22. Black, H. H. *et al. UBQLN2 restrains the domesticated retrotransposon PEG10 to maintain neuronal health in ALS*. http://biorxiv.org/lookup/doi/10.1101/2022.03.25.485837 (2022) doi:10.1101/2022.03.25.485837.
23. Ashley, J. *et al.* Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell* **172**, 262-274.e11 (2018).
24. Pastuzyn, E. D. *et al.* The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell* **172**, 275-288.e18 (2018).
25. Brandt, J., Veith, A. M. & Volff, J.-N. A family of neofunctionalized Ty3/gypsy retrotransposon genes in mammalian genomes. *Cytogenet. Genome Res.* **110**, 307–317 (2005).
26. Havecker, E. R., Gao, X. & Voytas, D. F. The diversity of LTR retrotransposons. *Genome Biol.* (2004).

27. Cordaux, R. & Batzer, M. A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **10**, 691–703 (2009).
28. Lux, A. *et al.* Human Retroviral *gag* - and *gag-pol* -like Proteins Interact with the Transforming Growth Factor-β Receptor Activin Receptor-like Kinase 1. *J. Biol. Chem.* **280**, 8482–8493 (2005).
29. Clark, M. B. *et al.* Mammalian Gene *PEG10* Expresses Two Reading Frames by High Efficiency –1 Frameshifting in Embryonic-associated Tissues. *J. Biol. Chem.* **282**, 37359–37369 (2007).
30. Steplewski, A. *et al.* MyEF-3, a Developmentally Controlled Brain-Derived Nuclear Protein Which Specifically Interacts with Myelin Basic Protein Proximal Regulatory Sequences. *Biochem. Biophys. Res. Commun.* **243**, 295–301 (1998).
31. Mohan, H. M. *et al.* RTL8 promotes nuclear localization of UBQLN2 to subnuclear compartments associated with protein quality control. *Cell. Mol. Life Sci.* **79**, 176 (2022).
32. Itakura, E. *et al.* Ubiquilins Chaperone and Triage Mitochondrial Membrane Proteins for Degradation. *Mol. Cell* **63**, 21–33 (2016).
33. *Structure and Physics of Viruses: An Integrated Textbook*. vol. 68 (Springer Netherlands, 2013).
34. Dodonova, S. O., Prinz, S., Bilanchone, V., Sandmeyer, S. & Briggs, J. A. G. Structure of the Ty3/Gypsy retrotransposon capsid and the evolution of retroviruses. *Proc. Natl. Acad. Sci.* **116**, 10048–10057 (2019).
35. Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409–413 (2000).
36. Zurowska, K., Alam, A., Ganser-Pornillos, B. K. & Pornillos, O. Structural evidence that MOAP1 and PEG10 are derived from retrovirus/retrotransposon Gag proteins. *Proteins Struct. Funct. Bioinforma.* **90**, 309–313 (2022).
37. Xie, T. *et al.* PEG10 as an oncogene: expression regulatory mechanisms and role in tumor progression. *Cancer Cell Int.* **18**, 112 (2018).
38. Zhang, M., Larsen, L. S., Irwin, B., Bilanchone, V. & Sandmeyer, S. Two-hybrid analysis of Ty3 capsid subdomain interactions. *Mob. DNA* **1**, (2010).

39. Lanman, J., Sexton, J., Sakalian, M. & Prevelige, P. E. Kinetic Analysis of the Role of Intersubunit Interactions in Human Immunodeficiency Virus Type 1 Capsid Protein Assembly In Vitro. *J. Virol.* **76**, 6900–6908 (2002).
40. Lanman, J. *et al.* Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat. Struct. Mol. Biol.* **11**, 676–677 (2004).
41. Yufenyuy, E. L. & Aiken, C. The NTD-CTD intersubunit interface plays a critical role in assembly and stabilization of the HIV-1 capsid. *Retrovirology* **10**, 29 (2013).
42. Constância, M. *et al.* Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* **417**, 945–948 (2002).
43. Moore, T. & Haig, D. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet.* **7**, 45–49 (1991).
44. Finn, E. H., Smith, C. L., Rodriguez, J., Sidow, A. & Baker, J. C. Maternal bias and escape from X chromosome imprinting in the midgestation mouse placenta. *Dev. Biol.* **390**, 80–92 (2014).
45. Armezzani, A., Varela, M., Spencer, T., Palmarini, M. & Arnaud, F. “Ménage à Trois”: The Evolutionary Interplay between JSRV, enJSRVs and Domestic Sheep. *Viruses* **6**, 4926–4945 (2014).
46. Passerini, L. D., Keckesova, Z. & Towers, G. J. Retroviral Restriction Factors Fv1 and TRIM5a Act Independently and Can Compete for Incoming Virus before Reverse Transcription. *J VIROL* **80**, 2100–2105 (2006).
47. Cottee, M. A. *et al.* Structure of a Ty1 restriction factor reveals the molecular basis of transposition copy number control. *Nat. Commun.* **12**, 5590 (2021).
48. Campillos, M., Doerks, T., Shah, P. & Bork, P. Computational characterization of multiple Gag-like human proteins. *Trends Genet.* **22**, 585–589 (2006).
49. Shahheydari, H. *et al.* Protein Quality Control and the Amyotrophic Lateral Sclerosis/Frontotemporal Dementia Continuum. *Front. Mol. Neurosci.* **10**, 119 (2017).
50. Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nat. Med.* **10**, S10–S17 (2004).

51. Ross, C. A. & Poirier, M. A. What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol.* **6**, 891–898 (2005).
52. Deng, H.-X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).
53. Fernandopulle, M. S. *et al.* Transcription Factor-Mediated Differentiation of Human iPSCs into Neurons: Rapid differentiation of iPSCs into neurons. *Curr. Protoc. Cell Biol.* **79**, e51 (2018).
54. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
55. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
56. Pettersen, E. F. *et al.* UCSF Chimera--A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

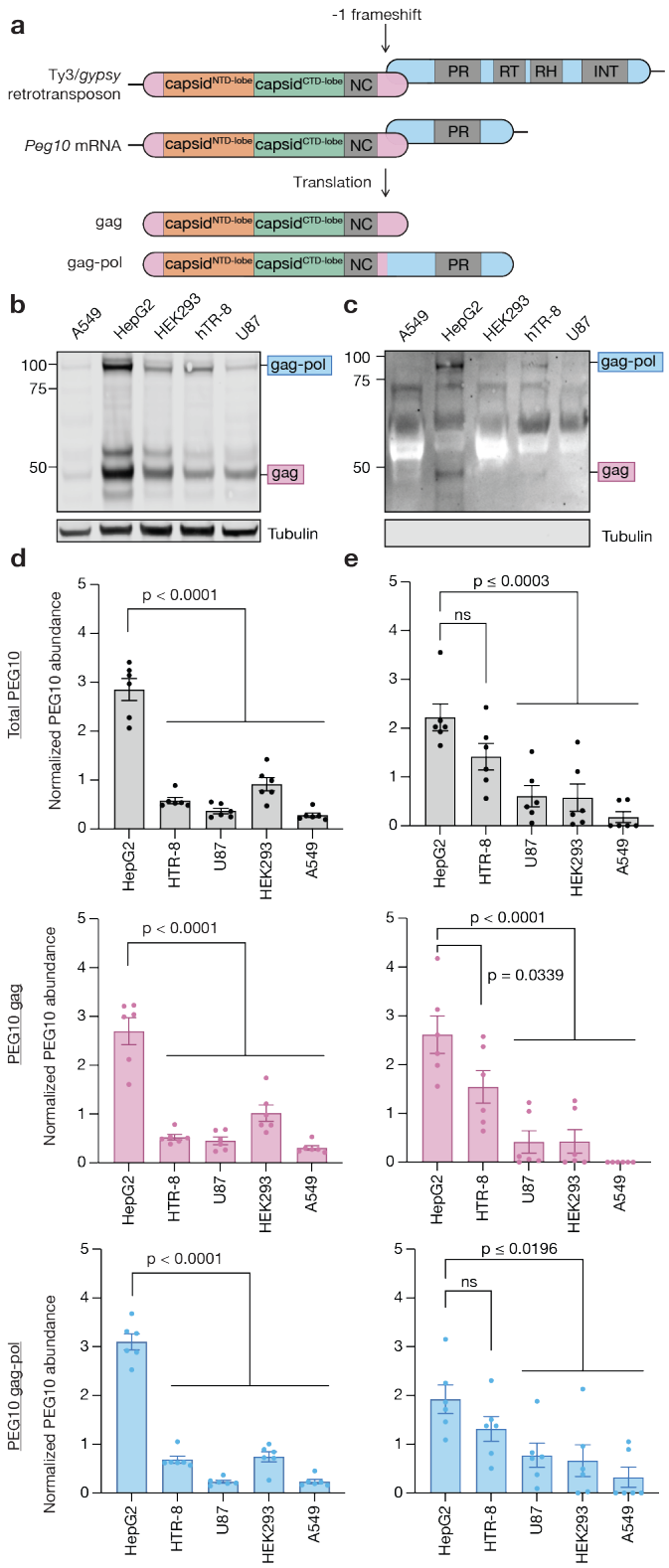


Figure 1: PEG10-derived VLPs are released by human cells. a) Schematic of the domesticated retrotransposon *PEG10*, aligned to its ancestral Ty3 retrotransposon. Upon use of a programmed -1 ribosomal frameshift, *PEG10* translates two different protein products, gag and gag-pol polyprotein. PEG10 gag contains the structural capsid forming two distinct lobes, capsid$^{NTD\text{-}lobe}$ and capsid$^{CTD\text{-}lobe}$, and nucleic acid-binding nucleocapsid (NC). PEG10 pol contains a retrovirus-like aspartic protease (PR).

b) Representative western blot showing PEG10 abundance in lysate from A549, HepG2, HEK293, hTR-8, and U87 cell lines. n = 6 representative experiments. **c)** Representative western blot showing PEG10 VLP production across cell lines. n = 6 representative experiments. **d)** Quantification of PEG10 abundance from Fig. 1b. (top) Total PEG10 was calculated as the sum of gag-pol and gag signal normalized to tubulin. gag (middle) and gag-pol (bottom) were analyzed as described in methods. n = 6 representative experiments. Data were analyzed by ordinary one-way ANOVA with column means compared to the highest abundance cell line, HepG2. **e)** Quantification of PEG10 VLP signal in Fig. 1c. Data are analyzed as in 1d.

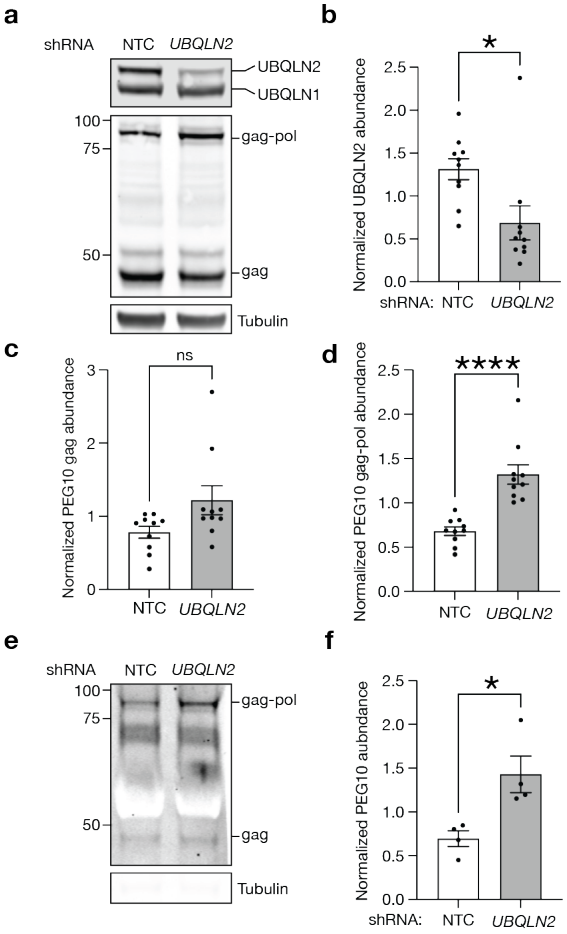


Figure 2: UBQLN2 regulates PEG10 gag-pol to restrict VLP abundance. a) Representative western blot showing PEG10 and UBQLN1/2 abundance in hTR-8 cells stably expressing non-targeting control (NTC) or *UBQLN2* knockdown (*UBQLN2*) shRNAs. **b)** Quantification of UBQLN2 abundance from Fig. 2a. **c)** Quantification of PEG10 gag abundance from Fig. 2a. **d)** Quantification of PEG10 gag-pol abundance from Fig. 2a. **e)** Representative western blot showing increased PEG10 VLP signal in hTR-8 sh*UBQLN2* cells compared to non-targeting control. n = 5 representative experiments. **f)** Quantification and analysis of PEG10 abundance from Figure 2e. PEG10 VLP signal was measured as the sum of gag-pol and gag band intensities normalized as described in methods. For b-d and f, data were analyzed by Student's t-test (*$p<0.05$, **** $p<0.0001$).

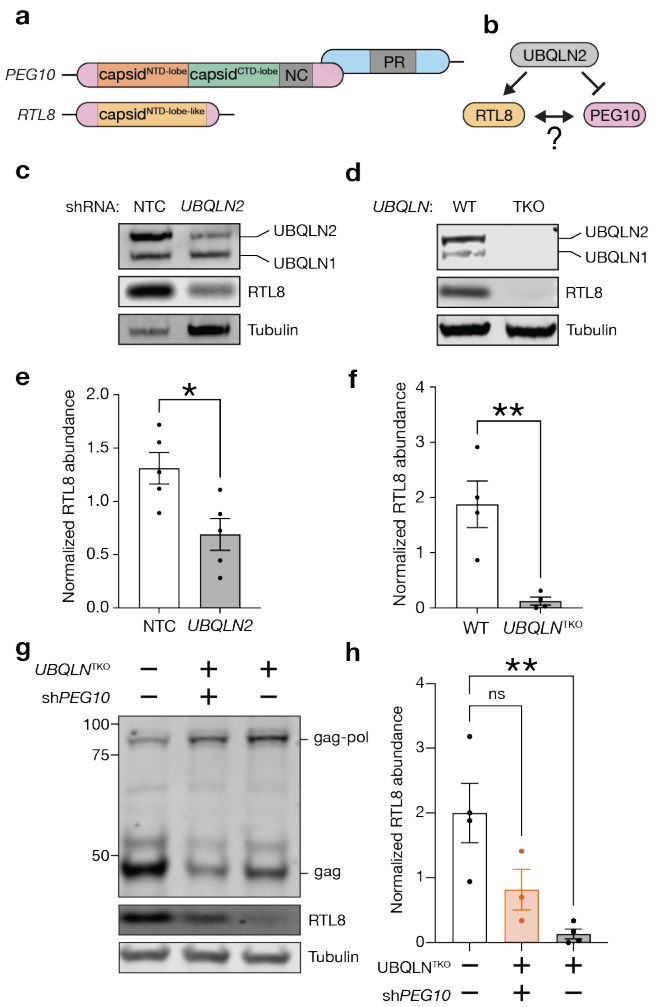


Figure 3: UBQLN2 regulates RTL8 levels in a PEG10-dependent manner. a) Cartoon schematic of PEG10 and RTL8 protein sequences. RTL8 shares high homology to PEG10 capsid^{NTD}. **b)** Cartoon highlighting the relationship between UBQLN2, PEG10, and RTL8. Expression of *UBQLN2* decreases PEG10 abundance and increases RTL8 abundance despite no changes in transcription or translation. The relationship between PEG10 and RTL8 is unknown. **c)** Representative western blot measuring changes to RTL8 abundance in upon knockdown of *UBQLN2* in hTR-8 cells. **d)** Representative western blot measuring changes to RTL8 abundance in upon genetic loss of *UBQLN1, 2*, and *4* in HEK293^{TKO} cells. **e)** Quantification of data in Fig. 3c. Data are analyzed by Student's t-test (*p<0.05, ** p<0.01). **f)** Quantification and analysis of data in Fig. 3d. Data are analyzed as in Fig. 3e. **g)** Representative western blot showing RTL8 abundance upon loss of *UBQLN2* and rescue upon subsequent knockdown of *PEG10*. **h)** Quantification of data in Figure 3g. Data were analyzed by ordinary one-way ANOVA with column means compared to HEK293^{WT} cells.

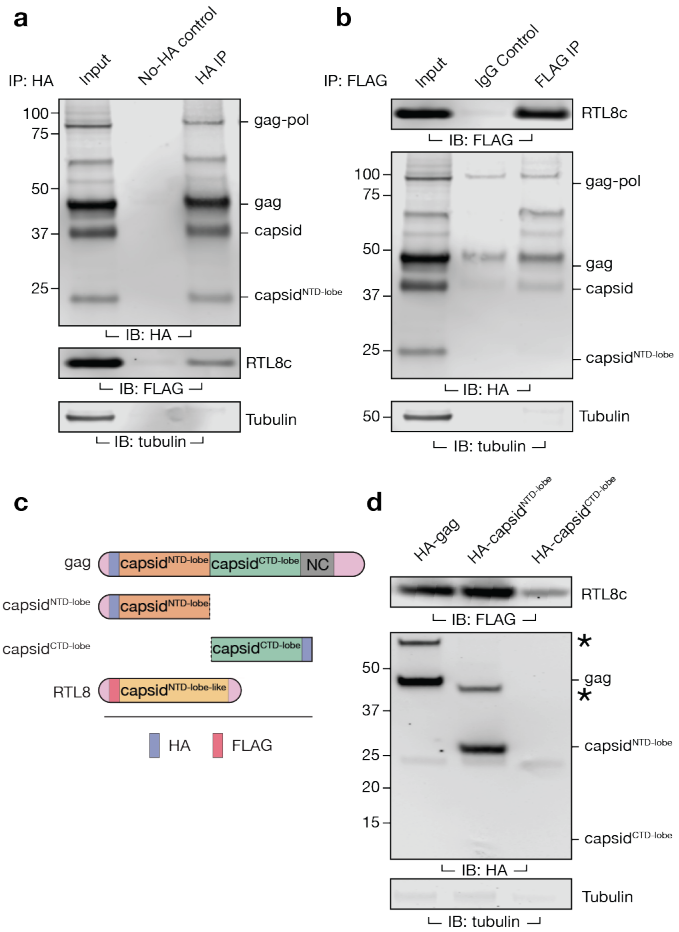


Figure 4: RTL8 interacts with the capsid$^{NTD\text{-}lobe}$ of PEG10. a) Crosslinking co-immunoprecipitation of RTL8c with PEG10. HA-tagged PEG10 and FLAG-tagged RTL8c were co-expressed, proteins crosslinked, cells lysed, and complexes immunoprecipitated using an HA antibody. No-HA control cells were transfected only with FLAG-RTL8c before HA-immunoprecipitation. **b)** Crosslinking co-immunoprecipitation of PEG10 with RTL8c. FLAG-tagged RTL8c and HA-tagged PEG10 were co-expressed and prepared as in Fig. 4a. Complexes were immunoprecipitated using a FLAG antibody or an IgG isotype control antibody. **c)** Diagram of PEG10 constructs including HA-capsid$^{NTD\text{-}lobe}$ and HA-capsid$^{CTD\text{-}lobe}$ truncations tested for interaction with FLAG-RTL8c. N-terminal HA tag is shown in blue and N-terminal FLAG tag is shown in red. **d)** Representative immunoprecipitations of FLAG-RTL8c showing co-immunoprecipitation of HA-gag and HA-capsid$^{NTD\text{-}lobe}$. Asterisk indicates an additional high molecular weight product positive for both HA and FLAG. Due to low relative expression of HA-capsid$^{CTD\text{-}lobe}$, an additional experimental replicate was performed with an overabundance of HA-capsid$^{CTD\text{-}lobe}$ lysate in Extended Data Fig. 5c.

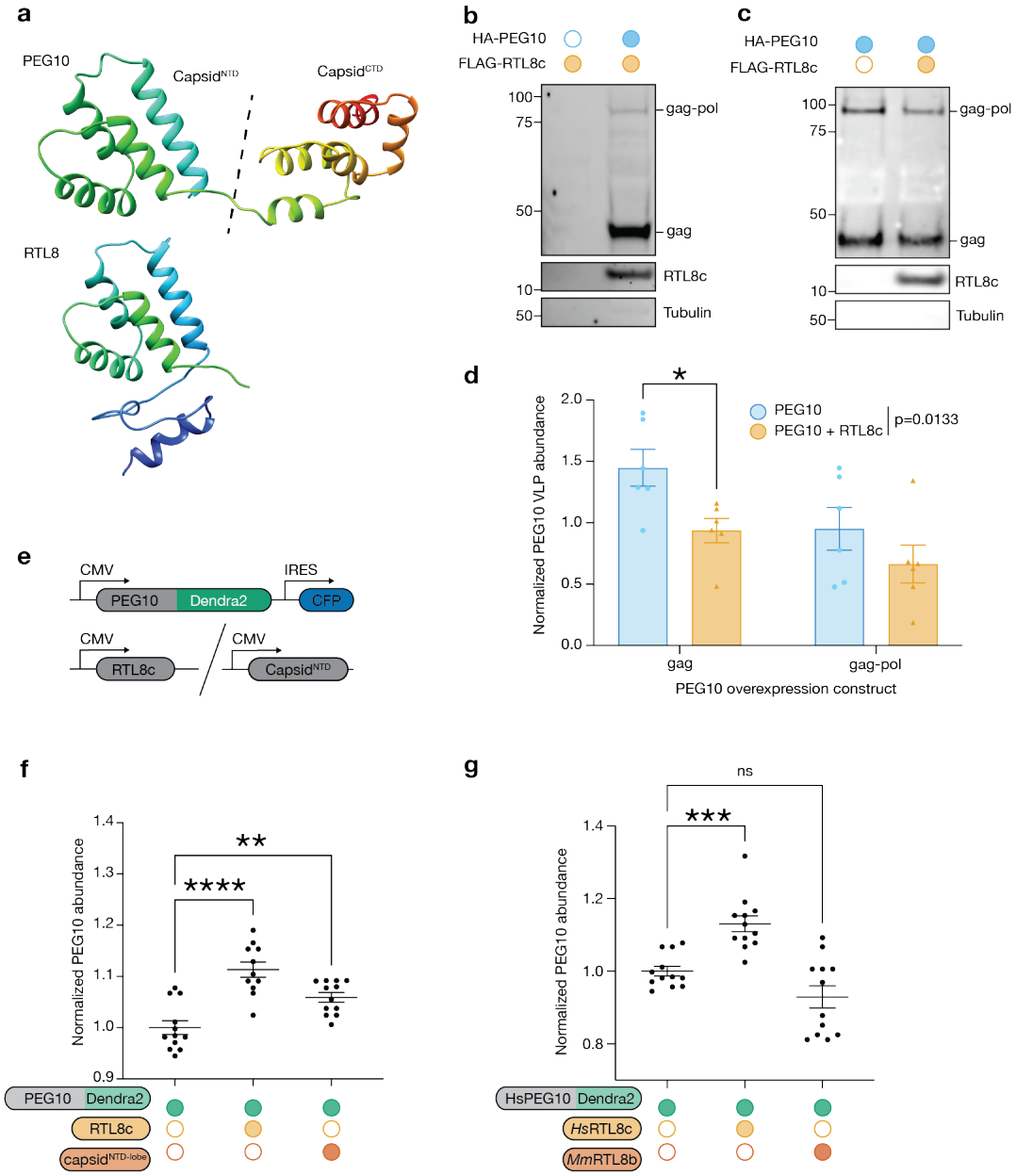


Figure 5: RTL8 incorporates into PEG10 VLPs and decreases the efficiency of release. a) (top) Predicted structure of PEG10 capsid shows the clear separation of NTD and CTD lobes. (bottom) The predicted structure of RTL8c, which bears striking similarity to capsid$^{NTD\text{-}lobe}$ of PEG10. **b)** Representative western blot measuring FLAG-RTL8c presence in the VLP fraction. FLAG-RTL8c was co-expressed without or with HA-PEG10. FLAG-RTL8c release in the VLP fraction is dependent upon co-expression of HA-PEG10. **c)** Representative western blot showing HA-PEG10 VLP signal without and with FLAG-RTL8c co-expression. **d)** Quantification of HA-PEG10 VLP signal with FLAG-RTL8c co-expression. For PEG10 gag-pol overexpression, gag and gag-pol signal was summed to calculate total PEG10 VLP signal. Data are analyzed by ordinary two-way ANOVA considering the effect of FLAG-RTL8c co-expression and HA-PEG10 overexpression construct. **e)** Design schematic of a flow assay assessing intracellular PEG10 signal. PEG10 is expressed as a fusion protein with the fluorescent protein Dendra2 at its C-terminus, followed by an IRES-CFP for transfection efficiency control. **f)** Flow cytometry

1 measurement of intracellular PEG10-Dendra2 abundance. PEG10-Dendra2 was co-expressed with
2 either a control vector (pcDNA3.1), FLAG-RTL8c, or HA-capsid$^{NTD\text{-}lobe}$, and PEG10-Dendra2 was
3 measured to determine its intracellular abundance. Data were analyzed by ordinary one-way ANOVA.
4 **g)** PEG10-Dendra2 was co-expressed with control, FLAG-*Hs*RTL8c, or FLAG-*Mm*RTL8b and
5 intracellular abundance was measured by flow cytometry. Data were analyzed as in Fig. 5f.

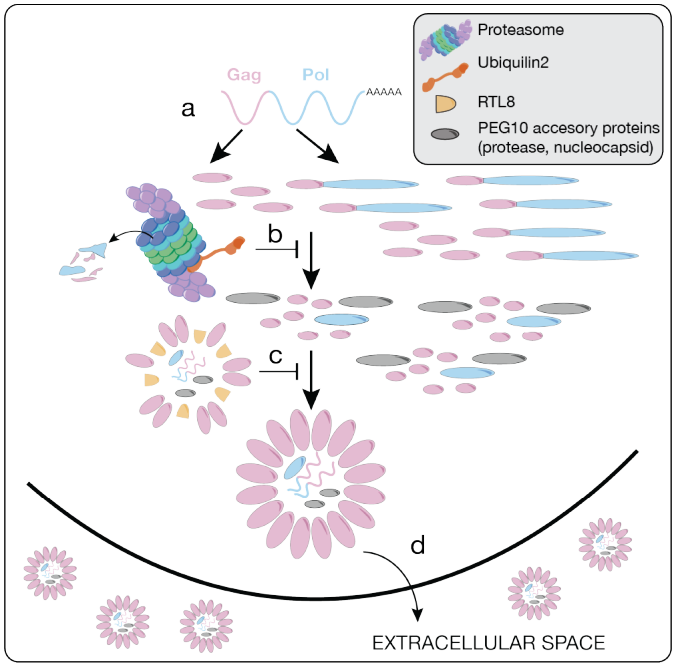
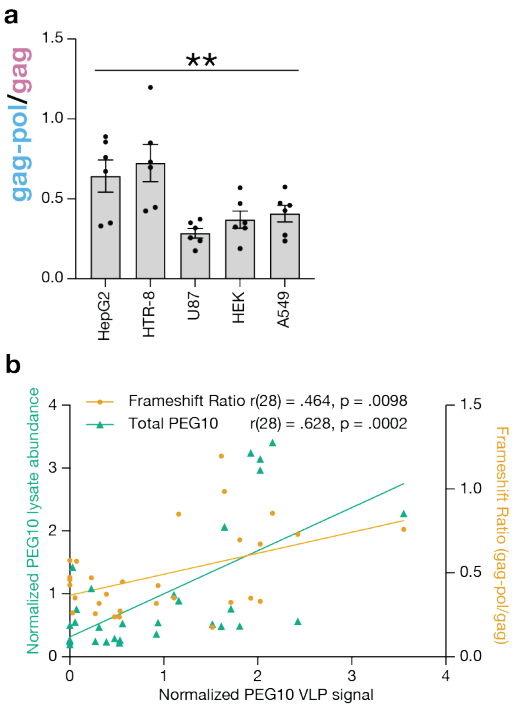
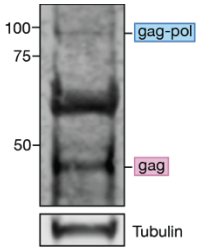


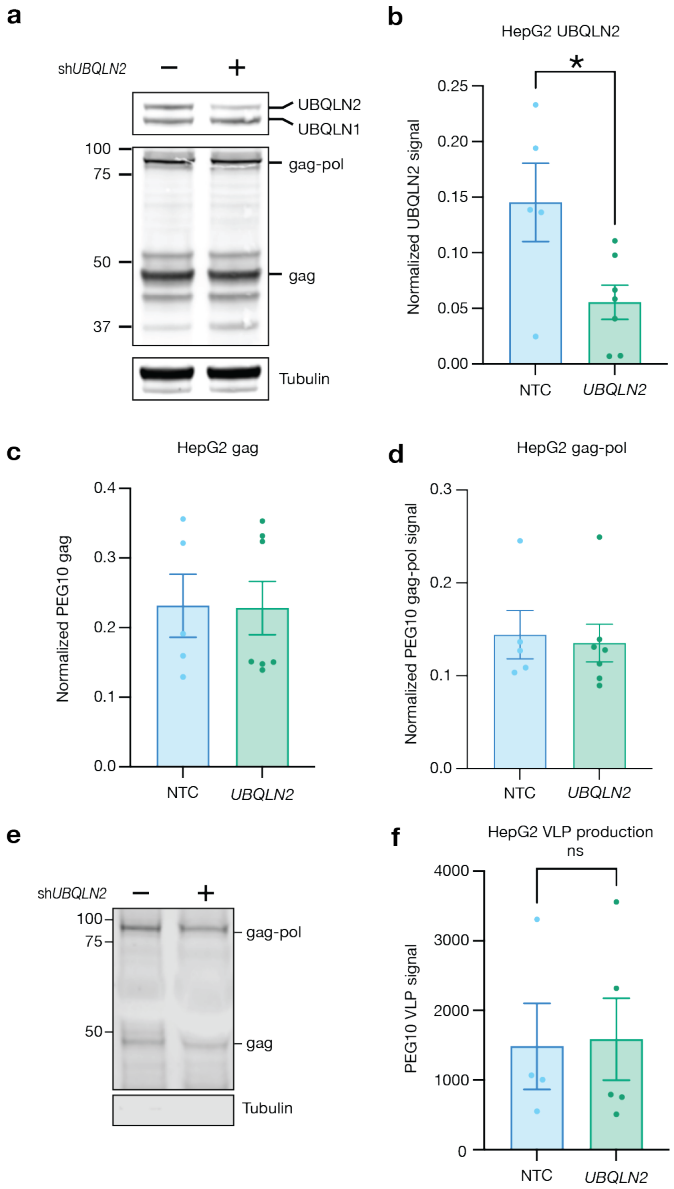
Figure 6: Model of UBQLN2 and RTL8 antagonism of PEG10 VLP abundance. a) *PEG10* mRNA is translated to produce gag and gag-pol polyprotein. **b)** PEG10 gag-pol polyprotein is recognized by UBQLN2 to mediate proteasomal degradation, decreasing VLP release. **c)** RTL8 competitively inhibits formation or release of PEG10 VLPs. **d)** Despite overlapping restriction mechanisms, PEG10 VLPs are exported from some human cell lines. More work is required to understand the delicate balance between restriction release, as well as downstream effects of PEG10 VLPs.



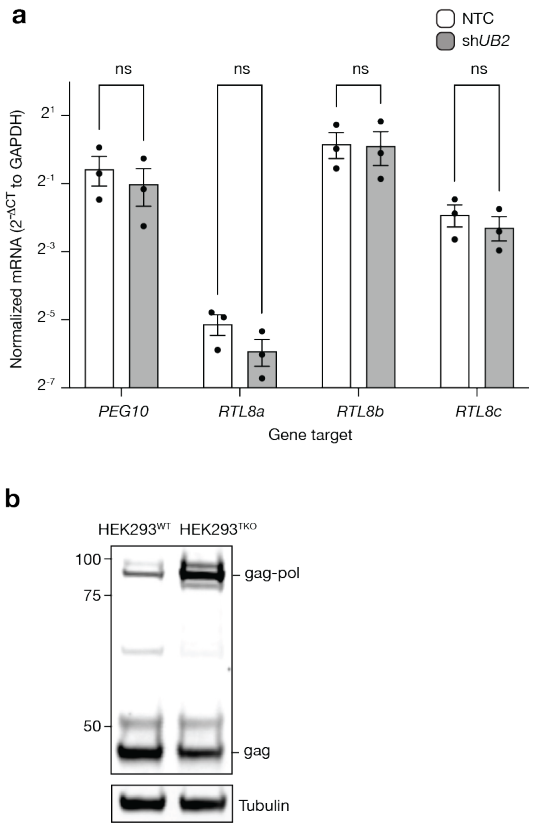
Extended Data Fig. 1: Putative descriptors of VLP release in human cell lines. a) Quantification of PEG10 frameshift ratio from Fig. 1b. Frameshift ratio is calculated as gag-pol signal over gag signal. Data are analyzed by ordinary one-way ANOVA. p-values are corrected for multiple comparisons by Dunnett's test. **b)** Pearson correlation between total PEG10 (left Y-axis) and frameshift ratio (right Y-axis) and PEG10 VLP signal.



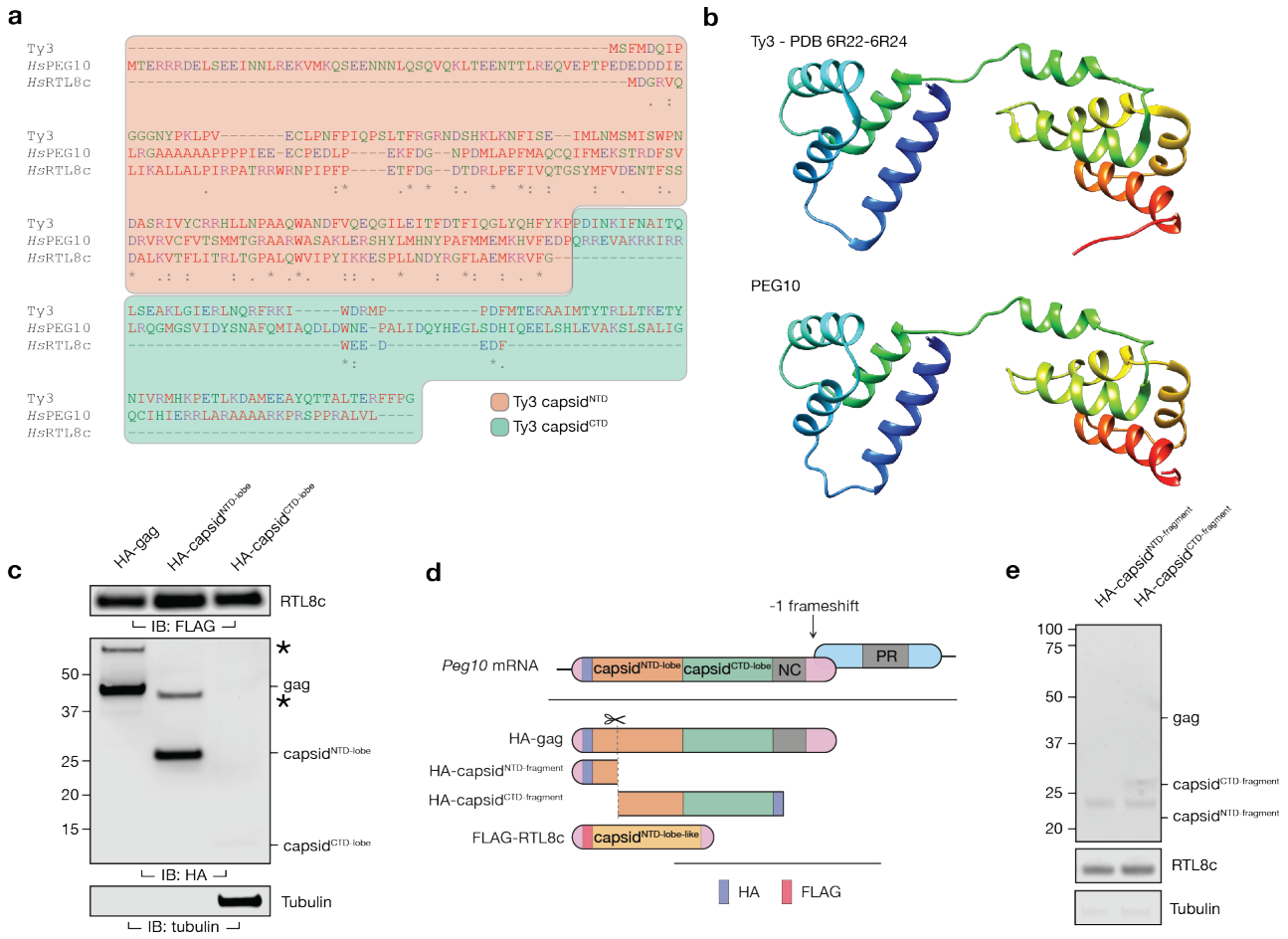
Extended Data Fig. 2: PEG10 VLPs are produced by induced human neurons. Neurons were differentiated and cultured as in Fernandopulle et al. 2018[53]. ~20mL cultured media was ultracentrifuged to generate the VLP fraction. Tubulin is observed in the VLP fraction, which may arise from Matrigel used during cell culture.



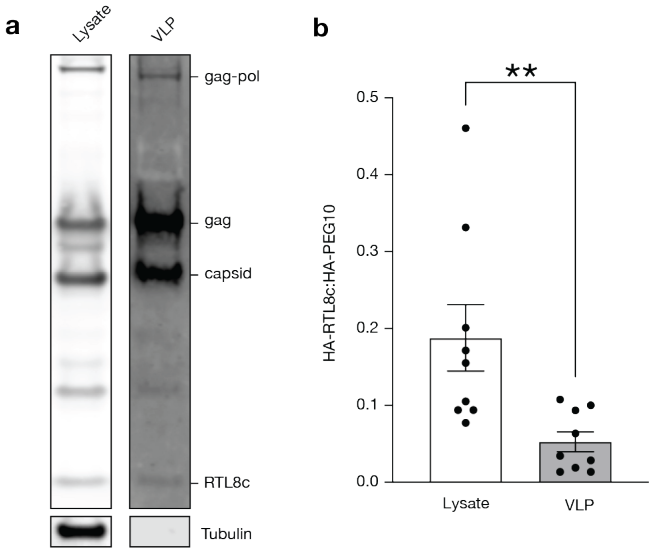
Extended Data Fig. 3: VLP abundance in HepG2 cells with loss of *UBQLN2* expression. a) Representative western blot showing PEG10 and UBQLN1/2 abundance in HepG2 cells stably expressing non-targeting control (NTC) or *UBQLN2* (*UBQLN2*) shRNA. **b)** Quantification of UBQLN2 abundance from Extended Data Fig. 3a. **c)** Quantification of PEG10 gag abundance from Extended Data Fig. 3a. **d)** Quantification and analysis of PEG10 gag-pol abundance from Extended Data Fig. 3a. **e)** Representative western blot showing no change to PEG10 VLP abundance in HepG2 sh*UBQLN2* cells compared to non-targeting control. n = 5 representative experiments. **f)** Quantification and analysis of PEG10 abundance from Extended Data Fig. 3e. PEG10 VLP abundance was calculated as in Fig. 2. All data were analyzed as in Fig. 2



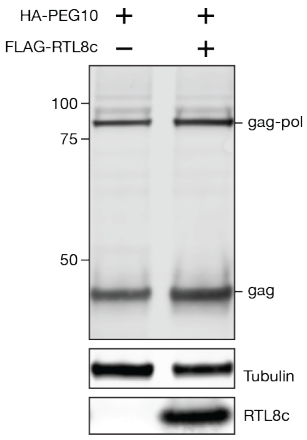
Extended Data Fig. 4: Attenuation of *UBQLN* expression does not influence transcript levels of *PEG10* or *RTL8*. a) qRT-PCR quantifying *PEG10*, *RTL8a*, *RTL8b*, and *RTL8c* expression in hTR-8 NTC and sh*UBQLN2* cells. n = 3 biological replicates with 3 technical replicates each. Data are analyzed by standard two-way ANOVA. **b)** Representative western blot showing PEG10 gag-pol accumulation in HEK293^{TKO} cells.



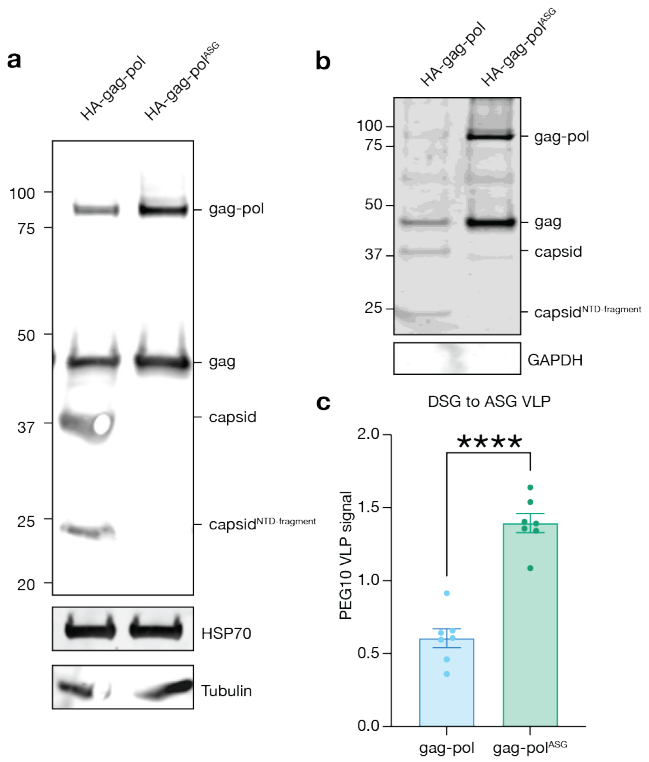
Extended Data Fig. 5: PEG10 closely resembles Ty3 and self-associates in a manner resembling the ancestral retrotransposon. a) Sequence alignment of the Ty3 and PEG10 capsid domains, and full length RTL8c. The demarcation between Ty3 capsid$^{NTD\text{-}lobe}$ and capsid$^{CTD\text{-}lobe}$ is indicated with shaded color background. **b)** (top) Structure of the Ty3 capsid[34]. (bottom) Predicted structure of PEG10 gag. **c)** CoIP of FLAG-RTL8c with HA-PEG10 gag, capsid$^{NTD\text{-}lobe}$, and capsid$^{CTD\text{-}lobe}$, as in Fig. 4d. Five times the total input protein was loaded in the capsid$^{CTD\text{-}lobe}$ immunoprecipitation to detect weak interactions with FLAG-RTL8c that may have been below the detection threshold in Fig. 4d. Input concentrations were high enough that nonspecific interactions with tubulin were detected, but capsid$^{CTD\text{-}lobe}$ remained below the limit of detection. **d)** Diagram of additional PEG10 constructs including HA-capsid$^{NTD\text{-}fragment}$ and HA-capsid$^{CTD\text{-}fragment}$ truncations to approximate the endogenous cleavage event within the capsid. **d)** Co-immunoprecipitation of FLAG-RTL8c with capsid cleavage products HA-capsid$^{NTD\text{-}fragment}$ and HA-capsid$^{CTD\text{-}fragment}$.



Extended Data Fig. 6: PEG10 VLPs released into media contain a low level of RTL8 incorporation. **a)** Representative western blots of lysate (left) and VLP (right) samples prepared from cells expressing HA-tagged PEG10 and HA-tagged RTL8c. HA antibody was used to detect proteins, which were identified as either PEG10 or RTL8 based on molecular weight. Data are representative of 5 independent experiments. **b)** Quantification of HA-RTL8c:HA-PEG10 ratio in lysate and VLP upon co-expression. RTL8 is less abundant in VLPs compared to lysate and is present in VLPs at less than 10% of PEG10 levels on average. Data were analyzed by Student's t-test. **$p<0.01$.

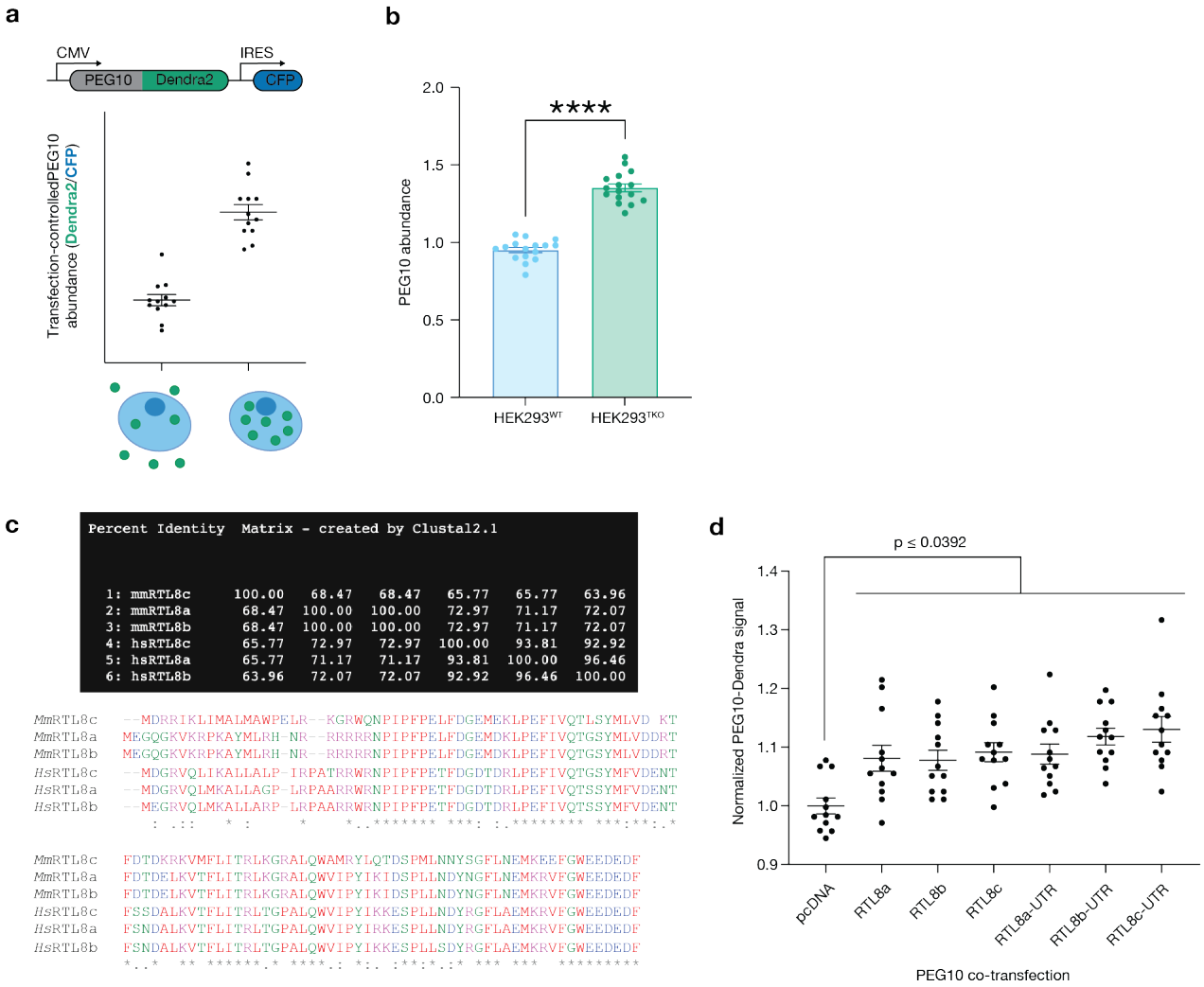


Extended Data Fig. 7: Co-expression of RTL8 with PEG10 does not measurably influence PEG10 levels in lysate by western blot. Representative western blot showing HA-PEG10 and FLAG-RTL8c levels upon co-expression.



1

2 **Extended Data Fig. 8: Loss of PEG10 protease activity increases VLP abundance. a)** Western blot
3 showing loss of HA-capsid and HA-capsid$^{NTD\text{-}fragment}$ proteins upon mutation of the PEG10 protease
4 active site aspartate to alanine. b) Representative western blot showing increased VLP production by
5 PEG10 HA-gag-pol^{ASG} relative to HA-gag-pol^{WT}. c) Quantification of VLP production by HA-gag-pol and
6 HA-gag-pol^{ASG}. Data were analyzed by Student's t-test. ****$p<0.0001$.



Extended Data Fig. 9: Co-expression of RTL8 causes intracellular retention of PEG10. a) Cartoon schematic of PEG10 abundance assay by flow cytometry (simulated data). **b)** Intracellular PEG10-Dendra2 abundance normalized to CFP expression in HEK293^{WT} and HEK293^{TKO} cells showing increased abundance upon loss of UBQLN2. Data were analyzed by Student's t-test. ****p<0.0001. **c)** (top) percent identity matrix and (bottom) protein sequence alignments between human and mouse RTL8a, b, c. **d)** Normalized PEG10-Dendra2 abundance upon mock (pcDNA empty) or FLAG-RTL8a, FLAG-RTL8b, or FLAG-RTL8c co-expression, comparing absence or presence of the 3'UTR. Data were analyzed by ordinary one-way ANOVA with column means compared to mock transfected.