

Evolutionary history of cotranscriptional editing in the paramyxoviral phosphoprotein gene

Jordan Douglas^{1,2,*}, Alexei J. Drummond^{1,2,3}, and Richard L. Kingston³

¹ Centre for Computational Evolution, University of Auckland, Auckland 1010, New Zealand

² School of Computer Science, University of Auckland, Auckland 1010, New Zealand

³ School of Biological Sciences, University of Auckland, Auckland 1010, New Zealand

* Correspondence: jordan.douglas@auckland.ac.nz

Abstract

The phosphoprotein gene of the paramyxoviruses encodes multiple protein products. The P, V, and W proteins are generated by transcriptional slippage. This process results in the insertion of non-templated guanosine nucleosides into the mRNA at a conserved edit site. The P protein is an essential component of the viral RNA polymerase, and is encoded by a direct copy of the gene in the majority of paramyxoviruses. However, in some cases the non-essential V protein is encoded by default and guanosines must be inserted into the mRNA in order to encode P. The number of guanosines inserted can be described by a probability distribution which varies between viruses. In this article we review the nature of these distributions, which can be inferred from mRNA sequencing data, and reconstruct the evolutionary history of cotranscriptional editing in the paramyxovirus family. Our model suggests that, throughout known history of the family, the system has switched from a P default to a V default mode four times; complete loss of the editing system has occurred twice, the canonical zinc finger domain of the V protein has been deleted or heavily mutated a further two times, and the W protein has independently evolved a novel function three times. Finally, we review the physical mechanisms of cotranscriptional editing via slippage of the viral RNA polymerase.

1 Introduction

The *Paramyxoviridae* are a family of nonsegmented, negative-sense, single-stranded RNA viruses, within the order *Mononegavirales* (Amarasinghe et al., 2019; Pringle, 1991; Rima et al., 2018). Type species include measles virus (MeV; genus: *Morbillivirus*), mumps virus (MuV; genus: *Orthorubulavirus*), Sendai virus (SeV; genus: *Respirovirus*), and Hendra virus (HeV; genus: *Henipavirus*). The *Paramyxoviridae* appear to infect most vertebrate species and are responsible for a number of serious diseases in both animals and humans.

The single-stranded RNA genome is bound to the viral nucleocapsid protein, forming a helical protein-nucleic acid complex which organises and protects the genome. The nucleocapsid acts as a template for both gene transcription and genome replication, processes that are executed by the viral RNA-dependent RNA-polymerase (RdRp; Whelan et al. (2004); Noton and Fearn (2015); Fearn and Plemper (2017); Guseva et al. (2019)). Each nucleocapsid protein binds six nucleotides of RNA (Gutsche et al., 2015; Alayyoubi et al., 2015; Jamin and Yabukarski, 2017; Webby et al., 2019), and paramyxoviral genomes always conform to the "rule of six" whereby genome length is some multiple of six (Kolakofsky et al., 1998; Calain and Roux, 1993; Kolakofsky et al., 2005). This is hypothesised to result from the requirement to position the promoter sequences required for initiation of RNA synthesis in the correct register, or phase, with respect to the nucleocapsid protein.

The phosphoprotein (P protein) is the non-catalytic subunit of the viral RdRP and mediates interaction between the polymerase and the nucleocapsid. The P protein is essential for translocation of the RdRP along its template, among other roles (Kingston et al., 2004; Du Pont et al., 2019; Bruhn et al., 2019; Milles et al., 2018; Guseva et al., 2019). The P gene, however, encodes multiple protein products through a combination of non-canonical transcription and translation events. This phenomenon, known as overprinting, is particularly common in viruses as it increases the utility of the genome, which is often under strong size constraints (Brandes and Linial, 2016; Sabath et al., 2012).

Cotranscriptional editing at the P gene governs production of the P, V, and W proteins. Production of multiple proteins results from insertion of one or more non-templated guanosine nucleosides (G) into the mRNA at a conserved edit site, through transcriptional slippage (Vidal et al., 1990b; Hausmann et al., 1999a). This process of base insertion stochastically shifts the reading frame. As a result, the P, V, and W proteins share a common N-terminal region (encoded by the gene sequence upstream of the edit site), but possess distinct C-terminal regions (encoded by the gene sequence downstream of the edit site), conferring differing functions on these three proteins. The mRNA encoding P, V, and W occur at different frequencies in different paramyxoviruses.

Although not the focus of this paper, operations at the translational level (Firth and Brierley, 2012), including leaky scanning (Giorgi et al., 1983; Shaffer et al., 2003), non-AUG initiation (Boeck et al., 1992; Curran and Kolakofsky, 1988), and ribosomal shunting (Latorre et al., 1998), facilitate production of yet more proteins from the P gene in many paramyxoviruses.

In this article we review cotranscriptional editing of the P gene in the *Paramyxoviridae*. We briefly discuss the assigned functions of the P, V, and W proteins, collate the available experimental information on the size distribution of guanosine nucleoside insertions, and propose a maximum parsimony model for evolution of the editing system. We also review what is known about transcriptional slippage and its connection with the genomic sequence at the edit site and displacement of nucleocapsid proteins from the genome.

2 Cotranscriptional editing in the *Paramyxoviridae*

Cotranscriptional editing at the P gene occurs through the insertion (or in certain mutants the deletion; Jacques et al. (1994)) of $m \in \{0, 1, 2, \dots\}$ guanosines, G_m , into the mRNA at a conserved edit site. A G_{3k+1} insertion ($m = 1, 4, 7, \dots$) shifts the reading frame downstream of the edit site by -1 (or alternatively +2). A G_{3k+2} nucleotide insertion ($m = 2, 5, 8, \dots$) shifts the reading frame by -2 (or alternatively +1). A G_{3k} insertion ($m = 0, 3, 6, \dots$) leaves the reading frame unaltered.

Proteins translated from a G_{3k+n} mRNA species are considered to be functionally equivalent across all values of $k \in \{0, 1, 2, \dots\}$. The mRNA flanking the edit site encodes an intrinsically disordered region of the protein (Habchi and Longhi, 2012; Guseva et al., 2019). Any extended sequence of G nucleotides is translated into polyglycine. While the conformational preferences of polyglycine are still not entirely established (Tran et al., 2008; Ohnishi et al., 2006), the homo-polymeric sequence will be disordered, and small variations in the length of this sequence are likely to be functionally neutral in this context.

The cotranscriptional editing system for the P gene operates in two different modes (**Figure 1**). In the $P \xrightarrow{G} V \xrightarrow{G} W$ mode, P is encoded by the unedited gene. This is the situation in MeV (Cattaneo et al., 1989) and SeV (Vidal et al., 1990a). In the $V \xrightarrow{G} W \xrightarrow{G} P$ mode, V is the default and a G_{3k+2} insertion is required to encode P. This is the situation in MuV (Paterson and Lamb, 1990).

We note that switching between these two edit modes requires a frameshift mutation in the genome, i.e. during genome replication. This mutation must occur at a position upstream of the edit site, but not so far upstream that it disrupts some other function of the encoded P protein. Moreover, due to the rule of six, any insertion or deletion (indel) must be rapidly compensated such that the genome length remains divisible by six. Otherwise the replication efficiency of the virus would be severely impacted (Calain and Roux, 1993; Kolakofsky et al., 2005; Skiadopoulos et al., 2003; Sauder et al., 2016). For example, a single nucleotide insertion upstream and proximal to the edit site, accompanied by a single nucleotide deletion elsewhere in the genome, would be sufficient to transit the system from $P \xrightarrow{G} V \xrightarrow{G} W$ to $V \xrightarrow{G} W \xrightarrow{G} P$.

2.1 P protein

The phosphoprotein has a range of functions. In complex with the viral large protein (L protein) it forms an integral part of RdRP. However, due to its greater relative abundance – a result of the “transcriptional gradient” that exists in all paramyxoviruses (Cattaneo et al., 1987; King et al., 2011) – it must also function independently of the L protein. The disordered N-terminal region of P is shared with V and W and contains the highly conserved soyuz1 and soyuz2 motifs (Karlin and Belshaw, 2012). These modules, together with internally located sequences, are involved in chaperoning the viral nucleocapsid protein monomers during replication (Yabukarski et al., 2014; Milles et al., 2018; Alayyoubi et al., 2015; Guryanov et al., 2016). This region also binds host proteins such as the STATs (Li et al., 2019; Devaux et al., 2007; Röthlisberger et al., 2010; Puri et al., 2009; Ciancanelli et al., 2009; Devaux et al., 2010; Chinnakannan et al., 2014), and its function is likely regulated by phosphorylation (Saikia et al., 2008; Sun et al., 2009; Sugai et al., 2012; Young et al., 2019; Pickar et al., 2014). The N-terminal region ranges from 109 aa (in APMV-3) to 570 aa (in GH-M74a) in length.

Manuscript has not been peer-reviewed

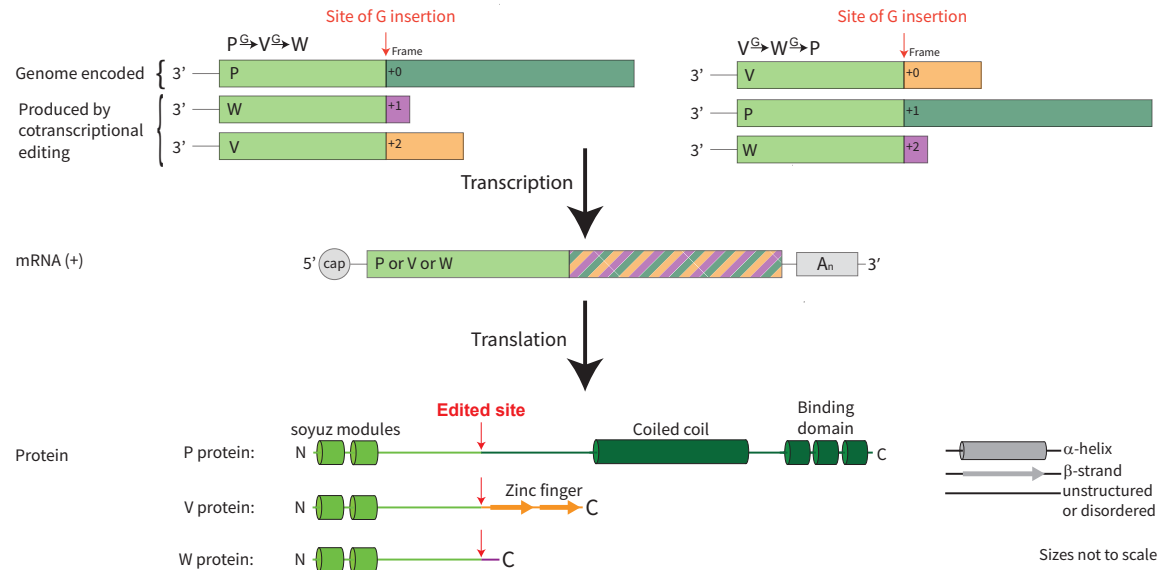


Figure 1: **Cotranscriptional editing of the P gene.** The two observed modes of editing are depicted: these are the $P \xrightarrow{G} V \xrightarrow{G} W$ and $V \xrightarrow{G} W \xrightarrow{G} P$ modes. A single transcript can encode one of P, V, or W depending on the number of guanosines stochastically inserted at the edit site during transcription. While the P, V, and W proteins share a common N-terminal region, their C-terminal regions are distinct.

Downstream of the edit site, the unique region of the phosphoprotein contains an oligomerisation domain (a coiled coil; Burmeister et al. (2000); Tarbouriech et al. (2000); Communie et al. (2013a); Cox et al. (2013); Bruhn et al. (2014)) and a nucleocapsid / large binding domain (the foot domain, or X domain; Johansson et al. (2003); Kingston et al. (2008); Yegambaram et al. (2013); Blanchard et al. (2004)) which are connected by a flexible linker. The C-terminal region of the phosphoprotein binds to both the large protein (Bruhn et al., 2019; Abdella et al., 2020) and the nucleocapsid (Kingston et al., 2004; Habchi et al., 2011; Communie et al., 2013b; Bloyet et al., 2016; Du Pont et al., 2019) and mediates their engagement. The phosphoprotein is therefore essential (Curran et al., 1991) and is encoded by all paramyxoviruses. P is the largest of the three P gene proteins. C-terminal regions range from 229 aa (in PIV-5) to 386 aa (in CPIV-3).

2.2 V protein

The V protein regulates viral genome replication (Horikami et al., 1996; Witko et al., 2006; Parks et al., 2006; Nishio et al., 2008; Sleeman et al., 2008; Yang et al., 2015) and interferes with the innate immune response to viral infection (see reviews: Fontana et al. (2008b); Ramachandran and Horvath (2009)). The latter is linked to increased viral virulence (Devaux et al., 2008; Alamares et al., 2010; Schaap-Nutt et al., 2010). V has been reported to bind a multitude of proteins involved in activation of the type I interferon response, sometimes in a genus specific fashion. These proteins include cytoplasmic pattern recognition receptors RIGI, MDA5 (Motz et al., 2013), and LGP2 (Ramachandran and Horvath, 2010; Mandhana et al., 2018; Sanchez-Aparicio et al., 2018; Andrejeva et al., 2004; Childs et al., 2007;

Manuscript has not been peer-reviewed

Rodriguez and Horvath, 2014), as well as TRIM25 (Sanchez-Aparicio et al., 2018), PP1 (Davis et al., 2014), MAVS (Sun et al., 2019), PLK1 (Ludlow et al., 2008), UBXN1 (Uchida et al., 2018), DDB1 (Li et al., 2006a; Salladini et al., 2017), various interferon regulatory factors (Takeuchi et al., 2003; Kitagawa et al., 2013; Palosaari et al., 2003; Lu et al., 2008), NF- κ B (Schuhmann et al., 2011), and the STAT proteins (Li et al., 2019; Didcock et al., 1999; Parisien et al., 2002). It has also been reported to interact with interferon stimulated gene products such as tetherin (Ohta et al., 2017, 2016) and other host proteins such as PKB/AKT1 (Sun et al., 2008).

The unique C-terminal region of V contains a highly conserved cysteine-rich zinc finger domain, which binds two zinc ions (Li et al., 2006a; Motz et al., 2013). The functions conferred by this unique C-terminal region are not always readily decoupled from the activities of the common N-terminal region. For instance, both of these regions bind STAT proteins (Li et al., 2019), and the structurally characterised interaction with DDB1 (Li et al., 2006a) is mediated by sequences from both the common N-terminal region and the unique C-terminal region of V.

V is the second largest of the P gene proteins: C-terminal regions range from 50 aa (in NiV) to 188 aa (in CPIV-3). While V aids viral replication, it is non-essential and is encoded by most but not all paramyxoviruses (Section 3.2).

2.3 W protein

A third protein may also be generated by contrascriptural editing. Unlike P and V, its unique C-terminal sequence is not conserved across, or even within, paramyxoviral genera and consequently this protein has been assigned many names (Fontana et al., 2008a) including W (Vidal et al., 1990a), D (Pelet et al., 1991; Galinski et al., 1992), PD (Wells and Malur, 2008), and I (Paterson and Lamb, 1990). For the purposes of this review, we use ‘W’ to denote the protein encoded by the reading frame that encodes neither P nor V. There is evidence that W has evolved a function within some paramyxoviral genera, though this is not always the case.

In Newcastle disease virus (NDV; genus: *Orthoavulavirus*), Hendra and Nipah virus (HeV and NiV; genus: *Henipavirus*), and Human parainfluenza virus 3 (HPIV-3; genus: *Respirovirus*), W accumulates in the nucleus (Karsunke et al., 2019; Yang et al., 2019; Shaw et al., 2005; Lo et al., 2009; Wells and Malur, 2008). Nuclear localisation signals can be identified in the unique region of the W protein (Shaw et al., 2005; Karsunke et al., 2019; Audsley et al., 2016a; Wells and Malur, 2008; Smith et al., 2018).

NDV sits alone, and we could not detect a homologous C-terminal region in the W protein of any other *Orthoavulaviruses*. A recent study showed that deleting the C-terminus of the W protein impaired NDV replication in cultured cells, and this effect was relieved when the full-length W protein was supplied in trans (Yang et al., 2019). However, no detailed function has been assigned to this protein.

The *Henipavirus* W protein influences the course of disease in animal models (Satterfield et al., 2015, 2016), and may play a direct role in subversion of the type I IFN response (Ciancanelli et al., 2009; Shaw et al., 2005). This protein also modulates host gene expression by interacting with the 14-3-3 family of regulatory proteins, an interaction that depends upon phosphorylation of the penultimate serine residue (Edwards et al. (2020); **Fig. 2**). However as with the V protein, it has proved difficult to experimentally decouple common functions conferred by the N-terminal region, with unique functions conferred by the C-terminal region.

Manuscript has not been peer-reviewed

For HPIV-3, in an early study, joint interruption of the V and W open reading frames attenuated viral replication (although individual interruptions had no effect; Durbin et al. (1999)). In interpreting this result, it should be noted that the V protein of HPIV-3 is abnormal, and likely to be expressed in truncated form (Section 3.2). A more recent study also suggests that the C-terminal region of the W protein promotes viral transcription and replication, and is potentially also involved in the downregulation of beta interferon expression (Roth et al., 2013). The C-terminal regions of HPIV-3, bovine parainfluenza virus 3 (BPIV-3; genus: *Respirovirus*), and caprine parainfluenza virus 3 (CPIV-3; genus: *Respirovirus*) W proteins have strong sequence similarity which is itself suggestive of shared function (**Figure 2**).

For remaining paramyxoviruses, the unique region of W may not necessarily possess any biological function at all, and is often very short (2 aa in SeV, 6 aa in MeV, 11 aa in MuV; Chinnakannan et al. (2014); Horikami et al. (1996); Curran et al. (1991); Paterson and Lamb (1990)). However the W protein could still potentially exert biological effects through its shared N-terminal region, with synthesis of W potentially being more rapid than the synthesis of either P or V.

Manuscript has not been peer-reviewed

W protein

Unique C-terminal region

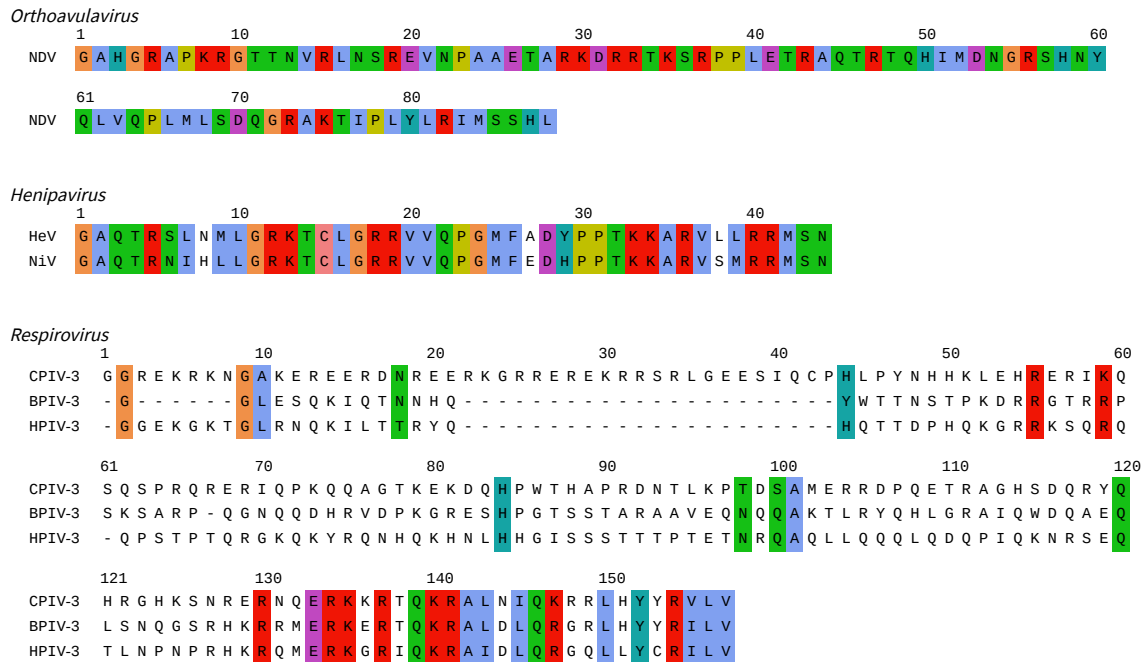


Figure 2: **W protein C-terminal regions.** For the displayed sequences there is experimental data regarding the cellular localisation or function of the W protein, or a W protein homolog in another virus. All numbering is relative to the start of the unique C-terminal region. Sites are coloured by amino acid characteristic if the characteristic is 100% conserved at the alignment position. Under the ClustalX colouring scheme hydrophobic residues are blue, positively charged residues – red, negatively charged residues – magenta, polar residues – green, cysteine – pink, glycine – orange, proline – yellow, and aromatic residues – cyan (Larkin et al., 2007).

Manuscript has not been peer-reviewed

3 Phylogeny of cotranscriptional editing

Let $p(G_m)$ be the empirical probability of the viral transcription machinery inserting m guanosines at the mRNA edit site. The most direct source of information on the nature of this probability distribution (i.e. the relative abundances of the transcripts) comes from sequencing the mRNA produced in virally infected cells. However, as Wignall-Fleming et al. (2019) have highlighted, if mRNA preparations are contaminated with anti-genomic RNA, the results may not faithfully reflect the actual abundance of mRNA, and the occurrence of the unedited transcript could be overestimated. Furthermore, several studies have noted that transcript abundance varies with time post-infection (Kulkarni et al., 2009; Qiu et al., 2016). In both cases the proportion of V and W transcripts increased as the infection progressed, though neither the mechanism nor functional implications are understood. Finally, it is noted that while mRNA abundances are often assumed to be related to encoded protein abundances, this may not hold in practice (Liu et al., 2016).

With these caveats noted, the experimentally derived probability distributions (or *edit patterns*) for 26 paramyxoviruses are displayed in **Figure 3**. The maximum observed insert size was G_{14} in NiV (Lo et al., 2009). Additional data on mRNA abundance, not displayed in the figure, can be found in the following publications – SeV: Pelet et al. (1991); Kato et al. (1997); NiV: Kulkarni et al. (2009); MeV: Liston and Briedis (1994); Millar et al. (2016); Donohue et al. (2019); NDV: Mebatsion et al. (2001); Yang et al. (2019); BeiV: Audsley et al. (2016b), TevPV: Johnson et al. (2019); Burroughs et al. (2015); HPIV-2: Ohgimoto et al. (1990); MuV: Takeuchi et al. (1990); CeMV: Bolt et al. (1995); PPRV: Mahapatra et al. (2003); PDV: Blixenkron-Möller et al. (1992); PorPV: Berg et al. (1992).

Fundamental differences that exist between these systems reflect evolutionary events which have occurred throughout the history of the family. The following events are minimally required to explain the data: i) gain of the editing system, ii) loss of the editing system, iii) evolution of the V protein zinc finger motif and gain of biological function, iv) loss of the V protein zinc finger and associated function, v) switching of the edit mode and adaptation of the edit pattern, and vi) acquisition of unique function by the W protein. We estimated the evolutionary history of the *Paramyxoviridae* and inferred the ancestral lineages where these events occurred as follows: for each event we imputed the occurrence of the event onto branches such that the number of events required to explain the states observed at the leaves in the tree is minimised (**Figure 4**). This is the maximum parsimony model. A limitation of this model is that it does not account for the full functional diversity of the V protein, which has multiple biological activities (Section 2.2).

Manuscript has not been peer-reviewed

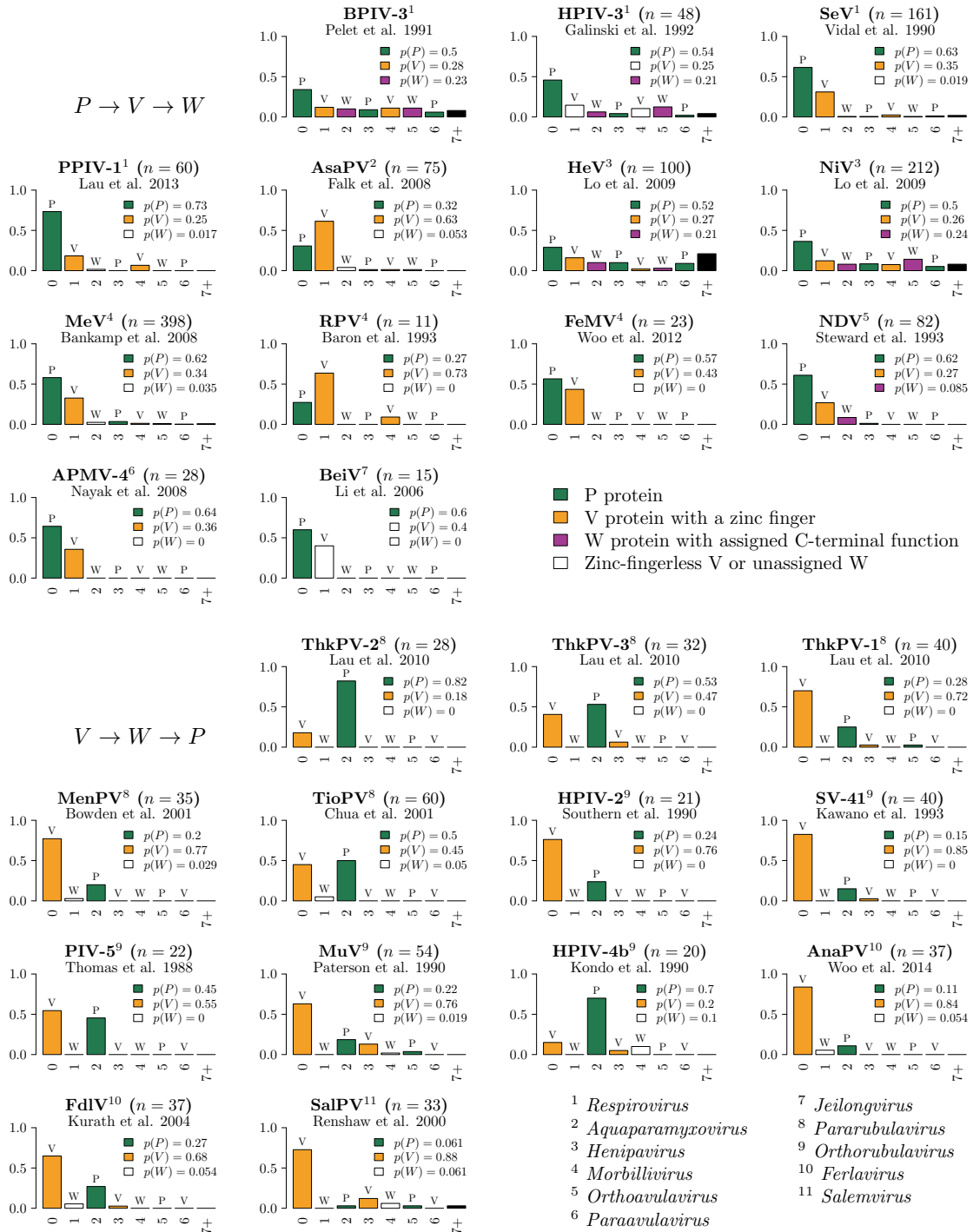


Figure 3: Empirical probability distributions (edit patterns) describing guanosine nucleotide insertion at the P gene edit site. The total proportion of transcripts which encode the three functionally distinct mRNA species are indicated for each experiment. The bulk of the experimental data was obtained by cDNA sequencing, for which the number of sequenced transcripts n is specified. Experimental data for BPIV-3 was obtained by a primer extension method acting directly on the mRNA population. Viral genera indicated in bottom right, see Section 6 for virus names.

Manuscript has not been peer-reviewed

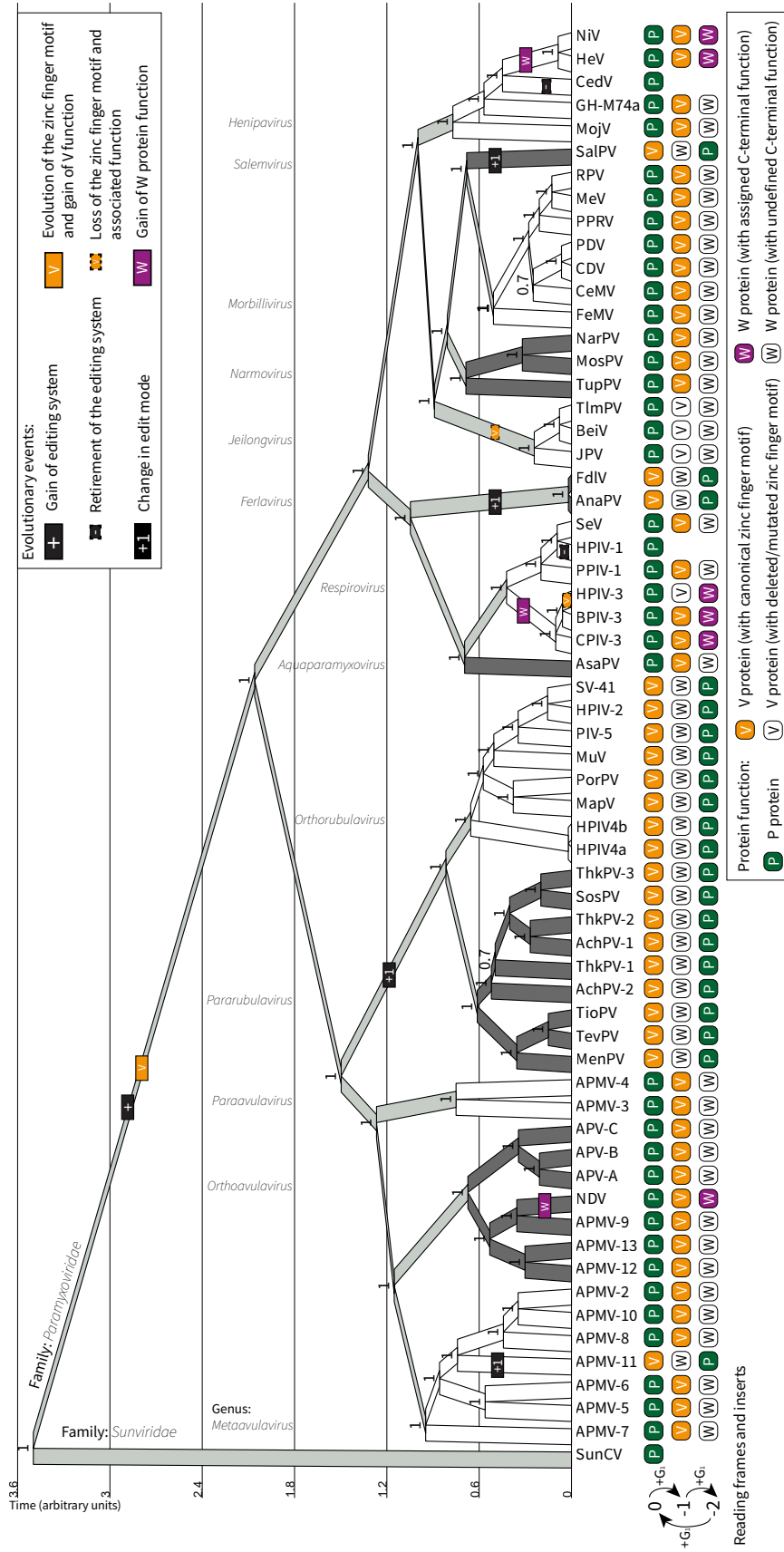


Figure 4: **Phylogeny of the Paramyxoviridae.** Tree created from an alignment of the viral L protein, with Sunshine coast virus (SunCV; Hyndman et al. (2012)) as an outgroup. Coloured rectangles on branches indicate a hypothesised evolutionary event occurring some time in that lineage. Clade posterior supports are shown on the internal nodes. Branch lengths are proportional to time such that there is an average of 1 amino acid substitution per unit of time. Clades are alternatingly coloured by viral genera. See Section 6 for virus names. Tree visualised using UglyTrees (Douglas, 2020).

Manuscript has not been peer-reviewed

3.1 The editing system and evolution of the V protein

This editing system has not been detected beyond the *Paramyxoviridae* (Hyndman et al., 2012). Therefore, the P gene editing system likely came into existence only once – in the lineage which led to the *Paramyxoviridae*. This event was coupled with the origin of the V protein; the evolution of its unique zinc binding motif; and the gain of many of its conserved functions (**Figure 4**). However the timing of these events cannot be resolved.

3.2 Partial or complete loss of the V protein

Under a maximum parsimony model, the V protein has been lost entirely on two independent occasions, both associated with the loss of the editing system (**Figure 4**). The C-terminal zinc binding domain has also been deleted, or significantly mutated, on two further occasions.

Loss of the V protein is associated with retirement of the cotranscriptional editing system. This occurred in both the lineage which lead to Human parainfluenza virus 1 (HPIV-1; genus: *Respirovirus*) and in the lineage which lead to Cedar virus (CedV; genus: *Henipavirus*). As these viruses once employed the $P \xrightarrow{G} V \xrightarrow{G} W$ mode, loss of the editing system was axiomatically coupled with loss of both V and W protein expression. It is possible that loss of V protein activity preceded loss of the edit system, but this cannot be determined. Retirement of the editing system appears impossible for viruses employing the $V \xrightarrow{G} W \xrightarrow{G} P$ edit mode because the P protein is essential for polymerase function.

In both HPIV-1 and CedV, the edit site is not identifiable using sequence analysis and edited transcripts could not be detected experimentally (Matsuoka et al., 1991; Marsh et al., 2012). In HPIV-1, the conserved V protein sequence is apparent in the genome however there is no clear mechanism for protein production due to the presence of multiple stop codons in the reading frame (Matsuoka et al. (1991); **Figure 5**). This suggests that loss of V occurred quite recently in evolutionary history and there has been insufficient time for the sequences to diverge, creating a pseudogene. For CedV, the V protein sequence is undetectable (Marsh et al., 2012).

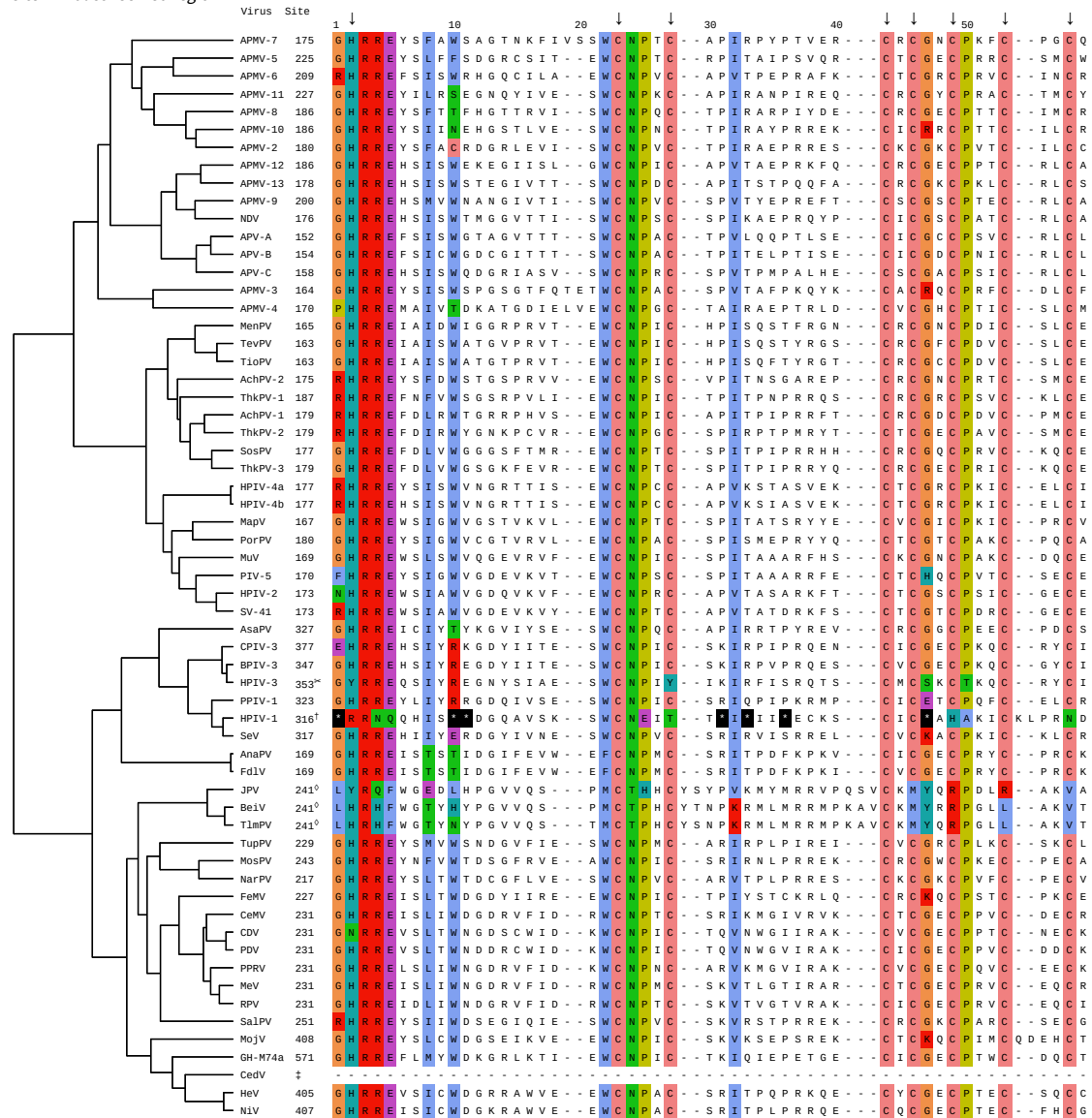
In the case of HPIV-3, the edit site is operational (Galinski et al., 1992) and the zinc finger motif is detectable in the genome by sequence analysis (**Figure 5**). However, several stop codons between the edit site and the zinc finger prohibit production of the full-length V protein (unless further non-canonical transcriptional or translational mechanisms are invoked (Galinski et al., 1992)). There are also two mutations in positions which are directly involved in zinc coordination (**Figure 5**). This suggests the V protein zinc finger is a pseudogene, similar to the situation in HPIV-1. In protein-based analysis of infected cells, the full V protein was not detected but a truncated variant which lacks the conserved C-terminal region was (Roth et al., 2013). Overall, current evidence suggests that the V protein of HPIV-3 is expressed in a truncated form lacking the canonical zinc binding motif. Its functional status is unclear.

Finally, in the case of the *Jeilongviruses*, the V protein C-terminal domain has been retained, but with mutation of several critical residues involved in zinc coordination (**Figure 5**). The C-terminal region does not interact with STAT1 or STAT2 (Audsley et al., 2016b), which is a conserved function of other paramyxoviral V proteins (Puri et al., 2009; Röthlisberger et al., 2010). Nonetheless, the Jeilongviral V protein has retained other functions, such as the ability to interact with MDA5 (Audsley et al., 2016b). This finding in particular highlights the multi-functional nature of the V protein, and the limitations of a nomenclature in which its multiple functionalities are not fully explicated.

Manuscript has not been peer-reviewed

V protein

C-terminal conserved region



- ⌘ Zinc finger domain is deleted due to the presence of stop codons in the sequence
- ◊ Zinc finger domain carries point mutations in critical zinc binding residues
- † Edit system has been lost, V protein sequence is a pseudogene
- ‡ Edit system has been lost, no sequence homologous to the V protein is detected

Figure 5: **Cysteine-rich C-terminal regions of the V protein.** The first amino acid in each aligned sequence is numbered, relative to the start of the V protein. The arrows at the top of the alignment indicate residues whose side chains directly coordinate bound zinc ions, according to a structural study on the PIV-5 V protein (Li et al., 2006a). Asterisks denote stop codons. Sites are coloured by amino acid group if a group is at least 70% conserved at the alignment position (colour scheme indicated in Figure 2). The tree is the same as that in Figure 4.

Manuscript has not been peer-reviewed

3.3 Switching of edit modes and adaption of edit patterns

$P \xrightarrow{G} V \xrightarrow{G} W$ was likely the edit mode of the last common ancestor of the *Paramyxoviridae*.

Under a maximum parsimony model, the editing system has switched to the $V \xrightarrow{G} W \xrightarrow{G} P$ mode four times during evolutionary history (**Figure 4**). These events occurred in the lineages that lead to: 1) Avian paramyxovirus 11 (APMV-11; genus: *Metaaavulavirus*), 2) the *Rubulavirinae* subfamily, 3) the *Ferlaviruses*, and 4) Salem virus (SalPV; genus: *Salemvirus*). Edit patterns have been experimentally investigated for three of these four clades: 10 rubulaviruses (Lau et al., 2010; Bowden et al., 2001; Chua et al., 2001; Southern et al., 1990; Ohgimoto et al., 1990; Kawano et al., 1993; Thomas et al., 1988; Paterson and Lamb, 1990; Takeuchi et al., 1990; Kondo et al., 1990), 2 *Ferlaviruses* (Woo et al., 2014; Kurath et al., 2004), and SalPV (Renshaw et al., 2000).

In general the edit patterns of viruses that retain the ancestral $P \xrightarrow{G} V \xrightarrow{G} W$ edit mode (**Figure 3**, top panel) are quite different to those of viruses that have subsequently adopted the $V \xrightarrow{G} W \xrightarrow{G} P$ edit mode (**Figure 3**, bottom panel). In the former, G_0 and G_1 insertions are most frequently observed, while in the latter, G_0 and G_2 insertions predominate. It seems clear that edit patterns have co-evolved with edit modes to maintain adequate production of P and V transcripts. In two clades (within the *Respirovirus* and *Henipavirus* genera), the edit patterns are long-tailed, and a significant fraction of the transcripts have more than 2 guanosine nucleotides inserted.

The edit pattern of SalPV (**Figure 3**, bottom panel) appears to be an outlier (Renshaw et al., 2000). The G_0 -centric distribution resembles those of viruses using the $P \xrightarrow{G} V \xrightarrow{G} W$ mode, and the relative abundance of P transcripts is very low. Given the taxonomic position of SalPV, as the most immediate outgroup of the *Morbilliviruses* (**Figure 4**), it could be that this is a virus that has switched edit mode but not yet adaptively evolved the edit pattern.

3.4 Acquisition of unique function by the W protein

Under our model, the W protein has evolved a novel function associated with its unique C-terminal region on 3 independent occasions (**Figures 2 and 4**): once for NDV (Yang et al., 2019; Karsunke et al., 2019), once for the henipaviral clade comprised of HeV and NiV (Shaw et al., 2005; Lo et al., 2009; Edwards et al., 2020), and once for the respiroviral clade composed of BPIV-3, HPIV-3, and CPIV-3 (Durbin et al., 1999; Pelet et al., 1991). There are varying levels of experimental evidence supporting the existence of a W protein function in these three clades (see Section 2.3). For the remaining paramyxoviruses, W has no known function. Rather, it is more likely that the expression of W is an inevitable byproduct of the editing system; an evolutionary spandrel (Gould and Lewontin, 1979).

For the most part, W transcripts are produced quite rarely (**Figure 3**). However, this does not appear to be the case for two clades where W has acquired function. Instead, the edit pattern is long-tailed, and the probability of producing a W transcript $\sum_k p(G_{3k+2})$ ranges from 21 to 24% in HeV, NiV, BPIV-3, and HPIV-3 (Lo et al., 2009; Pelet et al., 1991; Galinski et al., 1992), and sometimes even higher in temporal analyses (Kulkarni et al., 2009).

In contrast, production of W is not significantly elevated for NDV (Steward et al., 1993; Mebatsion et al., 2001). The overall proportion of W transcript in NDV is estimated at around 8-9% (Steward et al., 1993; Qiu et al., 2016; Yang et al., 2019) or as low as 2.4%

Manuscript has not been peer-reviewed

(Mebatsion et al., 2001). However, experiments studying the effects of W protein knockout on viral replication (Yang et al., 2019), suggest that these low transcript abundances are optimal for fulfilling the unknown biological function of the NDV W protein (Section 2.3).

4 Molecular mechanism of cotranscriptional editing

In the *Paramyxoviridae*, cotranscriptional editing results from transcriptional slippage. This same process facilitates overprinting in other viral families including the *Filoviridae* (Sanchez et al., 1996; Shabman et al., 2014) and *Potyviridae* (Olspert et al., 2015), as well as various prokaryotes (Larsen et al., 2000; Penno et al., 2015). Slippage sites can also rescue an organism from deleterious frameshift mutations (Tamas et al., 2008).

Transcription has been extensively studied, recently at the single-molecule level for the RdRP of bacteriophage $\phi 6$ (Dulin et al., 2015a,b) and DNA-dependent RNA polymerases of prokaryotes, eukaryotes, and DNA viruses (Abbondanzieri et al., 2005; Shaevitz et al., 2003; Dangkulwanich et al., 2013; Larson et al., 2012; Skinner et al., 2004; Douglas et al., 2020). These studies have provided significant insights into the mechanisms underlying transcription elongation.

In this final section, we discuss cotranscriptional editing in the *Paramyxoviridae* under the framework presented in the single-molecule literature, noting some additional complexities which arise from the viral genome being packaged within a nucleocapsid.

4.1 Transcription elongation by RNA polymerases and slippage

Under a simple Brownian ratchet model, transcription elongation can be modelled as a cycle involving three canonical steps (Bar-Nahum et al. (2005); Abbondanzieri et al. (2005); **Figure 6**, large arrows). First, RNA polymerase steps forward along the template from the pretranslocated to the posttranslocated state, which frees the enzyme's active site. Second, a complementary nucleoside triphosphate (NTP) binds to the active site. Third, the bound NTP is incorporated onto the 3' end of the mRNA and pyrophosphate is released, thus restoring the system to the pretranslocated state.

Through backtracking, where the polymerase translocates upstream along the template (Komissarova and Kashlev, 1997; Abbondanzieri et al., 2005), and hypertranslocation, where it translocates downstream (Yarnell and Roberts, 1999), the polymerase can arrive at a catalytically inactive state (**Figure 6**). These processes can lead to transcriptional pausing (Saba et al., 2018; Artsimovitch and Landick, 2000). In the case of paramyxoviruses, backtracking and hypertranslocation may be inhibited by the presence of genome-wrapped nucleoproteins acting as "roadblocks", analogous to the role played by nucleosomes in eukaryotes (Nudler, 2012).

Slippage involves the movement of one sequence in the product/template hybrid relative to the other, which can lead to imperfect basepairing. Slippage was hypothesised by Streisinger et al. (1966) as one of the primary mechanisms of indel events. The mechanism is thought to involve formation of a nucleotide bulge near the 3' end of the mRNA (Garcia-Diaz and Kunkel, 2006). If the bulge forms in the nascent strand, an insertion can result, whereas a bulge in the template strand can lead to a deletion.

By applying varying forces to individual dsDNA molecules, Kühner et al. (2007) and Neher and Gerland (2004) hypothesise that slippage occurs in three steps (**Figure 6**). First, a bulge forms on one side of the hybrid. This initial reaction must overcome a large Gibbs energy barrier. Second, the bulge diffuses along the hybrid. Diffusion is likely to be quite

Manuscript has not been peer-reviewed

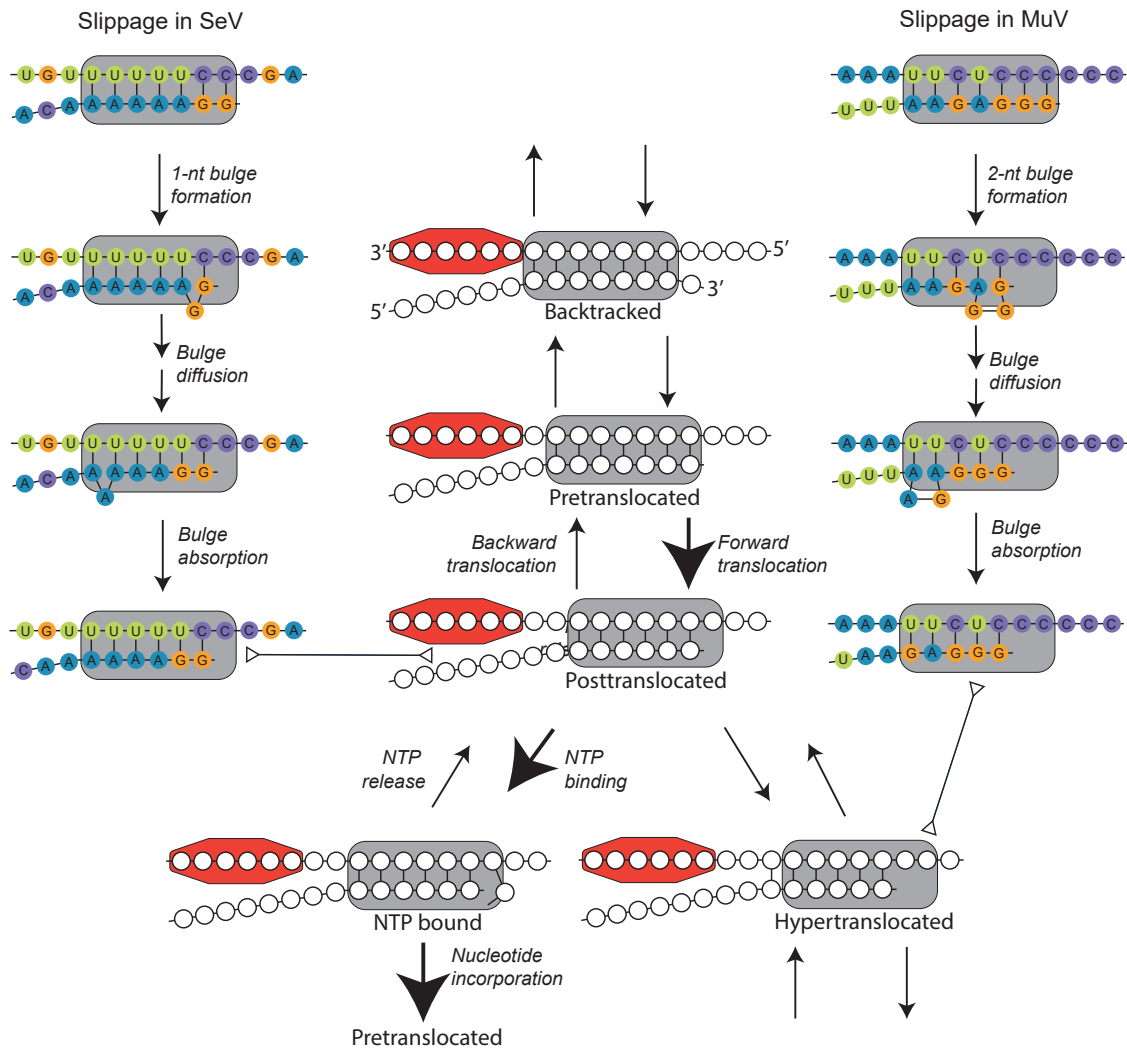


Figure 6: **State diagrams of Brownian ratchet and slippage models.** Plausible stuttering pathways for SeV (accession: AB039658; genomic position: 2783) and MuV (accession: EU884413; genomic position: 2432) are shown, with a RNA/mRNA hybrid of 7 bp in length. A nucleoprotein protomer bound to the viral genome (top strand) is depicted by the coloured octagon. Large arrows indicate the canonical transcription elongation pathway, double-ended triangular arrows denote equivalency between two connecting states, and unlabelled arrows describe translocation reactions. While slippage initialises in the pre-translocated state in this diagram, the actual state where this process initialises is unknown.

Manuscript has not been peer-reviewed

rapid (Woodson and Crothers, 1987), and favoured if Watson-Crick basepairing is maintained in the bulged hybrid. Third, the bulge is absorbed at the other end of the hybrid. During transcription, the next nucleotide could be incorporated before or after absorption, if absorption even occurs (Tippin et al., 2004).

In principle, transcriptional slippage could be initialised from any one of the states available to the polymerase (backtracked, pretranslocated, posttranslocated, or hypertranslocated), and it is not known if the process of bulge formation is linked to the translocation state.

4.2 Transcriptional slippage in paramyxoviruses

Through transcriptional slippage, a single templated nucleotide can be copied multiple times (stuttering). Stuttering is the apparent mechanism of cotranscriptional editing in paramyxoviruses and can explain many of the edit patterns presented in **Figure 3**.

The two distinct modes of editing (i.e. $P \xrightarrow{G} V \xrightarrow{G} W$ and $V \xrightarrow{G} W \xrightarrow{G} P$) are encoded by quite different sequences (**Figure 7**).

The edit sites among viruses which employ the $P \xrightarrow{G} V \xrightarrow{G} W$ edit mode are conserved. Using the PROSITE notation (Sigrist et al., 2002), this (genomic-sense) edit site motif can be described by U(3,6)-C(2,6). In SeV, the edit site sequence is transcribed into AAAAAGgG, where the lower-case g is the stutter site i.e. the site reiteratively transcribed from the template (Hausmann et al., 1999b,a; Vidal et al., 1990b). Under the stuttering model, inserts are added as follows (**Figure 6**, left hand side): 1) a bulge forms in the 3' mRNA of the RNA/mRNA hybrid. 2) The bulge is free to diffuse along the hybrid. Although the bulge is thermodynamically disfavoured, it can occur because of U/A and non-canonical U/G basepairing which are maintained throughout diffusion. 3) In no particular order, the bulge is absorbed at the 5' end and the lower-case g can be transcribed again. Each iteration of these three steps is associated with a G_1 insertion.

In contrast, the edit sites across the four clades of the $V \xrightarrow{G} W \xrightarrow{G} P$ group are quite distinct from one another. SalPV is anomalous, and its edit site sequence resembles the $P \xrightarrow{G} V \xrightarrow{G} W$ group (Renshaw et al., 2000). This could explain the relatively low amounts of P transcript produced (**Figure 3**). The *Ferlavirus* edit site is distinct from all other known edit sites (Woo et al., 2014; Kurath et al., 2004) and the mechanism of guanosine insertion is not clear. Through convergent evolution, APMV-11 and the *Rubulavirinae* subfamily have similar edit sites (PROSITE: A(3,4)-U(2)-C-U(1,2)-C(4,7); genomic-sense). In the case of MuV, the edit site (transcribed into UUUAAGAGGG) has been well characterised (Vidal et al., 1990b). Stuttering occurs in a similar fashion to SeV, however the edit sequence allows G_2 inserts (encoding the P protein) to occur at a greater frequency than G_1 inserts (encoding the W protein) due to the formation of a 2 nucleotide bulge (**Figure 6**, right hand side).

Slight variation in the edit site sequence perturbs stuttering of the viral RdRP. For instance, when the length of the poly(A) sequence at the SeV edit site was increased, from A(3)-G(6) to A(8)-G(1), the average number of inserts increased dramatically (Hausmann et al., 1999a). Similarly, when the SeV edit site sequence was mutated to resemble that of BPIV-3, its edit pattern changed correspondingly (Hausmann et al., 1999b). While slippage patterns can also be dependent on cellular environment (Donohue et al., 2019), these results speak to the primary importance of the genome sequence in governing polymerase stuttering.

The roles that nucleoprotein displacement and the rule of six play during cotranscriptional editing have been investigated (Iseni et al., 2002; Hausmann et al., 1996; Kolakofsky,

Manuscript has not been peer-reviewed

2016). Changing the nucleoprotein phase around the edit site sequence (of SeV) resulted in an apparent change in edit pattern (Iseni et al., 2002). We computed the expected nucleoprotein phase at the edit site of each virus under the rule of six model. Although nucleoprotein displacement may play a role in editing, the nucleoprotein phase at the edit site does not appear to be well conserved (**Figure 7**).

5 Conclusion

In this review, we compiled the genomic sequences of paramyxovirus edit sites (**Figure 7**) and, where available, their experimentally determined edit patterns (histograms of insert sizes; **Figure 3**). We estimated the evolutionary history of editing in the family (**Figure 4**). This analysis suggests that, among the characterised paramyxoviruses, the editing system has independently switched from the $P \xrightarrow{G} V \xrightarrow{G} W$ to the $V \xrightarrow{G} W \xrightarrow{G} P$ edit mode four times; the V protein zinc finger domain was deleted or significantly mutated twice, the W protein has evolved a known function in its unique region three times; and the cotranscriptional editing system has been lost entirely twice, leading to complete loss of both V and W expression. Although transcriptional slippage provides the mechanism for non-templated base insertion, it is currently unclear how this process is coordinated with either canonical or non-canonical steps of the elongation pathway.

6 Virus abbreviations

AchPV 1-2	Achimota viruses 1-2	AnaPV	Anaconda paramyxovirus
APMV 2-13	Avian paramyxoviruses 2-13	APV A-C	Antarctic penguin viruses A-C
AsaPV	Atlantic salmon paramyxovirus	GH-M74a	Ghanaian bat henipavirus
BeiV	Beilong virus	BPIV-3	Bovine parainfluenza virus 3
CDV	Canine distemper virus	CedV	Cedar virus
CeMV	Cetacean morbillivirus	CPIV-3	Caprine parainfluenza virus 3
FdlV	Fer de Lance virus	FeMV	Feline morbillivirus
HeV	Hendra virus	HPIV 1-4	Human parainfluenza viruses 1-4
JPV	J-virus	MenPV	Menangle virus
MeV	Measles virus	MojV	Mojiang virus
MosPV	Mossman virus	MapV	Mapuera virus
MuV	Mumps virus	NarPV	Nariva virus
NDV	Newcastle disease virus	NiV	Nipah virus
PDV	Phocine distemper virus	PIV-5	Parainfluenza virus 5
PorPV	Porcine rubulavirus	PPIV-1	Porcine parainfluenza virus 1
PPRV	Peste-des-petits-ruminants virus	RPV	Rinderpest virus
SalPV	Salem virus	SeV	Sendai virus
SosPV	Sosuga virus	SunCV	Sunshine coast virus
SV-41	Simian virus 41	TevPV	Teviot virus
ThkPV 1-3	Tuhoko viruses 1-3	TioPV	Tioman virus
TlmPV	Tailam virus	TupPV	Tupaia virus

Manuscript has not been peer-reviewed

P gene edit sites

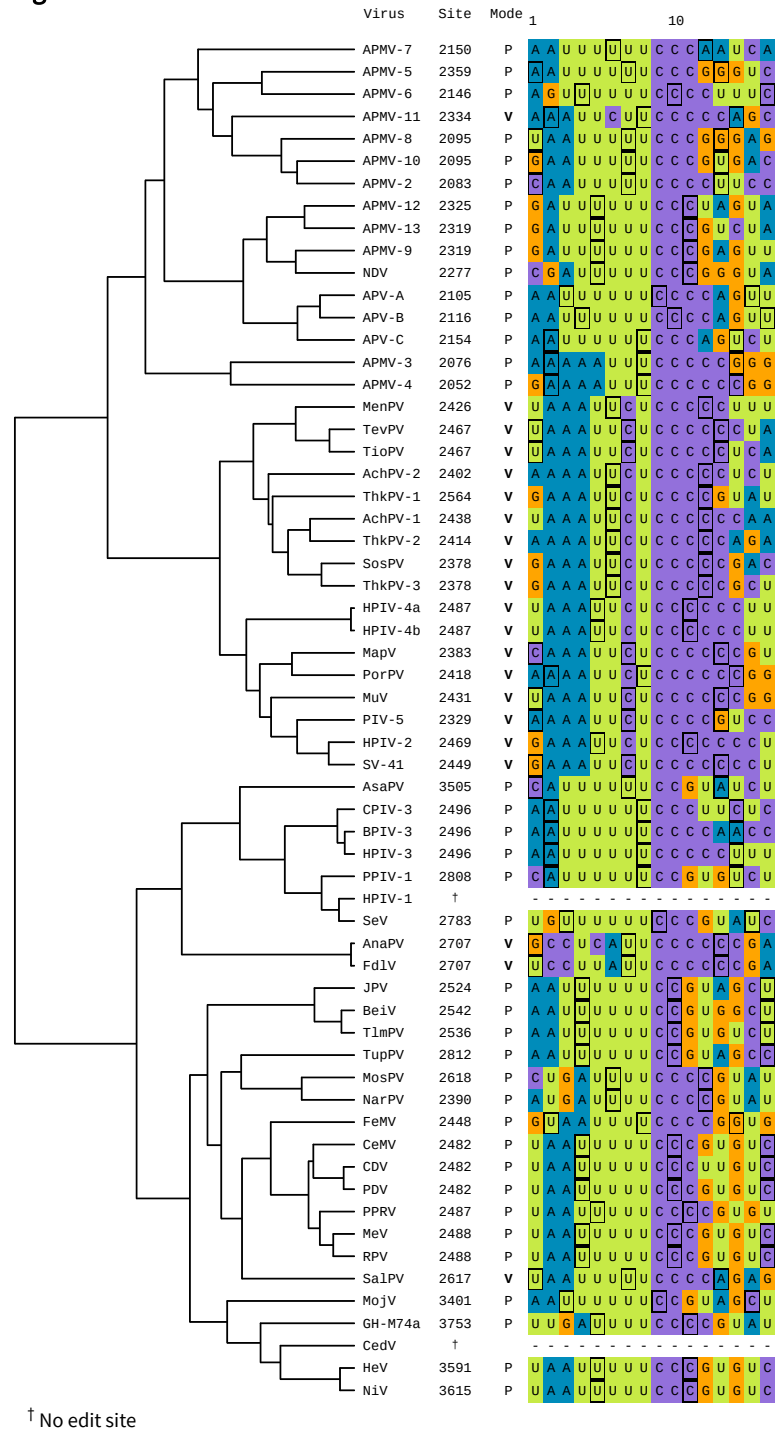


Figure 7: **Edit site sequences in the paramyxoviruses.** The sequences of the negative-sense (genomic) RNA are displayed. The numbers indicate the genomic position of the first displayed nucleotide. $P \xrightarrow{G} V \xrightarrow{G} W$ and $V \xrightarrow{G} W \xrightarrow{G} P$ edit modes are abbreviated to P and V respectively. Nucleoprotein phases are displayed; the first nucleotide within each nucleoprotein protomer is highlighted in black. This tree is the same as that in **Figure 4**.

Manuscript has not been peer-reviewed

7 Algorithms and data availability

Sequences were aligned by M-Coffee (Wallace, 2006) and treated with subsequent manual adjustment using AliView (Larsson, 2014). Phylogenetic tree built with BEAST 2 (Bouckaert et al., 2019) from an alignment of the L protein, and a relaxed clock model Drummond et al. (2006). P/V/W sequences, L alignment, and BEAST 2 input/output files are available at

<https://github.com/jordandouglas/ParamyxovirusSlippageEvolution>.

8 Funding

This work was supported by the University of Auckland Doctoral Scholarship.

References

- Elio A Abbondanzieri, William J Greenleaf, Joshua W Shaevitz, Robert Landick, and Steven M Block. Direct observation of base-pair stepping by rna polymerase. *Nature*, 438(7067):460–465, 2005.
- Ryan Abdella, Megha Aggarwal, Takashi Okura, Robert A Lamb, and Yuan He. Structure of a paramyxovirus polymerase complex reveals a unique methyltransferase-ctd conformation. *Proceedings of the National Academy of Sciences*, 117(9):4931–4941, 2020.
- Judith G Alamares, Subbiah Elankumaran, Siba K Samal, and Ronald M Iorio. The interferon antagonistic activities of the v proteins from two strains of newcastle disease virus correlate with their known virulence properties. *Virus research*, 147(1):153–157, 2010.
- Maher Alayyoubi, George P Leser, Christopher A Kors, and Robert A Lamb. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein–rna complex. *Proceedings of the National Academy of Sciences*, 112(14):E1792–E1799, 2015.
- Gaya K Amarasinghe, María A Ayllón, Yīmíng Bào, Christopher F Basler, Sina Bavari, Kim R Blasdel, Thomas Briese, Paul A Brown, Alexander Bukreyev, Anne Balkema-Buschmann, et al. Taxonomy of the order mononegavirales: update 2019. *Archives of virology*, 164(7):1967–1980, 2019.
- J Andrejeva, KS Childs, DF Young, TS Carlos, N Stock, S Goodbourn, and RE Randall. The v proteins of paramyxoviruses bind the ifn-inducible rna helicase, mda-5, and inhibit its activation of the ifn- β promoter. *Proceedings of the National Academy of Sciences*, 101(49):17264–17269, 2004.
- Irina Artsimovitch and Robert Landick. Pausing by bacterial rna polymerase is mediated by mechanistically distinct classes of signals. *Proceedings of the National Academy of Sciences*, 97(13):7090–7095, 2000.
- Michelle D Audsley, David A Jans, and Gregory W Moseley. Nucleocytoplasmic trafficking of nipah virus w protein involves multiple discrete interactions with the nuclear import and export machinery. *Biochemical and biophysical research communications*, 479(3):429–433, 2016a.

Manuscript has not been peer-reviewed

- Michelle D Audsley, Glenn A Marsh, Kim G Lieu, Mary Tachedjian, D Albert Joubert, Lin-Fa Wang, David A Jans, and Gregory W Moseley. The immune evasion function of j and beilong virus v proteins is distinct from that of other paramyxoviruses, consistent with their inclusion in the proposed genus jeilongvirus. *Journal of General Virology*, 97(3):581–592, 2016b.
- B Bankamp, EN Lopareva, JR Kremer, Y Tian, MS Clemens, R Patel, AL Fowlkes, JR Kessler, CP Muller, WJ Bellini, et al. Genetic variability and mrna editing frequencies of the phosphoprotein genes of wild-type measles viruses. *Virus research*, 135(2):298–306, 2008.
- Gil Bar-Nahum, Vitaly Epshtein, Andrei E Ruckenstein, Ruslan Rafikov, Arkady Mustaev, and Evgeny Nudler. A ratchet mechanism of transcription elongation and its control. *Cell*, 120(2):183–193, 2005.
- MD Baron, MS Shaila, and T Barrett. Cloning and sequence analysis of the phosphoprotein gene of rinderpest virus. *Journal of general virology*, 74(2):299–304, 1993.
- Mikael Berg, Bernt Hjertner, Jorge Moreno-López, and Tommy Linné. The p gene of the porcine paramyxovirus lpmv encodes three possible polypeptides p, v and c: the p protein mrna is edited. *Journal of general virology*, 73(5):1195–1200, 1992.
- Laurence Blanchard, Nicolas Tarbouriech, Martin Blackledge, Peter Timmins, Wilhelm P Burmeister, Rob WH Ruigrok, and Dominique Marion. Structure and dynamics of the nucleocapsid-binding domain of the sendai virus phosphoprotein in solution. *Virology*, 319(2):201–211, 2004.
- Merete Blixenkrone-Møller, Bhaskar Sharma, Tamas M Varsanyi, Aizhung Hu, Erling Norby, and Jan Kövamees. Sequence analysis of the genes encoding the nucleocapsid protein and phosphoprotein (p) of phocid distemper virus, and editing of the p gene transcript. *Journal of general virology*, 73(4):885–893, 1992.
- Louis-Marie Bloyet, Joanna Brunel, Marion Dosnon, Véronique Hamon, Jenny Erales, Antoine Gruet, Carine Lazert, Christophe Bignon, Philippe Roche, Sonia Longhi, et al. Modulation of re-initiation of measles virus transcription at intergenic regions by pxd to ntail binding strength. *PLoS pathogens*, 12(12):e1006058, 2016.
- Ronald Boeck, J Curran, Y Matsuoka, R Compans, and D Kolakofsky. The parainfluenza virus type 1 p/c gene uses a very efficient gug codon to start its c'protein. *Journal of virology*, 66(3):1765–1768, 1992.
- Gert Bolt, Søren Alexandersen, and Merete Blixenkrone-Møller. The phosphoprotein gene of a dolphin morbillivirus isolate exhibits genomic variation at the editing site. *Journal of general virology*, 76(12):3051–3058, 1995.
- Remco Bouckaert, Timothy G Vaughan, Joëlle Barido-Sottani, Sebastián Duchêne, Mathieu Fourment, Alexandra Gavryushkina, Joseph Heled, Graham Jones, Denise Kühnert, Nicola De Maio, et al. Beast 2.5: An advanced software platform for bayesian evolutionary analysis. *PLoS computational biology*, 15(4):e1006650, 2019.
- Timothy R Bowden, Marcel Westenberg, Lin-Fa Wang, Bryan T Eaton, and David B Boyle. Molecular characterization of menangle virus, a novel paramyxovirus which infects pigs, fruit bats, and humans. *Virology*, 283(2):358–373, 2001.

Manuscript has not been peer-reviewed

- Nadav Brandes and Michal Linial. Gene overlapping and size constraints in the viral world. *Biology Direct*, 11(1), May 2016. doi: 10.1186/s13062-016-0128-3. URL <https://doi.org/10.1186/s13062-016-0128-3>.
- Jessica F Bruhn, Katherine C Barnett, Jaclyn Bibby, Jens MH Thomas, Ronan M Keegan, Daniel J Rigden, Zachary A Bornholdt, and Erica Ollmann Saphire. Crystal structure of the nipah virus phosphoprotein tetramerization domain. *Journal of virology*, 88(1): 758–762, 2014.
- Jessica F Bruhn, Anne L Hotard, Christina F Spiropoulou, Michael K Lo, and Erica Ollmann Saphire. A conserved basic patch and central kink in the nipah virus phosphoprotein multimerization domain are essential for polymerase function. *Structure*, 27(4):660–668, 2019.
- Wilhelm P. Burmeister, Nicolas Tarbouriech, Joseph Curran, and Rob W.H. Ruigrok. *Nature Structural Biology*, 7(9):777–781, September 2000. doi: 10.1038/79013. URL <https://doi.org/10.1038/79013>.
- Amy L Burroughs, Mary Tachedjian, Gary Crameri, Peter A Durr, Glenn A Marsh, and Lin-Fa Wang. Complete genome sequence of teviot paramyxovirus, a novel rubulavirus isolated from fruit bats in australia. *Genome Announcements*, 3(2), 2015.
- Philippe Calain and Laurent Roux. The rule of six, a basic feature for efficient replication of sendai virus defective interfering rna. *Journal of virology*, 67(8):4822–4830, 1993.
- Roberto Cattaneo, Gabriela Rebmann, Anita Schmid, K Baczko, V Ter Meulen, and MA Billeter. Altered transcription of a defective measles virus genome derived from a diseased human brain. *The EMBO Journal*, 6(3):681–688, 1987.
- Roberto Cattaneo, Karin Kaelin, Knut Baczko, and Martin A Billeter. Measles virus editing provides an additional cysteine-rich protein. *Cell*, 56(5):759–764, 1989.
- Kay Childs, Nicola Stock, Craig Ross, Jelena Andrejeva, Louise Hilton, Michael Skinner, Richard Randall, and Stephen Goodbourn. mda-5, but not rig-i, is a common target for paramyxovirus v proteins. *Virology*, 359(1):190–200, 2007.
- Senthil K Chinnakannan, Barbara Holzer, Beatriz Sanz Bernardo, Sambit K Nanda, and Michael D Baron. Different functions of the common p/v/w and v-specific domains of rinderpest virus v protein in blocking ifn signalling. *Journal of General Virology*, 95(1): 44–51, 2014.
- Kaw Bing Chua, Lin-Fa Wang, Sai Kit Lam, Gary Crameri, Meng Yu, Terry Wise, David Boyle, Alex D Hyatt, and Bryan T Eaton. Tioman virus, a novel paramyxovirus isolated from fruit bats in malaysia. *Virology*, 283(2):215–229, 2001.
- Michael J Ciancanelli, Valentina A Volchkova, Megan L Shaw, Viktor E Volchkov, and Christopher F Basler. Nipah virus sequesters inactive stat1 in the nucleus via a p gene-encoded mechanism. *Journal of virology*, 83(16):7828–7841, 2009.
- G. Communie, T. Crepin, D. Maurin, M. R. Jensen, M. Blackledge, and R. W. H. Ruigrok. Structure of the tetramerization domain of measles virus phosphoprotein. *Journal of Virology*, 87(12):7166–7169, April 2013a. doi: 10.1128/jvi.00487-13. URL <https://doi.org/10.1128/jvi.00487-13>.

Manuscript has not been peer-reviewed

- Guillaume Communie, Johnny Habchi, Filip Yabukarski, David Blocquel, Robert Schneider, Nicolas Tarbouriech, Nicolas Papageorgiou, RW Ruigrok, Marc Jamin, Malene Ringkjøbing Jensen, et al. Atomic resolution description of the interaction between the nucleoprotein and phosphoprotein of hendra virus. *PLoS Pathog*, 9(9):e1003631, 2013b.
- Robert Cox, Todd J Green, Sangeetha Purushotham, Champion Deivanayagam, Gregory J Bedwell, Peter E Prevelige, and Ming Luo. Structural and functional characterization of the mumps virus phosphoprotein. *Journal of virology*, 87(13):7558–7568, 2013.
- J. Curran, R. Boeck, and D. Kolakofsky. The sendai virus p gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *The EMBO Journal*, 10(10):3079–3085, October 1991. doi: 10.1002/j.1460-2075.1991.tb07860.x. URL <https://doi.org/10.1002/j.1460-2075.1991.tb07860.x>.
- Joseph Curran and Daniel Kolakofsky. Ribosomal initiation from an acg codon in the sendai virus p/c mrna. *The EMBO journal*, 7(1):245–251, 1988.
- Manchuta Dangkulwanich, Toyotaka Ishibashi, Shixin Liu, Maria L Kireeva, Lucyna Lubkowska, Mikhail Kashlev, and Carlos J Bustamante. Complete dissection of transcription elongation reveals slow translocation of rna polymerase ii in a linear ratchet mechanism. *Elife*, 2:e00971, 2013.
- Meredith E Davis, May K Wang, Linda J Rennick, Florian Full, Sebastian Gableske, Annelies W Mesman, Sonja I Gringhuis, Teunis BH Geijtenbeek, W Paul Duprex, and Michaela U Gack. Antagonism of the phosphatase pp1 by the measles virus v protein is required for innate immune escape of mda5. *Cell host & microbe*, 16(1):19–30, 2014.
- Patricia Devaux, Veronika von Messling, Warangkhan Songsungthong, Christoph Springfield, and Roberto Cattaneo. Tyrosine 110 in the measles virus phosphoprotein is required to block stat1 phosphorylation. *Virology*, 360(1):72–83, 2007.
- Patricia Devaux, Gregory Hodge, Michael B McChesney, and Roberto Cattaneo. Attenuation of v-or c-defective measles viruses: infection control by the inflammatory and interferon responses of rhesus monkeys. *Journal of virology*, 82(11):5359–5367, 2008.
- Patricia Devaux, Andrew W Hudacek, Gregory Hodge, Jorge Reyes-del Valle, Michael B McChesney, and Roberto Cattaneo. Stat1-blind measles virus cannot control inflammation and is attenuated in rhesus monkeys. *Journal of Virology*, 2010.
- L Didcock, DF Young, S Goodbourn, and RE Randall. The v protein of simian virus 5 inhibits interferon signalling by targeting stat1 for proteasome-mediated degradation. *Journal of virology*, 73(12):9928–9933, 1999.
- Ryan C Donohue, Christian K Pfaller, and Roberto Cattaneo. Cyclical adaptation of measles virus quasispecies to epithelial and lymphocytic cells: To v, or not to v. *PLoS pathogens*, 15(2):e1007605, 2019.
- Jordan Douglas. UglyTrees: a browser-based multispecies coalescent tree visualiser. *Bioinformatics*, 2020.
- Jordan Douglas, Richard Kingston, and Alexei J Drummond. Bayesian inference and comparison of stochastic transcription elongation models. *PLoS computational biology*, 16(2):e1006717, 2020.

Manuscript has not been peer-reviewed

- Alexei J Drummond, Simon YW Ho, Matthew J Phillips, and Andrew Rambaut. Relaxed phylogenetics and dating with confidence. *PLoS biology*, 4(5):e88, 2006.
- Venice Du Pont, Yi Jiang, and Richard K Plemper. Bipartite interface of the measles virus phosphoprotein x domain with the large polymerase protein regulates viral polymerase dynamics. *PLoS pathogens*, 15(8):e1007995, 2019.
- David Dulin, Igor D Vilfan, Bojk A Berghuis, Susanne Hage, Dennis H Bamford, Minna M Poranen, Martin Depken, and Nynke H Dekker. Elongation-competent pauses govern the fidelity of a viral rna-dependent rna polymerase. *Cell reports*, 10(6):983–992, 2015a.
- David Dulin, Igor D Vilfan, Bojk A Berghuis, Minna M Poranen, Martin Depken, and Nynke H Dekker. Backtracking behavior in viral rna-dependent rna polymerase provides the basis for a second initiation site. *Nucleic acids research*, 43(21):10421–10429, 2015b.
- Anna P Durbin, Josephine M McAuliffe, Peter L Collins, and Brian R Murphy. Mutations in the c, d, and v open reading frames of human parainfluenza virus type 3 attenuate replication in rodents and primates. *Virology*, 261(2):319–330, 1999.
- Megan R Edwards, Mikayla Hoad, Sofiya Tsimbalyuk, Andrea R Menicucci, Ilhem Mes-saoudi, Jade K Forwood, and Christopher F Basler. Henipavirus w proteins interact with 14-3-3 to modulate host gene expression. *Journal of Virology*, 2020.
- K Falk, WN Batts, A Kvellestad, G Kurath, J Wiik-Nielsen, and JR Winton. Molecular characterisation of atlantic salmon paramyxovirus (aspv): a novel paramyxovirus associated with proliferative gill inflammation. *Virus research*, 133(2):218–227, 2008.
- Rachel Fearn and Richard K Plemper. Polymerases of paramyxoviruses and pneumoviruses. *Virus research*, 234:87–102, 2017.
- Andrew E Firth and Ian Brierley. Non-canonical translation in rna viruses. *The Journal of general virology*, 93(Pt 7):1385, 2012.
- Judith M. Fontana, Bettina Bankamp, and Paul A. Rota. Inhibition of interferon induction and signaling by paramyxoviruses. *Immunological Reviews*, 225(1):46–67, October 2008a. doi: 10.1111/j.1600-065x.2008.00669.x. URL <https://doi.org/10.1111/j.1600-065x.2008.00669.x>.
- Judith M Fontana, Bettina Bankamp, and Paul A Rota. Inhibition of interferon induction and signaling by paramyxoviruses. *Immunological reviews*, 225(1):46–67, 2008b.
- Mark S Galinski, Roberta M Troy, and Amiya K Banerjee. Rna editing in the phosphoprotein gene of the human parainfluenza virus type 3. *Virology*, 186(2):543–550, 1992.
- Miguel Garcia-Diaz and Thomas A Kunkel. Mechanism of a genetic glissando*: structural biology of indel mutations. *Trends in biochemical sciences*, 31(4):206–214, 2006.
- Colomba Giorgi, Benjamin M Blumberg, and Daniel Kolakofsky. Sendai virus contains overlapping genes expressed from a single mrna. *Cell*, 35(3):829–836, 1983.
- Stephen Jay Gould and Richard C Lewontin. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161):581–598, 1979.

Manuscript has not been peer-reviewed

- Sergey G Guryanov, Lassi Liljeroos, Prasad Kasaragod, Tommi Kajander, and Sarah J Butcher. Crystal structure of the measles virus nucleoprotein core in complex with an n-terminal region of phosphoprotein. *Journal of virology*, 90(6):2849–2857, 2016.
- Serafima Guseva, Sigrid Milles, Martin Blackledge, and Rob WH Ruigrok. The nucleoprotein and phosphoprotein of measles virus. *Frontiers in microbiology*, 10:1832, 2019.
- Irina Gutsche, Ambroise Desfosses, Grégory Effantin, Wai Li Ling, Melina Haupt, Rob WH Ruigrok, Carsten Sachse, and Guy Schoehn. Near-atomic cryo-em structure of the helical measles virus nucleocapsid. *Science*, 348(6235):704–707, 2015.
- Johnny Habchi and Sonia Longhi. Structural disorder within paramyxovirus nucleoproteins and phosphoproteins. *Mol. BioSyst.*, 8(1):69–81, 2012. doi: 10.1039/c1mb05204g. URL <https://doi.org/10.1039/c1mb05204g>.
- Johnny Habchi, Stéphanie Blangy, Laurent Mamelli, Malene Ringkjøbing Jensen, Martin Blackledge, Hervé Darbon, Michael Oglesbee, Yaoling Shu, and Sonia Longhi. Characterization of the interactions between the nucleoprotein and the phosphoprotein of henipavirus. *Journal of Biological Chemistry*, 286(15):13583–13602, 2011.
- Stéphane Hausmann, Jean-Philippe Jacques, and Daniel Kolakofsky. Paramyxovirus RNA editing and the requirement for hexamer genome length. *Rna*, 2(10):1033–1045, 1996.
- Stephane Hausmann, Dominique Garcin, Christophe Delenda, and Daniel Kolakofsky. The versatility of paramyxovirus rna polymerase stuttering. *Journal of virology*, 73(7):5568–5576, 1999a.
- Stéphane Hausmann, Dominique Garcin, Anne-Sophie Morel, and Daniel Kolakofsky. Two nucleotides immediately upstream of the essential a6g3 slippery sequence modulate the pattern of g insertions during sendai virus mrna editing. *Journal of virology*, 73(1):343–351, 1999b.
- Sandra M. Horikami, Sherin Smallwood, and Sue A. Moyer. The sendai virus v protein interacts with the NP protein to regulate viral genome RNA replication. *Virology*, 222(2):383–390, August 1996. doi: 10.1006/viro.1996.0435. URL <https://doi.org/10.1006/viro.1996.0435>.
- Timothy H Hyndman, Rachel E Marschang, James FX Wellehan Jr, and Philip K Nicholls. Isolation and molecular identification of sunshine virus, a novel paramyxovirus found in australian snakes. *Infection, Genetics and Evolution*, 12(7):1436–1446, 2012.
- Frédéric Iseni, Florence Baudin, Dominique Garcin, Jean-Baptiste Marq, Rob WH Ruigrok, and Daniel Kolakofsky. Chemical modification of nucleotide bases and mrna editing depend on hexamer or nucleoprotein phase in sendai virus nucleocapsids. *Rna*, 8(8):1056–1067, 2002.
- J.P. Jacques, S. Hausmann, and D. Kolakofsky. Paramyxovirus mRNA editing leads to g deletions as well as insertions. *The EMBO Journal*, 13(22):5496–5503, November 1994. doi: 10.1002/j.1460-2075.1994.tb06884.x. URL <https://doi.org/10.1002/j.1460-2075.1994.tb06884.x>.

Manuscript has not been peer-reviewed

- Marc Jamin and Filip Yabukarski. Nonsegmented negative-sense rna viruses—structural data bring new insights into nucleocapsid assembly. In *Advances in virus research*, volume 97, pages 143–185. Elsevier, 2017.
- Kent Johansson, Jean-Marie Bourhis, Valerie Campanacci, Christian Cambillau, Bruno Canard, and Sonia Longhi. Crystal structure of the measles virus phosphoprotein domain responsible for the induced folding of the c-terminal domain of the nucleoprotein. *Journal of Biological Chemistry*, 278(45):44567–44573, August 2003. doi: 10.1074/jbc.m308745200. URL <https://doi.org/10.1074/jbc.m308745200>.
- Rebecca I Johnson, Mary Tachedjian, Bronwyn A Clayton, Rachel Layton, Jemma Bergfeld, Lin-Fa Wang, and Glenn A Marsh. Characterization of teviot virus, an australian bat-borne paramyxovirus. *Journal of General Virology*, 100(3):403–413, 2019.
- David Karlin and Robert Belshaw. Detecting remote sequence homology in disordered proteins: Discovery of conserved motifs in the n-termini of mononegavirales phosphoproteins. *PLoS ONE*, 7(3):e31719, March 2012. doi: 10.1371/journal.pone.0031719. URL <https://doi.org/10.1371/journal.pone.0031719>.
- Julia Karsunke, Sandra Heiden, Magdalena Murr, Axel Karger, Kati Franzke, Thomas C. Mettenleiter, and Angela Römer-Oberdörfer. W protein expression by newcastle disease virus. *Virus Research*, 263:207–216, April 2019. doi: 10.1016/j.virusres.2019.02.003. URL <https://doi.org/10.1016/j.virusres.2019.02.003>.
- Atsushi Kato, Katsuhiko Kiyotani, Yuko Sakai, Tetsuya Yoshida, Tatsuo Shioda, and Yoshiyuki Nagai. Importance of the cysteine-rich carboxyl-terminal half of v protein for sendai virus pathogenesis. *Journal of Virology*, 71(10):7266–7272, 1997.
- Mitsuo Kawano, Masato Tsurudome, Naohiro Oki, Machiko Nishio, Hiroshi Komada, Haruo Matsumura, Shigeru Kusagawa, Hisataka Ohta, and Yasuhiko Ito. Sequence determination of the p gene of simian virus 41: presence of irregular deletions near the rna-editing sites of paramyxoviruses. *Journal of general virology*, 74(5):911–916, 1993.
- Andrew MQ King, Elliot Lefkowitz, Michael J Adams, and Eric B Carstens. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, volume 9. Elsevier, 2011.
- Richard L Kingston, Damon J Hamel, Leslie S Gay, Frederick W Dahlquist, and Brian W Matthews. Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proceedings of the National Academy of Sciences of the United States of America*, 101(22):8301–8306, 2004.
- Richard L Kingston, Leslie S Gay, Walter S Baase, and Brian W Matthews. Structure of the nucleocapsid-binding domain from the mumps virus polymerase; an example of protein folding induced by crystallization. *Journal of molecular biology*, 379(4):719–731, 2008.
- Yoshinori Kitagawa, Mayu Yamaguchi, Min Zhou, Machiko Nishio, Masae Itoh, and Bin Gotoh. Human parainfluenza virus type 2 v protein inhibits traf6-mediated ubiquitination of irf7 to prevent tlr7- and tlr9-dependent interferon induction. *Journal of virology*, 87(14):7966–7976, 2013.

Manuscript has not been peer-reviewed

Daniel Kolakofsky. Paramyxovirus rna synthesis, mrna editing, and genome hexamer phase: a review. *Virology*, 498:94–98, 2016.

Daniel Kolakofsky, Thierry Pelet, Dominique Garcin, Stéphane Hausmann, Joseph Curran, and Laurent Roux. Paramyxovirus rna synthesis and the requirement for hexamer genome length: the rule of six revisited. *Journal of virology*, 72(2):891–899, 1998.

Daniel Kolakofsky, Laurent Roux, Dominique Garcin, and Rob WH Ruigrok. Paramyxovirus mrna editing, the ‘rule of six’ and error catastrophe: a hypothesis. *Journal of general virology*, 86(7):1869–1877, 2005.

Natalia Komissarova and Mikhail Kashlev. Transcriptional arrest: Escherichia coli rna polymerase translocates backward, leaving the 3’ end of the rna intact and extruded. *Proceedings of the National Academy of Sciences*, 94(5):1755–1760, 1997.

Kunio Kondo, Hisanori Bando, Masato Tsurudome, Mitsuo Kawano, Machiko Nishio, and Yasuhiko Ito. Sequence analysis of the phosphoprotein (p) genes of human parainfluenza type 4a and 4b viruses and rna editing at transcript of the p genes: the number of g residues added is imprecise. *Virology*, 178(1):321–326, 1990.

Ferdinand Kühner, Julia Morfill, Richard A Neher, Kerstin Blank, and Hermann E Gaub. Force-induced dna slippage. *Biophysical journal*, 92(7):2491–2497, 2007.

Sachin Kulkarni, Valentina Volchkova, Christopher F Basler, Peter Palese, Viktor E Volchkov, and Megan L Shaw. Nipah virus edits its p gene at high frequency to express the v and w proteins. *Journal of virology*, 83(8):3982–3987, 2009.

Gael Kurath, William N Batts, Winfried Ahne, and James R Winton. Complete genome sequence of fer-de-lance virus reveals a novel gene in reptilian paramyxoviruses. *Journal of virology*, 78(4):2045–2056, 2004.

Mark A Larkin, Gordon Blackshields, Nigel P Brown, R Chenna, Paul A McGettigan, Hamish McWilliam, Franck Valentin, Iain M Wallace, Andreas Wilm, Rodrigo Lopez, et al. Clustal w and clustal x version 2.0. *bioinformatics*, 23(21):2947–2948, 2007.

Bente Larsen, Norma M Wills, Chad Nelson, John F Atkins, and Raymond F Gesteland. Nonlinearity in genetic decoding: homologous dna replicase genes use alternatives of transcriptional slippage or translational frameshifting. *Proceedings of the National Academy of Sciences*, 97(4):1683–1688, 2000.

Matthew H Larson, Jing Zhou, Craig D Kaplan, Murali Palangat, Roger D Kornberg, Robert Landick, and Steven M Block. Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of rna polymerase ii. *Proceedings of the National Academy of Sciences*, 109(17):6555–6560, 2012.

Anders Larsson. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22):3276–3278, August 2014. doi: 10.1093/bioinformatics/btu531. URL <https://doi.org/10.1093/bioinformatics/btu531>.

Patrizia Latorre, Daniel Kolakofsky, and Joseph Curran. Sendai virus y proteins are initiated by a ribosomal shunt. *Molecular and Cellular Biology*, 18(9):5021–5031, 1998.

Manuscript has not been peer-reviewed

- Susanna KP Lau, Patrick CY Woo, Beatrice HL Wong, Annette YP Wong, Hoi-Wah Tsoi, Ming Wang, Paul Lee, Huifang Xu, Rosana WS Poon, Rongtong Guo, et al. Identification and complete genome analysis of three novel paramyxoviruses, tuhoko virus 1, 2 and 3, in fruit bats from china. *Virology*, 404(1):106–116, 2010.
- Susanna KP Lau, Patrick CY Woo, Ying Wu, Annette YP Wong, Beatrice HL Wong, Candy CY Lau, Rachel YY Fan, Jian-Piao Cai, Hoi-Wah Tsoi, Kwok-Hung Chan, et al. Identification and characterization of a novel paramyxovirus, porcine parainfluenza virus 1, from deceased pigs. *Journal of General Virology*, 94(10):2184–2190, 2013.
- Pengfei Li, Zixiang Zhu, Xiangli Zhang, Wen Dang, Linlin Li, Xiaoli Du, Miaotao Zhang, Chunyan Wu, Qinghong Xue, Xiangtao Liu, et al. The nucleoprotein and phosphoprotein of peste des petits ruminants virus inhibit interferons signaling by blocking the jak-stat pathway. *Viruses*, 11(7):629, 2019.
- Ti Li, Xiujuan Chen, Kenneth C. Garbutt, Pengbo Zhou, and Ning Zheng. Structure of DDB1 in complex with a paramyxovirus v protein: Viral hijack of a propeller cluster in ubiquitin ligase. *Cell*, 124(1):105–117, January 2006a. doi: 10.1016/j.cell.2005.10.033. URL <https://doi.org/10.1016/j.cell.2005.10.033>.
- Zhuo Li, Meng Yu, Hong Zhang, Danielle E Magoffin, Philippa JM Jack, Alex Hyatt, Hai-Yan Wang, and Lin-Fa Wang. Beilong virus, a novel paramyxovirus with the largest genome of non-segmented negative-stranded rna viruses. *Virology*, 346(1):219–228, 2006b.
- Peter Liston and Dalius J Briedis. Measles virus v protein binds zinc. *Virology*, 198(1): 399–404, 1994.
- Yansheng Liu, Andreas Beyer, and Ruedi Aebersold. On the dependency of cellular protein levels on mrna abundance. *Cell*, 165(3):535–550, 2016.
- Michael K Lo, Brian H Harcourt, Bruce A Mungall, Azaibi Tamin, Mark E Peeples, William J Bellini, and Paul A Rota. Determination of the henipavirus phosphoprotein gene mrna editing frequencies and detection of the c, v and w proteins of nipah virus in virus-infected cells. *Journal of General Virology*, 90(2):398–404, 2009.
- Lenette L Lu, Mamta Puri, Curt M Horvath, and Ganes C Sen. Select paramyxoviral v proteins inhibit irf3 activation by acting as alternative substrates for inhibitor of κ b kinase ϵ (ikke)/tbk1. *Journal of Biological Chemistry*, 283(21):14269–14276, 2008.
- Louise E Ludlow, Michael K Lo, Jason J Rodriguez, Paul A Rota, and Curt M Horvath. Henipavirus v protein association with polo-like kinase reveals functional overlap with stat1 binding and interferon evasion. *Journal of virology*, 82(13):6259–6271, 2008.
- Madhuchhanda Mahapatra, Satya Parida, Berhe G Egziabher, Adama Diallo, and Tom Barrett. Sequence analysis of the phosphoprotein gene of peste des petits ruminants (ppr) virus: editing of the gene transcript. *Virus research*, 96(1-2):85–98, 2003.
- Roli Mandhana, Lily K Qian, and Curt M Horvath. Constitutively active mda5 proteins are inhibited by paramyxovirus v proteins. *Journal of Interferon & Cytokine Research*, 38(8): 319–332, 2018.

Manuscript has not been peer-reviewed

- Glenn A Marsh, Carol De Jong, Jennifer A Barr, Mary Tachedjian, Craig Smith, Deborah Middleton, Meng Yu, Shawn Todd, Adam J Foord, Volker Haring, et al. Cedar virus: a novel henipavirus isolated from australian bats. *PLoS pathogens*, 8(8):e1002836, 2012.
- Y Matsuoka, J Curran, T Pelet, D Kolakofsky, R Ray, and RW Compans. The p gene of human parainfluenza virus type 1 encodes p and c proteins but not a cysteine-rich v protein. *Journal of virology*, 65(6):3406–3410, 1991.
- Teshome Mebatsion, Stefan Versteegen, Leonarda TC De Vaan, Angela Römer-Oberdörfer, and Carla C Schrier. A recombinant newcastle disease virus with low-level v protein expression is immunogenic and lacks pathogenicity for chicken embryos. *Journal of virology*, 75(1):420–428, 2001.
- EL Millar, LJ Rennick, B Weissbrich, J Schneider-Schaulies, WP Duprex, and BK Rima. The phosphoprotein genes of measles viruses from subacute sclerosing panencephalitis cases encode functional as well as non-functional proteins and display reduced editing. *Virus research*, 211:29–37, 2016.
- Sigrid Milles, Malene Ringkjøbing Jensen, Carine Lazert, Serafima Guseva, Stefaniia Ivashchenko, Guillaume Communie, Damien Maurin, Denis Gerlier, Rob WH Ruigrok, and Martin Blackledge. An ultraweak interaction in the intrinsically disordered replication machinery is essential for measles virus function. *Science advances*, 4(8):eaat7778, 2018.
- Carina Motz, Kerstin Monika Schuhmann, Axel Kirchhofer, Manuela Moldt, Gregor Witte, Karl-Klaus Conzelmann, and Karl-Peter Hopfner. Paramyxovirus v proteins disrupt the fold of the rna sensor mda5 to inhibit antiviral signaling. *Science*, 339(6120):690–693, 2013.
- Baibaswata Nayak, Sachin Kumar, Peter L Collins, and Siba K Samal. Molecular characterization and complete genome sequence of avian paramyxovirus type 4 prototype strain duck/hong kong/d3/75. *Virology journal*, 5(1):124, 2008.
- Richard A Neher and Ulrich Gerland. Dynamics of force-induced dna slippage. *Physical review letters*, 93(19):198102, 2004.
- Machiko Nishio, Junpei Ohtsuka, Masato Tsurudome, Tetsuya Nosaka, and Daniel Kolakofsky. Human parainfluenza virus type 2 v protein inhibits genome replication by binding to the l protein: possible role in promoting viral fitness. *Journal of virology*, 82(13):6130–6138, 2008.
- Sarah L Noton and Rachel Fearn. Initiation and regulation of paramyxovirus transcription and replication. *Virology*, 479:545–554, 2015.
- Evgeny Nudler. Rna polymerase backtracking in gene regulation and genome instability. *Cell*, 149(7):1438–1445, 2012.
- Shinji Ohgimoto, Hisanori Bando, Mitsuo Kawano, Kousuke Okamoto, Kunio Kondo, Masato Tsurudome, Machiko Nishio, and Yasuhiko Ito. Sequence analysis of p gene of human parainfluenza type 2 virus: P and cysteine-rich proteins are translated by two mrnas that differ by two nontemplated g residues. *Virology*, 177(1):116–123, 1990.

Manuscript has not been peer-reviewed

Satoshi Ohnishi, Hironari Kamikubo, Masayoshi Onitsuka, Mikio Kataoka, and David Shortle. Conformational preference of polyglycine in solution to elongated structure. *Journal of the American Chemical Society*, 128(50):16338–16344, 2006.

K Ohta, H Goto, N Yumine, and M Nishio. Human parainfluenza virus type 2 v protein inhibits and antagonizes tetherin. *Journal of General Virology*, 97(3):561–570, 2016.

Keisuke Ohta, Yusuke Matsumoto, Morihiro Ito, and Machiko Nishio. Tetherin antagonism by v proteins is a common trait among the genus rubulavirus. *Medical microbiology and immunology*, 206(4):319–326, 2017.

Allan Olsper, Betty Y-W Chung, John F Atkins, John P Carr, and Andrew E Firth. Transcriptional slippage in the positive-sense rna virus family potyviridae. *EMBO reports*, 16(8):995–1004, 2015.

Heidi Palosaari, Jean-Patrick Parisien, Jason J Rodriguez, Christina M Ulane, and Curt M Horvath. Stat protein interference and suppression of cytokine signal transduction by measles virus v protein. *Journal of virology*, 77(13):7635–7644, 2003.

Jean-Patrick Parisien, Joe F Lau, Jason J Rodriguez, Christina M Ulane, and Curt M Horvath. Selective stat protein degradation induced by paramyxoviruses requires both stat1 and stat2 but is independent of alpha/beta interferon signal transduction. *Journal of virology*, 76(9):4190–4198, 2002.

Christopher L Parks, Susan E Witko, Cheryl Kotash, Shuo L Lin, Mohinder S Sidhu, and Stephen A Udem. Role of v protein rna binding in inhibition of measles virus minigenome replication. *Virology*, 348(1):96–106, 2006.

Reay G Paterson and Robert A Lamb. Rna editing by g-nucleotide insertion in mumps virus p-gene mrna transcripts. *Journal of virology*, 64(9):4137–4145, 1990.

Thierry Pelet, Joseph Curran, and Daniel Kolakofsky. The p gene of bovine parainfluenza virus 3 expresses all three reading frames from a single mrna editing site. *The EMBO journal*, 10(2):443–448, 1991.

Christophe Penno, Virag Sharma, Arthur Coakley, Mary O’Connell Motherway, Douwe van Sinderen, Lucyna Lubkowska, Maria L Kireeva, Mikhail Kashlev, Pavel V Baranov, and John F Atkins. Productive mrna stem loop-mediated transcriptional slippage: crucial features in common with intrinsic terminators. *Proceedings of the National Academy of Sciences*, 112(16):E1984–E1993, 2015.

Adrian Pickar, Pei Xu, Andrew Elson, Zhuo Li, James Zengel, and Biao He. Roles of serine and threonine residues of mumps virus p protein in viral transcription and replication. *Journal of virology*, 88(8):4414–4422, 2014.

C. R. Pringle. The order Mononegavirales. *Archives of Virology*, 117(1-2):137–140, March 1991. doi: 10.1007/bf01310499. URL <https://doi.org/10.1007/bf01310499>.

Mamta Puri, Ken Lemon, W Paul Duprex, Bertus K Rima, and Curt M Horvath. A point mutation, e95d, in the mumps virus v protein disengages stat3 targeting from stat1 targeting. *Journal of virology*, 83(13):6347–6356, 2009.

Manuscript has not been peer-reviewed

- X Qiu, Q Fu, C Meng, S Yu, Y Zhan, L Dong, T Ren, Y Sun, L Tan, C Song, et al. Kinetic analysis of rna editing of newcastle disease virus p gene in the early period of infection. *Acta virologica*, 60(1):71–77, 2016.
- A. Ramachandran and C. M. Horvath. Dissociation of paramyxovirus interferon evasion activities: Universal and virus-specific requirements for conserved v protein amino acids in MDA5 interference. *Journal of Virology*, 84(21):11152–11163, August 2010. doi: 10.1128/jvi.01375-10. URL <https://doi.org/10.1128/jvi.01375-10>.
- Aparna Ramachandran and Curt M Horvath. Paramyxovirus disruption of interferon signal transduction: Status report. *Journal of Interferon & Cytokine Research*, 29(9):531–537, 2009.
- Randall W Renshaw, Amy L Glaser, Hana Van Campen, Frank Weiland, and Edward J Dubovi. Identification and phylogenetic comparison of salem virus, a novel paramyxovirus of horses. *Virology*, 270(2):417–429, 2000.
- Bert Rima, Peter Collins, Andrew Easton, Ron Fouchier, Gael Kurath, Robert A Lamb, Benhur Lee, Andrea Maisner, Paul Rota, and Lin-Fa Wang. Problems of classification in the family paramyxoviridae. *Archives of virology*, 163(5):1395–1404, 2018.
- Kenny R Rodriguez and Curt M Horvath. Paramyxovirus v protein interaction with the antiviral sensor lgp2 disrupts mda5 signaling enhancement but is not relevant to lgp2-mediated rlr signaling inhibition. *Journal of virology*, 88(14):8180–8188, 2014.
- Jason P Roth, Joseph K-K Li, John D Morrey, Dale L Barnard, and Almut H Vollmer. Deletion of the d domain of the human parainfluenza virus type 3 (hpi3) pd protein results in decreased viral rna synthesis and beta interferon (ifn- β) expression. *Virus genes*, 47(1):10–19, 2013.
- Anne Röthlisberger, Dominique Wiener, Matthias Schweizer, Ernst Peterhans, Andreas Zurbriggen, and Philippe Plattet. Two domains of the v protein of virulent canine distemper virus selectively inhibit stat1 and stat2 nuclear import. *Journal of virology*, 84(13):6328–6343, 2010.
- Jason Saba, Xien Chua, Tatiana V Mishanina, Dhananjaya Nayak, Tricia A Windgassen, Rachel Anne Mooney, and Robert Landick. The elemental mechanism of transcriptional pausing. *bioRxiv*, page 422220, 2018.
- Niv Sabath, Andreas Wagner, and David Karlin. Evolution of viral proteins originated de novo by overprinting. *Molecular Biology and Evolution*, 29(12):3767–3780, July 2012. doi: 10.1093/molbev/mss179. URL <https://doi.org/10.1093/molbev/mss179>.
- Paramananda Saikia, M Gopinath, and MS Shaila. Phosphorylation status of the phosphoprotein p of rinderpest virus modulates transcription and replication of the genome. *Archives of virology*, 153(4):615–626, 2008.
- Edoardo Salladini, Vincent Delauzun, and Sonia Longhi. The henipavirus v protein is a prevalently unfolded protein with a zinc-finger domain involved in binding to ddb1. *Molecular BioSystems*, 13(11):2254–2267, 2017.

Manuscript has not been peer-reviewed

- Anthony Sanchez, Sam G Trappier, BW Mahy, Clarence J Peters, and Stuart T Nichol. The virion glycoproteins of ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proceedings of the National Academy of Sciences*, 93(8): 3602–3607, 1996.
- Maria T Sanchez-Aparicio, Leighland J Feinman, Adolfo García-Sastre, and Megan L Shaw. Paramyxovirus v proteins interact with the rig-i/trim25 regulatory complex and inhibit rig-i signaling. *Journal of Virology*, 92(6):e01960–17, 2018.
- Benjamin A Satterfield, Robert W Cross, Karla A Fenton, Krystle N Agans, Christopher F Basler, Thomas W Geisbert, and Chad E Mire. The immunomodulating v and w proteins of nipah virus determine disease course. *Nature communications*, 6:7483, 2015.
- Benjamin A Satterfield, Robert W Cross, Karla A Fenton, Viktoriya Borisevich, Krystle N Agans, Daniel J Deer, Jessica Graber, Christopher F Basler, Thomas W Geisbert, and Chad E Mire. Nipah virus c and w proteins contribute to respiratory disease in ferrets. *Journal of virology*, 90(14):6326–6343, 2016.
- Christian J Sauder, Vahan Simonyan, Laurie Ngo, Konstantinos Karagiannis, Yu Cong, Cheryl Zhang, Rong Wang, Wells W Wu, Tahir Malik, and Steven A Rubin. Evidence that a polyhexameric genome length is preferred, but not strictly required, for efficient mumps virus replication. *Virology*, 493:173–188, 2016.
- Anne Schaap-Nutt, Christopher D’Angelo, Margaret A Scull, Emerito Amaro-Carambot, Machiko Nishio, Raymond J Pickles, Peter L Collins, Brian R Murphy, and Alexander C Schmidt. Human parainfluenza virus type 2 v protein inhibits interferon production and signaling and is required for replication in non-human primates. *Virology*, 397(2):285–298, 2010.
- Kerstin M Schuhmann, Christian K Pfaller, and Karl-Klaus Conzelmann. The measles virus v protein binds to p65 (rela) to suppress nf- κ b activity. *Journal of virology*, 85(7): 3162–3171, 2011.
- Reed S Shabman, Omar J Jabado, Chad E Mire, Timothy B Stockwell, Megan Edwards, Milind Mahajan, Thomas W Geisbert, and Christopher F Basler. Deep sequencing identifies noncanonical editing of ebola and marburg virus rnas in infected cells. *MBio*, 5(6): e02011–14, 2014.
- Joshua W Shaevitz, Elio A Abbondanzieri, Robert Landick, and Steven M Block. Backtracking by single rna polymerase molecules observed at near-base-pair resolution. *Nature*, 426(6967):684–687, 2003.
- Jessica A Shaffer, William J Bellini, and Paul A Rota. The c protein of measles virus inhibits the type i interferon response. *Virology*, 315(2):389–397, 2003.
- Megan L Shaw, Washington B Cardenas, Dmitriy Zamarin, Peter Palese, and Christopher F Basler. Nuclear localization of the nipah virus w protein allows for inhibition of both virus-and toll-like receptor 3-triggered signaling pathways. *Journal of virology*, 79(10): 6078–6088, 2005.
- Christian JA Sigrist, Lorenzo Cerutti, Nicolas Hulo, Alexandre Gattiker, Laurent Falquet, Marco Pagni, Amos Bairoch, and Philipp Bucher. Prosite: a documented database using patterns and profiles as motif descriptors. *Briefings in bioinformatics*, 3(3):265–274, 2002.

Manuscript has not been peer-reviewed

- Mario H Skiadopoulos, Leatrice Vogel, Jeffrey M Riggs, Sonja R Surman, Peter L Collins, and Brian R Murphy. The genome length of human parainfluenza virus type 2 follows the rule of six, and recombinant viruses recovered from non-polyhexameric-length antigenomic cdnas contain a biased distribution of correcting mutations. *Journal of virology*, 77(1): 270–279, 2003.
- Gary M Skinner, Christoph G Baumann, Diana M Quinn, Justin E Molloy, and James G Hoggett. Promoter binding, initiation, and elongation by bacteriophage t7 rna polymerase a single-molecule view of the transcription cycle. *Journal of Biological Chemistry*, 279(5): 3239–3244, 2004.
- Katrina Sleeman, Bettina Bankamp, Kimberly B Hummel, Michael K Lo, William J Bellini, and Paul A Rota. The c, v and w proteins of nipah virus inhibit minigenome replication. *Journal of General Virology*, 89(5):1300–1308, 2008.
- Kate M Smith, Sofiya Tsimbalyuk, Megan R Edwards, Emily M Cross, Jyoti Batra, Tatiana P Soares da Costa, David Aragão, Christopher F Basler, and Jade K Forwood. Structural basis for importin alpha 3 specificity of w proteins in hendra and nipah viruses. *Nature communications*, 9(1):1–13, 2018.
- JA Southern, B Precious, and RE Randall. Two nontemplated nucleotide additions are required to generate the p mrna of parainfluenza virus type 2 since the rna genome encodes protein v. *Virology*, 177(1):388–390, 1990.
- Michael Steward, I Barry Vipond, Neil S Millar, and Peter T Emmerson. Rna editing in newcastle disease virus. *Journal of General Virology*, 74(12):2539–2547, 1993.
- George Streisinger, Yoshimi Okada, Joyce Emrich, Judith Newton, Akira Tsugita, Eric Terzaghi, and M Inouye. Frameshift mutations and the genetic code. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 31, pages 77–84. Cold Spring Harbor Laboratory Press, 1966.
- Akihiro Sugai, Hiroki Sato, Misako Yoneda, and Chieko Kai. Phosphorylation of measles virus phosphoprotein at s86 and/or s151 downregulates viral transcriptional activity. *FEBS letters*, 586(21):3900–3907, 2012.
- Dengyun Sun, Priya Luthra, Zhuo Li, and Biao He. Plk1 down-regulates parainfluenza virus 5 gene expression. *PLoS pathogens*, 5(7):e1000525, 2009.
- Minghao Sun, Sandra M Fuentes, Khalid Timani, Dengyun Sun, Chris Murphy, Yuan Lin, Avery August, Michael N Teng, and Biao He. Akt plays a critical role in replication of nonsegmented negative-stranded rna viruses. *Journal of virology*, 82(1):105–114, 2008.
- Yingjie Sun, Hang Zheng, Shengqing Yu, Yunlei Ding, Wei Wu, Xuming Mao, Ying Liao, Chunchun Meng, Zaib Ur Rehman, Lei Tan, et al. Newcastle disease virus v protein degrades mitochondrial antiviral signaling protein to inhibit host type i interferon production via e3 ubiquitin ligase rnf5. *Journal of virology*, 93(18):e00322–19, 2019.
- Kaoru Takeuchi, Kiyoshi Tanabayashi, Michiko Hishiyama, Yasuko K Yamada, Akio Yamada, and Akira Sugiura. Detection and characterization of mumps virus v protein. *Virology*, 178(1):247–253, 1990.

Manuscript has not been peer-reviewed

- Kaoru Takeuchi, Shin-ich Kadota, Makoto Takeda, Naoko Miyajima, and Kyosuke Nagata. Measles virus v protein blocks interferon (ifn)- α/β but not ifn- γ signaling by inhibiting stat1 and stat2 phosphorylation. *FEBS letters*, 545(2-3):177–182, 2003.
- Ivica Tamas, Jennifer J Wernegreen, Björn Nystedt, Seth N Kauppinen, Alistair C Darby, Laura Gomez-Valero, Daniel Lundin, Anthony M Poole, and Siv GE Andersson. Endosymbiont gene functions impaired and rescued by polymerase infidelity at poly (a) tracts. *Proceedings of the National Academy of Sciences*, 105(39):14934–14939, 2008.
- Nicolas Tarbouriech, Joseph Curran, Rob WH Ruigrok, and Wilhelm P Burmeister. Tetrameric coiled coil domain of sendai virus phosphoprotein. *Nature Structural & Molecular Biology*, 7(9):777, 2000.
- Sheila M Thomas, Robert A Lamb, and Reay G Paterson. Two mRNAs that differ by two non-templated nucleotides encode the amino coterminal proteins p and v of the paramyxovirus sv5. *Cell*, 54(6):891–902, 1988.
- Brigitte Tippin, Sawami Kobayashi, Jeffrey G Bertram, and Myron F Goodman. To slip or skip, visualizing frameshift mutation dynamics for error-prone dna polymerases. *Journal of Biological Chemistry*, 279(44):45360–45368, 2004.
- Hoang T Tran, Albert Mao, and Rohit V Pappu. Role of backbone- solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *Journal of the American Chemical Society*, 130(23):7380–7392, 2008.
- Shotaro Uchida, Ryo Horie, Hiroki Sato, Chieko Kai, and Misako Yoneda. Possible role of the nipah virus v protein in the regulation of the interferon beta induction by interacting with ubx domain-containing protein1. *Scientific reports*, 8(1):7682, 2018.
- Silvia Vidal, Joseph Curran, and Daniel Kolakofsky. Editing of the sendai virus p/c mRNA by g insertion occurs during mRNA synthesis via a virus-encoded activity. *Journal of virology*, 64(1):239–246, 1990a.
- Silvia Vidal, Joseph Curran, and Daniel Kolakofsky. A stuttering model for paramyxovirus p mRNA editing. *The EMBO journal*, 9(6):2017, 1990b.
- I. M. Wallace. M-coffee: combining multiple sequence alignment methods with t-coffee. *Nucleic Acids Research*, 34(6):1692–1699, March 2006. doi: 10.1093/nar/gkl091. URL <https://doi.org/10.1093/nar/gkl091>.
- Melissa N Webby, Matthew P Sullivan, Kavestri M Yegambaram, Mazdak Radjainia, Jeremy R Keown, and Richard L Kingston. A method for analyzing the composition of viral nucleoprotein complexes, produced by heterologous expression in bacteria. *Virology*, 527:159–168, 2019.
- Greg Wells and Achut Malur. Expression of human parainfluenza virus type 3 pd protein and intracellular localization in virus infected cells. *Virus Genes*, 37(3):358–367, 2008.
- S. P. J. Whelan, J. N. Barr, and G. W. Wertz. Transcription and replication of nonsegmented negative-strand RNA viruses. In *Current Topics in Microbiology and Immunology*, pages 61–119. Springer Berlin Heidelberg, 2004. doi: 10.1007/978-3-662-06099-5_3. URL https://doi.org/10.1007/978-3-662-06099-5_3.

Manuscript has not been peer-reviewed

- Elizabeth B Wignall-Fleming, David J Hughes, Sreenu Vattipally, Sejal Modha, Steve Goodbourn, Andrew J Davison, and Richard E Randall. Analysis of paramyxovirus transcription and replication by high-throughput sequencing. *Journal of virology*, pages JVI-00571, 2019.
- Susan E Witko, Cheryl Kotash, Mohinderjit S Sidhu, Stephen A Udem, and Christopher L Parks. Inhibition of measles virus minireplicon-encoded reporter gene expression by v protein. *Virology*, 348(1):107–119, 2006.
- Patrick CY Woo, Susanna KP Lau, Beatrice HL Wong, Rachel YY Fan, Annette YP Wong, Anna JX Zhang, Ying Wu, Garnet KY Choi, Kenneth SM Li, Janet Hui, et al. Feline morbillivirus, a previously undescribed paramyxovirus associated with tubulointerstitial nephritis in domestic cats. *Proceedings of the National Academy of Sciences*, 109(14):5435–5440, 2012.
- Patrick CY Woo, Susanna KP Lau, Paolo Martelli, Suk-Wai Hui, Candy CY Lau, Rachel YY Fan, Joseph M Groff, Emily WT Tam, Kwok-Hung Chan, and Kwok-Yung Yuen. Fatal systemic necrotizing infections associated with a novel paramyxovirus, anaconda paramyxovirus, in green anaconda juveniles. *Journal of clinical microbiology*, 52(10):3614–3623, 2014.
- Sarah A Woodson and Donald M Crothers. Proton nuclear magnetic resonance studies on bulge-containing dna oligonucleotides from a mutational hot-spot sequence. *Biochemistry*, 26(3):904–912, 1987.
- Filip Yabukarski, Philip Lawrence, Nicolas Tarbouriech, Jean-Marie Bourhis, Elise Delaforge, Malene Ringkjøbing Jensen, Rob W H Ruigrok, Martin Blackledge, Viktor Volchkov, and Marc Jamin. Structure of nipah virus unassembled nucleoprotein in complex with its viral chaperone. *Nature Structural & Molecular Biology*, 21(9):754–759, August 2014. doi: 10.1038/nsmb.2868. URL <https://doi.org/10.1038/nsmb.2868>.
- Yang Yang, James Zengel, Minghao Sun, Katrina Sleeman, Khalid Amine Timani, Jason Aligo, Paul Rota, Jianguo Wu, and Biao He. Regulation of viral rna synthesis by the v protein of parainfluenza virus 5. *Journal of virology*, 89(23):11845–11857, 2015.
- Yanling Yang, Yawen Bu, Jing Zhao, Jia Xue, Gang Xu, Yang Song, Ye Zhao, Huiming Yang, and Guozhong Zhang. Appropriate amount of w protein of avian avulavirus 1 benefits viral replication and w shows strain-dependent subcellular localization. *Virology*, 538:71–85, 2019.
- WS Yarnell and JW Roberts. Mechanism of intrinsic transcription termination and antitermination. *Science*, 284(5414):611–615, 1999.
- Kavestri Yegambaram, Esther MM Bulloch, and Richard L Kingston. Protein domain definition should allow for conditional disorder. *Protein Science*, 22(11):1502–1518, 2013.
- Dan F Young, Elizabeth B Wignall-Fleming, David C Busse, Matthew J Pickin, Jack Hankinson, Elizabeth M Randall, Amy Tavendale, Andrew J Davison, Douglas Lamont, John S Tregoning, et al. The switch between acute and persistent paramyxovirus infection caused by single amino acid substitutions in the rna polymerase p subunit. *PLoS pathogens*, 15(2):e1007561, 2019.