Molecular Plant Pathology

# *Cassava brown streak virus* Ham1 protein hydrolyses mutagenic nucleotides and is a necrosis determinant

| | |
|---|---|
| Journal: | *Molecular Plant Pathology* |
| Manuscript ID | MPP-OA-18-442.R2 |
| Manuscript Type: | Original Article |
| Date Submitted by the Author: | 14-Apr-2019 |
| Complete List of Authors: | Tomlinson, Katie; University of Bristol, School of Biological Sciences<br>Pablo-Rodriguez, José ; University of Bristol, School of Biological Sciences ; CINVESTAV Campus Irapuato<br>Bunawan, Hamidun; University of Bristol, School of Biological Sciences; Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia<br>Nanyiti, Sarah; Univeristy of Bristol, School of Biological Sciences; National Crops Resources Research Institute<br>Green, Patrick; Univeristy of Bristol, School of Biological Sciences<br>Miller, Josie; Univeristy of Bristol, School of Biological Sciences<br>Seal, Sue; NRI<br>Bailey, Andy; Univeristy of Bristol, School of Biological Sciences;<br>Foster, Gary; School of Biological Sciences |
| Keywords: | Virus, Cassava, Cassava brown streak virus, Ugandan cassava brown streak virus, Cassava brown streak disease, Food security |
| | |

# *Cassava brown streak virus* Ham1 protein hydrolyses mutagenic nucleotides and is a necrosis determinant

Authors: Katie R. Tomlinson[1*], José Luis Pablo-Rodriguez[1,2*], Hamidun Bunawan[1,3], Sarah Nanyiti[1,4], Patrick Green[1], Josie Miller[1], Susan E. Seal[5], Andy M. Bailey[1] and Gary D. Foster[1†].

1. School of Biological Sciences, University of Bristol, Life Sciences Building, 24 Tyndall Avenue Bristol, BS8 1TQ, UK.
2. CINVESTAV Campus Irapuato, Mexico.
3. Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor Darul Ehsan, Malaysia.
4. National Crops Resources Research Institute (NaCRRI), P.O. Box 7084, Kampala, Uganda.
5. Natural Resources Institute, Chatham Maritime, Kent ME4 4TB, UK.

†Corresponding Author: Prof. Gary Foster (gary.foster@bristol.ac.uk; +44 (0) 117 39 41178; ORCiD: 0000-0002-1979-1397).

* First two authors contributed equally to the work, and can be considered joint first authors.

Running head: CBSV Ham1 hydrolyses mutagens

Key words: Virus, Cassava, Cassava brown streak virus, Ugandan cassava brown streak virus, Cassava brown streak disease, Food security



**Summary**

**Cassava brown streak disease (CBSD) is a leading cause of cassava losses in East and Central Africa and is currently having a severe impact on food security. The disease is caused by two viruses within the *Potyviridae* family: *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV), which both encode atypical Ham1 proteins with highly conserved ITP pyrophosphohydrolase (ITPase) domains. ITPase proteins are widely encoded by plant, animal and archaea. They selectively hydrolyse mutagenic nucleotide triphosphates (NTP) to prevent their incorporation into nucleic acid and thereby function to reduce mutation rates. It has previously been hypothesised that U/CBSVs encode Ham1 proteins with ITPase activity to reduce viral mutation rates during infection. In this study, we investigate the potential roles of U/CBSV Ham1 proteins. We show that both CBSV and UCBSV Ham1 proteins have ITPase activities through *in vitro* enzyme assays. Deep-sequencing experiments found no evidence of the U/CBSV Ham1 proteins providing mutagenic protection during infections of *Nicotiana* hosts. Manipulations of the CBSV Tanza infectious clone were performed, including a Ham1 deletion, ITPase point mutations and UCBSV Ham1 chimera. Unlike severely necrotic wild-type CBSV Tanza infections, infections of *Nicotiana benthamiana* with the manipulated CBSV ICs do not develop necrosis, indicating that that the CBSV Ham1 is a necrosis determinant. We propose that the presence of U/CBSV Ham1 proteins with highly conserved ITPase motifs indicates that they serve highly selectable functions during infections of cassava and may represent a Euphorbia host adaptation that could be targeted in anti-viral strategies.**

**Introduction**

Cassava (*Manihot esculenta* Crantz, family *Euphorbiaceae*) is the second most important food crop in terms of *per capita* calories consumed in Africa (Nweke, 2004). It can withstand unpredictable rainfall and grows with minimal inputs on marginal land, so provides vital adaptation opportunities to projected climate change (Jarvis et al., 2012). Unfortunately, cassava production in sub-Saharan Africa (SSA) is currently limited by two viral diseases: cassava mosaic disease (CMD) and cassava brown streak disease (CBSD), which have severe impacts on food and economic security across the entire region (Patil et al., 2015; Tomlinson et al., 2017). CBSD is caused by at least two viral species: *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV), collectively termed U/CBSVs (Mbanzibwa et al., 2011; Monger et al., 2001a, 2001b; Winter et al., 2010; Amisse et al., 2019). In cassava, U/CBSVs cause foliar chlorosis and root necrosis and can cause yield losses of up to 100% in susceptible varieties (Kaweesi et al., 2014). Despite the importance of U/CBSVs, relatively little is understood about their fundamental biology and gene functions.

U/CBSVs belong to the *Ipomovirus* genus of the *Potyviridae* family; both viral species share several unusual genome features in that they encode a P1 protein, lack a HC-Pro protein and encode Maf/Ham1 like proteins (Mbanzibwa et al., 2009). Maf/Ham1 proteins belong to the ITP pyrophosphohydrolase (ITPase) protein family which in prokaryotes, eukaryotes and archaea, function to specifically hydrolyse non-canonical, mutagenic nucleotide triphosphates (NTP) and thereby reduce mutation rates (Simone et al., 2013; Waisertreiger et al., 2010; Zamzami et al., 2013). The non-canonical NTPs xanthine and inosine triphosphate (XTP/ITP) are formed as by-products of purine NTP biosynthesis or through oxidative deamination of canonical purine NTPs (Simone et al., 2013). If incorporated into nucleic acid, XTP and ITP can cause deleterious mispairing mutations, nucleic acid strand breaks and recombinations (Budke and Kuzminov, 2010; Burgis et al., 2003; Simone et al., 2013). ITPase proteins hydrolyse the pyrophosphate bonds in XTP and ITP to release xanthine and inosine monophosphates (XMP/IMP) respectively and a pyrophosphate (PPi) molecule. This prevents XTP/ITP incorporation into nucleic acid and thereby reduces mutation rates. U/CBSV Ham1 proteins contain the highly conserved Serine-Histidine-Arginine (SHR) signature motif, which in the human ITPase is involved with substrate binding and specificity (Gall et al., 2013; Stenmark et al., 2007).

*Potyviridae* genomes are comprised of single-stranded, positive sense RNA which are replicated by viral RNA dependant RNA polymerase (RdRp). Compared with DNA polymerases, RdRps have relatively low fidelity and lack proof-reading (Andino and Domingo, 2015; Duffy et al., 2008). This results in populations of closely related viral genome variants referred to as quasi-species that can rapidly adapt to changes in host environments (Andino and Domingo, 2015; Holmes, 2003). It has been reported that RNA viral mutation rates typically exist close to a critical error threshold, above which error catastrophe occurs whereby lethal numbers of mutations accumulate during each round of viral replication that cause a dramatic loss of viral viability (Crotty et al., 2001; Holmes, 2003). RNA viruses must therefore balance maximizing sequence diversity to enhance adaptability whilst conserving key sequences required for multiple essential functions. Proteins with conserved Maf/Ham1

domains are not commonly found in viral genomes, and have only been reported in two other viral species that also infect Euphorbia hosts: *Euphorbia ringspot virus* (EuRV: *Potyvirus* genus, *Potyviridae* family) and *Cassava torrado-like virus* (CsTLV: *Torrado* genus, *Secoviridae* family (Knierim et al., 2016; Jiménez Polo, 2018). Previously, it has been hypothesised that Ham1-like proteins in U/CBSV genomes may function to reduce viral mutation rates below a critical error threshold during infection to retain viral viability over long-term infections (Mbanzibwa et al., 2009). However, this function for U/CBSV Ham1 proteins is currently uncharacterised.

In addition to gene function, the U/CBSV genome sequence determinants that are associated with necrosis development during infection are unknown. Typically compared with UCBSV, CBSV causes more severe root necrosis in cassava and accumulates to higher titres (Kaweesi et al., 2014; Ogwok et al., 2014; Winter et al., 2010). In the model host *N. benthamiana*, CBSV tends to cause extremely severe systemic necrosis that results in plant death, whereas UCBSV tends to cause milder mosaic symptoms (Mohammed et al., 2012; Winter et al., 2010). UCBSV and CBSV Ham1 proteins typically share only 47% amino acid identity, suggesting that this region may be associated with differences in symptom development and viral accumulation during infection (Mbanzibwa et al., 2011; Winter et al., 2010).

In this study, we performed *in vitro* enzyme assays to investigate the ability of CBSV and UCBSV Ham1 proteins to selectively hydrolyse non-canonical, mutagenic NTPs. We then performed three sets of experiments to test the ability of U/CBSV Ham1 proteins to provide mutagenic protection. Firstly, 5-fluorouracil resistance assays were performed in yeast overexpressing the U/CBSV Ham1 proteins. Secondly, transgenic *N. tabacum* lines overexpressing the CBSV Ham1 sequence were infected *Potato virus Y* (PVY) and *Tobacco mosaic virus* (TMV) and viral genome diversity was measured using deep-sequencing. Thirdly, the recently constructed CBSV Tanza infectious clone (IC) (Duff-Farrier et al., 2018) was used to construct the CBSV_HKO IC, containing a Ham1 deletion. Deep-sequencing was then performed to compare viral genome diversity during CBSV_HKO with wild-type CBSV infections of *N. benthamiana*. Finally, we investigated the association of the Ham1 region with symptom development and viral accumulation. Our findings shed light on potential U/CBSV Ham1 protein functions that will be useful for further studies into the fundamental biology of these devastating viruses that could lead to novel anti-viral strategies.



**Results**

**ITPase motifs are highly conserved in CBSV and UCBSV Ham1 sequences**

To compare the Ham1 proteins, the CBSV_Tanza and UCBSV_Kikombe Ham1 amino acid sequences were aligned with 7 CBSV and 7 UCBSV isolates from the NCBI database. This demonstrated that both the CBSV Tanza and UCBSV Kikombe Ham1 proteins are 226 amino acids in length, located between the NIb and coat protein (CP) regions. Differences were found in the proteolytic cleavage sequences at the N' of CBSV and UCBSV Ham1 proteins. Whereas CBSV Ham1 proteins are predicted to be cleaved from the NIb peptide at the cleavage sequence: I-D-L-Q-/V, for UCBSV Ham1 proteins the sequence is: V-D-T-Q-/T, which suggests differential polyprotein processing of the two viral species (Fig. S1). The CBSV_Tanza and UCBSV_Kikombe Ham1 sequences were found to share 46% amino acid identity, which is comparable to identities of CBSV and UCBSV Ham1

proteins previously reported (Mbanzibwa et al., 2009; Winter et al., 2010). To investigate the level of ITPase conservation in U/CBSV Ham1 proteins, 12 CBSV and 20 UCBSV Ham1 sequences were aligned. Conserved ITPase motifs were identified in the Ham1 sequences in all isolates at amino acid positions: Ser192, His193, Arg194 (Fig. S2). The ExPASy ProtParam tool was used to predict the molecular weights for the Ham1 proteins according to their amino acid sequences. The CBSV_Tanza and UCBSV_Kikombe Ham1 proteins are predicted to be 25.5 and 25.1 KDa in size respectively.

**U/CBSV Ham1 proteins have *in vitro* ITPase activities**

To determine whether the CBSV_Tanza and UCBSV_Kikombe Ham1 proteins have ITPase activity, protein expression and *in vitro* enzyme assays were performed. Histidine tagged Ham1 proteins were expressed in *Escherichia coli* and purified using nickel chromatography columns. Purification of the Ham1 proteins was confirmed by SDS-PAGE analysis, which demonstrated that homogenous proteins of approximately 25 kDa had been purified, corresponding to the expected size of the CBSV and UCBSV Ham1 proteins (Fig. S3). The concentrations of the purified CBSV and UCBSV Ham1 proteins were estimated to be 0.5 and 0.3 mg/ml respectively, according to Bradford assay. To detect ITPase activity, enzyme assays were performed whereby 1.3 µg of the CBSV_Tanza or UCBSV_Kikombe Ham1 proteins were incubated with 0.2mM non-canonical NTPs: XTP and dITP, and eight canonical ribose and deoxyribose-NTPs in coupled reactions with 0.1 units of yeast inorganic pyrophosphatase (Thermo Fisher Scientific), according to Lin et al., (2001). Phosphate (Pi) concentration was measured using the PiColourLock kit (Innova Biosciences) recording $OD_{655}$ absorbance values in an iMark Microplate Absorbance Reader microtiter plate reader (Bio-Rad) with a 655 nm filter. Phosphate concentrations were calculated from $OD_{655}$ nm readings using linear regression of known phosphate concentrations.



Pi concentrations were negligible in negative control reactions whereby: 1) no Ham1 protein was added, 2) bovine serum albumin control protein was added or 3) no yeast inorganic pyrophosphatase was added. Both the CBSV and UCBSV Ham1 proteins showed significantly higher activity with XTP and dITP, followed by dGTP and GTP (Fig. 1). We found that the CBSV Ham1 protein has significantly higher activities with the non-canonical NTPs: XTP and dITP compared to canonical NTPs: UTP, dTTP, dATP, dCTP, CTP and ATP ($p$<0.05) and the UCBSV Ham1 protein has significantly higher activities with XTP and dITP compared to dTTP, dATP, dCTP, and ATP ($p$<0.05) (Table S1-S2). To test whether CBSV Ham1 ITPase activity could be heat inactivated, the protein was heated to 95°C for 10 mins and then added to reactions with dITP. Phosphate concentrations were 97% lower in reactions with heat inactivated protein, indicating that heat denaturation abolishes U/CBSV Ham1 ITPase activity (Fig. S4).

**Overexpression of U/CBSV Ham1 proteins in yeast does not provide resistance to 5-fluorouracil**

To test whether CBSV_Tanza and UCBSV_Kikombe Ham1 proteins can protect against mutagenic NTPs *in vivo*, resistance assays were performed in *Saccharomyces cerevisi*ae (yeast). Carlsson et al., (2013) have previously shown that overexpression of the yeast Ham1 gene in yeast protects against the exogenous supply of the non-canonical, mutagenic pyrimidine nucleoside 5-fluorouracil (5-FU). The yeast Ham1 gene can protect against several non-canonical nucleosides including 6-N-hydroxylaminopurine (Kozmin et al., 1998; Noskov et al., 1996),

5-fluorocytosine and 6-azauracil (Carlsson et al., 2013). To investigate whether U/CBSV Ham1 proteins can also provide protection from 5-FU, resistance assays were performed. The Ham1 sequences from CBSV Nampula (NCBI: MG019915), CBSV Tanza (NCBI: MG570022), UCBSV Kikombe (NCBI: KX753356.1) and yeast (NCBI: 853532) were cloned into pYES2 and transformed into the wild-type yeast BY4742 strain. Transformant yeast were cultured in liquid Yeast Synthetic Dropout Media (YSDM: 1.7 g/l yeast nitrogen base, 5 g/l ammonium sulphate, 0.77 g/l yeast uracil dropout mix (Sigma-Aldrich)), supplemented with 2% galactose to induce Ham1 overexpression and harvested at early log-phase. Transformant yeast were then plated onto YSDM agar plates, supplemented with 2% galactose and 10 µg/ml 5-FU, after 72 hours colonies were assessed for growth (Fig. S5). Liquid growth assays were also performed whereby yeast growth was compared in YSDM liquid media supplemented with 2% galactose and 10 µg/ml 5-FU (Fig. S6). In both plate and liquid assays, yeast transformed with U/CBSV Ham1 sequences showed low levels of growth, comparable to the negative control (empty pYES2), whereas transformants with the yeast Ham1 gene showed relatively high levels of growth. This suggests that unlike the yeast Ham1, the U/CBSV Ham1 proteins are unable to provide protection from mutagenic 5-FU and so may not function to reduce mutation rates.

**Effects of CBSV Ham1 on viral mutation rates during infections**

Experiments were then performed to test whether the ITPase activities of U/CBSV Ham1 proteins results in lower viral mutation rates during infections. Wild-type *N. tabacum* and transgenic *N. tabacum* lines expressing CBSV Nampula Ham1 were infected with TMV and PVY and deep-sequencing was performed on viral amplicons to compare the diversity of viral genomes. To generate transgenic lines, *N. tabacum* was transformed with an expression vector containing the CBSV Nampula Ham1 sequence (NCBI: MG019915). Expression of the transgene was confirmed in three separate lines by qPCR (Table S5). Wild-type and transgenic *N. tabacum* plants were inoculated with TMV (genus: *Tobamovirus*) and PVY (genus: *Potyvirus*). At 14 days post inoculation (dpi), RT-PCR was performed targeting two 500 bp regions of each viral genome. For TMV, the 126 kDa replicase and 54 kDa replicase regions were targeted and for PVY, the NIb and CP were targeted. Next generation sequencing (NGS) was performed on these amplicons, sequence reads were aligned to the reference sequences with a read depth of 100, 000 reads. The LoFreq algorithm was used to detect viral variants which occur at a low frequency (Wilm et al., 2012). No differences were found in the number of single nucleotide variants (SNVs) in TMV and PVY RT-PCR amplicons from Ham1-transgenic and wild-type *N. tabacum* plants (Fig. 2A). This indicates that CBSV Ham1 protein does not reduce TMV or PVY mutation rates in *N. tabacum*.

To investigate in greater depth whether the CBSV Ham1 protein specifically reduces the CBSV mutation rate, the complete Ham1 sequence was deleted from the CBSV_Tanza IC (Duff-Farrier et al., 2019) to produce the CBSV Tanza Ham1 knockout IC (CBSV_ HKO). *N. benthamiana* plants were then infected with either CBSV_Tanza or CBSV_HKO. At 10 dpi, RT-PCR was performed targeting 1200 bp of CBSV coat protein sequence. As before NGS and LoFreq analysis was performed on aligned reads. To determine whether the samples contained different viral titers, qPCR was performed which demonstrated that viral transcript abundance was higher in the wild-type CBSV_Tanza infections than in CBSV_HKO infections (Fig. 2B). Therefore, to account for SNVs in viral amplicons that are due to higher viral titers rather than changes to mutation rates, viral titers were included in

analyses. A one-way ANCOVA test found no significant difference in the number of SNVs in sequencing reads from wild-type CBSV_Tanza and CBSV_HKO, whilst accounting for variance in viral titers (Fig. 2C), $F(1, 0.008) = 1.423$, $p = 0.286$. This indicates that the CBSV Ham1 does not reduce CBSV mutation rate in *N. benthamiana*.

**Association of U/CBSV Ham1 proteins with necrosis development and viral accumulation**

*N. benthamiana* infected with CBSV_HKO developed dramatically different symptoms, compared with the IC of wild-type CBSV_Tanza (Fig. 3). Whilst wild-type CBSV_Tanza infections of *N. benthamiana* developed severe necrosis and plant death at 10 – 14 dpi, CBSV_HKO infections did not develop necrosis but instead developed strong systemic leaf curling, chlorotic mottling and stunting at 10 – 18 dpi (Fig. 3). Therefore this suggests that the CBSV Ham1 is not essential for infection and is associated with necrosis development during infection of *N. benthamiana*.

To further investigate the association of the CBSV_Tanza Ham1 protein with necrosis development, additional modifications to the CBSV_Tanza IC were performed. Firstly, to test whether the development of necrosis during CBSV_Tanza infections of *N. benthamiana* is specifically associated with the ITPase motifs within the CBSV Ham1 sequence, the CBSV_mutHam IC was constructed whereby the positively charged SH(+)R(+)motif was mutated to neutrally charged Serine-Alanine-Alanine (SAA). This mutation is predicted to alter electrostatic interactions of the CBSV Ham1 protein and thereby affect function. Gall et al., (2013) previously demonstrated that mutating the SHR motif to SAA in the human ITPase protein abolishes ITP substrate specificity or activity. Secondly, to also test whether the CBSV Tanza and UCBSV Kikombe Ham1 proteins are associated with differential symptom development during *N. benthamiana* infections, the chimeric CBSV_UHam IC was constructed, consisting of the CBSV Tanza genome with a UCBSV Kikombe Ham1 sequence replacement (Fig. S7). *N. benthamiana* plants were then agroinfiltrated with the modified ICs: CBSV_HKO, CBSV_mutHam and CBSV_UHam so that symptoms and viral titers could be compared with wild-type IC CBSV_Tanza infections. *N. benthamiana* infected with the different ICs developed a range of symptoms as illustrated in Fig. 3. As already seen with CBSV_HKO, infections with CBSV_mutHam and CBSV_UHam also lack necrosis and develop leaf curling, chlorotic mottling and stunting at 12 – 18 dpi. This further suggests that the CBSV Ham1 protein is associated with necrosis development during infection of *N. benthamiana*. To identify the effect of the manipulations on viral accumulation, qPCR was performed throughout infection. This revealed that CBSV_Tanza accumulates to higher titers during early infection (7 dpi) compared with CBSV_HKO, CBSV_mutHam and CBSV_UHam, suggesting that Ham1 may be involved with early viral accumulation (Fig. 4). Later in infection (10 – 18 dpi), CBSV_UHam accumulates to levels similar to CBSV_Tanza, suggesting potential complementation of the UCBSV Ham1 protein with CBSV. Titers remained low throughout CBSV_mutHam infections, demonstrating that mutating the SHR motif in the viral Ham1 protein has a large effect on viral accumulation. CBSV_HKO titers accumulated to high levels at 18 dpi, indicating that the Ham1 protein is not required for high levels of viral replication later in infection of *N. benthamiana*.

**Discussion**

This study has demonstrated that CBSV and UCBSV Ham1 proteins have ITPase activities with non-canonical, mutagenic NTPs. While this activity was detected *in vitro*, the *in vivo* functions of these proteins remain unclear. The ability of the CBSV IC lacking a Ham1 (CBSV_HKO) to infect and accumulate to high levels in *N. benthamiana* is surprising as RNA viruses typically have highly streamlined genomes that encode a small number of proteins with multiple, indispensable functions (Holmes, 2003). Genome regions which do not serve advantageous functions tend to be rapidly lost from viral genomes during infections. This is demonstrated by the rapid deletion of marker-gene sequences during infections with tagged viral ICs (Arazi et al., 2001; Beauchemin et al., 2005; Dawson et al., 1989; Guo et al., 1998). Willemsen et al., (2017) have also demonstrated that exogenous sequences are only maintained in *Tobacco etch virus* (TEV) genomes during infection if they provide advantageous functions for TEV. As all reported U/CBSV genome sequences encode Ham1 proteins with highly conserved ITPase motifs, they must serve advantageous functions during infection and/or transmission. We suggest that although the CBSV Ham1 protein was dispensable for CBSV infection of *N. benthamiana* in this study, further studies are required to determine whether U/CBSV Ham1 proteins are required specifically during cassava infections and/or vector transmission.

This is the first report that CBSV and UCBSV Ham1 proteins have ITPase activities and that these are significantly higher with non-canonical compared with canonical NTPs. The U/CBSV Ham1 ITPase activities are similar to those of previously characterised ITPase proteins from the archaea thermophile *Methanococcus jannaschii* (Cho et al., 1999), human (Lin et al., 2001) and *E. coli* (Burgis et al., 2003; Zheng et al., 2005). Compared with these proteins, U/CBSV Ham1 proteins have higher activities with the canonical nucleotide: (d)GTP. The reasons for this may include lower selective pressure on U/CBSV Ham1s to maintain binding cleft structures that are unfavourable to (d)GTP binding, or disruption of host processes, such as G-protein defence signalling (Brenya et al., 2016; Trusov and Botella, 2016), host DNA replication and expression or membrane trafficking (Inada and Ueda, 2014), which requires further investigation.

Despite finding that U/CBSV Ham1 proteins have ITPase activities with non-canonical NTPs *in vitro*, we found no evidence of mutagenic protection in 5-FU resistance assays in yeast. This result is difficult to explain but perhaps indicates that the viral ITPase does not target the same subcellular location as the native yeast Ham1. We were also unable to detect any significant decreases in viral genome variation during PVY and TMV infections of CBSV Ham1 transgenic *N. tabacum* lines or increases in CBSV genome variation during CBSV_HKO infections of *N. benthamiana*. These preliminary results indicate that U/CBSV Ham1 proteins do not lower viral mutation rates during infections of *Nicotiana* hosts. However the functions of U/CBSV Ham1 proteins may be complex and depend on other factors, including host and environment. Studies have found that different host environments can affect viral mutation rates, for instance *Cucumber mosaic virus* replicates with a lower fidelity during infections of pepper compared with tobacco (Pita et al., 2007). Therefore U/CBSV Ham1 proteins may have a bigger effect on viral mutation rates specifically during replication in cassava. U/CBSV Ham1 proteins may also play important functions during replication in wild hosts that can be found in close proximity to CBSD affected cassava fields (Amisse et al., 2018). If viral Ham1 sequences are found to serve as a Euphorbia or wild host adaptation, they could be targeted in anti-viral strategies such as post-transcriptional gene silencing.

In this study, CBSV IC manipulations demonstrated that the CBSV Ham1 is associated with the development of necrosis in *N. benthamiana*. Infections with CBSV ICs containing a deletion of the Ham1 sequence (CBSV_HKO), mutated SHR ITPase motif (CBSV_mutHam) and a UCBSV Ham1 chimera (CBSV_UHam) did not develop necrosis in *N. benthamiana*, but instead developed systemic leaf curling, chlorotic mottling and stunting. The CBSV Ham1 ITPase activity may be directly associated with necrosis development, however as *in vitro* enzyme assays demonstrated that the CBSV and UCBSV Ham1 proteins have similar ITPase activities and the CBSV_UHam infections lack necrosis, there may be alternative explanations. For instance the folded CBSV Ham1 protein may be recognised by the host which triggers severe, systemic necrosis. Examples of this in other systems includes recognition of the *Soybean mosaic virus* P3 protein by the host Rsv1 protein that leads to the induction of lethal systemic hyper sensitive response (Hajimorad et al., 2005). It should be noted that although the CBSV Ham1 protein is associated with the development of necrosis in *N. benthamiana*, this may be different to root necrosis development in cassava. We suggest that the ICs developed in this study should be used to investigate whether the CBSV Ham1 region is associated with root necrosis during cassava infections.

In addition to necrosis in *N. benthamiana*, the CBSV Ham1 protein also appears to be involved with viral accumulation during early infection. QPCR analysis of CBSV transcript abundance indicated that during early infection CBSV_Tanza accumulated to higher levels than CBSV_KO. However later in infection, CBSV_HKO increased to much higher levels than CBSV_Tanza, which is likely to be due to the lack of necrosis that limits wild-type CBSV_Tanza accumulation. Interestingly, CBSV_UHam transcripts accumulated to levels comparable to wild-type CBSV_Tanza, which may indicate that the UCBSV Ham1 is able to at least partially compensate for the absence of the CBSV Ham1. The potential synergistic and/or antagonist interactions between CBSV and UCBSV proteins are currently poorly understood, despite the frequent occurrence of mixed infections in the field (Kathurima et al., 2016; Mbanzibwa et al., 2011; Ogwok et al., 2014).

In line with previous research, we found that, outside of the conserved ITPase motifs, CBSV and UCBSV Ham1 proteins share relatively low amino acid sequence identity. They have different N' proteolytic cleavage sequences and so may form different intermediate fusion proteins with different specific functions (Adams et al., 2005), which requires further investigation.

To conclude, we have demonstrated that UCBSV and CBSV Ham1 proteins selectively hydrolyse non-canonical, mutagenic NTPs *in vitro*. We have also shown that the CBSV Tanza Ham1 protein is associated with the development of necrosis during infections of *N. benthamiana*. We found no evidence that U/CBSV Ham1 proteins function to reduce viral mutation rates during infections of experimental *Nicotiana* hosts. Further studies are needed to determine the functions of U/CBSV Ham1 proteins during infections of cassava and whether they serve as Euphorbia host adaptations. Information on the specific functions of U/CBSV Ham1 proteins during cassava infections could lead to novel anti-viral strategies.

**Experimental Procedures**

**U/CBSV Ham1 sequence alignments**

U/CBSV Ham1 amino acid sequences were downloaded from the NCBI database. Sequences were aligned using the T-coffee multiple sequence alignment package (http://tcoffee.crg.cat/apps/tcoffee/index.html), according to Notredame et al., (2000). Alignments were formatted using the Expasy Boxshade tool (https://embnet.vital-it.ch/software/BOX_form.html).

**U/CBSV Ham1 protein expression and purification**

The CBSV Tanza and UCBSV Kikombe Ham1 sequences were amplified from the CBSV_Tanza (NCBI: MG570022) and UCBSV_Kikombe (NCBI: KX753356) ICs respectively by PCR using the high-fidelity Phusion polymerase (Thermo Fisher Scientific). Primers were designed to include 21 bp of overlapping sequence with the POPINF plasmid vector (OPPF-UK). Primers used to amplify the CBSV_Tanza and UCBSV_Kikombe Ham1 sequences are provided in Table S3. The PCR products encoding the Ham1 sequences were cloned into the POPINF expression vector (OPPF-UK) using the InFusion cloning kit (Takara Bio USA, Inc.). The POPINF vector contains the 6X histidine tag, which is fused to the 3' of insert sequences. Plasmids were transformed into electro-competent TOP10 *E. coli* and blue/white screening was used to select colonies containing plasmids with inserts. Transformant colonies were cultured and plasmids were extracted using the GeneJet Plasmid Miniprep Kit (ThermoFisher Scientific). Plasmids were Sanger sequenced to confirm the presence of Ham1 sequences without mutation in both POPINF plasmids.

For expression, POPINF plasmids were transformed into the *Escherichia coli* strain: BL21 - DE3 (New England Biolabs). *E. coli* cultures were incubated in Luria Broth media containing 100 µg/ml carbenicillin and incubated at 37°C and 200 rpm until the $OD_{600}$ was approximately 0.4. Protein expression was induced through the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and cultures were incubated for a further 18 hours. The cultures were then centrifuged at 6,000 *g* for 30 mins at 4°C; cell pellets were resuspended in the resuspension buffer – pH 7.5 (150 mM NaCl, 20 mM Tris-HCl - pH 7.5) and centrifuged as before. The cell pellets were then resuspended in loading buffer – pH 7.5 (0.5 M NaCl, 20 mM Tris-base, 20 mM imidazole, 10% glycerol). Samples were homogenised and lysed using a french pressure cell press (Constant Systems). The lysate was centrifuged at 18,000 *g* for 40 mins at 4°C and the supernatant was filtered using a 0.22 µm syringe filter (Millipore). The lysate was loaded onto a HisTrap FF Crude histidine-tagged protein purification column (GE Healthcare). A Fast-Protein-Liquid Chromatography (ÄKTA) machine was used to set up an imidazole concentration gradient of 50% loading to elution buffer – pH 7.5 (0.5 M NaCl, 20 mM Tris-base, 1 M imidazole, 10% glycerol). Fractions corresponding to a peak in UV absorbance were analysed by SDS-PAGE to confirm the presence of a protein band with the expected size. Fractions containing protein were pooled dialysed using a SnakeSkin 10 kDa dialysis tube (Thermo Fisher Scientific) into the storage buffer – pH 7.5 (0.5 M NaCl, 50 mM Tris-HCl - pH 7.5, 1 mM dithiothreitol (DTT), 50 mM $MgCl_2$, 20% glycerol). To increase the concentration of the protein, the sample was transferred to a 15 ml Vivaspin column (GE Healthcare) with a molecular weight cut-off of 10 KDa. *In vitro* ITPase enzyme assays were performed as described in the 'Results' section. Nucleotide triphosphates were obtained from Thermo Fisher Scientific (dGTP, dCTP, dTTP, dATP, dITP), Invitrogen (GTP, CTP, ATP, UTP) and Sigma Aldrich (ITP and XTP).

**5-fluorouracil resistance assays in yeast**

The *S. cerevisiae* (yeast) Ham1 gene (NCBI: 853532) was amplified by PCR from the wild-type yeast strain BY4742/ Y10000 (Euroscarf). The U/CBSV Ham1 sequences were amplified by PCR from the CBSV_Tanza, CBSV Nampula and UCBSV_Kikombe ICs respectively, using the high-fidelity Phusion polymerase (Thermo Fisher Scientific). The Ham1 sequences were cloned into pYES2 plasmids, downstream of the inducible GAL1 promoter. Plasmids were Sanger sequenced to confirm the presence of Ham1 sequences without mutation. The plasmids were transformed into the wild-type yeast strain BY4742 (Y10000). Colony PCR confirmed the presence of the Ham1 genes in transformant yeast. Plate and liquid growth assays were performed as described in the 'Results' section.

**Reverse transcription PCR**

RNA was extracted from *Nicotiana* plants using the E.Z.N.A. Plant RNA kit (Omega Bio-Tek). Samples were DNAse I treated (Thermo Fisher Scientific), reverse transcription was performed 1 µg of RNA using the First Strand cDNA synthesis kit (Thermo Fisher Scientific) and the oligo d(T)18 primer. First-strand cDNA was used as template in PCR reactions using high-fidelity Phusion polymerase (Thermo Fisher Scientific) and primers to target specific regions of the viral genomes.

**Quantitative real-time PCR**

Quantitative real-time PCR (qPCR) was performed to detect expression level of CBSV Nampula Ham1 sequence in transgenic *N. tabacum*. Primers to detect CBSV Nampula Ham1 sequence were: qPCR_CHam1_Fw: CGAGTAGCTGCTGAACTTGTTGGAG/qPCR_CHam1_Rv: GATATGGCTCCACCAACTTGTATAG. The *Protein phosphatase 2A* (PP2A) (NCBI: TC21939/At1g13320) was used as an endogenous reference gene and targeted using primers designed and validated in Liu et al., (2012). QPCR was also performed to calculate relative viral transcript abundance in infected *N. benthamiana* material. Primers to detect the CBSV Tanza coat protein were: qPCR_CBSV_CP_Fw: ACTTCCTAGCCGAAGCACAA/ qPCR_CBSV_CP_Rv: GCACTAACATCCCGCGTAGT. The *N. benthamiana* F-box gene (NCBI: 24993647/At5g15710) was used as an endogenous reference gene and targeted using primers designed and validated in Liu et al., (2012). All primers were validated for amplification efficiency using a 1:10 serial dilution of template; only primers with an $R^2$ >0.99 were used. Inspection of the qPCR amplification peaks confirmed single amplification peaks. Reactions were set up using Maxima SYBR Green/ROX (Thermo Fisher Scientific) following the manufacturer's instructions and using a Stratagene MX3005 thermocycler. The following programme cycle was used: initial denaturation at 95°C for 10 mins, followed by 40 cycles of (denaturation at 95°C for 15 secs, primer annealing at 60°C for 30 secs and extension at 72°C for 30 secs). Data was acquired during the extension phase. The MxPro software calculated the threshold Ct values. Relative target gene expression in the three replicate plants were then calibrated with three control samples, using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

**Deep-sequencing of viral amplicons**

Viral amplicons were purified using the GeneJET gel extraction kit (Thermo Fisher Scientific) and used to prepare libraries according to the TruSeq Nano DNA Library Prep protocol (Illumina). Amplicons were fragmented and cleaned using the AmPure XP 0.8 clean-up procedure and two indexed adapters were ligated to paired ends of the fragments. PCR amplification was then performed to enrich for fragments containing both paired-end adaptors. Library sizes and purity was then validated using a DNA 1000 ScreenTape. Subsequently the samples were diluted to 10 pM with 25% PhiX spike-in and run in a MiSeq instrument using the MiSeq Reagent Kit v3 (2X 75 bp). FASTQ files generated from the MiSeq run were uploaded onto the Partek Flow server. The Illumina adapter sequences and low-quality reads with phred scores of less than 30 were then trimmed. Reads were aligned in Partek Flow to reference sequences using Bowtie 2-2.2.5, very-sensitive-local pre-sets. The LoFreq algorithm was used to detect low-frequency variants (<0.5%) from sequencing errors (Wilm et al., 2012).

**Infectious clone manipulations**

Infected plant material and viral infectious clones (IC) were used under the DEFRA license No. 51045/197610/2 and handled according to Brewer et al., (2018). IC manipulations were performed through homologous yeast recombination, according to the protocol outlined in Duff-Farrier et al., (2015, 2019). Briefly, the CBSV_Tanza IC (NCBI: MG570022) was linearized with the restriction enzyme *Bam* H1 (Thermo Fisher Scientific). Insert fragments were designed to contain 30 bp of homologous sequences at the 5' and 3' ends to enable recombination of overlapping sequences into linearized CBSV_Tanza IC. Insert fragments were amplified by PCR using the high-fidelity Phusion polymerase (Thermo Fisher Scientific). The conserved SHR motif was mutated through site-directed mutagenesis with primers encoding the SAA mutation. The UCBSV Ham1 sequence was amplified from the UCBSV_Kikombe IC (NCBI: KX753356) and the CBSV NIb – Ham1 proteolytic cleavage sequence was maintained according to the schematic in Fig. S7. All primers used in the construction of ICs are listed in Table S4. The *S. cerevisiae* YPH499 strain was transformed with linearized CBSV_Tanza IC and PCR insert fragments at a ratio of 4:1, according to the Gietz et al., (2002) protocol. Transformant cells were plated onto YSDM agar plates and incubated at 28°C for 48 hours. Plasmids were extracted from transformant yeast using the Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research, USA) and electroporated into the *E. coli* strain TOP10 (Thermo Fisher Scientific). Transformant *E. coli* colonies were cultured in LB broth containing 50 µg/ml kanamycin and plasmids were extracted using GeneJet Plasmid Miniprep Kit (ThermoFisher Scientific). Recombinant plasmids were analysed by restriction digest, PCR and Sanger sequencing to confirm correct construction.

**Plant cultivation**

Wild-type *N. tabacum* and *N. benthamiana* plants were cultivated from seed produced in-house and planted in Sinclair all-purpose potting compost. All plants were grown in a controlled growth cabinets at 28°C with a 16h/8h: light/dark cycle.

**Generation of transgenic plants**

CBSV_Ham1 transgenic *N. tabacum* lines were produced according the Gallois and Marinho, (1995) method. Briefly the CBSV_Nampula Ham1 sequence was cloned into a pCambia 2300 expression vector, containing a *Cauliflower mosaic virus* 35S promoter and tNOS terminator. Electrocompetent *Agrobacterium tumefaciens* strain: LBA4404 were transformed with the expression vector and grown in 20 ml LB media, supplemented with 50 µg/ml kanamycin and 20 µg/ml rifampicin. Liquid culture was diluted with Murashige and Skoog media (MSO) until $OD_{600}$ = 1. Fully expanded leaves of *N. tabacum* were collected 5 weeks after germination and sterilised by washing in a solution of bleach at 10 % v/v for 10 minutes and rinsed with sterile distilled water for 15 minutes. Rinsing was performed five times to eliminate residues of bleach from the surface of the leaves. Once leaves were sterilized, they were cut into small discs of 1 cm in diameter and immersed into agrobacterium MSO culture for 20 minutes. After inoculation, discs were placed on MSO shooting agar plates. Later plates were placed at 25°C in a photoperiod of 16 hours light and 8 hours dark for two days in order to adapt to the media and start to develop shoots. After the period of adaptation, disc leaves were removed from the first plate and placed on selective MSO shooting media supplemented with 50 µg/ml kanamycin to select transformed plants and 200 µg/ml timentin to eliminate agrobacteria. Leaf discs and callus cultures were transferred onto fresh media every three weeks. Plantlet shoots were excised and transferred to MSO rooting agar medium supplemented 50 µg/ml kanamycin and 200 µg/ml timentin. Once roots were established and leaves were fully developed, transgenic lines were assessed for the presence of the transgenic transcript by qPCR. Verified transgenic *N. tabacum* plants were micro-propagated by cutting stems with one to two nodes and placed into MSO rooting media.

**Viral infections**

*N. tabacum* and *N. benthamiana* plants were infected with CBSV_Tanza ICs through agroinfiltration, according to Voinnet et al., (2003). Electrocompetent agrobacterium strain LBA4044 were transformed with the CBSV_Tanza IC plasmids. Transformant colonies were cultured in LB broth containing kanamycin 50 µg/ml and rifampicin 20 µg/ml. After 48 hours the starter culture was used to inoculate LB broth containing kanamycin 50 µg/ml, rifampicin 20 µg/ml, 10 mM Morpholino Ethane Sulfonic acid buffer (MES) and 150 µM of acetosyringone, which was cultured for 20 hours. The cells were pelleted, washed and resuspended in infiltration buffer (10 mM $MgCl_2$; 10 mM MES and 150 µM acetosyringone) to an $OD_{600}$ of 1.0. The cells were left at room temperature for 5 hours. Plants were agroinfiltrated at four weeks; the abaxial surface of the youngest fully expanded leaf was infiltrated with 1 ml of agrobacterium suspension. *N. tabacum* plants were mechanically inoculated with PVY and TMV infected leaf material that was ground in liquid nitrogen and suspended in sterile deionised water. The adaxial surface of the youngest fully expanded leaf was dusted with 600 mesh carborundum powder (Fisher Scientific) and grounded tissue gently applied.

**Acknowledgements**

KRT was supported by BBSRC SWDTP programme, JLPR was funded by the Consejo Nacional de Ciencia y Tecnología (CONACyT) México, SN was funded by the Ugandan National Agricultural Research Organisation, East Africa Agricultural Productivity Project. Work was performed under DEFRA Plant Health License number 51045/197610/2. Deep-sequencing reactions were run by the University of Bristol Genomics facility.

**References**


Adams, M.J., Antoniw, J.F., and Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. Mol. Plant Pathol. *6*, 471–487.

Aguilar, E., Almendral, D., Allende, L., Pacheco, R., Chung, B.N., Canto, T., and Tenllado, F. (2015). The P25 protein of potato virus X (PVX) is the main pathogenicity determinant responsible for systemic necrosis in PVX-associated synergisms. J. Virol. *89*, 2090–2103.

Amisse, J.J.G., Ndunguru, J., Tairo, F., Boykin, L.M., Kehoe, M.A., Cossa, N., Ateka, E., Rey, C., and Sseruwagi, P. (2018). First report of Cassava brown streak viruses on wild plant species in Mozambique. Physiol. Mol. Plant Pathol. *105*, 88-95.

Amisse, J.J.G., Ndunguru J., Tairo, F., Ateka, E., Boykin, L.M., Kehoe, M.A., Cossa, N., Rey, C., SSeruwagi, P. (2019) Analyses of seven new whole genome sequences of cassava brown streak viruses in Mozambique reveals two distinct clades: evidence for new species. Plant Pathol. pre-print.

Andino, R., and Domingo, E. (2015). Viral quasispecies. Virology *479–480*, 46–51.

Arazi, T., Slutsky, S.G., Shiboleth, Y.M., Wang, Y., Rubinstein, M., Barak, S., Yang, J., and Gal-On, A. (2001). Engineering zucchini yellow mosaic potyvirus as a non-pathogenic vector for expression of heterologous proteins in cucurbits. J. Biotechnol. *87*, 67–82.

Beauchemin, C., Bougie, V., and Laliberté, J.-F. (2005). Simultaneous production of two foreign proteins from a polyvirus-based vector. Virus Res. *112*, 1–8.

van den Born, E., Omelchenko, M. V, Bekkelund, A., Leihne, V., Koonin, E. V, Dolja, V. V, and Falnes, P.Ø. (2008). Viral AlkB proteins repair RNA damage by oxidative demethylation. Nucleic Acids Res. *36*, 5451–5461.

Bratlie, M.S., and Drabløs, F. (2005). Bioinformatic mapping of AlkB homology domains in viruses. BMC Genomics *6*, 1.

Brenya, E., Trusov, Y., Dietzgen, R.G., and Botella, J.R. (2016). Heterotrimeric G-proteins facilitate resistance to plant pathogenic viruses in Arabidopsis thaliana (L.) Heynh. Plant Signal. Behav. *11*, e1212798.

Brewer, H.C., Hird, D.L., Bailey, A.M., Seal, S.E., and Foster, G.D. (2018). A guide to the contained use of plant virus infectious clones. Plant Biotechnol. J. *16*, 832–843.

Budke, B., and Kuzminov, A. (2010). Production of clastogenic DNA precursors by the nucleotide metabolism in *Escherichia coli*. Mol. Microbiol. *75*, 230–245.

Burgis, N.E., Brucker, J.J., and Cunningham, R.P. (2003). Repair system for noncanonical purines in Escherichia coli. J. Bacteriol. *185*, 3101–3110.

Carlsson, M., Gustavsson, M., Hu, G.-Z., Murén, E., and Ronne, H. (2013). A Ham1p-Dependent Mechanism and Modulation of the Pyrimidine Biosynthetic Pathway Can Both Confer Resistance to 5-Fluorouracil in Yeast. PLoS One *8*, e52094.

Cho, Y., Hwang, K.Y., Chung, J.H., Kim, S.-H., and Han, Y.S. (1999). Structure-based identification of a novel NTPase from Methanococcus jannaschii. Nat. Struct. Biol. *6*, 691–696.

Crotty, S., Cameron, C.E., and Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. Proc. Natl. Acad. Sci. U. S. A. *98*, 6895–6900.

Dawson, W.O., Lewandowski, D.J., Hilf, M.E., Bubrick, P., Raffo, A.J., Shaw, J.J., Grantham, G.L., and Desjardins, P.R. (1989). A tobacco mosaic virus-hybrid expresses and loses an added gene. Virology *172*, 285–292.

Duff-Farrier, C.R.A., Bailey, A.M., Boonham, N., and Foster, G.D. (2015). A pathogenicity determinant maps to the N-terminal coat protein region of the Pepino mosaic virus genome. Mol. Plant Pathol. *16*, 308–315.

Duff-Farrier, C.R.A., Mbanzibwa, D.R., Nanyiti, S., Bunawan, H., Pablo-Rodriguez, J.L., Tomlinson, K.R., James, A.M., Alicai, T., Seal, S.E., Bailey, A.M., et al. (2019). Strategies for the construction of cassava brown streak disease viral infectious clones. Mol. Biotechnol. *61*, 93-101.

Duffy, S., Shackelton, L.A., and Holmes, E.C. (2008). Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. *9*, 267–276.

Gall, A.D., Gall, A., Moore, A.C., Aune, M.K., Heid, S., Mori, A., and Burgis, N.E. (2013). Analysis of human ITPase nucleobase specificity by site-directed mutagenesis. Biochimie *95*, 1711–1721.

Gallois, P., and Marinho, P. (1995). Leaf Disk Transformation Using Agrobacterium tumefaciens-Expression of Heterologous Genes in Tobacco. In Plant Gene Transfer and Expression Protocols, (New Jersey: Humana Press), pp. 39–48.

Gietz, RD, Woods, and RA (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. *350*, 87–96.

Guo, H.S., López-Moya, J.J., and García, J.A. (1998). Susceptibility to recombination rearrangements of a chimeric plum pox potyvirus genome after insertion of a foreign gene. Virus Res. *57*, 183–195.

Hajimorad, M.R., Eggenberger, A.L., and Hill, J.H. (2005). Loss and gain of elicitor function of soybean mosaic virus G7 provoking Rsv1-mediated lethal systemic hypersensitive response maps to P3. J. Virol. *79*, 1215–1222.

Holmes, E.C. (2003). Error thresholds and the constraints to RNA virus evolution. Trends Microbiol. *11*, 543–546.

Inada, N., and Ueda, T. (2014). Membrane trafficking pathways and their roles in plant–microbe interactions. Plant Cell Physiol. *55*, 672–686.

Jarvis, A., Ramirez-Villegas, J., Herrera Campo, B.V., and Navarro-Racines, C. (2012). Is cassava the answer to African climate change adaptation? Trop. Plant Biol. *5*, 9–29.

Jiménez Polo, J. (2018). Identification of a torradovirus-encoded protein that complements the systemic movement of a potexvirus lacking the TGB3 gene. International Congress of Plant Pathology (ICPP) 2018: Plant Health in A Global Economy, (Boston).

Kathurima, T., Nyende, A., Kiarie, S., and Ateka, E. (2016). Genetic diversity and distribution of cassava brown streak virus and ugandan cassava brown streak virus in major cassava-growing regions in Kenya. Annu. Res. Rev. Biol. *10*, 1–9.

Kaweesi, T., Kawuki, R., Kyaligonza, V., Baguma, Y., Tusiime, G., Ferguson, M.E., Pennisi, E., Storey, H., Nichols, R., Legg, J., et al. (2014). Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. Virol. J. *11*, 216.

Knierim, D., Menzel, W., and Winter, S. (2016). Analysis of the complete genome sequence of Euphorbia ringspot virus, an atypical member of the genus Potyvirus. Arch. Virol. 1–3.

Kozmin, S., Leroy, P., and Pavbvle, Y. (1998). Overexpression of the yeast HAM1 gene prevents 6-JV-hydrosylaminopurine mutagenesis in Escherichia coli. *45*, 645-652

Lin, S., McLennan, A.G., Ying, K., Wang, Z., Gu, S., Jin, H., Wu, C., Liu, W., Yuan, Y., Tang, R., et al. (2001). Cloning, expression, and characterization of a human inosine triphosphate pyrophosphatase encoded by the itpa gene. J. Biol. Chem. *276*, 18695–18701.

Liu, D., Shi, L., Han, C., Yu, J., Li, D., and Zhang, Y. (2012). Validation of reference genes for gene expression studies in virus-infected Nicotiana benthamiana using quantitative real-time PCR. PLoS One *7*, e46451.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods *25*, 402–408.

Mbanzibwa, D.R., Tian, Y., Mukasa, S.B., and Valkonen, J.P.T. (2009). Cassava brown streak virus (Potyviridae) encodes a putative Maf/HAM1 pyrophosphatase implicated in reduction of mutations and a P1 proteinase that suppresses RNA silencing but contains no HC-Pro. J. Virol. *83*, 6934–6940.

Mbanzibwa, D.R., Tian, Y.P., Tugume, A.K., Mukasa, S.B., Tairo, F., Kyamanywa, S., Kullaya, A., and Valkonen, J.P.T. (2011a). Simultaneous virus-specific detection of the two Cassava brown streak-associated viruses by RT-PCR reveals wide distribution in East Africa, mixed infections, and infections in Manihot glaziovii. J. Virol. Methods *171*, 394–400.

Mbanzibwa, D.R., Tian, Y.P., Tugume, A.K., Patil, B.L., Yadav, J.S., Bagewadi, B., Abarshi, M.M., Alicai, T., Changadeya, W., Mkumbira, J., et al. (2011b). Evolution of cassava brown streak disease-associated viruses. J. Gen. Virol. *92*, 974–987.

Mohammed, I.U., Abarshi, M.M., Muli, B., Hillocks, R.J., and Maruthi, M.N. (2012). The symptom and genetic diversity of Cassava brown streak viruses infecting cassava in East Africa. Adv. Virol. *2012*, 795697.

Monger, W.A., Seal, S., Cotton, S., and Foster, G.D. (2001a). Identification of different isolates of cassava brown streak virus and development of a diagnostic test. Plant Pathol. *50*, 768–775.

Monger, W.A., Seal, S., Isaac, A.M., and Foster, G.D. (2001b). Molecular characterization of the cassava brown streak virus coat protein. Plant Pathol. *50*, 527–534.

Monger, W.A., Alicai, T., Ndunguru, J., Kinyua, Z.M., Potts, M., Reeder, R.H., Miano, D.W., Adams, I.P., Boonham, N., Glover, R.H., et al. (2010). The complete genome sequence of the Tanzanian strain of cassava brown streak virus and comparison with the Ugandan strain sequence. Arch. Virol. *155*, 429–433.

Noskov, V.N., Staak, K., Shcherbakova, P. V., Kozmin, S.G., Negishi, K., Ono, B.-C., Hayatsu, H., and Pavlov, Y.I. (1996). HAM1, the gene controlling 6-N-hydroxylaminopurine sensitivity and mutagenesis in the yeast Saccharomyces cerevisiae. Yeast *12*, 17–29.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. *302*, 205–217.

Nweke, F. (2004). New Challenges in the cassava transformation in Nigeria and Ghana. Int. Food Policy Res. *118*.

Ogwok, E., Alicai, T., Rey, M.E.C., Beyene, G., and Taylor, N. (2014). Distribution and accumulation of cassava brown streak viruses within infected cassava (Manihot esculenta Crantz) plants. Plant Pathol. *64*, 12343.

Patil, B.L., Legg, J.P., Kanju, E., and Fauquet, C.M. (2015). Cassava brown streak disease: a threat to food security in Africa. J. Gen. Virol. *96*, 956–968.

Pita, J.S., de Miranda, J.R., Schneider, W.L., and Roossinck, M.J. (2007). Environment determines fidelity for an RNA virus replicase. J. Virol. *81*, 9072–9077.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. J. Virol. *84*, 9733–9748.

Simone, P.D., Pavlov, Y.I., and Borgstahl, G.E.O. (2013). ITPA (inosine triphosphate pyrophosphatase): From surveillance of nucleotide pools to human disease and pharmacogenetics. Mutat. Res. Mutat. Res. *753*, 131–

146.

Snyder, L., Gold, L., and Kutter, E. (1976). c. Proc. Natl. Acad. Sci. U. S. A. *73*, 3098–3102.

Stenmark, På., Kursula, P., Flodin, S., Gräslund, S., Landry, R., Nordlund, P., and Schüler, H. (2007). Crystal structure of human inosine triphosphatase. J. Biol. Chem. *282*, 3182–3187.

Sztuba-Solińska, J., Urbanowicz, A., Figlerowicz, M., and Bujarski, J.J. (2011). RNA-RNA Recombination in Plant Virus Replication and Evolution. Annu. Rev. Phytopathol. *49*, 415–443.

Tomlinson, K.R., Bailey, A.M., Alicai, T., Seal, S., and Foster, G.D. (2017). Cassava brown streak disease: historical timeline, current knowledge and future prospects. Mol. Plant Pathol. *19*, 1282-1294.

Trusov, Y., and Botella, J.R. (2016). Plant G-Proteins come of age: breaking the bond with animal models. Front. Chem. *4*, 24.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. *33*, 949–956.

Waisertreiger, I.S.-R., Menezes, M.R., Randazzo, J., and Pavlov, Y.I. (2010). Elevated levels of DNA strand breaks induced by a base analog in the human cell line with the P32T ITPA variant. J. Nucleic Acids *2010*.

Willemsen, A., Zwart, M.P., Ambrós, S., Carrasco, J.L., and Elena, S.F. (2017). 2b or not 2b: experimental evolution of functional exogenous sequences in a plant RNA Virus. Genome Biol. Evol. *9*, 297–310.

Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., and Nagarajan, N. (2012). LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. *40*, 11189–11201.

Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M., and Butgereitt, A. (2010). Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. J. Gen. Virol. *91*, 1365–1372.

Zamzami, M.A., Duley, J.A., Price, G.R., Venter, D.J., Yarham, J.W., Taylor, R.W., Catley, L.P., Florin, T.H., Marinaki, A.M., and Bowling, F. (2013). Inosine Triphosphate Pyrophosphohydrolase (ITPA) polymorphic sequence variants in adult hematological malignancy patients and possible association with mitochondrial DNA defects. J. Hematol. Oncol. *6*, 24.

Zheng, J., Singh, V.K., and Jia, Z. (2005). Identification of an ITPase/XTPase in Escherichia coli by structural and biochemical Analysis. Structure *13*, 1511–1520.



**Figure legends**

Figure 1: CBSV Tanza and UCBSV Kikombe Ham1 proteins have higher ITPase activities with non-canonical, mutagenic nucleotide triphosphates (NTP): XTP and ITP, compared with canonical NTPs. Purified CBSV Tanza and UCBSV Kikombe Ham1 proteins (1.3 µg) were incubated with 0.2 mM substrates at 37°C for 20 min in 50 mM Tris-HCl, pH 8.5, 1 mM DTT, 50 mM MgCl2, according to Lin et al., (2001). Activity was measured colourimetrically according to the 'Experimental Procedures'. One-way Games Howell ANOVA analyses demonstrated that the CBSV Ham1 protein has significantly higher activities with the non-canonical nucleotide triphosphates XTP and dITP compared to canonical nucleotide triphosphates UTP, dTTP, dATP dCTP, CTP and ATP (*= $p < 0.05$) and that the UCBSV Ham1 protein has significantly higher activities with XTP and dITP compared to dTTP, dATP dCTP, and

ATP (*= p<0.05). Results are provided in Tables S1 – S2. Each result is the mean phosphate concentration from three separate experiments (n = 3) ± S.E.

Figure 2: A) Results from deep-sequencing experiment indicate that the transgenic expression of CBSV Ham1 in *Nicotiana tabacum* does not significantly affect PVY or TMV genome diversity. The number of single nucleotide variations (SNVs) per million nucleotides of aligned sequence reads from PVY and TMV infections of wild-type (WT) *N. tabacum* and three transgenic *N. tabacum* lines expressing the CBSV Nampula Ham1 sequence (Ham 1.1, Ham 1.2 and Ham 1.3) are shown . B) QPCR of relative CBSV transcript abundance in CBSV_Tanza infections (C61 – 4) and CBSV_HKO infections (KO1 – 4) used in deep-sequencing experiment. C) Results from deep-sequencing experiment indicate that the deletion of CBSV Ham1 does not significant increase CBSV genome diversity. The number of SNVs per million nucleotides of aligned sequenced reads from CBSV_Tanza infections (C61 -4) and CBSV_HKO infections (KO1 – 4) are shown.

Figure 3 . CBSV Ham1 is associated with the development of necrosis during *Nicotiana benthamiana* infection. Symptom development during *N. benthamiana* infections with CBSV_Tanza ICs at 18 dpi. Wild-type (WT) CBSV_Tanza infections develop severe systemic necrosis, chlorosis, stunting and plant death. Whereas infections with the Ham1 knockout IC: CBSV_HKO, UCBSV Ham1 replacement: CBSV_UHam and mutated SHR motif: CBSV_mutHam develop systemic leaf curling, chlorotic mottling and stunting at 10 – 18 dpi, compared with non-inoculated plants. This suggests that the CBSV Ham1 protein is not essential for infection and is associated with necrosis development during *N. benthamiana* infection. Scale bar = 5 cm.

Figure 4: CBSV Ham1 is involved with viral accumulation during *Nicotiana benthamiana* infections. *N. benthamiana* was agroinfiltrated with the following ICs: wild-type CBSV_Tanza, Ham1 deletion (CBSV_HKO), mutated Ham1 ITPase SHR motif (CBSV_mutHam) and the chimera IC containing a UCBSV Ham1 swap (CBSV_UHam). QPCR was performed at 7, 10, 14 and 18 days post inoculation (dpi). Each result is the mean from three replicate plants (n = 3) ± S.E. Results demonstrate that CBSV_Tanza accumulates to higher titers during early infection (7 dpi) compared with CBSV_HKO, CBSV_mutHam and CBSV_UHam, suggesting that Ham1 may be involved with early viral accumulation. Results were consistent in three separate experiments.

**Supporting information legends**

Figure S1: Relevant sections of a T-coffee alignment of 8 CBSV and 8 UCBSV Ham1 amino acid sequences showing differences in proteolytic cleavage sequences. The junctions where the NIb and CP proteins end are highlighted in yellow. N' Ham1 proteolytic cleavage sequences between the NIb – Ham1 proteins are different in CBSV isolates (green) and UCBSV isolates (pink), indicating potential differences in proteolytic processing. Whereas the C' Ham1 proteolytic cleavage sequence between the Ham1 – CP proteins is conserved in CBSV and UCBSV sequences (blue). The cleave sites where the are shown in yellow Highly conserved regions (>90%) are highlighted in black. Sequences were obtained from the NCBI database; accession numbers are provided for each sequence.

Figure S2: Relevant section of T-coffee alignment of 12 CBSV and 20 UCBSV Ham1 amino acid sequences. The ITPase signature Serine-Histidine-Arginine (SHR) motif is highly conserved and was found in all sequences at positions 192 – 194 (yellow). Highly conserved regions (>90%) are highlighted in black. Sequences were obtained from the NCBI database; accession numbers are provided for each sequence.

Figure S3: SDS-PAGE of purified CBSV_Tanza (gel A) and UCBSV_Kikombe (gel B) Ham1 proteins (25 kDa). Lanes in gel A correspond separate fractions B4 – B13 that were eluted from the AKTA machine during protein purification at a range of imidazole concentrations. Lane D in gel B refers to UCBSV Ham1 protein which had been dialysed into the storage buffer. To prepare the protein samples for loading, 10 µL of loading buffer (4% SDS, 0.25 M Tris-HCl - pH 6.8, 20% glycerol, 0.004% bromophenol blue, 10% β-mercaptoethanol), 1 µL of protein sample and 9 µL of water were mixed and heated at 95°C for 5 mins. A TruPAGE Precast Gel (Sigma Aldrich) was loaded with 10 µL of each prepared sample and 10 µL of PageRuler Protein Ladder (Thermo Fisher Scientific). The gel was run at 220V for 40 mins. The gel was stained with 20 ml InstantBlue Protein Stain (Sigma Aldrich) and analysed under white light using the ChemDoc Bio-Rad System and images were taken using the Quantity One 1D software (Bio-Rad).

Figure S4: Enzyme assay results from the heat inactivation experiment to test for loss of CBSV Tanza Ham1 ITPase activity. Incubation of 0.2 mM dITP with active CBSV Tanza Ham1 protein (1.3 µg) resulted in a phosphate concentration of 136 µM. Heating the CBSV Tanza Ham1 at 95°C for 10 mins – 1 hour resulted in a 41 – 43% reduction in phosphate concentration, indicating inactivation of its pyrophosphohydrolase activity. Control assays were set where BSA protein (1.3 µg) was added, which produced a comparable phosphate concentration to assays where CBSV Tanza Ham1 had been heat inactivated, indicating that BSA could be used as a control for addition of protein to assay samples. Low background phosphate concentrations were found in negative controls: 1) containing 0.2 mM dITP in reaction buffer with the addition of 0.1 units of yeast inorganic pyrophosphatase, 2) 0.2 mM dITP in reaction buffer and 3) water.



Figure S5: 5-FU resistance agar plate growth assays. The wild-type yeast strain BY4742 was transformed with pYES2 plasmids containing Ham1 sequences from CBSV Nampula, CBSV Tanza, UCBSV Kikombe and yeast. Transformant yeast were cultured and plated onto SD media as ten-fold serial dilutions onto test plates containing 2% galactose and 10 µg/ml 5-FU or control plates containing 2% galactose only. Colony growth was imaged after 72 hours. Results were consistent in three separate experiments.

Figure S6: 5-FU resistance liquid growth assays. The wild-type yeast strain BY4742 was transformed with pYES2 plasmids containing Ham1 sequences from CBSV Nampula, CBSV Tanza, UCBSV Kikombe and yeast. Transformant yeast were cultured liquid SD media containing 2% galactose and 10 µg/ml 5-FU. The cell density ($OD_{600}$) was taken at 4, 8, 12, 24, 48 and 72 hours. Yeast transformed with yeast Ham1 sequence demonstrated relatively high levels of growth. Whereas as yeast transformed with U/CBSV Ham1 sequences showed low levels of growth, comparable to the negative control (empty pYES2 plasmid). This indicates that unlike the yeast Ham1 sequence, the U/CBSV Ham1 sequences were unable to protect against mutagenic 5-FU. Each result is the mean $OD_{600}$ value from three replicate samples (n = 3) ± S.E. Results were consistent in three separate experiments.

Figure S7: Schematic showing the replacement of CBSV Tanza Ham1 sequence (blue) with UCBSV Ham1 sequence (red) in the CBSV_UHam1 IC. To ensure proteolytic cleavage of UCBSV Ham1 sequence from the CBSV Tanza polyprotein the NIb – Ham1 protease cleavage sequence I-D-L-Q-V was maintained at the start of the UCBSV Ham1 sequence: T-K-D.

Table S1: Games-Howell one-way ANOVA tests to compare mean phosphate concentration in enzyme assay reactions with CBSV Tanza Ham1 incubated with the non-canonical nucleotides XTP and dITP and a range of canonical nucleotides.

Table S2: Games-Howell one-way ANOVA tests to compare mean phosphate concentration in enzyme assay reactions with UCBSV Kikombe Ham1 incubated with the non-canonical nucleotides XTP and dITP and a range of canonical nucleotides.

Table S3: Primers used to amplify CBSV and UCBSV Ham1 sequences during cloning into the POPINF vector. Sequences overlapping with the POPINF vector are shown in red.

Table S4: Primers used to amplify PCR fragments to construct the CBSV_mutHam and CBSV_UHam ICs. The nucleotide sequence encoding the SHR to SAA mutation are shown in blue. The nucleotide sequence encoding the UCBSV Ham1 are shown in red.

Table S5: Relative expression of the CBSV Nampula Ham1 gene in three transgenic *Nicotiana tabacum* lines compared with wild-type; detected by qPCR as described in the 'Experimental Procedures'.

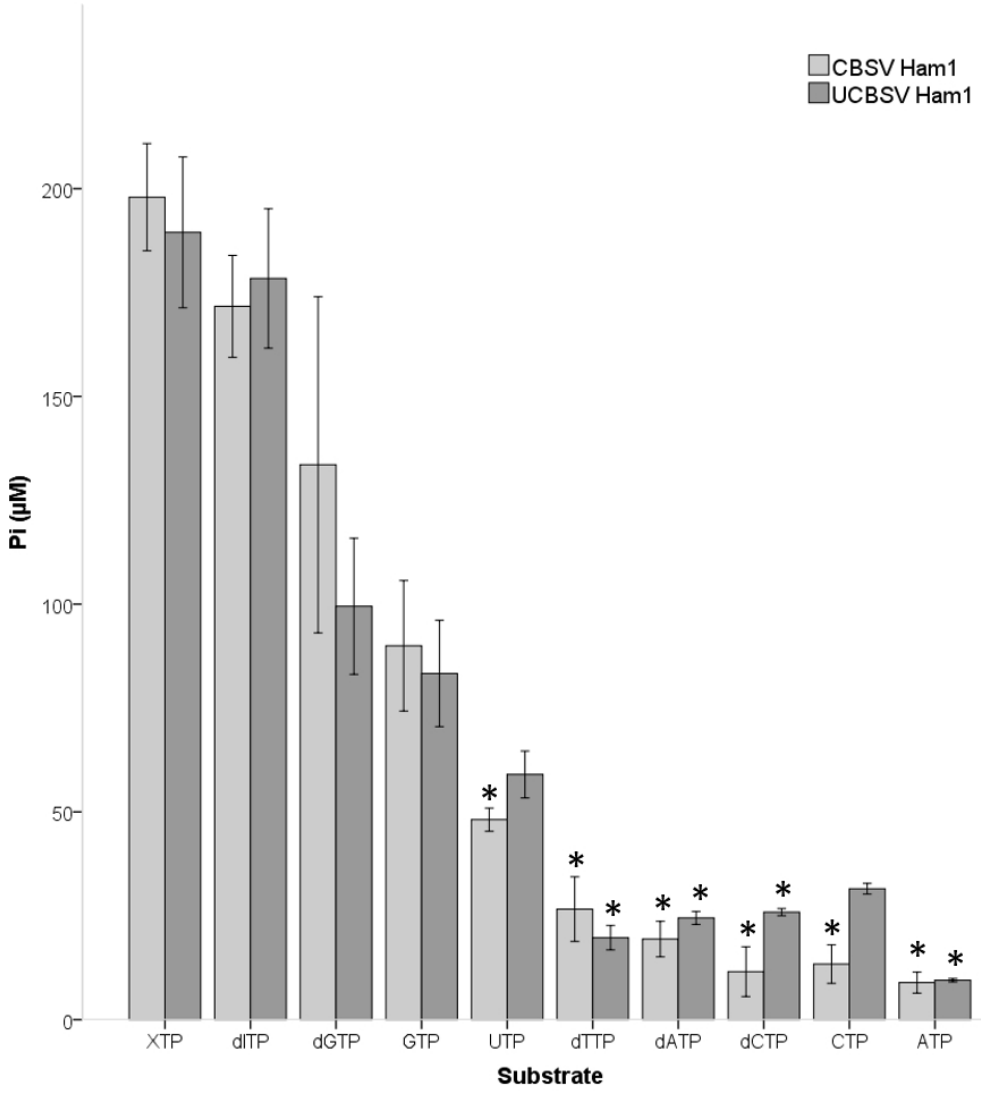


Figure 1: CBSV Tanza and UCBSV Kikombe Ham1 proteins have higher ITPase activities with non-canonical, mutagenic nucleotide triphosphates (NTP): XTP and ITP, compared with canonical NTPs. Purified CBSV Tanza and UCBSV Kikombe Ham1 proteins (1.3 µg) were incubated with 0.2 mM substrates at 37°C for 20 min in 50 mM Tris-HCl, pH 8.5, 1 mM DTT, 50 mM MgCl2, according to Lin et al., (2001). Activity was measured colourimetrically according to the 'Experimental Procedures'. One-way Games Howell ANOVA analyses demonstrated that the CBSV Ham1 protein has significantly higher activities with the non-canonical nucleotide triphosphates XTP and dITP compared to canonical nucleotide triphosphates UTP, dTTP, dATP dCTP, CTP and ATP (*= $p < 0.05$) and that the UCBSV Ham1 protein has significantly higher activities with XTP and dITP compared to dTTP, dATP dCTP, and ATP (*= $p<0.05$). Results are provided in Tables S1 – S2. Each result is the mean phosphate concentration from three separate experiments (n = 3) ± S.E.

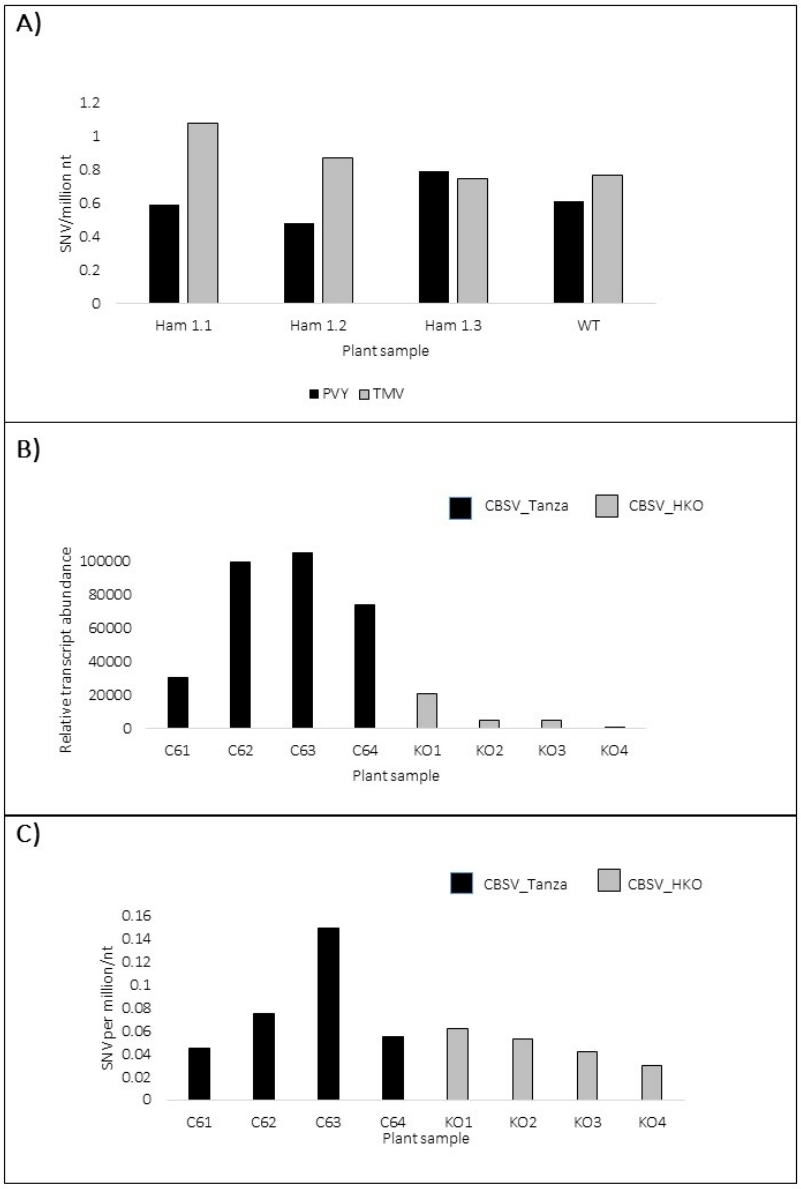


Figure 2: A) Results from deep-sequencing experiment indicate that the transgenic expression of CBSV Ham1 in *Nicotiana tabacum* does not significantly affect PVY or TMV genome diversity. The number of single nucleotide variations (SNVs) per million nucleotides of aligned sequence reads from PVY and TMV infections of wild-type (WT) *N. tabacum* and three transgenic *N. tabacum* lines expressing the CBSV Nampula Ham1 sequence (Ham 1.1, Ham 1.2 and Ham 1.3) are shown . B) QPCR of relative CBSV transcript abundance in CBSV_Tanza infections (C61 – 4) and CBSV_HKO infections (KO1 – 4) used in deep-sequencing experiment. C) Results from deep-sequencing experiment indicate that the deletion of CBSV Ham1 does not significant increase CBSV genome diversity. The number of SNVs per million nucleotides of aligned sequenced reads from CBSV_Tanza infections (C61 -4) and CBSV_HKO infections (KO1 – 4) are shown.

190x275mm (96 x 96 DPI)

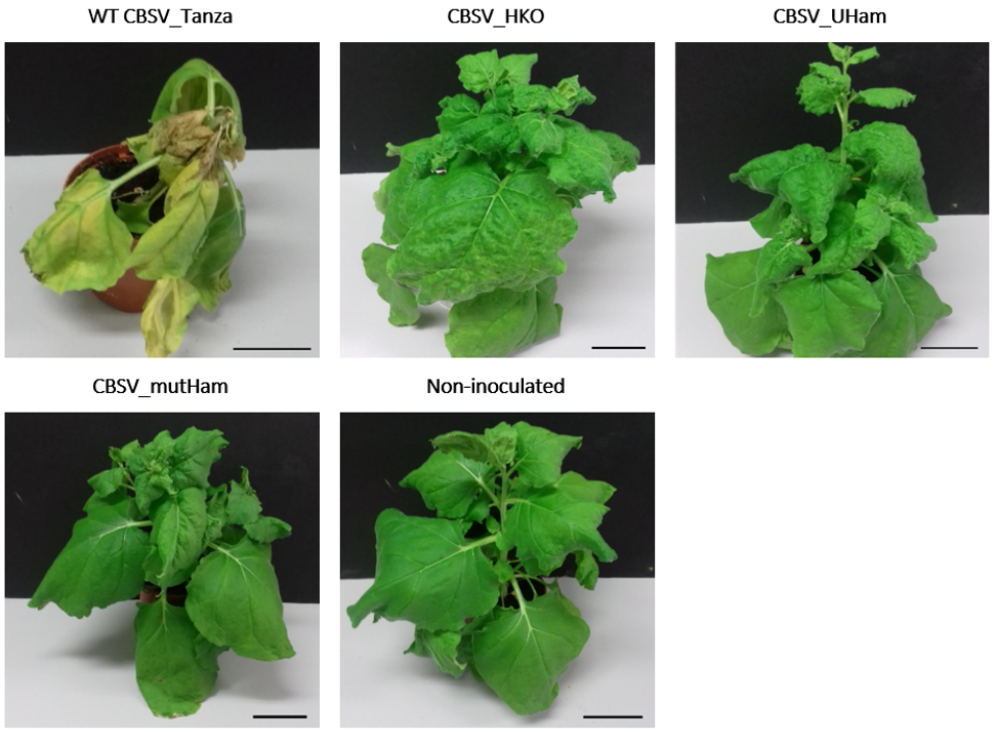


Figure 3 . CBSV Ham1 is associated with the development of necrosis during *Nicotiana benthamiana* infection. Symptom development during *N. benthamiana* infections with CBSV_Tanza ICs at 18 dpi. Wild-type (WT) CBSV_Tanza infections develop severe systemic necrosis, chlorosis, stunting and plant death. Whereas infections with the Ham1 knockout IC: CBSV_HKO, UCBSV Ham1 replacement: CBSV_UHam and mutated SHR motif: CBSV_mutHam develop systemic leaf curling, chlorotic mottling and stunting at 10 – 18 dpi, compared with non-inoculated plants. This suggests that the CBSV Ham1 protein is not essential for infection and is associated with necrosis development during *N. benthamiana* infection. Scale bar = 5 cm.

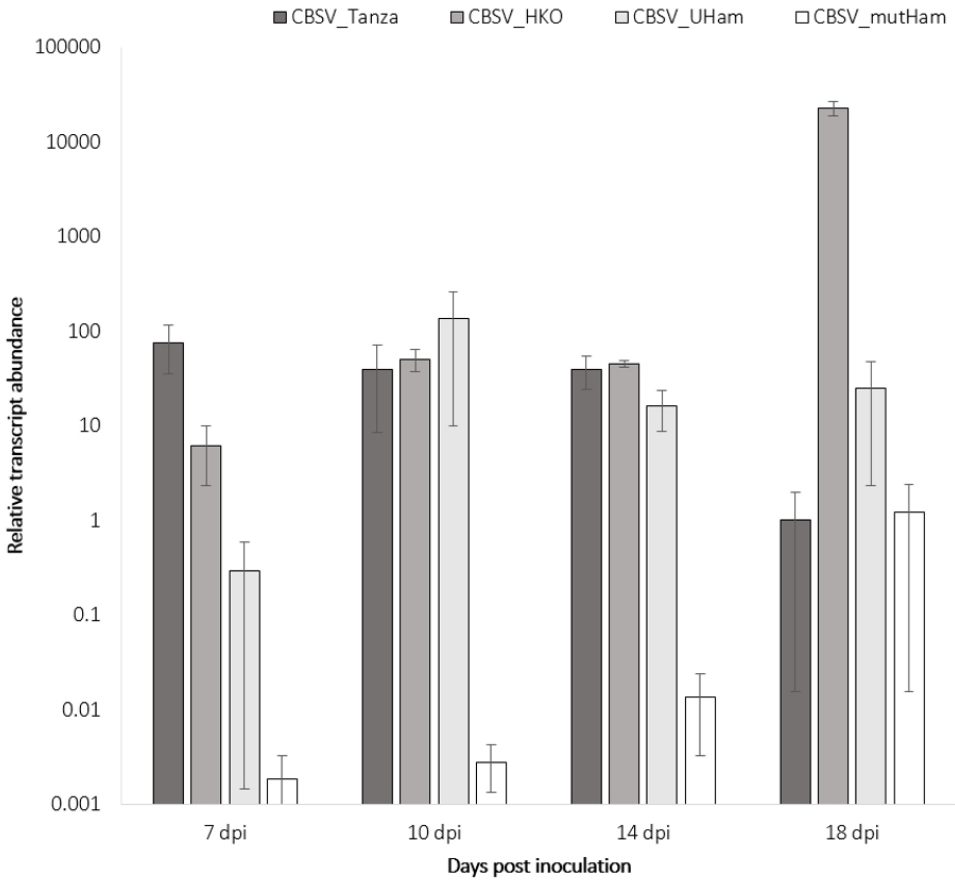


Figure 4: CBSV Ham1 is involved with viral accumulation during *Nicotiana benthamiana* infections. *N. benthamiana* was agroinfiltrated with the following ICs: wild-type CBSV_Tanza, Ham1 deletion (CBSV_HKO), mutated Ham1 ITPase SHR motif (CBSV_mutHam) and the chimera IC containing a UCBSV Ham1 swap (CBSV_UHam). QPCR was performed at 7, 10, 14 and 18 days post inoculation (dpi). Each result is the mean from three replicate plants (n = 3) ± S.E. Results demonstrate that CBSV_Tanza accumulates to higher titers during early infection (7 dpi) compared with CBSV_HKO, CBSV_mutHam and CBSV_UHam, suggesting that Ham1 may be involved with early viral accumulation. Results were consistent in three separate experiments.