Author:
Groenhof, Sophie

Title:
Studies Towards the Elucidation of the Function of the Cassava Brown Streak Virus HAM1h Protein

STUDIES TOWARDS THE ELUCIDATION OF THE FUNCTION OF THE *CASSAVA BROWN STREAK VIRUS* HAM1H PROTEIN

Sophie Groenhof

School of Biological Sciences

University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Science in the Faculty of Life Sciences

Submitted: September 2018 Word Count: *ca*. 26,000

ABSTRACT

The re-emergence of Cassava brown streak disease (CBSD) in the last twenty years is posing a huge threat to the cassava crop in sub-Saharan Africa with its rapid spread towards Western Africa, causing crop losses of up to 70% of the region's second most important carbohydrate source. The causal agents of this disease were found to be *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV), with CBSV being the more virulent. The viruses were found to be the only two of the *Ipomovirus* genus to contain putative HAM1 pyrophosphatases with the function of the viral HAM1h proteins currently unknown. With breeding for resistant cassava cultivars not bringing much success in the field, discovery of the function of these unique proteins may provide a specific target for the development of resistance to CBSD.

Work carried out in this thesis aimed to try to elucidate the function of the CBSV HAM1h. The viral *HAM1* was first transformed into wild type and *ham1* knockout yeast and used in spot assays against increasing concentrations of mutagens, to ascertain if the viral HAM1h does display the same phenotype as the overexpression of the yeast HAM1. Results from these show that the CBSV HAM1h does appear to have functionality. Secondly, the *HAM1* transcript levels in *Nicotiana benthamiana* when infected with CBSV, were investigated to identify why the virus has its own *HAM1*, with results showing that the plant *HAM1* is not downregulated upon CBSV infection. This suggests that the viral HAM1 acts against RNA non-canonical nucleotides and works in addition to the plant HAM1 to afford protection to high rates of mutagenesis. Further studies will need to be conducted to conclude if the CBSV HAM1h does display pyrophosphatase activity and whether this can be targeted to induce resistance to CBSD.

Word Count: 298

DEDICATION AND ACKNOWLEDGEMENTS

DEDICATION:

I would like to dedicate this to my mum and dad.

ACKNOWLEDGEMENTS:

Firstly, I would like to express my thanks to my supervisors, Gary Foster and Andy Bailey. I am extremely grateful for their support and guidance throughout the two years. At times I know I tested their patience, but they never gave up on believing in me and that I could do it, and for that I will be eternally grateful. I would like to thank them both for tirelessly reviewing several drafts and aiding me in the direction to produce work I am proud of.

I would like to thank all members of Lab 321 in providing a friendly atmosphere in which to work and where everyone was happy to help. You were all so supportive and it was such a pleasure to get to know you all. I would particularly like to extend my sincere gratitude to Katie and Luis for their help in all aspects of my project; from the initial understanding to helping me learn all the protocols. Thank you both for always being there to help in every way you could, for being supportive and for answering all the questions I had.

I would also like to thank my dad for proof-reading my many drafts, supporting me when I thought I would not be able to complete it, and for always offering his help. Thank you to both of my parents for always having belief in me and just really for everything you do.

Thank you also to Rebecca, Beth and Ashley for going through the motions of starting a postgraduate course with me and for always being there when I needed any advice and help; to Nia and Matt for going through the ups and downs of an MSc degree with me; and to all the girls at the football club, who have become my second family, for thinking I was silly to undertake such a task but who always knew I could do it and always supported me when the going was tough.

Finally, I would like to thank Lily, Chris, Liam and Coral for always being there no matter what, and thank you to Em, Nat, Laura, Alice and Beth for always having open ears, being supportive of me and giving me perspective on things. Thank you all for keeping me sane throughout it and for being such great friends.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in this dissertation are those of the author.

SIGNED…… **DATE**…….

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

bp	base pair
CBSD	Cassava brown streak disease. Comprises of the effects of both CBSV and UCBSV
CBSV	*Cassava brown streak virus*
CMD	*Cassava mosaic disease*
DEPC	diethyl pyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
(d)NTPase	(deoxy)nucleoside triphosphate pyrophosphatase
EDTA	ethylenediaminetetraacetic acid
GAL-	without supplemented galactose
GAL+	with 20% supplemented galactose
x g	gravitational force
HAM1	nucleoside triphosphate pyrophosphatase found in *Saccharomyces cerevisiae* that hydrolyses non-canonical nucleotides into their constituent monophosphates
ham1	mutant *S. cerevisiae* strain lacking the *HAM1* gene
HAM1h	HAM1 homologue found in CBSV and UCBSV
ITPA	inosine triphosphate pyrophosphatase found in humans that hydrolyses non-canonical nucleotides into their constituent monophosphates
ITP	inosine triphosphate
ITPase	inosine triphosphate pyrophosphatase. Superfamily of NTPases
kbp	kilo base pairs
MES	2-(*N*-morpholino)ethanesulfonic acid
NTP	nucleoside triphosphate
OD	optical density
RdgB	inosine triphosphate pyrophosphatase found in *Escherichia coli* that hydrolyses non-canonical nucleotides into their constituent monophosphates
RNA	ribonucleic acid
rpm	revolution per minute
SDW	sterile deionised water
TAE	Tris-acetate-EDTA

TMV	*Tobacco mosaic virus*
UCBSV	*Ugandan cassava brown streak virus*
URA-	without supplemented uracil
w/v	weight per volume
YPD	Yeast Peptone Dextrose
YPDA	Yeast Peptone Dextrose Agar
YSDM	Yeast Synthetic Dropout Media (URA-)

CHAPTER 1 – INTRODUCTION

1.1 FOOD SECURITY

With the global population expected to reach 11.2 billion by the turn of the next century, the need for increased crop yields has never been more pressing (Johnson *et al*., 2017). It is estimated that 80% of the world's population are currently malnourished and with the majority of those being in developing countries (von Braun, 2010; Abass *et al*., 2018), food production as it currently stands would have to be doubled by 2050 to accommodate for a rising population and ensure that everyone is adequately nourished (von Braun, 2010). Yet, with crop production yield increases reaching a plateau after the great advances made over the last sixty years (Johnson *et al*., 2017) and with approximately 25-40% of reduced crop production due to pests and pathogens (Myers *et al*., 2017), it is evident that crop cultivars resistant to pests and pathogens will be vital in the search for global food security (Johnson *et al*., 2017).

Plant diseases have caused many famines globally that have devastated communities and indeed countries. One of the most well-known of these events is the Irish Potato Famine that occurred during the mid-nineteenth century when the oomycete pathogen, *Phytophthora infestans*, caused Potato late blight disease and decimated the crop causing famine across the whole country. As a result, almost one eighth of the population died and a further one million people emigrated. Such devastating losses were mainly accountable due to the fact that the population relied on one crop and that propagation of the monoculture was through the continued planting of apparently healthy tubers. The continued propagation of the tubers created a limited gene pool, making the crop highly vulnerable to new diseases, and thus the devastation of *P. infestans* infection was deemed inevitable (Scholthof, 2007).

It is estimated that at least 10% of all global food crop production is lost due to plant disease alone, and with only fourteen crops providing the staple food source to diets worldwide (Strange and Scott, 2005), understanding crop disease mechanisms is becoming imperative to ensure sufficient yields to feed a growing population.

1.2 CASSAVA

1.2.1 BACKGROUND

Cassava (*Manihot esculenta* Crantz; Family Euphorbiaceae) is a woody perennial, growing to heights of between two and four metres (Fig. 1.1) (Pandey *et al*., 2000), and is farmed for its carbohydrate-rich roots (Patil *et al*., 2015) across the tropical regions of Africa, Asia and Latin America (Patil *et al*., 2011), where it provides food for an estimated 800 million people globally (Fondong, 2017; Wilson *et al*., 2017). Originating from Brazil where it has been cultivated for the past 9,000 years, its introduction into Africa in the sixteenth century following Portuguese conquests (Howeler *et al*., 2013; Fondong, 2017) and with its expansion throughout the continent, the cassava crop is now Africa's second most important food crop after maize (Abarshi *et al*., 2010). Subsequently, Africa is the largest producer of cassava globally (Abarshi *et al*., 2010), producing 158 million tonnes in 2013 (Bennet, 2015) and providing the staple food source for 67% of the poorer households in Eastern Uganda and Western Kenya alone (Taylor *et al*., 2012).

Figure 1.1. View of the cassava crop in the field (reproduced from http://keralacultivation.blogspot.com/2010/12/tapioca.html)

In Asia however, following on from its introduction in the eighteenth century (Pandey *et al*., 2000), the cassava crop is mainly used in industrial processes and as a promising source of biofuel, particularly from the cassava peels (Ye *et al*., 2017). Indonesia, the Philippines, Cambodia and East Timor all cultivate the crop for human consumption, while in Vietnam and China, large amounts of the cassava crop are used for animal feed, either for their own farms or as an export (Howeler, 2006). Intercropping of cassava is common in Asia where plots of land are small, to help increase yields of crops that are able to make use of the different nutrients available without exhausting them. Rice, maize, legumes and vegetables are all used in intercropping, with these planted in between rows of

cassava, and also provide the farmer with several crops over the year. Cassava can also be used as an intercrop itself for young trees like rubber and cashew. Very few economically important diseases affect the cassava crop in Asia, with the exception of the Indian sub-continent, so few steps need to be taken to reliably obtain a good yield (Howeler, 2006). The lack of precautionary methods taken to therefore ensure a broad gene pool, mean that the cassava crop may be particularly vulnerable if the diseases found in Africa are transmitted to Asia; with the crop being at risk of rapid decimation similar to what was seen when *P. infestans* caused the Irish Potato Famine.

Cassava is particularly important in sub-Saharan Africa where the starchy roots comprise over half the calorie intake for more than fifty percent of the rural and urban populations in those countries (Abarshi *et al*., 2010). It is the continent's second most important food staple in terms of the *per capita* calories consumed (Nweke, 2004; Tomlinson *et al*., 2017). Many subsistence farmers favour cassava production as the crop is tolerant to drought conditions, can be grown in marginal land and harvested throughout the year, is cheap to grow as it is propagated through stem cuttings (Abarshi *et al*., 2010), and can provide a high yield through its ability to grow large, starchy roots with little labour input (Wilson *et al*., 2017). These roots, which are high in carbohydrate content but low in protein and nutrient content (Wilson *et al*., 2017), form the primary carbohydrate source for the growers themselves and provide a source of income through the sale of the roots and the sale of flour processed from the roots (Olsen and Schaal, 1999; Tomlinson *et al*., 2017). Furthermore, as three quarters of the world's ultra-poor (defined as living on less than fifty cents a day) are situated in sub-Saharan Africa (von Braun, 2010), the cassava crop is vital to their survival and forms a huge part of the local economies. The cassava leaves are also a diet staple as they provide an important source of protein in several African countries (Lozano, 1986).

Cassava has become a staple food security crop with its ability to remain in the ground for up to three years and produce harvest throughout the year and thus is able to outperform cereal crops which are quickly exhausted (Legg *et al*., 2014; Tomlinson *et al*., 2017). Cassava is certainly a suitable crop for the future with rising temperatures and extreme climatic events affecting the production of many cereal crops while cassava, on the other hand, is able to withstand these changing environmental conditions and high carbon dioxide concentrations. It has since become the second most important source of carbohydrate in sub-Saharan Africa, and over 105 counties are now recognising its advantages (Legg *et al*., 2014).

1.2.2 Cassava diseases

Despite the high-yielding production, the cassava crop is affected by both biotic and abiotic factors that limit the yield of the starchy storage roots to fifteen tonnes per hectare worldwide as opposed to

a potential yield of fifty tonnes (Kuria *et al*., 2017). The ability of the plant to grow on malnourished soils is just one limiting factor to its productivity and in Asia, the use of irrigation in commercial settings has enabled productivity to quadruple (Blagbrough *et al*., 2010); a system that is not available to subsistence farmers in Africa.

One of the limiting biotic factors is the crop's susceptibility to disease; over thirty diseases, from bacterial to viral, are known to affect production (Owolade *et al*., 2005). Of these, cassava bacterial blight (CBB), caused by the bacterium *Xanthomonas axonopodis* pv. *manihotis*, devastated the crop in Africa and Latin America in the 1970s. It not only caused losses to the fresh roots of up to 75%, but also affected the protein-rich leaves, inducing losses of up to 80%. This disease was spread through the propagation of infected stem cuttings and then through the use of contaminated tools, as well as naturally through rain splashing (Lozano, 1986). Cassava anthracnose disease (CAD), furthermore, caused by the fungus *Colletotrichum gloeosporioides* f.sp. *manihotis*, induced widespread reductions in crop yield and extensive dieback across much of Africa, with 80-90% of cultivars showing high susceptibility, and was again spread through the propagation of infected stem cuttings as well as via the insect vector, *Pseudotheraptus devastans*. Both diseases were brought under control through the propagation of healthy cuttings, phytosanitation methods, and most successfully through the selection and breeding of resistant cultivars (Lozano, 1986; Fokunang *et al*., 2001; Owolade *et al*., 2005; López and Bernal, 2012). Moreover, with the delay to the cassava planting season, CBB is now a disease of only minor importance (Lozano, 1986).

There are at least twenty viral diseases (Patil and Fauquet, 2009) that can infect the cassava crop, with cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) being the most devastating in Africa (Patil *et al*., 2011). Furthermore, the cultivation of the crop as a monoculture and the ease of planting of cassava from stem cuttings, forming its primary source of propagation, enables infection to persist and cause such devastating losses, as seen in Eastern Africa (Hillocks and Jennings, 2003).

1.2.3 Cassava mosaic disease

1.2.3.1 Background

Cassava mosaic disease (CMD) is caused by several cassava mosaic geminiviruses (Family *Geminiviridae*; Genus *Begomovirus*) and is widespread throughout Africa and the Indian subcontinent (Patil and Fauquet, 2009; Vanderschuren *et al*., 2009). Cassava mosaic geminiviruses consist of two circular single-stranded DNA parts, DNA A and DNA B. DNA A constitutes the viral coat protein (AV1), a silencing suppressor that targets the post-transcriptional gene-silencing response of the plant (AV2), suppressors of host-mediated gene silencing (AC2 and AC4) and viral replication genes (AC1 and AC3).

DNA B on the other hand, encodes two genes that are involved with the movement of the virus within the host (BC1 and BV1) (Singh *et al*., 2015; Kuria *et al*., 2017).

There are eleven recognised species of cassava mosaic geminiviruses and the disease is transmitted by the whitefly *Bemisia tabaci* on a persistent basis, as well as through the propagation of infected stem cuttings (McCallum *et al*., 2017). CMD, having first been described in Tanzania in 1894, is now found in all cassava-cultivating countries of Africa (Hillocks and Thresh, 2000; Legg *et al*., 2014) and is responsible for not only economic losses of between 1.9 and 2.7 billion US Dollars annually, but also for the loss of millions of lives as a result of malnutrition following the disease pandemic in Central and Eastern Africa (Scholthof *et al*., 2011).

Symptoms of CMD are characterised by the chlorotic mosaic patterning on leaves, deformed leaves and the stunting of shoot growth (McCallum *et al*., 2017). Not only does this disease severely affect the cassava crop foliage, which is also consumed as it provides a source of protein (Howeler *et al*., 2013), it also affects the root tuber yield causing average losses of 82% (McCallum *et al*., 2017) but with some highly susceptible cassava varieties experiencing 100% losses in root yield (Fondong, 2017).

1.2.3.2 Resistance to CMD

Some protection has been afforded against CMD through the breeding of naturally resistant cassava varieties and through the planting of healthy stem cuttings. In countries such as the Democratic Republic of Congo, where the leaves are consumed as nutritious vegetables, there has been selection for cassava cultivars showing mild foliar CMD symptoms, instead of those expressing no symptoms. This has hindered some of the control of CMD through the planting of CMD-susceptible varieties (Thresh and Cooter, 2005), yet the use of CMD-resistant varieties has successfully been used to control devastating epidemics in Madagascar, Nigeria and Uganda (Thresh and Cooter, 2005). Work over the years has shown that the planting of virus-free cuttings, the use of improved resistant varieties and the wide-spread use of phytosanitation methods, can effectively control CMD across large areas (Thresh and Cooter, 2005). There are currently three types of CMD resistant cassava strategies that have been deployed for use in the field. The first (CMD1) was introgressed from *Manihot glaziovii* Muell. Arg. (ceara rubber) but is recessive and polygenic. The second (CMD2) is derived from a single genetic locus and is found in various strains of West African *Manihot esculenta* landraces. This form of resistance is highly heritable and is also able to confer stable resistance to a broad spectrum of cassava mosaic geminiviruses, and thus has been used in breeding programmes in Africa and Latin America to produce highly CMD resistant cultivars. The third cultivar (CMD3) is highly resistant to CMD, with less than 1% disease incidence, and contains the same genetic locus as the second resistant cultivar, but also another locus in the same linkage group (Kuria *et al*., 2017). A recent study by Beyene

et al. (2016) showed that field level CMD2 resistance may be at threat, however, with the regeneration of plants through somatic embryogenesis showing susceptibility to CMD in field trials. Furthermore, this generation and wide-spread distribution of CMD-resistant cassava varieties across much of Africa may have helped spread CBSD as many of these cultivars show susceptibility to the disease (Vanderschuren *et al*., 2012).

With no known cultivars conferring resistance to both CMD and CBSD to date, the objective remains to generate a cultivar that is tolerant to the two diseases. Following advances in sequencing and biotechnology, work has been done on antiviral RNA silencing associated defence mechanisms, with the loss of cassava mosaic virus RNA from infected plants having been seen (Kuria *et al*., 2017). In the years to come, this could provide a promising source of resistance that could also incorporate resistance to CBSD.

1.3 CASSAVA BROWN STREAK DISEASE

1.3.1 HISTORY OF THE DISEASE

Cassava brown streak disease (CBSD) was first described in 1936 in the Amani district of Tanzania (Storey, 1936) and by 1950 was found in the lowlands (less than 1000 metres above sea level) along the East African coast from Mozambique to Kenya and inland to include the lowlands of Malawi and Uganda (Nichols, 1950; Mbanzibwa *et al*., 2009a). This first discovery of CBSD identified symptoms on the lower leaves and rot of the tuber roots (Storey, 1936). It wasn't until a study by Nichols in 1950 that the observation that the disease can affect all parts of the plant, was made. Subsequently, two main foliar symptoms of CBSD have been described: firstly, the feathery chlorosis along secondary veins, and secondly, mottled chlorosis that appears distinct to the feathering along the veins (Nichols, 1950; Tomlinson *et al*., 2017). Nichols (1950) also discovered that the severity of symptoms can vary on biotic factors, such as the cultivar and age of the cassava plant when infected, and on abiotic factors such as the environmental conditions. Moreover, the symptoms and severity of the disease are affected by the viral strain (Nichols, 1950), with many isolates having since been identified (Monger *et al*., 2001b). These differences in symptom severity, ranging from those plants exhibiting severe foliar and root symptoms, to severe root necrosis and only mild leaf chlorosis or vice versa, to those exhibiting both mild leaf and root symptoms, make detection of CBSD in the field extremely difficult (Kaweesi *et al*., 2014). Most farmers are thus unaware that the crop is infected until they come to harvest the root tubers (Legg *et al*., 2015). The initial spread of CBSD along the lowlands of the East African coast and inland, furthermore, is thought to have been through the planting of infected stem cuttings from the Amani district, particularly with the spread of the disease to Uganda (Nichols, 1950; Jameson, 1964; Tomlinson *et al*., 2017). The spread of infection further afield was prevented through

the lack of observation of any vector-borne transmission at higher altitudes (Nichols, 1950; Jennings, 1960; Tomlinson *et al*., 2017), keeping this disease confined to the lowland regions of East Africa for the next seventy years (Mbanzibwa *et al*., 2009a).

1.3.2 Initial methods to control CBSD

As a consequence of the spread of CBSD, mechanisms of control were established. Initial breeding of cultivated cassava varieties with wild relatives, such as *Manihot glaziovii*, *M. dichotoma*, *M. catingea*, *M. saxicola* and *M. melanobasis*, all species which showed greater resistance to CBSD, began in 1937 with the view to creating tolerant cultivars (Jennings, 1957; Nichols, 1946; Kawuki *et al*., 2016; Tomlinson *et al*., 2017). Tolerance for CBSD is defined as those cassava varieties displaying no root necrosis or those displaying root rot at a late stage of growth (Legg *et al*., 2011). Hybrids generated from this cross-breeding programme that were subsequently introduced to the local farmers included the first cassava cultivar to be produced which showed relatively high levels of tolerance to CBSD (Kaweesi *et al*., 2014; Tomlinson *et al*., 2017). This CBSD tolerant cultivar was the *M. esculenta-M. glaziovii* hybrid, known as 'Namikonga' in Tanzania and 'Kaleso' in Kenya, but it was not widely distributed to farmers due to its susceptibility to CMD (Hillocks and Jennings, 2003; Kaweesi *et al*., 2014; Tomlinson *et al*., 2017).

Following on from the increase in CBSD incidence observed in the 1990s along East Africa, virus-free tolerant CBSD cultivars were made available to farmers in Mozambique who relied heavily on CBSD susceptible cassava plants (Hillocks and Jennings, 2003; Tomlinson *et al*., 2017). At this time, CMD was the more devastating disease, causing wide-spread famine across Africa with many farmers abandoning the crop, and so methods to control and prevent the spread of CBSD were of little importance as it seemed to remain confined to the lowlands of East Africa (Thresh and Cooter, 2005; Tomlinson *et al*., 2017). As a result, CMD resistant cultivars were widely distributed amongst farmers (Legg and Thresh, 2000; Tomlinson *et al*., 2017), but as these varieties show varying levels of susceptibility to CBSD, it is thought this may have helped spread CBSD (Legg *et al*., 2006; Tomlinson *et al*., 2017). The development of CBSD resistant cultivars through traditional breeding methods has proved difficult and is limited by the availability of resistance genes to the causal viruses in existing cassava varieties (Taylor *et al*., 2016), but cultivars created from the initial breeding programme started in 1937 still form an important genepool for the continued search for CBSD resistance (Kaweesi *et al*., 2014).

1.3.3 Re-emergence of CBSD

In late 2004, incidences of CBSD were observed in central Uganda, and marked the first time it had been seen in Uganda since the initial introduction of the disease from Tanzania in the 1940s (Alicai *et*

al., 2007). Moreover, this re-emergence of CBSD was now found at higher altitudes, over 1000 metres above sea level, and over 1000km inland. No longer was CBSD restricted to the lowlands of coastal East Africa and the surrounding lowlands of Lake Malawi, but the disease was now spreading westwards to Central Africa (Fig. 1.2). The current CBSD epidemic in East Africa is responsible for most of the cassava losses and it is feared with its increasing spread of infection, the disease will reach the cassava-producing countries in Central and West Africa, which are some of the world's leading, and where CBSD is currently not found. CBSD has already spread to neighbouring Rwanda, Burundi, The Democratic Republic of Congo, South Sudan and even on into Congo (Alicai *et al*., 2007; 2016; Legg *et al*., 2014; Patil *et al*., 2015).

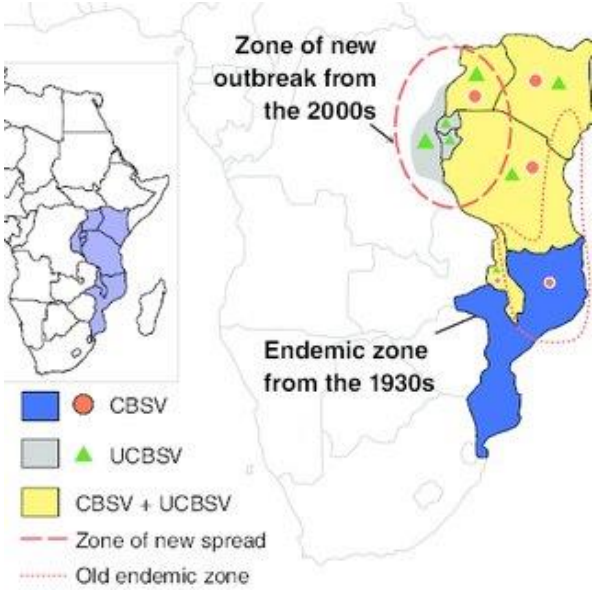

Figure 1.2. Map showing the spread of CBSD in the 1930s, with the two viral species being found in separate and overlapping areas, and the re-emergence of the disease in 2004 (reproduced from Legg *et al*., 2015).

This new incidence of CBSD in Central Uganda is thought to either have arisen from a new introduction of the disease from the planting of infected stem cuttings, or that the disease had persisted, albeit at low levels, from its initial introduction in the 1940s (Alicai *et al*., 2007; Tomlinson *et al*., 2017). Confirmation of the disease was carried out using RT-PCR from plants displaying CBSD symptoms at these higher altitudes in Uganda (Alicai *et al*., 2007). Comparison of the coat protein sequence from this viral isolate to those from CBSD-infected plants in Mozambique and Tanzania, revealed only 77.0 to 82.9% sequence identity (Alicai *et al*., 2007; Tomlinson *et al*., 2017) suggesting that two genetically different viral strains cause CBSD (Mbanzibwa *et al*., 2009a).

Initial studies thought that the causal agent was a member of the carlaviruses, based on the particle size of around 650nm and its flexuous rod shape, but the discovery of pin wheel inclusion bodies made it typical of a potyvirus infection (Lennon *et al*., 1985; Legg *et al*., 2015). Furthermore, raising an antiserum to the *Cowpea mild mottle virus* (family: *Betaflexiviridae*, genus: *Carlavirus*) in cassava

plants revealed that this virus was not the causal agent of CBSD (Lennon *et al*., 1985; Legg *et al*., 2015). It was not until 2001 when work by Monger *et al*. (2001a) demonstrated that a virus consistently associated with CBSD showed similarities to members of the genus *Ipomovirus* within the *Potyviridae* family, with its coat protein showing 43.2% sequence identity to that of *Sweet potato mild mottle virus* (SPMMV), the then only sequenced virus belonging to the *Ipomovirus* genus. Further work by Monger *et al*. (2001b) comparing coat protein sequences between isolates from Mozambique and Tanzania, varying by 8% and 6% at the nucleotide and amino acid levels respectively, revealed that there may be several different strains of CBSD-causing viruses, with the appearance of variable symptoms on different cassava cultivars (Hillocks *et al*., 1996; Monger *et al*., 2001b) and experimental hosts, such as *Nicotiana benthamiana* (Monger *et al*., 2001b), also supporting this observation (Legg *et al*., 2015).

It was confirmed that CBSD is caused by two phylogenetically related species of positive-sense single-stranded RNA viruses belonging to the genus *Ipomovirus* in the *Potyviridae* family (Mbanzibwa *et al*., 2009b; Winter *et al*., 2010; Patil *et al*., 2011). Symptoms on cassava plants from Uganda varied to those grown in different environmental conditions in the coastal, lowland regions of Tanzania and Mozambique, and subsequent sequence analysis of the isolated viral fragments revealed on average only approximately 70% nucleotide homology and 74% polypeptide homology, thus confirming the appearance of two distinct species (Mbanzibwa *et al*., 2009a; 2011a; Monger *et al*., 2010; Winter *et al*., 2010). *Cassava brown streak virus* (CBSV) was responsible for the first CBSD infection in the lowland and coastal regions of East Africa, and *Ugandan cassava brown streak virus* (UCBSV) was responsible for the spread of infection to highland regions and the Lake Victoria basin (Patil *et al*., 2011), though both species are now found across the highlands and lowlands of this East African region (Fig. 1.2), where co-infection of CBSV and UCBSV is not uncommon (Ndunguru *et al*., 2015). Indeed, this co-infection now forms 34-50% of CBSD cases in Kenya (Kathurima *et al*., 2016), Tanzania (Mbanzibwa *et al*., 2011b) and Uganda (Ogwok *et al*., 2014; Tomlinson *et al*., 2017). Whether the two viruses act synergistically in the cassava hosts is still unknown (Tomlinson *et al*., 2017), though Vanderschuren *et al*. (2012) did not find any evidence of such in the field.

1.3.4 Detection and propagation of CBSD

Detection of CBSD infection above ground is near impossible with new growth and immature leaves showing no symptoms (Fig. 1.3.A), and with the infection symptoms varying with the cassava cultivar and the environmental conditions (Nichols, 1950; Monger *et al*., 2001b). The study by Mohammed *et al*. (2012) has highlighted the differences in symptom severity on various cassava varieties, with the variety Albert showing high susceptibility to both CBSV and UCBSV, with other varieties showing greater susceptibility to CBSV only. Most symptoms are present in the older leaves where oxidative

stress becomes the main driving force of causing early senescence in the cassava crop (Mbanzibwa *et al*., 2009b) with the crop then exhibiting the tendency to shed these older symptomatic leaves, especially during long dry periods, making detection of the infection even less obvious (Monger *et al*., 2001a; Abarshi *et al*., 2010). Foliar chlorotic symptoms only become visible as the leaf matures as opposed to when the leaf unfurls, as is observed in cassava infected with CMD, and the chlorosis differs from CMD in that it appears as feathering along the veins (Fig. 1.3.B). With CBSD, there is also a lack of leaf distortion and furthermore, the symptoms of CBSD can often be masked by other diseases, such as CMD, or by pests like the cassava green mite (*Mononychellus tanajoa*) (Abarshi, *et al*., 2010). CBSD can be detected in all parts of the cassava plant, however, whether exhibiting symptoms or not (Abarshi *et al*., 2010), and an RT-PCR test has been developed that can detect the virus even in the young cassava leaves that are not yet showing symptoms (Monger *et al*., 2001b). Use of this RT-PCR test has also been used to demonstrate that virus titre is lowest in the young leaves, and highest in the mature leaves (Ogwok *et al*., 2014).

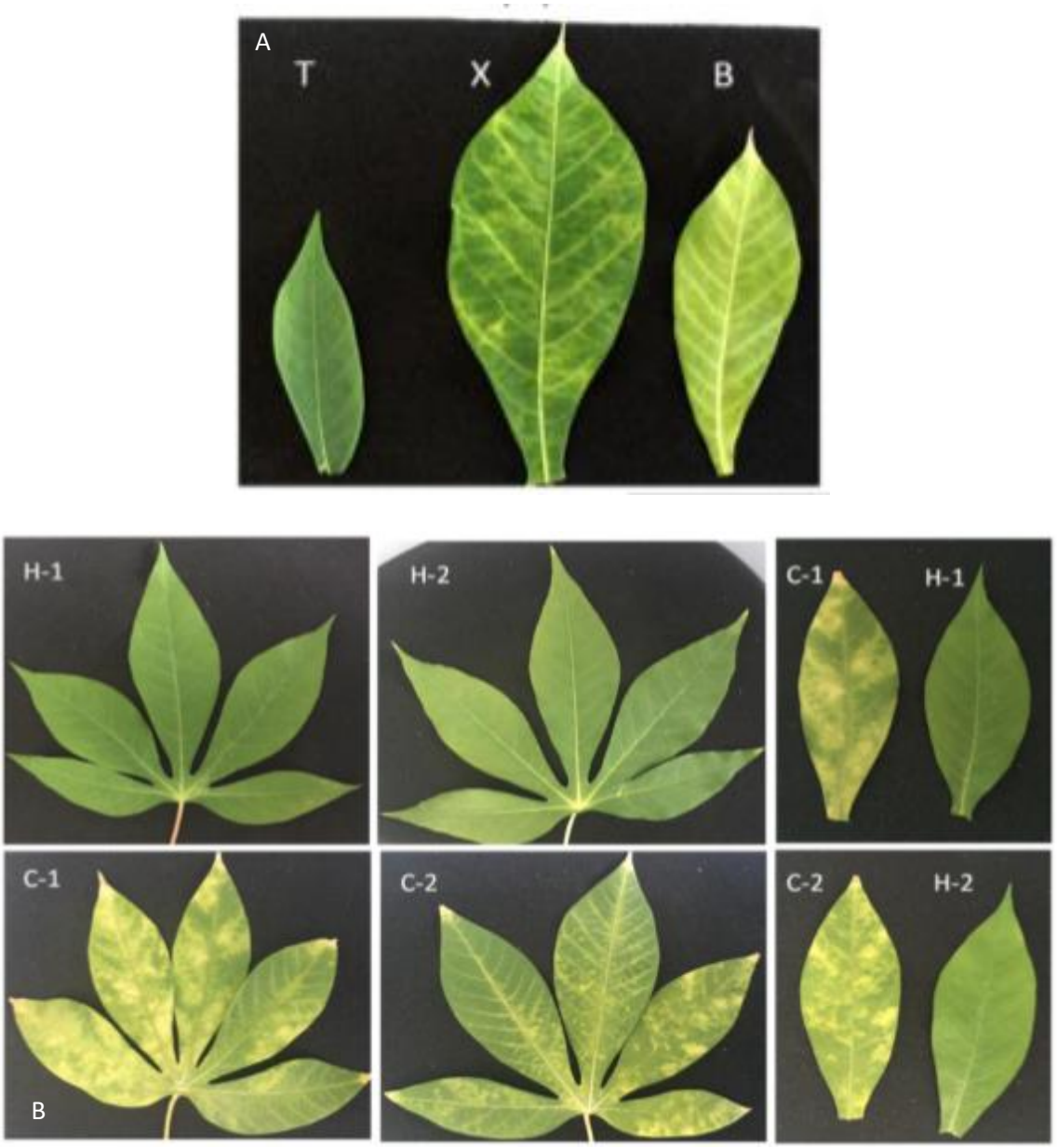

Figure 1.3. A. Cassava infected with CBSV (variety Kibandameno) with leaves from positions top (T), middle (X) and bottom (B) of plant showing symptom severity increasing with maturity of leaves. B. H-1 and H-2 were asymptomatic for CBSV whilst C-1 and C-2 show the foliar chlorosis symptoms of CBSV (variety Kibandameno) (reproduced from Saggaf *et al*., 2018).

The characteristic feature of CBSD infection which gives the disease its name, is the brown, streaky lesions that can be found on the stems beneath the bark and which appear as purple or brown lesions on the exterior of young stems (Fig. 1.4). Despite this being the characteristic feature of the disease, however, it is not always present (Hillocks and Jennings, 2003). Outbreaks of CBSD were found to occur around three to twelve years after increases in whitefly numbers, with these whiteflies also being able to thrive above 1000 metres above sea level (Alicai *et al*., 2007; Legg *et al*., 2011; Tomlinson *et al*., 2017). It was confirmed that CBSD is also transmitted by whitefly vectors, with *Bemisia afer* originally thought to be the vector, but later evidence found that the CMD vector *B. tabaci* also

transmits CBSD (Maruthi *et al*., 2005). Spread of CBSD infection through the insect vector is only semi-persistent, with infection being spread by *B. tabaci* approximately 22% of the time, as has been observed in the field and in laboratory conditions by Maruthi *et al*. (2005). This suggests that the disease is spread further afield by the whitefly vector, but that the disease persists through the propagation of asymptomatic infected stem cuttings (Patil *et al*., 2011; Yadav *et al*., 2011; Tomlinson *et al*., 2017).

Figure 1.4. Brown streaks observed on the stems of the cassava plant (reproduced from Patil *et al*., 2015).

1.3.5 Economic Impact

The spread of CBSD at such a pace is concerning as this disease has up to 100% incidence and in the most severe cases, can cause decreases of up to 70% in root weight (Hillocks *et al*., 2001; Maruthi *et al*., 2005; Ateka *et al*., 2017). It is considered to be more of a threat than CMD as not only does it cause a reduction in overall yield (Fig. 1.5.A), but CBSD causes the dry, necrotic rot of the root tubers (Fig. 1.5.B) which appears around six months post planting, and renders them unfit for human consumption and for sale (Hillocks *et al*., 2001; Maruthi *et al*., 2005; Patil, *et al*., 2011; Mohammed *et al*., 2012), severely affecting those that rely on this staple crop. Furthermore, the extent of the damage to the roots caused by CBSD does not become apparent until they are harvested (Maruthi *et al*., 2005), where they subsequently suffer rapid deterioration (Monger *et al*., 2001a). As a result, farmers suffer further reduced yields through the premature harvesting of the crop before the root necrosis takes hold (Kaweesi *et al*., 2014).

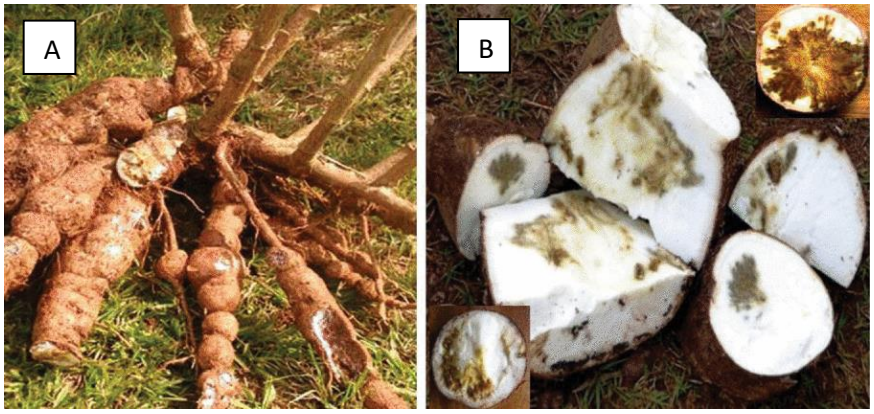

Figure 1.5. A. Constrictions on the root surface of cassava which reduce the root yield. B. The dry, necrotic rot of the cassava root (reproduced from Patil *et al*., 2015).

1.3.6 CBSV AND UCBSV

Advances in sequencing technology have enabled more isolates of CBSV and UCBSV to be identified and sequenced. Complete genome analysis of CBSV and UCBSV isolates have revealed that the UCBSV genome is larger, with a length of 9,069 nucleotides for its MLB3 isolate, compared to that of the CBSV isolate TZ: Kor6: 08, which is 8,995 nucleotides in length (Ndunguru *et al*., 2015). Analysis also showed that genetic diversity is greater among the CBSV isolates, with 79.3-95.5% sequence identity, compared to the UCBSV isolates which show 86.3-99.3% sequence identity (Ndunguru *et al*., 2015). Indeed, these differences in isolate identity have suggested that there may be more than two species causing CBSD, with CBSV and UCBSV forming two clades, with the potential of three species within the UCBSV clade alone (Fig. 1.6) (Ndunguru *et al*., 2015; Tomlinson *et al*., 2017).

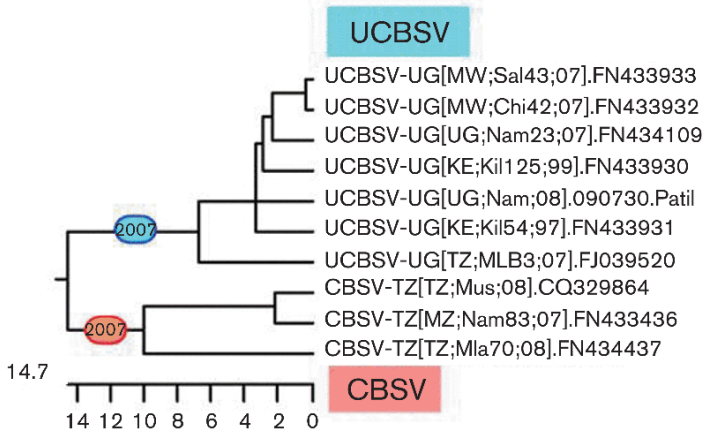

Figure 1.6. Phylogenetic tree of the full genome sequences of isolates of CBSV and UCBSV, highlighting how there is the potential for species within the CBSV and UCBSV clades (reproduced from Patil *et al*., 2015).

Differences in genome sequence may be behind the variability in symptom severity caused by CBSV and UCBSV. Cassava plants infected with CBSV suffer from more severe root necrosis and foliar symptoms coalesce to form chlorotic blotches on the leaves, whereas with UCBSV infection, both symptoms appear less severe (Tomlinson *et al*., 2017). CBSV also accumulates higher titres in cassava

plants (Kaweesi *et al*., 2014) and indicator plant hosts (Mohammed *et al*., 2012) compared to infections with UCBSV, demonstrating that CBSV is the more virulent of the two (Ndunguru *et al*., 2015; Tomlinson *et al*., 2017). The greater diversity of CBSV isolates may also be the reason why resistant cassava cultivars are hard to generate as the viral isolates can rapidly adapt to outsmart the plant defence mechanisms that are selected for (Tomlinson *et al*., 2017).

1.4 Family – *Potyviridae*

1.4.1 Background

With at least 176 species currently described, the *Potyviridae* is the largest viral plant pathogen family and contains many viral diseases that cause catastrophic losses both economically and agriculturally. The family comprises eight genera: *Brambyvirus*, *Bymovirus*, *Ipomovirus*, *Macluravirus*, *Poacevirus*, *Rymovirus*, *Tritimovirus* and *Potyvirus*, which is the largest genus. Viruses belonging to the *Potyviridae* family have single-stranded RNA genomes that are encapsulated in flexuous filamentous particles of length 650-900nm. The genera are further categorised based on their vector transmission and genomic relatedness as shown in Table 1.1 (Adams *et al*., 2005a; 2005b; Dombrovsky *et al*., 2014).

Table 1.1. Table of the different genera belonging to the *Potyviridae* family showing their genome organisation and methods of vector transmission (Adams *et al*., 2005a; 2005b; Dombrovsky *et al*., 2014).

Genus	Number of Species	Genomic Organisation	Transmission Vector and Manner
Potyvirus	146	Monopartite	Aphids Non-persistent
Ipomovirus	6	Monopartite	Whiteflies Semi-persistent
Macluravirus	6	Monopartite	Aphids Non-persistent
Poacevirus	2	Monopartite	Mites
Rymovirus	3	Monopartite	Eriophyid *Abacarus* mites Semi-persistent
Tritimovirus	5	Monopartite	Eriophyid *Aceria* mites Semi-persistent
Brambyvirus	1	Monopartite	-
Bymovirus	6	Bipartite	Plasmodiophorid protist, *Plymyxa graminis* Persistent

1.4.2 GENOME ORGANISATION

With the exception of viruses belonging to the *Bymovirus* genus, the genomes of viruses within the *Potyviridae* family encode a large polyprotein that is cut by three virus-encoded proteases to form ten mature proteins: P1, HC-Pro, P3, 6K1, CI, 6K2, VPg, NIa-Pro, NIb and CP (Adams *et al*., 2005b). This genome organisation is highly conserved amongst the family (Fig. 1.7). In the case of the *Bymovirus* genus, however, which has a bipartite genomic construction, RNA1 encodes eight of the proteins found in the other genera viral genomes from P3 onwards, and the smaller RNA2 part of the *Bymovirus* genome, which has less similarity to the 5'-section of the other genera in the family, the C-terminal of the *Bymovirus* P2-1 gene resembles the C-terminal part of the helper component proteinase (HC-Pro) gene found in the monopartite viruses (Adams *et al*., 2005a; 2005b). All members of the *Potyviridae* family are characterised by the production of pinwheel-shaped crystalline cylindrical inclusion bodies in the cytoplasm of infected cells, encoded by the viral CI protein, with these pinwheel-shaped inclusion bodies having been found in CBSD infected plants (Mbanzibwa *et al*., 2009a).

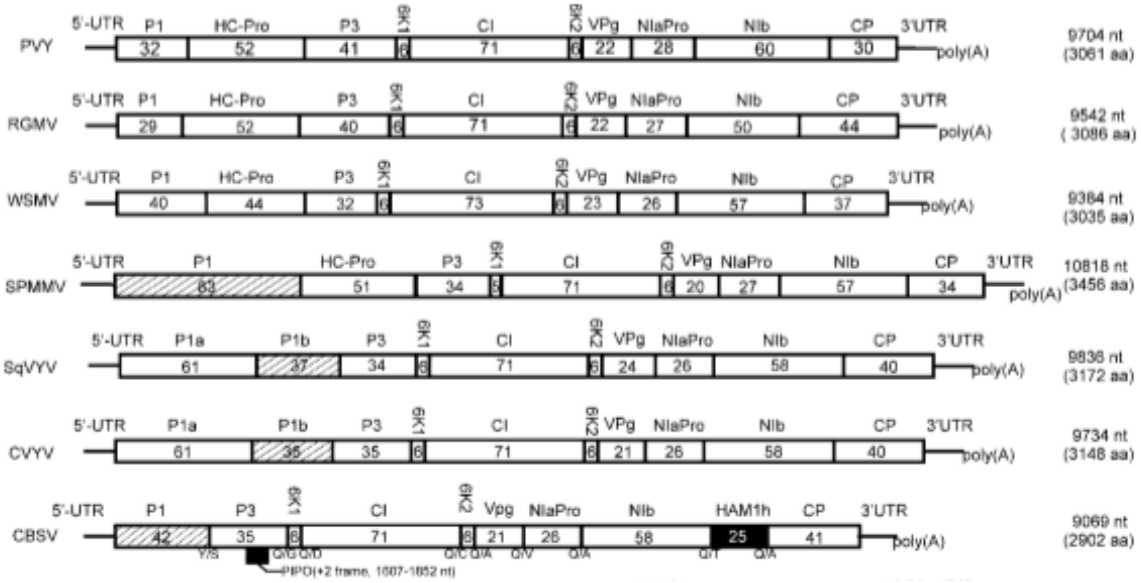

Figure 1.7. Comparison of the viral genomes of four genera of the family *Potyviridae*: *Potyvirus – Potato virus Y* (PVY), *Rymovirus – Ryegrass mosaic virus* (RGMV), *Tritimovirus – Wheat streak mosaic virus* (WSMV) and *Ipomovirus – Sweet potato mild mottle virus* (SPMMV). Further comparison between the genome structures of members of the *Ipomovirus* genus are shown, with *Squash vein yellowing virus* (SqVYV), *Cucumber vein yellowing virus* (CVYV) and *Cassava brown streak virus* (CBSV). Numbers indicate the estimated relative molecular weights of the proteins (kDa) (reproduced from Mbanzibwa *et al*., 2009b).

1.5 GENUS – *IPOMOVIRUS*

Ipomoviruses are transmitted by whiteflies and despite there only being six viral species within this genus, they are able to infect a range of plants from four families; from sweet potato (*Ipomoea batatas*, Family: Convolvulaceae), watermelon (*Citrullus lanatus*, Family: Cucurbitaceae), cucumber (*Cucumis sativus*, Family: Cucurbitaceae), squash (Family: Cucurbitaceae), tomato (*Solanum lycopersicum*, Family: Solanaceae) to cassava (*M. esculenta* Crantz Family: Euphorbiaceae), and cause serious losses to those crops. The single-stranded positive-sense RNA genomes of the *Ipomovirus* genus range in size from 9,069 nucleotides in length to 10,818 for CBSV and *Squash vein yellowing virus* (SqVYV) respectively. Like the rest of the *Potyviridae* family, *Ipomovirus* genomes encode a singular open reading frame (ORF) with the addition of a small protein known as PIPO that is produced occasionally when frameshift allows for the expression of a short overlapping ORF within the encoded P3 protein (Mbanzibwa *et al*., 2009b; 2011a; Dombrovsky *et al*., 2014). Coat protein (CP) sequence analysis of CBSV revealed that it shared 43.2% homology with the CP sequence of isolates of *Sweet potato mild mottle virus* (SPMMV), the type species of the *Ipomovirus* genus, confirming the virus' status in the genus (Monger *et al*., 2001a; Mbanzibwa *et al*., 2009a; 2011a).

Ipomoviruses are highly variable in the protein-encoding 5' region of their genomes (Fig.1.8) where the genus can be split into three groups depending on the presence and structure of the P1 and HC-Pro proteins. Only SPMMV and *Tomato mild mottle virus* (TomMMoV) contain the HC-Pro protein, which usually acts as a suppressor of RNA silencing-mediated host defence among other functions in other *Potyviridae* members (Webster and Adkins, 2012), and a P1 proteinase which is an enhancer of the HC-Pro RNA silencing suppressor activity (Webster and Adkins, 2012). Where this P1 protein is lacking in the SqVYV and *Cucumber vein yellowing virus* (CVYV) genomes, its role is instead covered by a second P1 proteinase, P1b, though its mode of action is significantly different to that of the P1 of SPMMV (Valli *et al*., 2008; Giner *et al*., 2010; Webster and Adkins, 2012). In the CBSV and UCBSV genomes, however, the HC-Pro protein is not present and nor are there the two P1 proteinases. These two viruses are thus the first members of the *Potyviridae* family to encode a single P1 proteinase and no HC-Pro (Mbanzibwa *et al*., 2009b). Sequence analysis by Mbanzibwa *et al*. (2009b) revealed that the P1 of CBSV is most closely related to the P1 of SPMMV and the P1b proteins of SqVYV and CVYV. Despite the CBSV P1 protein only showing sequence identity of 31 and 30% to the P1b proteins of CVYV and SqVYV respectively, the conserved histidine, aspartic acid and serine motif in the HDS triad of the P1 protein was observed in the CBSV genome, suggesting that the P1 protein of CBSV also functions as a suppressor of RNA silencing (Mbanzibwa *et al*., 2009b; 2011a). It is thought that the high variability in structure at the 5'-end of *Ipomovirus* genomes implies that recombination has been

vital to the P1 evolution and has aided in the facilitation of the viruses to adapt to a wide range of hosts (Valli *et al*., 2007; Mbanzibwa, *et al*., 2009b; 2011a; Dombrovsky *et al*., 2014).

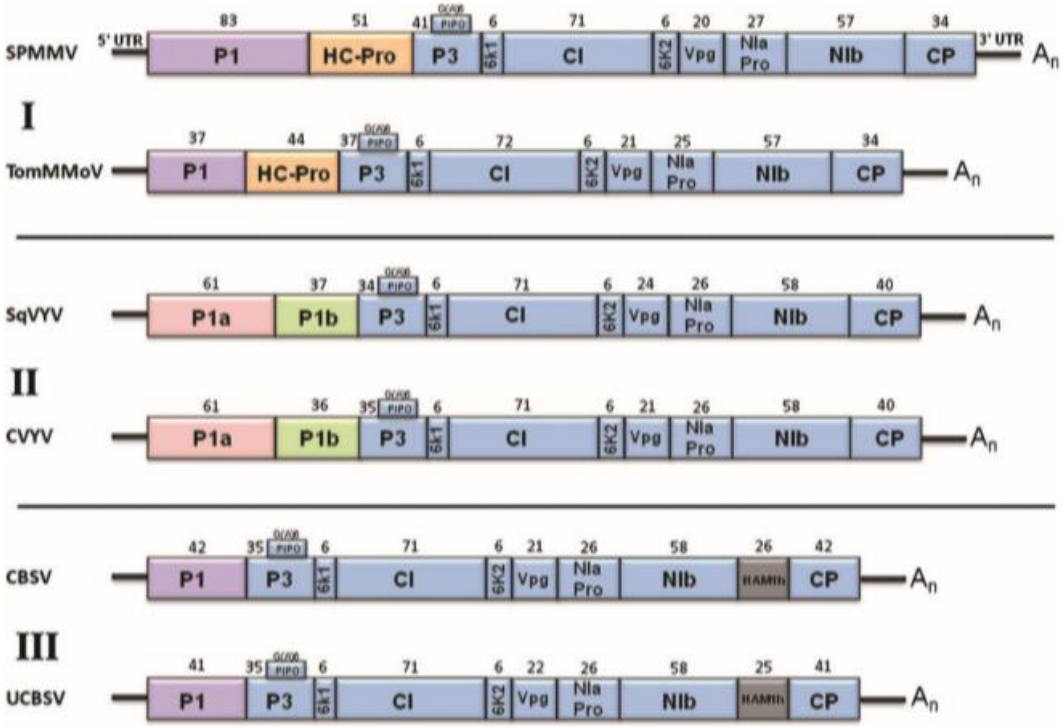

Figure 1.8. Diagram showing the different genomic structures of the viruses belong to the *Ipomovirus* genus and how they can be separated into three groups based on the genomic organisation (reproduced from Dombrovsky *et al*., 2014).

1.6 CBSV AND UCBSV VIRAL PROTEINS

The *Ipomovirus* RNA genomes are translated as polyproteins and then cleaved by the encoded proteases to form ten mature proteins (Mbanzibwa *et al*., 2011a). The following descriptions of the proteins and their functions are based on those belonging to other viruses in the *Potyviridae* family that have been studied and outlined, but the functions of the proteins found in the CBSV and UCBSV genomes are likely to be similar.

1.6.1 P1 PROTEIN

The P1 protein, the most variable part of the genome, is a serine protease that cleaves itself from the polyprotein at its C-terminus to function *in trans* to stimulate genome amplification, with it also being shown that *in vitro*, the P1 proteinase can interact with nucleic acids (Adams *et al*., 2005a; 2005b; Dombrovsky *et al*., 2014; Revers and García, 2015). It also functions as a suppressor to RNA silencing, enhancing the action of HC-Pro, but the function of P1 is dependent on its separation from HC-Pro (Valli *et al*., 2006; Mbanzibwa *et al*., 2009b; Revers and García, 2015). The P1 protein is the most divergent potyviral protein due to large variability in its N-terminal region and subsequently,

negatively regulates P1 self-cleavage (Adams *et al*, 2005a; Valli *et al*., 2007; Pasin *et al*., 2014; Revers and García, 2015). Where the P1 protein is not present, two evolutionary diversified proteins exist in the genomes of SqVYV and CVYV, P1a and P1b (Carbonell *et al*., 2012; Tatinemi *et al*., 2012).

1.6.1.1 P1A

Sequence identity analysis has revealed that the P1a proteins of SqVYV and CVYV show high identity with the P1 proteins of the aphid transmitted potyviruses (Webster and Adkins, 2012), suggesting that this protein may function in the viral interactions with the insect vector.

1.6.1.2 P1B

Sequence analysis has shown that the P1b proteins of SqVYV and CVYV show they are more related to the P1 proteins of Tritimoviruses (Webster and Adkins, 2012), and as described in 1.4, function as suppressors to RNA silencing.

1.6.2 HELPER COMPONENT PROTEINASE (HC-PRO)

The HC-Pro is a multifunctional protein with its name deriving from its first discovered function as the Helper Component (HC) for aphid transmission (Govier *et al*., 1977; Revers and García, 2015). It is a cysteine protease that, like P1, cleaves itself at its C-terminus (Revers and García, 2015). The protein can be divided into three regions: the N-terminal (approximately 100 amino acids), the central region (approximately 250 amino acids) and the C-terminal (approximately 100 amino acids). The N-terminal is responsible for functions involving aphid transmission and the C-terminal contains the protease activity that allows for the self-cleavage at the C-terminus (Dombrovsky *et al*., 2014; Revers and García, 2015). Two domains, domain A and domain B, have been found in the central region of the protein which independently bind RNA and where most of the protein functions are based. Domain B contains highly conserved motifs that are implicated in RNA amplification and the systemic movement of the virus, with mutations of these motifs preventing such actions (Cronin *et al*., 1995; Urcuqui-Inchima *et al*., 2001; Revers and García, 2015). Along with its role of suppressing RNA silencing, enhanced by the P1 protein (Valli *et al*., 2006), the HC-Pro is required to stabilise the viral coat protein, increase viral titre and ensure the infectivity of viral progeny (Revers and García, 2015).

1.6.3 P3 PROTEIN

Little is still known about the mechanisms behind the function of the P3 protein, however, it is known that this protein is required for viral replication. The P3 protein does not seem to bind with viral RNA, but instead targets the membrane of the endoplasmic reticulum and forms inclusions associated with the Golgi apparatus. The P3 protein has also been found to associate with the pinwheel-shaped

inclusion bodies produced by the viral CI protein that are characteristic of the Potyviruses (Revers and García, 2015).

1.6.3.1 PIPO

PIPO is formed through the plus two frameshift of an overlapping ORF within the P3 gene sequence. The P3N-PIPO complex is a 25kDa protein that appears to play a role in viral cell-to-cell movement through association with the CI protein. The P3N-PIPO and CI proteins are transported to the plasmodesmata via the secretory pathway, where they coordinate to form conical structures that associate with the plasmodesmata to facilitate viral intercellular movement. Furthermore, CI mutations that impair the ability of the protein to associate with plasmodesmata, have prevented viral intercellular movement (Wei *et al*., 2010; Revers and García, 2015).

1.6.4 6K1 AND 6K2 PROTEINS

6K1 and 6K2 are small proteins, both only 52 amino acids long and each weighing 6kDa, and like the P3 protein, they do not bind to any viral RNA (Mbanzibwa *et al*., 2009b; Revers and García, 2015). 6K1 is hypothesised to be involved in viral infectivity and that the combination of P3-6K1 and its subsequent proteolytic process, may have a regulatory role specifically for the multiplication of the potyvirus (Revers and García, 2015).

The 6K2 protein, when a part of the 6K2-VPg-NIaPro complex, plays a crucial role in viral RNA replication (Revers and García, 2015).

1.6.5 CYLINDRICAL INCLUSION PROTEIN (CI)

The CI protein is responsible for forming the characteristic pinwheel-shaped inclusion bodies of Potyviral infection. The CI protein is also responsible in viral replication by providing ATPase and RNA helicase activities and plays an important role in viral cell-to-cell movement. The CI protein associates with the P3N-PIPO complex to aid viral movement, though the mechanism is not yet known. The CI protein also binds to the end of the viral genome where it is thought to provide a motor function to facilitate viral movement through the plasmodesmata (Revers and García, 2015).

1.6.6 NUCLEAR INCLUSION PROTEIN (NIA)

The NIa is the largest protein found in *Ipomovirus* genomes and forms a crystalline inclusion in the host, most commonly in the nucleus but they can also appear in the cytoplasm of infected host cells. NIa in its whole form can bind to the 5'-end of the *Ipomovirus* genome, localising both in the cytoplasm and the nucleus, and exhibits NTPase activity. The function of this NTPase activity is still unknown (Revers and García, 2015).

The protein NIa can also be cleaved to produce two proteins: VPg at the amino (N) terminus and NIaPro at the carboxy (C) terminus. Cleaving is quite inefficient and so implies that both intact NIa protein and the processed VPg and NIaPro proteins coexist in the infected cells (Martínez *et al*., 2016).

1.6.6.1 Viral Genome-Linked Protein (VPg)

The VPg protein is able to interact with many different proteins, including itself, and thus is involved in many processes including viral replication, translation and movement (Revers and García, 2015; Martínez *et al*., 2016). The VPg protein on its own, forms a covalent link with the 5'-end of the *Ipomovirus* genome by means of a Tyr residue and consequently binds a host translation initiation factor that is required for infection (Martínez *et al*., 2016). When the VPg protein forms part of the 6K2-VPg-NIaPro complex, it instead targets viral membranous factories where it plays a vital role in viral RNA replication. There has also been evidence that VPg could undergo phosphorylation and this post-translational modification could provide the regulation for the various roles this protein is involved in (Revers and García, 2015).

1.6.6.2 NIaPro

NIaPro is a serine protease that cleaves itself from the remainder of the viral polyprotein before cleaving the viral polyprotein in both the cis and trans directions at specific conserved motifs to generate most of the viral proteins. It also exhibits RNA binding associating with the NIb protein, the viral RNA-dependent RNA polymerase, and is involved in viral replication (Martínez *et al*., 2016). Furthermore, NIaPro provides DNase activity and it is thought that when the NIa protein is in the host nucleus, this could provide some regulation in host gene expression which is beneficial for the viral infection (Revers and García, 2015).

1.6.7 Nuclear Inclusion Protein b (NIb)

The NIb protein is the RNA-dependent RNA polymerase that is crucial for the replication of the viral genome. NIb interacts with VPg and NIaPro as part of the VPg-6K2-NIaPro complex on membranous structures where it uridylylates VPg to prime for viral RNA synthesis and start the process of viral replication. Further interaction of the NIb protein with selected host proteins, allows for the creation of replication complexes contributing to the process of the viral RNA replication. NIb also forms nuclear inclusion bodies with the NIa protein where it is thought that a possible interaction with a host enzyme may bring about the regulation of NIb activity in the nucleus while also allowing NIb to make the environment in the cell more favourable for viral replication, although the role that the NIb protein plays in the nucleus is still largely unknown (Revers and García, 2015).

1.6.8 HAM1H PROTEIN

The presence of a 226 amino acid sequence, the HAM1 homologue (HAM1h), between the viral replicase (NIb) and coat protein (CP) in the CBSV and UCBSV genomes is unique among the *Ipomovirus* genus. Only one other virus is known to contain a HAM1h protein and that is the *Euphorbia ringspot virus* (EuRSV). Despite this virus also infecting a *Euphorbiacea* plant, EuRSV, however, belongs to the *Potyvirus* genus. The HAM1h proteins found in these viral species are homologous to the Maf/HAM1 superfamily found in prokaryotes and eukaryotes where these proteins function as nucleoside triphosphate pyrophosphatases (NTPases), reducing the rates of mutation during nucleic acid replication by hydrolysing non-canonical bases in nucleotide precursor pools. The HAM1h function in these viruses remains unknown, but it is suggested they function in reducing mutation rates of the viral RNA (Monger *et al*., 2010; Mbanzibwa *et al*., 2011a).

1.6.9 COAT PROTEIN (CP)

The coat protein's primary role is to encapsulate the viral genome with around 2,000 helical subunits required to enclose the Potyviral RNA. The coat protein also functions in aphid transmission, viral assembly and viral cell-to-cell movement (Revers and García, 2015; Urcuqui-Inchima *et al*., 2001). The central region of Potyvirus coat proteins is highly conserved while the N- terminal is highly variable and the C- terminal variable, and it is these terminals that interact together in the initiation of viral assembly. The N-terminal is exposed at the viral particle surface, and as this region is highly variable, may allow for the different methods of vector transmission among Potyvirus species. Furthermore, it has been shown that the coat protein can undergo post-translational modifications, among them phosphorylation, and it is hypothesised that these changes may regulate the allocation of viral RNA for translation, replication or propagation. Moreover, the coat protein has NTPase activity that may also regulate viral replication and propagation (Revers and García, 2015).

1.7 CBSV AND UCBSV HAM1H

1.7.1 BACKGROUND

The HAM1h protein found in the CBSV and UCBSV genomes, situated between the viral replicase (NIb) protein and coat protein (CP) in the C-proximal part of the polyprotein, is unique to the *Ipomovirus* genus. EuRV is the only other virus to contain such a protein and it also infects plants in the *Euphorbiaceae* family. Together, CBSV, UCBSV and EuRV, represent a small proportion of viruses that are able to infect *Euphorbiacea* plants, suggesting that the HAM1 homologues are a euphorbia host adaptation (Monger *et al*., 2010; Tomlinson *et al*., 2017). Sequence identity between the CBSD-causing viruses and EuRV (34% amino acid identity) is lower than when compared to eukaryotic HAM1

homologues (Monger *et al*., 2010; Winter *et al*., 2010; Tomlinson *et al*., 2017). One suggestion is that the integration of the HAM1 homologue from a eukaryotic host into these viruses occurred as two separate events. It is possible, however, that the uptake of the HAM1 homologue was an ancient event that occurred in a common ancestor and that EuRV and the CBSVs evolved separately. Similarly, as CBSV and UCBSV show relatively high levels of divergence, especially in the P1 and HAM1h regions where they share 59% and 47% amino acid identities respectively, the same two scenarios can be suggested for the evolution of the two CBSD viruses (Monger *et al*., 2010; Winter *et al.,* 2010; Tomlinson *et al*., 2017). Like the CBSV and UCBSV HAM1h proteins, the function of the EuRV HAM1h is unknown (Knierim *et al*., 2016; Tomlinson *et al*., 2017).

As the two viruses causing CBSD are only found in Africa, it is assumed that they evolved within East Africa in an unidentified species and subsequently jumped from the native host into cassava. Furthermore, as two species of CBSD-causing viruses exist, it may suggest that two jumps occurred independently, or they may represent the evolution of CBSV in cassava (Monger *et al*., 2010; Tomlinson *et al*., 2017). CBSV has been detected in the wild perennial species *M. glaziovii* and it has been shown that CMD was transferred from *M. glaziovii* to cassava through transmission by the whitefly vector, and the same could have occurred with CBSVs (Mbanzibwa *et al*., 2011b; Ogwok *et al*., 2014). It is thought there are other CBSD hosts that could serve as sources of viral inoculum that could then go on to adapt to infect cassava again (Monger *et al*., 2010; Tomlinson *et al*., 2017).

Sequence analysis has shown that the HAM1h proteins in the CBSV and UCBSV genomes are homologous to the Maf/HAM1 superfamily of NTPases found in organisms from bacteria, frogs, fish and humans, sharing some 33 highly conserved amino acid residues, and thus has given the viral proteins their name (Mbanzibwa *et al*., 2009a; 2009b; 2011a). The HAM1h proteins in CBSV and UCBSV are thought to have come from cellular origin through the uptake of cellular RNAs from host plants into the viral genome – a function that was fundamental in the evolution of viruses in the *Closteroviridae* family – and provides evidence that viruses of the *Potyviridae* family are capable of cellular RNA recombination in their genomes (Mbanzibwa *et al*., 2011a). Although the CBSV and UCBSV species are the only two *Ipomovirus* species to have this protein, the Ham1h is the most diverse region between the two species, sharing approximately only 47% and 51.1% sequence identity, depending on the isolates (Winter *et al*., 2010; Mbanzibwa *et al*., 2011a). Diversity of this region is 30% higher than the coat protein sequences which is the second-most diverse region between the two viral species (Mbanzibwa *et al*., 2011a), with around 80% sequence identity between CBSV and UCBSV (Winter *et al*., 2010).

Adaptive evolution between CBSV and UCBSV has been proposed through positive selection on the UCBSV HAM1h protein sequences and thus may be the reason why CBSD has been able to spread to new areas and may continue to do so over the coming years (Mbanzibwa *et al*., 2011a). Variation in the nucleotide sequences between isolates of CBSV show 79.3-95.5% identity and 86.3-99.3% identity between UCBSV isolates, though there has been no evidence of recombination between CBSV and UCBSV isolates despite the identification of putative homologous recombination sites (Mbanzibwa *et al*., 2011a; Ndunguru *et al*., 2015; Tomlinson *et al*., 2017)

The function of the CBSV and UCBSV HAM1h proteins are yet to be established, but they are thought to play a similar role to the NTPases found in prokaryotes and eukaryotes and prevent viral RNA mutation during replication. Evidence of purifying selection on the HAM1h sequences found in both CBSV and UCBSV indicate a desired retention of this protein, suggesting that the protein does play a key role in the viral infection. Based on the evidence that CBSV is the more virulent of the two viruses causing CBSD, work carried out in this project was directed towards studying the CBSV HAM1h.

1.7.2 Previous studies on the function of HAM1 proteins

The model organisms *S. cerevisiae* and *E. coli* have provided opportunities for the study of not only their own HAM1 and HAM1 homologue RdgB, respectively, but how their functions may elucidate the mechanisms behind a number of other HAM1 homologues from a variety of organisms, most notably humans. The yeast HAM1 and its homologues belong to the inosine triphosphate pyrophosphatases (ITPases), a superfamily of highly conserved nucleoside triphosphate pyrophosphatases (NTPases) (Galperin *et al*., 2006). ITPases have been shown to act upon non-canonical nucleotides in nucleotide precursor pools that have built up as a result of oxidative stress and prevent their incorporation into replicating nucleic acids, helping to reduce the rate of mutagenesis and ensure high-fidelity genomic replication (Savchenko *et al*., 2007; Stepchenkova *et al*., 2009b). The importance of these proteins is shown through their highly conserved functionality from bacteria to eukaryotes (Stepchenkova *et al*., 2009b).

Studies on the viral HAM1h homologue in *S. cerevisiae*, HAM1, have shown that the yeast *HAM1* gene does encode a nucleoside triphosphate pyrophosphatase and thus, hydrolyses non-canonical nucleotides. It was previously thought that the yeast HAM1 only hydrolyses non-canonical purine nucleotides, yet Carlsson *et al*. (2013) showed that the HAM1 specificity is broader than first thought, providing resistance to pyrimidine analogues as well. Further testing of the *E. coli* HAM1 homologue, RdgB, revealed that the highly conserved Serine-Histidine-Arginine (SHR) motif found in all HAM1 homologues, is located within the cleft of the protein and suggests that this motif may be forming part of the active site of all ITPases (Savchenko *et al*., 2007).

Such studies have drastically increased the knowledge of how the HAM1 and its homologues function in model organisms through the use of plate assays. Most assays conducted have involved the use of both wild type and HAM1h mutant forms of the model organisms to observe the effect of the applied mutagens on their respective phenotypes. The purine base analogue 6-*N*-hydroxylaminopurine (HAP) has often been used in the past as it is a potent mutagen for both prokaryotic and eukaryotic organisms (Carlsson *et al*., 2018). Studies with *S. cerevisiae* in particular, showed that omission of the *HAM1* gene conferred hypersensitivity to HAP, but that the HAM1 function can be restored through overexpression of the yeast HAM1 (Noskov *et al*., 1996; Kozmin *et al*., 1998).

As the CBSV HAM1h shows homology with these ITPases, it has been hypothesised that the viral protein also targets these non-canonical nucleotides. With the cassava hosts also expressing its own HAM1, it is thought that perhaps the viral HAM1h specifically targets RNA non-canonical nucleotides to prevent viral RNA mutagenesis, though no research into the interactions of the host and viral HAM1h proteins has been conducted (Monger *et al*., 2010; Mbanzibwa *et al*., 2011a). For CBSV, the HAM1h may be playing a vital role as the virus titre is highest in the mature cassava leaves and where the oxidative stress induced as a response to infection causes the early senescence of these leaves (Mbanzibwa *et al*., 2009b; Ogwok *et al*., 2014). The high level of oxidative stress within these cells causes damage to the nucleotides and generates high levels of non-canonical bases in the precursor pools. The CBSV HAM1h, therefore, may act upon the RNA non-canonical nucleotide bases to prevent their incorporation into the replicating viral RNA.

1.8 Resistance to CBSD

With sub-Saharan Africa experiencing some of the fastest population growths in the world, feeding this ever-increasing population has become an urgent issue (Legg *et al*., 2011). CBSD is currently managed through the propagation of disease-free stem cuttings and through the spread of CBSV-tolerant varieties, yet no robust source of resistance has been established (Blagbrough *et al*., 2010; Yadav *et al*., 2011). The lack of highly resistant CBSD cassava cultivars means that the risk of further CBSD spread remains high. Despite the generation and distribution of tolerant CBSD cultivars which display foliar symptoms but lack or exhibit mild root necrosis, they still remain susceptible to CBSD (Hillocks and Jennings, 2003; Tomlinson *et al*., 2017). Furthermore, the selection and breeding of resistant cassava varieties has proven a time-consuming and difficult process through the challenges encountered in resistant traits being introgressed into the farmer- and consumer-favoured germplasm (Jennings, 2003; Yadav *et al*., 2011). The deployment of a clean seed system is therefore imperative to prevent the spread of CBSD, but no such clean cassava seed systems exist in most East African countries (Tomlinson *et al*., 2017).

Post-transcriptional gene silencing (PTGS) using small interfering (si)RNAs is a promising biotechnology tool for the control of CBSD. It uses the plant's innate defence mechanism to bring about resistance to the selected viruses. In transgenic plants, this PTGS is induced through the action of RNA interference (RNAi) which triggers sequence-specific RNA degradation with the use of an artificial double-stranded (ds)RNA intermediate. This dsRNA intermediate, known as a hairpin, is homologous to a selected viral sequence and can be easily created through the fusion of the respective sequence in the sense and antisense orientations in one transcript (Vanderschuren *et al*., 2007; Yadav *et al*., 2011; Tatinemi *et al*., 2012). The hairpin is then cleaved by Dicer proteins and processed into 21-25 nucleotide siRNAs by the RNA-induced silencing complex (RISC). These siRNA molecules bind to the Argonaute (AGO) protein complex to direct the degradation of RNA molecules with sequences homologous to the siRNAs, thus conferring plant resistance to the virus (Yadav *et al*., 2011; Tatinemi *et al*., 2012; Kuria *et al*., 2017). This RNAi mechanism has already been applied and generated resistance to *Cucumber mosaic virus*, *Zucchini yellow mosaic virus*, *Watermelon mosaic virus 2*, *Potato leaf roll virus*, *Potato virus Y* and *Potato virus X*, *Papaya ring spot virus*, as well as *Plum pox virus* (Yadav *et al*., 2011).

Viral CP sequences are often used to induce plant resistance to the virus because of its highly conserved nature and the multiple roles it plays in the viral life cycle (Monger *et al*., 2001a; Vanderschuren *et al*., 2007; Yadav *et al*., 2011). Furthermore, it has been shown that the 3' end of the CP region is highly conserved amongst CBSV isolates (Monger *et al*. 2001b). The use of this RNAi-mediated resistance was initially demonstrated in *Nicotiana benthamiana*, a systemic host for CBSV infection, where one third of transgenic plants containing siRNAs against full length regions of CP sequences showed resistance to diverse isolates of CBSV and UCBSV (Patil *et al*., 2011). RNAi-mediated resistance has since been used in successfully generating transgenic cassava expressing siRNAs against the UCBSV CP sequence, which showed 100% resistance to UCBSV in replicated graft inoculation experiments (Yadav *et al*., 2011). Subsequently, The Virus Resistant Cassava for Africa (VIRCA) was established in 2012 to deliver cassava cultivars with enhanced CBSD resistance to farmers in East Africa through the use of RNAi mechanisms (Taylor *et al*., 2012; Wagaba *et al*., 2017).

Furthermore, Wagaba *et al*. (2017) demonstrated field level RNAi-mediated resistance to CBSD through expressing coat protein sequences from CBSV and UCBSV in the transgenic cassava cultivar TME 204. Not only did the transgenic cassava remain asymptomatic for CBSD, but the results demonstrated that this RNAi-mediated resistance could work across all agro-ecological locations where cassava is grown, and across the vegetative cropping cycle. This therefore provides a promising source for the development of further resistance to both CBSV and UCBSV.

Another aspect that could be used to limit CBSD infection within the cassava host is through the deposition of callose at the plasmodesmata, with susceptible cassava varieties showing reduced accumulation. Increased deposition of callose at the plasmodesmata was enough to restrict viral intercellular movement, though it is dependent on the cassava cultivar and viral isolate (Anjanappa *et al*., 2017). Furthermore, the development of full-length infectious clones of CBSV and UCBSV, which is ongoing, will enable the site-directed mutagenesis of key viral gene sequences and will aid in the elucidation of the protein functions. Ultimately, use of these clones will then be used to identify targets for the restriction of viral infection (Tomlinson *et al*., 2017).

1.9 AIMS OF THE RESEARCH

Previous work had demonstrated the ability of using model organisms to test the expression of ITPases on different strain phenotypes to determine the protein function. The use of spot assays had enabled mutations in specific genes to be identified that made the organism susceptible to the mutagenic effects of the accumulation of non-canonical nucleotides. Results of these assays can then, therefore, also provide a genetic source of resistance to the selected factors. In terms of disease management, knowledge of a viral protein function can be used to induce resistance in the host to combat infection. Further studies to elucidate the function of the CBSV HAM1h are therefore warranted as the spread of CBSD increases.

In order to determine the function of the CBSV HAM1h, one method was to test how its functionality compared to known HAM1 homologues using previously described spot assays. Firstly, the CBSV HAM1h will be overexpressed in wild type *S. cerevisiae* to test its effect on conferring increased resistance to a range of mutagens, and secondly, the effect of its expression in *ham1* mutant *S. cerevisiae* will be used to test if the HAM1 function of yeast can be complemented by the viral HAM1h, again when exposed to mutagens.

Given that the host plant of CBSV infection also contains a HAM1 protein, the reasoning behind the uptake and retention of the viral HAM1h remains to be fully understood. To date, no studies looking at the interaction of the host and viral *HAM1* expression levels have been conducted. This project thus presents a novel investigation into the effects of the transcript levels of the host *HAM1* upon infection with CBSV in the model host plant, *Nicotiana benthamiana*.

Determination of the role the CBSV HAM1h performs will allow for better understanding of where the viral HAM1h originated and why it has been retained. More importantly, such discoveries could potentially provide a target for resistance to CBSD, such as the generation of improved RNAi-mediated resistance.

CHAPTER 2 – MATERIALS AND METHODS

All plastic, glassware and media were autoclaved at 121°C for fifteen minutes before use. Sterile deionised water (SDW) was used to prepare all solutions unless otherwise stated. Chemicals and media used in this study were purchased from Melford, Sigma-Aldrich or Thermo Fisher Scientific unless otherwise stated. Primers were purchased from Integrated DNA Technologies. Commercial kits were purchased from Thermo Fisher Scientific or Zymo Research. All microbial culturing was performed in a class II biosafety cabinet using a Bunsen burner and sterilised material.

2.1 STRAINS

2.1.1 FUNGAL STRAINS

2.1.1.1 *SACCHAROMYCES CEREVISIAE*

Saccharomyces cerevisiae strain BY4742 (*MATalpha his3Δ leu2Δ lys2Δ ura3Δ*) was used as the wild type strain in all experiments and for use in yeast recombination experiments.

S. cerevisiae ham1 knockout deletion strain YJR069C MATalpha, in the background of BY4742 (*ham1*), was obtained from the Yeast Knockout Collective (YKO) from GE healthcare. The strain was used as the *ham1* mutant strain in all experiments and used for yeast recombination experiments.

2.1.2 BACTERIAL STRAINS

2.1.2.1 *ESCHERICHIA COLI*

Escherichia coli strain DH5α was used for the propagation of plasmids.

2.1.2.2 *AGROBACTERIUM TUMEFACIENS*

Agrobacterium tumefaciens strain LBA1126 was used for the propagation of the CBSV infectious clone and subsequent infection of *Nicotiana benthamiana* via agroinfiltration.

2.2 PLASMIDS

2.2.1 PYES2 VECTOR PLASMIDS

The plasmid pYES2 was used for its ability to express inserted genes in yeast, and for the presence of the URA3 gene to restore the URA function in *ura3* deletion strains. Plasmids containing the *HAM1* genes from CBSV (Nampula strain) and *S. cerevisiae* were generated by digestion of pYES2 and pJET_HAM1_Sc or pJET_HAM1_CBSV plasmids with *HindIII* and *SphI* restriction enzymes, then

subsequent ligation of the *HAM1* gene regions with ligated pYES2 using T4 DNA ligase, produced the plasmids pYES2_HAM1_Sc and pYES2_HAM1_CBSV. Plasmids were then transformed into electrocompetent *E. coli* DH5α cells. Recovery of the plasmids from the *E. coli* cells and subsequent digestion analysis, confirmed the correct plasmid construction.

2.2.2 C6 *CASSAVA BROWN STREAK VIRUS* INFECTIOUS CLONE

CBSV infected cassava material from Mikocheni Rd, Tanzania was amplified to produce six overlapping RT-PCR fragments covering the full CBSV Tanza genome. Instability of this clone when attempting to transform *E. coli*, meant that three plant introns were inserted into the P3, CI and NIb regions to provide sufficient sequence stability for it to be cloned successfully in yeast before propagation into the *E. coli* strain OverExpress C43. The full-length infectious clone was then cloned into a yeast adapted pCAMBIA0380 vector, which included the CaMV 35S promoter and the tNOS terminator to enable *in vivo* transcription (made by C. Duff-Farrier, PhD student). This infectious CBSV clone was used to transform the *A. tumefaciens* LBA1126 strain before use in agroinfiltration to inoculate *N. benthamiana* plants.

2.3 MEDIA

2.3.1 *SACCHAROMYCES CEREVISIAE* MEDIA

Yeast Peptone Dextrose (YPD) Broth (10g/L yeast extract, 20g/L Bacto-peptone, 20g/L D-glucose, 0.04g/L adenine sulphate) was used for the cultivation and maintenance of *S. cerevisiae*. YPD Agar (YPDA) was prepared with YPD and the addition of 20g/L agar, and used for the cultivation and maintenance of *S. cerevisiae*.

Yeast Synthetic Dropout Media (YSDM) Broth (1.7g/L yeast nitrogen base, 5g/L ammonium sulphate, 20g/L D-glucose, 0.77g/L yeast drop-out mix: -URA, pH 6.8) was used for the selection, cultivation and maintenance of the transformed *S. cerevisiae*. YSDM Agar was prepared with YSDM and the addition of 20g/L agar, and was used as a selection medium for the transformed *S. cerevisiae* and in the yeast plate-based assays. Transformation of the *ura3* deficient BY4742 and *ham1* knockout strains with the pYES2.1 plasmid that contains the orotidine 5-phosphate decarboxylase (URA3 gene), enabled their selection on media lacking uracil.

2.3.2 *ESCHERICHIA COLI* AND *AGROBACTERIUM TUMEFACIENS* MEDIA

Luria-Bertani (LB) Broth (25g/L Luria-Bertani) was used for the cultivation and maintenance of *E. coli* and *A. tumefaciens.*

Luria-Bertani (LB) Agar (25g/L Luria-Bertani, 15g/L agar) was used for the cultivation and maintenance of *E. coli* and *A. tumefaciens*.

2.3.3 ANTIBIOTIC SELECTION

Antibiotics were added to supplement the media to allow for the selection of transformants. The antibiotics selected depended on the plasmids used for transformation and the resistance genes they contained. Table 2.1 shows the antibiotics used, where they were purchased and how they were prepared.

Table 2.1. Details on the antibiotics used in this study.

Antibiotic	Supplier	Dissolved in	Stock (mg/ml)	Working (µg/ml)
Ampicillin	Sigma Aldrich	SDW	10	100
Kanamycin	Sigma Aldrich	SDW	5	50
Rifampicin	Melford	Methanol	20	20

Ampicillin was added to LB media to provide selection for transformed *E. coli* DH5α cells (LB Amp+), while Rifampicin and Kanamycin were added to LB media to provide selection for transformed *A. tumefaciens* LBA1126 cells (LB Rif+ Kan+).

2.4 CULTURE MAINTENANCE AND STORAGE

2.4.1 *SACCHAROMYCES CEREVISIAE* MAINTENANCE

S. cerevisiae were grown on YPDA, inverted and incubated at 28°C for two to three days. Liquid cultures were established in YPD by inoculating with a single colony taken from an agar plate using a flame-sterilised inoculating loop and incubated at 28°C overnight with 200rpm shaking. Transformed wild type and *ham1* knockout cultures were established and maintained in the same way, instead using YSDM.

2.4.2 *ESCHERICHIA COLI* MAINTENANCE

E. coli DH5α cultures were grown on LB plates with the appropriate antibiotic selection, inverted and incubated overnight at 37°C. Liquid cultures were grown by inoculating LB with the appropriate antibiotic selection, with a single colony taken from an agar plate using a flame-sterilised inoculating loop, and incubated at 37°C overnight with 250rpm shaking.

2.4.3 *AGROBACTERIUM TUMEFACIENS* MAINTENANCE

Agrobacterium tumefaciens cultures were grown on LB plates with the appropriate antibiotic selection, inverted and incubated overnight at 28°C. Liquid cultures were grown by inoculating LB with the appropriate antibiotic selection, with a single colony taken from an agar plate using a flame-sterilised inoculating loop, and incubated at 28°C overnight with 200rpm shaking.

2.4.4 *S. CEREVISIAE*, *E. COLI* AND *A. TUMEFACIENS* STORAGE

Cultures on agar plates or in liquid media were stored at 4°C for up to four weeks. Glycerol stocks (20%) were prepared for long-term storage through the addition of 500µl of 40% glycerol to 500µl of fresh microbial liquid culture, before rapidly frozen in liquid nitrogen and stored at -80°C.

2.4.5 VIRAL STORAGE

Infected *N. benthamiana* tissue samples were rapidly frozen in liquid nitrogen and stored at -80°C.

2.5 BUFFERS

2.5.1 AGROINFILTRATION MES BUFFER

A solution of 10mM morpholino ethane sulfonic acid (MES) buffer, with the pH adjusted to 5.7 with the addition of NaOH, was filter sterilised.

2.6 MOLECULAR METHODS

2.6.1 AGAROSE GEL ELECTROPHORESIS

2.6.1.1 AGAROSE GEL PREPARATION

A 1% (w/v) agarose solution was prepared in TAE buffer (40mM Tris acetate, 1mM EDTA, pH 8.0) for the visualisation of DNA fragment bands, and a 2% (w/v) agarose solution was prepared in TAE buffer for the visualisation of RNA fragment bands. Both solutions were heated until fully dissolved then cooled to between 50 and 60°C before the addition of Midori Green Nucleic Acid Staining Solution (Bulldog Bio). The solutions were poured into a UV-transparent gel tray (Bio-Rad), enclosed in a gel cast (Bio-Rad) and wells were formed using a gel comb. The gel combs were removed once the gels had set, the gels were placed in the electrophoresis tanks (Bio-Rad) and covered with TAE buffer.

2.6.1.2 LOADING AND RUNNING GELS

With the exception of PCR products generated using DreamTaq® DNA Polymerase (Thermo Fisher Scientific), a 5x loading dye (Bioline) was used for all nucleic acid samples. Hyperladder 1 (Bioline) (Fig. 2.1) was used as a molecular weight marker to deduce the size of the nucleic acid fragments and concentration, with 5µl loaded onto the gel alongside the samples. 5-10µl of each nucleic acid sample

was then loaded onto the gel and gels were typically run at 120V for thirty minutes on 1% agarose gels, or 80V for forty minutes on 2% agarose gels.

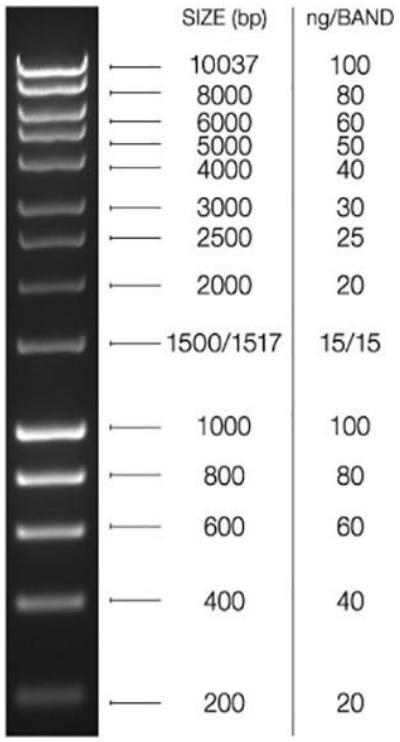

Figure 2.1. Hyperladder 1 (Bioline) is used as a molecular weight marker in gel electrophoresis (reproduced from https://www.bioline.com/uk/hyperladder-1kb.html).

2.6.1.3 VISUALISATION OF NUCLEIC ACIDS

Midori Green intercalates with the nucleic acids and fluoresces green under UV light, allowing for their visualisation. Visualisation was achieved using the Bio-Rad Gel Doc 2000 Imaging System.

2.6.2 PLASMID EXTRACTION

2.6.2.1 *ESCHERICHIA COLI*

A single colony of transformed *E. coli* was inoculated into 5ml LB with the appropriate antibiotic selection and grown at 37°C overnight with 250rpm shaking. The plasmid extraction was carried out in accordance with the Thermo Fisher Scientific GeneJET Plasmid Miniprep Kit, where the cultures were lysed, centrifuged, and the plasmid DNA, contained in the lysate, was bound to silica membranes under high salt conditions. The plasmid DNA was then purified through a series of washing steps before elution from the column.

2.6.2.2 *SACCHAROMYCES CEREVISIAE*

A single colony of transformed *S. cerevisiae* was inoculated into 5ml YSDM and grown at 28°C overnight with 200rpm shaking. The plasmid extraction was carried out in accordance with the Zymoprep™ Yeast Plasmid Miniprep I Kit with some minor adjustments: 3ml of the overnight culture

was used in the process instead of the stated 0.5-1.0ml; and following the addition of Zymolase, the mixtures were incubated for two hours instead of the stated 15-60 minutes.

2.6.3 Restriction enzyme digestion

Restriction enzyme digestion was carried out on the purified plasmids to verify the plasmids obtained, to check no transposons had been inserted and for use in yeast homologous recombination. Restriction enzymes were purchased from Thermo Fisher Scientific. 1µl of the chosen restriction enzyme was added to 4µl of the extracted plasmid. 2µl of the corresponding buffer solution was added and then 13µl of sterilised water. The solution was incubated at 37°C for thirty minutes, before being electrophorised on a gel at 100V for forty minutes to visualise the plasmid fragments.

2.6.4 Low fidelity Polymerase Chain Reaction (PCR)

Plasmids were used in PCR to check the plasmid had been taken up correctly in transformants and to amplify any genes of interest. Green DreamTaq DNA polymerase (Thermo Fisher Scientific) was used in analytical PCR. A 2X DreamTaq master mix was prepared containing 2ml of 10X buffer, 400µl dNTPs (40µl of each 10mM dNTP mixed thoroughly with 240µl SDW), 400µl DreamTaq DNA polymerase (5U/µl), made up to 10ml with SDW. All designed primers were ordered from Integrated DNA Technologies at stock concentrations of 100µM. Primers were diluted to 10µM with SDW prior to use in PCR.

DreamTaq PCR reaction volumes per sample were set up as follows:

- 10µl 2X DreamTaq
- 6µl SDW
- 1µl Forward 10µM Primer
- 1µl Reverse 10µM Primer

With the addition of 2µl of purified plasmid diluted 1:10 with SDW per sample.

DreamTaq PCR cycle:

- Initial denaturation – 2 minutes at 95°C
- Denaturation – 30 seconds at 94°C
- Annealing – 30 seconds at 55°C
- Extension – 1kb/min at 72°C

x 35

- Final extension – 10 minutes at 72°C

PCR products were then run on an agarose gel to visualise the products.

To screen *A. tumefaciens* cultures for the presence of transformant plasmids, the initial denaturation time at 95°C was increased from two minutes to ten minutes.

2.6.5 TRANSFORMATION OF ELECTROCOMPETENT *AGROBACTERIUM TUMEFACIENS* LBA1126

An aliquot of electrocompetent *A. tumefaciens* LBA1126 cells was removed from the -80°C stores and thawed on ice. Up to 2µl of low-salt DNA sample was added. The *A. tumefaciens*-DNA mix was transferred to an ice-cold sterile 0.2cm Bio-Rad electroporation cuvette. Cells were then electroporated using 400 Ohms, 2.5V in the Bio-Rad GenePulser. 900µl LB was then added to the cuvette and the mixture transferred to a 1.5ml Eppendorf tube. The solution was incubated for three hours in an orbital shaker at 28°C. An appropriate volume was then spread onto LB agar plates containing the appropriate antibiotic selection, inverted and incubated at 28°C for 48 hours.

2.6.6 AGROINFILTRATION

Three liquid cultures of the transformed *A. tumefaciens* LBA1126 were grown overnight. PCR was then carried out on these solutions to determine how many had successfully taken up the plasmid, with PCR carried out as described above. Once it had been established which of the three overnight cultures would be best for use in agroinfiltration, a 50µl aliquot of the culture was added to a fresh universal containing 10ml LB with 200µl 10mM MES buffer, 10µl Kanamycin, 10µl Rifampicin and 15µl acetosyringone. The solution was then incubated at 28°C with 200rpm shaking for 12-20 hours. 10ml of the culture solution was placed into a 50ml Falcon tube then centrifuged at 4°C at 6,000rpm for ten minutes. The supernatant was decanted, and the pellet washed with a mixture of 10ml SDW and 200µl MES buffer. The solution was centrifuged again at 4°C at 6,000rpm for ten minutes. The supernatant was decanted, and the pellet resuspended in a solution of 50ml SDW, 200µl MES buffer, 50µl $MgCl_2$ and 15µl acetosyringone. This solution was added a few millilitres at a time until the solution became less cloudy and about to clear. The culture was then incubated at room temperature for two to five hours. After incubation, the solution was then infiltrated onto the underside of *Nicotiana benthamiana* leaves using a 1 or 3ml syringe.

2.6.7 *SACCHAROMYCES CEREVISIAE* TRANSFORMATION

The LiAc/SS carrier DNA/PEG method, as described by Gietz and Schiestl (2007), for yeast homologous recombination was used to transform the *S. cerevisiae* cells.

Once the plasmid was extracted from *E. coli* and confirmed as correct through a plasmid restriction enzyme digestion, it was then used to transform *S. cerevisiae* BY4742 and *ham1* knockout cells. BY4742 and *ham1* knockout yeast liquid cultures were grown overnight at 28°C. The yeast starter cultures were poured into 40ml YPD and incubated for four and a half hours at 28°C with 200rpm

shaking. The cultures were then transferred to a 50ml Falcon tube and centrifuged for five minutes at 3,000 x g. The supernatant was discarded, and the cells resuspended in 25ml SDW and centrifuged again at 3,000 x g for five minutes. The supernatant was discarded, and the cells resuspended in 1ml 0.1M LiOAc and transferred to a 1.5ml Eppendorf tube. The cells were pelleted at 13,000rpm for fifteen seconds, and the supernatant removed. The cells were resuspended in 400µl of 0.1M LiOAc, vortexed and 50µl aliquots were transferred to new tubes, one per transformation. Cells were pelleted for a further fifteen seconds and the supernatant removed. Components were added to each transformation tube in the following order: 240µl 50% polyethylene glycol 3350 (PEG), 36µl 1M LiOAc, 50µl salmon-sperm DNA (boiled at 95°C for five minutes and then chilled on ice) and 34µl DNA fragments (10µl DNA:24µl SDW). The cells were mixed in this solution by pipetting for one minute. The cells were then incubated at 30°C for thirty minutes, followed by incubation at 42°C for a further thirty minutes. The cells were then pelleted at 6,000rpm for fifteen seconds and the supernatant removed. Cells were resuspended in 500µl sterile water by pipetting, and 100µl of the mixtures were spread onto YSDM plates before being inverted and incubated at 28°C for two to three days. Confirmation of transformation was established through growth of transformed cells on the YSDM agar plates.

2.6.8 RNA EXTRACTION USING TRIZOL

The TRIzol protocol was established by Chomczynski and Sacchi in 1987 and is now a common method of extracting total RNA from cells. It takes longer than column-based methods but yields more RNA that is generally of a better quality.

The frozen healthy and infected tissue samples were ground in sterile pestles and mortars with liquid nitrogen to a fine powder. The ground samples were vortexed with 0.75ml TRIzol per 0.25ml of sample tissue. The homogenised samples were incubated for five minutes at room temperature to enable the complete dissociation of the nucleoprotein complex. 0.2ml of chloroform was added per 1ml TRIzol added and shaken for fifteen seconds before incubation at room temperature for three minutes. The samples were centrifuged at 12,000 x g for fifteen minutes at 4°C. The upper aqueous phase was transferred to a new Eppendorf tube with 0.5ml of ice-cold isopropanol, and gently inverted before being incubated at room temperature for ten minutes. The addition of chloroform, centrifugation and removal of the upper aqueous phase was repeated to maximise the amount of RNA extracted. The incubated samples were centrifuged with isopropanol at 12,000 x g for ten minutes at 4°C. The isopropanol was removed, and the pellet washed with 1ml of 75% ethanol. The sample was vortexed, then centrifuged at 7,500 x g for five minutes at 4°C. The pellets were air dried for ten minutes then resuspended in 50µl DEPC treated water.

Extracted RNA (5µl) was run with 3µl of loading dye in a gel electrophoresis to check for the presence of RNA following the RNA extraction method.

2.6.9 First-strand cDNA synthesis

2.6.9.1 RNA concentration determination

To determine the concentration of extracted RNA in the sample, a nanodrop was used. The nanodrop measures the amount of ultraviolet light that passes through the nucleic acid sample, the optical density, and thus determines the concentration of nucleic acid in the sample. To obtain the concentration of RNA, 1µl of the sample was pipetted onto the nanodrop pedestal and the measurement taken. The DEPC water that the RNA pellet was resuspended in, was used as the control measurement. The nanodrop also provides measurements of the nucleic acid purity in the sample.

2.6.9.2 Removal of genomic DNA

In order to progress to the synthesis of first-strand cDNA, the genomic DNA was removed from the sample in accordance with the Thermo Fisher Scientific protocol to ensure cDNA was synthesised from the RNA extracted from the tissue samples. To remove the genomic DNA, 0.5 µl of DNase and 0.5µl of buffer solution was added to 4µl of the extracted RNA samples. This solution was incubated at 37°C for one hour. 0.5µl of EDTA was then added to the solution before incubation at 65°C for ten minutes.

2.6.9.3 First-strand cDNA synthesis

The concentration of RNA obtained from the nanodrop measurements was required in this process to ascertain how much had to be added in order that a similar concentration of cDNA was produced for all the samples. The reaction was conducted in accordance with the Thermo Fisher Scientific protocol:

- 2-5µl RNA (dependent on the sample concentration)
- 1µl oligo(dT) primer
- Nuclease-free water to a volume of 11µl

The solution was incubated at 65°C for five minutes. The following was then added:

- 4µl reaction buffer
- 1µl RiboLock RNase inhibitor
- 2µl 10mM dNTP mix
- 2µl reverse transcriptase

The solution was then incubated at 37°C for one hour.

2.6.9.4 cDNA CONCENTRATION DETERMINATION

The concentration of cDNA synthesised in each sample was measured using the nanodrop. The same process was carried out as above with determining the concentration of RNA in the sample, but nuclease-free water was used as the control measurement instead of the DEPC water used for RNA samples.

2.6.10 QUANTITATIVE POLYMERASE CHAIN REACTION (QPCR)

2.6.10.1 POLYMERASE CHAIN REACTION

To further check the cDNA samples synthesised and to check the affinity of the primers for the housekeeping gene and gene of interest, PCR with the selected primers was carried out on the cDNA samples.

The DreamTaq polymerase reaction was set up as detailed above with the addition of 2µl cDNA sample, and the PCR cycle carried out as follows:

- Initial denaturation – 2 minutes at 95°C
- Denaturation – 30 seconds at 94°C
- Annealing – 30 seconds at 52°C
- Extension – 25 seconds at 72°C

x 35

- Final extension – 10 minutes at 72°C

The PCR products were then visualised following gel electrophoresis on 2% agarose gel to allow for better separation of the DNA fragments.

2.6.10.2 QUANTITATIVE POLYMERASE CHAIN REACTION (QPCR)

Quantitative Polymerase Chain Reaction (qPCR) is used to ascertain the relative transcription levels of a gene of interest compared to that of a housekeeping gene. qPCR uses fluorescent reporters to quantify the generation of DNA in samples, with SYBR Green I dye being the most commonly used. The dyes fluoresce when intercalated with double-stranded DNA. The Thermo Fisher Scientific Maxima SYBR Green/ROX qPCR Master Mix (2X) was used, and the reaction set up as follows per sample with the addition of appropriate control samples:

- 12.5µl Maxima SYBR Green/ROX qPCR Master Mix (2X)
- 0.3µM Forward Primer
- 0.3µM Reverse Primer
- ≤500ng Template DNA
- Nuclease-free water up to 25µl

The PCR tubes were gently mixed, and the thermal cycler programmed as detailed below:

- Initial denaturation – 10 minutes at 95°C
- Denaturation – 30 seconds at 95°C
- Annealing – 1 minute at 52°C
- Extension – 1 minute at 72°C

x 40

- 1 minute at 95°C
- 30 seconds at 55°C
- 30 seconds at 95°C

The samples were placed in the cycler and the programme started. Data is acquired during the extension stage and is recorded as the cycle threshold (Ct) values on amplification plots.

qPCR products were then run on a 2% agarose gel to check the size of the DNA fragments and that they were of the expected size given the primers used.

2.7 PLATE-BASED ASSAYS

2.7.1 MUTAGEN STOCK CONCENTRATIONS

Mutagenic compounds were ordered from Sigma-Aldrich. Stock concentrations were made using dimethyl sulfoxide (DMSO) (purchased from Sigma-Aldrich) as the solvent, see Table 2.2. Solutions were made in a fume hood as highly toxic chemicals were used.

Table 2.2. Stock and working concentrations of mutagenic compounds.

Mutagen	Dissolved in	Stock (mg/ml)	Working (µg/ml)
Cycloheximide	DMSO	0.1	0.045, 0.080, 0.1, 0.13
5-fluoruracil	DMSO	10	10, 40, 60, 80
5-flucytosine	DMSO	10	20, 40, 60, 100
Methotrexate	DMSO	10	10, 40, 60, 80
Bromodeoxyuridine	DMSO	10	76, 153, 307, 460
6-azauracil	DMSO	10	500, 1000, 2000, 4000

2.7.2 PRODUCTION OF MEDIA

YSDM plates were made up, half of which contained 20% galactose which was added following the autoclaving of the media. The corresponding volume of stock solution for each mutagen was added to the autoclaved media to generate plates with the desired working concentration.

2.7.3 YEAST SPOTTING

Single colonies of the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed wild type and *ham1* knockout yeast were grown in YSDM overnight at 28°C with 200rpm shaking for 16-18 hours until all colonies showed stationary phase growth (OD_{600} > 1.5). Overnight cultures were diluted in fresh media to a final optical density of 0.1 (OD_{600} = 0.1) in 250ml conical flasks and were incubated again for a further 3-5 hours at 28°C with 200rpm shaking until the cultures were grown to mid-log phase (OD_{600} = 0.4-0.6). The cells were harvested by centrifugation for five minutes at 3,000 x g. The pellets were resuspended in 25ml SDW and centrifuged again as before. The pellets were then resuspended in 1ml SDW and transferred to 1.5ml Eppendorf tubes. The supernatant was discarded following centrifugation for fifteen seconds at 13,000rpm and the pellets resuspended in 1ml SDW. Four ten-fold serial dilutions were performed, giving the final cell concentrations of 1, 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}. 5µl aliquots of the serial yeast cell dilutions were spotted onto the YSDM plates (Fig. 2.2) with the corresponding concentrations of mutagen, and with or without the addition of 20% galactose.

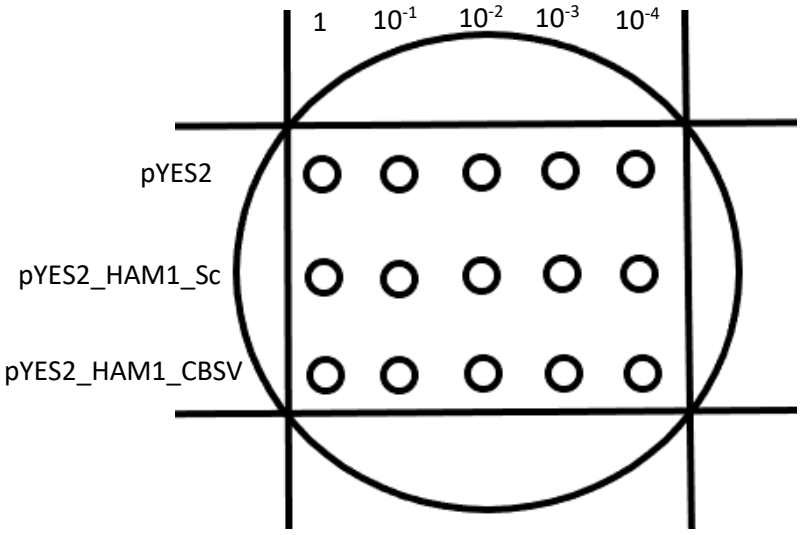

Figure 2.2. Diagram of the placement of the transformed yeast spots, from a cell concentration of 1 to a cell concentration of 10^{-4} dilution, on the YSDM plates.

2.8 PLANT METHODS

2.8.1 PLANT MAINTENANCE

N. benthamiana plants were maintained at 25°C in growth chambers on a 16-hour light/8-hour dark cycle and watered every three days.

2.8.2 Plant inoculation

2.8.2.1 Inoculation of *Cassava brown streak virus*

Inoculation of two-week-old *N. benthamiana* seedlings with *Cassava brown streak virus* was carried out using the Agroinfiltration method as described above. Detection of viral infection was monitored up to two weeks following inoculation by recording symptoms.

2.8.2.2 Inoculation of *Tobacco mosaic virus*

Inoculation of two-week-old *N. benthamiana* seedlings with *Tobacco mosaic virus* was carried out via sap inoculum. A TMV infected *N. benthamiana* leaf sample was ground to a fine powder in a pestle and mortar using liquid nitrogen. SDW was then added and mixed with the ground sample. Carborundum was dusted onto the selected healthy *N. benthamiana* leaf, and gloved fingers were used to gently rub the sap inoculum onto the leaf. The leaf was left to dry for ten minutes before excess carborundum was rinsed off using SDW.

2.8.3 Collection of RNA samples

Tissues samples were taken from healthy and infected *N. benthamiana* plants eighteen days post inoculation, using sterile gloves, and placed into labelled sterile Eppendorf tubes. The samples were immediately rapidly frozen in liquid nitrogen and stored at -80^{o}C.

CHAPTER 3 – INVESTIGATING THE FUNCTION OF THE CBSV HAM1H BY EXPRESSION IN YEAST

3.1 BACKGROUND INFORMATION

All organisms are subject to the effects of exogenous and endogenous nucleic acid damaging agents, which when accumulated, cause replication infidelity and genomic instability. One such causal agent of nucleic acid damage, is the accumulation of non-canonical nucleotides in cellular precursor pools (Menezes *et al*., 2012) brought about through the oxidation, deamination and other modifications of nucleotides (Kozmin *et al*., 2000; Galperin *et al*., 2006; Savchenko *et al*., 2007). Most often these non-canonical nucleotides are created naturally in the cell as by-products in the pathways of purine and pyrimidine synthesis, or they arise as a result of the cell being under oxidative stress (Simandan *et al*., 1998, Kozmin *et al*., 2000). This intracellular production of non-canonical (deoxy)nucleoside triphosphates ((d)NTPs), such as (deoxy)inosine triphosphate ((d)ITP) and (deoxy)xanthosine triphosphate ((d)XTP), can be incorporated into DNA and RNA during replication or transcription by the DNA and RNA polymerases (Savchenko *et al*., 2007). This incorporation of such bases is highly mutagenic due to the mispairing of nucleotides and can result in cell death as a result of the many nicks in the nucleic acid caused by the cell repair mechanisms (Kozmin *et al*., 2000). Indeed, this base mispairing can go unrecognised by the cell repair systems in both *E. coli* and *S. cerevisiae* (Kozmin *et al*., 1998). The mutagenic non-canonical nucleotides, therefore, must be removed from the precursor pools before their incorporation into synthesising nucleic acids.

Non-canonical NTPs are removed from the cellular nucleotide precursor pools by 'house-cleaning' enzymes, nucleoside triphosphate pyrophosphatases (NTPases). One such superfamily of NTPases is the highly conserved inosine triphosphate pyrophosphatases (ITPases), which contain the HAM1 protein and its homologues (Galperin *et al*., 2006). The importance of the role these proteins play in the organisms to ensure high fidelity genome amplification is demonstrated through their highly conserved functionality from bacteria and archaea, through to eukaryotes (Stepchenkova *et al*., 2009b).

Previous experiments have demonstrated the functionality of theses ITPases when tested against potent mutagens with the deletion or overexpression of the gene of interest using spot assay methodologies. In most of these investigations, 6-*N*-hydroxylaminopurine (HAP), a known mutagen for both prokaryotic and eukaryotic organisms, was used (Pavlov *et al*. 1996). HAP is an adenine base

analogue and due to its toxicity to bacteria, yeast and eukaryotic cells, it has been termed a universal mutagen (Kozmin *et al*., 2000). Initial studies carried out in *S. cerevisiae* showed that a *ham1* mutant strain was hypersensitive to HAP, though it was shown that the HAM1 function could be restored through overexpression of the yeast HAM1, and due to their highly conserved functionality, even as a result of the overexpression of HAM1 homologues from *E. coli* and humans (Noskov *et al*., 1996; Kozmin *et al*., 1998; Hwang *et al*., 1999; Stepchenkova *et al*., 2009b). Furthermore, Kozmin *et al*. (1998) showed that the overexpression of the yeast *HAM1* gene functions in *E. coli* just as it does naturally in yeast and that this overproduction of yeast HAM1 brought about a significant reduction in sensitivity to HAP. Work by Carlsson *et al*. (2013; 2018) showed that overexpression of the yeast *HAM1* in yeast *ham1* mutant strains, was able to confer increased resistance to pyrimidine analogues as well as purine analogues. Similarly, RdgB, which has been identified as the ITPase homologue in *E. coli*, does not discriminate between deoxyribonucleotides and ribonucleotides, and has demonstrated high nucleotide pyrophosphatase activity against three non-canonical NTPs (ITP, dITP and dXTP) while maintaining low pyrophosphatase activity against some eight canonical nucleotides (Savchenko *et al*., 2007). This discrimination was also found through the action of those ITPases found in archaea and eukaryotes, and thus suggests that the functional part of these proteins is highly conserved (Savchenko *et al*., 2007; Carlsson *et al*., 2013). Most ITPases in the HAM1 family contain a signature Serine-Histidine-Arginine (SHR) motif, and Savchenko *et al*. (2007) have already demonstrated that this signature HAM1 motif has been found inside a cleft that they propose forms part of the active site of the RdgB protein in *E. coli*. The CBSV HAM1h has been shown to contain this highly conserved SHR motif (Mbanzibwa *et al*., 2009b) and thus is speculated to function in much the same way as other ITPases.

3.2 AIMS

The aim of this investigation was to determine if the CBSV HAM1h functions in the same was as other ITPases by providing a direct comparative study to other known HAM1 homologues. It was proposed to use spot assays, as described previously by Carlsson *et al*. (2013), to provide a suitable experimentation technique to observe the effects of the expression and overexpression of the CBSV *HAM1h* on the phenotypes displayed in *ham1* mutant and wild type strains of *S. cerevisiae* respectively, when exposed to increasing concentrations of mutagens. Yeast was selected as the host model organism despite the fact that *E. coli* is more sensitive to a greater range of purine and pyrimidine analogues than eukaryotic cells (Kozmin *et al*., 1998), as the correct mutation in *E. coli* for use in this study was not available. In addition, HAP was no longer commercially available due to its highly toxic effects and so this mutagen could not be used in the spot assays. Instead, work by Carlsson *et al*. (2013) showed that the yeast HAM1 was able to confer increased resistance to a pyrimidine

analogue, 5-fluorouracil. Two other known mutagens that disrupt nucleic acid synthesis and replication, cycloheximide and methotrexate, were also used.

3.3 PLASMID DESIGN AND CONFIRMATION

HAM1 genes from CBSV and yeast were inserted into the pYES2 plasmid downstream of the inducible GAL1 promoter by ligation of sticky ends following digestion with restriction enzymes (supplied by J. Miller, PhD student). The correct insertion of the HAM1 genes in the pYES2_HAM1_Sc and the pYES2_HAM1_CBSV plasmid constructs were confirmed previously through use of PCR analysis with specific primer pairs for the respective genes, and through plasmid digests with selected restriction enzymes which have unique sites in the constructs. Stores of the plasmid constructs transformed into electrocompetent *E. coli* DH5α were used to subsequently extract the plasmid constructs using the Thermo Scientific GeneJet Plasmid Miniprep Kit according to the manufacturer's protocol. To further confirm the correct construction of the plasmids and to ensure that no contamination had occurred, a restriction digest was carried out. The digest uses restriction enzymes to cut the plasmid at specific points to generate fragments of certain sizes depending on the plasmid construct. For this, the restriction enzyme *ScaI* was used as it contains three specific sites in the pYES2_HAM1_Sc plasmid construct (Fig. 3.1.A), and only two in the pYES2_HAM1_CBSV plasmid construct (Fig. 3.1.B). The plasmid digest was also carried out the pYES2 empty vector construct, which will provide a control in the yeast spot assays, and like the pYES2_HAM1_CBSV plasmid construct, contains two sites specific to *ScaI*.

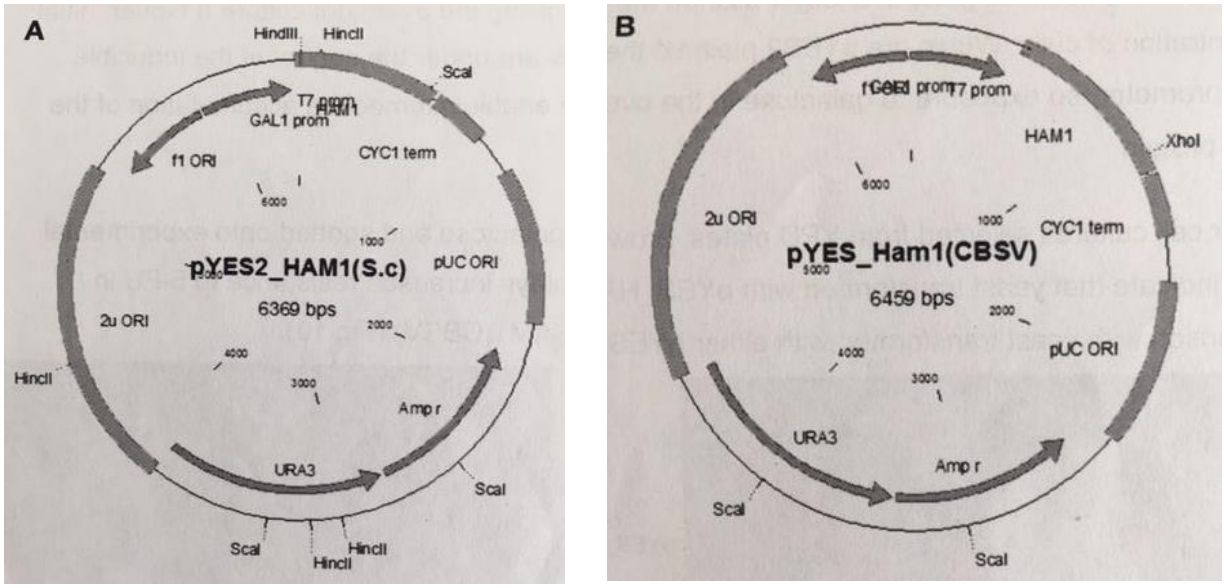

Figure 3.1. Plasmid maps of the constructs generated from the ligation of pYES2 and the *HAM1* genes from either *S. cerevisiae* or CBSV to give A. pYES2_HAM1_Sc and B. pYES2_HAM1_CBSV. Restriction enzymes used for the digestion of the plasmid constructs are shown, with *ScaI* being common to both.

Following digestion of the plasmid constructs with the restriction enzyme *ScaI*, the fragments were visualised using gel electrophoresis with Midori Green staining. Visualisation of the fragments after the subsequent digest of the pYES2_HAM1_Sc plasmid verified that three fragments were produced of sizes 3631 base pairs (bps), 1845 bps and 893 bps (Fig. 3.2). Furthermore, visualisation of the fragments following the digestion of the pYES2_HAM1_CBSV plasmid verified that two fragments were produced of sizes 5566 bps and 893 bps (Fig. 3.2).

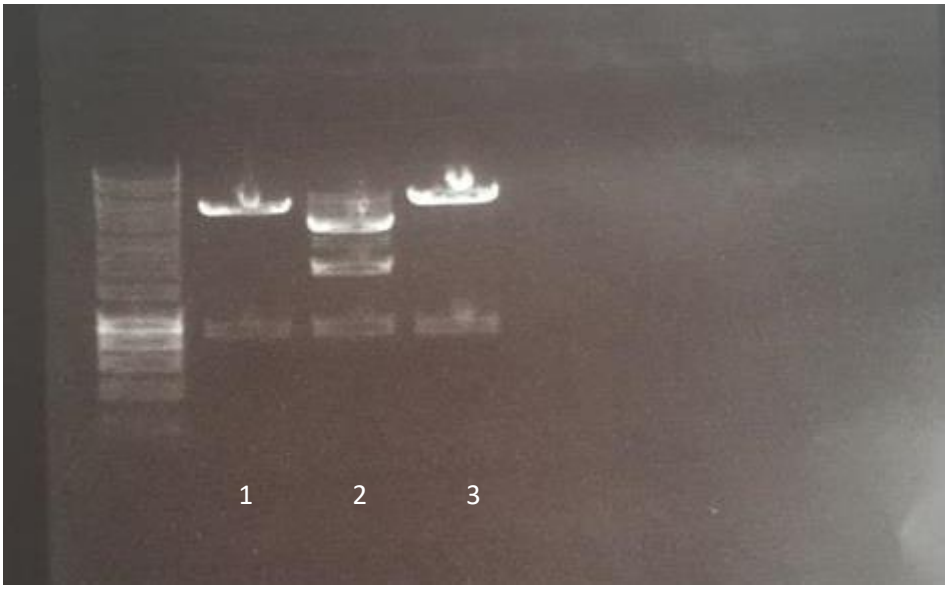

Figure 3.2. Gel electrophoresis of the fragments generated from the plasmid digest with the *ScaI* restriction enzyme of the pYES2 empty vector (lane 1), pYES2_HAM1_Sc (lane 2) and pYES2_HAM1_CBSV (lane 3) plasmid constructs recovered from transformed *E. coli*.

Following the confirmation that the plasmid constructs are all correct and have not been contaminated, the plasmids were transformed into wild type and *ham1* mutant *S. cerevisiae* strains. Confirmation of the transformation of the yeast cells was determined by the ability of the transformants to grow on synthetic drop-out media (YSDM) which lacks uracil. Only those yeast cells that had successfully taken up the plasmid constructs would be able to grow as the pYES2 vector contains the *URA3* gene. Yeast not expressing a *URA3* gene would not be able to grow on the YSDM. Successful transformants were then used in the yeast spot assays following four series of ten-fold cell dilutions.

3.4 SCORING OF THE SPOT ASSAYS

The serial dilutions of the transformed yeast cells were then spotted onto plates of YSDM media containing the varying concentrations of the chosen mutagens (GAL-), with half supplemented with 20% galactose (GAL+). Addition of galactose will induce the expression of the HAM1 genes, if present, transformed into the yeast cells, leading to either the overexpression of HAM1 in the wild type yeast cells, or expression of a HAM1 gene in the *ham1* mutant cells. Results of the yeast spot assays were then scored according to the number of colonies present and the intensity of growth. The number of cells present were scored from 1-5, with one representing the growth of one colony up to a number of five, and five representing the growth of visibly more than thirty colonies. The intensity of growth was scored from 1-4, with the differentiation as labelled in Fig. 3.3. The values scored for the growth of the transformed yeast strains at each serial dilution with increasing concentrations of the applied mutagens, were multiplied together to give one combined value that was more sensitive to the phenotypes displayed. This more sensitive method of scoring enabled subtle differences to become apparent that may otherwise have gone unnoticed.

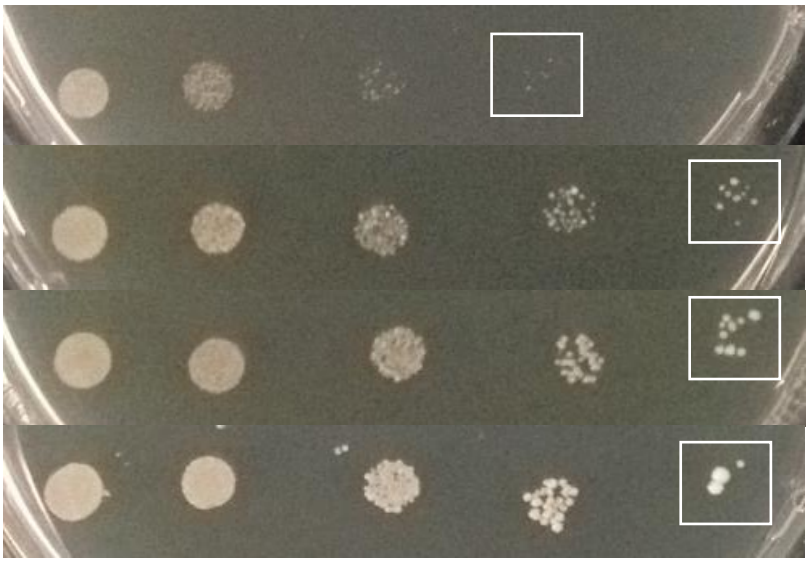

Figure 3.3. Key showing the scoring system from 1-4 of the growth intensities of the yeast cells, highlighted by the white boxes, used in the analysis of the yeast spot assay study.

Following on from the scoring of the yeast growth in the plate assays, statistical analyses were carried out to confirm if there were any significant differences between the growth of the yeast cells when grown in conditions with or without the supplemented galactose. Multiple two-way ANOVA analyses were determined to be most appropriate for these results.

3.5 Cycloheximide resistance assay

Cycloheximide (CHX) was first discovered as an anti-fungal compound produced by the gram-positive bacterium *Streptomyces griseus*, but has been later described as a glutarimide antibiotic. It is a cell-permeable molecule that specifically binds to the 60S ribosomal subunit where it inhibits eukaryotic cytosolic translation (Alamgir *et al*., 2010; Buchanan *et al*., 2016). As a result, the potent inhibition of the ribosome prevents translation elongation causing defects in protein synthesis, and when exposed to high levels of CHX, *S. cerevisiae* has been shown to cease protein synthesis and cell growth (Gerlinger *et al*., 1997; Alamgir *et al*., 2010; Schneider-Poetsch *et al*., 2010). At lower concentrations of CHX, however, it can be used to test the function of the yeast HAM1 and HAM1 homologues (Alamgir *et al*., 2010).

3.5.1 Overexpression of HAM1

At high cell densities, at the 10^{-2} cell dilution (Fig. 3.4), there was no readily apparent difference in growth caused by the overexpression of either the yeast or CBSV *HAM1* genes when induced by the supplemented galactose. Despite an increase in growth of those wild type cells overexpressing *HAM1* genes from yeast and CBSV under the influence of galactose from CHX concentrations of 80, 100 and 130ng/ml, cells transformed with the empty pYES2 vector, and thus not overexpressing a *HAM1* gene, also showed increases in growth. Indeed, at the CHX concentration of 80ng/ml, all the transformed cells showed significant differences in cell growth with the added galactose (pYES2, $p = 0.0161$; HAM1_Sc, $p = 0.0161$; HAM1_CBSV, $p = 0.0399$). At the CHX concentration of 100ng/ml, however, none of the transformed cells showed a significant difference in growth (pYES2, $p = 0.374$; HAM1_Sc, $p = 0.116$; HAM1_CBSV, $p = 0.374$). There was, however, one significant result that was not also observed in the pYES2 transformed cells. At the CHX concentration of 130ng/ml, a significant difference in the cell growth of the pYES2_HAM1_Sc transformed cells was observed with the addition of the 20% galactose ($p < 0.001$). Despite this one significant result, however, there seems to be no readily apparent increase in resistance to the toxic effects of CHX from overexpressing the yeast or CBSV *HAM1* gene at this cell dilution.

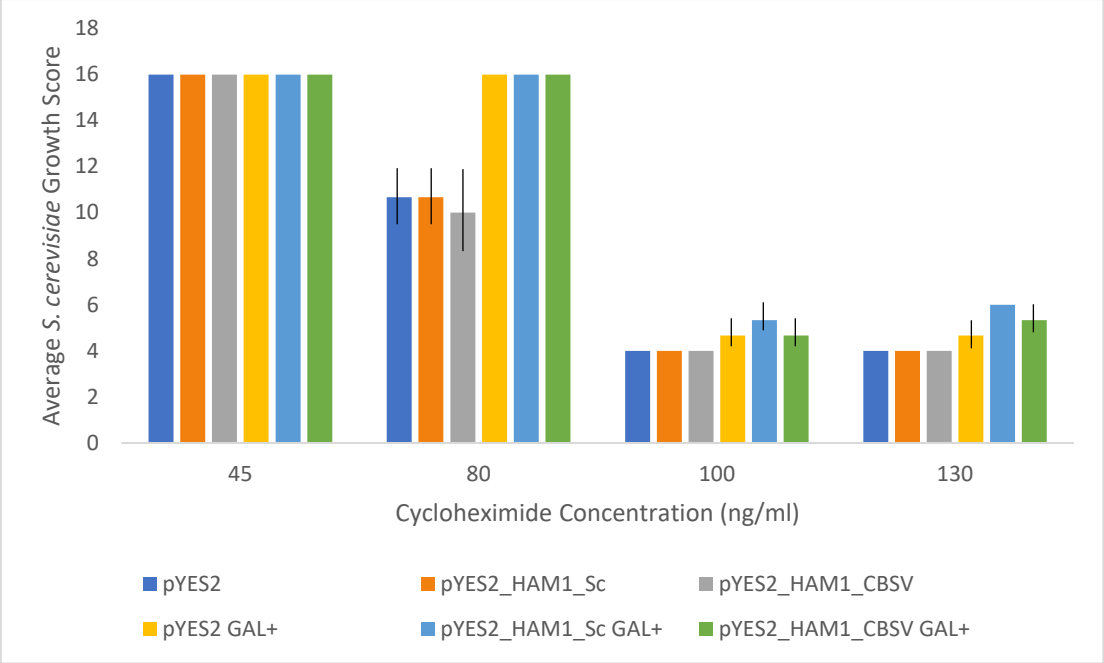

Figure 3.4. Bar chart showing the average growth scores (n=3) at the 10^{-2} cell dilution of the wild type *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids with increasing concentration of cycloheximide compound, with and without supplemented galactose.

At lower cell densities, when considering resistance to CHX at the concentration of 45ng/ml, it becomes apparent that those yeast cells overexpressing the *HAM1* genes from yeast and CBSV show increased growth, and therefore increased resistance to the mutagen, compared to those transformed with the empty pYES2 vector at the 10^{-3} and 10^{-4} cell dilutions (Fig. 3.5), which saw no effect on their growth upon the addition of galactose to the media. Indeed, statistical analysis of the results has shown that the growth of the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed cells was significantly different with the addition of galactose at these cell dilutions (HAM1_Sc, $p < 0.001$; HAM1_CBSV, $p < 0.001$). These results thus show that overexpression of a yeast *HAM1* gene does confer increased resistance to the toxic effects of CHX and that the CBSV HAM1 shows similar functionality to the yeast HAM1.

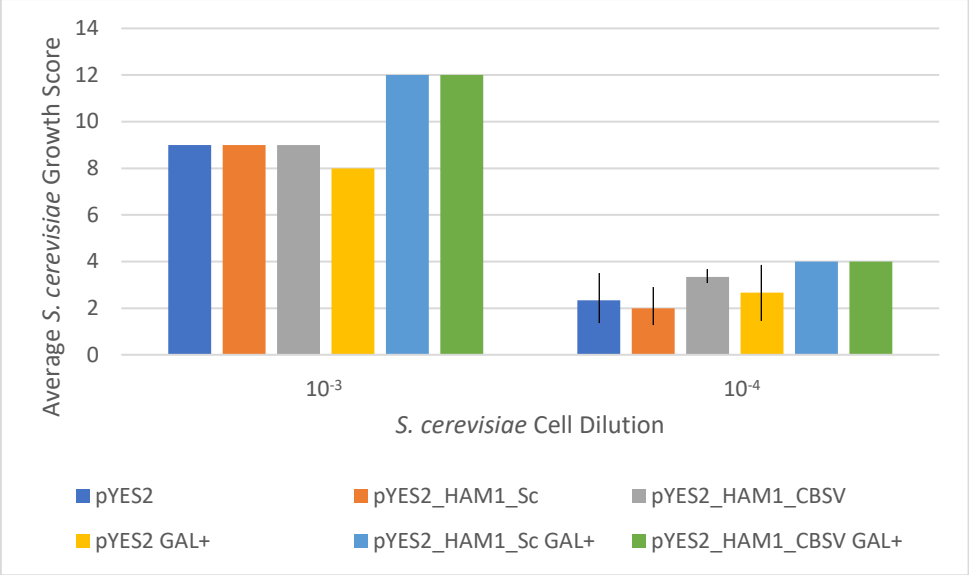

Figure 3.5. Bar chart showing the average growth scores (n=3) of the wild type *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} and 10^{-4} cell dilutions at a concentration of 45ng/ml cycloheximide, with and without supplemented galactose.

Focusing on just the 10^{-3} cell dilution of yeast (Fig. 3.6), the wild type yeast cells overexpressing the *HAM1* genes from *S. cerevisiae* and CBSV under the induction of galactose, show an increase in growth across all concentrations of CHX, with the differences in growth being significant at the CHX concentration of 45ng/ml (HAM1_Sc, $p < 0.001$; HAM1_CBSV, $p < 0.001$). At the CHX concentration of 80ng/ml, however, the marked increase in growth recorded for cells overexpressing the yeast *HAM1* gene is also observed for those cells transformed with the empty pYES2 vector, thus, no apparent effect of overexpressing a *HAM1* gene can be deduced, with any differences in growth analysed as not being significant (pYES2, $p = 0.0161$; HAM1_Sc, $p = 0.0907$; HAM1_CBSV, $p = 0.238$). Similarly, at the CHX concentration of 100ng/ml, there is an increase in growth recorded for the cells not overexpressing a *HAM1* gene with the addition of galactose, as well as an increase in growth of those cells expressing the CBSV *HAM1* gene, although these differences are not significant (pYES2, $p = 0.374$; HAM1_CBSV, $p = 0.374$). Yet as this CHX concentration, there is a significant difference in growth observed for those cells now overexpressing the yeast *HAM1* (HAM1_Sc, $p < 0.001$). Again, no apparent effect of the overexpression of the CBSV *HAM1* gene can be deduced for this serial dilution at this CHX concentration. At the CHX concentration of 130ng/ml, an effect of the overexpression of the yeast and CBSV *HAM1* genes can be seen with increases in growth score observed for both these transformed cells under the induction of galactose and comparably, there is no increase for the wild type cells transformed with the empty pYES2 vector. The increases in growth observed by those cells overexpressing the yeast and CBSV *HAM1* genes at the CHX concentration of 130ng/ml, are much reduced than at the CHX concentration of 45ng/ml and that is most likely due to the various mechanisms in which CHX can exert its toxic effects and thus, the action of HAM1 alone at these higher

concentrations of CHX may not be enough to afford full protection to the mutagen, hence why the increases in growth scores were likely to not be significant (HAM1_Sc, p = 0.116; HAM1_CBSV, p = 0.374). This effect was also seen at the 10^{-4} cell dilution across all concentrations of CHX.

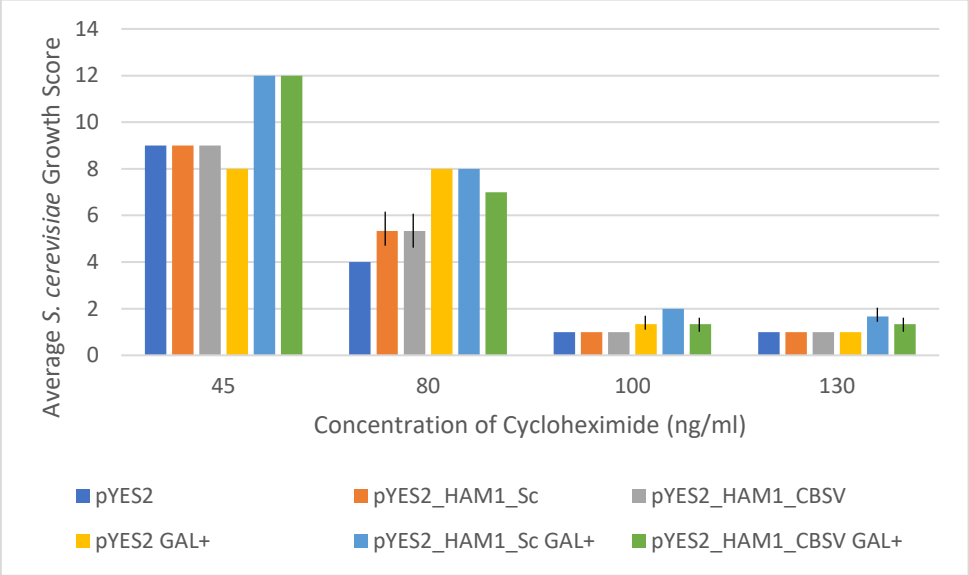

Figure 3.6. Bar chart showing the average growth scores (n=3) of the wild type *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} cell dilution across all concentrations of cycloheximide, with and without supplemented galactose.

3.5.2 Complementation of ΔHAM1

Addition of 20% galactose to the *ham1* knockout cells had no readily apparent effect at the CHX concentration of 45ng/ml with no increases in growth observed as a result of the expression of the yeast or CBSV *HAM1* genes (Fig. 3.7.A). An increase in growth was seen for the *ham1* knockout cells transformed with the pYES2 and pYES2_HAM1_CBSV plasmids at the 10^{-2} and 10^{-3} cell dilutions at 80ng/ml CHX concentration (Fig. 3.7.B), yet these results show no apparent effect on resistance to CHX when expressing a *HAM1* gene (HAM1_CBSV, p = 0.374 and p = 0.116, respectively). Increases in growth were also seen for those cells transformed with the pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} cell dilution at 100ng/ml CHX concentration, and for the pYES2_HAM1_Sc transformed cells at the 10^{-4} cell dilution also at this CHX concentration (Fig. 3.7.C). A large increase was seen in the growth of the pYES2_HAM1_Sc transformed cells at the 10^{-4} cell dilution at the CHX concentration of 130ng/ml with the addition of 20% galactose, with no growth observed for the pYES2 transformed cells (Fig. 3. 7.D).

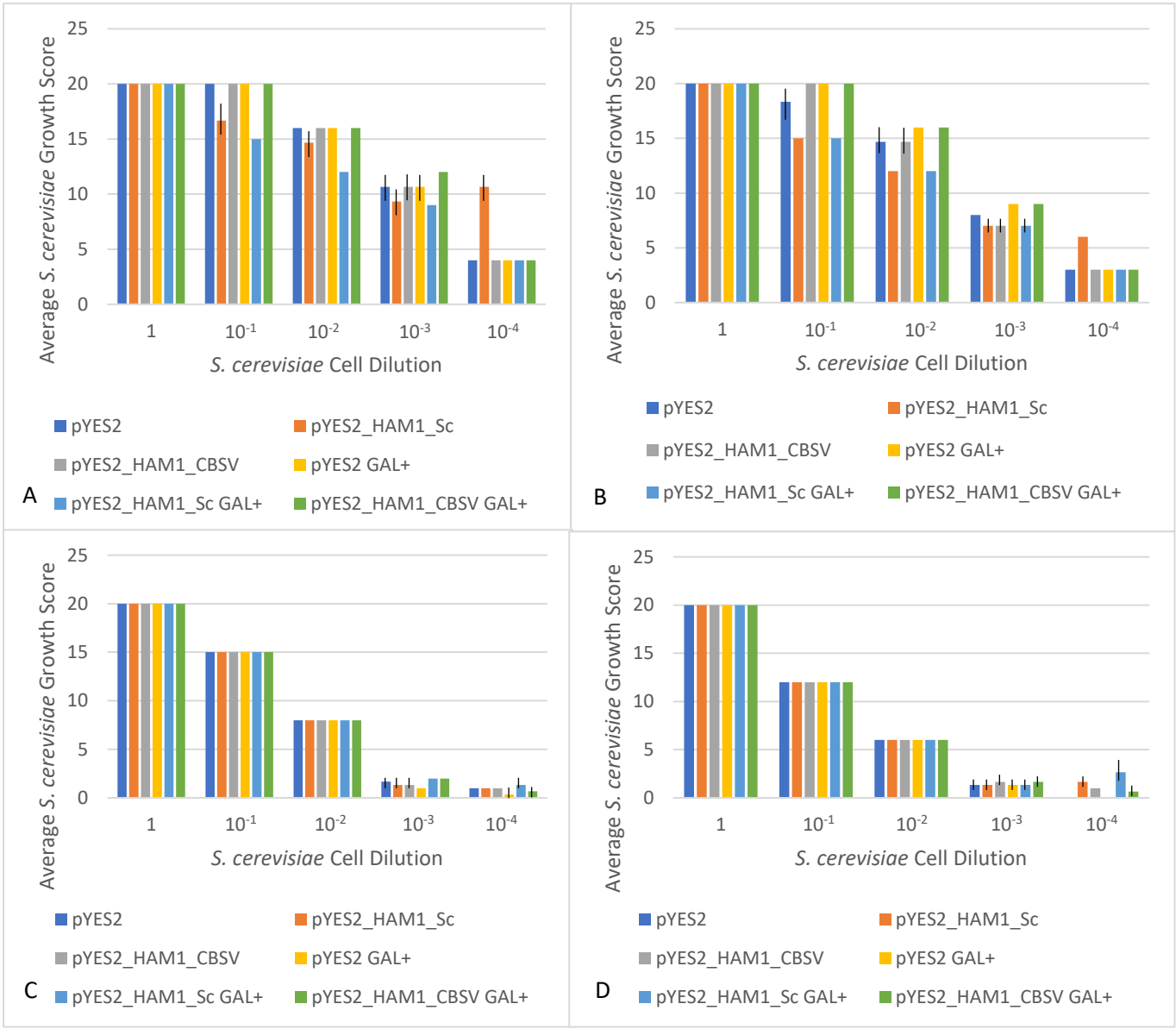

Figure 3.7. Bar charts showing the average growth scores (n=3) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids across all serial dilutions at (A) 45ng/ml (B) 80ng/ml (C) 100ng/ml and (D) 130ng/ml cycloheximide concentration, with and without supplemented galactose.

At the 10^{-3} cell dilution at a CHX concentration of 100ng/ml, cells transformed with the pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids showed an increase in growth scores on addition of 20% galactose to the agar (Fig. 3.8), though these differences were not significant (HAM1_Sc, p = 0.116; HAM1_CBSV, p = 0.116). Although not significant, these increases in growth scores were not seen for those *ham1* knockout cells transformed with the empty pYES2 vector, where there was a decrease in growth scores with the galactose supplemented agar, and thus shows that the expression of a *HAM1* gene, whether from yeast or CBSV, does suggest an increased resistance to CHX under this condition. Further replicates in this study, however, would need to be conducted to confirm this trend. The increase in growth scores for both the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed cells, furthermore, suggests that the CBSV HAM1 does complement the yeast HAM1.

At the 10^{-4} cell dilution at the CHX concentration of 100ng/ml, the pYES2_HAM1_Sc transformed cells showed an increase in growth under the influence of the supplemented galactose (Fig. 3.8). Although not significant (HAM1_Sc, p = 0.374), this result again shows that with more replicates, the expression of the *HAM1* gene does seem to confer an increased resistance to CHX. Cells transformed with the empty pYES2 vector showed no reaction to the addition of galactose and growth decreased to an average score of 0.3 (pYES2, p = 0.116). Despite a reduction in average growth score also being observed for the pYES2_HAM1_CBSV transformed cells with the addition of galactose (HAM1_CBSV, p = 0.374), this reduction in growth was not to the same extent as those cells transformed with the empty vector, thus showing that the expression of the CBSV *HAM1* gene does still afford some protection against the cytotoxic effects of CHX.

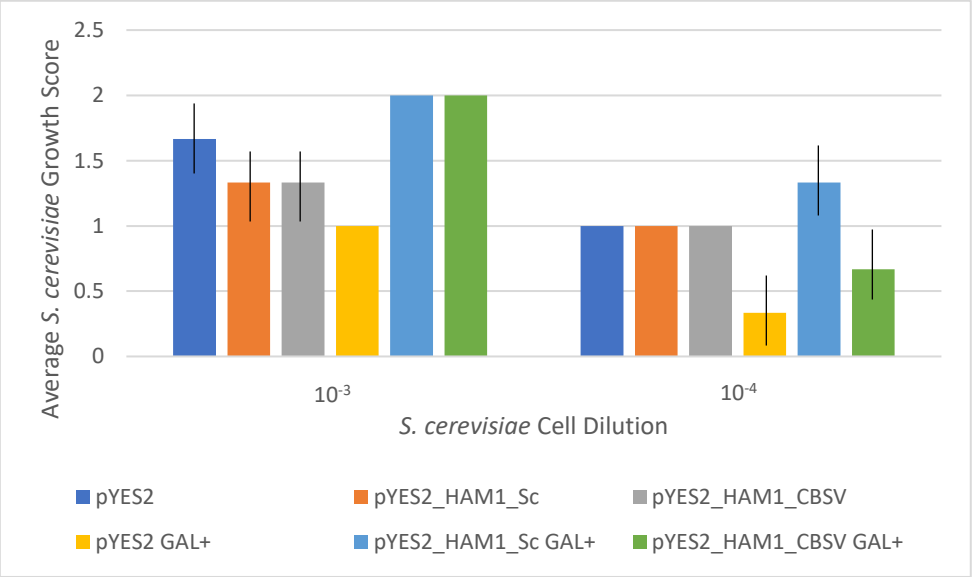

Figure 3.8. Bar chart showing the average growth scores (n=3) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} and 10^{-4} cell dilutions at 100ng/ml cycloheximide concentration, with and without supplemented galactose.

No growth was recorded for the *ham1* knockout cells transformed with the empty pYES2 vector at the 10^{-4} cell dilution at the CHX concentration of 130ng/ml in both GAL+ and GAL- conditions, despite growth being recorded for both the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed cells in the GAL- conditions (Fig. 3.9). With the addition of galactose, the pYES2_HAM1_Sc transformed cells showed an increase in growth (HAM1_Sc, p = 0.588). Although not a significant difference, most likely due to a lack of replicates, this result suggests a trend where the expression of the yeast *HAM1* gene does confer greater resistance to the toxic effects of CHX. As was seen at the 10^{-4} cell dilution at 100ng/ml CHX concentration, the addition of galactose at the 130ng/ml CHX concentration also saw a decrease in growth scores for the pYES2_HAM1_CBSV transformed cells, yet this decrease was insignificant (HAM1_CBSV, p = 0.374). Without any observed growth for the pYES2 transformed cells,

it is not possible to see if the scores for the cells expressing the CBSV *HAM1* gene are still greater than those cells not expressing a *HAM1* gene.

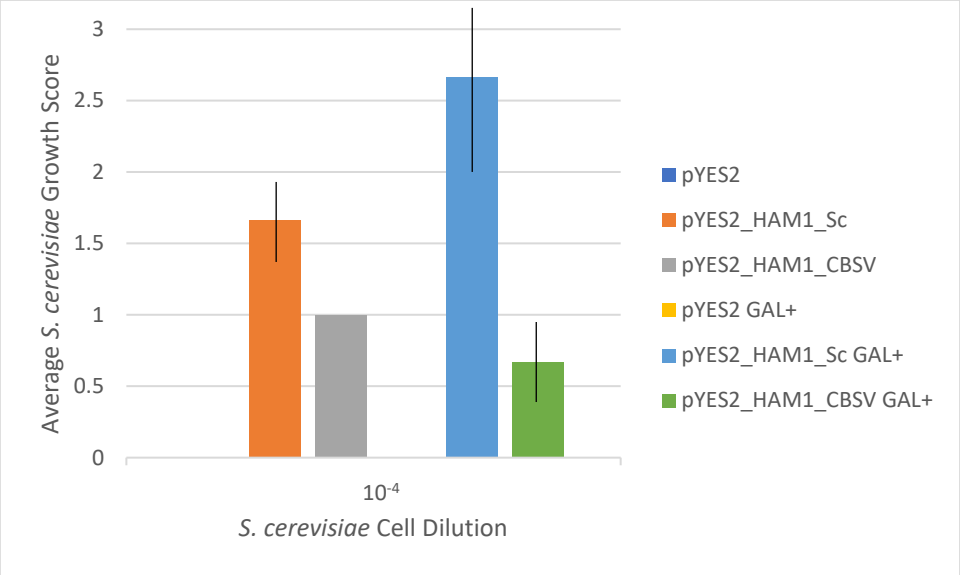

Figure 3.9. Bar chart showing the average growth scores (n=3) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-4} cell dilution at 130ng/ml cycloheximide concentration, with and without supplemented galactose.

Overall, most of the effects on increased resistance to CHX were observed at the two highest concentrations of CHX or at the greatest cell dilutions (10^{-3} and 10^{-4}) of the lower concentrations of CHX, most notably 45ng/ml. This is probably because at the higher concentrations, CHX is now exhibiting greater cytotoxic effects and thus the expression or overexpression of the CBSV and yeast *HAM1* genes becomes more readily apparent. Indeed, at the greater cell dilutions, the effects of those cells expressing or overexpressing *HAM1* genes also become more readily apparent. Increases in growth scores observed for those wild type cells overexpressing the yeast and CBSV *HAM1* genes at CHX concentrations of 45ng/ml and 130ng/ml, corroborate the suggestion that the CBSV HAM1 shares functionality with the yeast HAM1. Furthermore, expression of these *HAM1* genes in the *ham1* mutant yeast strains at 100ng/ml and 130ng/ml CHX concentration, show that the CBSV HAM1 can complement the yeast HAM1.

3.6 Methotrexate Resistance Assay

Methotrexate (MTX) is a folate analogue metabolic inhibitor which ceases the production of tetrahydrofolate (THF). The lack of THF prevents the formation of the CH_2THF complex and subsequently inhibits the biosynthesis of thymidine through the inability of dUMP to be broken down into dTMP. Furthermore, the lack of THF prevents purine biosynthesis, and thus, MTX has drastic effects on DNA synthesis (Longley *et al*., 2003; Carlsson *et al*., 2013).

3.6.1 Overexpression of HAM1

At the 10^{-3} cell dilution at an MTX concentration of 10µg/ml, both the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed wild type yeast cells show greater growth than those transformed with the empty vector (Fig. 3.10). The same scores are also observed with the supplementation of the agar with galactose across all three transformants and show that there was no effect on MTX resistance from the overexpression of the yeast or CBSV *HAM1* genes. It is likely that no effect was seen upon the addition of galactose as there were sufficient levels of leaky expression in the GAL- conditions to support increased resistance with either the expression of the yeast or CBSV *HAM1* genes at MTX concentrations of 10-60µg/ml. This was not observed at a concentration of 80µg/ml, however, where the MTX concentration may have been too high to observe such an effect of any *HAM1* expression.

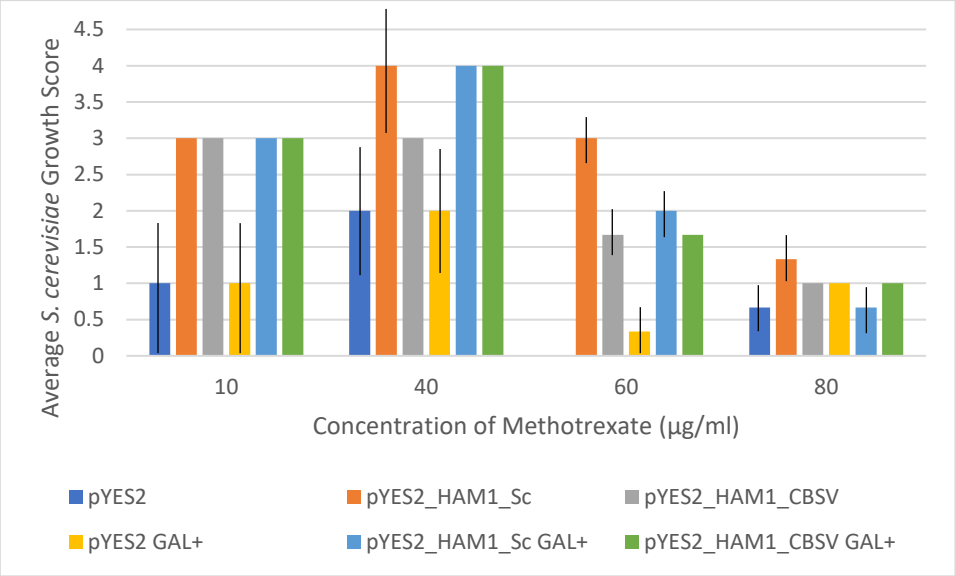

Figure 3.10. Bar chart showing the average growth scores (n=3) of the wild type *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} cell dilution across all concentrations of methotrexate, with and without supplemented galactose.

At the MTX concentration of 40µg/ml, again the growth scores of the cells transformed with the empty pYES2 vector remain the same when under the influence of galactose, a result that was expected (Fig. 3.10). As with the MTX concentration of 10µg/ml, no apparent effect of the overexpression of the yeast *HAM1* was observed. There was, however, an increase in growth observed when the *HAM1* gene from CBSV was overexpressed (HAM1_CBSV, $p = 0.374$). Although not a significant difference, the result suggests that the action of the CBSV HAM1 is conferring some increased resistance to MTX, but more sensitive assays, and indeed replicates, would need to be conducted to confirm such an effect.

Despite an insignificant reduction in growth scores for those cells overexpressing the yeast *HAM1* gene at the 10^{-3} cell dilution at the MTX concentration of 60µg/ml (HAM1_Sc, $p = 0.374$), and with no

apparent effect of the overexpression of the CBSV *HAM1* gene as well, their growth was much greater than those wild type cells not overexpressing a *HAM1* gene (Fig. 3.10). However, the decreases in growth scores for cells overexpressing the yeast or CBSV HAM1 genes means that little can be inferred of their effect on resistance to MTX.

At the MTX concentration of 80µg/ml, there seems to be very little effect on the growth of the wild type *S. cerevisiae* cells through the overexpression of either the yeast or CBSV *HAM1* genes (Fig. 3.10). This result is most likely due to the inability of the HAM1 protein alone to be able to continue to afford protection to the toxic effects of MTX at that high a concentration.

3.6.2 Complementation of ΔHAM1

Addition of galactose to the *ham1* knockout yeast cells at the 10^{-3} cell dilution for all concentrations of MTX (Fig. 3.11) seemed to have little effect on the growth scores of any of the transformed cells and in most cases, the growth scores declined, with the results being significantly different for those pYES2_HAM1_CBSV ($p = 0.050$) and pYES2_HAM1_Sc ($p < 0.001$) transformed cells at the 40µg/ml and 60µg/ml MTX concentrations, respectively. Some improvement in growth was expected to be seen with the addition of the galactose, due to the induction of expression of the yeast and CBSV *HAM1* genes, yet none was observed. A possible explanation why no effect was seen could be that the concentration of MTX was not great enough at this cell concentration to really exert its cytotoxic effects, although an effect was seen in the wild type yeast assay. Why an effect was seen on the growth of the wild type yeast at 80µg/ml MTX but not on the growth of the *ham1* knockout yeast, cannot be explained.

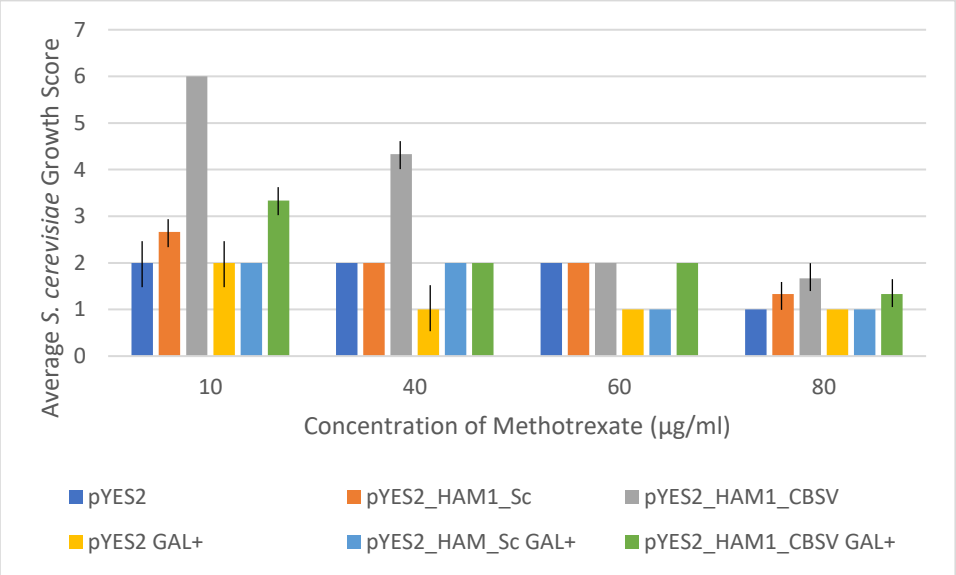

Figure 3.11. Bar chart showing the average growth scores (n=3) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-3} cell dilution across all concentrations of methotrexate, with and without supplemented galactose.

Under the induction of galactose, the expression of the *HAM1* genes from yeast and CBSV do seem to have an effect at the 10^{-4} dilution at the MTX concentration of 10µg/ml (Fig. 3.12), where both transformants showed increases in growth scores, though neither results are significantly different (HAM1_Sc, p = 0.422; HAM1_CBSV, p = 0.643). This result, however, corroborates the suggestion that the CBSV HAM1 does complement the yeast HAM1 by providing increased resistance to MTX.

Increased resistance to the effects of MTX through the expression of the CBSV *HAM1* gene was also observed at this 10^{-4} cell dilution across all concentrations of MTX. The effect of the CBSV *HAM1* expression was most notably seen at the MTX concentration of 60µg/ml, where there was a marked decrease with the addition of galactose in the growth of cells not expressing any *HAM1*, those transformed with the pYES2 empty vector, compared to those expressing the CBSV *HAM1* which saw its growth score double (Fig. 3.12). Due to a lack of replicates, however, neither result was significantly different compared to their respective growth scores in GAL- conditions (pYES2, p = 0.374; HAM1_CBSV, p = 0.519).

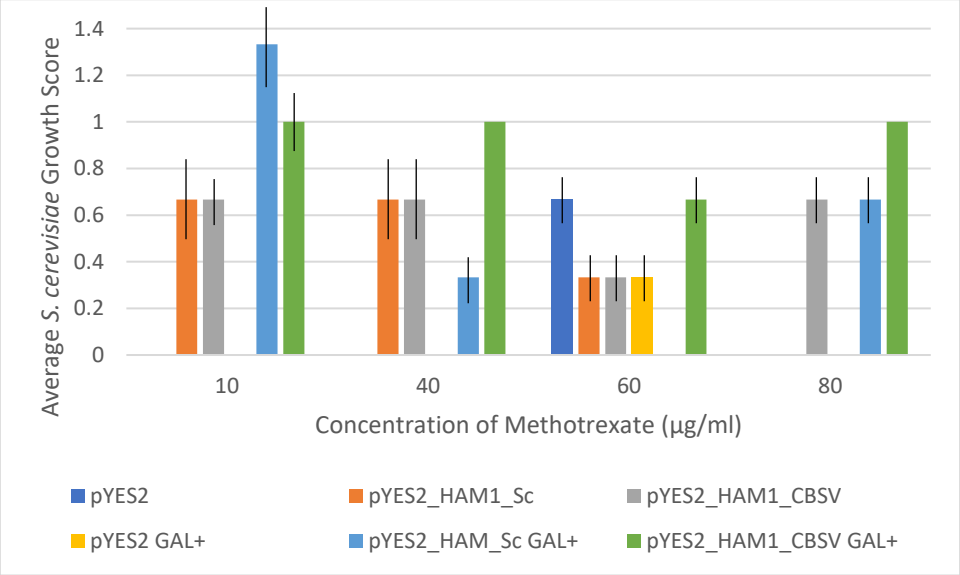

Figure 3.12. Bar chart showing the average growth scores (n=3) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids at the 10^{-4} cell dilution across all concentrations of methotrexate, with and without supplemented galactose.

The effects of the expression of the yeast *HAM1* gene cannot be determined at the two greater concentrations of MTX as no growth was observed, either in conditions without galactose or with the addition of galactose.

The only effects on increased resistance to MTX have been observed through the expression of the CBSV and yeast *HAM1* genes at the greatest cell serial dilutions (10^{-4}). At these cell concentrations, the assay seems to have become sensitive enough to observe the subtle differences in phenotypes and so to fully confirm if the CBSV HAM1 can complement the yeast HAM1, assays need to be

conducted at this cell concentration and perhaps, even lower cell concentrations. Effects of the overexpression of the CBSV HAM1, with the only result observed at the 10^{-3} cell dilution at 40µg/ml MTX, was not enough to determine if the CBSV HAM1 shares functionality with the yeast HAM1 and so further assays would need to be conducted to achieve conclusive results.

3.7 5-FLUOROURACIL RESISTANCE ASSAY

5-fluorouaracil (5-FU) is a pyrimidine analogue that was designed as an antimetabolite drug to target tumour cells for a variety of cancers, from colorectal cancers to breast cancer (Carlsson *et al*., 2013; 2018), and thus has been extensively studied especially since the development of resistance to this nucleotide analogue within human cells (Carlsson *et al*., 2013). Antimetabolite drugs usually exert their effect by inhibiting essential biosynthetic pathways, or through their uptake into DNA or RNA and preventing their normal function as a result of increased levels of mutagenesis. 5-FU carries out both; with the misincorporation of fluoronucleotides into RNA and DNA, and the inhibition of thymidylate synthase (TS), a nucleotide synthetic enzyme.

5-FU has the same structure as uracil with a fluorine atom in place of a hydrogen atom at the C-5 position, with the molecule being taken up into the cell in the same way as uracil. Within the cell, however, 5-FU is then converted into three active metabolites: fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP) and fluorouridine triphosphate (FUTP) – all of which disrupt RNA synthesis and inhibit the action of thymidylate synthase (TS) (Fig. 3.13) (Longley *et al*., 2003; Carlsson *et al*., 2013).

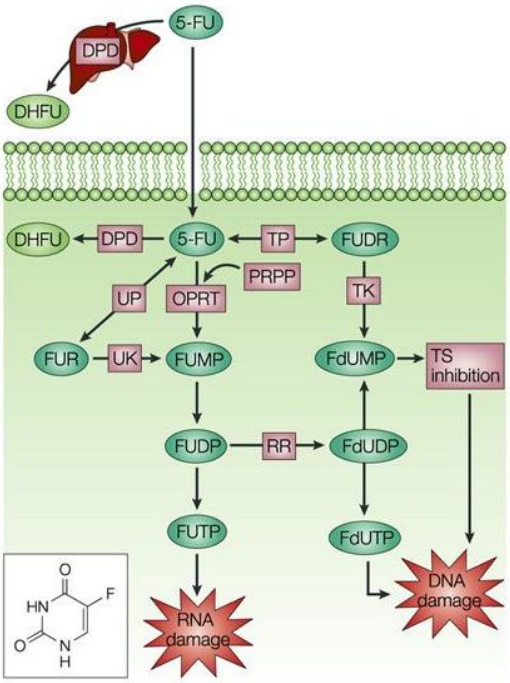

Figure 3.13. Action of the three active metabolites of 5-fluorouracil (FdUMP, FdUTP and FUTP) in the cell that causes nucleic acid damage and thymidylate synthase (TS) inhibition (reproduced from Longley *et al*., 2003).

The action of TS is crucial for DNA replication and repair within a cell. TS acts through the catalysis of the reductive methylation of deoxuyridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP) which provides the sole intracellular production of thymidylate. With the presence of 5-FU in the cell, however, the 5-FU FdUMP metabolite instead binds to the nucleotide-binding site of TS and coupled with CH_2THF, forms a stable ternary complex which prevents the binding of dUMP. This blocking of the binding of dUMP inhibits the synthesis of dTMP which would normally go on to form the thymidylate required for DNA replication and repair (Fig. 3.14). Furthermore, a lack of dTMP leads to a depletion in deoxythymidine triphosphate (dTTP) levels which subsequently causes imbalances in cellular deoxynucleotide levels, further disrupting DNA synthesis and repair (Fig. 3.14) (Longley *et al*., 2003).

The increased levels of dUMP brought about through the inhibition of TS action leads to increases of deoxyuridine triphosphate (dUTP) and together with the 5-FU metabolite FdUTP, can be misincorporated into DNA, a process known as thymineless death. Subsequent repair of the 5-FU containing DNA through the action of the nucleotide excision repair enzyme, uracil-DNA-glycosylase (UDG), only causes further false-nucleotide incorporation. Eventually, the continual cycle of misincorporation and excision results in DNA strand breaks and then cell death. One way to bring about resistance to 5-FU is through the action of thymidine kinase which salvages thymidylate from thymidine and thus alleviates the effects of TS inhibition (Longley *et al*., 2003; Carlsson *et al*., 2013).

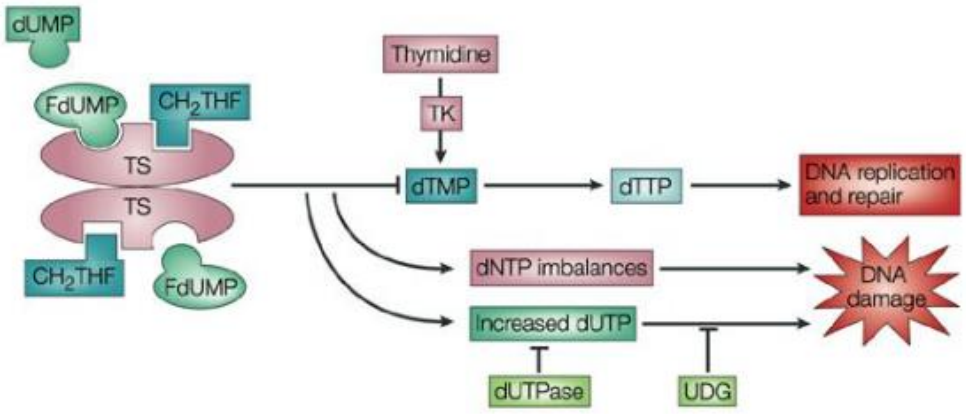

Figure 3.14. The binding of dUMP to thymidylate synthase leads to the inhibition of thymidylate, which is required for DNA replication and repair, and the increase of nucleotide imbalances which go on to cause damage to the nucleic acids (reproduced from Longley *et al*., 2003).

Effects of 5-FU on RNA processing and function is just as pronounced as its effects on DNA, with disruption being brought about through the extensive incorporation into RNA of the 5-FU metabolite FUTP. This misincorporation into RNA inhibits the processing of pre-rRNA into mature rRNA, affects the formation of snRNA/protein complexes which then inhibits the splicing of pre-mRNA, as well as inhibiting polyadenylation of mRNA even at low concentrations. With these disruptions to RNA processing, 5-FU causes severe effects to cellular metabolism and viability, and these effects appear to contribute significantly to the cytotoxic effects of the compound (Longley *et al*., 2003; Carlsson *et al*., 2013).

3.7.1 Overexpression of HAM1

At the 5-FU concentration of 10µg/ml and from the 10^{-1} cell dilution, the addition of galactose consistently caused a reduction in the growth scores of all transformants across all subsequent serial dilutions. The same effect was seen for the 5-FU concentration of 40µg/ml from the 10^{-2} cell dilution onwards.

The only increases in growth for the pYES2_HAM1_Sc transformed cells observed upon the addition of galactose were seen at the 10^{-2} yeast cell dilution at a concentration of 60µg/ml 5-FU (Fig. 3.15), however, the increase in growth scores of those wild type *S. cerevisiae* cells transformed with the empty pYES2 vector, suggest that there was no effect on resistance to 5-FU from the overexpression of the yeast *HAM1*, nor was there any effect through the overexpression of the CBSV *HAM1*. Indeed, the increase in growth scores observed at this cell dilution were insignificant (pYES2, $p = 0.423$; HAM1_Sc, $p = 0.423$). Similarly, at the 10^{-4} cell dilution, the lack of increase in growth scores for the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed cells in the presence of the supplemented galactose, shows that again, the overexpression of the *HAM1* genes did not confer increased

resistance to 5-FU. It may be, however, that there is simply not enough HAM1 produced by the cells to overcome the cytotoxic effects of 5-FU.

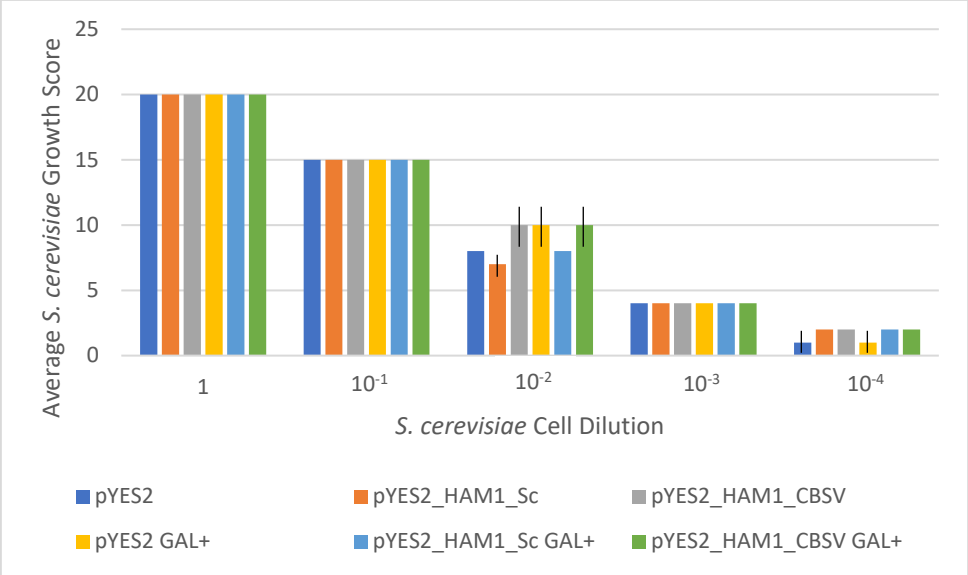

Figure 3.15. Bar chart showing the average growth scores (n=2) of the wild type *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids across all serial dilutions at the 5-fluorouracil concentration of 60µg/ml, with and without supplemented galactose.

Addition of galactose generally caused the transformed wild type yeast cells to show a reduction in growth, an effect that was also found by Carlsson *et al*. (2013), so no effect on the overexpression of the yeast or CBSV *HAM1* genes on resistance to 5-FU could be deduced.

3.7.2 Complementation of ΔHAM1

At the 10^{-2} cell dilution at the 5-FU concentration of 10µg/ml, upon addition of galactose, those *ham1* knockout cells transformed with the pYES2 vector and not expressing a *HAM1* gene, show a decrease in growth (pYES2, p = 0.423) (Fig. 3.16). This is contrary to what was seen for those cells expressing the *HAM1* gene from *S. cerevisiae* which showed an increase in growth with the supplemented galactose, though this result was not significant due to a lack of replicates (HAM1_Sc, p = 0.423). Despite there being no increase in the growth scores for the pYES2_HAM1_CBSV transformed cells, the level of growth remained the same and was the greatest compared to the other two transformants. This result shows that expression of the *HAM1* gene, either from *S. cerevisiae* or CBSV, does provide protection against the pyrimidine analogue and suggests that the CBSV HAM1 can complement that of yeast.

At the 10^{-3} cell dilution at this same 5-FU concentration, both the pYES2_HAM1_Sc and pYES2_HAM1_CBSV transformed cells showed an increase in growth upon addition of galactose, whereas the pYES2 transformed cells showed a decrease (Fig. 3.16). None of these results were significant due to a lack of replicates (pYES2, p = 0.423; HAM1_Sc, p = 0.423; HAM1_CBSV, p = 0.423),

but the results do provide evidence to suggest that the expression of the *HAM1* genes from yeast and CBSV can confer increased resistance to the toxic effects of the pyrimidine analogue.

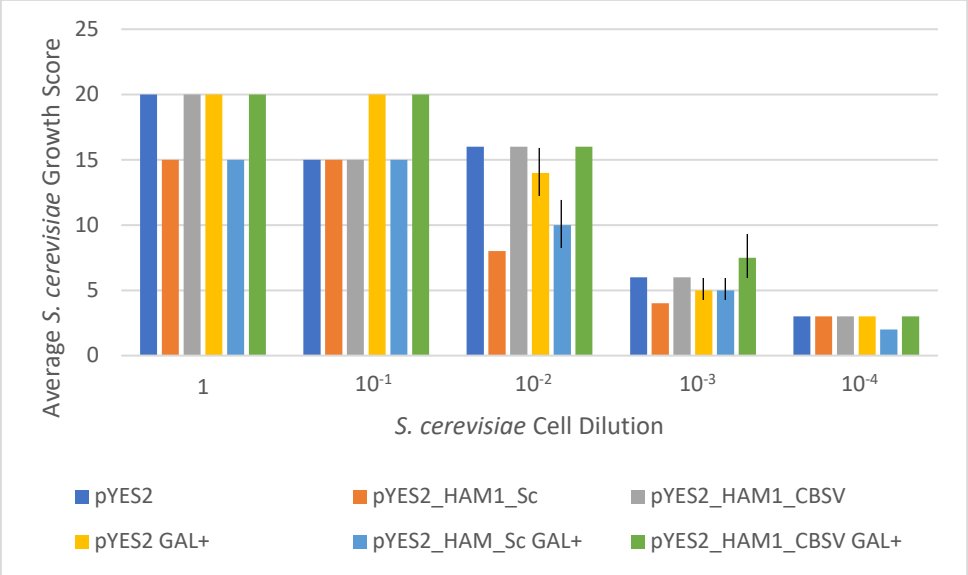

Figure 3.16. Bar chart showing the average growth scores (n=2) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids across all serial dilutions at a 5-fluorouracil concentration of 10µg/ml, with and without supplemented galactose.

At the 10^{-2} cell dilution at a concentration of 40µg/ml 5-FU, the same result was seen upon the addition of galactose as was observed at the 10^{-2} cell dilution with 5-FU at 10µg/ml (Fig. 3.17), yet the increase in growth score observed for the pYES2_HAM1_Sc transformed cells was significant ($p < 0.001$). Otherwise, at the other serial dilutions at this 5-FU concentration, the growth scores of the pYES2_HAM1_CBSV transformed cells increase when the CBSV *HAM1* is being expressed, but these results are not conclusive that the CBSV HAM1 complements the function of the yeast HAM1 due to no readily apparent differences to the pYES2 transformed cells and due to the fact that the increases in growth scores are not significant ($p = 0.3$, $p = 0.423$). At the 10^{-1} cell dilution, this increase in the growth scores of the pYES2_HAM1_CBSV transformed cells ($p = 0.423$) was also reflected in those cells transformed with the empty pYES2 vector upon addition of galactose ($p < 0.001$), thus, at this dilution, there seems to be no apparent effect on the ability of the *ham1* knockout cells to display resistance to 5-FU through the expression of the yeast or CBSV *HAM1* genes.

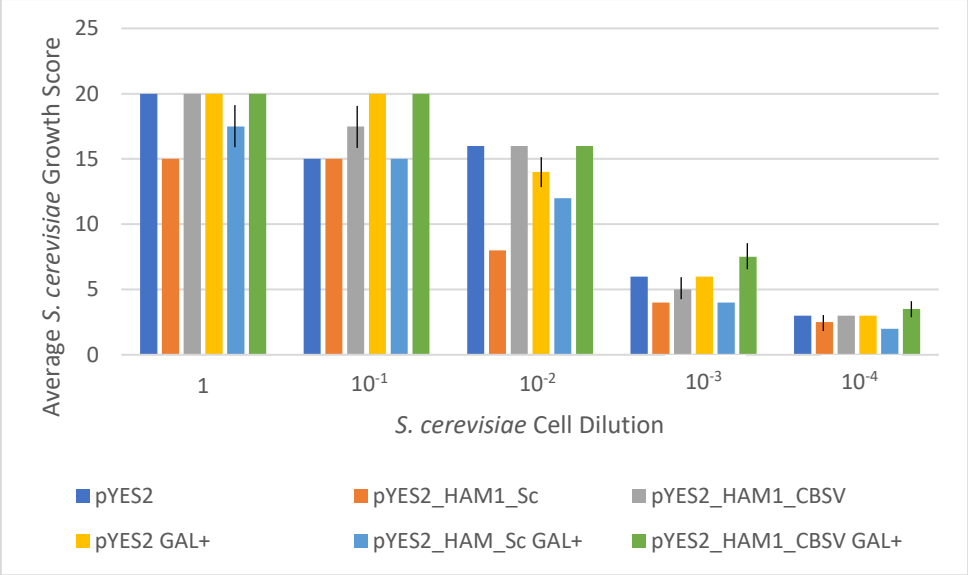

Figure 3.17. Bar chart showing the average growth scores (n=2) of the *ham1* knockout *S. cerevisiae* cells transformed with the pYES2, pYES2_HAM1_Sc and pYES2_HAM1_CBSV plasmids across all serial dilutions at a 5-fluorouracil concentration of 40µg/ml, with and without supplemented galactose.

At a 5-FU concentration of 60µg/ml, the addition of galactose did little to affect the growth of the cells; with the only increase in growth observed for the pYES2_HAM1_Sc transformed cells at the 10^{-4} cell dilution ($p < 0.001$). Similarly, at the highest concentration of 5-FU tested (80µg/ml), there was no increased resistance to the pyrimidine analogue brought about through the expression of the yeast or CBSV *HAM1* genes. At this greatest concentration of 5-FU, it might be too concentrated for the function of HAM1 alone to afford increased protection to the effects of the molecule.

Thus, the only effect seen on resistance to 5-FU through the expression of the yeast or CBSV *HAM1* genes was at two cell dilution concentrations at the lowest 5-FU concentration. These results alone are not enough to confirm whether the CBSV HAM1 can complement the yeast *ham1* mutant, or whether the viral HAM1h shares functionality with the yeast HAM1.

3.8 Bromodeoxyuridine, 5-flucytosine and 6-azauracil resistance assays

Similar plate-based assays were attempted with other base analogues such as bromodeoxyuridine, 5-flucytosine and 6-azauracil, however, these inhibitors proved to be far more mutagenic than 5-FU. Despite evaluating a range of concentrations of inhibitor and yeast cell dilution, no clear pattern of results could be obtained.

CHAPTER 4 – INVESTIGATING THE TRANSCRIPT LEVELS OF THE *NICOTIANA BENTHAMIANA HAM1* WHEN INFECTED WITH *CASSAVA BROWN STEAK VIRUS* AND *TOBACCO MOSAIC VIRUS*

4.1 BACKGROUND

As with bacteria, fungi and animals, plants also appear to have HAM1 proteins that along with other eukaryotic HAM1 proteins, function in the nucleus of the cells, yet their NTPase activity has yet to be fully investigated. In *Arabidopsis thaliana*, there are several HAM-like proteins that primarily function as histone acetyltransferases (HATs) and regulate gene expression to control such things as cell determination, flowering times and the formation of gametophytes (Earley *et al*., 2007; Latrasse *et al*., 2008; Xiao *et al*., 2013; Zhou *et al*., 2015). These HAM proteins belong to the Maf family, which show NTPase activity, and their overexpression is shown to inhibit cell division through the arrest of septum formation, likely as a result of their action on non-canonical nucleotides which lowers the concentrations of GTP and ATP (Galperin *et al*., 2006). As these plant proteins appear to exhibit NTPase activity alongside their primary function, they may provide an as yet unrecognised plant defense mechanism with the ability to protect against mutation inducing oxidative stress when the plant is infected. The potential that the plant could down regulate its *HAM1* and thus provide a highly mutagenic environment for the pathogen within its cells, was the hypothesis behind this novel investigation. Studies looking into the effect of CBSV infection on the expression of the cassava *HAM1* have not been conducted, nor has any such investigation been carried out on model host plant species. It was thought, therefore, that by investigating the effects of CBSV infection on the plant *HAM1*, whether the gene is up or down regulated as a result, may provide further elucidation into the function of the CBSV HAM1h.

With the time constraints and the lengthy life cycle of the cassava plant, it was not possible to carry out the investigation within the CBSV natural host species. Due to its ability as a systemic host for CBSV infection (Monger *et al*., 2001b) and its susceptibility to many plant viruses (Ogwok *et al*., 2010), however, *Nicotiana benthamiana* plants were used in this study. In order to measure the effect that infection with CBSV had on the expression of the plant *HAM1*, changes in the plant *HAM1* transcript

levels from four different parts of the plant would be compared to those of a reference gene, known as a housekeeping gene, that is expressed constitutively. These results would also be compared to the transcript levels of the housekeeping and *HAM1* genes in healthy *N. benthamiana* plants and ones infected with *Tobacco mosaic virus* (TMV). The four parts of the plant tested would be the shoot tip, stem, mature leaves and young leaves. It was further proposed that the *HAM1* transcript levels in each part of the plant would show differences; with the shoot tip and young leaves being areas of rapid cell division and thus, should present greater *HAM1* transcript levels when healthy, with the effect of CBSV and TMV infection being unknown. Quantitative PCR (qPCR) is the most common method for determining gene expression patterns and comparing gene transcript levels due to its high specificity which provides accurate, sensitive and reproduceable results (Liu *et al*., 2012). qPCR works through the use of fluorescent dyes that bind to the synthesised dsDNA with the number of cycles required to produce a detectable amount above the cycle threshold (Ct) level, recorded. The Ct values obtained for the gene of interest, in this case *HAM1*, and the housekeeping gene, are compared and the analysis of these results determines if the gene of interest has been up or down regulated through an increase or decrease in transcript levels, respectively.

There are four *HAM-like* genes (*HAM1-HAM4*) found in *A. thaliana*. Deletion of the *A. thaliana* HAM1 amino acids 117 through to 230, showed that it then prevented the protein's ability to interact with the WUSCHEL (WUS) protein that is required for the control of stem cell populations in the shoot apical meristem (SAM). This revealed that the N-terminal of the HAM1 protein is essential for its function, and it was shown that this region is highly conserved amongst the *A. thaliana* HAM proteins and in other plants (Zhou *et al*., 2015). Discovery of the HAM1 protein in *N. benthamiana*, however, required bioinformatic analysis in order to design primers for use in this investigation.

4.2 BIOINFORMATIC ANALYSIS

A Basic Local Alignment Search Tool (BLAST) search using the NCBI database of the yeast HAM1 protein sequence revealed 40% identity to a sequence in *N. benthamiana*. No described sequences for this HAM1 protein in *N. benthamiana* are currently described in the NCBI database, and so the yeast HAM1 sequence was used in a BLAST search in the Sol Genomics Network database against the *N. benthamiana* genome. Results of this search found the same sequence in *N. benthamiana*; a protein of length 227 amino acids was identified as a 'non-canonical purine NTP pyrophosphatase (ITPase-like)' (Appendix Fig. 1). Further BLAST searches of this *N. benthamiana* HAM1 in the Sol Genomics Network found two other proteins within the *N. benthamiana* genome with 96% and 66% sequence identity and they were identified as the 'inosine triphosphate pyrophosphatase-like' protein and the

'non-canonical purine NTP pyrophosphatase', respectively, highlighting that three apparent HAM1 proteins are found within the *N. benthamiana* genome.

The *N. benthamiana* HAM1 sequence that showed greatest homology with the yeast HAM1 (40% sequence identity) was then used in a BLAST protein search in the NCBI databases and revealed approximately 49% and 50% sequence identity with the UCBSV and CBSV HAM1h proteins respectively. All four HAM1h protein sequences analysed contained the highly conserved SHR motif (Fig. 4.1).

```
Query_10001   1    MAAAARTVGS    GVVLPRPVTFVTGNAKKLEEVRAILGQSIP   FQSLKLDLPELQGE-PEDISKEKARIAAKEVNG  72
Query_10002   1    ----------    --MSNNEIVFVTGNANKLKEVQSILTQEVD[7]LINEALDLEELQDTdLNAIALAKGKQAVAALGK  68
Query_10003   1    VVDRPQSLNV[11]RMGIEAPITFVTGNAQKLKEVKQIFGPTIP   IVSRKIDLPESQGT-VEEIIKEKARVAAELVGG  83
Query_10004   1    ----------    --QMKFPVTFVTGNLGKLAEVKSILGISSD   VMARNIDLPEVQGT-PDEIVIKKAQLAVKMTNS  60

Query_10001   73   --PVLVEDTCLCFNALKGLPGTQ[20]CKWFLQKIGHEGLNNLLLAYEDKTAYAMCVFSLALGPNSEPLTFVGKTLGRIV  166
Query_10002   69   gkPVFVEDTALRFDEFNGLPGAY    IKWFLKSMGLEKIVKMLEPFENKNAEAVTTICFADSR-GEYHFFQGITRGKIV  143
Query_10003   84   --PVLVEDTSLCFDALNGLPGPY    IKWFLEGIGLEGLYKLVEPYQNRMASALCVFAFVNKVGDDPIIFKGVLRGEIV  157
Query_10004   61   --PVLVEDTCLCFNAFNGLPGPY    IKWFLKELGLEGVVKMLSAFGDKSAYALCTFAYVHNELSDPVVFKGVVNGEIV  134

Query_10001   167  PARGPNDFGWDPIFQP-HGYDQTYAEMPKEGKNKISHRGKALELVKSHFAEARYTFQTDTTA--------  227
Query_10002   144  PSRGPTTFGWDSIFEPFDSHGLTYAEMSKDAKNAISHRGKAFAQFKEYLYQNDF----------------  197
Query_10003   158  IPRGPNSFGWDPIFQPLD-WKRTFAEMMIEEKNMISHRFRALSLVRDFLKSSSYFSFAKGLDRDIFIDVQ  226
Query_10004   135  PPRGNNGFGWDPIFKP-DECSCTFAEMPSSIKNDFSHRRRALEKVKLFLDNLVVKQEKKKAGVALTIDVQ  203
```

Figure 4.1. Results of the BLAST search and alignment showing sequence identity, highlighted in red, between the *N. benthamiana* HAM1 (Query_10001), *S. cerevisiae* HAM1 (Query_10002), CBSV HAM1h (Query_10003) and UCBSV HAM1h (Query_10004). All have the conserved SHR motif that is found in all HAM1 homologues.

4.3 Primer Design

Further analysis on the Sol Genomics Network database of this BLAST search revealed that the genomic DNA sequence of the *N. benthamiana* HAM1 is 6,987 base pairs long. Comparisons of the DNA sequence with the amino acid sequence (Appendix Fig. 2) allowed for the identification of the exons which were subsequently used for the design of the primers (Appendix Fig. 3). qPCR Primers were required to be designed for the *N. benthamiana* HAM1 such that the size of the product would be of a similar size to the housekeeping gene product, and that the conditions would be appropriate to amplify both the *N. benthamiana* HAM1 and the housekeeping genes. One limiting factor to the technique is the selection of a suitable housekeeping reference gene. This reference gene must be stably expressed within the tissues being tested regardless of any treatments applied. As these housekeeping genes provide the point of reference to which the target gene is compared, selection of an unsuitable reference gene can lead to the results being interpreted incorrectly (Liu *et al*., 2012).

Protein phosphatase 2A (*PP2A*) was identified as a suitable housekeeping gene for use in the qPCR experiment due to its stability in virus-infected *N. benthamiana* tissue (Liu *et al*., 2012). Primers

designed to this *PP2A* gene amplified a product of 123 base pairs under PCR conditions of between 50 and 55^{o}C (Table 4.1). The nucleotide sequence of the *N. benthamiana* HAM1 gene was entered into the GenScript website and primer pairs were designed in view of the conditions required for amplification and the size of the housekeeping gene product. Two primer pairs were designed; the first pair generated a PCR product of 124 base pairs and was named HAM11, and the second pair generated a PCR product of 126 base pairs and was named HAM12 (Table 4.1). Subsequent testing of the two HAM1 primer pairs and the PP2A primer pair using PCR analysis was carried out to establish the optimum conditions for both. Furthermore, to ensure the *N. benthamiana* HAM1 primers selected were only going to amplify the *N. benthamiana HAM1* gene and not the viral *HAM1* gene in the CBSV infected plants, the HAM11 and HAM12 sequences were compared to those of a primer pair designed to amplify the CBSV *HAM1* gene, CBSVHAM1 (Table 4.1). Comparisons of the primer sequences revealed some similarities between the pairs, with the HAM12 forward primer sharing some 66% sequence identity to the CBSVHAM1 equivalent. Most importantly, however, it is the bases at the 3' end of primer sequences that confer specificity to the target gene and mismatches of even one or two bases can drastically reduce primer efficiency (Whiley and Sloots, 2005). Despite the similarities in the primer sequences, there are few similarities at the 3' end and thus, this ensures that the *N. benthamiana* HAM1 primer pairs are indeed specific only to the plant *HAM1* gene.

Table 4.1. List of primer pair sequences for the *N. benthamiana HAM1* gene, the CBSV *HAM1* gene and the three *N. benthamiana* housekeeping genes: *PP2A*, *F-BOX* and *GBP*.

Primer pair	Gene accession number	Primer sequences (5'-3') forward/reverse
PP2A	TC21939 (At1g13320)	GACCCTGATGTTGATGTTCGCT/GAGGGATTTGAAGAGAGATTTC
HAM11	Niben101Scf18106g00001.1	AAGGGTCTCCCAGGTACTCA/TGTTGAGGCCTTCATGACCA
HAM12	Niben101Scf18106g00001.1	GCCAGAACTTCAAGGTGAGC/GGGAGACCCTTAAGAGCGTT
CBSVHAM1	GQ169759	ACCAGAACCACAAGGGACAG/TGGCATCACAATCTCACCTC
F-BOX	Niben.v0.3.Ctg24993647 (At5g15710)	GGCACTCACAAACGTCTATTTC/ACCTGGGAGGCATCCTGCTTAT
GBP	TC20872 (At5g59840)	GGAACTGGATTCGCAACATAGA/GACCCTTGGAAGTTGGCACAGC

4.4 PRIMER OPTIMISATION

Healthy and CBSV infected *N. benthamiana* were grown before tissue samples were collected, RNA extracted and cDNA synthesised, to test the primers for the target and housekeeping genes. The two primer pairs for the *N. benthamiana HAM1* were tested through PCR, with the second pair (HAM12) proving to produce greater amplification of the gene. For all subsequent PCR and qPCR reactions, this

HAM12 primer pair was used. Initially, the housekeeping gene *PP2A* was proposed to be used in the experiments, however, on use in PCR, this gene did not perform well at the suggested annealing temperature of 55°C (Fig. 4.2) and thus, amplification of the *PP2A* housekeeping gene was deemed not to be strong enough for use in qPCR. Reduction of the annealing temperature to 52°C saw amplification of *HAM1* increase, but again, amplification of *PP2A* was too low (Fig. 4.3).

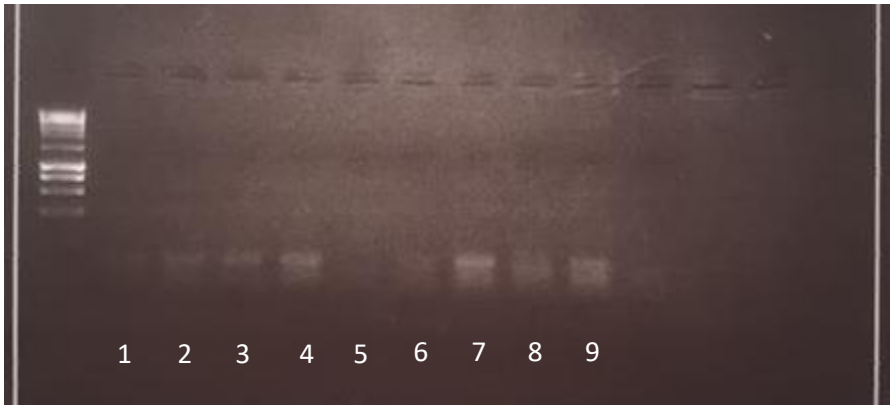

Figure 4.2. Gel electrophoresis showing amplification of the *PP2A* gene in lanes 1-4 and of the *HAM1* gene in lanes 6-9 extracted from healthy *N. benthamiana* tissue following PCR with an annealing temperature of 55°C.

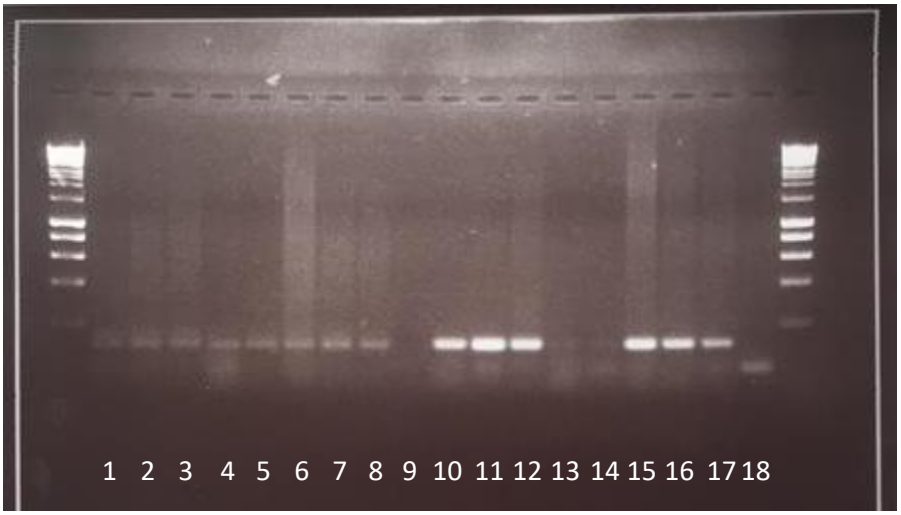

Figure 4.3. Gel electrophoresis showing amplification of the *PP2A* gene in lanes 1-8 and of the *HAM1* gene in lanes 10-17 taken from healthy and CBSV infected *N. benthamiana* tissue following PCR with an annealing temperature of 52°C. It is shown here that the reduction in annealing temperature allowed for greater amplification of the *HAM1* gene but did not produce sufficient amounts of the *PP2A* gene for use in qPCR.

Instead, primers for a further second and third *N. benthamiana* housekeeping genes which also showed stability in CBSV infected tissue, the *GTP binding protein* (*GBP)* and *F-box protein* (*F-BOX*), were used (Table 4.1) (Liu *et al*., 2012). In order to achieve sufficient amplification of a housekeeping gene, primers for the *GBP* and *F-BOX* genes were used in PCR. Initial PCR procedures carried out at a temperature of 55°C proved too high for sufficient *GBP* and *F-BOX* amplification (Fig. 4.4), although amplification was higher for the *GBP* gene at this annealing temperature compared to that of the other housekeeping gene *F-BOX*. Reduction of the annealing temperature, again to 52°C, provided the ideal

conditions for sufficient amplification of *GBP* (Fig. 4.5), and so this housekeeping gene was selected for use in qPCR with the HAM12 primers with the annealing temperature set to 52°C.

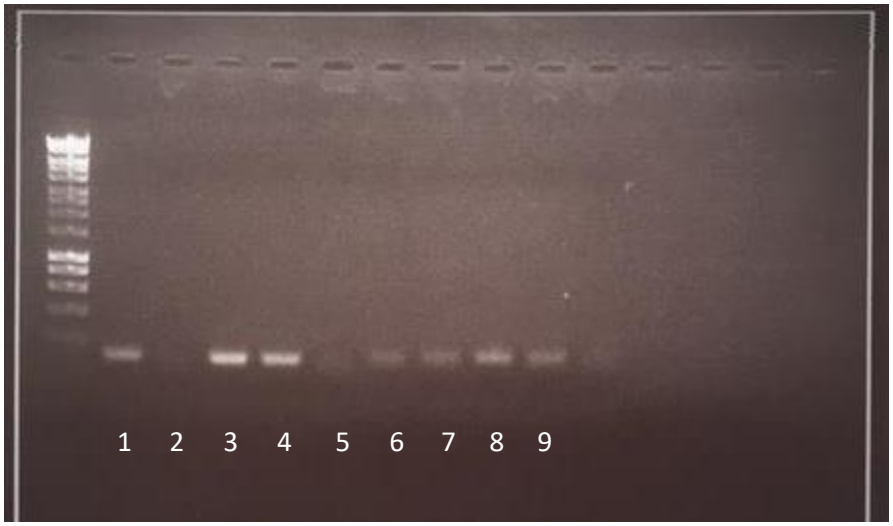

Figure 4.4. Gel electrophoresis showing the amplification of the *GBP* housekeeping gene in lanes 1-4 and of the *F-BOX* housekeeping gene in lanes 6-9, following PCR with an annealing temperature of 55°C. *GBP* showed greater amplification than the *F-BOX* housekeeping gene, so *GBP* was selected for further optimisation tests to find the annealing temperature to produce sufficient levels of amplification for use in qPCR.

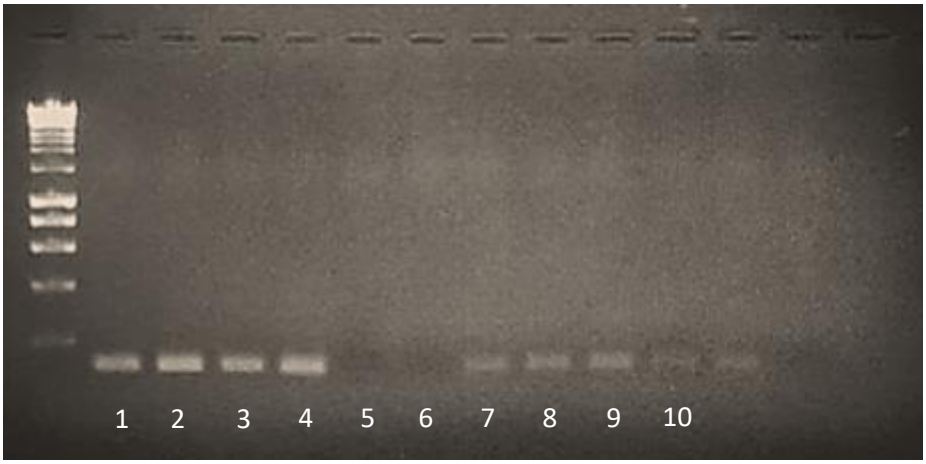

Figure 4.5. Gel electrophoresis showing the amplification of the *GBP* housekeeping gene in lanes 1-4 and of the *F-BOX* gene in lanes 7-10, following PCR with an annealing temperature of 52°C. Sufficient amplification of *GBP* was produced for use in qPCR.

4.5 Quantitative PCR (qPCR)

Healthy *N. benthamiana* plants were grown in the same conditions (16h/8h light/dark cycle at 25°C and watered when required) as those infected with CBSV via agroinfiltration with the CBSV Tanza infectious clone, or with TMV via sap inoculation. Symptoms of CBSV were observed at fourteen days post-inoculation (Fig. 4.6). Tissue samples were collected from the shoot tips, stems, mature leaves and young leaves of the healthy and virus-infected plants eighteen days post-inoculation. RNA was extracted from the tissue samples using the TriZol protocol, before complementary (c)DNA was

produced using first-strand cDNA synthesis following DNase treatment of the extracted RNA. PCR was then carried out on the cDNA produced from the CBSV infected plant tissue to confirm the presence of the virus (Fig. 4.7).

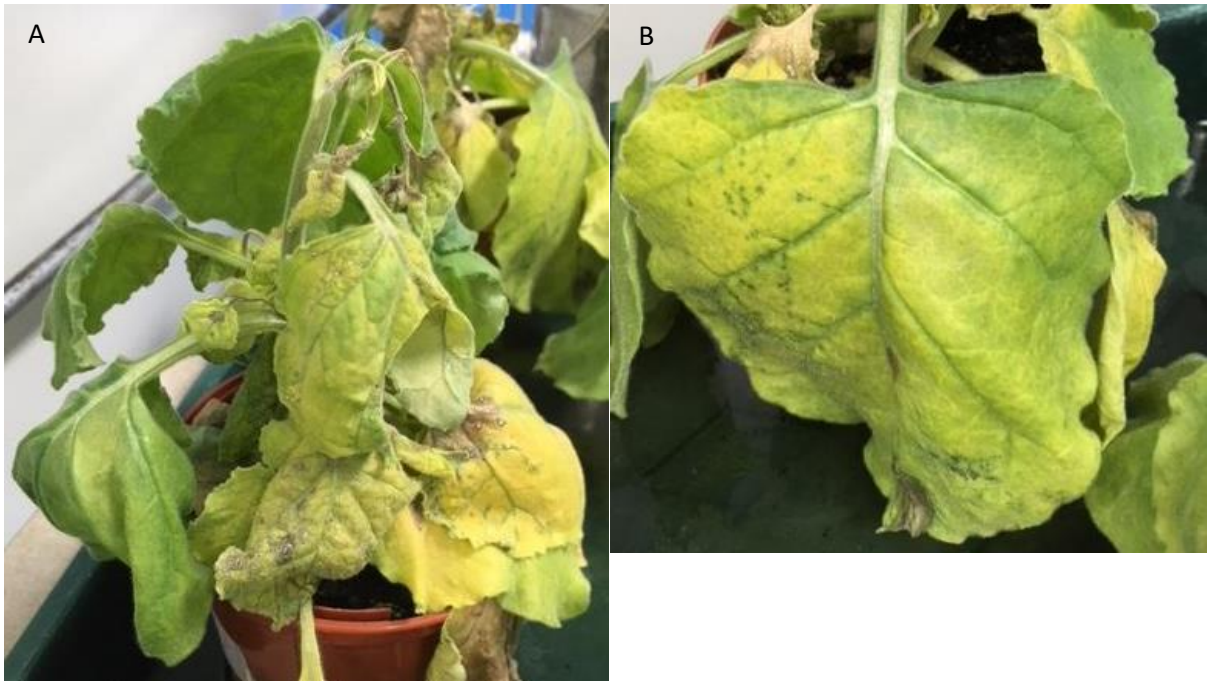

Figure 4.6. Symptoms of CBSV observed on *N. benthamiana* plant (A) and *N. benthamiana* leaf (B) fourteen days post inoculation.

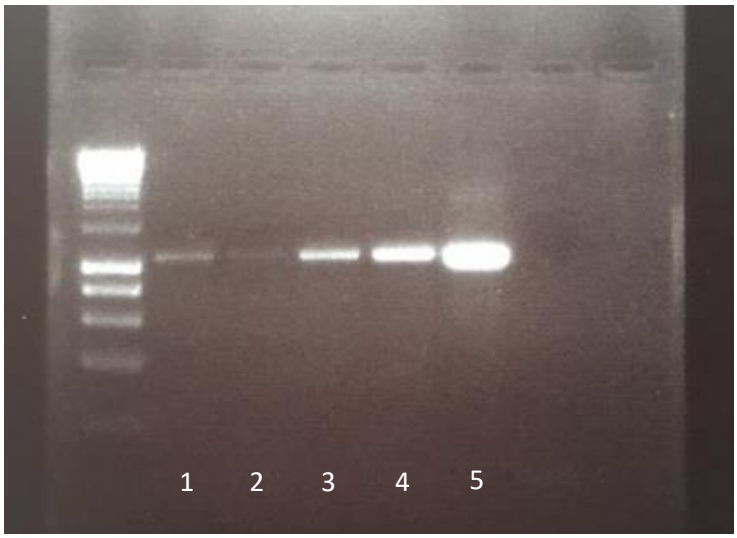

Figure 4.7. Confirmation of viral infection by reverse-transcription PCR (RT-PCR) using primer pairs for section 1 (1kbp) of the CBSV infectious clone, with samples from the shoot tip (1), stem (2), mature leaves (3) and young leaves (4). Variations in band strength were due to the amount of cDNA synthesised from each RNA sample which were not adjusted before use in the PCR. The CBSV infectious clone (5) band, used as a control, is much brighter as the sample was at a much higher concentration.

Following confirmation that the plants were infected with CBSV, qPCR could then be carried out. The concentration of the cDNA synthesised from the purified RNA for each sample was determined by nanodrop and adjusted through dilution, so all samples contained 500ng cDNA. The cDNA from the

healthy, CBSV infected and TMV infected samples were then used in the qPCR procedure with the *HAM1* and *GBP* primer pairs and SYBR Green I fluorescent dye. The end-point qPCR products were then run on an agarose gel to check the size of the fragments produced to ensure the correct genes were amplified (Fig. 4.8).

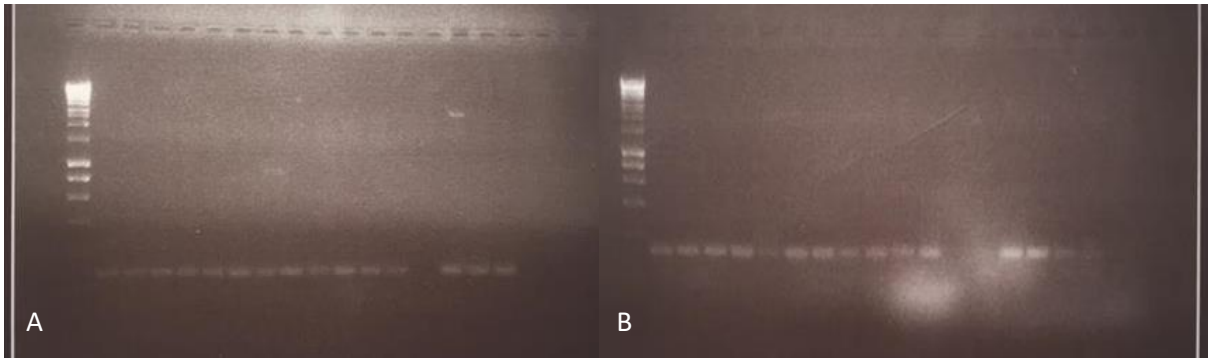

Figure 4.8. Gel electrophoresis of the qPCR products with the *GBP* fragments (A) and the *HAM1* fragments (B). Controls containing nuclease-free water are in the last two lanes (from left-to right) of each gel. The bands of the qPCR products in each lane were below the 200bp band on the hyperladder as the fragment sizes produced for the *GBP* and *HAM1* genes were approximately 120bp long.

The qPCR result output is an amplification plot which shows the Ct values of each sample. These Ct values were used in $2^{-\Delta\Delta Ct}$ analysis to establish the relative expression of the *HAM1* transcript levels in the different parts of the healthy, CBSV infected or TMV infected plant compared to the housekeeping *GBP* transcript levels (Fig. 4.9).

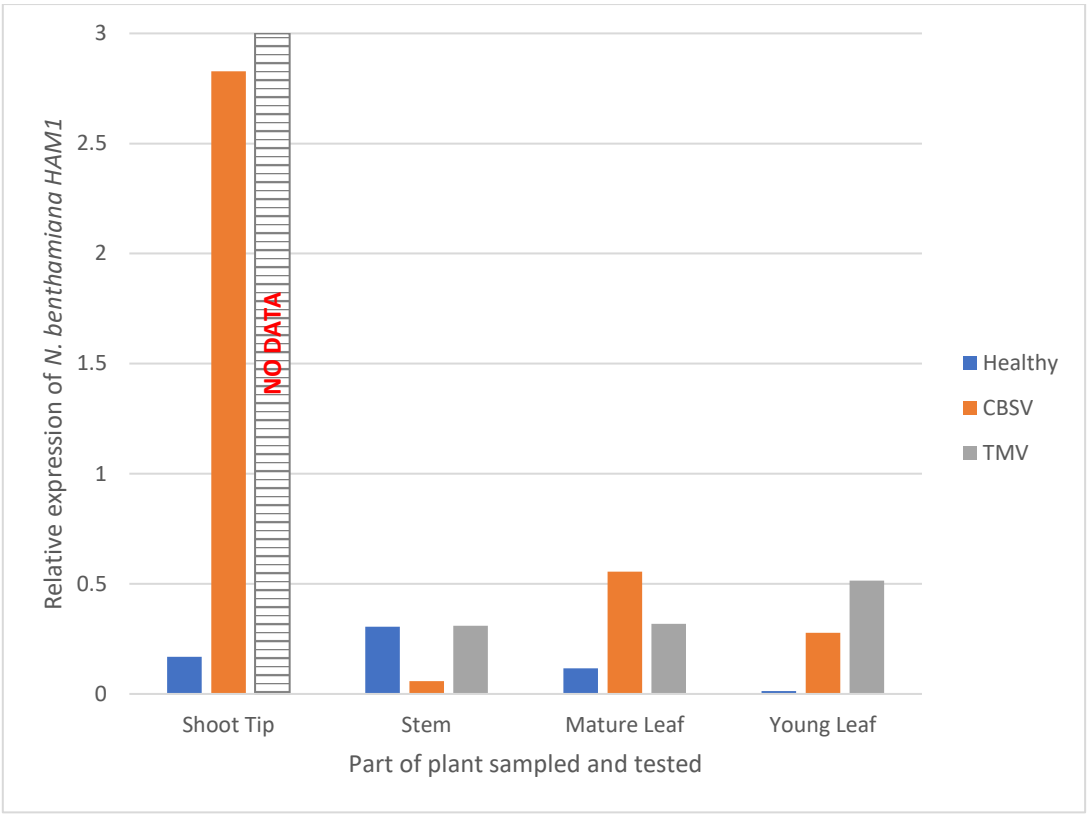

Figure 4.9. Bar chart showing the relative expression of the *N. benthamiana HAM1* transcript levels compared to that of the housekeeping gene *GBP*, in different parts of the plant when healthy, infected with CBSV or infected with TMV (n=1). No value was obtained from the qPCR for the shoot tip sample of the TMV infected *N. benthamiana* plants.

Following the analysis of the Ct values, it is observed that the *HAM1* transcript levels are relatively low in all parts of a healthy *N. benthamiana* plant. However, in areas where rapid cell division occurs, such as in the shoot tip and young leaves, there seems to be no apparent increase in *HAM1* expression. There are significantly higher levels of expression of the *N. benthamiana HAM1* in the shoot tip of the CBSV infected plant compared to the *HAM1* transcript levels in the shoot tip of a healthy plant. Whether this is a plant response specific to CBSV is not known as no result was obtained from the TMV infected *N. benthamiana* plants due to RNA degradation.

The samples taken from the stems of healthy and TMV infected plants both show greater levels of HAM1 expression than stem samples from CBSV infected plants. For those CBSV infected plants, it could be that the infection has yet to induce an oxidative stress response from the plant in the stem, as perhaps the virus remains undetected in the phloem. Why the *HAM1* transcript levels are lower than those in the healthy plant, however, remains to be explained.

In both the mature and young leaf samples, the greatest *HAM1* transcript levels are seen in the CBSV and TMV infected plants. It was expected that there would be higher *N. benthamiana HAM1* transcript levels in the infected plants as the plant would have been subject to oxidative stress in those areas

and therefore would have produced greater levels of HAM1 to protect against the increase in non-canonical nucleotides.

As no prior work has been done on the transcript levels of the *N. benthamiana HAM1* when infected with CBSV, it is important to highlight that this investigation was very much a pilot study to ascertain the possible effect that CBSV infection may have. Furthermore, due to time constraints, there were no repeats carried out in this study and so this can only act as a guide to inform future experiments.

CHAPTER 5 – DISCUSSION AND CONCLUSION

CBSV is one of two viruses belonging to the *Ipomovirus* genus in the *Potyviridae* family to contain a HAM1h protein. Bioinformatic analysis has shown it shares some 33 amino acid residues that are highly conserved across all HAM1 homologues from a variety of organisms, from prokaryotes to eukaryotes, yet the precise function of the viral protein remains unknown (Mbanzibwa *et al*., 2009a; 2011a). The aim of this thesis was to help establish what the function of the viral protein could be, with this being investigated in two ways; firstly, the functionality of the protein was tested in both wild type and *ham1* mutant strains of *S. cerevisiae*, and secondly, the expression of the *HAM1* gene in *N. benthamiana* was investigated to see what effect infection with CBSV would have.

5.1 CBSV HAM1H FUNCTIONALITY IN YEAST

Extensive work has been carried out on ITPase function in humans with respect to cancer biology and acquired resistance to chemotherapy drugs, with many studies being conducted using both *E. coli* and *S. cerevisiae* as model organisms and the functionality of their respective ITPases when exposed to mutagens (Carlsson *et al*., 2013). The ITPase family, in which HAM1 and its homologues can be found, represent a highly conserved superfamily of NTPases (Galperin *et al*., 2006) which act upon (d)ITP and (d)XTP to prevent the incorporation into replicating nucleic acids of hypoxanthine or xanthine respectively (Savchenko *et al*., 2007). ITPases thus perform a vital function to ensure high-fidelity genomic replication and furthermore, their importance is shown through their highly conserved functionality from bacteria and archaea, through to eukaryotes (Stepchenkova *et al*., 2009b).

Studies have shown that the yeast HAM1 and its homologues in *E. coli*, *Methanococcus jannaschii* and in humans, are able to dephosphorylate the non-canonical purine and pyrimidine base analogues in the cellular nucleotide precursor pools, preventing their incorporation into the replicating DNA and RNA, and due to the high level of conservation, can compensate for mutant defects in each other (Kozmin *et al*., 1998; Hwang *et al*., 1999; Stepchenkova *et al*., 2009b). As the viral HAM1 protein sequence shares 33 conserved amino acid residues with the yeast HAM1 and its homologues (Mbanzibwa *et al*., 2009a; 2011a), it was hypothesised that the viral protein would also show similar protective activity. Spot assays were conducted with both wild type and *ham1* mutant *S. cerevisiae* strains to establish if through overexpression of the viral *HAM1* gene, a similar phenotype would be observed as those overexpressing the yeast *HAM1*, and whether through expression of the viral HAM1 in the *ham1* mutant strains, if the yeast HAM1 function can be complemented. Due to HAP being

unavailable, cycloheximide (CHX), methotrexate (MTX) and the pyrimidine base analogue, 5-fluorouracil (5-FU), were used.

Work carried out by Kozmin *et al*. (1998) showed that HAP-induced mutagenesis in *E. coli*, through the mispairing of HAP base pairs, is not recognised by the cell repair systems nor is it proof-read. This was also observed for HAP-induced mutagenesis in *S. cerevisiae* and highlights the need for cells to eliminate the mutagenic non-canonical nucleotides before their incorporation into the replicating DNA. Additionally, a double mutant in the *E. coli* HAM1 homologue, RdgB, elicits the formation of a lesion in the DNA that ultimately makes the cell inviable through degradation of the chromosome, and further corroborates the findings that RdgB functions to prevent the incorporation of these non-canonical nucleotides into the replicating DNA (Clyman and Cunningham, 1987; Burgis *et al*., 2003). The HAM1 homologue of *E. coli*, RdgB, has been extensively studied. Findings by Savchenko *et al*. (2007) show it forms a homodimer containing 12 α-helices and 6 β-sheets, with a large cavity between the two protein lobes. Found within this cavity is the Serine-Histidine-Arginine (SHR) motif that is common to all HAM1 homologues and therefore suggests that this highly conserved sequence forms part of the active site of not only the *E. coli* HAM1 homologue, but of all ITPases (Savchenko *et al*., 2007). Bioinformatic analysis (see section 4.2) of the HAM1 sequences from *S. cerevisiae*, CBSV and UCBSV, showed that they all contained this highly conserved SHR motif, providing further evidence to suggest that the viral HAM1h proteins could also display nucleoside pyrophosphatase activity.

These prior studies have mostly used 6-*N*-hydroxylaminopurine (HAP), a purine analogue (Carlsson *et al*., 2018), as it is a potent mutagen for both prokaryotic and eukaryotic organisms and thus has allowed for the extensive study of mutant phenotypes in both *E. coli* and *S. cerevisiae*. Most notably, *S. cerevisiae ham1* mutant strains demonstrate hypersensitivity to this potent mutagen, yet studies have shown the HAM1 function can be complemented (Noskov *et al*., 1996; Kozmin *et al*., 1998). Work by Kozmin *et al*. (1998) showed that the HAM1 protein does not function in the repair of DNA but rather has enzymatic activity that breaks down the non-canonical nucleotides into their constituent monophosphates. This was later supported in the findings by Carlsson *et al*. (2018) that suggested the HAM1 protein does have pyrophosphatase activity. It was first thought that the HAM1 protein only provided specific resistance to HAP (Pavlov, 1986), yet further studies have shown it also confers resistance to several pyrimidine analogues, including 6-azauracil and 5-bromodeoxyuridine (Takayama *et al*., 2007; Carlsson *et al*., 2013), as well as exhibiting weaker resistance to the folate analogue metabolic inhibitor, MTX (Carlsson *et al*., 2013). In the current study, however, no such evidence of protective activity towards MTX was displayed in the phenotypes of yeast cells expressing or overexpressing either the yeast or CBSV *HAM1* genes. It is thought, therefore, that perhaps the GAL induced overexpression of the *HAM1* gene was too much of a burden on the cells and thus did not

confer any increased resistance, or indeed that the phenotype is very condition-dependent and suitable growth conditions were not created in these assays. Growth of *ham1* mutant yeast expressing the yeast *HAM1* gene was not always observed at some cell dilutions, particularly at 60µg/ml MTX concentration, thus comparisons with the activity of the viral HAM1h could not be made but does suggest that the CBSV HAM1h can provide a similar level of resistance to MTX as the yeast HAM1, as was seen at other MTX concentrations at the same cell serial dilutions. More assays would need to be conducted, however, to reliably ascertain the phenotypic responses.

As HAP is a nucleotide analogue, it thus provides a direct assay to investigate the function of the HAM1 protein and HAM1 homologues. Due to its toxicity, however, HAP is no longer commercially available and as a consequence of it not being available for use in this investigation, there were constraints in the chemicals that could be used to test the HAM1 function. Like HAP, 5-FU is a 'pure' base analogue that only induces mutagenesis through its misincorporation into the replicating nucleic acid and causes further damage through the subsequent repair mechanisms to remove these mismatched bases. These 'pure' nucleotide analogues do not damage other cellular components and their lethal effects can be attributed to mutations in specific vital genes (Pavlov *et al*., 1991), thus providing a direct assay for the testing of the mutant strains.

Surprisingly, however, no effect was seen on either the overexpression of the yeast or CBSV *HAM1* genes on the resistance to 5-FU. There was, however, increases in growth observed for both the pYES2_HAM1_Sc and pYES2_HAM1_CBSV *ham1* knockout *S. cerevisiae* transformed cells compared to the growth of cells not expressing any *HAM1* at the two lowest concentrations of 5-FU, and thus suggests that expression of the CBSV *HAM1* gene can provide a similar level of resistance to 5-FU as the yeast *HAM1*. At the higher concentrations, the expression of either *HAM1* gene was possibly not sufficient to confer resistance to the pyrimidine analogue and suggests that these concentrations of 5-FU were too high for use in the spot assays. A study by Carlsson *et al*. (2013) found that a deletion of the *HAM1* gene had little effect on the yeast sensitivity to 5-FU and thus, the effects of the deletion would not be significantly observable in these spot assays. Furthermore, Carlsson *et al*. (2013) only produced significant results between the wild type and *ham1* mutant strains when the yeast cells were forced to use the 5-FU as their uracil source, which was not the case in these spot assays. Taken together, these studies support at least in part, that the CBSV HAM1h is indeed a functional HAM1 as similar phenotypes are seen if overexpressing the yeast or CBSV HAM1. For a definitive outcome, more sensitive assays would need to be conducted, preferably using HAP as the mutagen as the *HAM1* deletion in yeast provides a hypersensitive phenotype that is clearly observed (Noskov *et al*., 1996).

The other mutagens studied, CHX and MTX, are not base analogues and thus do not provide a direct assay; instead, they target other cellular components and protein synthesis which subsequently causes imbalances in the nucleotide precursor pools (Longley *et al*., 2003; Alamgir *et al*., 2010; Buchanan *et al*., 2016). Their effect, therefore, is complex and indirect, with no one mutation in specific vital genes attributable to their toxic effects (Pavlov *et al*., 1991), thus making their phenotypes much harder to observe. The HAM1 protein, however, can still afford some protection against these molecules through their indirect action of causing oxidative stress within the cells which leads to the production of non-canonical nucleotides. Results from the CHX spot assay, moreover, show that the overexpression of both the yeast and CBSV *HAM1* genes did confer greater resistance to the effects of CHX at the greatest cell serial dilutions at the CHX concentration of 45ng/ml, and more importantly, that the CBSV HAM1h appears to confer similar protective activity to CHX as the yeast HAM1; a result that was also observed at the 10^{-4} cell dilution at the CHX concentration of 130ng/ml. The spot assays conducted to determine whether the *HAM1* genes from CBSV and yeast can restore resistance to CHX revealed no apparent effect on their expression at the two lowest CHX concentrations studied. It might be that this concentration was too low for the slight phenotypic effects of the expression of the *HAM1* genes to be observed. At the CHX concentration of 100ng/ml, however, at the two greatest cell serial dilutions, an effect due to the expression of both the yeast and CBSV *HAM1* genes was noted and an increase in resistance to CHX was confirmed. Similarly, this result could be inferred at the 10^{-4} cell dilution at a concentration of 130ng/ml, though it could not be confirmed through the lack of results obtained for cells not expressing a *HAM1* gene.

Overall, results from the CHX spot assays suggest that the CBSV HAM1h is likely to show pyrophosphatase activity, replicating the function of the yeast HAM1 in conferring some increased resistance to CHX, however, further studies would need to be conducted on the effects of the viral HAM1 protein to a series of purine and pyrimidine analogues. In the future, these studies would preferably be carried out in an *E. coli* background as the prokaryotic cells are more sensitive to a broader range of nucleotide analogues (Kozmin *et al*., 1998). For this study, however, the *E. coli* strain with the correct mutation was not available.

5.2 Effect of CBSV infection on plant *HAM1* transcript levels

Plants use a variety of mechanisms to defend themselves against pathogens, with most comprising preventative measures to inhibit the entry of pathogens. With viruses, however, plants are unable to deploy these mechanisms as their preventative defences have already been breached. Once the virus has become a part of the cell, the plant can only respond by inducing oxidative stress. This oxidative stress provides two vital roles in the plant's fight against the pathogen; firstly, it elicits localised

necrosis of infected tissue and containment of the pathogen, and secondly, provides a diffusible signal to stimulate antioxidant and disease response genes (Demidchik, 2015; Hernández *et al*., 2016). These responses were first noted in the defence against bacterial and fungal pathogens but have also been observed in virus-infected plants. (Hernández *et al*., 2016).

A result of this oxidative stress, however, is the accumulation of non-canonical nucleotides as canonicals bases in the precursor pools are damaged, and the likely subsequent increase in the rate of mutation through their misincorporation into synthesising DNA. It is therefore imperative that the plant produces its own NTPase, also known as HAM1, to protect itself from the incorporation of the non-canonical nucleotides if it is to overcome infection. Through expression of its own *HAM1* gene when infected with viruses, however, the plant would simultaneously be providing protection to the virus and with the cycloheximide spot assays highlighting that the CBSV HAM1h does likely show pyrophosphatase activity, the question was posed why CBSV needs a HAM1 protein when the plant produces its own.

It was therefore hypothesised that perhaps the plant downregulates its *HAM1* when it is infected with a virus to create hostile conditions for the virus, with the increases in non-canonical nucleotides providing high rates of mutagenesis that may render the virus inviable and cause an extinction in the population (Mbanzibwa *et al*., 2009b). This hypothesis would provide a reason why CBSV has acquired and retained its own *HAM1* gene. Alternatively, the plant may not downregulate its *HAM1* when infected with CBSV in order to protect itself from those non-canonical DNA nucleotides that have arisen as a result of oxidative stress. The virus may then have acquired its own HAM1 as the plant ITPase may afford insufficient protection against the RNA non-canonical nucleotides that have formed in the precursor pools and that the virus will depend upon to replicate. Most eukaryotes NTPases act against DNA non-canonical bases and to a lesser extent on non-canonical RNA nucleotides, and thus the CBSV HAM1h may be specific to RNA non-canonical nucleotides. If this is the case, perhaps the virus requires the plant HAM1 in addition to its own to overcome the stresses of those non-canonical RNA bases.

In order to establish whether the plant *HAM1* is either downregulated or overexpressed upon infection with CBSV, quantitative PCR (qPCR) was carried out on *N. benthamiana* plants when infected with CBSV or *Tobacco mosaic virus* (TMV) and compared to healthy plants. *N. benthamiana* has been shown to be a suitable host for the study of the infection of CBSV (Bua and Namara, 2009; Ogwok *et al*., 2010), with plants showing severe symptoms leading to necrosis (Patil *et al*., 2015) across all parts of the plant which is unlike what is seen in cassava plants where the majority of foliar symptoms, if any, are only observed in the mature leaves (Mbanzibwa *et al*., 2009b; Abarshi *et al*., 2010).

Results of the qPCR investigation, following analysis, showed that the *N. benthamiana HAM1* is clearly not being downregulated when infected with either CBSV or TMV. This is shown by the increases in *N. benthamiana HAM1* transcript levels in the mature and young leaves when infected. The greater level of plant *HAM1* expression in these parts of the plant will likely be in response to the oxidative stress that is being induced from the developing infection from both CBSV and TMV – shown through the yellowing of the leaves.

A very large increase in the *N. benthamiana HAM1* transcript levels was observed in the shoot tip of those plants infected with CBSV. Samples were collected from the plants eighteen days post-inoculation, and with CBSV infection causing the shoot tip to become necrotic about fourteen days post-inoculation, this part of the plant would have been under lots of oxidative stress at the time. These high levels of oxidative stress would thus induce increased expression of the plant *HAM1* to try and combat this increase in non-canonical nucleotide levels. It is unclear, however, if the same would have been seen for those *N. benthamiana* infected with TMV as no result was obtained for the shoot tip sample. As TMV does not cause necrosis to shoot tips, it could be that the *HAM1* levels in the plants infected with TMV would show a similar level of *HAM1* expression as the other tissues taken from the TMV infected plants. Yet, as no result was obtained for the shoot tip sample due to RNA degradation, further experiments will be required to determine if the increase in plant *HAM1* expression is a result of only CBSV infection or indeed a generalised response for any viral infection.

The decrease in *N. benthamiana HAM1* transcript levels in the stems of those infected with CBSV are likely due to the low levels of oxidative stress occurring as the virus is yet to take hold in the stems. Similarly, the same level of *HAM1* transcript was observed in *N. benthamiana* plants infected with TMV as in the healthy plants, and suggests that there was little, if any, response from the plant to induce oxidative stress in the stems. Therefore, the lack of oxidative stress meant that there was no need for increased *HAM1* expression.

Experiments for this investigation were only able to be carried out once due to time constraints, therefore, this was a pilot study looking into the effects CBSV infection may have on *N. benthamiana HAM1* transcript levels. Overall, the results from this investigation suggest that as the *N. benthamiana HAM1* expression is not downregulated upon infection with CBSV, the CBSV HAM1 could be functioning to afford protection for itself against non-canonical RNA nucleotides and ensure the virus is not subject to any detrimental mutations. As no study of the effect of viral infection on plant *HAM1* expression has been conducted before, these results present a foundation for further study into the interactions of the plant and viral HAM1 proteins, and indeed for the effect on cassava *HAM1* transcript levels. Cassava is a longer living plant than *N. benthamiana* and thus the effects of the viral

HAM1 may be more pronounced in the actual viral host. Interestingly, a similar result was found by Martelli *et al*. (2007) where the *Blackberry virus Y* exhibited an AlkB domain in its P1 proteinase that conferred repair mechanisms to methylation damage of nucleic acids, specifically RNA. Furthermore, as this virus infects perennial, woody plants, it was suggested that this RNA repair was advantageous in providing stability for the virus in longer-lived hosts and through changing environmental conditions (Martelli *et al*., 2007; Mbanzibwa *et al*., 2009b).

5.3 CONCLUSION

Evidence produced through this project supports the hypothesis that the CBSV HAM1h does show pyrophosphatase activity, just like its homologues. Upon infection in the plant, which induces high levels of oxidative stress in response to the CBSV infection, the virus is likely to experience high levels of mutagenesis as a result of the incorporation of the non-canonical nucleotides into its replicating RNA. This will cause high-infidelity genomic replication, and lead to the virus no longer being viable and eventually lead to extinction for the virus population (Mbanzibwa *et al*., 2009b). Therefore, as with other organisms containing a HAM1, it is imperative for the virus to be able to combat the high levels of non-canonical nucleotides and prevent their incorporation into its replicating RNA. The results of the effect of CBSV infection on the plant *HAM1* transcript levels, shows that the viral HAM1h is likely to target non-canonical RNA nucleotides.

With the advances being made in generating full-length infectious clones of CBSV, the ability to cause specific mutations within viral genes will enable further investigations to take place looking into the function of the CBSV HAM1h. In the future, the targeting of this gene could provide a new source of CBSD resistance through the use of RNAi mechanisms, or through the mutation of this gene such that the virus can no longer afford protection for itself against the non-canonical nucleotides.

REFERENCES

Abarshi, M. M., Mohammed, I. U., Wasswa, P., Hillocks, R. J., Holt, J., Legg, J. P., Seal, S. E. and Maruthi, M. N. 2010. Optimization of diagnostic RT-PCR protocols and sampling procedures for the reliable and cost-effective detection of *Cassava brown streak virus*. *Journal of Virological Methods* **163**(2):353-359

Abass, A. B., Awoyale, W. Alenkhe, B., Malu, N., Asiru, B. W., Manyong, V. and Sanginga, N. 2018. Can food technology innovation change the status of a food security crop? A review of cassava transformation into "bread" in Africa. *Food Reviews International* **34**(1):87-102

Adams, M. J., Antoniw, J. F. and Fauquet, C. M. 2005a. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Archives of Virology* **150**:459-479

Adams, M. J., Antoniw, J. F. and Beaudoin, F. 2005b. Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. *Molecular Plant Pathology* **6**(4):471-487

Alamgir, M., Erukova, V., Azizi, A. and Golshani, A. 2010. Chemical-genetic profile analysis of five inhibitory compounds in yeast. *BMC Chemical Biology* **10**:6

Alicai, T., Omongo, A. Maruthi, M. N., Hillocks, R. J., Baguma, Y., Kawuki, R., Bua, A., Otim-Nape, G. W. and Colvin, J. 2007. Re-emergence of Cassava Brown Streak Disease in Uganda. *Plant Disease* **91**(1):24-29

Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., Kiiza, L., Kubatko, L., Kehoe, M. A. and Boykin, L. M. 2016. *Cassava brown streak virus* has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. *Scientific Reports* **6**:36164 doi:10.1038/srep36164

Anjanappa, R. B., Mehta, D., Okoniewski, M. J., Szabelska, A., Gruissem, W. and Vanderschuren, H. 2017. Early transcriptome analysis of the brown streak virus-cassava pathosystem provides molecular insights into virus susceptibility and resistance. *BioRxiv* doi:10.1101/100552

Ateka, E., Alicai, T., Ndunguru, J., Tairo, F., Sseruwagi, P., Kiarie, S., Makori, T., Kehoe, M. A. and Boykin, L. M. 2017. Unusual occurrence of a DAG motif in the Ipomovirus *Cassava brown streak virus* and implication for its vector transmission. *PLoS ONE* **12**(11):e0187883

Bennet, B. 2015. Guest editorial: smallholder cassava production and the cassava processing sector in Africa. *Food Chain* **5**:1–3

Beyene, G., Chauhan, R. D., Wagaba, H., Moll, T., Alicai, T., Miano, D., Carrington, J. C. and Taylor, N. J. 2016. Loss of CMD2-mediated resistance to cassava mosaic disease in plants regenerated through somatic embryogenesis. *Molecular Plant Pathology* **17**(7):1095-1110

Blagbrough, I. A., Bayoumi, S. A. L., Rowan, M. G. and Beeching, J. R. 2010. Cassava: An appraisal of its phytochemistry and its biotechnological prospects. *Phytochemistry* **71**(17-18):1940-1951

Bua, B. and Namara, J. 2009. Reaction of *Nicotiana* species to Cassava Brown Streak Virus from Uganda. In *9th African Crop Science, Conference Proceedings, Cape Town, South Africa, 28 September - 2 October 2009* (pp. 647-650). African Crop Science Society

Buchanan, B. W., Lloyd, M. E., Engle, S. M. and Rubenstein, E. M. 2016. Cycloheximide Chase Analysis of Protein Degradation in *Saccharomyces cerevisiae*. *Journal of Visualized Experiments* **110**:53975

Burgis, N. E., Brucker, J. J. and Cunningham, R. P. 2003. Repair System for Noncanonical Purines in *Escherichia coli*. *Journal of Bacteriology* **185**(10):3101-3110

Carbonell, A., Dujovny, G., García, J. A. and Valli, A. 2012. The *Cucumber vein yellowing virus* Silencing Suppressor P1b Can Functionally Replace HCPro in *Plum pox virus* Infection in a Host-Specific Manner. *Molecular Plant-Microbe Interactions* **25**(2):151-164

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**(10):e52094

Carlsson, M., Hu, G-Z. and Ronne, H. 2018. Gene dosage effects in yeast support broader roles for the *LOG1*, *HAM1* and *DUT1* genes in detoxification of nucleotide analogues. *PLoS ONE* **13**(5): e0196840

Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**(1):156-159

Clyman, J. and Cunningham, R. P. 1987. *Escherichia coli* K-12 Mutants in Which Viability Is Dependent on *recA* Function. *Journal of Bacteriology* **169**(9):4203-4210

Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M. C. and Carrington, J. C. 1995. Long-distance movement factor: a transport function of the potyvirus helper component proteinase. *Plant Cell* **7**:549-559

Demidchik, V. 2015. Mechanisms of oxidative stress in plants: From classical chemistry to cell biology. *Environmental and Experimental Botany* **109**:212-228

Dombrovsky, A., Reingold, V. and Antignus, Y. 2014. *Ipomovirus* – an atypical genus in the family *Potyviridae* transmitted by whiteflies. *Pest Management Science* **70**(10):1553-1567

Earley, K. W., Shook, M. S., Brower-Toland, B., Hicks, L. and Pikaard, C. S. 2007. *In vitro* specificities of *Arabidopsis* co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. *The Plant Journal* **52**:615-626

Fokunang, C. N., Dixon, A. G. O., Ikotun, T., Asiedu, R., Tembe, E. A. and Akem, C. N. 2001. *In vitro*, greenhouse and field assessments of cassava lines for resistance to anthracnose disease caused by *Colletotrichum gloeosporioides* f.sp. *manihotis*. *Mycopathologia* **154**:191-198

Fondong, V. N. 2017. The Search for Resistance to Cassava Mosaic Geminiviruses: How Much We Have Accomplished, and What Lies Ahead. *Frontiers in Plant Science* **8**:408

Galperin, M. Y., Moroz, O. V., Wilson, K. S. and Murzin, A. G. 2006. House cleaning, a part of good housekeeping. *Molecular Microbiology* **59**:5-19

Gerlinger, U-M., Gückel, R., Hoffmann, M., Wolf, D. H. and Hilt, W. 1997. Yeast Cycloheximide-resistant *crl* Mutants Are Proteasome in Protein Degradation. *Molecular Biology of the Cell.* **8**(12):2487-2499

Gietz, R. D. and Schiestl, R. H. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**:31-34

Giner, A., Lakatos, L., García-Chapa, M., López-Moya, J. J. and Burgyán, J. 2010. Viral Protein Inhibits RISC Activity by Argonaute Binding through Conserved WG/GW Motifs. *PLoS Pathogens* **6**(7):e1000996

Govier, D. A., Kassanis, B. and Pirone, T. P. 1977. Partial purification and characterization of the potato virus Y helper component. *Virology* **78**:306-314

Hernández, J. A., Gullner, G., Clemente-Moreno, M. J., Künstler, A., Juhász, C., Díaz-Vivancos, P. and Király, L. 2016. Oxidative stress and antioxidative responses in plant-virus interactions. *Physiological and Molecular Plant Pathology* **94**:134-148

Hillocks, R. J., Raya, M. and Thresh, J. M. 1996. Factors affecting the distribution, spread and symptom expression of Cassava brown streak disease in Tanzania. *International Journal of Pest Management* **42**:285-289

Hillocks, R. J., Raya, M. D., Mtunda, K. and Kiozia, H. 2001. Effects of Brown Streak Virus Disease on Yield and Quality of Cassava in Tanzania. *Journal of Phytopathology* **149**(7-8):389-394

Hillocks, R. J. and Jennings, D. L. 2003. Cassava brown streak disease: a review of present knowledge and research needs. *International Journal of Pest Management* **49**(3):225-234

Hillocks, R. J. and Thresh, J. M. 2000. Cassava Mosaic and Cassava Brown Streak Virus Diseases in Africa: A comparative guide to symptoms and aetiologies. *Roots* **7**(1) The Southern Africa (SARRNET) and The East Africa Root Crops Research Network (EARRNET)

Howeler, R. 2006. Cassava in Asia: trends in cassava production, processing and marketing. *Workshop on "Partnership in modern science to develop a strong cassava commercial sector in Africa and appropriate varieties by 2020*." Bellagio, Italy.

Howeler R., Lutaladio N. and Thomas G. 2013. Save and grow: cassava – a guide to sustainable production intensification. *Rome, Italy: Food and Agriculture Organization of the United States of America.*

Hwang, K. Y., Chung, J. H., Kim, S-H., Han, Y. S. and Cho, Y. 1999. Structure-based identification of a novel NTPase from *Methanococcus jannaschii*. *Nature Structural & Molecular Biology* **6**:691-696

Jameson, J. D. 1964. Cassava mosaic disease in Uganda. *East African Agricultural and Forestry Journal* **29**(3)

Jennings, D. L. 1957. Further studies in breeding cassava for virus resistance. *The East African Journal* **22**(4):213-219

Jennings, D. L. 1960. Observations on virus diseases of cassava in resistant and susceptible varieties. II. Brown streak disease. *Empire Journal of Experimental Agriculture* **28**(111):261-270

Jennings, D. L. 2003. Historical perspective on breeding for resistance to cassava brown streak virus disease. *Cassava Brown Streak Virus Disease: Past, Present, and Future* pp.27-30

Johnson, S. N., Karley, A. J., Gregory, P. J. and Brennan, R. M. 2017. Editorial: Crop Traits for Defence against Pests and Disease: Durability, Breakdown and Future Prospects. *Frontiers in Plant Science* **8**:209

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. *Annual Research & Review in Biology* **10**(5):1-9

Kaweesi, T., Kawuki, R., Kyaligonza, V., Baguma, Y., Tusiime, G. and Ferguson, M. E. 2014. Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. *Virology Journal* **11**:216-230

Kawuki, R. S., Kaweesi, T., Esuma, W., Pariyo, A., Kayondo, I. S., Ozimati, A., Kyaligonza, V., Abaca, A., Orone, J., Tumuhimbise, R., Nuwamanya, E., Abidrabo, P., Amuge, T., Ogwok, E., Okao, G., Wagaba,

H., Adiga, G., Alicai, T., Omongo, C., Bua, A., Ferguson, M., Kanju, E. and Baguma, Y. 2016. Eleven years of breeding efforts to combat Cassava brown streak disease. *Breeding Science* **66**(4):560-571

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*. *Archives of Virology* **162**(1):291-293

Kozmin, S. G., Leroy, P. and Pavlov, Y. I. 1998. Overexpression of the yeast *HAM1* gene prevents 6-*N*-hydroxylaminopurine mutagenesis in *Escherichia coli*. *Acta Biochimica Polonica* **45**(3):645-652

Kozmin, S. G., Pavlov, Y. I., Dunn, R. L. and Schaaper, R. M. 2000. Hypersensitivity of *Escherichia coli Δ(uvrB-bio)* Mutants to 6-Hydroxylaminopurine and Other Base Analogs Is Due to a Defect in Molybdenum Cofactor Biosynthesis. *Journal of Bacteriology* **182**(12):3361-3367

Kuria, P., Ilyas, M., Ateka, E., Miano, D., Onguso, J., Carrington, J. C. and Taylor, N. J. 2017. Differential response of cassava genotypes to infection by cassava mosaic geminiviruses. *Virus Research* **227**:69-81

Latrasse, D., Benhamed, M., Henry, Y., Domenichini, S., Kim, W., Zhou, D-X. and Delarue, M. 2008. The MYST histone acetyltransferases are essential for gametophyte development in Arabidposis. *BMC Plant Biology* **8**:121

Legg, J. P. and Thresh, J. M. 2000. Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus Research* **71**(1):135-149

Legg, J. P., Owor, B., Sseruwagi, P. and Ndunguru, J. 2006. Cassava Mosaic Virus Disease in East and Central Africa: Epidemiology and Management of A Regional Pandemic. *Advances in Virus Research* **67**:355-418

Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G., Bouwmeester, H., Bigirimana, S., Tata-Hangy, W., Gashaka, G., Mkamilo, G., Alicai, T. and Lava Kumar, P. 2011. Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Research* **159**(2):161-170

Legg, J. P., Attiogbevi Somado, E., Barker, I., Beach, L., Ceballos, H., Cuellar, W., Elkhoury, W., Gerling, D., Helsen, J., Hershey, C., Jarvis, A., Kulakow, P., Kumar, L., Lorenzen, J., Lynam, J., McMahon, M., Maruthi, G., Miano, D., Mtunda, K., Natwuruhunga, P., Okogbenin, E., Pezo, P., Terry, E., Thiele, G., Thresh, M., Wadsworth, J., Walsh, S., Winter, S., Tohme, J. and Fauquet, C. 2014. A global alliance declaring war on cassava viruses in Africa. *Food Security* **6**(2):231-248

Legg, J.P., Kumar, L. and Kanju, E. E. 2015. Virus Diseases of Tropical and Subtropical Crop. *Chapter 4; CAB International*, Wallingford, United Kingdom ISBN: 9781780644264

Lennon, A., Aiton, M. and Harrison, B. 1985. Cassava viruses from Africa: third country quarantine. *Report for 1985. Scottish Crop Research Institute, Invergowrie, UK* pp. 168-169

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**(9):e46452

Longley, D. B., Harkin, D. P. and Johnston, P. G. 2003. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer* **3**:330-338

López, C. E. and Bernal, A. J. 2012. Cassava Bacterial Blight: Using Genomics for the Elucidation and Management of an Old Problem. *Tropical Plant Biology* **5**:117-126

Lozano, J. C. 1986. Cassava Bacterial Blight: A Manageable Disease. *Plant Disease* **70**(12):1089-1093

Martelli, G. P., Adams, M. J., Kreuze, J. F. and Dolja, V. V. 2007. Family *Flexiviridae*: A Case Study in Virion and Genome Plasticity. *Annual Review of Phytopathology* **45**:73-100

Martínez, F., Rodrigo, G., Aragonés, V., Ruiz, M., Lodewijk, I., Fernández, U., Elena, S. F. and Daròs, J-A. 2016. Interaction network of tobacco etch potyvirus NIa protein with the host proteome during infection. *BMC Genomics* **17**:87

Maruthi, M. N., Hillocks R. J., Mtunda, K., Raya, M. D., Muhanna, M., Kiozia, H., Rekha, A. R., Colvin, J. and Thresh, J. M. 2005. Transmission of *Cassava brown streak virus* by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology* **153**(5):307-312

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., Kullaya, A. and Valkonen, J. P. T. 2009a. Genetically distinct strains of *Cassava brown streak virus* in the Lake Victoria basin and the Indian Ocean coastal area of East Africa. *Archives of Virology* **154**(2):353-359

Mbanzibwa, D. R., Tian, Y., Mukasa, S. B. and Valkonen, J. P. T. 2009b. *Cassava Brown Streak Virus* (*Potyviridae*) Encodes a Putative Maf/HAM1 Pyrophosphatase Implicated in Reduction of Mutation and a P1 Proteinase That Suppresses RNA Silencing but Contains No HC-Pro. *Journal of Virology* **83**(13):6934-6940

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., Muli, M. B., Mukasa, S. B., Tairo, F., Baguma, Y., Kyamanywa, S., Kullaya, A., Maruthi, M. N., Fauquet, C. M. and Valkonen, J. P. T. 2011a. Evolution of cassava brown steak disease-associated viruses. *Journal of General Virology* **92**:974-987

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., Kullaya, A. and Valkonen, J. P. T. 2011b. 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*. *Journal of Virological Methods* **171**(2):394-400

McCallum, E. J., Anjanappa, R. B. and Gruissem, W. 2017. Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). *Current Opinion in Plant Biology* **38**:50-58

Menezes, R. M., Waisertreiger, I. S-R., Lopez-Bertoni, H., Luo, X. and Pavlov, Y. I. 2012. Pivotal Role of Inosine Triphosphate Pyrophosphatase in Maintaining Genome Stability and the Prevention of Apoptosis in Human Cells. *PLoS ONE* **7**(2): e32313

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. *Advances in Virology* **2012**:795697

Monger, W. A., Seal, S., Isaac, A. M. and Foster, G. D. 2001a. Molecular characterization of the *Cassava brown streak virus* coat protein. *Plant Pathology* **50**:527-534

Monger, W. A., Seal, S., Cotton, S. and Foster, G. D. 2001b. Identification of different isolates of *Cassava brown streak virus* and development of a diagnostic test. *Plant Pathology* **50**:768-775

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. and Smith, J. 2010. The complete genome sequence of the Tanzanian strain of *Cassava brown streak virus* and comparison with the Ugandan strain sequence. *Archives of Virology* **155**(3):429-433

Myers, S. S., Smith, M. R., Guth, S., Golden, C. D., Vaitla, B., Mueller, N. D., Dangour, A. D. and Huybers, P. 2017. Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition. *Annual Review of Public Health* **38**:259-277

Ndunguru, J., Sseruwagi, P., Tairo, F., Stomeo, F., Maina, S., Dijnkeng, A., Kehoe, M. and Boykin, L. M. 2015. Analyses of Twelve New Whole Genome Sequences of Cassava Brown Streak Viruses and Ugandan Cassava Brown Streak Viruses from East Africa: Diversity, Supercomputing and Evidence for Further Speciation. *PLoS ONE* **10**(10):e0141939

Nichols, R. F. W. 1946. Breeding cassava for virus resistance. *The East African Agricultural Journal* **12**(3):184-194

Nichols, R. F. W. 1950. The brown streak disease of cassava. *The East African Agricultural Journal* **15**(3):154-160

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

Nweke, F. 2004. New challenges in the cassava transformation in Nigeria and Ghana. *International Food Policy Research Institute* **118**

Ogwok, E., Patil, B. L., Alicai, T. and Fauquet, C. M. 2010. Transmission studies with *Cassava brown streak Uganda virus* (*Potyviridae: Ipomovirus*) and its interactions with abiotic and biotic factors in *Nicotiana benthamiana*. *Journal of Virological Methods* **169**:296-304

Ogwok, E., Alicai, T., Rey, M. E. C., Beyene, G. and Taylor, N. J. 2014. Distribution and accumulation of cassava brown streak viruses within infected cassava (*Manihot esculenta*) plants. *Plant Pathology* **64**(5):1235-1246

Olsen, K. M. and Schaal, B. A. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences of the United States of America* **96**:5586-5591

Owolade, O. F., Dixon, A. G. O., Adeoti, A. A. and Osunlaja, S. O. 2005. Sources of resistance to cassava anthracnose disease. *African Journal of Biotechnology* **4**(6):570-572

Pandey, A., Soccol, C. R., Nigam, P., Soccol, V. T., Vandenberghe, L. P. S. and Mohan, R. 2000. Biotechnological potential of agro-industrial residues. II: cassava bagasse. *Bioresource Technology* **74**(1):81-87

Pasin, F., Simón-Mateo, C. and García, J. A. 2014. The hypervariable amino-terminus of P1 protease modulates potyviral replication and host defense responses. *PLoS Pathogens* **10**:e1003985

Patil, B. L. and Fauquet, C. M. 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. *Molecular Plant Pathology* **10**(5):685-701

Patil B. L., Ogwok, E., Wagaba, H., Mohammed, I. U., Yadav, J. S., Bagewadi, B., Taylor, N. J., Kreuze, J. F., Maruthi, M. N., Alicai, T. and Fauquet, C. M. 2011. RNAi-mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. *Molecular Plant Pathology* **12**(1):31-41

Patil, B. L., Legg, J. P., Kanju, E. and Fauquet, C. M. 2015. Cassava brown streak disease: a threat to food security in Africa. *Journal of General Virology* **96**:956-968

Pavlov, Y. I. 1986. *Saccharomyces cerevisiae* mutants highly sensitive to the mutagenic action of 6-*N*-hydroxylaminopurine. *Soviet Genetics* **22**:1099-1107

Pavlov, Y. I., Noskov, V. N., Lange, E. K., Moiseeva, E. V., Pshenichnov, M. R. and Khromov-Borisov, N. N. 1991. The genetic activity of N^6-hydroxyadenine and 2-amino-N^6-hydroxyadenine in *Escherichia coli*, *Salmonella typhimurium* and *Saccharomyces cerevisiae*. *Mutation Research* **252**:33-46

Pavlov, Y. I., Suslov, V. V., Shcherbakova, P. V., Kunkel, T. A., Ono, A., Matsuda, A. and Schaaper, R. M. 1996. Base analog N^6-hydroxylaminopurine mutagenesis in *Escherichia coli*: genetic control and molecular specificity. *Mutation Research* **357**:1-15

Revers, F. and García, J. A. 2015. Molecular Biology of Potyviruses. *Advances in Virus Research* **92**:101-199

Saggaf, M. H., Ndunguru, J., Tairo, F., Sseruwagi, P., Ascencio-Ibáñez, J. T., Kilalo, D. and Miano, D. W. 2018. Immunohistochemical localization of *Cassava brown streak virus* and its morphological effect on cassava leaves. *Physiological and Molecular Plant Pathology* (In press)

Savchenko, A., Proudfoot, M., Skarina, T., Singer, A., Litvinova, O., Sanishvili, R., Brown, G., Chirgadze, N. and Yakunin, A. F. 2007. Molecular Basis of the Antimutagenic Activity of the House-Cleaning Inosine Triphosphate Pyrophosphatase RdgB from *Escherichia coli. Journal of Molecular Biology* **374**(4):1091-1103

Schneider-Poetsch, T., Ju, J., Eyler, D. E., Dang, Y., Bhat, S., Merrick, W. C., Green, R., Shen, B. and Liu, J. O. 2010. Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. *Nature Chemical Biology* **6**(3):209-217

Scholthof, K-B. G. 2007. The disease triangle: pathogens, the environment and society. *Nature Reviews Microbiology* **5**:152-156

Scholthof, K-B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C. and Foster, G. D. 2011. Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology* **12**(9):938-954

Simandan, T., Sun, J. and Dix, T. A. 1998. Oxidation of DNA bases, deoxyribonucleosides and homopolymers by peroxyl radicals. *Biochemical Journal* **335**:233-240

Singh, A., Taneja, J., Dasgupta, I. and Mukherjee, S. K. 2015. Development of plants resistant to tomato geminiviruses using artificial trans-acting small interfering RNA. *Molecular Plant Pathology* **16**(7):724-734

Stepchenkova, E. I., Kozmin, S. G., Alenin, V. V. and Pavlov, Y. I. 2009a. Genetic Control of Metabolism of Mutagenic Purine Base Analogs 6-Hydroxylaminopurine and 2-Amino-6-Hydroxylaminopurine in Yeast *Saccharomyces cerevisiae*. *Russian Journal of Genetics* **45**(4):409-414

Stepchenkova, E. I., Tarakhovskaya, E. R., Spitler, K., Frahm, C., Menezes, M. R., Simone, P. D., Kolar, C., Marky, L. A., Borgstahl, E. O. and Pavlov, Y. I. 2009b. Functional Study of the P32T ITPA Variant Assocaited with Drug Sensitivity in Humans. *Journal of Molecular Biology* **392**:602-613

Storey, H. H. 1936. Virus diseases of East African plants: VI, A progress report of studies of the diseases of cassava. *East African Agricultural Journal* **2**:34-39

Strange, R. N. and Scott, P. R. 2005. Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology* **43**:83-116

Takayama, S., Fujii, M., Kurosawa, A., Adachi, N. and Ayusawa, D. 2007. Overexpression of *HAM1* gene detoxifies 5-bromodeoxyuridine in the yeast *Saccharomyces cerevisiae*. *Current Genetics* **52**(5-6):203-211

Tatinemi, S., Qu, F., Li, R., Morris, T. J. and French, R. 2012. *Triticum mosaic poacevirus* enlists P1 rather than HC-Pro to supress RNA silencing-mediated host defense. *Virology* **433**(1):104-115

Taylor, N. J., Halsey, M., Gaitán-Solis, E., Anderson, P., Gichuki, S., Miano, D., Bua, A., Alicai, T. and Fauquet, C. M. 2012. The VIRCA Project. Virus resistant cassava for Africa. *GM Crops & Food. Biotechnology in Agriculture and the Food Chain* **3**(2):93-103

Taylor, N. J., Sekabira, H. A., Sibiko, K. W., Bua, A. and Lynam, J. K. 2016. Disease-resistant GM cassava in Uganda and Kenya during a pandemic. *Analyses: Africa's Future…Can Biosciences Contribute?* **B4FA**:76-85

Thresh, J. M. and Cooter, R. J. 2005. Strategies for controlling Cassava mosaic virus disease in Africa. *Plant Pathology* **54**(5):587-614

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. *Molecular Plant Pathology* **19**(5):1282-1294

Urcuqui-Inchima, S., Haenni, A-L. and Bernardi, F. 2001. Potyvirus proteins: a wealth of functions. *Virus Research* **74**(1-2):157-175

Valli, A., Martín-Hernández, A. M., López-Moya, J. J. and García, J. A. 2006. RNA Silencing Suppression by a Second Copy of the P1 Serine Protease of *Cucumber Vein Yellowing Ipomovirus*, a Member of the Family *Potyviridae* That Lack the Cysteine Protease HC-Pro. *Journal of Virology* **80**(20):10055-10063

Valli, A., López-Moya, J. J. and García, J. A. 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. *Journal of General Virology* **88**:1016-1028

Valli, A., Dujovny, G. and García, J. A. 2008. Protease activity, self interaction, and small interfering RNA binding of the silencing suppressor p1b from *Cucumber vein yellowing ipomovirus*. *Journal of Virology* **82**(2):974-986

Vanderschuren, H., Stupak, M., Fütterer, J., Gruissem, W. and Zhang, P. 2007. Engineering resistance to geminiviruses – review and perspectives. *Plant Biotechnology Journal* **5**:207-220

Vanderschuren, H., Alder, A., Zhang, P. and Gruissem, W. 2009. Dose-dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. *Plant Molecular Biology* **70**:265-272

Vanderschuren, H., Moreno, I., Anjanappa, R. B., Zainuddin, I. M. and Gruissem, W. 2012. Exploiting the Combination of Natural and Genetically Engineered Resistance to Cassava Mosaic and Cassava Brown Streak Viruses Impacting Cassava Production in Africa. *PLoS ONE* **7**(9):e45277

von Braun, J. 2010. Food insecurity, hunger and malnutrition: necessary policy and technology changes. *New Biotechnology* **27**(5):449-452

Wagaba, H., Beyene, G., Aleu, J., Odipio, J., Okao-Okuja, G., Deepika Chauhan, R., Munga, T., Obiero, H., Halsey, M. E., Ilyas, M., Raymond, P., Bua, A., Taylor, N. J., Miano, D. and Alicai, T. 2017. Field Level RNAi-Mediated Resistance to Cassava Brown Streak Disease across Multiple Cropping Cycles and Diverse East African Agro-Ecological Locations. *Frontiers in Plant Science* **7**:2060

Webster, C. G. and Adkins, S. 2012. Low genetic diversity of *Squash vein yellowing virus* in wild and cultivated cucurbits in the U.S. suggests a recent introduction. *Virus Research* **163**(2):520-527

Wei, T., Zhang, C., Hong, J., Xiong, R., Kasschau, K. D., Zhou, X., Carrington, J. C. and Wang, A. 2010. Formation of Complexes at Plasmodesmata for Potyvirus Intercellular Movement Is Mediated by the Viral Protein P3N-PIPO. *PLoS Pathogens* **6**(6):e1000962

Whiley, D. M. and Sloots, T. P. 2005. Sequence variation in primer targets affects the accuracy of viral quantitative PCR. *Journal of Clinical Virology* **34**(2):104-107

Wilson, M. C., Mutka, A. M., Hummel, A. W., Berry, J., Chauhan, R. D., Vijayaraghavan, A., Taylore, N. J., Voytas, D. F., Chitwood, D. H. and Bart, R. S. 2017. Gene expression atlas for the food security crop cassava. *New Phytologist* **213**:1632-1641

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. *Journal of General Virology* **91**(5):1365-1372

Xiao, J., Zhang, H., Xing, L., Xu, S., Liu, H., Chong, K. and Xu, Y. 2013. Requirement of histone acetyltransferases HAM1 and HAM1 for epigenetic modification of *FLC* in regulating flowering in *Arabidopsis*. *Journal of Plant Physiology* **170**(4):444-451

Yadav, J. S., Ogwok, E., Wagaba, H., Patil, B. L., Bagewadi, B., Alicai, T., Gaitan-Solis, E., Taylor, N. J. and Fauquet, C. M. 2011. RNAi-mediated resistance to *Cassava brown streak Uganda virus* in transgenic cassava. *Molecular Plant Pathology* **12**(7):677-687

Ye, F., Li, Y., Lin, Q. and Zhan, Y. 2017. Modeling of China's cassava-based bioethanol supply chain operation and coordination. *Energy* **120**:217-228

Zhou, Y., Liu, X., Engstrom, E. M., Nimchuk, Z. L., Pruneda-Paz, J. L., Tarr, P. T., Yan, A., Kay, S. A. and Meyerowitz, E. M. 2015. Control of plant stem cell function by conserved interacting transcriptional regulators. *Nature* **517**:377-380

APPENDIX

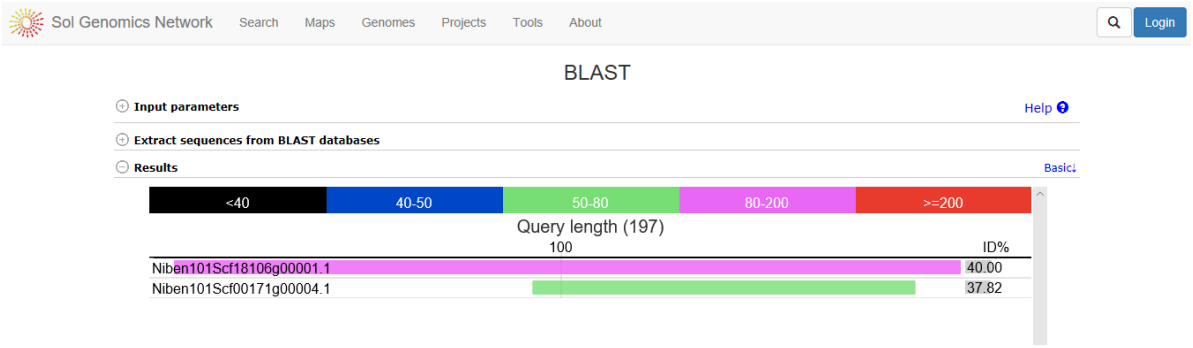

Figure 1. BLAST protein search of the yeast HAM1 revealed 40% sequence identity with the *N. benthamiana* protein, 'non-canonical purine NTP pyrophosphatase (ITPase-like)'.

>Niben101Scf18106g00001.1 sp|Q8TV07|NTPA_METKA ***- Non-canonical purine NTP pyrophosphatase IPR029001 (Inosine triphosphate pyrophosphatase-like) GO:0009143 (nucleoside triphosphate catabolic process), GO:0016787 (hydrolase activity), GO:0047429 (nucleoside-triphosphate diphosphatase activity)

MAAAARTVGSGVVLPRPVTFVTGNAKKLEEVRAILGQSIPFQSLKLDLPELQGEPEDISK
EKARIAAKEVNGPVLVEDTCLCFNALKGLPGTQHLLLGVVLGTVKVYYRKCFPCKWFLQK
IGHEGLNNLLLAYEDKTAYAMCVFSLALGPNSEPLTFVGKTLGRIVPARGPNDFGWDPIF
QPHGYDQTYAEMPKEGKNKISHRGKALELVKSHFAEARYTFQTDTTA

Figure 2. The amino acid sequence of the *N. benthamiana* HAM1.

>Niben101Scf18106 Niben101Scf18106:2833..9819 (- strand) class=mRNA length=6987

CAAAATTCAATTAGGAACAATGAAAATTTATCGTGGCTTCTATCAAATAAGATTAATTTGTTTAAGTTGAGAGT
GGGAACTTATTTTTAATTGAGTTGCGCACTGTAAGCTTTCACTCCTAAAACCCAAAAACTATAAATCTCCTCCCG
TTGAACGTGCAGATTTTCCGGTGATATAAAGGACAATGGCGGCTGCGGCGAGGACAGTGGGATCAGGGGTG
GTGTTGCCGCGTCCGGTGACGTTTGTTACCGGAAACGCTAAGAAGTTGGAGGAAGTTAGGGCCATTCTTGGC
CAGTCCATTCCCTTTCAGTCCCTTAAGCTTGATTGTAAGTCCACAATTTTTTTTATTTTTTAAAAGTATTTAACTT
GTGTTTAATTTCTATCATTTAAAAAAATAAATAAAATTAACCCCTTTATGTCAAAATTTTATTTTTAGTGCCAGA
ACTTCAAGGTGAGCCTGAAGATATCTCCAAAGAAAAAGCAAGAATTGCTGCCAAAGAGGTAATAAACCCTGC
CTTTATTGCACCACATATGTATAATTGTCATTTTGTTAGTTTTCTTTGTCCATTTATGGGATAAAGTTTTAAATTT
AATGTTTATGTACTTTTATATTATGATTTTAGCTGAGTTAGAAAAATTGCTTAGGTTAATGGACCAGTGTTAGTT

GAGGATACCTGTCTATGCTTCAACGCTCTTAAGGGTCTCCCAGGTACTCAGCATTTGCGTACCTCAATATTTGT
ATATGTGCTTTTTAGAGTTGATAAAAAAATCATAATGAACTAAGGTGCTGAAATTGTTTTTGCAGGACCTTACA
TGTAAGTTCATCTCCTCTTTTATCATTCTGCTCGAGTTAAGTGTGTTCTTCTTTATGTGCTTCCTATATTATATTTG
TGTGATTTTTGGTGATGCTTTCCTGGTAATATCGTATTAAGCAATTCTGCTAATTGCAATTTAGAACAATGAGG
AGGTGTGAGAGGCTGGTCTTGGAGGGCCCAAGGAGAGGTAGAGGTAGGCCAAAAAAGTATTAGGAAGAGG
TGATTTGGCAAGATATGGCGATGCTCCAGCTTACTGAGGACATAACCTTAGATAGGAAGGTGTGGAGGTCGA
GAATTAAGGTAGAGGGTTAGCTAGCCTAGCGTTGTCTTTCCGTGTTCGTGTAGTAGTGCATCACTTGGTGTTA
GTCTTGCGTTTTCTTATTCTTTAGATTTCTGTTATTAATTGTTTTTTCACATCGCTTTCTGATTGCACCACTAAGG
GTTAGTCTTGTATTATCACTTCGGTTTTCTGATTGGTTGCTTGTTGTTGCTTCATTTTATCTTTCTTGAGCTGAGG
GTCTATTAGAAACAATCTCTCTGCCTTCCTATCGGTAGGGATAAGGTCTGCGTACACGGTACCCTCCCCAAACC
CCACTTGTGGAATTACACTGAGTTTTTGTTGTTGTTGTTGTTGTTGTTGCAATTTAGAACAAAGTTATAAACTAC
AGTGGCATTTCCCTAAGATGAAGTCTATAGTTCTCGATTTGTTTATGAAATATAGGTGAATTAGACTAATAAAT
TTTACTGAGTGTGAAATATTGGGTTTATTTAAGGTGATAACTTGTTTTGCTCAATATGCATACTTTCTTGAGGCA
TGTAGTTTTTGAAATTTAATGTTGTCAGTATTAGGAGTTGTGCTTGGCACGGTTAAAGTATATTACAGAAAATG
TTTTCCTTGGTAATAGCATTTTCTACTGTTTGGATGTTGTTAAGATGTGGATAATGTTTTTGCGAAGGAAAACA
TGTTTTCACCAAATGGGGGTAATTGACTTACTTTGATGCACGTGAATCATTTTTCTTTACAATTGTTAACTCATC
CACTTTCTTGCTTCAGTCAATGTATAACAGTTTCCATCGTTATACTCAATATTATCACAAACCACCTCCCGGATG
TCCCAATCCCACATTATACTATTGTGATGTGTTGTAAATTGTTTCCTGAAACAATATTTCCATGTTCACTCAAAA
AACTATTTTGTAGATCATGTTAGAGTATTGGGTGATAAGATAGTAACTTATATCATAAGTAAAATTTAGGGGAT
GCGGCACAAGGTTTCAACTTGTCAGAATGAAAACTTTATAGGAATAGGCAATATTCAATAAAGTTTTCATGTC
AAAAAAGAGCATAACGATGTGCTGATGAACCACATTTTGGGAACTTATGCATTCCCTTTGACTTTGAGAGAAT
TTCTACTCTCACAGGGAAGTTCTCTTTTTTCATGTTTGGCGCTATTTATATATTTTTGTACTAATTGATCAACTAG
CTGTATCTGCAACTTTATCTATTTCTTTTCCTTTCTTGTGGGGTGAGTTTTCGGGGTGGTGTGGGGAAGTGGAA
CTAATGTTGATGGAGGAAGTGTATTAAGAACTTTCAGTTGCTGGGAATCTCTCGGCTTATTCTTACTCATTGGA
GCTTTTTGCAAGTTGCAAGATCACCCTTTAGATCAGGGAGTAAAGACAGAACATCTAATTTCCTTTGTAGAACC
TCTAGAAAAGTTGAACATGCAGCTTAAGCTTGAATTTGTTTCTTTTTTTCCCTTTCTTTATTCGGAAAAGGGTCA
GAACTGCCCCTAATGTATTGGCGCCAGGTTAAAAATGCCCTCCGTTAACCTATTTTGAAGAAAATGCCCCTACC
GTTAAAATTCTGGCATATTTGTGCCCCTAACGTATTGACTCCAGTTTAAAAATGCCCCTACCGTTAACCTATTTT
GAAAAAAATGCCCCTACCGTTTTAATTATAAGTACAGTCAGTTGTTAGACTTAGGGGCACAAATATGCCACATT
TTTAACGGTAAGGGCATTTTCTTCAAAATAGGTTAACGGAGGGCATTTTTAACCTGGATTCAATACGTTAGGG
GCAGTTCTGATCCTTTTCCCTTCTTTATTTGATGAAGGCAGTCTAAGCTAGAAAAATTGTATCCTTTTGTATGAA
GTAGGTTCTTAGAGATATTCTTGGACAATGGTGTTGTTGGGACAGCTTGTGCGCACGTTGAATATTCCACCAT
ATACCAGCTCCTTGAAATAGGTTCCTAGATATTATTGCTCCAATGTAAAGGATTTTCTTTCAGTAAGAGAATCCT
CTTGCTTGATACAAGAATCTTATTCATCATGCTTCATGTTACCGGTGATTGGTCCATGGAGCTAGAGAGGACAT

GACATTTTCATATTAGATGACCATAGAAACACACATCTGCTTCAATAATCCCTTAACAAAGTGAAAGGAGAAAT
ATTGTAAACTAAAGTTACCAAATAGTCATGTGTCTTGTTCCCACCTTTGAAGAGAGTGGTGGTTGGGTGAACTT
AAACTACACACTTGCGAAGAGGAAAATATAGTTGTTACAGTTGCTAGTCAGTGGCATGAAAATTTGGGAGAG
GGTGGTAGAATTGAGGGTGAGGAGGGCAGTATTTAGACAACCAGTTCGGGCTCATGGCGGGGCGTCCTACC
ACGGAAGCTTCCACCTTATTCGGAGGTTGGTGGAATAATATAGGGAAAGAAAGAAGGAGAGGTTGTGAGTTC
GAGTCTGCCCAAGAGCAAGGTGGGAAGTTCTTGGAGGGAAGGATGCCGGGGGTCTATTTGGAAACAGCCTC
TCTACCCAAGGGTAGGGGTAAGGTCTGCGTACATACTACCCTCCCGAGACCCCACTAAGTGAGATTATACTGG
GTTGTTGTTGTTGTTGTTGAAGAAGGATTTGCACATGGTGTTAATTGATTTGGAGAAAGCCTATGACAAGGTT
CCTAGAGAGGTTCTATGGGAGTGCCTGAAGGCATAAAGTGTGTCGGTTGCTTACATTAGGGTGATTAAGGAC
ATGTATGATGGAGCTAAGACTCTGGTTTGGACAGTGGGAGGCGACTCGGAGCATTTACCAGTTGTTATGGGG
TTACACCAAGGGTCTGCGCTCAGACCATTTCTTTTTGCCCTGGTAATGGATGCACTGACACACCATATTCAAGG
AGATGTGCCATGGTGTATGTTATTTGCGGATGACATAGTGCTGATTGATGAGACGCGTGGCGTTGTCAATGA
GTGGCTAGAGGTTTGGGGACAGACTCTTGAGTCTAAGGGTTTCAAGTTGAGTAGGACTAAGACGGAGTACTT
AGAGTGCAAGTTCAAGTGTTGAGACAACGGAAACGGGATTGGACGTGAGACTTGATTCTCAAGTCATCCCTA
AGAAGGTCAGTTTCAAGTACCTTGGGTCGGTTATTCAGGGGGATGGGGAGATCGATAAGGATGTAACACATC
GTATAAGAGTTGGGTGCATGAAGTGGAGGTTAGCGTCTGGAGTCCTGTGTGACAAGAAGGTTAGTTTCAAGT
ACCTTGGGTCGGTTATTCAGGGGGATGGGGAGATCGATAAGGATGTAACACATCGTATTGGAGTTGGGTGGA
TGAAGTGGAGGTTAGCGTCTGGAGTCCTATGTGACAAGAAAGTGCCTCAGATTCTCAAAGGTAAATGTTATA
GAGCGGTGGTTAGATCGGCCATGTTGTATGGGGCTGAGTGCTGGCTAGTCAAAAACTCTCATATCCAAAAGA
TGAAAGTAGCTGAAATGAGGATGTTGAGGTGGATGTGCGGGCACACCAAGATAGATTAGATTAGGAATAAA
GTTGTTTAGGAGAAGGTGGGTGTCGCCCTCATGGATGACAAGTTGAGGGAAGCAAGACTCAGATGGTTTGG
GCACGTCCAGAGGAGAAGCTCATATGCGCCGGTGAGGAGGTATGAGCGGCTGGCTTTGGTGGGCATGAGGA
GAGGTAGAGGGCGGCAAGTATTGCGGAGAGGTGATCAGACAGGACATGGTGAGGCTTCAGATTATTGAGGA
CATGGCCCTAGACAGGAAGTTGTGGAGATCGGGCATCAAGGTTGTAGGTTAGGTGGTAGTTGAGGATATCTC
TATGGCGCCGCAGAGCGAGGCTAGTCTGTTAGGACTTGATCTTTGACTGTTAGTGTTTTATGTTGAGTTCACAC
TACTTTTTCATTCCTCTTTGATGTGCTTTATCTATTGGTTATTGTTTTGCTTTTCACTTATTTTCTGGTTGTGGTGA
TGTTGTTCTTTGTTCTATGGTTTCTATTGAAGGCACTGATAGATTGCCTCTTTTTGTCTTCTTTTGTCTTCTTGAG
CTGAGGGTCTTTCGGTAACCCTTTCTACCCCCGCGGGGTAGGGGTAAGGTCTGCGTACACACTACCCTCCCCA
GACCCCACTAGTAGAATTTTACTGGGTCGTTATTGTTGTTGTAGTTGCTAGTCAGTGGACTTTGCAGGATCTTT
TATTAACACACAAAAGAAGTCCCAACCTTAAGTTAATACCTTTAACCAGTCAAAATATGACGCTAAAATGGATA
GAGTATTGGAACTTAACCGATTGAACTTGTTAGAAATTTGGAATTATGTCCAAAGATTTGGTTTACTTTCTTGA
ATCATACGCAATCTTTAATTTTAATAGAATGTTCTGTAATTCACATATCGTGGACATGCCAAATAGTTTTTTTGT
AATTCTTGTATATGCTGTTATTTTCCAATGTTGCTAACTATTATCTTTTGCCTTTAACATAATATACTTTGTAATTC
TTGCTTTATGCTGGGGATTGCACCAATTTACGTTTTGCAGCAAGTGGTTTCTGCAAAAGATTGGTCATGAAGGT

AATCGCTTAGCTTGCTTGTTCTCATGTATCACTCCCTTGAAATTTATGTCCTCCATTTCTGTGAAATCACTCTGTT
TGGTTTATTTTATTTTTTAGAAAAAAATTTCTGACATCTTATAGGCCTCAACAACTTACTCTTGGCTTATGAGGA
TAAAACAGCATATGCCATGTGTGTGTTTTCTCTAGCTCTTGGGCCAAATTCAGAACCATTGACTTTTGTTGGAA
AAACACTGGTAAGTGTTTGCTTCTGTAGTCCATGAATATTTCCTTTAGACTTTGTATTTCCATTAATCTTTCTGCT
GATGAGTCTTATGTTTCAAACTAGAATGATCCTCGGTGGGTATTGCGTGAATATTTTGCTGAAGTAATACCAAG
ACTGCTGGTTAATATTTTTTTCAAAATTAAGTACCAGAATGCTGGTTATTTGTGTATTTATATGAAAAGAAAAT
GGAAAAATAGTTCCTATTTGGCTTCTTACGATGTTACTTCTTATTTATTACTCGTATCTACTATTTAGGCAAAAT
AACTCGCCCATTCTTACTAAGAATTGAAAGAAGACTCAGAAACAGAAGAAAGGATGGAAAGTTCGTTTAACAT
ACTTGTAAAGATAATCAGTGTGGATACTATCCTAACATGAAGAACTTGCGAAACTTGATAGATTTTCTTTTATT
ATTTCTTTATTGAACTATGCAAACAATAGCACTAACATTGCGAACATGGCTCTTCAGGGCAGGATAGTTCCAGC
GAGGGGACCCAATGATTTCGGATGGGATCCAATATTTCAACCTCATGGCTATGACCAGACGTAAATCTCCTGT
CCTCTCAATATTTGCCTATCGAACCTCTTGACATGAAGAGTACTAACTGATAGAACTGTTCTTGCAGTTATGCTG
AGATGCCTAAGGAAGGAAAGAACAAAATTTCTCATCGGGGTAAGGCTCTTGAACTAGTGAAGTCACATTTTGC
TGAGGCTCGATACACTTTCCAGACGGACACAACTGCCTAAGACTAAGGTGTATGCTCTGTAGCGACATAATTT
TCATGTTTTTGCTCAAGAACTACAATTGATGACAGTGTAAGCTTCTTTTATTCCCCCAGATAAAATAGGTTGGA
GGATGAGATATATGAAATATGAGGACTTGCTTTGTTGTTATAGAGCCTAGATCATAGATGAGACACAAAATAT
TCTATTGTGAATGTACTATATGTATTTTCGGATTAGGAGCTTATTCTCCTAAAGATTCCAATGTCACTTATCTTAT
GATAAGTTGAAGTTCCACCAACAGCACTGATTTTTTGCTTATTTTGCATATTCTTAGC

Figure 3. Genomic DNA sequence of the *N. benthamiana HAM1* with exons highlighted and the stop codon underlined and highlighted in red.