Doctoral Thesis

Molecular and transcriptomic characterization of natural resistance to cassava brown streak viruses in cassava

Author(s):
Anjanappa, Ravi B.

Publication Date:
2015

Permanent Link:
https://doi.org/10.3929/ethz-a-010572922

Rights / License:
In Copyright - Non-Commercial Use Permitted
Molecular and transcriptomic characterization of natural resistance to cassava brown streak viruses in cassava
(Manihot esculenta, Crantz)

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

Presented by

Ravi Bodampalli Anjanappa

Master of Science in Plant Biotechnology, University of Agricultural Sciences, Bangalore

Born 12th August 1979
Citizen of India

Accepted on the recommendation of

Prof. Dr. Wilhelm Gruissem, examiner
Prof. Dr. Hervé Vanderschuren, co-examiner
Dr. Maruthi M. N. Gowda, co-examiner

2015
# Table of Contents

Abstract 1

Zusammenfassung 3

Abbreviations 5

General Introduction 7

Chapter 1: Natural Resistance to Plant Viruses 13

Chapter 2: Characterization of brown streak virus-resistant cassava 55

Chapter 3: Transcriptome modulation in susceptible and resistant cassava varieties inoculated with cassava brown streak viruses 82

Chapter 4: Identification and characterization of a resistance-breaking Cassava brown streak virus isolate 131

Chapter 5: Discussion and perspectives 159

Acknowledgements 169

Curriculum Vitae 170
Abstract

Cassava (Manihot esculenta Crantz) production in eastern and central African countries is adversely impacted by cassava brown streak disease (CBSD), thereby threatening food security. CBSD is caused by two viral species, Cassava brown streak virus (CBSV) and Ugandan Cassava brown streak virus (UCBSV) which are collectively termed as cassava brown streak viruses (CBSVs). Both are (+) single stranded (ss) RNA viruses of the genus Ipomovirus; family Potyviridae. The prevalence of mixed infections of the two ipomoviruses CBSV and UCBSV along with cassava mosaic geminiviruses (CMGs) that cause cassava mosaic disease (CMD) is now well known. In the present study, 14 elite cassava cultivars (selected from the cassava germplasm and breeding programmes) were assayed using top cleft grafting procedures for identifying novel sources of resistance against the multiple CBSV and UCBSV isolates. These infection assays identified cassava varieties with contrasting levels of virus resistance. KBH 2006/18 and KBH 2006/26, two elite breeding lines from the IITA breeding programme were identified as highly resistant as they remained symptom-free and did not support replication for both CBSV and UCBSV isolates. Importantly, challenging with a mixed infection of CBSVs and East African cassava mosaic virus (EACMV-Ug) could not break the virus resistance. Independent grafting experiment using scions resistant KBH 2006/18 and KBH 2006/26, moderately resistant TMS 30001 and susceptible 60444 varieties was performed over a time course (16, 22 and 28 dag) to identify early molecular interactions. As expected, these assays revealed variable CSV titers across the different time points. Comparative transcriptome analysis of resistant KBH 2006/18 and susceptible 60444 varieties were performed using Illumina Hiseq2000 from both CBSV-infected and virus-free scions collected at 28 dag. The RNA-seq analysis revealed that an average of 83.5% of the total reads mapped to phytozome cassava reference genome v9 and identified 26,206 genome transcripts. A total of 853 (60.3% of them up-regulated) and 334 (35.6% of them up-regulated) differentially expressed genes (DEGs) were identified in susceptible 60444 and resistant KBH 2006/18 varieties, respectively. Further functional characterization of the DEGs using the MapMan software to identify processes and pathways, revealed 27 and 24 bins MapMan bins or functional classes in susceptible and resistant varieties, respectively. In variety 60444, signaling (MapMan bin 30) was modulated with majority of the DEGs (85 %) being up-regulated while in contrast KBH 2006/18 only 27% of the DEGs were modulated in signaling pathway. In KBH 2006/18, 90% of DEGs related to stress (MapMan bin 20) were up-regulated while 70% of DEGs were up-regulated in susceptible variety. In both the varieties DEGs implicated in photosystem were down-regulated (100%). Further, de novo assembly of unmapped reads determined novel cassava (host) genes, in addition to complete genome sequences of CBSV (TAZ:DES:01) (8,975 nt) and partial sequence of TAZ:DES:02 (6,645 nt). Moreover, analysis
of host genes using RT-qPCR, revealed that four DEGs (*RDRP1, NTL9, NIMIN2* and *BG3*) where up-regulated in early time points samples (16 and 22 dag) in susceptible 60444 variety. Further, expression analysis of these genes in moderately resistant TMS 30001 variety also confirmed consistent up-regulation in presence of the virus.

CBSV-like symptoms were first reported in resistant variety KBH 2006/18 in a field trail in the Chambezi province in Tanzania in 2012. Field-collected material was multiplied in the greenhouse and used to inoculate KBH 2006/18 scions. CBSD was graft-transmissible to healthy KBH 2006/18, which developed typical CBSD symptoms. We used conserved sequences between CBSV and UCBSV genome sequences to develop primers and amplify fragments of the isolates infecting KBH 2006/18. We detected only one CBSV isolate in the field-collected material that we named CBSV TAZ: Cham:14. Comparing the resistant breaking TAZ:Cham:14 isolate to other isolates (TAZ:DES:01, TAZ:DES:02, and UG:Kab4-3:07) revealed specific mutations in the proteins from the TAZ:Cham:14 isolate with the exception of the 6K1 protein which was identical between TAZ:DES:01 and TAZ:Cham:14. As expected, the sequence diversity was greater between the two CBSVs species (~74.5% identical) than within CBSV isolates (88% identical). This is the first report on CBSV isolates differing in virulence and further efforts are being made to develop infectious clones and to identify molecular virulence determinants of TAZ:Cham:14 in the KBH 2006/18 line.
Zusammenfassung

hochreguliert waren. In beiden Sorten waren die ins Photosystem involvierten DEGs runterreguliert (100%). Des Weiteren konnten durch die de novo Assemblierung von nicht zugeordneten Fragmenten, nebst der kompletten Genomsequenz von CBSV (TAZ:DES:01) (8975 nt) und der partiellen Sequenz von TAZ:DES:02 (6645 nt), auch neue Maniok Gene bestimmt werden. Darüber hinaus ergab die Analyse der Wirts-Gene mittels RT-qPCR, dass vier DEGs (RDRP1, NTL9, NIMIN2 und BG3) an frühen Zeitpunkten (16 und 22 TNP) in 60444 hochreguliert sind. Expressionsanalysen dieser Gene in der mässig resistenten Sorte TMS 30001 bestätigten eine konsistente Hochregulation über alle drei Zeitpunkte.

Abbreviations

6K1  6K1 protein
6K2  6K2 protein
AA   Amino acid
ACMV African cassava mosaic virus
Avr  Avirulence
BG3  β 1,3 Glucanase
BLAST Basic local alignment search tool
bp   Base pairs
CBSD Cassava brown streak disease
CBSVs Cassava brown streak viruses
CBSV Cassava brown streak virus
CMD  Cassava mosaic disease
CC   Coiled-coil
cDNA Complementary deoxynucleic acid
CI   Cylindrical inclusion
CMGs Cassava mosaic geminiviruses
CP   Coat protein
CTAB Cetyltrimethylammonium bromide
dag  Days after grafting
DEGs Differentially expressed genes
EACMV-Ug East-African cassava mosaic virus-Uganda
EDTA Ethylenediaminetetraacetic acid
eIF4E Eukaryotic initiation factor 4E
EMS  Ethyl methanesulfonate
FAOSTAT Food and Agriculture Organisation Statistics
HAM1h HAM1 homolog
HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ICTV International Committee on Taxonomy of Viruses
kb   Kilobase
kDa  Kilodalton(s)
MAMPs Microbe-associated molecular patterns
MgCl2 Magnesium chloride
NaCl Sodium chloride
NBS-LRRs Nucleotide-binding site leucine-rich repeat
NCBI National Center for Biotechnology Information
Ni  Nuclear inclusion a
Nlb  Nuclear inclusion b
nt   Nucleotide
NTL9 NAC transcription like-9
ORF  Open reading frame
P1   First Protein
P3   Third Protein
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAMPs</td>
<td>Pathogen-associated molecular patterns</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Pd</td>
<td>Plasmodesmata</td>
</tr>
<tr>
<td>PDLP</td>
<td>Plasmodesmata-located proteins</td>
</tr>
<tr>
<td>PME</td>
<td>Pectin methylesterase</td>
</tr>
<tr>
<td>PR</td>
<td>Pathogenesis-related proteins</td>
</tr>
<tr>
<td>PRRs</td>
<td>Pattern recognition receptors</td>
</tr>
<tr>
<td>PTGS</td>
<td>Post-transcriptional gene silencing</td>
</tr>
<tr>
<td>PTI</td>
<td>PAMP-Triggered Immunity</td>
</tr>
<tr>
<td>RDRP</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>RLKs</td>
<td>Receptor-like kinases</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse transcription-quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RYMV</td>
<td>Rice yellow mottle virus</td>
</tr>
<tr>
<td>SOD</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>TIFs</td>
<td>Translation initiation factors</td>
</tr>
<tr>
<td>TIR</td>
<td>Toll and Interleukin-1 Receptor</td>
</tr>
<tr>
<td>TME</td>
<td>Tropical <em>Manihot esculenta</em></td>
</tr>
<tr>
<td>TMS</td>
<td>Tropical Manihot selection</td>
</tr>
<tr>
<td>TMV</td>
<td><em>Tobacco mosaic virus</em></td>
</tr>
<tr>
<td>TOM</td>
<td>Tobamovirus multiplication</td>
</tr>
<tr>
<td>ToMV</td>
<td><em>Tomato mosaic virus</em></td>
</tr>
<tr>
<td>TYLCV</td>
<td><em>Tomato yellow leaf curl virus</em></td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>Vpg</td>
<td>Viral genome-linked protein</td>
</tr>
<tr>
<td>wag</td>
<td>Weeks after grafting</td>
</tr>
<tr>
<td>YSL3</td>
<td>Yellow stripe like 3</td>
</tr>
</tbody>
</table>
General Introduction

Importance of cassava

Cassava (*Manihot esculenta* Crantz), a member of the *Euphorbiaceae* family is a semi-perennial woody shrub that produces edible starchy tuberous roots. It serves as an important food crop for around a billion people in 105 countries where the root provides as much as one third of daily calories and also has many industrial applications (Henry, 2000; FAO, 2008). It is the 3rd most important crop after rice and maize, with estimated production of 276 million tons in 2013 (FAOSTAT, 2013). Cassava is sometimes considered a ‘food security’ crop, especially in Africa, because it produces stable yields under drought-prone conditions, low fertile soils and less labor inputs. In addition to these good agronomic traits, its in-ground storability of cassava roots allows its progressive harvest (Burns *et al.*, 2010; Fermont *et al.*, 2010).

Cassava brown streak disease

Cassava production is constrained by several pests and diseases. In Africa, viral diseases are amongst the major factors limiting cassava production. Of the 20 known viral diseases affecting cassava, cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are the most widespread and severe diseases on the African continent (Patil and Fauquet, 2009; Legg *et al.*, 2011). CBSD was first reported in eastern Africa during the 1930’s (Storey, 1936). CBSD was initially confined to low attitudes (i.e. <1000 meters above mean sea level (masl)) in coastal Africa (Tanzania) but has been recently reported at higher altitudes (>1000 masl) (Alicai *et al.*, 2007). CBSD causes losses up to 70% in cassava root yield (Hillocks *et al.*, 2001). CBSD symptoms may vary depending on the variety, age of the plant, virus isolate and the prevailing environmental conditions (Mohammed *et al.*, 2012). Typical CBSD symptoms are leaf chlorosis, brown streaks on stems and dry hard rot in roots. In some varieties only leaf symptoms are exhibited without any root symptoms and vice versa, while sensitive varieties exhibit foliar, stem and root symptoms (Hillocks and Jennings, 2003). CBSD spreads mainly by propagation of infected stem cuttings. Of late, the CBSD is being reported from new geographical regions (Bigirimana *et al.*, 2011; Mulimbi *et al.*, 2012) clearly indicate the need to control the disease.

Cassava brown streak viruses

The viral nature of CBSD was first characterized in the 50’s (Lister, 1959). Recent molecular studies have revealed at least two distinct species, *Cassava brown streak virus* (CBSV) and *Ugandan Cassava brown streak virus* (UCBSV), collectively termed CBSVs, are causing
CBSD. CBSVs are positive-sense single-stranded RNA viruses taxonomically grouped into the genus *Ipomovirus*, family *Potyviridae* (Mbanzibwa et al., 2009; Winter et al., 2010). CBSV has a genome size of ~9 kb with absences of potyviruses -HC-Pro protein but has a HAM1h-like protein and a single P1 protein (serine proteinase) that performs silencing suppressor activity (Mbanzibwa et al., 2009). The CBSV genome encodes a large polyprotein from a single open reading frame, which is cleaved into 10 viral proteins (Figure 1) (Mbanzibwa et al., 2009; Winter et al., 2010). The viruses are transmitted between plants by whitefly *Bemisia tabaci*, though the rate of CBSV transmission by whiteflies is low (25 - 40.7%) (Maruthi et al., 2005; Mware et al., 2009). The recent reports on the evolutionary nature of CBSV and UCBSV along with the presence of virus in *Manihot glaziovii*, a wild species of cassava (Mbanzibwa et al., 2009; 2011) clearly suggests the occurrence of severe forms of CBSVs might lead to another disease outbreak under favorable conditions.

**Figure 1:** Schematic representation of cassava brown streak virus genome (adapted from Mbanzibwa et al., 2009). P1, a serine proteinase (the first protein); P3, the third protein; 6K1 and 6K2, the two 6-kDa proteins; CI, cylindrical inclusion protein; VPg, viral genome-linked protein; N1a-Pro, viral proteinase; N1b, replicase; HAM1h, HAM1 homolog; CP, coat protein and 5’ and 3’ UTR; untranslated regions.

**Availability of CBSD resistance in cassava germplasm**

Screening for CBSD resistant genotypes started at Amani research station in northern Tanganyika (now Tanzania) as early as 1937 (Nichols, 1947; Jennings, 1957) and since then continued efforts are being made to identify and/or develop varieties resistant to CBSD (Childs, 1957; Abubaker, 1989). Early breeding efforts to develop CBSD resistant cassava varieties were based on interspecific crosses between *Manihot esculenta* and *Manihot glaziovii* or *Manihot melanobasis* (Hillocks, 2003; Hillocks and Jennings, 2003). The term resistance, refers to range of plant reactions to virus infection and is often used vaguely by plant breeders, virologists, physiologists or plant pathologists (Buddenhagen and Toenniessen, 1981). Based on plant’s ability to limit virus multiplication, the terms extreme resistance (complete failure of the virus to establish) -tolerance (reduced damage in presence of a virus) susceptibility (plants do not impair virus infection) are described (Lecoq et al., 2004; Fraile and García-Arenal, 2010). Recent studies on screening CBSD resistance from selected cultivars indicate that the lines do not possess stable and robust resistance to CBSVs (Kaweessi et al., 2014; Pariyo et al., 2015). One of the major limitations and challenges in breeding for disease resistance is the non-availability of robust resistant sources and also lack of efficient and reliable phenotyping in cassava (Yadav et al., 2011; Ceballos et al., 2012). Development of CBSD under field
conditions are influenced by inoculum source and whitefly populations and usually requires six-nine months to develop 100% CBSD incidence (Kaweesi et al., 2014; Gwandu et al., 2015). Therefore it is essential to identify and characterize CBSD resistance in the cassava germplasm to provide new means to control the disease in the field. Presence of mixed infections of CBSVs, its spread to new geographical regions and co-occurrence with cassava mosaic geminiviruses (CMGs) are further threatening cassava production (Alicai et al., 2007; Legg et al., 2011; Mbanzibwa et al., 2011; Abarshi et al., 2012) and pose additional new challenges. Hence the identified varieties should possess robust resistance to CBSD as well as to CMD, so that they could be immediately used as an efficient and easiest control measure. Other control measures include planting early maturing cultivars and use of tolerant varieties along with use of disease free stem cutting for propagation, reducing the virus inoculum sources by removal of infected plants from the field, reducing the whitefly population by timely spraying insecticides can reduces the disease incidences (Gwandu et al., 2015). Studying molecular responses at early stage of viral infection would reveal host factors that could potentially be considered as molecular markers for virus resistance.

**Aims of the study**

1. To identify new sources of resistance/tolerance to CBSVs in cassava.

2. To study transcriptome modulation in cassava varieties contrasting for CBSV resistance in order to characterize host response to CBSV infection and to identify potential markers for virus resistance.

3. To characterize CBSV isolates breaking natural CBSV resistance identify in Aim 1.
References


Natural resistance to plant viruses

Ravi B. Anjanappa¹, M. N. Maruthi², Wilhelm Gruissem¹ and

Hervé Vanderschuren¹,*

¹ Department of Biology, Plant Biotechnology, ETH Zurich, Zurich, Switzerland

² Natural Resources Institute (NRI), University of Greenwich, Chatham Maritime, Kent ME4 4TB, United Kingdom

* Corresponding author

Hervé Vanderschuren

Email: hvanderschuren@ethz.ch

Phone: Office: +41 44 632 87 25

Author contributions:

Review written: Ravi B. Anjanappa

Edits and corrections: Hervé Vanderschuren, M. N. Maruthi and Wilhelm Gruissem
Chapter 1
Natural resistance to plant viruses

Abstract
Natural resistance to plant viruses is the most economical, durable and an efficient option for controlling and limiting the devastating damages caused by viral diseases. Over the years, considerable efforts have been made to understand the mode of native innate immunity operating within the plant to restrict virus establishment. Understanding the molecular responses and the host factors involved during virus infection would point to potential molecular markers that could be exploited in resistant breeding programs geared towards development of durable broad spectrum resistance. Future identification and deployment of resistant cultivars should also account for host-virus-environmental interactions.

Introduction
Plant diseases cause major economic losses with devastating consequences on the livelihoods of large populations around the world. Well-known examples include the 1840’s potato blight caused by Phytophthora infestans (Nowicki et al., 2011), brown spot fungus disease in rice that led to the Great Bengal Famine of 1943 (Padmanabhan, 1973) and the southern corn leaf blight epidemic of 1970–1971 in the USA (Ullstrup, 1972). Global distribution of pathogens is a dynamic processes involving host availability (susceptibility and abundance), suitability of climatic conditions and evolutionary changes either in host and/or in the pathogen (Shaw and Osborne, 2011).

About 900 species of plant viruses are listed by the International Committee for the Taxonomy of Viruses (King et al., 2012). Most of the studied viruses are from cultivated crops though recently efforts are being made to study viruses present in wild species (Sserubombwe et al., 2008; Tugume et al., 2013; Wylie et al., 2013). Plant viruses are non-cellular or obligate intracellular parasites with approximately 80% of them depending on insect vectors (aphids, thrips, whiteflies, leafhoppers, planthoppers, treehoppers, mites, nematodes and also fungi) for their transmission (horizontal) between plants (Hohn, 2007). Plant viruses are considered to be major primary pathogens causing severe losses in staple food crops like cereals, grain legumes, root and tuber crops (Table 1) (Hull, 2014). The most suitable approach to control viral diseases is through deployment of natural resistant varieties as often other control measures are limited or ineffective and (Hanssen et al., 2010). Natural plant resistance to viruses comes from large a number of traits that are broadly grouped into ‘non-host resistance’ and ‘host resistance’, that
restrict virus entry, translation, replication and movement of viruses (Kobayashi et al., 2014). Non-host resistance is comprised of mostly physical barriers against most invading plant pathogens (Mysore and Ryu, 2004), while host resistance is based on direct or indirect interactions between pathogen effectors and the corresponding plant resistant (R)-gene products (Flor, 1971; Hammond-Kosack and Jones, 1997). During the process of natural co-evolution, plant-virus interactions exert selective pressure on each other to evolve and adapt to newer characters that assist in overcoming the limitation imposed on each other (Rausher, 2001). Many viral variants are generated by natural mutations, errors during viral genome replication, natural recombination and re-assortment mechanisms (Roossinck, 1997; Drake and Holland, 1999; Bonnet et al., 2005; Elena et al., 2008). Changes in the genetic diversity of viral populations to overcome plant resistance represents an important route/pathway in the emergence of new pathogenicity factors for viruses (Elena et al., 2014). The rate of spontaneous mutation per genome is variable among microorganisms, favorable mutations are very rare, while most mutations are deleterious (Drake et al., 1998). Hence the use of varieties with resistant traits that confers broad spectrum resistance to a large number of strains still remains the most efficient and sustainable means of controlling plant viral disease outbreaks.
Table 1: Major plant viral diseases of five major staple food crops.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Disease</th>
<th>Virus</th>
<th>Virus genome</th>
<th>Transmitted by</th>
<th>Loss %</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize (Zea mays)</td>
<td>Maize Streak disease</td>
<td>Maize streak virus</td>
<td>Mastrevirus (ssDNA)</td>
<td>Leafhoppers</td>
<td>100</td>
<td>(Shepherd et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>Mosaic</td>
<td>Wheat yellow mosaic virus</td>
<td>Bymovirus (+ss RNA)</td>
<td>Soil-borne fungus Polymyxa graminis</td>
<td>Up to 70</td>
<td>(Liu et al., 2005)</td>
</tr>
<tr>
<td>Rice (Oryza sativa)</td>
<td>Rice tungro disease</td>
<td>Rice tungro spherical virus, Rice tungro bacilliform virus</td>
<td>Waikavirus (ssRNA) and dsDNA</td>
<td>Leafhoppers</td>
<td>100</td>
<td><a href="http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/tungro">http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/tungro</a></td>
</tr>
<tr>
<td></td>
<td>Southern rice black-streaked dwarf virus</td>
<td>Southern rice black-streaked dwarf virus</td>
<td>Fijivirus (dsRNA)</td>
<td>Whitebacked planthopper</td>
<td>Up to 100</td>
<td>(Zhang et al., 2013)</td>
</tr>
<tr>
<td></td>
<td>Rice yellow mottle virus</td>
<td>Rice yellow mottle virus</td>
<td>Sobemovirus (+ss RNA)</td>
<td>Beetle (Sesselia pusilla)</td>
<td>Up to 97</td>
<td>(Awoderu, 1991)</td>
</tr>
<tr>
<td>Wheat (Triticum aestivum, Triticum turgidum,)</td>
<td>Mosaic</td>
<td>Triticum mosaic virus</td>
<td>Tritimovirus (+ss RNA)</td>
<td>Mite</td>
<td>70</td>
<td>(Byamukama et al., 2013)</td>
</tr>
<tr>
<td></td>
<td>Wheat streak mosaic virus</td>
<td>Triticum mosaic virus</td>
<td>Tritimovirus (+ss RNA)</td>
<td>Mite</td>
<td>100</td>
<td>(Edwards and McMullen, 1988)</td>
</tr>
<tr>
<td>Cassava (Manihot esculenta)</td>
<td>Mosaic</td>
<td>African cassava mosaic virus, East African cassava mosaic virus, Indian cassava mosaic virus, South African cassava mosaic virus, Sri Lankan cassava mosaic virus</td>
<td>Begomovirus (ssDNA)</td>
<td>White flies</td>
<td>15-90 %</td>
<td>(Hahn et al., 1980b; Legg and Fauquet, 2004);</td>
</tr>
<tr>
<td>Plant</td>
<td>Virus/Phytoplasma</td>
<td>Virus/Phytoplasma Type</td>
<td>Vector</td>
<td>Incidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cassava</td>
<td><em>Cassava Brown streak Virus; Uganda Cassava Brown streak virus</em></td>
<td>Ipomovirus (+ss RNA)</td>
<td>White flies</td>
<td>Up to 70%</td>
<td>(Hillocks et al., 2001)</td>
<td></td>
</tr>
<tr>
<td>Frogskin</td>
<td>16SrIII-L phytoplasma / Rice ragged stunt virus?</td>
<td>Oryzavirus (dsRNA)</td>
<td>Unknown</td>
<td>Up to 100%</td>
<td>(Calvert et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>Potato</td>
<td><em>Potato leafroll virus</em></td>
<td>Potyvirus (+ss RNA)</td>
<td>Aphids</td>
<td>29-98</td>
<td>(Killick, 1979; van der Wilk et al., 1991)</td>
<td></td>
</tr>
<tr>
<td>Mosaic</td>
<td><em>Potato virus X, Potato virus Y</em></td>
<td>Potexvirus (+ss RNA)</td>
<td>Aphids</td>
<td>Up to 80%</td>
<td>(Vaikonen, 1994)</td>
<td></td>
</tr>
</tbody>
</table>
Plant breeding programs across the world are constantly striving to improve disease resistance in crop plants by identifying new sources of resistance either in cultivated or related wild species. A large number of disease resistant genes have been identified and more so in wild relatives. In cassava for instance, resistance against cassava mosaic disease (CMD), was not found in cultivated cassava species (Manihot esculenta) but identified in wild relatives of cassava (Manihot glaziovii MUELLER VON ARGAU) (Nichols, 1947) and African landraces (Fregene et al., 2001; Akano et al., 2002). Currently, cassava breeding programs against CMD are based on identified natural resistance sources that are either polygenic or due to a single dominant gene, CMD2. A number of cultivars (TMS 30752, TMS 30395, TMS 30337 and TMS 60142) obtained from the interspecific crosses are reported to have a higher level of tolerance and have been released as varieties (Otim-Nape et al., 1998). Similarly, several sources of resistance to viruses have been identified from wild species such as Solanum pimpinellifolium, S. peruvianum, S. chilense, and S. habrochaites (Stevens et al., 1991; Vidavski et al., 2008; Anbinder et al., 2009). These wild species have been used to breed resistant lines. For example, crosses between S. habrochaites and S. lycopersicum yielded a resistant tomato hybrid hybrid (Vidavsky and Czosnek, 1998) to Tomato yellow leaf curl virus (TYLCV), which causes yield losses up to 100% (Picó et al., 1996). To date, several resistant genes/loci against TYLCV have been identified in wild tomato relatives, Ty-1 from S. chilense (Zamir et al., 1994), Ty-3 (Ji et al., 2007) and Ty-4 (Ji et al. 2009), Ty-2 from S. habrochaites (Hanson et al. 2006), and Ty-5 from S. peruvianum (Anbinder et al. 2009). S. peruvianum, PI 126944 is also the source of Sw-5, which provides resistance to Tomato spotted wilt virus (TSWV) (Picó et al., 2002).

Stacking/pyramiding a two or more resistant genes into a single cultivar has been the strategy of choice to not only increase disease resistance but also enhance its durability. Besides conventional breeding, much effort is also being made touse biotechnological approaches to engineer virus resistant into crop plants. Various approaches including but not limited to expression of viral proteins (coat protein, movement protein etc.)or RNAs (sense RNAs, anti-sense RNAs, hairpin RNAs), non-viral genes (RNases, antiviral inhibitor proteins, plant-antibodies, plant defence response elicitors) and various host-derived genes (dominant and recessive resistance genes)(Cillo and Palukaitis, 2014) are being used to generate virus resistance. For example transgenic lines TME 7-Hp-11 and TME 7-Hp-12 generated by expressing hairpin RNA homologous to the 3′-end of Cassava brown streak virus (CBSV)-coat protein (CP) in TME 7 cultivar (CMD-resistant variety under field conditions) background. These transgenic lines were resistant to CBSV as well as to East African cassava mosaic virus (EACMV)-Ugandan isolate (Vanderschuren et al., 2012), clearly indicating that coupling natural resistance with biotechnological approaches can lead to the development of durable broad spectrum resistance.
Plant virus replication in hosts

Based on the Baltimore classification, viruses are classified into seven families, depending on genomic composition and their mode of replication (Baltimore, 1971). The viral genome comprises of either RNA or DNA and the Baltimore system designates a virus as (+) or (−) sense depending on its relationship to the viral mRNA sequence. The mRNA sequence is designated as the positive (+) strand because it is the template for translation of viral proteins. RNA viruses with a genome complementary to their mRNA sequence are designated (−) sense. For a DNA virus, the strand complementary to the mRNA is the (+) strand. DNA viruses can be double stranded (+/−) or single stranded (−). Viruses encode very few proteins and require a large number of host proteins to successfully survive in the host (Ahlquist et al., 2003). The outcome of these interactions between virus and host proteins determine if a virus infection is successful or not; the identification of such interactions remains a challenge (Whitham and Wang, 2004). Plant viruses utilize a variety of replication mechanisms. The replication cycle is a multi-phased process that starts from virus entry into the host cell, uncoating of the virus particle, transcription and translation of the genome followed by viral genome assembly.

DNA virus cycle

Plant viruses are obligate parasites that intricately associate with living host systems for their translation, replication, systemic movement and genome assembly (Niehl and Heinlein, 2011; Schoelz et al., 2011; Ueki and Citovsky, 2011; Manabayeva et al., 2013). Most DNA viruses require a vector for transmission between hosts. Geminiviruses, that comprise a large part of the plant virus family, possess a monopartite or bipartite, (each 2.5–3 kb) circular single-stranded DNA (ssDNA). Replication of the DNA takes place via a rolling-circle replication mechanism from the intermediates formed by conversion of ssDNA to dsDNA in the host nucleus, followed by production and encapsidation of mature single-stranded DNA into viral particles (Gutierrez et al., 2004). Geminiviruses depend on host DNA-dependent RNA polymerase II for genome transcription (Caracuel et al., 2012; Saxena and Lomonossoff, 2014). Among the 6 to 8 proteins produced by geminiviruses, the Replication initiator protein (Rep), contains highly conserved domains (Rizvi et al., 2014) and is a multifunctional protein that is not only essential, but also sufficient, for viral DNA replication (Elmer et al., 1988; Hanley-Bowdoin et al., 1990). Rep mediates the recognition of origin of replication, DNA cleavage/ligation and acts as a terminase, thus determining release of complete genome-sized particles (Laufs et al., 1995). A homologue of Rep, Replication protein A (RepA)(Heyraud-Nitschke et al., 1995) also plays an important role in geminivirus replication (Luque et al., 2002; Castillo et al., 2003; Singh et al., 2007). Viral replication enhancer, REnc (also called AC3, C3, or L3) interacts with the Rep protein to enhance viral DNA accumulation (Stanley et
Viral C4 protein is not directly involved in either replication or translation but interact with the movement protein, MP (Rojas et al., 2001). Viral C4 protein also represses a plant DNA methyltransferase, METHYLTRANSFERASE 1 (MET1) and reduces de novo plant DNA methylation (Rodríguez-Negrete et al., 2013). In addition, Transcriptional Activator Protein or TrAP (also called AL2, AC2, or C2) is required for the accumulation of ssDNA (Hayes and Buck, 1989). TrAP expression in turn activates expression of coat protein (CP) and the nuclear shuttle protein (NSP). NSP binds to viral DNA and moves it across the nuclear envelope, where the movement protein (MP) traffics it across the plasmodesmata. Circular ssDNA that are encapsidated by coat protein into virions are then available to be spread around the host and also acquired by insect vectors (Hanley-Bowdoin et al., 2013).

Viral proteins in addition to interacting with host cellular factors also interact with each other to complete their life cycle (Novoa et al., 2005). Rep protein interacts with the replication protein A32 (RPA32), which represses Rep nicking while promoting its ATPase activity (Singh et al., 2007). Silencing on RPA32 increases virus (TYLCV) accumulation in N. benthamiana (Czosnek et al., 2013). Rep interacts with replication factor C (RFC) to recruit DNA polymerase delta to the primer-template (Luque et al., 2002). Rep also interacts with minichromosome maintenance 2 (MCM2) (Suyal et al., 2013) and proliferating cell nuclear antigen (PCNA) (Bagewadi et al., 2004). Wheat dwarf virus (WDV)-RepA interacts with GRAB2 (Geminivirus Rep A-binding), a NAC-containing protein that might have differential role in geminivirus infections (Czosnek et al., 2013). REn protein interacts with host proliferating cell nuclear antigen (PCNA) for its replication and retinoblastoma-related protein (pRBR) during the infection process (Settlage et al., 2005). C4 protein interacts with host RKP, a RING finger E3 ligase and regulates cell-cycle by degrading ICK/KRP proteins (Lai et al., 2009). Viral C4 protein also induces a severe developmental phenotype by the disruption of multiple hormonal pathways (Mills-Lujan and Deom, 2010).

All plant dsDNA viruses are considered to be plant pararetroviruses since they replicate their genome by transcription in the nucleus followed by reverse transcription in the cytoplasm and accumulate as episomes (i.e., nonintegrated in the host genome) (Ménissier et al., 1983; Hohn and Rothnie, 2013) and/or as endogenous (i.e., integrated) forms in the host genome (Staginnus et al., 2009). The pararetroviruses that integrate into the host germinal cells are fixed by natural selection and/or genetic drift in plant populations (Hull and Covey, 1995). Plant pararetroviruses consist of six genera (Caulimovirus, Soymovirus, Cavemovirus, Petuvirus, Badnavirus and Tungrovirus) and each of them vary in having differing strategies for generation of sub-genomic RNAs, number of translation initiation sites per mRNA, arrangement of their open reading frames and the degree of ORF-fusion (Hohn and Rothnie,
Plant pararetroviruses encode several genes that are driven by two promoters, the 35S and 19S promoter. Transcription of pre-genomic RNA/polycistronic mRNA is driven by the 35S promoter, and the 19S promoter drives the mRNA encoding transactivator-viroplasmin (TAV) (Covey and Hull, 1981). Many ORFs are produced due to alternative splicing and are known to be efficiently translated and have regulatory functions (Kiss-László et al., 1995; Froissart et al., 2004). Translation occurs basically in two process, shunting (Fütterer et al., 1993) and TAV mediated polycistronic translation (Bonneville et al., 1989). The pre-genomic RNA is reverse transcribed using a methionine initiator tRNA, which may serve as a primer binding site for minus-strand DNA synthesis (Huang and Hartung, 2008; Gayral et al., 2010).

Coat protein (CP), which is processed into two major forms, is involved in docking of the virions at the nuclear pore and release of DNA into the nucleus (Karsies et al., 2002).

Viral proteins interact with many plant factors for efficient replication. For example, viral-TAV interacts with plant protein kinase TOR (target of rapamycin), S6 kinase (S6K), ribosomal subunits (L13, L18 and L24), eukaryotic initiation factors (eIF) 4B and 3 (subunit g) along with re-initiation supporting protein (RISP) (Bureau et al., 2004; Thiébeauld et al., 2009; Schepetilnikov et al., 2011; Schepetilnikov et al., 2013).

**RNA virus cycle**

Replication of positive-strand RNA viruses starts with the disassembly of coat protein subunits to release genomic RNA for translation and replication. The disassembly of viral coat protein is an active process that proceeds from the 5′ end of the viral RNA. By involvement of host ribosomes, substantial overlap develops between translation and disassembly. This is termed co-translational disassembly (Wilson, 1984; Barton et al., 1999). The released viral polyproteins are co- and post-translationally processed by virus-encoded proteinases (Dougherty et al., 1989). Replication proceeds through synthesis of an intermediate complementary RNA strand followed by assembly of viral particles (reviewed (Saxena and Lomonossoff, 2014). The viral proteins then redirect the viral genome from functioning as an mRNA towards synthesizing complementary negative-strand RNA by RNA-dependent RNA polymerase (RdRP) in a 3′-to-5′ direction by removal of coat protein subunits from the 3′ end by viral replicate protein (Wu and Shaw, 1997). Various viral proteins interact with host factors (Table 2) and with host membrane structures (like chloroplast outer membranes, endoplasmic membrane, peroxisomes, golgi bodies, vascular membranes or mitochondria as their site of replication) forming viral replication complex (VRC) (Ahlquist et al., 2003). The complementary strand that is synthesized is then capped on the 5′ end for stability by either synthesizing its own cap structure or by acquisition of host cellular mechanisms for capping (Decroly et al., 2012). In addition to 5′ capping, 3′ ends are polyadenylated and mature
mRNAs are released into the cytoplasm to serve as template for generating multiple copies of positive-strand RNAs (Kawamura-Nagaya et al., 2014). Negative-sense (-) RNA viruses on the other hand have an alternate mode of replication in which the viral genome encodes its own RdRP to transcribe its negative-sense RNA. The positive-sense RNA that is consequently formed serves as template for viral protein synthesis and genome replication (Palese et al., 1996). This process is especially susceptible to mutation since the RdRps encoded by RNA viruses lack proof-reading activity. As a result the high rates of mutation in RNA virus genomes may provide higher adaptive capability (Elena and Sanjuán, 2005). Additionally, the mutations can be selectively eliminated by lowering translation of replication proteins from the defective genomes due to the inherent capacity of virus cis-acting elements for replication (Kawamura-Nagaya et al., 2014). dsRNA viruses upon entering the host cell, encode viral RdRP to transcribe their negative-sense strand into a positive-sense strand that serves as a template for viral protein synthesis and genome replication (Ahlquist, 2006). Finally, encapsidation of RNA viruses is essential for systemic spread of viral infection and sometimes for cell-to-cell movement (Saxena and Lomonossoff, 2014), which is highly specific and controlled by specific packaging signal (Rochon and Siegel, 1984; Reade et al., 2010). Some of the host-virus interacting factors are presented in Table 2.
Table 2: Some of the host factors involved in plant virus establishment.

<table>
<thead>
<tr>
<th>Viral cycle</th>
<th>Sl. No.</th>
<th>Viral gene</th>
<th>Virus</th>
<th>Host factors</th>
<th>Host</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>RdRP</td>
<td>ZYMV</td>
<td>poly-(A) binding protein (PABP)</td>
<td><em>Cucumis sativus</em></td>
<td><em>(Wang et al., 2000)</em></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>RdRP</td>
<td>TuMV</td>
<td>heat shock cognate 70-3 (Hsc70-3) and poly(A)-binding (PABP)</td>
<td><em>Arabidopsis thaliana</em></td>
<td><em>(Dufresne et al., 2008)</em></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>RdRp and VPg-protease (VPg-Pro)</td>
<td>TuMV</td>
<td>eEF1A</td>
<td><em>Arabidopsis thaliana</em></td>
<td><em>(Thivierge et al., 2008)</em></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>RdRP</td>
<td>TMV</td>
<td>eEF1A</td>
<td><em>N. tabacum</em></td>
<td><em>(Yamaji et al., 2006)</em></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>RdRP</td>
<td>TMV</td>
<td>eEF1B</td>
<td><em>N. tabacum</em></td>
<td><em>(Hwang et al., 2013)</em></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Vpg</td>
<td>TuMV and LMV</td>
<td>eIF4E</td>
<td><em>A. thaliana</em></td>
<td><em>(Duprat et al., 2002)</em></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Ren</td>
<td>TYLCV</td>
<td>PCNA (proliferating cell nuclear antigen)</td>
<td><em>A. thaliana</em></td>
<td><em>(Castillo et al., 2003)</em></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>P6</td>
<td>CaMV</td>
<td>CHUP1</td>
<td><em>A. thaliana</em></td>
<td><em>(Angel et al., 2013)</em></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>NSP</td>
<td>CalCuV</td>
<td>AtNSI</td>
<td><em>A. thaliana</em></td>
<td><em>(McGarry et al., 2003)</em></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>C4</td>
<td>BSCTV</td>
<td>RKP, a RING finger protein</td>
<td><em>A. thaliana</em></td>
<td><em>(Lai et al., 2009)</em></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>C4</td>
<td>BCTV</td>
<td>Shaggy-related protein kinase AtSKeta</td>
<td><em>A. thaliana</em></td>
<td><em>(Piroux et al., 2007)</em></td>
</tr>
<tr>
<td>No.</td>
<td>Category</td>
<td>Virus</td>
<td>Protein/Entity</td>
<td>Host Plant</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Translation and its enhancers</td>
<td>TMV</td>
<td>TOM1/TOM2A</td>
<td>A. thaliana</td>
<td>(Hagiwara et al., 2003; Tsujimoto et al., 2003)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>TMV</td>
<td>eIF1A</td>
<td>Cowpea</td>
<td>(Matsuda and Dreher, 2004)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>TyMV</td>
<td>eIF4F</td>
<td>N. benthamiana</td>
<td>(Gallie, 2002)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>TyMV</td>
<td>eIF4F</td>
<td>Cowpea</td>
<td>(Matsuda et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Vpg</td>
<td>LMV</td>
<td>eIF4E</td>
<td>Capsicum annuum</td>
<td>(Ruffel et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Vpg</td>
<td>TuMV and LMV</td>
<td>eIF(iso)4E</td>
<td>A. thaliana</td>
<td>(Wittmann et al., 1997a)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Vpg</td>
<td>TuMV and LMV</td>
<td>eIF4E</td>
<td>A. thaliana</td>
<td>(Duprat et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Movement</td>
<td>TMV</td>
<td>MBP2C</td>
<td>N. benthamiana</td>
<td>(Kragler et al., 2003)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Movement</td>
<td>TMV</td>
<td>Plasmodesmata-located proteins</td>
<td>N. benthamiana</td>
<td>(Amari et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Movement</td>
<td>TMV</td>
<td>pectin methylesterase (PME)</td>
<td>N. benthamiana</td>
<td>(Chen and Citovsky, 2003a)</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**
- **TMV:** Tobacco mosaic virus
- **TuMV:** Turnip mosaic virus
- **LMV:** Lettuce mosaic virus
- **TyMV:** Turnip yellow mosaic virus
- **TMV:** Turnip yellow mosaic virus
- **BCTV:** Beet curly top virus
- **ZYMV:** Zucchini yellow mosaic virus
- **TYLCV:** Tomato yellow leaf curl Sardinia virus
- **CaMV:** Cauliflower mosaic virus
- **CaLCuV:** Cabbage leaf curl virus
- **BSCTV:** Beet severe curly top virus
Common responses to plant virus infection

In a susceptible system, host-virus interaction leads to a visible manifestation of disease symptoms that determines severity while in a resistant system, interactions are associated with either by the hypersensitive response (HR)-independent (resistance activated early in the infection cycle) or HR-dependent response (resistance activated later in the infection cycle) (Bendahmane et al., 1999). In potato, extreme resistance provided by dominant R gene, Rx against potato virus X, was reported to be a non-necrotic resistance response (Bendahmane et al., 1999) while Tobacco mosaic virus (TMV) induces necrotic local lesions and is restricted to inoculated leaves of N. tabacum having N resistance genes (Whitham et al., 1994).

Gene expression profiling in virus-infected plants indicate a significant impact of viral infection on a wide array of cellular processes (Golem and Culver, 2003), altering changes in basal plant metabolism, protein synthesis, developmental and photosynthetic processes (Agudelo-Romero et al., 2008; Babu et al., 2008). Some of the common host genes that are expressed upon viral infection include defence-associated genes PR1, PR2 (1,3 GLUCANASE), PR4, PR5 (Thaumatin like protein) and heat shock genes (HEAT SHOCK PROTEIN 70) that are coordinately regulated (Whitham et al., 2003; Chen et al., 2008; Gorovits et al., 2013). Activation of defence signaling pathways elicit salicylic acid (SA) which in turn increases expression of defence-related genes PR1 and BGL2 (β 1,3-GLUCANASE 2) (Huang et al., 2005).

Heat shock proteins are largely synthesized as a general response to accumulation and aggregation of unfolded or misfolded viral proteins in the cytoplasm (Aparicio et al., 2005). Hsp70 associates with membrane-bound potyvirus viral replication complex and regulates the amount of coat protein (CP) during virus replication (Hafrán et al., 2010). In addition, many stress- and defence-associated signaling genes including regulatory genes PHYTOALEXIN DEFICIENT 3 & 4 (PAD) are involved during plant defence responses (Jirage et al., 1999; Zhou et al., 1999).

Another remarkable immune response to viral infection is the production of plant small RNAs (sRNA) that are 20-24 nucleotides long. sRNAs are broadly classified into microRNAs (miRNAs) and small interfering RNAs (siRNAs) depending on the source of their precursors. These are diverse in size, sequence, distribution and function (Qu et al., 2008; Ruiz-Ferrer and Voinnet, 2009). In plants, sRNA biogenesis is triggered in response to viral infection by the production of virus derived–dsRNA or hairpin (hp) RNA molecules. For plant RNA viruses, dsRNA and hpRNAs are formed either through direct hybridization of complementary strands during the RNA-viral replication, the action of host RdRPs on viral mRNA or the self-annealing of viral RNA to form hairpin structures. In ssDNA viruses, dsRNA structures originate from
annealing of sense and antisense transcripts (Chellappan et al., 2004) with the 35S leader sequence of dsDNA virus has also been implicated in the process (Moissiard and Voinnet, 2006).

In general, RNase III family dicer-like enzymes cleave dsRNA or intramolecular hairpins within viral genomic RNA to produce siRNAs of varying lengths of siRNAs (Hammond, 2005; Molnár et al., 2005; Ding and Voinnet, 2007). In plants, four forms of DCL proteins, DCL1, 2, 3 and 4 are reported. DCL1 produces miRNAs while DCL2, 3 and 4 produce 22, 24 and 21 nt siRNAs, respectively (Blevins et al., 2006). Mature siRNAs are then incorporated into a complex called RNA-induced silencing complex (RISC), via argonaute proteins (AGO1-10) (Pantaleo, 2011). The siRNAs are also amplified by RdRps (Wassenegger and Krczal, 2006). The siRNA duplex is then separated into a guide and a passenger strand through helicase activity. The guide strand remains bound to the RISC and guides it to complementary regions on the viral mRNA or viral genome (for RNA viruses) which are then degraded by the AGO proteins (Matranga et al., 2005; Ambrus and Frolov, 2009; Pumplin and Voinnet, 2013). To counteract these defence pathways, viruses encode proteins that suppress RNA silencing. Viral suppressors of RNA silencing (VSR) commonly bind and interfere with key components involved in the RNA silencing machinery (Burgyán and Havelda, 2011). Turnip crinkle virus (TCV)-CP strongly suppresses post-transcriptional gene silencing (PTGS) by interfering with the function of the Dicer-like RNase in plants (Qu et al., 2003). The TCV-P38 protein is known to suppress RNA silencing by binding to AGO1 (Thomas et al., 2003). TMV-encoded 126-kDa replicase subunit (P126) interferes with HEN1-mediated methylation of small RNAs (Vogler et al., 2007). The TYLCV-V2 protein interacts directly with SISGS3 and inhibits its role in RNA silencing (Glick et al., 2008). In response to VSR, plants induce resistance mechanisms that can effectively suppress VSRs leading to extreme resistance. In Nicotiana sp ‘extreme resistance’ is displayed against Tomato bushy stunt virus (TBSV) infection by sensing and antagonizing the effects of TBSV-P19 suppressor and this immunity is salicylate- and ethylene-dependent (Sansregret et al., 2013). Another important characteristic of antiviral silencing via RNAi is the phenomenon of silencing amplification. Several siRNAs generated in response to viral infection are termed secondary siRNAs and are derived from the activity of host RdRPs as compared to ‘primary’ siRNAs, which are products of virus RNA templates (Pumplin and Voinnet, 2013).

**Nature of plant resistance**

Plant pathogens have to overcome many layers of complex defence to successful infect the host (De Ronde et al., 2014). Resistance to diseases is broadly grouped into non-host and host resistance (Fraser, 1992). Non-host resistances which is regulated by multiple genes is
multilayered and durable against a broad range of pathogens (Thordal-Christensen, 2003; Ham et al., 2007); However, the underlying mechanism for non-host resistance is not completely understood (Mysore and Ryu, 2004). Proteins located on plant plasma membranes recognize pathogens via so called microbe-associated molecular patterns (MAMPs) or PAMPs (Pathogen-associated molecular patterns) which activates defences at the primary infection site leading to rapidly induced programmed cell death, resulting in a visible necrotic local lesion, known as the hypersensitive response (HR) (Goldbach et al., 2003). Recent studies have indicated that during plant-virus interactions, PAMP recognition by pattern recognition receptors (PRRs) might possibly involve intracellular receptor-like kinases (RLKs) (Kørner et al., 2013).

Host resistance or vertical resistance (also known as race-specific or gene-for-gene resistance) is based on direct or indirect interaction between pathogen avirulence (Avr) factors and plant resistance or R proteins. This kind of resistance is heritable (Revers and Nicaise, 2001) and results in prevention of viral replication or systemic spread of the virus. The consequence is either a highly localized HR or even a complete absence of any disease symptom (Bendahmane et al., 1999). Host resistance is determined by the specific interaction between pathogen effectors and the corresponding plant R-proteins and can be isolate-specific (Flor, 1971; Hammond-Kosack and Jones, 1997). Dominant resistance (R) genes encode nucleotide-binding site leucine-rich repeat type proteins consisting of a Nucleotide Binding Site (NBS) region, Leucine Rich Repeat (LRR) and Coiled-coil (CC) or Toll and Interleukin-1 like Receptor (TIR) domains that allow for the recognition of a vast range of pathogens (Moffett, 2009). These are one of the largest and most variable gene families found in plants (Clark et al., 2007).

**Dominant resistance**

Pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) response in plants results from both host and non-host resistance. PTI include many disease induced responses such as callose deposition at infection sites, production of reactive oxygen species (ROS) and phenolic compounds, and induction of hormone-mediated signaling leading to the activation of defence genes (Nicaise et al., 2009). Suppression of PTI by plant pathogens is achieved by encoding effector proteins that bind to PRRs and interfere with substrate recognition by PRRs (De Ronde et al., 2014). Such a successful infection process is generally termed as effector-triggered susceptibility (ETS). Several of the resistance genes have been successfully used for disease management but they are limited and not always durable as the pathogens continuously evolve to overcome them (Lo Iacono et al., 2013). For instance, in commercial tomato cultivars the Tm-2 gene that confers resistance to TMV and its close relative Tomato mosaic virus (ToMV) was overcome by mutations in the viral MP gene (Weber et al., 1993). Similarly, Sw-
which confirms resistance to TSWV was overcome by two different single amino acid substitutions (C118Y or T120N) in TSWV-movement protein, NSm (López et al., 2011). Some examples of dominant resistance genes that have been studied are N, Rx1, Rx2, Sw5, Tm2\(^2\), HRT, RTM1, RTM2 and RCY1 (for review see Galvez et al., 2014).

**Recessive resistance**

Recessive resistance is the consequence of mutations resulting in the loss of a host factor specifically required for successful infection. In general, nearly 50% of the known plant resistance genes to viruses behave in a recessive manner with resistance to potyviruses in particular, to be recessive in nature (Revers and Nicaise, 2001). Recently, a rice recessive resistance gene RYMV2 has been associated with a gene homologous to the Arabidopsis CPR5 gene, a known defence mechanism regulator against *Rice yellow mottle virus* (Orjuela et al., 2013). In addition, recessive resistance can be generated by identifying the host factors that are essential for virus establishment. For example, proteins TOBAMOVIRUS MULTIPLICATION (TOM) 1 and its homologs TOM3 and TOM2A interact with each other to facilitate formation of the tobamovirus replication complex that is essential for multiplication of tobamoviruses (Tsujimoto et al., 2003; Yamanaka et al., 2000). Replication of TMV and other tobamoviruses has been reported to be completely inhibited in tom1, tom3, and tom2A mutants of *N. tabacum* without affecting plant growth (Asano et al., 2005). Many cloned recessive resistance genes are reported to be translation initiation factors of the 4E or 4G family (eIF4E/eIF4G) (Truniger and Aranda, 2009). Some examples of recessive genes are sbm2, pvr1, pot-1, rymv1, bc-3 and for more examples see review (Galvez et al., 2014). Recessive resistance (eIF4E-mediated resistance) can be overcome by favorable mutations on virus proteins. For example, mutations observed in Potato Virus Y-viral genome-linked protein (at specific amino acids position 101,105,115, 119 and 123) resulted in the adaptation to recessive resistance (Charron et al., 2008). Similarly, mutation in Lettuce mosaic virus-cylindrical inclusion protein at position 621 (serine is substituted by a threonine) allowed the virus to overcome resistance offered by mol1 and mol2 genes in lettuce (Abdul-Razzak et al., 2009).

**Cross-protection**

Cross-protection is induced in a host plant upon infection with milder virus strains that can subsequently trigger resistance against from challenge a more virulent strain of the same virus (Zhou and Zhou, 2012). Cross-protection arises when several virus species compete in one host for important host factors (Roossinck, 2005). This results in host-mediated exclusion through silencing where siRNAs derived from infection by the first virus degrade the genome of the second infecting virus (Ratcliff et al., 1999). For example, a mild strain of *Zucchini yellow mosaic virus*-WK (ZYMV-WK), containing the FRNK-to-FINK mutation on the conserved
region of ZYMV HC-Pro protein, has been successfully used for cross-protection against virulent ZYMV strains infecting Zucchini (Lecoq, 1991). This technique has also been successfully employed in grapevine against Grapevine fanleaf virus (Komar et al., 2008).

**Durability of plant resistance**

Durable resistance is defined as resistance that remains effective while being extensively used in agriculture for a long period in an environment conducive to the disease (Johnson, 1983). Durable resistance is often quantitative, resulting from additive effects of many genes (Parlevliet, 2002). Increasing the durability of crop resistance to plant pathogens is one of the key factors of virulence management (Parlevliet, 2002). Durable and broadened resistance can be developed by stacking several R genes (Pink, 2002) and also by the cultivation of mixtures (multiline cultivars and cultivar mixtures) of plants (Mundt, 2002) that greatly reduce the evolutionary pressure on pathogens. The breakdown of genetic resistance against plant pathogen interactions is a major limit to the genetic control of crop disease (Palloix et al., 2009). Resistance genes may be overcome through passive means such as spontaneous mutation or genetic diversity (Whitham and Wang, 2004). For example, polygenic resistance in pepper cultivar pvr³ was overcome by PVY-SON41p isolates with a single mutation in VPg protein (Palloix et al., 2009). The number of mutations required for virulence acquisition is another parameter that is related to resistance durability (Harrison, 2002). In order to overcome polygenic durable resistant alleles plant pathogenic viruses will quite plausibly require several mutations (Ayme et al., 2007; Quenouille et al., 2013). When breeding for resistant cultivars, breeders must consider the ability of the virus to respond to a progressive selection despite the fact that greater durability of polygenic resistance is observed relative to monogenic resistance (Palloix et al., 2009). A recent study indicates that infection with an additional compatible plant virus (CMV) can compromise resistance confirmed by the single monogenic resistant Ty-1 allele against TYLCV (Butterbach et al., 2014).

Relatively few studies have been performed on polygenic inherited resistance or tolerant responses to plant viruses, and even fewer cases exist where recessive resistance has been transferred into crop varieties (Harrison, 2002). Two homozygous recessive pea genes (sbm1 and sbm2) have provided particularly durable resistance to Pea seed-borne mosaic potyvirus (Keller et al., 1998; Johansen et al., 2001).

**Impact of climate change on host-virus interaction**

The classical disease triangle used by plant pathologists involving a host, pathogen, and the environment and their interactions determine the extent of disease development (Pautasso et al., 2012). A virulent pathogen cannot infect a susceptible host without a favorable
environment, or similarly, a susceptible host in a favorable environment will not develop disease without the virulent pathogen (Parker and Gilbert, 2004). Agriculture intensification by humans along with the climatic changes will impact plant-vector-virus pathosystems (Coakley et al., 1999; Jones, 2009). Climate change is a dynamic and complex process that could have either a positive, negative or no impact on yield, efficacy of disease management strategies or the geographical distribution of plant diseases (Chakraborty et al., 2000). Currently, much of the available literature suggests that climate change can directly affect the micro and macro environment of plants and plant communities even in the absence of pathogens but may also bring about changes in plants that will affect their interactions with pathogens (Garrett et al., 2006). Climate change such as increased temperatures, quantity and pattern of precipitation, increased carbon dioxide (CO₂) and ozone (O₃) levels, drought affects plant diseases together with other components of global change, i.e., long-distance introduction of exotic species (Bradley et al., 2011). This might further influence the co-evolution of plants and their pathogens (Garrett et al., 2006; Eastburn et al., 2011).

Elevated CO₂ alters the resistance against PVY infection by altering plant secondary metabolites (Matros et al., 2006). Plant could also benefit from the climate change. For example under elevated CO₂ levels, the incidence and severity of TYLCV infection on tomato was reduced by altered plant defence strategies (Huang et al., 2012b). Soybean mosaic virus (SMV) under elevated O₃ treatment slowed systemic infection and disease development by inducing a nonspecific resistance against SMV (Bilgin et al., 2008). On the contrary, elevated CO₂ increases the preference of green peach aphid for Potato leafroll virus-infected plants over uninfected plants (Srinivasan et al., 2006).

The effects of elevated temperature on plants tend to vary greatly during the year. Elevated temperatures, under predicted climate change scenarios and extended seasons (over-summering and overwintering) tend to create stress that makes plant more susceptible to pathogens (Garrett et al., 2006). Under temperature increases from 22°C to 30°C, resistant N. benthamiana plants became partially susceptible to Potato virus X (PVX) and its aphid elicitor (Wang et al., 2009). Silencing against a geminivirus was 10-fold greater in plants grown in cycles of 30/26°C (day/night) than 22/18°C (Tuttle et al., 2008). High temperature can also induce higher antiviral RNA silencing which is inhibited by low temperature (Qu et al., 2005; Zhang et al., 2012). Higher disease incidences was observed under drought stress for Beet yellows virus (Clover et al., 1999) and Maize dwarf mosaic virus (Olson et al., 1990).

Climate change can also indirectly affect virus distribution by influencing the vector distribution over a geographical area which indirectly influences the spread of plant viruses. For example, the incidence and spread of CMD in cassava directly correlated to the whitefly
population (Legg et al., 2011). Another example include the increased survival of aphids during warmer/mild winters which can cause more severe Barley yellow dwarf potyvirus (BYDV) infections (Coakley et al., 1999). Additional examples for effects of abiotic variables on plant viruses are presented in Table 3.

**Approaches to investigate natural resistance**

Development of resistant cultivars is a long and costly process with the selected cultivars often lacking either broad spectrum resistance or durable resistance (Lecoq et al., 2004). The majority of resistance genes used in plant breeding originate from exotic gene pools such as old cultivars or non-agricultural species (Leroy et al., 2014) or from wild relatives of crop plants. These resistance genes have been successfully introgressed into elite varieties (Bent, 1996). Phenotypic selection of natural resistance under field conditions is largely inefficient due to highly variable environmental conditions and many plants may escape infection (Sood et al., 2009). Currently many improved methods, both mechanical and non-mechanical transmission of viruses, are used for identifying natural resistance in the germplasm under controlled conditions. For example, natural vectors (Bemisia tabaci) were selectively fed on TYLCV infected plants (forced-feeding) and viruliferous vectors then used for infecting test plants under controlled conditions (cage inoculation). This resulted in 100% infection and was used to identify cultivars with variable levels of resistance (Picó et al., 1998).
### Table 3: Effect of climatic variables on plant virus pathogenicity.

<table>
<thead>
<tr>
<th>Sl no.</th>
<th>Virus</th>
<th>Genus</th>
<th>Host</th>
<th>Variable</th>
<th>Effect on plants</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Turnip mosaic virus (TMV)</td>
<td>Potyvirus</td>
<td>Arabidopsis thaliana</td>
<td>Both heat and drought stress</td>
<td>Deactivation of defence responses and increased susceptibility</td>
<td>(Prasch and Sonnewald, 2013)</td>
</tr>
<tr>
<td>2</td>
<td>Cymbidium ringspot virus (CymRSV)</td>
<td>Potexvirus</td>
<td>Nicotiana benthamiana</td>
<td>Low temperature (15°C)</td>
<td>Inhibits RNA silencing and exhibits higher accumulation of virus</td>
<td>(Szittya et al., 2003)</td>
</tr>
<tr>
<td>3</td>
<td>Tobacco mosaic virus (TMV)</td>
<td>Tobamovirus</td>
<td>Nicotiana tabacum</td>
<td>Higher temperature (28°C)</td>
<td>Overcomes N-mediated resistance</td>
<td>(Király et al., 2008)</td>
</tr>
<tr>
<td>4</td>
<td>Tobacco mosaic virus (TMV)</td>
<td>Tobamovirus</td>
<td>Nicotiana tabacum</td>
<td>Higher temperature (33°C)</td>
<td>Overcome Tm-1 resistance gene</td>
<td>(Fraser and Loughlin, 1982)</td>
</tr>
<tr>
<td>5</td>
<td>Tomato Spotted Wilt Tospovirus (TSWV)</td>
<td>Tospovirus</td>
<td>Capsicum chinense</td>
<td>Higher temperature (32°C)</td>
<td>Overcome Tsw mediated resistance</td>
<td>(Király et al., 2008)</td>
</tr>
<tr>
<td>6</td>
<td>Tomato yellow leaf curl virus (TyLCV)</td>
<td>Begomovirus</td>
<td>Lycopersicon esculentum</td>
<td>Higher temperature (35°C)</td>
<td>More mortality of viruliferous whiteflies than non-viruliferous whiteflies</td>
<td>(Pusag et al., 2012)</td>
</tr>
<tr>
<td>7</td>
<td>Potato virus Y (PVY)</td>
<td>Potyvirus</td>
<td>Nicotiana tabacum</td>
<td>Elevated CO2</td>
<td>Delay the viral spread</td>
<td>(Ye et al., 2010)</td>
</tr>
<tr>
<td>8</td>
<td>Cucumber mosaic virus (CMV)</td>
<td>Cucumovirus</td>
<td>Nicotiana tabacum</td>
<td>Elevated CO2</td>
<td>Delay the viral spread</td>
<td>(Matros et al., 2006)</td>
</tr>
<tr>
<td>9</td>
<td>Barley yellow dwarf virus (BYDV)</td>
<td>Ipomovirus</td>
<td>Avena sativa</td>
<td>Elevated CO2</td>
<td>Increased water use efficiency (48-174%) and enhances the viral reservoir</td>
<td>(MalmstrÖM and Field, 1997)</td>
</tr>
</tbody>
</table>
Similarly, cassava genotypes were screened for resistance against CBSV by using a grafting technique (Moreno et al., 2011; Mohammed et al., 2012; Wagaba et al., 2013). Improvements in mechanical transmission techniques such as rubbing the sap collected from infected leaves onto the upper leaf surface of the test plant and agro-inoculation techniques could also be efficiently used for screening the resistant hosts (Hull, 2005). Plant breeders utilize naturally resistant genes in virus resistance breeding programs to transfer them by crossing to compatible species as well as to elite cultivars to impart resistance. Protoplast fusion has also been used to transfer resistance genes between incompatible species (Thach et al., 1993). Hence in breeding programs, precise identification of resistance genotypes in the given population is a prerequisite for release of new resistant cultivars (Rubio et al., 2014).

Another approach for identifying genes of interest is by mutation studies. Analyzing mutagenized (site-directed and random mutagenesis) populations with novel screening methods has made it relatively easy to identify newer traits that are being developed and characterized at the molecular level (Sikora et al., 2011). While recessive mutations are often recovered, dominant mutations are recovered only very rarely (McConnell et al., 2001; Østergaard and Yanofsky, 2004). Mutagenesis results in an allelic series, allowing for strong, intermediate and weak alleles of a given gene, in addition to loss-of-susceptibility (Bowman et al., 1991; Gawehns et al., 2013). For example, screening ethyl methanesulphonate (EMS)-mutagenized Arabidopsis mutants for cum1-1 (cucumovirus multiplication 1 locus that encodes translation initiation factor 4E) obtained mutants exhibiting resistance to Clover yellow vein virus (CIYVV) (Sato et al., 2005). Similarly, loss-of-susceptibility to potyviruses 1 (lsp1) mutant (LSP1 gene encodes eukaryotic initiation factor 4E) displayed reduced susceptibility to Turnip mosaic potyvirus (TuMV) in Arabidopsis (Lellis et al., 2002). The mutated genes of interest are then obtained by either molecular or genetic markers or by developing a mapping population that segregates for the mutation of interest (collectively termed as map-based cloning, often called positional cloning) (Østergaard and Yanofsky, 2004). In addition to EMS-induced mutation, other approaches like T-DNA insertion mutation (Krysan et al., 1999) and mutation by transposable element insertion (Tadege et al., 2008) have been effectively used in functional analyses of genes. Only a small proportion of the knockout lines generated by induced mutation exhibit obvious phenotypes at the whole-plant level (Bouché and Bouchez, 2001). Hence this approach is a tedious and time-consuming process requiring sequence-based screening of large-scale mutant populations or transgenic plants to detect mutations in a gene of interest (Ramegowda et al., 2013).

TILLING technique (targeting induced local lesions in genome) is a reverse genetic tool that involves screening for variations created by mutations (either natural or by use of a potent chemical mutagenic agent in existing genes (Nicais and Revers, 2001). TILLING is suitable
for both small- and large-scale screening (McCallum et al., 2000). This approach can also be applied to small genomes such as Arabidopsis as well to very large genomes such as wheat. TILLING technique assists in identification of novel alleles (silent, missense, nonsense, and splice site mutations) but a disadvantage being that it requires prior DNA sequence information of a target gene (Barkley and Wang, 2008). The TILLING technique has been successfully employed in Capsicum annuum for the identification of novel resistance gene (eIF4E alleles: pvr210, pvr211, pvr212, pvr213 and pvr214) against PVY (Ibiza et al., 2010). Currently, the availability of new genetic tools permit the selection of genes using reliable molecular markers such as random-amplified polymorphic DNAs (RAPDs), restriction fragment length polymorphisms (RFLPs), microsatellites and PCR-based DNA markers such as sequence characterized amplified regions (SCARs), sequence-tagged sites (STS) and inter-simple sequence repeat amplification (ISA), amplified fragment length polymorphic DNAs (AFLPs) and amplicon length polymorphisms (ALPs). These methods can be used for selection using F2 and back-cross populations, near-isogenic lines, doubled haploids and recombinant inbred lines and substantially reduce the time and cost of producing superior resistant varieties by avoiding the need for screening large populations (Mohan et al., 1997; Panthee et al., 2013). For example, the novel PCR-based locus-specific molecular marker NCTm-019, associated with Tm2⁰, which confers resistance to ToMV can be efficiently used to develop ToMV-resistant cultivars (Panthee et al., 2013). Marker assisted selection was performed to identify tomato lines possessing Ty-2 and Ty-3 genes that confirms broad-spectrum resistance to tomato (yellow) leaf curl disease (TYLCD/ToLCD) (Prasanna et al., 2014). Similarly, resistant cassava lines obtained by introgression of CMD2 gene were identified through marker-assisted selection (MAS) (Okogbenin et al., 2007; Rabbi et al., 2014).

Another functional genomics tool to identify and study the function of either an individual gene or multiple genes in a single plant is by silencing target genes through virus-induced gene silencing (VIGS). VIGS suppresses the expression of a targeted endogenous plant gene(s) and this technique has been successfully employed to identify many plant genes that are involved in developmental process as well as study plant pathogen interaction (reviewed in (Huang et al., 2012a)). For example, silencing of RAR1, EDS1, and NPR1/NIM1-like genes in N-mediated resistance to TMV using a TRV based VIGS approach suggested that these genes confer resistance to TMV (Liu et al., 2002).

With the advent of next generation sequencing (NGS), transcriptome profiling of pathosystems that is independent from available genomic resources has led to the discovery of novel genes that are implicated in host-pathogen interactions. Sequencing siRNA upon infection by TMV-Cg, a crucifer infecting strain, revealed the roles of host RdRp1 and RdRp6 in virus-host
interaction and also revealed two predicted host proteins (cleavage and polyadenylation specificity factor, CPSF30 and an unknown protein similar to translocon-associated protein alpha that are targeted by the virus (Qi et al., 2009). Genome-wide transcript profile for systemic symptom development in tobacco plants infected with CMV determined molecular mechanisms of the tobacco symptom development process and also indicated that the plant innate immunity process were significantly enhanced during the transient recovery process (Lu et al., 2012). NGS technologies have been utilized in plant virology to discover and characterize new viruses in different plant host species, including wild species, as well as in different insect vectors without depending on prior knowledge of antibody or sequence of the virus or viroid of interest (Barba et al., 2014). For examples, previously undescribed Cereal yellow dwarf virus (CYDV) genomes were identified in Dactylis glomerata leaves (Pallett et al., 2010). Similarly, complete viral (Sweet potato feathery mottle virus, Sweet potato chlorotic stunt virus) and novel viral (causing sweet potato virus disease) sequences infecting sweet potato were revealed by NGS (Kreuze et al., 2009). Also complete genome sequences were determined for both CBSV and UCBSV strains using NGS (Monger et al., 2010). Differential gene expression studied after Plum pox virus (PPV, sharka disease) infection in peach revealed gene encoding S-adenosylmethionine (SAM), that may be critical for virus infection in addition to many significant alternative splicing events and single nucleotide polymorphisms (SNPs) (Rubio et al., 2014b). Similarly, analysis of differential gene expression in response to TYLCV infection in resistant (R) and susceptible (S) genotypes revealed that NBS-LRR resistance gene (Solyc05g009760.1) is involved in resistance to TYLCV along with a series of genes involved in cell wall reorganization, transcriptional regulation, defence response, ubiquitination, metabolite synthesis and in addition to many R genes (Chen et al., 2013). Similarly, comparative transcriptome studies in contrasting CBSD resistant cassava cultivars revealed overexpression of genes encoding NAC proteins in resistant cultivar Kaleso (Maruthi et al., 2014). Thus the new developments in NGS along with advances in bioinformatics has greatly enhanced our understanding of the complex molecular interactions.

**Conclusion**

This review serves as a synthesis of existing knowledge about virus replication and host factors involved, as well as recent developments in, and new technologies for identifying natural resistance mechanisms. Understanding the mode of resistance to plant viruses is critical for developing durable virus-resistant crops. The widespread use of NGS and VIGS technologies can greatly augment the identification of new sources of virus resistance. Identified resistance sources can then be brought into other crop varieties and species through induced mutagenesis. Mutagenesis can also be used to generate new sources of resistance, which can be screened for
by technologies such as TILLING. The identification of new resistance genes also opens the
door for next generation breeding technologies (such as CRISPR/Cas9, TALENs and Zinc
Finger Nucleases) for engineering resistance in susceptible varieties. Proven biotechnological
approaches such as RNAi should be adopted in cases where natural resistance sources are
insufficient or unavailable. Biotechnological methods can also be coupled with natural
resistances to confer immunity to multiple virus species. Further work needs to be done to
identify and classify common responses to virus infections. The definition of a clear effect-
response distinction at the transcriptome level would simplify the identification of genuine plant
immune responses versus symptomatic responses to a disease. Identified immune responses
serve as a database of potential sources of virus resistance. Knowledge about how plants
respond to viruses in the context of climate change must also be utilized in existing breeding
and research programs to screen for hardier varieties of plants that retain resistance under low
or high temperatures and changing atmospheric CO$_2$ content. Viruses are a fast and continually
evolving plant pathogen that can break existing resistance sources relatively quickly. This
presents breeders with additional challenges since they must constantly mine the germplasm
for new resistance sources. To meet this challenge, scientists and breeders must harness all
available, emerging technologies for achieving future food security.

**Acknowledgements**

We gratefully acknowledge the support of our work by ETH Zurich. Ravi B. Anjanappa is
supported by the Research Fellow Partnership Programme (ETH North – South Centre). The
authors thank Devang Mehta (ETH Zurich), Wilfred Elegba (ETH Zurich), Glen Uhrig (ETH
Zurich) and Vamsi J Nalam (Indiana University-Purdue University Fort Wayne) for useful
comments and editing English on the review.
References


Huang, L., Ren, Q., Sun, Y., Ye, L., Cao, H., and Ge, F. 2012b. Lower incidence and severity of tomato virus in elevated CO2 is accompanied by modulated plant induced defence in tomato. Plant Biology 14:905-913.


Hwang, J., Oh, C.-S., and Kang, B.-C. 2013. Translation elongation factor 1B (eEF1B) is an essential host factor for Tobacco mosaic virus infection in plants. Virology 439:105-114.


Nicaise, V., and Revers, F. 2001. Plant resistance to infection by viruses. eLS.


Okogbenin, E., Porto, M.C.M., Egesi, C., Mba, C., Espinosa, E., Santos, L.G., Ospina, C., Marín, J., Barrera, E., Gutiérrez, J., Ekanayake, L., Iglesias, C., and Fregene, M.A. 2007. Marker-Assisted Introgression of Resistance to Cassava Mosaic Disease into Latin American Germplasm for the Genetic Improvement of Cassava in Africa All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher. Crop Sci. 47:1895-1904.


Revers, F., and Nicaise, V. 2001. Plant resistance to infection by viruses. eLS.


Ueki, S., and Citovsky, V. 2011. To Gate, or Not to Gate: Regulatory mechanisms for intercellular protein transport and virus movement in plants. Molecular Plant 4:782-793.


Characterization of brown streak virus-resistant cassava

(Submitted to Molecular Plant-Microbe Interactions)

Ravi B. Anjanappa¹, M. N. Maruthi², Edward Kanju³, Wilhelm Gruissem¹ and
Hervé Vanderschuren¹,*

¹ Department of Biology, Plant Biotechnology, ETH Zurich, Zurich, Switzerland
² Natural Resources Institute (NRI), University of Greenwich, Chatham Maritime, Kent ME4 4TB, United Kingdom
³ International Institute of Tropical Agriculture (IITA), P.O Box 34441, Dar es Salaam, Tanzania

* Corresponding author
Hervé Vanderschuren
Email: hvanderschuren@ethz.ch
Phone: Office: +41 44 632 87 25
Chapter 2

Characterization of brown streak virus-resistant cassava

Abstract

Cassava brown streak disease (CBSD) has become a major constraint to cassava production in East and Central Africa. A major obstacle in screening cassava germplasm for CBSD resistance is the absence of virus infectious clones and the lack of an efficient transmission method. A stringent infection method based on top cleft grafting combined with precise virus titers quantitation was utilized to screen cassava cultivars and elite breeding lines. When inoculated with mixed infections of *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV), all susceptible control plants developed CBSD foliar symptoms at 4 weeks. KBH 2006/18 and KBH 2006/26 elite breeding lines remained symptom-free during a prolonged period of virus inoculation. Virus titers quantitation showed that CBSV and UCBSV could not replicate to detectable levels in resistant elite breeding lines. CBSD resistance in elite breeding lines was not altered by the co-inoculation of CBSV, UCBSV and cassava geminiviruses. However, analysis of virus titers showed that both CBSV and UCBSV move at undetectable levels in scions from KBH 2006/18 and transgenic CBSD resistant 60444 (60444-Hp 9 line), suggesting that systemic movement of CBSVs does not differ significantly in the susceptible and resistant varieties characterized in the present study. Our molecular characterization of CBSD resistance in cassava offers a robust virus–host system to further investigate the molecular determinants of CBSD resistance.

Introduction

Cassava is considered as an important food security crop in Africa due to its relative good performance under difficult growing conditions as well as its in-ground storability that allows progressive harvest (Fermont *et al.*, 2010). Despite its superior agronomic performance under adverse conditions cassava production is severely constrained by viral diseases in Africa. Cassava is affected by at least 20 different viral diseases, of which cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are the most widespread and severe diseases on the African continent (Patil and Fauquet, 2009a). While CMD has long been viewed as the main constraint to cassava production in Africa (Seif, 1982; OtimNape *et al.*, 1997), CBSD has recently become a major threat due to its re-emergence in East Africa and its rapid spread into new geographical areas in Central Africa (Alicai *et al.*, 2007; Bigirimana *et al.*, 2011; Mulimbi *et al.*, 2012).
Viral disease management in cassava fields has mostly relied on the identification of existing sources of virus resistance, the introgression of virus resistance into farmer-preferred cultivars and the deployment of virus-resistant varieties in the field (Thresh and Cooter, 2005). This strategy has been particularly important for mitigating the impact of CMD particularly in the CMD pandemic regions of Africa (Legg et al., 2006). Even though sources of CBSD resistance were initially identified in the breeding programmes performed in the 1940’s and 1950’s at Amani Research Station, research and breeding efforts were focused on controlling CMD, which was more widely distributed (Hillocks et al., 2001). The CMD-resistant cultivars and landraces deployed in CMD-affected regions were not originally tested for resistance against CBSD. They later appeared to be susceptible to CBSD and this might have facilitated the spread of CBSD in eastern and central Africa during the last decades. Thus, urgent measures for identifying, characterizing and preserving CBSD resistance in cassava germplasm are required as part of long-term and sustainable disease management strategies. Adequate characterization of CBSD resistance today can be performed based on technical advances allowing precise assessment of virus titers in host plants as well as a better knowledge of the diversity of the CBSD causal agents.

Cassava brown streak virus (CBSV) was confirmed to be the causal agent of CBSD over half a century ago (Lister, 1959). However it is only recently that CBSV was taxonomically grouped into the genus Ipomovirus (family Potyviridae) (Monger et al., 2001) and that its full genome was released (Mbanzibwa et al., 2009a). Sequencing efforts of CBSD infected cassava samples from different geographical regions have subsequently allowed distinguishing the two viral species causing CBSD (Mbanzibwa et al., 2009; Monger et al., 2010; Winter et al., 2010) that are now referred to as Cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV). Despite their initial identification from distinct geographical zones, co-occurrence of CBSV and UCBSV in both zones has been reported (Legg et al., 2011; Mbanzibwa et al., 2011a). Under controlled conditions CBSV and UCBSV also appear to differ in virulence on cassava hosts (Winter et al., 2010). However, recent outbreaks of CBSD are not uniquely associated with a particular virus species (Legg et al., 2011).

CBSD resistance, originating from Manihot glaziovii, Manihot melanobasis and a few cassava varieties of Brazilian origin, has been initially reported in the cassava improvement programmes at Amani Research Station throughout the 1940’s and 1950’s (Hillocks et al., 2001; Jennings and Iglesias, 2002). However, recent evaluations of the cassava germplasm have only identified CBSD tolerance but no high levels of CBSD resistance (Legg et al., 2011). It remains unclear whether CBSD resistance was lost when breeding activities became focused on controlling CMD during the following decades (Hillocks et al., 2001). A wide range of plant responses to viral pathogens exists (Cooper and Jones, 1983) and inconsistent use of the
terminology could also, at least partially, explain the apparent loss of CBSD resistance. Leaf chlorosis triggered by CBSD appears in a feathery pattern, initially developing along the margins of the secondary and tertiary veins. The leaf symptoms may eventually develop into chlorotic blotches (Hillocks and Jennings, 2003). The level of leaf symptoms does not always correlate with the extent of symptoms observed in the roots (Hillocks et al., 1996a; Mohammed et al., 2012).

The diversity of leaf chlorosis symptoms (Winter et al., 2010; Mohammed et al., 2012) makes a robust assessment of CBSD resistance level by visual observation difficult in cassava. Molecular tools available today for detecting and discriminating CBSV species offer new opportunities to better determine the tolerance and resistance levels to CBSVs as well as investigate cassava response to CBSD (Abarshi et al., 2010; Mbanzibwa et al., 2011a; Moreno et al., 2011; Abarshi et al., 2012a; Tomlinson et al., 2013). Importantly, such tools should also be used to investigate the robustness of CBSD resistance when exposed to mixed infections of CBSVs, as well as in combination with EACMV-Ug, a cassava mosaic geminivirus (CMG), that do occur in several cassava growing regions (Alicai et al., 2007; Legg et al., 2011).

Here we report the screening of selected cassava farmer preferred varieties and elite breeding lines for CBSD resistance using a robust and reproducible inoculation method in combination with selected CBSV and UCBSV isolates. We monitored viral replication in order to determine the level of resistance and performed co-inoculation of CBSV and UCBSV isolates with a severe CMG isolate to test the robustness of CBSD resistance under mixed virus infection conditions. This first comprehensive characterization of CBSD resistance in cassava under controlled conditions opens new perspectives of investigating the molecular mechanism of CBSD resistance and screening for resistant elite breeding lines and farmer-preferred cassava varieties.

**Material and methods**

**Plant material and virus isolates**

Disease-free varieties and elite cassava breeding lines used for the study were obtained from international and national research institutes in Africa as well as lines from ongoing breeding programs (Supplementary Table 1). Cassava plants were grown under greenhouse conditions (27°C, 16h light, 60% humidity). The virus isolates were obtained from field infected cassava plants. Mixed infections were generated through grafting and subsequent propagation of scions carrying mixed infections.
**Virus inoculation method**

Individual plants were assessed for their viral resistance by using the top cleft grafting procedure (Supplementary Figure 1) and side grafting method as previously described (Mohammed *et al.*, 2012). Mock plants consisted of disease-free scions grafted onto disease-free rootstocks. The double grafting procedure involved establishment of the first scion for 14 weeks and subsequent grafting of a second scion following the top cleft grafting procedure. Lower and upper stem samples from 1st scion corresponded to the stem sections 15 – 20 centimeters and 20 – 25 centimeters above the point of grafting, respectively. For detection of CBSVs in the lower and upper stem sections of KBH 2006/18, 60444–Hp 9 and 60444 (Figure 5), and the bark of the stems were removed to have a larger fraction of vascular tissues in the samples.

**Virus titer quantitation**

Total RNA was extracted from leaf samples using a protocol modified from a pine tree RNA extraction protocol (Chang *et al.*, 1993; Moreno *et al.*, 2011). First strand cDNA was synthesized according to the manufacturer instructions (Fermentas) with random hexamer primers mix and 1 µg of total RNA in a final reaction volume of 20 µl. Real-time PCR reactions were performed with the 7500 Fast Real Time PCR System (Applied biosystems, Foster City, CA) using the SDS software. Virus titers were quantitated relative to internal control *MePP2A* as previously described (Moreno *et al.*, 2011). All primers used for virus detection and internal control are listed in Supplementary table 4.

**Results**

**Identification of cassava genotypes resistant to CBSD**

Criteria for the selection of farmer-preferred cassava varieties included their geographical origin as well as reported CMD resistance (Supplementary Table 1). Disease-free scions were grafted onto cassava variety 60444 rootstocks carrying a mixed infection of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02]. The top cleft grafting method (Supplementary Figure 1) previously established in our laboratory (Moreno *et al.*, 2011; Vanderschuren *et al.*, 2012) resulted in a highly successful graft rate (over 90%) and 100% CBSD infection in control 60444 scions. CBSD foliar symptoms appeared in scions from susceptible varieties at 4 weeks after grafting (wag) (Table 1). The symptom severity differed between susceptible varieties. In particular variety MTAI 25 was highly susceptible with early CBSD symptoms followed by dieback of the scions, an observation previously made in the field on highly susceptible cassava varieties (Hillocks *et al.*, 1996; Hillocks *et al.*, 2001).
Table 1: Summary table of the CBSD resistance screening on selected cassava varieties and elite breeding lines

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Accession</th>
<th>N° of symptomatic scions / N° of grafts (% of infection)</th>
<th>CBSD symptom appearance (wag)</th>
<th>CBSD foliar symptom severity *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60444</td>
<td>3/3 (100%)</td>
<td>4 – 6</td>
<td>+++</td>
</tr>
<tr>
<td>2</td>
<td>KBH 2002/363</td>
<td>3/3 (100%)</td>
<td>6 – 8</td>
<td>++</td>
</tr>
<tr>
<td>3</td>
<td>KBH 2006/12</td>
<td>3/3 (100%)</td>
<td>6 – 8</td>
<td>++</td>
</tr>
<tr>
<td>4</td>
<td>KBH 2006/18</td>
<td>0/3 (0%)</td>
<td>No symptom</td>
<td>No symptom</td>
</tr>
<tr>
<td>5</td>
<td>KBH 2006/26</td>
<td>0/3 (0%)</td>
<td>No symptom</td>
<td>No symptom</td>
</tr>
<tr>
<td>6</td>
<td>TMS 30572</td>
<td>2/2 (100 %)</td>
<td>4 – 6</td>
<td>+++</td>
</tr>
<tr>
<td>7</td>
<td>TMS 30001</td>
<td>1/3 (33 %)</td>
<td>7 – 9</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>TME 3</td>
<td>4/4 (100 %)</td>
<td>6 – 8</td>
<td>++</td>
</tr>
<tr>
<td>9</td>
<td>TME 7</td>
<td>4/4 (100 %)</td>
<td>6 – 8</td>
<td>++</td>
</tr>
<tr>
<td>10</td>
<td>Albert</td>
<td>3/3 (100 %)</td>
<td>4 – 6</td>
<td>+++</td>
</tr>
<tr>
<td>11</td>
<td>UMUCASS 33</td>
<td>4/4 (100 %)</td>
<td>4 – 6</td>
<td>++</td>
</tr>
<tr>
<td>12</td>
<td>SC 8</td>
<td>3/3 (100%)</td>
<td>4 – 6</td>
<td>+++</td>
</tr>
<tr>
<td>13</td>
<td>Ebwanatereka</td>
<td>3/3 (100%)</td>
<td>4 – 6</td>
<td>+++</td>
</tr>
<tr>
<td>14</td>
<td>M TAI 25</td>
<td>3/3 (100%)</td>
<td>4 – 6</td>
<td>++++</td>
</tr>
</tbody>
</table>

* Visual assessment of CBSD symptoms on first two fully expanded leaves of grafted scions at 8 wag. CBSD symptoms represent up to 10% (+), 10 – 30% (++), 30 – 60% (+++), above 60% (++++) of leaf surface. KBH (Kibaha); TMS (Tropical Manihot species); TME Tropical Manihot esculenta), SC (South China).

Varieties, SC 8, 60444 and TMS 30572 also showed an early onset of CBSD symptoms (Supplementary Figure 2A) but the scions survived during the observation period of 16 weeks. The varieties TME 3, TME 7, UMUCASS 33, KBH 2006/12 and KBH 2002/363 developed mild CBSD symptoms between 4 to 8 wag. Symptom appearance on scions of TMS 30001, a variety that developed few symptoms of restricted distribution when infected with CGMs (Thresh and Cooter, 2005), was either delayed or not observed in all scions and could only be classified as CBSD symptoms at 8 wag. No CBSD symptoms could be detected in scions of KBH 2006/18 and KBH 2006/26 even at 16 wag. This is consistent with unpublished results from the so called Great Lakes Cassava Initiative, in which the two elite breeding lines KBH 2006/18 and 2006/26 were identified as CBSD resistant breeding lines in the field (http://www.docstoc.com/docs/129028752/Summary-Information). However, no further data have been reported on these varieties.
The grafting experiment was repeated with the susceptible 60444 variety and the varieties best performing in the initial virus resistance screen (i.e. TMS 30001, KBH 2006/18 and KBH 2006/26) using rootstocks of AR34.2 with mixed infections of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02]. AR34.2 rootstocks develop stronger CBSD symptoms and accumulate higher levels of CBSV (Moreno et al., 2011). Scions from both KBH 2006/18 and KBH 2006/26 remained symptom free and the delay previously recorded for CBSD symptom development between 60444 and TMS 30001 scions was consistently seen in this experiment. All data from the top cleft grafting experiment on CBSV [TAZ:DES:01] + UCBSV [TAZ:DES:02] infected rootstocks are summarized in Table 1.

At completion of an independent grafting experiment (14 wag), scions from 60444, KBH 2006/18 and KBH 2006/26 were propagated via stem cuttings in the greenhouse. Cuttings displayed a higher establishment rate in KBH 2006/18 (84%) and KBH 2006/26 (92%) rate compared to 60444 (60%) for mixed infection of CBSVs (Supplementary Table 2). All 60444 cuttings that had grown showed CBSD symptoms while all propagated KBH 2006/18 and KBH 2006/26 cuttings remained symptom-free (Supplementary Table 2).

**No or reduced viral replication in CBSD resistant cassava**

We earlier reported a correlation between CBSD symptom severity and viral titers (Moreno et al., 2011). Virus titers were monitored in KBH 2006/18 and KBH 2006/26 as well as 60444 to investigate whether viral replication occurs in scions free of CBSD symptoms. We used quantitative real-time reverse transcription-PCR (RT-qPCR) using a cassava reference gene as internal control to quantify virus titers (Moreno et al., 2011). Quantitation of virus titers in 60444 rootstocks confirmed that scions were exposed to high levels of infection pressure (Supplementary Figure 3). Neither CBSV [TAZ:DES:01] nor UCBSV [TAZ:DES:02] were detectable in KBH 2006/18 and KBH 2006/26 (Figure 1 A&B). Both CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] were replicating in 60444 scions that showed typical CBSD symptoms. As previously observed (Vanderschuren et al., 2012), UCBSV on average accumulated to higher titers than CBSV in 60444 scions.
Figure 1: RT-qPCR quantitation of virus titers in scions from selected accessions grafted on rootstocks carrying mixed CBSV and UCBSV infection. **A.** Detection of CBSV [TAZ-DES-01], Y-axis values are average ratios (3 technical replicates) of CBSV CP / MePP2A. **B.** Detection of UCBSV [TAZ-DES-02], Y-axis values are average ratios (3 technical replicates) of CBSV HAM1h / MePP2A transcripts. Numbers following the accessions indicate the biological replicates.

**Identified CBSD resistance is independent of grafting procedure**

Top cleft grafting on virus-infected rootstocks has proven highly effective for CBSV and UCBSV transmission with 100% infection rates in control plants ((Moreno et al., 2011; Vanderschuren et al., 2012); see also Table 2 in this study). In order to investigate virus replication in the roots, we performed top grafting of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] infected 60444 scions on virus-free rootstocks from a resistant and a susceptible variety. CBSD symptoms appeared in susceptible 60444 rootstocks at regrowth following stem propagation while KBH 2006/18 remained symptom-free. RT-qPCR quantitation of virus titers in tuberous roots of inoculated rootstocks confirmed that both CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] could not replicate in KBH 2006/18 rootstocks (Figure 2). Propagation of the inoculated rootstock plants via stem cuttings resulted in low percentage of cuttings establishment for 60444 compared to KBH 2006/18 (Supplementary Table 2). All established 60444 cuttings developed typical CBSD symptoms while cuttings from KBH 2006/18 remained CBSD symptom-free (Supplementary Table 2).

**CBSD resistance holds against different CBSV and UCBSV isolates**

Resistance of KBH 2006/18 and KBH 2006/26 to CBSD was further tested using UCBSV [UG:Kab4-3:07], a virus isolate from the epidemic area of Kabanyoro, Uganda (Mbanzibwa et al., 2011; Mohammed et al., 2012). Top cleft grafting of 60444 scions displayed typical CBSD symptoms at 4 wag. UCBSV [UG:Kab4-3:07] titers were quantitated in the rootstock plants (Supplementary Figure 4) and in the scions (Supplementary Figure 5A). High levels of
Figure 2: RT-qPCR quantitation of virus titers in roots from selected accessions grafted on rootstocks carrying mixed CBSV and UCBSV infection. A. Detection of CBSV [TAZ-DES-01], Y-axis values are average ratios (3 technical replicates) of CBSV CP / MePP2A. B. Detection of UCBSV [TAZ-DES-02], Y-axis values are average ratios (3 technical replicates) of CBSV HAM1h / MePP2A transcripts. Numbers following the accession indicate the biological replicates.

Table 2: Summary table of the grafting experiments

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Top cleft grafting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accessions</td>
<td>N° of symptomatic scions / N° of grafts ( % of infection)</td>
<td>N° of symptomatic scions / N° of grafts ( % of infection)</td>
<td>N° of symptomatic plants / N° of grafts ( % of infection)</td>
</tr>
<tr>
<td>60444</td>
<td>34/34 (100%)</td>
<td>3/3 (100%)</td>
<td></td>
</tr>
<tr>
<td>KBH 2006/18</td>
<td>0/30 (0%)</td>
<td>0/3 (0%)</td>
<td>0/4 (0%)</td>
</tr>
<tr>
<td>KBH 2006/26</td>
<td>0/30 (0%)</td>
<td>0/3 (0%)</td>
<td>0/4 (0%)</td>
</tr>
<tr>
<td>TMS 30001</td>
<td>11/24 (45%)</td>
<td>0/3 (0%)</td>
<td>3/3 (100%)</td>
</tr>
<tr>
<td>Mock</td>
<td>0/34 (0%)</td>
<td>0/4 (0%)</td>
<td></td>
</tr>
<tr>
<td>Albert</td>
<td></td>
<td>2/2 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

UCBSV [UG:Kab4-3:07] titers could be detected in all susceptible scions but remained undetectable in the scions from KBH 2006/18 and KBH 2006/26 (Supplementary Figure 5A). Additional testing was done according to Mohammed and colleagues (2012) to assess resistance of KBH 2006/18 and KBH 2006/26 to mixed CBSV [MZ:Nam1-1:07] and UCBSV [UG:Kab4-3:07] infections. Accumulation of CBSV [MZ:Nam1-1:07] was detected in the susceptible 60444 plants at 8 wag (Supplementary Figure 5B). No virus replication occurred in the resistant varieties even at 16 wag (Supplementary Figure 5B).
KBH 2006/18 is resistant to mixed CBSVs and CGM infection

Despite the prevalence of CBSD in CMD tolerant cultivars, co-occurrence of CBSD and CMD has been reported in the field (Thresh et al., 1994; Alicai et al., 2007). However, field data do not support synergism between CBSVs and CGMs (Legg et al., 2011). In order to further characterize the CBSD resistant elite breeding lines, we inoculated KBH 2006/18 with single CGM species, ACMV-NOg and EACMV-Ug, by using the top-grafting method. KBH 2006/18 scions remained symptom-free while control 60444 scions displayed CMD symptoms at 2 wag for both viral species. In order to compare KBH 2006/18 performance to a CMD tolerant landrace, we also inoculated TME 7 scions. Contrary to KBH 2006/18 scions that remained symptom-free after CMD infection until the final observation stage at 24 wag, CMD symptoms appeared on the first emerging leaves of TME 7 scions followed by a recovery phenotype typical of CMD tolerant cultivars. Subsequent viral DNA quantitation revealed that ACMV-NOg was detectable in both control 60444 and TME 7 scions while EACMV-Ug viral particles were only detectable in 60444 control scions (Supplementary Figure 6). Both ACMV-NOg and EACMV-Ug could not be detected in KBH 2006/18 scions.

In order to evaluate resistance of KBH 2006/18 line to a mixed CBSD and CMD infection, disease-free KBH 2006/18 and 60444 scions were grafted on 60444 rootstocks carrying two combinations of CBSVs and CMGs: 1) EACMV-Ug + CBSV [TAZ:DES:01] + UCBSV [TAZ:DES:02], 2) EACMV-Ug + UCBSV [UG:Kab4-3:07]. Both CBSD and CMD symptoms appeared in 60444 scions at 3 wag. CMD symptoms were prominent over CBSD symptoms. KBH 2006/18 scions did not show either CMD or CBSD symptoms. Virus detection at 8 wag in inoculated scions revealed that both EACMV-Ug and CBSVs could replicate in 60444 scions (Figure 3). EACMV-Ug and CBSVs titters varied between scions. We also noticed that CBSV [TAZ:DES:01] was not detectable in the 60444 scions grafted on 60444 rootstocks carrying the EACMV-Ug + CBSV [TAZ:DES:01] + UCBSV [TAZ:DES:02] mixed infection (Figure 3). Viruses used in the mixed infections were below detection limits in KBH 2006/18 scions. We therefore conclude that CBSD resistance in the elite breeding line KBH 2006/18 remains stable even when co-inoculated with CBSVs and EACMV-Ug.
Figure 3: Virus detection in KBH 2006/18 and 60444 scions grafted on 60444 rootstocks carrying mixed CBSVs and geminivirus infection at 8 wag. “+” and “-” indicate absence and presence of listed virus isolates in 60444 rootstocks. A. Multiplex PCR with EACMV-AC1 and PP2A primers, B. RT-PCR with CBSDDF2 and CBSDDR primers.

**CBSVs are transmitted through KBH 2006/18 scions**

In order to test if the elite breeding line KBH 2006/18 is restricting viral movement, we performed double-grafting experiments (Figure 4). KBH 2006/18 and 60444 scions (referred to as 1st scions) were grafted on susceptible 60444 rootstocks infected with CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02]. Following graft establishment and stem hardening, uninfected 60444 scions (referred to as 2nd scions) were grafted onto the established first scions. Appearance of CBSD symptoms was monitored in the second scions. When grafted on 60444 first scions, the susceptible second scions developed typical CBSD symptoms as early as 4 wag. When grafted on KBH 2006/18 first scions CBSD symptoms in the uninfected 60444 second scions only appeared at 10 wag. Quantitation of virus titers in stems indicated that CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] accumulated in 60444 first scions (Figure 4). CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] were below detection levels in stems of KBH 2006/18 first scions (Figure 5), but infections of 60444 second scions clearly indicate that CBSVs can persist and move at low levels in this elite breeding line as well.

The detection of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] in the 60444 second scions grafted on 60444 first scions was consistent with symptom development at 4 wag. Despite detection of UCBSV [TAZ:DES:02] at 4 wag in some of the 60444 second scions grafted on KBH 2006/18 first scions, they did not display CBSD symptoms. Eventually, CBSD symptoms were observed on 60444 second scions from all the grafts at 10 wag (Figure 4). The double
grafting experiment clearly indicates that CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] movement occurs in both susceptible and resistant varieties.

### Figure 4: Summary table of the double grafting experiments. “+” and “−” indicate the positive and negative values for each biological replicate (CP/PP2A ratios are indicated below), ND (not detected), NM (not measured).

<table>
<thead>
<tr>
<th>SCIONS</th>
<th>1st scion</th>
<th>2nd scion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBSD symptoms (at 10 wag)</td>
<td>NONE</td>
<td>+++</td>
</tr>
<tr>
<td>CBSD symptoms (at 4 wag)</td>
<td>NONE</td>
<td>++</td>
</tr>
<tr>
<td>detection CBSV[TAZ:DES:01]</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>detection UCBSV[TAZ:DES:02]</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>detection CBSV[TAZ:DES:01] (leaves, 1st scion (14 wag), 2nd scion (4 wag))</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>detection UCBSV[TAZ:DES:02] (leaves, 1st scion (14 wag), 2nd scion (4 wag))</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>detection CBSV[TAZ:DES:01] (stems, 14 wag)</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>detection UCBSV[TAZ:DES:02] (stems, 14 wag)</td>
<td>NM</td>
<td>NM</td>
</tr>
</tbody>
</table>

### ROOTSTOCKS

| detection CBSV[TAZ:DES:01] (leaves, 14 weeks after multiplication) | +++ |
| detection UCBSV[TAZ:DES:02] (leaves, 14 weeks after multiplication) | +++ |

Systemic movement of CBSVs occurs at undetectable levels in KBH 2006/18 and transgenic CBSD resistant 60444–Hp 9 lines

In order to determine whether KBH 2006/18 systemically transmits CBSVs at a lower rate than 60444, we performed an independent grafting experiment with scions from KBH 2006/18 and 60444–Hp 9, a transgenic 60444 line resistant to CBSVs (Vanderschuren et al., 2012). Both KBH 2006/18 and 60444–Hp 9 scions remained free of CBSD symptoms at 14 wag. Control 60444 scions displayed CBSD symptoms at 4-6 wag. CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] could only be detected in the leaves as well as lower and upper parts of the stems of the control 60444 scions (Figure 5). Noticeably, both CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] were detected at similar titers in leaves but CBSV [TAZ:DES:01] was the predominant isolate in the stems 60444 susceptible scion (Figure 5). Our results indicate that viral titers detected in the stems of the control 60444 scions are primarily derived from viral replication occurring in the stem and leaves of the susceptible scion. The delayed onset of CBSD symptoms observed in 60444 second scions grafted on KBH 2006/18 first scion (Figure 5) can be explained by the lower amounts of virions transmitted by KBH 2006/18 first scion compared to 60444 first scion. We could confirm that 60444–Hp 9 scions also transmit CBSVs.
by detecting UCBSV [TAZ:DES:02] at 6 wag in the 60444 susceptible second scions (data not shown). Based on the CBSVs detection methods presented in this study, our results indicate that there are no detectable differences in the systemic transport of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] virions in KBH 2006/18 and 60444 varieties.

Our results also suggest that KBH 2006/18 and 60444-Hp 9 have similar resistance levels to CBSVs since both CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02] remained undetectable in the leaves of KBH 2006/18 and 60444-Hp 9 scions (Figure 5).

**Figure 5:** Detection of CBSV and UCBSV in stems and leaves of KBH 2006/18, 60444-Hp 9 and 60444 scions. Numbers following the variety indicate the biological replicates.

**Discussion**

Natural CBSD resistance is key to control CBSD in African regions where it has become a major constraint to cassava production. This is also critical to minimize the threat of dissemination to regions where CBSD is currently absent. Diagnostics and precise characterization of virus resistance require standard procedures and terminology that need to be commonly accepted to allow selection and utilization of plant material by virologists and plant breeders. The use of terms for plant responses to virus inoculation has long been debated (Cooper and Jones, 1983). Here we used a stringent top cleft grafting method for inoculation of selected cassava cultivars and elite breeding lines with CBSVs. Our virus inoculation method resulted in 100% infection rates in the susceptible 60444 scions in all experiments. Infection
rates obtained by the top grafting method are more consistent and reproducible compared to other inoculation methods reported to date (Maruthi et al., 2005; Mohammed et al., 2012; Ogwok et al., 2012). Because the top grafting method provides a constant virus inoculum from the infected rootstock to the scion, it also allows the assessment of resistance over several weeks of inoculation. Using a mixed CBSV and UCBSV infection we identified two cassava elite breeding lines, KBH 2006/18 and KBH 2006/26 that remained symptom-free even at 16 wag. Cultivar TMS 30001, which was previously reported to be CMD tolerant (Thresh and Cooter, 2005), only developed inconspicuous CBSD symptoms but CBSV accumulated in scions developing disease symptoms. CBSD infected susceptible varieties usually develop a dry brown-black necrotic rot of the tuberous roots. Despite evidence that CBSVs accumulate in symptomatic and non-symptomatic root tissues (Abarshi et al., 2010; Moreno et al., 2011), the role of root organs in CBSV replication and cycle has not yet been elucidated. Studies in other plant-virus systems suggest that virus accumulation is not homogenous in root systems and that primary roots can sustain high level of virus replication (Dalmay et al., 2000; Valentine et al., 2002). Side grafting and top grafting experiments with CBSV-infected 60444 scions on virus-free KBH 2006/18 rootstocks confirmed that the KBH 2006/18 rootstocks are also resistant to CBSVs. Together, our data show that the top grafting method is suitable for identification of CBSD resistance and that resistance against the mixed CBSV – UCBSV infection used in our screen was robust in two elite breeding lines.

CBSV and UCBSV differ in their virulence on cassava cultivars (Winter et al., 2010; Mohammed et al., 2012). We therefore used different combinations of CBSV and UCBSV isolates to assess the stability of the CBSD resistance. KBH 2006/18 and KBH 2006/26 remained symptom-free with all combinations of virus isolates and inoculation methods.

Successful virus disease management of vegetatively propagated crop requires the selection of symptomless cultivars that do not support virus replication and accumulation (van den Bosch et al., 2007). Characterization of plant responses to virus diseases therefore requires molecular quantitation of virus titers in inoculated plants. CBSV and UCBSV were near or below the detection limit in KBH 2006/18 and KBH 2006/26, indicating that these elite breeding lines qualify as resistant (Cooper and Jones, 1983). Comparison of Ct values obtained in RT-qPCR assays of mock controls and inoculated scions of KBH 2006/18 suggests that CBSV and UCBSV do not move into or replicate in leaves of this variety. Our results also demonstrate that co-inoculation of CBSVs with a severe CGM isolate does not break CBSD resistance in resistant elite breeding lines. Moreover, co-inoculation does not affect the ratio of the CBSD virus isolates initially present in the susceptible cassava 60444 line, suggesting that both types of viruses do not interfere with their respective replication mechanisms.
Based on the results of our double-grafting experiment, it appears that KBH 2006/18 can transmit CBSVs. The susceptibility of a plant to virus infection depends on both the ability of the virus to gain access to the phloem long-distance transport (Wang et al., 1999; German-Retana et al., 2000) as well as host factors, such as restricted TEV movement (RTM) proteins (Chisholm et al., 2001; Cosson et al., 2010) that permit or inhibit viral movement in vascular tissues. However, the comparison between KBH 2006/18 and a CBSD resistant 60444 line could not reveal any significant difference in the vascular transport of CBSVs. It suggests that the main constituent of CBSD resistance in KBH 2006/18 is not based on restricted or blocked systemic transport of CBSVs particles. Future characterization of the CBSD resistance in KBH 2006/18 and KBH 2006/26 will require the development of CBSV and UCBSV infectious clones to study their replication in protoplasts. CBSV and UCBSV infectious clones will also be instrumental in further advancing large scale virus resistance screening methods as well as in determining viral mutations that can overcome the CBSD resistance reported in the present study.

The identification and characterization of elite breeding lines resistant to virulent CBSV isolates represent an important advance for further work of dissecting the molecular mechanisms involved in CBSD resistance. Analysis of cassava varieties with contrasting CBSD resistance using genome sequencing and transcriptome profiling could be particularly instrumental in identifying genes and their expression patterns that are key for compatible and incompatible interactions in the CBSV – cassava pathosystem. Furthermore, isolation of CBSV isolates capable of breaking CBSD resistance in KBH 2006/18 and KBH 2006/26 elite breeding lines will also provide further insights into the virulence factors of CBSVs.

Our results highlight the importance of developing a standardized and robust method for characterization of virus resistance to allow rapid screening of cassava varieties that are preferred by farmers and used in breeding programmes. Considering the importance of environmental conditions that affect host responses such as RNA silencing (Szittya et al., 2003) and plasmodesmata aperture during virus infections (Burch-Smith and Zambryski, 2012), standardized inoculation methods, diagnostics and virus resistance assessments would be preferable in large-scale virus screening programmes.

**Acknowledgements**

We gratefully acknowledge the support of our work by ETH Zurich. Ravi B. Anjanappa is supported by the Research Fellow Partnership Programme (ETH North – South Centre). The authors thank Irene Zurkirchen (ETH Zurich) for special care of the cassava plants in the
glasshouse. Simon Bull (ETH Zurich), Adrian Alder (ETH Zurich), Devang Mehta (ETH Zurich) and Joel Kuon (ETH Zurich) are acknowledged for useful comments on the manuscript.

The authors thank Haidari Nawabu (IITA) for collecting plant material in the field.
References


Supplementary Figures

Supplementary Figure 1: Schematic representation of top cleft grafting method

Supplementary Figure 2: Diversity of CBSD symptoms on cassava accessions. A. Leaves from scions inoculated with CBSV [TAZ:DES:01] + UCBSV [TAZ:DES:02]. B. Leaves from scions inoculated with UCBSV [UG:Kab4-3:07]
**Supplementary Figure 3:** RT-qPCR quantitation of virus titers on rootstocks carrying mixed CBSV and UCBSV infection. Detection of CBSV [TAZ:DES:01] and UCBSV [TAZ:DES:02]. Numbers following the variety indicate the biological replicates.

**Supplementary Figure 4:** RT-qPCR quantitation of UCBSV [UG:Kab4-3:07] on rootstocks. Y-axis values are average ratios (3 technical replicates) of CBSV CP / MePP2A transcripts.
Supplementary Figure 5: RT-qPCR quantitation of virus inoculated plant. A. RT-qPCR quantitation of UCBSV [UG:Kab4-3:07] detection in scion inoculated by top grafting method on UCBSV [UG:Kab4-3:07] infected rootstocks. B. RT-qPCR quantitation of CBSV [MZ:Nam1-1:07] detection at 8 and 24 wag in plant inoculated by side grafting method with CBSV [MZ:Nam1-1:07] and UCBSV [UG:Kab4-3:07] infected scions. Y-axis values are average ratios (3 technical replicates) of CBSV CP / MePP2A transcripts.
Supplementary Figure 6: RT-qPCR quantitation of geminivirus titers in scions inoculated by top grafting method on geminivirus infected rootstocks. Y-axis values are average ratios (3 technical replicates) of CGM AC1 / MePP2A gene copies.
**Supplementary Table 1:** Summary table of cassava lines and varieties used in the study

<table>
<thead>
<tr>
<th>Accession</th>
<th>Source</th>
<th>Response to CMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60444</td>
<td>CMD susceptible</td>
</tr>
<tr>
<td>2</td>
<td>KBH 2002/363</td>
<td>CMD resistant (Kanju, 2014)*</td>
</tr>
<tr>
<td>3</td>
<td>KBH 2006/12</td>
<td>CMD resistant (Kanju, 2014)*</td>
</tr>
<tr>
<td>4</td>
<td>KBH 2006/18</td>
<td>CMD resistant (Kanju, 2014)</td>
</tr>
<tr>
<td>5</td>
<td>KBH 2006/26</td>
<td>CMD resistant (Kanju, 2014)</td>
</tr>
<tr>
<td>6</td>
<td>TMS 30572</td>
<td>CMD resistant (Fregene et al., 2000)</td>
</tr>
<tr>
<td>7</td>
<td>TMS 30001</td>
<td>CMD resistant (Hahn et al., 1980a)</td>
</tr>
<tr>
<td>8</td>
<td>TME 3</td>
<td>CMD resistant (Fregene et al., 2000)</td>
</tr>
<tr>
<td>9</td>
<td>TME 7</td>
<td>CMD resistant (Fregene et al., 2000)</td>
</tr>
<tr>
<td>10</td>
<td>Albert</td>
<td>CMD resistant (Hillocks et al., 2001)</td>
</tr>
<tr>
<td>11</td>
<td>UMUCASS 33</td>
<td>CMD resistant (Okogbenin et al., 2013)</td>
</tr>
<tr>
<td>12</td>
<td>SC 8</td>
<td>CMD susceptible (Bi et al., 2010)</td>
</tr>
<tr>
<td>13</td>
<td>Ebwanateraka</td>
<td>CMD susceptible (Otim-Nape et al., 1998)</td>
</tr>
<tr>
<td>14</td>
<td>MTAI 25</td>
<td>Not available</td>
</tr>
</tbody>
</table>

CIAT (International Center for Tropical Agriculture), IITA (International Institute of Tropical Agriculture), MARI (Mikocheni Agricultural Research Institute), NRCRI (National Root Crops Research Institute), SCCB (Shanghai Center for Cassava Biotechnology), KBH (Kibaha); TMS (Tropical Manihot species); TME Tropical Manihot esculenta), SC (South China).


* A few cuttings collected in the field displayed typical CMD symptoms.
**Supplementary Table 2:** Stem propagation of scions inoculated by top grafting method

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N° of established cuttings / N° of planted cuttings (% of symptomatic cuttings)</td>
<td>N° of established cuttings / N° of planted cuttings (% of symptomatic cuttings)</td>
</tr>
<tr>
<td>60444</td>
<td>29 / 48 (100%)</td>
<td>15 / 24 (93%)</td>
</tr>
<tr>
<td>KBH 2006/18</td>
<td>32 / 38 (0%)</td>
<td>17 / 21 (0%)</td>
</tr>
<tr>
<td>KBH 2006/26</td>
<td>38 / 41 (0%)</td>
<td>14 / 17 (0%)</td>
</tr>
<tr>
<td>TMS 30001</td>
<td>Not tested</td>
<td>16 / 23 (0%)</td>
</tr>
<tr>
<td>Mock</td>
<td>28 / 32 (0%)</td>
<td>21 / 24 (0%)</td>
</tr>
</tbody>
</table>

**Supplementary Table 3:** Stem propagation of rootstocks inoculated by top grafting of infected scions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N° of established cuttings / N° of planted cuttings (% of symptomatic cuttings)</td>
<td>N° of established cuttings / N° of planted cuttings (% of symptomatic cuttings)</td>
</tr>
<tr>
<td>60444</td>
<td>3 / 9 (100%)</td>
<td>1 / 10 (100%)</td>
</tr>
<tr>
<td>KBH 2006/18</td>
<td>14 / 14 (0%)</td>
<td>17 / 17 (0%)</td>
</tr>
</tbody>
</table>
**Supplementary Table 4:** List of primers

<table>
<thead>
<tr>
<th>Target sequence</th>
<th>Primer name</th>
<th>Forward</th>
<th>Primer name</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP2A transcript</td>
<td>PP2A-F</td>
<td>5’-TGCAAGGCTCACACTTCATC–3’</td>
<td>PP2A-R</td>
<td>5’-CTGAGCGTAAGCAGAGGAAG–3’</td>
</tr>
<tr>
<td>CBSV CP</td>
<td>CBSV-CP3F</td>
<td>5’-GAAGTTGAGAATGGGGCCATC–3’</td>
<td>CBSV-CP3R</td>
<td>5’-CAGCTCTCCACGATTTCTCATT–3’</td>
</tr>
<tr>
<td>CBSV CP</td>
<td>TAZ01_CP3F</td>
<td>5’-TCAAGTGAAGACGAGAGAGCA–3’</td>
<td>TAZ01_CP3R</td>
<td>5’-ATCATTTCCACCACAGCTC–3’</td>
</tr>
<tr>
<td>UCBSV CP</td>
<td>UCBSV-CPF</td>
<td>5’-CCATTTGAGGCTAGGAGATTGA–3’</td>
<td>UCBSV-CPR</td>
<td>5’-ACTTCCCATCATGTTCTC–3’</td>
</tr>
<tr>
<td>UCBSV HAM1h</td>
<td>UCBSV-HAM1hF</td>
<td>5’-AAGCCTGACGCTTGTTGTTG–3’</td>
<td>UCBSV-HAM1hF</td>
<td>5’-GCCTCTTCTCTCCTGTTCACC–3’</td>
</tr>
<tr>
<td>CBSVs CP</td>
<td>CBSDDDF2</td>
<td>5’-GCTMGAATGCGRTAYAACA–3’</td>
<td>CBSDDR</td>
<td>5’-GGATATGGAGAAGRKCTCC-3</td>
</tr>
<tr>
<td>CBSVs CP</td>
<td>CBSVs-F</td>
<td>5’-CCTCCATCWCATGTATAGACA-3</td>
<td>CBSDDR</td>
<td>5’-GGATATGGAGAAGRKCTCC-3</td>
</tr>
<tr>
<td>ACMV AC1</td>
<td>APA9 qPCR_FP2</td>
<td>5’-CAATTTCACCCCAACATCA–3’</td>
<td>APA9 qPCR_RP2</td>
<td>5’-GCCTAAGCATATTGCTGAT–3’</td>
</tr>
<tr>
<td>EACMV AC1</td>
<td>EACMV_AC1 F</td>
<td>5’-TGCGTCTGATCATACCTTG–3’</td>
<td>EACMV_AC1 R</td>
<td>5’-ATTCAAGGCTAAATCATCGT–3’</td>
</tr>
<tr>
<td>EACMV AC1</td>
<td>EACMV_AC1_qP</td>
<td>5’-GGGTCTGGATTGAGGAGAGTAG–3’</td>
<td>EACMV_AC1_qPCR R1</td>
<td>5’-TTGATGACGTGCACCCCCACTAC–3’</td>
</tr>
<tr>
<td>PP2A genomic</td>
<td>PP2A_g F</td>
<td>5’-CGCTGTGAAATATGGCATC–3’</td>
<td>PP2A_g R</td>
<td>5’-CTGGCTCAAACGAGATC–3’</td>
</tr>
</tbody>
</table>
Transcriptome modulation in susceptible and resistant cassava varieties inoculated with cassava brown streak viruses

Ravi B. Anjanappa¹, Alicja Szabelska², Michal J Okoniewski³, Wilhelm Gruissem¹ and Hervé Vanderschuren¹,*

¹ Department of Biology, Plant Biotechnology, ETH Zurich, Zurich, Switzerland
² Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, 60-637 Poznan, Poland
³ ID Scientific IT Services, ETH Zurich, Zurich, Switzerland

* Corresponding author

Hervé Vanderschuren

Email: hvanderschuren@ethz.ch

Phone: Office: +41 44 632 87 25

Author contributions:

Design of experiment: Hervé Vanderschuren and Ravi B. Anjanappa

RNA-seq data analysis: Alicja Szabelska, Michal J Okoniewski, Hervé Vanderschuren and Ravi B. Anjanappa

Manuscript written: Ravi B. Anjanappa

Edits and corrections: Hervé Vanderschuren and Wilhelm Gruissem
Chapter 3

Transcriptome modulation in susceptible and resistant cassava varieties inoculated with cassava brown streak viruses.

Abstract

Cassava brown streak viruses (CBSVs) cause major economic losses in Africa. In a recent study we characterized two elite breeding lines (i.e. KBH 2006/18 and KBH 2006/26) displaying high resistance against CBSVs. To study early molecular responses, a time course CBSV-infection assay using a top cleft graft method to inoculate the susceptible variety (var.) 60444, the resistant elite breeding lines KBH 2006/18 and KBH 2006/26 as well as the tolerant var. TMS 30001 with a mixed infection of Cassava brown streak virus (CBSV TAZ:DES:01) and Ugandan cassava brown streak virus (UCBSV TAZ:DES:02). At 28 days after grafting (dag), both CBSV TAZ:DES:01 and UCBSV TAZ:DES:02 could be detected in the var. 60444 scions. Comparative transcriptome analysis of virus-inoculated scions from var. 60444 and KBH 2006/18 at 28 dag was performed using Illumina RNA-seq. We identified 853 and 334 differentially expressed genes (DEGs) respectively in var. 60444 and an elite breeding line KBH 2006/18. The de novo assembly of unmapped reads allowed the additional identification of 865 and 515 transcripts in the transcriptomes of respectively var. 60444 KBH 2006/18. Furthermore de novo assembly was instrumental to reconstruct the full genome sequence of CBSV TAZ:DES:01 and the partial genome of UCBSV TAZ:DES:02. Functional characterization of the DEGs revealed that transcripts from the phenylpropanoid pathway and post-transcriptional gene silencing (PTGS) were significantly up-regulated in var. 60444 upon virus inoculation. In the resistant KBH 2006/18, DEGs related to stress were significantly up-regulated and DEGs related to cell wall maintenance were down-regulated. RT-qPCR analysis of selected DEGs in var. 60444 confirmed the regulation of genes encoding the RNA-DEPENDENT RNA POLYMERASE 1 (RDRP1), the NAC transcription factor like 9 (NTL9), the SUPEROXIDE DISMUTASE 2 (SOD2), the NIM interacting protein (NIMIN) 2, the β-1,3-GLUCANASE 3 (BG3) and the YELLOW STRIPE LIKE 3 (YSL3) in virus-inoculated scions from var. 60444. Enzymatic assays confirmed that accumulation of β-1,3-GLUCANASE 3 (BG3) transcripts translates into increased BG3 enzymatic activity. We further showed that the increased BG3 activity in virus-inoculated var. 60444 scions leads to reduced callose detection at the plasmodesmata, suggesting that CBSVs alter the size exclusion limit. RT-qPCR analysis also revealed that RDRP1, NTL9 and BG3 transcripts are highly regulated the scions
from the tolerant var. TMS 30001 which supported virus replication in our inoculation assay, indicating that activation of these genes is part of a generic response to CBSV and UCBSV infections. Comparison of the DEGs from the present study with other potyvirus pathosystems revealed that only few genes are commonly regulated across various pathosystems. Our RNA-Seq highlights the specificity of cassava response against ipomovirus infection.

Introduction

Cassava brown streak disease (CBSD) is one of the most damaging cassava diseases in Africa causing losses of up to 70% in susceptible varieties (Hillocks et al., 2001; Legg et al., 2011). The characteristic symptoms of CBSD includes leaf chlorosis and dry hard necrosis in roots, thus affecting both the quality and quantity of edible storage roots. CBSD is caused by cassava brown streak viruses (CBSVs)—positive ssRNA viruses belonging to the Ipomovirus genus; family Potyviridae. CBSVs consist of a ~ 9 kb genome that encodes a large polyprotein from a single open reading frame which is eventually cleaved into 10 viral proteins (Mbanzibwa et al., 2009; Winter et al., 2010). CBSVs comprise of two phylogenetically distinct species i.e. Cassava brown streak virus (CBSV) and Ugandan Cassava brown streak virus (UCBSV) (Monger et al., 2001b; Winter et al., 2010). Both viral species have been reported to co-occur in the field (Mbanzibwa et al., 2011). CBSV and UCBSV differ in virulence in controlled and field conditions (Winter et al., 2010; Mohammed et al., 2012; Ogwok et al., 2015). CBSVs are transmitted by whiteflies in a semi-persistent manner, though at a very low rate (Maruthi et al., 2005; Mware et al., 2009).

Identification of CBSD resistance in the cassava germplasm and the development of locally adapted resistant cultivars has attained greater importance after the re-emergence and outbreak of CBSD in East Africa (Alicai et al., 2007). The recent spread of CBSD into central Africa (Bigirimana et al., 2011; Mulimbi et al., 2012) indicates that CBSD could soon become the major constraint to cassava production in all Sub-Saharan cassava producing regions. Early breeding efforts initiated in the 1930’s to develop CBSD resistant cassava varieties were based on interspecific crosses between Manihot esculenta and M. glaziovii or M. melanobasis (Hillocks and Jennings, 2003). One of these varieties resulting from the backcross of the initial M. esculenta x M. glaziovii cross, is called Kaleso (formerly known as 46106/27 and genetically identical to the variety Namikonga) and is often considered to be resistant to CBSD (Childs, 1957; Pariyo et al., 2013). However, the variety Kaleso does not exhibit complete resistance as it expresses early CBSD symptoms and also supports replication of both CBSV species (Kulembeka et al., 2012; Pariyo et al., 2013; Maruthi et al., 2014). The parental line of Kaleso, M. glaziovii, also supports replication of CBSVs (Mbanzibwa et al., 2011). Recent field trials to assess CBSD resistance in 11 genotypes from Uganda and Tanzania showed contrasting...
levels of virus replication but none displayed resistance *sensu stricto* (Kaweesi *et al*., 2014). Hence identification of CBSD resistance in the cassava germplasm and wild relatives remains an important objective. We recently reported the characterization of CBSD resistance in elite breeding lines KBH 2006/18 and KBH 2006/26 (Anjanappa *et al*., under review). Noticeably, the two elite breeding lines did not support replication of CBSV and UCBSV isolates using the stringent top-grafting inoculation method (Anjanappa *et al*., under review). The identification of cassava accessions with contrasting levels of CBSV infection provides a useful tool to characterize the molecular determinants underlying CBSV replication in the cassava host and to ultimately understand the mechanisms of CBSD resistance in cassava.

The release of the cassava genome (Prochnik *et al*., 2012) has opened new opportunities to perform large-scale characterization of the cassava proteome and transcriptome (Allie *et al*., 2014; Maruthi *et al*., 2014; Vanderschuren *et al*., 2014). Cassava varieties contrasting in resistance against cassava geminiviruses (CMGs) were recently used in a time course experiment to identify pathways differentially regulated upon infection with *South African cassava mosaic virus* (SACMV) (Allie *et al*., 2014). Similarly, a comparative transcriptome analysis was performed to identify transcripts differentially expressed in Albert and Kaleso varieties (Maruthi *et al*., 2014).

Studying gene modulation during early stages of viral infection can reveal host factors that are required for virus translation, replication and movement, and also potential host defence responses or adaptations of the host cells to a modified metabolic balance, as well as the negative consequences of symptom expression (Maule *et al*., 2000). Depletion of host susceptibility factors at later stages of virus infection have been suggested to help prevent the formation of new replication complexes by inducing a shift from the translation/replication cycle to the encapsidation (Nagy, 2015). Hence, studying early stages of infection helps identifying the factors and pathways preventing the transmission, the establishment and the replication of the virus in the host.

In the present work we established a time course experiment with four cassava genotypes previously reported to support contrasting levels of CBSV and UCBSV replication to select samples at the early phase of virus replication. We subsequently performed a transcriptome analysis of samples from cassava genotypes contrasting for virus replication. Our work provides a thorough characterization of the cassava – CBSV pathosystem and identifies candidate genes to develop resistance strategies against CBSV species in cassava.
Material and Methods

Plant material and virus inoculation

Elite cassava breeding lines KBH 2006/18, KBH 2006/26 and var. TMS 30001 were obtained from IITA and grown along with the susceptible control var. 60444 under greenhouse conditions (27°C, 16h light, and 60% humidity). Infected stem cuttings carrying a mixed infection of TAZ:DES:01 CBSV isolate and TAZ:DES:02 UCBSV isolate were collected in the field in Tanzania and used to inoculate var. 60444 scions via the previously described top grafting method (Moreno et al., 2011). Inoculated var. 60444 scions were propagated in soil and used as rootstock to inoculate cassava KBH 2006/18, var. TMS 30001 and var. 60444 plants via the top grafting method. KBH 2006/18, var. TMS 30001 and var. 60444 plants were also top grafted on virus-free 60444 rootstocks as mock controls. Individual leaves were collected from scions of CBSV-infected and mock plants at 16, 22 and 28 dag.

Real Time-quantitative PCR

RNA was extracted from cassava leaves using a previously reported RNA extraction protocol (Chang et al., 1993; Moreno et al., 2011). One μg of DNaseI-treated total RNA was reverse-transcribed according to the manufacturer instructions (RevertAid First Strand cDNA synthesis kit, Thermo Scientific) with random hexamer primers. RT-qPCR was performed with SYBR green using 7500 Fast Real Time PCR System and analyzed with SDS software (Applied Biosystems, Foster City, CA). Three independent biological replicates were analyzed per sample. The expression level of each sample was calculated using MePP2A (cassava4.1_012251m) as internal control (Moreno et al., 2011). Gene-specific primers were designed using Primer Express 3.0 (Life Technologies). Primers were then blasted against the Phytozome cassava genome v4.1 and selected based on their specificity. Primers used in this study are summarized in Supplementary Table 1.

Illumina RNA-seq and data analysis

Library preparation and sequencing was performed as follows: 5 μg of total RNA was DNase treated as per instructions from manufacturer (RNase-free DNase Kit, Qiagen) and subsequently column purified (RNeasy Plant Mini Kit, Qiagen). The quality was assessed with a Qubit® (1.0) Fluorometer (Life Technologies, California, USA) and Bioanalyzer 2100 (Agilent, Waldbronn, Germany) (Schroeder et al., 2006). The samples with RNA integrity number (RIN) above 6 were used for sequencing. One μg of total RNA was polyA enriched and mRNA libraries were synthesized using TruSeq RNA Sample Prep Kit v2 (Illumina, Inc,
California, USA). Cluster generation was done using 10 pM of pooled normalized libraries on the cBOT with TruSeq PE Cluster Kit v3-cBot-HS (Illumina, Inc, California, USA) and subsequently, Illumina HiSeq 2000 sequencing was performed to generate reads.

Sequence assembly, mapping to cassava genome and identification of differentially expressed genes (DEGs)

Fastq files were aligned to the cassava genome (Phytozome v9) using tophat aligner v1.4.1 (Trapnell et al., 2009) with default parameters. The alignments were counted using rnaSeqMap Bioconductor library (Lesniewska and Okoniewski, 2011) to the gene-level features defined in the Phytozome GTF/GFF3 annotation file (Prochnik et al., 2012). The count tables for all samples were analyzed using DESeq Bioconductor library (Anders and Huber, 2010). The default criteria for differential expression were set to false discovery rate <0.05 (controlled by Benjamini-Hochberg correction (Benjamini and Hochberg, 1995)), minimum 50 reads aligned to the gene and absolute log2 fold change ≥1. Novel transcript discovery was performed with Cufflinks v2 (Trapnell et al., 2012) and the aforementioned criteria were applied to identify DEGs.

de novo assembly and gene annotation

The de novo assembly of unmapped reads was performed using Trinity software package (r2013_08_14) with default parameters except that the kmer coverage was set to 2 (Haas et al., 2013). The de novo assemblies were constructed separately for each cassava variety and were used as a reference genome for yet another round of alignment with tophat and transcript discovery with cufflinks. The assembled sequences were BLASTed against the NCBI nucleotide database using BLASTN from CLC Workbench software v7.5.1 (CLC Bio-Qiagen, Aarhus, Denmark) (http://www.clcbio.com).

Enzymatic assays

β-1,3-glucanase enzymatic assay

Leaf samples from cassava scions (3 biological replicates) were collected and soluble proteins were extracted from 100 mg of leaf powder in 300 µl of extraction buffer (20mM HEPES, pH 8.0, 5mM MgCl₂ and 1 tablet / 50 ml buffer of EDTA-free protease inhibitor) and quantitated with a BCA kit (Pierce BCA Protein Assay kit) as previously described (Owiti et al., 2011). The amount of glucose released upon the activity of β-1,3-glucanase from a known amount of soluble proteins in the presence of laminarin (Laminaria digitata, Sigma) was then measured spectrophotometrically at 540 nm (Ramada et al., 2010; Owiti et al., 2011).
Pectin methylesterase (PME) enzymatic assay

Non-soluble proteins were extracted from leaf samples using non-soluble protein extraction buffer (20 mM HEPES pH 8, 1 M NaCl, 0.1 % CHAPS, and 1 tablet/50 ml buffer of EDTA-free protease inhibitor) following an established procedure (Owiti et al., 2011) and quantitated with a BCA kit (Pierce BCA Protein Assay kit). PME activity was measured following the established standard method (Hagerman and Austin, 1986).

Quantification of Plasmodesmata (Pd)-associated callose

Leaf sections from mock and inoculated scions were prepared according to standard procedure (Guenoune-Gelbart et al., 2008) with modifications. Leaf samples were stained overnight with Aniline blue (Merck, Switzerland). Stained leaf samples were then examined with a Zeiss AxioObserver.Z1 microscope equipped with a Zeiss LSM780 laser module confocal unit. Zeiss LSM 780 laser confocal microscope with a 40X water-immersion objective. Callose accumulation was detected at an excitation of 405 nm and emission between 413 and 563 nm. Callose quantification was performed on callose dots around Pds from 3 individual cells randomly selected on each picture (2048 pixel x 2048 pixel). Fluorescence intensity values were calculated by deducting the minimum (background) value from the maximum grey value (actual) around each plasmodesmata from the digital images using ImageJ 1.47v software. In total, nine measurements from two independent biological replicates were used for callose quantification.

Results

Time course infection of CBSVs in cassava

Fully expanded individual leaves from KBH 2006/18, KBH 2006/26, TMS 30001 and 60444 scions grafted on CBSVs (TAZ:DES:01 and TAZ:DES:02)-infected 60444 rootstocks were collected at 16, 22 and 28 dag. Scions from the aforementioned cassava genotypes did not display CBSD symptoms during the time course experiment. Virus titers in 60444 and KBH 2006/18 scions were monitored using RT-qPCR and primers that detect both the CBSV species (Moreno et al., 2011) on the second emerging leaf of the scion. Analysis of CBSV titers revealed relatively homogenous infection and replication in 60444 scions at 28 dag (Figure 1). Therefore samples from 28 dag were selected for transcriptome analysis. The time course experiment also confirmed that both viral species could not replicate in KBH 2006/18 and in KBH 2006/26 scions as previously reported (Anjanappa et al., submitted). Despite homogenous viral load when using primer pairs amplifying both viral species, analysis of the titers of each
viral species revealed a less homogenous pattern of infection (Supplementary Figure 1). In moderately resistant cultivar, TMS 30001, virus (TAZ:DES:01) titers were negligible at 16 dag and highly variable virus titers were observed at 28 dag (Supplementary Figure 2). Three biological replicates from 60444 and KBH 2006/18 inoculated and mock scion samples collected at 28 dag were used for RNA-seq analysis.

**Figure 1:** RT-qPCR quantitation of virus titers in scions (16, 22 and 28 dag) from susceptible 60444 and resistant KBH 2006/18 grafted on rootstocks carrying mixed CBSV and UCBSV infections. Y-axis values are average ratios (3 technical replicates except for 16 dag) of CBSV CP / MePP2A transcripts. Numbers following the accession indicate biological replicates.

**Transcriptome modulation upon CBSV inoculation in susceptible and resistant cassava**

Global gene expression in 60444 and KBH 2006/18 from CBSV-inoculated scions collected at 28 dag was determined using Illumina HiSeq 2000 sequencing. A total of 963,897,018 reads with 100-bp paired-ends were generated from the 12 samples, ranging from 71 to 100 million reads per sample (Supplementary Table 2). On average, 83.5% of the total reads were successfully mapped to the Phytozome cassava reference genome v4.1. A total of 26,206 unique transcripts were classified as expressed from the 12 samples, representing 85% of the complete cassava transcriptome (Supplementary Table 3). The number of unique transcripts detected in the present study exceeded previous cassava transcriptome RNA-seq studies (Allie et al., 2014; Maruthi et al., 2014). Unmapped reads were used to perform de novo transcript assembly which allowed the identification of 1,868 additional expressed transcripts. A greater number of de
novo transcripts were identified in 60444 samples (865 transcripts) as compared to KBH 2006/18 samples (515 transcripts) (Supplementary Table 4).

A total of 334 genes in KBH 2006/18 and 853 genes in 60444 were significantly modulated at 28 dag reflecting a higher number of modulated genes in a compatible host–virus interaction (Figure 2A and Supplementary Table 3). Over 60% of all differentially expressed genes were up-regulated in variety 60444 while only 35% were up-regulated in KBH 2006/18. Sixty and 28 genes showed a similar regulation pattern, respectively up-regulated and down-regulated, in var. 60444 and KBH 2006/18.

![Figure 2: Differentially expressed genes between KBH 2006/18 and 60444. Significant genes were defined according to basemean ≥ 50, Log2 foldchange ≥1 and FDR ≤ 0.05. (A) Venn diagram to depict specific and common DEGs in resistant KBH 2006/18 and susceptible 60444. (B) The numbers in the columns indicate the number of significant genes.](image)

In 60444, the genes showing the strongest up-regulation encode defence related leucine-rich repeat (LRR) protein kinase family proteins, basic pathogenesis-related protein 1 (PR1) and stress-related (peroxidase superfamily protein, NAC transcription factor-like 9 (NTL9)) and gamete expressed protein 1 as well as three DEGs with unknown functions. In KBH 2006/18, the genes showing the strongest up-regulation included a homolog of carrot EP3-3 chitinase, gibberellin 2-oxidase 8, heat shock proteins (HSP20-like, HSP17.6) as well as lysine decarboxylase family protein and NAC domain containing protein 2 (Supplementary Table 3).

The de novo assembly of reads that were not mapped to the cassava genome allowed the identification and quantitation of transcripts that are not annotated in the cassava genome. Among the de novo assembled transcripts, only two genes encoding a glycine rich protein (GRP) and S-adenosylmethionine synthetase (SAM) were found to be significantly up-regulated in 60444 while none was deregulated in KBH 2006/18 upon infection with CBSVs.
(Supplementary Table 4). In addition full and partial genome sequence of respectively TAZ:DES:01 (CBSV) and TAZ:DES:02 (UCBSV) could be retrieved from the \textit{de novo} assembled sequences (Supplementary Data 1 and 2).

**Functional analysis of DEGs in CBSVs resistant and susceptible cassava**

Cassava transcript identifiers with log\textsubscript{2} fold value changes were used to map DEGs onto pathways using the MapMan software (v3.2.1R2) (Thimm \textit{et al.}, 2004). Of the 35 MapMan bins, DEGs from 60444 and KBH 2006/18 were functionally categorized into 27 and 24 bins, respectively. Both varieties shared a total of 22 MapMan bins, each bin comprising varying proportions of up-regulated and down-regulated genes (Figure 3). The five bins exclusively comprising DEGs from 60444 variety were glycolysis, N-metabolism, polyamine metabolism, nucleotide metabolism and biodegradation of xenobiotics. Each of these bins contained fewer than five DEGs in each bin. Noticeably, the signaling bin had the highest number of DEGs (107 DEGs) in 60444 and of these 86\% were up-regulated. In contrast, the signaling bin only comprised 22 DEGs in KBH 2006/18. The most up-regulated DEGs in the signaling bin included LRRs, one calmodulin-binding family protein, one calcium-transporting ATPase and a guanyl nucleotide binding / signal transducer. Among the common bins between 60444 and KBH 2006/18, the highest number of DEGs were assigned to RNA metabolism, with 83 and 41 DEGs, respectively (Figure 3). However a large fraction (64\%) of KBH 2006/18 DEGs assigned to RNA metabolism bin were down-regulated in contrast with only 29\% of DEGs in 60444.

**Gene Set Enrichment Analysis using MapMan**

Gene set enrichment analysis performed using MapMan identified 11 bins as significantly overrepresented (pvalue <0.05) in the susceptible variety and four bins in the resistant variety. In 60444, the signaling bin was the most significantly overrepresented followed by photosynthesis, transport, protein degradation and post-translational modification, lipid metabolism and stress. In KBH 2006/18, signaling is not significantly overrepresented while the photosystem II, terpenoid metabolism, cell wall major carbohydrate metabolism are the significantly overrepresented bins (Supplementary Table 5).
Figure 3: Functional categorization of differentially regulated genes using MapMan software. Log2 fold changes obtained from DESeq output were used to represent expression changes. The value on x-axis represents the number of genes. The red bars represent up-regulated and the green bars represent the down-regulated genes.

Comparison and validation of RNA-seq data by RT-qPCR

The expression patterns of six genes (four DEGs and two non-DEG) were confirmed by RT-qPCR on the same set of samples (28 dag) that were used for Illumina RNA-Seq analysis. RNA-
seq data and the RT-qPCR validation showed similar trend in gene expression indicating the consistency of the transcriptomic study (Supplementary Figure 3).

We selected another six DEGs to analyze the expression pattern over all three time points (16, 22 and 28 dag): RNA-DEPENDENT RNA POLYMERASE 1 (RDRP1), SUPEROXIDE DISMUTASE 2 (SOD2), NIM INTERACTING PROTEIN 2 (NIMIN2), YELLOW STRIPE LIKE 3 (YSL3), NAC transcription factor like 9 (NTL9) and HOMOLOG OF ANTI-OXIDANT 1 (ATX1). In 60444, RDRP1 and NTL9 gene expression increased from 16 to 28 dag (Figure 4A and 4C). The expression of metal carrier YSL3 and ATX1 was down-regulated in CBSVs-infected 60444 throughout all the three time points (Figure 4E and 4F). Further, the most up-regulated genes RDRP1, NTL9 and BG3 were investigated in the moderately resistant TMS 30001 variety (Figure 5). NTL9 was found to be consistently up-regulated in CBSVs-infected TMS 30001 scions (Figure 5C).
Figure 4: RT-qPCR analysis of six genes in susceptible variety 60444. Each time point consisted of three biological replicates from both non-inoculated (mock) (green bars) and CBSVs-infected (purple bars) samples. *(t-test, P<0.05). Error bars represent the standard deviation for three biological replicates.
Figure 5: RT-qPCR analysis of RNA-seq data in moderately tolerant TMS 30001. Each time-point consisted of three biological replicate from both non-inoculated (mock) (green bars) and CBSVs-infected (Blue bars). Error bars represent the standard deviation for three biological replicates. * (t-test, P<0.05).

CBSVs infection modulates photosynthetic and starch regulatory genes

Thirty-eight photosynthetic genes encoding proteins of the light-harvesting complexes from both photosystem I and II were down-regulated in variety 60444 while only 17 genes were down-regulated in KBH 2006/18. Of the 17 DEGs in KBH 2006/18, 16 DEGs were also down-regulated in 60444. The most down-regulated genes in both 60444 and KBH 2006/18 were cassava4.1_014055m (LHCB2.1; chlorophyll binding) and cassava4.1_017616m (PSAN; calmodulin binding). In 60444, starch biosynthesis genes including starch synthase (cassava4.1_003884m) and glucose-1-phosphate adenyltransferase (cassava4.1_005446m, cassava4.1_005409m) were down-regulated while KBH 2006/18 showed a marked increase in transcripts involved in starch degradation. The gene encoding a major starch degrading enzyme, disproportionating enzyme (DPE1) was found to be up-regulated in both KBH 2006/18 and
Two AAA-ATPase genes (cassava4.1_005875m.g and cassava4.1_023408m.g) displayed a contrasting expression profile with a two fold up-regulation in 60444 and a ten fold down-regulation in KBH 2006/18 (Supplementary Table 3). The down-regulation of photosynthetic genes clearly indicates a decrease in the efficiency of photosynthesis and up-regulation of genes encoding starch degrading enzyme might increase sugars that mediate plant defence responses (Coué’e et al., 2006).

Host factors involved in systemic virus infection display contrasting regulation in susceptible and resistant cassava.

RNA virus translation, replication and movement in plants require host factors. Some of these host factors include PLASMODESMATA LOCATED PROTEIN 1 (PDLPL1), β-1,3-glucanase (BG3), pectin methylesterase (PME) 3, and elongation initiation factor (eIFs).

BG3 (cassava4.1_010983m.g) transcripts were found to be 10 fold up-regulated in inoculated 60444 but down-regulated in KBH 2006/18. However, BG3 transcripts fluctuate over time and we found that they were significantly down-regulated in the susceptible cultivar at 22 dag (Figure 6A). Further analysis of BG3 transcripts in 60444 and KBH 2006/18 scions graft-inoculated with CBSV TAZ:DES:01 and UCBSV TAZ:DES:02 revealed consistent significant differences between susceptible and resistant varieties (Figure 6B). BG3 transcripts in symptomatic leaves, accumulated to high levels in inoculated 60444 scions but remained at low levels in inoculated scions from KBH 2006/18 (Figure 6B). Enzymatic assays demonstrated that increase in BG3 transcripts correlated with increased BG3 activity in inoculated 60444 scions (Figure 6). The mock 60444 non-infected scions and the resistant elite breeding line KBH 2006/18 did not display altered BG activity (Figure 6C). Analysis of BG3 transcript in moderately resistant TMS 30001, revealed a consistent increase (1.8-2.8 fold) at early stages of infection (Figure 4). Investigation of the accumulation of callose in 60444 and KBH 2006/18 mock and virus inoculated scions showed a decrease in callose accumulation in inoculated 60444 scions consistent with the increased BG3 activity (Figure 6E). Glucan synthase-like 4 (GSL04) (cassava4.1_000082m), another gene involved in callose synthesis (Maeda et al., 2014), was found to be up-regulated in 60444. In addition, a novel gene (discovered through de novo assembly of unmapped transcripts) encoding a Glycine-rich protein (GRP) was down-regulated (23 fold) in 60444. A cadmium induced GRP has previously been reported to increase callose content in N.tabacum and induce resistance to turnip vein-clearing tobamovirus (TVCV) (Ueki & Citovsky, 2002).

PDLPL1 (cassava4.1_012804m) transcripts was up-regulated in susceptible 60444 (2.4 fold). RT-qPCR expression analysis revealed that at 16 and 28 dag, PDLPL1 transcripts increased in
the range of 1.3-3.3 fold in inoculated 60444. In KBH 2006/18, no significant regulation was observed both in RNA-Seq data and RT-qPCR analysis (Figure 6D).

Figure 6: Transcript quantitation and enzymatic activities in mock and inoculated (CBSV [TAZ:DES:01] + UCBSV [TAZ-DES-02]) scions. A. RT-qPCR on β-1,3 GLUCANASE transcripts from 60444 (three time points), B. RT-qPCR on β-1,3 GLUCANASE transcripts (independent experiment with symptomatic leaves), C. Enzymatic activity of β-1,3 glucanase, D. RT-qPCR on PDLPI transcripts E. Callose accumulation. * (t-test, P<0.05).

A host enzyme, pectin methylesterase (PME) was previously shown to interact with viral proteins and also involved in systemic infection (Chen et al., 2000; Chen and Citovsky, 2003). RT-qPCR analysis of the expression profile of PECTIN METHYLESTERASE 3 (PME3) (AT3G14310) in inoculated scions from KBH 2006/18 and 60444 varieties revealed that PME3 transcript levels were variable with no significant differences between the control and inoculated plants in both 60444 and KBH 2006/18 (Figure 7A). PME3 enzymatic activity was up to 4-fold different between biological replicates and no significant differences were observed between control and inoculated plants in both susceptible and resistant varieties (Figure 7B). In our RNA-seq experiment, PECTIN METHYLESTERASE INHIBITOR (PMEI), a negative regulator of PME3, was significantly up-regulated in 60444 (Supplementary Table 3).

In the present study, out of the eight eIFs annotated in the cassava genome, only cassava4.1_016601m (eIF(iso)4E) was down-regulated (2.2 fold) in 60444 and cassava4.1_033254m.g (eIF3f) was up-regulated (2.7 fold) while the other eIFs were not
significantly regulated. Comparing the amino acid sequences of eIF4E (cassava4.1_013223m), eIF(iso)4E-1 (cassava4.1_016601m) and eIF(iso)4E-2 (Supplementary Figure 4, Supplementary Data 3). Genes encoding ER resident chaperones such as CRT3 (cassava4.1_008376m) and lumenal binding protein (BiP) (cassava4.1_003144m) were up-regulated in KBH 2006/18 but not modulated in 60444.

**Figure 7:** Transcript quantitation and enzymatic activities in mock and inoculated (CBSV [TAZ:DES:01] + UCBSV [TAZ-DES-02]) scions. A. RT-qPCR on PME3 transcripts B. Enzymatic activity of PME3

**Defense pathways are regulated in response to inoculation with CBSVs.**

Eighteen genes encoding WRKY transcription factors (TFs) were found to be significantly regulated by CBSVs infection in cassava. Of the 18 regulated TFs, 15 were up-regulated and one gene was down-regulated in 60444. In KBH 2006/18, only two WRKY genes, cassava4.1_010539m and cassava4.1_032017m were up- and down-regulated, respectively. In variety 60444, **WRKY 70** (cassava4.1_012154m), a positive regulator of salicylic acid (SA) biosynthesis, was up-regulated (4.3 fold), while it was not modulated in KBH 2006/18. In addition, the genes **JASMONATE RESISTANT 1** (JAR1) (cassava4.1_004446m) and **JASMONATE INSENSITIVE1/MYC2** (JIN1/MYC2) (cassava4.1_024442m, cassava4.1_027174m) involved in jasmonate (JA) signaling (Bari and Jones, 2009) were down-regulated (Figure 8).
RNA-dependent RNA polymerase (RDRP) 1 (cassava 4.1_021449m) is an important component of post-transcriptional gene silencing (PTGS) immune response to virus infection (Voinnet, 2005; He et al., 2010; Yu et al., 2003) and can also act as a susceptible factor (Ying et al., 2010). In our RNA-seq analysis, the gene encoding RDRP1 was up-regulated in 60444 (37.5 fold) and not regulated in KBH 2006/18. Our RT-qPCR analysis across the three time points revealed that the \textit{RDRP1} transcripts were consistently up-regulated in 60444 (Figure 4). The RT-qPCR analysis of \textit{RDRP1} transcripts in moderately resistant TMS 30001 confirmed that \textit{RDRP1} regulation is associated with CBSVs infection (Figure 4). Other components of PTGS including argonaute proteins (AGO2, AGO5), SILENCING DEFECTIVE 5 (SDE5) and SUPPRESSOR OF GENE SILENCING 3 (SGS3) appeared to be significantly up-regulated in variety 60444 (Figure 9, Supplementary Table 3).
Figure 9: Schematic representation of RNA silencing pathway in cassava. Values in red indicate up-regulation. Values represents fold change @FDR <0.05 (> 2.0, considered as DEGs). NS indicate non-significant

Genes encoding for key enzymes involved in the phenylpropanoid and lignin biosynthesis pathways such as phenylalanine ammonia-lyase 1 (PAL) (cassava4.1_034377m), 4-coumarate: coenzyme A ligase (4CL) (cassava4.1_005014m), S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (CCoAOMT1) (cassava4.1_024269m) and cinnamyl alcohol dehydrogenase 9 (CAD) (cassava4.1_010316m) were only up-regulated in the susceptible var. 60444 scions (Figure 10).
**Global gene regulation pattern in CBSVs–cassava pathosystem shares few commonalities with other potyvirus-pathosystems**

We retrieved transcriptome analysis data from four additional natural plant host – potyvirus pathosystems to investigate common patterns of gene regulation upon virus infection. The
analysis included transcriptome data from Arabidopsis – Turnip mosaic virus (TuMV) and Arabidopsis – Tobacco etch virus (TEV)–at17 (Rodrigo et al., 2012), soybean – Soybean mosaic virus (SMV) (Babu et al., 2007) and the recent study on cassava – cassava brown streak virus (Maruthi et al., 2014). In order to allow a comparison between the different species, the Arabidopsis orthologs corresponding to the differentially expressed genes were retrieved for each pathosystem. Each pathosystem had number of significantly expressed genes ranging from 246 genes in the soybean – Soybean mosaic virus pathosystem to 2,391 genes in the Tobacco etch virus–Arabidopsis pathosystem (Supplementary Table 6). Most of the genes were unique to each pathosystem and none of the DEGs was found to be commonly regulated in all five pathosystems. Only 11 DEGs were regulated in the present study as well as in three other pathosystems and 29 DEGs were regulated in three pathosystems including the present study. Nine genes were exclusively regulated in the other pathosystems (Table 1). Among the pathosystems compared, the CBSV-Cassava system (Maruthi et al., 2014) had the most DEGs in common (66 genes) with the present CBSV+UCBSV-Cassava pathosystem (Supplementary Table 7).

Table 1: Comparing different susceptible host-potyvirus pathosystems.

<table>
<thead>
<tr>
<th>TAIR Ids</th>
<th>CBSV +UCBSV-Cassava</th>
<th>CBSV-Cassava</th>
<th>TEV-Arabidopsis</th>
<th>TMV-Arabidopsis</th>
<th>SMV-Soybean</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G53540</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>HSP20-like chaperones superfamily protein</td>
</tr>
<tr>
<td>AT3G09270</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>glutathione S-transferase TAU 8</td>
</tr>
<tr>
<td>AT1G75900</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>GDSL-like Lipase/Acylhydrolase superfamily protein</td>
</tr>
<tr>
<td>AT3G22740</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>homocysteine S-methyltransferase 3</td>
</tr>
<tr>
<td>AT5G48380</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>BAK1-interacting receptor-like kinase 1</td>
</tr>
<tr>
<td>AT1G08830</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>copper/zinc superoxide dismutase 1</td>
</tr>
<tr>
<td>AT2G18680</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>unknown protein</td>
</tr>
<tr>
<td>AT3G18280</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Bifunctional inhibitor/lipid-transfer protein superfamily protein</td>
</tr>
<tr>
<td>AT5G54190</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>protochlorophyllide oxidoreductase A</td>
</tr>
<tr>
<td>AT1G64660</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>methionine gamma-lyase</td>
</tr>
<tr>
<td>AT3G09390</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>metallothionein 2A</td>
</tr>
<tr>
<td>AT4G38620</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>myb domain protein 4</td>
</tr>
<tr>
<td>AT1G75250</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>RAD-like 6</td>
</tr>
<tr>
<td>AT1G70740</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>HSP20-like chaperones superfamily protein</td>
</tr>
<tr>
<td>AT1G77760</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>nitrate reductase 1</td>
</tr>
<tr>
<td>AT2G25625</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>unknown protein</td>
</tr>
<tr>
<td>AT1G80160</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Lactoylglutathione lyase/glyoxalase I family protein</td>
</tr>
<tr>
<td>AT4G34950</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Major facilitator superfamily protein</td>
</tr>
<tr>
<td>AT2G26560</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>phospholipase A 2A</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Expression</td>
<td>Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G19010</td>
<td>+</td>
<td>- 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G12735</td>
<td>+</td>
<td>unknown protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G37370</td>
<td>+</td>
<td>cytochrome P450, family 81, subfamily D, polypeptide 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G52870</td>
<td>+</td>
<td>Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G29420</td>
<td>+</td>
<td>glutathione S-transferase tau 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G68830</td>
<td>+</td>
<td>STT7 homolog STN7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G32280</td>
<td>+</td>
<td>indole-3-acetic acid inducible 29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G23010</td>
<td>+</td>
<td>Cupredoxin superfamily protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G12735</td>
<td>+</td>
<td>NAD(P)-binding Rossmann-fold superfamily protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G22160</td>
<td>+</td>
<td>VQ motif-containing protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G44735</td>
<td>+</td>
<td>PHYTOSULFOKINE PRECURSOR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G53450</td>
<td>+</td>
<td>OBP3-responsive gene 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G56360</td>
<td>+</td>
<td>unknown protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G64310</td>
<td>+</td>
<td>arabinogalactan protein 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G48300</td>
<td>+</td>
<td>unknown protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G74940</td>
<td>+</td>
<td>Protein of unknown function (DUF581)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G26250</td>
<td>+</td>
<td>3-ketoacyl-CoA synthase 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G62720</td>
<td>+</td>
<td>Integral membrane HPP family protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G12520</td>
<td>+</td>
<td>copper chaperone for SOD1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G78380</td>
<td>+</td>
<td>glutathione S-transferase TAU 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G47640</td>
<td>+</td>
<td>basic helix-loop-helix (bHLH) DNA-binding superfamily protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G52930</td>
<td>-</td>
<td>Ribosomal RNA processing Brix domain protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G02610</td>
<td>-</td>
<td>Ribosomal L29 family protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G28900</td>
<td>-</td>
<td>outer plastid envelope protein 16-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G31200</td>
<td>-</td>
<td>actin depolymerizing factor 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G47510</td>
<td>-</td>
<td>unknown protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G33250</td>
<td>-</td>
<td>eukaryotic translation initiation factor 1K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G64460</td>
<td>-</td>
<td>Phosphoglycerate mutase family protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G21750</td>
<td>-</td>
<td>PDI-like 1-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G29290</td>
<td>-</td>
<td>NAD(P)-binding Rossmann-fold superfamily protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(+/-) indicates significantly expressed genes, (-) indicates non-significantly expressed genes, CBSV-Cassava brown streak virus, UCBSV-Ugandan Cassava brown streak virus, TEV-Tobacco etch virus, TMV-Turnip Mosaic virus and SMV-Soybean mosaic virus

**Comparison of DEGs during early and late CBSVs-infection in cassava identifies few commonalities in gene regulation**

A direct comparison of a previous study reporting on transcriptome regulation at the late stage of CBSV infection (Maruthi et al., 2014) revealed that very few genes were regulated in the
same manner at early (pre-symptomatic phase, ~1 month after infection; this study) and late (symptomatic phase, 12 months after infection) stage of infections, in both susceptible and resistant varieties (Supplementary Table 7)

In total 1,417 genes were significantly regulated in both the susceptible varieties (i.e. 60444 and Albert) upon CBSVs infection. (1,417 is the sum of significantly regulated genes in both datasets.) Of these, only 66 genes were found in common across both data sets; 24 of these were up-regulated and 24 were down-regulated in both 60444 and Albert. Eighteen genes were found differently regulated in 60444 and Albert, indicating a differential response of these genes at early and late stages of infection. MapMan analysis of genes similarly regulated across both systems (48 genes) revealed that six of these genes have a role in protein post-translational modifications and degradation. Other classes of genes similarly regulated were photosynthesis related (1 gene), cell wall related (2 genes), biotic stress response (1 gene), RNA regulation of transcription (3 genes) and transport related (5 genes). Seventeen genes were not assigned to any bin. For the genes that were regulated differently at early and late time points, two photosynthesis related genes and two genes related to amino acid metabolism were down-regulated at early stage and up-regulated at late stage of infection. Abiotic stress genes (two temperature related genes) and one calcium and one light signaling gene were all up-regulated at early stage of infection and down-regulated at the late stage.

In the resistant varieties (KBH 2006/18 and Kaleso) 1,522 genes were significantly regulated in total. Of these, only 73 genes were found in common; with 16 up-regulated and 23 down-regulated in both. Thirty-four genes were differently regulated between the two systems, indicating a differential response of these genes at early and late stages of infection. MapMan analysis of genes similarly regulated across both systems (39 genes) revealed that most of these genes, which were assigned to bins, are abiotic stress (heat) related (5 genes) followed by signaling (4 genes) and cell wall degradation (3 genes). For the genes regulated differently in both systems, transport (6 genes), signaling (4 genes) and photosynthesis (3 genes) related genes, were all down-regulated at the early stage and up-regulated at the late stage of infection.

**Discussion**

**Detecting early host response to CBSVs infection**

Transmission and establishment of virus infection by top-grafting is an efficient and well-established method (Moreno et al., 2011). We performed a time course experiment in order to determine the early establishment of CBSVs in cassava and select samples with relatively homogenous virus establishment. We sampled the second emerging leaf from the scion in order to study the first mature organ encountered by the virus for replication. Our time course
experiment also confirmed that CBSVs do not establish and replicate in scions from the elite breeding lines KBH 2006/18 and KBH 2006/26 (Anjanappa et al., submitted). Both viral species fully established in the susceptible scions at 28 dag. Quantitation of virus titers in 60444 and TMS 30001 varieties as scions on 60444 CBSVs-infected rootstocks over a time course (16, 22 and 28 dag) had variable virus titers across growing periods, similar to previous reports in other cassava varieties (Kaweesi et al., 2014). Moreover, the accumulation of CBSVs is higher in leaves on lower position of infected cassava plant (Ogwok et al., 2015). Quantitation of virus titer in KBH 2006/18 and KBH 2006/26 (Additional Figures at the end of the chapter) varieties showed low or undetectable virus titers similar to our previous study (Anjanappa et al., submitted). Susceptible 60444 had consistently higher CBSVs titers as compared to moderately resistant TMS 30001 (Supplementary Figure 1 and 2).

Comparative transcriptome analysis was performed to understand innate immune responses that occur in systemically inoculated resistant and susceptible cassava leaf tissue in response to virus infection. The larger number of DEGs in susceptible scions compared to the resistant scions is in line with previous cassava transcriptomics studies comparing transcriptome regulations upon geminivirus infection in susceptible and tolerant cassava varieties (Allie et al., 2014). However, the number of DEGs detected in the CBSV tolerant variety Kaleso upon CBSVs infection was previously reported to be higher than in the susceptible variety Albert (Maruthi et al., 2014). Importantly, the study by Allie and colleagues (2014) characterizing early stage geminivirus infection (12, 32 and 67 days after inoculation) also found a higher number of DEGs in the susceptible cultivar; contrary to the study by Maruthi and colleagues (2014) who reported a larger number of DEGs in the tolerant variety at a late stage of CBSVs infection. This observation suggests that several biological processes are perturbed due to a successful virus infection, in a susceptible variety, at the early stage (Postnikova and Nemchinov, 2012; Allie et al., 2014).

Identification and regulation of non-annotated transcripts

In addition to the quantitative data on annotated genes, RNA-Seq analysis allows the identification of novel transcripts (Kratz & Carninci, 2014) that are currently not reported in the cassava genome. Following the mapping of Illumina reads, we retrieved the unmapped reads to perform de novo assembly using the Trinity pipeline (Grabherr et al., 2011). The generated contigs were used to generate individual reference genomes to quantitate expression levels of unmapped reads in 60444 and KBH 2006/18. The number of contigs obtained from de novo assembly was larger in 60444 compared to KBH 2006/18, possibly reflecting the large number of unmapped reads from 60444 samples (Table 1). These assembled contigs were not present/annotated in the Phytozome cassava genome v4.1 and therefore represents additional
biological information. Statistical analysis of their expression levels revealed two DEGs in 60444. According to BLAST analysis, the two DEGs correspond to a gene coding for a glycine-rich family protein (GRP) and a S-adenosylmethionine synthetase (SAM). GRPs are cell wall proteins induced by virus infections (Fang et al., 1991; Chang-Jin et al., 2001). A cadmium induced GRP in *N. tabacum* has been shown to reduce the systemic spread of turnip vein-clearing tocamovirus (TVCV) by enhancing callose deposition in the cell (Ueki and Citovsky, 2002,2005). In infected 60444, GRP was down-regulated (23 fold), and this might have contributed to the observed decrease in callose content (Figure 7).

**Comparison with other natural host–potyvirus pathosystems reveals a poorly conserved plant response against potyviruses.**

Comparing significantly expressed genes among five different susceptible natural host-virus pathosystems revealed that no genes were common across all the pathosystems. In a meta-analysis of microarray data on 11 different virus infections in *Arabidopsis*, Postnikova and Nemchinov (2012) showed that number of common genes affected by all viruses was very limited. Similarly, in our study only a few genes were found to be common among three or four of the five pathosystems (Table 2); all the other genes were unique to each virus-host interaction. However, since we used *Arabidopsis* gene orthologs for our comparison, it is possible that some commonalities were lost.

A direct comparison of the present study with a recently reported transcriptomics study of the CBSV – cassava pathosystem (Maruthi et al., 2014) showed that only ORG1 transcript (cassava4.1_003020m), a gene coding for a protein with unknown function (Kang et al., 2003) was down-regulated in both studies, and all four cassava varieties (Supplementary Table 5). As suggested, this small number of common genes might reflects important differences in host components required for virus infection and replication (Rodrigo et al., 2012) as well as in the experimental designs. Comparison studies within the susceptible and resistant pathosystems from the two studies revealed few genes significantly regulated in both. Of these, photosynthesis related genes were down-regulated at the early stage of infection and were up-regulated at late stage of infection. Abiotic stress related genes were only up-regulated in the resistant pathosystem at early stage of infection. A contrasting response was observed in signaling related genes: up-regulated in the susceptible pathosystem while down-regulated in the resistant pathosystem during the early stage of infection. Also the major carbohydrate metabolism genes were only regulated in the resistant pathosystem and not in the susceptible pathosystem in both studies.
Transcriptional regulation of photosynthesis components and starch metabolism

Characteristic features of viral infection in compatible hosts are reduced photosynthesis, decreased starch accumulation and an increase in soluble sugars (Tecsi et al., 1994; Shalitin and Wolf, 2000). Virus infection in plants leads to many metabolic perturbations accompanied by symptoms such as chlorosis and necrosis associated with changes in the chloroplast structure and function (Rahoutei et al., 2000). In the present study, 100 % of all the DEGs associated with photosystems were down-regulated in both 60444 and KBH 2006/18 CBSVs-infection.

Role of host eIFs in the CBSVs – cassava pathosystem

Virus translation requires cap binding proteins, eukaryotic translation initiation factor 4E (eIF4E), eIF(iso)4E, and novel cap binding protein (nCBP) that are provided by the host plant (Lellis et al., 2002). Interaction between the viral genome-linked potyviral protein with eIF(iso)4E is, for example, essential for the translation of the potyviral genomes (Wittmann et al., 1997). In the present study, we only identified one elongation factor isoform eIF(iso)4E displaying down-regulation in susceptible 60444. Based on previous studies investigating potyvirus – plant pathosystems, the regulation of the eukaryotic translation initiation factor 4E (eIF4E) does not appear to be an integrated response against potyvirus infection, despite the importance of this factor for viral translation and replication (Contreras-Paredes et al., 2012). However, multiple mutations in the cap recognition pocket of different eIF4Es have been reported to cause recessive resistance to potyviruses in several species, including lettuce (Nicaise et al., 2003) and pepper (Ruffel et al., 2006). The complete sequence identity between the three different eIF4Es in 60444 and KBH 2006/18 rules out eIF4E mediated resistance as the resistance mechanism employed by the resistant cultivar (Supplementary Figure 4, Supplementary Data 3).

CBSVs-infection alters transcript levels of cell-wall related enzymes in susceptible cassava

Maintenance of plant cell wall integrity is an essential component of the biotic stress response (Hamann, 2012). In 60444, the up-regulation of genes encoding key enzymes in the phenylpropanoid pathway such as phenylalanine ammonia-lyase (PAL), 4-coumarate: coenzyme A ligase (4CL), S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (CCoAOMT1) (cassava4.1_024269m) and cinnamyl alcohol dehydrogenase 9 (CAD) which are involved in lignin synthesis (Pellegrini et al., 1994) indicates a hypersensitive response to virus infection. Furthermore, the accumulation of phenylpropanoid compounds and their derivatives might interfere with cell-to-cell spread of the virus (Massala et al., 1980; La Camera et al., 2004). Further, up-regulation of genes in the phenylpropanoid pathway resulted in accumulation of higher levels of hydroxycinnamic acid
ester /chlorogenic acid (CGA) and enhanced resistance as reported for TMV infection in tobacco (Felton et al., 1999; Shadle et al., 2003; Sade et al., 2015). In contrast, suppression of PAL in tobacco plants produced four fold lower salicylic acid (SA) and the systemically infected leaves lacked PR proteins (Pallas et al., 1996). Up-regulation of PAL, 4CL, CAD and OMT enhanced lignification of cell walls and contributed to the plant defence response (Bhuiyan et al., 2009; Sade et al., 2015).

A host enzyme, pectin methylesterase (PME) was previously shown to interact with viral proteins and also involved in systemic infection (Chen et al., 2000; Chen and Citovsky, 2003). In the RNA-seq analysis, PME expression was not significantly regulated in both 60444 and KBH 2006/18. PME transcript levels (based on qRT-PCR) and enzymatic activity were also not significantly regulated by the virus infection. This seems to indicate that the PME-viral protein interaction, while important in other plant-virus pathosystems, is not significant in CBSV infections in cassava.

**CBSVs-infection alters callose deposition at the plasmodesmata in susceptible cassava**

An important factor for systemic virus infection is the ability of the virus to efficiently move between plant cells (short and long distance) via the plasmodesmata (Pd) channels (Beachy and Heinlein, 2000; Heinlein and Epel, 2004). The deposition of callose at the Pd is known to regulate the trafficking of nutrients, and signal molecules by controlling the size exclusion (SE) limit of the channels. Callose deposition thus plays an important role in many plant development processes and in response to multiple biotic and abiotic stresses. Pd permeability is therefore controlled by callose synthases (CalS) or glucan synthase like proteins (GSL)) that synthesize callose as well as β-1,3-glucanase 3 (BG3) that degrades callose (Verma and Hong, 2001; Chen and Kim, 2009). In our study, we found that BG3 transcripts are up-regulated in the susceptible variety, 60444 and that this up-regulation translates into increased BG3 enzymatic activity. We also found that a callose synthase, GSL4 is up-regulated in 60444. This increase in enzymes with opposing functions in callose deposition can be understood in the context of host-virus interplay. In the susceptible variety 60444, viruses likely increase Pd permeability by up-regulating BG3, a callose degrading enzyme and the plant responds by over-expressing callose synthases to limit Pd permeability (Figure 7). Modulation of the callose-related genes did not occur in the scions from the resistant KBH 2006/18 elite breeding line. Noticeably, the accumulation of callose in CBSVs-inoculated KBH 2006/18 scions was higher than that in non-inoculated plants (Figure 6E). In accordance with our observations, an increase in expression of BG3 transcripts was previously reported in another susceptible cassava cultivar, Albert (Maruthi et al., 2014)
PLASMODESMATA LOCATED PROTEIN 1 (PDLP1a), a type I membrane protein, contributes functionally to symplastic communication as well as potentially modulating cell-to-cell trafficking (Thomas et al., 2008). It has been hypothesized that PDLP1a interacts with viral proteins and assists their movement across the cell membrane and cell wall through the Pd (Amari et al., 2010; Rodriguez et al., 2014). The increase in PDLP1 transcripts in 60444 suggests that CBSVs may exploit PDLP1a as a necessary endogenous factor for movement across the Pd (Figure 7). The regulation of AtGSL04 and SA biosynthesis pathway (Figure 7) might reflect an attempt by the susceptible cassava host to limit Pd permeability by increasing callose deposition, to counteract the induction of BG3 and PDLP1 by CBSVs.

Endoplasmic reticulum stress response

During cellular perturbations, a large number of unfolded proteins accumulate in the endoplasmic reticulum (ER) leading to a condition known as the unfolded protein response (UPR) (Aparicio et al., 2005). As an adaptive mechanism for UPR, plants activate altered transcriptional programs through inositol-requiring enzyme 1 (IRE1) and bZIP17/28/60 and also induce quality control (QC), translational attenuation, ER-associated degradation (ERAD) and ER stress-induced apoptosis. This increase in protein processing mechanisms might also benefit plant RNA virus infection, since many viruses rely on the same mechanisms for generating viral proteins and for replication (Moreno et al., 2012; Wahyu Indra Duwi et al., 2013; Verchot, 2014).

The upregulation of various heat shock genes in KBH 2006/18 as observed in the RNA-seq analysis could be a general response to unfolded or misfolded viral proteins as they accumulate and aggregate in the cytoplasm. CALRETICULIN (CTR) impairs cell-to-cell movement of *Tobacco mosaic virus* (TMV) by redirecting the TMV-movement proteins (MP) to protein degradation rather than to direct the MP to plasmodesmata (Chen et al., 2005). The increased levels of CTR transcripts observed in the resistant KBH 2006/18 upon virus inoculation (Supplementary Table 3) suggest that CBSV resistance could be based, at least partly, on reduced virus movement. BiP, an essential component of the translocation machinery playing a role in retrograde transport across the ER membrane by binding and directing the aberrant proteins for degradation (Gething, 1999), appeared up-regulated in KBH 2006/18.

Immune responses to CBSVs infection

WRKY proteins belong to a large family of transcriptional regulators involved in various developmental and defence responses (Johnson et al., 2002; Eulgem and Somssich, 2007). For instance, WRKY70 is implicated in positively regulating SA-dependent defences and
negatively regulating JA-dependent defence pathways (Jing et al., 2004; Li et al., 2004; Li et al., 2006). Another TF, AtMYC2 that encodes a nuclear localized helix-loop-helix-leucine zipper bHLH-type transcription factor and Ethylene Response Factor1 (ERF1) are antagonistically regulated in JA signaling (Lorenzo et al., 2003; Lorenzo et al., 2004). We observed that AtMYC2 was down-regulated while ERF1 was up-regulated, suggesting that jasmonate (JA)-mediated systemic responses to wounding are suppressed in 60444. These changes in CBSV-infected 60444 along with up-regulation of both WRKY70 and RDRP1 together indicate that SA-dependent defence is activated (Figure 8). Genes encoding NAC domain transcription factor family members were all up-regulated, more specifically NAC transcription factor-like 9 (NTL9) (cassava4.1_015961m, cassava4.1_023870m and cassava4.1_028212m) were the most up-regulated in 60444. NTL9 is a membrane-associated NAC TF induced by osmotic stress and mediates signaling during leaf senescence (Hye-Kyung et al., 2008). The fact that NTL9 is induced by CBSVs infection in both susceptible 60444 (Figure 4C) and moderately resistant TMS 30001 varieties (Figure 5A) and was previous reported in susceptible Albert (Maruthi et al., 2014) suggests that leaf senescence in CBSVs infected plants is mediated through up-regulation NTL 9.

Another universal and major defence response used by plants against viruses is RNA silencing machinery via the production of virus-specific short interfering RNA (siRNA) (Pumplin and Voinnet, 2013). Expression of genes encoding argonaute proteins (AGO2, AGO5), RNA-DEPENDENT RNA POLYMERASE 1 (RDRP1), SILENCING DEFECTIVE 5 (SDE5) and SUPPRESSOR OF GENE SILENCING 3 (SGS3) were up-regulated in 60444 but not in KBH 2006/18. SA levels up-regulate the WRKY70 transcription factor, which in turn regulates SA-responsive genes including PRI (Figure 9) (Li et al., 2004).

The up-regulation of RDRP1 transcripts in susceptible 60444 and moderately resistant TMS 30001 varieties in the presence of CBSVs indicates that RDRP1 might be induced by SA, as shown previously (Xie et al., 2001). SA-dependent defence mechanisms are likely activated in 60444 and this was also evident from the higher expression of RDRP1 (cassava4.1_021449m), WRKY70 and PRI. RDRP1 has a differential role in regulating plant viruses and increases the turnover of viruses but is not essential for RNA silencing (Yu et al., 2003; Rakhshandehroo et al., 2009; Ying et al., 2010). In the present pathosystem, RDRP1 transcripts in the infected 60444 are likely to be induced both by CBSVs infection and also by SA-mediated defence mechanisms, similar to that observed in Tobacco mosaic virus and Potato virus X infections in Arabidopsis (Yu et al., 2003). Up-regulation of the PTGS pathway observed in var. 60444 may not be sufficient for virus resistance and might be actively suppressed by the CBSV-encoded P1 protein, a known silencing suppressor (Mbanzibwa et al., 2009). We did not observe regulation of SA-dependent and PTGS pathways in KBH 2006/18 suggesting that other
mechanisms are likely responsible for the high resistance or the responses of these pathways occurred at the time point earlier than those sampled in the experiment.

Conclusion

Understanding early events during virus-host interaction are essential to determine host innate immune responses that limit systemic infection. We have attempted to study these events in CBSVs-susceptible (60444) and resistant (KBH 2006/18) cassava varieties upon infection by CBSVs, through comparative transcriptome analysis using Illumina® HiSeq. We have identified differences in gene expression in several important pathways, including callose regulation, the SA defence pathway, and the PTGS pathways—which may provide hints regarding the resistance mechanisms employed by both varieties. From our observation that a compatible virus infection results in the detection of a larger number of transcripts, as well as a larger number of up-regulated genes, as compared to resistant plants, we hypothesise that global gene expression is increased in the early stages of a successful virus infection. This fits other observations in cassava and Arabidopsis where successful infections resulted in the detection of a larger number of transcripts and up-regulated genes in susceptible plants. The fact that 60444 and KBH 2006/18 have identical eIF4E proteins (alleles of which are known to confer potyvirus resistance in several other plant species), leads us to conclude that KBH 2006/18 does not employ this mode of resistance to infection by CBSVs. Observations on PD-associated callose content, transcripts governing callose accumulation and BG3 enzymatic activity indicate that in the susceptible variety, there is a complex interplay in regulating callose content between CBSV and the host. Interestingly, the same genes and enzymes are not regulated in the resistant plant upon CBSV inoculation. However, an increase in callose content at the plasmodesmata is observed. This restriction on the movement of CBSVs could possibly play a role in the resistance offered by KBH 2006/18. The data on genes involved in the SA defence pathways leads us to believe that this pathway is preferred by the susceptible variety over the JA defence pathway. However, KBH 2006/18 mediated plant resistance is independent of the SA and JA pathways as indicated by the absence of their regulation. Similarly, while the PTGS pathway is clearly used by the susceptible plant as a CBSV defence mechanism, resistance offered by KBH 2006/18 seems independent of this mechanism. In conclusion, it is surprising that KBH 2006/18 does not display any HR response, and has no regulation in standard virus resistance pathways (SA, JA and PTGS). A possible explanation for this could be that these pathways might be involved only at an even earlier stage of virus infection and that they, together with increased callose content successfully localize and eliminate virus infection at the inoculation site. In order to test this hypothesis, the expression of immune pathway genes would have to be checked at earlier time points, such as during the first
emergence of a scion leaf. Alternatively, if infectious agro-clones of CBSVs were developed, the infection could be monitored at even earlier time-points, for example within hours after infection. Extensive transcriptome analysis at earlier time-points could also reveal if a hitherto unknown mechanism is responsible for this extreme resistance.

Acknowledgements

We gratefully acknowledge the support of our work by ETH Zurich. Ravi B. Anjanappa is supported by the Research Fellow Partnership Programme (RFPP. ETH North – South Centre). The authors thank Irene Zurkirchen (ETH Zurich) for special care of the cassava plants in the glasshouse. Devang Mehta (ETH Zurich) and Emily J. McCallum (ETH Zurich) are acknowledged for useful comments on the manuscript. The authors also thank Asuka Kuwabara (ETH Zurich) for guidance in the quantification of callose.
References


prevent pepper veinal mottle virus infection of pepper. Journal of General Virology. 87: 2089-2098


Supplementary Figure 1: RT-qPCR quantitation of virus titers in scions from susceptible 60444 grafted on rootstocks carrying mixed CBSV and UCBSV infection. Detection of CBSV [TAZ-DES-01] and UCBSV [TAZ-DES-02]. Y-axis values are average ratios (3 technical replicates). dag, days after grafting.
Supplementary Figure 2: RT-qPCR with CBSV (TAZ:DES:01) detection in TMS 30001 scions grafted on 60444 rootstocks carrying mixed CBSVs infection. Error bars represent standard deviation from three technical replicates. DAG, days after grafting.
Supplementary Figure 3: RT-qPCR analysis of six genes from 60444 and KBH 2006/18 scion samples collected at 28 days after grafting from both non-inoculated (mock) (green bars) and CBSV-infected (purple bars). Error bars represent the standard deviation for three biological replicates. *(t-test, P<0.05). Values in the box represent the fold change from RNA-seq data.
Supplementary Figure 4: Comparison of the sequence of eukaryotic translational initiation factor [eIF(iso)4E-2] observed in CBSV-resistant and susceptible cassava varieties.

Additional Figure A. CBSV quantitation on KBH 2006/26 scion at 22 and 28 days after grafting. RT-qPCR primers used detect both CBSV and UCBSV.

B. Detection of CBSVs in TMS 30001 scion at 16, 22 and 28 days after grafting. The primers used differentiate CBSV (TAZ:DES:01) and UCBSV (TAZ:DES:02).
Supplementary Table 1: List of primers used in the study

<table>
<thead>
<tr>
<th>Gene Id</th>
<th>Gene</th>
<th>Primer</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBSVs coat protein</td>
<td>CP</td>
<td>CBSVs_CP and CBSDDR</td>
<td>5'-CCTCCCATCRCATGTATAGACA-3'</td>
<td>5'-GGATATGGAGAAAGRKCTCC-3'</td>
</tr>
<tr>
<td>CBSV-CP3</td>
<td>CP</td>
<td>CP3</td>
<td>5'-GAAGTTGAATTGGGCCATC-3'</td>
<td>5'-CAGCTTCCACGATTICATT-3'</td>
</tr>
<tr>
<td>TAZ-DES:01_CBSV</td>
<td>CP</td>
<td>TAZ01_CP</td>
<td>5'-TTCGACTCAAACAGAAGTATATGG-3'</td>
<td>5'-GCGGTTGAAGCTCTTCCGA-3'</td>
</tr>
<tr>
<td>TAZ-DES:02_UCBSV</td>
<td>HAM1h-like</td>
<td>TAZ02-HAM1h</td>
<td>5'-AAGCCCTACGCGTTGTTGT-3'</td>
<td>5'-GCGTTTCTCTTCTTCTGACC-3'</td>
</tr>
<tr>
<td>cassava4.1_015961m.g /cassava4.1_028212m.g</td>
<td>NTL9</td>
<td>MeNTL9</td>
<td>5'-CAATCGGGCAAAAAGGACAA-3'</td>
<td>5'-ACACCTTTTGGAGCACTTGAT-3'</td>
</tr>
<tr>
<td>cassava4.1_015578m</td>
<td>SOD2</td>
<td>MeSOD2</td>
<td>5'-TCTCTAAAGGCACCTCAAATGT-3'</td>
<td>5'-GCGTATCCATCTTGTTGTCA-3'</td>
</tr>
<tr>
<td>cassava4.1_019789m</td>
<td>NIMIN2</td>
<td>MeNIMIN2</td>
<td>5'-TCGAGAAGAATGCGCCAGAAG-3'</td>
<td>5'-AACCCCTATTATCACCAGAACA-3'</td>
</tr>
<tr>
<td>cassava4.1_020156m</td>
<td>ATX1</td>
<td>MeATX1</td>
<td>5'-CACACACAAACCGTCAAAACA-3'</td>
<td>5'-CACAGCCCCAACAACACCTT-3'</td>
</tr>
<tr>
<td>cassava4.1_010134m /cassava4.1_003249m</td>
<td>YSL3</td>
<td>MeYSL3</td>
<td>5'-CTGCAACTCCTCAACTATTCA-3'</td>
<td>5'-GCCTCCACAGTAAACCA-3'</td>
</tr>
<tr>
<td>cassava4.1_012251m</td>
<td>PP2A</td>
<td>MePP2A</td>
<td>5'-TGCAAGGCTCACAACCTTCAT-3'</td>
<td>5'-CTGAGCGTAAAGCGGGAAG-3'</td>
</tr>
<tr>
<td>cassava4.1_021449m</td>
<td>RDRP1</td>
<td>MeRDRP1</td>
<td>5'-CAAATCAAAAACACCGCTAGA-3'</td>
<td>5'-GCCTCCTCTGCGTTGTAATGG-3'</td>
</tr>
<tr>
<td>cassava4.1_012804m</td>
<td>PDLP1</td>
<td>MePDLP1</td>
<td>5'-AAAGGTGAAGCAGGTCAATTCCAA-3'</td>
<td>5'-CAAGCGACACAGCAAGGAA-3'</td>
</tr>
<tr>
<td>cassava4.1_028920m</td>
<td>TOM1</td>
<td>MeTOM1</td>
<td>5'-ACACCGGCTAGCATGATAT-3'</td>
<td>5'-GACCAGGGCTGAGCATAAA-3'</td>
</tr>
<tr>
<td>cassava4.1_001970m</td>
<td>CDC48 (ATPase, AAA-type)</td>
<td>MeCDC48</td>
<td>5'-TGACAGAAATGGATGCGATGAC-3'</td>
<td>5'-GAAAGCGGAGGCTTCTCATCT-3'</td>
</tr>
<tr>
<td>cassava4.1_008453m</td>
<td>CRT1b</td>
<td>MeCRT1b</td>
<td>5'-AAAGGTGAGCAGGCTCAATCCCCA-3'</td>
<td>5'-CACTTGCACACTAATACCA-3'</td>
</tr>
<tr>
<td>cassava4.1_008376m</td>
<td>CRT3</td>
<td>MeCRT3</td>
<td>5'-CATGGAGGCGCCAAGAAGTAATCT-3'</td>
<td>5'-GACCTTACCAGTCCACA-3'</td>
</tr>
<tr>
<td>cassava4.1_004835m</td>
<td>AAP3</td>
<td>MeAAP3</td>
<td>5'-ATCGCCAACGTCGCCATA-3'</td>
<td>5'-GGAACAGGAAAGTGGAGT-3'</td>
</tr>
<tr>
<td>cassava4.1_001583m</td>
<td>HSP17.6</td>
<td>MeHSP17.6</td>
<td>5'-GGGAAAAGAGGGGTGCAAGTT-3'</td>
<td>5'-TGCAATGGCTTCGCTTTTTT-3'</td>
</tr>
<tr>
<td>cassava4.1_0010983m</td>
<td>BG3</td>
<td>MeBG3 F</td>
<td>5'-GCCGAGGAGCTTGGAGATAT-3'</td>
<td>5'-CAGTTGGTTTCCAGGCTCTT-3'</td>
</tr>
<tr>
<td>cassava4.1_004724m</td>
<td>PME3</td>
<td>MePME3 F</td>
<td>5'-TGACGCCAGCACCACAT-3'</td>
<td>5'-GTAAATGCTATGGTCTGATCT-3'</td>
</tr>
</tbody>
</table>
**Supplementary Table 2.** Summary of RNA-seq reads mapping to cassava reference genome.

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Accession-treatment</th>
<th>Total reads</th>
<th>Phytozome reference genome</th>
<th>Trinity de novo assembly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>% reads mapped</td>
<td>% properly paired</td>
</tr>
<tr>
<td>1</td>
<td>60444_non-infected_1</td>
<td>59,327,915</td>
<td>91.8</td>
<td>84.1</td>
</tr>
<tr>
<td>2</td>
<td>60444_non-infected_2</td>
<td>86,783,140</td>
<td>94.8</td>
<td>85.6</td>
</tr>
<tr>
<td>3</td>
<td>60444_non-infected_3</td>
<td>78,621,381</td>
<td>92.5</td>
<td>84.3</td>
</tr>
<tr>
<td>4</td>
<td>60444_CBSVs-infected_1</td>
<td>80,539,310</td>
<td>89.3</td>
<td>81.0</td>
</tr>
<tr>
<td>5</td>
<td>60444_CBSVs-infected_2</td>
<td>10,020,4172</td>
<td>93.7</td>
<td>86.3</td>
</tr>
<tr>
<td>6</td>
<td>60444_CBSVs-infected_3</td>
<td>93,917,976</td>
<td>93.6</td>
<td>85.0</td>
</tr>
<tr>
<td>7</td>
<td>KBH 2006/18_Non-infected_1</td>
<td>58,497,523</td>
<td>85.1</td>
<td>78.6</td>
</tr>
<tr>
<td>8</td>
<td>KBH 2006/18_Non-infected_2</td>
<td>76,340,939</td>
<td>89.4</td>
<td>82.3</td>
</tr>
<tr>
<td>9</td>
<td>KBH 2006/18_Non-infected_3</td>
<td>71,831,760</td>
<td>92.6</td>
<td>84.2</td>
</tr>
<tr>
<td>10</td>
<td>KBH 2006/18_CBSVs-infected_1</td>
<td>97,920,169</td>
<td>94.4</td>
<td>85.5</td>
</tr>
<tr>
<td>11</td>
<td>KBH 2006/18_CBSVs-infected_2</td>
<td>75,221,225</td>
<td>87.3</td>
<td>80.7</td>
</tr>
<tr>
<td>12</td>
<td>KBH 2006/18_CBSVs-infected_3</td>
<td>84,691,508</td>
<td>93.2</td>
<td>83.6</td>
</tr>
</tbody>
</table>
**Supplementary Data 3**

eIF4E amino acid sequences in 60444 and KBH 2006/18

Sequences were identical in both varieties.

> CASSAVA4.1_013223M (eIF4E)

```
MDTLGQINLDLSPLSLSITYRTESKKQQQAAEKINAKMAAEEPLKSTTEITPNPNLNSNPRAOVDNYDDE

MDTLGQINLDLSPLSLSITYRTESKKQQQAAEKINAKMAAEEPLKSTTEITPNPNLNSNPRAOVDNYDDE
```

> CASSAVA4.1_016601M (eIF(ISO)4E-1)

```
MATETATEGsATATGVEKPLQHKLKRWFTWFDNQSPKQGAAGWTSRVRKVTYFTDFTEEEFWCLY
```

> CASSAVA4.1_016620M (eIF(ISO)4E-2)

```
MASETAlEgTAATATGVEKPLQHKLKRWFTWFDNQSPKQGAAGWTSRVRKVTYFTDFTEEEFWCLY
```

130
Identification and characterization of a resistance-breaking Cassava brown streak virus isolate

Ravi B. Anjanappa¹, M. N. Maruthi², Wilhelm Gruissem¹ and Hervé Vanderschuren¹,*

¹ Department of Biology, Plant Biotechnology, ETH Zurich, Zurich, Switzerland
² Natural Resources Institute (NRI), University of Greenwich, Chatham Maritime, Kent ME4 4TB, United Kingdom

* Corresponding author
Hervé Vanderschuren
Email: hvanderschuren@ethz.ch
Phone: Office: +41 44 632 87 25

Author contributions:
Design the experiment: Ravi B. Anjanappa and Hervé Vanderschuren
Manuscript written: Ravi B. Anjanappa
Edits and corrections: Hervé Vanderschuren, M. N. Maruthi and Wilhelm Gruissem
Chapter 4

Identification and characterization of a resistance-breaking Cassava brown streak virus isolate

Abstract
Cassava brown streak viruses (CBSVs) are (+ve) single stranded (ss) RNA viruses belonging to the genus Ipomovirus, family Potyviridae that cause cassava brown streak disease (CBSD) in cassava. Among select cassava genotypes that were infected with single (UG:Kab:07) and mixed CBSV isolates (TAZ:DES:01 and TAZ:DES:02) in the greenhouse, KBH 2006/18, an elite breeding line was found to be resistant. KBH 2006/18 displayed CBSD-like symptoms in a field trail in Chambezi province of Tanzania. Field-infected KBH 2006/18 stems with CBSD symptoms were propagated under greenhouse conditions and used to successfully inoculate virus-free KBH 2006/18 including the mock controls, confirming its susceptibility to the CBSV (TAZ:Cham:14) isolates. Independent assays with different CBSVs isolates revealed that TAZ:Cham:14 isolate was the only CBSV isolate to infect and induce the foliar symptoms in resistant KBH 2006/18 line, while all the other CBSVs failed to establish in the resistant KBH 2006/18, clearly indicating that TAZ:Cham:14 was a resistance breaking isolate. Further to characterize the resistance breaking isolate, the complete genome sequences of CBSVs (TAZ:DES:01 and TAZ:DES:02) were determined by de novo assembly of unmapped reads from Illumina High-Throughput Sequencing, while the TAZ:Cham:14 genome was sequenced by amplifying cDNA with degenerate primers and Sanger sequencing. The sequence for isolate UG:Kab:07 was retrieved from NCBI Genbank. Genetic diversity analyses based on the polyprotein amino acid sequence, revealed that CBSV TAZ:Cham:14 had a larger polyprotein (2916 AA) than TAZ:DES:01 (2912 AA), TAZ:DES:02 and UG:Kab:07 (both had 2902 AA). The TAZ:Cham:14 isolate had 87% identity with TAZ:DES:01 and 74% identity with TAZ:DES:02 and UG:Kab:07. The sequence coverage only included partial 5’ and 3’ UTRs. As expected the sequence diversity was greater between the two CBSVs species (~74.5% identical) than within CBSV isolates (88% identical). Further studies are required to demonstrate unique substitutions observed in TAZ:Cham:14-VPg proteins could likely alter the interactions with the eukaryotic initiation factors (eIFs) and other host proteins, contributing to its ability to overcoming the resistance of KBH 2006/18 line.

Introduction
Cassava brown streak disease (CBSD) and cassava mosaic disease (CMD) are major threats to cassava (Manihot esculenta) production in Africa (Legg and Raya, 1998; Hillocks and Jennings, 2003; Legg et al., 2011). CBSD causes losses up to 70% in cassava root yield because
the root quality is reduced due to the necrotic inclusions, making it unfit for consumption (Hillocks et al., 2001). In highly susceptible cultivars, stem necrosis results in shoot die back (Hillocks et al., 1996b; Hillocks and Jennings, 2003). CBSD symptoms were first described in detail by Nichols (1950). The disease was initially confined to low attitude <1000 meters above mean sea level (masl) in coastal Africa (Tanzania) but is now present in other regions as well as at higher altitudes (>1000 masl) (Alicai et al., 2007). CBSD is caused by two different virus species, the Cassava brown streak virus (CBSV) and the Ugandan Cassava brown streak virus (UCBSV) which are collectively termed as cassava brown streak viruses (CBSVs). Both are (+) single stranded (ss) RNA viruses of the genus Ipomovirus; family Potyviridae (Mbanzibwa et al., 2009b). The typical CBSD symptoms are highly variable (Hillocks and Jennings, 2003) and manifested on leaves (vein clearing/chlorosis to yellowing of complete leaf), the stem (brown streaks) as well as roots (necrotic inclusions or dry rot) (Hillocks & Thresh, 2000). Also, CBSV does not produce leaf distortion phenotypes and the virus can be present in the plant without obvious symptoms (Hillocks et al., 1996b). CBSVs are transmitted either by the propagation of infected stems (vegetative), grafting (experimentally) and by whitefly (Bemisia tabaci) vectors, although at a very low efficiency (Maruthi et al., 2005).

In addition to CBSVs, the genus Ipomovirus comprises of Sweet potato mild mottle virus (SPMMV), Cucumber vein yellowing virus (CVYV), Squash vein yellowing virus (SqVVY), and probable new entries such as Tomato mild mottle virus (TomMMoV) and an Israeli isolate TomMMoV-IL (Dombrovsky et al., 2014). Similar to other members of Ipomoviruses, CBSVs are non-enveloped, filamentous particles of 650 nm in length and have a genome size of ~9 kb, encoding a large polyprotein that is cleaved to produce 10 mature proteins(Mbanzibwa et al., 2009; Winter et al., 2010). In addition to these proteins, the pretty interesting Potyviridae ORF (P3N-PiPO) protein is created from +2 frame shift in the N-terminal region of the virus P3 protein (Chung et al., 2008). The mature proteins processed from the N terminus of the polyprotein are P1 (serine proteinase) protein, P3 protein, 6-Kilodalton protein 1 (6K1), cylindrical inclusion (CI), 6-Kilodalton protein 2 (6K2), viral genome-linked protein (Nia-VPg), nuclear inclusion a (Nia-Pro viral proteinase), nuclear inclusion b (Nib; replicase), HAM1h-like (pyrophosphatase) and coat protein (CP) (Mbanzibwa et al., 2009; Winter et al., 2010). The P1 protein is cleaved by the P1-serine proteinase (Verchot et al., 1992) while all the other proteins are cleaved by the Nia-proteinase (Carrington and Dougherty, 1987). Among the Ipomoviruses, CBSVs have a relatively small genome size of ~9070 nucleotides (nt) and encode a polyprotein of 2,902-2,916 amino acids (AA) compared to SPMMV (~10,818 nt; 3,456 AA), CVYV (9,734 nt; 3,148 AA), Squash vein yellowing virus (SqVVY; 9,836 nt; 3,172 AA). CBSVs lack a HC-Pro (helper component) but encode the HAM1h-like protein (226 AA) between the Nlb and coat protein sequences. CBSVs also have a smaller P1 protein compared
to other members of the ipomoviruses (Mbanzibwa et al., 2009b). The CBSV genome is smaller than the UCBSV genome but encodes a larger polyprotein (Monger et al., 2001; Monger et al., 2010). Under controlled condition, both CBSVs differ in their virulence (Winter et al., 2010; Mohammed et al., 2012).

Plant virus variations are constantly created by diverse mechanisms, including an error-prone high replication rate (mutation), recombination and re-assortment of the genetic material within and between species. These variations assist viruses in constantly evolving and adapting to host/environmental changes and challenges, although very little is known about the impact of selective pressure on their evolution (Roossinck, 2011). During the continuous evolution of plant viruses, some of the variants among the quasispecies might overcome natural selection (both biotic and abiotic) and also acquire new biological functions that allows them to infect new hosts and overcome host resistant genes (Lacroix et al., 2010). In cassava mosaic geminiviruses (CMGs), re-assortment and recombination events have been found between CMG species (Padidam et al., 1999; Ndunguru et al., 2005), while recombination events appear to occur only within and not between CBSV species (Mbanzibwa et al., 2011).

Plants possess a wide range of both host and non-host resistance mechanisms that protect them against viral diseases. The reported variation among resistance and non-resistance breaking isolates in other viruses points to specific single or multiple mutations in the polyprotein. For example, resistance breaking isolates of Potato virus Y (PVY) encode a glutamate at position 1,948 in the VPg protein while the non-resistance breaking isolates encode a lysine or arginine (Masuta et al., 1999). Similarly, a single amino acid mutation in the cylindrical inclusion (CI) protein of TuMV breaks the dominant resistance in Brassica napus TuRB01 (Jenner 2000). A mutation in the C-terminus of the Lettuce mosaic virus (LMV) CI protein can overcome the recessive mol1 resistance in lettuce (Tavert-Roudet et al., 2012).

In the present study, we have identified contrasting CBSV isolates (TAZ:DES:01 and TAZ:Cham:14) differing in their virulence to infect resistant cassava KBH 20006/18, elite breeding line. Subsequently, characterization of these resistant breaking and non-resistant breaking CBSV isolates are also presented.

**Material and methods**

*Plant material and virus isolates*

CBSV-infected (TAZ:DES:01, TAZ:DES:02 and UG:Kab:07 isolates) cuttings were obtained from Dr. Maruthi Gowda (Natural Resources Institute (NRI), UK) and CBSV- TAZ:Cham:14 infected KBH 2006/18 stem cuttings were collected by Dr. Herve Vanderschuren (ETH Zurich).
from Tanzania. Disease free cassava varieties were obtained from international and national research institutes in Africa as well as lines from ongoing breeding programs. Variety 60444 plants infected with CBSVs (mixed infection of TAZ:DES:01 and TAZ:DES:02 isolates, variety 60444 UCBSV infected (UG:Kab:07 isolate) and KBH 2006/18 infected CBSV (TAZ:Cham:14 isolates) were used as rootstock for top grafting procedure to test the level of resistance to all CBSVs isolates as described by (Moreno et al., 2011). Variety 60444 scions were grafted on to non-infected (mock control) as well as CBSV-infected rootstock.

All the cassava plants were grown at 27°C, 16h light and 60% humidity. Individual plants challenged with the four CBSVs isolates were assessed for symptom development for 14 weeks, and all the plants that displayed symptoms were classified as susceptible and non-symptomatic plants were classified as resistant plants. Only the non-symptomatic plants were tested for the presence of CBSVs. Further, the scions from CBSV-challenged plants were multiplied in the greenhouse to monitor the CBSD symptoms.

**Determination of complete sequences of 3 CBSVs isolates**

*De novo* assembly of reads (that did not map to the cassava genome) from Illumina RNA-seq of infected and non-infected cassava plants was attempted to generate full length sequences of TAZ:DES:01 and TAZ:DES:02. Further, degenerate primers (Supplementary Table 1) were designed on conserved sequences obtained by aligning all published CBSV genomes (Supplementary Table 2). cDNA from CBSV-infected samples was prepared using random (hexamers) primers according to manufacturer’s instructions (RevertAid Reverse Transcriptase (Thermo scientific, USA). Amplicons of ~1500 nt obtained from amplifying cDNA from infected tissue were cloned into pJET 1.2 and Sanger sequenced from both directions (Microsynth AG). The same primers were used to obtain amplicons from CBSV-infected KBH 2006/18 tissues which were similarly cloned and sequenced. Confirmation of the degenerate primer binding regions was done by re-sequencing the primer binding regions with specific primers designed from previous sequencing results. The sequences were submitted to [http://web.expasy.org/translate/](http://web.expasy.org/translate/) to obtain the deduced protein sequences. In addition, full-length CBSV sequences for other isolates including UG:Kab:07 were obtained from NCBI Genbank. Multiple sequence alignments of the deduced AA sequences were done with ClustalW2 ([http://www.ebi.ac.uk/Tools/msa/clustalw2/](http://www.ebi.ac.uk/Tools/msa/clustalw2/)) as well as on CLC genomics 7.5.1 version (CLC Bio-Qiagen, Aarhus, Denmark) ([http://www.clcbio.com](http://www.clcbio.com)) using default settings.
Results

Cassava cultivars challenged with CBSV isolates by graft inoculation

Scions from 60444 and KBH 2006/18 cultivars successfully established on the 60444 CBSV-infected and non-infected rootstock. In the present study, only one (CBSV-TAZ:Cham:14) of the four isolates developed typical CBSD symptoms on KBH 2006/18, an elite breeding line. KBH 2006/18 scions grown on 60444 CBSV-infected rootstocks were non-symptomatic for three isolates (mixed infection of TAZ:DES:01 and TAZ:DES:02 and single infection of UG:Kab:07). CBSV DNA could not be detected in non-symptomatic KBH 2006/18, demonstrating that KBH 2006/18 has high resistance to the three isolates. On the contrary, KBH 2006/18 scions were symptomatic when grafted on (field collected) KBH 2006/18 CBSV-infected (TAZ:Cham:14) rootstocks. Thus only the TAZ:Cham:14 isolate produced clear CBSD symptoms and viral DNA was detected in KBH 2006/18 (Figure 1).

Figure 1: A. Development of typical CBSD leaf symptoms upon challenging KBH 2006/18 with CBSV isolates by top grafting method. The four virus isolates used in the study are indicated parenthesis. B. RT-PCR detection of CBSVs from the KBH 2006/18 scion samples after challenging with CBSVs. Generic primers were used to differentiate CBSV (705 nt) and UCBSV (800 nt).

Sequencing of full-length genomes of CBSVs

Total leaf RNA from 60444 infected with mixed infection TAZ:DES:01 and TAZ:DES:02 was sequenced using Illumina RNA-seq. The fastq files were aligned to published full-length sequences of CBSVs (Supplementary table 2) using tophat aligner v1.4.1 (Trapnell et al., 2009)
with the default parameters. The complete sequences of TAZ:DES:01 and TAZ:DES:02 could not be obtained from the RNA reads, though the some reads mapped to P3, CI, Nb and CP viral genes. All the reads not mapping to the cassava genome v9 (Prochnik et al., 2012) were therefore de novo assembled to generate contigs of varying length. The contigs that were present only in the CBSV-infected samples were analyzed using BLAST-N, which resulted in the assembly of partial TAZ:DES:01 and TAZ:DES:02 genome sequences. The TAZ:DES:01 genome was obtained as a single contig (8,975 nt) with 29 nt missing from the 5’ end compared to the KOR6 isolate genome (GU563327.1). A partial TAZ:DES:02 genome was constructed from two contigs of 4,092 nt and 2,575 nt with a 22 nt overlapping sequence. The TAZ:DES:01 and TAZ:DES:02 sequences were confirmed by Sanger sequencing of amplicons obtained from CBSV-infected leaf tissues. Full length sequences from RNA-seq de novo assembly and Sanger sequencing were aligned and found to be similar. Analysis of the newly assembled genome of TAZ:DES:01 using BLAST-N revealed that the isolate was 92% identical (99% query coverage) to previously published TZ:NAL:07 isolates (HG965221.1) (Beyene et al., direct submission). The assembled TAZ:DES:02 genome was 93% identical (100% query coverage) to the Ug_23 isolate (FN434109.1) (Winter et al., 2010). Similarly conserved genome sequences between CBSV and UCBSV were used to develop degenerate primers and to amplify genomic fragments of TAZ:Cham:14 isolate. Subsequently, TAZ:Cham:14 isolate was found to be 94% identical with 100% query coverage to previously published Tanzanian strain (GQ329864.1) (Monger et al., 2010).

Phylogenetic tree derived based on deduced AA (amino acid) sequence of the polyprotein from these 13 CBSVs including the 3 new isolates, TAZ:DES:01, TAZ:DES:02 TAZ:Cham:14 clearly formed into 2 distinct species specific clades (Figure 2). The TAZ:Cham:14 and TAZ:DES:01 isolates were placed in the CBSV clade (Groups I and II) and TAZ:DES:02 was placed in the UCBSVs clade (Group III) (Figure 2). Groups, upon analysis, were also found to have different numbers of amino acids: Group I had 2916 AA, Group II had 2912 AA and Group III had 2902 AA. Comparison of complete polyprotein sequence indicated that the two distinct species of CBSVs (CBSV and UCBSV) shared only ~74% identity, while the identities ranged between 92–99% among the UCBSVs isolates and 87 – 98% identity among the CBSV isolates (Supplementary Table 3).
Figure 2: Phylogenetic tree created based on the deduced AA of cassava brown streak viruses. The isolates sequenced in this study are indicated in dark green box. The tree was constructed by Neighbor joining method and was based on multiple sequence alignment using Clustal Omega. Bootstrap values (%) based on 1000 replication are given at each node. Scale bar shows amino acid substitutions per site.

Each of the viral proteins has specific functions in replication, movement (long or short distance) as well as interaction with the host proteins either as individual proteins or in complex with other host and viral proteins (Table 1). Comparison between isolates TAZ:Cham:14 and TAZ:DES:01 revealed that P1 protein was more variable (68.9% identity) followed by P3N-PiPO (69.2% identity) while 6K1 protein was identity between the CBSV isolates. Similarly comparing the CBSV (TAZ:Cham:14 and TAZ:DES:01) and UCBSV (TAZ:DES:02) isolates revealed that both the species have only 74% identity. The lowest identity was found to be in CBSV-HAM1h-like protein (45.1-46.0%) followed by P3N-PiPO proteins (60.5 -61.4%), while the highest identity was found for the 6K1 (88%) and CI proteins (84%). Further, the genetic diversity at the remaining individual protein level was also high between the isolates (Table 2).
Table 1: Description of *Cassava brown streak virus* (CBSV) proteins and their functions based on other potyviruses.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Properties (no. of AA, molecular weight, domains)</th>
<th>Functions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>358-362 AA; 42 kDa; H-7x-D-34x-S; basic IxKA domain (L-10x-D) and Zinc finger domain (C-2x-C-13x-C-2x-C); serine residue (GWSG)</td>
<td>proteolytic (serine proteinase) activity; RNA silencing suppression activity</td>
<td>(Valli <em>et al</em>., 2007; Valli <em>et al</em>., 2008; Mbanzibwa <em>et al</em>., 2009b)</td>
</tr>
<tr>
<td>P3</td>
<td>294 AA; 35 kDa</td>
<td>host determinants; Replication</td>
<td>(Urcuqui-Inchima <em>et al</em>., 2001; Hjulsager <em>et al</em>., 2006)</td>
</tr>
<tr>
<td>P3N-PiPO</td>
<td>77 AA; 9 kDa; (T/C)G1 A1-6 domain; +2 frameshifting</td>
<td>Cell-cell movement</td>
<td>(Vijayapalani <em>et al</em>., 2012)</td>
</tr>
<tr>
<td>6K1</td>
<td>52 AA; 6 kDa</td>
<td>Cell-cell movement; Replication</td>
<td>(Wei <em>et al</em>., 2010; Abegci <em>et al</em>., 2013)</td>
</tr>
<tr>
<td>CI</td>
<td>630 AA; 71 kDa; helicase domain; NTP binding (GGVGGSKST)</td>
<td>Replication, cell-to-cell movement and long-distance movements</td>
<td>(Kadaré and Haenni, 1997; Rodriguez-Cerezo <em>et al</em>., 1997; Sorel <em>et al</em>., 2014)</td>
</tr>
<tr>
<td>6K2</td>
<td>52 AA; 6 kDa</td>
<td>Replication, cell-to-cell movement and long-distance movements; anchor viral replication complex to host membranes</td>
<td>(Restrepo-Hartwig and Carrington, 1994; Urcuqui-Inchima <em>et al</em>., 2001; Spetz and Valkonen, 2004)</td>
</tr>
<tr>
<td>N1α-VPg</td>
<td>185 -186 AA; 22 kDa</td>
<td>Binding translation initiation factor eIF4E; enhances viral translation; cell-to-cell movement</td>
<td>(Schaad <em>et al</em>., 1996; Roudet-Tavert <em>et al</em>., 2007; Eskelin <em>et al</em>., 2011)</td>
</tr>
<tr>
<td>N1α-Pro</td>
<td>234 AA; 26 kDa; H-34x-D-71x-GDCG (His-34x-Asp-71x-Gly-Asp-Cys-Gly)</td>
<td>proteolytic activity; replication</td>
<td>(Dougherty and Carrington, 1988; Carrington <em>et al</em>., 1993)</td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
<td>Function</td>
<td>References</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>N1b</td>
<td>502 AA; 58 kDa; SG-3x-T-3x-NT-28x-GDD domain; RNA-dependent RNA polymerase</td>
<td>Replication</td>
<td>(Hong et al., 1995; Hong and Hunt, 1996)</td>
</tr>
<tr>
<td>HAM1h</td>
<td>226 AA; 25 kDa</td>
<td>Reduction of the mutation rate of viral RNA</td>
<td>(Mbanzibwa et al., 2009b)</td>
</tr>
<tr>
<td>CP</td>
<td>367-378 AA; 42 kDa</td>
<td>Replication, cell-cell movement, genome encapsidation and long-distance movement</td>
<td>(Mahajan et al., 1996; Rodriguez-Cerezo et al., 1997; Urcuqui-Inchima et al., 2001)</td>
</tr>
</tbody>
</table>
Table 2: Number of amino acids encoded in each protein and their sequence identity among the three CBSVs isolates.

<table>
<thead>
<tr>
<th>Protein</th>
<th>No. of AA (amino acids)</th>
<th>% identity (based on AA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAZ:DE S:01</td>
<td>TAZ:Cham:14</td>
</tr>
<tr>
<td>P1</td>
<td>358</td>
<td>362</td>
</tr>
<tr>
<td>P3</td>
<td>294</td>
<td>294</td>
</tr>
<tr>
<td>6KI</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>CI</td>
<td>630</td>
<td>630</td>
</tr>
<tr>
<td>6K2</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>VPg</td>
<td>186</td>
<td>186</td>
</tr>
<tr>
<td>NIA</td>
<td>234</td>
<td>234</td>
</tr>
<tr>
<td>NIB</td>
<td>502</td>
<td>502</td>
</tr>
<tr>
<td>HAM1h</td>
<td>226</td>
<td>226</td>
</tr>
<tr>
<td>CP</td>
<td>378</td>
<td>378</td>
</tr>
<tr>
<td>P3N-PiPO</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td>Polyprotein</td>
<td>2912</td>
<td>2916</td>
</tr>
</tbody>
</table>
In summary, comparison between CBSV and UCBSV clearly indicates that the HAM1-like, P3N-PiPO, P1, P3, VPg, 6K2 and coat protein, in the same order are the most divergent proteins. The most divergent protein sequence within the two CBSV isolates was P1 protein followed by P3N-PiPO, P3, VPg, 6K2 and coat protein. Interestingly, 6K1 protein, which is the smallest protein of CBSVs, is 100% identical between TAZ:DES:01 and TAZ:Cham:14. The largest protein of CBSV, CI (630 AA) also shows high similarity between the isolates (95.5% identity) but low (84.7% identity) between the CBSV and UCBSV strains. Similarly another large protein, NIb (502 AA) shows high similarity between the isolates (91.4% identity) and low similarity between the species (~82% identity) (Table 2).

**Unique amino acid changes observed in the resistance breaking TAZ:Cham:14 isolate**

Amino acid sequences of all the four isolates (2 CBSV and 2 UCBSV isolates) tested against the resistant line KBH 2006/18 were aligned and compared to identify unique changes in the resistance breaking TAZ:Cham:14 isolate. AA substitutions were considered as unique when the same AA was conserved in TAZ:DES:01, TAZ:DES:02 and UG:Kab:07, and if a change in the corresponding AA was observed in the TAZ:Cham:14 isolate. Based on the pairwise comparisons of the deduced AA, TAZ:Cham:14 had 80 unique AA substitutions in its polyprotein. Among the proteins, P1 protein, known to be involved in suppressing silencing was found to be highly diverse (20 AA), followed by P3 protein (15 AA) and CI protein (13 AA). No unique substitutions were observed on 6K1 protein and HAM1-like protein had the only one unique AA substitution among the isolates in the study. Similarly alignment of AA sequences of TAZ:Cham:14 isolate with all the 15 CBSVs including the TAZ:DES:01 and TAZ:DES:02, revealed 25 unique AA substitutions that were spread across the polyprotein. Most changes were observed in P1 and P3 protein with seven unique substitution. In addition to these AA substitution, many unique substitutions were noted in the polyprotein of TAZ:Cham:14 when compared to all isolates belonging CBSV specific phylogenetic clades (Table 3).
Table 3: Unique AA changes observed in the resistance breaking TAZ:Cham:14 CBSVs isolates as compared to CBSVs. The numbers indicated are starting from the N-terminal of the polyprotein (except for P3N-PiPO protein). The alphabet in front of the number indicates the amino acids reported (TAZ:DES:01) and alphabet following to the number is the amino acid change observed in TAZ:Cham:14.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Changes in TAZ:Cham:14 compared with TAZ:DES:01, TAZ:DES:02 and UG:Kab:07 (change only in TAZ:Cham:14 and other isolates have the same AA)</th>
<th>Changes in TAZ:Cham:14 compared with other CBSVs</th>
<th>Changes in TAZ:Cham:14 compared other (only) CBSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>6KI</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>6K2</td>
<td>R1353K, A1367S, L564C, A1369V, M1378L</td>
<td>A1369V</td>
<td>A1369V</td>
</tr>
<tr>
<td>VPg</td>
<td>N1450D, H1457R, T1479S, V1489I, N1488D, D1505G, Y1500F, F1509L</td>
<td>N1450D</td>
<td>N1450D</td>
</tr>
<tr>
<td>HAM1h</td>
<td>R2439K</td>
<td>M2440T, W2488L, R2489E</td>
<td>M2440T, W2488L, R2489E</td>
</tr>
<tr>
<td>CP</td>
<td>R2563K, V2578E, Q2586K, K2652R</td>
<td>V2578E</td>
<td>V2578E, P2626S</td>
</tr>
<tr>
<td>P3N-PiPO</td>
<td>V51, S13A, L45V, C58L, F64I, R77S</td>
<td>Not determined</td>
<td>Not determined</td>
</tr>
</tbody>
</table>

Some of the unique AA substitutions observed in TAZ:Cham:14 isolate were present in highly conserved motifs. For examples, in P1 protein of TAZ:Cham:14, a mutation was observed on the highly conserved IxFG motif, in which Isoleucine (I) was replaced by Valine (V). Similarly unique AA substitutions were also observed on other conserved motifs (Pasin et al., 2014) (Figure, 3). No specific substitutions were observed on the highly conserved helicase domains of the Cylindrical Inclusion (CI) protein, although two unique substitutions observed were reported to affect replication in TEV (Sorel et al., 2014) (Figure 4). In the viral genome linked
(VPg) protein known to interact with the eukaryotic initiation factors (eIFs) as many as six out of the nine substitutions that were unique to TAZ:Cham:14 isolate corresponded to the central domain that is interacting with eIFs (Figure 5). As expected considerable variation was found between the CBSV isolates (48-70% identity) (Supplementary Table 3).

**Figure 3:** Amino acid substitutions observed on P1 protein of CBSV isolates used in this study, green box indicate some of the conserved motifs according to (Pasin et al., 2014). a. indicate the conserved motif (IxFG), b. LRR domain (L-10x-D), c. C-2x-C-13-C domain, d. ISI conserved motif, e. VELI conserved motif, f. FxxLE conserved motif, g. GWSG serine residue and h. protease conserved QG dipeptide. Amino acids are labelled from start of the P1 protein and * indicate the AA substitution in TAZ:Cham:14.
Figure 4: Amino acid substitution observed on cylindrical inclusion (CI) of CBSV isolates used in this study, green box indicate the conserved helicase domains (based on Sorel et al., 2014) and black box indicate the AA substitution known to affect viral replication (corresponding AA in TEV, based on Carrington et al., 1998). Amino acids are labelled from start of the CI protein and * indicate the AA substitution in TAZ:Cham:14.
Figure 5: Amino acid substitution observed on viral genome linked (VPg) protein of CBSV isolates used in this study, green box indicate the central domain and black box indicate putative NLS binding sites. Light green box indicate the AA substitution observed in PVY, known to affect viral replication (Charron et al., 2008). Amino acids are labelled on top indicate the AA number from start of the polyprotein and * indicate the AA substitution in TAZ:Cham:14.

Discussion

The continuous evolution of viral genetic diversity is a major limitation for sustained resistance in plants. Significant viral diversity is continuously evolving because viruses have compact genomes, enormous population sizes, short generation times, large selection coefficients and higher mutation rates (Elena et al., 2008). Genetic diversity in RNA viruses is caused by the high replication errors arising from their RNA-dependent RNA polymerases (RdRp) that lack proofreading activity (Drake and Holland, 1999). RNA viruses have a higher mutational rate of $10^{-4}$ and $10^{-6}$ mutations per bp per generation compared to dsDNA viruses $10^{-6}$ and $10^{-8}$ mutations per bp per generation (Sanjuán et al., 2010). These modification in the genomes creates heterogeneous virus populations of closely related genomes or “quasispecies” that coexist in the host plant (Biebricher and Eigen, 2006). The evolution or survival of these divergent isolates are greatly influenced by the environmental conditions (Carbonell et al., 2013). Natural recombination and re-assortment between viruses also creates genetic diversity, and this is reported in many viruses (Simon-Loriere and Holmes, 2011). In CBSVs, the recombination events are reported at 3’ end of the virus (HAM1h-CP-3’UTR) between the CBSVs isolates (Mbanzibwa et al., 2011). Recombination in SPMMV was observed to be occurring on both the 5’ and 3’ ends of viral genome (Tugume et al., 2010). Thus the viral variant resulting from diverse mechanisms creates relatively robust isolates leading to both disease emergence and also emergence of resistance breaking isolates. Similarly, many molecular pathogenicity determinants are identified across all potyvirus proteins and each of these proteins play an important role in viral infection process (Revers et al., 1999; Carbonell...
For example in *Turnip mosaic virus* (TuMV), changes in the C-terminal half of P3 protein observed between the two isolates UK1 and CND1 was responsible differential general symptoms on host *Brassica juncea* A. Further, a single amino acid in the UK1 isolate-P3 protein was responsible for the disease causing ability on *B. napus* lines known for possessing a variety of dominant resistance genes (Jenner et al., 2003). Analysis of N-terminal region of CP of *Plum pox virus* (PPV)-D and PPV-R (the two isolates from strain D of PPV that differ in host specificity) suggested that CP acts as host-specific pathogenicity determinants (Carbonell et al., 2013). Also a single AA change (mutation) on CI protein of *Lettuce mosaic virus* resulted in overcoming of *mol*¹ and *mol*² induced resistance (Abdul-Razzak et al., 2009). Induced single point synonymous mutation in the TuMV P3N-PiPO protein rendered the virus to a non-infectious form (Chung et al., 2008). In cucurbits, tolerance-breaking ability of the aggressive Zucchini yellow mosaic virus (ZYMV) (ZYMV-H and CIF06 isolates) was attributed to single amino acid difference in the P3 protein (Glasa et al., 2007). Mutation in the conserved (G12A17) domain of P3N-PiPO rendered the virus incapable of systemic infection (Chung et al., 2008). So these studies suggest that some of the mutations occurring in the viral proteins might results in new capabilities to overcome resistant genes.

A recent report suggests that CBSVs-VPg and host eukaryotic translation initiation factor 4E (eIF4E and eIF(iso)4E) interact and that specific mutations on eIF4E can abolish it’s interaction with VPg protein (https://pag.confex.com/pag/xxiii/webprogram/Paper17082.html). This suggests that recessive resistance to CBSVs could be generated in cassava.

The central domain of TAZ:Cham:14–VPg protein at position 98-114 corresponds with domain that interacts with host elongation factors as observed in *Lettuce mosaic virus* (LMV) VPg proteins, (Roudet-Tavert et al., 2007). A Glycine (G) to Aspartic Acid (D) substitution was observed in TAZ:Cham:14 at AA position 115 and this position corresponds to position 105 in PVY which is within the PVY VPg-eIF4E interaction domain and mutations on this domain have been observed in resistance breaking isolates of PVY. (Charron et al., 2008; Moury et al., 2014). Many other substitutions observed in the putative eIF4E interacting domain of the TAZ:Cham:14–VPg protein likely suggests that the sequence diversity in the TAZ:Cham:14–VPg isolates compared to other isolates might alter complex interaction between the viral proteins and host-eIFs and might actively replicate in KBH 2006/18. Similarly, VPg protein of TAZ:DES:01 isolate had a substitution at position 90 where Glutamic acid (E) (observed in other isolates) was replaced by Aspartic acid (D), the same position corresponded to position 77 in the identified domain within the TuMV-VPg that is involved in the interactions and this interaction affects the viral virulence (Léonard et al., 2000). Given that two other UCBSV isolates known to be less virulent and carry the same substitution (Glutamic acid (E)) as that of the TAZ:Cham:14, this change might likely be less significant in CBSV. Recessive resistance
can also be overcome by point mutation on other viral proteins, for example recessive resistance
in lettuce to LMV is known to be overcome by a resistance breaking isolate due to a stronger
interaction of viral CI protein and its VPg with eIF4E (Tavert-Roudet et al., 2012). Specific
mutations on the C terminus of LMV CI protein, an helicase that is required for replication
(Carrington et al., 1998) together with the VPg can overcome mol resistance in lettuce (Tavert-
Roudet et al., 2012). Similarly on the TAZ:Cham:14 CI protein, some substitutions were
observed on the conserved motifs including a substitution at position 7, G is replaced by D
(corresponding to position 11 of TEV-CI protein, that is known to affect its replication
(Carrington et al., 1998), indicating that the TAZ:Cham:14 CI protein may contribute to
overcome the possible recessive resistance in KBH 2006/18. In pea, a specific mutation on P1
protein of Clover yellow vein virus is reported to be involved overcoming eIF4E-mediated
recessive resistance (Nakahara et al., 2010). The interactions between the P3, P1, CI and VPg
proteins are known (Merits et al., 1999), and are also involved in the resistance breaking
(Hjulsager et al., 2002; Hjulsager et al., 2006). It is demonstrated that point mutation on P3 and
CI, the pathogenicity determinants of Turnip mosaic virus (TuMV) isolates, are required to
overcome two resistant genes in Brassica napus (Jenner et al., 2002). Soybean mosaic virus-
P3 protein was also known to contain a pathogenicity determinant (Wen et al., 2011) and that
a single AA substitution would result in virulence and overcome the resistance offered in Rsv-
4 genotype soybean (Wang et al., 2014b). The high number of specific mutations observed in
the P1 protein of TAZ:Cham:14 might likely show increased silencing suppressor activities as
suggested (Winter et al., 2010) and observed in Tobacco etch potyvirus (TEV) HC-Pro protein
(Torres-Barceló et al., 2008). In TAZ:Cham:14-P3, other unique AA substitutions were also
observed but none of changes could be compared to other studies and hence their significance
remains unknown. Also, the recently identified protein P3N-PiPO (resulting from a +2 or -1
frame shift ORF) that has a highly conserved G_{1:2}A_{6:7} motif (Chung et al., 2008) was found to
differ among the CBSVs, and the motif was observed to be either G_{1:7} or T_{1:7} (Chung et al.,
2008; Monger et al., 2010; Vijayapalani et al., 2012). Changes in the length of the P3N-PiPO
proteins did not affect potyvirus virulence but affected spread and accumulation (Hillung et al.,
2013). Also changes in P3N-PiPO were responsible for overcoming cyvl1-mediated recessive
resistance in pea (Choi et al., 2013). The sequence diversity on viral proteins (either on
individual proteins or in combination) in the TAZ:Cham:14 isolates compared to other isolates
alters complex virus-host interactions. Functional validation is required to substantiate and
identify specific mutations that confers the resistant breaking and to characterize molecular
determinants of virulence in the CBSV-cassava pathosystem.
**Conclusion**

A key discovery of our study was the identification of CBSV isolate (TAZ:Cham:14) which can infect KBH 2006/18, elite breeding line that has resistance to three CBSVs isolates tested. CBSV genetic diversity analysis between four isolates having differential ability to infect variety 60444 and KBH 2006/18 an elite breeding line, revealed that the isolates, TAZ:Cham:14 and TAZ:DES:01 belonged to CBSV clade (Group I and II, respectively) while the isolates, TAZ:DES:02 and UG:Kab:07 belonged to Subgroup III. Neither TAZ:DES:01 and TAZ:DES:02 together nor UCBSV (UG:Kab:07) alone could infect the resistant KBH 2006/18 while TAZ:Cham:14 was able to break its resistance. But further functional validation is required to identify specific single or multiple mutation responsible for resistance-breaking in CBSV isolates.

**Acknowledgements**

We gratefully acknowledge the support of our work by ETH Zurich. R.B.A. is supported by the Research Fellow Partnership Programme (ETH North – South Centre)
References


Lacroix, C., Glais, L., Kerlan, C., Verrier, J.L., and Jacquot, E. 2010. Biological characterization of French Potato virus Y (PVY) isolates collected from PVY-susceptible or -
resistant tobacco plants possessing the recessive resistance gene va. Plant Pathology 59:1133-1143.


Wang, Y., Khatabi, B., and Hajimorad, M.R. 2014. Amino acid substitution in P3 of Soybean mosaic virus to convert avirulence to virulence on Rsv4-genotype soybean is influenced by the genetic composition of P3. Molecular Plant Pathology:n/a-n/a.


**Supplementary Table 1**: Primer used to determine the CBSVs nucleotide sequences. Primers 1-5 are designed on CBSVs consensus sequences. Primers 6-9 designed to confirm the primer binding sites on KBH18 CBSV.

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Primers</th>
<th>Primer</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CBSV-P1-P3 primers</td>
<td>Forward</td>
<td>ATAAAARATGAYATARGAAWACAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>CCGTAMTGKTTTGTGATGAGGCTC</td>
</tr>
<tr>
<td>2</td>
<td>CBSV-P3-CI primers</td>
<td>Forward</td>
<td>GAGCTCATCAAACARCAWTACGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>TCAGCCABTTCTTHACHTCCAT</td>
</tr>
<tr>
<td>3</td>
<td>CBSV-CI-NIa primers</td>
<td>Forward</td>
<td>ATGGADGTDAAGAAVTGGGCTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>GAAAHYHCRCRKCATRGGTGGATTBA</td>
</tr>
<tr>
<td>4</td>
<td>CBSV-NIa-NIb primers</td>
<td>Forward</td>
<td>ATGGGWCAYTATCAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>AACCATGTGATATCC</td>
</tr>
<tr>
<td>5</td>
<td>CBSV-NIb-HAM1h</td>
<td>Forward</td>
<td>CAAGTGGTGAGCAGACCTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>CCTCCTAGTCATCATCTCAG</td>
</tr>
<tr>
<td>6</td>
<td>KBH18-CBSV-HAM1h</td>
<td>Forward</td>
<td>CAGTGGCTCGTGTGTCAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>CACATCAACTGGAGCACCCTG</td>
</tr>
<tr>
<td>7</td>
<td>KBH18-CBSV-NIb</td>
<td>Forward</td>
<td>GGACAAGCGGACTGAAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>CACATCAGTCATCGTGG</td>
</tr>
<tr>
<td>8</td>
<td>KBH18-CBSV-Nla-VPG</td>
<td>Forward</td>
<td>GRCATCGACCTGAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>GTCTCAGTCACCTCTT</td>
</tr>
<tr>
<td>9</td>
<td>KBH18-CBSV-CI</td>
<td>Forward</td>
<td>CACACGAGTCTAGCAGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse</td>
<td>CGGAACTCATCAACCTGTC</td>
</tr>
<tr>
<td>10</td>
<td>CBSV-HAM1h-3’UTR</td>
<td>CBSVF2</td>
<td>GGRCCATACATYARTGTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBSVR1</td>
<td>AAYARAAAGGATATGGAGAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mohammed et al., 2012))</td>
</tr>
</tbody>
</table>
**Supplementary Table 2: Cassava brown streak virus and Ugandan cassava brown streak virus isolates used for the genetic diversity analysis along with the 3 isolates sequenced in this study.**

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Isolate Name</th>
<th>CBSV species</th>
<th>Accession protein Id.</th>
<th>Country</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TAZ:DES:01</td>
<td>CBSV</td>
<td>This study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CBSV-TZ,Nal,07</td>
<td>CBSV</td>
<td>CDO33928</td>
<td>Tanzania</td>
<td>Beyene et al., (direct submission)</td>
</tr>
<tr>
<td>3</td>
<td>KOR6</td>
<td>CBSV</td>
<td>YP_007027011</td>
<td>Tanzania</td>
<td>(Mbanzibwa et al., 2011)</td>
</tr>
<tr>
<td>4</td>
<td>Tan_70</td>
<td>CBSV</td>
<td>CBA13345</td>
<td>Tanzania</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>5</td>
<td>Tan Z</td>
<td>CBSV</td>
<td>ACT78701</td>
<td>Tanzania</td>
<td>(Monger et al., 2010)</td>
</tr>
<tr>
<td>6</td>
<td>Mo 83</td>
<td>CBSV</td>
<td>CBA13344</td>
<td>Mozambique</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>7</td>
<td>TAZ.Cham:14</td>
<td>CBSV</td>
<td>This study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>TAZ:DES:02</td>
<td>UCBSV</td>
<td>This study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Ug_23</td>
<td>UCBSV</td>
<td>CBA18486</td>
<td>Uganda</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>10</td>
<td>Ma_43</td>
<td>UCBSV</td>
<td>CBA13051</td>
<td>Malawi</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>11</td>
<td>Ma_42</td>
<td>UCBSV</td>
<td>CBA13050</td>
<td>Malawi</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>12</td>
<td>Ugandan</td>
<td>UCBSV</td>
<td>ACN50007</td>
<td>Uganda</td>
<td>(Monger et al., 2010)</td>
</tr>
<tr>
<td>13</td>
<td>Ke_125</td>
<td>UCBSV</td>
<td>CBA13048</td>
<td>Kenya</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>14</td>
<td>UG.Kab:07</td>
<td>UCBSV</td>
<td>CDO33929</td>
<td>Uganda</td>
<td>Beyene et al., (direct submission)</td>
</tr>
<tr>
<td>15</td>
<td>Ke_54</td>
<td>UCBSV</td>
<td>CBA13049</td>
<td>Kenya</td>
<td>(Winter et al., 2010)</td>
</tr>
<tr>
<td>16</td>
<td>MLB3</td>
<td>UCBSV</td>
<td>ACM48176</td>
<td>Tanzania</td>
<td>(Mbanzibwa et al., 2009)</td>
</tr>
</tbody>
</table>
**Supplementary Table 3:** AA diversity (% identity) of the polyproteins in different CBSV species. Boxes marked in yellow indicate the isolates sequenced in the present study.

Chapter 5

Discussion and Perspectives

Cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are the most devastating cassava diseases in Africa (Patil and Fauquet, 2009; Legg et al., 2011). Although CBSD was first reported in 1930’s, its re-emergence and spread to newer cassava growing regions coupled with co-occurrence of the disease with CMD is threatening food security in Africa (Alicai et al., 2007; Bigirimana et al., 2011; Legg et al., 2011; Abarshi et al., 2012). At beginning of this study, no robust resistance sources to CBSD had been identified (Yadav et al., 2011). This prompted us to search for new resistance sources to help deploying virus resistant cassava to affected areas as an immediate effort to control the disease and to ensure stable food production. This is the first study to identify and characterize natural resistance sources to CBSD in cassava. This study also provides information about the performance of cassava cultivars from regions where the disease is not yet reported. The 14 varieties used in the study were initially selected based on resistance/tolerance to CMD, so that identified CBSD resistance lines could be directly used to mitigate the effects of both CMD and CBSD. The selected varieties and elite breeding lines were screened for CBSD resistance using a robust and reproducible top cleft grafting inoculation method. In all the independent experiments the susceptible variety 60444 developed CBSV infection, therefore providing a more robust assessment of CBSD resistance as compared to previously reported methods (Maruthi et al., 2005; Mohammed et al., 2012; Wagaba et al., 2013). Two elite breeding lines, KBH 2006/18 and KBH 2006/26 displayed a high level of resistance, as they remained symptom-free while all the other varieties developed typical leaf symptoms 4-8 weeks after grafting (wag). Another variety TMS 30001 was moderately resistant, as only inconspicuous CBSD symptoms developed 8 wag. Additionally, neither co-inoculation of CBSV, UCBSV and EACMV-Ug nor the mode of infection (top cleft or side-grafting) altered CBSD resistance in KBH 2006/18, indicating robust resistance. Precise quantitation of CBSVs using RT-qPCR revealed that the symptomless varieties had low or undetectable virus titers. In the susceptible variety 60444, CBSV titers were highly variable among the different tissues (leaf, stem and root) as has been recently reported (Ogwok et al., 2015). A double grafting experiment using virus-free 2nd scions from susceptible variety 60444 grafted onto 1st scions from resistant KBH 2006/18 and susceptible variety 60444 revealed that systemic movement of CBSV occurs at low levels in resistant KBH 2006/18. A similar systemic movement of viruses was also reported in plum pox virus-resistant apricot varieties (Ion-Nagy et al., 2006). The appearance of early symptoms in 2nd scions with a susceptible middle scion confirms differential virus movement in vascular
tissues of susceptible variety 60444 versus the resistant KBH 2006/18, similar to that reported in the pepper golden mosaic virus resistant (BG-3821-R) and the susceptible Capsicum chinense Jacq plants (BG-3821-S) (García-Neria and Rivera-Bustamante, 2010). The appearance of late symptoms in 2nd scions grafted onto resistant lines suggests that a low concentration of CBSV titers is transported even in resistant varieties. This concurs with observations in the TMV-N. tabacum pathosystem with a transgenic, resistant middle scion (Arce-Johnson et al., 1997). However, further investigation needs to be conducted in order to determine if KBH 2006/18 supports or inhibits CBSV replication or if it only supports movement across tissues, without replication. One way to measure the viral replication process is to quantitate the negative strand of the virus by using strand-specific RT-qPCR or tag-specific primers (Plaskon et al., 2009; Pasin et al., 2014). (Tag specific primers are strand-specific primers which add specific sequence tags to the 5’ ends of transcribed cDNA. The sequence tags are used as primer binding sites during qPCR to differentiate + and – strand derived cDNA.) Another method to check for virus replication is to introduce virus infectious clones, or less commonly, virions into leaf-discs or protoplasts and measuring virus titers over a period of time (Czosnek et al., 1993; Fernández-Calvino et al., 2015; Tian et al., 2015). However, in the present CBSV-cassava pathosystem, the unavailability of an infectious clone was a major limitation for measuring CBSV replication in protoplasts of resistant KBH 2006/18 and susceptible 60444, although efforts were made in this direction. Further efforts are also being made to establish a CBSV virion-protoplasts pathosystem to determine the level of virus replication in contrasting varieties. In conclusion, the present study reduces the impact of potential damages and threats posed by CBSD as well as CMD by identifying new resistant elite breeding lines. The graft inoculation methods developed can readily be used for efficient screening of CBSD resistant accessions under controlled conditions. The top grafting method and double grafting method adopted in this study can also be extended to characterize resistant varieties in other crops where the pathogens are graft transmissible. Further, resistant KBH 2006/18 and KBH 2006/26 lines and the moderately tolerant TMS 30001 variety identified in the study can be further screened under field conditions as well as in multi-location field trials to study the level of resistance to CBSV isolates present in different locations. In addition, since KBH 2006/18 is also resistant to CMGs, it can be utilized as a parent in virus (CBSVs and CMGs) resistance breeding programs. Introgression of resistance gene(s) from KBH 2006/18 to other farmer preferred varieties would greatly benefit cassava growers especially in CBSD endemic areas. Analysis of mapping populations derived from crosses involving KBH 2006/18 would help to better understand the genetics of CBSD resistance. Further, currently available whitefly resistant genotypes (Omongo et al., 2012) could be used to transfer the whitefly resistance trait to these virus resistant lines so that the durability of resistance could be
enhanced. CBSVs and CMGs are known to co-exist and the possible synergism between them could be further studied.

In this work we also performed transcriptome analysis to decipher molecular responses to CBSVs infection in the identified susceptible and resistant varieties, since no molecular mechanisms for either resistance or susceptibility were reported in CBSV-cassava pathosystems. We performed a time course CBSV-infection assay with cassava varieties (KBH 2006/18, KBH 2006/26, TMS 30001 and 60444 varieties) using a top grafting method as previously described. Comparative transcriptome modulations in resistant KBH 2006/18 and susceptible 60444 were first analyzed to identify potential candidate genes involved in early defense responses to CBSV infection. These selected genes were also validated (RT-qPCR) in a moderately tolerant (TMS 30001) variety. RNA-seq results revealed 853 DEGs in susceptible 60444 CBSV infected samples as compared to mock (non-CBSV infected) samples. In resistant KBH 2006/18, 334 DEGs were detected in inoculated samples as compared to mock samples. Susceptible variety 60444 had a higher number of DEGs than the resistant line, KBH 2006/18; this was in contrast to a recent study where the number of DEGs in CBSV-resistant variety Kaleso was higher than susceptible variety Albert (Maruthi et al., 2014). This could likely be due to the experimental set up as well as sampling time, ~1 month (this study) vs 12 months after CBSV-inoculation (Maruthi et al., 2012). The higher number of DEGs in susceptible variety 60444 could likely redirect several biological processes for virus establishment (Postnikova and Nemchinov, 2012; Allie et al., 2014). Our RNA-seq data and subsequent RT-qPCR analysis of selected candidate genes in the early time points (16, 22 and 28 DAG) in moderately resistant TMS 30001 and susceptible 60444 revealed that NTL9 and RDRP1 transcripts were up-regulated in the CBSV-infected plants. Comparing early stage (present study) with late stages or steady state defense responses to CBSV infection (Maruthi et al., 2014) revealed that RDRP1 is only expressed in the early stage of infection. Hence the precise role of RDRP1 expression in CBSV-resistance has to be further elucidated by infecting CBSV in RDRP1 silenced plants and subsequently quantitating the virus titers. Also by developing modified CBSVs (infectious clones) lacking the P1 protein (silencing suppressor), we could further study the role of RDRP1 along with DCL2 and AGO2 that were found to be up-regulated in the present pathosystem.

The RNA-seq analysis also revealed genes encoding 12 disease resistance family protein including TIR-NBS-LRR, dirigent-like and NB-ARC domain-containing disease resistance proteins in 60444 CBSV infected samples. In KBH 2006/18, no DEGs for disease resistance family proteins were found. The genes encoding TIR-NBS-LRR were up-regulated in 60444. Thus it would be interesting to determine their expression during early time point as well as in symptomatic leaves to ascertain their role in CBSV resistance.
The RNA-seq data and subsequent analysis revealed a potential candidate for engineering resistance to Ipomoviruses in general. We observed that activity of β 1,3-glucanase 3 (BG3) in susceptible CBSV-infected 60444 was increased. This increase in activity led to reduction in the callose deposition around the plasmodesmata. This reduction in callose increases the size exclusion limit of the plasmodesmata allowing the movement of the virus from cell to cell (Zavaliev et al., 2013). By reducing the activity of BG3, host susceptibility to virus can be decreased, as reported earlier with transgenic tobacco plants expressing antisense constructs against BG3 (Beffa et al., 1996). The modulation of translation initiation factor 4E (eIFiso4E) in only the susceptible variety 60444 indicates an involvement in virus translation. Further, recessive resistance to viruses can be developed by silencing eIF4E members essential for viral infection (Wang et al., 2013). Alternatively, EMS can be used to generate induced mutation population for the specific eIF4E gene and the resulting mutant population can be screened for virus resistance (Piron et al., 2010). These resistant lines can then be analyzed to identify specific mutations on eIF4E that are responsible for recessive resistance. It is also important to elucidate the role of other significant genes from this dataset to determine their role in disease resistance. This could lead to identification of additional resistant genes that subsequently can be used in developing molecular markers linked to CBSD resistance. Thus the identified molecular markers can assist in identifying new disease resistant varieties from cassava germplasm as well as from mapping populations in resistance breeding programs. Use of molecular markers can also greatly reduce the time required for developing new varieties. We also discovered new transcripts by de novo assembly of Illumina RNA-seq reads; this has a potential application in improving annotations to the cassava genome (Denoeud et al., 2008; Trapnell et al., 2010; Roberts et al., 2011).

KBH 2006/18, an elite breeding line identified by us as resistant to several CBSV isolates (TAZ:DES:01, TAZ:DES:01, UG:Kab:07, MZ:Nam1:07) recently showed typical CBSD symptoms in a field trial performed by IITA – Dar Es Salaam at the Chambezi research station in Tanzania (Edward Kanju, personal communication). The infected KBH 2006/18 stem cuttings were collected and propagated in the greenhouse for further experiments. Virus-free KBH 2006/18 grafted on to field-infected KBH 2006/18 rootstocks under greenhouse conditions confirmed that KBH 2006/18 was susceptible to this field isolate. The complete genome sequence of this isolate was obtained using degenerate primers and Sanger sequencing. We named this new isolate TAZ:Cham:14 and phylogenetic sequence analysis revealed that it belonged to the CBSV clade. Comparing the amino acid (AA) sequences of resistance breaking CBSV-TAZ:Cham:14 with non-resistance breaking CBSV-TAZ:DES:01 and UCBSV-TAZ:DES:02 isolates revealed specific mutations. No specific mutations were identified in the viral 6K1 protein suggesting that 6K1 might not be a virulence factor contributing to resistance.
breaking in KBH 2006/18. The large number of specific mutations observed in P1 protein (a suppressor of PTGS) of TAZ:Cham:14 isolate could hypothetically cause increased suppression of silencing as suggested by (Winter et al., 2010) and observed in Tobacco etch potyvirus (TEV) HC-Pro protein (Torres-Barceló et al., 2008). Specific mutations in viral VPg, CI and P3 proteins are also reported in resistance breaking potyvirus isolates (Jenner et al., 2002; Abdul-Razzak et al., 2009; Carbonell et al., 2013). The mutations observed on the TAZ:Cham:14-VPg protein might show complex adaptations to a still unknown resistance mechanism in KBH 2006/18 which could be monogenic or polygenic. Resistance breaking similar to that observed in four pepper genotypes carrying different alleles at the pvr2 locus for potato virus Y (Ayme et al., 2007) could explain the observed susceptibility of KBH 2006/18. In addition multiple AA substitutions were also observed in the P3 and CI proteins of the TAZ:Cham:14 isolate, which might also be involved in overcoming KBH 2006/18 resistant genes as observed in soybean against the Soybean mosaic virus (Chowda-Reddy et al., 2011).

In conclusion, this is the first study to identify CBSV isolates with a differing ability to infect cassava varieties. The degenerate primers designed in the study can be used to sequence both CBSV and UCBSV isolates. Further study on the observed mutations in the CBSV-TAZ:Cham:14 isolate is required to determine its molecular virulence factor(s). The identification and whole genome sequencing of an isolate with contrasting virulence represents an important first step towards determining virulence determinants in cassava-infecting ipomoviruses. Our phylogenetic analysis of CBSVs sequences revealed that isolates segregated into three groups. UCBSVs formed one group, designated Group III; CBSV isolates separated into two distinct groups, with the group containing TAZ:Cham:14 named Group I and TAZ:DES:01, Group II. To characterize the role of each mutations/substitution present in the TAZ:Cham:14 isolates, the first step would be to access more number of isolates belonging to each group and to challenge the elite breeding line KBH 2006/18 with these isolates so that additional isolates that might have differential virulence are identified. Further comparing the sequence diversity among the relatively large number of resistance breaking and non-resistance breaking isolates would pinpoint molecular determinants of virulence in cassava-infecting ipomoviruses. It is imperative to develop infectious clones to study these viral genome modification by creating chimeric sequences or by domain swapping of isolate specific AA motifs to determine their role in pathogenicity. The first step in the development of infectious clones is to obtain a large amount of full length infectious RNA or cDNA obtained by reverse transcription (RT) of viral RNA; PCR amplification of overlapping 5’ and 3’ cDNA fragments or by concatenating smaller fragments of cDNA (Chapman, 2008); or by gene synthesis. With recent advances, robust methods are available for generation of full length infectious clone and with different promoters like CaMV 35S promoter, T7 bacteriophage RNA polymerases that drive the full length cDNA clones (Weber et al., 1992). Infectivity assays with full length virus
cDNA clones are performed by symptom observation as well as by quantitating viral transcripts, after inoculation either by mechanical rubbing or infiltration into leaves of susceptible/test cultivars (Burgyan et al., 1990; Ryabov, 2008). Alternatively, protoplast cultures offer a quicker means to test the infectivity of infectious clones by quantitation of the viral transcript over a time course to study viral replication. Additionally, the creation of a modified virus tagged with green fluorescent protein (GFP) as a reporter would be useful to characterize viral movement in resistant and susceptible varieties.

In summary, the present work has identified new CBSV-resistance sources. The study has also contributed to improve the existing knowledge on molecular mechanisms of pathogenesis and defense responses to CBSV infections. Identification of a resistant breaking CBSV isolate has clearly indicated the need for newer methodologies to mitigate the impact of CBSV on cassava production. The data generated in the present study can also be utilized in future CBSD-resistant breeding programs.
References


Piron, F., Nicolaï, M., Minoïa, S., Piednoir, E., Moretti, A., Salgues, A., Zamir, D., Caranta, C., and Bendahmane, A. 2010. An Induced Mutation in Tomato eIf4E Leads to Immunity to Two Potyviruses. PLoS ONE 5:e11313.


Acknowledgements

I would like to first and foremost thank Prof. Wilhelm Gruissem for giving me the opportunity to pursue my PhD in his laboratory. His advice, support and mentorship have been invaluable for the successful completion of my project.

I am very grateful to Dr Hervé Vanderschuren for supervising me during my PhD. His advice, support and substantial help are greatly acknowledged.

I am also thankful to Dr Maruthi Gowda for providing virus isolates, hosting me in his laboratory and assisting me with critical portions of my research.

I greatly acknowledge the constructive comments and improvements to grammar provided by Devang, Adrian, Joel, Simon, Emily, Glen, Wilfred, Charles and Vamsi. I thank Ima, Eze, Kuan-Te and entire cassava group for constructive advice and suggestions. My heart-felt gratitude towards Irene for helping me keeping my plants happy, in spite of the virus infection.

Endless thanks to Doris, Kim, Serena and all the group members for keeping the atmosphere in lab supportive. A very special thanks to Daniela for all the administrative help.

Last but not least, my very best thanks go to my wife, Bhavana and my daughter, Sadrushya for being there with love and full support when most needed.

A very special thanks to the RFPP program from the North-South Centre, Zurich for financial support.
CURRICULUM VITAE

Ravi Bodampalli Anjanappa  
PhD candidate  
Plant Biotechnology Group  
ETHZ, Zurich  
Switzerland  
E-mail: aravi@ethz.ch  
Phone: +41-786381981

RESEARCH EXPERIENCE:

1. Worked in Genome Characterization Laboratory, Division of Plant Genetic Resources, Indian Institute of Horticultural Research, Bangalore, India from November 2003 to March 2005 in pursue of completing M.Sc research work

2. Worked as Senior Research Fellow in Indian Institute of Horticultural Research, Bangalore, India in the DBT project entitled “Development of mapping population and identification of molecular markers linked to tomato leaf curl virus resistance in tomato (ToLCV) (Lycopersicon esculentum Mill) from August 2005 to March 2007.
   ➢ Identifying stable sources of resistance to ToLCV in tomato.
   ➢ Identifying molecular markers linked to ToLCV resistance in tomato.
   ➢ Identifying pathogen related proteins.
   ➢ Identification of viral strain

3. Worked as Consultant for Global Blooms, Agri-tech consultancy services from March 2007 to March 2008

4. Worked as Research Associate in Project entitled "Genetic enhancement of Moth Bean (Phaseolus aconitifolia Jacq.) involving molecular markers" in Plant Molecular Biology Unit, Kirkhouse Trust funded lab, Department of Genetics and Plant Breeding, GKVK, UAS, Bangalore, India.

EDUCATIONAL QUALIFICATION:

<table>
<thead>
<tr>
<th>Name of the Examination</th>
<th>University</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ph. D</td>
<td>Swiss Federal Institute of Technology, Zurich (ETHZ, Zurich)</td>
<td>2010-</td>
</tr>
<tr>
<td>Master of Science (Agri) in Plant biotechnology</td>
<td>University of Agricultural Sciences, Bangalore</td>
<td>2002-2005</td>
</tr>
<tr>
<td>Bachelor of Science (Agriculture)</td>
<td>University of Agricultural Sciences, Bangalore</td>
<td>1997-2001</td>
</tr>
</tbody>
</table>
PUBLICATIONS:


Publications (in-preparation):
1. Characterization of brown streak virus-resistant cassava
2. Transcriptome modulation in susceptible and resistant cassava varieties inoculated with cassava brown streak viruses.
3. Identification and characterization of resistance-breaking Cassava brown streak virus isolate

REFERENCES

Prof. Dr. Wilhelm Gruissem
Plant Biotechnology Group
Universitatstrasse 2
8092, Zurich
Switzerland
Phone: +41 44 632 08 57
Fax: +41 44 632 10 79
E-Mail: wgruisse@ethz.ch

Dr. Herve Vanderschuren
Group Leader
Plant Biotechnology Group
Universitatstrasse 2
8092, Zurich
Switzerland
Phone: +41 44 632 87 25
E-Mail: hvanderschuren@ethz.ch

Dr. Maruthi M N Gowda
Reader in Molecular Plant Pathology
Natural Resources Institute
University of Greenwich
Kent, ME4 4TB, UK
Tel: +44 (0) 1634 883957
E-mail: m.n.maruthi@gre.ac.uk