SYNTHETIC BIOLOGY FOR PLANT VIRAL DIAGNOSTICS: APPLICATION TO MAIZE LETHAL NECROSIS DISEASE

by

Susan Wangechi Gachara

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Plant and Soil Sciences

Spring 2020

SYNTHETIC BIOLOGY FOR PLANT VIRAL DIAGNOSTICS:

APPLICATION TO MAIZE LETHAL NECROSIS DISEASE

by

Susan Wangechi Gachara

Approved: _____

Randall J. Wisser, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Ervin Erik, Ph.D.
Chair of the Department of Plant and Soil Science

Approved: _____

Mark Rieger, Ph.D.
Dean of the College of Agriculture and Natural Resources

Approved: _____

Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

ACKNOWLEDGMENTS

My appreciation and respect to my advisor and mentor, Randall Wisser, PhD, for your relentless support, encouragement, guidance and patience throughout this research; you have inculcated in me skills to become a better scientist and researcher. Thank you to my committee members, Dr. Donofrio, Dr. Fondong and Dr. Runo for your guidance, direction and feedback throughout this research. Thank you also to Dr. Green who is part of the team that designed the toehold switch technology and whose assistance has been vital in this research.

My utmost gratitude to Jim and Marcia Borel for funding my master's program at UD and making my dream to study abroad possible. Much gratitude to the College of Agriculture and Natural Resources Seed Grant who also funded this research.

I give my love and thanks to Mum and Dad, for their love, support and encouragement throughout my life, to my siblings; Edwin, Peris, Hilda and Rodah for offering me their help whenever I need it, to my husband, Alex for helping me maintain my sanity during my studies abroad, and to my son, Jabali for giving me a new meaning in this life.

My gratitude to all the Wisser lab members, Michael, Scott, Heather, Zhu, Felix, Isaac and Tecle for the comradeship in and out of the lab. To my friends and the Plant and Soil Science family for making me feel at home away from home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Plant viruses are a risk to many economically important crops, including maize which serves as a staple food for people throughout the world. Indeed, a current epidemic of maize lethal necrosis (MLN) is affecting food security across eastern and central Africa. The disease results from synergistic interaction between two viruses; maize chlorotic mottle virus (MCMV) and sugarcane mosaic virus (SCMV; or other potyviruses). Rapid detection of disease causal agents is imperative for disease management. Applications in synthetic biology have enabled the development of new, reliable and affordable molecular diagnostic tools that overcome the challenges of the current diagnostic tools used in agriculture. In 2016, Toehold switch technology was demonstrated for the detection of Zika virus in the developing world where resources are limited. This technology can potentially be used for detection of plant viruses as it can easily be designed to detect any RNA sequence. Using publicly available whole-genome sequences for 43 and 73 global isolates of MCMV and SCMV, respectively, an *in silico* filtering pipeline was used to identify five MCMV and six SCMV sequences that match the canonical structure of a toehold switch and form the basis for toehold switch construction. An *in vitro* screen was used to assess the functionality of the MCMV and SCMV toehold switches. In addition, three criteria: performance, sensitivity and stability were used to assess the merit for moving forward in the development of field-based sensors for MLN-viral detection.

Chapter 1

INTRODUCTION

1.1 Overview

Synthetic biology is a rapidly emerging field of research that integrates principles of engineering and biology to imitate the functions of biological systems with the aim of innovating novel mechanisms to use in diagnostics, discoveries, improve human health, address environmental issues and use in industries (Ceroni, 2015; Way et al., 2014). Synthetic biologists have developed and adapted new molecular diagnostic technology for pathogen detection. A good example is Quick et al., (2016) who used MinIoN Nanopore technology in surveillance of Ebola virus in West Africa. Another example is Pardee et al., (2016) who used programmable biomolecular sensors which utilize toehold switches technology in the detection of Zika virus. In agriculture, most viral diagnostics are limited to laboratory tests which are expensive, tedious and time-consuming. Toehold switch technology offers a promising but untested solution for plant pathogen detection as it can be rapidly designed to sense any RNA sequence.

A toehold switch is a synthetically constructed molecular sensor designed to detect a predefined trans-acting RNA trigger sequence, such as part of a viral genome or RNA (encoded by genes) from bacterial or fungal pathogens (Green et al., 2014). Based on the confirmation of a secondary hairpin structure, the toehold switch exists

in either an off-state or on-state (Figure 1.1). The switch RNA mimics messenger RNA. It consists of an exposed single-stranded region called the toehold sequence, a hairpin structure which sequesters the ribosomal binding site (RBS) and a start codon, and repressed reporter gene. The trigger RNA strand is a complementary sequence to the toehold sequence. In the absence of the trigger sequence, the secondary hairpin conformation renders the switch to an off-state. The trigger sequence binds to the toehold domain which destabilizes the hairpin structure thereby exposing the RBS and the start codon which consequently expresses the repressed gene giving off a signal.

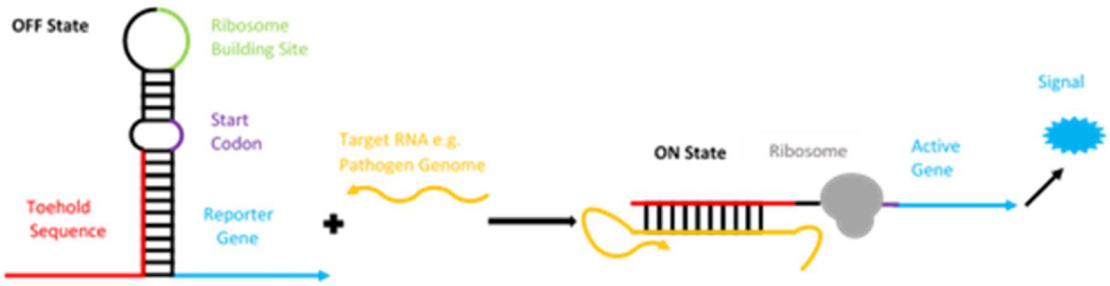


Figure 1.1: A schematic representation of the toehold switch system

The Zika toehold switches detected a minimum concentration of 2.8 fM (femtomolar) from plasma samples (Pardee et al., 2016). Toehold switches have also been developed for the detection of norovirus on a paper-based cell-free platform (Ma et al., 2018). This study used toehold switches together with synbody-based enrichment of the virus and was able to detect norovirus GII.4 Sydney at a concentration of 270 zM (zeptomolar) from stool sample.

1.2 Maize production

Maize is one of the most important cereal crops in the world, utilized as food, animal feed, biofuel and in industries. However, its production varies greatly between countries (Table 1). In 2017/2018, global maize production was estimated at 1,100,237,000 tons, with the USA, China and Brazil yielding 33%, 19% and 8% of the total maize produced respectively (US Department of Agriculture & USDA Foreign Agricultural Services, 2018). In Sub-Saharan Africa (SSA), maize is a popular staple crop occupying approximately 36 million ha of land with an annual production of 72 million metric tons (MT) (Abate et al., 2017). In Kenya, more than 85% of the population rely on maize as their main staple crop. Small-holder farmers are the main maize producers for both subsistence and commerce which accounts for 65% of maize harvested in the country (Abate & Wawa, 2015). The total maize harvested annually is roughly 2.8 million MT grown in 1.7 million ha of land in all agro-ecological zones (Gitonga, 2017). This translates to 1.6 tons per hectare which is low compared to the global average of 5.0 tons per hectare.

The low maize production is attributed to poor weather, lack of adoption of high yielding varieties, inadequate extension services to smallholder farmers, poor rural infrastructure, pests and diseases in farm and in storage (Abate & Wawa, 2015; Onono, Wawire, & Ombuki, 2013). In 2011, a new maize disease identified as maize lethal necrosis disease (MLN) was reported in Rift Valley Province of Kenya (Wangai et al., 2012), and it has since spread to other maize growing regions in the country and

across SSA. This was the first report of the disease in Africa; however, the disease has been a threat to maize growing regions in most parts of the world, including the U.S.A., South America and China (Mahuku et al., 2015). Following its initial identification in Kenya, the disease has spread to neighboring countries and has been confirmed in Tanzania (Makumbi & Wangai, 2012), Uganda (Frenken, 2013), Rwanda (Adams et al., 2014), the Democratic Republic of Congo (Lukanda et al., 2014) and Ethiopia (Mahuku et al., 2015)

Table 1.1: Global Maize Production. Source: US Department of Agriculture and USDA Foreign Agriculture Service, 2018

Country	**Maize production in 2017/2018 (in 1,000 metric tons)**
United States of America	370 960
China	215 891
Brazil	94 500
Ethiopia	65 000
European Union	61 139
Argentina	36 000
India	27 150
Mexico	26 500
Ukraine	24 115

Canada	14 100
Russia	13 229
South Africa	13 000
Indonesia	11 350
Nigeria	10 500
Philippines	8 200
Egypt	6 000
Others	102 603

1.3 Biology of MLN

MLN is a viral disease resulting from synergistic interaction of maize chlorotic mottle virus (MCMV: genus Machlomovirus, family Tombusviridae) and a Potyvirus: maize dwarf mosaic virus (MDMV), wheat streak mosaic virus (WSMV) or sugarcane mosaic virus (SCMV) (Niblett & Claflin, 1978; UYEMOTO et al., 1981; Wangai et al., 2012). MLN was first reported in the states of Kansas and Nebraska in the United States of America (USA) in 1976 as mixed infections of MCMV and MDMV or WSMV (Niblett & Claflin, 1978). It was then identified in Hawaii, still in the USA (Jiang et al., 1985), China (Xie et al., 2011), Sub-Saharan Africa (Mahuku et al., 2015) and Ecuador (Alvarez & Mendoza, 2016). Consistent with initial reports of MLN, in Africa, MLN is associated with coinfection of MCMV and SCMV (Wangai et al.,

2012), and yield losses of up to 90% have been reported (Mahuku et al., 2015). SCMV has always been a problem in agriculture worldwide, the occurrence of MCMV in regions where SCMV, MDMV or WSMV are present triggers the development of MLN.

MCMV is vectored by six species of Chrysomelid beetles namely, the cereal leaf beetle (*Oulema melanopa*), the corn flea beetle (*Chaetocnema pulicaria*), the flea beetle (*Systena frontalis*), the southern corn rootworm (*Diabrotica undecimpunctata*), the northern corn rootworm (*D. longicornis*) and the western corn rootworm (*D. virgifera*) in larvae and adult forms (Nault, 1978). Additionally, maize thrips (*Frankliniella williamsi*) and western flower thrips (*Frankliniella occidentalis*) transmit MCMV in both larvae and adult forms (Cabanas et al., 2013; Jiang et al., 1992; Zhao et al., 2014). MCMV transmission by thrips and Chrysomelid beetles is similar in that the virus is transmitted shortly after acquisition with no latent period and is retained up to six days thereafter (Cabanas et al., 2013; Nault, 1978). SCMV as other Potyviruses is transmitted by aphids in non-persistent manner therefore the transmission ranges from minutes to hours.

MCMV and SCMV are also transmitted through seed. In a recent study, 72% of maize seeds from MCMV infected plants tested positive for MCMV through RT-PCR (Mahuku et al., 2015). Jensen et al., (1991) reported 0.04% of seeds from MCMV infected plants transmitting MCMV to maize seedlings. Although this is a very low transmission rate, seed- based transmission forms the initial virus source of infection, which together with transmission by insect vectors cause significant effect

on MCMV epidemics. Healthy seeds planted in MCMV contaminated soil has been reported to develop disease (Mahuku et al., 2015).

When the virus infects a susceptible plant, it alters the plant cell structure and functions to gain access to the phloem for long-distance transport to the rest of the plant (Gilbertson & Lucas, 1996). Viruses move within the plants using cell-to-cell movement through the plasmodesmata and long distance via the phloem's sieve tube network. The virions are too large to pass through the plasmodesmata microchannels therefore, viral genomes interact with host cellular factors to effect viral infection (Kumar et. al., 2015). Viruses encode movement proteins (MP) and coat proteins (CP) that target and modify the size exclusion limit (SEL) of the plasmodesmata to aid in passage of the viral genome (Gilbertson & Lucas, 1996). The CP also protects the viral nucleic acid from degradation as it moves within the plant. Plant viruses recruit the host proteins to help in the viral entry into the phloem by forming a stable transport-competent complex. While in the phloem, the virus mixes with the phloem sap for transportation to distant sites and exits the phloem to infect new sites and distribute throughout the plant (Hipper et. al., 2013). MLN is characterized by chlorotic mottling of leaves, necrosis developing from leaf margins to the mid rib, "dead heart" symptom in young leaves leading to the development of small cobs and in some cases plant death (Wangai et al., 2012).

1.4 MLN causing viruses in Africa

1.4.1 Maize chlorotic mottle virus

Maize chlorotic mottle virus (MCMV) is a major maize virus, causing significant yield losses in maize worldwide. MCMV was first reported in South America in Peru in 1974 (Castillo & Hebert, 1974) then in North America and most recently in China and East and Central Africa (Figure 1.2) (Mahuku et al., 2015). MCMV can cause disease on its own or synergistically with either maize dwarf mosaic virus (MDMV), wheat streak mosaic virus (WSMV) or sugarcane mosaic virus (SCMV) from Potyviridae family.

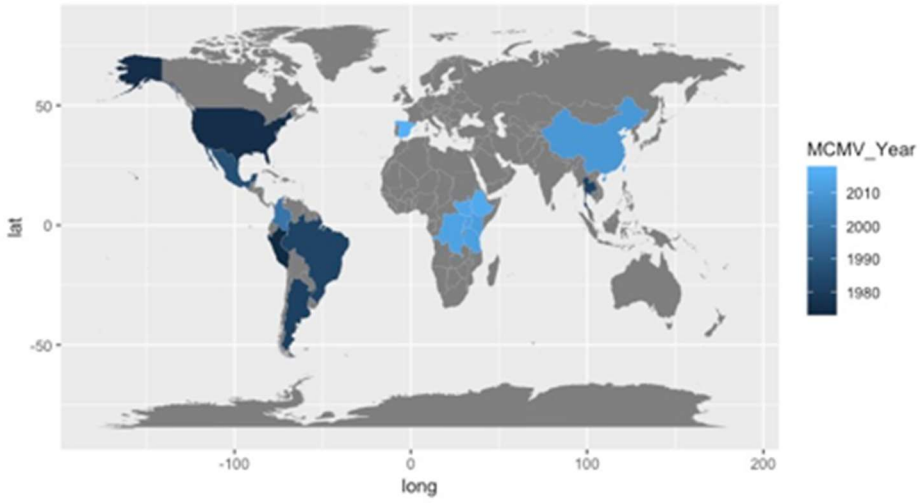


Figure 1.2: Distribution of MCMV worldwide. The more the lighter the shade of blue the more recent the infection. Map generated using package ggplot in R.

MCMV is the only member of Machlomovirus (family: Tombuviridae). It has a positive sense single-stranded linear RNA (ssRNA) genome of approximately 4.4 kb encased in icosahedral virion of 30nm in diameter (Nutter et al., 1989). MCMV has

two sub-genomic RNAs (sgRNA) at the 3' end of the genome, however, it lacks both the viral protein genome-linked (VPg) and 3' polyA tail (Nutter et al., 1989). The MCMV genome has six overlapping open reading frames (ORFs) (Figure 1.3) distributed across genomic RNA (gRNA) and one sgRNA (Scheets, 2016). The gRNA encodes three proteins: p32 of unknown function, replicase proteins, p50 and its read-through p111 (Scheets, 2016). Although p32 is neither required for virus infection nor spread, its absence significantly reduces virus accumulation and symptoms in the host (Scheets, 2016). This indicates that p32 could be supporting the infections by suppressing plant defense. Replicase proteins, p50 and p111, are the only proteins required for virus replication in the host and they act at low levels *in trans* (Scheets, 2016). The first sgRNA is approximately 1467 nucleotides in size and encodes three proteins: p7 (p7a and p7b), required for cell to cell movement, p31 which is involved in long distance systemic movement of the virus in the plant, and coat protein (CP), which enhance the cell-to-cell movement of the virus (Scheets, 2000, 2016). Scheets, (2016) showed that hindering the production of CP considerably impairs the virus infection *in vivo*. The second sgRNA is 337 nucleotides long and is not known to encode any protein.

Figure 1.3: Genome map of Maize chlorotic mottle virus based on the refence sequence X14736.2. The open reading frames (ORFs) are indicated by the cylinders. The numbers indicate the start and stop positions of the nucleotide and the arrow heads indicate the stop codons read through of p50 and p7. (Adapted from Wamaitha, et al., 2018).

MCMV infection modifies the maize starch grains in the chloroplasts cell into small starch grains thereby inhibiting photosynthesis (Wang et al., 2017). Mitochondria play an essential role in plant defense by integrating and amplifying defense signals in plant cells, such as salicylic acid, nitrite oxide, reactive oxygen species (ROS) or pathogen elicitors. These signals promote expression of defense genes to block pathogen from spreading to the whole plant through programmed cell death (Colombatti et al,. 2014). Wang et al., (2017) reported that MCMV infections damage mitochondria, disrupting their functions thus hindering plant defense mechanisms. Additionally, infected plants develop multi-vesicular bodies (MVBs), which are thought to act as replication sites for MCMV (Wang et al., 2017). This results in the phenotypic characteristics of chlorotic mottling of leaves to severe stunting growth, premature leaf senescing and lack or partially filled ears (UYEMOTO et al., 1981).

1.4.2 Sugarcane mosaic virus

Sugarcane mosaic virus, a member of Potyvirus (family Potyviridae), causes significant yield losses to a wide range of plants belonging to the Poaceae family, particularly maize, sugarcane and sorghum (Wu et. al., 2012). SCMV was first identified in sugarcane fields in Puerto Rico in 1916 as a result of introduction of infected cuttings (Brandes, 1919). In Kenya, SCMV was reported in 1969 in maize at the Kenya agriculture and livestock research organization (KALRO) in Kitale and was later confirmed in neighboring countries (Kulkarni, 1973).

SCMV as well as other potyviruses has positive sense, single stranded linear RNA genome. SCMV has an approximately 9.7 kb genome, which is covalently connected to a viral protein genome-linked (VPg) at the 5' end and a polyadenylated (poly A) tail at the 3' end and is enclosed in flexuous filamentous particles of 700-750 nm in length (Tennant et. al., 2018; Wu et al., 2012). The genome has a long open reading frame (ORF) that is translated into a large polyprotein (Figure 1.4) that is later cleaved by viral encoded proteinases (P1, HC-Pro and Nla – Pro) to produce 11 functional proteins, namely protein 1 (P1), helper component proteinase (HC-Pro), protein 3 (P3), 6K1, cylindrical inclusion protein (CI), 6K2, viral protein genome-linked (VPg), major protease of the small nuclear inclusion protein, Nla (Nla-Pro), large nuclear inclusion protein (Nlb), coat protein (CP) and PIPO protein (ORF2) (Krczal & Wassenegger, 2013; Valli, 2013). The SCMV protein functions are listed in Table 1.2 (Urcuqui-Inchima et. Al., 2001).

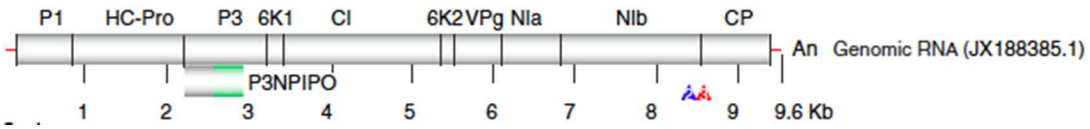


Figure 1.4: Sugarcane mosaic virus genome map based on the Ohio isolate sequence JX188385.1. The open reading frames (ORFs) are indicated by the cylinders. (Adapted from Wamaitha, et al., 2018).

Table 1.2: Functions of SCMV Proteins.

SCMV proteins	Functions
P1	Viral genome amplification
HC-Pro	Interacts with aphids and virions
	Helps in viral genome amplification
	Aids in cell-to-cell movement of the virus and systemic movement
	Suppresses gene silencing
	Promotes synergism and symptom development
P3	Interacts with CI for virus amplification
	Associated with the wilting phenotype together with other proteins
P3NPIPO	Recruits CI to the plasmodesmata
6K1	Aids in the formation of replication vesicles
CI	Interacts with 6K2 and serves as a docking point for the intercellular
	movement of virus replication vesicles
6K2	Remodels the endoplasmic reticulum (ER) membrane into convolutional
	structures

	Induces the virus to form replication vesicles for intracellular and intercellular movement of the potyvirus
VPg	Associated with virus replication, translation and movement Interacts with host translation initiation factor, eIF4E or its isoform eIF(iso)4E, for infection.
NIa-Pro	Interacts with viral RNA-dependent RNA polymerase (NIb protein) Involved in viral replication
NIb	RNA-dependent RNA polymerase
CP	Enables aphid transmission Facilitates cell-to-cell and systemic movement of the virus Provides encapsidation of the viral RNA Regulates viral RNA amplification

Plants employ a defense mechanism, referred to as RNA silencing, against viral infection. During SCMV infection, the HC-Pro protein is involved in cell to cell and systemic movement of the virus, replication and acts as a suppressor of post-transcriptional gene silencing (Xia et al., 2016). Researchers have reported increased accumulations of MCMV in mixed infections of MCMV and SCMV which could be due to the functions of HC-Pro protein (Xia et al., 2016).

1.5 Current MLN diagnostic methods

The first stage in management and treatment of any disease is the correct identification of the causal agent. Many methods have been developed for the

detection of plant viruses. Below are the current methods used in the detection of MLN-causing viruses in Kenya.

1.5.1 Enzyme-linked immunosorbent assay (ELISA)

ELISA is a qualitative/quantitative serological technique developed for the detection of antigens such as hormones, peptides, nucleic acids, plant secondary metabolites and proteins among others (Clark & Adams, 1977). This method is based on the concept of antibody-antigen reaction in which the antigens are identified by their specific antibodies giving off a colorimetric signal (Lima et al., 2009; Sakamoto et al., 2018). Briefly, ELISA is carried out in a 96-well microtiter plate which contain antibodies in each well; a secondary antibody linked to an enzyme catalyzes the reaction resulting in color change. The samples are prepared by grinding the tissue and diluting it with the General Extraction Buffer (GEB) provided in the kit. The samples are added to the test wells and incubated for 2 hours or overnight at 2-8 °C to allow the antigen to react with the specific antibody. The presence of the antigens is detected by the development of color while the negative samples remain colorless. The degree of color change indicates the degree of reactivity and can be read by the ELISA plate reader. MLN is detected by ELISA using MCMV ELISA kits and Potyvirus ELISA kits. A minimum concentration of MCMV can be detected at 25 pg/ml using triple antibody sandwich ELISA (Wu et al., 2013).

1.5.2 Immunostrips

Immunostrips also called lateral flow assay is a qualitative diagnostic technique for the detection of antibody, as from the sample in question with immobilized antigens resulting in antibody-antigen complexes (Posthuma-Trumpie et al., 2009). The method is based on capillary flow across a control strip that captures all molecules to ensure the strip and reaction are working and a viral antibody specific strip that captures specific viral antibodies. The assembled strips are divided into four parts: a lower casing, an absorbent panel, a reactive membrane panel, and an upper casing. A sample is placed on the reactive membrane panel and allowed to flow through the absorbent pad, with the target analyte captured in the membrane. The results are easy to read and interpreted, i.e., single band for negative and double bands for positive.

There are commercially available MCMV-specific and potyvirus immunostrip kits. The advantage of immunostrip kits is that they are simple to use and thus do not require expert knowledge, they can be used in non-laboratory environments and are relatively inexpensive. The disadvantage of immunostrips is that they are not locally available in Kenya therefore have to be shipped from overseas making them expensive. They are also designed for individual tests thus cannot be used for high throughput screening.

1.5.3 Polymerase Chain Reaction (PCR)

PCR is an in vitro technique that allows amplification of genes and DNA/RNA templates for sequencing, detecting the presence or absence of a pathogen and for forensic DNA profiles (Maheaswari et al., 2016). PCR assay requires template DNA (the sample in question), primers, DNA polymerase, free nucleotides called dNTPs and buffer. The primers, forward and reverse primer, are short DNA sequences and are designed to bind to either side of the DNA template to initiate a PCR reaction. The free nucleotides (A, C, G and T) are added by the DNA polymerase to construct the new DNA strand. The buffer contains magnesium chloride to provide the right conditions for the reaction. The DNA template is extracted and added to the reaction mix in a microfuge tube or a 96-well plate then placed in a thermal cycler that allows repeated cycles of heating and cooling for DNA amplification. PCR takes place in three main stages: denaturation, which separates the DNA template into two single strands by heating to 94-95 °C; annealing, at 50-58 °C to allow primers to attach to the DNA template at specific sites; extension, which allows the DNA polymerase to synthesize the new complementary DNA strand at 72 °C. This process is repeated up to 35 times. After completion of the amplification process, the size of the amplified product (amplicon) can be detected by electrophoresis. Since MCMV and SCMV are RNA viruses, reverse transcription PCR (RT-PCR) is used for the detection of the viruses. This method uses a primer which anneals to the RNA of interest and reverse transcriptase enzyme to synthesize a single-stranded cDNA from the RNA (Rodriguez et. al., 2009). The cDNA is then used as a template for PCR.

PCR is highly sensitive and specific; however, it requires technical expertise to set up the reaction, it is expensive and the sample can easily be contaminated resulting in false positive or negative results. Minimum SCMV concentrations of 100 pg/uL have been detected by conventional RT-PCR method (Keizerweerd et al., 2015). The lowest detection limit of MCMV using conventional reverse transcription PCR method is 80.7 fg/uL (Zhanmin et al., 2016).

Chapter 2
EXPERIMENTAL MATERIALS AND METHODS

2.1 Bacterial strains and plasmids

Properties of the bacterial strains and plasmids used for this research are listed in the table 2.1.

Table 2.1: Bacterial Strains and Plasmids Used in this Study.

Plasmid or *E. coli*	Characteristics	Source
pET_15b	N/A	Addgene
ZIKV_Sensor_27B_LacZ	Codes *lacZ* gene	Addgene
ZIKV_Sensor_32B_LacZ	Codes *lacZ* gene	Addgene
NEB 5 Alpha	Compatible with Gibson assembly components	NEB
***lacZ* deficient BL21-Gold-dLacZ (DE3)**	T7 express strain Lac operon deletion	Addgene

2.2 Culture conditions

E. coli strains were grown at 37°C with shaking at 250 – 275 rpm in liquid LB medium or in solid LB/agar at 37°C. The growth medium was supplemented with either kanamycin, ampicillin or both antibiotics when necessary. To induce expression of the T7 promoter of the recombinant strains, isopropyl ß-D-1-thiogalactopyranoside (IPTG) was added to the culture medium. Blue –white selection substrate, 5-Bromo-4-chloro-3-inolyl-ß-D-galactopyranoside (X-gal), was also added to the culture medium.

2.3 Validation of toehold switch function

As a proof of concept, we validated the toehold switch technology using the previously published two sets (27B and 32B) of toehold switches for Zika virus (Pardee et al., 2016). The toehold switches were tested using two-plasmid system (Green, 2017). The sensor RNAs were expressed from ZIKV_Sensor_27B_LacZ and ZIKV_Sensor_32B_LacZ (Addgene, catalog numbers: 75006 and 75007) plasmids which were constructed from a medium copy plasmid pCOLA Duet vector bearing ColA origin and kanamycin resistance (Pardee et al., 2016). Both plasmids regulate *lacZ* as a reporter gene. The trigger plasmids were assembled in the lab and were expressed from pET15b plasmids having ColE1 origin and ampicillin resistance. The expression of the sensor and trigger plasmids is driven by T7 RNA polymerase.

2.4 *In silico* design of the toehold switches

Toehold switch structure has three main domains, the sensor domain, the loop domain and the reporter domain (Green et al., 2014). The sensor domain is defined by the complementary sequences of the target virus genome. The loop domain includes a ribosomal binding site (RBS) and the start codon. The reporter domain includes a reporter gene. This study used series B toehold switch described by Pardee et al. (2016) as it was reported to reduce gene leakage in the OFF state. Initially, previously published toehold switches for Zika virus were used as validation (Pardee et al. 2016). Subsequently, toehold switches were redesigned for MCMV and SCMV viruses. These modifications are described below.

Complete genomes of 43 MCMV and 73 SCMV isolates were downloaded from GenBank (10/2/2017) (Appendix A). Conserved regions among the isolate genomes of each species with 36 nt and longer were identified through multiple sequence alignment (Kearse et al., 2012) for species wide viral detection. All the possible 36-mers sequences for each virus were extracted using R (R Core Team, 2013) which formed the candidate sensor domain sequences (Table 2.2). Sequences with homology to the maize genome and other related viruses were filtered using BLAST, highly similar sequences criteria. Additionally, sequences with stop codons after the start codon were filtered out to prevent premature gene translation (Green et al., 2014).

A nucleic acid system design software called Nucleic Acid Package (NUPACK) (Zadeh et al., 2011) was used in the design and optimization of the toehold switch structure. NUPACK calculates the minimum free energy (MFE) of the RNA secondary structure to determine the stability of the switch assuming the secondary structures are pseudoknot free (Green et al., 2014; Zadeh et al., 2011). Using the NUPACK web platform, the switch sequences in Table 2.2 were selected as conforming to the toehold switch structure.

Table 2.2: MCMV and SCMV Toehold Switch Sequences and their Cognate Trigger Sequences

Trigger name	**Target region in the MCMV and SCMV genome (Trigger sequence 5'-3')**	**Alignment start site**	**Alignment end site**	**Sensor name**	**Sensor sequence (5' – 3')**
MC_T.1	CAGGTGGAACA GGCTATGGAAC	893	928	MC_S.1	CATCGCTCCATTCTGT TCCATAGCCTGTTCCA

	AGAATGGAGCGATG				CCTGGGACTTTAGAACAGAGGAGATAAAGATGCAGGTGGAACAAAACCTGGCGGCAGCGCAAAAGATGACCATG
MC_T.2	GGTGGAACAGGCTATGGAACAGAATGGAGCGATGAT	895	930	MC_S.2	ATCATCGCTCCATTCTGTTCCATAGCCTGTTCCACCGGACTTTAGAACAGAGGAGATAAAGATGGGTGGAACAGGAAACCTGGCGGCAGCGCAAAAGATGACCATG
MC_T.3	CAATACCCACTGGGGGTGGAGGATGTTGCCAGGATCG	3784	3819	MC_S.3	CGATCCTGGCAACATCCTCCACCCCAGTGGGTATTGGGACTTTAGAACAGAGGAGATAAAGATGCAATACCCACTAAACCTGGCGGCAGCGCAAAAGATGACCATG
MC_T.4	AATACCCACTGGGGTGGAGGATGTTGCCAGGATCGT	3785	3820	MC_S.4	ACGATCCTGGCAACATCCTCCACCCCAGTGGGTATTGGACTTTAGAACAGAGGAGATAAAGATGAATACCCACTGCAACCTGGCGGCAGCGCAAAAGATGACCATG
MC_T.5	GGAGGAACGGAAGATGACGGTGGTCCAAGCCTTGAG	192	231	MC_S.5	CTCAAGGCTTGGACCACCGTCATCTTCCGTTCCTCCGGACTTTAGAACAGAGGAGATAAAGATGGGAGGAACGGACAACCTGGCGGCAGCGCAAAAGATGACCATG
SC_T.1	AGATTTGGATGCTAGTGTTGAAGAGCAGTGCAAAGT	6808	6843	SC_S.1	ACTTTGCACTGCTCTTCAACACTAGCATCCAAATCTGGACTTTAGAACAGAGGAGATAAAGATGAGATTTGGATGTAACCTGGCGGCAGCGCAAAAGATGACCATG
SC_T.2	GATTTGGATGCTAGTGTTGAAGAGCAGTGCAAAGTT	6809	6844	SC_S.2	AACTTTGCACTGCTCTTCAACACTAGCATCCAAATCGGACTTTAGAACAGAGGAGATAAAGAT

					GGATTTGGATGCAAAC CTGGCGGCAGCGCAA AAGATGACCATG
SC_T.3	ATTTGGATGCTA GTGTTGAAGAG CAGTGCAAAGT TA	6810	6845	SC_S.3	TAACTTTGCACTGCTC TTCAACACTAGCATCC AAATGGACTTTAGAAC AGAGGAGATAAAGAT GATTTGGATGCTTAAC CTGGCGGCAGCGCAA AAGATGACCATG
SC_T.4	CCATGGACTGTT GGGATAACAAA ATTCTATGGAG GT	7493	7528	SC_S.4	ACCTCCATAGAATTTT GTTATCCCAACAGTCC ATGGGGACTTTAGAAC AGAGGAGATAAAGAT GCCATGGACTGTAAAC CTGGCGGCAGCGCAA AAGATGACCATG
SC_T.5	CATGGACTGTTG GGATAACAAAA TTCTATGGAGGT T	7494	7529	SC_S.5	AACCTCCATAGAATTT TGTTATCCCAACAGTC CATGGGACTTTAGAAC AGAGGAGATAAAGAT GCATGGACTGTTTAAC CTGGCGGCAGCGCAA AAGATGACCATG
SC_T.6	ACAAAATTCTAT GGAGGTTGGAA TCGCTTACTTGA G	7511	7546	SC_S.6	CTCAAGTAAGCGATTC CAACCTCCATAGAATT TTGTGGACTTTAGAAC AGAGGAGATAAAGAT GACAAAATTCTACAAC CTGGCGGCAGCGCAA AAGATGACCATG

2.5 Construction of toehold switches for MCMV and SCMV

Toehold switches for MCMV and SCMV were constructed using Gibson assembly (Gibson et al., 2009). This method uses two or more linear PCR products that are overlapping and join them in a single isothermal step. PCR primers (Table 2.3) were used to amplify and consequently add 30-bp homologous regions on the

vector backbones and on the plasmid inserts for Gibson assembly (Green, 2017). The primers used in this study were modified from the primers listed in RNA Nanostructures chapter (Green, 2017). These primer sequences were constructed using 30-bp homology domains which enable the addition of homology domains to both the plasmids and the inserts for Gibson assembly. The DNA plasmid inserts were purchased from Integrated DNA Technologies (IDT) as single strands. The switch DNA templates contain a conserved T7 promoter sequence, the switch sequence, 21-bp linker region and the first 9-bp of reporter gene sequence (*lacZ*) (Green, 2017). The trigger DNA templates contain the T7 promoter sequence, trigger sequence and 30-bp T7 terminator sequence.

Table 2.3: Primer Sequences Used in this Study

Primer Name	**Sequence (5'-3')**	**Description**
SFP.0	CGGCGAAACAAAGCTAAGAG AACCCCAACAGCGCTAATAC GACTCACTATAGGG	Amplify sensor sequences and adds an overhang at the 5' end for Gibson assembly
SRP.0	CATGGTCATCTTTTGCGCTG CCGCCAGGTT	Universal reverse primer for the sensors
TFP.0	GTAGTCTGCCAGTGGTTCGC GATATAGCTCGCGCTAATAC GACTCACTATAGGG	Amplify trigger sequences and adds an overhang at the 5' end for Gibson assembly
TRP.0	CCCGTTTAGAGGCCCCAAGG GGTTATGCTA	Universal reverse primer for the triggers

FP_pET_15b	TAGCATAACCCCTTGGGGC	Amplify pET_15b plasmid upstream region of homology
RP_pET_15b	GAGCTATATCGCGAACCACT GGCAGACTACCGAGATCTCG ATCCTCTACGC	Amplify pET_15b plasmid upstream region of homology
FP_ pCOLADuet	AACCTGGCGGCAGCGCAAAA GATGACCATGATTACGGATT CACTGGCC	Amplify pCOLADuet plasmid upstream region of homology
RP_ pCOLADuet	TGTTGGGGTTCTCTTAGCTT TGTTTCGCCGCATAAGGGAG AGCGTCGAGATC	Amplify pCOLADuet plasmid upstream region of homology

Standard PCR reactions utilized the Q5 High-Fidelity DNA Polymerase (NEB), NEBNext High-Fidelity 2X PCR Master Mix (NEB), 1 μM of each primer, 1 μM of sensor and trigger DNA templates, and 1-4 ng of plasmid DNA. The sensor and trigger reaction conditions comprised of initial denaturation at 98°C for 30 seconds followed by 35 cycles of 5 seconds of denaturation at 98°C, 15 seconds for annealing at 59°C, 15 seconds of extension at 72°C. The PCR products were then incubated for 5 minutes at 72°C for final extension. The PCR products were measured using 2% agarose gel. Amplification of the plasmids consisted of initial denaturation at 98°C for 30 seconds, followed by 35 cycles of 5 seconds of denaturation at 98°C, 15 seconds for annealing at 57°C, 2 minutes of extension at 72°C and a final extension at 72°C for 8 minutes. The PCR products were confirmed on a 0.8% agarose gel.

Gibson assembly master mix (NEB) was used to ligate the two PCR products by incubating the reaction at 50°C for 30 minutes to form circular sensor and trigger plasmids (Figure 2.1). Each recombinant plasmid was transformed into NEB 5 Alpha chemically competent cells (NEB) and grown on solid LB/agar medium supplemented with either kanamycin (50 mg/mL) or ampicillin (100 mg/mL) for 12 – 18 hours. The plasmids were then harvested using PureLink Quick Plasmid Miniprep Kit (Invitrogen).

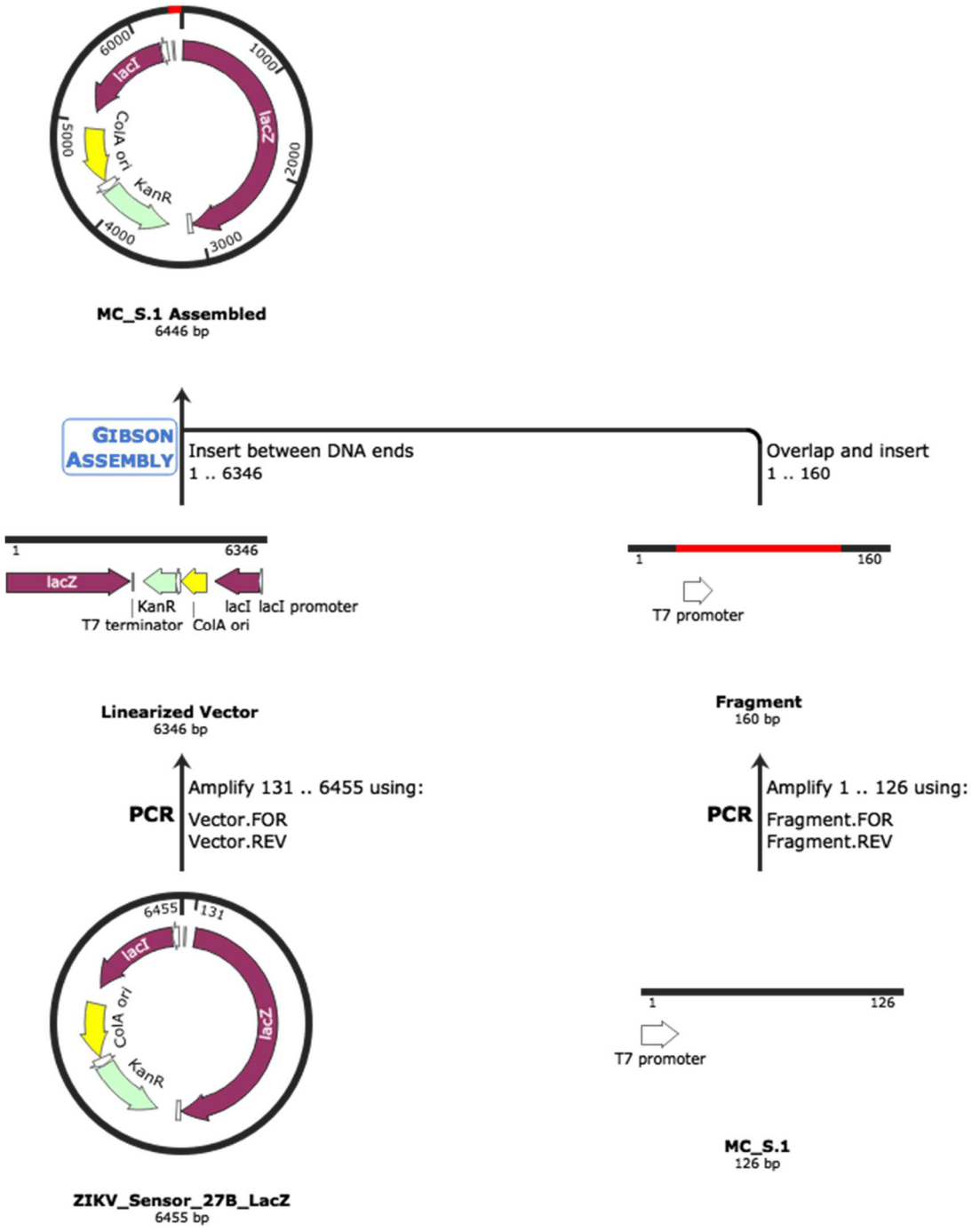


Figure 2.1: Schematic illustration of Gibson assembly procedure. The vector is linearized by PCR and consequently adds homologous regions on both the vector and insert. Source: SnapGene version 4.2.6.

2.6 Screening of toehold switches

The switch RNA regulates *lacZ* gene as the reporter protein which releases ß-galactosidase enzyme upon translation. X-gal, a soluble colorless analog of lactose compound, was added to the growth media to detect the presence of ß-galactosidase enzyme (Juers, Matthews, & Huber, 2012). X-gal is cleaved by ß-galactosidase to form 5-bromo-4-chloro-indoxyl which spontaneously dimerizes and oxidizes to form a blue insoluble pigment, 5,5'-dibromo-4,4'-dochloro-indigo. As a result of this reaction, cells with functional ß-galactosidase enzyme form blue colonies.

Every set of sensor and trigger plasmids were transformed into BL21-LacZ deficient chemically competent cells (Addgene) to evaluate the performance of each sensor. The transformed *E. coli* cells were spread on solid LB/agar medium with kanamycin (50 mg/mL) for the switch plasmids, ampicillin (100 mg/mL) for the trigger plasmids and kanamycin and ampicillin for the double transformant. The growth medium was supplemented with 0.1 mM IPTG to induce expression of *lacZ* gene and 20 mg/mL X-gal to detect the presence of ß-galactosidase enzyme. The cells were incubated at 37 °C for 12-16 hours.

2.7 Evaluation of selected toehold switches

Toehold switches were evaluated by measuring the concentration of ß-galactosidase enzyme in the *E. coli* cells (Griffith & Wolf, 2002). Ortho-Nitrophenyl-ß-galactosidase (ONPG), a colorimetric substrate, was used to detect ß-galactosidase activity. The ß-galactosidase enzyme hydrolyzes the ONPG substrate into galactose

and ortho-nitrophenol (ONP) compound which has a yellow color. The intensity of the yellow color produced was used to measure the enzymatic rate at 420 nm wavelength. The amount of ß-gal activity was calculated as Miller units using the formula;

$$\text{Miller units} = 1000 \times \frac{(A420-(1.75 \times A550))}{(t \times v \times A60\ )}$$

A420 is the absorbance of the yellow product ONP concentration; A550 is the scatter from the cell debris which when multiplied by 1.75 approximates the scatter observed at A420; t is reaction time in minutes; v is volume of cell lysate in milliliters and A600 reflects the cell concentration before lysis.

One milliliter of LB medium containing antibiotics was introduced in a 2.2 mL 96-well plate. Three colonies from each plate were randomly sampled and used for this experiment. Each well was inoculated with a single colony of either the switch or double transformant of the switch and trigger. Wells containing only LB medium with antibiotics served as negative controls. The cultures were covered with adhesive plate seals and incubated overnight at 37 °C with shaking at 250 rpm. Five microliters of the overnight cultures were transferred into a 96-well plate containing fresh 1 mL LB medium with antibiotics and incubated on a shaker at 37 °C until the OD600 of the cultures reached 0.1-0.2. IPTG (0.1 mM) was then added to each well to induce expression and incubated for 1.5 hours at 37 °C with shaking at 250 rpm. Cell growth was then arrested by adding 20 uL of chrolomphenical (5 mg/mL).

The culture density which is proportional to the protein content in the cell was measured in a microplate reader (Griffith & Wolf, 2002). Fifty microliters of culture were dispensed into a flat-bottomed microtiter plate containing 150 uL of ultrapure water in each well. The bacterial concentration was determined by measuring the A600 of the diluted culture using SpectraMax M2 microplate reader. The cells were then permeabilized by adding 100 uL of cell cultures from the original plate into a second 96-deep well plate containing 1 mL of Z buffer (60 mM Na2HPO4. 7H2O, 40 mM NaH2 PO4.H2O, 10 mM KCl, 1 mM MgSO4. 7H2O, 50 mM ß-mercaptoethanol), 0.1% sodium dodecyl sulfate (SDS) and 40 uL of chloroform. The mixture was carefully mixed by pipetting up and down and left to stand for chloroform to settle at the bottom.

Aliquots of 100 uL of permeabilized cells were transferred to two duplicate flat-bottomed microtiter plates. The reaction was started by adding 20 uL of ONPG (4 mg/ml) to each well using a repeater pipette. The reaction was incubated for 10 minutes and terminated by adding 50 uL of 1M Na_2CO_3. The A420 and A550 were then measured using SpectraMax M2 microplate reader. These values were transferred to R software and the Miller units calculated for switch alone and switch with trigger.

Toehold switches were evaluated using three criteria: performance, sensitivity and stability. The performance of the toehold switches was estimated by the fold change in ß-gal activity which was computed by dividing the Miller units for the switch with trigger by the Miller units of the switch alone. A Student's t-test was used

to calculate the statistical significance of the toehold switches in the presence and absence of the trigger at P=.05. The sensitivity of the sensors was measured by the mean differences of the Miller units (Miller units of the switch propagated with the corresponding trigger minus switch alone) and Cohen's d effect size. Cohen's d was calculated by the mean difference divided by the pooled standard deviation (Cohen, 1988). In his book, Cohen suggested that an effect size of 0.2 to be considered a small effect size, 0.5 a medium effect size and 0.8 a large effect size. The stability of the sensors was predicted by the gene leakiness (ß-gal activity of the switch when propagated alone) and by estimating the variability across the replicates. Simple regression was also used to determine the relationship between the MFE of the switches and the gene leakiness.

Chapter 3

RESULTS

3.1 Proof of concept

To confirm that the functional screen used previously for Zika toehold sensors could be independently implemented for MLN viruses (MCMV and SCMV), validation tests were performed using published Zika sensors (Pardee et al., 2016). When switch plasmids, ZIKV_Sensor_27B_LacZ and ZIKV_Sensor_32B_LacZ, were co-transformed with their corresponding trigger plasmids in the presence of X-gal and IPTG, as expected, blue colonies were formed. This indicated that the trigger RNA sequence unlocked the toehold switch into an on-state for translation of the *lacZ* reporter gene (Figure 3.1). In the absence of the trigger sequence, white or very light blue colonies were detected for both toehold switches, indicating minor instability of the sensors hairpin structure which is consistent with prior results (Pardee et al., 2016).

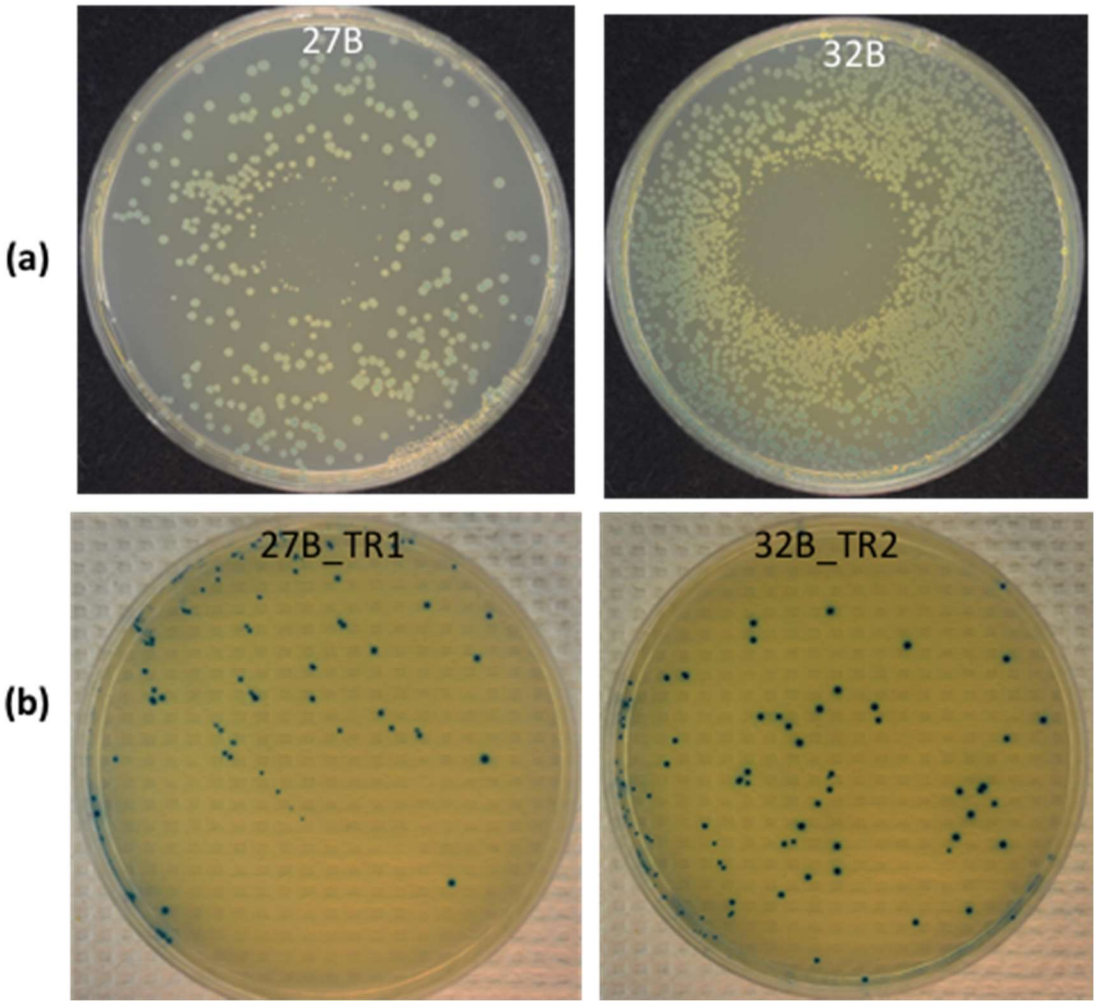


Figure 3.1: Zika toehold switches on LB agar medium containing X-gal and IPTG. (a) Each toehold-switch plasmid propagated without the corresponding trigger plasmid. (b) Co-transformation of each toehold-switch plasmid with the corresponding trigger plasmid.

3.2 Toehold switch design for MCMV and SCMV

The trigger sequence for a toehold switch is 36-nts long. A whole-genome sequence alignment for all 43 isolates of MCMV identified 19 non-overlapping conserved regions of 36-nts and longer. No such conserved regions were detected

among the 73 SCMV isolates. Therefore, nine SCMV isolates from Africa were used which contained 17 conserved regions of 36-nts and longer were used (Appendix A).

The conserved sequences were analyzed using the filter steps in Table 3.1. All possible 36-mers of each virus were extracted using a custom R script resulting to 380 MCMV and 433 SCMV sequences. There was no sequence homology between MCMV sequences and other viruses; however, 224 SCMV sequences were filtered due to perfect match with other related viruses in a BLAST analysis. These sequences formed the trigger sequences and their complementary sequences formed the toehold domain. The toehold domain sequence was analyzed for complementarity to *lacZ* gene sequence to prevent formation of secondary structures. There were no sequences filtered for both viruses. The toehold switch sequence was assembled and analyzed for stop codons after the start codon to avoid premature gene translation which resulted to 283 MCMV and 209 SCMV sequences (Table S1). The integrity of the switch structure and the interaction between the trigger and the switch of the resulting sequences (Table S1) were analyzed using the NUPACK web platform (Zadeh et al., 2011). NUPACK identified 5 MCMV (Table 2.2) and 51 SCMV sequences (Appendix B) that conform to the toehold switch structure.

Table 3.1: Number of MCMV and SCMV Sequences After the Filter Steps.

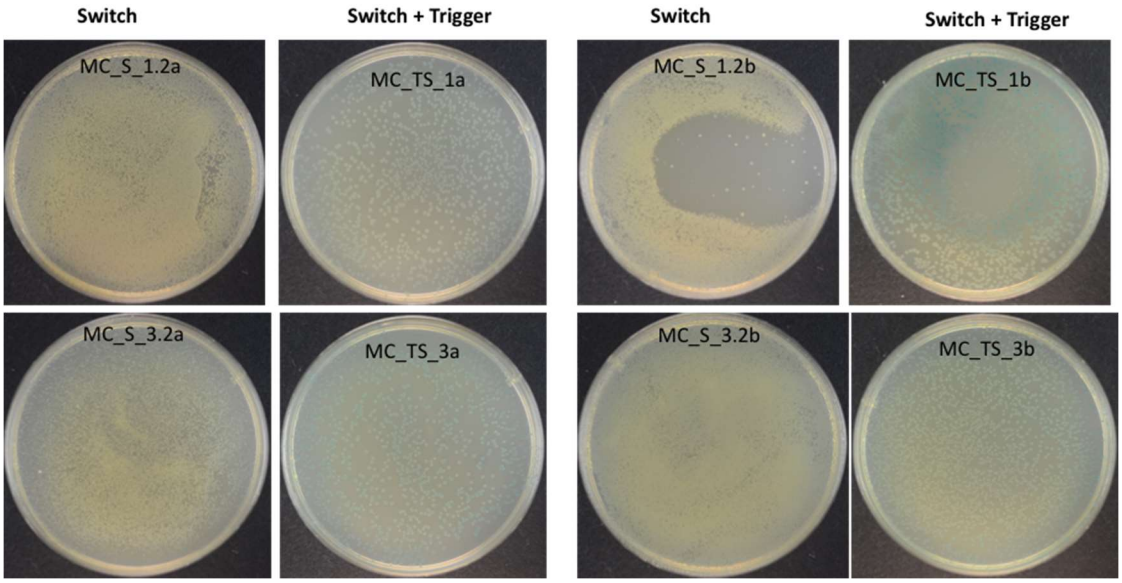
MLN Sequences Filter Process	**Number of Sequences**	
	MCMV	SCMV
Isolates	43	9
Conserved regions	19	17
36-nt triggers	380	433
BLAST analysis	380	263
Complementarity of toehold domain to *lacZ* gene sequence	380	263
Presence of stop codons after the start codon	283	209
NUPACK analysis	5	51

Five MCMV and six SCMV toehold switches (Table 2.2) were evaluated using an *in vitro* screen and quantitatively evaluated using three criteria: performance, sensitivity and stability.

3.3 *In vitro* screen for functional toehold switches

The candidate trigger and switch sequences were amplified in two replicates and subsequently analyzed independently. The switch and trigger plasmids were evaluated using blue white screen in two biological replicates. True to form, all trigger plasmids formed white colonies when propagated alone as they do not encode the *lacZ* gene. White colonies were observed in three MCMV sensors propagated with the

switch plasmid alone. This was expected due to the hairpin structure in the switch which prevents translation of the *lacZ* gene. Two of the MCMV sensors formed light blue colonies showing instability of the hairpin structure. When propagated with the corresponding trigger plasmids, all but one (M2) of the MCMV sensors formed blue colonies indicating translation of the reporter gene. The blue white screen of one MCMV sensor (M4) was not consistent between replicates as one replicate formed white colonies while the other formed light blue colonies in the singly transformed cells (Figure 3.2, Appendix C).



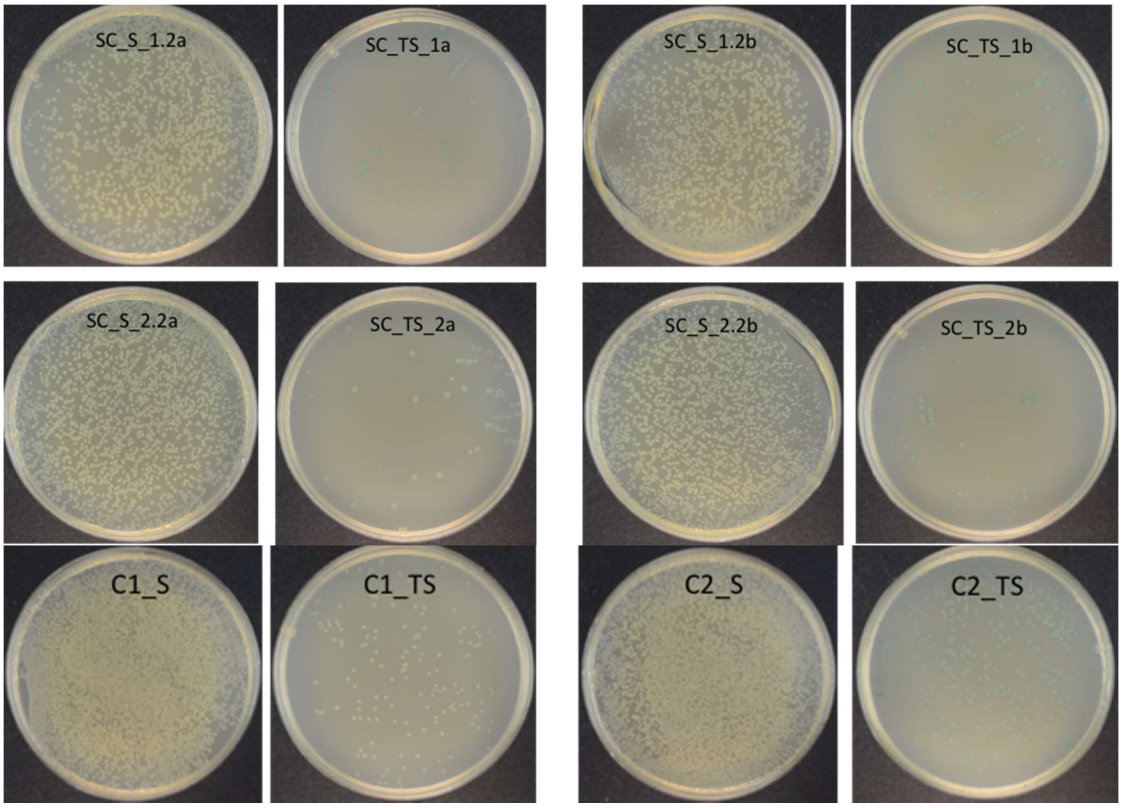


Figure 3.2: MCMV and SCMV toehold sensors on LB agar media containing X-gal and IPTG. Each toehold- switch plasmid propagated without the corresponding trigger plasmid (switch) and co-transformed with the corresponding trigger plasmid (switch + trigger). Sensor pairs are placed side by side.

All but one of the SCMV sensors propagated with the switch plasmid alone developed a light blue pigment indicating instability of the hairpin structure. Consequently, all SCMV sensors except one (S4a) formed blue colonies when the switch plasmids were co-transformed with their corresponding trigger plasmids. One of the SCMV sensors (S4) had inconsistent results in the doubly transformed cells as one of the replicates formed white colonies while the other formed blue colonies (Appendix C).

The number of colonies formed in single transformations and in double transformations differed greatly. Double transformations yielded fewer colonies than the single transformations since only 10 - 20 % of the competent cells are able to take up two or more plasmid molecules in a single cell (Weston et. al., 1979).

3.4 Toehold switch sensors performance

The performance of the toehold switch sensors is equivalent to the ß-Gal activity of the sensors measured in Miller units and the amount of fold change. Miller (1972) explained that uninduced *E. coli* cell produces approximately one Miller unit while a fully induced *E. coli* cell produces approximately 1000 Miller units. The performance of each sensor with its corresponding trigger RNA were quantitatively evaluated in single transformed and co-transformed *E. coli* cells using three randomly sampled colonies from each replicate. The previously published Zika toehold switches (Pardee et al., 2016) were used as positive controls. Both controls had no ß-Gal activity in the absence of the trigger RNAs although the blue white screen showed light blue colonies. In the presence of their cognate trigger RNAs, C1 produced more than 1078 Miller units and a fold change of 11while C2 produced 2262 Miller units and a fold change of 44 (Figure 3.3, Figure 3.4).

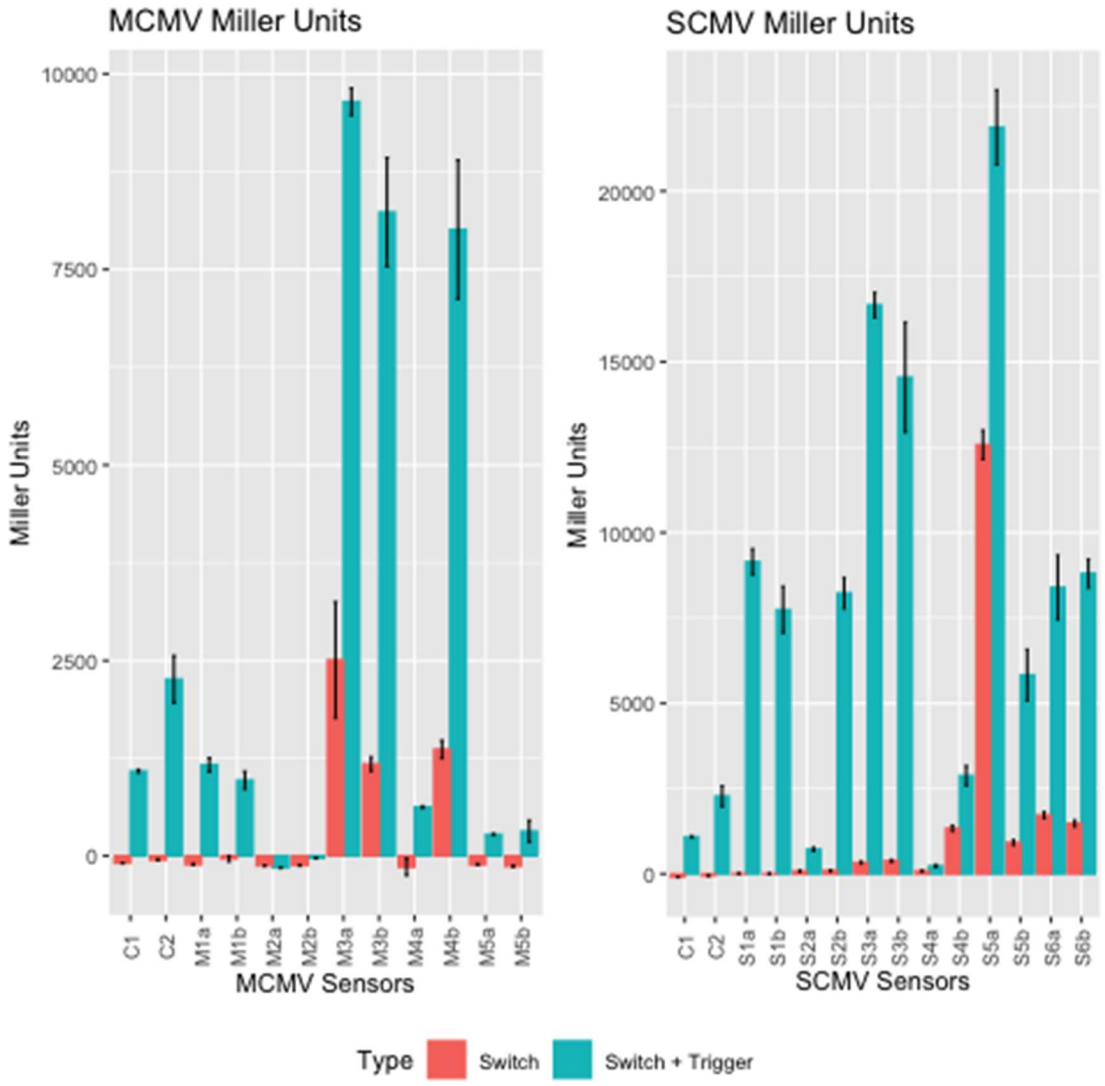


Figure 3.3: MCMV and SCMV sensors ß-galactosidase activity average measurements (in Miller units) in the absence (switch) and presence (switch + trigger) of their cognate trigger RNA. The means are calculated from ß-galactosidase activity of three single colonies of each sensor in the absence and presence of trigger RNAs. Error bars represent standard error of the mean.

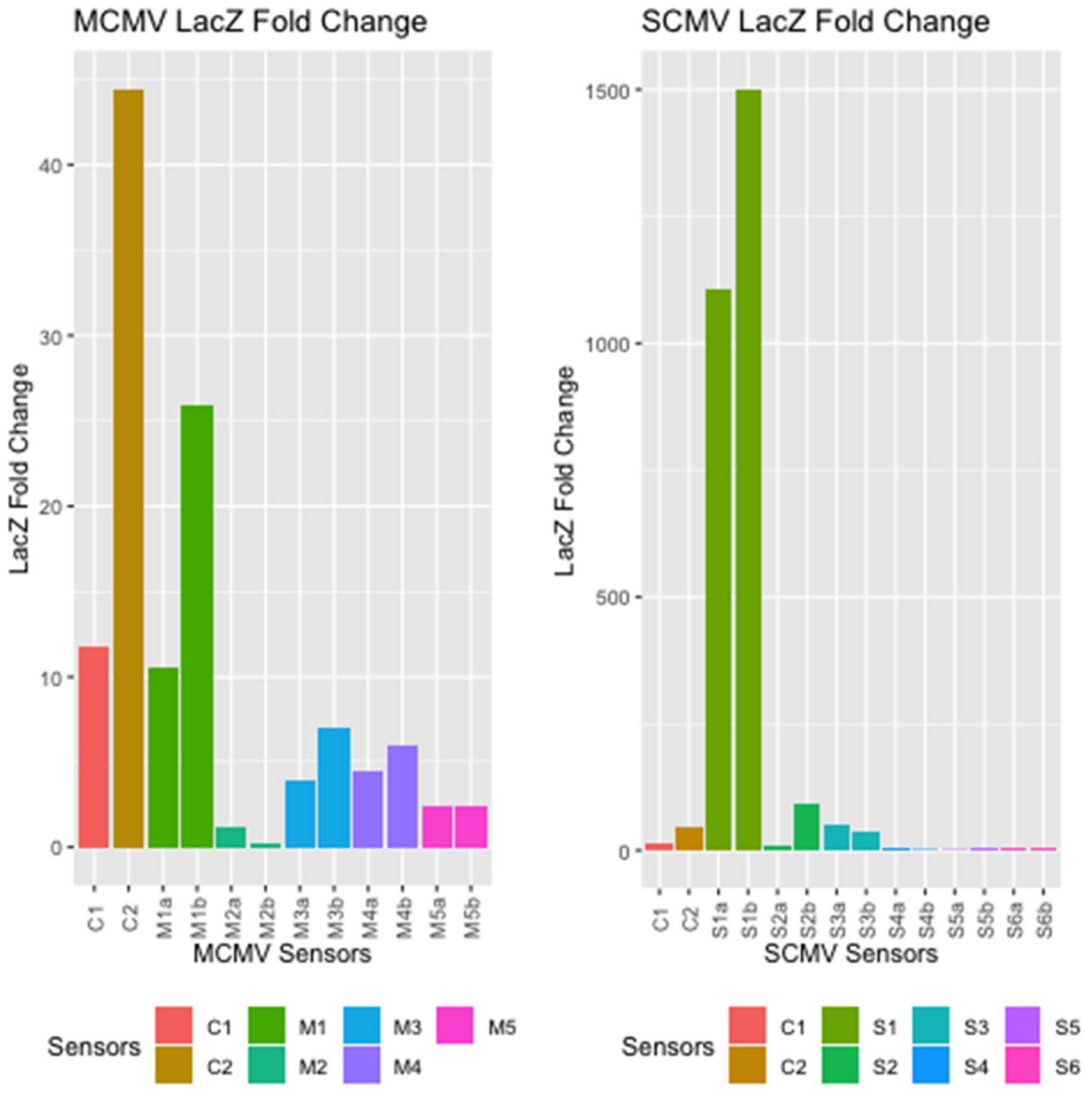


Figure 3.4: MCMV and SCMV sensors fold change determined by the ratio of ß-galactosidase activities in the presence and absence of trigger RNAs.

Two of the five MCMV sensors (M3 and M4) had ß-Gal activity in the absence of their cognate trigger RNAs indicating the switch hairpin structure was not

stable. All but one of the MCMV sensors showed ß-Gal activity in the presence of their cognate trigger RNA (Figure 3.3). Sensor M2 showed no ß-Gal activity. This was in line with the visual assessment of the sensors. In comparison to the fold changes, taking the sensor with less activation ration between the two biological replicates, all but one MCMV sensors had 5-fold change or less (Figure 3.4). The best performing MCMV sensor was M1 as it had the highest fold change of 10 and had no ß-Gal activity in the absence of the trigger RNA. A Student's t-test showed that there was significant difference between the sensor performance means in presence and absence of trigger RNAs at 95% confidence level for all of the MCMV and SCMV sensors apart from M5b sensor (Table 3.1).

All SCMV sensors had ß-Gal activity when treated with the sensor alone (Figure 3.3). This was expected as the blue white screen yielded light blue colonies for all sensors without the trigger RNAs. However, two sensors (S1 and S2) had low ß-Gal activity of less than 100 Miller units. The best performing sensor, S1, had a fold change of more than 1000 and produced more than 7700 Miller units in the presence of trigger RNA, compared to a fold change of 9 for S2 sensor. All the other SCMV sensors had a fold change of 6 or below apart from S3 which had a fold change of more than 30 (Figure 3.4).

Table 3.2: MCMV Sensors Data Summary

Name	Mean (SEM)	Mean Background	Mean Difference	MFE_Switch (kcal/mol)	Coefficient of Variation	P-Values	Effect Size
M1a	1157.2 (88.4) **	-109.3	1266.5	-19.3	13.2	0.005	16.0
M1b	962.0 (113.0) *	-37.1	999.1	-19.3	20.4	0.02	7.5
M2a	-151.0 (3.7) *	-124.5	-26.5	-20.4	4.3	0.03	5.2
M2b	-29.2 (3.6) **	-120.8	91.5	-20.4	21.3	0.005	13.1
M3a	9643.1 (173.1) **	2508.7	7134.4	-18.4	3.1	0.01	8.6
M3b	8232.2 (695.2) **	1167.4	7064.8	-18.4	14.6	0.01	10.1
M4a	620.6 (10.7) *	-141.5	762.1	-18.4	3.0	0.02	3.3
M4b	8010.2 (888.0) *	1354.4	6655.7	-18.4	19.2	0.02	7.6
M5a	272.3 (9.5) **	-111.1	383.4	-23.1	6.0	0.0004	25.5
M5b	312.2 (138.3)	-132.3	444.5	-23.1	76.7	0.09	2.8
C1	1078.3 (20.8) **	-91.8	1170.2	-19.9	3.3	0.0003	26.0
C2	2262.0 (300.4) *	-50.9	2312.9	-19.9	23.0	0.02	8.4

* significant at <p 0.05; ** significant at <p 0.01

Table 3.3: SCMV Sensors Data Summary

Name	Mean (SEM)	Mean Background	Mean Difference	MFE_Switch (kcal/mol)	Coefficient of Variation	P-Values	Effect Size
S1a	9145.0 (377.1) **	8.3	9136.7	-15.4	7.1	0.002	3.6
S1b	7742.7 (681.6) **	5.2	7737.5	-15.4	15.2	0.008	12.8
S2a	716.2 (53.4) **	76.1	640.1	-17.1	12.9	0.009	9.2
S2b	8228.8 (456.2) **	91.5	8137.3	-17.1	9.6	0.003	8.7
S3a	16662.6 (365.4) **	329.4	16333.1	-16.5	3.8	0.0004	20.6

S3b	14546.8 (1611.2) **	379.0	14167.8	-16.5	19.2	0.01	8.8
S4a	229.8 (39.2) *	81.3	148.5	-19.3	29.6	0.04	1.0
S4b	2861.4 (287.7) *	1326.0	1535.4	-19.3	17.4	0.03	3.5
S5a	21866.6 (1093.5) *	12570.6	9296.0	-17.3	8.7	0.02	7.0
S5b	5830.6 (749.3) *	911.2	4919.4	-17.3	22.3	0.02	2.0
S6a	8391.5 (944.0) *	1712.4	6679.1	-12.5	19.5	0.02	4.5
S6b	8800.6 (417.9) **	1470.4	7330.2	-12.5	8.2	0.004	15.5
C1	1078.3 (20.8) **	-91.8	1170.2	-19.9	3.3	0.0003	26.0
C2	2262.0 (300.4) *	-50.9	2312.9	-18.8	23.0	0.02	8.4

* significant at <p 0.05; ** significant at <p 0.01

3.5 Sensitivity of the toehold switches

Here, sensitivity is defined as a measure of the effect of the trigger on the switch determined by the mean difference (calculated as Miller units for the switch subtracted from Miller units for the switch and trigger) and the Cohen's d effect size. The mean difference for the MCMV sensors ranged from -26.5 to 7134.4 Miller units for M2a and M3a respectively (Table 3.1). The low mean difference observed in M2 sensor was attributed to lack of ß-gal activity when the switch was treated with its cognate trigger. The mean difference for SCMV sensors ranged from 148.5 to 16333.1 Miller Units for S4a and S3a respectively (Table 3.2).

The calculation of Cohen's d comparing sensors treated with their cognate triggers and without the triggers showed large effect sizes for all MCMV and SCMV sensors (Table 3.1; Table 3.2) using the classification suggested by Cohen (Cohen, 1988). M5a sensor had the largest effect size amongst MCMV sensors of 25.5

however, its mean difference was low (383.4 Miller units) compared to M3a. S3a had the largest effect size amongst SCMV sensors of 20.6 and the highest mean difference therefore, we can conclude that it was the most sensitive SCMV sensor.

3.6 Stability of the toehold switches

Stability of the toehold switches is the ability of the switches to remain in the off-state and was determined by gene leakage and variability across replicates of the switches. The mean background (gene leakage) for MCMV sensors ranged from 141.5 to 2508.7 Miller units for M4a and M3a respectively (Table 3.1). Only two MCMV sensors, M3 and M4b had gene leakage indicating that the switch did not remain in the off-state. The SEM for MCMV sensors ranged from 3.6 (M2b) to 888.0 (M4b). M2 had the least variability as it remained in the off-state even in the presence of the cognate trigger. Coefficient of variation (CV) was also used to determine the variability of the sensors. The CV for MCMV sensors ranged from 3% (M4a) to 76.7% (M5b). This differed greatly between the two biological replicates for each sensor apart from M1.

All SCMV sensors had ß-gal activity in the absence of the cognate trigger (mean background) ranging from 5.2 to 12570.6 Miller units for S1b and S5a respectively indicating gene leakage for all sensors (Table 3.2). However, S1and S2 had a low mean background of less than 100 Miller units as mentioned in section 3.4. Similarly, S4a had a low mean background of 81.3 Miller units but S4b had a high mean background of 1326.0 Miller units. The SEM for SCMV sensors ranged from 39.2 (S4a) to 1093.5 (S5a). In addition, the CV for SCMV sensors ranged from 29.6%

(S4a) to 3.8%(S3a). Both controls did not have any gene leakage in contrast to the blue white results of light blue colonies. C1 had less variability than C2.

The regression analysis showed a positive correlation between the minimum free energy (MFE) of the switches and the average background effect (estimate of stability) (Figure 3.5). One extreme value (S5a) was removed from the analysis.

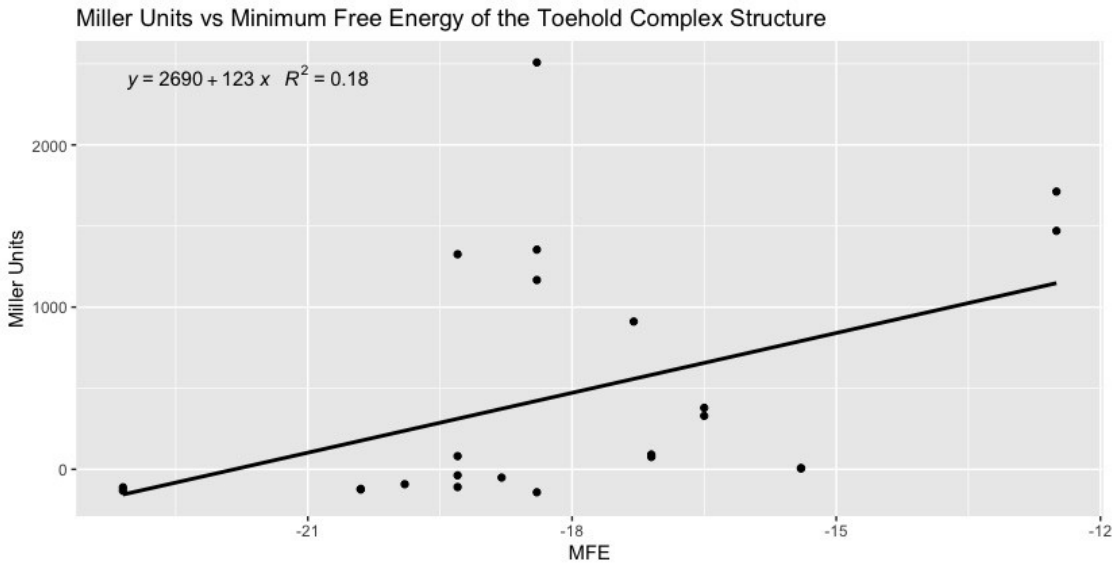


Figure 3.5: Scatterplot showing relationship between Miller units of the toehold switches and the Minimum Free Energy (MFE) of the toehold switches structure. The line represents the regression line.

Chapter 4
DISCUSSION AND CONCLUSIONS

This study demonstrated the efficacy of the synthetic biology construct called toehold switch riboregulators ("sensors") for the detection of the MCMV and SCMV viruses that cause MLN. After in silico characterization of hundreds of candidate trigger sequences that target conserved sections of the MCMV and SCMV genomes, five MCMV and six SCMV sensors were designed and evaluated using a cell-based screen (Table 2.2). Three criteria, performance, sensitivity and stability were used to assess the merit for moving forward in the development of field-based sensors for MLN-viral detection.

In reference to the genome maps of the MLN causing viruses published by Wamaitha, et al. (2018), the MCMV target sequences, M1, M2 and M5, mapped on two overlapping ORFs, p32 and p50, of MCMV genome. The target sequences, M3 and M4 mapped on other two overlapping regions, p31 and capsid protein (Figure 4.1). The SCMV target sequences, S1, S2 and S3 mapped between N1a and N1b regions of the SCMV genome while S4, S5 and S6 mapped in the N1b region (Figure 4.2).

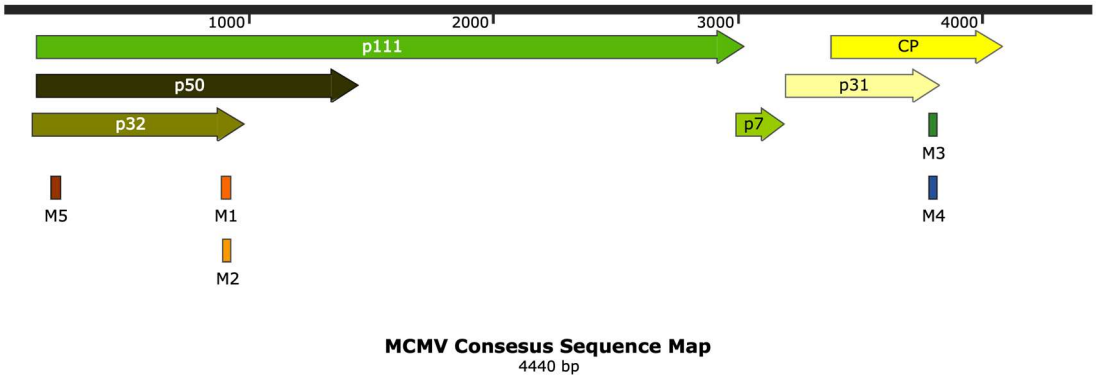


Figure 4.1: Schematic representation of the MCMV genome and the locations of the MCMV sensors. The open reading frames (ORFs) are represented by the arrows and are labelled while the sensors are indicated by the boxes. The map is generated using SnapGene version 5.1.0.

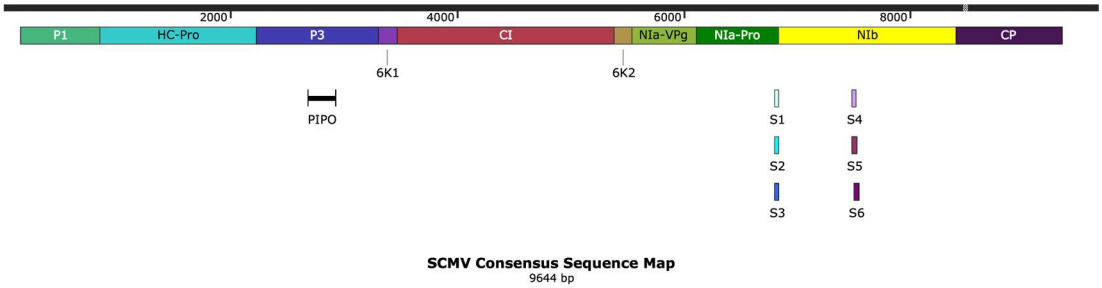


Figure 4.2: Illustration of SCMV sensors mapped on the SCMV genome. The open reading frames (ORFs) are indicated by the open boxes and are labelled. The closed boxes represent the sensors mapped on the SCMV consensus sequence. The map is generated using SnapGene version 5.1.0.

Taken together, the results indicate that M1 was the best performing and most stable sensor amongst the MCMV sensors. M3 was the most sensitive sensor;

however, it did not remain in the off-state in the absence of the trigger. Similarly, S1 was the best performing, most sensitive and stable sensor amongst SCMV sensors.

Further research is needed to test the MLN toehold switches in a cell-free system to validate their use and determine their sensitivity in terms of detectable concentrations of the viruses. Sensitivity analysis can be done using different concentrations of the trigger RNA equivalent to the viral RNA concentrations in an infected maize plant. A study by Wu et al., (2013) detected a concentration of 256.0 ng/mL of purified MCMV from an infected maize leaf. There is no documentation of SCMV concentration in infected maize plants. However, minimum SCMV concentrations of 100 pg/uL have been detected by conventional reverse transcription PCR method (Keizerweerd et al., 2015). Using these concentrations as threshold values, serial dilutions can be made to determine the sensitivity of MLN toehold switches.

The major advantages of the toehold switch platform over the current MLN diagnostic techniques are: (i) the rapid design and redesign for new pathogens or strains; (ii) it can be designed to detect multiple pathogens; (iii) it is highly sensitive; and (iv) for the application setting, it is portable and requires limited expertise. Toehold switches have been designed to detect Zika virus (Pardee et al., 2016), Ebola virus (Pardee et al., 2014), *Clostridium difficile* bacteria (Takahashi et al., 2018) and most recently norovirus (Ma et al., 2018). The capability for rapid redesign of new sensors makes the toehold system attractive for MCMV and SCMV and other RNA viruses due to their relatively high mutation rate (Duffy, 2018). Developing a sensor in

a cell-free system would take approximately five days as proposed by Pardee et al. (2014). Green et al., (2014) demonstrated multiplexed regulation of 12 toehold switches which were independently activated by their cognate trigger RNAs. The Zika virus toehold switch detected the virus from plasma at a concentration of 2.8 fM which makes the system highly sensitive. In addition, the toehold switch system coupled with genome editing technologies such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) can be used to more precisely discriminate between pathogen lineages (Pardee et al., 2016). The virus specific toehold switches can be freeze-dried on cell-free paper that enables a platform for visual detection of the virus that is inexpensive and field-deployable (Pardee, et al., 2014).

The toehold switch technology could be used to detect virtually any RNA sequence; therefore, it can be designed to detect any pathogen from plant tissue extracts. For instance, toehold sensors can be developed to detect avirulence genes in the plant fungus. This technology offers a promising solution for rapid, low-cost detection of plant pathogens in remote settings. However, there is need for optimization of sample preparation for RNA isolation to enable the technology to be fully field-deployable.

REFERENCES

Abate, T., Fisher, M., Abdoulaye, T., Kassie, G. T., Lunduka, R., Marenya, P., & Asnake, W. (2017). Characteristics of maize cultivars in Africa: How modern are they and how many do smallholder farmers grow? *Agriculture and Food Security*, *6*(1), 1–17. https://doi.org/10.1186/s40066-017-0108-6

Abate, T., & Wawa, B. (2015). September 2015 About the Bulletin Maize in Kenya : Chance for Getting Back to Former Glory ?, *4*(3), 1–4. https://doi.org/10.1007/s12571-015-0488-z

Adams, I. P., Harju, V. ., Hodges, T., Hany, U., Skelton, A., Rai, S., … Boonham, N. (2014). First report of maize lethal necrosis disease in Rwanda. *New Disease Reports*, *29*, 22. https://doi.org/10.5197/j.2044-0588.2014.029.022 First

Alvarez, R. A., & Mendoza, A. A. (2016). Occurrence of maize lethal necrosis in Ecuador : a disease without boundaries ? *European Journal of Plant Pathology*, *1914*, 705–710. https://doi.org/10.1007/s10658-016-0943-5

Brandes, E. W. (1919). MOSAIC DISEASE OF SUGAR CANE.

Cabanas, D., Watanabe, S., Higashi, C. H. V, Bressan, A., Bockelman, D. L., Uyemoto, J. K., … Zhou, X. P. (2013). Dissecting the mode of maize chlorotic mottle virus transmission (Tombusviridae: Machlomovirus) by Frankliniella williamsi (Thysanoptera: Thripidae). *Journal of Economic Entomology*, *106*(1), 16–24. https://doi.org/10.1603/ec12056

Castillo, J., & Hebert, T. T. (1974). A new virus disease of maize in Peru. *Fitopatologia*, *9*(2), 79–84. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19751319459

Ceroni, F. (2015). Editorial – Synthetic biology : engineering complexity and refactoring cell capabilities, *3*(August), 1–2. https://doi.org/10.3389/fbioe.2015.00120

Clark, M. F., & Adams, A. N. (1977). Characteristics of the Microplate Method of Enzyme-Linked Immunosorbent Assay for the Detection of Plant Viruses. *Journal of General Virology*, *34*(3), 475–483.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.

Colombatti, F., Gonzalez, D. H., & Welchen, E. (2014). Plant mitochondria under pathogen attack: A sigh of relief or a last breath? *Mitochondrion*, *19*(PB), 238–244.

https://doi.org/10.1016/j.mito.2014.03.006

Duffy, S. (2018) Why are RNA virus mutation rates so damn high? *PLOS Biology 16(8)*: e3000003. https://doi.org/10.1371/journal.pbio.3000003

Duo Ma, Luhui Shen, Kaiyue Wu, Chris W Diehnelt, Alexander A Green. (2018). Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology*, 3(1), ysy018, https://doi.org/10.1093/synbio/ysy018

Frenken, A. (2013). Maize Lethal Necrosis Disease (MLND) A snapshot (FSNWG), *7*(7), 2.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/NMETH.1318

Gilbertson, R. L., & Lucas, W. J. (1996). How do viruses traffic on the 'vascular highway'? *Trends in Plant Science*, *1*(8), 260–268. doi: https://doi.org/10.1016/1360-1385(96)10029-7

Gitonga, K. (2017). 2017 Kenya Corn, Wheat and Rice Report Grain and Feed Annual Kenya, 6. Retrieved from https://gain.fas.usda.gov/Recent GAIN Publications/Grain and Feed Annual_Nairobi_Kenya_3-15-2017.pdf

Green, A. A. (2017). RNA Nanostructures, *1632*, 285–302. https://doi.org/10.1007/978-1-4939-7138-1

Green, A. A., Silver, P. A., Collins, J. J., & Yin, P. (2014). Toehold switches: De-novo-designed regulators of gene expression. *Cell*, *159*(4), 925–939. https://doi.org/10.1016/j.cell.2014.10.002

Griffith, K. L., & Wolf, R. E. (2002). Measuring [illegible] -Galactosidase Activity in Bacteria : Cell Growth , Permeabilization , and Enzyme Assays in 96-Well Arrays, *402*, 397–402. https://doi.org/10.1006/bbrc.2001.6152

Hipper, C., Brault, V., Ziegler-Graff, V., & Revers, F. (2013). Viral and cellular factors involved in Phloem transport of plant viruses. *Frontiers in plant science*, *4*, 154. https://doi.org/10.3389/fpls.2013.00154

Huang, L., Yuan, Z., Liu, P., & Zhou, T. (2015). Effects of promoter leakage on

dynamics of gene expression, 1–12. https://doi.org/10.1186/s12918-015-0157-z

Jensen, S. G., Wysong, D. S., Ball, E. M., & Higley, P. M. (1991). Seed Transmission of Maize Chlorotic Mottle Virus. Plant Disease. https://doi.org/10.1094/PD-75-0497

Jiang, X. Q., Meinke, L. J., Wright, R. J., Wilkinson, D. R., & Campbell, J. E. (1992). Maize chlorotic mottle virus in Hawaiian-grown maize : vector relations , host range and associated viruses, 248–254.

Juers, D. H., Matthews, B. W., & Huber, R. E. (2012). LacZ b -galactosidase : Structure and function of an enzyme of historical and molecular biological importance, *21*. https://doi.org/10.1002/pro.2165

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keizerweerd, A. T., Chandra, A., & Grisham, M. P. (2015). Development of a reverse transcription loop-mediated isothermal amplification (RT-LAMP) assay for the detection of Sugarcane mosaic virus and Sorghum mosaic virus in sugarcane. *Journal of Virological Methods*, *212*, 23–29. doi: 10.1016/j.jviromet.2014.10.013

Krczal, G., & Wassenegger, M. (2013). Chapter 492 - Potyvirus Helper Component Proteinase. In N. D. Rawlings & G. Salvesen (Eds.), *Handbook of Proteolytic Enzymes (Third Edition)* (Third Edit, pp. 2187–2192). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-382219-2.00491-9

Kulkarni, H. Y. (1973). Notes on East African Plant Virus Diseases. *East African Agricultural and Forestry Journal*, *39*(2), 158–164. https://doi.org/10.1080/00128325.1973.11662633

Kumar, D., Kumar, R., Hyun, T.K. et al. J Plant Res (2015) 128: 37. https://doi-org.udel.idm.oclc.org/10.1007/s10265-014-0683-6

Li, L., Wang, X., & Zhou, G. (2011). Effects of seed quality on the proportion of seed transmission for *Sugarcane mosaic virus* in maize. *Cereal Research Communications*, *39*(2), 257–266. https://doi.org/10.1556/CRC.39.2011.2.9

Lima, J. A. A., Nascimento, A. K. Q., Radaelli, P., & Purcifull, D. E. (2009). Serology Applied to Plant Virology. *Serological Diagnosis of Certain Human, Animal and Plant Diseases*, 72–94. https://doi.org/10.5772/2140

Lukanda, M., Owati, A., Ogunsanya, P., Valimunzigha, K., Katsongo, K., Ndemere, H., & Kumar, P. L. (2014). First Report of Maize chlorotic mottle virus Infecting Maize in the Democratic Republic of the Congo. *Plant Disease*. https://doi.org/10.1094/PDIS-05-14-0484-PDN

Maheaswari, R., Kshirsagar, J. T., & Lavanya, N. (2016). Polymerase chain reaction: A molecular diagnostic tool in periodontology. *Journal of Indian Society of Periodontology*, *20*(2), 128–135. doi:10.4103/0972-124X.176391

Mahuku, G., Wangai, A., Sadessa, K., Teklewold, A., Wegary, D., Ayalneh, D., … Prasanna, B. M. (2015). First Report of *Maize chlorotic mottle virus* and Maize Lethal Necrosis on Maize in Ethiopia. *Plant Disease*. https://doi.org/10.1094/PDIS-04-15-0373-PDN

Mahuku, George, Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., … Redinbaugh, M. G. (2015). Maize Lethal Necrosis (MLN), an Emerging Threat to Maize-Based Food Security in Sub-Saharan Africa. *Phytopathology*, *105*(7), 956–965. https://doi.org/10.1094/PHYTO-12-14-0367-FI

Makumbi, D., & Wangai, A. (2012). Maize lethal necrosis (MLN) disease in Kenya and Tanzania: Facts and actions. URL http://www.fao.org/fileadmin/user_upload/drought/docs/1 Maize Lethal Necrosis-(MLN) Facts and Actions.pdf (6.4.16), 4. Retrieved from http://www.fao.org/fileadmin/user_upload/drought/docs/1 Maize Lethal Necrosis-(MLN) Facts and Actions.pdf

Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor (New York): Cold Spring Harbor Laboratory Press. 466 p.

Nault, L. R. (1978). Transmission of Maize Chlorotic Mottle Virus by Chrysomelid Beetles. *Phytopathology*, *68*(7), 1071. https://doi.org/10.1094/Phyto-68-1071

Niblett, C. L., & Claflin, L. E. (1978). Corn lethal necrosis - a new virus disease of corn in Kansas. *Plant Disease Reporter*, *62*(1), 15–19. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19781343650

Nutter, R. C., Scheets, K., Panganiban, L. C., & Lommel, S. A. (1989). The complete nucleotide sequence of the maize chlorotic mottle virus genome. *Nucleic Acids Research*, *17*(8), 3163–3177. https://doi.org/10.1093/nar/17.8.3163

Onono, P. A., Wawire, N. W. H., & Ombuki, C. (2013). The response of maize production in Kenya to economic incentives, *2*(2), 530–543.

Padmanabhan, S., Banerjee, S., & Mandi, N. (2011). Screening of Bacterial Recombinants: Strategies and Preventing False Positives. *Molecular Cloning - Selected Applications in Medicine and Biology*. https://doi.org/10.5772/22140

Pardee, K., Green, A.A., Ferrante, T., Cameron, D.E., DaleyKeyser, A., Yin, P., and Collins, J.J. (2014). Paper-based synthetic gene networks. Cell 159, 940–954.

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., … Collins, J. J. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, *165*(5), 1255–1266. https://doi.org/10.1016/j.cell.2016.04.059

Posthuma-Trumpie, G.A., Korf, J. & van Amerongen, A. Anal Bioanal Chem (2009) 393: 569. https://doi.org/10.1007/s00216-008-2287-2

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., … Michel, J. (2016). Ebola surveillance. *Nature*, *530*(7589), 228–232. https://doi.org/10.1038/nature16996

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.r-project.org/

Rodríguez, R. A., Pepper, I. L., & Gerba, C. P. (2009). Application of PCR-based methods to assess the infectivity of enteric viruses in environmental samples. *Applied and environmental microbiology*, *75*(2), 297–307. https://doi.org/10.1128/AEM.01150-08

Sakamoto, S., Putalun, W., Vimolmangkang, S., & Phoolcharoen, W. (2018). Enzyme - linked immunosorbent assay for the quantitative / qualitative analysis of plant secondary metabolites. *Journal of Natural Medicines*, *72*(1), 32–42. https://doi.org/10.1007/s11418-017-1144-z

Scheets, K. (2000). Maize chlorotic mottle machlomovirus expresses its coat protein from a 1.47-kb subgenomic RNA and makes a 0.34-kb subgenomic RNA. *Virology*, *267*(1), 90–101. https://doi.org/10.1006/viro.1999.0107

Scheets, K. (2016). Analysis of gene functions in Maize chlorotic mottle virus. *Virus Research*, *222*, 71–79. https://doi.org/10.1016/j.virusres.2016.04.024

Takahashi, M.K., Tan, X., Dy, A.J. *et al.* (2018). A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. *Nat Commun* **9,** 3347 doi:10.1038/s41467-018-05864-4

Tennant, P., Gubba, A., Roye, M., & Fermin, G. (2018). Chapter 6 - Viruses as Pathogens: Plant Viruses. In P. Tennant, G. Fermin, & J. E. Foster (Eds.), *Viruses* (pp. 135–156). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-811257-1.00006-1

Urcuqui-Inchima, S., Haenni, A.-L., & Bernardi, F. (2001). Potyvirus proteins: a wealth of functions. *Virus Research*, *74*(1-2), 157–175. doi: 10.1016/s0168-1702(01)00220-9

US Department of Agriculture, & USDA Foreign Agricultural Service. (2018). World corn production by country 2017/18 (in 1,000 metric tons)*. In Statista - The Statistics Portal. Retrieved October 1, 2018, from https://www-statista-com.udel.idm.oclc.org/statistics/254292/global-corn-production-by-country/.

UYEMOTO, J. K., CLAFLIN, L. E., & WILSON, D. L. (1981). Maize Chlorotic Mottle and Maize Dwarf Mosaic Viruses: Effect of Single and Double Inoculations on Symptomatology and Yield. *Plant Disease*, *65*(1), 39–41. https://doi.org/10.1094/PD-65-39

Valli, A. (2013). Chapter 689 - Potyvirus P1 Proteinase. In N. D. Rawlings & G. Salvesen (Eds.), *Handbook of Proteolytic Enzymes (Third Edition)* (Third Edit, pp. 3130–3133). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-382219-2.00689-X

Wamaitha, M.J., Nigam, D., Maina, S. *et al.* Metagenomic analysis of viruses associated with maize lethal necrosis in Kenya. *Virol J* **15,** 90 (2018) doi:10.1186/s12985-018-0999-2

Wang, Q., Zhang, C., Wang, C., Qian, Y., Li, Z., Hong, J., & Zhou, X. (2017). Further characterization of Maize chlorotic mottle virus and its synergistic interaction with Sugarcane mosaic virus in maize. *Scientific Reports*, *7*(1), 39960.https://doi.org/10.1038/srep39960

Wangai, A., Redinbaugh, M. G., Kinyua, Z. M., Miano, D. W., Leley, P. K., Kasina, M., … Jeffers, D. (2012). Rapid #: -12032214.

Way, J. C., Collins, J. J., Keasling, J. D., & Silver, P. A. (2014). Review Integrating Biological Redesign : Where Synthetic Biology Came From and Where It Needs to Go. *Cell*, *157*(1), 151–161. https://doi.org/10.1016/j.cell.2014.02.039

Weston, A., Humphreys, G. O., Brown, M. G. M., & Saunders, J. R. (1979). Simultaneous transformation of Escherichia coli by pairs of compatible and incompatible plasmid DNA molecules. *MGG Molecular & General Genetics*, *172*(1), 113–118. doi: 10.1007/bf00276222

Wu, L., Zu, X., Wang, S., & Chen, Y. (2012). Sugarcane mosaic virus - Long history but still a threat to industry. *Crop Protection*, *42*, 74–78. https://doi.org/10.1016/j.cropro.2012.07.005

Wu, J. X., Wang, Q., Liu, H., Qian, Y. J., Xie, Y., & Zhou, X. P. (2013). Monoclonal antibody-based serological methods for maize chlorotic mottle virus detection in China. *Journal of Zhejiang University. Science. B*, *14*(7), 555–562. doi:10.1631/jzus.B1200275

Xia, Z., Zhao, Z., Chen, L., Li, M., Zhou, T., Deng, C., … Fan, Z. (2016). Synergistic infection of two viruses MCMV and SCMV increases the accumulations of both MCMV and MCMV-derived siRNAs in maize. *Scientific Reports*, *6*. https://doi.org/10.1038/srep20520

Xie, L., Zhang, J., Wang, Q., Meng, C., Hong, J., & Zhou, X. (2011). Characterization of Maize Chlorotic Mottle Virus Associated with Maize Lethal Necrosis Disease in China, *193*, 191–193. https://doi.org/10.1111/j.1439-0434.2010.01745.x

Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., … Pierce, N. A. (2011). NUPACK: Analysis and Design of Nucleic Acid Systems. *Journal of Computational Chemistry*, *32*(1), 170–173. https://doi.org/10.1002/jcc.21596

Zhanmin Liu, Luxi Zhang, Cuiyun Yang & Xueying Xia | Fatih Yildiz (Reviewing Editor) (2016) Development of real-time reverse transcription PCR for detection of *Maize chlorotic mottle virus* based on a novel molecular marker, Cogent Food & Agriculture, 2:1, DOI: 10.1080/23311932.2016.1224047

Zhao, M., Ho, H., Wu, Y., He, Y., & Li, M. (2014). Western Flower Thrips (Frankliniella occidentalis) Transmits Maize Chlorotic Mottle Virus. *Journal of Phytopathology*, *162*(7–8), 532–536. https://doi.org/10.1016/j.wneu.2015.12.089

Appendix A

MCMV AND SCMV GENBANK ACCESSION NUMBERS

There were 43 complete genomes of MCMV used in this study. Their GenBank accession numbers are; EU358605.1, GU138674.1, JQ982468.1, JQ982469.1, JQ982470.1, KF010583.1, KJ782300.1, KP851970.3, MF510219.1, MF510220.1, MF510221.1, MF510222.1, MF510223.1, MF510224.1, MF510225.1, MF510226.1, MF510227.1, MF510228.1, MF510229.1, MF510230.1, MF510231.1, MF510232.1, MF510233.1, MF510234.1, MF510235.1, MF510236.1, MF510237.1, MF510238.1, MF510239.1, MF510240.1, MF510241.1, MF510242.1, MF510243.1, MF510244.1, MF510245.1, MF510246.1, MF510247.1, MF510248.1, MF510249.1, MF510250.1, MF510251.1, NC_003627.1, X14736.2

Accession numbers of 73 SCMV complete genomes used in this study. JX237863.1, JX237862.1, JX047381.1, JX047382.1, JX047383.1, JX047384.1, JN021933.1, JX047385.1, JX047386.1, JX047387.1, JX047388.1, JX047389.1, JX047390.1, JX047391.1, AY042184.1, JX047392.1, JX047393.1, AJ297628.1, JX047414.1, JX047415.1, JX047416.1, KP860935.1, KP860936.1, JX047417.7, JX047418.1, JX047419.1, KR108212.1, KR108213.1, JX047420.1, JX047421.1, JX047422.1, JX047423.1, JX047424.1, JX047425.1, JX047394.1, JX047395.1, KY006657.1, NC_003398.1, KT895080.1, JX047396.1, JX047397.1, JX047398.1, JX047399.1, JX047400.1, JX047401.1, JX047402.1, JX047403.1, JX047404.1, JX047405.1, JX188385.1, KP772216.1, AY149118.1, AY569692.1, EU091075.1, JX185303.1, JX047406.1, JX047407.1, AF494510.1, JX047408.1, JX047409.1, JX047410.1, JX047411.1, JX047412.1, JX047413.1, JX047426.1, JX047427.1, JX047428.1, KT895081.1, JX047429.1, JX047430.1, JX047431.1.

Accession numbers of 9 SCMV genomes used in this study.

KM926615.1, KF744392.1, KM926614.1, KP860935.1, KP860936.1, KM926613.1, KM926616.1, KF744390.1, KF744391.1

Appendix B

SCMV TOEHOLD SEQUENCES

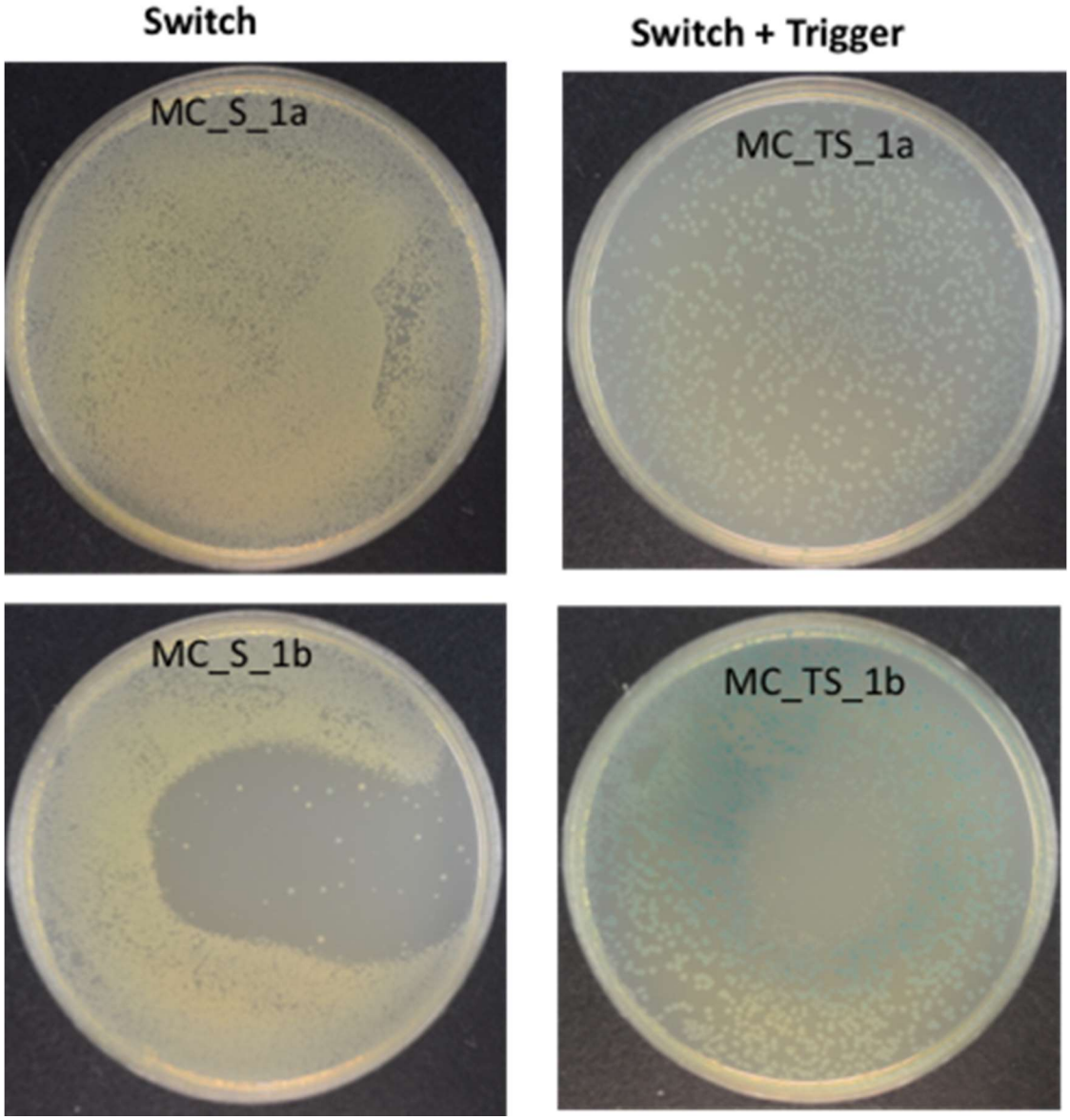
1.	TGAAAATTTTCCACAGTATGATAATTCTGCCCATGAGGACTTTAGAACAGAGG AGATAAAGATGTCATGGGCAGA
2.	CTGAAAATTTTCCACAGTATGATAATTCTGCCCATGGGACTTTAGAACAGAGG AGATAAAGATGCATGGGCAGAA
3.	GCTGAAAATTTTCCACAGTATGATAATTCTGCCCATGGACTTTAGAACAGAGG AGATAAAGATGATGGGCAGAAT
4.	GCCATATTGCTGAAAATTTTCCACAGTATGATAATTGGACTTTAGAACAGAGG AGATAAAGATGAATTATCATAC
5.	CGCCATATTGCTGAAAATTTTCCACAGTATGATAATGGACTTTAGAACAGAGG AGATAAAGATGATTATCATACT
6.	ACGCCATATTGCTGAAAATTTTCCACAGTATGATAAGGACTTTAGAACAGAG GAGATAAAGATGTTATCATACTG
7.	CACGCCATATTGCTGAAAATTTTCCACAGTATGATAGGACTTTAGAACAGAG GAGATAAAGATGTATCATACTGT
8.	ACACGCCATATTGCTGAAAATTTTCCACAGTATGATGGACTTTAGAACAGAG GAGATAAAGATGATCATACTGTG
9.	CACACGCCATATTGCTGAAAATTTTCCACAGTATGAGGACTTTAGAACAGAG GAGATAAAGATGTCATACTGTGG
10.	ACACACGCCATATTGCTGAAAATTTTCCACAGTATGGGACTTTAGAACAGAG GAGATAAAGATGCATACTGTGGA
11.	AACACACGCCATATTGCTGAAAATTTTCCACAGTATGGACTTTAGAACAGAG GAGATAAAGATGATACTGTGGAA
12.	GAACACACGCCATATTGCTGAAAATTTTCCACAGTAGGACTTTAGAACAGAG GAGATAAAGATGTACTGTGGAAA
13.	TGAACACACGCCATATTGCTGAAAATTTTCCACAGTGGACTTTAGAACAGAG GAGATAAAGATGACTGTGGAAAA
14.	ACTTTGCACTGCTCTTCAACACTAGCATCCAAATCTGGACTTTAGAACAGAGG AGATAAAGATGAGATTTGGATG
15.	AACTTTGCACTGCTCTTCAACACTAGCATCCAAATCGGACTTTAGAACAGAGG AGATAAAGATGGATTTGGATGC
16.	TAACTTTGCACTGCTCTTCAACACTAGCATCCAAATGGACTTTAGAACAGAGG AGATAAAGATGATTTGGATGCT
17.	GTAACTTTGCACTGCTCTTCAACACTAGCATCCAAAGGACTTTAGAACAGAGG AGATAAAGATGTTTGGATGCTA
18.	AATATAAATTGGTTTTGCATACTTTGATATGTCTTTGGACTTTAGAACAGAGG AGATAAAGATGAAAGACATATC

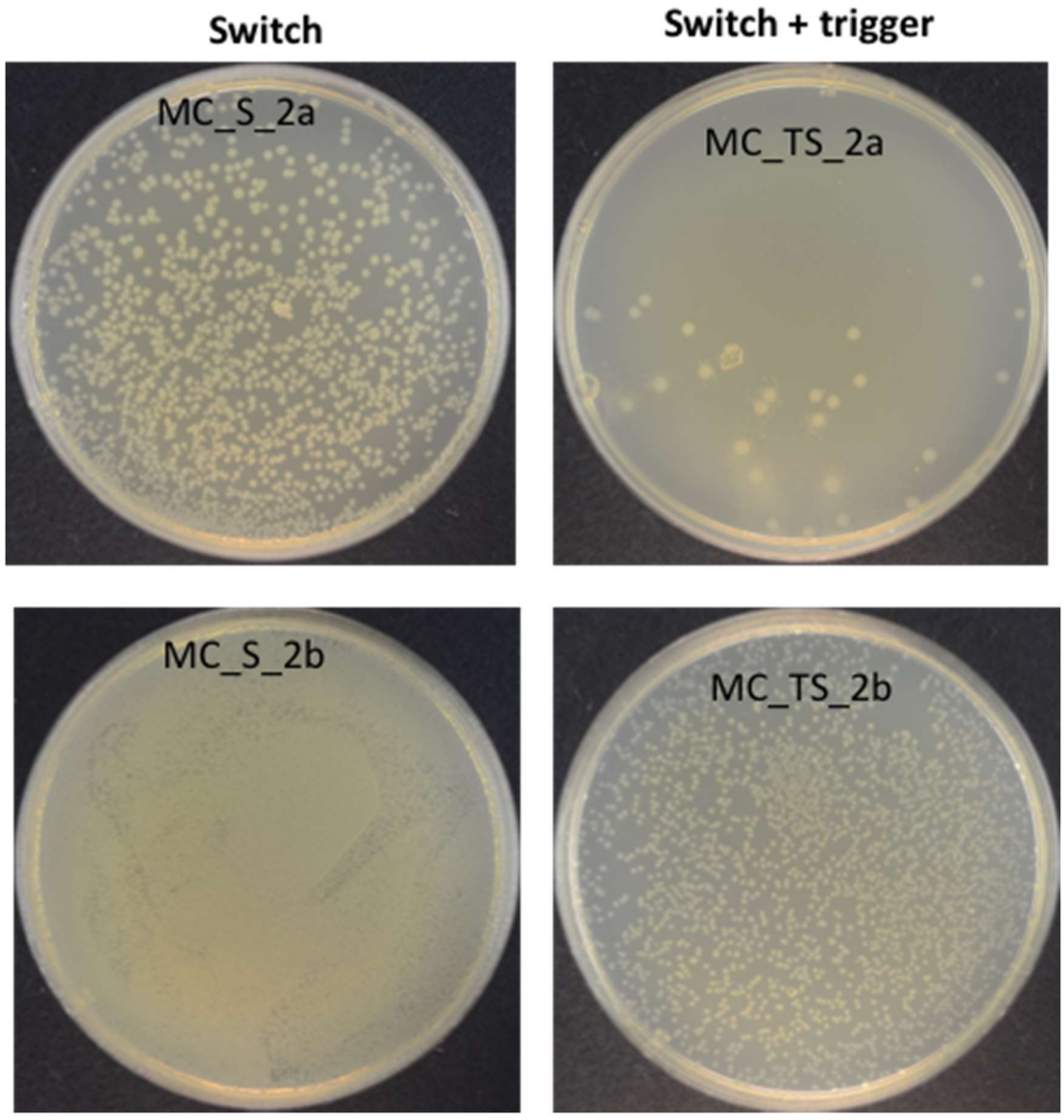
19.	GTAATCTTTCTTCTTTCCTGTGTACAGTGCTCCGACGGACTTTAGAACAGAGG AGATAAAGATGGTCGGAGCACT
20.	AGTAATCTTTCTTCTTTCCTGTGTACAGTGCTCCGAGGACTTTAGAACAGAGG AGATAAAGATGTCGGAGCACTG
21.	CCAAGTTGTCCATTGTAAATGCGTTCACATGATCTCGGACTTTAGAACAGAGG AGATAAAGATGGAGATCATGTG
22.	TTTCCTCCAAGCAAAGTTTCCAATGGAGCTGCTGTAGGACTTTAGAACAGAG GAGATAAAGATGTACAGCAGCTC
23.	TTTTCCTCCAAGCAAAGTTTCCAATGGAGCTGCTGTGGACTTTAGAACAGAG GAGATAAAGATGACAGCAGCTCC
24.	CTTTTCCTCCAAGCAAAGTTTCCAATGGAGCTGCTGGGACTTTAGAACAGAG GAGATAAAGATGCAGCAGCTCCA
25.	CACACACACTTTTCCTCCAAGCAAAGTTTCCAATGGGGACTTTAGAACAGAGG AGATAAAGATGCCATTGGAAAC
26.	CCACACACACTTTTCCTCCAAGCAAAGTTTCCAATGGGACTTTAGAACAGAGG AGATAAAGATGCATTGGAAACT
27.	TCCACACACACTTTTCCTCCAAGCAAAGTTTCCAATGGACTTTAGAACAGAGG AGATAAAGATGATTGGAAACTT
28.	GTCCACACACACTTTTCCTCCAAGCAAAGTTTCCAAGGACTTTAGAACAGAGG AGATAAAGATGTTGGAAACTTT
29.	TCGTCCACACACACTTTTCCTCCAAGCAAAGTTTCCGGACTTTAGAACAGAGG AGATAAAGATGGGAAACTTTGC
30.	ATCGTCCACACACACTTTTCCTCCAAGCAAAGTTTCGGACTTTAGAACAGAGG AGATAAAGATGGAAACTTTGCT
31.	AATCGTCCACACACACTTTTCCTCCAAGCAAAGTTTGGACTTTAGAACAGAGG AGATAAAGATGAAACTTTGCTT
32.	AAATCGTCCACACACACTTTTCCTCCAAGCAAAGTTGGACTTTAGAACAGAGG AGATAAAGATGAACTTTGCTTG
33.	CTCCATAGAATTTTGTTATCCCAACAGTCCATGGGCGGACTTTAGAACAGAGG AGATAAAGATGGCCCATGGACT
34.	CCTCCATAGAATTTTGTTATCCCAACAGTCCATGGGGGACTTTAGAACAGAGG AGATAAAGATGCCCATGGACTG
35.	ACCTCCATAGAATTTTGTTATCCCAACAGTCCATGGGGACTTTAGAACAGAGG AGATAAAGATGCCATGGACTGT
36.	AACCTCCATAGAATTTTGTTATCCCAACAGTCCATGGGACTTTAGAACAGAGG AGATAAAGATGCATGGACTGTT
37.	CAACCTCCATAGAATTTTGTTATCCCAACAGTCCATGGACTTTAGAACAGAGG AGATAAAGATGATGGACTGTTG
38.	CCAACCTCCATAGAATTTTGTTATCCCAACAGTCCAGGACTTTAGAACAGAGG AGATAAAGATGTGGACTGTTGG

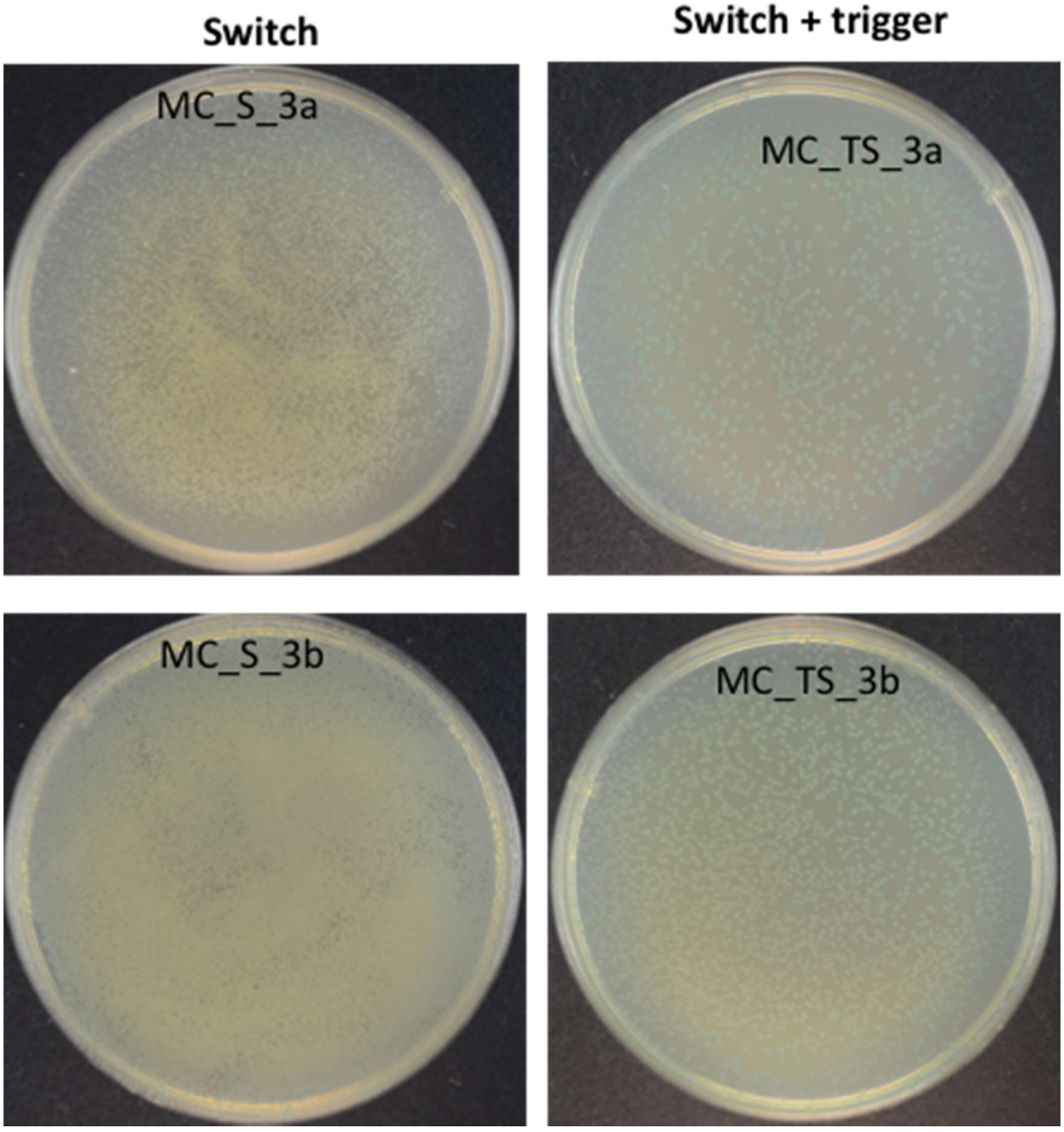
39.	TCCAACCTCCATAGAATTTTGTTATCCCAACAGTCCGGACTTTAGAACAGAGG AGATAAAGATGGGACTGTTGGG
40.	TTCCAACCTCCATAGAATTTTGTTATCCCAACAGTCGGACTTTAGAACAGAGG AGATAAAGATGGACTGTTGGGA
41.	ATTCCAACCTCCATAGAATTTTGTTATCCCAACAGTGGACTTTAGAACAGAGG AGATAAAGATGACTGTTGGGAT
42.	GATTCCAACCTCCATAGAATTTTGTTATCCCAACAGGGACTTTAGAACAGAGG AGATAAAGATGCTGTTGGGATA
43.	TCAAGTAAGCGATTCCAACCTCCATAGAATTTTGTTGGACTTTAGAACAGAGG AGATAAAGATGAACAAAATTCT
44.	CTCAAGTAAGCGATTCCAACCTCCATAGAATTTTGTGGACTTTAGAACAGAGG AGATAAAGATGACAAAATTCTA
45.	TCTCAAGTAAGCGATTCCAACCTCCATAGAATTTTGGGACTTTAGAACAGAGG AGATAAAGATGCAAAATTCTAT
46.	TTCTCAAGTAAGCGATTCCAACCTCCATAGAATTTTGGACTTTAGAACAGAGG AGATAAAGATGAAAATTCTATG
47.	CTTCTCAAGTAAGCGATTCCAACCTCCATAGAATTTGGACTTTAGAACAGAGG AGATAAAGATGAAATTCTATGG
48.	GACATCTTTTCCTTTTGCTTTCGGTAAGCGCATTTTGGACTTTAGAACAGAGG AGATAAAGATGAAAATGCGCTT
49.	AGACATCTTTTCCTTTTGCTTTCGGTAAGCGCATTTGGACTTTAGAACAGAGG AGATAAAGATGAAATGCGCTTA
50.	AAGACATCTTTTCCTTTTGCTTTCGGTAAGCGCATTGGACTTTAGAACAGAGG AGATAAAGATGAATGCGCTTAC
51.	CAAGACATCTTTTCCTTTTGCTTTCGGTAAGCGCATGGACTTTAGAACAGAGG AGATAAAGATGATGCGCTTACC

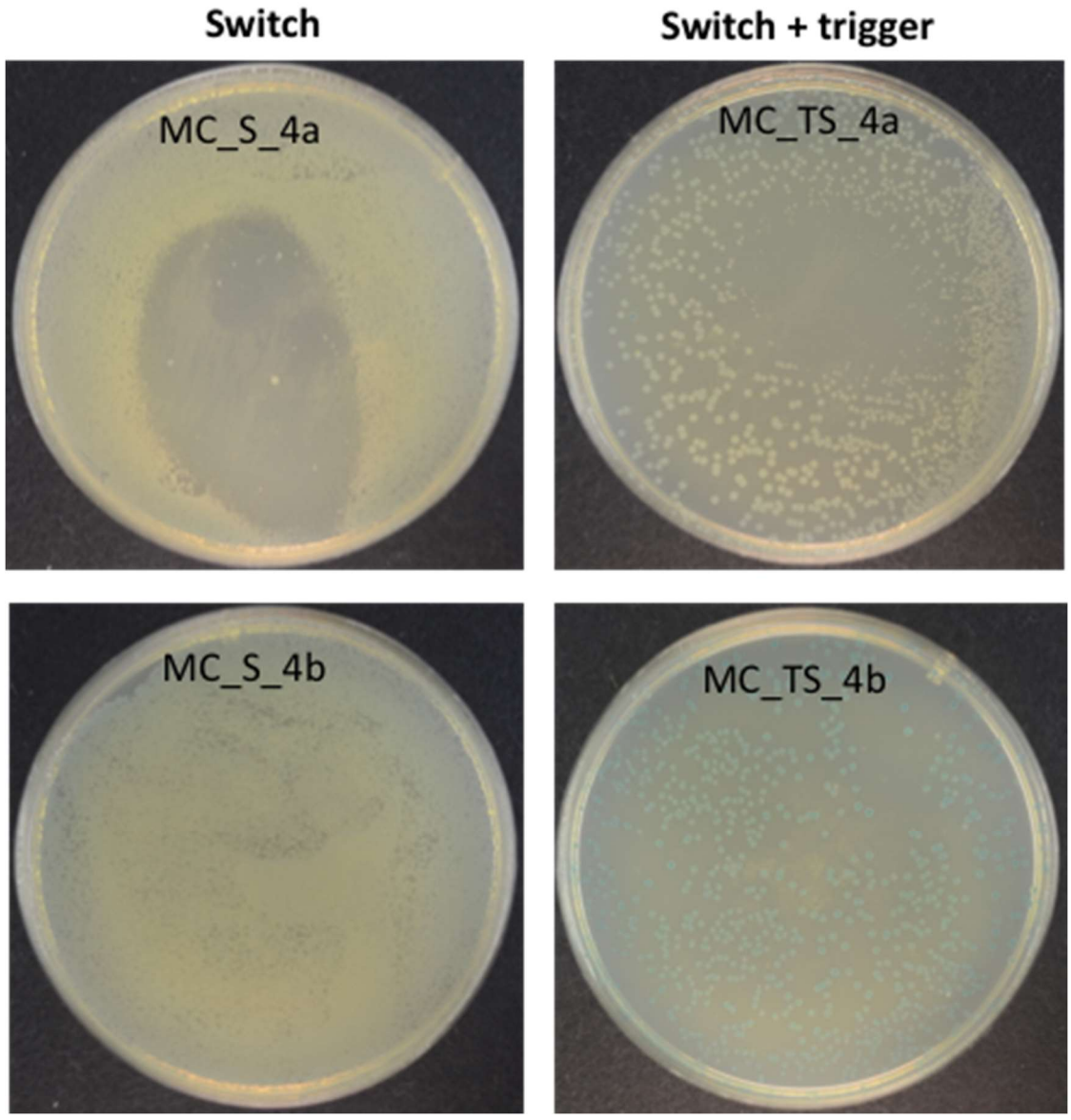
Appendix C

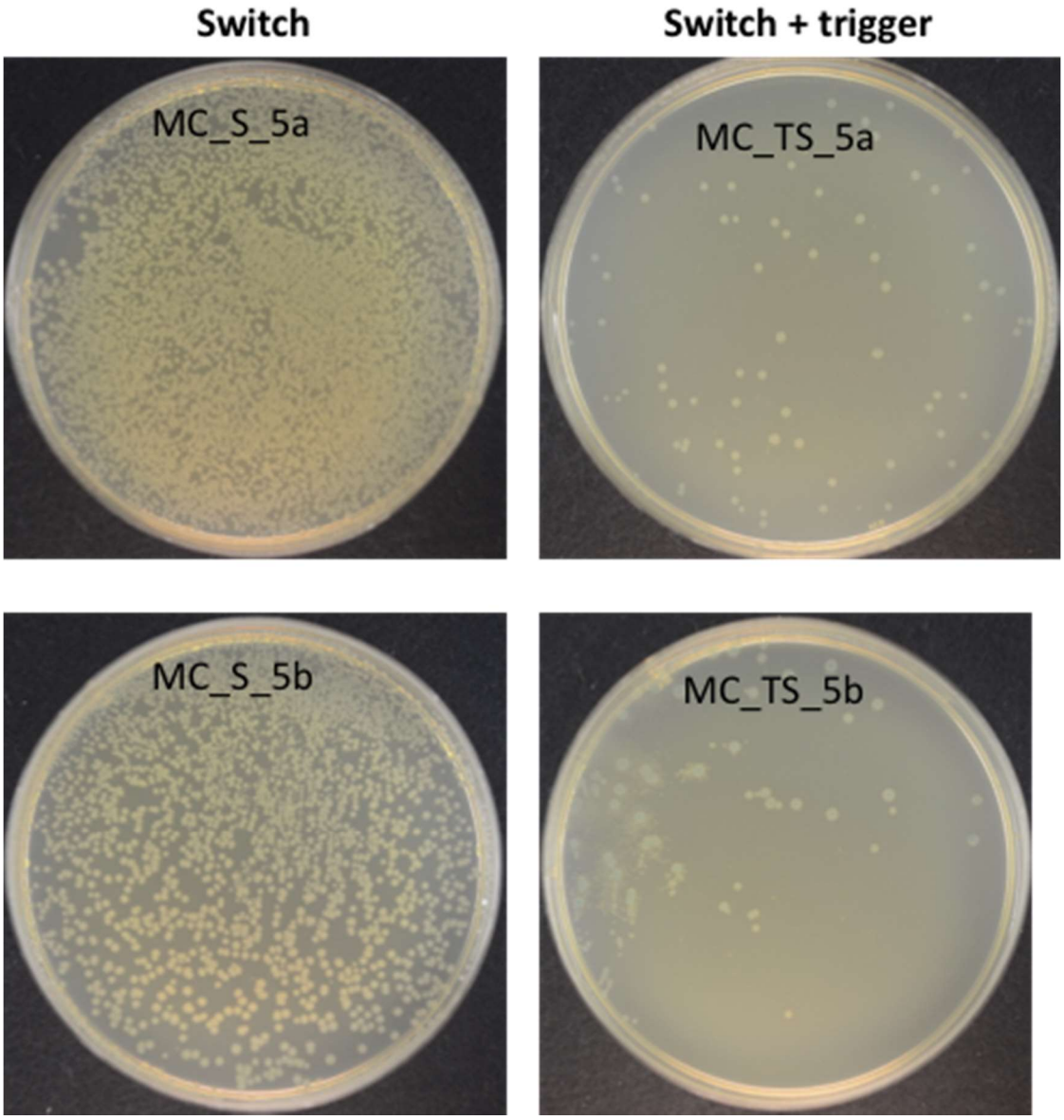
BLUE WHITE SCREEN ENLARGED PLATES

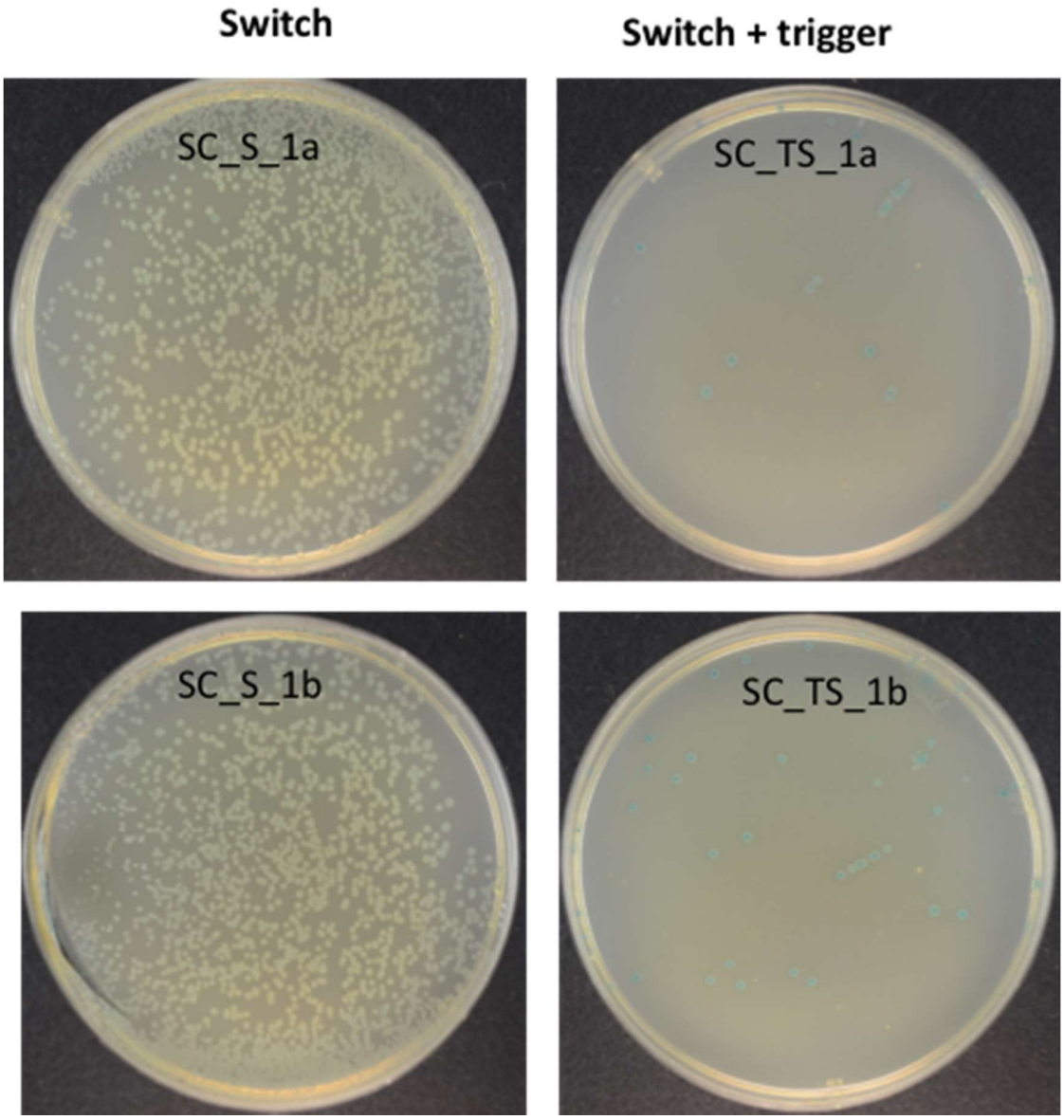


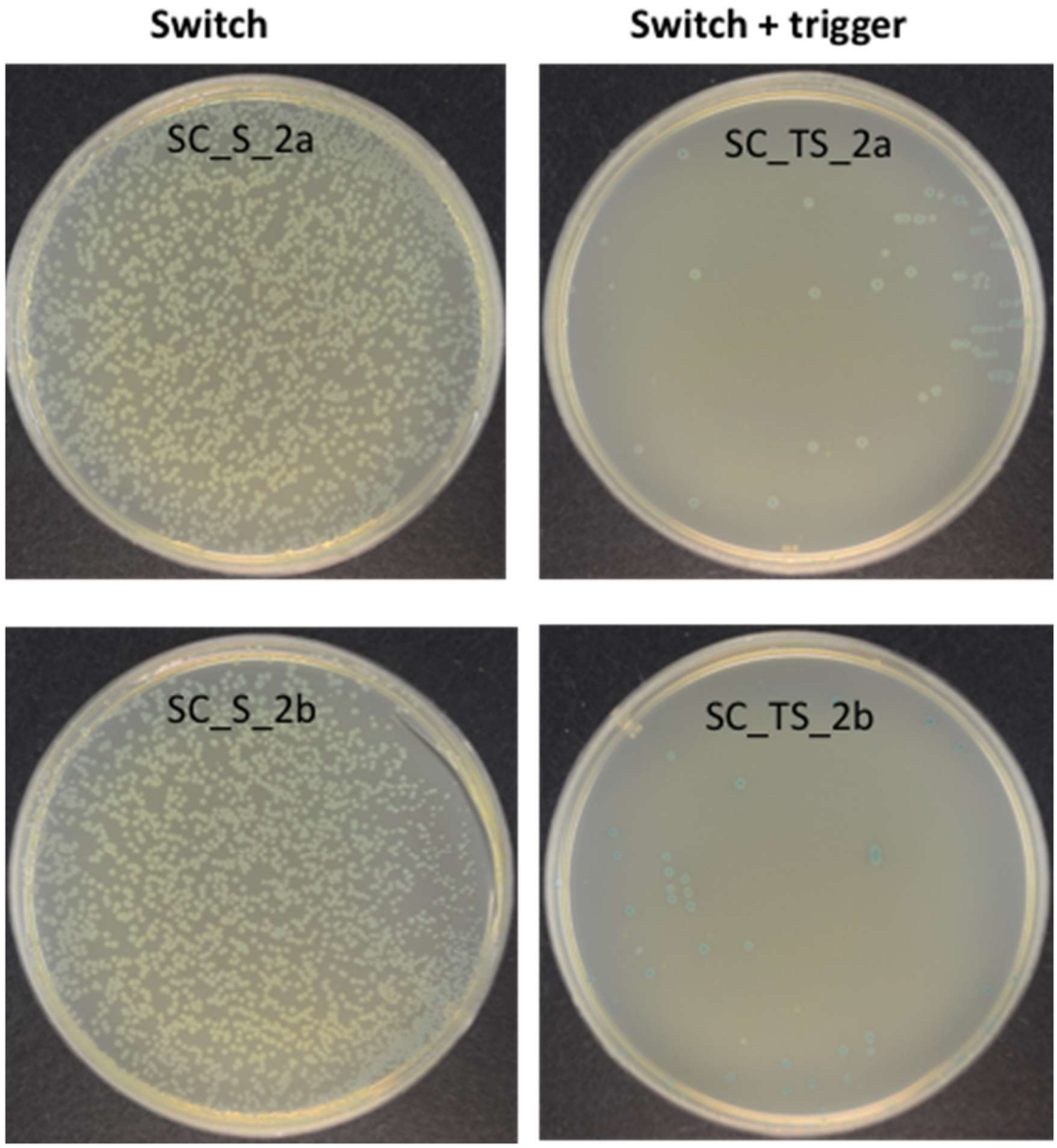

Switch
Switch + trigger
MC_S_2a
MC_TS_2a
MC_S_2b
MC_TS_2b

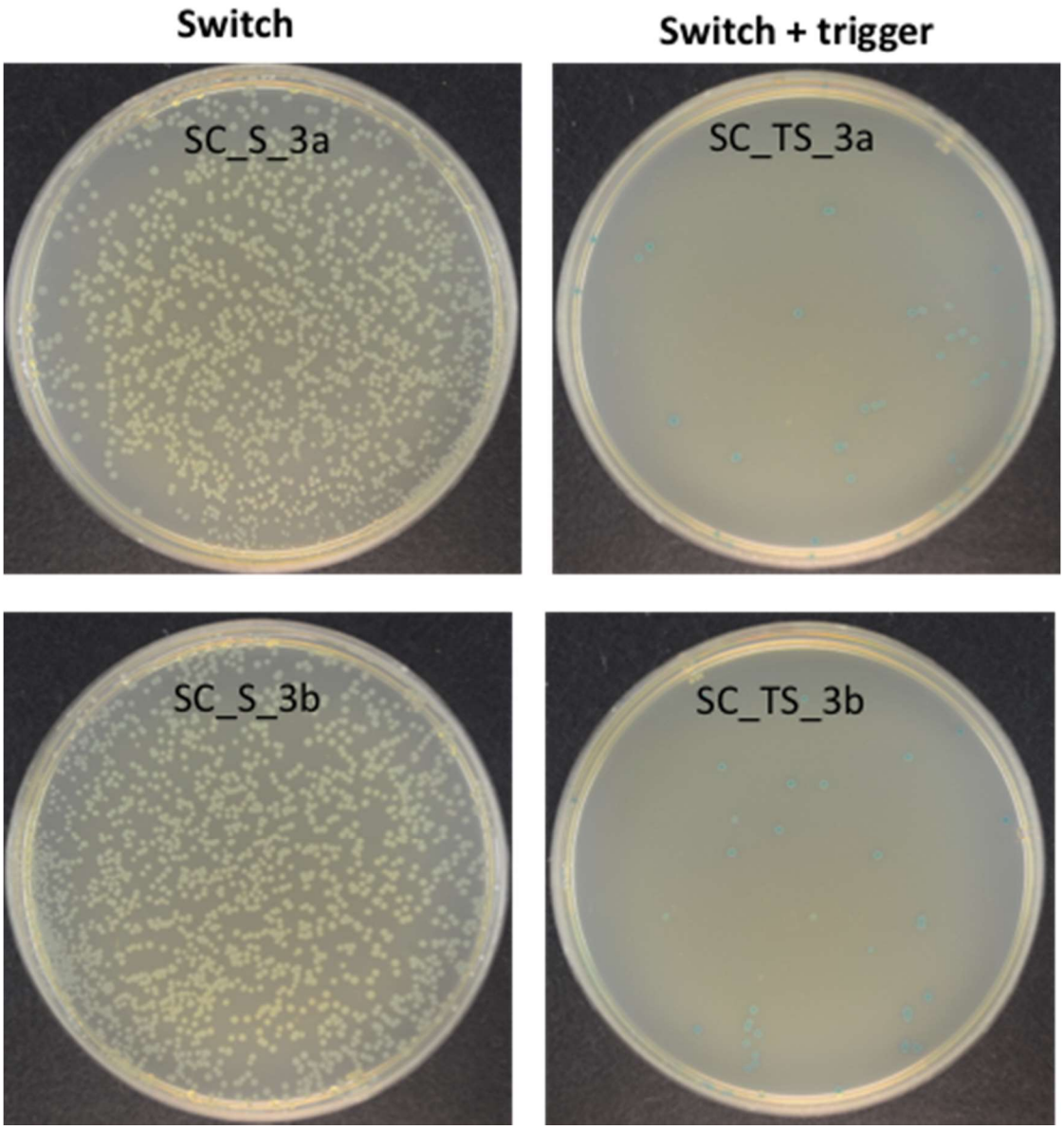

Switch
Switch + trigger
MC_S_3a
MC_TS_3a
MC_S_3b
MC_TS_3b

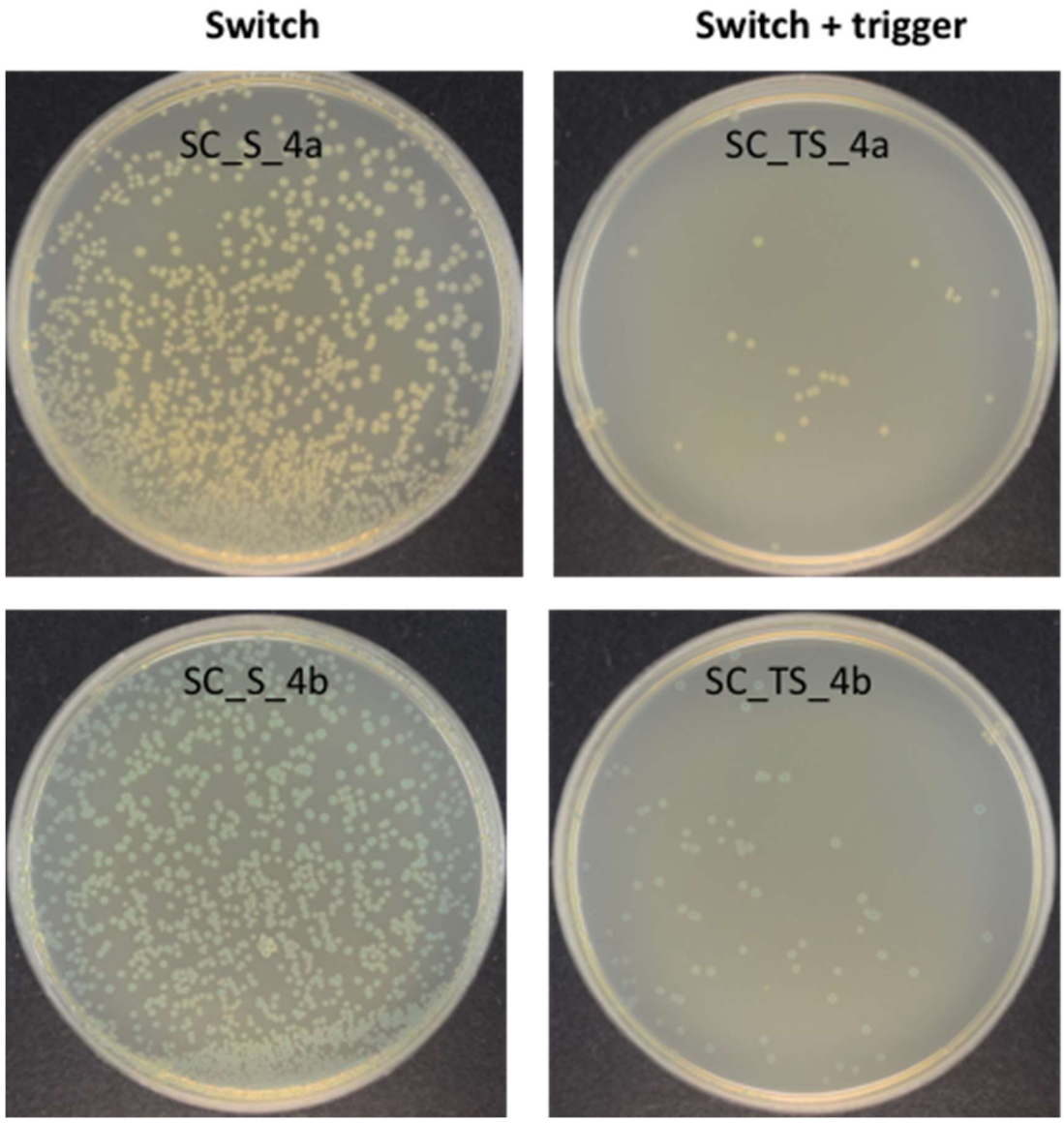

Switch
Switch + trigger
MC_S_4a
MC_TS_4a
MC_S_4b
MC_TS_4b

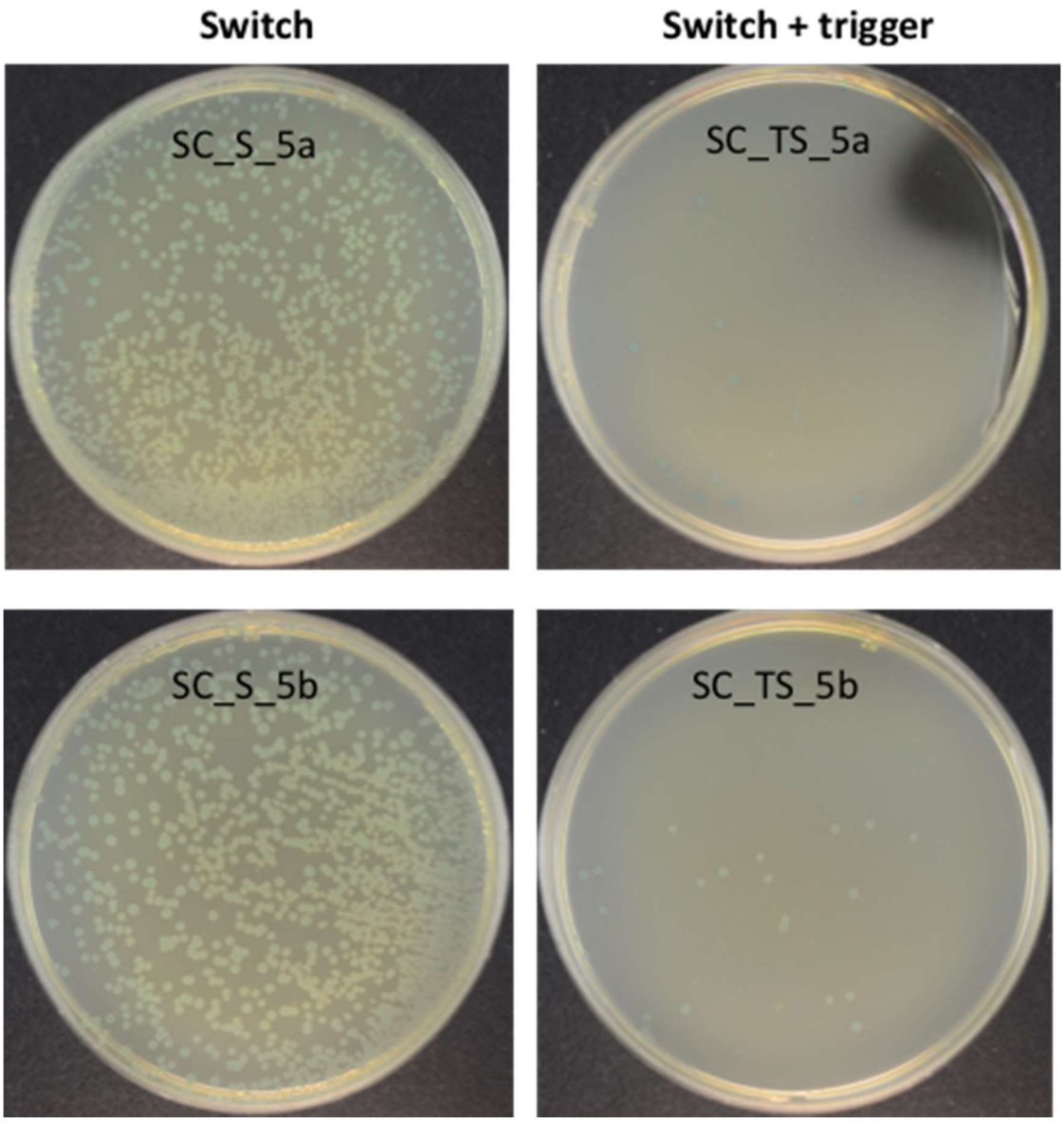

Switch
Switch + trigger
MC_S_5a
MC_TS_5a
MC_S_5b
MC_TS_5b

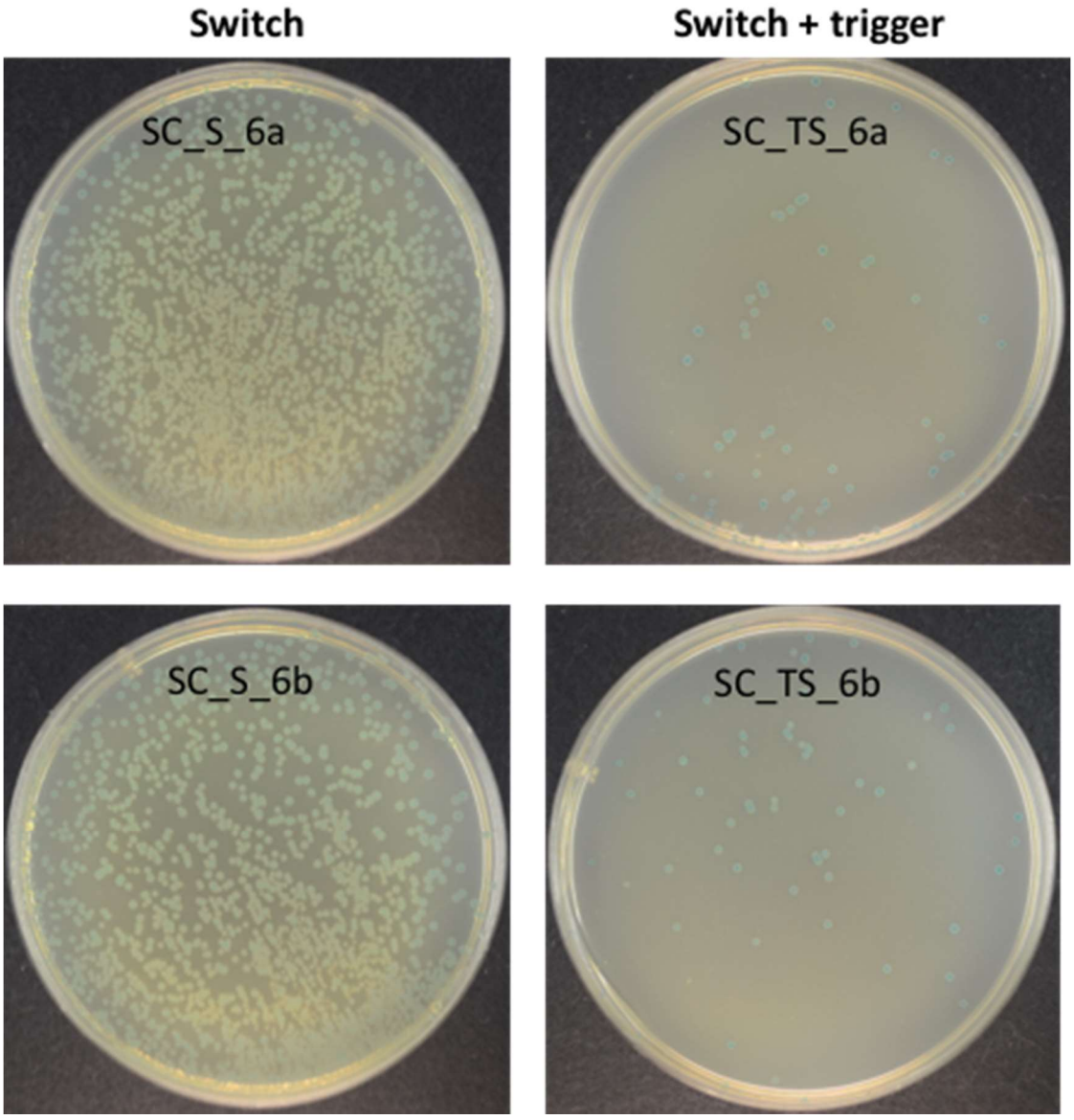

Switch
Switch + trigger
SC_S_1a
SC_TS_1a
SC_S_1b
SC_TS_1b

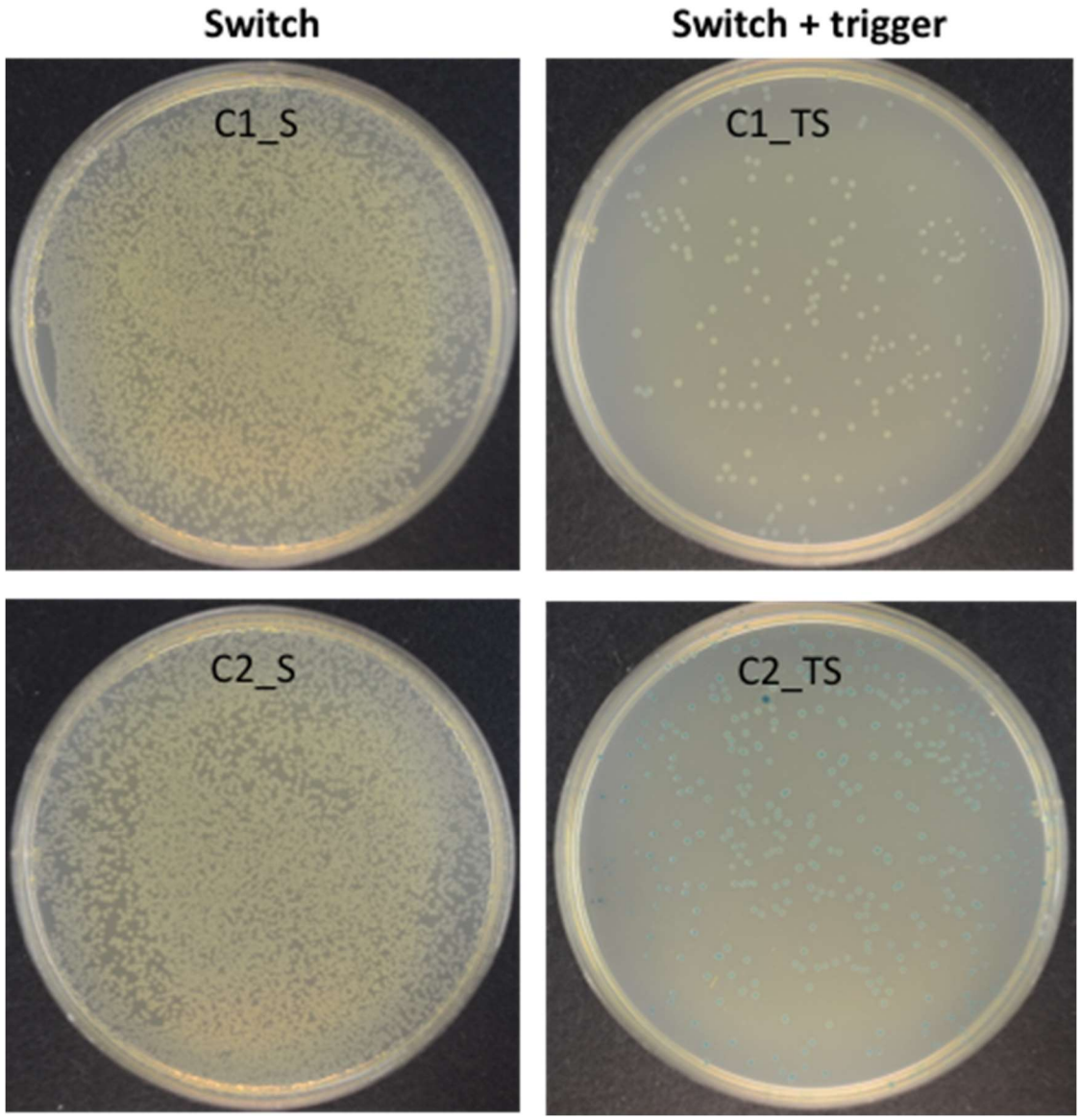

Switch
Switch + trigger
SC_S_2a
SC_TS_2a
SC_S_2b
SC_TS_2b


Switch
Switch + trigger
SC_S_3a
SC_TS_3a
SC_S_3b
SC_TS_3b


Switch
Switch + trigger
SC_S_4a
SC_TS_4a
SC_S_4b
SC_TS_4b


Switch
Switch + trigger
SC_S_5a
SC_TS_5a
SC_S_5b
SC_TS_5b


Switch
Switch + trigger
SC_S_6a
SC_TS_6a
SC_S_6b
SC_TS_6b


Switch
Switch + trigger
C1_S
C1_TS
C2_S
C2_TS