Development of biological recognition elements for the detection of the genome of cocoa swollen shoot virus

BERTRAND M. MONNIER

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy (PhD).

Faculty of Applied Science, University of the West of England, Bristol

June 2024

Word count: 42,049

Abstract.

Cacao swollen shoot virus (CSSV) is the main threat to chocolate production in West Africa. Infection causes a decrease in bean production yield and can cause tree death within three years, causing socio-economic damage to small-holding farmers.

CSSV control measures relying on cutting out infected and surrounding trees have proven ineffective. Novel strategies are needed to stem the pathogen propagation in the endemic area by identifying and removing asymptomatic CSSV infected trees. This work assessed if a cost-effective easy to operate biosensor could be developed to detect CSSV DNA. Two biological recognition elements (BRE) deriving from CSSV, serving as an interface between the user and the targeted analyte, were designed, produced, and tested with CSSV-infected *T. cacao* extracts. First, a BRE based on a chimeric zinc finger protein targeting CSSV genome was functionally tested using an electrophoretic mobility shift assay, a double filter binding assay and a superparamagnetic particles pull-down assay. The second BRE was constructed with an oligonucleotide lattice following *de novo* sequencing of a CSSV UWE strain. Using these nucleotides, a temperature-independent, "cold", hybridisation assay was developed and tested by high resolution melt and dot blotting. Simple DNA extraction methods were developed to isolate CSSV genome from tree and sap samples. Both DNA and protein based BREs were able to interact with the genome of CSSV with higher specificity observed with the oligonucleotide lattice. Successful simple CSSV DNA extraction was performed from leaves but also from vascular tissues. This combined with a cold hybridisation method could be integrated in a CSSV biosensing procedure.

A cost effective and user-friendly biosensor to detect CSSV in an area of low economic capital is feasible using the biological recognition elements, the DNA extractions and hybridisation assay developed in this project. The system could be adapted to monitor other analytes as indicators of tree health to help achieve sustainable agriculture targets.

Acknowledgements.

This thesis is dedicated to my late parents Mireille and Jacques.

Many people have contributed to this work, and I would like to thank all my fellow PhD students, especially Dr Arwa Flemban and Dr Niamh Brannelly and my colleagues that have accompanied me on this *long* journey. Thanks to all the staff at the University of the West of England, in the administration and the facilities.

First, I would like to thank my director of study, Associate Professor Joël Allainguillaume, Dr Jackie Barnett and Professor Luxton for your precious guidance and support throughout this project. Joël, I am very grateful for your patience, the great ideas you have given me and your unconditional "bonne humeur," which really helped to relieve the stress that comes with that type of endeavor. Jackie, I cannot thank you enough for your continual support since the very beginning of our work with Sue and Owen, what a great team this was! Your constant encouragement, your kindness and your advice have convinced me to pursue this chapter of my life. You were crucial to this achievement. Thank you again! Richard, thank you for giving me the opportunity to work at UWE and introducing me to the fascinating world of biosensing. Thank you also for all your precious advice while working on this project.

I would also like to thank Associate Professor Saliha Saad, for your kindness, your support, and the conversations we have had throughout my time at UWE.

This work would not have been possible without the help, the kindness, and the understanding of my family and all my friends, on both sides of the channel, much love to you all. Thank you to both my sisters, Béa and Bénédicte, for bearing with me while I was busy doing this work. A special thanks to Tara and Ken for the wonderful retreat. Thank you all for believing in me.

Pambu, thank you for being who you are, and for the support you have given me. You played an important part in this journey, and I hope I can return the courtesy soon. Our challenging conversations and all the laughter kept my brain working when I needed it.

Much love to Monty, who has been a close companion while I was writing this, your purrs and meows kept me relaxed and grounded.

Finally, thank you Houdie and Charlie. I would like to thank Houdie for supporting me, in both the English and the French meaning (and supplying me with countless vegan chocolate bars). Your kindness, attention and resilience are an inspiration to me. I would not have made it this far without you.

Table of Contents.

List of Tables.

List of Figures.

List of abbreviations.

AB	Annealing Buffer
BLAST	Basic local alignment search tool.
BRE	Biological recognition element
CaMMV	Cacao mild mosaic virus
CSSV	Cocoa swollen shoot virus
CSSTAV	Cocoa swollen shoot virus Togo A virus
CSSCEV	Cocoa swollen shoot CE virus
CSSGNV	Cocoa swollen shoot Ghana N virus
CSSGQV	Cocoa swollen shoot Ghana Q virus
CSSGRV	Cocoa swollen shoot Ghana R virus
CYVBV	Cocoa yellow vein-banding virus
ORF	Open reading frame
PCR	Polymerase chain reaction
HRM	High resolution melting
KDa	KiloDalton
Kb	Kilobase
LFT	Lateral flow Test
Mb	Megabase
T. cacao	*Theobroma cacao*
SELEX	Systematic evolution of exponential enrichment
XNA	Xeno-nucleic acid
ZFP	Zinc finger protein

Chapter 1 : Introduction.

1.1 History of *Theobroma cacao* use.

Chocolate is one of the world's favourite confectionary products. It is derived from the fruit of the cacao tree, *Theobroma cacao* (*T. cacao*). Now cultivated in all tropical areas of the world (America, Africa, and Asia), it has a neotropical origin (Central America, the Caribbean and South America).

Evidence from an archaeological study in a preserve of the ancient Mayo-Chinchipe culture points to the earliest use of cacao in Santa Ana-La Florida, in the southeast of Ecuador (Zarrillo *et al.*, 2018a), 5300 years ago (Figure 1.1). The presence of *Theobroma cacao* was evidenced through the identification of theobromine and starch by ultra-performance-liquid-chromatography tandem mass spectrometry and ancient DNA by quantitative polymerase chain reaction (qPCR) on ornamented vessels or more rudimentary, non-elaborate pottery such as stone bowls, mortars, and stirrup-spout bottles used by the Mayo-Chinchipe.

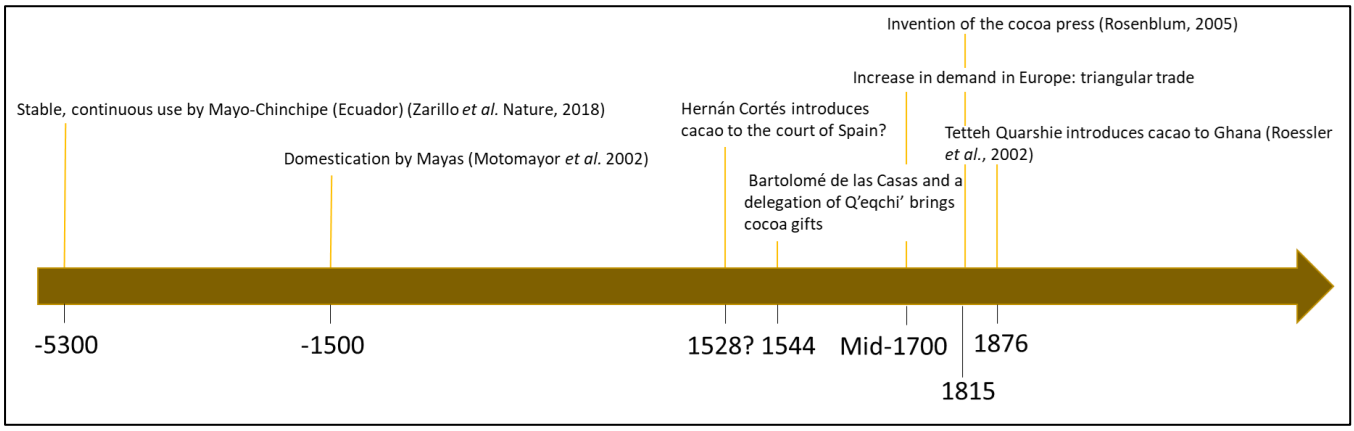


Figure 1.1: A simplified chronology of the use of cacao.

Additional archaeological evidence has indicated usage of cacao by Olmec circa 1800 to 1000 BC (Powis *et al.*, 2011) and Mayan people circa 600 to 900 AD (Ford, Williams, and de Vries, 2022). Cacao

was of particular value, being used as currency and was found associated with deities on a range of artifacts (Figure 1.2).

Figure 1.2: Depiction of a diving deity holding a cacao pod from the post classic Mayan era. (Kerr, 2003)

Over time "kakau" (Figure 1.3) was used as a flavouring in food, as a drink, and a medicine; cacao was also found to have stimulant properties and was also used in ceremonies (Zarrillo *et al.*, 2018b). One of many stimulant plants used in the New World, cacao made its way as a drink to the court of Spain in the middle of the sixteenth century (see Figure 1.1), and from there reached the aristocracy all over Europe within the next century (Squicciarini, 2016). The recipe favoured by South Americans was adapted to accommodate the taste of the European population. From the cold, water-based brew drunk by the Aztecs, Spanish people added cane sugar, vanilla, and cinnamon to mitigate the bitterness, while heating the preparation to consume it hot. Importantly, the conquistadores also produced tablets made from ground cacao to facilitate its transportation (Coe, 1996).

Figure 1.3: Mayan glyph depicting the word cacao. (Soparamens, 2017. Under CC0 1.0 Licence)

Enjoyed for its culinary and medicinal properties, the commercial demand for cacao increased and led to the cultivation of *T. cacao* by European colonial powers. France established plantations in the Caribbean, while Spain developed their own in the Philippines (Dillinger *et al.*, 2000). By the end of the nineteenth century, the plant reached mainland Africa, in Ghana, where a blacksmith and farmer named Tetteh Quarshie (Agyei-Ohemeng, Opare and Obeng-Ofori, 2018) brought a seed from a voyage in the then Spanish-ruled island of Fernando Po, in Equatorial Guinea and started cultivating the tree. Soon followed by other farmers, the colonial British administration saw the potential of the crop and developed the cultivation of *T. cacao* trees in Ghana. Similarly, cultivation flourished in the neighbouring Cote d`Ivoire, then under French rule. Today it is cultivated in several countries in West Africa. Expansion of the use of cacao in Europe was made possible through several technical advances, that led to the creation of chocolate as a transformed product. The industrialisation of chocolate coincides with the Industrial Revolution of the 19th century (Squicciarini, 2016) that introduced original production processes. For instance, Coenraad Van Houten in the Netherlands invented the cocoa press in 1815 (Rosenblum, 2005) , that extracts the fatty portion from the cocoa mass, resulting in the separation of cocoa butter from cocoa powder. Thereby, it was possible to create a new product, which was chocolate readily edible. In 1847, the grandson of the Bristol-based chocolate maker Joseph Fry developed a process to combine cocoa powder, sugar and melted cocoa butter which resulted in a paste that could be cast in moulds to produce chocolate bars (Goldstein, 2015).

1.2 *Theobroma cacao* description.

Theobroma cacao belongs to the order of the Malvales, the family of the Malvacae, the subfamily of the Byttneriodeae and the genus *Theobroma*. The Malvacae family contains 250 genera and 4200 species including herbs, shrubs, and trees. Other economically notable members of the family include the cotton plant (*Gossypium spp*.), the jute plant (*Corchorus spp.*), the cola tree (*Cola nitida*), okra (*Abelmoschus esculentus*), the durian tree (*Durio zibethinus*) and the baobab tree (*Adansonia digitate*) (Simpson, 2010). *Theobroma cacao* is a six to twelve meters evergreen tree, with twelve to sixty centimetres ovate-oblong leaves (Figure 1.4) and white/yellow or pink fifteen millimetres diameter cauliflorous flowers (Figure 1.5) (Claus *et al.*, 2018).

Figure 1.4: Photography of a *Theobroma cacao* tree in Cameroon (Hiobson, 2017. Under CC 4.0 Licence). A scale bar was added to provide an approximate size.

Figure 1.5: The pentaflower of *Theobroma cacao*, attached to a branch (Vinayaraj, 2018. Under CC 4.0 Licence). A scale bar was added to provide an approximate size.

Pollinated trees start producing their berries after three years, reaching a maximum yield after eight to nine years. The fruit is a seventeen- to twenty-centimetre-long drupe (Nigam, 1999), with a three-centimetre leathery or smooth rind depending on the variety (Thi *et al.*, 2016). Pods bear twenty to forty seeds called beans, enclosed within a mucilaginous pulp (Figure 1.6).

Figure 1.6: (Left) A cacao pod on a *Theobroma cacao* tree in Ecuador (Bernard Gagnon, 2014. Under CC 3.0 Licence). (Right) botanical drawing of the cacao berry ("pod"), showing the arrangement of the seeds ("beans") within the mucilaginous pulp.

Whole seeds and sections are presented on the right-hand side of the drawing (Köhler, 1887). A scale bar was added to provide an approximate size.

Theobroma cacao is diploid with ten chromosomes (2n=20) (Argout *et al.*, 2011). The genome is within the media size of the green plants (Kress *et al.*, 2022), with genotypes size ranging from 411Mb to 494Mb. Analysis of the Belizean Criollo B97-61/B2 genome revealed 21,437 protein-coding genes, 2229 noncoding genes and 1165 pseudogenes (Argout *et al.*, 2017). There are three main traditional cultivars or groups of cacaos based on geographic origin and morphology: Criollo, Forastero and a hybrid Criollo x Forastero named Trinitario (Cornejo *et al.*, 2018). A microsatellite analysis using 106 markers from samples of a germplasm collection in South America further divided *Theobroma cacao* into ten genetic clusters: Amelonado, Contanama, Criollo, Curaray, Guiana, Iquitos, Marañon, Nacional, Nanay, and Purús (Motamayor *et al.*, 2008) (Figure 1.7).

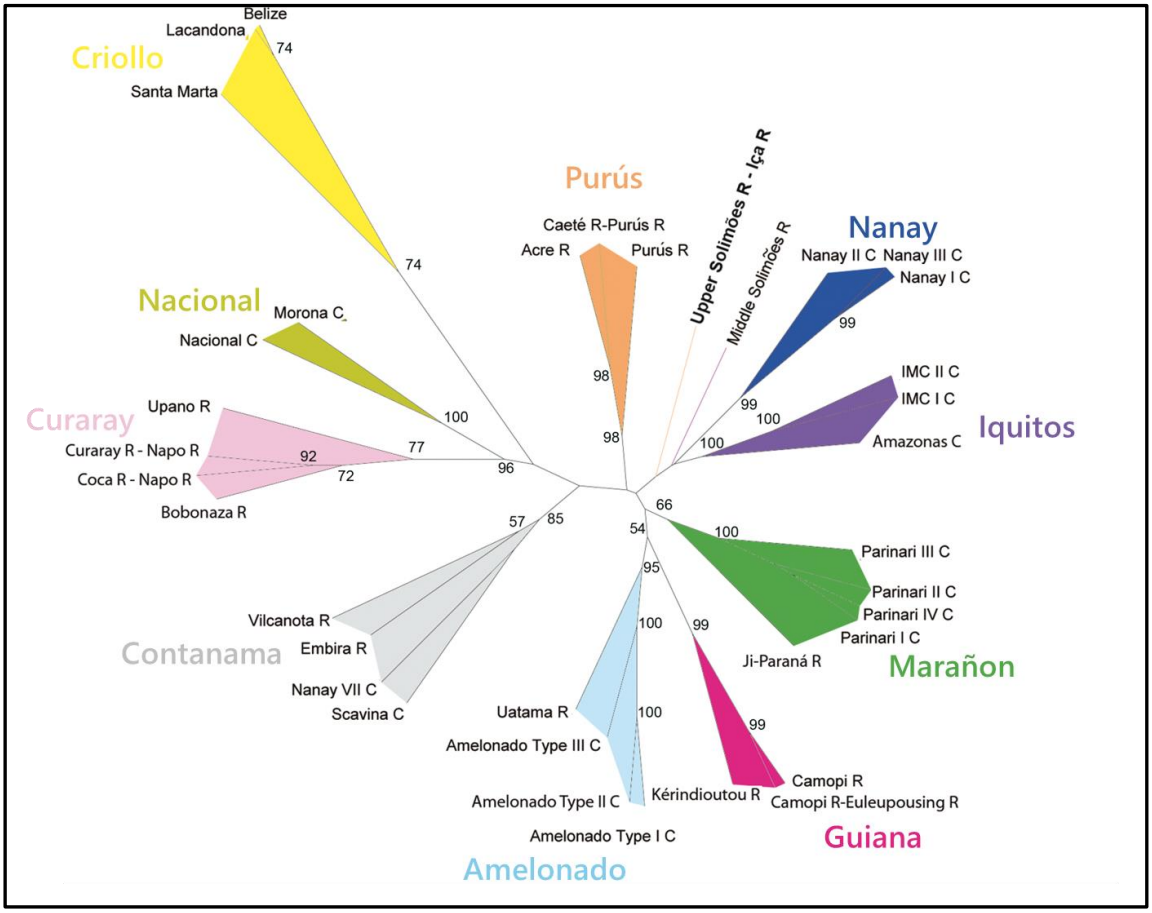


Figure 1.7: Neighbour joining tree inferred from Cavalli-Sforza and Edwards genetic distance presenting the ten main genetic clusters identified in a South American *T. cacao* germplasm.

Figure adapted from Motamayor, J.C., Lachenaud, P., Schnell, R.J., 2008. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L). Under CC BY 4.0 Licence.

The exact origin of this tree is still controversial and is an ongoing topic of discussion within the scientific community. The difference in opinion resides in whether cacao originated in Mexico (Figure 1.8), Central America or in the region delineated by the Amazon and the Orinoco basins (Figure 1.9). A third hypothesis suggests a separated, simultaneous origin. Cuatrecasas (Cuatrecasas, 1964) suggested that a separation brought by the Isthmus of Panama would have yielded two subspecies, based on the important difference in the plant's morphology observed between the two regions. Both locations would have therefore yielded their own subspecies, having evolved separately, *Theobroma cacao* subspecies *cacao*, the Criollo variety, and *Theobroma cacao* subspecies *sphaerocarpum*, the Forastero variety.

Figure 1.8: Map of Mexico.

The black circle delineates the regions of Tabasco and Chiapas that are thought to be the meso/central American Theobroma cacao place of origin (Miranda, 1962).

Figure 1.9: Map of the South American continent.

The pink line delineates the location of the Orinoco basin and the green line the Amazon basin, two areas believed to be the origin of the cacao tree (Hall, 1914).

Large scale genetic studies have identified the Amazon area to bear the highest *Theobroma cacao* genetic diversity. Thomas *et al*. (Thomas *et al.*, 2012) used 96 SSR markers to genotype 939 cacao trees from Central and South America and evidenced a higher species diversity in the upstream areas of the western Amazon, reinforcing the South American origin hypothesis.

The *T. cacao* tree introduced by Tetteh Quarshie in Ghana was from the Amelonado cultivar type, and the majority of the plants currently cultivated on the continent derived from this variety (Opoku *et al.*, 2007). Furthermore, the trees subsequently introduced to Africa by colonial powers came predominantly from Amelonado plants from Brazil (Motamayor *et al.*, 2003) which resulted in reduced genetic diversity on the continent (Aikpokpodion et al., 2009 ; Gopaulchan et al., 2019).

1.3 *Theobroma cacao* cultivation.

Today, 67% of the crop production is in West Africa (Figure 1.10) (FAO, 2023), mainly in Ivory Coast and Ghana with smaller production in Togo, Cameroon, and Nigeria (Figure 1.10). Cacao beans are now one of the main cash crops contributing significantly to the GDP of these countries, with an annual production of 5.58 million metric tons (FAO, 2021), involving 14 million workers worldwide (Beg *et al.*, 2017).

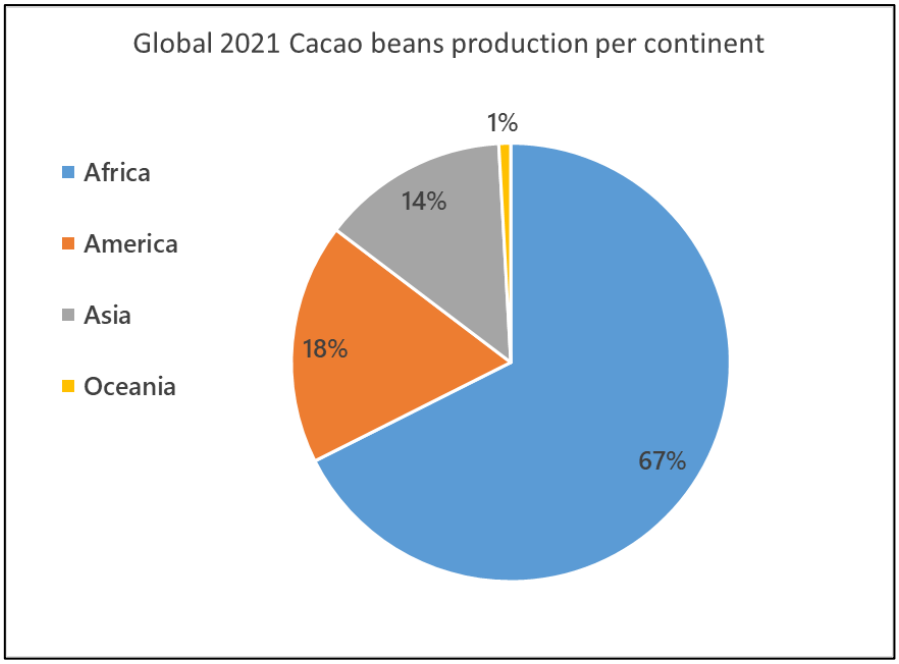


Figure 1.10: Global cacao beans production per continent in 2021.

The percentage of the production from each continent is shown in comparison to the global production (Figure created with data from the Food and Agricultural Organisation of the United Nations (FAO, 2023)).

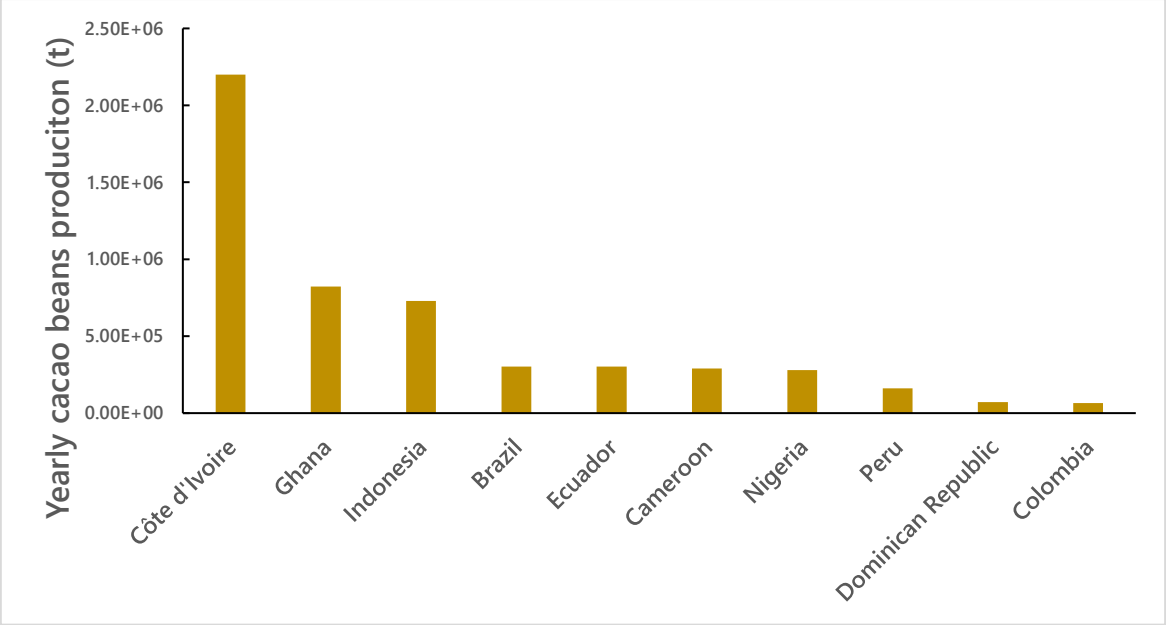


Figure 1.11: Cocoa beans are mainly produced in the west of Africa.

Data is the yearly production of cacao beans of 2021 for the first highest ten producers (Figure created with data from the Food and Agricultural Organisation of the United Nations (FAO, 2023)).

Smallholder farmers from four west African countries account for 70% of the global production (Beg *et al.*, 2017). In Côte d'Ivoire, the average farmland is 5.9 ha, with 3.8 ha dedicated to the culture of cacao (Callebaut, 2023). A report from Fairtrade international (Carla Veldhuyzen, 2019) estimated that 800kg of beans/hectare was the "sustainable yield" for farmers to produce to achieve a "high probability" of earning a living income.

After harvest, the beans are gathered and commonly covered with banana leaves for three to seven days (Figure 1.12) where a succession of fermentation reactions by yeast, lactic acid bacteria and acetic acid bacteria transform the carbohydrate portion of the cocoa pulp, producing chemical compounds associated with flavour such as alcohols, aldehydes, esters, lactic and acetic acids (De Vuyst and Leroy, 2020) . After fermentation, the beans are air-dried, collected by cooperatives, and then exported for further processing. Later stages of chocolate manufacturing for beans originating from West Africa, which feed the bulk chocolate production still take place mainly in Europe or North America.

Figure 1.12: Cacao beans fermenting on banana leaves in a plantation in Ghana (Allainguillaume, J., 2016, personal communication).

Most of the production is exported to Europe and the United Stated of America, but China and India are now seen as growing markets, increasing the pressure to maintain the cacao production yield as high as possible.

1.4 Threats to the cacao production.

Cocoa production requires a set of optimal climatic conditions (Kosoe and Ahmed, 2022), that includes precipitation, temperature, light exposure, and carbon dioxide levels (Aytehgiza and Gebresilasie, 2019). Any disruption of these parameters has the potential to significantly affect the production yields (Bomdzele and Molua, 2023). Furthermore, episodic climate event such as droughts (Gateau-Rey *et al.*, 2018), floods or dry spells (Okoffo, Mensah and Fosu-Mensah, 2016) are known to influence the growth of the tree.

Besides these abiotic factors caused by climate change, *T. cacao* is exposed to a range of pathogenic biotic stressors wherever it is cultivated (Adeniyi and Adeniyi, 2019). *T. cacao* trees are, like any living organisms, sensitive to infection by pathogens. It was estimated that up to 38% of the global

production is lost due to diseases (ICCO, 2017). In its American cradle, the tree has been infected with a variety of fungi-like eukaryotes, such as *Moniliophthora perniciosa,* an oomycete responsible for the witches' broom disease and several species of *Phytophthora* (Marelli *et al.*, 2019) that causes the black rot pod disease. *Moniliophthora roreri* is a basidiomycete closely related to *Moniliophthora perniciosa* that causes the cacao frosty pod rot in South America. *M. roreri* infects the pods only and can result in 90% losses if not treated (Bailey *et al.*, 2018). Several viruses such as the cocoa necrotic virus (Kenten, 1972) and the cacao yellow mosaic virus (CYMV) (Cooper and Jones, 2006) can infect *T. cacao* along with cacao swollen shoot virus (CSSV). To avoid the spread of diseases by particular viruses across areas of *T. cacao* production around the world, quarantine procedures have been developed to assess the presence of pathogens in material moved between countries. Material can be moved far away from *T. cacao* production to areas such as the International Cocoa Quarantine Centre at the University of Reading (ICQC-R) and tested for the presence of pathogens. Their main method of CSSV screening relies on grafting the assessed plant onto *T. cacao* rootstocks of the Amelonado variety and recording the potential development of symptoms over a period of three years. Quarantine screening has also enabled the discovery of novel viruses in asymptomatic trees such as Cacao Mild Mosaic Virus (CaMMV) and Cacao Yellow Vein-Banding Virus (CYVBV) in Cocoa (*Theobroma cacao*) discovered in the germplasm collection at Reading University (Ullah *et al.*, 2021).

1.5 Cacao swollen shoot virus (CSSV).

When introduced in the west coast of Africa, *T. cacao* plants were infected with local pathogens (Attafuah and Glendinning, 1965) through a host shift mechanism. CSSV, the agent of the endemic viral disease (Muller *et al.*, 2018) was found to infect local hosts trees prior to the implantation of *Theobroma cacao* such as *Cola chlamydantha, Bombax buonopozense, Hildegardia barteri or Adansonia digitata* (Ameyaw, 2019), all *Malvacea* plants. CSSV was found to infect other non-*Malvacaea* plants, notably in Ghana where a host range study identified the *Sterculiaceae, Bombacaceae* and *Tiliaceae* families as being natural hosts of the infection (Posnette, Robertson and

Todd, 1950). The transmission of CSSV can be documented by mechanical infectivity experiments where the mealybug vector is placed on suspected alternative hosts and the development of symptoms is then monitored and visually confirmed (Brunt and Kenten, 1962). More recent techniques to study the diversity of alternative plant hosts involves molecular techniques such as PCR and phylogenetic relationship analysis (Abrokwah *et al.*, 2016).

Cacao swollen shoot virus (CSSV) disease is one of the most important factors limiting cacao production in West Africa. This disease, which was described for the first time in Ghana by Steven in 1936, is primarily confined to the countries of Ivory Coast, Ghana, Nigeria, Sierra Leone, Cameroon, and Togo. Susceptible cacao plants exhibit a reduction of fruit yield of up to 25%, a year after infection and eventually die within 5 years depending on the strain of CSSV. However, symptoms of infection can take up to three years to be visible (Ofori *et al.*, 2022). A study looking at the impact of diseases on the crop in 2005 revealed that CSSV alone resulted in a 50,000 metric tons loss every year.

In comparison, the annual world production is estimated to be around 4,000,000 metric tons and epidemiology studies predict that the virus will have an increased impact in the future (Ploetz, 2006).

Upon the onset of the infection by CSSV, symptoms may be observed in several parts of the tree, including leaves, stems, roots and on the pods. Young leaves may develop red vein-banding (Figure 1.13), stems may develop swelling (Figure1.14) and pods can exhibit deformation (Posnette, 1943).

Figure 1.13: Young *T. cacao* leaf showing red vein-banding associated with an infection with CSSV in Ghana (Allainguillaume, J. personal communication). A scale bar was added to provide an approximate size.

Figure 1.14: Swollen *T. cacao* stem associated with a CSSV infection in grove Ghana. (Allainguillaume, J. personal communication). A scale bar was added to provide an approximate size.

CSSV belongs to the genus *Badnavirus* and together with the *Tungrovirus* forms the family of *Caulimoviridae*. Members of this family infect plants of agricultural interest, such as Sugarcane Bacilliform Virus, Sweet Potato Pakakuy virus, Taro Bacilliform Virus, Jujube Mosaic-associated Virus, and Banana Streak Virus (ICTV, 2021) (Figure 1.15).

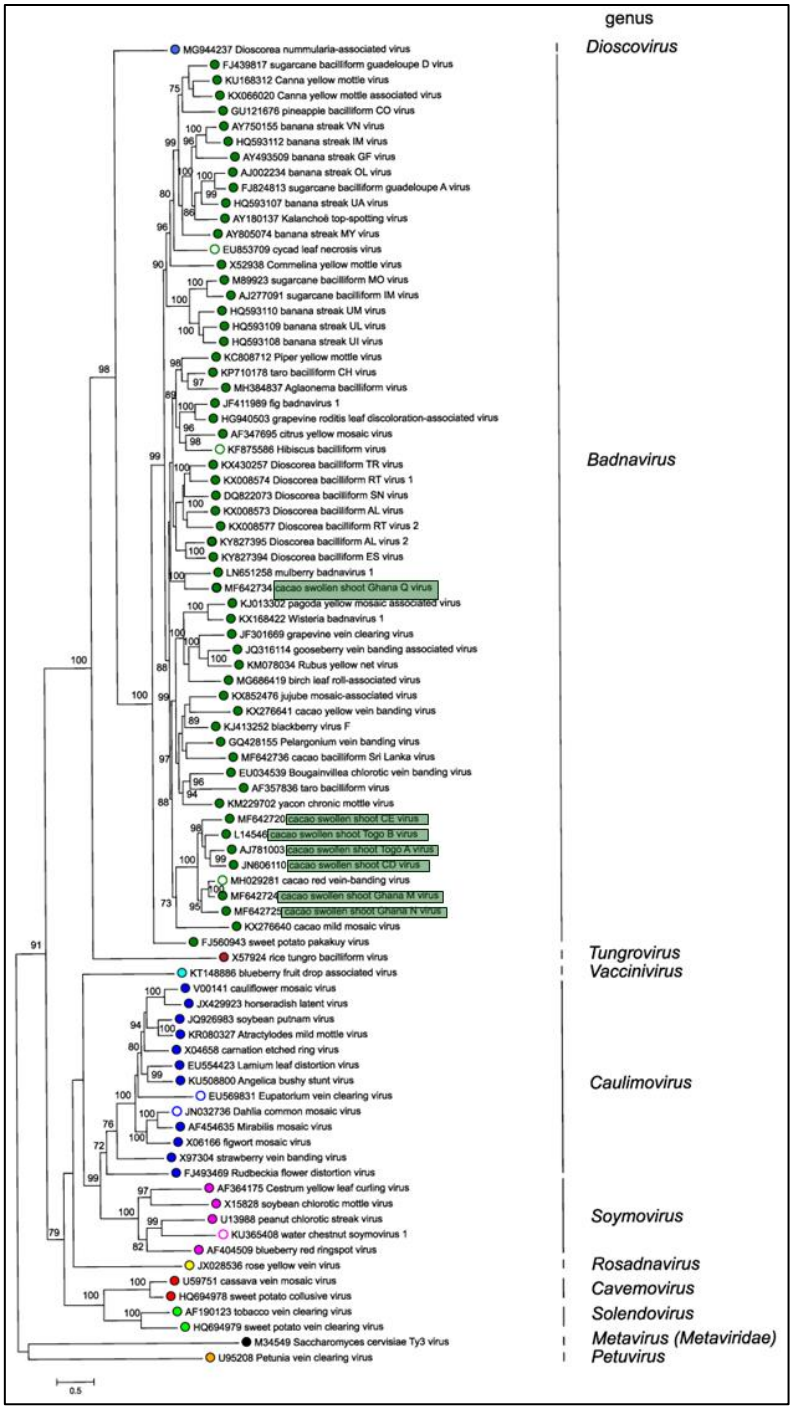


Figure 1.15: Phylogenetic tree of the family *Caulimoviridae*. Members of the CSSV group are highlighted in green. Adapted from (Teycheney *et al.*, 2020) under CC BY-SA 4.0 Licence.

The name Badnavirus derives from the characteristic bacillus shape of the virus particles and their double stranded DNA genome. CSSV genome is a 7 to 9.2 kilobases circular double stranded DNA molecule. CSSV appears as bacilliform non-enveloped particles of about 120-130 nm long and 28 nm in diameter (Figure 1.16).

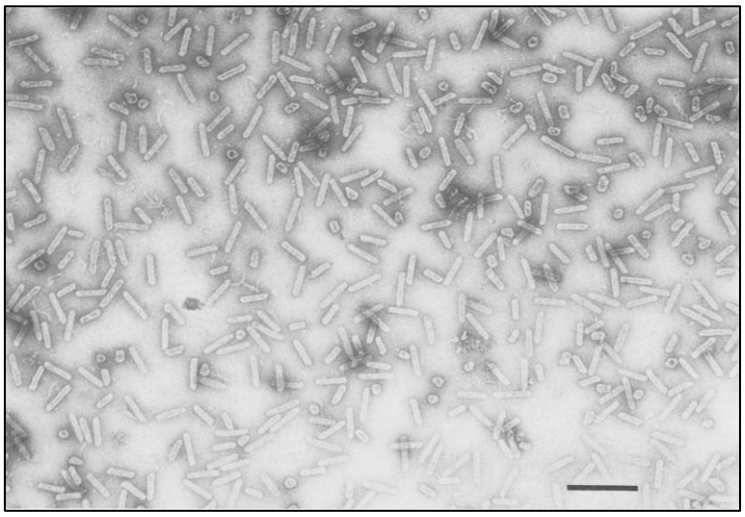


Figure 1.16: Purified preparation of CSSV obtained after sucrose density gradient centrifugation.

Bar represents 200 nm. Image adapted from (Lot, Djiekpor and Jacquemond, 1991) under agreement with CCC Press (with the license 1400237-1).

The virulent strain of CSSV, Agou1 from Togo was sequenced in 1993 (Figure 1.17) (Hagen *et al.*, 1993) and was identified as a double stranded DNA genome of 7161 bp. Another five genomes were obtained in 2004 by Muller (Muller and Sackey, 2005). In 2018, a study aiming at characterizing the high variability of the CSSV population in West Africa based on next generation sequencing enabled the reconstruction of twenty extra genomes sequences from field and collection samples (Muller *et al.*, 2018).

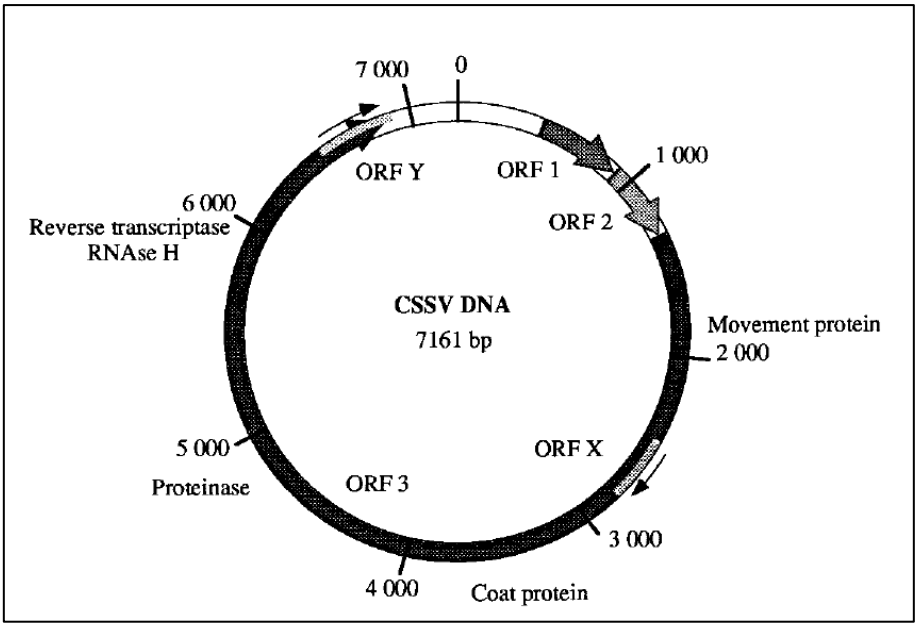


Figure 1.17: Genome map of CSSV strain Agou 1.

The arrows delineate the Open Reading Frames (ORFs) on the positive strand. The function of the encoded proteins are highlighted. Adapted from (Jacquot *et al.*, 1996) under licence CCC 5721321033888.

A phylogenetic analysis of reverse transcriptase and ribonuclease H gene sequences defined ten groups of CSSV. Group A corresponded to cacao swollen shoot virus Togo A virus (CSSTAV), as described in 2005 (Muller and Sackey, 2005), group B was the sequence of the cacao swollen shoot virus described for the first time in 1993 (Hagen *et al.*, 1993), group D was similar to cacao swollen shoot CD virus defined in 2012 (Kouakou and Kébé, 2012). Seven additional groups were also identified these are groups E, M, N, Q, R, S and T. Since whole genome data was not available for groups S and T and it was not possible to determine if they were new species. However, E, M, N, Q and R were respectively named cacao swollen shoot CE virus (CSSCEV), cacao swollen shoot Ghana M virus (CSSGMV), cacao swollen shoot Ghana N virus (CSSGNV), cacao swollen shoot Ghana Q virus (CSSGQV) and cacao swollen shoot Ghana R virus (CSSGRV). Interestingly, both studies by Muller included sampled sequences from Trinidad and Sri Lanka originating from plants displaying symptoms similar to an infection with CSSV. Analysis of the genomic data placed these viruses in a clade distinct from CSSV for the samples from Sri Lanka or in the same one but very far from CSSV for the samples from Trinidad. The viruses from Trinidad, cacao mild mosaic virus (CaMMV) and cacao yellow vein-banding

virus (CYVBV) were identified in 2017 (Chingandu *et al.*, 2017). Next generation sequencing of isolates showed variation in the size of the CSSV genome in ranging from 6985 to 7412bp.

Five Open Reading Frames (ORFs) were identified upon analysis of the viral sequences. ORF1 encodes a protein of unknown function of 138 to 161 amino acids depending on the isolate. ORF2 is translated into a 98 to 149 amino acids long nucleic acid-binding protein. ORF3 is the largest ORF and encodes a 1772 to 2018 amino acids polyprotein further cleaved into a capsid protein, three enzymes (Aspartyl proteinase, Reverse transcriptase and Ribonuclease H) and a possible cell-to-cell movement protein. ORFX was not found in all isolates and its size varies from 91 to 145 amino acids. In contrast, ORFY was found in all isolate sequences (115 to 146 amino acids). The role of both ORF X and Y is still unclear (Hagen *et al.*, 1993).

Caulimoviruses provided insights towards the understanding of the replication mechanisms of the genera. Upon entering the infected cell, the double stranded DNA is ejected from the viral particle and imported in the nucleus of the host (Figure 1.18).

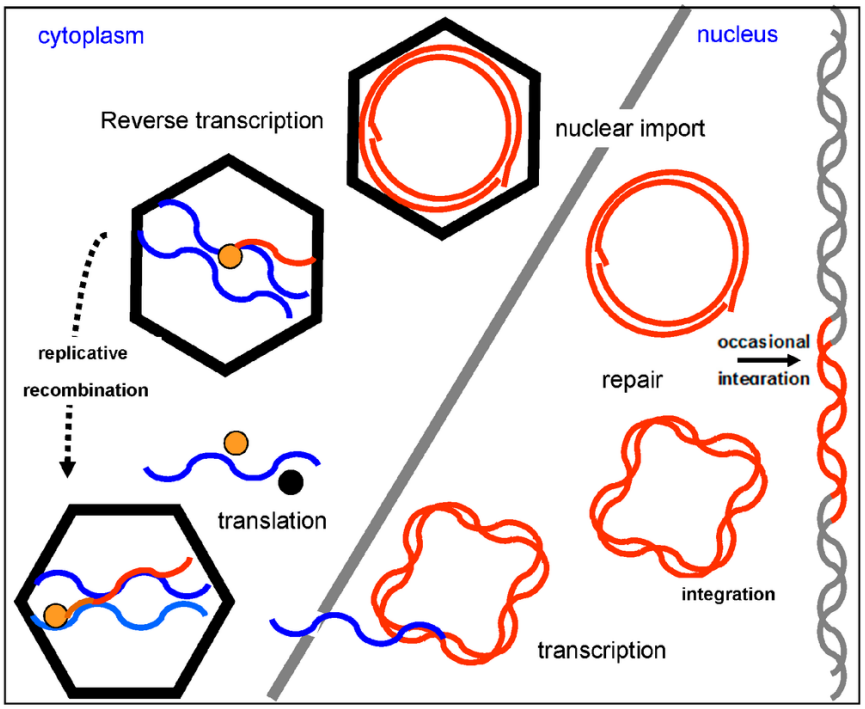


Figure 1.18: Replication cycle of a *Caulimoviridae* virus, *Cauliflower mosaic virus* (CaMV), that provides a model for the replication of badnaviruses.

Red circles represent the DNA of the virus, blue lines are the transcribed RNA copies. Figure adapted from (Bhat, Hohn and Selvarajan, 2016) under CC BT licence.

The viral DNA molecules have a nick present on both plus and minus strands, these are repaired during the first phase of the viral replication cycle. Repaired DNA is now supercoiled and then associated with histone proteins to form minichromosomes (Hull, 2001). Theses scaffolds constitute the transcription template for the host DNA-dependent RNA polymerase II. The transcription relies on the binding of the host tRNA initiator (tRNAmetinit) to the (-)-DNA strand while the (+)-strand is attached through a redundant purine-rich region. These two priming sites constitute the nicked ends of the circular molecules. The resulting transcripts have a dual function: transferred into the cytoplasm of the host, they are used as a polycistronic template for translation and as a pregenome molecule. In the second phase of the cycle, in the plant cytoplasm, the messenger RNA is translated. Two RNA copies are then packaged into viral particles and the reverse transcription occurs using the viral RT to yield a circular open double stranded DNA molecule (Bhat, Hohn and Selvarajan, 2016). During the repair phase, integration of viral sequences can occur through illegitimate recombination. The phenomenon has been documented in pararetroviruses (Harper *et al.*, 2003). Recently, the presence of badnaviral sequences within *T. cacao* genome has been confirmed by *in silico* and molecular analysis (Muller *et al.*, 2021a). This has a direct impact on the implementation of detection and control strategies for CSSV.

1.6 CSSV transmission.

CSSV is naturally transmitted by mealybugs (Figure 1.19) belonging to the *Pseudococcidae* family (Cotterell, 1943). They mainly live in the vicinity of the leaves and feed by puncturing the plant tissue to access the sap present in the phloem.

Figure 1.19: Mealybug *Planococcus citri*, one of the species able to transmit CSSV (Cranshaw, 2007)**. Image under CC BY 3.0 licence. A scale bar was added to provide an approximate size.**

There are sixteen species of mealybugs (*Hemiptera*: *Coccoidea*: *Pseudococcidae*) which have been reported in West Africa to act as vectors of CSSV (Wetten, Campbell and Allainguillaume, 2015). *Formicococcus njalensis* and *Planococcus citri* (Risso) are thought to be the most important viral vectors of cacao and are the predominant mealybugs on the crop in West Africa (Wetten, Campbell and Allainguillaume, 2016). Transmission of CSSV through seeds was also demonstrated on experimentally infected plants (Quainoo, Wetten and Allainguillaume, 2008), but was not reproduced since.

Symptoms of a CSSV infection are seen in leaves, stem, and roots and on the pods of the affected plants. Symptomatic observations of the leaves such as red vein banding in young leaves, yellow banding, a vein clearing producing a fern-like pattern or flocking and mottling due to chlorosis is used to identify the infection. Crucially, symptoms tend to only be observed in the later stage of infection once the disease might have already spread to neighbouring trees (Gilmour, Crozier and Kouamé, 2015).

1.7 Control of CSSV.

Since the discovery of CSSV in 1940 in Ghana, numerous attempts at controlling the virus have been made. The aim is to eradicate the virus and to do so, sources of infection must be identified and eliminated.

From 1946 onwards, Ghana launched several CSSV Eradication and Rehabilitation Programmes using the “cutting out method”, where infected and surrounding trees were cut down. This was combined with the use of synthetic insecticides targeting the mealybugs and the selection of virus-tolerant or resistant trees. In Ghana alone, 200 million cocoa trees were removed, representing 150 000 hectares of plantation (Amoah, Dzahini-Obiatey and Owusu, 2010). The main reason for the failure of this strategy comes from the difficulty of identifying infected trees. Detection and preferably early detection of plant pathogens is therefore critical to implement an appropriate response and limit the spread of the diseases. On the onset of infection by CSSV, the range of symptoms can be observed by inspecting the plants directly in the field. However, direct observation of symptoms does not always enable a definitive diagnostic. Similar symptoms can arise from both abiotic and biotic agents. They may also appear once the pathology has spread through the organism and already had a chance to contaminate other plants.

Lab-based techniques provide a more suitable mean of identifying infection by pathogens. These techniques detect either nucleic acid or an antigen from the pathogen. Polymerase chain reaction (PCR) (Sackey, ST; Bartels, PK; Amponsah, 1992) (Muller, Jacquot and Yot, 2001) is the main technique used to detect CSSV infection. First applied using primers deriving from other badnaviruses, specific CSSV markers were designed following the publication of the genome sequence of CSSV. Other nucleic acid-based detection techniques including dot blot hybridization (Sackey and Hull, 1992) and Southern blot (Sackey, S. T.; Lowor, S. T. & Dzahini-Obiatey, 1995) were also developed. Electron microscopy following purification of infected samples enabled observation and identification of CSSV, using negative staining procedures with phosphotungstate (Brunt, Kenten, and Nixon, 1964).This however

is a difficult procedure which requires highly purified preparation of CSSV extract and that has since only been repeated successful twice (Lot, Djiekpor and Jacquemond, 1991; Adomako, 1985). Production of highly purified CSSV samples has also made possible the development of antisera in animals. Antibodies-mediated molecular methods were successfully applied to study and identify CSSV. These techniques include enzyme-linked immunosorbent assay (ELISA) (Sagemann *et al.*, 1983), virobacterial agglutination tests (Hughes and Ollennu, 1993) , immunosorbent electronic microscopy (Adomako *et al.*, 1983) and immunocapture polymerase chain reaction (ICPCR) (Hoffmann *et al.*, 1997a).

These methods all have in common a requirement for further development and optimisation and more importantly substantial operational costs, making them unsuitable as routine diagnostic procedures for CSSV detection in the west of Africa. PCR was shown to be inhibited by leaf extracts (Hoffmann *et al.*, 1997b). Furthermore, a recent survey of the ten most used primers for CSSV diagnostic evidenced high variability in the efficiency of their detection rate, with the best performing primers pairs exhibiting 23% only (Ameyaw *et al.*, 2022).

1.8 Biosensing.

Biosensors, in contrast, offer the possibility to bring laboratory capabilities directly into the field and require minimal training of the personnel. Biosensors are measurement devices enabling the qualification or quantitation of a molecule of biological interest from a target sample. The conceptual basis of a biosensor was laid in the beginning of the 20th century (Cremer, 1906) when it was established that the concentration of lactic acid in a liquid could be correlated to the electric potential created between liquids separated by a glass membrane. Following the discovery of the concept of pH and later the possibility of immobilising enzymes on solid matrixes, the potential to measure relevant human biomolecules appeared. The first biosensor invention is credited to Leland C. Clark Jr in 1956 to measure oxygen (Heineman and Jensen, 2006).

Biosensors are analytical devices typically composed of a biological recognition element (BRE), a biomolecule, usually a protein or a nucleic acid, that can recognise and bind to or interact with the analyte of interest. This binding event is then recorded through a signal transducer. The function of the signal transducer is to convert the recognition of the biomolecule into a signal that can be measured at a detector. The third element captures the measurable signal and displays it to the user as shown below (Fig 1.20).

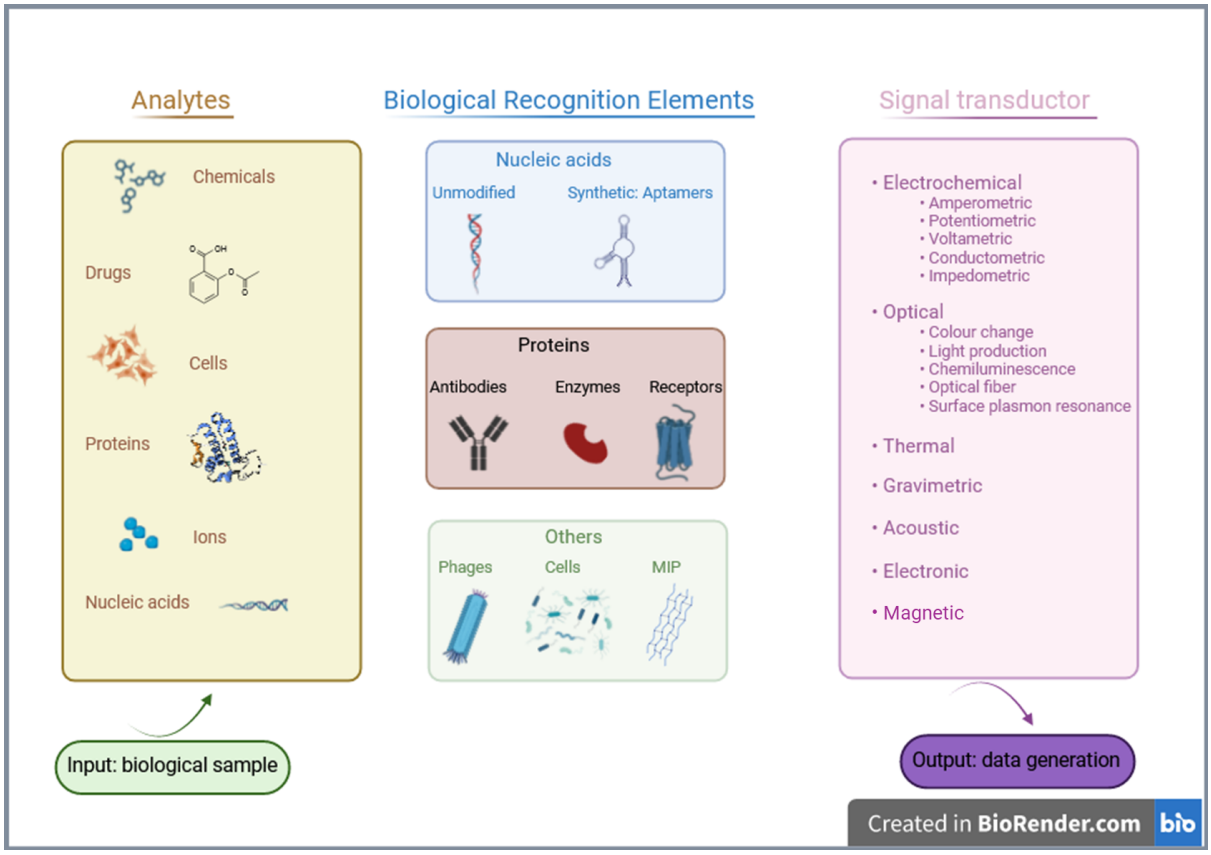


Figure 1.20: Diversity of analytes detected biological recognition elements and signal transduction of biosensors.

1.9 Biological recognition elements.

The BRE, or bioreceptor, is the interface between the user and the analyte targeted. BRE are usually nucleic acids, including modified versions. BREs also include proteins, bacteriophages, whole cells, or less frequently molecular imprinted polymers (Nawaz *et al.*, 2021). The next sections will focus on nucleic acids and protein-based bioreceptors.

1.9.1 Nucleic acid BREs.

Nucleic acid based BREs utilise the ability of nucleic acids molecules to assemble to form double stranded molecules. The molecular basis of DNA inter-strand bonding was established alongside the elucidation of DNA structure by Rosalind Franklin and her team (Franklin and Gosling, 1953) who produced an X-ray diffraction pattern of the molecule in 1953, followed by the demonstration of the double helix model by Watson and Crick in 1953 (Watson and Crick, 1953). DNA is composed of a backbone made of an alternation of phosphate and 2- deoxyribose (a pentose) residues, held together through phosphodiester bonds. On this backbone are attached the two types of nitrogenous bases included in DNA, the purines, adenine and guanine, and pyrimidines, cytosine, and thymine. A unit composed of the sugar and phosphate moieties, and the associated base is called a nucleotide and represents the monomeric component of DNA. A remarkable feature of the polymeric form is the fact that it has a directionality. The strands are anti-parallel, and the direction is provided by the number of the chemical moiety present on the end of the molecule: on the 3' end sits the hydroxyl group of the ribose and on the 5' end is found the phosphate group. DNA strands interact via hydrogen bonding between the nucleotide bases in a way defined by the rule of Chargaff, after the name of the Austrian biochemist who discovered this principle in 1952 (Elson and Chargaff, 1952) : "the amount of adenine (A) is usually similar to the amount of thymine (T), and the amount of guanine (G) usually approximates the amount of cytosine (C). In other words, the total amount of purines (A + G) and the total amount of pyrimidines (C + T) are usually nearly equal" (Pray, 2008). It has since been found that the adenine interacts with thymine and guanine with cytosine. The chemical nature of the guanine and the cytosine allows the formation of three hydrogen bonds between them, where the adenine and thymine interaction involve two hydrogen bonds only. Therefore, the sequence of a given DNA molecule will strongly influence its stability. This type of interaction, that is positive for the stability of the duplex DNA, is known as the Watson and Crick hydrogen bonding or base pairing. H-bonds are not the only forces involved in the interaction. Indeed, the stability of the molecule is the sum of the forces enabling bonding between the strands and the negative forces repelling the strands (Horton, 2006).

Thermodynamics studies (Horton, 2006) have established that the other main force stabilising the DNA helix structure are the stacking interactions that are created between the pi-orbitals of the aromatic ring from the nucleotide bases and are of hydrophobic nature (Morozov *et al.*, 2005). Using the affinity and specificity of the base pairing, DNA or RNA molecules can be selected to bind to a targeted genomic region of an organism of interest. The biosensor will then be able to detect the hybridisation of the BRE to the complementary sequence from the organisms targeted.

1.9.2 Aptamers.

Nucleic acid based BRE may be derived from naturally occurring sequences, or alternatively custom made. Aptamers are synthesised DNA, RNA or XNA (xeno-nucleic acid construct) that emulate the concept of monoclonal antibodies for antigen recognition, as they are selected to recognise a biomolecule of interest, with high specificity and affinity. They have been used in a range of fields such as therapeutical research (Bunka, Platonova and Stockley, 2010), environmental contamination surveillance (Mishra, Sharma, and Mishra, 2018; Kim et al., 2009; Li et al., 2009) and food safety (Gaudin, 2020; Reybroeck *et al.*, 2007). Aptamers are identified by the systematic evolution of ligand selection by exponential enrichment (SELEX), a method that relies on the iterated selection of nucleic acid sequences flanked by primers from a randomised pool by binding them to the desired ligands. While conventional nucleic acids such as DNA and RNA are used as aptamers for biosensing, further developments were introduced to create XNA. Originally designed for the fundamental study of nucleic acids structure and evolution, (Herdewijn and Marlière, 2009) their improved characteristics lead to their inclusion in biosensors, as biological recognition elements. Enhanced features of the XNA are introduced by chemical modifications to any part of the nucleobase. For instance, PNA, or peptide nucleic acids, are XNA where the phosphate and deoxyribose backbone have been replaced by a peptide and have demonstrated increased thermodynamic stability, because they possess the property of being electrically neutral (Ma *et al.*, 2016), which abolishes inter-strand repulsion. In addition, the modification improves the resistance of the PNAs to nuclease action (Lundin *et al.*, 2006).

These structural modifications also further enabled the design of label-free assays for the detection of protein in solution (Xu *et al.*, 2018) , micro RNAs (Li *et al.*, 2016) or DNA (Keighley *et al.*, 2008).

1.10 Proteins BREs.

The two main types of proteins used as BREs are antibodies and enzymes. As with DNA BREs they rely on the specificity of their interaction with the analyte they are targeting.

Immunoassays have been used for more than sixty years since the seminal work of Yalow and Berson using radio-labelled antibodies (Yalow and Berson, 1959). Antibody-based biosensors enable detection of the target without its prior purification, and biological samples such as blood, saliva or urine may be used directly (Chambers *et al.*, 2008). Furthermore, monoclonal antibodies are specific and possess high affinity for their ligands. A critical parameter when antibodies are used as BREs is their immobilization on the platform (Sharma, Byrne and O'Kennedy, 2016), to enable the epitope/paratope interaction: incorrect orientation of the antibody can result in a 200-fold decrease in the signal compared to an adequate positioning (Trilling, Beekwilder and Zuilhof, 2013). Usual strategies to immobilise the antibodies rely either on non-covalent reactions, using electrostatic interaction or entrapment (John *et al.*, 1991) or by covalent association, using an amine or thiol groups (Sharma, Byrne and O'Kennedy, 2016). Other immobilisation strategies take advantage of the structure of the fragment crystallizable (Fc) region of the antibodies and use the affinity-based interaction with protein G or A, as intermediate proteins, to minimise the potential damages to the BRE associated with the chemical/physical immobilisation procedures (De Juan-Franco *et al.*, 2013): protein A or G can be first coated on the surface of the biosensor, providing an adequate orientation of the antibody through the specific interaction of these proteins with the Fc region. Biosensors based on antibodies have been developed for the detection a range of analytes from 2,4,6-trinitotoluene (TNT) (Park *et al.*, 2010), to human viruses like HIV (Encarnação *et al.*, 2007) or plant viruses such as the cucumber mosaic virus (CMV) (Jiao, Sun and Zhang, 2000a) or the bacterium *Pantoea stewartii* (Zhao *et al.*, 2014), a threat to the culture of corn, jackfruit, sugarcane and rice.

According to a review on Scopus conducted for the period from 2010 to 2018, protein BREs based on enzymes were ranked first by the number of articles published, before antibodies based BREs (Nguyen *et al.*, 2019). This class of biosensor, first used in 1967 (Mehrotra, 2016) uses the catalytic activity of an enzyme to transform an analyte into a detectable signal, through electrochemistry (Kilic *et al.*, 2023) . While the immobilisation of the enzyme is, as with antibody-based biosensors, a delicate step, this type of sensor presents the benefit to be reusable (Katchalski-Katzir, 1993), because the enzymes are not depleted during usage. Another advantage of using enzyme BREs is the wide variety of catalytic bioreactions covered by these proteins, enabling in return a great flexibility in the type of signal that can be captured by the user. This diversity is reflected in the range of applications developed that utilise enzymes. Such sensors are used for medical applications, with detection of cholesterol (Alvi *et al.*, 2013) , lactate (Hernández-Ibáñez *et al.*, 2016) , secretory phospholipase group2-IIA (Mansor *et al.*, 2018), a marker of bacterial sepsis, and the most famous glucose sensor for the management of diabetes (Yoo and Lee, 2010), or for environmental analyte monitoring with an assay detecting catechol, a phenolic pollutant (Kurbanoglu and Ozkan, 2018). Enzyme biosensors have also been developed for the assessment of food products, with assays designed for the detection of tyramine, a toxic derivative of tyrosine (Sánchez-Paniagua López, Redondo-Gómez and López-Ruiz, 2017), L-ascorbic acid (Ibupoto *et al.*, 2011) and glucose content in fruits (Ang, Por, and Yam, 2015). Volatile organic compounds can be markers of pathogen infection in plants and assays using enzymes were developed for their detection, for instance, methylsalicylate that is produced by plants in response to a biotic stress (Fang, Umasankar and Ramasamy, 2016) using alcohol oxidase and horseradish peroxidase as enzymatic BREs.

1.11 BRE integration to biosensing platforms.

The interaction of the BRE with the target analyte produces a measurable physicochemical change that a specific transducer will be able to record. Biosensors may be classified depending on the nature of the transducers. Electrochemical transducers-based biosensing platforms are designed to respond

to electricity. Thus, amperometric transducers rely on a change in electric current, potentiometric transducers on voltage signal variation, impedimetric biosensors on impedance changes and on conductivity changes through conductive polymers for conductimetric transducers all resulting from the BRE/analyte binding event (Dyussembayev *et al.*, 2021), often via the reaction between a labelled enzyme and an electroactive substrate. Alternatively, transducers may react to the modification of the mass of analyte biomolecule detected by the BRE. Quartz crystal microbalance transducers take advantage of the piezoelectric properties of quartz crystals (Bragazzi *et al.*, 2015). A thin section of quartz crystal is assembled to an electrode, providing a current enabling the quartz to resonate. A mechanical pressure will elicit a modification of the resonance frequency that can be recorded and associated with the binding of the BRE to the target analyte. Another type of mass modification transducer relies on microfabricated cantilevers (Fauver *et al.*, 1998), wherein the BRE is attached to the tip of the cantilever and upon interaction with the target analyte, the resulting deflection of the cantilever is detecting by optical measurements mediated by a laser diode.

These technologies have been used for the development of several types of biosensors for the detection of pathogens in crops of economic interest (Table 1.2).

Table 1.2: Examples of biosensors developed for the detection of plant viral infection.

Techniques	Target pathogen(s)	Crop	Limit of detection	Reference
Voltammetric enzyme-based biosensor	Cucumber mosaic virus	Cucumber	0.5 ng/mL	(Jiao, Sun, and Zhang, 2000b)
Amperometry (polypyrrole nanoribbons on gold microelectrode array)	Cucumber mosaic virus	Cucumber	10 ng/mL	(Chartuprayoon *et al.*, 2013)
Surface plasmon resonance biosensor	Cymbidium mosaic virus	Orchids	48 pg/mL	(Lin *et al.*, 2014)
DNA hybridization voltammetric system	*Candidatus* Phytoplasma cynodontis	Sugar cane	4.7 pg/mL	(Wongkaew *et al.*, 2014)
Electrochemical based biosensor	Banana streak virus, Banana bunchy top virus	Banana	50 fmol	(Tang *et al.*, 2007a)
Quartz crystal microbalance biosensor	Potato virus x	Potato	2 ng/mL	(Drygin *et al.*, 2012)
Impedimetric Immunosensor	Prunus necrotic ringspot virus	Genus *Prunus*, Rose, hops	250 ng/mL	(Jarocka *et al.*, 2013)
Lateral Flow Assay (Gold nanoparticles tag)	Potato virus x	Potato	$2x10^3$ pg/mL	(Razo *et al.*, 2018)
Lateral flow immunoassay	*Pantoea stewartii subsp. Stewartii*	Maize	5.38 pg/mL	(Feng *et al.*, 2015)
Magnetic microsphere assay	watermelon silver mottle virus	Watermelon	$20.5x10^3$ pg/mL	(Charlermroj *et al.*, 2013)

1.11.1 Optical biosensors.

Optical biosensors are the most common type of biosensors (Damborský, Švitel and Katrlík, 2016), because they can yield fast results directly readable without requiring complex equipment. This enables a reduction in the cost associated with the fabrication of the sensor (Qin *et al.*, 2022). They can be classified as either label or label-free detection, depending on how the binding event is monitored.

Label-based sensors involve the production of light, fluorescence or colour change using a reporter molecule following the interaction between the BRE and the analyte of interest. Label-free assays, on the other hand, do not require addition of a reporter molecule. Surface plasmon resonance (SPR), described four decades ago (Liedberg, Nylander and Lunström, 1983) is based on the production of plasmons, a measurable oscillation of quasiparticles generated by electron movements occurring at the interface between glass and a liquid upon illumination with a polarized light at a set angle, that is captured on the surface of the sensor (Damborský, Švitel and Katrlík, 2016)and does not necessarily require additional labelled probes. Instead, a quantifiable signal can be created by the modification of the angle of reflection of the polarized light consecutive to the binding of the analyte, for instance an avian influenza virion (Bai *et al.*, 2012), or by the change in colour due to aggregation of metallic nanoparticles for the detection of the bacteria *Vibrio fischeri* (Shin *et al.*, 2018).

1.11.2 A biosensor to detect CSSV "in field"?

For the '*in field*' detection of CSSV in cocoa in West Africa, an ideal diagnostic device would utilise a biosensor that is specific, sensitive, fast, cheap to produce and easy to use by non-qualified personal. With these specifications in mind, an optical biosensor would be ideal because of their low cost and because the outcome of the test can be visualised in the field (Eksin and Erdem, 2023). Furthermore, this type biosensor is reliable, sensitive, and robust (Chen and Wang, 2020).

1.11.3 DNA optical biosensor to detect CSSV.

Targeting the DNA of CSSV using an optical biosensor through DNA hybridization could be a suitable approach. Since the production costs should be kept low, strategies involving complex transducers such as SPR, optical fibres or advanced signal detection equipment to detect chemiluminescence or fluorescence would not be appropriate. Instead, the strategy should rely on the production of a directly observable signal.

One of the main challenges arising for the development of a DNA hybridisation biosensor using current protocols lies in the control of hybridization temperature on a portable device, whilst keeping the production costs low. Traditional laboratory hybridization techniques for nucleic acids such as Southern, northern or dot blot rely on temperature modification and long incubation times (up to 12-16 hours); these are typically conducted in a laboratory set-up (Ausubel *et al.*, 2003). Traditionally, probes are oligonucleotides tagged with either a radioactive nucleotide or with a biotin moiety to enable the use of safer chemiluminescent detection systems. These methods give information about the presence or absence of given targeted sequences within a genome or a sub- genome. The biological principle underlying this technique is the ability of the DNA strand to bind to its complementary sequence in a specific manner. Modification of temperature and ionic strength is currently used to denature double stranded DNA and then enable hybridisation. However, controlling the temperature in the field is not ideal because it would increase the cost of the device. An alternative approach is a “cold hybridization” which would enable hybridisation of target DNA and oligonucleotide probe at room temperature.

1.11.4 Protein optical biosensor to detect CSSV.

A protein based BRE could be used as well to immobilise the DNA of CSSV from infected *T. cacao* samples. At the time of the design of this project, no antibodies to CSSV were available. However, zinc finger proteins (ZFP) have the potential to bind to any nucleic acid sequence (Grover *et al.*, 2010) and as such represent an interesting avenue to design bespoke biological recognition elements to target

any genome region of CSSV. ZFP are metalloproteins involved in a range of physiological pathways including DNA recognition and transcriptional activation (Krishna, Majumdar and Grishin, 2003a). A more detailed account of zinc finger structure and function is detailed in the second chapter. Besides, this class of protein has been successfully used in optical biosensors (Shim *et al.*, 2021)

Aims and objectives of the thesis.

The aim of this project is the design and functional testing of integrated tools for the development of a biosensor to detect CSSV in the field. The design is required to accommodate with the specific context of a usage in West Africa and requirements are set to limit both the production costs and the training required by the field workers to operate the biosensor.

The first objective is to design, produce and test the functionality of a synthetic zinc finger protein based BRE to detect the virus. This will be achieved by first selecting a conserved locus within the *de novo* sequenced genome of an isolate of the strain New Juaben (NJ) of CSSV, a Ghanaian strain maintained at the envirotron at UWE.

The second objective will focus on developing a nucleic acid based BRE that relies on a combination of DNA probes aimed at hybridising with the CSSV genome of the same New Juaben (NJ) strain. The ability of the BRE to interact with its target will be assess by dot blotting and High Resolution Melt (HRM) analysis.

The third objective is to create a simple DNA extraction method from either CSSV-infected *T. cacao* leaves or vascular tissues that is compatible with the assay used in the biosensor and to test the BRE developed on CSSV infected *T. cacao* tissue extracts. Detection of the target viral sequence will occur through a newly developed “cold” hybridization method, to limit the equipment required in the field.

.

Chapter 2 : Production of a DNA-binding protein as a Biological Recognition Element to capture CSSV DNA.

2.1 Introduction

Proteins are one of the biomolecule classes commonly used as the biological recognition element (BRE) of biosensors. Conveniently produced and offering high specificity, they constitute an instrument of choice to target an analyte of interest. Proteins traditionally used as the biological recognition element of a biosensor are antibodies or enzymes.

Immobilised, modified, or native antibodies have been successfully employed to detect plants of agronomic relevance such as cowpea mosaic virus (Torrance *et al.*, 2006), tobacco mosaic virus (Dubs, Altschuh and Van Regenmortel, 1992) and orchid viruses (Eun *et al.*, 2002) using laboratory-based biosensing platforms. Antibodies may be used without requiring an additional label, depending on the signal transduction system they are used with. Surface plasmon resonance-based sensors for instance enable the detection of the antibodies' target by measuring a difference in the refractometry of the sensor surface. Likewise, antibodies immobilised on a quartz crystal microbalance will allow the measurement of changes in the resonance oscillation frequency following binding of the targeted antigen. Orientation of the immunoglobulins is important to elicit appropriate binding of the cognate antigen and particular care must be given to the chemistry used for the attachment of the biomolecule to the surface of the sensor. The most usual substrates to which the antibodies have to bind to are inorganic (metal, glass) (Skottrup, Nicolaisen and Justesen, 2008). Antibodies are either directly physically adsorbed to the surface or chemically covalently attached to the substrate. Several companies are now providing ELISA and lateral flow tests for detecting plant pathogens of crops of agricultural interest. While providing fast results, these tests are not cost-effective; for example, fire blight diagnostics for apple orchards cost between $4.25 and $8.50 per sample, and the limit of

detection of the devices was a 1000-fold lower than a system relying on nucleic acid amplification: the minimum CFU detected was $1x10^6$ for the LFT and $1x10^3$ for the NAT (Singh *et al.*, 2021). Similarly, LFTs developed to detect the Potato Virus Y were found to be 1000-fold less sensitive than loop-mediated isothermal amplification assay (Przewodowska et al., 2015).

Enzymes are also highly specific biomolecules that interact selectively with their substrate to catalyse a biochemical reaction. The first blood glucose test strip, developed in 1964 by Anton H. Clemens (Hirsch, 2018) used a mixture of immobilised glucose oxidase, peroxidase and a chromogen, the latter providing a signal indicative of the concentration of blood glucose in a dose-dependent manner (Krynski and Logan, 1967). Enzymatic reactions are now used coupled to electrochemical transduction platforms by utilising the bio-electrocatalytic abilities of selected enzymes immobilised on an electrode to produces a current when the target analyte is recognised. This system has enabled the detection of plant pathogens, although indirectly, by detecting biotic stress molecules such as volatile organic compounds (VOC). A portable biosensor was developed using three enzymes, esterase, salicylate hydroxylase and tyrosinase immobilised on a screen-printed electrode for detecting methyl salicylate, which is produced by the infected plant in response to tobacco mosaic virus infection (Fang and Ramasamy, 2018).

While interesting, neither of these types of proteins are commercially available to detect cacao swollen shoot virus. Purified, readily available antibody does not exist and could prove arduous to produce cost effectively. Recently, antibodies to CSSV have been produced *in vitro* (Allainguillaume *et al.*, 2019). However, the efficacy of this antibody needs to be confirmed on a wider range of plant material. Also, the pathogenic mechanisms consecutive to a CSSV infection are not well understood yet, which prevent the selection of relevant targets for a suitable enzyme-based biosensor system.

In contrast, DNA-binding proteins constitute a potential, yet unexplored avenue for capturing analytes of interest. They theoretically possess both selectively and sensitivity, which are required characteristics of an adequate biological recognition element. Amongst DNA-binding protein, zinc

finger proteins (ZFP) are a major class, with an estimated 3-10% of the human and murine genome encoding them (Andreini et al., 2005; Andreini et al., 2006) and as such have been extensively studied. First revealed with a study on the transcription factor IIIA (TFIIIA) (Miller, McLachlan, and Klug, 1985), a eukaryotic nucleic acid binding protein binding to both DNA and RNA (Shastry, 1996) from *Xenopus laevis* oocytes, their three-dimensional structure has since been resolved.

This type of metalloprotein is well defined and is characterised by the presence of one or multiple (thirty to a hundred amino acid) zinc finger domains bound to a zinc cation (Zn^{2+}). Resolution of the structure of TFIIIA enabled the understanding of the structure of the zinc finger domains. The structure relies on the cysteine (C) and histidine (H) residues that coordinate the zinc cation, stabilising the structure of the domain (Figure 2.1).

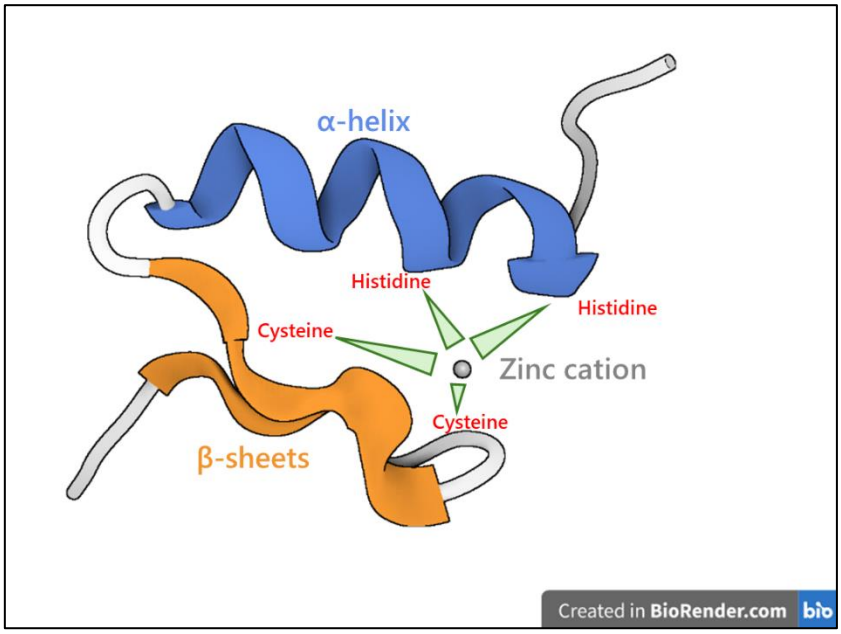


Figure 2.1: C2H2 Zinc finger protein domain. Two histidine residues on the α-helix and two cysteine residues on the β-sheets interact with a zinc cation to stabilise the structure of the domain.

This arrangement gave its name to the C2H2 or CCHH motif which is a signature of the zinc finger domains. When the histidine residues are substituted, TFIIIA loses its ability to regulate transcription (Del Rio and Setzer, 1993). The consensus sequence of the C2H2 domains is O-X-C-X_{2-4}-C-X_3-O-X_5-O-X_2-H-X_{3-4}-H-X_{2-6}, where X is any amino acid and O is a hydrophobic residue. The three-dimensional

structure of the zinc finger domains is composed of two β-sheets containing both conserved cysteines of the canonical motifs, connected to an α- helix with both histidines involved in the binding to the zinc cation. C2H2-like domains are the most prevalent DNA-binding domains found in metazoan proteins (Persikov *et al.*, 2015). Typically, several of these domains are arranged to bind to their cognate DNA molecule, with each domain binding specifically to three consecutive nucleotides on the same strand, 3' to 5'. A potential fourth nucleotide may be bound to the domain, establishing a cross-strand connection. Contact is made with the major grove of the DNA molecule through canonical positions located on the α-helix of each domain (Wolfe, Nekludova and Pabo, 2003). Because of the wide variety of domains existing, any DNA molecule can theoretically be bound to a given combination of C2H2 units. This potential has been used to design genome editing tools by targeting site-specific double-strand breaks using a fusion of the DNA-binding domains of the Zinc finger proteins to the bacterial type IIS *Fok*I endonuclease. This led to the modification of plants of agronomic interest (Townsend *et al.*, 2009), analysis of the model organisms *Drosophila melanogaster* and *Caenorhabditis elegans* (Bostan and Elibuyuk, 2010) and *Xenopus laevis* oocytes (Bibikova *et al.*, 2001). Genome editing procedures now mainly rely on the Clustered-Regularly Interspaced Palindromic Repeats/CRISPR-associated proteins (CRISPR/Cas) system. However, zinc finger proteins, with their relative ease of use and production cost, are an interesting tool for specifically targeting a region of interest within a genome.

The mechanism underlying the recognition of specific DNA sequences by corresponding zinc finger domains has been well studied and has led to the creation of web-based bioinformatic tools such as Zinc Finger Tools (http://www.zincfingertools.org) (Mandell and Barbas, 2006). Authors have considered known cognate DNA triplets of each zinc finger domain to design a corresponding alphabet of sixty-four domains to create "polydactyl" recombinant synthetic proteins able to bind unique DNA sequences (Beerli *et al.*, 1998). Characterization of the domains is based on structure determination, site-directed mutagenesis, and phage-display selection (Mandell and Barbas, 2006; Dreier, Segal, and Barbas, 2000).

2.1.1 Aim and objectives of this chapter.

This chapter describes the design, fabrication, and functionality testing of a synthetic zinc finger protein to be used as a BRE for detecting CSSV's genomic DNA.

Objectives are set as follows:

- The feasibility of designing an 18 bp unique target sequence from CSSV NJ will be evaluated. This will be achieved by screening the genome sequence of CSSV NJ using the zinc finger tool website (http://www.zincfingertools.org) to search for potential zinc finger binding sites and select a candidate protein.

- The candidate zinc finger protein will be produced. The gene corresponding to the protein will be cloned into a bacterial expression system and the recombinant produced. This process will be monitored by SDS-PAGE and Western blotting.

- The functionality and specificity of the zinc finger will be assessed by a combination of biochemical tests. The Double Filter Binding Assay (DFBA) will determine if the protein binds to a synthetic short oligonucleotide containing the targeted sequence. Electrophoretic Mobility Shift Assay (EMSA) will determine the binding specificity of the protein to two genome amplicons. HRM will be assessed to evaluate if it can be used to inform about the binding of the protein to oligomeric cognate DNA sequence.

- The feasibility of using the protein as BRE will be investigated by pull-down assay using both magnetic particles with DNA extracts from CSSV-infected *Theobroma cacao* samples. Superparamagnetic particles modified with nickel will be used to capture the recombinant zinc finger protein pre-incubated with DNA extracts from infected *T. cacao* samples.

2.2 Material and Methods.

2.2.1 Zinc Finger production.

2.2.1.1 Zinc finger design.

The genome sequence of CSSV NJ, a strain of CSSV originally from Ghana and maintained at the UWE envirotron, was copied in the search field of the "zinc finger tools" website (Barbas and Mandell, 2006). The website scans the target genome for ZFP binding sites and returns candidate regions based on the set parameters. Selected regions are associated with a set of binding zinc finger domains. The tool also provides the DNA sequence corresponding to the protein so the gene can be cloned.

The size of the targeted region was set to 18 base pairs, because this is the shortest theoretical unique sequence within a genome (Choo and Isalan, 2000). Parameters selected for the search were total library of zinc fingers, all triplets, search on contiguous sites with a minimum target size of 18 base pairs, not on both strands. The website then returns a list of potential sites with a score for each. Several scoring options are available on the site and the "base score" that relies on an ELISA-based screening of each domain affinity towards a given DNA triplet to provide the most optimal combination of domains, was selected.

Candidate ZFP binding sites were aligned on the sequence of the CSSV New Juaben isolate used in this study and a blast search against the non-redundant nucleotide database was performed.

2.2.1.2 ZFP2143K Expression and purification.

The 528bp synthetic gene coding for the candidate zinc finger protein 2143K was commercially synthetised by assembling synthetic oligonucleotides and PCR products by Invitrogen, cloned into a pMK-RQ vector, and delivered in *Escherichia coli* K12 DH10B T1R cells (Invitrogen, UK).

The gene was then subcloned in a pET-21(+) expression vector (Novagen, UK) that enables high level expression under the T7 promoter and adds an hexahistidine tag at the C-term of the protein. Cloned plasmids were transformed in *E. coli* DH5α cells. Cells were screened by extracting the plasmids and

digesting by *Bam*HI/*Nde*I and *Nde*I/*Xho*I to verify the correct sequence insertion. Plasmids from positive clones were transformed by heat shock into *E. coli* BL21(DE3) strain for expression. Resulting clones were screened by PCR and positive clones inoculated onto fresh media for propagation.

Overnight cultures from positive clones were diluted and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for three hours when the optical density of the culture reached the exponential phase. Cells were then collected by centrifugation and the pellet frozen. Thawed pellets were weighed and resuspended in ZFP Lysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, 10 µM $ZnCl_2$, pH 8 with NaOH, supplemented with a protease inhibitor cocktail (Sigma-Aldrich, UK) and 1mg/mL lysozyme (Thermo, UK). Lysates were incubated for 30 min on ice then pulse-sonicated (6x10s with 10 s pauses, at 240 W). Lysed samples were then centrifuged at 10,000x g for 20 min at 4°C. Supernatant fractions were collected to be further purified by immobilised metal affinity chromatography (IMAC). Pellets were also collected to evaluate the solubility of the recombinant protein.

The soluble supernatant fraction was incubated with a 50% slurry of Ni-NTA agarose (Qiagen) that was previously equilibrated in wash buffer (ZFP Lysis buffer supplemented with 20 mM Imidazole) for 1h at 4°C on a spinning wheel. Incubated lysate was spun 2 min at 1000x g to collect the agarose beads. Supernatant was kept and labelled Flow-through, for analysis. Agarose beads are washed twice in three times their volume of wash buffer and washed fractions were kept too. The recombinant protein was then eluted with elution buffer (Lysis buffer supplemented with 250 mM Imidazole) and kept at 4°C until further use.

2.2.1.3 SDS-PAGE and western blotting.

- SDS-PAGE.

Fractions from the purification were analysed by SDS-PAGE using Tris Glycine pre-cast polyacrylamide gels (Biorad, UK) in 1x Tris/Glycine/SDS buffer (Biorad, UK) (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3). Samples were mixed with Laemmli sample buffer (Biorad, UK) and 2,5% β-mercaptoethanol

and incubated at 95°C for 5 min and run on a Mini-Protean electrophoresis cell (Biorad, UK). Precision plus marker was run alongside the samples (Biorad, UK). Gels were stained after the run with Bio-Safe Coomassie Stain (Biorad, UK).

- Western blotting.

Gels were transferred to nitrocellulose membranes (Amersham), with a semi-dry transfer cell (Transblot SD Biorad) in Tris-Glycine-Methanol buffer (48 mM Tris, 39 mM glycine, 20% methanol), pH 9.2 for 60 min at 10V.

Membranes were subsequently blocked in PBS+ 0.1% Tween 20 + 5% Non-fat dried milk or 5% Bovine serum albumin (BSA) (Sigma, UK) for 60 min at room temperature. Horseradish peroxidase (HRP)-conjugated Anti-6X His tag (Abcam, UK) diluted to 1/7500 was then added and incubated for 60 min at room temperature. Membranes were then washed three times in PBS + 0.1% Tween 20.

Membranes were then incubated with HRP chemiluminescent substrate (Pierce, UK), and signal acquired using a Li-Cor Odyssey imaging system (LICORbio, UK).

2.2.1.4 ZFP2143K Quantification.

Purified recombinant ZFP2143K was quantified using a bicinchoninic acid (BCA) Protein Assay Kit (Pierce) following the manufacturer's instructions.

2.2.2 Zinc Finger functionality testing.

2.2.2.1 Oligonucleotides and primers.

Two types of DNA material were used as a substrate for the recombinant DNA binding protein ZFP2143K: genomic fragments of CSSV NJ and oligonucleotide probes.

The double-stranded sequence recognised by ZFP2143K computed by the zinc finger tools website was modified by adding respectively 3 and 4 bp to each side of the 18 bp target sequence, to facilitate the binding of the recombinant, and the resulting oligonucleotide was named 2143ROI. The complementary, antisense version of the probe was also synthesised and named 2143ROIC. A

biotinylated version of theses probes was als-o used for detection of fragments using streptavidin-HRP.

Fragment S1, a genomic portion of CSSV containing the target 18bp sequence targeted by ZFP2143K was amplified with the primer pair S1F and S1R. Another fragment S6 was amplified with the primer pair S6F and S6R, did not contain the target sequence and was used as a control. The sequences of the primers are listed in Table 2.1. The oligonucleotides used in this section are derived from the sequence of CSSV New Juaben (CSSV NJ) that was sequenced for this work (Chapter 3). HPLC or HPSF-purified DNA oligonucleotides, including biotinylated probes, listed in Table 2.1, were obtained from Eurofins (Germany).

Table 2.1: Oligomeric probes and primers used in this chapter.

Primer name	Sequence 5′-3′	Size (bp)
2143ROI	ATAAGAGAACACACTGTAGCAGTCT	25
2143ROIC	AGACTGCTACAGTGTGTTCTCTTAT	25
S1Forward	GCATTTCCATCAGATTCCATGA	22
S1Reverse	TTTAAGAATCTTGGCACGTGTG	22
S6Forward	AATCAGGCACTGTTATTGATCC	22
S6Reverse	AATTTTCCACTAGCGTAGGCAC	22

2.2.3 ZFP2143K DNA substrate synthesis and purification

Polymerase chain reactions (PCR) to produce CSSV genome fragments used for analysis were set-up as follows: 5 µL of CSSV NJ-infected *T. cacao* DNA extract was used as a template and combined with 5 µL Ampligold 10x Buffer (Applied Biosystems), 1 µL 10 mM dNTP solution, 3 µL $MgCl_2$, 0.25 µL AmpliTaq polymerase (Applied Biosystems), 10 µL 0.2 µM forward primer, 10 µL 0.2 µM reverse primer and DNAse-free water to 50 µL.

PCRs were set up in a Rotor-Gene Q (Qiagen) using the following programme: 95°C for 10 s, then 35 cycles: 30 s at 94°C, 30 s at 58°C, and 60 s at 72 °C, with a final 300 s elongation step at 72°C.

Subsequently, fragments were run on a 0.8% agarose gel, excised using a sterile scalpel blade and purified using the GeneJET PCR purification kit (Thermofisher). Eluates from the kit were pooled and quantified using a NanoDrop 1000 UV-Vis spectrophotometer (Thermofisher).

2.2.4 Agarose gel Electrophoretic Mobility Shift Assay.

The Electrophoretic Mobility Shift Assay (EMSA) is a well-established technique (Fried and Crothers, 1981) to determine the degree of protein/nucleic acid interactions (Altschuler, Lewis and Wuttke, 2013) , and can be used to evaluate the functionality of DNA-binding protein. DNA sample and the studied protein were mixed in physicochemical conditions allowing the binding of one to another. Here a PCR-amplified fragment of CSSV containing the cognate substrate to the zinc finger protein was incubated with the recombinant in a formulated buffer. DNA/Protein binding reactions were set to 20 µL by mixing reagents as follows: 2 µL 10X Binding Buffer (100 mM Tris, 500 mM KCl, 10 mM DTT, pH 7.5), DNA (1 to 5 µL), ZFP2143K (0 to 5 µL) and DNAse-free water to 20 µL. When required, binding reactions were supplemented with extra DNA molecules: 1 µL Salmon Sperm DNA solution at 50 ng/µL (ThermoFischer). Binding reactions were incubated at 25°C for 30 min at room temperature to allow interaction of the protein with the DNA and then stopped using 5µL of 5X loading buffer (EMSA LightShift kit, ThermoFisher). Subsequently, the binding reactions were deposited in the wells of a 0.8 to 1.2% TAE (Tris-Acetate-EDTA) agarose gels and run at 120 V for 45-60 min. A detection system labelling either the protein or the nucleic acid part of the reactants then enables the electrophoretic pattern of the samples to be determined by being run on a gel.

The gels were then post-stained with Sybr Safe (ThermoFisher, S33102) using the manufacturer's instructions that targets the DNA, enabling the visualisation of the binding pattern of the recombinant. Gels were then imaged using a camera coupled to a UV transilluminator.

2.2.4.1 EMSA Signal acquisition and measurement.

Fluorescence signal from the PCR fragments was measured using the Fiji imaging software (Schindelin *et al.*, 2012b) as follows: pictures were opened in Fiji and a rectangular section the size of the

fluorescent dye-stained DNA, defining a region of interest (ROI) was created and saved. The function "Analyze>Gels>Select first lane" was used to label the signal in the first lane of the gel. The process was iterated with all lanes of the gel analysed using "Analyze>Gels>Select next lane". The "Plot Gels" function was used to render the signal corresponding to the samples measured. A line was manually drawn on the top of each bell-shaped curve produced by the software using the straight-line pointer. The magic wand pointer was then selected to measure the area defined by the straight line and the curve. This was repeated for all samples and the corresponding pixel density measurements were copied for analysis in an excel spreadsheet.

2.2.5 Double filter binding assay.

The Double Filter Binding Assay (DFBA) summarised in Figure 2.2, was inspired from a method developed by Wong and Lohman (Wong and Lohman, 1993). In this assay, binding reactions containing the DNA binding protein and biotin-labelled DNA oligonucleotides previously diluted to 2 µM in AB1X and heated to 95°C for 5 min and left to anneal at RT until cooled down targets were mixed, and incubated as follows: 5 µL of (biotinylated oligonucleotide, 5 µL of ZFP2143K, 2 µL 10X Binding Buffer (see EMSA 2-3), 0.1 µL Salmon Sperm DNA (50 ng/µL final) and DNase-free water to 20 µL final reaction volume. Binding reactions were incubated for 30 min at 25°C. In the meanwhile, a vacuum manifold dot blotting apparatus (Bio-Rad, Bio-Dot SF Microfiltration Apparatus 1703938) was assembled with three filter papers at the bottom.

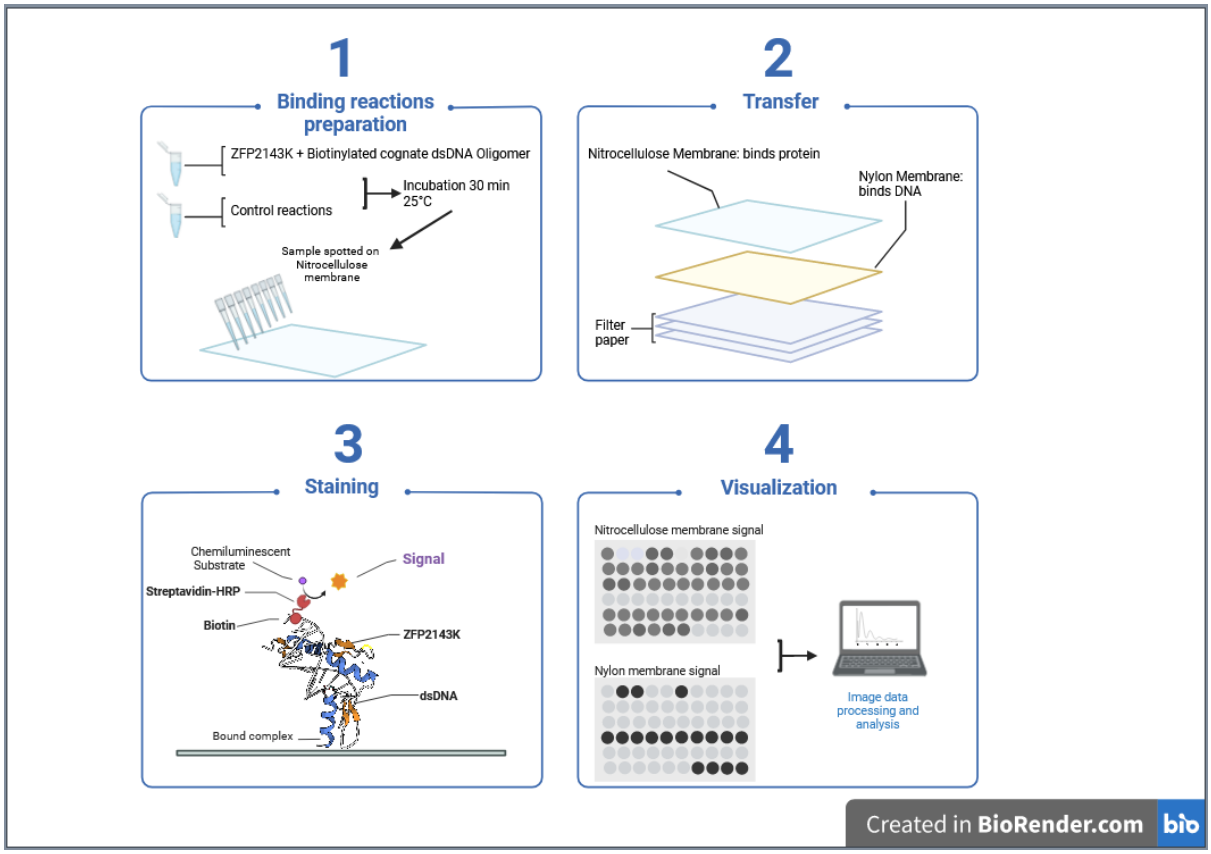


Figure 2.2: Principles of the Double Filter Binding Assay (DFBA).

A stack of membrane and filter paper are assembled and placed in a vacuum-manifold dot blotter. Binding reactions (ZFP2143K+DNA) are loaded on the top of the stack and the differential capture of either DNA/protein fractions, or the DNA/protein complexes is observed.

Reactions were then loaded onto a stack of membranes. The top layer was a nitrocellulose (NC) membrane (ThermoFischer, 77010) pre-incubated overnight at RT in Annealing Buffer 1X with 50 µM $ZnCl_2$ at RT that has the property to bind proteins but not double stranded DNA. The bottom layer that is placed just below the NC membrane was a nylon (NY) membrane which can bind nucleic acid. Nylon membranes (Biodyne™, Thermo Scientific 770160) were previously incubated in 0.4 M KOH for 10 min at room temperature and then rinsed in distilled water until the pH went neutral again, then incubated overnight in Annealing Buffer 1X (AB1x: 10 mM Tris pH 8, 50 mM NaCl, 1 mM EDTA) supplemented with 50 µM Zinc chloride $ZnCl_2$ at room temperature (RT). Both membranes were then placed on wet Whatman filter paper and placed in a dot blotter to apply the binding reactions. Wells were first rinsed with 100 µL AB1X then the samples were applied, and vacuum filtered. Wells were then washed twice with 100 µL AB1X and the membranes were collected. In this assay, the DNA targeted by the ZFP are

biotinylated oligonucleotides, to allow their detection. Streptavidin coupled to HRP is used to detect the DNA: the signal generated in this assay therefore derives from the biotinylated double-stranded oligonucleotides and is used to trace where the binding reactions are located (top or bottom membranes). Both membranes were blocked in PBS+0.1% Tween 20 supplemented with 2,5% Non-Fat dried milk for overnight. Membranes were washed twice in PBS+0.1% Tween 20 for 5 min and a conjugated streptavidin-horseradish peroxidase (Invitrogen) 1/500 dilution was applied on the membranes in PBS+0.1% Tween 20. Membranes were washed twice in PBS+0.1% Tween 20, chemiluminescent substrate applied (Pierce) and imaged using a Li-Cor Odyssey.

A gradient of ZFP2143K was combined with fixed amount of labelled, annealed biotinylated oligonucleotide (2143ROI + 2143ROIC = the DNA sequence targeted by ZFP2143K). The inverse experiment was also conducted, where fixed amount of protein was mixed with a range of ZFP2143K. Controls were set-up as follows: No protein control, No DNA control, Competitor unlabelled DNA: equimolar amount of non-biotinylated oligonucleotide mixed with labelled oligonucleotide and ZFP.

It was anticipated that the double stranded oligonucleotides cannot bind the NC membrane. Therefore, any signal on the NC membrane means that the immobilised protein was able to form complexes with the DNA. Furthermore, if the protein does not bind to the DNA, the oligonucleotides would go through the NC membrane and bind to the NY membrane located underneath.

2.2.6 ZFP2143K-CSSV high resolution melting analysis.

A ZFP/CSSV high resolution melting assay (HRM) was designed to address whether the interaction between the DNA binding protein and its substrate could be monitored using HRM analysis. DNA fragments were mixed with the fluorescent dye SYTO 13 (Invitrogen) and then combined with increasing amount of ZFP2143K.

Reaction containing 2 µM solution of the 2143ROI/AS oligomers were annealed in AB2X (Annealing Buffer 2X: 20 mM Tris, 100 mM NaCl, 2 mM EDTA, pH 8) or Binding Buffer 1x (10 mM Tris, 50 mM KCl,

1 mM DTT; pH 7.5) in the Rotor-Gene Q thermocycler using the following programme: 95°C for 300s then ramp down from 95°C to 30°C in 45 min. Annealed samples were then kept on ice until processed.

Binding reactions were prepared by mixing 5 µL annealed 2143ROI/AS to 5 µL ZFP2143K (at 866 µg/mL) in 10 µL AB2X or BB1X. Reactions were incubated for 30 min at 30°C in a dry block. 5 µL of SYTO13 dye diluted to 50 µM were then added and the samples placed in the Rotor-Gene Q for HRM analysis.

Samples were run on the following program: HRM cycle, 30 to 90°C, rising by 0.1 degree each step, wait for 90 s of pre-melt conditioning on first step, wait for 2 s for each step afterwards. Gain was optimised before melt on all tubes, with the gain giving the highest fluorescence less than 95 being selected.

Data was acquired and analysed using the Rotor-Gene Q manufacturer' software: Rotor-Gene Q Pure detection version 2.3.5. Fluorescence levels were compared between samples with or without ZFP during the period preceding the melt and the temperature of melt was also recorded. Graphs were made using Minitab (Minitab LLC, 2022), for the normalised fluorescence melting curve profiles and MS Excel version 2304 for the fluorescence plots.

2.2.7 Pull down assay of CSSV using Histidine-modified magnetic particles.

Superparamagnetic particles modified with nickel were used to target the hexahistidine tag that was inserted at the C-terminal end of the protein. This tag is normally used for the immobilized metal affinity chromatography purification step in the production of the recombinant. The chromatography takes advantage of the affinity of the protonated nickel bound on the resin for the nitrogen of the imidazole ring present on the histidine residues of the recombinant protein. The nickel ions have six ligand binding sites in their coordination sphere, with four used to bind the nitriloacetic acid, and two available to interact with two histidine residues (Crowe *et al.*, 1994). The bond is disrupted traditionally by adding buffers containing imidazole at concentrations that competes with the imidazole ring from the tag, creating a displacement of the affinity from the resin to the buffer,

releasing the recombinant from the purification matrix into the media. In this experiment, the particles are repurposed with the aim of capturing viral DNA-bound ZFP2143K from binding reactions set-up with infected plant extracts.

HisPur™ Ni-NTA Magnetic Beads superparamagnetic beads (ThermoFisher) were placed in four 200 µL volumes of equilibration buffer (PBS1x, 0.05% Tween 20, 30 mM Imidazole, pH 8), vortexed 10 s, placed on the magnetic stand to collect the beads and the supernatant removed and resuspended in 800 µL of equilibration buffer.

In parallel, DNA was extracted from CSSV NJ-infected *T. cacao* leaves using liquid nitrogen and Plant DNeasy kit (Qiagen).

Calibrated leaf sections were cut using a coring tool producing 7 mm diameter discs. Resulting samples were weighed. 25 mg sections were frozen in liquid nitrogen and broken-down using tweezers. Whole DNA was then extracted using the guanidine salts method with the DNeasy plant mini kit (Qiagen, 69104), eluted in TE buffer, and quantified with a Nanodrop UV Vis spectrophotometer (Thermo). 45 µL of the DNA solution was combined with 10 µL of a range of ZFP2143K and 345 µL equilibration buffer and incubated 30 min on a rotating shaker at RT. DNA/protein binding reactions were added to 50 µL of the equilibrated beads and incubated for 30 min on a rotating shaker at RT. Beads were then either washed twice in 500 µL wash buffer (Phosphate Buffer Saline1x (PBS), 0.05% Tween 20, 50 mM Imidazole, pH 8) or twice in 2x Sodium Citrate Saline(SSC)+Sodium dodecyl sulphate (SDS) 0.1%. Beads were placed on the magnetic rack; wash buffer was decanted and 50 µL TE buffer was added to the preparations. Samples were eluted by heat treatment for 5 min at 95°C.

2.2.7.1 qPCR on DNA eluted from the magnetic particles.

The efficacy of the recombinant to extract viral DNA from plant samples processed with the magnetic beads was evaluated by quantitative PCR (qPCR) using primers targeting the reverse transcriptase (RT) of CSSV. A nuclear genome locus positioned next to the simple sequence repeat (SSR) mTcCIR25 from *T. cacao* was assessed via qPCR alongside all samples, to quantify the quantity of plant DNA co-purified

with the viral sequences. mTcCIR25 is a single nuclear locus enabling quantification of plant genome number. In parallel, a set of synthetic RTase target consisting of a range of increasing copy number (500, 5000, 50000 and 500000) and amplified using the RTase assay was used to determine the copy number for both viral and nuclear genomes.

Table 2.2: Primers, probes, and reference CSSV fragment sequences used in the PCR quantification of the particles extracted DNA.

Description	Forward primer	Reverse primer	Probe
CSSV Reverse Transcriptase	F-74, CTGAAGCGAGTAGGCAACAA	R-151, CAGTCCAAGGGATGGACTCT	CSSV RTase probe [FAM]TCCATCAGGTTGCCATGGCA[BHQ1]
Theobroma cacao genome Single locus SSR (Y16997)	MTcCIR25F CAGATAAGGAAAGGTGGAGTTTGG	MTcCIR25R CAAGAATGTCTCCTACATTCACTACG	MTcCIR25P [FAM]TTCCCGTAAGCTTCGTCCCAGATGC[BHQ1]
Description	**Sequence**		
Rtase synthetic fragment standard	CTGAAGCGAGTAGGCAACAAAAAAATTTTCAGCAAGTTTGATTTAAAATCAGGCTTCCATCAGGTTGCCATGGCAGAAGAGTCCATCCCTTGGACTG		

TaqMan qPCR was then performed on 5 µl of DNA samples using both a CSSV primer/probe assay and a plant primer/probe mix (Table 2.2) in Sensifast No RoxMaster Mix (Bioline). The qPCR reaction mix primer/probe mix was prepared by combining probe (2.5 µl of 100 pmol/ µl), forward primer (10 µl of 100 pmol/ µl), reverse primer (10 µl of 100 pmol/ µl) and 77.5 µl of 10 mM Tris-Cl. 0.5 mM EDTA; pH 9.0. Each 20 µl reaction consisted of 10 µl Sensifast No RoxMaster Mix (Bioline), 0.8 µl primer/probe mix, 4.2 µl nuclease free water and 5 µl DNA.

2.3 Results.

2.3.1 CSSV New Juaben site selection.

A search on the Zinc Finger tools website using the contig sequence from the *de novo* 7024 bp sequencing of CSSV New Juaben provided 376 potential ZFP binding sites. Amongst these sites, 10.9% (41) were given a score above 60, which is considered a high score in the website's ranking system, since the scoring ranges from 1 (worst) to 64 (best). A candidate target site at position 2143 of the CSSV contig sequence with a high score was selected.

The specificity of the ZFP target sequence was assessed by performing several blast searches with various election criteria, using the nucleotide collection database (nt= GenBank+EMBL+DDBJ+PDB+RefSeq sequences, excluding EST, STS, GSS, WGS, TSA, patent sequences as well as phase 0, 1, and 2 HTGS sequences and sequences longer than 100 Mb). Blast search was conducted using all sequences, CSSV sequences only or *T. cacao* sequences only. Over the hundred results of a blast search conducted with the target sequence 2143 in the nucleotide collection (nt), 85 were from a range of isolates of CSSV Togo B (CSSTBV) with 100% coverage and identity, and no gaps. The fifteen other hits were unrelated organisms such as the fish *Gobius niger*, the cactus mouse *Peromyscus eremicus* or the moth *Cosmia pyralina*. CSSTBV is the name of the species that the New Juaben strain belongs to (Gyamera, Domfeh and Ameyaw, 2023). When repeating the blast with the taxon restricted to CSSV, a hundred accessions of multiple CSSV isolates was returned. A search restricted to *Theobroma cacao*'s genome returned no hits for the whole 18pb sequence, however, multiple sequences of up to 15 bp were found on several sites on eight of the chromosomes.

The candidate target site 2143 sequence is 5'-GAG AAC ACA CTG TAG CAG-3'. The corresponding zinc finger protein sequence, ZFP2143K computed by the website is:

LEPGEKPYKCPECGKSFSRADNLTEHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRND ALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGEKPYKCPEC GKSFSRSDNLVRHQRTHTGKKTS that uses the following architecture: Fixed N-terminal – (N-terminal backbone – Recognition helix – C-terminal backbone – ZF linker)$_n$ – Fixed C-terminal, where Fixed N and C-terminals, located at both ends of the protein are derived from the sequence of SP1 human transcription factor, the ZF linker is the portion of protein between the units and is a consensus sequence obtained from the analysis of a transcription factors database and n is the number of units within the protein. The N-terminal sequence includes the β-sheets while the C-terminal portion of the α-helix is in the C-terminal backbone part. Figure 2.3 highlights the sequence and the architecture of

the candidate protein ZFP2143K protein and the corresponding nucleotide triplets targeted by each of the six domains.

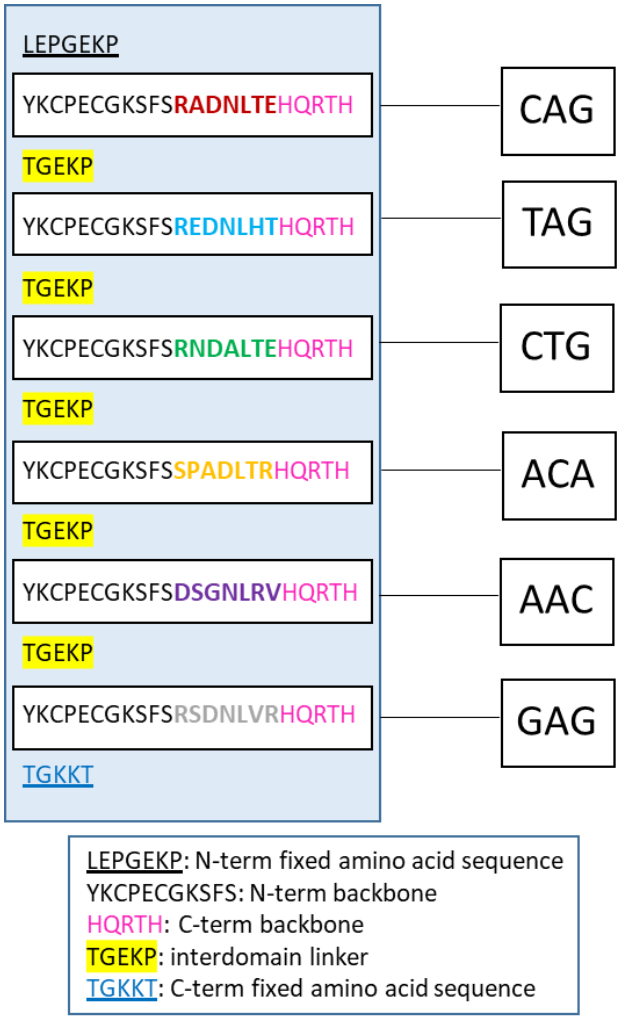


Figure 2.3: Architecture of the ZFP2143K protein (Left).

ZFP2143K is composed of six domains, each recognising a triplet of nucleic acid. Sections of the domains interacting with the targeted codons are depicted in red, blue, green, gold, purple and grey.

Bioinformatic analysis of the predicted protein derived from the sequenced plasmids using the ProtParam algorithm shows that ZFP2143K is a non-polar, 184 amino-acids long protein with a molecular weight of 21.1 kDaltons and has an isoelectric point of 9.05.

2.3.2 Expression and purification of ZFP2143K.

Expression of ZFP2143K was analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). An accumulation of recombinant protein was observed on all eluted fractions (Lanes E1

to E2, Figure 2.4), at 21.1 kDa, in agreement with the predicted molecular weight computed by bioinformatic analysis. The elution conditions did not allow to dislodge the totality of the produced recombinant as a significant band is still visible on the resin fraction. Furthermore, a band of similar size is visible in the insoluble fraction, suggesting that part of the expressed protein precipitated as inclusion bodies in the cytoplasm of the bacteria.

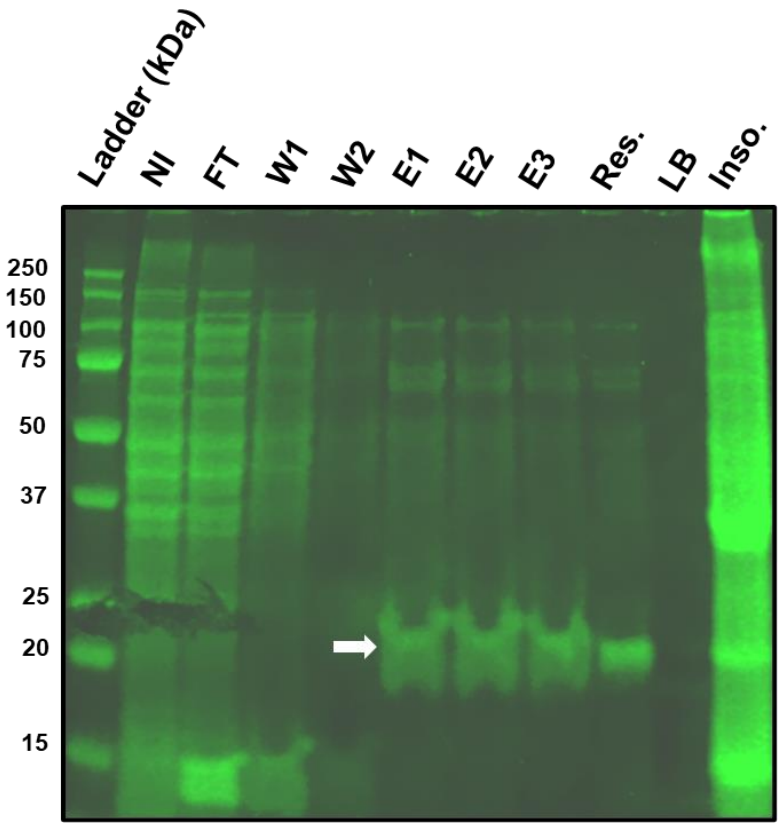


Figure 2.4: SDS-PAGE Analysis of ZFP2143K Expression and Purification.

Ladder(kDa): Protein molecular weight marker (values on the left), NI: Non-induced culture whole cell lysate, FT: Flow-through, Cytosolic fraction unbound to the column, W1: First column wash, W2: Second column wash, E1: First eluate, E2: Second eluate, E3: Third eluate, Res.: Resin post wash and elution, LB: Loading buffer only, Inso.: Cytosolic insoluble fraction. The white arrow indicates the position of the ZFP2143K recombinant protein, present in the eluted fraction and on the agarose resin.

Western blot using an antibody directed against the hexahistidine tag used for purification evidenced the specificity of the purification process. The signal observed on the membrane indicates that the accumulated protein seen on the SDS-PAGE is the recombinant protein bearing the hexahistidine tag (Lanes E1 to E3, Res. Inso. Figure 2.5). The eluted fractions one to three produced an intense signal,

and so did both the resin and the insoluble fraction, confirming the observations from the SDS-PAGE. Interestingly, it was noticed that the non-induced lane also produced a signal.

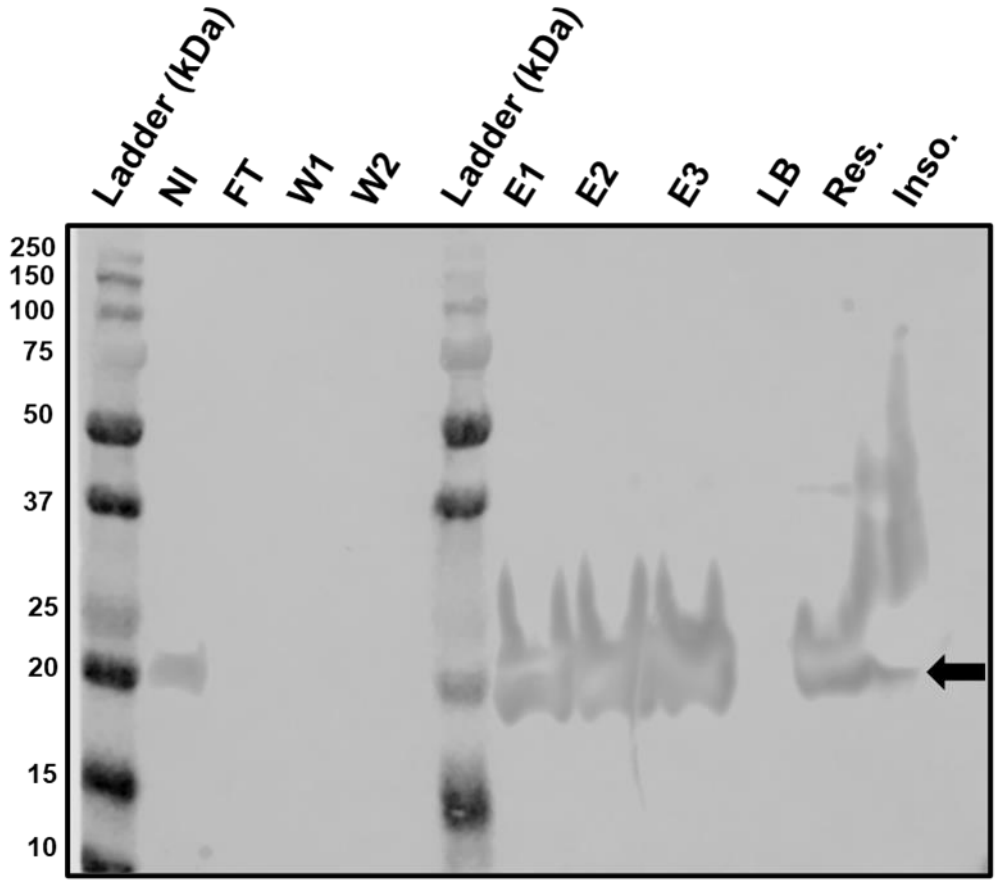


Figure 2.5: Western Blot of ZFP2143K Expression and Purification.

Ladder(kDa): Protein molecular weight marker (values on the left), NI: Non-induced culture lysate, FT: Flow-through, Cellular lysate unbound to the column, W1:First column wash, W2:Second column wash, Ladder(kDa): Protein molecular weight marker E1:First eluate,E2:Second eluate,E3:Third eluate, LB: Loading buffer only, Res.: Resin post wash and elution, Inso.: Cytosolic insoluble fraction. The black arrow indicates the position of the ZFP2143K recombinant protein, present in the eluted fraction, the agarose resin and in the insoluble fraction. The protein was also expressed in the non-induced culture (NI lane).

2.3.3 Assessment of ZFP2143K functionality using the double filter binding assay (DFBA).

The ability of the candidate ZFP2143K to bind DNA was assessed using the double filter binding assay (Figure 2.6). Two quantities of biotinylated probe were tested against a range of ZFP2143K concentrations (28 nM to 1400 nM), to first evaluate the system parameters.

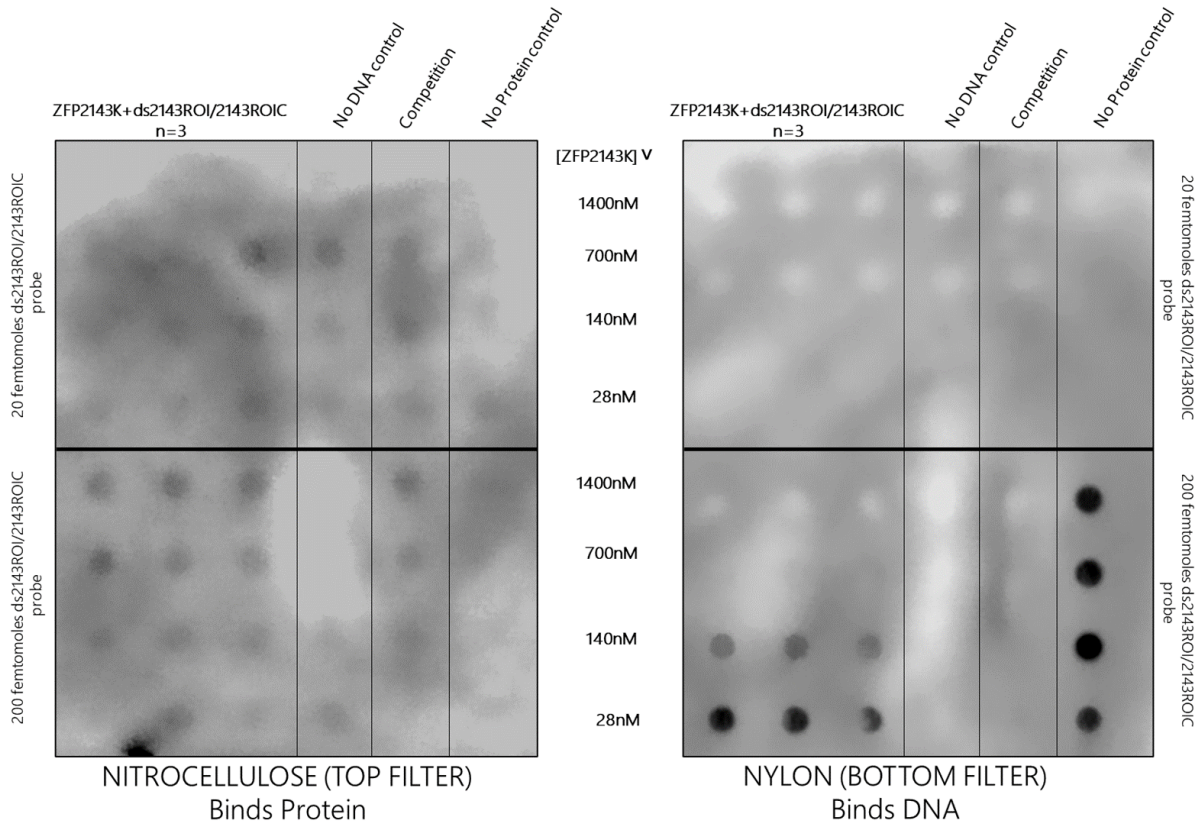


Figure 2.6: Double filter binding assay with 20 or 200 femtomoles of ds 2143ROI/2143ROIC incubated with a range of [ZFP2143K], in triplicates.

Binding reactions containing a range of ZFP2143K and either 20 or 200 femtomoles pre-annealed target dsDNA were loaded on a stack of filters. Control reactions included a no DNA control, a competition binding reaction with equimolar of labelled and unlabelled target sequence and a no protein control. When using 200 femtomoles of ds2143ROI/2143ROIC, it was observed that the BRE candidate ZFP2143K binds in a dose dependent manner, as observed on the nylon filter at the bottom of the stack.

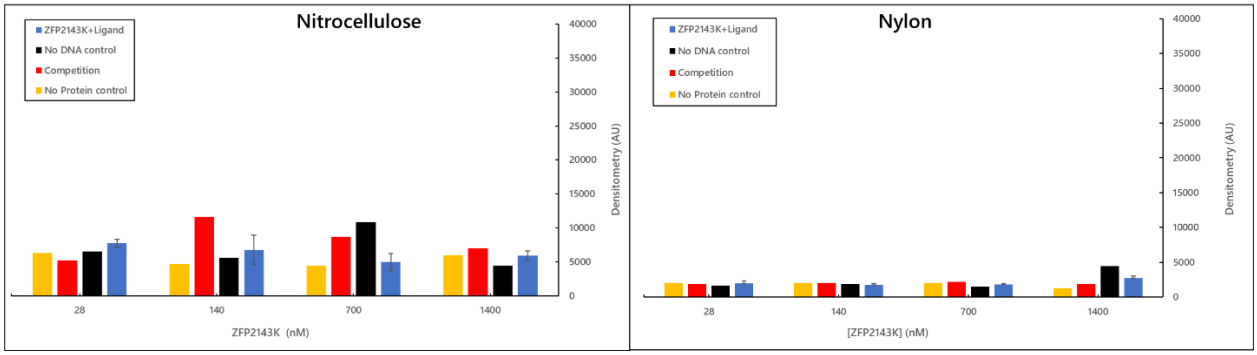


Figure 2.7: Chemiluminescent signal generated with 20 femtomoles of the biotinylated dsDNA cognate substrate incubated with a range of [ZFP2143K] measured by pixel density, derived from the image from Figure 2.5.

Blue bars are the average signal of three replicates. Controls with No DNA or no protein and competition were tested as single replicates for all protein concentration tested. The signal obtained was close to background levels for all sample tested.

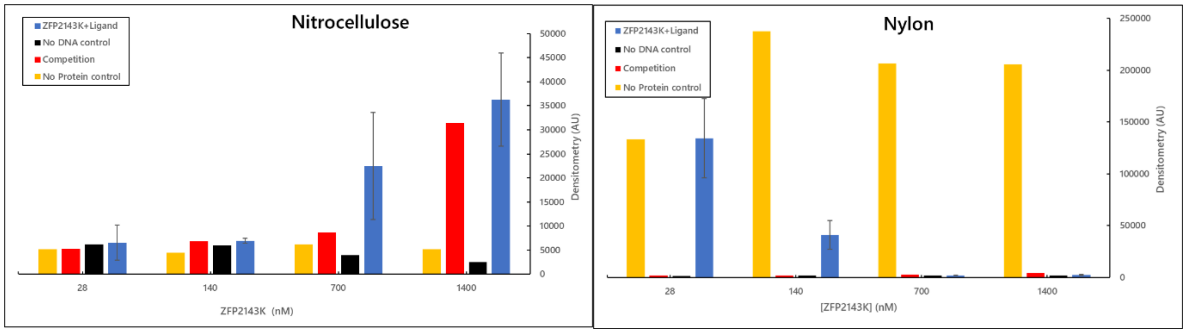


Figure 2.8: Chemiluminescent signal generated with 200 femtomoles of the biotinylated dsDNA cognate substrate incubated with a range of [ZFP2143K] measured by pixel density, derived from the image from Figure 2.5.

Blue bars are the average signal of three replicates. Controls with No DNA or no protein and competition were tested as single replicates for all protein concentration tested. The signal detected with a constant 200 fmol of labelled oligonucleotide on the nylon membrane decreases with the concentration of ZFP2143K (blue bars), suggesting that the recombinant protein bound to the cognate DNA is dose-dependant.

Two quantities of probe, 200 and 20 femtomoles of biotinylated double-stranded 2143ROI/2143ROIC oligomeric probe were tested against a dilution range of the ZFP2143K recombinant. The signal produced was deduced from the pixel density analysis derived from the pictures of the membranes.

200 fmol of biotinylated probe per reaction produced a detectable signal (Figure 2.6 and 2.8) but was close to background signal at 20 femtomoles per reaction (Figures 2.6 and 2.7).

Increasing amount of ZFP2143K in the binding reactions with 200 femtomoles of the double stranded probe generated an increasing signal on the nitrocellulose membrane, while the signal decreased on the nylon membrane, demonstrating that the recombinant protein was able to retain its cognate DNA molecule on the top layer of the membrane stack (Figures 2.6 and 2.8). In contrast, control samples containing no recombinant failed to produce a signal on the nitrocellulose membrane while producing a consistent signal on the nylon side. Samples containing an equimolar amount of biotinylated and non-labelled double-stranded DNA always produced a signal inferior to the samples containing the biotinylated oligomers only. This demonstrates that unlabelled probe competed with the biotinylated probe and reduced the signal as a result. Negative control samples containing no DNA did not produce any signal (Figures 2.6 and 2.8).

2.3.3.1 Titration curves of ZFP2143K.

The experiment was repeated by preparing a set of binding reactions containing a fixed 200 femtomoles of the biotinylated double stranded cognate DNA substrate 2143ROI/2143ROIC combined to a range of concentrations (13.5 nM to 3478.7 nM) of the recombinant 2143K protein (Figure 2.9).

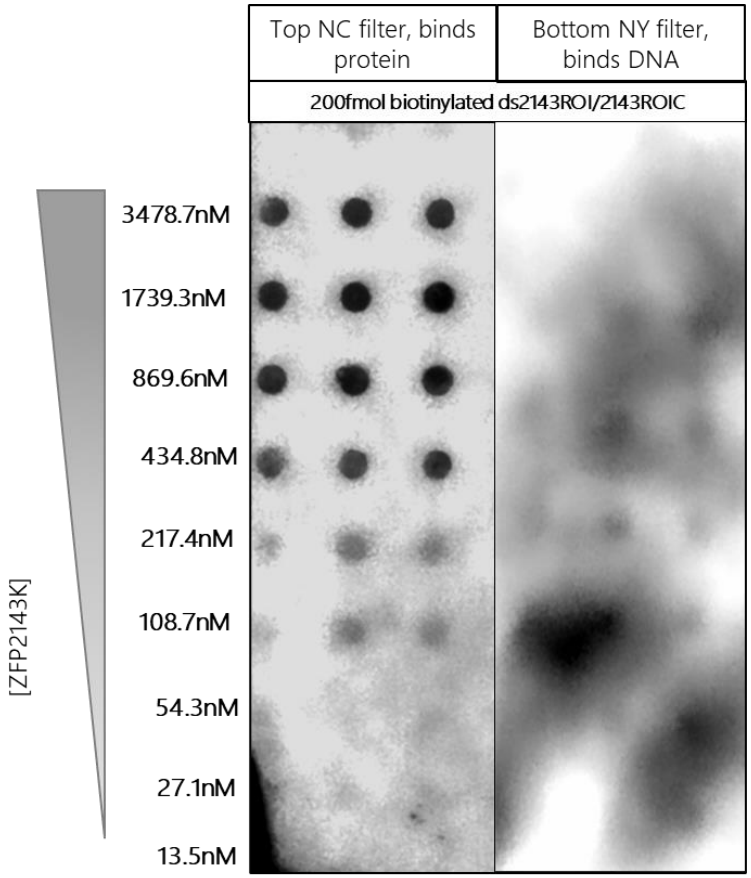


Figure 2.9: DFBA using a constant 200 fmol ds2143ROI/2143ROIC biotinylated probe combined with a range of concentration of the ZFP2143K recombinant protein (13.5 nM to 3478.7 nM).

The signal on the nitrocellulose membrane (left picture) decreased steadily with decreasing amount of protein, suggesting that binding of the labelled dsDNA to the recombinant ZFP2143K was dose-dependant.

When 200 fmol of the probe were mixed with decreasing concentration of ZFP2143K, a signal was detected on the nitrocellulose membrane until the protein concentration decreased to 108.7 nM at which point the signal was undetectable over the background (Figure 2.9) indicating that the limit of sensitivity of the system was reached. This is consistent with the first experiment where 140 nM was the lowest concentration producing a signal detected. The signal measured on the nylon membrane increased when the concentration of protein decreased. Signal close to background levels was

observed for 1739.4, 869.7 and 434.8 nM, then increased four times above background levels at 217.4 nM, before reaching the maximum signal at 54.4 nM. The data obtained from the nylon membrane displayed an inverse correlation with the data obtained from the curve for the nitrocellulose membrane. Furthermore, the curves intersected within the range of 217.4 and 434.8 nM ZFP2143K (Figure 2.10).

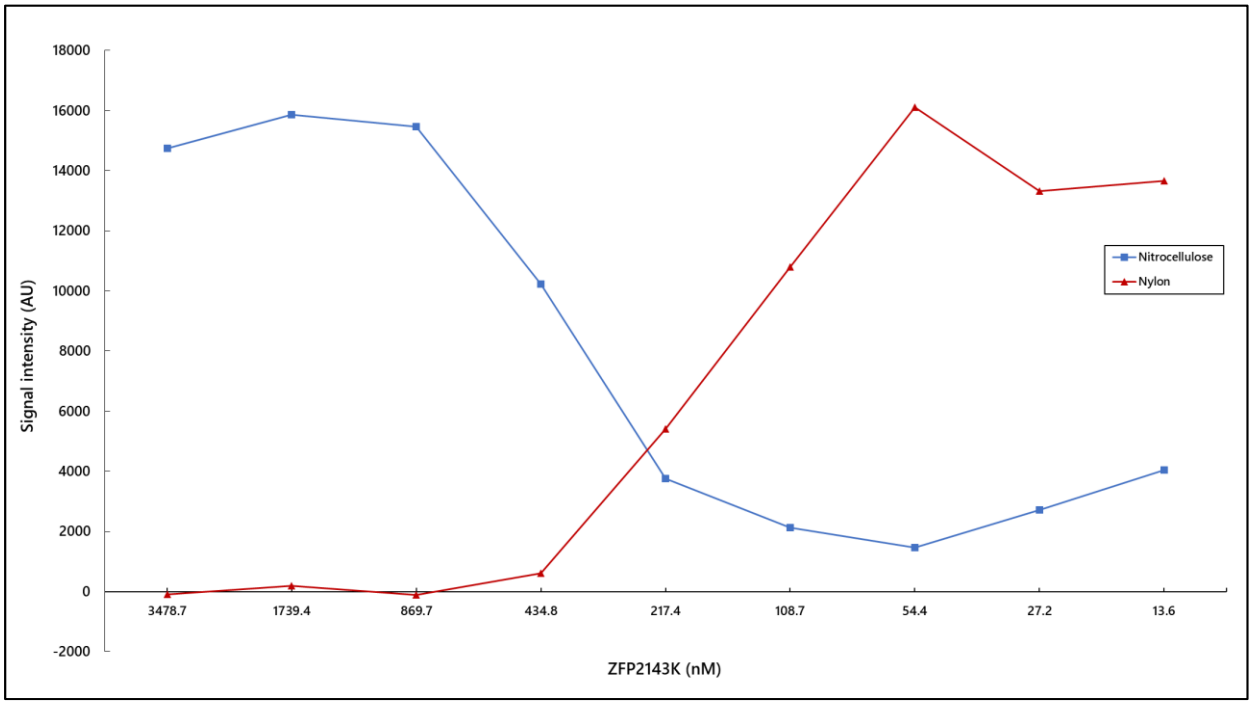


Figure 2.10: Signal on the nitrocellulose and nylon membrane intercept between 217.4 and 434.8 nM ZFP2143K when titrated with 200 femtomoles of the ds2143ROI/AS biotinylated oligomeric probe.

Signal measured on the nitrocellulose membrane was divided by ten to set both curves on the same scale. Outliers datapoint caused by excessive background signal were removed from the analysis.

2.3.4 HRM analysis of the interaction of ZFP2143K with a synthetic double stranded target DNA.

HRM analysis was evaluated for analysing the binding of ZFP2143K to its DNA ligand. The aim of the experiment was to assess if this rapid and accurate technique could be used to study the protein-DNA interactions.

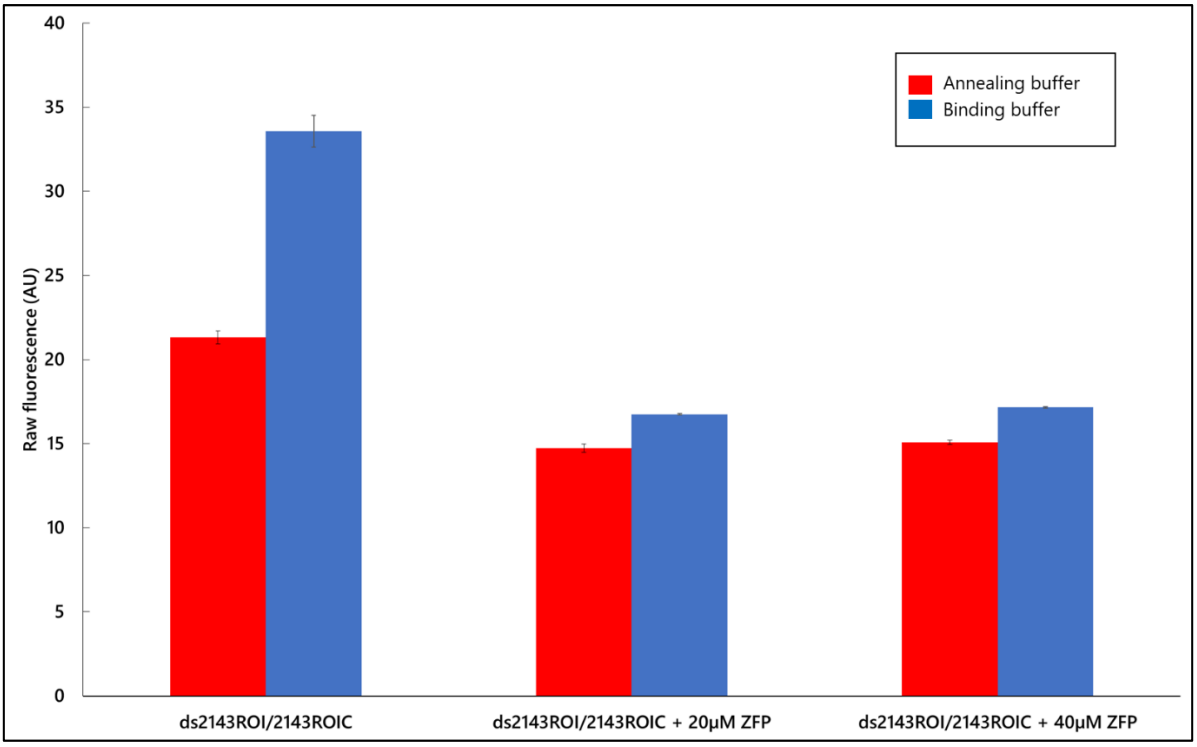


Figure 2.11: Influence of adding ZFP2143K on the raw fluorescence released by the double stranded fragment 2143ROI/2143ROIC in the melting phase from 30 to 37°C.

Data presented is the averaged SYTO13 raw fluorescence measurements of the 303 datapoints recorded by the Rotor-Gene Q of triplicate samples from 30 to 37°C. Error bars correspond to two standard errors.

Binding reactions performed in annealing buffer showed little difference between samples containing ZFP2143K or none (Figure 2.11), regardless of the concentration of protein used. In contrast, the averaged raw fluorescence of samples incubated in binding buffer in absence of ZFP2143K was higher than the samples incubated with the recombinant (Figure 2.11). Doubling the quantity of ZFP2143K from 20 µM to 40 µM had a marginal effect on the fluorescence measured for samples incubated in binding buffer.

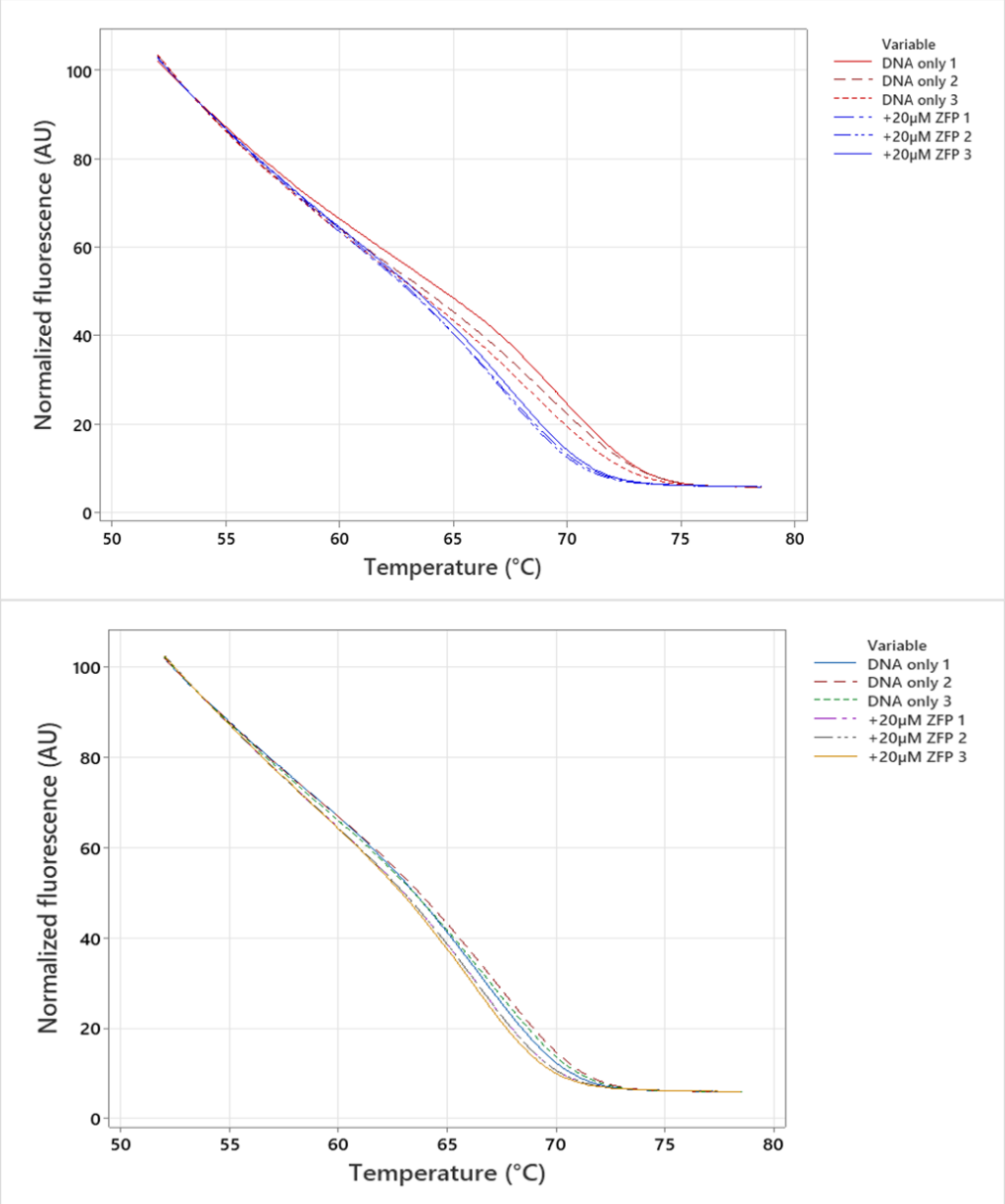


Figure 2.12: Influence of buffer composition on the melting curves of ZFP2143K binding to its cognate DNA sequence.

Top chart: binding reactions were incubated in Binding Buffer; Bottom chart: binding reactions were incubated in Annealing Buffer. Replicates samples are shown.

A difference in the melting profile was observed between the double-stranded 2143ROI/2143ROIC DNA samples containing 20 μM ZFP2143K and the samples containing the DNA alone when the binding reactions were incubated in binding buffer, as compared to annealing buffer (Figure 2.12). The three replicates incubated with the ZFP2143K in binding buffer grouped together and melted at 73°C before the samples without protein, at 75°C. This was not replicated with the samples incubated in annealing buffer, where the melting curves are very close, regardless of the presence of protein in the samples.

2.3.5 EMSA analysis of ZFP2143K binding.

The specificity of the binding of ZFP2143K was evaluated using the electrophoretic mobility shift assay. A 654 bp CSSV amplicon containing the 18 bp target sequence of ZFP2143K, fragment S1, was synthetised by PCR, purified, and incubated with a range of concentrations of the recombinant protein (Figure 2.13).

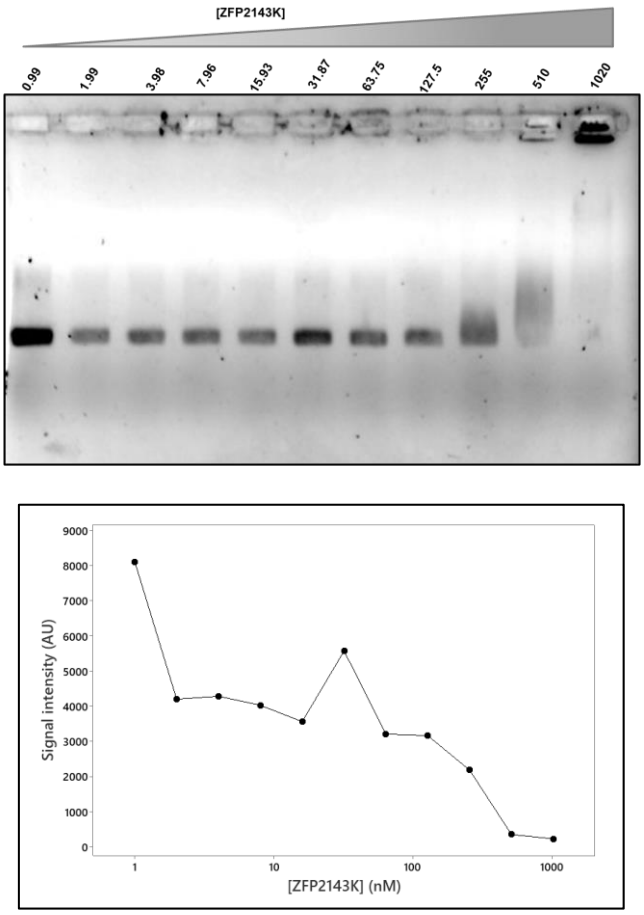


Figure 2.13: EMSA of 200 ng of Fragment S1 with a range of concentrations of ZFP2143K.

The quantity of recombinant in each binding reaction, in nM, is indicated above each corresponding lane. The graph underneath plots the signal intensity derived from the gel picture as a function of the quantity of ZFP2143K in each binding reactions. The abscissa axis was transformed to a logarithmic scale.

The signal intensity decreased with increasing amount of the recombinant and was at its lowest with 1020 nM of ZFP2143K, the highest quantity tested on this experiment. A higher value than expected was measured with 31.87 nM ZFP2143K, that is possibly artefactual. A shift was observed at 510 and 255 nM of ZFP2143K. However, the DNA-protein complex did not migrate with 1020 nM ZFP2143K.

The same procedure was made with another 906 bp CSSV amplicon, fragment S6 that does not contain the target sequence. A shift was observed with 1180 µM ZFP2143K for both fragments (Figure 2.14) since the fragments did not migrate, as observed in the previous experiment.

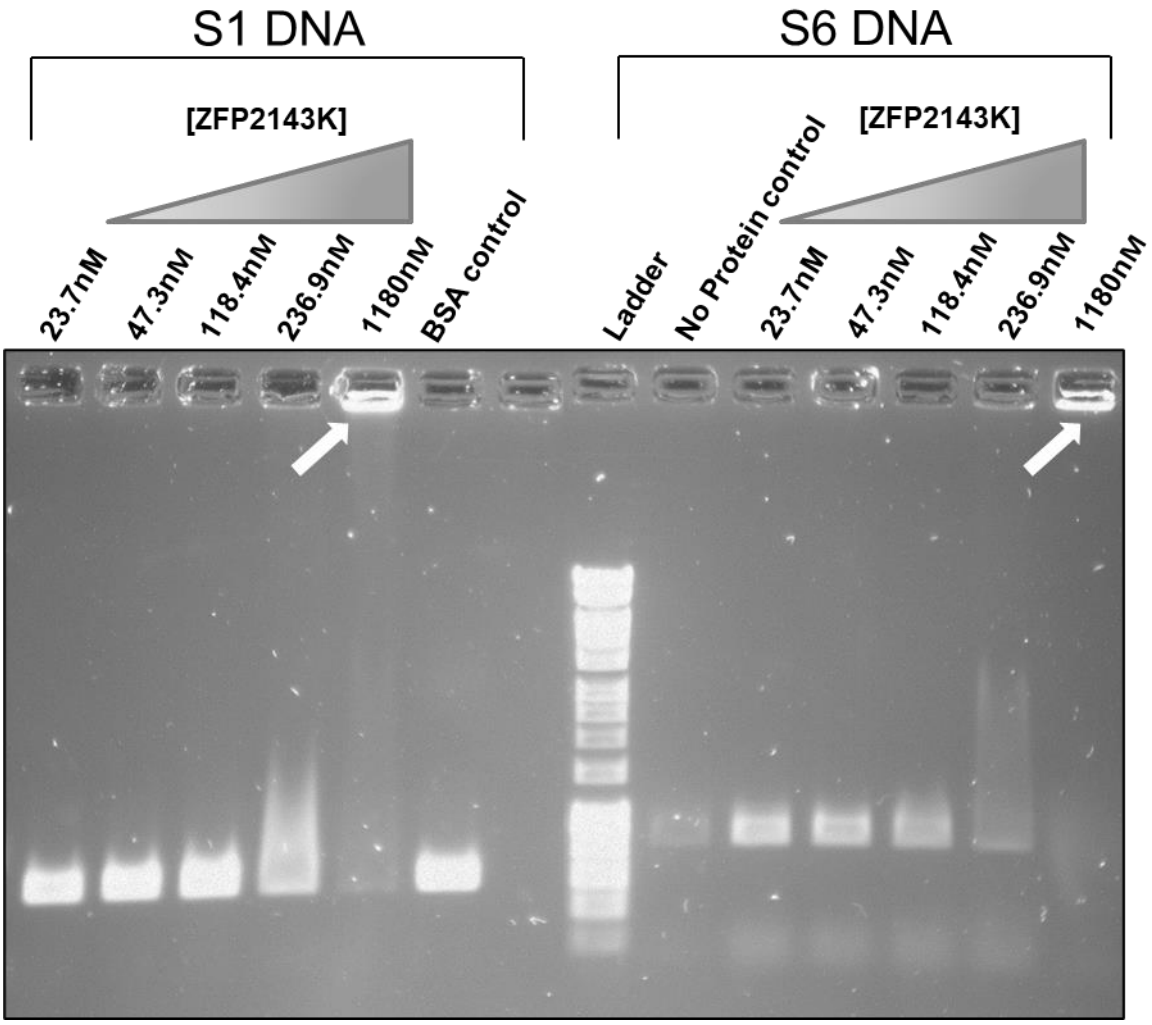


Figure 2.14: ZFP2143K shifts the electrophoretic migration of the Fragment S1 containing its target sequence.

The recombinant protein also shifted the migration of the fragment S6. White arrows are indicating material that did not migrate.

In order to mitigate non-specific interactions, the experiment was repeated with salmon sperm added to both samples. Adding non-specific DNA in the binding reactions has been shown to improve the specificity of the reaction (Ayala *et al.*, 2014). There was no improvement to the specificity of the reaction when competing DNA was added to the binding reaction (Data not shown). Fragment S6 was still able to bind to ZFP2143K in the presence or absence of salmon sperm DNA.

2.3.6 Pull-down assay targeting CSSV DNA sequence from infected plant samples using magnetic particles bound-ZFP2143K.

The capacity of ZFP2143K to act as a BRE for integration to a biosensor was assessed using a pull-down assay. DNA was extracted from CSSV-infected *T. cacao* leaves, column-purified, mixed, and incubated with 0, 2, 20 or 40 µM ZFP2143K and then with nickel-modified magnetic particles targeting the hexahistidine tag on the recombinant. Eluates from the particles were used as a template for a qPCR targeting both a region nearby a SSR nuclear marker for quantifying the plant host- DNA and the RTase from CSSV for quantifying captured viruses.

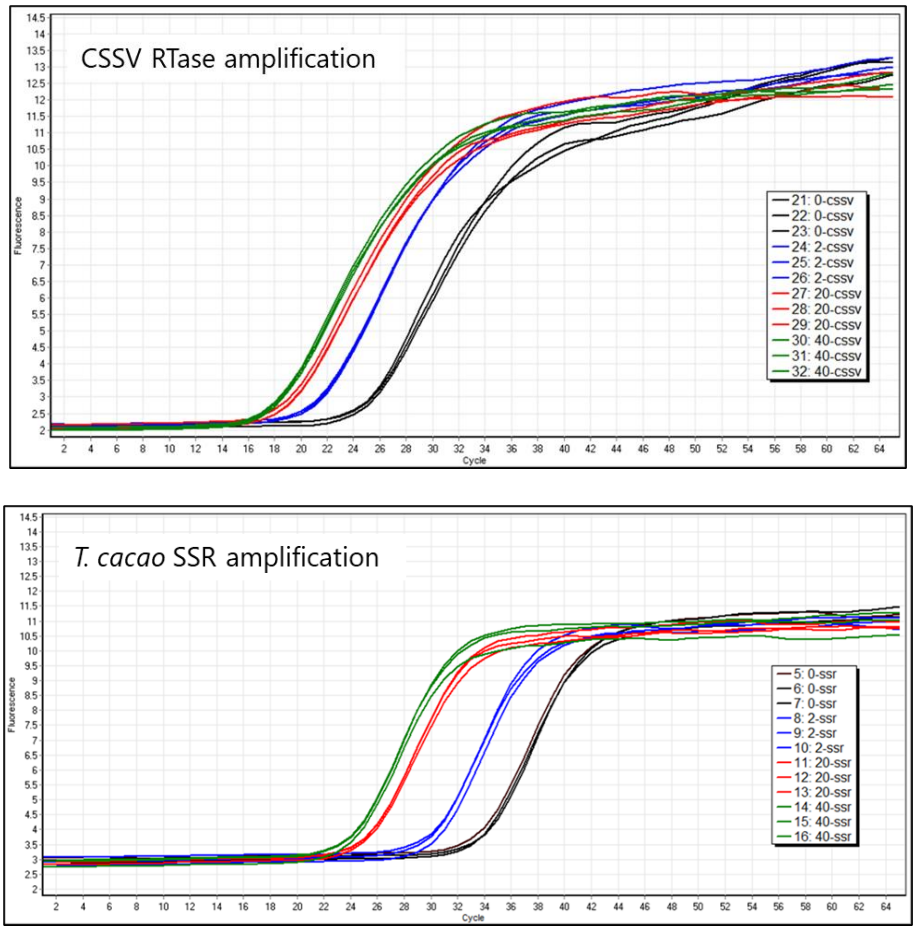


Figure 2.15: qPCR amplification plots targeting the RTase of CSSV (top plot) and an SSR of *T. cacao* (bottom plot) of samples incubated with 0, 2, 20, or 40 µM of the recombinant protein ZFP2143K.

Increasing the concentration of ZFP2143K resulted in an increase quantity of DNA amplified both CSSV and from *T. cacao*. Both graphs are showing triplicated amplified DNA samples eluted from the superparamagnetic nickel-modified beads.

Increasing the concentration of ZFP2143K incubated with the extracted DNA resulted in a gradual rise in the quantity of CSSV DNA captured by the magnetic particles (Figure 2.15, top plot, and Figure 2.16,

red bars). It was observed that the quantity of *T. cacao* co-captured also increased with incremental amount of the recombinant protein (Figure 2.15, bottom plot, and Figure 2.16, blue bars). CSSV was however preferentially bound by the particles regardless of the quantity of ZFP2143K used (Figure 2.16), but the ratio bound CSSV DNA/bound *T. cacao* DNA did not improve with increasing quantity of the recombinant protein (Figure 2.17).

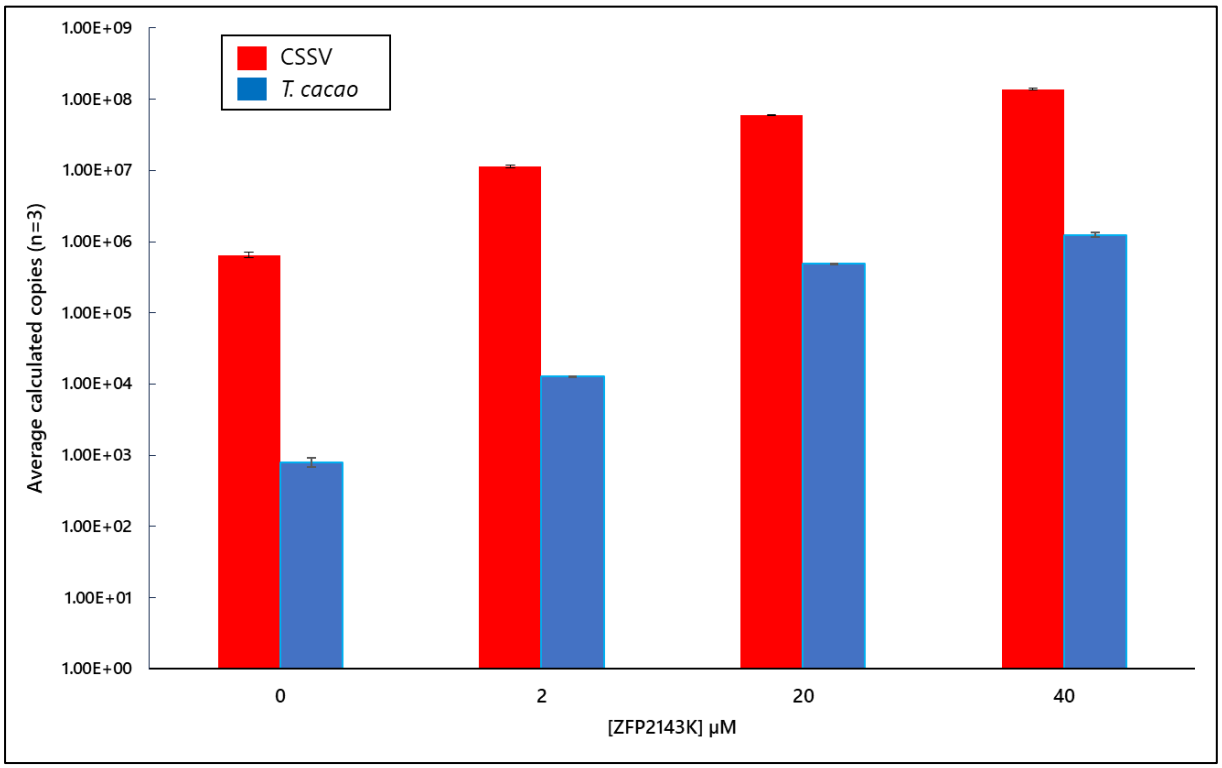


Figure 2.16: ZFP2143K binds to the targeted region on CSSV in a dose-dependent manner.

Mean (n=3) of calculated copy numbers from a qPCR targeting CSSV and *T. cacao*, as a function of the concentration of ZFP2143K used in the binding reactions. The ordinate axis was transformed to a logarithmic scale.

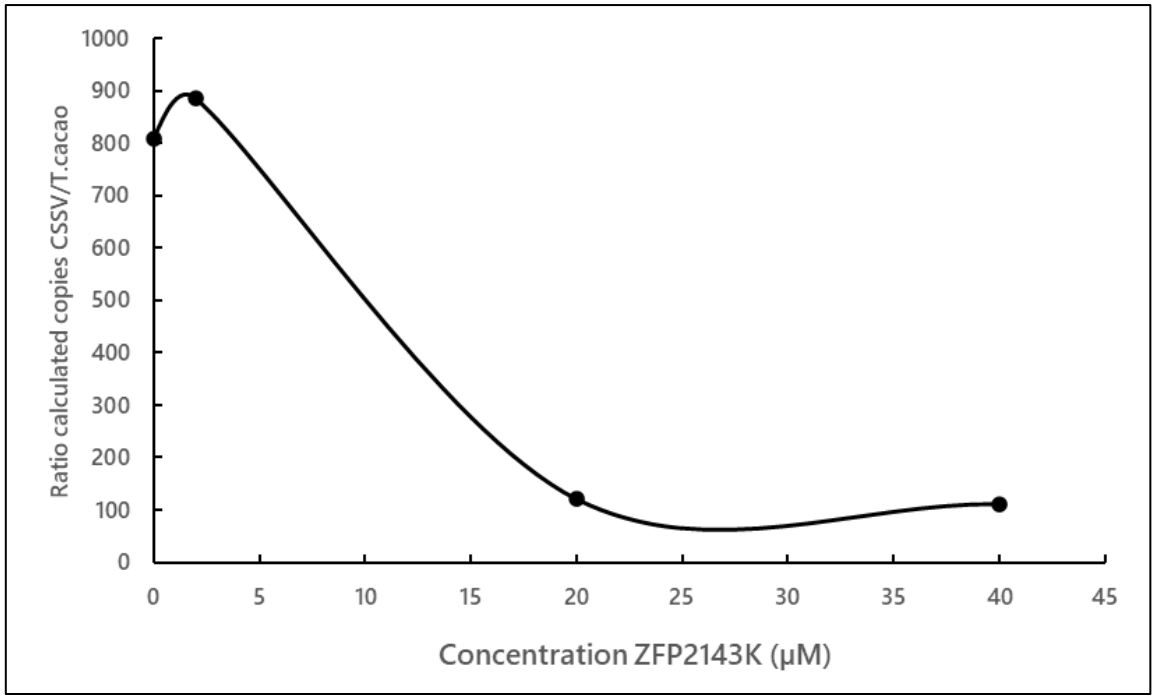


Figure 2.17: Ratio of calculated copies of CSSV/calculated copies *T. cacao* as a function of the concentration of ZFP2143K derived from the data presented in Figure 2.16.

Ratios were calculated by dividing the average of three replicates of CSSV amplification by T. cacao amplifications.

To assess if there is an optimum concentration of ZFP to favour binding to the viral DNA instead of the genomic DNA, a second experiment was performed with a lower ZFP concentration range of 0.02, 0.2 and 2 µM. To reduce non-specific binding to plant genomic DNA (Figure 2.18), the particles were washed off with two round of stringency washes including two low stringency washes (2x SSC, 0.1% SDS) followed by two high stringency washes (2x 0.1 SSC, 0.1% SDS).

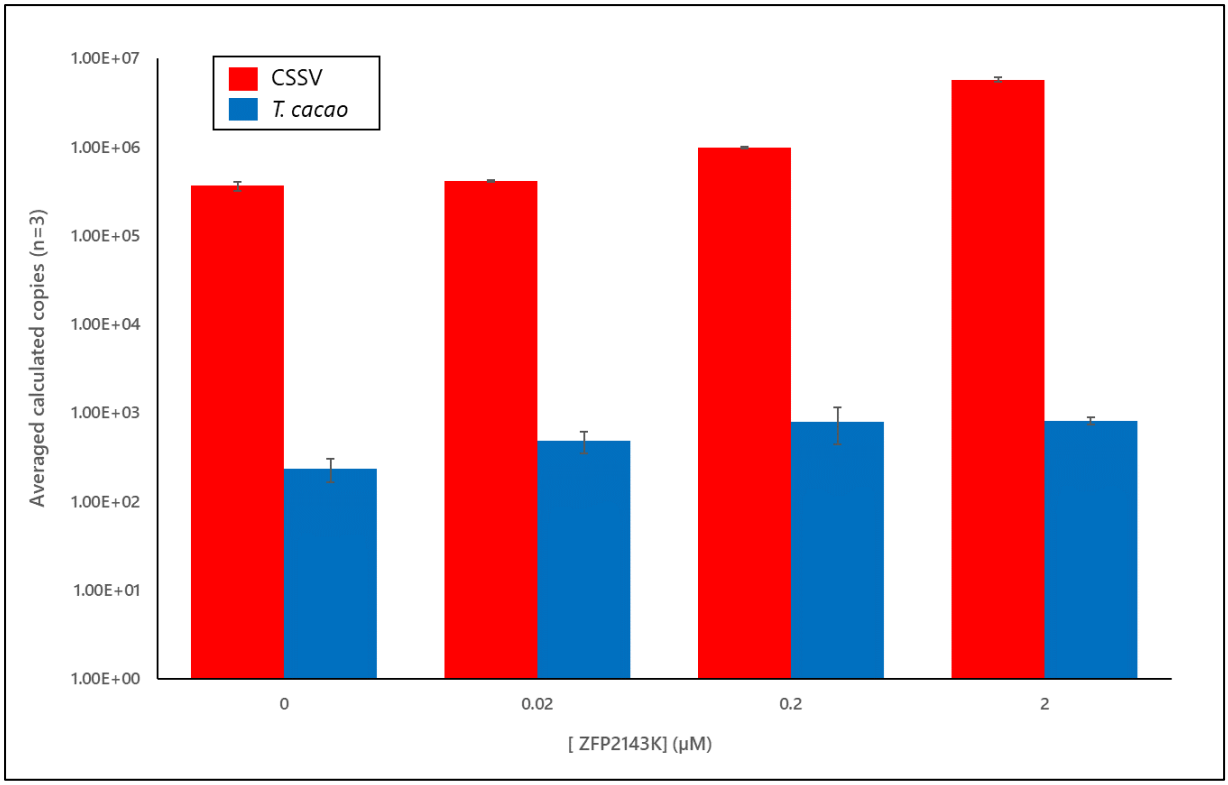


Figure 2.18: Stringency washes improve the CSSV/*T. cacao* calculated copies ratio when 2µM ZFP2143K are incubated with *T. cacao*.

Mean (n=3) of calculated copy numbers from a qPCR targeting CSSV and *T. cacao*, as a function of the concentration of ZFP2143K used in the binding reactions. Stringency washes improved the specificity of the reaction by reducing the amplification of the plant DNA. The ordinate axis was transformed to a logarithmic scale.

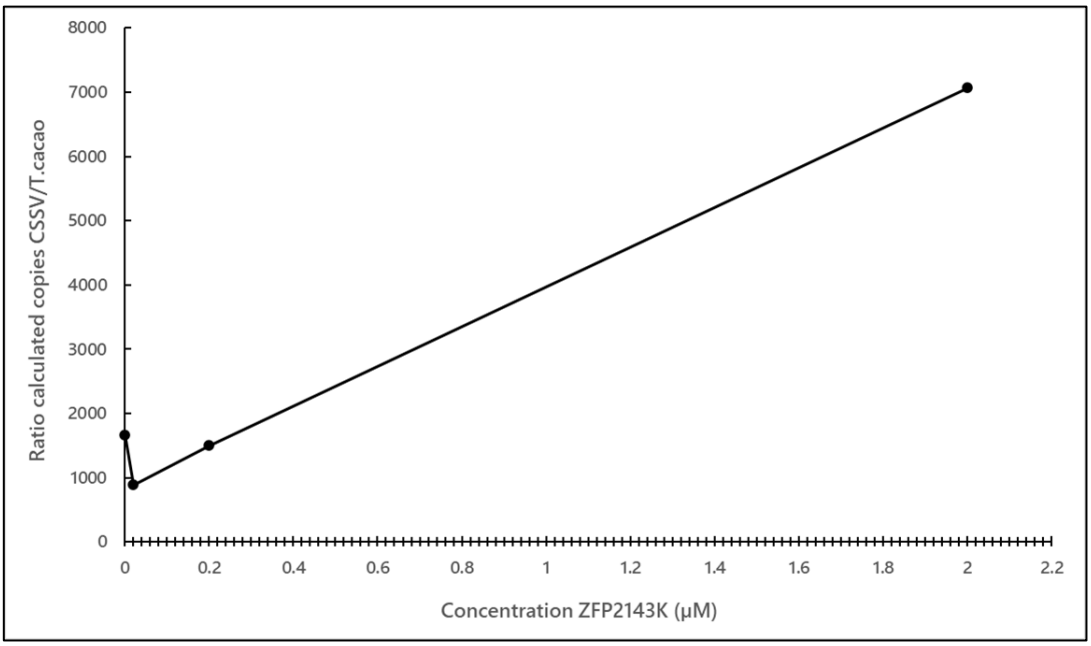


Figure 2.19: Ratio of calculated copies of CSSV/calculated copies *T. cacao* as a function of the concentration of ZFP2143K derived from the data presented in Figure 2.18.

Ratios were calculated by dividing the average calculated copies number from three PCR replicates.

Similarly, to what was observed in the previous experiment (Figure 2.16), an increase in recombinant was associated with a higher amount of DNA captured by the magnetic particles for both CSSV and *T. cacao*. Adding 0.02 µM of ZFP2143K increased the amount of CSSV DNA captured but decreased the CSSV/*T. cacao* immobilised DNA ratio by half. 0.2 µM ZFP2143K resulted in a higher amount of CSSV DNA amplification, but the ratio did not improve compared to the control with the recombinant (Figure 2.19). At a concentration of 2 µM, the ratio was significantly higher. CSSVDNA was in more than a 7000-fold excess compared to *T. cacao,* highlighting the prevalence of viral DNA in the sample.

In order to improve the signal to background further, the binding reactions time between the recombinant protein and the DNA preparation was decreased by half to 15 min in order to potentialize the binding of the recombinant to the shorter viral genomic material (Figure 2.20). In addition, two washing procedures were compared. Samples bound to the magnetic particles were either washed with stringency washes (twice 2xSSC with 0.1% SDS at RT followed by twice 0.1xSSC with 0.1%SDS at 40°C) or with the PBS1X with 50 mM Imidazole and 0.05% Tween 20 used in the first experiment.

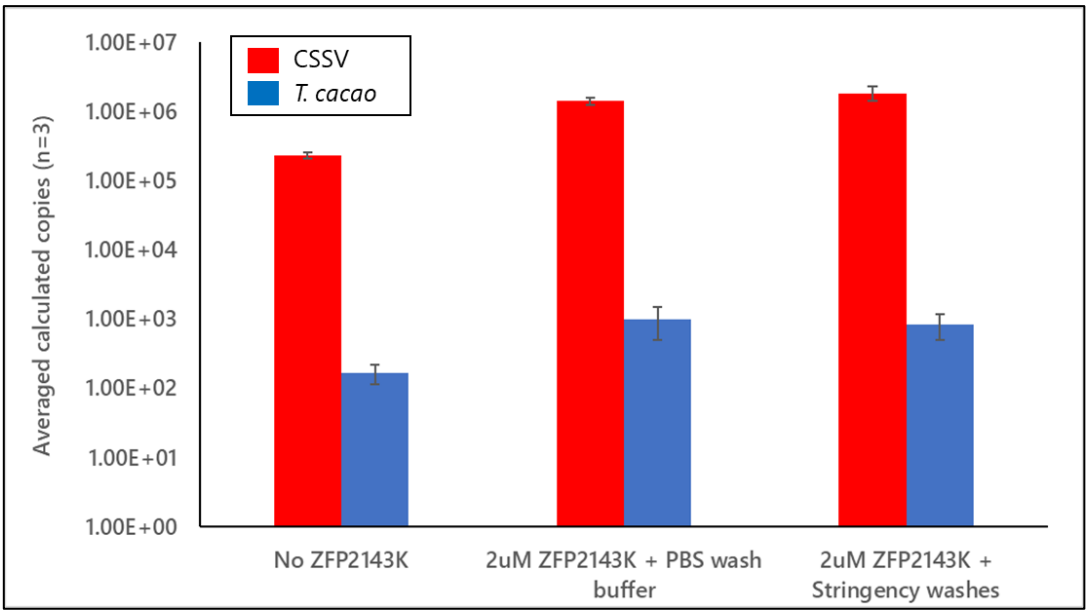


Figure 2.20: Decreasing the ZFP2143K-DNA incubation does not improve the ratio of CSSV/*T. cacao* bound on the particles.

Mean (n=6) of calculated copy numbers from a qPCR targeting CSSV and *T. cacao.* The binding reaction time was reduced compared to the experiment presented in Figure 2.18 and two washing procedures of the particles were tested. The ordinate axis was transformed to a logarithmic scale.

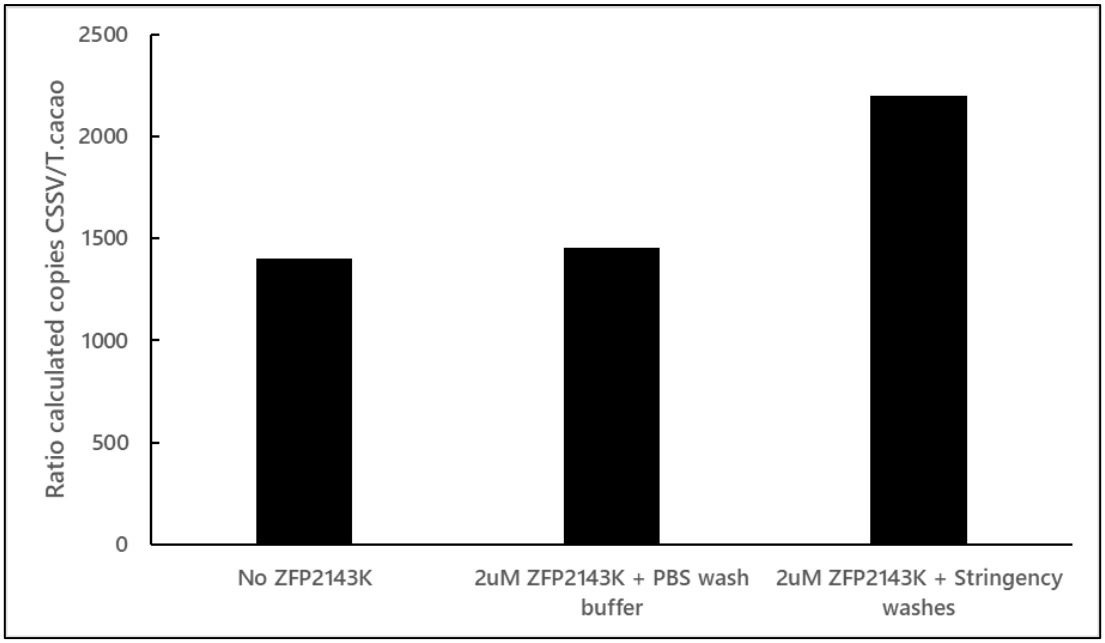


Figure 2.21: Ratio of calculated copies of CSSV/calculated copies *T. cacao* derived from the data presented in Figure 2.20, analysing the effect of two washing procedure and a shorter reaction time between ZFP2143K and the DNA.

Ratios were calculated by dividing the average of six replicates of CSSV amplification by *T. cacao* amplifications.

The mean ratio of eluted CSSV/*T. cacao* DNA was not improved by a reduced recombinant/DNA incubation time (Figure 2.21). CSSV was still predominantly amplified, but the mean ratio measured (1452.5) was nearly identical to the no recombinant control (1403.6).

The stringency washes procedure slightly improved the ratio of eluted CSSV/*T. cacao* DNA (2199.6) however, the effect was not as marked as the data obtained with a longer protein-DNA incubation time (7049, Figure 2.19).

2.4 Discussion and conclusion.

A bespoke synthetic protein targeting a DNA sequence from CSSV was designed using the zinc finger tool website. The website offers a user-friendly, simple mean of screening CSSV genome to identify potential binding sites. The selected ZFP target site, 2143, returned hits predominantly on isolate of CSSV by BLAST search (85%), which validated the site and enabled production of the corresponding synthetic ZFP. The binding site of ZFP2143K was designed to measure 18 bp, with the aim of being a unique sequence within the genome of CSSV. *T. cacao* genome is approximately 70.000 times larger

than CSSV. The size difference suggests that there is a possibility that the target sequence may also be found within the plant's genome. However, when considering a CSSV-infected *T. cacao* sample, the target sequences would be more frequently associated with the virus rather than with the plant due to the virus copy number. Production of the 528 bp-long 2143 gene corresponding to the Zinc finger protein was made by a commercial third-party company because of time constraints and for the relatively low cost of the procedure. However, it would have been possible to synthetise the gene in-house using PCR and standard cloning techniques.

The synthetic gene was subcloned in a prokaryotic expression system, using the pet-21(+) plasmid, a generic expression vector. The protein ZFP2143K was produced soluble in quantity compatible with subsequent functional tests without requiring specific optimisation of the expression. It was noticed that the non-induced culture also produced a faint signal, suggesting that the *lac*UV5 promoter of the gene partially leaked. However, this is expected from this expression system and high titer growth of the bacteria and production of the recombinant shows that the protein was not detrimental to the cells. Further optimisation of the expression could have improved the ratio soluble/insoluble fractions of the recombinant. This was not done deemed necessary because the biochemical tests used to assess the functionality of the protein did not require large amounts. Optimisation could however be envisaged in a situation where the recombinant would require large scale production, for integration on a biosensing testing platform.

The functionality of ZFP2143K was evaluated using a set of biochemical tests. First, the ability of the recombinant to bind to its cognate DNA partner was measured by the double binding filter assay, using a short, pre-annealed, biotinylated oligonucleotide. The method was successful in proving that ZFP2143K binds to the probe that was designed to contain the sequence the protein was raised against. This experiment also proved that the protein was binding to the DNA in a dose-dependent manner (Figure 2.6 and 2.8). Optimisation of the blocking steps could have led to a better signal-to-noise ratio, which is an issue that was encountered while performing the experiment. This could be

achieved by further testing combinations of blocking reagent, blocking time, and washing procedures (Heda *et al.*, 2020).

High Resolution Melt Analysis was explored to confirm if it was possible to detect the binding of the ZFP to the targeted sequence, using a pre-annealed double stranded oligonucleotide. The raw fluorescence measurements sampled during the pre-melting phase (30°C-37°C) when ZFP2143K is expected to be functional revealed always lower intensity levels in samples containing the recombinant compared to the non-protein controls. This suggests that binding of ZFP2143K would have prevented some of the dye to bind to the double-stranded oligonucleotides, since the fluorescent molecule SYTO13 was added following DNA-protein incubation, in a similar way as a DNA-binding protein would prevent DNA enzymatic digestion in a nuclease protection assay (Brenowitz *et al.*, 1986). Also, binding of the protein would explain why the fragment melting temperature measured was lower in samples incubated with the recombinant in comparison to the samples not containing the ZFP. More Syto13 dye remained in the control samples compared to the one exposed to the Zinc finger protein. The difference was particularly marked when the reactions were set-up in binding buffer, compared to annealing buffer. A main difference in the buffer composition is the presence of 1 mM Dithiothreitol (DTT) in the binding buffer. DTT is a commonly used reducing agent that prevents cysteines from proteins to form inter and intra-molecules bonding. The presence of this molecule in the reactions ran in the binding buffer may have facilitated binding of ZFP2143K to the double stranded DNA partner molecule (Branza-Nichita *et al.*, 2002) .

To further explore the specificity of the protein to the target in larger DNA sequences, EMSA was carried out with PCR amplicons. The first experiment involved binding the protein to a fragment, S1, containing the target sequence (Figure 2.13). As the concentration of ZFP2143K increases, the PCR fragment signal faded, and a smear was eventually observed when the binding reactions were set with ZFP2143K at 236 nM. When applying higher concentrations, the DNA fragment did not leave the gel well. This contrast with expected EMSA results for single protein when a shift of positioning is

observed between protein free DNA product and protein bound product. The result suggested instead and under the conditions tested, a range of DNA fragment sizes rather than a single product migrated, leaving this characteristic smear pattern which could be explained by each finger domains binding to distinct sites within the amplicon, creating a scaffold of DNA molecules.

To confirm this, the EMSA was run with another fragment (S6) not containing the target region (Figure 2.14). The same results were observed confirming that the protein binds to multiple sites. While not containing the target sequence, S6 contains the binding sites of two contiguous domains targeting the sequence CAA ACA (Figure 2.22).

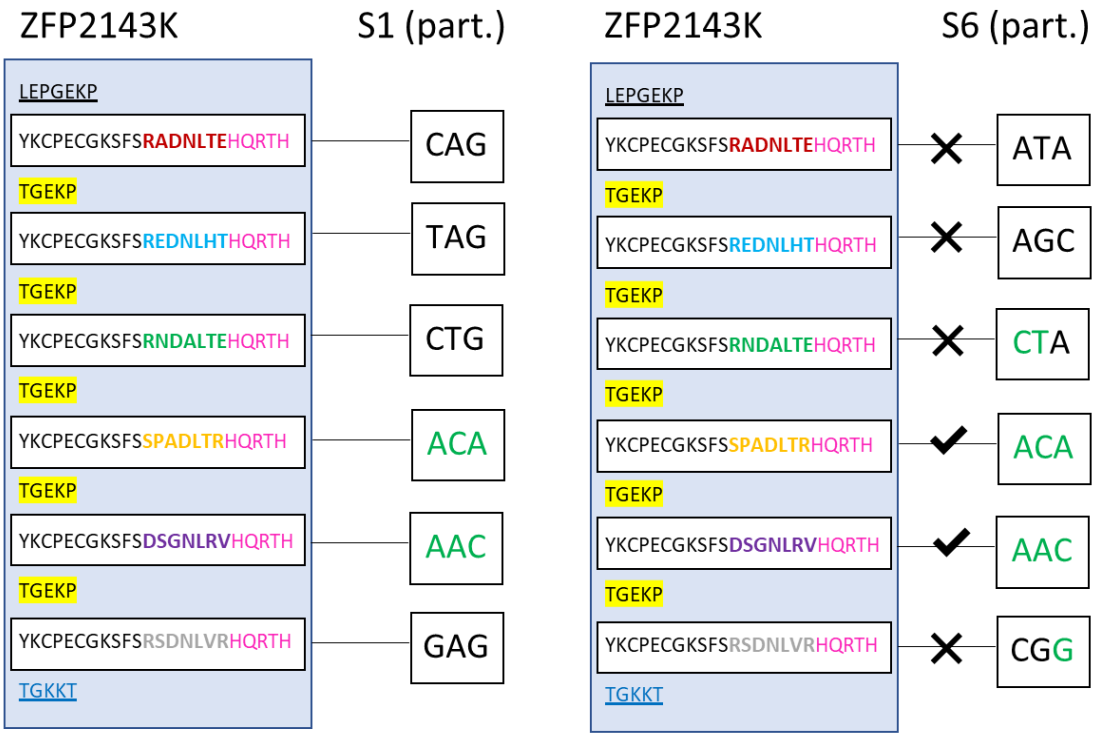


Figure 2.22: A 9 bp segment of the targeted region of interest is present on the PCR amplicon S6.

Left: primary structure of ZFP2143K with the corresponding codon from the S1 amplicon bound by each domain. Right: primary structure of ZFP2143K with the corresponding codon from the S6 amplicon: two contiguous codons recognised by two domains of ZFP2143K are present in the sequence of S6.

The data observed suggests that the interaction of two domains with the corresponding two codons was enough for ZFP2143K to bind the whole fragment.

Furthermore, an accumulation of fluorescence was observed, in wells containing higher concentrations of protein, suggesting that even larger DNA-protein aggregates prevented migration. It is however difficult to ascertain if the precipitation was due to a low affinity of the recombinant to the targeted DNA fragment or because of interfering substances. The buffer selected for the IMAC purification did not contain detergent because the bacterial cell lysis relied on a gentle sodium phosphate buffer disruption. However, it is possible that the imidazole present in the elution buffer may have interfered with the binding, as it was reported to inhibit DNA-BP/RNA interaction previously (Fuellgrabe et al., 2011). A dialysis of the protein solution may have been beneficial to the assay, by removing the imidazole.

So far, the experiments were using short, synthetic DNA sequences either annealed oligonucleotides fragments or PCR amplicons. The final experiments focused on testing the bespoke recombinant with whole DNA extracted from CSSV infected *Theobroma cacao* leaves samples, in order to emulate field conditions. ZFP2143K was used in pull-down assays where pre-incubated binding reactions were combined with paramagnetic particles surface-modified with nickel. Subsequent qPCR results confirmed that the recombinant was able to bind to the viral DNA in a dose-dependent manner (Figure 2.16, 2.18 and 2.20). However, it was also found that high amount of *T*. *cacao* DNA was binding to the particles. This is consistent with the data from the EMSA experiments, where it was established that some of the contiguous domains from the recombinant protein were able to bind to fragments of DNA only containing part of the targeted sequence. In infected cacao leaves, a high density of binding sites is present with full target sequences observed on the CSSV genome and a range of potential binding sites for some of the domains of the recombinant ZFP. Using magnetics beads combined with the ZFP, it was possible to capture the CSSV target efficiently with higher copy number observed when using higher protein concentration level. However, corroborating the EMSA results, the protein also bound readily to the plant genomic DNA. Furthermore, if the protein concentration was increased in the binding reactions, the ratio between CSSV and nuclear genome capture decreased.

It is possible that the whole six domains of the recombinant ZFP bound to its cognate target within CSSV's genome and simultaneously to other sites on either the virus or the plant, creating large aggregates of DNA from both source that were then co-eluted from the magnetic particles.

In order to mitigate this, stringency washes were included in the particle's treatment procedure. Protein-DNA interaction is a complex process, involving hydrogen bonding between selected amino acids and nucleic acids bases, Van der Waals contact, electrostatic forces, and salt bridges (Guo *et al.*, 2023);(Vo *et al.*, 2021). Stringency washes, employing buffer with defined osmolarity, are commonly used in DNA/DNA hybridizations studies in order to improve the signal-to-noise ratio of the reaction(Sterchi, 2008). It traditionally consists of a non-stringent wash to remove non-specifically bound DNA, followed by a more stringent wash to only permit specific interactions. The stringency is normally controlled by both salt concentration in the washing buffer and the temperature of the washes. Because it was anticipated that the recombinant protein would not be structurally active at 65°C, the stringency was here only modulated by the concentration of the SSC buffer. The ratio of CSSV DNA/*T. cacao* bound on the particles was improved when stringent washes were applied (Figure 2.18 and 2.19). Further, a better ratio was observed when the binding reaction time between the zinc finger protein and the nucleic acids was set to thirty minutes as opposed to fifteen minutes.

Another possibility is that ZFP2143K attached to more than one site, both as a whole protein but also which each domains acting independently. The zinc finger protein ZFP2143K is a C2H2 multidomain protein and while the "zftools" website model computed the identified target sequence to be the most likely binding partner, each domain may bind to other sequence, with a different affinity. It has been established that each domain of C2H2 zinc finger multidomain proteins can bind sequences independently (Iuchi, 2001). Binding pattern mechanisms of already characterized zinc fingers are complex, and more research is required to fully understand how synthetic, engineered zinc finger are attaching to their cognate DNA sequences. For example, the Ikaros protein (Xia *et al.*, 2021), a six domains zinc finger transcription factor encoded by the *IKZF1* gene and involved in lymphopoiesis and

tumorigenesis in humans, binds to DNA using the four N-terminals fingers while the two C-terminals fingers bound to each other, forming a homodimer that strengthen the bond of the protein to its target.

The best bound CSSV/*T. cacao* ratio was observed with 2 µM of ZFP2143K (Figure 2.17 and 2.19). The reaction could be possibly improved by increasing the contact time between DNA and the recombinant and then applying stringent washes. Another option would be to pre-treat the particles with blocking reagents to limit the amount of DNA bound on particles, as it was found that DNA from each origin was able to bind without protein present in the reaction.

Fabricating and testing the zinc finger protein led to promising results and with some further optimisation, using this technology and field biosensor is an avenue that could be pursued. Others (Kim *et al.*, 2011) have used an array of two synthetic ZFP immobilised on a hydrogel-coated surface to detect bacterial DNA. The bespoke proteins were each fused with halves of a β-lactamase, that would be re-assembled upon binding of the ZFP to its DNA target, a method called sequence-enabled reassembly (SEER). The successful reassembly, implying that the ZFP was bound to the target, was signalled by the addition of a chromogenic substrate to the β-lactamase, nitrocefin, enabling the visual observation of the binding event. This system was able to achieve a limit of detection of 5nM of the targeted DNA oligonucleotide, which is in line with other plant virus biosensors (Tang *et al.*, 2007b). Such a system could be envisaged for detecting CSSV with a careful selection and thorough testing of the ZFP to maximise the specificity of the binding. Ideally, a bioinformatic analysis could pre-select several ZFPs deriving from the zinc finger tool website. The candidates would be cloned and expressed in small scale then tested simultaneously for affinity and specificity to their target. Amongst the successful recombinant, combinations could be assessed to optimise the binding further. The produced proteins would then be integrated as the biological recognition element of the biosensor.

Chapter 3 : Design and chemical-mediated manipulation of the structure of a DNA biological recognition element to detect CSSV.

3.1 Introduction.

Symptoms specific to cacao swollen shoot virus (CSSV) can be identified in the field by trained personnel, however such clear manifestations of infection by CSSV happen late after the onset of the disease, which gives the pathogen an opportunity to spread to neighbouring trees. Being able to detect the virus as early as possible would therefore be beneficial to the cacao sector. Current methodologies to accurately detect CSSV in asymptomatic *T. cacao* trees require the use of research laboratory facilities with often variation on PCR targeting the CSSV genome used as the main platform of detection. Since 67% of the cacao production occurs in West African countries with limited technical and financial resources, a cost-effective and simple method to tackle CSSV would be advantageous. A PCR-based detection requiring rapid and accurate temperature changes would be challenging and expensive to transpose in the field. A loop-mediated isothermal amplification (LAMP) assay has been developed by the company Swiss DeCode (Swiss DeCode, 2021) and relies on the amplification of target loci on the CSSV genome. This test has been developed to target strains commonly found in West Africa. While effective, the cost of the assay is likely to be fairly high since it requires the use of a combination of engineered enzymes. Another issue regarding amplification of nucleic acid target has emerged recently since it was discovered that the *T. cacao* genome contains a range of viral sequences integrated during its recent evolution which bear strong similarities to the CSSV genome. (Muller *et al.*, 2021b). Named eTcBV for endogenous *T. cacao* bacilliform virus, the sequences are present in asymptomatic trees, and genetic analysis suggests that the integration occurred prior to the export of the tree in other continents. They were identified following the PCR screening of several loci found on ORF3 of material from The International Cocoa Quarantine Centre,

at the University of Reading (ICQC, R) (Allainguillaume, J., personal communication). This material was known to be CSSV free which raises the possibility of false CSSV positive results on asymptomatic trees wherever they are grown with assays relying on sequence amplification. A membrane-based biosensor targeting the nucleic acid of the virus could therefore mitigate the potential false positive detection. It might be easier to identify regions of CSSV not found in *T. cacao* genome. Unlike PCR approaches which require two primers and sometimes an internal probe thus covering a larger genomic area, a hybridisation assay could work with a short unique sequence. Furthermore, hybridisation assay relies on the detection of actual copy numbers of a targeted sequence of interest. In a CSSV-infected tree, actual viral copies of a targeted candidate sequence would outnumber a putative integrated sequence that would be present as two copies per cell in the case of a single insertion event. This means that a probe-based hybridisation approach would be less likely to cause a false positive signal. Assays relying on a membrane hybridisation could meet the cost and limited personnel training requirements even if LAMP assays are fairly user friendly. A "cold" membrane-based assay involving chemicals to manipulate the structure of the double-stranded DNA molecules and enable binding of the labelled oligonucleotide to the targeted viral sequence could replace temperature-induced denaturation and renaturation of the nucleic acid.

DNA denaturation is the process where the complex bond between two DNA fragments of a double-stranded DNA molecule is disrupted, yielding two single-stranded molecules. It is naturally occurring during normal molecular processes such as DNA replication. DNA helicases catalyse the unwinding of the double helix, allowing the DNA polymerase to reach its substrate, which is a single-stranded specific sequence. It has been established that the DNA molecule operates a phase transition, following a helix to coil model. During this process, the organized, tightly packed polymeric molecule loses its three-dimensional structure and eventually, produces two single disorganised strands (Hammouda *et al.*, 2006). Alongside these enzymatic reactions, temperature induced denaturation has been extensively studied and provided the basis for the development of analytical methods. Such work has made a technique such as polymerase chain reaction possible. Indeed, each DNA molecule

is characterized by its melting temperature (Tm), that is the temperature where 50% of the molecules of a given DNA species are single stranded.

As an alternative to temperature, a range of chemicals has been used to alter the structure of DNA. Some of them have been routinely used to decrease the melting temperature of DNA in hybridization experiments, such as dimethylformamide (Blake and Delcourt, 1996a) or to relax secondary structures of DNA for PCR enhancement with dimethyl sulfoxide (Hardjasa *et al.*, 2010). Alkali denaturation of DNA has been studied since the early 1970's (Ageno, Dore and Frontali, 1969a) and this has led to the development of the alkali extraction procedure of plasmid DNA also known as a mini prep by Birnboim (Birnboim and Doly, 1979) from recombinant *Escherichia coli*. In this study, the authors use the differential melting abilities of high molecular weight DNA (the genomic DNA) and comparatively lower weight (the plasmid DNA) to purify the latter. Sodium hydroxide is therefore used to both lyse the cells the plasmid is extracted from and to denature DNA.

Using chemicals to modulate the strength of the interaction between a target DNA molecule and a reporter probe in "cold" conditions is therefore a promising approach to avoid the requirement for temperature control of the assay. Early methods to study DNA denaturation relied on physico-chemical properties such as sedimentation coefficients, viscosity, and light-scattering (Eigner and Doty, 1965). Since then, study of DNA melting curves was implemented, and when coupled to spectrophotometric techniques, provided the most common means of monitoring DNA denaturation. This technique relies on a peculiar property of the DNA molecule which is its hyperchromicity. This feature of the DNA molecule was experimentally observed (Thomas, 1993) together with the discovery of its structure in the early 50's.

The stability of the DNA molecule being strongly influenced by its sequence, useful genomic data can therefore be derived from DNA denaturation experiments. High Resolution Melt (HRM) Analysis was historically developed to study genetic variants (Wittwer *et al.*, 2003). It is based on the parallel development of two new types of equipment, namely the Lightcycler that enables the real-time

monitoring of DNA quantity with fluorescent dyes (Lay and Wittwer, 1997) and the development of new fluorescent dyes (Wittwer *et al.*, 2003) known as third generation dyes like Evagreen. In this technique, melting of the DNA fragments is achieved by increasing the temperature of the samples incrementally, until all fragments are temperature denatured. HRM analysis can monitor this phenomenon in real time thanks to theses dyes: while unbound dye does not produce fluorescence, a fluorescent signal is emitted when the dye is bound to DNA and is detected in real time. The other main difference derives from the equipment itself. HRM devices record up to two hundred times more data points compared to traditional Lightcyclers (Herrmann *et al.*, 2007). Besides, HRM equipment can process tens of samples at the same time in volumes as little as 25µL, reducing the amount of sample needed compared to the ultraviolet-visible spectroscopy. Since its development, HRM has been used for numerous applications, including pre-sequence screening (Provaznikova et al., 2008), determination of the methylation status of DNA in cancer (Ehrich *et al.*, 2006) or quantification of methylation in plants (Rodriguez *et al.*, 2010), single nucleotide polymorphism (SNP) detection (Garritano *et al.*, 2009) and as an alternative to gel electrophoresis (Vossen et al., 2009). HRM could therefore be able to monitor subtle structural changes of the DNA molecule used as a BRE for the detection of CSSV.

3.1.1 Aims and objectives of this chapter.

This chapter describes the design of an oligonucleotide lattice composed of oligonucleotide probes derived from the sequence of an isolate of the CSSV strain New Juaben available in *T. cacao* plants growing in the envirotron at UWE Bristol and its assessment as a hybridisation probing system.

Objectives will cover:

- The *de novo* sequencing of an isolate of the New Juaben strain of CSSV. This is achieved by designing set of primers to amplify overlapping fragments covering the whole DNA sequence of CSSV followed by Sanger sequencing.

- The assessment of the ability of an oligonucleotide lattice used to detect a synthetic fragment derived from CSSV's genome sequence. This is tested using a dot blot assay and HRM.
- The evaluation of HRM as a tool to study the chemically controlled denaturation and renaturation of the lattice.
- The development of a competition assay to model the interaction of the oligonucleotides with the targeted region of interest by HRM.

3.2 Material and Methods

3.2.1 CSSV sourcing.

Viral material was sourced from CSSV-infected *T. cacao* saplings maintained at the UWE envirotron. The plants were originally kindly provided by Dr Andy Wetten through the University of Reading. The saplings were produced from seedling infected with CSSV at the cocoa research institute of Ghana (CRIG) and transferred to the University of Reading. The viral source was maintained by Dr Wetten through mealybug infection of healthy cacao cultivars. This isolate is derived from the strain New Juaben from Ghana is used as the reference material. Healthy cacao cultivars were also kept at the UWE envirotron, to be used as a reference negative material.

3.2.2 CSSV sequencing.

For the design of the assays an up-to date genomic sequence of the CSSV strain present in the UWE plants was required to determine sections of CSSV genome that were the less variable across the 8 accessions available when the project began (2015), to optimise the chance of the assay to detect the virus across multiple strains. The entire sequence of the strain available was therefore sequenced *de novo* from saplings using Sanger sequencing. The New Juaben strain sequence AJ608931 (Muller and Sackey, 2005) deemed to be the most similar to the UWE strain was used to design a set of primers covering the whole 7024 base pairs (bp) genomic sequence of the virus. Primer pairs were designed to amplify 11 overlapping fragments ranging from 536 to 910 bp. An additional seven primer pairs

were added to obtain a complete cover of the genome following unsuccessful amplifications in the first round of amplification (Table 3.1).

Table 3.1: Primer sequence, name and theoretical size of the PCR amplified fragments used for *de novo* CSSV sequencing.

Fragment ID	Theoretical amplicon size (bp)	Forward primer Name and sequence	Reverse primer Name and sequence
S1	655	NewJ-S1F GCATTTCCATCAGATTCCATGA	NewJ-S1R TTTAAGAATCTTGGCACGTGTG
S2	701	NewJ-S2F GGTGACACACTGATGACACCTG	NewJ-S2R AAGACTCCTACTTCATTGCCTG
S3	792	NewJ-S3F TTCTAGGAACAGAGCTCAGTGA	NewJ-S3R AGATAGGGGCACACATGAAGCA
S4	867	NewJ-S4F TGCTTGCACCAATGGGAGCACA	NewJ-S4R GTAGAAGGTGATTGTGTGACCT
S5	711	NewJ-S5F CCGGATCCCGTACTGTTATGCC	NewJ-S5R CGCTTCAGGATGGTCTGTATAC
S6	907	NewJ-S6F AATCAGGCACTGTTATTGATCC	NewJ-S6R AATTTTCCACTAGCGTAGGCAC
S7	830	NewJ-S7F-6255 TCGCTATCCCACCAGCAAGATG	NewJ-S7R ATCGTCCAAACTAGGGTTGGAC
S8	834	NewJ-S8F AACTGGGCCAGTGATAGACTGC	NewJ-S8R GAGCTAGTTGGTTGTTGGTCAC
S9	761	NewJ-S9F AGTATCCAGGAATGGTATGAGA	NewJ-S9R CAAAATTTTGTAAGGATCTTTG
S10	794	NewJ-S10F CCATCCAGGAGAAAGGTGAAGC	NewJ-S10R TTCTCATATCCTCGTGTGAGGA
S11d	536	NewJ-S11F GTYRTACCRRAYAYYATGATGAC	NewJ-S11R GTTYCCRTTRSYRGAYTCYTCCCATAC
S1a	458	NewJ-S1aF TCTATCTTCGCACAAATGGAGA	NewJ-S1aR TCCTCGTTGTAATGGGGTCT
S4a	524	NewJ-S4aF CAATCTTCCTGAGGGTTTCG	NewJ-S4aR GCCTTAGATCTGCAGCTTCC
S5a	588	NewJ-S5aF CAGCGGTTAACAGAGGCAGT	NewJ-S5aR CAGCGGTTAACAGAGGCAGT
S8a	823	NewJ-S8aF GCTATCCCACCAGCAAGA	NewJ-S8aR CCAAACTAGGGTTGGACAGC
S91a	919	NewJ-S91aF CGTGATTAACGCCATTATCG	NewJ-S91aR CTCGGAACTGCTTGGTTAGC
S92a	500	NewJ-S92aF ACCAAGCAGTTCCGAGAAAA	NewJ-S92aR TGTTCCTTTTGGCTCTGGTT
S93a	387	NewJ-S93aF TGAAGAACTGGTGCAAGATCA	NewJ-S93aR TCAACCTTCTGGCATTCCTT

3.2.3 PCR amplifications.

CSSV DNA was extracted from leaf samples using the GeneJet KIT Plant DNA extraction kit (ThermoFisher Scientific) using the manufacturer' instructions. PCR reactions were set up in a final volume of 40 µL containing: 4 µL CSSV DNA, 20 µL DreamTaq Buffer Master Mix 2X (ThermoFisher Scientific), 2 µL of each sequencing primers pairs (2 µM) and 12 µL DNase/RNase Distilled water (Invitrogen). All reactions were run in triplicate on an Applied Biosystem GeneAmp PCR System 9700 thermocycle instrument using the following parameters: Hold 94°C 2 min; Cycle: 94°C 30 s, 56°C 30 s, 72°C 60 s; Final extension 72°C 300 s. Cycles were allowed to repeat 40 times.

PCR reactions were loaded on a 2% agarose tris-acetate-ethylenediaminetetraacetic acid (TAE) gel and the gels viewed in a UV gel documentation system to assess amplification success. Single bands amplicons were column-purified using the GeneJet PCR purification kit (ThermoFisher Scientific) following manufacturer protocol and sent for Sanger sequencing to Aberystwyth University. Sequencing was performed using both forward and reverse primers for each product. Sequencing traces were curated using the sequences analysis software Geneious prime (Kearse *et al.*, 2012), and aligned to the CSSV New Juaben sequence AJ608931.1 chosen as a reference that was previously published (Muller and Sackey, 2005).

3.2.4 CSSV genome sequence assembly.

Sequencing traces readouts from both primers were aligned for all fragments using the multialin webpage (Corpet, 1988). Segments of sequence containing sequencing errors or ambiguities were removed. A consensus sequence was constructed for each fragment amplified. Consensus sequences were assembled into a contig using the sequence analysis software CodonCode Aligner (CodonCode Corporation, 2002). The consensus sequence of the contig was retained as a reference sequence.

3.2.5 Biological recognition element oligonucleotide lattice design.

CSSV genomic sequences sets were downloaded from the public database NCBI. Whole genome sequences were aligned using the Geneious prime software (Anon., 2019). Areas of lesser variability

were pre-selected to conduct local alignments to refine the region of interest. Candidate regions were selected, and the process of alignment reiterated. A segment of the open reading frame 1 (ORF1) of 76 base pairs was chosen to develop the model and was named ORF1T (Table 3.2 and Figure 3.1). The complementary strand was also synthetized and named ORF1TC (Table3.2 and Figure 3.1). Two probes covering both halves of ORF1TC were synthesised and named cap1 and cap2 (Table 3.2 and Figure 3.1). Either cap1 or cap2 could be adsorbed to the surface of the biosensor to detect ORF1T, with ORF1T able to create a bridge between cap1 to cap2, resulting in a DNA oligonucleotide lattice structure. Biotinylated versions of these probes were also purchased for dot blotting analysis. All experiments were performed on synthetic DNA oligonucleotides ordered from Eurofins as high purity salt free grade freeze-dried preparations. Oligonucleotides were resuspended in DNase/RNase-free distilled water (Thermofisher) at a concentration of 100 µM and stored at -20C until use.

Table 3.2: Name, sequence and length of the oligonucleotides used as an oligonucleotide lattice to detect CSSV.

Primer name	Sequence 5'-3'	Size (bp)
cap1	GAGTACCTTGACCTGGCCTCTACCAGTAAAGTGACCAA	38
cap2	CAACCAACTAGCTCATAACCTCGCAGTAACCTTTGATA	38
ORF1T	TATCAAAGGTTACTGCGAGGTTATGAGCTAGTTGGTTGTTGGTCA CTTTACTGGTAGAGGCCAGGTCAAGGTACTC	76
ORF1TC	GAGTACCTTGACCTGGCCTCTACCAGTAAAGTGACCAACAACCA ACTAGCTCATAACCTCGCAGTAACCTTTGATA	76

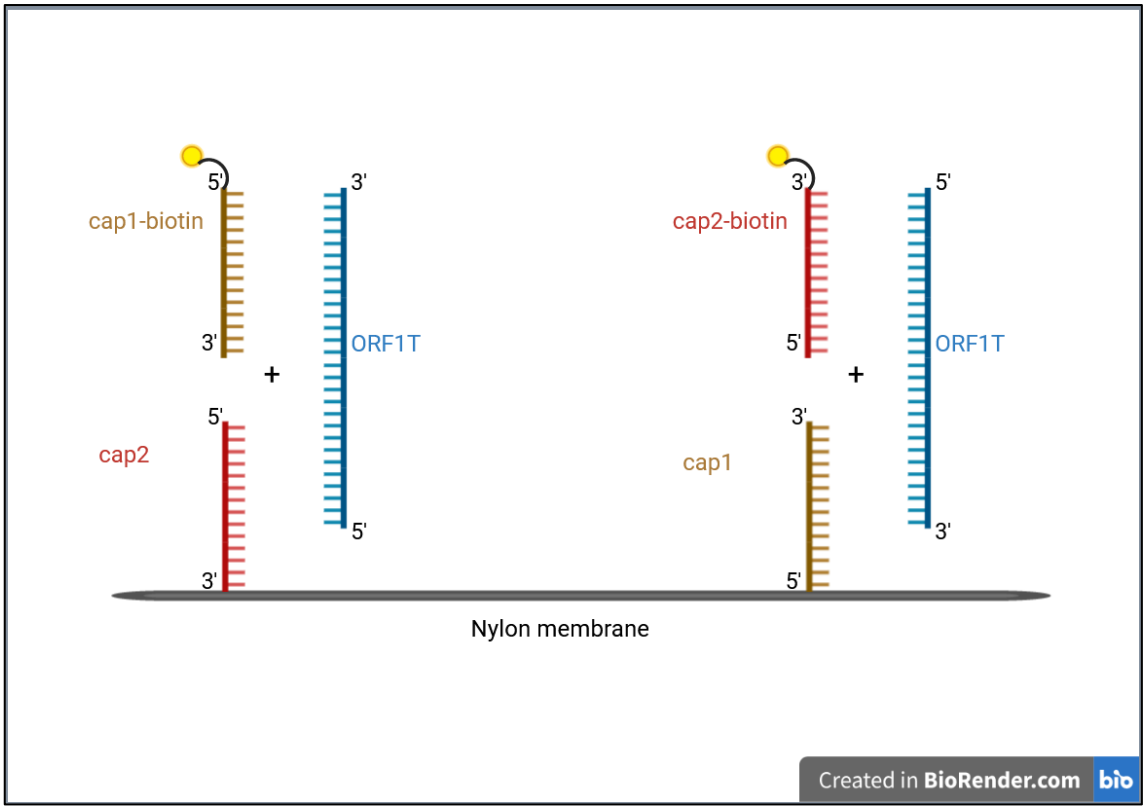


Figure 3.1: Schematic description of the lattice assembly tested by dot blot.

Cap1 and cap2 are used either as the oligonucleotide anchoring ORF1T to the surface of the nylon membrane or as the probe producing a signal.

3.2.6 Oligonucleotide lattice interaction validation by dot blotting.

Dot blot experiments were performed to verify the ability of the oligonucleotide probes to detect ORF1T. Serial dilutions of the synthetic oligonucleotides cap1 or cap2 prepared from the 100 µM stock were manually spotted in 2X Saline Sodium citrate (SSC) buffer (300 mM NaCl, 30 mM Sodium citrate) (ThermoFisher Scientific, UK) on a Biodyne B Nylon 0.45 µm membrane (ThermoFisher Scientific, UK). 2 µL of a solution containing either 50, 10 or 5 ng/µL were applied on the membrane and left to air dry for 10 min at room temperature. Negative controls spots containing only 2X SSC buffer were also applied to the membrane. Membranes were then incubated for 1h30 at 65°C in prehybridization buffer (5X SSC, 5X Denhardt's solution, 10 µg/mL Salmon sperm DNA, 0.5% SDS (ThermoFisher, UK). After this incubation, prehybridization buffer was removed and replaced with prehybridization buffer containing 1 µg of the ORF1T oligonucleotide and either 1 µg of the biotinylated cap1 probe (for the detection of spotted cap2) or 1 µg of the biotinylated cap2 probe (for the detection of cap1).

Hybridization was carried out at 65°C in a hybridization oven for 12 hours. Following this step, the hybridization solution was removed, and stringency washes were performed at room temperature as follows: 2X SSC 5 min twice, then 0.2X SSC and 0.1% SDS (ThermoFisher, UK) 5 min twice. Membrane was then blocked 1 hour at room temperature in PBS and 1% BSA (Sigma Aldrich, UK). After blocking, the membrane was incubated for 1 hour at room temperature with 1/500 streptavidin-peroxidase (ThermoFisher Scientific, UK) in PBS and 1% BSA. Streptavidin-peroxidase solution was then removed, and the membrane washed three times for 5 min at room temperature in PBS and 0.1% Tween 20. 2 mL of SuperSignal™ West Pico PLUS Chemiluminescent substrate was applied to the membranes (Pierce, UK) and signal was measured using a Li-Cor Odyssey (LICORBIO, UK) imaging system.

Densitometry analysis of the pictures was performed using the Fiji software (Schindelin *et al.*, 2012a). An area corresponding to the signal observed was set and the density of the signal was acquired with the software. The intensity of the control buffer spots was deducted from all sample results. The resulting intensity per area was calculated and plotted against the amount of oligonucleotide probe.

3.2.7 Study of the oligonucleotide lattice by high resolution melt analysis.

HRM was used to model the interaction between the oligonucleotide forming the prospective biological recognition element of the biosensor. Mixtures of nucleotides with the Syto9 fluorescent dye were placed in a Rotor-Gene Q instrument and cycled in HRM mode. Curve shape and data were then analysed.

- HRM reaction mix.

  The reaction mixture was prepared by mixing the following reagents per sample: 2.5 µL 10x Annealing buffer (100 mM TrisHCl pH7.7, NaCl 50 mM, EDTA 10 mM), 2.5µL Syto9 50 µM (Life technologies), 15 µL DNase/RNase-free distilled water (Gibco, USA) and 5 µL DNA Matrix at 2 µM, for a total sample volume of 25 µL.

- HRM conditions.

All assays were run on the Rotor-Gene Q real time PCR instrument (Qiagen). The temperature was increased from 25 to 95 °C, with 0.1 °C increments steps held for 2 seconds.

- Data Processing.

Samples were run as four replicates and averaged. The HRM curves were obtained by selecting two normalization regions, one at the start of the run before the samples are melted and one at the end of the cycle, once the strands are fully separated. This step enables the normalization of the data points from the raw channel and the subsequent comparison of the different samples analysed. Temperature and fluorescent data from the melting curves was exported to MS Excel for calculations and plotting the data measured.

The significance of observed differences between treatments in melting temperature and altered curve shape was calculated by principal component analysis (PCA) and sample cluster analysis (SCA) using the Rotor-Gene Q ScreenClust (Reja *et al.*, 2010)HRM Software (Qiagen). The software calculates the optimal number of sample clusters and allocates each sample into the most appropriate cluster, indicating, in this case, which combination of oligonucleotides and buffer compositions generates significantly different HRM curves. The software also provides the probability that each sample belongs to the assigned (and others) cluster(s) and the typicality of samples within its allocated cluster, which indicates how well the sample fits within its allocated cluster. To maximize the discriminatory power, a preliminary analysis was run for all experiments by fixing the maximum number of possible clusters for each specific experiment and then comparing to the optimal number of clusters generated automatically by the software. Only groups of samples separated by both types of analysis and with probabilities (P) >0.999 and typicalities (T) >0.5 were deemed statistically different (Rodriguez et al., 2010).

3.2.8 Analysis of the interaction of the oligonucleotides forming the biological recognition element.

To evaluate HRMA as an analytical tool, normal physicochemical conditions required for the formation of double stranded oligonucleotide were established. 2 µM of each complimentary oligonucleotides ORF1T and ORF1TC were prepared and diluted in either DNase/RNase-free distilled water or in annealing buffer (100 mM TrisHCl pH7.7, NaCl 50 mM, EDTA 10 mM). The fluorescent Syto9 dye was added at a final concentration of 50 µM, and the mixes placed in the Rotor-Gene Q.

In subsequent experiments, oligonucleotides are diluted to a final concentration of 2 µM, in annealing buffer 1X.

3.2.8.1 Competition Assay between the cap1, 2 and ORF1TC oligonucleotides.

Cap1 or 2 oligonucleotides are complementary to ORF1T and therefore will be competing with the natural fragments to bind ORF1T when field samples will be applied to the sensor. To analyse the interaction between these two types of fragments, a competition assay was run by preparing several groups of oligonucleotides mixtures.

Samples containing all the oligonucleotides were prepared by mixing ORF1T, cap1 and cap2, all at a final concentration of 2 µM. The ORF1TC sequence fragment was added diluted either 2, 10, 50 or 100 times. Controls containing ORF1T and ORF1TC only or ORF1T + cap1+ cap2 were also prepared. A master mix was prepared by combining 175 µL of 10x annealing buffer, 1050 µL of DNase/RNase-free distilled water and 175 µL 50µM Syto9. The DNA matrixes were prepared separately by diluting the respective oligonucleotides to 2 µM, except for the samples where the ORF1TC sequence was further diluted 2, 5, 10, 50 or 100 times.

3.2.8.2 Sodium Hydroxide-mediated denaturation of the oligonucleotide lattice.

Sodium hydroxide was used to disrupt the bond between ORF1T and ORF1TC. The ordering of introduction of the reagents in the reaction tube was controlled to prevent potential damage of the fluorescent dye by exposure to the sodium hydroxide.

Oligonucleotides (2 µM) were first added, then exposed to a range of concentration of sodium hydroxide (4, 6, 8, 10, 11, 12, and 100 mM) for 5 min and then placed in the Rotor-Gene Q. A master mix was prepared by mixing 180 µL of 10x annealing buffer, 360 µL of oligonucleotide mix (equimolar mixture of ORF1T, ORF1TC, cap1 and cap2 at 2 µM), 180 µL DNase/RNase-free distilled water and 180 µL 50 µM Syto9. Then 100 µL of this mix was placed in 100 µL of a sodium hydroxide at twice the concentration tested (respectively 8, 12, 20, 22, 24 and 200 mM). 50 µL of each sample was then analysed as four replicates.

Denaturation/Renaturation experiments samples were prepared by exposing a mixture of ORF1T, ORF1TC sequence and cap1 at 2 µM to 100 mM NaOH. Two groups of samples were prepared. The first group of samples was incubated in a tube containing a sodium hydroxide solution at 100 mM then the pH was lowered with 2.5 N of acetic acid. Then, DNase/RNase-free distilled water and Syto9 at a final concentration of 50 µM was added. The reagents were prepared as follows: 40 µL 100 mM NaOH, 9 µL 2.5 N acetic acid, 20 µL DNA (ORF1T, ORF1TC and Cap1), 121 µL DNase/RNase-free distilled water, 10 µL Syto9.

The second group of samples was incubated with 100 mM NaOH, DNase/RNase-free distilled water and Syto9 for 5 min to denature the DNA. Then, 2.5 N Acetic acid was added to enable renaturation of the fragments. The reagents were prepared as follows: 40 µL 100 mM NaOH, 20 µL DNA (Target, ORF1TC and Cap1), 121 µL DNase/RNase-free distilled water, 10 µL Syto9. After 5 min incubation at room temperature, 9 µL of 2.5 N acetic acid were added to neutralise the solution. All samples were then analysed in 50 µL volume, as four replicates.

3.3 Results

3.3.1 CSSV New Juaben sequencing

Eleven PCR primer pairs designed to allow the amplification of overlapping sections of the entire CSSV NJ genome were used to run PCR amplification reactions. An aliquot from each reaction was migrated

on a 2% agarose gel. Analysis of the gel electrophoresis revealed that the fragments migrated at the theoretical size calculated from the reference sequence AJ608931.1, except for fragments S7 and S11d that were not amplified (Figure 3.2).

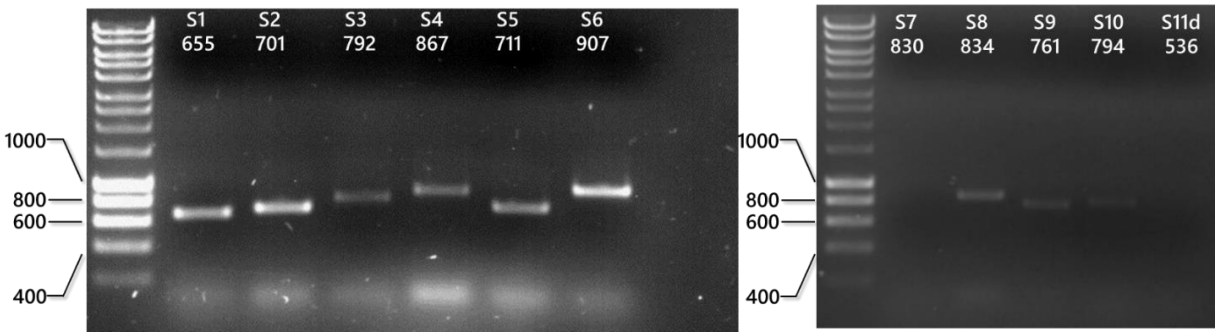


Figure 3.2: Electrophoresis of PCR amplification products of CSSV New Juaben fragments covering the whole genome.

The name and theoretical size of the fragments is delineated above each amplicon. Hyperladder (Bioline) standard fragment sizes in base pairs are indicated on the left of the gels.

Another set of primers were designed to cover the segments of the genome that were not amplified. Primers were also made to amplify overlapping fragments to confirm the sequence of the isolate. All primers pairs produced an amplicon, that migrated at the theoretical size (Figure 3.3).

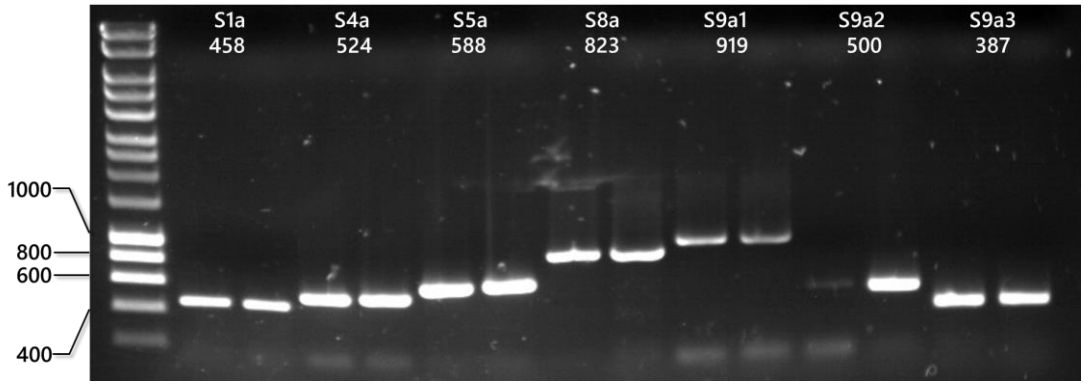


Figure 3.3: Electrophoresis of PCR amplification products of CSSV New Juaben fragments covering areas not amplified with the first set of primers or targeting overlapping segments.

Replicated reactions were loaded on the gel. The name and theoretical size of the fragments is delineated above each amplicon. DNA standard fragment sizes in base pairs are indicated on the left of the gels.

Amplicons were excised from the gels, purified, and sequenced using Sanger sequencing by the Aberystwyth University. Traces of the sequencing data were aligned *in silico* to form a contig spanning the genome of CSSV (Figure 3.4).

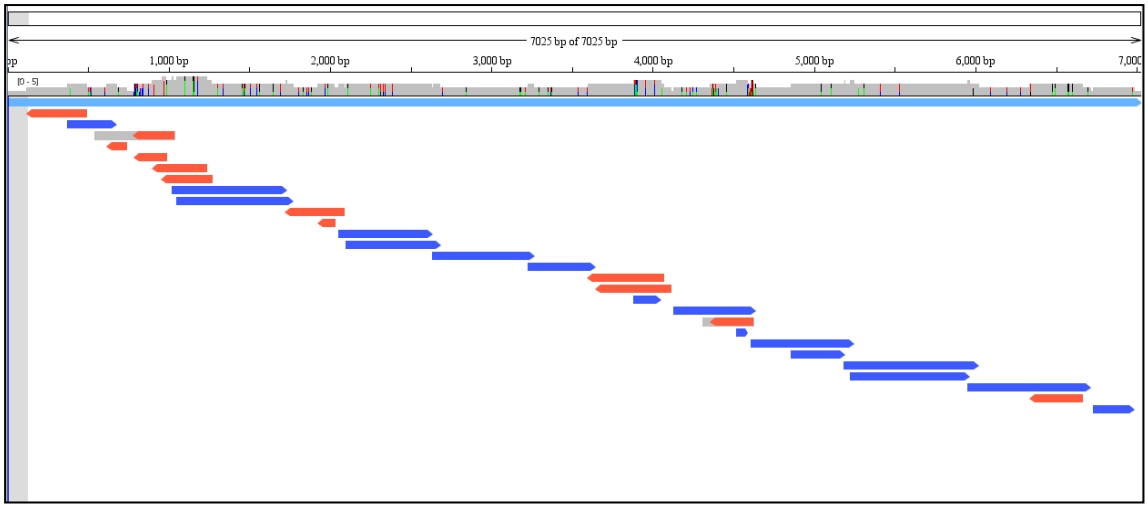


Figure 3.4: Graphical representation of the *de novo* sequenced contig aligned to the reference sequence AJ608931.1. Red segments represent sequenced fragments using forward primers, blue segments, reverse primers.

The fragments obtained were overlapping to cover the whole genome sequence. Further analysis of the sequence revealed 123 nucleotide differences with the reference sequence AJ608931.

3.3.2 Construction of the DNA-based biological recognition element.

A set of eight whole genome sequences of CSSV deriving from West African isolates were aligned to delineate regions of lesser genetic variability (Figure 3.5).

Figure 3.5: Alignment of eight CSSV whole genome sequences from the Geneious prime software.

The second line delineates the degree of identity along the alignment. Green areas represent mean pairwise identity of 100%, green/brown at least 30% and under 100% and red areas 0%. The yellow box delineates the position of the ORF1 protein.

The region open reading frame 1 (ORF1) encoding a 16.7 kDa protein of unknown function, was selected because of the conservation of this portion of the genome amongst the isolates (Figure 3.6A and 3.6B).

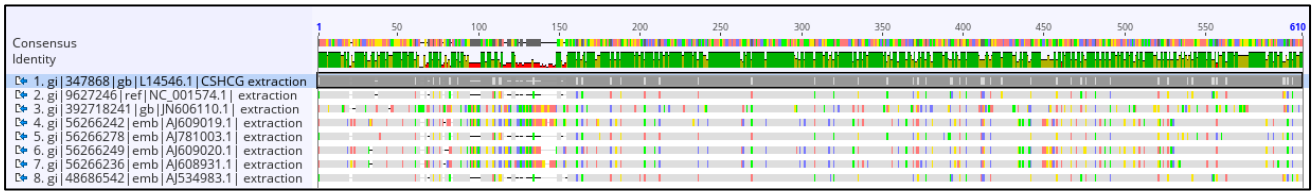


Figure 3.6A: Alignment of eight CSSV ORF1 sequences from the Geneious prime software.

The second line delineates the degree of identity along the alignment. Green areas represent mean pairwise identity of 100%, green/brown at least 30% and under 100% and red areas 0%.

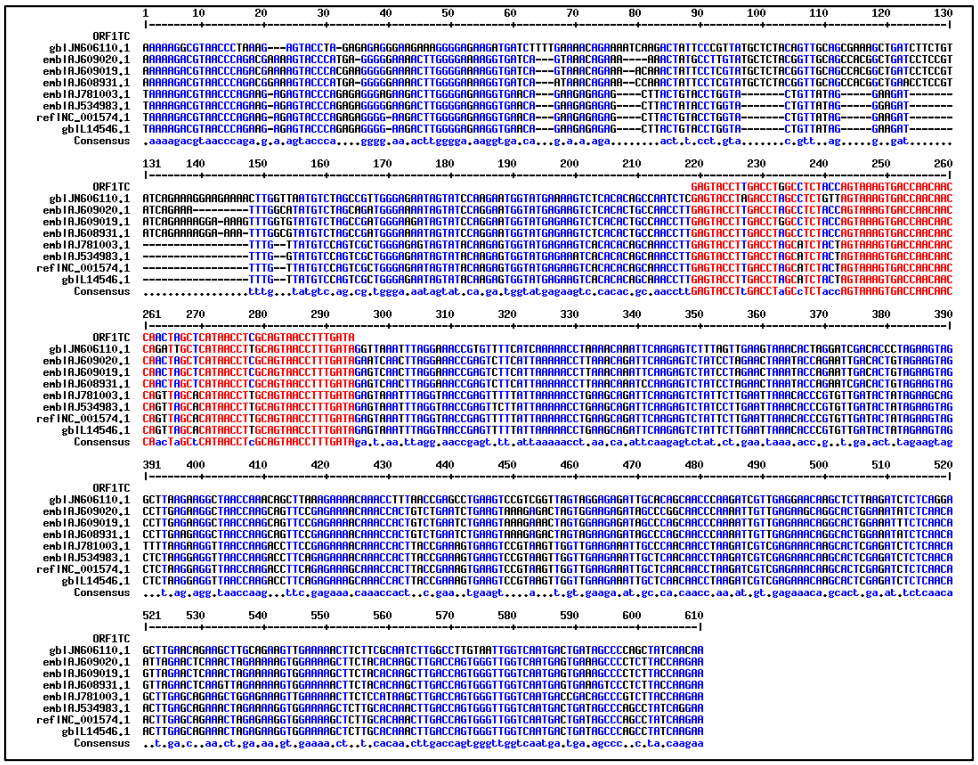


Figure 3.6B: Alignment of ORF1T with eight CSSV ORF1 sequences using the multialin algorithm (Corpet, 1988)**.**

Red letters represent a consensus of 90% amongst all sequences selected at the position considered. Blue letters represent a consensus of 50% between the sequences.

3.3.3 Verification of the ability of the cap1 or 2 oligonucleotides to detect the ORF1T oligonucleotide by dot blotting.

10, 20 or 100 ng of the oligonucleotides cap1 or cap2 (Table 3.2 and Figure 3.1) were spotted on a nylon membrane and then incubated with 1 µg of both the ORF1T oligonucleotide and either a biotinylated version of cap2 (with cap1 was spotted) or cap1 (with cap2 was spotted) (Figure 3.7) to assess if these probes were able to form a bridge with the synthetic ORF1T fragment from the CSSV genome.

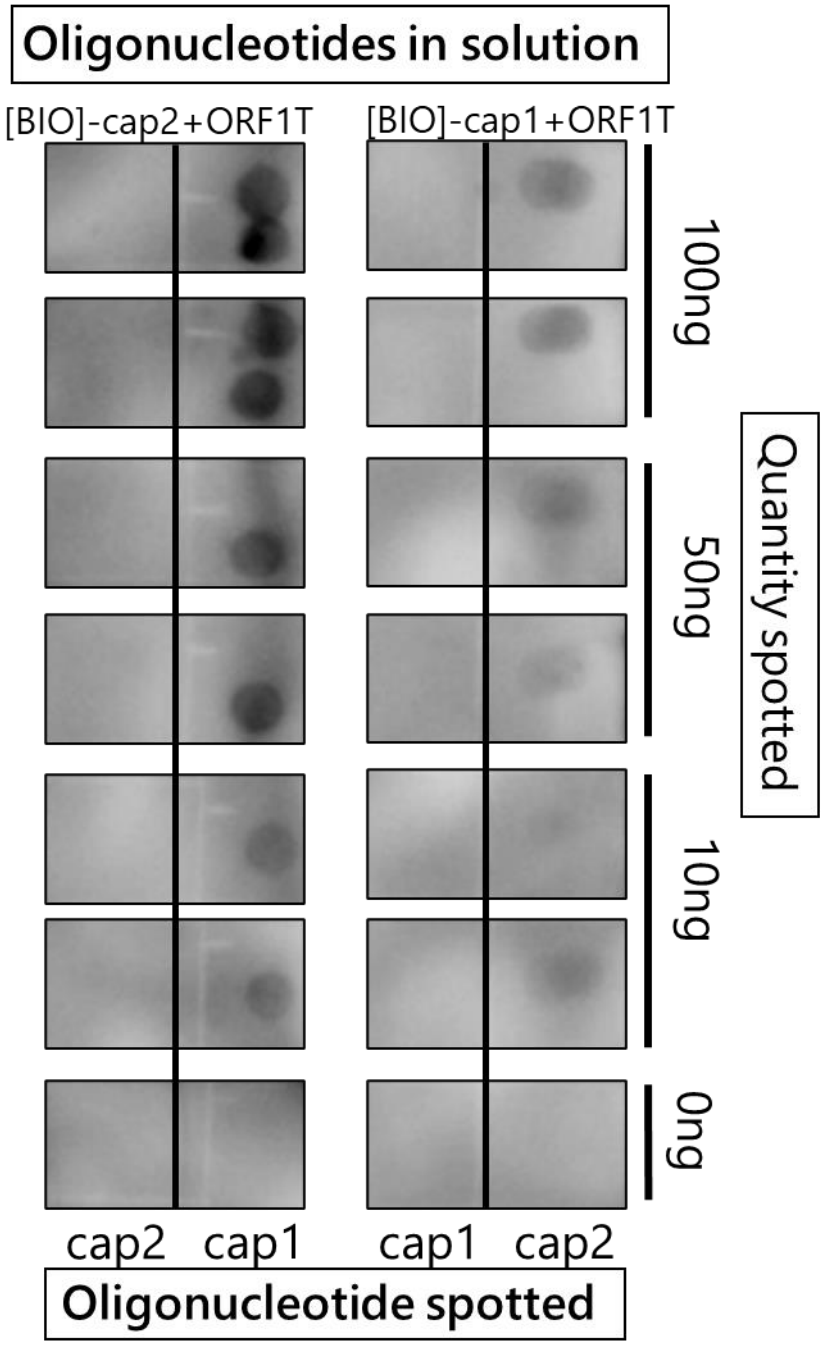


Figure 3.7: Dot blotting using either biotinylated cap1/ORF1T or biotinylated cap2/ORF1T in solution.

Decreasing amount of oligonucleotide (0, 10, 50 or 100ng) was spotted and incubated with a constant concentration of biotinylated probe/ORF1T (1µg of biotinylated probe + 1µg ORF1T). Negative spots were incubated with buffer only. The signal detected on the membrane represents the lattice presented in Figure 3.1.

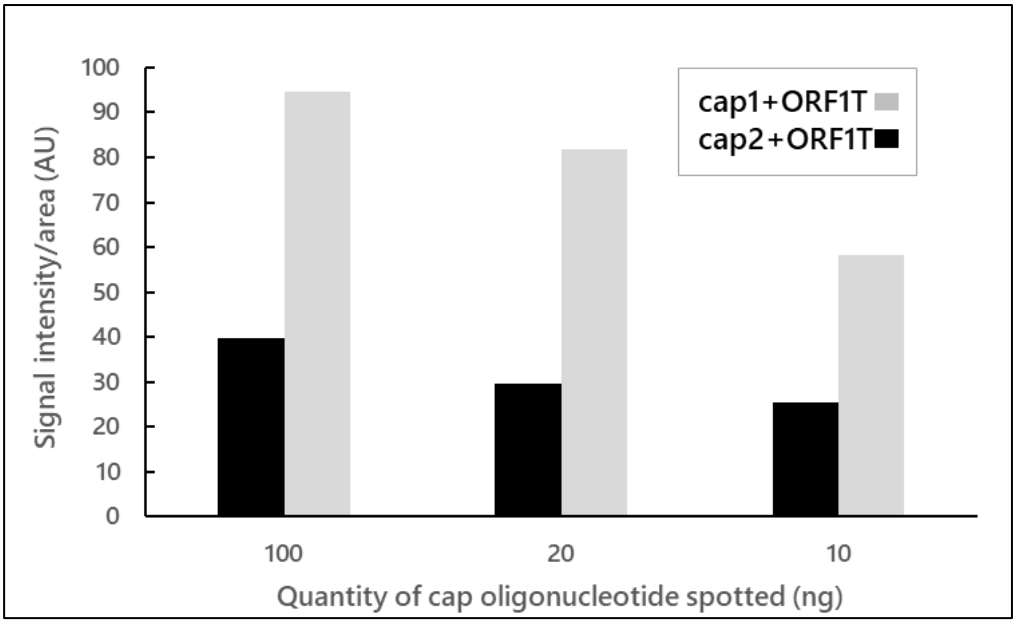


Figure 3.8: Dot blot signal intensity measurement derived from Figure 3.7.

When cap1 was spotted, biotinylated cap2 in solution was used to detect ORF1T. When cap2 was spotted, cap1 was used in solution to detect ORF1T. Pixel density was divided by the surface covered by each spot.

At the concentrations tested, both immobilised, capturing probes, cap1 and cap2, successfully detected the ORF1T sequence present in solution, allowing the formation of a bridge with the biotinylated oligonucleotide used as a detection probe: biotinylated cap2 when cap1 was immobilised and biotinylated cap1 when cap2 was spotted on the membrane (Figure 3.7). Control reactions were designed to assess the specificity of the reaction by spotting cap1 on the membrane and then incubating the membrane with a mixture of biotinylated cap1 and ORF1T. Similarly, cap2 was spotted on the membrane, and the membrane was then incubated with a solution containing a mixture of biotinylated cap2 and ORF1T. Both control experiments yielded background signals. Furthermore, negative control spots containing only buffer (0 ng, Figure 3.7) did not produce signals above the background levels, regardless of the biotinylated probe (cap1 or cap2) used for detection.

When 100 ng of cap1 was spotted, the signal to surface ratio was 94.7 while the signal to surface ratio with cap2 spotted was 39.6. At 20 ng, the values were respectively 81.86 and 29.68 and at 10 ng, 58.15 and 25.15, showing a consistently higher signal when using cap1 was used as an anchoring probe and cap2 as a detection probe at the concentrations tested (Figure 3.8).

3.3.4 HRM analysis of the biological recognition element.

3.3.4.1 Monitoring of the base-pairing formation between ORF1T and ORF1TC.

Equimolar amounts (2 µM) of the ORF1T and ORF1TC oligonucleotides were mixed in either DNase/RNase free water or in Annealing buffer. Theoretically, oligonucleotides cannot bind to one another in water. The annealing buffer was formulated to control pH and ionic conditions favourable to the formation of a double-stranded DNA molecule. Samples containing either ORF1T or the complementary fragment ORF1TC alone in the two solutions tested were also analysed by HRM as single-stranded controls (Figure 3.9). When the ORF1T plus ORF1TC samples were placed in the annealing buffer that contains 50 mM of NaCl, a steady, progressive decrease of the fluorescence was observed as the temperature increases, until no fluorescence could be detected. In contrast, the fluorescence of the same fragments analysed in water decreased sharply soon after the temperature started rising.

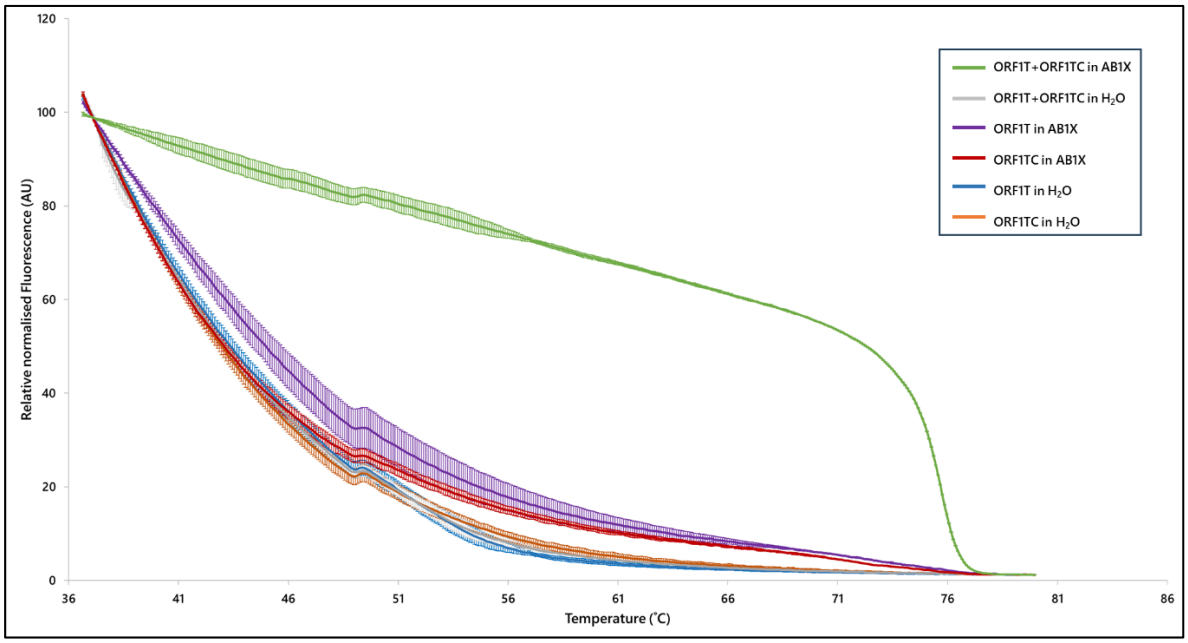


Figure 3.9: Effect of the buffer composition on the fragments thermostability.

Normalized high resolution melting curve of the 76bp oligonucleotide double-stranded ORF1T+ORF1TC fragments or single-stranded ORF1T or ORF1TC in either annealing buffer (AB1X) or water (H20). ORF1T+ORF1TC placed in H_2O were unable to anneal to each other and their melting profile were comparable to samples containing single oligonucleotide. Each curve is an average of three samples, represented with standard deviation bars.

Samples containing only one type of oligonucleotide (ORF1T or ORF1TC alone in water or annealing buffer) exhibited a comparable behaviour to the ORF1T plus ORF1TC samples in water, with a sharp decrease of fluorescence observed as soon as the temperature increased in the medium. Principal component analysis (PCA) and sample cluster analysis (SCA) using the ScreenClust software was run to verify if samples would cluster in different groups to reflect their structure. A first run was performed in unsupervised mode with a cluster count set to 2 and the discrimination of the samples done by the software (Figure 3.10). ORF1T and ORF1TC samples in annealing buffer clustered separately (Blue circle on Figure 3.10) from all other samples which grouped all together (P=1, Typicality=0.513) both above the acceptance threshold set for the experiment in section 3.2.7.

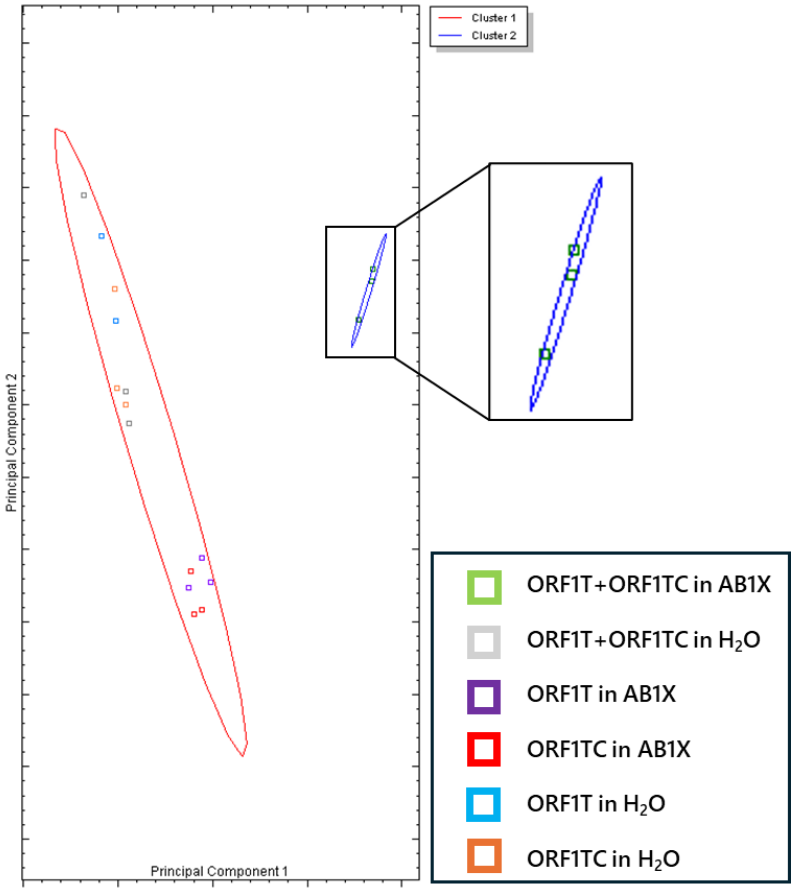


Figure 3.10: Unsupervised Principal Component Analysis (PCA) of the HRM run presented in Figure 3.9, using the software Screenclust.

The cluster count was manually set to two by the user prior to running the PCA, to reflect the predicted two type of fragments: either single or double stranded. Double stranded DNA samples in annealing buffer clustered separately to all other samples. Each square represents single replicates from the HRM run.

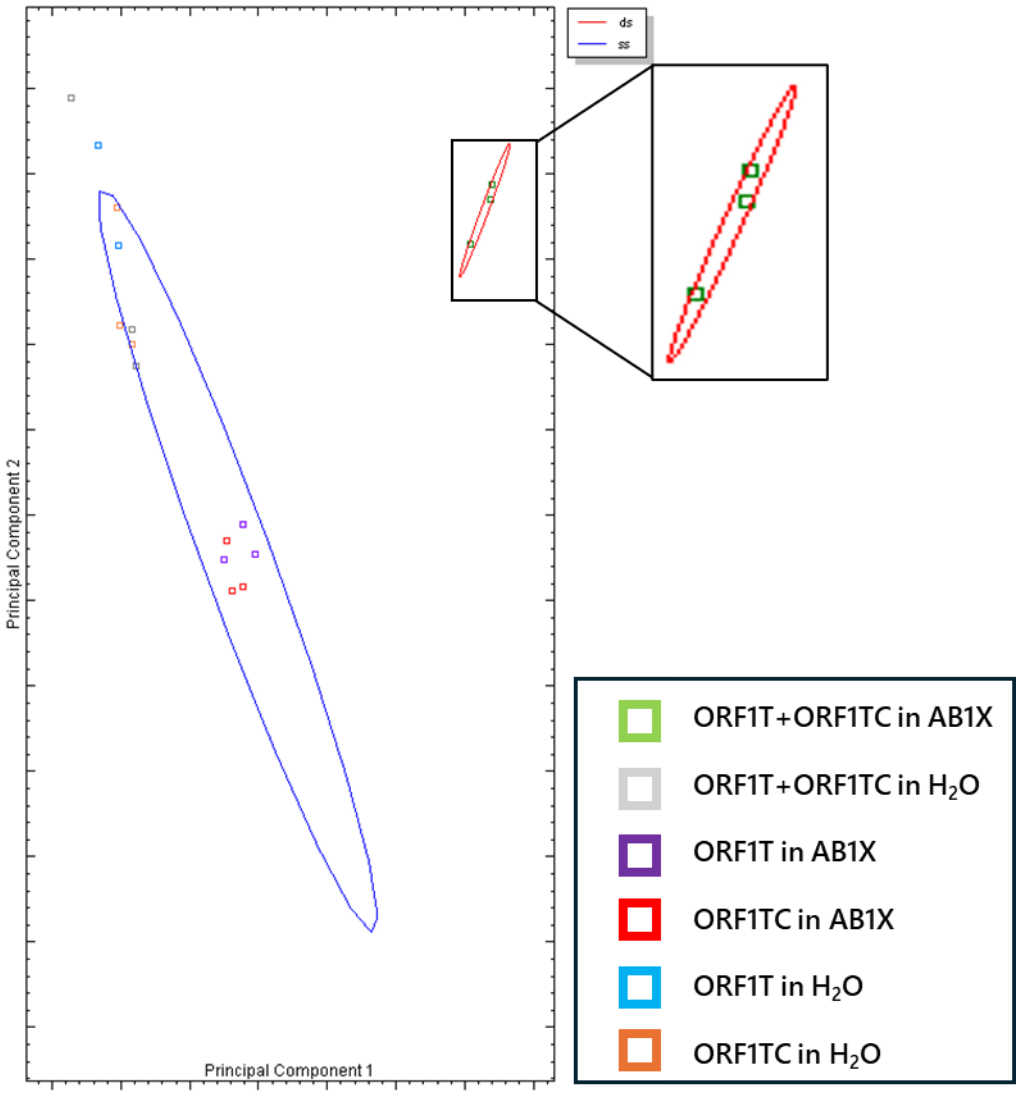


Figure 3.11: Supervised PCA of the HRM run represented in Figure 3.9, using the software Screenclust.

Samples were manually placed in either single stranded fragment or double stranded fragment clusters (Blue circle: Single stranded fragments, red circle: Double stranded fragments).

When the PCA was performed in supervised mode (Figure 3.11), all samples were associated with one of two specific groups. All samples except ORF1T plus ORF1TC in Annealing Buffer were placed in one group labelled single-stranded while ORF1T plus ORF1TC in Annealing Buffer was placed in the double-stranded DNA group. This PCA generated the same results as the first analysis, with double stranded fragments in annealing buffer found to cluster separately from the remaining samples, with a probability equal to 1 and a typicality of 0.53, both above the acceptance threshold set for the experiment in section 3.2.7. Both analyses also demonstrated that double stranded and single

stranded fragments in water were not distinct in clustering confirming that oligonucleotide could not hybridize in water (Average typicality of single stranded fragments was 0.454, P=1; Average typicality of double- stranded fragments was 0.633, P=1 in Unsupervised mode and respectively 0.445, P=1 and 0.557, P=1. In Supervised mode).

3.3.4.2 Evaluation of the bonding strength between the different oligonucleotides of the lattice.

Oligonucleotides forming the lattice for use as biological recognition element were mixed in annealing buffer with the Syto9 dye to investigate their interaction by HRM.

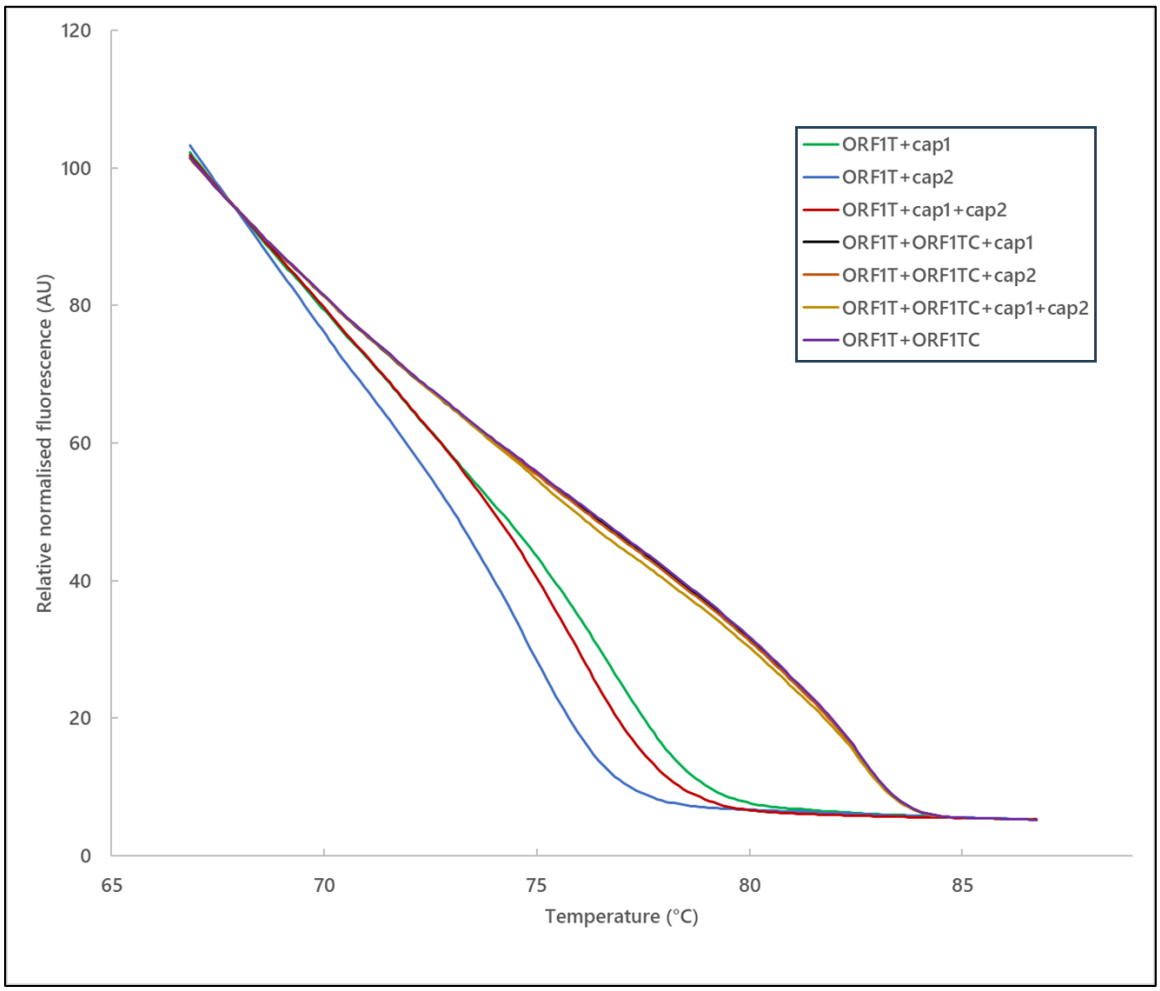


Figure 3.12: HRM curves analysis of the oligonucleotide lattice in AB1X.

A set of duplicated samples containing combinations of ORF1T, ORF1TC and cap1 and cap2 fragments were prepared, annealed and their melting curves produced. The fluorescence values of the normalised curves of two replicates were plotted against the temperature. The difference in thermostability of the fragment combination is associated with decreasing melting temperatures, with ORF1T+ORF1TC being the most stable fragment combination.

The strength of the interaction between the oligonucleotides was tested by analysing the melting profile of combinations of oligonucleotides by HRM (Figure 3.12). The ORF1T sequence was mixed with either the cap1 probe or with the cap2 or with both. Another group was prepared by mixing the ORF1T and the ORF1TC sequence with either cap1 or cap2 and an additional one containing both.

A group of curves formed that contained ORF1T associated with the shortest oligonucleotides: ORF1T/cap2, then ORF1T/cap1/cap2 and ORF1T/cap1, with lower melting temperatures, under 77°C. The samples containing the ORF1T sequence with cap 2 exhibited the lowest melting temperature of all samples tested (Figure 3.12).

Fluorescence measurements between 75 and 80°C (Figure 3.13) confirmed the order in which the sample combination melted.

Sample composition	Average fragment retained fluorescence between 75-80°C (AU)
ORF1T+cap2	11.93
ORF1T+cap1+cap2	18.07
ORF1T+cap1	21.88
ORF1T+ORF1TC+cap1+cap2	42.25
ORF1T+ORF1TC+cap2	43.36
ORF1T+ORF1TC+cap1	43.64
ORF1T+ORF1TC	43.94

Figure 3.13: Fluorescence retained within the oligonucleotide assemblies between 75 and 80°C. 50 fluorescence data measurements from the melting curves were averaged in the temperature range between 75 and 80°C for each sample.

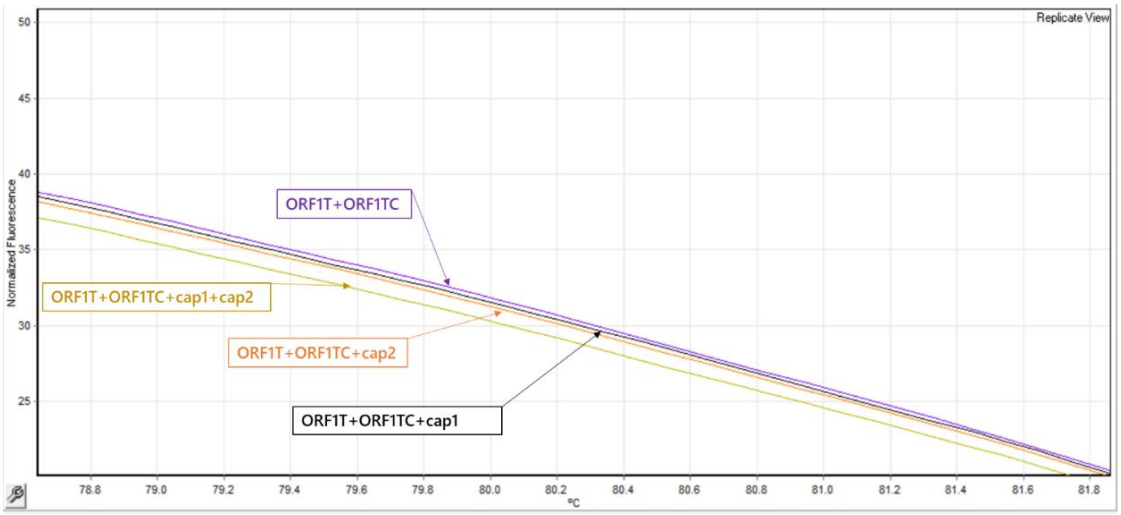


Figure 3.14: Detailed view of the highest melting temperature group of curves from the HRM run presented in Figure 3.12.

The arrangement of the melting temperature curves is displayed between in a region of the melt between 78.8 and 81.8°C. ORF1T+ORF1TC was the most thermostable combination of oligonucleotides.

Samples containing ORF1T and ORF1TC alone or supplemented with either cap1 or cap2 were in a group displaying the highest melting temperature of all samples tested, above 82°C (Figure 3.14).

3.3.4.3 Analysis of a competition between cap1 and cap2 and different proportions of ORF1TC to bind to ORF1T by HRM.

To assess the amount of capture element required to successfully immobilise the CSSV target sequence, a competition experiment was run. The aim of this experiment was to evaluate the proportion of the extra complementary sequence required to displace the equilibrium from the two large fragments (ORF1T/ORF1TC) to obtaining a profile similar to what is observed with ORF1T+cap1+cap2. When this point would be reached, this would mean that most of the fragments bound to ORF1T are the detection probes cap1 and cap2, rather than ORF1TC and would place the samples in conditions which would replicate a CSSV assay, where the detecting probes would compete with the complementary sequence of the region of interest to bind to the targeted region.

This was assessed by preparing samples containing all fragments (ORF1T, ORF1TC, cap1, cap2) in equimolar proportions at 2 µM and then comparing them to samples containing respectively 2 (1 µM), 5 (0.4 µM), 10 (0.2 µM), 50 (0.04 µM) or 100 (0.02 µM) times less ORF1TC fragment. Reference mixes included samples of known profile like ORF1T plus ORF1TC in equimolar proportions and ORF1T with

cap1 and cap2. The HRM profiles of the three groups of oligonucleotide mixture were compared (Figure 3.15).

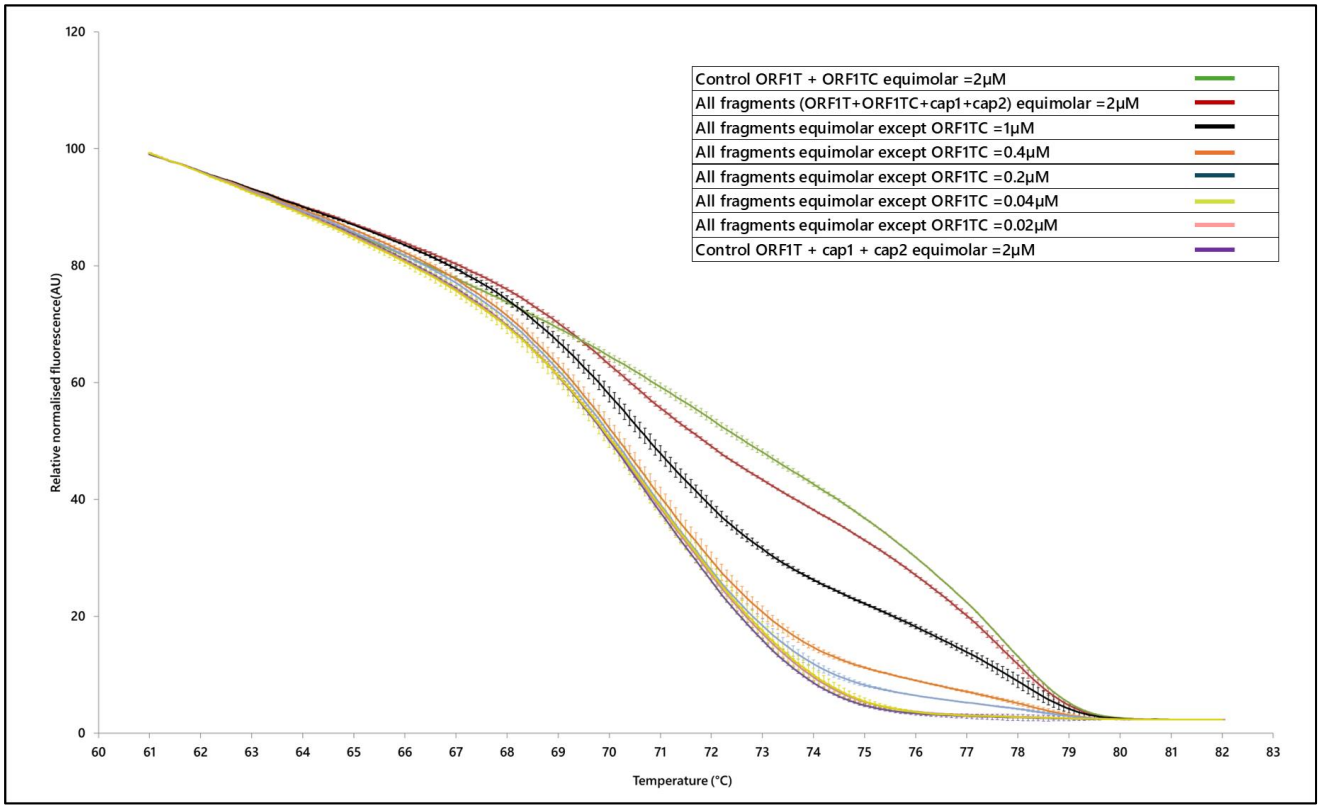


Figure 3.15: High resolution melt curves of competition experiment between ORF1TC and the detection probes cap1 and cap2.

As the ORF1TC concentration in samples containing ORF1T+ORF1TC+cap1+cap2 was progressively reduced, the melting curve profile shifted from the most stable green curve (ORF1T+ORF1TC) to the least stable purple curve (ORF1T+cap1+cap2). The samples are presented in the legend box by decreasing melting temperature. Normalised melting curves were calculated from four replicated samples and are displayed with standard error bars.

The sample group containing ORF1T + ORF1TC in equimolar proportions exhibited the highest melting temperature. The decrease of the slope of the corresponding curve, which reflects the emission of fluorescence, was the smallest among all the samples tested. This indicates that the ORF1T+ORF1TC samples had the highest resistance to heat denaturation. In contrast, the samples containing ORF1T+cap1+cap2 had the lowest melting temperature. When cap1 and cap2 were added to ORF1T+ORF1TC, the overall curve was closer to the ORF1T+ORF1TC samples.

As the proportion of ORF1TC in the mixture was gradually reduced 2, 5, 10, 50 and 100-fold, the melting temperature shifted accordingly towards the profile observed for ORF1T+cap1+cap2. With the

complementary sequence diluted 2, 5 or 10-fold, intermediate melting temperatures were observed. At 50 and 100-fold dilutions, the curves merged with the ORF1T + cap1 + cap2 curve, showing that at this concentration, most of the ORF1T sequence was bound to cap1 and 2 rather than to ORF1TC.

This means that in an assay using CSSV-infected *T. cacao* extracts, the concentration of the capturing probes cap1 and cap2 should be set to be 50 times greater than the ORF1T targeted sequences originating from the infected plant samples. This proportion would allow the detection of most of the targeted region of interest of CSSV present in the samples.

This was confirmed by a PCA run in unsupervised mode (Figure 3.16): samples where the complement was diluted 50 or 100-fold all segregated with the ORF1T+capture1+capture2 (Cluster3, Green circle), with an average typicality respectively of 0.86 and 0.48.

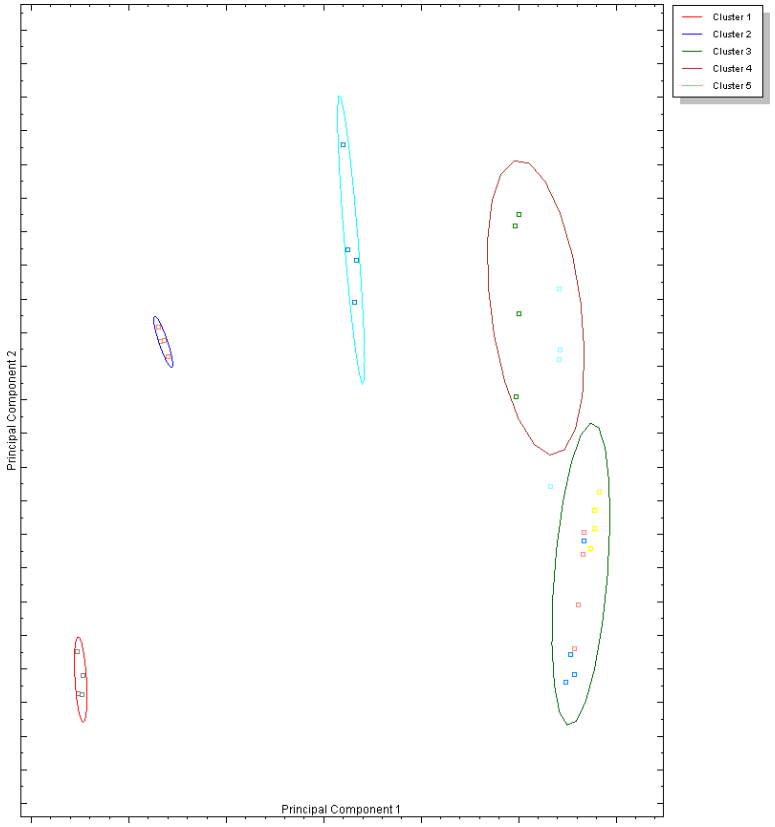


Figure 3.16: Screenclust unsupervised PCA of the competition HRM run from figure 3.15.

Cluster number was set to 5. ORF1T+cap1+cap2 samples (yellow squares) and All fragments+ORF1TC/50 (red squares) and All fragments+ORF1TC/100(deep blue) samples segregated in the same cluster (Cluster 3, Green).

3.3.4.4 Chemical-mediated oligonucleotide lattice structure modification analysis by HRM.

- Sodium hydroxide-mediated denaturation of the oligonucleotide lattice analysis by HRM.

Sodium hydroxide was assessed as a means of manipulating the structure of the oligonucleotide lattice. A range of sodium hydroxide concentration (0, 4, 6, 8, 10, 11 or 12 mM) were added to samples containing ORF1T, ORF1TC, cap1 and cap2 in equimolar amounts and analysed by HRM (Figure 3.17).

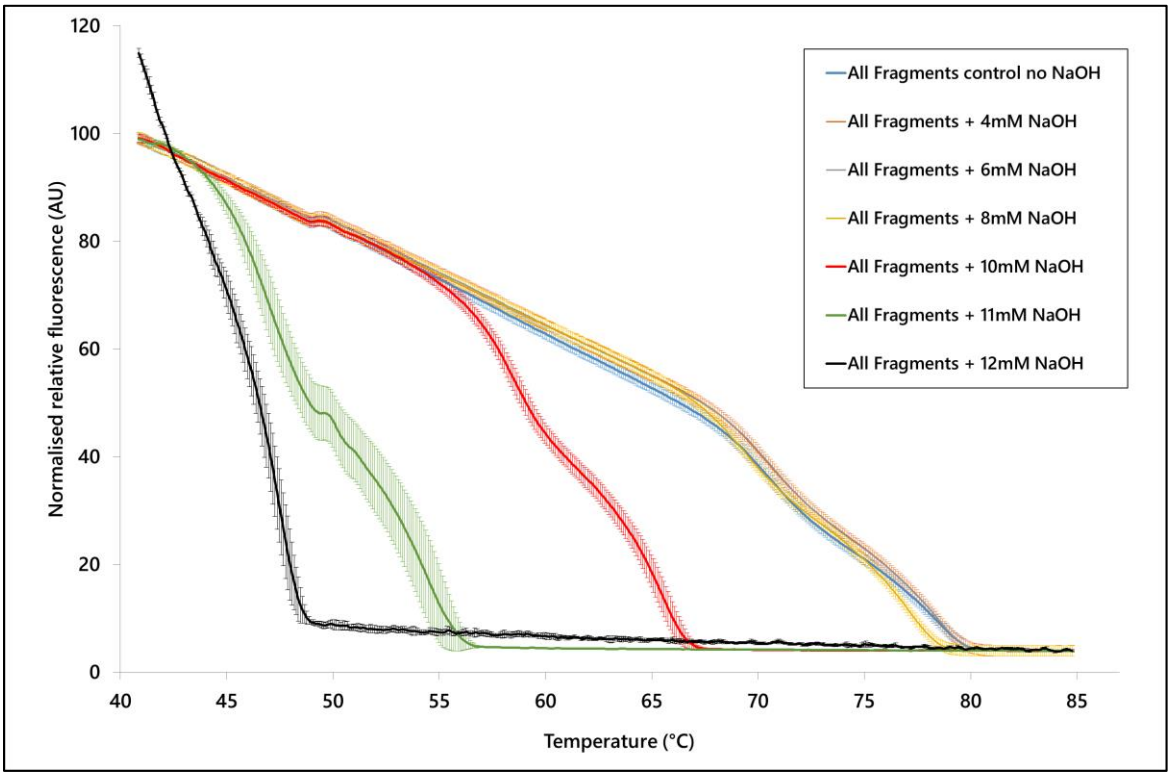


Figure 3.17: Effect of sodium hydroxide on the oligonucleotide lattice thermostability.

Equimolar sample mixtures of all fragments containing ORF1T+ORF1TC+cap1+cap2 were incubated with either 4, 6, 8, 10, 11 or 12 mM sodium hydroxide and placed in the Rotor-Gene Q to assess the thermostability of the oligonucleotide lattice to the chemical. The thermostability of the lattice decreased when the concentration of sodium hydroxide reached 10 mM (red curve). Curves are normalized melting curves of four replicates with standard error bars.

Several groups were observed based on the offset from the control curve. The melting curves when the DNA was exposed to 4, 6 or 8 mM NaOH were superimposed to the untreated control, with the 8 mM NaOH curve diverging slightly when the temperature was reaching 80°C. When the concentration of sodium hydroxide in the medium was set to 10 mM, an offset was observed, and the melting temperature of the lattice was reduced from 80˚C to 67˚C. A further decrease of fluorescence was then observed at the higher NaOH concentrations of 12 mM. The profile for this melting curve was

comparable to what was observed when single stranded DNA molecules are analysed as shown on Figure 3.9, suggesting that either the fluorescence was immediately lost or that the double-stranded conformation could not be achieved under these conditions.

A PCA analysis was set to verify if the differences observed between the melting groups were significant (Figure 3.18).

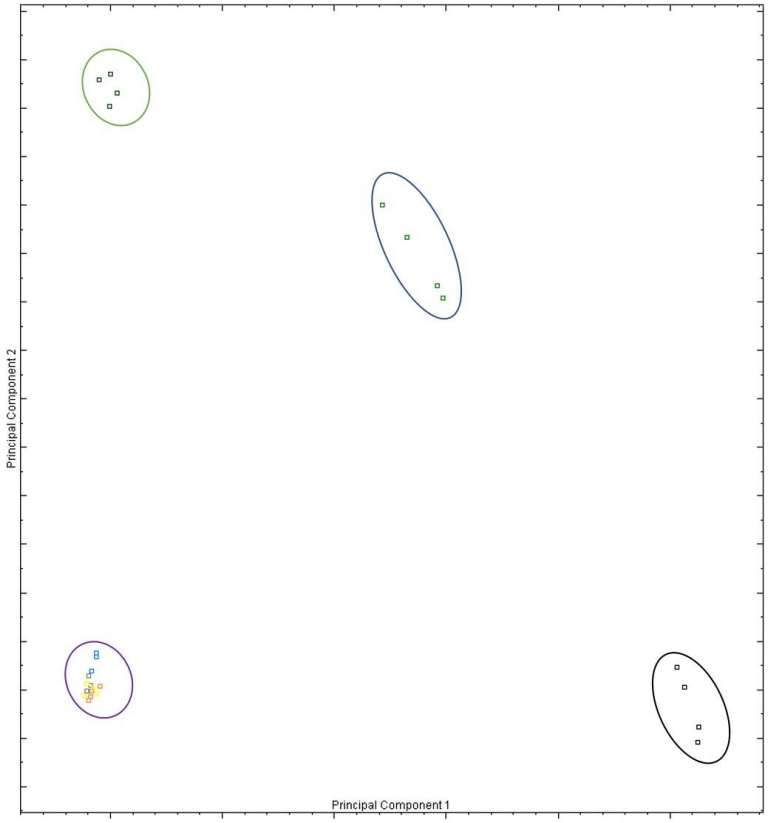


Figure 3.18: Screenclust Supervised PCA of sodium hydroxide-treated samples.

A cluster was set to contain untreated control samples. Black circle contains samples treated with 12 mM of sodium hydroxide; Blue circle: 11 mM, green circle: 10 mM and purple circle: All remaining samples (0, 4, 6 and 8 mM). The clustering observed matched the melting curves presented in Figure 3.17.

The analysis was performed in supervised mode, with one cluster allocated to control samples not exposed to NaOH. Samples containing either 0 (typicality=0.999), 4 (typicality=0.989), 6 (typicality=0.994) or 8 mM NaOH (typicality=0.991) all separated in this same cluster (Purple circle, Figure 3.18). The other groups of samples grouped separated in clusters according to the concentration of sodium hydroxide used. The clustering therefore reflects the grouping observed on the melting curves (Figure 3.17).

3.3.4.4.1 Denaturation and renaturation of the oligonucleotide lattice using sodium hydroxide and acetic acid.

Acetic acid was used to renature sodium hydroxide-denatured oligonucleotides mixture. The inter-strand bond between the fragments in solution was disrupted by adding 100 mM sodium hydroxide to the samples, followed by addition of a 2.5 N acetic acid solution to buffer the medium and therefore re-enabling the formation of hydrogen bonding between the complementary strands. This was achieved by adding the cap1 oligonucleotide to the ORF1T and ORF1TC mixture in equimolar proportion, to mimic the configuration of the nucleic acids on the surface of the biosensor. High resolution melting curves of the same samples exposed to either a high concentration of sodium hydroxide or a buffered medium were produced using the Rotor-Gene Q (Figure 3.19).

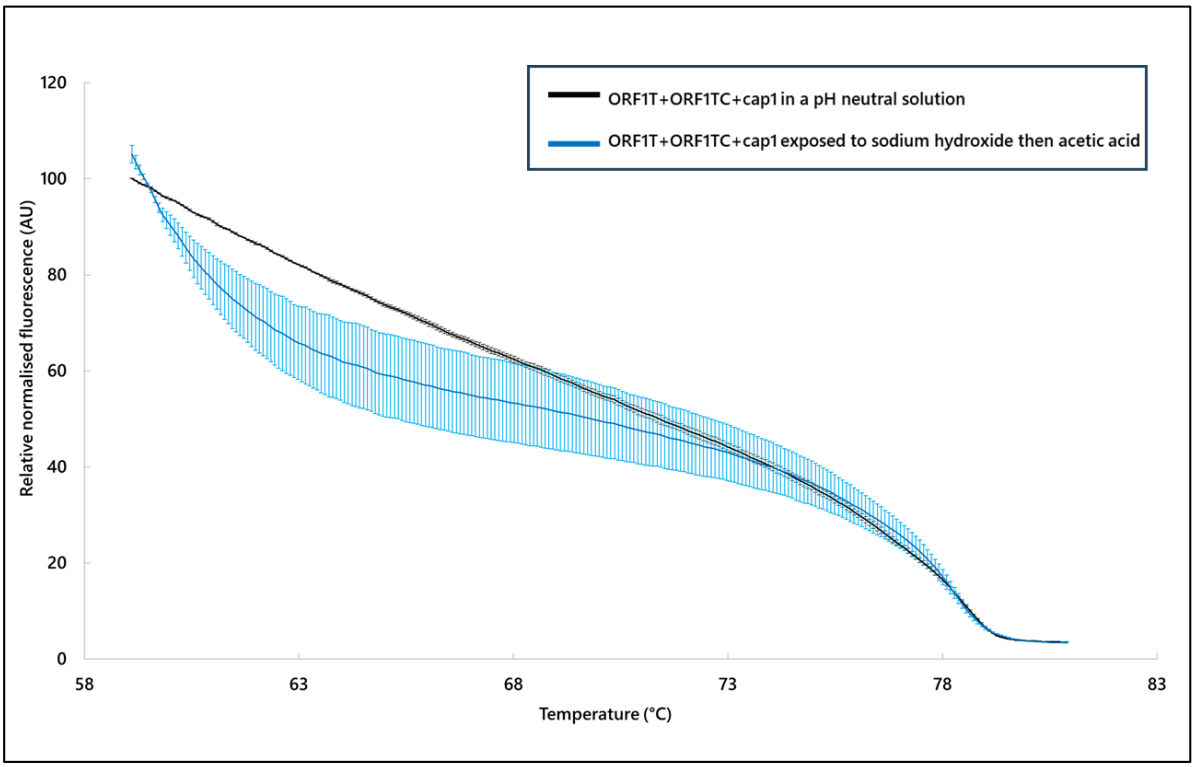


Figure 3.19: Denaturation-renaturation of ORF1T+ORF1TC with sodium hydroxide and acetic acid.

High resolution melt curves of ORF1T+ORF1TC+cap1 placed in a pH neutral solution or incubated with sodium hydroxide (100mM) and acetic acid (2.5N). A combination of sodium hydroxide and acetic acid can be used to denature then renature the oligonucleotide lattice with a limited impact on the melting temperature. The curves presented are normalized melting curves of four replicates with standard error bars.

The two set of samples that were analysed exhibited distinct melting curve shapes. The melting curve replicates of the samples exposed to a sodium hydroxide and buffered with acetic acid were more

variable as demonstrated by the higher value of the error bars than the samples placed in a pH neutral solution, and the fluorescence was lower before reaching 68°C. Both type of samples melted at the same temperature. This shows that the sodium hydroxide-denatured fragments were able to re-anneal one to another and melt at the same temperature.

3.4 Discussion and conclusion

The genome of the CSSV isolate maintained at the envirotron at UWE was sequenced *de novo* using eighteen primer pairs. This step was performed to ensure that the sequence used to design the nucleic acid biological recognition element part of the biosensor was the same as the sequence of the virus infecting *T. cacao* trees used as samples. This was verified and alignments of the eight genome sequences available at the time of the study identified the ORF1 of CSSV as a suitable region of interest for detection of the virus using a biosensor. Numerous genome sequences have since been published (Muller *et al.*, 2018; Chingandu *et al.*, 2017) and have deepened the knowledge of the genetic diversity and variability amongst the different CSSV isolates. The 76 pb region of interest, ORF1T, selected for this study enabled the development of a prospective BRE based on CSSV's nucleic acid sequence and further adjustments to increase or decrease the specificity of the sensing platform would not be a labour-intensive endeavour. Indeed, procuring DNA oligonucleotides to match the different strains currently circulating in West Africa but with no close match in *T. cacao* genome would not be an issue.

Two probes deriving from the *de novo* sequencing of ORF1T, cap1 and cap2 were designed as the biological recognition elements of the sensor. Biotinylated versions of the probes were synthesized and tested with the synthetic targeted viral region ORF1T and its complementary sequence ORF1TC in a dot blot assay, to assess the ability of the probes to detect a region of interest of CSSV onto the surface of a membrane. To do so, the two short oligonucleotide probes covering the targeted sequence were synthetized and spotted on a nylon membrane. Then, a bridge between the immobilised anchoring sequence and the detection probe was created by incubating the ORF1T sequence together with the probes. Regardless of the probe used, it was possible to detect ORF1T

with this system. Densitometry analysis of the dot blot (Figure 3.8) revealed that when cap1 is used as the capturing element (as opposed to being the probing element), the signal obtained was more intense than when cap2 was used for this purpose. This difference in the capturing abilities of the probe could be explained by the GC-content of the two probes, respectively 50% and 42.1% for cap1 and cap2 since GC bonds are more stable than AT bonds. The experiment proved the specificity of the binding of both probes to ORF1T, in a dose-dependent manner. Both GC content and probe/target proportion could be a factor in the decision-making process of selecting a probe to include in the biosensing platform. The apparent differences in binding efficacy between cap1 and cap2 probes should however be confirmed in further development steps, to ensure that the difference observed between these probes to detect ORF1T can be replicated in the assay conditions chosen for the sensor. The dot blotting technology is well established (Stott, 1989) and is a robust method to assess if oligonucleotide fragments can bind to their specific target. However, it is a lengthy procedure and is therefore not ideal for developing a cold hybridization system. Instead, an approach using high resolution melt analysis was implemented. Indeed, HRM run times are shorter and less labour intensive than dot blot experiments. Furthermore, the samples are analysed in aqueous solution which makes the regulation of the pH easier.

The candidate, oligonucleotide-based, biological recognition element study presented in this chapter was further analysed using HRM analysis. HRM was repurposed as a tool to characterize the formation of an oligonucleotide lattice between the cap1 and cap2 probes and the ORF1T fragment of CSSV. This methodology was used to assess the stability of the lattice, explore the ratio of capture/detection probe needed to immobilise the target sequence and to monitor the controlled denaturation and renaturation of the oligonucleotides. First, the physicochemical conditions enabling the formation of double stranded DNA molecules were verified by incubating complementary fragments in either water, that will not enable inter-strand bonding formation, or an annealing buffer where the nucleotides can bind one to another. As a control, single stranded samples were incubated in the same experiment. Double-stranded samples in water produced a melt curve similar to single stranded,

demonstrating that they were not able to associate in these conditions. In contrast, the same fragments placed in annealing buffer created a typical melting curve, specific to the oligonucleotide association, where the fluorescence dye is release progressively and the signal is lost at a higher melting temperature (Figure 3.9). Tan and Chen, (2006) demonstrated that ions have a crucial role in nucleic acid stability. Indeed, annealing experiments performed in water exhibited profiles of single-stranded molecules when compared to samples placed in a buffer saline solution containing the ions allowing the thermodynamic arrangement of the DNA as double-stranded helix (Figure 3.9). Further experiments aimed at investigating the dynamic of interaction between the region of interest and the oligonucleotide to be used as anchoring or detection probes were caried out. Subtle differences between the fragment's association strength were observed using HRM (Figure 3.12). The association of the two large fragments, ORF1T+ORF1TC sample group exhibited the highest stability of all the samples group, which is the result expected provided that these sequences could form the most bonds, and therefore possessed the highest thermostability. In contrast, the samples containing ORF1T and cap1+ cap2 released the fluorescence dye first. This data is consistent with the dot blot assay result, where cap1 provided a stronger signal than cap2 (Figure 3.7). The difference observed between the combination of long fragments ORF1T/ORF1TC and ORF1T with the shorter cap1 and cap2 can be explained by the nick present in the middle of the double-stranded molecule that occurs when the two probes are bound to ORF1. While the sequence of the two types of samples was identical, the resulting thermostability was significantly affected (Figure 3.13) and HRM analysis could account for this structural change information.

Experiments were conducted to quantitatively assess the dynamic of interaction between the region of interest ORF1T, ORF1TC and the two probes cap1 and cap2. In the context of a sensor analysing real-life samples, a competition would occur for binding to ORF1T between the naturally occurring complementary sequence ORF1T and cap1+cap2. Knowing the respective melting curve profiles of all combination of fragments, HRM was utilised to determine the ratio of competing ORF1TC required for shifting the profile from a two long fragment shape to one long fragment with two smaller fragments

profile. Analysis of samples containing a range of dilution of ORF1TC revealed that when the complementary fragment was diluted 50 and 100-fold compared to the other fragments, the curve overlapped with the curve of the control samples containing ORF1T with the probes cap1 and cap2 (Figure 3.15). These theoretical values would need to be confirmed since discrepancies can be observed between the ideal synthetic samples and actual viral DNA plant samples. For example, Eun *et al.* (2002) reported a limit of detection of 1 ng of purified nucleic acid from two orchid viruses which was increased to 10 ng of crude sap when using a piezoelectric DNA sensor.

The HRM approach was further used to investigate a chemical-mediated denaturation and renaturation of the oligonucleotide biological recognition element. Using chemicals to manipulate the DNA structure could replace temperature-based denaturation/renaturation and therefore reduce the cost of the sensor. HRM experiments were performed to determine the precise amount of sodium hydroxide necessary to disrupt the oligonucleotide lattice (Figure 3.17). 4, 6 or 8 mM did not modify the melting temperature of the lattice significantly. However, when the concentration of NaOH reached 10 mM, a 10˚C difference was observed compared to the control, and when the concentration was set to 11 mM, a further 10˚C melting temperature was measured. At 12 mM, the fluorescence dropped to zero almost instantly, showing that full denaturation of the fragments was obtained. This result contrasts with previous observations (Wang, Lim, and Son, 2014), where 10 mM sodium hydroxide denatured only 5% of the test fragment throughout the experiment. In their study, the author used a range concentration of NaOH (10, 100 and 1 M) amongst other methods to denature an 86bp PCR product at room temperature and analysed the optical density using a NanoDrop UV-Vis spectrophotometer. The discrepancy observed could be caused by the difference in the respective experimental procedure. In this chapter, the DNA fragments analysed by HRM are heated, which reduce the thermostability of inter-strand bonding while in the spectrophotometer experimentation, measurements are performed at room temperature where the fragments are more stable and less sensitive to chemical denaturation. Furthermore, spectrophotometry may not be the adequate means to accurately measure subtle changes in the oligonucleotide lattice structure: experiments performed

to observe the hyperchromicity of the single stranded fragments ORF1T or ORF1TC were inconclusive (Data not shown).

Another key aspect of the study was to assess if it was possible to reverse the sodium hydroxide-mediated denaturation of a mixture of DNA fragments mimicking the conformation of the biological recognition element of the CSSV biosensor. To do so, DNA samples containing the ORF1T, ORF1TC sequence and cap1 fragments were placed in a high concentration of sodium hydroxide (0.1 M), then acid acetic (2.5 N) was added to neutralise the pH. The subsequent HRM analysis showed that although the curve had a different shape compared to the untreated sample, the melting temperature remained unchanged (Figure 3.19). This shows that the denaturation of the lattice in these conditions is reversible since the fragments can re-anneal after having been exposed to a high concentration of sodium hydroxide for 5 minutes. It is known that exposure to sodium hydroxide can induce irreversible denaturation of the DNA (Ageno, Dore and Frontali, 1969a) but the conditions tested here did not seemingly damaged the nucleic acid fragments. This result is important for the development of the sensor. Sodium hydroxide will be tested in the subsequent chapter to extract the DNA of CSSV from field samples. The protocol to prepare the samples would therefore be simplified. Based on this data, an alkali extraction buffer could be used, allowing both the lysis of the plant cells and denaturation of the viral sequence in a single step, enabling their hybridization to an immobilised capturing probe. Then, a neutralising solution based on acetic acid could induce the annealing to the probe.

Chapter 4 Evaluation of the DNA BRE to detect CSSV from *Theobroma cacao* samples.

4.1 Introduction.

The interaction between the oligonucleotides cap1, cap2 and the synthetic targeted region of interest within CSSV genome ORF1T and ORF1TC was studied using high resolution melt analysis in the previous chapter. Using this technique, it was proven that sodium hydroxide and acetic acid could be used to manipulate the interaction between the oligonucleotides forming a lattice that could be used as a biological recognition element of a biosensor targeting CSSV. A PCR-based detection requiring rapid and accurate temperature changes would be challenging to transpose in the field. While an isothermal assay that was developed by the company Swiss DeCode (Swiss DeCode, 2021) is now available, the technology relies on the amplification of a section of CSSV genome. It was however recently discovered that CSSV sequence fragments could be found integrated within *T. cacao* genome (Muller *et al.*, 2021b). Named eTcBV for endogenous *T. cacao* bacilliform virus, the sequences are present in asymptomatic trees, and genetic analysis suggests that the integration occurred prior to the export of the tree in other continents. This raises the possibility of false CSSV positive results on asymptomatic trees wherever they are grown with assays relying on sequence amplification. A membrane-based biosensor targeting the nucleic acid of the virus could therefore mitigate the potential false positive detection. Furthermore, hybridisation assay relies on the detection of actual copy numbers of a targeted sequence of interest. In a CSSV-infected tree, actual viral copies of a targeted candidate sequence would outnumber a putative integrated sequence that would be present as two copies per cell in the case of a single insertion event. This means that a probe-based hybridisation approach would be less likely to cause a false positive signal. Assay relying on a membrane hybridisation could also feature the cost and limited personnel training requirements. A

“cold” membrane-based assay involving chemicals to manipulate the structure of the double-stranded DNA molecules and enable binding of the labelled oligonucleotide to the targeted viral sequence could replace temperature-induced denaturation and renaturation of the nucleic acid.

The first challenge to the development of such a biosensor is the preparation of the sample that will be introduced to the platform by separating the nucleic acids from the other cell components. CSSV DNA extraction is currently performed for the purpose of fundamental research and relies primarily on the utilisation of commercial kit designed for plant DNA extraction. DNA extraction protocols solely targeting the virus are not likely to be applicable in a field set up since they require extensive laboratory work, involving enrichment steps to concentrate the viral particles (Xie and Hu, 1995). Besides, most of the DNA extraction protocols in the literature are designed to produce samples suitable for downstream molecular applications such as PCR. For these, the purity of the nucleic acid is critical for reliable DNA exponential amplification. These extraction methods tend to include complex protocols with often hazardous chemicals. Commercial kit procedures were developed to produce the best yield and obtain the purest DNA possible from plant samples and use the guanidine thiocyanate extraction created by Ulrich *et al.* in 1977. This detergent lyses the cells and denatures proteins, hence protecting nucleic acids from nucleases degradation during the extraction process (Boom, 1990). Typical DNA extraction methods combine a chaotropic agent such as a detergent and alkaline lysis to destabilise the cellular structure of the plant cells and create physicochemical conditions where the nucleic acids are selectively separated from the other cell components. Sodium hydroxide in solution, by locally raising the pH of the extracted plant cells, creates the conditions to enable DNA extraction (Kashyap *et al.*, 2016). Furthermore, sodium hydroxide can break the hydrogen bonds between two DNA strands, creating single stranded molecules, which would enable binding of the probes to detect CSSV without resorting to temperature.

4.1.1 Aim and objectives of this chapter.

This chapter describes the feasibility of integrating the nucleic acid lattice described and characterised by HRM in chapter 3 as the biological recognition element to detect CSSV using T. *cacao* extracts in a temperature-independent assay.

Objectives will cover:

- The feasibility of extracting total DNA from CSSV-infected *T. cacao* samples. This is addressed by developing chemical and mechanical methods to disrupt plant tissues to collect both CSSV and *T. cacao* DNA. The quantity of the DNA extracted will be subsequently evaluated using qPCR. The method will focus on the leaves and branches of the tree.

- The ability of the biotinylated oligomeric probes developed in chapter 2 to detect CSSV from infected plant samples. This will be evaluated with a pull-down assay using streptavidin-coated magnetic particles to retrieve nucleic acid-bound probes.

- The development of a "cold" temperature-independent assay using the oligomeric probes to detect CSSV DNA. This will be tested with a membrane-based hybridisation assay.

- The ability of sodium hydroxide to denature the whole DNA for enabling the binding of oligonucleotide probes in samples from CSSV-infected *T. cacao* extracts will also be investigated using a pull-down assay and qPCR.

4.2 Material and Methods.

4.2.1 Material sampling.

T. cacao plants were kept at the UWE envirotron and were the source for all the material tested in this work. CSSV negative and CSSV positive plants were kept in distinct sections of the green house.

4.2.1.1 Leaf sampling.

Calibrated samples were cut from a single leaf, targeting a region of the lamina close to the midrib, using a coring tool to produce *T. cacao* discs of 1 cm diameter. Discs were weighed and 25 mg were placed in a 1.5 mL microcentrifuge tube and stored at -20°C until processing or used extemporaneously.

4.2.1.2 Phloem sampling from branches.

4.2.1.2.1 Phloem sampled from sectioned branches.

Figure 4.1: Sampling of *T. cacao* branches for phloem extraction.

50 mm (+/-10%) pieces of *T. cacao* branches were sectioned with dissection scissors. The bark was removed using a scalpel and both the bark and the branch were placed in a 50 mL tube and stored at 2-8°C for 12 hours (Figure 4.1).

4.2.1.2.2 Phloem sampled directly from *T. cacao* tree branches.

T. cacao branches were sectioned with a scalpel and sectioned 15 mL conical centrifuge tubes containing 5 mL deionised water were placed on the branches, to collect the phloem directly (Figure 4.2).

Figure 4.2: Direct T. cacao phloem collection.

Conical centrifuge tubes were placed overnight on CSSV-infected T. cacao trees to collect phloem.

Tubes were left for 12 hours on the trees and the deionised water containing the phloem was collected in a fresh 15 mL centrifuge tube.

4.2.2 DNA extraction methods development

4.2.2.1 Extraction from *T. cacao* leaves.

The extractions procedures developed in this chapter (Figure 4.3) were compared to a commercial, standard method and evaluated by quantifying the DNA extracted by qPCR.

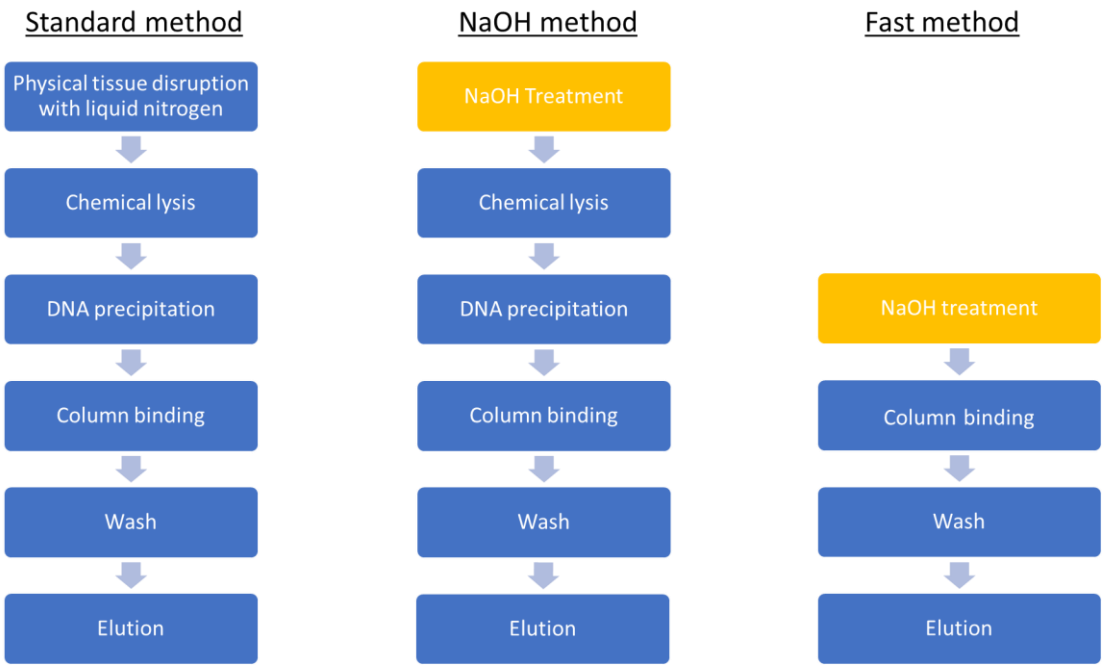


Figure 4.3: Diagram with summarised steps of the DNA extraction methods tested.

The Standard method is used as a reference protocol the NaOH and Fast methods are compared to.

In comparison to the standard kit extraction protocol, a second procedure called "NaOH method" was tested. Plant samples were incubated in sodium hydroxide for one minute to replace the liquid nitrogen step tissue disruption step from the kit protocol. Samples were then processed as per the manufacturer instructions.

Another method named "Fast method" was developed where the samples are treated with sodium hydroxide and then introduced in the standard, kit procedure from the column binding step onwards. The aim of this procedure was to evaluate if a simple sodium hydroxide treatment was sufficient to extract CSSV's DNA. Since the extraction efficiency is evaluated by a subsequent qPCR, a clean-up step was required to remove the sodium hydroxide that would interfere with the polymerase.

The benefit of adding the detergent sodium dodecyl sulphate (SDS) in the samples was also evaluated for both the NaOH and the Fast methods. This is an adaptation of a method developed by Dellaporta (Dellaporta, 1983) that relies on the SDS ability to precipitate cellular proteins and lipids while keeping nucleic acid in solution, allowing their differential isolation.

4.2.2.1.1 Standard DNA extraction procedure.

Pre-weighted stored samples were immersed in liquid nitrogen and leaves samples were disrupted with forceps. Ground samples were processed using GeneJET Plant Genomic DNA Purification Kit (Thermo) as a standard method, following the manufacturer' protocol.

4.2.2.1.2 Sodium hydroxide extraction procedure.

This extraction includes a combination of chemical lysis using sodium hydroxide and mechanical disruption using sterile glass sand. 200 mg of sterile glass sand were added to each tube containing calibrated leave discs. 350 µL of NaOH 0.2 M was added and the sample was ground with forceps for 1 min. 350 µL of Lysis Buffer A from the GeneJET Plant Genomic DNA Purification Kit (Thermo) was added to the samples and the kit procedure resumed.

4.2.2.1.3 Fast extraction method

This procedure uses the same treatment as described in the sodium hydroxide extraction procedure, but the chemical lysis and DNA precipitation were omitted. 200 mg of sterile glass sand are added to each tube containing calibrated leaves discs. 350 µL of NaOH 0.2 M were added and the sample ground for 1 min with forceps. Samples were centrifuged 5 min at 16,000x g. Supernatant was transferred to a fresh microcentrifuge tube and centrifuged again 1 min at 16,000x g. 400 µL of binding buffer were added to the supernatant and the mix was applied to the purification column from the GeneJet Kit, and the manufacturer' protocol was then followed.

4.2.2.1.4 Influence of SDS

The influence of SDS was evaluated by adding the detergent with the sodium hydroxide in the Sodium hydroxide and Fast methods. In the initial step of the procedure, 1% SDS was added to 350 µL of NaOH 0.2 M and then the remaining of the protocol was carried out similarly.

4.2.2.2 Extraction from *T. cacao* phloem.

4.2.2.2.1 Extraction from sectioned branches.

Sections of *T. cacao* branches were sectioned with dissection scissors and placed in a tube containing 30mL of a solution of either 1 M NaCl, 25 mM NaOH or deionised H_2O. Samples were macerated for 60 min at RT on a rotating wheel. After incubation, 150 µL samples were extracted using the DNeasy plant mini kit (Qiagen) using the manufacturer' instructions.

4.2.2.2.2 Direct extraction from branches.

150µL from the overnight samples collected from the sectioned tubes placed on the trees were extracted using the DNeasy plant mini kit (Qiagen) using the manufacturer' instructions.

4.2.3 Oligomeric probe functionality testing.

Biotinylated oligomeric probes targeting a region of interest located on the ORF1 of CSSV's genome are tested for their ability to detect the virus in solution.

4.2.3.1 Preliminary testing.

Prior to testing the probes with extracted plant DNA, streptavidin-coated superparamagnetic particles (SA-MP) (Pierce) were tested for their ability to detect synthetic CSSV DNA using a combination of single strand DNA probes (Figure 4.4). The first combination aimed at detecting the 76 bp ORF1T region of interest located on the ORF1 of CSSV, using the set of probes cap1 and cap2 (Table 4.1). Cap1 is complementary to the first 38 bp of ORF1T while cap2 is complementary to the second 38 bp of ORF1T (Figure 3.1). Stock SA-MP were washed twice with two volumes of buffer B (100 mM Tris, 1 M LiCl, Tween 20 0.1% pH 8) and resuspended with one volume biotin incubation buffer (5 mM Tris, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween 20, pH 8). 50 pmol of the modified oligonucleotides diluted in TE buffer (10 mM Tris-HCl, pH 9.0, 0.5 mM EDTA) were then added to the particles and incubated for 30min on a rotating wheel at RT in 500 µL Hybridisation buffer (750 mM NaCl, 150 mM Sodium citrate, pH 7.4). Solutions containing 5 pmol of the ATTO 565 nm fluorescent oligonucleotide (Eurofins) and 2.5 pmol of the synthetic single stranded ORF1T region of interest were incubated with the particles

at RT for 30 min, protected from light. Binding reactions were then washed with three volumes of Hybridization buffer using a magnetic rack. Reactions were then eluted from the SA-MP with 50 mM NaOH for 15 min at RT. 50 mM HCl were added to the preparation and the eluates were collected, transferred to a 96 well plate, and imaged using a Li-Cor Odyssey imaging system.

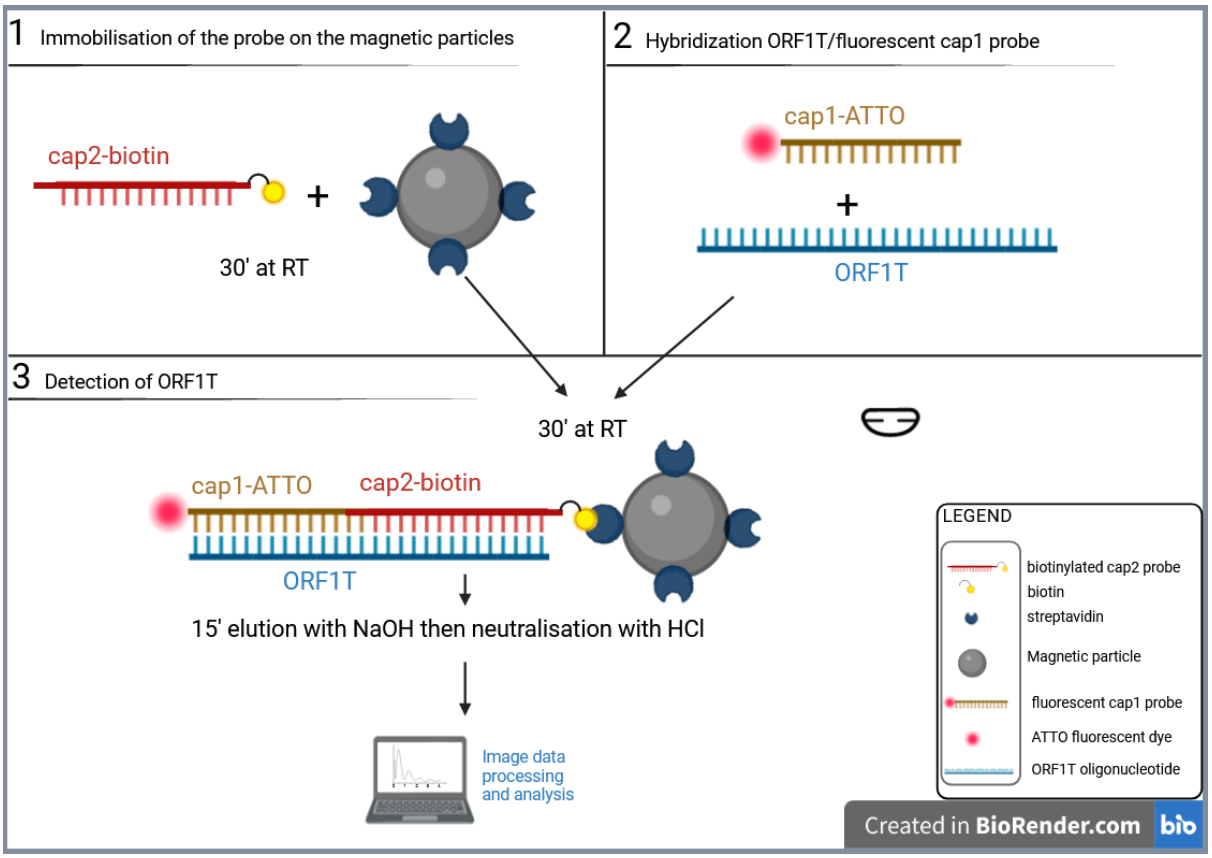


Figure 4.4: Validation of streptavidin coated magnetic particles for detecting the synthetic DNA fragment ORF1T of the ORF1 of CSSV using the oligonucleotide probe cap1 and cap2.

The second combination tested used the 2143ROI oligonucleotide probe that is targeting a sequence on the ORF3 of CSSV. The probe was immobilised on the SA-MP and the complementary sequence probe, 2143ROIC, labelled with the ATTO fluorescent dye, was used as a reporter (Table 4.1).

Table 4.1: Oligomeric probes and synthetic oligonucleotides used in this chapter.

Primer name	Sequence 5′-3′	Size (bp)
cap1	GAGTACCTTGACCTGGCCTCTACCAGTAAAGTGACCAA	38
cap2	CAACCAACTAGCTCATAACCTCGCAGTAACCTTTGATA	38
2143ROI	ATAAGAGAACACACTGTAGCAGTCT	25
2143ROIC	AGACTGCTACAGTGTGTTCTCTTAT	25
ORF1T	TATCAAAGGTTACTGCGAGGTTATGAGCTAGTTGGTTGTTGGTCACTTTACTGGTAGAGGCCAGGTCAAGGTACTC	76
ORF1TC	GAGTACCTTGACCTGGCCTCTACCAGTAAAGTGACCAACAACCAACTAGCTCATAACCTCGCAGTAACCTTTGATA	76

4.2.3.2 Pull-down assay with CSSV-infected *Theobroma cacao* extracts.

The objective of the pull-down was to verify that the single stranded probe can detect CSSV DNA in solution. Biotinylated probes were first annealed with CSSV DNA. SA-MP were then added to the samples, incubated at room temperature on a rotating wheel to capture the probes bound to the DNA. Whole DNA was then eluted from the SA-MP and quantified by qPCR.

4.2.3.2.1 Leaf extracts.

150 µL of CSSV-infected *T. cacao* leaf DNA extracts were mixed with 2.5 µL of 2 µM biotinylated oligonucleotides probes. Whole DNA samples were then heated at 95˚C for 10 min in the Rotor-Gene Q (Qiagen) to denature the double stranded molecules. Probe and whole DNA were annealed by incubating the samples 2°C below the Tm of the probes for 5 min and then placing them overnight on ice or alternatively left to cool to RT on the bench. 4 mg/mL SA-MP were added to the samples and incubated for 30 min at RT on a rotating wheel. Reactions were placed on a magnetic rack, the SA-MP washed twice in two volumes of washing buffers. Several wash buffers were tested. Default washing buffer was buffer B/W (20 mM Tris pH 1.5, 1 M NaCl, 1 mM EDTA, 0.0005% Triton X-100). For the stringent wash testing, SA-MP were first washed twice for 5 min with 2X SSC, 0.1% SDS at RT and twice

with 0.1X SSC, 0.1% either at RT, 40, 50 or 60°C. After the washes, SA-MP were incubated 5 min at 95°C in TE buffer to elute the DNA. Eluted DNA was use as a template for qPCR reactions.

4.2.3.2.2 Phloem extracts.

- Phloem from sectioned branches.

24 mL of the macerated branches samples from section 2.2.2.a were mixed with 6 mL of 5X B/W buffer and 2 µM of biotinylated cap1 and cap2 oligonucleotides. Oligonucleotides and extracts were incubated for 30 min on a rotating wheel. 50 µL of SA-MP at 4 mg/mL were added to the samples and incubated with the extracts for 30 min on a rotating wheel. Samples were placed in the fridge for 18 hours on the magnetic rack to allow SA-MP to be collected. Supernatant was removed, and the SA-MP washed twice with 1 mL B/W buffer, followed by stringency washes with 1 mL 2X SSC, 0.1% SDS at RT twice and twice with 0.1X SSC, 0.1% at RT. Samples were eluted with 50 µL TE buffer, heated to 95°C for 5 min and amplified by qPCR.

- Phloem collected directly from the branches.

2mL from the overnight samples were transferred in a bijou container and mixed with 500 µL 5X B/W buffer, 15 µL biotinylated cap1 and 15 µL biotinylated cap2 at 2 µM, and 100 µL SA-MP at 4 mg/mL. Samples were incubated for 30 min at RT on a rotating wheel. Samples were then diluted with 1-3 mL of B/W buffer until the consistency of the samples allows the SA-MP to bind to the side of the microcentrifuge tube when placed on the magnetic rack. SA-MP were then washed with 1 mL 2X SSC, 0.1% SDS at RT twice and twice with 0.1X SSC, 0.1% at RT. Samples were eluted with 50 µL TE buffer, heated to 95°C for 5 min and amplified by qPCR.

4.2.4 qPCR amplifications.

Two qPCR were used in this chapter. A SYBR Green qPCR assay was employed to evaluate the efficacy of the CSSV DNA extraction methods from *T. cacao* leaves. A TaqMan assay was performed for

analysing the extraction of CSSV from the tree's phloem and for assessing the performance of the biotinylated oligonucleotides in the pull-down assay.

The SYBR Green qPCR were set using primers targeting a region of the CSSV genome overlapping ORF2 and ORF3 to produce a 387 base pairs fragment (Table 4.2). Each 20 µl reaction consisted of 2 µl forward primer "93" at 2 µM, 2 µl reverse primer "93" at 2 µM, 10 µL Sensimix Sybr Master Mix 2X (Bioline) and 1 µL DNase/Rase free H_2O (Invitrogen) and 5 µL template DNA. All reactions were run in triplicate on a Rotor-Gene Q instrument using the following parameters: Hold 95°C 10 min, Cycle: 95°C 10 s, 54°C 15 s, 72°C 20 s. Cycles were allowed to repeat 55 times. A melt step was added following the amplification cycles, to verify the specificity of the reaction: ramp from 67°C to 95°C, rising by 1°C each step, waiting for 5 s after each step. The efficiency of the extractions was evaluated by comparing the cycle threshold values (Ct values) of the method tested to the Ct values obtained with the standard reference extraction. Extract samples from the Standard method were pooled (normalised at 20 ng/µL) and diluted 10, 100 and 1000-fold and used as standards for evaluating the efficiency of both the NaOH and the Fast methods. Triplicates qPCRs were performed on each four extraction replicates. Triplicates Ct values were averaged, and two standard error of the mean (2SE) calculated for the 4 biological replicates.

The efficacy of the phloem CSSV DNA extraction and the ability of the biotinylated oligonucleotides to retrieve viral DNA from plant samples processed with the magnetic beads were evaluated by TaqMan PCR assay using primers targeting the reverse transcriptase of CSSV (RTase). A nuclear genome locus positioned next to the simple sequence repeat (SSR) mTcCIR25 from *T. cacao* was assessed via qPCR alongside all samples, to quantify the quantity of plant DNA co-purified with the viral sequences. mTcCIR25 is a single nuclear locus enabling quantification of plant genome number. In parallel, a set of synthetic RTase targets consisting of a range of increasing copy number (500, 5000, 50000 and 500000) and amplified using the RTase assay was used to determine the copy number for both viral and nuclear genomes.

Table 4.2: Primers, probe, and reference CSSV fragment sequences used in the qPCR quantifications.

			Utilisation		
Description	**Function**	**Sequence 5' - 3'**	**Leaf extraction qPCR**	**Phloem extraction TaqMan**	**Biotinylated probe Pull-down assay**
F-74	Forward primer - CSSV Reverse Transcriptase	CTGAAGCGAGTAGGCAACAA	×	✓	✓
R-151	Reverse primer - CSSV Reverse Transcriptase	CAGTCCAAGGGATGGACTCT	×	✓	✓
CSSV RTase	Taqman Assay probe	[FAM]TCCATCAGGTTGCCATGGCA[BHQ1]	×	✓	✓
MTcCIR25F	Forward primer - *T. cacao* single locus SSR (Y16997)	CAGATAAGGAAAGGTGGAGTTTGG	×	✓	✓
MTcCIR25R	Reverse primer - *T. cacao* single locus SSR (Y16997)	CAAGAATGTCTCCTACATTCACTACG	×	✓	✓
MTcCIR25P	Taqman Assay probe	[FAM]TTCCCGTAAGCTTCGTCCCAGATGC[BHQ1]	×	✓	✓
F-93	Forward primer - CSSV ORF2-3 Overlap	TGAAGAACTGGTGCAAGATCA	✓	×	×
R-93	Reverse primer - CSSV ORF2-3 Overlap	TCAACCTTCTGGCATTCCTT	✓	×	×
Rtase Synthetic fragment	Standard sequence	CTGAAGCGAGTAGGCAACAAAAAAATTTTCAGCAA GTTTGATTTAAAATCAGGCTTCCATCAGGTTGCCAT GGCAGAAGAGTCCATCCCTTGGACTG	✓	✓	✓

TaqMan qPCR was then performed on 5 µl of DNA samples using both a CSSV primer/probe assay and a plant genome primer/probe mix (Table 4.2) in Sensifast No RoxMaster Mix (Bioline). The qPCR reaction mix primer/probe mix was prepared by combining probe (2.5 µl of 100 pmol/ µl), forward primer (10 µl of 100 pmol/ µl), reverse primer (10 µl of 100 pmol/ µl) and 77.5 µl of 10 mM Tris-Cl. 0.5 mM EDTA; pH 9.0. Each 20 µl reaction consisted of 10 µl Sensifast No RoxMaster Mix (Bioline), 0.8 µl primer/probe mix, 4.2 µl nuclease free water and 5 µl DNA template.

4.2.5 Development of a membrane-based cold neutral hybridisation assay.

4.2.5.1 Synthetic oligonucleotides.

The aim of this experiment was to evaluate if CSSV's DNA can be detected in a temperature independent assay. Nylon membranes were spotted with synthetic ORF1T with ORF1TC spotted as a negative control. DNA extracted from CSSV-infected *T. cacao* leaves using the NaOH method was mixed with biotinylated cap1 and cap2 oligonucleotides probes and applied to the membranes, at RT. Membranes were washed, SA-HRP incubated on the membranes and the chemiluminescent signal measured with a Li-Cor Odyssey.

Biodyne nylon membranes B (Thermo, UK) were pre-incubated in AB2X (200 mM Tris HCl pH 8, 500 mM NaCl, 100 mM EDTA) for 120 min. 50 or 100 ng Synthetic ORF1T and 100 ng ORF1TC were applied on the membrane using a vacuum manifold dot blotting apparatus (Bio-Rad, UK). Membranes were

blocked with 100 µg/mL pre-denatured salmon sperm DNA (Thermo, UK) in AB2X. Biotinylated cap1 and cap2 probe diluted in AB1X were added at a range of concentration from 0.48 to 1000 ng/mL. For competition assay, ORF1T and ORF1TC were added to the probe mix and applied to the membranes at a range of concentrations from 5 nM to 500 nM. Membrane was washed twice in AB2X. Membranes were then removed from the dot blot apparatus and blocked in PBS + 1% BSA (Sigma, UK) for 60 min at RT. Streptavidin-horseradish peroxidase conjugate (Invitrogen, UK) at 1/500 dilution was applied to the membrane in PBS + 1% BSA for 45 min. Membranes were washed three times in PBS+0.1% Tween 20. Chemiluminescent substrate was applied to the membranes (Pierce, UK) and signal was measured using a Li-Cor Odyssey. Resulting chemiluminescent signal from digital pictures of the membranes was measured using the Fiji imaging software (Schindelin *et al.*, 2012b) with the protein analyser module and corresponding pixel density measurements copied for further analysis in an excel spreadsheet.

4.2.5.2 Cold hybridization with CSSV-infected plant samples

A competition was conducted adding biotinylated cap1 and cap2 probe to *T. cacao* extracts positive or negative to CSSV.

Biodyne nylon membranes B pre-incubated with AB2X. ORF1T and ORF1TC were spotted on the membranes in triplicates. Membranes blocked with 100 µg/mL pre-denatured salmon sperm DNA (Thermo) in AB2X. Strips of the membrane were placed in bijous containers containing 1 µg/mL biotinylated cap1 and cap2 in AB1X and either 1400 ng of CSSV negative *T. cacao* whole DNA extracts or 1400 ng of CSSV positive *T. cacao* whole DNA samples (as defined in section 4.2.1) extracted using the NaOH method. Bijous were placed on a tube roller for 45 min at RT. DNA solutions were decanted and the membrane strips washed three times with PBS+0.1% Tween 20. Strips were blocked for 18 hours at 4°C in PBS+1% BSA (Sigma). Streptavidin-horseradish peroxidase conjugate (Invitrogen) at 1/500 dilution was applied to the membrane in PBS + 1% BSA for 45 min. Membranes were washed three times in PBS+0.1% Tween 20. Chemiluminescent substrate was applied to the membranes (Pierce) and signal was measured using a Li-Cor Odyssey.

4.2.6 Pull-down assay following sodium hydroxide-mediated DNA denaturation.

Whole DNA extracted from either CSSV positive, or negative *T. cacao* leaves were treated with sodium hydroxide. The aim of the procedure was to open the double-stranded DNA without the intervention of temperature, to allow access of the region of interest to the biotinylated probes.

Calibrated leaves sections were cut using a coring tool and the resulting samples were weighed. 25 mg sections were frozen in liquid nitrogen and broken-down using tweezers. Whole DNA was then extracted using the guanidine salts method with the DNeasy plant mini kit (Qiagen, UK), and eluted in TE buffer. Triplicates 45 µL samples from either CSSV positive or negative plants were mixed with 2.5 µL of the cap1 and cap2 biotinylated oligonucleotide probes. 15 µL of 100 mM sodium hydroxide was added to each sample and incubated for 5min at room temperature. Solutions were neutralised with 1.25 µL 1.5 N acetic acid and incubated for 5min at room temperature. 400 µL of B/W buffer (200 mM Tris-HCl pH 7.5, 1 M NaCl, 1 mM EDTA, 0.0005% Triton X-100) was added to the samples and mixed with 50 µL of 4 mg/mL SA-MP pre-equilibrated in reaction buffer. Samples was incubated 30 min at RT on a rotating mixer. SA-MP were collected by placing the sample tubes on a magnetic rack and washed twice with 500 µL B/W buffer. DNA was eluted by heating the SA-MP in 50 µL TE buffer and quantified by qPCR.

4.3 Results

4.3.1 Evaluation of the DNA extraction methods from leaves material.

DNA samples extracted from *T. cacao* leaves using either a standard method using a commercial kit and two modified procedures including NaOH treatment, the Fast and the NaOH methods were evaluated by qPCR (Figure 4.5).

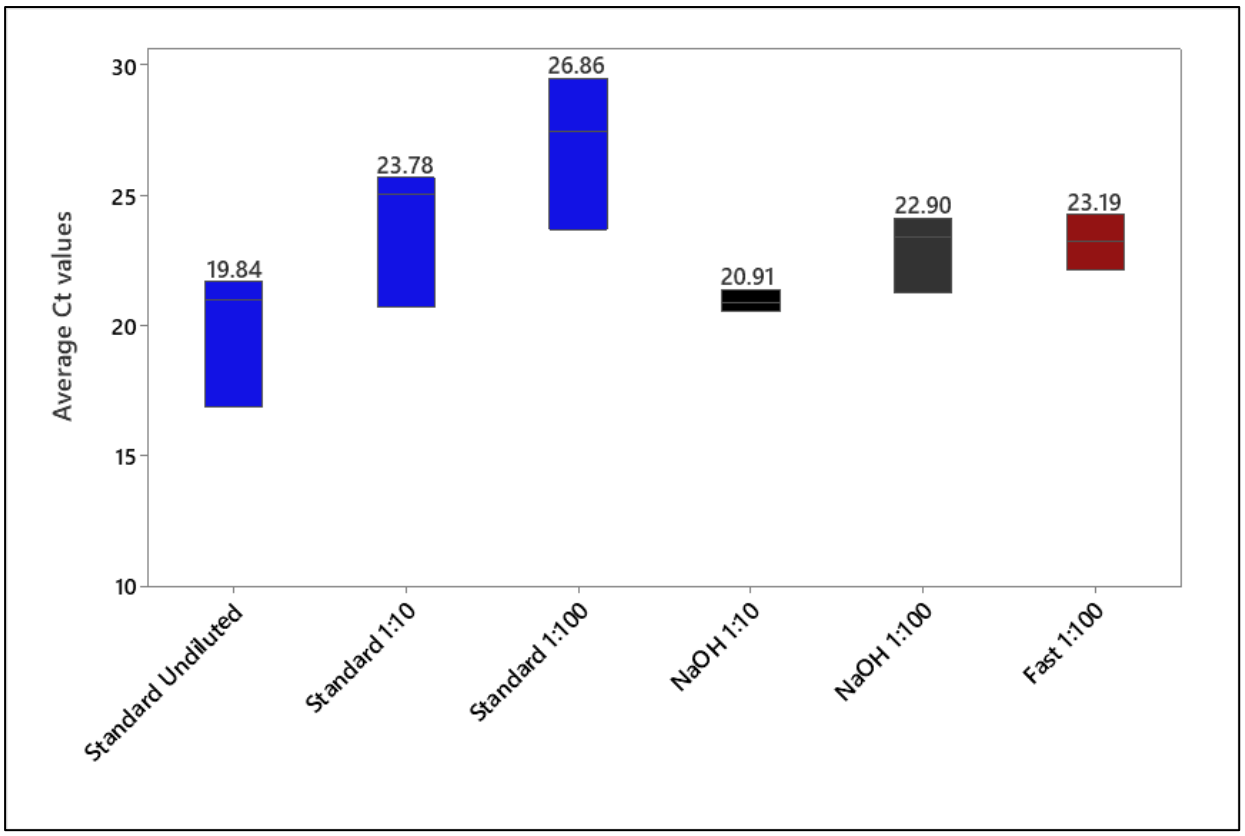


Figure 4.5: Boxplot of the Ct values of the qPCR amplification of a 387 bp CSSV amplicon comparing the three extraction procedures.

The figure above each box is the mean value of three extraction each run as three replicates qPCR.

Only samples from the standard extraction samples resulted in a qPCR amplification for undiluted samples (Figure 4.5). For samples diluted 10 times, extracts from the NaOH method were first to amplify on average, followed by the Standard method. Samples from the Fast method did not amplify. At a 1 in 100 dilution, NaOH method extracts amplified first, followed by the Fast method extracts samples. Samples from the Standard kit extraction amplified then last.

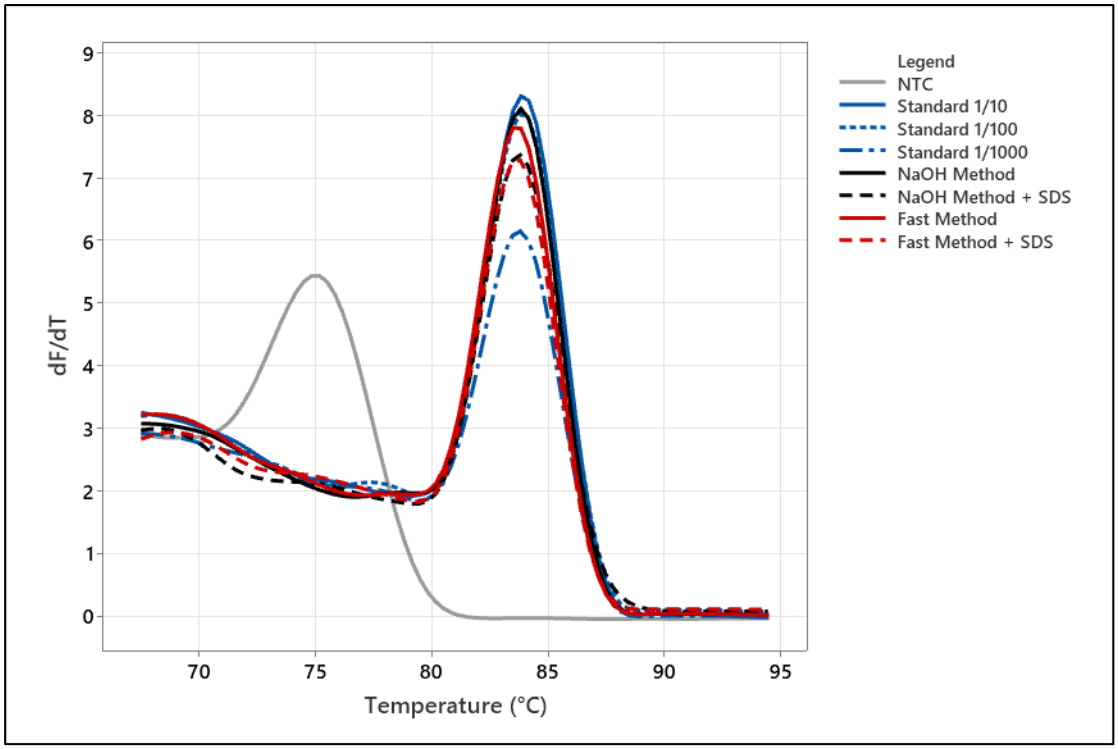
Since the dilution of the eluate had influenced the observed Ct values of the extractions, a further dilution, 1 in 1000 was tested (Table 4.3). In this run, eluates from the Standard method were used as standards and run diluted 1 in 10, 1 in 100 and 1 in 1000. In addition, the benefit of adding SDS to improve extraction of genomic DNA was tested by adding 0.1% SDS to the sodium hydroxide solution used in the first step of the procedure.

Table 4.3: Average Ct values obtained in two qPCR amplifications of a 387bp CSSV amplicon obtained with the tested DNA extractions methods. Each Ct value of the samples from the method tested (NaOH, NaOH+SDS, Fast, Fast+SDS) was calculated from four biological replicates of the extraction, and three mechanical qPCR replicates from each sample.

Run 1			
Method	Dilution	Average Ct	2 Standard Errors
Standard	1/10	17.72	
	1/100	29.08	
	1/1000	35.74	
NaOH	1/1000	29.99	1.705
NaOH+SDS	1/1000	29.74	0.704
Fast	1/1000	29.14	0.297
Fast+SDS	1/1000	30.94	1.444
NTC	N/A	47.13	5.657

Run 2			
Method	Dilution	Average Ct	2 Standard Errors
Standard	1/10	18.10	
	1/100	22.69	
	1/1000	35.37	
NaOH	1/1000	32.14	1.082
NaOH+SDS	1/1000	32.33	1.810
Fast	1/1000	29.26	0.760
Fast+SDS	1/1000	29.71	0.855
NTC	N/A	45.49	2.385

Ct values of the 1/1000 diluted eluates were 35.74 for the Standard method on the first qPCR and 35.37 on the second. Ct values for the Fast method were 29.14 in the first run and 29.26 in the second. For the NaOH method, the Ct values were 29.99 and 32.14 in the first and the second run, respectively. Adding SDS had no noteworthy effect on the Ct values obtained with the NaOH method (29.99 without SDS, 29.74 with SDS in the first run, 32.14 without SDS, 32.33 with SDS in the second run). For the Fast method, adding SDS increased the Ct values: 29.14 without SDS, 30.94 with SDS in the first run; 29.26 without SDS, 29.71 with SDS in the second run.



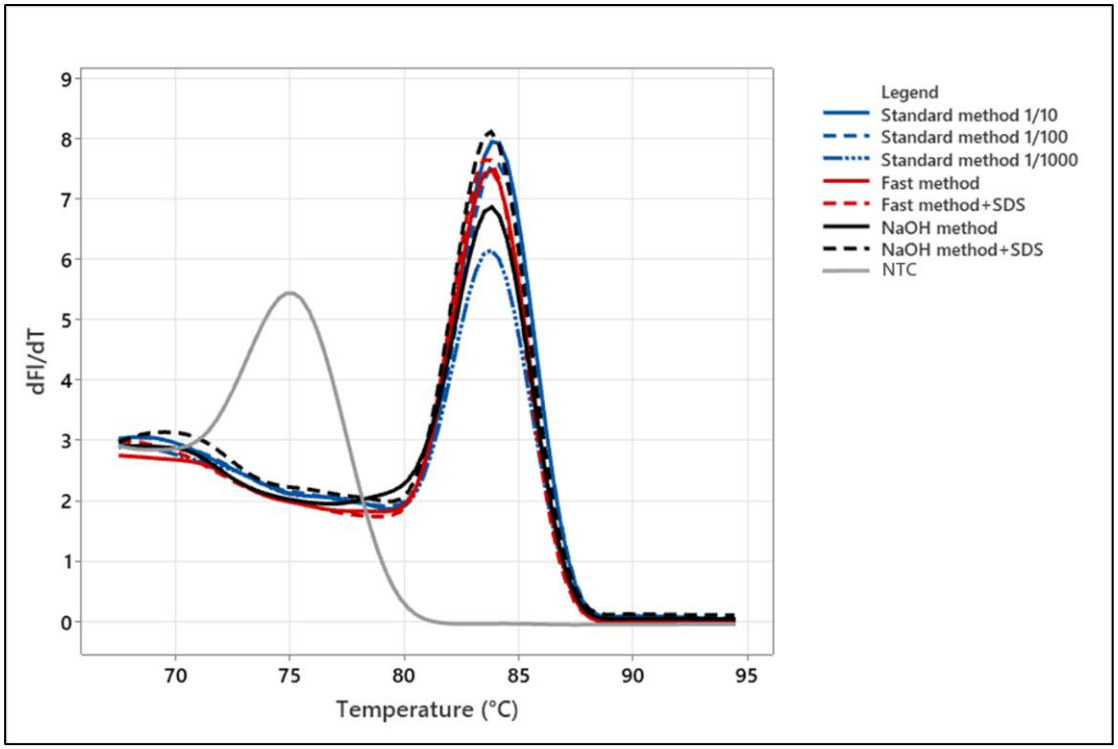


Figure 4.6: Melt curve profile after qPCR amplifications of the extracts.

Top graph represents the melt of the first qPCR run, the bottom graph represents the melt of the second qPCR run. In both instances, the amplified fragment displayed a single melting temperature, suggesting that a single product was amplified, regardless of the extraction method used. A single replicate is displayed.

A melt step was added after the amplification to verify the specificity of the amplification for both qPCR amplifications (Figure 4.6). A single peak at 84°C was observed for all samples tested, apart from the no template control. NTC samples melt resulted in a peak at 75°C corresponding to primer dimers.

4.3.2 Evaluation of the DNA extraction methods from phloem samples.

DNA was extracted from T. cacao phloem from macerated, sectioned branch using either deionised water, 25mM sodium hydroxide or a 1M NaCl solution, followed by a commercial kit extraction (Table 4.4).

Table 4.4: TaqMan qPCR calculated copies of DNA extraction from macerated CSSV-infected *T. cacao* branch sections in several solutions. The values are the average amplified copies calculated from three qPCR replicates from a single extraction.

	Young branch sample Tree "61"			Control Extracts
	Deionised water	NaCl	NaOH	
Amplified *T. cacao* copies	1.42E+02	N/A	8.11E+02	2.19E+06
Amplified CSSV copies	3.38E+04	1.10E+02	1.29E+04	1.65E+08
Ratio CSSV/*T. cacao*	4.08E+02	N/A	1.59E+01	7.54E+01

Samples incubated in 1 M NaCl produced the lowest amplification for the viral marker, with $1.10x10^2$ copies produced on average. The nuclear marker from *T. cacao* was not amplified. Samples extracted in sodium hydroxide yielded $1.29x10^4$ CSSV copies on average and $8.11x10^2$ *T. cacao* copies, for a ratio CCSV/*T. cacao* of 15.94. Deionised water resulted in the highest CSSV copies extracted, $3.38x10^4$ on average, and $1.42x10^2$ copies of *T. cacao* were measured on average, for an average ratio of 407.97, which was a 5.41 increase compared to the control kit leaves extracts ratio for this tree.

A PCR was run on phloem extracted directly in deionised water contained in a centrifuge conical tube attached overnight to a sectioned branch of CSSV-infected *T. cacao* trees, as described in 2.1.2.b (Table 4.5).

Table 4.5: TaqMan qPCR calculated copies of DNA phloem direct extractions from CSSV-infected *T. cacao* branches in deionised water.

	Tree "61"		Tree "8"	Tree "10"
	Branch 1 (bottom)	Branch 2 (top)		
Amplified *T. cacao* copies	4.26E+01	6.45E+00	1.92E+01	1.17E+01
Amplified CSSV copies	6.30E+03	5.57E+04	1.53E+04	1.09E+04
Ratio CSSV/*T. cacao*	1.48E+02	4.75E+11	6.56E+02	9.40E+02

Two tubes were placed on tree “61”, one at the top of the tree and one at the bottom. On the top tube, $5.57x10^4$ copies of CSSV were amplified on average. Amplification of *T. cacao* from this tube was low, with an average of 6.45 copies measured. In the bottom tube, $4.27x10^1$ copies of *T. cacao* were detected on average and $7.64x10^3$ copies of CSSV on average, resulting in an average CSSV/*T. cacao* ratio of 179.06.

On tree “8”, $1.53x10^4$ copies of CSSV were detected on average. *T. cacao* was detected on average at $1.92x10^1$ copies. In the last tree sampled, “10”, three qPCR replicates targeting *T. cacao* resulted on average in an amplification of $1.17x10^1$ copies. On the same sample, CSSV amplification yielded $1.09x10^4$ copies.

4.3.3 Evaluation of biotinylated oligonucleotide probes to detect CSSV in solution.

The biotinylated oligonucleotide probes cap1, cap2, 2143ROI and 2143ROIC were evaluated for detecting a single region of interest within the genome of CSSV.

A preliminary experiment was conducted to verify if the combination of biotinylated cap1 immobilised on streptavidin-coated magnetic particles (SA-MP) and an ATTO-labelled cap2 could bind to the synthetic CSSV region of interest ORF1T. 2143ROI is an alternative biotinylated probe targeting a 25bp sequence in the ORF3 of CSSV. 2143ROI binding to the SA-MP was monitored using an ATTO-labelled oligonucleotide, 2143ROIC which is the complementary sequence to 2143ROI.

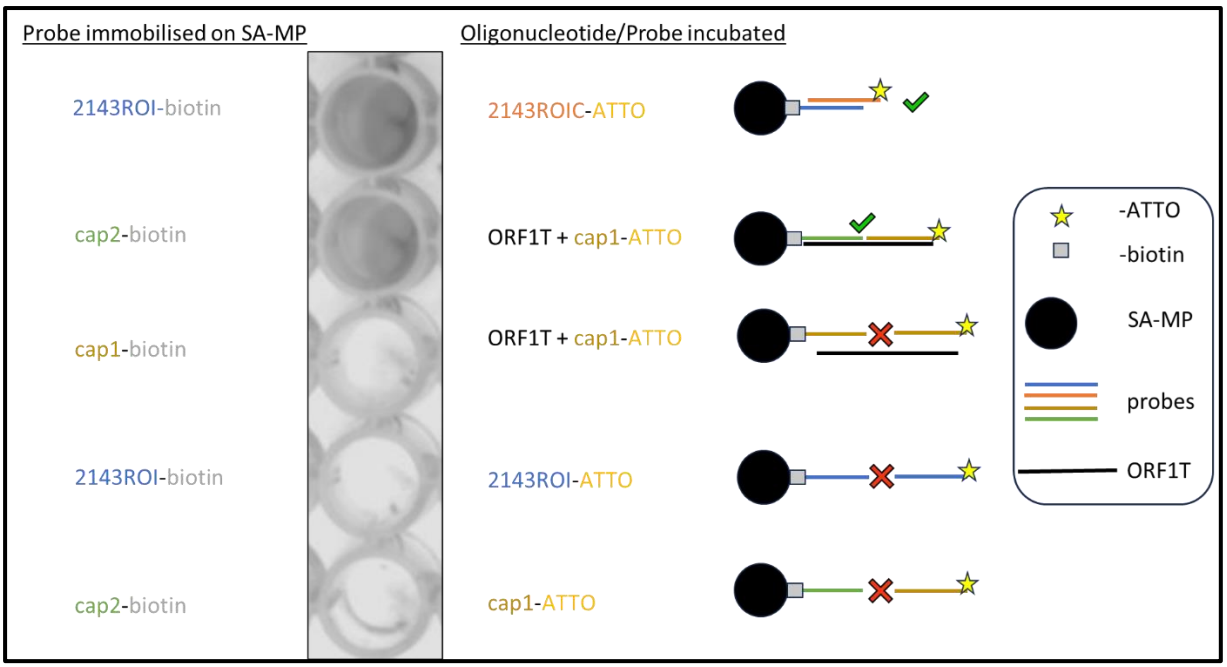


Figure 4.7: Validation of streptavidin coated magnetic particles for detecting the synthetic DNA fragment ORF1T of the ORF1 of CSSV using the oligonucleotide probe cap1 and cap2 or the probe 2143ROI.

A set of biotinylated probes (cap1, cap2 or 2143ROI) were immobilised on streptavidin magnetic particles (SA-MP). Cognates probes labelled with the ATTO dye (2143ROIC or cap1) were added to the particles with or without ORF1T. Eluted DNA was transferred to a 96-wells plate and imaged using a Li-Cor Odyssey imaging system.

The cap2 biotinylated oligonucleotide successfully bound to the SA-MP and was able to bind to the complementary sequence of the synthetic region of interest ORF1T. A signal mediated by the binding of the ATTO-modified cap1 oligonucleotide was produced (Figure 4.7). No signal was observed when the biotinylated cap2 oligonucleotide immobilised on the particles was incubated with cap1-ATTO in absence of ORF1T. Similarly, no signal was observed when the biotinylated cap1 oligonucleotide immobilised on the particles was incubated with ORF1T and cap1-ATTO.

The alternative 2143ROI oligonucleotide probe bound to the SA-MP produced a signal when the complementary 2143ROIC-ATTO was used, but not when 2143ROI-ATTO was used instead, demonstrating the specificity of the reaction.

A calibration curve was drawn by incubating diluted cap1-ATTO + synthetic ORF1T with a constant concentration of biotinylated cap2 immobilised on SA-MP (Figure 4.8).

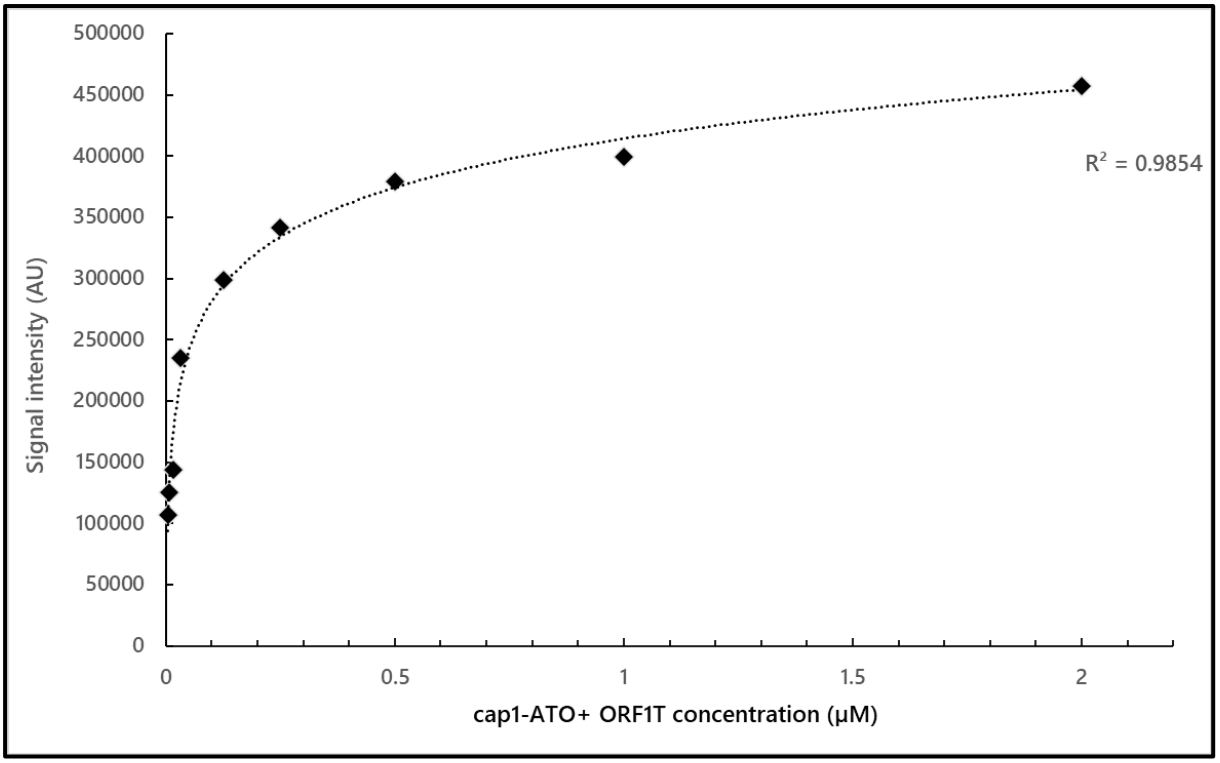


Figure 4.8: Biotinylated cap2 (2 µM) was immobilised on SA-MP. Increasing concentration of cap1-ATTO+ synthetic ORF1T was added to the SA-MP and produced a dose-response curve.

After incubation and washes, DNA eluted from the SA-MP was transferred to a 96-well plate the signal produced was measured using a Li-Cor Odyssey imaging system. A logarithmic line of best fit was added.

When increasing amounts of cap1-ATTO and ORF1T were incubated with a constant amount of biotinylated cap2 immobilised on the particles, an increasing fluorescent signal was measured on the eluates (Figure 4.8).

4.3.4 Evaluation of biotinylated oligonucleotide probes to detect CSSV from *T. cacao* leaf extracts.

Oligonucleotide probes were mixed with CSSV-infected *T. cacao* extract, the samples heated and then incubated with SA-MP. The magnetic particles were then washed, the DNA eluted, and qPCR targeting both CSSV and *T. cacao* was used to quantify the DNA retained by the particles.

qPCR was run on both the eluted DNA and on the SA-MP (Figure 4.9). The cap1 and cap2 biotinylated oligonucleotides pulled down by the SA-MP could bind the DNA of CSSV. The DNA from *T. cacao* also bound to the particles. The PCR were run using either the eluted whole DNA from the SA-MP or using the particles as a template. $4.74x10^6$ copies of CSSV and $1.02x10^5$ copies of *T. cacao* were amplified

when using the eluates as a template, on average. The ratio of amplified CSSV/amplified T. cacao was 46.59. There was a lower amplification of DNA on the particles compared to the eluted fractions. Specifically, $2.24x10^4$ copies of CSSV and 981 copies of *T. cacao* were measured. The ratio CSSV/*T. cacao* was 227.99 and higher than on the eluate fractions on average.

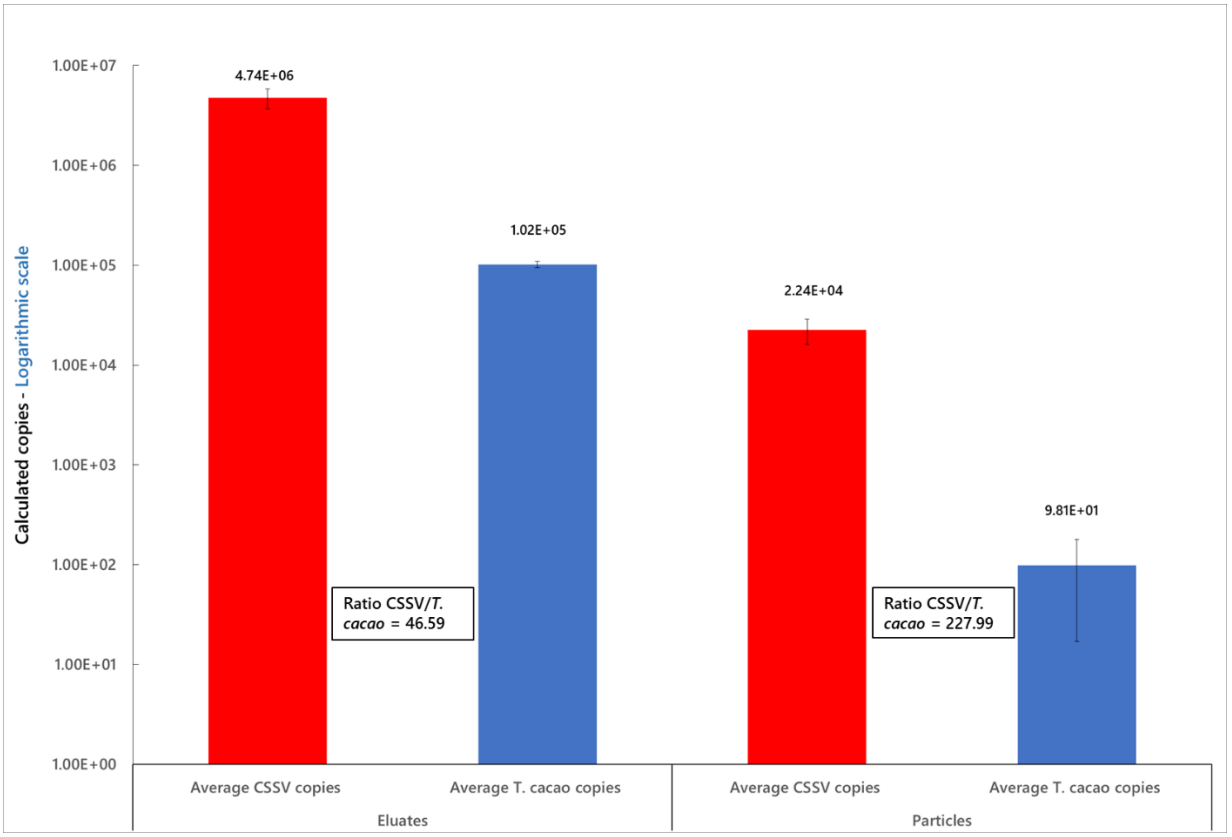


Figure 4.9: TaqMan qPCR analysis was performed on both the eluted fractions and the particles to quantify the CSSV and *T. cacao* DNA captured by the biotinylated cap1 and cap2 oligonucleotides probes using streptavidin-coated particles.

Calculated copies on the y-axis are represented on a logarithmic scale. Error bars represent two standard errors of the mean.

The ability of the probes to act synergistically was evaluated by combining cap1, cap2 and either 2143ROI or 2143ROIC with CSSV-infected T. cacao extracts. Probes were annealed with the DNA extracts in the Rotor-Gene Q 2°C below their melting temperature and then placed on ice for 18hours. SA-MP were then added to the mixture, incubated, and washed. Whole DNA was eluted, and a qPCR was run on the eluates (Figure 4.10).

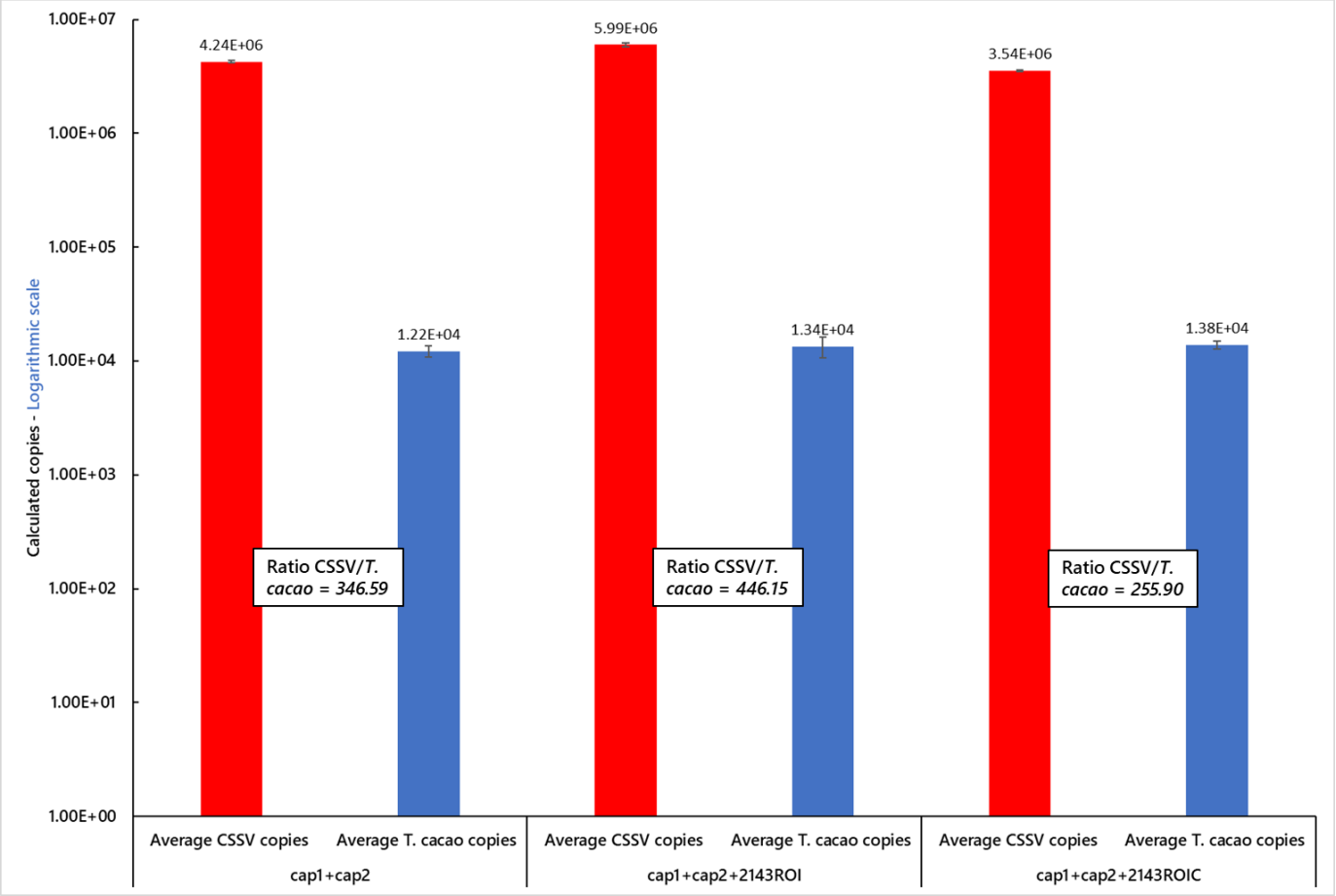


Figure 4.10: Evaluation of probe combinations to detect CSSV genome.

TaqMan qPCR analysis was performed on eluted fractions from SA-MP to quantify the CSSV and *T. cacao* DNA captured by combinations of biotinylated cap1, cap2, 2143ROI and 2143ROIC oligonucleotides probes. Calculated copies on the y-axis are represented on a logarithmic scale. Error bars represent two standard error of the mean from three qPCR replicates.

When using cap1 with cap2, 4.26×10^6 copies of CSSV were amplified, and 1.22×10^4 copies of *T. cacao*, for a ratio of 346.59. Adding the probe 2143ROI resulted in an increase in amplified CSSV copies to 5.99×10^6, and to an increase in *T. cacao* copies to 1.34×10^4. The ratio was improved to 446.15. Using the 2143ROIC probe, that is the complementary sequence to 2143ROI, with cap1 and cap2 resulted in the least copies of CSSV amplified of the set tested with 3.53×10^6 copies measured. In contrast, *T. cacao* copies were the highest, with 1.38×10^4 copies detected, resulting in the lowest ratio, 255.9, of the three combinations.

In order to improve the ratio CSSV/*T. cacao* amplified copies further, the washing procedure was modified with the aim was to reduce non-specific binding of *T. cacao* DNA to the oligonucleotide

probes. Stringent washes utilising saline sodium citrate buffer (SSC) with SDS were compared to the control procedure that relies on Tris, EDTA and NaCl buffer (B/W buffer) (Figure 4.11).

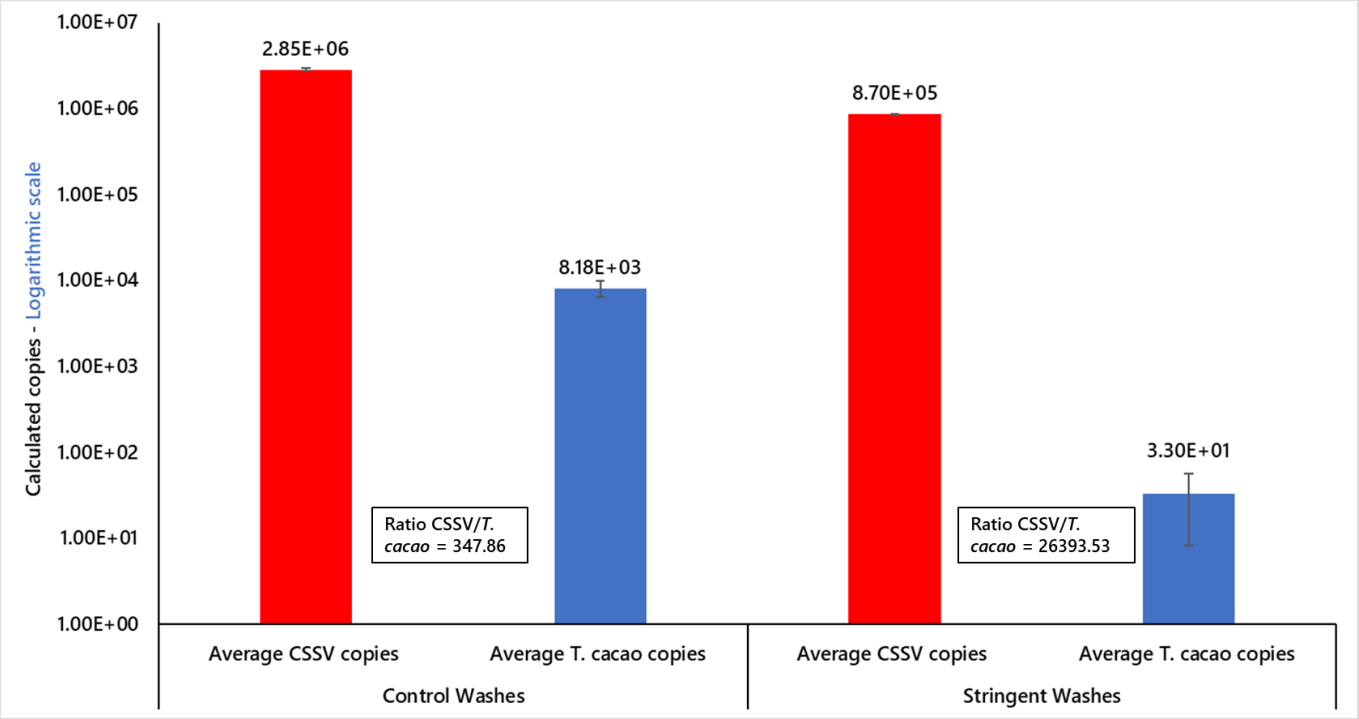


Figure 4.11: Influence of the washing procedure on the DNA binding specificity to the SA-MP.

TaqMan qPCR analysis was performed on eluted fractions from SA-MP to quantify CSSV and *T. cacao* DNA captured by the biotinylated cap1 and cap2 oligonucleotides probes. Calculated copies on the y-axis are represented on a logarithmic scale. Error bars represent two standard errors of the mean from three qPCR replicates.

$2.85x10^6$ copies of CSSV and $8.18x10^3$ copies of *T. cacao* were obtained with the control wash procedure. The CSSV/*T. cacao* ratio was 347.86. When the stringent washes procedure was applied, $8.7x10^6$ copies of CSSV were eluted from the particles and 330 copies of *T. cacao* were recovered. The CSSV/*T. cacao* amplified copies ratio was 26393.53, which is over 75 times greater than with the control washing procedure.

4.3.5 Detection of synthetic ORF1T with the cap 1 and cap2 oligonucleotide probes in a cold, membrane-based hybridisation assay.

A cold hybridisation assay was developed to enable the detection of CSSV without modifying the temperature.

The synthetic region of interest from the genome of CSSV, ORF1T was spotted on a nylon membrane. The complementary sequence ORF1TC was also spotted as a negative control. The biotinylated oligonucleotide probes were added at a concentration ranging from 1µg/mL to 0.48ng/mL (Figure 4.12).

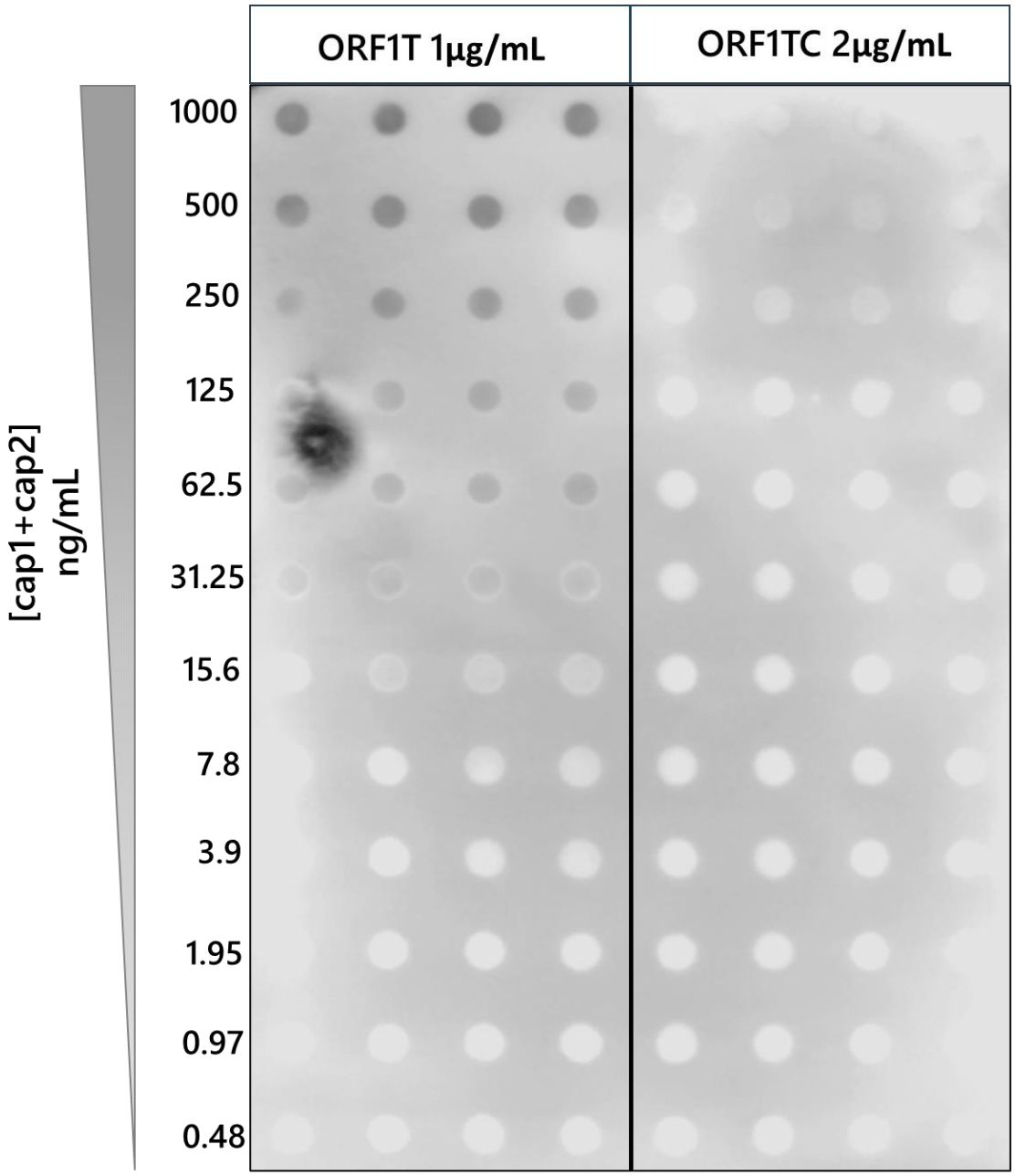


Figure 4.12: Membrane assay in "cold" conditions, performed at 25°C (+/- 3°C) showing a calibration curve of the biotinylated cap1 and cap2 probes.

Synthetic ORF1T (left rows) or ORF1TC (right rows) were spotted on a nylon membrane and incubated with a decreasing concentration of biotinylated cap1+cap2 oligonucleotides.

The intensity of the signal on the membrane was measured and plotted (Figure 4.13). The signal on the lanes spotted with ORF1T decreased incrementally when the quantity of probe in solution was reduced. The signal measured at 15.6ng/mL decreased to 6.64% of the value calculated at 31.35ng/mL. Further decrease of the concentration of probe led to a linear diminution of the signal until it reached a plateau at a concentration of 0.97ng/mL.

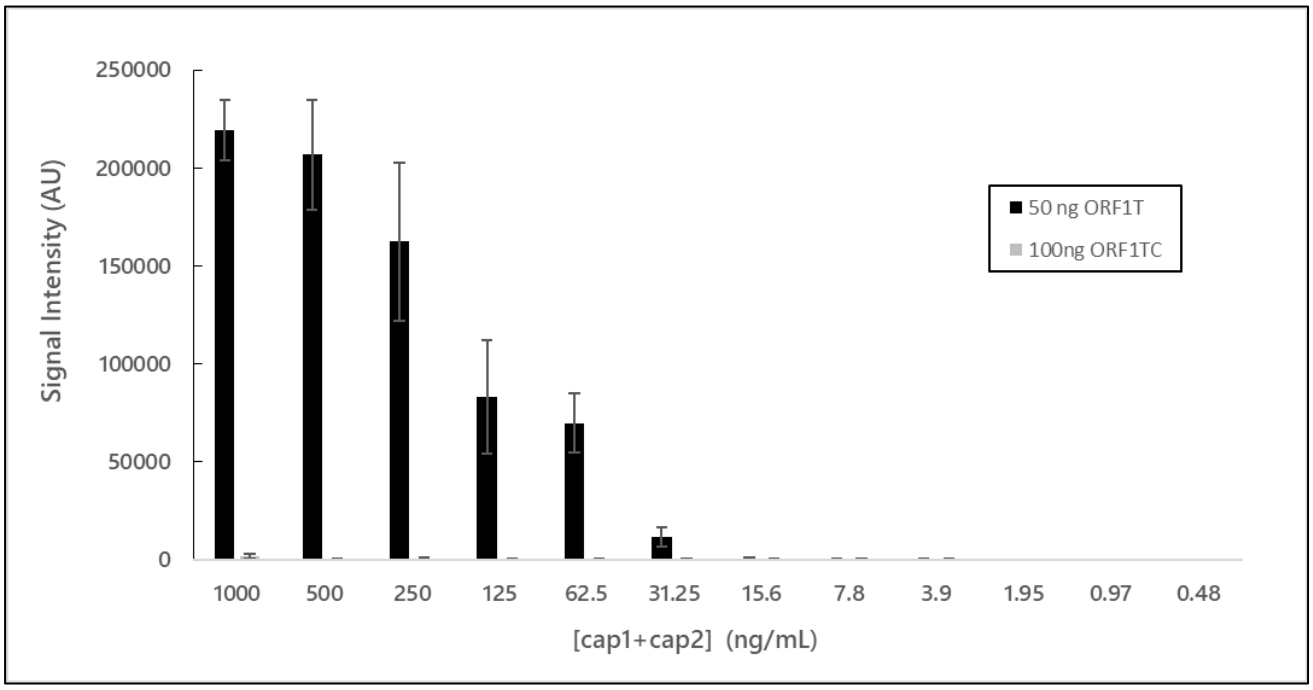


Figure 4.13: Probe calibration signal intensity measurement derived from the "cold" dot blot experiment presented in Figure 4.12.

The signal measured decreased proportionally to the concentration of biotinylated probe used. Error bars represent the standard deviation of four measurements.

In order to emulate the interaction between the probe and the viral sequences, a competition assay was performed with synthetic fragments of ORF1T and ORF1TC (Figures 4.14 and 4.15).

A set amount of ORF1T was spotted on a nylon membrane. The complementary sequence ORF1TC was spotted on at the same concentration, as a negative control. Both biotinylated probes cap1 and cap2 were added on the membrane, together with increasing concentration of either ORF1T or ORF1TC. The aim of the experiment was to evaluate the ratio of ORF1T and ORF1TC in solution necessary to inhibit the signal produced by the probes. ORF1T or ORF1T+ORF1TC were added either

at the same concentration than cap1 and cap2 or in 2, 10, 50 or 100-fold in excess. As a positive control and a reference, a lane was incubated with only cap1 and cap2.

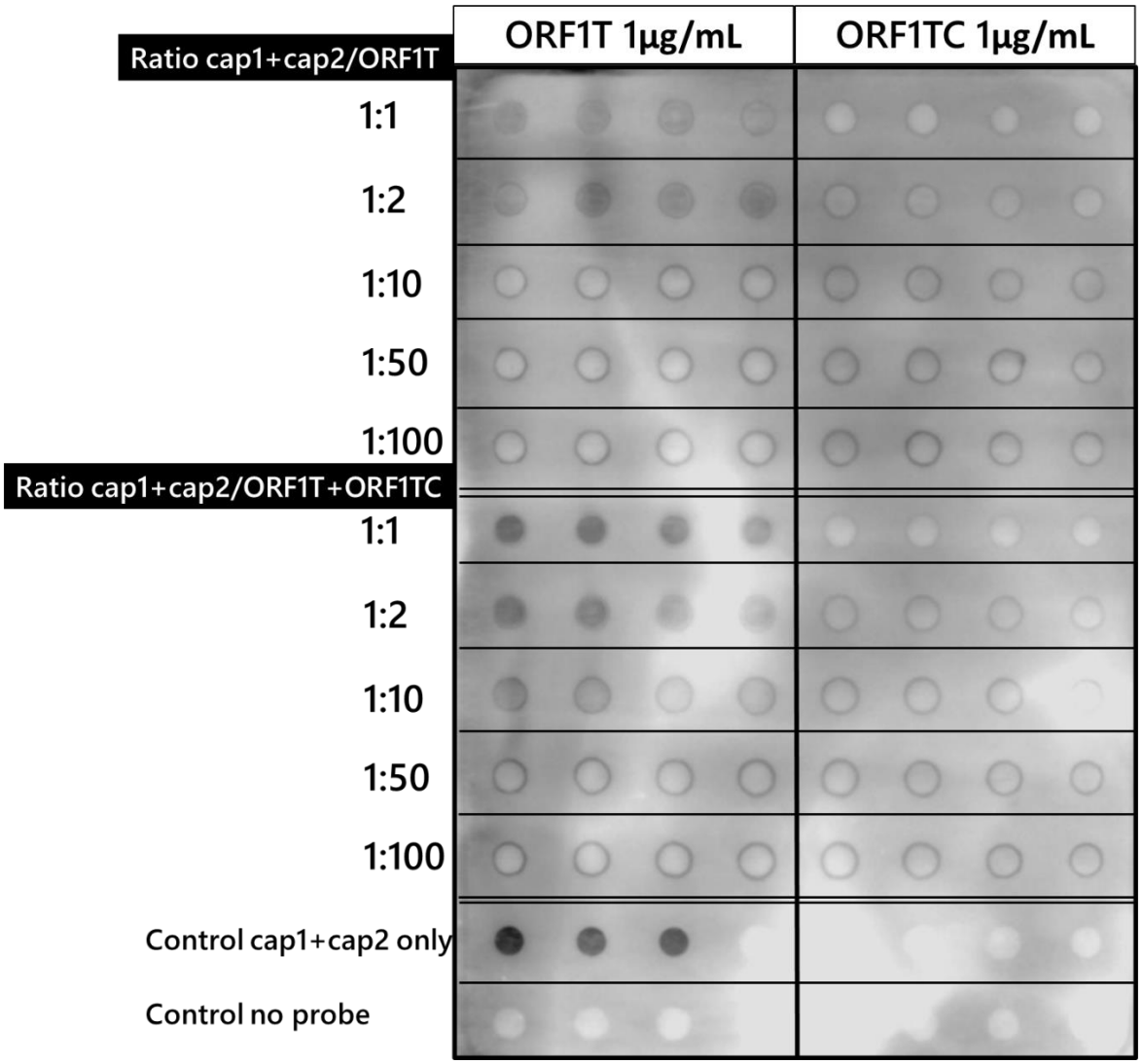


Figure 4.14: Membrane competition assay in "cold" conditions.

Synthetic ORF1T (left rows) or ORF1TC (right rows) were spotted as four replicates on a nylon membrane and incubated with either biotinylated cap1+cap2 oligonucleotides and synthetic ORF1T at a range of concentration or with biotinylated cap1+cap2 oligonucleotides and synthetic ORF1T and the synthetic ORF1TC at a range of concentration to determine the concentration required to inhibit the signal generated by the biotinylated probes visible in the control cap1+cap2 only lane.

The signal produced by the biotinylated cap1 and cap2 oligonucleotide was inhibited by adding ORF1T in solution with the probes. The signal increased when ORF1T was two times in excess compared to the probes. At ten times in excess, the signal was 4.19% of the control, 5.91% for fifty time in excess and 5.14% for hundred times in excess.

Adding ORF1TC with ORF1T in solution resulted in a slower decrease in the signal produced by the probes. When the extra oligonucleotides were added at the same concentration as the probes, the signal was reduced to 51.6%. For two times in excess, the signal was 32.84% of the control, 13.91% for ten times excess, 7.13% for fifty times and 7.11% for a hundred-time excess synthetic target and complementary sequence.

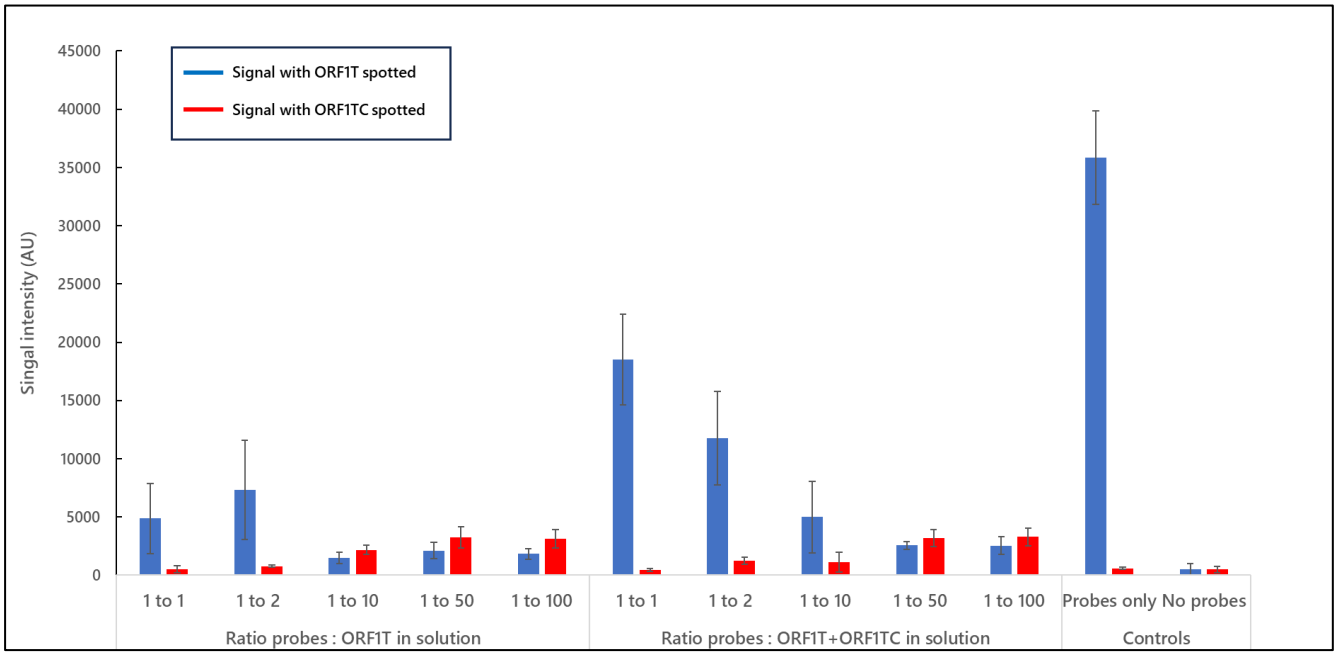


Figure 4.15: Membrane competition assay signal in "cold" conditions intensity measurement derived from Figure 4.14.

The signal generated by the binding of the biotinylated cap1+cap2 oligonucleotides to synthetic ORF1T or ORF1TC was measured using a Li-Cor Odyssey imaging system. Error bars represent two standard error of the mean of four replicates.

The combination of ORF1T and ORF1TC required a higher concentration to inhibit the signal from the biotinylated probes.

The signal-to-background ratio varied from 3.85 to 36.09 for the lanes were ORF1T was spotted. In comparison, the signal on the lanes where the negative control ORF1TC was spotted varied from 1.91 to 7.55.

4.3.6 CSSV competitive, cold, membrane-based hybridisation assay using *T. cacao* leaf extracts.

The competitive assay was repeated with extracts from CSSV-infected *T. cacao* DNA samples (Figure 4.17). The DNA was prepared using the NaOH method, where the plant leaf samples are ground with sterile sand in sodium hydroxide then introduced in the commercial DNA extraction kit (Figure 4.3).

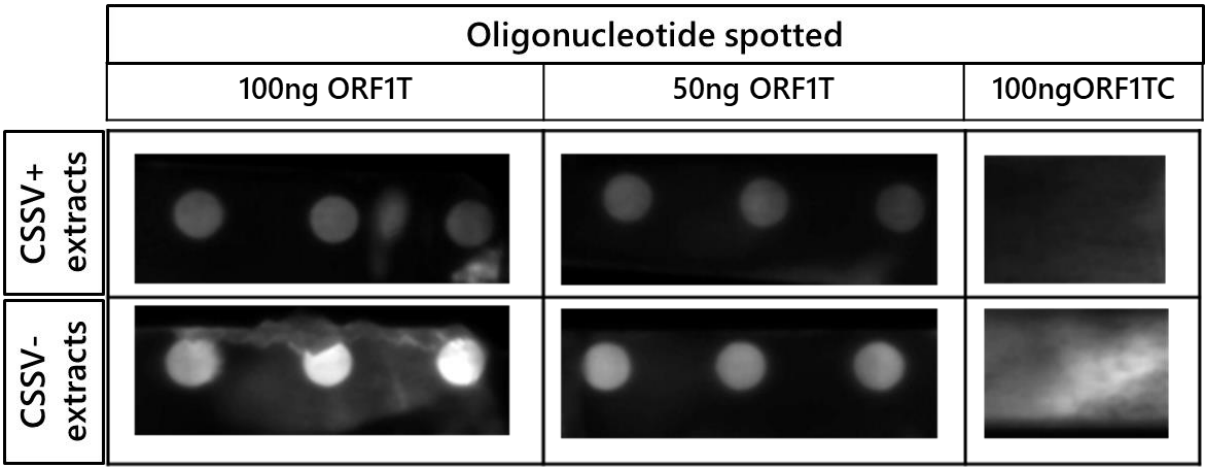


Figure 4.16: Membrane competitive hybridisation assay in "cold" conditions with *T. cacao* extracts.

Synthetic ORF1T and ORF1TC were spotted on a nylon membrane. Extracts from CSSV positive or negative *T. cacao* extracts were mixed with the cap1 and cap2 biotinylated probes, incubated on the membrane with streptavidin-peroxidase and imaged using a Li-Cor Odyssey imaging system. The signal was successfully inhibited in the CSSV positive extracts. Colours were inverted compared to those in Figures 4.12 and 4.14.

The signal produced on the membrane incubated with CSSV positive samples was lower than the signal measured on the membrane incubated with CSSV negative samples (Figure 4.16 and 4.17), for both concentrations of spotted ORF1T tested.

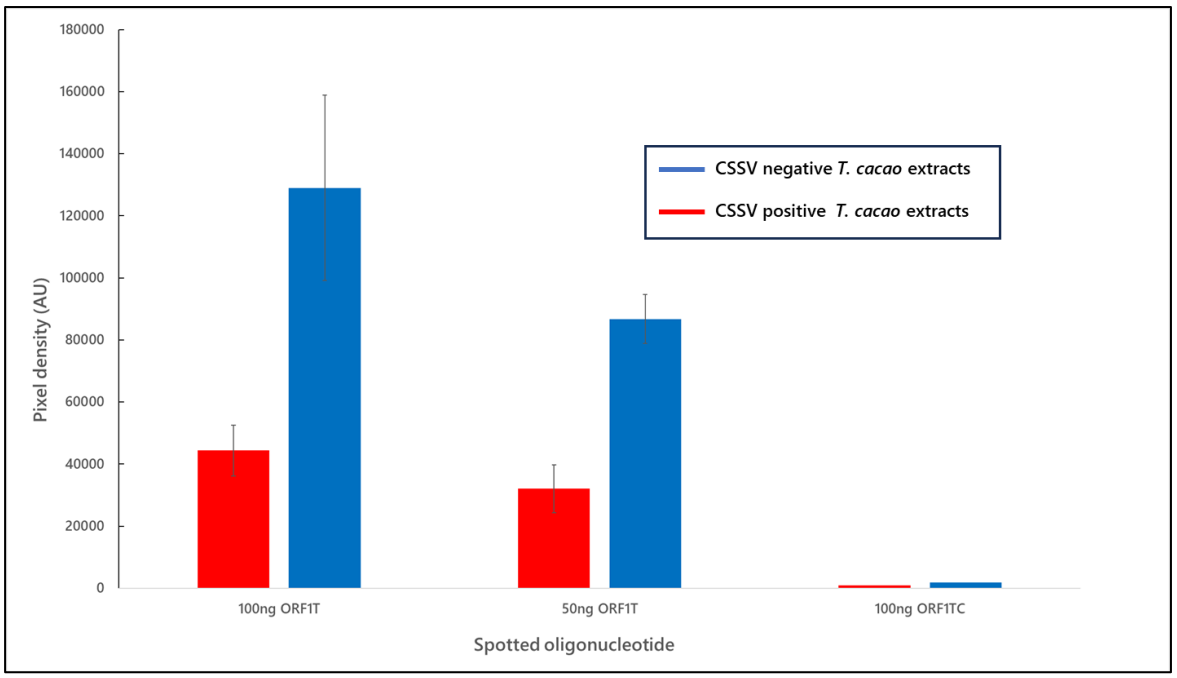


Figure 4.17: Competitive "cold" hybridisation assay with *T. cacao* extracts signal measurement derived from Figure 4.16.

The signal produced by the biotinylated cap1 and cap2 oligonucleotides was inhibited when using CSSV positive extracts. Error bars represent the standard error of measurements of three replicate spots for ORF1T. ORF1TC was spotted in duplicate.

Samples containing the extracts and probes did not produce a signal on the spot containing ORF1TC.

4.3.7 Cold hybridisation assay in solution with *T. cacao* extracts.

4.3.7.1 CSSV-infected *T. cacao* leaf samples.

Kit extracted *T. cacao* DNA samples were mixed with biotinylated cap1 and cap2 and incubated with sodium hydroxide at RT for five minutes. Samples were then neutralised with acetic acid, diluted with B/W buffer, and incubated with SA-MP. Eluates from SA-MP were used as a template qPCR amplification targeting both *T. cacao* and CSSV markers. Extracts not incubated with the SA-MP were added as a control reaction (Figure 4.18).

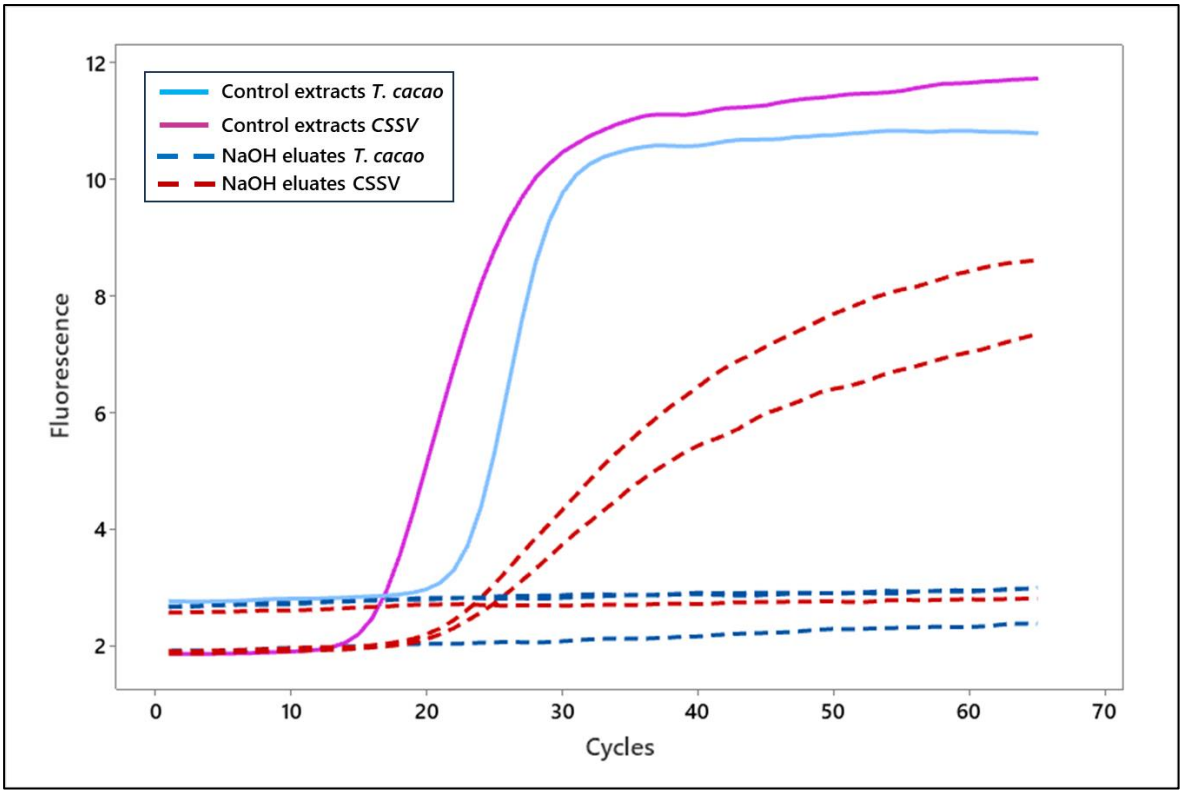


Figure 4.18: Sodium hydroxide mediated annealing of the cap1 and cap2 oligonucleotide probes.

Whole DNA extracts from CSSV-infected *T. cacao* samples were treated with sodium hydroxide, incubated with the biotinylated oligo cap1 and cap2, detected with SA-MP and amplified by qPCR in triplicates. Controls qPCR reactions were ran using either CSSV positive or negative *T. cacao* extracts with no prior incubation with the SA-MP.

No amplification was observed in the samples targeting the SSR from *T. cacao*. In contrast, two out of three samples targeting CSSV produced Ct values of 17.37 and 17.17, corresponding to 5.23 and $6.27x10^6$ copies amplified, respectively.

4.3.7.2 CSSV-infected *T. cacao* phloem extracts.

SA-MP were incubated with two types of phloem extracts mixed with biotinylated cap1 and cap2 and a qPCR was performed on the eluted fractions from the particles (Table 4.6).

Magnetic particles were added to the phloem extracted from branches macerated in either deionised water, 25mM sodium hydroxide or 1M NaCl solutions prepared in section 4.2.1.2.1. Two branches, a young branch and an older one, based on morphological characteristics, were selected on tree “61”.

Table 4.6: SA-MP detection of CSSV and *T. cacao* from phloem macerated in several extraction solutions.

	Younger branch			Older branch		
	di H20	NaCl	NaOH	di H21	NaCl	NaOH
Amplified *T. cacao* copies	7.41E+01	N/A	2.04E+02	N/A	N/A	7.84E+02
Amplified CSSV copies	6.11E+02	N/A	1.87E+03	N/A	N/A	1.23E+04
Ratio CSSV/*T. cacao*	8.34E+00	N/A	9.24E+00	N/A	N/A	1.67E+01

On the younger branch, $6.11x10^2$ copies of CSSV were amplified in samples extracted in deionised water on average. $7.41x10^1$ copies of *T. cacao* were detected on average in the same samples. The average ratio was 8.34. Samples extracted in 1M sodium chloride did not produce detectable amplicons. Samples from the younger branch extracted in 25mM sodium hydroxide produced $1.87x10^3$ copies of CSSV and $2.04x10^2$ copies of *T. cacao* on average for an average ratio of 9.24.

Samples extracted from the maceration of the older branch produced a qPCR detectable amplicon only with the sodium hydroxide extracts. $1.23x10^4$ CSSV copies were detected on average and $7.84x10^2$ copies of *T. cacao* were measured on average. The average amplified CSSV/*T. cacao* copies was 16.69.

Biotinylated cap1 and cap2 were also incubated with phloem extracted directly from the trees, as described in section 4.2.2.2.2., retrieved using SA-MP and amplified by qPCR (Table 4.7).

Table 4.7: SA-MP detection of CSSV and *T. cacao* from phloem sampled directly from *T. cacao* branches in deionised water. The calculated copies are derived from the average of three qPCR replicates from a single extraction.

	Tree "61"		Tree "8"	Tree "10"
	Branch 1 (bottom)	Branch 2 (top)		
Amplified *T. cacao* copies	5.51E+01	3.59E+01	1.44E+02	8.55E+00
Amplified CSSV copies	3.30E+03	3.75E+02	9.00E+02	1.07E+03
Ratio CSSV/*T. cacao*	5.99E+01	1.05E+01	6.27E+00	1.25E+02

Two tubes were placed on tree “61”. The average copies of CSSV amplified from the first branch near the base of the tree was $3.3x10^3$. Amplification of the *T. cacao* marker yielded $5.51x10^1$ copies. The ratio was $5.99x10^1$ on average. On the second branch located at the top of the tree, $3.75x10^2$ copies of CSSV were detected by the particles together with $3.59x10^1$ copies of *T. cacao*. The resulting ratio was 5.73 lower compared to the other branch, at 10.46. On tree “8”, $9x10^2$ copies of CSSV were detected on average and $1.44x10^2$ copies of *T. cacao*. The ratio CSSV/*T. cacao* was 8.41. An average of $1.07x10^3$ CSSV copies were amplified in the sample from tree “10”. 8.55 copies of *T. cacao* were detected in the same sample, yielding a CSSV/*T. cacao* ratio of 1070.

4.4 Discussion and conclusion.

A simple, cost effective CSSV DNA extraction method from *T. cacao* leaves was developed and compared to a standard, commercial kit procedure. The method tested used a simple mechanical grinding step using sterile sand and chemical tissue disruption using 200mM sodium hydroxide. For testing the extraction efficiency by qPCR amplification, samples were then cleaned up using the last steps of a commercial kit. Undiluted eluates produced with the Fast method did not produce an amplification signal, unlike the extracts produced from the commercial kit (Figure 4.5). This was possibly due to the presence of inhibitory substances from the plant in the eluates. PCR is a sensitive technique that relies on the enzymatic action of the DNA polymerase, which is affected by a range of physicochemical factors. Plant samples are known to produce PCR inhibitory chemicals such as polyphenol compounds, polysaccharides, or humic acid (López *et al.*, 2008). This hypothesis was confirmed by the qPCR results with the thousand-fold dilutions of the extracts. The measured Ct values were on average slightly lower with the Fast method to those obtained with the standard kit method, showing that the target DNA was in higher concentration in samples extracted with that procedure. In summary, extracted CSSV DNA was of better quality with samples from the commercial kit, but the short sodium hydroxide procedure extracted higher quantity of DNA. The Fast method could therefore be used as procedure to obtain CSSV DNA samples for using in an enzyme-independent assay. The

grinding in sodium hydroxide step could be adapted to fit a field use by replacing the microcentrifuge tube with purposedly designed pouches containing the sodium hydroxide lysis solution and the sterile sand. The company Agdia (Agdia, 2023) supplies lateral flow plant pathogen assay test with extraction bags containing a specially formulated buffer. Plant leaf samples are added to the pouch that is lined with a mesh structure. Cellular content of the plant is extracted by macerating the sample in the pouch and applying pressure using a blunt object such as a pen or a marker through the lining. A similar system could be applied to *T. cacao* leaves to extract the DNA of CSSV.

Extracting CSSV directly from phloem is not currently an established method but could have useful application in the detection of CSSV in *T. cacao*. Direct assessment of stems could be a useful tool for quarantine studies with current methods still involving grafting on CSSV sensitive material and assessment of symptoms after 3 months confirm with PCR from leaves extract (Ullah *et al.*, 2021). In the context of a field biosensor, a requirement in the design of the extraction method is the simplicity of the procedure and its cost effectiveness. The formulation of the extraction solutions from branches was developed with these prerequisites in mind. The first method used calibrated sectioned branches from *T. cacao* trees that were macerated in either deionised water, sodium hydroxide or sodium chloride solution with the aim to collect all the virus present in the sample (Table 4.4). The second method involved extraction directly from the trees using sectioned tubes containing deionised water attached to the branches (Table 4.5). These chemicals are relatively inexpensive reagents. Besides, extracting the samples with sodium hydroxide would allow to influence the denaturation of the DNA and enable to use the preparation directly with the cold hybridisation assay. Using a concentrated sodium chloride solution aimed at creating a differential in osmotic pressure to accelerate the exit of the sap from the phloem. qPCR using samples extracted in sodium chloride were however negative for both CSSV and *T. cacao,* suggesting that either no material was extracted or that the PCR was inhibited (Table 4.4). Salts are known PCR inhibitors and can affect the DNA polymerase from a concentration as little as 40 mM, depending on the enzyme (Davalieva and Efremov, 2010). It is possible that the sodium chloride was not thoroughly eliminated during the extraction step and

subsequently interfered with the qPCR. In contrast, the highest CSSV copy number was achieved from the samples macerated in deionised water, which produced 2.61 times more copies than with the sodium hydroxide extraction. Furthermore, the *T. cacao* copy number measured within these samples was also the lowest, resulting in a CSSV/*T. cacao* ratio 25.58 times higher than with sodium hydroxide, which makes deionised water a suitable candidate for CSSV extraction from branches. This is the first evidence of detection of CSSV from stems in *T. cacao*. Besides, when extracting sap directly from branches, the variability of copy numbers observed between samples from the same tree was high which suggest differences in virus load across the tree (Table 4.5). This could be caused by experimental variation from the volume of phloem collected across branches but could also indicate true differences previously observed when testing different leaves from the same tree. For example, the difference in CSSV copy number between the lowest and highest averaged qPCR replicates from tree "61" was 12.76. Movement of CSSV within *T. cacao* is currently not well characterised and it has been hypothesised that the virus can form vesicles (Allainguillaume, J., personal communication) using plant lipids that would transport its non-enveloped DNA and ultimately spread the infection within the tree. Higher levels of CSSV obtained from SA-MP on crude phloem samples in water compared to sodium hydroxide suggest that CSSV travels through the vascular tissue as DNA and not in capsid.

While the quantity of viral DNA collected from the phloem may not be compatible with an amplification-independent biosensor, it is still valuable to know that the virus can be extracted directly from branches. Both phloem sampling procedures were simple and did not require complex equipment, and the amount recovered can be detected by PCR. Having a methodology to sample phloem from calibrated sections of branch, enabling the quantification of the virus load at any given location on the tree is therefore a beneficial tool to understand the physiopathology of CSSV. Besides, the sap collection procedure could be potentially improved by using the EDTA-facilitated exudation method. This method, developed in 1974 (King and Zeevaart, 1974) aims at enhancing the accumulation of sap by preventing the self-repair mechanisms (Oparka and Santa Cruz, 2003) from the plant consecutive to the wound generated at the point of sampling.

The ability of a set of biotinylated oligonucleotides targeting several loci within the genome of CSSV were tested using a pull-down assay. DNA extracted from leaf samples were incubated with the biotinylated probes, then streptavidin-coated magnetic particles were used to retrieve the DNA captured by the oligonucleotide assembly. All probes tested were successful in detecting the virus, using either samples from leaves or from phloem extracts. Using the pull-down assay, it was possible to enrich the original extracts several thousand-fold, under certain washing and annealing conditions (Figure 4.11). It was however noticeable that if the reaction parameters were not set precisely, unspecific binding of DNA can occur, as demonstrated by the calculated ratio amplified CSSV/*T. cacao* copies (Figure 4.10). Interestingly, cap1 and cap2 returned no hits when NCBI blast search were performed on *T. cacao* genome, unlike the probes 2143ROI and 2143ROIC, that were found in chapter 2 to bind to both the virus and the tree. It is possible that a portion of the DNA from both organisms interacted with the SA-MP directly in a non-specific manner, rather than with the probes. This was verified when the stringent wash procedures were introduced, which resulted in hundred- or thousand-fold increases in the detected virus/plant copies ratio (Figure 4.11). The pull-down assay was here utilized to prove that the oligonucleotides selected were able to bind to CSSV extracted from plant samples and the results obtained proved that the oligonucleotides could bind to the virus more specifically.

A membrane-based cold neutral hybridisation assay depending on the binding of the biotinylated probes cap1 and cap2 to CSSV was developed (Figure 4.12). The aim was to provide the foundation for the biological recognition element of a biosensor to detect CSSV. The oligonucleotides were first proven to bind specifically and, in a dose-dependent manner, to the synthetic region of interest from the virus. A competition assay was then designed. The assay relied on the signal generated by the binding of single-stranded, virus specific oligonucleotides, cap1 and cap2, in solution to a single-stranded synthetic region of interest from the genome of CSSV, ORF1T, spotted on a membrane. A competition assay wherein this signal is inhibited by adding increasing amount of ORF1T in solution was successfully implemented. Using synthetic nucleic acid, it was found that the signal was decreased

by more than 90% when the targeted region and its complementary sequence in solution was a hundred-fold in excess compared to the spotted region of interest. Similarly, the signal was reduced by more than 60% when both competing oligonucleotides were two-fold in excess in solution, demonstrating that the signal reduction was dose-dependent (Figures 4.14 and 4.15). In chapter 3, a high-resolution melt analysis model was developed to investigate the interaction between cap1, cap2, ORF1T and ORF1TC. The four nucleotides were mixed either in equimolar proportions or with decreasing ORF1TC concentrations. It was then found that a difference in the melting profile could be observed starting from a two-fold dilution of ORF1TC, and the profile gradually shifting to the curve shape comparable to the control containing ORF1T and ORF1TC alone when ORF1TC was diluted fifty times. The results observed in the competitive assay confirms that the displacement of the ORF1T/ORF1TC equilibrium can also be measured using a membrane-based procedure.

The membrane assay was then performed with *T. cacao* whole extracts and signal inhibition was observed with infected samples but not with PCR-confirmed CSSV-free extracts. Using either 50 or 100ng of spotted ORF1T, the signal produced by the cap1 and cap2 biotinylated probes was successfully inhibited (Figure 4.16 and 4.17): this indicated that a portion of the labelled cap1 and cap2 probe bound to the viral sequences from the samples and not on the synthetic ORF1T spotted on the membrane. One of the challenges from using the biotin system with the nylon membrane was inconsistent background signal. Patchy unspecific signal was observed in several instances, causing variability that was detrimental to the quality and consistency of the measurements. It would be beneficial to further improve both the blocking and the washing procedure, by testing several blocking agent and formulation of washing buffers(Heda *et al.*, 2020). Besides, the sensitivity of the assay could be increased by using fluorescent oligonucleotides in place of biotin. Preliminary work was undertaken to either using cap1 and cap2 labelled with the ATTO fluorophore, or by using a self-assembled "Christmas tree" scaffold of cap1 to generate an amplified signal.

A variation of the pull-down assay using sodium hydroxide was conducted using plant extracts (Figure 4.18). Sodium hydroxide was added to a mix of CSSV-infected *T. cacao* extracts and the biotinylated probes with the aim to prove that the chemical would denature the DNA enough to allow the oligonucleotides to reach their target when the solution was neutralised with acetic acid. SA-MP were then added to the mixture and a qPCR was run using the eluates. CSSV was amplified in two out of three replicates, while the nuclear marker from the tree was not. This demonstrated that a solution of diluted sodium hydroxide can be used in lieu of temperature to enable denaturation of the DNA and the subsequent hybridisation of oligonucleotide probes to the targeted region. This experiment confirmed the data obtained in chapter 2, where it was demonstrated that incremental increase in sodium hydroxide concentration was able to destabilise the cap1/cap2/ORF1T/ORF1TC scaffold, shifting from a double-stranded structure to a single-stranded profile melt (Figure 3.17). Importantly, the destabilisation could be reverted by neutralising the medium with acid, showing that the action of the sodium hydroxide, while effective, is not detrimental to the quality of the DNA in this concentration range. Together with the data from the leaf's extraction experiments, this proves that it is possible to first extract the DNA of CSSV using sodium hydroxide, while preparing the sample for hybridisation with labelling probes in one single step for using in a competitive assay.

Furthermore, the method was tested with phloem samples, obtained either by macerating branches sections in a range of solutions, or by attaching centrifuge tubes containing deionised water directly to the infected tree branches (Tables 4.6 and 4.7). The main challenge of using the magnetic particles was the consistency of the sap. The samples had a gelatinous consistency that impeded the dispersion of the particles in the sample. In a pull-down assay, SA-MP are usually collected rapidly by placing a microcentrifuge tube on a magnetic rack. This was not possible with the samples extracted from the phloem, and the sap had to be incubated with the magnet for 18hours so that the particles would gather on the wall of the tube. While the copy number recovered from phloem with the SA-MP were lower than with the Kit extracted leave samples, it was still possible to detect the virus from unprocessed preparation which indicate that the approach can work on crude extracts. Even if the

approach is not sufficient to be used as a solely hybridisation detection system, it could enable sole detection of CSSV as part of a PCR approach, avoiding genomic integrated nonspecific targets.

Chapter 5 General discussion and conclusions.

5.1 Introduction

Cacao swollen shoot virus is the main threat to the cacao production in west Africa, where most of the production is occurring. Since the identification of the causal agent more than eighty years ago, attempts to eradicate the virus have failed. Plans to contain the infection involve cutting the affected tree and the trees surrounding met little success and the virus is still epidemic in the area. The reasons for these failures are multifactorial. First, socioeconomical factors of the growers is a key issue. 70% (Carla Veldhuyzen, 2019) of the growers are subsistence farmers. Governmental compensation may be delayed and/or not cover all the expenses associated with removing the trees affected by the disease. Methods to detect the virus on the field primarily rely on visual identification by trained personnel. While effective, several factors may hamper the timely identification of the virus.

The objective of this work was to develop biological recognition elements of a biosensor to detect CSSV in the field, together with a simplified DNA extraction procedure and a cold hybridisation assay to detect the DNA. The overall design of the biosensor was created with two main key requirements. The biosensor needed to be cost-effective, to fit with the limited resources available in the virus endemic area, and simple to operate, so that training needs of the personnel involved could be minimal. At the time of the design of the project, eight genome sequences of CSSV had been published and no antibodies to the virus were commercially available. In consequence, the genome of the isolate of the CSSV New Juaben strain maintained at UWE was sequenced de novo and specific BREs were developed.

5.2 Review of the results and perspectives for further work

Two biological recognition elements (BRE), or bioreceptors were developed in this thesis. The BRE is a critical part of a biosensor that acts as an interface between the user and the targeted analyte. The role of the BRE is to physically bind to a selected biological analyte that is specifically associated with a condition that the user require to investigate. The biological analytes vary vastly in nature (Loo et al., 2023), and can be nucleic acids, including DNA, RNA or microRNAs, proteins, ions, or drugs. BRE on the other hand are either protein-based or nucleic acid-based. For this project, both types of BRE were developed.

5.2.1 ZFP BRE

The protein-based molecule selected as BRE to detect CSSV in this thesis was a synthetic chimeral zinc finger protein (ZFP). ZFP are well characterised regulatory proteins with one or multiple domains each able to bind to a trinucleotide target. *In silico* models were developed to decipher the binding alphabet generated by the diversity of zinc finger domains. Online tools enable the generation of unique proteins potentially able to bind to any DNA molecule. A candidate BRE based on such protein was produced and characterized in this project with the aim to detect CSSV. A region of interest within the genome of CSSV was identified and used to screen an online library of ZFP domains. A gene encoding the candidate ZFP, ZFP2143K, was then selected, cloned, expressed in *E. coli,* and purified using immobilised affinity chromatography. The functionality of the resulting ZFP2143K was characterized by electrophoresis mobility shift assay (EMSA), double filter binding assay (DFBA) and in a CSSV-infected *T. cacao* DNA extracts pull-down assay using streptavidin-coated superparamagnetic particles followed by qPCR. Selection and production of the recombinant protein was relatively easy to implement, unlike the functionality characterization process, that was longer and required optimisation of a range of techniques. ZFP2143K was found to detect the targeted CSSV region of interest, that spanned a region overlapping ORF2 and ORF3 on the genome of CSSV. However, pull-down assays using kit extracted CSSV DNA showed that the protein was also binding to the *T. cacao*

genome. EMSA data also demonstrated that the protein could bind to a DNA fragment containing the sequence targeted by two of the six domains of the recombinant. The functionality of the protein could have potentially been improved by removing the hexahistidine tag, that is required for the purification step, as it may interfere with the three-dimensional structure of the recombinant (Zhao and Huang, 2016; Chant *et al.*, 2005) which in turn may impact the binding ability of the ZFP. To investigate further, several strategies could be employed: the hexahistidine tag could be placed on the N-terminal end of the protein, since the position of the tag may also have an influence on the structure and the function of a recombinant (Aslantas and Surmeli, 2019; Booth *et al.*, 2018). Another possibility could have been to use a different tag for purification or use a strategy whereby the tag is removed after purification. One of the difficulties when choosing a tag is that it is not possible to predict the impact it would have on the structure and functionality of the protein, either positive or negative, and this has to be investigated on a case-by-case basis.

In order to optimise the chances to select potential ZFP for biosensing, a strategy could involve selecting several candidates using cross-referenced data from the online tools available. Concomitant recent advancements in artificial intelligence and transcription factor genomic data have led to the development of machine learning models (Ichikawa *et al.*, 2023) able to deepen the understanding of the mechanistic underlying the interaction of zinc finger protein with their cognate DNA partners. While primarily developed for integration with gene-editing tools such as clustered regularly interspaced short palindromic repeats/ CRISPR associated protein (CRISPR/Cas) or Transcription activator-like effector nucleases (TALEN) (Bhardwaj & Nain, 2021), ZFP designed to bind CSSV specific sequences are a promising and versatile type of BRE prospect. A mid-throughput ZFP expression platform could be designed to first screen candidates ZFP *in silico* based on the viral sequences of interest (Figure 5.1).

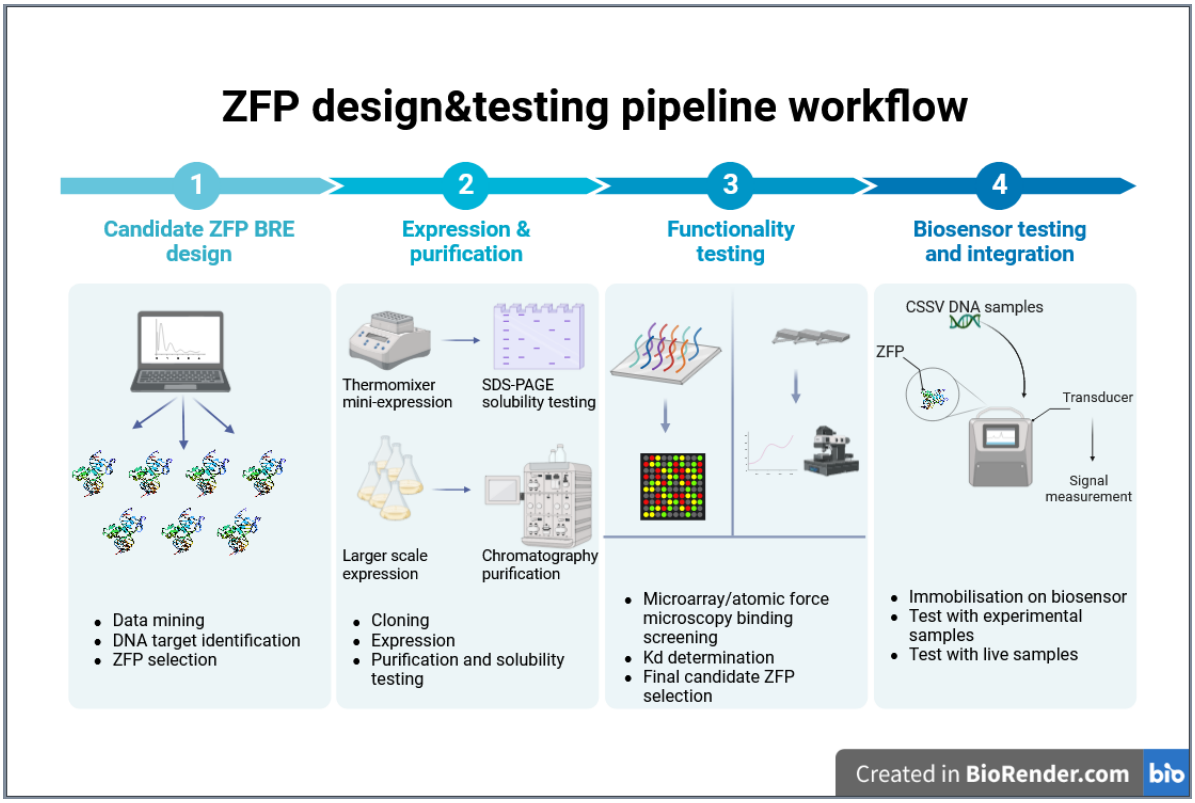


Figure 5.1: Prospective ZFP design and production workflow.

Alternatively, to an algorithmic-based method, candidate ZFP could also be selected *in vitro*, using phage display technology, in which a library of zinc finger proteins cloned into the bacterial viruses that are able to express the recombinant protein on their surface is screened. Phage display has been successfully used to select ZFPs (Isalan, Klug and Choo, 2001). When choosing the *in silico* ZFP selection procedure, a set of candidates could then be cloned in rapid cloning and expressed in the *E. coli* BL21(DE3) cells that were used in this project, as they are adequate to express a range of ZFP from a variety of organisms (Lenz *et al.*, 2021; Lata, Afsar and Ramachandran, 2017; Qu *et al.*, 2014; Ericsson *et al.*, 2006; Archdeacon *et al.*, 2000). Solubility of the chimeric proteins would be analysed by SDS-PAGE (Taylor, Denson, and Esposito, 2017) to select candidates suitable for larger scale production with the expression system considered. Soluble recombinant ZFP would subsequently be tested for their functionality, which is their ability to bind to the targeted viral region of interest (Cassandri *et al.*, 2017; Klug, 2010; Vrana *et al.*, 1988). The key factors to consider are the specificity and the affinity of the DNA-protein interaction. Theses parameters would require mid- to high- throughput screening.

As was demonstrated in the current project, EMSA, DFBA and pull-down assays would be too fastidious for a study of this scale. Amongst the alternative screening methods, protein microarrays could potentially curb both scale and time constraints. Protein binding microarrays (Berger *et al.*, 2006; Mukherjee *et al.*, 2004) are high-throughput assays aimed at characterising DNA-protein interactions. They have been used successfully to analyse the binding of ZFP to transcription factors (Bulyk et al., 2001; Zeitler et al., 2019). The arrays would feature immobilisation of targeted viral sequences on the surface of the chip, including variants to either test the affinity of the ZFP candidate affinity spectrum towards a particular CSSV sequence or a range of specificity towards CSSV isolates. Washes would remove unspecific interaction, and a fluorescent-modified antibody targeting a tag introduced on the recombinant ZFP would allow the measurement of the strength of the interaction between DNA and protein. Bulyk et al. (Bulyk *et al.*, 2001) were able to correlate the fluorescence measured with the equilibrium dissociation constant (Kd) of the proteins, which informs about the suitability of the ZFP for detecting the virus. Atomic force microscopy (AFM) could be used as a second line of characterisation of the interaction between the protein and the DNA. The technique is based on a microcantilever attached to a spring mechanism. Biological analytes are absorbed at the tip of the cantilever. Upon binding of the analyte to its ligand, a tension is created in the cantilever that can be measured through a laser beam reflecting on a photodiode (Johnson, Hilal, and Bowen, 2009). AFM allows to visualise the interaction of the two biological molecules and can also provide parametric information such as dissociation constants (Yang, Wang, and Erie, 2003) that are of particular interest to characterise a biological recognition element. This technique has been used (Mawhinney *et al.*, 2018) to characterise the binding of the CTCF transcription factor, a eukaryotic eleven zinc finger domains protein involved in several cellular processes such as chromatin architecture regulation or transcriptional activation (Kim, Yu and Kaang, 2015), to its cognate DNA sequence. While AFM may not be adapted to screen multiple ZFP, it could be used on a pre-selection of protein. Furthermore, efforts are made to increase the analytical capabilities of AFM, and novel multiplexed instruments using arrayed cantilevers are in development (Yang *et al.*, 2019) and could offer an interesting means

to study the binding abilities of several selected variants of DNA target simultaneously. Candidate ZFPs with adequate functional performances could then be brought to the next step of the process, which is testing the protein with live samples within the biosensing platform. Interestingly, Kim (Kim *et al.*, 2011) found that the association of ZFP with the highest affinity does not necessarily result in the highest sensitivity in a live assay. This is due to the complex interaction between these DNA-binding proteins. Therefore, deployment of a set of ZFP able to bind to the targeted viral sequence with a range of affinities as determined by the microarray/AFM would be highly beneficial to select an optimal combination of protein BREs.

Interestingly, a zinc finger motif has been commonly described in many badnaviruses (Sharma, Kumar, and Baranwal, 2014). As an alternative to artificial ZFP, an internal peptide containing a gag-knuckle domain was identified during this project in the genome of CSSV by *in silico* analysis of the strain available at UWE. The peptide was commercially synthesized and analysed by circular dichroism (CD) at the university of Bristol to analyse its three-dimensional structure. EMSA were carried out using the peptide with fragments of CSSV to identify potential interactions, but none could be evidenced in the conditions tested (data not shown). It could be interesting to identify the cognate partner of this peptide to improve our knowledge of the pathogenicity mechanism associated with CSSV infection., since gag-knuckle domains may be involved in nucleic acid binding (Krishna, Majumdar and Grishin, 2003b). Furthermore, the peptide could be used as a BRE, providing that the cognate partner of the peptide is a viral element, either a protein, another peptide, or a nucleic acid.

5.2.2 DNA BRE

The second class of BRE developed in this project were nucleic acid fragments derived from the genome of CSSV. The whole genome sequence of the isolate of CSSV used in this study was amplified by PCR using eighteen primer pairs to produce overlapping fragments. Amplified fragments were analysed by agarose gel electrophoresis, excised from the gels, purified, and sent off-site for Sanger sequencing. This step was not challenging and provided a framework to design a DNA biological

recognition element to detect CSSV. Using DNA molecules as the BRE of a biosensor presents considerable advantages. First, production of short nucleic acid probes is cost effective, which is a key factor when resources are of importance. Besides, they are easy to produce, allowing fine tuning of the spectrum of sequences targeted conveniently: CSSV isolates circulating in west Africa exhibit genetic variability associated with different degree of pathogenicity (Muller *et al.*, 2018) and being able to rapidly adapt to the variants of interest is crucial.

The interaction between the oligonucleotides composing the DNA BRE was analysed by high resolution melt analysis (HRM). HRM is usually utilised for the genotyping of DNA polymorphisms. Here, it was used to model the interaction of oligonucleotide probes to develop cold hybridization assays. HRM experimental data produced in this project were reproducible and accurate to determine the strength of the binding between the probes and the targeted viral region. HRM confirmed a dot blotting assay which showed that the bond between the cap1 probe and ORF1T was stronger that between cap2 and ORF1T. Melting curves were established for each combination of oligonucleotides and were well replicated within the runs and in between runs. Furthermore, the Rotor-Gene Q permitted the analysis of sodium hydroxide-mediated fragment denaturation. Preliminary fluorescent quantification experiments using the instrument demonstrated that exposure to sodium hydroxide was not detrimental to the fluorescent Syto9 dye use in the tests at the concentration tested (data not shown) and differences in the concentration of sodium hydroxide correlated with modification in the thermostability of the double stranded fragments. HRM experiments also showed that fragments denatured with sodium hydroxide and renatured using acetic acid would exhibit the same melting temperature than fragments not exposed to the chemicals.

5.2.3 Cold hybridisation procedure

HRM data suggested that a “cold” hybridization, that relies on chemical rather than temperature to manipulate the structure of double stranded DNA fragments was feasible. Biosensors relying on DNA hybridisation have been developed for a range of detection purpose. Opting for a DNA

hybridisation-based assay was in part dictated by the cost requirements of the biosensor developed for this project. While PCR is a faster method, it comes with higher costs, and could produce false positive results when detecting CSSV from infected *T. cacao* plants because of the presence of integrated viral sequences in the tree (Muller *et al.*, 2021b). The recent discovery of the existence of CSSV-related sequences integrated with *T. cacao* genome was an essential finding in the understanding of the detection of the virus. Inadequate selection of the DNA-based BRE could indeed result in false positives detections, impeding the success of containment of a CSSV infection in the field.

Furthermore, hybridisation assays have been found to perform as well or better than PCR methods for the detection of pathogens in some cases (Potter, Nakhla, and Maxwell, 2003; Scaletsky et al., 2002). For these reasons, a membrane hybridisation strategy was chosen for developing a CSSV biosensor. In this project, the eighteen hours-long probe/target incubation time usually associated with DNA hybridisation assays was shortened by developing a so called “cold” hybridisation assay whereby denaturation/renaturation of the DNA is manipulated using chemicals rather than temperature. A range of chemicals can be used to influence DNA denaturation: dimethyl sulfoxide (DMSO)(Xu *et al.*, 2022), formamide (Blake & Delcourt, 1996) or sodium chloride (Rizzi et al., 2017) all interfere with DNA structure and the melting profiles of the double-stranded molecules. The alkali-induced DNA denaturation used in this project has been previously demonstrated (Ageno, Dore and Frontali, 1969b) and presented the advantage of doubling as a DNA extraction agent, to simplify the biosensing procedure. Experimental data gave strong indication that an alkali induced denaturation of the samples could be used instead of temperature modification for modifying the structure of the DNA of CSSV. Sodium hydroxide (NaOH) treated CSSV-infected *T. cacao* kit extracts were successfully detected by the cap1 and cap2 oligonucleotides probe using streptavidin-coated paramagnetic particles (SA-MP) and qPCR, while *T. cacao*’s DNA was not. Furthermore, it was possible to detect CSSV from crude NaOH extracts samples from vascular tissues of the tree using the oligonucleotide probes captured by the SA-MP. This demonstrated that sodium hydroxide can be used for denaturing the DNA

of CSSV and hybridising an oligonucleotide probe without relying on temperature modification. The "cold" hybridisation assay that was developed in this project using HRM was tested successfully on synthetic single stranded probe and target DNA, and on CSSV-infected *T. cacao* extracts using the "NaOH method" (sodium hydroxide lysis followed by a kit column clean-up). Because of time constraints, it was not possible to test crude extracts with this hybridization procedure. Further work is therefore required to validate the assay fully. Interestingly, the DNA used in the "cold" competition used *T. cacao* extracted with the "NaOH procedure" and was not heat-denatured. Nevertheless, a difference between CSSV positive and negative samples was observed, regardless of the quantity of ORF1T target spotted (Figure 4.16) which suggests that the detection probes cap1 and cap2 were able to access the complementary sequence to ORF1T within the genome, inducing a reduction of the signal in the CSSV positive samples compared to the CSSV negative samples. For the probe to bind to the viral sequence, the DNA of CSSV must be single stranded. At this stage, it is difficult to explain the presence of single stranded DNA in the samples. This could be due to the sodium hydroxide introduced during the extraction and would imply that the column clean-up step did not enable renaturation of the DNA. Alternatively, it is possible that CSSV DNA is present as a single stranded molecule in the leaves of infected *T. cacao*, which is suspected but has not been formally demonstrated (Allainguillaume, J., personal communication). This hypothesis is supported by the cold hybridisation performed with crude extracts from vascular tissues of *T. cacao*. Two types of extraction were carried out: one using calibrated sections of branch from the tree macerated in either deionised water, a solution of sodium chloride or a solution of sodium hydroxide; the second using tubes connected directly to a live tree to collect the sap in deionised water. The cap1 and cap2 probes were added to the samples with no other treatment and incubated at room temperature. SA-MP were then added to the samples and recovered using a magnet. Subsequent PCR on the SA-MP eluates were able to detect CSSV on all samples analysed, except for the sodium chloride extracts where no amplification was observed. Hybridisation of the capturing probes in samples extracted in deionised water suggests that CSSV DNA was present as a single stranded molecule. This has important implications for the

further development of the assay to be included in a CSSV biosensor, since prior denaturation of the samples may not be necessary. Deepening the knowledge of the pathophysiology of the CSSV infection would be useful to determine the proportion of single to double stranded DNA molecules in an infected tree throughout the course of the disease.

5.2.4 Extraction methods

Extractions methods were designed with the same requirements as the assay. The procedure had to be cost effective and relatively simple to accommodate the technical limitations of a use in the field. One of the first methods tested for this project was inspired by the work of Ikeda (Ikeda *et al.*, 2012), that developed a fast method to extract DNA from rice leaves for a subsequent PCR microsatellite analysis. The procedure consists of immersing "a few pieces" of leaves in a microcentrifuge tube in Tris-EDTA (TE) buffer, grinding with a disposable pipette tip, placing the tube in boiling water for 15 minutes, diluting in more TE buffer, centrifuging the cellular debris, and using the supernatant as a PCR template. The procedure was applied to *T. cacao* leaves, but the PCR was negative (data not shown). It was found during this experiment that *T. cacao* leaves were resistant to the disposable pipette tip grinding procedure. Indeed, *T. cacao* leaves are leathery and in the subsequent experimentations, a sterile sand grinding step was added to provide sufficient mechanical force to destabilise the tissue structure. The combination of sodium hydroxide with sterile sand could consistently extract the DNA from CSSV from the leaves of the tree. The yield of the extraction was compared to standard kit method. Several variations of the kit method were tested. All these methods started with a liquid nitrogen lysis of the leaf sections, followed by incubation with RNase A and for some variation of the procedure with proteinase K, a broad-spectrum serine protease of the family of the subtilisins (Saenger, 2013) that destabilises the cellular structure, and ended with the detergent lysis provided by the kit. These procedures were used as a golden standard of extractions, aiming at extracting as much DNA as it was possible. The "fast method" developed to produce sodium hydroxide denatured *T. cacao*/CSSV DNA extracts compared to the golden standard methods (Table 4.3) and could be envisaged as a routine extraction method to produce samples for CSSV detection.

Furthermore, titration experiments were performed to measure the quantity of acetic acid required to neutralise sodium hydroxide solutions. The titration measurements were then transposed to evaluate the quantity of acetic acid to add to samples containing synthetic ORF1T and ORF1TC. The pH of the neutralised solution was then verified by blotting drops of the solution on pH indicator and verifying the calculation of the titration. This would be useful for designing the procedure of renaturation destined for a biosensing kit in the field.

Absolute quantification of CSSV in the extracted samples was challenging first because no calibrated standards are commercially available yet and also because the extractions employed sodium hydroxide to produce crude samples, that contain a range of plant biochemicals with some such as polyphenols or acidic polysaccharides known to interfere with PCR (Rezadoost, Kordrostami and Kumleh, 2016; Demeke and Jenkins, 2010) which is used for quantification. Using a quantification technique, qPCR, potentially hampered by the chemical used or produced during the extraction development was an issue. Samples were therefore diluted and quantified indirectly by comparing with extracts obtained from commercial kit extraction methods.

As an alternative to leaf samples, this project demonstrated that CSSV could be extracted from *T. cacao* vascular tissues. CSSV was recovered from calibrated branch sections, or from sap exudates directly collected from the tree in a centrifuge tube. So far, CSSV has been detected by PCR from leaf samples (Muller, Jacquot and Yot, 2001; Jacquot *et al.*, 1999), pollen (Ameyaw *et al.*, 2013) and seeds (Length, 2010; Quainoo, Wetten and Allainguillaume, 2008) but not from vascular tissue. This result has implications for the CSSV quarantine procedure. The current protocol involves grafting a branch of the suspected tree to a PCR-confirmed CSSV negative tree and waiting for the symptoms to eventually appear on the grafted tree. This procedure is complex and lengthy, and being able to screen samples directly using biotinylated oligonucleotide probes and magnetic particles would simplify the method and represent a significant saving of time. The data from chapter 4 (Table 4.4 and 4.5) demonstrated that CSSV was detected on infected trees 61, 10 and 8. Preliminary data (Table 4.5)

suggested that different quantities of the virus were detected depending on the branch sampled; it would therefore be advisable to sample several areas of the tree to maximise the chances of detection. Besides, the method where sap exudate was collected was not optimised during this project, because of time constraints. Indeed, cutting a branch elicits a response from the tree, and repair mechanisms are put in place to preserve the integrity of the organism, such as callose synthesis, whereby callose, a β-1,3-glucan polysaccharide, is deposited on the cells wall of wounded tissues (Chen and Kim, 2009), or P-protein plugging, where proteins from the phloem form aggregates in the sieve plates of the tree, to prevent further sap loss (Azizpor *et al.*, 2022). These mechanisms are put in place rapidly upon sectioning of the branch and it is probable that the collection of *T. cacao* sap using deionised water in this project was not optimal. A range of chemicals have been used to optimise sap collection, such as citric acid, β-aminoethyl ether and EDTA(King and Zeevaart, 1974). EDTA, by chelating divalent metallic anions, can prevent the self-repair mechanisms (Dinant *et al.*, 2010) and allow collection of sap for a prolonged period while inhibiting nucleases that could compromise the integrity of the nucleic acids recovered. The main experimental challenge encountered while processing the sap exudate from *T. cacao* was the viscous consistency of the samples, that was apparent during collection of the sample (Figure 5.2, left). This was an issue when using the magnetic particles because the viscosity of the samples greatly hampered their dispersion in the samples (Figure 5.2, right).

Figure 5.2: (Left) CSSV-infected *T. cacao* exudate collection from a sectioned branch showing the viscosity of a representative sample. (Right) Streptavidin-coated superparamagnetic particles cushion formed by the viscosity of the sap exudate sample.

Dilution of the samples and vigorous mixing eventually permitted the dispersion of the particles, however their collection using a magnet was difficult, and while usually a rapid process, this step required an incubation of several hours. A protease/hydrolases treatment prior to incubation with the SA-MP could improve the fluidity of the samples by degrading the protein and polysaccharidic fraction of the samples responsible for the excessive viscosity. The sap exudate collection for screening tree samples during quarantine procedure could be further developed. Sap extraction devices have been constructed previously for the analysis of the vascular tissue content of wheat(Veselova *et al.*, 2006) or cucumber (Li *et al.*, 2022), with a difference in the quantity and quality of the analytes recovered with the instrument compared to the EDTA collection. A procedure adapted to *T. cacao* could be designed whereby a puncture device such as needle pre-treated with EDTA to prevent protein-G coagulation is connected to a pump and introduced in the vascular tissue of the trees, at distinct locations, to collect the sap into a centrifuge tube. The tube could be detached, and the sap diluted in a specifically formulated applied onto the surface of a filter spotted with anchoring oligonucleotide specific to CSSV using a syringe. Filters could then be hybridised with a detection probe to visualise the presence of the virus directly.

5.3 Biosensor integration.

Extractions methodologies, two types of BREs and a competition assay were developed for the design of biosensor detecting CSSV during this chapter. The next step would focus on the reporting system, to visualise and quantify the presence of the virus.

The nucleic acid hybridisation was visualised using biotinylated oligonucleotide(s) and a conjugate streptavidin-horseradish peroxidase for detection with a chemiluminescent substrate using a Li-Cor imaging system. While this system is fast, sensitive, and safe for laboratory use, it cannot be transposed to a field context. Preliminary work was undertaken to replace the biotin-labelled detecting oligonucleotide probe with an ATTO-modified probe. ATTO dyes are photostable fluorescent markers (Sun *et al.*, 2022) that can be attached to oligonucleotides and be detected by photoelectric sensors. No differences were observed between CSSV-negative and -positive samples and background signal hindered the data collection. Because the sensitivity of the assay is a crucial requirement of this biosensor, the detection strategy could rely on signal amplification instead of amplifying the targets.

Such an approach was initiated at the end of this project and DNA origami-based probes were constructed. DNA origami is a technique invented in 2006 (Rothemund, 2006) that enables creation of DNA superstructures (Bush *et al.*, 2020; Saccà and Niemeyer, 2012) that relies on the self-assembly properties of the DNA molecule to build scaffold assembly of several thousands of nucleotides (Chen *et al.*, 2018). The rationale for the design of such a scaffold was to increase the signal generated by the detection probes. It was hypothesized that combining nucleotide monomers capable of self-assembly with fluorescent probe(s) would result in a "Christmas tree" type signal amplification (Fahrlander & Klausner, 1988). Unfortunately, the attempts were unsuccessful and could not be pursued further due to lack of time available. This is however a potential avenue to generate a signal from the hybridisation of CSSV since DNA-based probes and biosensors have been developed using DNA origami previously (Wang *et al.*, 2020). Also, DNA origami nucleotides are inexpensive which is an important parameter in this project. CSSV extracts could therefore be analysed by a competitive

hybridisation assay using DNA origami-constructed fluorescent probes and detected with a photoelectric sensor module.

Alternatively, a visible light optical detection strategy could be selected, to reduce the production costs of the biosensor further. Metallic nanoparticles (NPs) have been used with success thanks to their improved sensitivity and affordability (Song, Wei, and Qu, 2011; Stewart et al., 2008; Soper et al., 2006) to capture the DNA from pathogen using nucleic acids as biological recognition elements. In particular, gold NPs have been used to detect a range of virus such as influenza (Le *et al.*, 2017), hepatitis B virus (Lu *et al.*, 2013), herpesvirus (Mancuso *et al.*, 2013), and human papillomavirus genotype 16 and 18 (Kumvongpin *et al.*, 2016). Silver nanoparticles (AgNps) combined with a peptide nucleic acid (PNA) probe were the basis for optical detection of a multiplexed paper-based biosensor targeting Middle East respiratory syndrome–related coronavirus, *mycobacterium tuberculosis*, and human papillomavirus A (Teengam *et al.*, 2017) (Figure 5.3). Utilising metallic NPs resulted in limits of detection (LOD) ranging from as little as thirty virions, corresponding to 10 fg of the DNA of the cyprinid herpesvirus-3 targeted (Saleh and El-Matbouli, 2015), 10 copies for human papillomavirus genotype 18 (Kumvongpin *et al.*, 2016), and 0.12 µM of DNA of the dengue virus (Abdul Rahman *et al.*, 2014) .

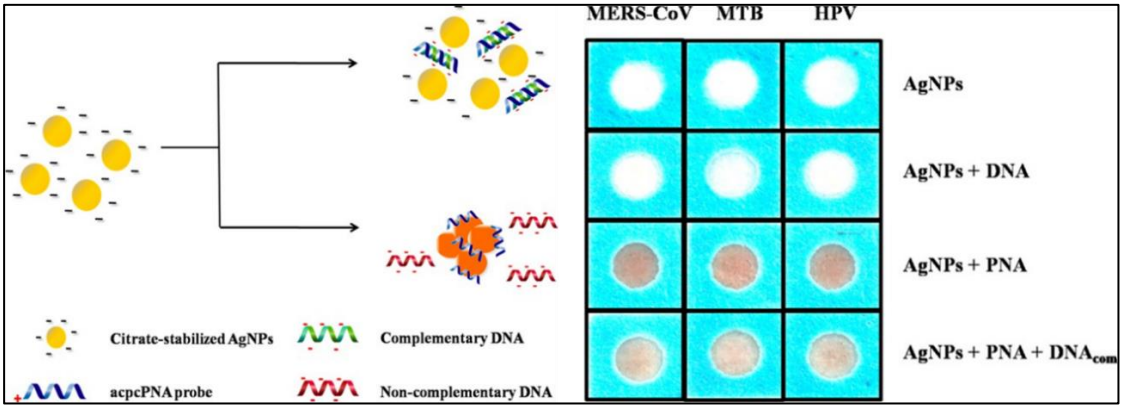


Figure 5.3: Multiplexed silver nanoparticles optical paper biosensor to detect MERS-CoV, MTB and HPV.

Left panel describes the detection mechanism mediated by the AgNps to produce a visible signal. The basal electrostatic state of the NPs enables them to be dispersed in solution and colourless. Upon addition of the PNA probe (acpcPNA), the particles aggregate and turn red. When the complementary sequence to the probe is further added to the mix, the DNA-probe interaction outcompetes the PNA-NP binding, de-aggregating the particles and changing the colour to yellow, proportionally to the quantity of complementary DNA detected. Adapted from Teengam et al., (2017) with permission.

Based on the CSSV copy number retrieved from crude extracts (figure 4.4, 4.5) the LOD permitted with the metallic NPs in a hybridisation biosensor would be compatible with the quantity of CSSV circulating in the *T. cacao* trees used as a model for this project, and therefore, the method used by Teengam *et. al* could be transposed for the detection of CSSV.

A different approach to circumvent the issue of low viral copy number would be to build up viral material by using the filter-based sap exudate collection described in the previous section, since accumulation of the virus on the support would provide an amount of CSSV genetic material compatible with the sensitivity of optical detection systems. Sampling devices could be installed on several trees, on several branches, and the sap collected overnight in a tube containing a buffer formulated to reduce the viscosity. The day after, samples could be collected, place in a syringe and filtered through a disposable membrane containing immobilised anchoring probes. The detection of CSSV could then be visualised using metallic NPs attached to a detection probe, creating an oligonucleotide lattice as described in the third chapter of this project (Figure 3.1).

Similarly, to the system imagined for the DNA BRE, a fluorescent detection probe was thought to be used as a visualisation technique when employing a zinc finger protein as a BRE. The targeted region of interest could be first recognised and immobilised by the ZFP on the surface of a membrane, and a detection DNA probe would be then used to visualise the binding of CSSV. Alternatively, the sequence-enabled reassembly of β-lactamase (SEER-LAC) is an encouraging direction of research as an optical detection system. The method (Ooi et al., 2006) follows an original strategy wherein the capturing and the signalling parts of the biosensor are combined. The detection of the targeted DNA is devolved to two zinc finger proteins, each recognising a distinct, close-by sequence of interest. The ZFP used in the study were engineered to be fused with the catalytical domain of a β-lactamase from *E. coli*. Upon binding to the double stranded DNA, the ZFPs would come into contact and consequently resume the functionality of the enzyme. The chromogenic substrate nitrocefin which turns from yellow to red upon action of the enzyme is added to the samples, enabling a direct visual observation. The specificity

was measured by testing the protein complex with a modified DNA target sequence bearing a single nucleotide mutation. The authors then observed a ten-fold reduction in the signal to background levels when the mutation was on one of the targeted sites, and a 36% reduction when the mutation was on the second site targeted. This integrated signal-generation method could be incorporated to the ZFP production pipeline described in figure 5.1.

5.4 Conclusion and perspectives.

In this project, a DNA extraction method, two type of BREs and a hybridisation assay were developed to construct and operate a cost-effective and simple biosensor for the detection of CSSV, the main threat to the global production of chocolate. The demand for this crop is increasing everywhere in the world and is driven in particular by the economic development of emerging countries such as China and India (ICCO, 2017). Concomitantly the effect of abiotic factors caused by climate change have current and predicted negative impact on the production, due to a lack of precipitation (Schroth *et al.*, 2016), increase in the temperature during the dry season (Schroth *et al.*, 2016) and modification of the evapotranspiration during the growing season (Schroth *et al.*, 2016). Because current CSSV control measures are not adequate, an additional tool to monitor the propagation of the virus in the cultivation areas is desirable. The findings of this project could be integrated into a biosensor used as a monitoring device. Furthermore, the biological recognition elements on which the biosensor is based could be easily tuned to accommodate other pathogenic targets of *T. cacao*.

Biotic stresses like viral infection are not the only threat to the cacao sustainable supply chain. Reducing the quantity of agricultural inputs such as fertilizers, pesticides, growth regulators and soil conditioners is required to displace agricultural towards more sustainable practices to avoid further deforestation of primary forest in West Africa (Hadavi and Ghazijahani, 2018);(Wijerathna-Yapa and Pathirana, 2022). An avenue to address this challenge would be to monitor plant health more closely and provide more personalised care to trees in plantations. While the detection of DNA provides the user with a qualitative information on the presence or absence of, for example, a pathogen, plant

small RNAs (sRNAs), give information about the regulation of a range of processes occurring in a plant at a given time point, and draws a snapshot of the metabolic state of a plant under a set of environmental conditions. sRNas can be classified in two type of molecules, micro RNAs (miRNAs) and small interfering RNAs (siRNAs), based on how they are synthesised and their functional utilisation by the plant (Ivanova *et al.*, 2022). They are involved in all regulatory processes occurring in the organism, which includes response to abiotic stress, such as heat, cold, salinity, nutrient deficiency, heavy metals, or drought (Yu *et al.*, 2019) and interaction with pathogens(Huang *et al.*, 2019). These regulatory molecules are mobilised by the plant using several mechanisms such as extracellular vesicles, that possess specific markers (Huang *et al.*, 2019) which can be identified and monitored. Also, sRNAs can move freely as naked molecules or associated to RNA binding proteins (RBPs)(Tang *et al.*, 2022). Interestingly, RBPs have been found in the phloem compartments of plants (Yan *et al.*, 2020; Ham and Lucas, 2017). Seen as "systemic signalling agents" (Ham and Lucas, 2017), sRNAs represent a promising class of molecules to provide accurate information about the health of plants. In consequence, a biosensor could be constructed to detect relevant sRNAs. The assay could rely on a nucleic acid probe or a zinc finger protein, since they are able to bind RNA (Hall, 2005; Lu, Searles and Klug, 2003), such as the BREs developed in this project. A sap collection contraption could be installed on the plants of interest and the assay would provide dynamic data to manage the nutrient, pesticide, or fertiliser usage in a more efficient manner. This would benefit both farmers and agricultural companies by facilitating the health management of their plants while reducing the cost associated with pesticide and fertiliser usage while government agencies would progress towards their sustainability targets.

References.

Abdul Rahman, S., Saadun, R., Azmi, N.E., Ariffin, N., Abdullah, J., Yusof, N.A., Sidek, H. and Hajian, R. (2014) Label-free dengue detection utilizing PNA/DNA hybridization based on the aggregation process of unmodified gold nanoparticles. *Journal of Nanomaterials* [online]. 2014. [Accessed 18 September 2023].

Abrokwah, F., Dzahini-Obiatey, H., Galyuon, I., Osae-Awuku, F. and Muller, E. (2016) Geographical Distribution of Cacao swollen shoot virus Molecular Variability in Ghana. *Plant disease* [online]. 100 (10), pp. 2011–2017. Available from: https://pubmed.ncbi.nlm.nih.gov/30682997/ [Accessed 29 September 2023].

Adeniyi, D. and Adeniyi, D. (2019) Diversity of Cacao Pathogens and Impact on Yield and Global Production. *Theobroma Cacao - Deploying Science for Sustainability of Global Cocoa Economy* [online]. Available from: https://www.intechopen.com/chapters/65131 [Accessed 29 September 2023].

Adomako, D. (1985) *Observations on the purification and electron microscopy of cocoa swollen shoot virus*. In: *Proceedings of 9th International Cocoa Research Conference*. pp. 327–331.

Adomako, D., Lesemann, D. -E, Paul, H.L. and Owusu, G.K. (1983) Improved methods for the purification and detection of cocoa swollen shoot virus. *Annals of Applied Biology* [online]. 103 (1), pp. 109–116. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1744-7348.1983.tb02746.x [Accessed 18 February 2023].

Agdia (2023) *Agdia Sample extraction bag SEB1*. 2023.

Ageno, M., Dore, E. and Frontali, C. (1969a) The Alkaline Denaturation of DNA. *Biophysical Journal* [online]. 9 (11), pp. 1281–1311.

Ageno, M., Dore, E. and Frontali, C. (1969b) The alkaline denaturation of DNA. *Biophysical journal* [online]. 9 (11), pp. 1281–1311. Available from: https://pubmed.ncbi.nlm.nih.gov/4982056.

Agyei-Ohemeng, J., Opare, P. and Obeng-Ofori, D. (2018) Inventory of exisiting and exploitable knowledge and technologies on cocoa in Ghana. *International Journal of Management and Commerce Innovations* [online]. 6 (1), pp. 1911–1921. Available from: https://www.researchpublish.com/papers/inventory-of-existing-and-exploitable-knowledge-and-technologies-on-cocoa-in-ghana [Accessed 28 September 2023].

Aikpokpodion, P.O. *et al.* (2009) Genetic diversity assessment of sub-samples of cacao, Theobroma cacao L. collections in West Africa using simple sequence repeats marker. *Tree Genetics & Genomes* [online]. 5, pp. 699–711. [Accessed 29 September 2023].

Allainguillaume, J., Barnett, J.M., Kiely, J. and Luxton, R. (2019) *CSSV biosensor (WO2020089607-PAMPH-20200507-8395)* [online]. Available from: https://uwe-repository.worktribe.com/output/7440964 [Accessed 6 July 2023].

Altschuler, S.E., Lewis, K.A. and Wuttke, D.S. (2013) Practical strategies for the evaluation of high-affinity protein/ nucleic acid interactions. *J Nucleic Acids Investig* [online]. 4 (1).

Alvi, N. ul H., Gómez, V.J., Rodriguez, P.E.D.S., Kumar, P., Zaman, S., Willander, M. and Nötzel, R. (2013) An InN/InGaN Quantum Dot Electrochemical Biosensor for Clinical Diagnosis. *Sensors (Basel,*

Switzerland) [online]. 13 (10), p. 13917. Available from: /pmc/articles/PMC3859099/ [Accessed 21 September 2023].

Ameyaw, G.A. (2019) Management of the Cacao Swollen Shoot Virus (CSSV) Menace in Ghana: The Past, Present and the Future. In: *Plant Diseases - Current Threats and Management Trends* [online]. IntechOpen. [Accessed 30 January 2023].

Ameyaw, G.A., Domfeh, O., Armooh, B., Boakye, A.Y. and Arjarquah, A. (2022) Inconsistent PCR detection of Cacao swollen shoot virus (CSSV) is linked to the occurrence of different variants across the cocoa regions of Ghana. *Journal of Virological Methods* [online]. 300, p. 114400. [Accessed 9 May 2023].

Ameyaw, G.A., Wetten, A., Dzahini-Obiatey, H., Domfeh, O. and Allainguillaume, J. (2013) Investigation on Cacao swollen shoot virus (CSSV) pollen transmission through cross-pollination. *Plant Pathology* [online]. 62 (2), pp. 421–427. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-3059.2012.02640.x [Accessed 17 September 2023].

Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2005) Counting the Zinc-Proteins Encoded in the Human Genome. *Journal of Proteome Research* [online]. 5 (1), pp. 196–201.

Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2006) Zinc through the three domains of life. *Journal of Proteome Research* [online]. 5 (11), pp. 3173–3178.

Ang, L.F., Por, L.Y. and Yam, M.F. (2015) Development of an Amperometric-Based Glucose Biosensor to Measure the Glucose Content of Fruit. *PLOS ONE* [online]. 10 (3), p. e0111859. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0111859 [Accessed 21 September 2023].

Anon. (2002) [online]. Available from: https://www.codoncode.com/aligner/download.htm.

Anon. (2019) [online]. pp. 9–25. Available from: https://manage.geneious.com/free-trial [accessed 10 May 2016].

Archdeacon, J., Bouhouche, N., O'Connell, F. and Kado, C.I. (2000) A single amino acid substitution beyond the C2H2-zinc finger in Ros derepresses virulence and T-DNA genes in Agrobacterium tumefaciens. *FEMS Microbiology Letters* [online]. 187 (2), pp. 175–178. Available from: https://dx.doi.org/10.1111/j.1574-6968.2000.tb09156.x [Accessed 16 September 2023].

Argout, X. *et al.* (2011) The genome of Theobroma cacao. *Nature Genetics* [online]. 43 (2), pp. 101–108. [Accessed 27 October 2022].

Argout, X., Martin, G., Droc, G., Fouet, O., Labadie, K., Rivals, E., Aury, J.M. and Lanaud, C. (2017) The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. *BMC genomics* [online]. 18 (1). Available from: https://pubmed.ncbi.nlm.nih.gov/28915793/ [Accessed 27 October 2022].

Aslantas, Y. and Surmeli, N.B. (2019) Effects of N-Terminal and C-Terminal Polyhistidine Tag on the Stability and Function of the Thermophilic P450 CYP119. *Bioinorganic Chemistry and Applications* [online]. 2019. Available from: /pmc/articles/PMC6610755/ [Accessed 17 September 2023].

Attafuah, A. and Glendinning, D.K. (1965) Studies on resistance and tolerance to cocoa viruses in Ghana. *Annals of Applied Biology* [online]. 56 (2), pp. 219–225. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1744-7348.1965.tb01229.x [Accessed 28 September 2023].

References.

Ausubel, F.M. *et al.* (2003) *Current Protocols in Molecular Biology Molecular Biology* [online]. 1.

Ayala, E., Downey, J.S., Mashburn-Warren, L., Senadheera, D.B., Cvitkovitch, D.G. and Goodman, S.D. (2014) A Biochemical Characterization of the DNA Binding Activity of the Response Regulator VicR from Streptococcus mutans. *PLOS ONE* [online]. 9 (9), p. e108027. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0108027 [Accessed 19 June 2023].

Aytehgiza, M.A. and Gebresilasie, G. (2019) Effect of Climate Change on Agricultural Output Growth in Ethiopia: Co-Integration and Vector Error Correction Model Analysis. *Budapest International Research in Exact Sciences (BirEx) Journal* [online]. 1 (4), pp. 132–143. Available from: https://bircu-journal.com/index.php/birex/article/view/461 [Accessed 29 September 2023].

Azizpor, P., Sullivan, L., Lim, A. and Groover, A. (2022) Facile Labeling of Sieve Element Phloem-Protein Bodies Using the Reciprocal Oligosaccharide Probe OGA488. *Frontiers in Plant Science* [online]. 13. Available from: /pmc/articles/PMC8867008/ [Accessed 18 September 2023].

Bai, H., Wang, R., Hargis, B., Lu, H. and Li, Y. (2012) A SPR Aptasensor for Detection of Avian Influenza Virus H5N1. *Sensors (Basel, Switzerland)* [online]. 12 (9), p. 12506. Available from: /pmc/articles/PMC3478855/ [Accessed 21 September 2023].

Bailey, B.A., Evans, H.C., Phillips-Mora, W., Ali, S.S. and Meinhardt, L.W. (2018) Moniliophthora roreri, causal agent of cacao frosty pod rot. *Molecular Plant Pathology* [online]. 19 (7), p. 1580. Available from: /pmc/articles/PMC6638017/ [Accessed 28 September 2023].

Barbas, C.F. and Mandell, J.G. (2006) *zinc finger tools*. 2006 [online]. Available from: http://www.zincfingertools.org [Accessed 2 June 2023].

Beerli, R.R., Segal, D.J., Dreier, B. and Barbas, C.F. (1998) Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 95 (25), p. 14628. Available from: /pmc/articles/PMC24500/ [Accessed 29 May 2023].

Beg, M.S., Ahmad, S., Jan, K. and Bashir, K. (2017) Status, supply chain and processing of cocoa - A review. *Trends in Food Science & Technology* [online]. 66, pp. 108–116. [Accessed 29 September 2023].

Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W. and Bulyk, M.L. (2006) Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nature Biotechnology 2006 24:11* [online]. 24 (11), pp. 1429–1435. Available from: https://www.nature.com/articles/nbt1246 [Accessed 15 September 2023].

Bhardwaj, A. and Nain, V. (2021) TALENs—an indispensable tool in the era of CRISPR: a mini review. *Journal of Genetic Engineering & Biotechnology* [online]. 19 (1). Available from: /pmc/articles/PMC8380213/ [Accessed 15 September 2023].

Bhat, A.I., Hohn, T. and Selvarajan, R. (2016) Badnaviruses: The Current Global Scenario. *Viruses 2016, Vol. 8, Page 177* [online]. 8 (6), p. 177. Available from: https://www.mdpi.com/1999-4915/8/6/177/htm [Accessed 15 February 2023].

Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.-G. and Chandrasegaran, S. (2001) Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.

Molecular and Cellular Biology [online]. 21 (1), p. 289. Available from: /pmc/articles/PMC88802/ [Accessed 29 May 2023].

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* [online]. 7 (6), pp. 1513–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/388356 [Accessed 9 July 2017].

Blake, R. and Delcourt, S.G. (1996a) Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Research* [online]. 24 (11), pp. 2095–2103. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/24.11.2095 [Accessed 9 July 2017].

Blake, R.D. and Delcourt, S.G. (1996b) Thermodynamic Effects of Formamide on DNA Stability. *Nucleic Acids Research* [online]. 24 (11), pp. 2095–2103. Available from: https://dx.doi.org/10.1093/nar/24.11.2095 [Accessed 17 September 2023].

Bomdzele, E. and Molua, E.L. (2023) Assessment of the impact of climate and non-climatic parameters on cocoa production: a contextual analysis for Cameroon. *Frontiers in Climate* [online]. 5, p. 1069514. [Accessed 29 September 2023].

Booth, W.T. *et al.* (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega* [online]. 3 (1), pp. 760–768. Available from: https://pubs.acs.org/doi/full/10.1021/acsomega.7b01598 [Accessed 17 September 2023].

Bostan, H. and Elibuyuk, I.O. (2010) A simple and rapid nucleic acid preparation method for reverse transcription polymerase chain reaction (RT- PCR) in dormant potato tubers. *African Journal of Biotechnology* [online]. 9 (13), pp. 1982–1986.

Bragazzi, N.L., Amicizia, D., Panatto, D., Tramalloni, D., Valle, I. and Gasparini, R. (2015) Quartz-Crystal Microbalance (QCM) for Public Health: An Overview of Its Applications. *Advances in Protein Chemistry and Structural Biology* [online]. 101, pp. 149–211. [Accessed 21 September 2023].

Branza-Nichita, N., Lazar, C., Durantel, D., Dwek, R.A. and Zitzmann, N. (2002) Role of disulfide bond formation in the folding and assembly of the envelope glycoproteins of a pestivirus. *Biochemical and Biophysical Research Communications* [online]. 296 (2), p. 470. Available from: /pmc/articles/PMC7111099/ [Accessed 27 September 2023].

Brenowitz, M., Senear, D.F., Shea, M.A. and Ackers, G.K. (1986) Quantitative DNase footprint titration: a method for studying protein-DNA interactions. *Methods in enzymology* [online]. 130 (C), pp. 132–181. Available from: https://pubmed.ncbi.nlm.nih.gov/3773731/ [Accessed 27 September 2023].

Brunt, A.A. and Kenten, R.H. (1962) Mechanical transmission of cocoa swollen-shoot virus to and from cocoa and other hosts. *Annals of Applied Biology* [online]. 50 (4), pp. 749–754. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1744-7348.1962.tb06075.x [Accessed 29 September 2023].

Brunt, A.A., Kenten, R.H. and Nixon, H.L. (1964) SOME PROPERTIES OF COCOA SWOLLEN-SHOOT VIRUS. *Journal of general microbiology* [online]. 36, pp. 303–309. Available from: https://pubmed.ncbi.nlm.nih.gov/14195653/ [Accessed 28 September 2023].

Bulyk, M.L., Huang, X., Choo, Y. and Church, G.M. (2001) Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc. Natl. Acad. Sci. USA* [online]. 98 (13), pp. 7158–7163. [Accessed 15 September 2023].

References.

Bunka, D.H.J., Platonova, O. and Stockley, P.G. (2010) Development of aptamer therapeutics. *Current Opinion in Pharmacology* [online]. 10 (5), pp. 557–562. [Accessed 20 September 2023].

Bush, J., Singh, S., Vargas, M., Oktay, E., Hu, C.H. and Veneziano, R. (2020) Synthesis of DNA Origami Scaffolds: Current and Emerging Strategies. *Molecules (Basel, Switzerland)* [online]. 25 (15). Available from: https://pubmed.ncbi.nlm.nih.gov/32722650/ [Accessed 18 September 2023].

Callebaut, B. (2023) *Barry Callebaut Agrilogic White Paper* [online]. Available from: https://www.barry-callebaut.com/system/files/2023-05/Barry Callebaut Agrilogic White Paper 2023_1.pdf.

Carla Veldhuyzen (2019) *Fairtrade Living Income Reference Prices for Cocoa* [online]. Available from: https://files.fairtrade.net/2019_RevisedExplanatoryNote_FairtradeLivingIncomeReferencePriceCocoa.pdf [Accessed 29 September 2023].

Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G. and Raschellà, G. (2017) Zinc-finger proteins in health and disease. *Cell Death Discovery 2017 3:1* [online]. 3 (1), pp. 1–12. Available from: https://www.nature.com/articles/cddiscovery201771 [Accessed 16 September 2023].

Chambers, J.P., Arulanandam, B.P., Matta, L.L., Weis, A. and Valdes, J.J. (2008) Biosensor Recognition Elements. *Current Issues in Molecular Biology 2008, Vol. 10, Pages 1-12* [online]. 10 (1–2), pp. 1–12. Available from: https://www.mdpi.com/1467-3045/10/1/1 [Accessed 20 September 2023].

Chant, A., Kraemer-Pecore, C.M., Watkin, R. and Kneale, G.G. (2005) Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein expression and purification* [online]. 39 (2), pp. 152–159. Available from: https://pubmed.ncbi.nlm.nih.gov/15642465/ [Accessed 17 September 2023].

Charlermroj, R. *et al.* (2013) Multiplex detection of plant pathogens using a microsphere immunoassay technology. *PloS one* [online]. 8 (4). Available from: https://pubmed.ncbi.nlm.nih.gov/23638044/ [Accessed 28 September 2023].

Chartuprayoon, N., Rheem, Y., Ng, J.C.K., Nam, J., Chen, W. and Myung, N. V. (2013) Polypyrrole nanoribbon based chemiresistive immunosensors for viral plant pathogen detection. *Analytical Methods* [online]. 5 (14), pp. 3497–3502. Available from: https://pubs.rsc.org/en/content/articlehtml/2013/ay/c3ay40371h [Accessed 28 September 2023].

Chen, C. and Wang, J. (2020) Optical biosensors: an exhaustive and comprehensive review. *Analyst* [online]. 145 (5), pp. 1605–1628. Available from: https://pubs.rsc.org/en/content/articlehtml/2020/an/c9an01998g [Accessed 21 September 2023].

Chen, X., Wang, Q., Peng, J., Long, Q., Yu, H. and Li, Z. (2018) Self-Assembly of Large DNA Origami with Custom-Designed Scaffolds. *ACS Applied Materials and Interfaces* [online]. 10 (29), pp. 24344–24348. Available from: https://pubs.acs.org/doi/abs/10.1021/acsami.8b09222 [Accessed 18 September 2023].

Chen, X.Y. and Kim, J.Y. (2009) Callose synthesis in higher plants. *Plant Signaling & Behavior* [online]. 4 (6), p. 489. Available from: /pmc/articles/PMC2688293/ [Accessed 18 September 2023].

References.

Chingandu, N., Zia-ur-rehman, M., Sreenivasan, T.N., Surujdeo-Maharaj, S., Umaharan, P., Gutierrez, O.A. and Brown, J.K. (2017) Molecular characterization of previously elusive badnaviruses associated with symptomatic cacao in the New World. *Archives of Virology* [online]. 162 (5), pp. 1363–1371.

Choo, Y. and Isalan, M. (2000) Advances in zinc finger engineering. *Current Opinion in Structural Biology* [online]. 10, pp. 411–416. Available from: https://ac.els-cdn.com/S0959440X0000107X/1-s2.0-S0959440X0000107X-main.pdf?_tid=58e754a2-d6b3-11e7-a1e3-00000aab0f02&acdnat=1512145324_54205cdd254539b38ed6c471e889ab92 [Accessed 1 December 2017].

Claus, G. *et al.* (2018) Challenges in Cocoa Pollination: The Case of Côte d'Ivoire. In: *Pollination in Plants* [online]. IntechOpen.

Coe, S.D.C. and M.D. (1996) *The True History of Chocolate. . Thames and Hudson, New York, 1996. 280 pp.,*.

Cooper, I. and Jones, R.A.C. (2006) Wild Plants and Viruses: Under-Investigated Ecosystems. *Advances in Virus Research* [online]. 67, pp. 1–47. [Accessed 29 September 2023].

Cornejo, O.E. *et al.* (2018) Population genomic analyses of the chocolate tree, Theobroma cacao L., provide insights into its domestication process. *Communications Biology* [online]. 1 (1). Available from: /pmc/articles/PMC6191438/ [Accessed 21 October 2022].

Corpet, F. (1988) *Multialin Multiple sequence alignment Nucl. Acids Res. 16 (22)*. 1988 [online]. pp. 10881–10890. Available from: http://multalin.toulouse.inra.fr/multalin/.

Cranshaw, W. (2007) *Whitney Cranshaw*. 2007 [online]. Available from: https://commons.wikimedia.org/wiki/File:Planococcus_citri_1455198.jpg [Accessed 5 May 2023].

Cremer, M. (1906) Über die Ursache der elektromotorischen Eigenschaften der Gewebe, zugleich ein Beitrag zur Lehre von den polyphasischen Elektrolytketten. *Z. Biol.* 47, pp. 562–608.

Crowe, J., Döbeli, H., Gentz, R., Hochuli, E., Stüber, D. and Henco, K. (1994) 6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification. *Methods in molecular biology (Clifton, N.J.)* [online]. 31, pp. 371–387. Available from: https://link.springer.com/protocol/10.1385/0-89603-258-2:371 [Accessed 26 June 2023].

Cuatrecasas, J. (1964) Cacao and Its Allies, a Taxonomic Revision of the Genus <I>Theobroma</I>. In: *Systematic Plant Studies* [online]. Smithsonian Institution Press, pp. 379–614. [Accessed 22 October 2022].

Damborský, P., Švitel, J. and Katrlík, J. (2016) Optical biosensors. *Essays in biochemistry* [online]. 60 (1), pp. 91–100. Available from: https://pubmed.ncbi.nlm.nih.gov/27365039/ [Accessed 27 September 2023].

Davalieva, K. and Efremov, G.D. (2010) Influence of salts and PCR inhibitors on the amplification capacity of three thermostable DNA polymerases. *Macedonian Journal of Chemistry and Chemical Engineering* [online]. 29 (1), pp. 57–62. Available from: https://www.researchgate.net/publication/216175309_Influence_of_salts_and_PCR_inhibitors_on_the_amplification_capacity_of_three_thermostable_DNA_polymerases [Accessed 4 September 2023].

Demeke, T. and Jenkins, G.R. (2010) Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Analytical and*

Bioanalytical Chemistry [online]. 396 (6), pp. 1977–1990. Available from: https://link.springer.com/article/10.1007/s00216-009-3150-9 [Accessed 18 September 2023].

Dillinger, T.L., Barriga, P., Escárcega, S., Jimenez, M., Lowe, D.S. and Grivetti, L.E. (2000) Food of the Gods: Cure for humanity? A cultural history of the medicinal and ritual use of chocolate. *Journal of Nutrition* [online]. 130 (8 SUPPL.), pp. 2057–2072.

Dinant, S., Bonnemain, J.-L., Girousse, C. and Kehr, J. (2010) Plant biology and pathology / Biologie et pathologie vé gé tales Phloem sap intricacy and interplay with aphid feeding Complexite´deComplexite´de lasèvephloèmienne et impact sur l'alimentation des pucerons Comptes Rendus Biologies. *Biologies* [online]. 333, pp. 504–515. [Accessed 18 September 2023].

Dreier, B., Segal, D.J. and Barbas, C.F. (2000) Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains. *Journal of molecular biology* [online]. 303 (4), pp. 489–502. Available from: https://pubmed.ncbi.nlm.nih.gov/11054286/ [Accessed 29 June 2023].

Drygin, Y.F., Blintsov, A.N., Grigorenko, V.G., Andreeva, I.P., Osipov, A.P., Varitzev, Y.A., Uskov, A.I., Kravchenko, D. V. and Atabekov, J.G. (2012) Highly sensitive field test lateral flow immunodiagnostics of PVX infection. *Applied Microbiology and Biotechnology* [online]. 93 (1), pp. 179–189. Available from: https://link.springer.com/article/10.1007/s00253-011-3522-x [Accessed 28 September 2023].

Dubs, M.C., Altschuh, D. and Van Regenmortel, M.H.V. (1992) Interaction between viruses and monoclonal antibodies studied by surface plasmon resonance. *Immunology Letters* [online]. 31 (1), pp. 59–64.

Dyussembayev, K., Sambasivam, P., Bar, I., Brownlie, J.C., Shiddiky, M.J.A. and Ford, R. (2021) Biosensor Technologies for Early Detection and Quantification of Plant Pathogens. *Frontiers in Chemistry* [online]. 9, p. 636245. [Accessed 21 September 2023].

Ehrich, M., Field, J.K., Liloglou, T., Xinarianos, G., Oeth, P., Nelson, M.R., Cantor, C.R. and Van Den Boom, D. (2006) Cytosine methylation profiles as a molecular marker in non-small cell lung cancer. *Cancer Research* [online]. 66 (22), pp. 10911–10918.

Eigner, J. and Doty, P. (1965) The native, denatured and renatured states of deoxyribonucleic acid. *Journal of Molecular Biology* [online]. 12 (3), pp. 549–580. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0022283665803126 [Accessed 11 September 2017].

Eksin, E. and Erdem, A. (2023) Recent Progress on Optical Biosensors Developed for Nucleic Acid Detection Related to Infectious Viral Diseases. *Micromachines* [online]. 14 (2). Available from: /pmc/articles/PMC9966969/ [Accessed 21 September 2023].

Elson, D. and Chargaff, E. (1952) On the desoxyribonucleic acid content of sea urchin gametes. *Experientia* [online]. 8 (4), pp. 143–145. Available from: http://link.springer.com/10.1007/BF02170221 [Accessed 29 June 2017].

Encarnação, J.M., Rosa, L., Rodrigues, R., Pedro, L., da Silva, F.A., Gonçalves, J. and Ferreira, G.N.M. (2007) Piezoelectric biosensors for biorecognition analysis: Application to the kinetic study of HIV-1 Vif protein binding to recombinant antibodies. *Journal of Biotechnology* [online]. 132 (2), pp. 142–148. [Accessed 20 September 2023].

Ericsson, A., Faria, L., Cruz, W., Martins de Sá, C. and Lima, B. (2006) TcZFP8, a novel member of the Trypanosoma cruzi CCHC zinc finger protein family with nuclear localization. *Genetics and Molecular*

Research [online]. 5 (3), pp. 553–563. Available from: www.funpecrp.com.br [Accessed 16 September 2023].

Eun, A.J.-C., Huang, L., Chew, F.-T., Fong-Yau Li, S. and Wong, S.-M. (2002) Detection of two orchid viruses using quartz crystal microbalance-based DNA biosensors. *Phytopathology* [online]. 92 (6), pp. 654–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19628008.

Fahrlander, P.D. and Klausner, A. (1988) Amplifying DNA Probe Signals: A 'christmas Tree' Approach. *Bio/Technology 1988 6:10* [online]. 6 (10), pp. 1165–1168. Available from: https://www.nature.com/articles/nbt1088-1165 [Accessed 18 September 2023].

Fang, Y. and Ramasamy, R.P. (2018) A Portable Electrochemical System for Plant Volatile Detection. *ECS Transactions* [online]. 85 (13), pp. 1359–1367.

Fang, Y., Umasankar, Y. and Ramasamy, R.P. (2016) A novel bi-enzyme electrochemical biosensor for selective and sensitive determination of methyl salicylate. *Biosensors and Bioelectronics* [online]. 81, pp. 39–45. [Accessed 21 September 2023].

FAO (2023) *FAO*. 2023. [Accessed 14 May 2023].

FAO (2021) *FAOSTAT Crops and livestock products*. 2021 [online]. Available from: https://www.fao.org/faostat/en/#data/QCL.

Fauver, M.E., Dunaway, D.L., Lilienfeld, D.H., Craighead, H.G. and Pollack, G.H. (1998) Microfabricated cantilevers for measurement of subcellular and molecular forces. *IEEE transactions on bio-medical engineering* [online]. 45 (7), pp. 891–898. Available from: https://pubmed.ncbi.nlm.nih.gov/9644898/ [Accessed 21 September 2023].

Feng, M., Kong, D., Wang, W., Liu, L., Song, S. and Xu, C. (2015) Development of an Immunochromatographic Strip for Rapid Detection of Pantoea stewartii subsp. stewartii. *Sensors (Basel, Switzerland)* [online]. 15 (2), p. 4291. Available from: /pmc/articles/PMC4367412/ [Accessed 28 September 2023].

Ford, A., Williams, A. and de Vries, M.S. (2022) New light on the use of Theobroma cacao by Late Classic Maya. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 119 (40), p. e2121821119. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2121821119 [Accessed 28 September 2023].

Franklin, R.E. and Gosling, R.G. (1953) *Molecular Configuration in Sodium Thymonucleate Nature* [online]. 171 (4356), pp. 740–741. Available from: http://www.nature.com/doifinder/10.1038/171740a0.

Fried, M. and Crothers, D.M. (1981) Nucleic Acids Research~~~~~~ Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research*. 9, p. 1981. [Accessed 31 May 2023].

Fuellgrabe, M.W., Boonrod, K., Jamous, R., Moser, M., Shiboleth, Y., Krczal, G. and Wassenegger, M. (2011) Expression, purification and functional characterization of recombinant Zucchini yellow mosaic virus HC-Pro. *Protein Expression and Purification* [online]. 75 (1), pp. 40–45. [Accessed 19 June 2023].

Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D., Lesueur, F., Landi, S. and Tavtigian, S. V (2009) Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *BMC genetics* [online]. 10, p. 5.

References.

Gateau-Rey, L., Tanner, E.V.J., Rapidel, B., Marelli, J.P. and Royaert, S. (2018) Climate change could threaten cocoa production: Effects of 2015-16 El Niño-related drought on cocoa agroforests in Bahia, Brazil. *PLOS ONE* [online]. 13 (7), p. e0200454. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0200454 [Accessed 29 September 2023].

Gaudin, V. (2020) The Growing Interest in Development of Innovative Optical Aptasensors for the Detection of Antimicrobial Residues in Food Products. *Biosensors* [online]. 10 (3). Available from: /pmc/articles/PMC7146278/ [Accessed 20 September 2023].

Goldstein, D. (2015) *The Oxford Companion to Sugar and Sweets The Oxford Companion to Sugar and Sweets* [online]. Oxford University Press. [Accessed 29 September 2023].

Gopaulchan, D., Motilal, L.A., Bekele, F.L., Clause, S., Ariko, J.O., Ejang, H.P. and Umaharan, P. (2019) Morphological and genetic diversity of cacao (Theobroma cacao L.) in Uganda. *Physiology and Molecular Biology of Plants* [online]. 25 (2), p. 361. Available from: /pmc/articles/PMC6419697/ [Accessed 29 September 2023].

Grover, A., Pande, A., Choudhary, K., Gupta, K. and Sundar, D. (2010) Re-programming DNA-binding specificity in zinc finger proteins for targeting unique address in a genome. *Systems and Synthetic Biology* [online]. 4 (4), p. 323. Available from: /pmc/articles/PMC3065587/ [Accessed 28 September 2023].

Guo, A. Di *et al.* (2023) Spatiotemporal and global profiling of DNA–protein interactions enables discovery of low-affinity transcription factors. *Nature Chemistry 2023 15:6* [online]. 15 (6), pp. 803–814. Available from: https://www.nature.com/articles/s41557-023-01196-z [Accessed 26 September 2023].

Gyamera, E.A., Domfeh, O. and Ameyaw, G.A. (2023) Cacao Swollen Shoot Viruses in Ghana. [online]. Available from: https://doi.org/10.1094/PDIS-10-22-2412-FE.

Hadavi, E. and Ghazijahani, N. (2018) Closed and Semi-closed Systems in Agriculture. [online]. pp. 295–310. Available from: https://link.springer.com/chapter/10.1007/978-3-319-99076-7_10 [Accessed 24 September 2023].

Hagen, L.S., Jacquemond, M., Lepingle, A., Lot, H. and Tepfer, M. (1993) *Nucleotide sequence and genomic organization of cacao swollen shoot virus. Virology* [online]. 196 (2), pp. 619–28. Available from: http://www.sciencedirect.com/science/article/pii/S0042682283715187.

Hall, C. Van (1914) *Cocoa, by Dr. C.J.J. Van Hall* [online]. London: Macmillan. [Accessed 12 January 2018].

Hall, T.M.T. (2005) Multiple modes of RNA recognition by zinc finger proteins. *Current opinion in structural biology* [online]. 15 (3), pp. 367–373. Available from: https://pubmed.ncbi.nlm.nih.gov/15963892/ [Accessed 24 September 2023].

Ham, B.K. and Lucas, W.J. (2017) Phloem-Mobile RNAs as Systemic Signaling Agents. *https://doi.org/10.1146/annurev-arplant-042916-041139* [online]. 68, pp. 173–195. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev-arplant-042916-041139 [Accessed 24 September 2023].

References.

Hammouda, B., Worcester, D., Becker, M., Parsegian, V.A. and Abramchuk, S. (2006) The denaturation transition of DNA in mixed solvents. *Biophysical journal* [online]. 91 (6), pp. 2237–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16815902 [Accessed 9 July 2017].

Hardjasa, A., Ling, M., Ma, K. and Yu, H. (2010) Investigating the Effects of DMSO on PCR Fidelity Using a Restriction Digest-Based Method. *Journal of Experimental Microbiology and Immunology (JEMI)*. 14, pp. 161–164. [Accessed 7 September 2023].

Harper, G., Hull, R., Lockhart, B. and Olszewski, N. (2003) VIRAL SEQUENCES INTEGRATED INTO PLANT GENOMES. *https://doi.org/10.1146/annurev.phyto.40.120301.105642* [online]. 40, pp. 119–136. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev.phyto.40.120301.105642 [Accessed 16 February 2023].

Heda, G.D., Shrestha, L., Thapa, S., Ghimire, S. and Raut, D. (2020) Optimization of western blotting for the detection of proteins of different molecular weight. *Biotechniques* [online]. 68 (6), p. 318. Available from: /pmc/articles/PMC7333534/ [Accessed 26 September 2023].

Heineman, W.R. and Jensen, W.B. (2006) Leland C. Clark Jr. (1918–2005). *Biosensors and Bioelectronics* [online]. 21 (8), pp. 1403–1404. [Accessed 19 February 2023].

Herdewijn, P. and Marlière, P. (2009) Toward safe genetically modified organisms through the chemical diversification of nucleic acids. *Chemistry and Biodiversity* [online]. 6 (6), pp. 791–808. [Accessed 20 September 2023].

Hernández-Ibáñez, N., García-Cruz, L., Montiel, V., Foster, C.W., Banks, C.E. and Iniesta, J. (2016) Electrochemical lactate biosensor based upon chitosan/carbon nanotubes modified screen-printed graphite electrodes for the determination of lactate in embryonic cell cultures. *Biosensors & bioelectronics* [online]. 77, pp. 1168–1174. Available from: https://pubmed.ncbi.nlm.nih.gov/26579934/ [Accessed 21 September 2023].

Herrmann, M.G., Durtschi, J.D., Wittwer, C.T. and Voelkerding, K. V. (2007) Expanded Instrument Comparison of Amplicon DNA Melting Analysis for Mutation Scanning and Genotyping. *Clinical Chemistry* [online]. Available from: http://clinchem.aaccjnls.org/content/early/2007/06/07/clinchem.2007.088120 [Accessed 9 July 2017].

Hirsch, I.B. (2018) Introduction: History of Glucose Monitoring. *ADA Clinical Compendia* [online]. pp. 1–1.

Hoffmann, K., Sackey, S.T., Maiss, E., Adomako, D. and Veiten, H.J. (1997a) Immunocapture Polymerase Chain Reaction for the Detection and Characterization of Cacao Swollen Shoot Virus 1 A Isolates. *Journal of Phytopathology* [online]. 145 (5–6), pp. 205–212. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1439-0434.1997.tb00387.x [Accessed 17 February 2023].

Hoffmann, K., Sackey, S.T., Maiss, E., Adomako, D. and Vetten, H.J. (1997b) Immunocapture polymerase chain reaction for the detection and characterization of cacao swollen shoot virus 1 A isolates. *J. Phytopathology*. 145, pp. 205–212.

Horton, H.R. (2006) *Principles of biochemistry* [online]. Pearson Prentice Hall. [Accessed 12 July 2017].

References.

Huang, C.Y., Wang, H., Hu, P., Hamby, R. and Jin, H. (2019) Small RNAs - Big Players in Plant-Microbe Interactions. *Cell host & microbe* [online]. 26 (2), pp. 173–182. Available from: https://pubmed.ncbi.nlm.nih.gov/31415750/ [Accessed 24 September 2023].

Hughes, J. d. A. and Ollennu, L.A. (1993) The virobacterial agglutination test as a rapid means of detecting cocoa swollen shoot virus. *Annals of Applied Biology* [online]. 122 (2), pp. 299–310.

Hull, R. (2001) Caulimoviridae (Plant Pararetroviruses). [online]. Available from: www.els.net.

Ibupoto, Z.H., Usman Ali, S.M., Khun, K. and Willander, M. (2011) L-Ascorbic Acid Biosensor Based on Immobilized Enzyme on ZnO Nanorods. *L-Ascorbic Acid Biosensor Based on Immobilized Enzyme on ZnO Nanorods. J Biosens Bioelectron* [online]. 2, p. 110. [Accessed 21 September 2023].

ICCO (2017) *ICCO Quarterly Bulletin of Cocoa Statistics. Issue No. 3-Volume XLIV.* 2017. [Accessed 12 June 2023].

Ichikawa, D.M. *et al.* (2023) A universal deep-learning model for zinc finger design enables transcription factor reprogramming. *Nature Biotechnology 2023 41:8* [online]. 41 (8), pp. 1117–1129. Available from: https://www.nature.com/articles/s41587-022-01624-4 [Accessed 15 September 2023].

ICTV (2021) *International Committee on Taxonomy of Viruses: ICTV https://ictv.global/taxonomy*. 2021.

Ikeda, N., Bautista, N.S., Yamada, T., Kamijima, O. and Ishii, T. (2012) Ultra-simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. *Plant Molecular Biology Reporter* [online]. 19 (1), pp. 27–32. Available from: http://link.springer.com/10.1007/BF02824075.

Isalan, M., Klug, A. and Choo, Y. (2001) A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nature biotechnology* [online]. 19 (7), p. 656. Available from: /pmc/articles/PMC2677679/ [Accessed 16 September 2023].

Iuchi, S. (2001) *Review Three classes of C 2 H 2 zinc finger proteins CMLS Cellular and Molecular Life Sciences Based on the number and the pattern of the fingers, most of the proteins can be classified into one of three groups: triple-C 2 H 2 , multiple-adjacent-C 2 H 2 , and separated-pair-ed-C 2 H 2 finger proteins. In contrast to proteins with triple-C 2 H 2 fingers, proteins with multiple-adjacent-C 2 H 2 fingers can bind multiple CMLS, Cell. Mol. Life Sci*. 58.

Ivanova, Z., Minkov, G., Gisel, A., Yahubyan, G., Minkov, I., Toneva, V. and Baev, V. (2022) The Multiverse of Plant Small RNAs: How Can We Explore It? *International Journal of Molecular Sciences* [online]. 23 (7). Available from: /pmc/articles/PMC8999349/ [Accessed 24 September 2023].

Jacquot, E., Hagen, L.S., Jacquemond, M. and Yot, P. (1996) The Open Reading Frame 2 Product of Cacao Swollen Shoot Badnavirus Is a Nucleic Acid-Binding Protein. *Virology* [online]. 225 (1), pp. 191–195. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0042682296905875.

Jacquot, E., Hagen, L.S., Michler, P., Rohfritsch, O., Stussi-Garaud, C., Keller, M., Jacquemond, M. and Yot, P. (1999) In situ localization of cacao swollen shoot virus in agroinfected Theobroma cacao. *Archives of virology* [online]. 144 (2), pp. 259–271. Available from: https://pubmed.ncbi.nlm.nih.gov/10470252/ [Accessed 17 February 2023].

Jarocka, U., Radecka, H., Malinowski, T., Michalczuk, L. and Radecki, J. (2013) Detection of Prunus Necrotic Ringspot Virus in Plant Extracts with Impedimetric Immunosensor based on Glassy Carbon

Electrode. *Electroanalysis* [online]. 25 (2), pp. 433–438. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/elan.201200470 [Accessed 28 September 2023].

Jiao, K., Sun, W. and Zhang, S.S. (2000a) Sensitive detection of a plant virus by electrochemical enzyme-linked immunoassay. *Fresenius' Journal of Analytical Chemistry* [online]. 367 (7), pp. 667–671. [Accessed 20 September 2023].

Jiao, K., Sun, W. and Zhang, S.S. (2000b) Sensitive detection of a plant virus by electrochemical enzyme-linked immunoassay. *Fresenius' Journal of Analytical Chemistry* [online]. 367 (7), pp. 667–671. Available from: https://link.springer.com/article/10.1007/s002160000423 [Accessed 28 September 2023].

John, R., Spencer, M., Wallace, G.G. and Smyth, M.R. (1991) Development of a polypyrrole-based human serum albumin sensor. *Analytica Chimica Acta* [online]. 249 (2), pp. 381–385. [Accessed 20 September 2023].

Johnson, D., Hilal, N. and Bowen, W.R. (2009) Basic Principles of Atomic Force Microscopy. *Atomic Force Microscopy in Process Engineering: An Introduction to AFM for Improved Processes and Products* [online]. pp. 1–30. Available from: https://nyuscholars.nyu.edu/en/publications/basic-principles-of-atomic-force-microscopy-2 [Accessed 16 September 2023].

De Juan-Franco, E., Caruz, A., Pedrajas, J.R. and Lechuga, L.M. (2013) Site-directed antibody immobilization using a protein A-gold binding domain fusion protein for enhanced SPR immunosensing. *The Analyst* [online]. 138 (7), pp. 2023–2031. Available from: https://pubmed.ncbi.nlm.nih.gov/23400028/ [Accessed 20 September 2023].

Kashyap, A., Autebert, J., Delamarche, E. and Kaigala, G. V. (2016) Selective local lysis and sampling of live cells for nucleic acid analysis using a microfluidic probe. *Scientific Reports* [online]. 6. Available from: /pmc/articles/PMC4944176/ [Accessed 15 August 2023].

Katchalski-Katzir, E. (1993) Immobilized enzymes--learning from past successes and failures. *Trends in biotechnology* [online]. 11 (11), pp. 471–478. Available from: https://pubmed.ncbi.nlm.nih.gov/7764401/ [Accessed 20 September 2023].

Kearse, M. *et al.* (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* [online]. 28 (12), p. 1647. Available from: /pmc/articles/PMC3371832/ [Accessed 8 September 2023].

Keighley, S.D., Estrela, P., Li, P. and Migliorato, P. (2008) Optimization of label-free DNA detection with electrochemical impedance spectroscopy using PNA probes. *Biosensors and Bioelectronics* [online]. 24 (4), pp. 906–911. [Accessed 20 September 2023].

Kenten, R.H. (1972) The purification and some properties of cocoa necrosis virus, a serotype of tomato black ring virus. *Annals of Applied Biology* [online]. 71 (2), pp. 119–126. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1744-7348.1972.tb02946.x [Accessed 29 September 2023].

Kerr, J. (2003) *Mayavase.com*. 2003. [Accessed 5 May 2023].

Kilic, N.M., Singh, S., Keles, G., Cinti, S., Kurbanoglu, S. and Odaci, D. (2023) Novel Approaches to Enzyme-Based Electrochemical Nanobiosensors. *Biosensors 2023, Vol. 13, Page 622* [online]. 13 (6), p. 622. Available from: https://www.mdpi.com/2079-6374/13/6/622/htm [Accessed 28 September 2023].

References.

Kim, M., Um, H.J., Bang, S., Lee, S.H., Oh, S.J., Han, J.H., Kim, K.W., Min, J. and Kim, Y.H. (2009) Arsenic removal from Vietnamese groundwater using the arsenic-binding DNA aptamer. *Environmental science & technology* [online]. 43 (24), pp. 9335–9340. Available from: https://pubmed.ncbi.nlm.nih.gov/20000526/ [Accessed 20 September 2023].

Kim, M.S., Stybayeva, G., Lee, J.Y., Revzin, A. and Segal, D.J. (2011) A zinc finger protein array for the visual detection of specific DNA sequences for diagnostic applications. *Nucleic Acids Research* [online]. 39 (5), p. e29. Available from: /pmc/articles/PMC3061069/ [Accessed 12 July 2023].

Kim, S., Yu, N.K. and Kaang, B.K. (2015) CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental & Molecular Medicine 2015 47:6* [online]. 47 (6), pp. e166–e166. Available from: https://www.nature.com/articles/emm201533 [Accessed 16 September 2023].

King, R.W. and Zeevaart, J.A.D. (1974) Enhancement of Phloem Exudation from Cut Petioles by Chelating Agents. *Plant Physiology* [online]. 53 (1), pp. 96–103. Available from: https://dx.doi.org/10.1104/pp.53.1.96 [Accessed 5 September 2023].

Klug, A. (2010) The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. *Q Rev Biophys* [online]. 43 (1), pp. 1–21. [Accessed 16 September 2023].

Köhler, H.A. (1887) *Köhler's Medizinal-Pflanzen*.

Kosoe, E.A. and Ahmed, A. (2022) Climate change adaptation strategies of cocoa farmers in the Wassa East District: Implications for climate services in Ghana. *Climate Services* [online]. 26, p. 100289. [Accessed 29 September 2023].

Kouakou, K. and Kébé, B.I. (2012) Geographical Distribution of Cacao swollen shoot virus Molecular Variability in Côte d'Ivoire. [online]. Available from: http://dx.doi.org/10.1094.

Kress, W.J., Soltis, D.E., Kersey, P.J., Wegrzyn, J.L., Leebens-Mack, J.H., Gostel, M.R., Liu, X. and Soltis, P.S. (2022) Green plant genomes: What we know in an era of rapidly expanding opportunities. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 119 (4), p. e2115640118. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2115640118 [Accessed 9 November 2022].

Krishna, S.S., Majumdar, I. and Grishin, N. V. (2003a) Structural classification of zinc fingers. *Nucleic Acids Research* [online]. 31 (2), pp. 532–550.

Krishna, S.S., Majumdar, I. and Grishin, N. V. (2003b) SURVEY AND SUMMARY: Structural classification of zinc fingers. *Nucleic Acids Research* [online]. 31 (2), p. 532. Available from: /pmc/articles/PMC140525/ [Accessed 29 September 2023].

Krynski, I.A. and Logan, J.E. (1967) Dextrostix as a Quantitative Test for Glucose in Whole Blood. *Canad. Med. Ass. J*. 97, pp. 1006–1011.

Kumvongpin, R. *et al.* (2016) High sensitivity, loop-mediated isothermal amplification combined with colorimetric gold-nanoparticle probes for visual detection of high risk human papillomavirus genotypes 16 and 18. *Journal of virological methods* [online]. 234, pp. 90–95. Available from: https://pubmed.ncbi.nlm.nih.gov/27086727/ [Accessed 18 September 2023].

Kurbanoglu, S. and Ozkan, S.A. (2018) A Novel Enzymatic Biosensor for the Detection of Catechol Using Multi-walled Carbon Nanotubes and Gold Nanowires. *Electrocatalysis* [online]. 9 (2), pp. 252–257.

Available from: https://link.springer.com/article/10.1007/s12678-017-0408-4 [Accessed 21 September 2023].

Lata, K., Afsar, M. and Ramachandran, R. (2017) Biochemical characterization and novel inhibitor identification of Mycobacterium tuberculosis Endonuclease VIII 2 (Rv3297). *Biochemistry and Biophysics Reports* [online]. 12, pp. 20–28. [Accessed 16 September 2023].

Lay, M.J. and Wittwer, C.T. (1997) Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. *Clinical Chemistry* [online]. 43 (12). Available from: http://clinchem.aaccjnls.org/content/43/12/2262?ijkey=2b400b26b9cb670b7bd4d2c1633b7eab62d11bf8&keytype2=tf_ipsecsha [Accessed 13 July 2017].

Le, T.T., Chang, P., Benton, D.J., McCauley, J.W., Iqbal, M. and Cass, A.E.G. (2017) Dual Recognition Element Lateral Flow Assay Toward Multiplex Strain Specific Influenza Virus Detection. *Analytical chemistry* [online]. 89 (12), pp. 6781–6786. Available from: https://pubmed.ncbi.nlm.nih.gov/28558471/ [Accessed 18 September 2023].

Length, F. (2010) (CSSV) inoculated cocoa seeds and the discovery of the cotyledons of the resultant plants as rich sources of CSSV. *Journal of Biotechnology* [online]. 9 (5), pp. 593–603.

Lenz, J. *et al.* (2021) Ush regulates hemocyte-specific gene expression, fatty acid metabolism and cell cycle progression and cooperates with dNuRD to orchestrate hematopoiesis. *PLoS Genetics* [online]. 17 (2). Available from: /pmc/articles/PMC7891773/ [Accessed 16 September 2023].

Li, L., Li, B., Qi, Y. and Jin, Y. (2009) Label-free aptamer-based colorimetric detection of mercury ions in aqueous media using unmodified gold nanoparticles as colorimetric probe. *Analytical and bioanalytical chemistry* [online]. 393 (8), pp. 2051–2057. Available from: https://pubmed.ncbi.nlm.nih.gov/19198811/ [Accessed 20 September 2023].

Li, Q., Xu, H., Zhang, W., Sun, J. and Yue, Y. (2022) Extraction of cucumber phloem sap based on the capillary-air pressure principle. *BioTechniques* [online]. 72 (6), pp. 233–243. Available from: https://www.future-science.com/doi/10.2144/btn-2021-0101 [Accessed 22 September 2023].

Li, X., Song, J., Chen, B.L., Wang, B., Li, R., Jiang, H.M., Liu, J.F. and Li, C.Z. (2016) A label-free colorimetric assay for detection of c-Myc mRNA based on peptide nucleic acid and silver nanoparticles. *Science Bulletin* [online]. 61 (4), pp. 276–281. [Accessed 20 September 2023].

Liedberg, B., Nylander, C. and Lunström, I. (1983) Surface plasmon resonance for gas detection and biosensing. *Sensors and Actuators* [online]. 4 (C), pp. 299–304. [Accessed 21 September 2023].

Lin, H.Y., Huang, C.H., Lu, S.H., Kuo, I.T. and Chau, L.K. (2014) Direct detection of orchid viruses using nanorod-based fiber optic particle plasmon resonance immunosensor. *Biosensors and Bioelectronics* [online]. 51, pp. 371–378. [Accessed 28 September 2023].

Loo, J.F.C., Ho, A.H.P. and Mak, W.C. (2023) Printed microfluidic biosensors and their biomedical applications. *Microfluidic Biosensors* [online]. pp. 1–40. [Accessed 14 September 2023].

López, M.M., Llop, P., Olmos, A., Marco-Noales, E., Cambra, M. and Bertolini, E. (2008) Are Molecular Tools Solving the Challenges Posed by Detection of Plant Pathogenic Bacteria and Viruses? *Current Issues in Molecular Biology 2009, Vol. 11, Pages 13-46* [online]. 11 (1), pp. 13–46. Available from: https://www.mdpi.com/1467-3045/11/1/2 [Accessed 4 September 2023].

Lot, H., Djiekpor, E. and Jacquemond, M. (1991) Characterization of the genome of cacao swollen shoot virus. *Journal of General Virology* [online]. 72 (7), pp. 1735–1739. Available from:

https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-72-7-1735 [Accessed 14 February 2023].

Lu, D., Searles, M.A. and Klug, A. (2003) Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* [online]. 426 (6962), pp. 96–100. Available from: https://pubmed.ncbi.nlm.nih.gov/14603324/ [Accessed 24 September 2023].

Lu, X., Dong, X., Zhang, K., Han, X., Fang, X. and Zhang, Y. (2013) A gold nanorods-based fluorescent biosensor for the detection of hepatitis B virus DNA based on fluorescence resonance energy transfer. *The Analyst* [online]. 138 (2), pp. 642–650. Available from: https://pubmed.ncbi.nlm.nih.gov/23172079/ [Accessed 18 September 2023].

Lundin, K.E., Good, L., Strömberg, R., Gräslund, A. and Smith, C.I.E. (2006) Biological activity and biotechnological aspects of peptide nucleic acid. *Advances in genetics* [online]. 56, pp. 1–51. Available from: https://pubmed.ncbi.nlm.nih.gov/16735154/ [Accessed 20 September 2023].

Ma, Q., Lee, D., Tan, Y.Q., Wong, G. and Gao, Z. (2016) Synthetic genetic polymers: advances and applications. *Polymer Chemistry* [online]. 7 (33), pp. 5199–5216. Available from: https://pubs.rsc.org/en/content/articlehtml/2016/py/c6py01075j [Accessed 20 September 2023].

Mancuso, M., Jiang, L., Cesarman, E. and Erickson, D. (2013) Multiplexed Colorimetric Detection of Kaposi's Sarcoma Associated Herpesvirus and Bartonella DNA using Gold and Silver Nanoparticles. *Nanoscale* [online]. 5 (4), p. 1678. Available from: /pmc/articles/PMC3581344/ [Accessed 18 September 2023].

Mandell, J.G. and Barbas, C.F. (2006) Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Research* [online]. 34 (suppl_2), pp. W516–W523. Available from: https://dx.doi.org/10.1093/nar/gkl209 [Accessed 29 June 2023].

Mansor, N.N.N. *et al.* (2018) An Amperometric Biosensor for the Determination of Bacterial Sepsis Biomarker, Secretory Phospholipase Group 2-IIA Using a Tri-Enzyme System. *Sensors (Basel, Switzerland)* [online]. 18 (3). Available from: /pmc/articles/PMC5876737/ [Accessed 21 September 2023].

Marelli, J.P., Guest, D.I., Bailey, B.A., Evans, H.C., Brown, J.K., Junaid, M., Barreto, R.W., Lisboa, D.O. and Puig, A.S. (2019) Chocolate Under Threat from Old and New Cacao Diseases. *Phytopathology* [online]. 109 (8), pp. 1331–1343. Available from: https://pubmed.ncbi.nlm.nih.gov/31115251/ [Accessed 27 October 2022].

Mawhinney, M.T., Liu, R., Lu, F., Maksimoska, J., Damico, K., Marmorstein, R., Lieberman, P.M. and Urbanc, B. (2018) CTCF-Induced Circular DNA Complexes Observed by Atomic Force Microscopy. *Journal of molecular biology* [online]. 430 (6), pp. 759–776. Available from: https://pubmed.ncbi.nlm.nih.gov/29409905/ [Accessed 16 September 2023].

Mehrotra, P. (2016) Biosensors and their applications – A review. *Journal of Oral Biology and Craniofacial Research* [online]. 6 (2), p. 153. Available from: /pmc/articles/PMC4862100/ [Accessed 28 September 2023].

Miller, J., McLachlan, A.D. and Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *The EMBO journal* [online]. 4 (6), pp. 1609–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4040853 [Accessed 22 March 2019].

Minitab LLC (2022) *Minitab*.

References.

Miranda, F. (1962) Wild Cacao in the Lacandona Forest, Chiapas, Mexico. *CATIE: Costa Rica*. 7:7.

Mishra, G.K., Sharma, V. and Mishra, R.K. (2018) Electrochemical Aptasensors for Food and Environmental Safeguarding: A Review. *Biosensors* [online]. 8 (2). Available from: /pmc/articles/PMC6022872/ [Accessed 20 September 2023].

Morozov, V.F., Mamasakhlisov, E.S., Grigoryan, A. V, Badasyan, A. V, Hayryan, S. and Hu, C.-K. (2005) Helix–coil transition in closed circular DNA ARTICLE IN PRESS. *Physica A* [online]. 348, pp. 327–338. Available from: www.elsevier.com/locate/physa [Accessed 13 July 2017].

Motamayor, J.C., Lachenaud, P., Da Silva E Mota, J.W., Loor, R. and Kuhn, D.N. (2008) Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (Theobroma cacao L). *PLoS ONE* [online]. 3 (10), p. 3311. Available from: www.plosone.org [Accessed 19 February 2023].

Motamayor, J.C., Risterucci, A.M., Heath, M. and Lanaud, C. (2003) Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity* [online]. 91 (3), pp. 322–330. Available from: https://doi.org/10.1038/sj.hdy.6800298.

Mukherjee, S., Berger, M.F., Jona, G., Wang, X.S., Muzzey, D., Snyder, M., Young, R.A. and Bulyk, M.L. (2004) Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat. Genet.* [online]. 36 (12), pp. 1331–1339. [Accessed 15 September 2023].

Muller, E. *et al.* (2018) Next generation sequencing elucidates cacao badnavirus diversity and reveals the existence of more than ten viral species. *Virus Research* [online]. 244, pp. 235–251. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168170217306743 [Accessed 31 January 2023].

Muller, E., Jacquot, E. and Yot, P. (2001) Early detection of cacao swollen shoot virus using the polymerase chain reaction. *Journal of virological methods* [online]. 93 (1–2), pp. 15–22.

Muller, E. and Sackey, S. (2005) Molecular variability analysis of five new complete cacao swollen shoot virus genomic sequences. *Archives of Virology* [online]. 150 (1), pp. 53–66. Available from: http://link.springer.com/10.1007/s00705-004-0394-8 [Accessed 1 July 2017].

Muller, E., Ullah, I., Dunwell, J.M., Daymond, A.J., Richardson, M., Allainguillaume, J. and Wetten, A. (2021a) Identification and distribution of novel badnaviral sequences integrated in the genome of cacao (Theobroma cacao). *Scientific Reports 2021 11:1* [online]. 11 (1), pp. 1–13. Available from: https://www.nature.com/articles/s41598-021-87690-1 [Accessed 16 February 2023].

Muller, E., Ullah, I., Dunwell, J.M., Daymond, A.J., Richardson, M., Allainguillaume, J. and Wetten, A. (2021b) Identification and distribution of novel badnaviral sequences integrated in the genome of cacao (Theobroma cacao). *Scientific reports* [online]. 11 (1). Available from: https://pubmed.ncbi.nlm.nih.gov/33859254/ [Accessed 9 September 2023].

Nawaz, N., Abu Bakar, N.K., Muhammad Ekramul Mahmud, H.N. and Jamaludin, N.S. (2021) Molecularly imprinted polymers-based DNA biosensors. *Analytical Biochemistry* [online]. 630, p. 114328. [Accessed 20 September 2023].

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D. and Kim, M. (2019) Immobilized Enzymes in Biosensor Applications. *Materials* [online]. 12 (1). Available from: /pmc/articles/PMC6337536/ [Accessed 20 September 2023].

Nigam, P. (1999) COCOA AND COFFEE FERMENTATIONS. *Encyclopedia of Food Microbiology* [online]. pp. 466–473. [Accessed 21 October 2022].

References.

Ofori, A., Padi, F.K., Ameyaw, G.A., Dadzie, A.M., Opoku-Agyeman, M., Domfeh, O. and Ansah, F.O. (2022) Field evaluation of the impact of cocoa swollen shoot virus disease infection on yield traits of different cocoa (Theobroma cacao L.) clones in Ghana. *PLoS ONE* [online]. 17 (1). Available from: /pmc/articles/PMC8775274/ [Accessed 27 September 2023].

Okoffo, E.D., Mensah, M. and Fosu-Mensah, B.Y. (2016) Pesticides exposure and the use of personal protective equipment by cocoa farmers in Ghana. *Environmental Systems Research* [online]. 5 (1). [Accessed 29 September 2023].

Ooi, A.T., Stains, C.I., Ghosh, I. and Segal, D.J. (2006) SEquence-Enabled Reassembly of β-Lactamase (SEER-LAC): a Sensitive Method for the Detection of Double-Stranded DNA†. *Biochemistry* [online]. 45 (11), p. 3620. Available from: /pmc/articles/PMC2688710/ [Accessed 18 September 2023].

Oparka, K.J. and Santa Cruz, S. (2003) THE GREAT ESCAPE: Phloem Transport and Unloading of Macromolecules1. *https://doi.org/10.1146/annurev.arplant.51.1.323* [online]. 51, pp. 323–347. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev.arplant.51.1.323 [Accessed 5 September 2023].

Opoku, S., Bhattacharjee, R., Kolesnikova-Allen, M., Motamayor, J.C., Schnell, R.J., Ingelbrecht, I., Enu-Kwesi, L. and Adu-Ampomah, Y. (2007) Genetic Diversity in Cocoa (Theobroma cacao L.) Germplasm Collection from Ghana. *Journal of Crop Improvement* [online]. 20.

Park, M., Cella, L.N., Chen, W., Myung, N. V. and Mulchandani, A. (2010) Carbon nanotubes-based chemiresistive immunosensor for small molecules: Detection of nitroaromatic explosives. *Biosensors & bioelectronics* [online]. 26 (4), p. 1297. Available from: /pmc/articles/PMC3108850/ [Accessed 20 September 2023].

Persikov, A. V., Wetzel, J.L., Rowland, E.F., Oakes, B.L., Xu, D.J., Singh, M. and Noyes, M.B. (2015) A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Research* [online]. 43 (3), pp. 1965–1984.

Posnette, A. (1943) The diagnosis of swollen shoot disease of cacao. *Tropical Agriculture (Tri)* [online]. Available from: https://www.cabdirect.org/cabdirect/abstract/20057005983 [Accessed 27 September 2023].

Posnette, A.F., Robertson, N.F. and Todd, J.M. (1950) Virus diseases of cacao in West Africa. V. Alternative host plants. *Ann Appl Biol*. 37, pp. 229–240.

Potter, J.L., Nakhla, M.K. and Maxwell, D.P. (2003) PCR and DNA Hybridization Methods for Specific Detection of Bean-Infecting Begomoviruses in the Americas and Caribbean. [Accessed 17 September 2023].

Powis, T.G., Cyphers, A., Gaikwad, N.W., Grivetti, L. and Cheong, K. (2011) Cacao use and the San Lorenzo Olmec. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 108 (21), pp. 8595–8600. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1100620108 [Accessed 28 September 2023].

Pray, L.A. (2008) *Discovery of DNA Double Helix: Watson and Crick Nature Education 1(1):100*. 2008 [online]. Available from: https://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397 [Accessed 29 June 2017].

Provaznikova, D., Kumstyrova, T., Kotlin, R., Salaj, P., Matoska, V., Hrachovinova, I. and Rittich, S. (2008) High-resolution melting analysis for detection of MYH9 mutations. *Platelets* [online]. 19 (6), pp. 471–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18925516.

Przewodowska, A., Zacharzewska, B., Chołuj, J. and Treder, K. (2015) A One-Step, Real-Time Reverse Transcription Loopmediated Isothermal Amplification Assay to Detect Potato Virus Y. *American Journal of Potato Research* [online]. 92 (3), pp. 303–311. [Accessed 13 July 2023].

Qin, J., Wang, W., Gao, L. and Yao, S.Q. (2022) Emerging biosensing and transducing techniques for potential applications in point-of-care diagnostics. *Chemical Science* [online]. 13 (10), pp. 2857–2876. Available from: https://pubs.rsc.org/en/content/articlehtml/2022/sc/d1sc06269g [Accessed 14 March 2023].

Qu, J., Kang, S.G., Wang, W., Musier-Forsyth, K. and Jang, J.C. (2014) Arabidopsis thaliana tandem zinc finger 1 (AtTZF1) protein in RNA binding and decay. *The Plant journal : for cell and molecular biology* [online]. 78 (3), p. 452. Available from: /pmc/articles/PMC4026020/ [Accessed 16 September 2023].

Quainoo, A.K., Wetten, A.C. and Allainguillaume, J. (2008) Transmission of cocoa swollen shoot virus by seeds. *Journal of Virological Methods* [online]. 150, pp. 45–49. [Accessed 16 June 2023].

Razo, S.C., Panferov, V.G., Safenkova, I. V., Varitsev, Y.A., Zherdev, A. V. and Dzantiev, B.B. (2018) Double-enhanced lateral flow immunoassay for potato virus X based on a combination of magnetic and gold nanoparticles. *Analytica Chimica Acta* [online]. 1007, pp. 50–60. [Accessed 28 September 2023].

Reja, V., Kwok, A., Stone, G., Yang, L., Missel, A., Menzel, C. and Bassam, B. (2010) ScreenClust: Advanced statistical software for supervised and unsupervised high resolution melting (HRM) analysis. *Methods (San Diego, Calif.)* [online]. 50 (4), pp. S10-4. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1046202310000563 [Accessed 13 September 2017].

Reybroeck, W., Ooghe, S., De Brabander, H. and Daeseleire, E. (2007) Validation of the tetrasensor honey test kit for the screening of tetracyclines in honey. *Journal of Agricultural and Food Chemistry* [online]. 55 (21), pp. 8359–8366. [Accessed 20 September 2023].

Rezadoost, M.H., Kordrostami, M. and Kumleh, H.H. (2016) An efficient protocol for isolation of inhibitor-free nucleic acids even from recalcitrant plants. *3 Biotech* [online]. 6 (1), pp. 1–7. Available from: /pmc/articles/PMC4752943/ [Accessed 18 September 2023].

Del Rio, S. and Setzer, D.R. (1993) The role of zinc fingers in transcriptional activation by transcription factor IIIA. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 90 (1), p. 168.

Rizzi, G., Dufva, M. and Hansen, M.F. (2017) Two-dimensional salt and temperature DNA denaturation analysis using a magnetoresistive sensor. *Lab on a Chip* [online]. 17 (13), pp. 2256–2263. Available from: https://pubs.rsc.org/en/content/articlehtml/2017/lc/c7lc00485k [Accessed 17 September 2023].

Rodriguez, C.M., Guzma, B., Lloyd, A.J. and Wilkinson, M.J. (2010) Direct Detection and Quantification of Methylation in Nucleic Acid Sequences Using High-Resolution Melting Analysis. 82 (21), pp. 9100–9108.

Rosenblum, Mort. (2005) *Chocolate: a bittersweet saga of dark and light*. North Point Press. [Accessed 29 September 2023].

Rothemund, P.W.K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* [online]. 440 (7082), pp. 297–302. Available from: https://pubmed.ncbi.nlm.nih.gov/16541064/ [Accessed 18 September 2023].

Saccà, B. and Niemeyer, C.M. (2012) DNA origami: the art of folding DNA. *Angewandte Chemie (International ed. in English)* [online]. 51 (1), pp. 58–66. Available from: https://pubmed.ncbi.nlm.nih.gov/22162047/ [Accessed 18 September 2023].

Sackey, S. and Hull, R. (1992) *The use of dot blot hybridisation methods to detect cocoa swollen shoot virus isolates. Rep. Cocoa Res. Inst. Ghana*.

Sackey, S. T.; Lowor, S. T. & Dzahini-Obiatey, H. (1995) *Classification of CSSV isolates by PCR and DNA hybridisation.*

Sackey, ST; Bartels, PK; Amponsah, R. (1992) *Clone PCR amplification products for identification of CSSV 1A and Nsaba by DNA hybridisation.*

Saenger, W. (2013) Proteinase K. *Handbook of Proteolytic Enzymes* [online]. 3, pp. 3240–3242. [Accessed 17 September 2023].

Sagemann, W., Paul, H.L., Adomako, D. and Owusu, G.K. (1983) The Use of Enzyme-Linked Immunosorbent Assay (ELISA) for Detection of Cacao Swollen Shoot Virus (CSSV) in Theobroma cacao. *Journal of Phytopathology* [online]. 106 (3), pp. 281–284. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1439-0434.1983.tb00053.x [Accessed 17 February 2023].

Saleh, M. and El-Matbouli, M. (2015) Rapid detection of Cyprinid herpesvirus-3 (CyHV-3) using a gold nanoparticle-based hybridization assay. *Journal of virological methods* [online]. 217, pp. 50–54. Available from: https://pubmed.ncbi.nlm.nih.gov/25738211/ [Accessed 18 September 2023].

Sánchez-Paniagua López, M., Redondo-Gómez, E. and López-Ruiz, B. (2017) Electrochemical enzyme biosensors based on calcium phosphate materials for tyramine detection in food samples. *Talanta* [online]. 175, pp. 209–216. Available from: https://pubmed.ncbi.nlm.nih.gov/28841980/ [Accessed 21 September 2023].

Scaletsky, I.C.A., Fabbricotti, S.H., Aranda, K.R., Morais, M.B. and Fagundes-Neto, U. (2002) Comparison of DNA Hybridization and PCR Assays for Detection of Putative Pathogenic Enteroadherent Escherichia coli. *Journal of Clinical Microbiology* [online]. 40 (4), p. 1254. Available from: /pmc/articles/PMC140355/ [Accessed 17 September 2023].

Schindelin, J. *et al.* (2012a) Fiji: an open-source platform for biological-image analysis. *Nature Methods 2012 9:7* [online]. 9 (7), pp. 676–682. Available from: https://www.nature.com/articles/nmeth.2019 [Accessed 10 September 2023].

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T. and Cardona, A. (2012b) *Fiji ImageJ (1.54f)*. pp. 676–682.

Schroth, G., Läderach, P., Martinez-Valle, A.I., Bunn, C. and Jassogne, L. (2016) Vulnerability to climate change of cocoa in West Africa: Patterns, opportunities and limits to adaptation. *Science of The Total Environment* [online]. 556, pp. 231–241. [Accessed 24 September 2023].

Sharma, S., Byrne, H. and O'Kennedy, R.J. (2016) Antibodies and antibody-derived analytical biosensors. *Essays in Biochemistry* [online]. 60 (1), p. 9. Available from: /pmc/articles/PMC4986469/ [Accessed 20 September 2023].

Sharma, S.K., Kumar, P.V. and Baranwal, V.K. (2014) Immunodiagnosis of episomal Banana streak MY virus using polyclonal antibodies to an expressed putative coat protein. *Journal of Virological Methods* [online]. 207, pp. 86–94. [Accessed 29 September 2023].

Shastry, B.S. (1996) Transcription factor IIIA (TFIIIA) in the second decade. *Journal of cell science* [online]. 109 (Pt 3) (3), pp. 535–539. Available from: https://pubmed.ncbi.nlm.nih.gov/8907699/ [Accessed 6 July 2023].

Shim, J., Williams, L., Kim, D., Ko, K. and Kim, M.S. (2021) Application of Engineered Zinc Finger Proteins Immobilized on Paramagnetic Beads for Multiplexed Detection of Pathogenic DNA. *Journal of Microbiology and Biotechnology* [online]. 31 (9), p. 1323. Available from: /pmc/articles/PMC9705829/ [Accessed 28 September 2023].

Shin, W.R., Sekhon, S.S., Rhee, S.K., Ko, J.H., Ahn, J.Y., Min, J. and Kim, Y.H. (2018) Aptamer-Based Paper Strip Sensor for Detecting Vibrio fischeri. *ACS Combinatorial Science* [online]. 20 (5), pp. 261–268. Available from: https://pubs.acs.org/doi/abs/10.1021/acscombsci.7b00190 [Accessed 21 September 2023].

Simpson, M. (2010) *Plant systematics: Second edition Plant Systematics: Second Edition*. Academic Press.

Singh, J., Cobb-Smith, D., Higgins, E. and Khan, A. (2021) Comparative evaluation of lateral flow immunoassays, LAMP, and quantitative PCR for diagnosis of fire blight in apple orchards. *Journal of Plant Pathology* [online]. 103 (Suppl 1), p. 131. Available from: /pmc/articles/PMC7456764/ [Accessed 13 July 2023].

Skottrup, P.D., Nicolaisen, M. and Justesen, A.F. (2008) Towards on-site pathogen detection using antibody-based sensors. *Biosensors and Bioelectronics* [online]. 24 (3), pp. 339–348.

Song, Y., Wei, W. and Qu, X. (2011) Colorimetric biosensing using smart materials. *Advanced materials (Deerfield Beach, Fla.)* [online]. 23 (37), pp. 4215–4236. Available from: https://pubmed.ncbi.nlm.nih.gov/21800383/ [Accessed 18 September 2023].

Soper, S.A. *et al.* (2006) Point-of-care biosensor systems for cancer diagnostics/prognostics. *Biosensors & bioelectronics* [online]. 21 (10), pp. 1932–1942. Available from: https://pubmed.ncbi.nlm.nih.gov/16473506/ [Accessed 18 September 2023].

Squicciarini, M.P. (2016) *The economics of chocolate*. Oxford University Press.

Sterchi, D.L. (2008) Molecular Pathology—In Situ Hybridization. *Theory and Practice of Histological Techniques* [online]. p. 537. Available from: /pmc/articles/PMC7310933/ [Accessed 26 September 2023].

Stewart, M.E., Anderton, C.R., Thompson, L.B., Maria, J., Gray, S.K., Rogers, J.A. and Nuzzo, R.G. (2008) Nanostructured plasmonic sensors. *Chemical reviews* [online]. 108 (2), pp. 494–521. Available from: https://pubmed.ncbi.nlm.nih.gov/18229956/ [Accessed 18 September 2023].

Stott, D.I. (1989) Immunoblotting and dot blotting. *Journal of Immunological Methods* [online]. 119 (2), p. 153. Available from: /pmc/articles/PMC7172508/ [Accessed 16 September 2023].

Sun, Z., Li, J., Yang, Y., Tong, Y., Li, H., Wang, C., Du, L. and Jiang, Y. (2022) Ratiometric Fluorescent Biosensor Based on Self-Assembled Fluorescent Gold Nanoparticles and Duplex-Specific Nuclease-Assisted Signal Amplification for Sensitive Detection of Exosomal miRNA. *Bioconjugate chemistry*

[online]. 33 (9), pp. 1698–1706. Available from: https://pubmed.ncbi.nlm.nih.gov/35960898/ [Accessed 18 September 2023].

Swiss DeCode (2021) *DNAFoil CSSV*. 2021. [Accessed 2 May 2023].

Tan, Z.-J. and Chen, S.-J. (2006) Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophysical journal* [online]. 90 (4), pp. 1175–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16299077%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1367269 [Accessed 12 July 2017].

Tang, Y. bing, Xing, D., Zhu, D. bin and Liu, J. feng (2007a) An improved electrochemiluminescence polymerase chain reaction method for highly sensitive detection of plant viruses. *Analytica Chimica Acta* [online]. 582 (2), pp. 275–280. [Accessed 28 September 2023].

Tang, Y. bing, Xing, D., Zhu, D. bin and Liu, J. feng (2007b) An improved electrochemiluminescence polymerase chain reaction method for highly sensitive detection of plant viruses. *Analytica Chimica Acta* [online]. 582 (2), pp. 275–280. [Accessed 12 July 2023].

Tang, Y., Yan, X., Gu, C. and Yuan, X. (2022) Biogenesis, Trafficking, and Function of Small RNAs in Plants. *Frontiers in Plant Science* [online]. 13, p. 825477. [Accessed 24 September 2023].

Taylor, T., Denson, J.P. and Esposito, D. (2017) Optimizing expression and solubility of proteins in E. coli using modified media and induction parameters. *Methods in Molecular Biology* [online]. 1586, pp. 65–82. Available from: https://link.springer.com/protocol/10.1007/978-1-4939-6887-9_5 [Accessed 16 September 2023].

Teengam, P., Siangproh, W., Tuantranont, A., Vilaivan, T., Chailapakul, O. and Henry, C.S. (2017) Multiplex Paper-Based Colorimetric DNA Sensor Using Pyrrolidinyl Peptide Nucleic Acid-Induced AgNPs Aggregation for Detecting MERS-CoV, MTB, and HPV Oligonucleotides. *Analytical Chemistry* [online]. 89 (10), pp. 5428–5435. Available from: /pmc/articles/PMC7077925/ [Accessed 18 September 2023].

Teycheney, P.Y. *et al.* (2020) ICTV Virus taxonomy profile: Caulimoviridae. *Journal of General Virology* [online]. 101 (10), pp. 1025–1026.

Thi, L. *et al.* (2016) Characterization of leaf, flower, and pod morphology among vietnamese cocoa varieties (Theobroma cacao l.). *Pak. J. Bot*. 48 (6), pp. 2375–2383. [Accessed 21 October 2022].

Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G. and van Etten, J. (2012) Present spatial diversity patterns of Theobroma cacao L. in the neotropics reflect genetic differentiation in pleistocene refugia followed by human-influenced dispersal. *PloS one* [online]. 7 (10). Available from: https://pubmed.ncbi.nlm.nih.gov/23112832/ [Accessed 27 October 2022].

Thomas, R. (1993) The denaturation of DNA. *Gene* [online]. 135 (1), pp. 77–79. Available from: http://www.sciencedirect.com/science/article/pii/0378111993900514 [Accessed 13 July 2017].

Torrance, L., Ziegler, A., Pittman, H., Paterson, M., Toth, R. and Eggleston, I. (2006) Oriented immobilisation of engineered single-chain antibodies to develop biosensors for virus detection. *Journal of Virological Methods* [online]. 134 (1–2), pp. 164–170.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K. and Voytas, D.F. (2009) High frequency modification of plant genes using engineered zinc finger nucleases. *Nature* [online]. 459 (7245), p. 442.

References.

Trilling, A.K., Beekwilder, J. and Zuilhof, H. (2013) Antibody orientation on biosensor surfaces: a minireview. *The Analyst* [online]. 138 (6), pp. 1619–1627. Available from: https://pubmed.ncbi.nlm.nih.gov/23337971/ [Accessed 20 September 2023].

Ullah, I., Daymond, A.J., Hadley, P., End, M.J., Umaharan, P. and Dunwell, J.M. (2021) Identification of cacao mild mosaic virus (Cammv) and cacao yellow vein-banding virus (cyvbv) in cocoa (theobroma cacao) germplasm. *Viruses* [online]. 13 (11). Available from: /pmc/articles/PMC8623607/ [Accessed 29 September 2023].

Veselova, S. V., Farkhutdinov, R.G., Veselov, D.S. and Kudoyarova, G.R. (2006) Role of cytokinins in the regulation of stomatal conductance of wheat seedlings under conditions of rapidly changing local temperature. *Russian Journal of Plant Physiology* [online]. 53 (6), pp. 756–761. Available from: https://link.springer.com/article/10.1134/S1021443706060057 [Accessed 22 September 2023].

Vo, T.D., Schneider, A.L., Wilson, W.D. and Poon, G.M.K. (2021) Salt bridge dynamics in protein/DNA recognition: a comparative analysis of Elk1 and ETV6. *Physical chemistry chemical physics : PCCP* [online]. 23 (24), p. 13490. Available from: /pmc/articles/PMC8233815/ [Accessed 26 September 2023].

Vossen, R.H.A.M., Aten, E., Roos, A. and Den Dunnen, J.T. (2009) *High-resolution melting analysis (HRMA) - More than just sequence variant screening Human Mutation* [online]. 30 (6), pp. 860–866.

Vrana, K.E., Churchill, M.E., Tullius, T.D. and Brown, D.D. (1988) Mapping functional regions of transcription factor TFIIIA. *Mol Cell Biol* [online]. 8 (4), pp. 1684–1696. [Accessed 16 September 2023].

De Vuyst, L. and Leroy, F. (2020) Functional role of yeasts, lactic acid bacteria and acetic acid bacteria in cocoa fermentation processes. *FEMS Microbiology Reviews* [online]. 44 (4), pp. 432–453. [Accessed 29 September 2023].

Wang, S., Zhou, Z., Ma, N., Yang, S., Li, K., Teng, C., Ke, Y. and Tian, Y. (2020) DNA Origami-Enabled Biosensors. *Sensors (Basel, Switzerland)* [online]. 20 (23), pp. 1–18. Available from: /pmc/articles/PMC7731452/ [Accessed 18 September 2023].

Wang, X., Lim, H.J. and Son, A. (2014) Characterization of denaturation and renaturation of DNA for DNA hybridization. *Environmental health and toxicology* [online]. 29, p. e2014007. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4168728&tool=pmcentrez&rendertype=abstract.

Watson, J.D. and Crick, F.H.C. (1953) *Molecular structure of nucleic acids Nature* [online]. 171 (4356), pp. 737–738. Available from: http://www.nature.com/physics/looking-back/crick/\nhttp://www.ncbi.nlm.nih.gov/pubmed/13054692.

Wetten, A., Campbell, C. and Allainguillaume, J. (2016) High-resolution melt and morphological analyses of mealybugs (Hemiptera: Pseudococcidae) from cacao: Tools for the control of Cacao swollen shoot virus spread. *Pest Management Science* [online]. 72 (3), pp. 527–533. [Accessed 19 February 2023].

Wijerathna-Yapa, A. and Pathirana, R. (2022) Sustainable Agro-Food Systems for Addressing Climate Change and Food Security. *Agriculture 2022, Vol. 12, Page 1554* [online]. 12 (10), p. 1554. Available from: https://www.mdpi.com/2077-0472/12/10/1554/htm [Accessed 24 September 2023].

References.

Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G. and Pryor, R.J. (2003) High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. *Clinical Chemistry* [online]. 49 (6). Available from: http://clinchem.aaccjnls.org/content/49/6/853.full [Accessed 9 July 2017].

Wolfe, S.A., Nekludova, L. and Pabo, C.O. (2003) DNA Recognition by Cys2His2 Zinc Finger Proteins. *https://doi.org/10.1146/annurev.biophys.29.1.183* [online]. 29, pp. 183–212.

Wong, I. and Lohman, T.M. (1993) A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interactions (DNA-protein interactions/equilibrium binding). *Biochemistry*. 90, pp. 5428–5432.

Wongkaew, P., Poosittisak, S., Wongkaew, P. and Poosittisak, S. (2014) Diagnosis of Sugarcane White Leaf Disease Using the Highly Sensitive DNA Based Voltammetric Electrochemical Determination. *American Journal of Plant Sciences* [online]. 5 (15), pp. 2256–2268. Available from: http://www.scirp.org/Html/8-2601490_47637.htm [Accessed 28 September 2023].

Xia, R., Cheng, Y., Han, X., Wei, Y. and Wei, X. (2021) Ikaros Proteins in Tumor: Current Perspectives and New Developments. *Frontiers in Molecular Biosciences* [online]. 8, p. 788440. [Accessed 12 July 2023].

Xu, M., Dai, T., Wang, Y. and Yang, G. (2022) The incipient denaturation mechanism of DNA. *RSC Advances* [online]. 12 (36), p. 23356. Available from: /pmc/articles/PMC9383117/ [Accessed 17 September 2023].

Xu, M., Xing, S., Xu, X., Fu, P., Xu, W. and Zhao, C. (2018) Label-free colorimetric aptasensor for highly sensitive and selective detection of proteins by using PNA/DNA hybrids and a cyanine dye. *Analytical Methods* [online]. 10 (31), pp. 3824–3829. Available from: https://pubs.rsc.org/en/content/articlehtml/2018/ay/c8ay01071d [Accessed 20 September 2023].

Yalow, R.S. and Berson, S.A. (1959) Assay of plasma insulin in human subjects by immunological methods. *Nature* [online]. 184 (Suppl 21) (4699), pp. 1648–1649. Available from: https://pubmed.ncbi.nlm.nih.gov/13846363/ [Accessed 28 September 2023].

Yan, Y., Ham, B.K., Chong, Y.H., Yeh, S.D. and Lucas, W.J. (2020) A Plant SMALL RNA-BINDING PROTEIN 1 Family Mediates Cell-to-Cell Trafficking of RNAi Signals. *Molecular Plant* [online]. 13 (2), pp. 321–335. Available from: http://www.cell.com/article/S1674205219303971/fulltext [Accessed 24 September 2023].

Yang, Q. *et al.* (2019) Array atomic force microscopy for real-time multiparametric analysis. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 116 (13), pp. 5872–5877. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1813518116 [Accessed 16 September 2023].

Yang, Y., Wang, H. and Erie, D.A. (2003) Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods* [online]. 29 (2), pp. 175–187. [Accessed 16 September 2023].

Yoo, E.H. and Lee, S.Y. (2010) Glucose Biosensors: An Overview of Use in Clinical Practice. *Sensors (Basel, Switzerland)* [online]. 10 (5), p. 4558. Available from: /pmc/articles/PMC3292132/ [Accessed 21 September 2023].

References.

Yu, Y., Zhang, Y., Chen, X. and Chen, Y. (2019) Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Annual review of cell and developmental biology* [online]. 35, pp. 407–431. Available from: https://pubmed.ncbi.nlm.nih.gov/31403819/ [Accessed 24 September 2023].

Zarrillo, S. *et al.* (2018a) The use and domestication of Theobroma cacao during the mid-Holocene in the upper Amazon. *Nature ecology & evolution* [online]. 2 (12), pp. 1879–1888. Available from: https://pubmed.ncbi.nlm.nih.gov/30374172/ [Accessed 29 September 2023].

Zarrillo, S. *et al.* (2018b) The use and domestication of Theobroma cacao during the mid-Holocene in the upper Amazon. *Nature Ecology & Evolution 2018 2:12* [online]. 2 (12), pp. 1879–1888.

Zeitler, B. *et al.* (2019) Allele-selective transcriptional repression of mutant HTT for the treatment of Huntington's disease. *Nature Medicine 2019 25:7* [online]. 25 (7), pp. 1131–1142. Available from: https://www.nature.com/articles/s41591-019-0478-3 [Accessed 15 September 2023].

Zhao, D. and Huang, Z. (2016) Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorganic Chemistry and Applications* [online]. 2016. Available from: /pmc/articles/PMC4971304/ [Accessed 17 September 2023].

Zhao, Y., Liu, L., Kong, D., Kuang, H., Wang, L. and Xu, C. (2014) Dual amplified electrochemical immunosensor for highly sensitive detection of Pantoea stewartii sbusp. stewartii. *ACS Applied Materials and Interfaces* [online]. 6 (23), pp. 21178–21183. Available from: https://pubs.acs.org/doi/full/10.1021/am506104r [Accessed 20 September 2023].