University of Reading

Identification of a novel polerovirus in cocoa (Theobroma cacao) germplasm and development of molecular methods for use in diagnostics

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Ullah, I. ORCID: https://orcid.org/0000-0002-9367-6741, Kamran, M. and Dunwell, J. M. ORCID: https://orcid.org/0000-0003-2147-665X (2023) Identification of a novel polerovirus in cocoa (Theobroma cacao) germplasm and development of molecular methods for use in diagnostics. Pathogens, 12 (11). 1284. ISSN 2076-0817 doi: 10.3390/pathogens12111284 Available at https://centaur.reading.ac.uk/113813/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.3390/pathogens12111284

Publisher: MDPI

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

pathogens

MDPI

Article

Identification of a Novel Polerovirus in Cocoa (*Theobroma cacao*) Germplasm and Development of Molecular Methods for Use in Diagnostics

Ihsan Ullah [1], Muhammad Kamran [2] and Jim M. Dunwell [1,*]

[1] School of Agriculture, Policy and Development, University of Reading, Reading RG6 6EU, UK; i.ullah@reading.ac.uk
[2] Plant Pathology Research Institute, Ayub Agricultural Research Institute, Faisalabad 38850, Pakistan; mkamran.uaf.pk@gmail.com
* Correspondence: j.m.dunwell@reading.ac.uk; Tel.: +44-1183786313

Abstract: The cocoa crop (*Theobroma cacao* L.) is known to be a host for several badnaviruses, some of which cause severe disease, while others are asymptomatic. Recently, the first preliminary evidence has been published concerning the occurrence of a polerovirus in cacao. We report here the first near-complete genome sequence of cacao polerovirus (CaPV) by combining bioinformatic searches of cacao transcript databases, with cloning from the infected germplasm. The reported novel genome has all the genome features known for poleroviruses from other species. Pairwise identity analyses of RNA-dependent RNA polymerase and coat protein indicates < 60% similarity of CaPV with any reported poleroviruses; hence, we propose that the polerovirus isolate reported in this study is a novel polerovirus. The genome sequence information was also used to develop a multiplex RT-PCR assay, which was applied to screen a selected range of germplasms and to identify several infected clones. Although there is no evidence that this virus causes any severe disease, this new information, together with a robust diagnostic assay, are of strategic importance in developing protocols for the safe international transfer of cacao germplasms.

Keywords: cacao; polerovirus; genome; RT-PCR; NGS; ORF

check for updates

Citation: Ullah, I.; Kamran, M.; Dunwell, J.M. Identification of a Novel Polerovirus in Cocoa (*Theobroma cacao*) Germplasm and Development of Molecular Methods for Use in Diagnostics. *Pathogens* **2023**, *12*, 1284. https://doi.org/10.3390/pathogens12111284

Academic Editors: Sudeep Bag, Albert K. Culbreath and Rajagopalbabu Srinivasan

Received: 11 October 2023
Revised: 23 October 2023
Accepted: 25 October 2023
Published: 26 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cacao (*Theobroma cacao*) holds significant economic importance for both producing and consuming countries. This tropical evergreen tree is predominantly grown in West Africa, Central and South America, and Southeast Asia. The cacao beans, also referred to as cocoa, represent an important ingredient for food, pharmaceutical and cosmetics industries. In 2022, the global chocolate confectionery market alone generated a revenue of approximately 0.22 trillion U.S. dollars [1]. The sustainable production of cocoa is facing numerous challenges. Lack of access to improved planting materials, poor agricultural practices, impoverished soils and nutrient deficiencies, climate change and losses due to pests and diseases influence cacao productivity.

Fungal and other diseases like black pod rot, caused by the *Phytophthora* species, and witches' broom of cacao, caused by *Moniliophthora perniciosa*, cause significant loss in cacao yields. Similar to other crop species, cacao is also susceptible to viral infections. Until recently, most viral diseases of cacao were believed to be caused by the members of the genus *Badnavirus*. On the basis of their impact on plant growth and productivity, cacao badnaviral diseases can be categorized into two groups. The first, cacao swollen shoot virus disease (CSSVD), is a serious threat to cacao trees. CSSVD is caused by seven species related to the genus *Badnavirus* [2] and is confined to the cacao growing region of West Africa, which contributes over 74% of global cacao production [3]. The second group of badnaviral diseases are caused by the species *cacao bacilliform Sri Lanka virus* (CBSLV), *cacao*

mild mosaic virus (CaMMV) and *cacao yellow vein banding virus* (CYVBV). The latter two species have been reported to infect cacao plantations in multiple countries [4–8], whereas CBSLV has been detected in Sri Lanka [2]. To date, none of these infections have been associated with any effect on plant growth or productivity. Badnaviruses have also been reported as endogenized in the cacao genome [9,10].

Solemoviridae is a family of plant infecting viruses that can significantly influence yield and quality in crops such as cotton [11]. The family consists of four genera: *Enamovirus*, *Polemovirus*, *Polerovirus* and *Sobemovirus* [12]. The 4–6 kb non-enveloped, positive-strand RNA genome comprises 4–10 open reading frames (ORFs). The genome contains a virus protein genome-linked (VPg) cap at the 5′ terminus, whereas the 3′ terminus is non-polyadenylated but is protected by a stable stem-loop or tRNA-like structure [13]. The members of these four genera vary regarding nucleotide sequence, genome size and number and arrangement of ORFs. Poleroviruses are closer to the enamoviruses in terms of the structure and expression of the genome [14]. Enamoviruses and poleroviruses contain a non-coding intergenic region, which is absent in the genomes of polemoviruses and sobemoviruses. Enamoviruses and poleroviruses vary in terms of nucleotide sequence and coding regions, with a major difference being the absence of the movement protein in enamoviruses [15].

Recently, two studies have reported the detection of virus species from the genus *Polerovirus* in cacao trees in the Americas. The first study detected polerovirus-like partial sequences, designated as cacao leafroll virus, in cacao germplasms imported into USDA-ARS-SHRS cacao quarantine facilities. The plants exhibited leaf distortion with downward rolling of leaves, discoloration and yellow speckling or mottling [16]. The second study reported the infection of cotton leafroll dwarf virus (CLRDV) in cacao trees in Bahia, Brazil, exhibiting virus-like symptoms such as leaf mosaicism, vein clearing and yellow spots [17]. Illumina sequencing and RT-PCR of the cacao leaf samples yielded two highly similar contigs, which represented ~29% of the CLRDV genome. The contigs showed >85% identity with CLRDV isolates previously reported in South America and the USA. Interestingly, 14 out of 30 symptomatic plants tested negative. It was argued that the symptoms on the negative plants could have been caused by other viral groups or abiotic stress.

In terms of host range, Polerovirus is a diverse genus infecting a wide range of plants, from monocots to eudicots. Poleroviruses are phloem-limited viruses, which are believed, under usual conditions, to be solely transmitted by aphids. The infection may cause stunt, interveinal yellowing, discoloration of the main leaf vein, leaf curling and leaf necrosis. The International Committee on Taxonomy of Viruses (ICTV) at present has recognised 26 member species in the genus. Six hundred and two complete genome sequences for species have been recognised by ICTV, as well as unrecognised poleroviruses that have been reported in NCBI (https://www.ncbi.nlm.nih.gov/nuccore; accessed on 20 September 2023), out of which 42 genomes have been recognised as reference sequences. These reported genomes are ascribed to over 90 host plant species. The poleroviruses genome length in 26 exemplar species listed in the Virus Metadata Resource (VMR; https://ictv.global/vmr; last accessed on 30 September 2023) ranged from 5612 (maize yellow dwarf virus-RMV; NC_021484) to 6244 nucleotides (NC_075422; pepper vein yellows virus).

The polerovirus genome consists of six major ORFs. A non-conserved 5′-proximal ORF0 encodes an RNA silencing suppressor [18]. The ORF1 encodes a polyprotein P1, a protease- and genome-linked protein. The Polerovirus genome encodes the RNA-dependent RNA polymerase (RdRP) by −1 frameshift (−1 rfs) translational fusion of ORF1 through ORF2. The ORF2 is translated from genomic RNA using a ribosomal leaky scanning strategy [19]. The −1 rfs signal of the polerovirus consists of the slippery sequence 5′-GGGAAAC-3′ followed by a putative small pseudoknot. Other slippery heptanucleotide variants can also be found in polerovirus genomes [20,21]. The 3′-proximal ORFs are separated from 5′-proximal ORFs by a non-coding intergenic region. These ORFs are translated from sub-genomic RNA(s). The ORF3 and ORF4 encode the coat protein (CP),

and movement protein (MP). The ORF4 overlaps ORF3 but in a different reading frame. The ORF3 and ORF5 combine to form the coat protein-readthrough domain fusion protein (P3-P5). In the process, ribosomes skip the stop codon at the end of ORF3 and continue to translate ORF5 in the same frame [22]. A short non-AUG-initiated ORF3a, which encodes a different MP (P3a), has also been described in poleroviruses [23].

Though polerovirus infections have been found in cacao, and though partial nucleotide sequences of the RdRP gene, along with CP and MP genes, have been reported, the genome of cacao polerovirus has not been sequenced. Here, we report the first near-complete sequence of a novel polerovirus with the proposed name of cacao polerovirus (CaPV). We also report a multiplex diagnostic RT-PCR test for the detection of this virus.

2. Materials and Methods

2.1. Analysis of Transcriptome Datasets

The Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra; last accessed on 30 June 2023) was searched for publicly available RNA sequence (RNA-Seq) datasets of *T. cacao*. Initially, the datasets were searched with the BLASTn tool for cacao polerovirus, using a reference database consisting of five cacao leafroll virus sequences (GenBank accessions ON745771-75). The raw read FastQ format files of the RNA-Seq dataset in BioProject PRJNA558793 were downloaded from the SRA for comprehensive analysis. The OmicsBox 3.0.30 software package (available at https://www.biobam.com/omicsbox; last accessed on 30 September 2023) was employed, to align the read to the reference databases using a BWA mapper. The Trinity assembler software 2.14.0 in the OmicsBox was used to construct an RNA-Seq de novo transcriptome assembly from the RNA-Seq dataset of cacao accession IMC60 (SRR9938230). A BLASTx search was conducted in the functional analysis module of the OmicsBox, using the de novo transcriptome assembly as a query against the non-redundant protein database (nr) of the *Solemoviridae* family.

2.2. Sequence Analysis and Phylogeny

The de novo contig identified in the BLASTx search, with enquiry set to Solemoviruses, was subjected to sequence analysis. The putative ORFs and predicted amino acid sequences were identified using the CLC sequence viewer (Qiagen, Manchester, UK). The protein domain features were predicted using the InterPro database (https://www.ebi.ac.uk/interpro/; last accessed on 30 September 2023). The amino acid sequences of the RdRP and CP of 43 representative species in *Solemoviridae* family were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/; last accessed on 30 September 2023). The obtained sequences were used as the source data for pairwise comparisons, using Sequence Demarcation Tool version 1.2 (SDT v1.2) [24], and the construction of phylogenetic trees using the maximum likelihood method in the RAxML, with a JTT substitution model [25]. One thousand bootstrap samplings were carried out. Trees were drawn in MEGA11 [26].

2.3. RNA Extraction and cDNA Synthesis

A fully expanded young leaf was sampled for RNA extraction from each of 26 cacao accessions from the germplasm held at International Cocoa Quarantine Centre, Reading, UK (ICQC-R). The selected germplasm was imported from the International Cacao Collection (IC3) at Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Costa Rica, between 1988 and 2021.

Four leaf discs of 8 mm diameter were ground with liquid nitrogen in a 2 mL Eppendorf tube in the presence of ceramic beads, using a TissueLyser II (Qiagen, UK). Total RNA was isolated by a modified CTAB method, followed by purification with the Spectrum™ plant total RNA kit (Merck, Gillingham, UK) according to the manufacturer's instructions. DNA digestion was performed with an On-Column DNase I Digestion Set (Merck, Gillingham, UK) for the removal of trace amounts of genomic DNA. The quality and quantity of isolated RNA was determined using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Horsham, UK). Aliquots of RNA were also run on 1.2% agarose gel for

quality analysis. Isolated total RNA (300–1000 ng) was transcribed into cDNA using the SuperScript™ IV VILO™ Master Mix with ezDNase™ Enzyme (Thermo Fisher Scientific, UK). The cDNA was diluted 5-fold with RNase- and DNase-free water.

2.4. Primer Design and RT-PCR Amplification

The primer sets were designed with the online tool Primer3Plus (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi; last accessed on 30 September 2023). The cacao polerovirus-specific diagnostic primers were designed from ORFs encoding RdRP, MP and P5 proteins in the identified contig. The *T. cacao* acyl carrier protein 1 gene (ACP1; LOC18599903) was selected as the internal control (IC), as it constitutively expresses across the various cacao tissues [27]. A primer set was designed from exon two and three of the APC1 gene to be used as the IC in diagnostic multiplex RT-PCR assays. Primer sets were designed from the 5 and 3′ UTR region for the amplification of a near-complete genome of CaPV.

For the diagnostic analysis, a 20 µL single or multiplex RT-PCR reaction consisting of 10 µL of Platinum Hot Start PCR Master Mix, 1.5 µL each of 5 µM primers, 5 µL of cDNA template and x µL of PCR water was performed, with thermal cycling of one cycle of 94 °C for 2 min, followed by thirty five cycles of 94 °C for 10 s, 60 °C for 10 s and 68 °C for 25 s.

The Phusion Flash PCR Master Mix was used to amplify the genome sequence which was harbouring coding sequences. The RT-PCR reaction, which contained 12.5 µL of PCR Mix, 2 µL each of 5 µM forward and reverse primers, 5 µL of cDNA template and 4.5 µL of PCR water, was run at one cycle of 98 °C for 10 s, followed by thirty five cycles of 98 °C for 1 s, 60 °C for 5 s and 72 °C for 150 s. Final extension was performed at 72 °C for 3 min. The RT-PCR reactions were performed in a thermal cycler (Veriti, Applied Biosystems). The amplicon was cloned with a Zero Blunt™ TOPO™ PCR Cloning Kit for Sequencing. The cloned RT-PCR products were sequenced using Sanger technology (Source Bioscience, Nottingham, UK). The sequencing data were assembled using SeqMan Ultra contig assembly software Version 17.3 (DNASTAR, Madison, USA). The RT-PCR products were resolved on 1–1.7% agarose gel and stained with ethidium bromide. Information about the primers used in this study, including primer sequences, specific targets, coordinates and product size, is described in Table 1.

Table 1. Primers used in cloning and detection of cacao polerovirus. RdRP, MP and P5 denote RNA-directed RNA polymerase, movement protein and readthrough domain protein, respectively. ACP1 denotes *Theobroma cacao* acyl carrier protein 1.

Specificity	Primer Location	Primer Name	Primer Sequence	Coordinates (nt)	Amplicon Size (bp)
Cacao polerovirus	5′ UTR	CLN_F	CGACCGTATCATTGCTTGTG	33–52	5794
	3′ UTR	CLN_R	CCATGCCTGTGGGTTTTAAG	5807–5826	
	RdRP (ORF2)	P2_F	GTACCCTCCTAGCCCAGACC	3106–3125	150
		P2_R	CTGGGGATTCTAAGGCATCA	3236–3255	
	MP (ORF4)	P4_F	AGGGAGCAACAGCTTCACA	3954–3972	201
		P4_R	GAGATGGAGCCTGATGTCGT	4135–4154	
	P5 (ORF5)	P5_F	CAAGGAGTGCTGGTCGTACA	4990–5009	212
		P5_R	GGGGTTTGATTTACCGGTTT	5182–5203	
Theobroma cacao	ACP1_Exon2	ACP_F	GAAGCAGCCTCTCATTTAATTTG	1148–1170	* 90
	ACP1_Exon3	ACP_R	CACATACCTTATCCACAGTCTCT	1754–1776	** 629

* Size from mRNA (XM_007030091.2); ** Size from gene (LOC18599903).

3. Results

3.1. Detection of Polerovirus Sequences in Transcriptome Datasets

A recent study on the detection of partial sequences of a polerovirus (designated as cacao leafroll virus) in an imported cacao germplasm at the USDA-ARS-SHRS cacao

quarantine facilities [16] led us to screen the available next generation sequencing (NGS) data in public data depositories against the reported sequences. The SRA database search of *T. cacao* RNA-Seq data found 954 datasets, generated from 100 cacao accessions in 18 BioProject studies (Table S1). The BLASTn search for the reported partial sequences identified 33 datasets of 11 cacao accessions in BioProject PRJNA558793 [28], with a considerable number of reads similar to the query sequences. This particular BioProject was conducted on 40 cacao accessions and contains 276 cacao RNA-Seq datasets. Multiple sets of RNA-Seq data were produced for 31 accessions, representing four cacao genetic groups. One set comprised the data produced from leaf samples collected from these 31 accessions from the IC3 at CATIE. Cuttings from these trees were imported and vegetatively propagated at The Pennsylvania State University (PSU), USA, in 2020. Subsequently, mature leaves were sampled and subjected to RNA-Seq, to generate another set of transcriptome data [29]. The data from 22 runs, representing positive datasets for 11 accessions at both locations, were downloaded and analysed. Information on accession name, genetic group, geolocation, run number, data size, number of mapped reads and coverage depth is provided in Table 2. Although the number of mapped reads and coverage depth varied considerably among the cacao accessions, similar patterns were observed at both locations. Coverage depth ranged from 2.4 to 14.1-fold and from 0.8 to 13.8-fold when considering the data of 11 cacao accessions, each from CATIE and PSU, respectively. The data for the number of mapped reads varied between the locations, primarily due to differences in sequencing depth. Based on the prevalence of the mapped reads, accession IMC60 had the maximum amount of virus titre, followed by PA 16. The accessions SPEC 54/1 and AMAZ 12 had a low amount of virus.

Table 2. List of the RNA-Seq datasets in BioProject PRJNA558793 that contained a significant number of reads mapped to partial cacao polerovirus sequences. Regular and italic font styles represent accessions held at Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Costa Rica, and Pennsylvania State University (PSU), USA, respectively.

Run	Station/GPS Coordinates	Accession	Genetic Group	Size (Mb)	Reads Mapped	Coverage (X)
SRR9938230 *SRR23566402*	CATIE/9.87675, −83.65673333 *Experimental Greenhouse J, PSU*	IMC 60 *IMC60_S62*	Iquitos	1008 *544*	153 *93*	14.1 *13.8*
SRR9938245 *SRR23566388*	CATIE/9.876666667, −83.65666667 *Experimental Greenhouse J, PSU*	PA 16 *PA16_S70*	Maranon	845 *456*	135 *75*	12.0 *13.9*
SRR9938323 *SRR23566397*	CATIE/9.876466667, −83.65663333 *Experimental Greenhouse J, PSU*	NA 34 *NA34_S80*	Iquitos	1010 *548*	64 *39*	5.7 *6.4*
SRR9938250 *SRR23566375*	CATIE/9.87575, −83.65738333 *Experimental Greenhouse J, PSU*	GU 123/V *GU123V_S57*	Guiana	1168 *634*	54 *39*	5.0 *5.1*
SRR9938307 *SRR23566396*	CATIE/9.8771, −83.65805 *Experimental Greenhouse J, PSU*	NA 70 *NA70_S81*	Nanay	990 *538*	51 *37*	4.6 *5.7*
SRR9938308 *SRR23566398*	CATIE/9.877016667, −83.6576 *Experimental Greenhouse J, PSU*	NA 33 *NA33_S79*	Nanay	2443 *1290*	51 *21*	4.4 *4.2*
SRR9938223 *SRR23566373*	CATIE/9.8757, −83.65738333 *Experimental Greenhouse J, PSU*	GU 257/E *GU257E_S52*	Guiana	1068 *576*	41 *31*	3.6 *4.5*
SRR9938227 *SRR23566379*	CATIE/9.875166667, −83.65623333 *Experimental Greenhouse J, PSU*	SPEC 54/1 *SPEC54/1_S63*	Iquitos	1099 *596*	32 *28*	2.9 *3.2*
SRR9938231 *SRR23566405*	CATIE/9.876783333, −83.65706667 *Experimental Greenhouse J, PSU*	AMAZ 12 *AMAZ12_S64*	Iquitos	1139 *613*	30 *9*	2.8 *1.2*
SRR9938234 *SRR23566403*	CATIE/9.8769, −83.65676667 *Experimental Greenhouse J, PSU*	IMC 57 *IMC57_S67*	Iquitos	1016 *522*	28 *7*	2.5 *0.8*
SRR9938218 *SRR23566391*	CATIE/9.87695, −83.65716667 *Experimental Greenhouse J, PSU*	NA 916 *NA916_S77*	Nanay	955 *518*	26 *20*	2.4 *2.4*

3.2. Discovery of Novel Cacao Polerovirus Genome

The detection of a considerable number of cacao leafroll virus sequence reads in the RNA-Seq data of multiple cacao accessions triggered a search for the complete genome of the virus. The sequencing dataset SRR9938230, which showed the maximum number of virus reads, was used to construct a de novo assembly. This specific dataset, comprising 12.1 M paired-end reads, was generated from a leaf sample library of cocoa accession IMC60 from CATIE. The BLASTx searches of the de novo assembled contigs against the NCBI RefSeq protein database of the *Solemoviridae* family found a contig of 5942 nucleotides (nt), similar to members of the genus *Polerovirus*. A total of 820 reads were mapped to the identified contig, covering the entire sequence. Multiple sequence alignment of the contig with the previously reported 1439 bp cacao leafroll virus sequences (ON745771–ON745775) found >99.2% similarity (Figure S1). A BLASTn search of the nucleotide sequence of the contig against the RefSeq genome database revealed 69% identity (score: 553; E-value: 1×10^{-154}; query coverage: 28%) with the potato leaf roll virus (NC_001747), a member of the *Polerovirus* genus in the *Solemoviridae* family. We therefore regarded this contig to be a near-complete genome of a novel species of the genus *Polerovirus* and propose to name it cacao polerovirus (CaPV; GenBank accession number OR605721).

3.2.1. Genome Organization of CaPV

The NGS-derived near-complete genome of CaPV possesses the typical genome organization of previously characterized poleroviruses. It contains six ORFs, flanked by 57 and 137 nt long 5′ and 3′ untranslated regions (UTRs), respectively. The ORFs 0, −1 and −2 are separated from ORFs 3, −4 and −5 by a 223 nt non-coding intergenic region. The FSFinder2 webtool (http://wilab.inha.ac.kr/fsfinder2/; last accessed on 30 September 2023) predicted the slippery heptanucleotide (5′-GGGAAAA-3′) at nt 1845. A linearized representation of the genome organization is provided in Figure 1. The 687 nt long predicted ORF0 (nt 90–776) encodes a protein (P0) of 26.7 kDa. The P0 (IPR006755) is a suppressor of RNA-mediated gene silencing [17]. The predicted ORF1 is found to be 2058 nt in length (nt 217–2274) and encodes the P1 protein. This 76.2 kDa protein contains a peptidase S39b domain (IPR000382), which is putatively involved in the cleavage of the polyprotein. The CaPV genome encodes the RNA-dependent RNA polymerase (RdRP; IPR001795) of 125.4 kDa by −1 frameshift translational fusion of ORF1 (nt 217–1845) through ORF2 (1845–3551 nt). The 603 nt long ORF3 (nt 3778–4380) and 402 nt long ORF4 (nt 3863–4264) encode the P3 protein (CP; IPR001517) of 22.2 kDa and P4 protein (MP; IPR001964) of 15.3 kDa, respectively. The ORF4 is initiated by a CTG start codon and overlaps ORF3, but in a different reading frame. The ORF5 (4384–5802 nt) encodes the readthrough protein (PLRV readthrough domain; IPR002929), which is translated via a readthrough of the termination codon of ORF3, and hence allows the transcription of ORF5 to translate an extended coat protein.

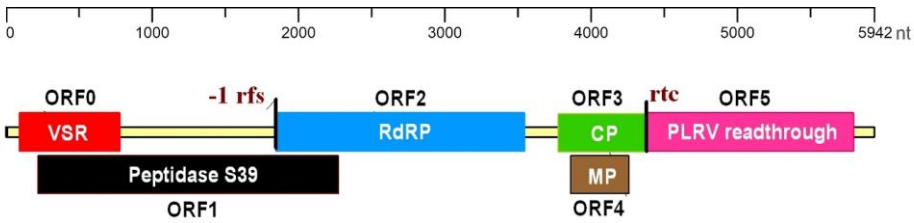

Figure 1. A schematic representation of the genome organisation of the cacao polerovirus (CaPV). Horizontal bars represent the arrangement of the predicted open reading frames (ORFs). The figure is drawn to scale. The bars are labelled with the ORF numbers and the conserved family/domain predicted in the gene products, which include P0 suppressor of silencing (VSR), peptidase S39, RNA-directed RNA polymerase (RdRP), coat protein (CP), movement protein (MP) and potato leaf roll virus readthrough protein (PLRV readthrough). Vertical lines indicate the positions of the predicted −1 ribosomal frameshift (−1 rfs) and predicted readthrough stop codon (rtc). nt denotes nucleotide.

3.2.2. Similarity Analyses of CaPV and Selected Solemovirids

The amino acid sequences of CP and RdRP proteins of 43 selected species in the *Solemoviridae* family, listed as exemplar species in the Virus Metadata Resource by the ICTV, were downloaded and compared with the respective proteins in CaPV. It includes twenty two, eighteen, two and one species of the genera *Polerovirus*, *Sobemovirus*, *Enamovirus* and *Polemovirus*, respectively.

The pairwise comparison of CP revealed 40–57% similarity of CaPV to the selected polerovirus species and showed the highest sequence similarity with cotton leafroll dwarf virus (YP_003915151) and melon aphid borne yellows (YP_001949873). The degree of identity of the selected enamoviruses and sobemoviruses to CaPV ranged from 32 to 38% and 18 to 29%, respectively. The CaPV was found to be 25% similar to *Poinsettia latent virus* (PnLV), a Polemovirus (Table S2).

The pairwise comparison, based on the amino acid sequence of the RdRP protein of CaPV, also showed high similarity with poleroviruses, with a range of 43–58%. The selected sobemoviruses and enamoviruses were found to be 28–39% and 33–34% similar to CaPV, respectively (Table S3). The sequence of RdRP protein encoded by CaPV was found to be 62% similar to that of PnLV, which is an expected result, as PnLV was found to be a chimeric virus with a closer relationship of its RdRP genes and CP genes to those of poleroviruses and sobemoviruses, respectively [30].

The genome organization and pairwise identity of CP and RdRP proteins indicate that CaPV is a member of the genus *Polerovirus*. The ICTV species demarcation criterion for the genus *Polerovirus* is based on >10% differences in the amino acid sequence identity of any gene product [31]. Pairwise identity analyses of CP and RdRP proteins indicate < 60% similarity of CaPV with any reported poleroviruses; hence, we propose that the polerovirus isolate reported in this study is a novel polerovirus.

3.2.3. Phylogenetic Analysis

Amino acid sequences of CP or RdRP proteins, of the CaPV isolate discovered in this study and the 43 species mentioned above, were used to construct phylogenetic trees, using the maximum likelihood method in the RAxML with a JTT substitution model. The tree, constructed with RdRP sequences, separated 44 sequences into three distinct clades, representing the *Polerovirus, Enamovirus*, and *Sobemovirus* genera, with >90% bootstrap support (Figure 2a). The CaPV isolate was grouped with the genus *Polerovirus*. A similar pattern was recorded in the tree constructed with CP sequences (Figure 2b). As expected in these phylogenetic comparisons, Poinsettia latent virus, from the genus *Polemovirus*, was grouped with the genera *Polerovirus* and *Sobemovirus* in CP- and RdRP-based trees, respectively.

3.3. Confirmation of NGS-Assembled Genome

The near-complete NGS genome of CaPV was verified by RT-PCR amplification. The RNA-Seq libraries in the PSU study that we analysed were constructed from leaf samples obtained in 2018 from the cacao accessions held at CATIE. Hence, we decided to use the germplasm at ICQC-R, imported from CATIE, for the detection of CaPV. We initially selected three recently imported accessions, namely ARF 2, GU 139/A and TSH 1188, together with RB 37, which was imported in 2013 but exhibited prominent yellow vein symptoms (Figure 3a). The accessions were tested with three CaPV-specific primers sets, and a cacao-specific ACP1 primer set. All four primer sets amplified discrete bands of the expected sizes, of 150, 201, 212 and 90 bp, from CaPV genome regions encoding RdRP (P2), MP (P4) and readthrough domain proteins (P5), and the cacao ACP1 gene, respectively. No amplification was found in the No-Reverse Transcriptase (-RT) control (Figure 3b).

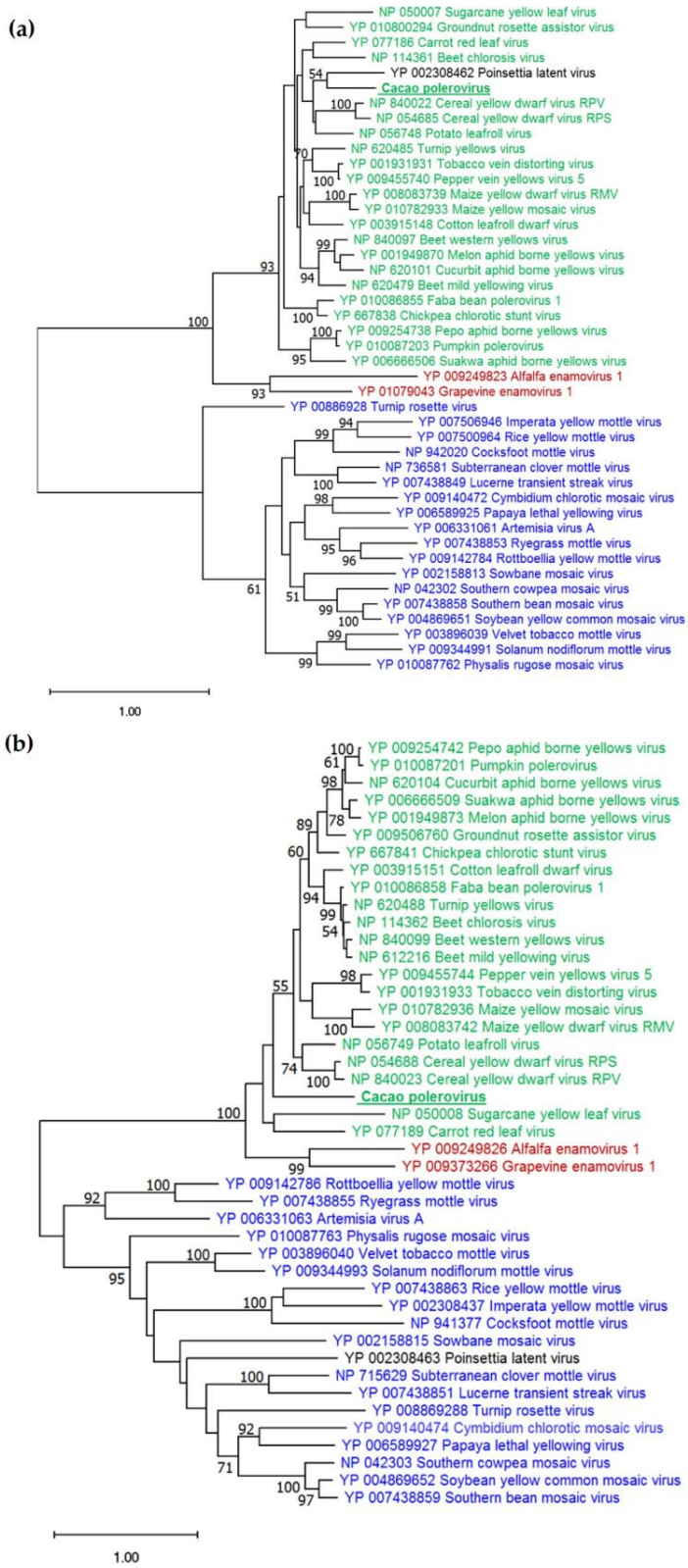

Figure 2. Phylogenetic analysis of selected virus species from four genera of the *Solemoviridae* family, based on amino acid sequences of coat protein (**a**) and RNA-directed RNA polymerase (**b**), using the maximum likelihood method. The numbers near the branches represent ultrafast bootstrap values, derived from 1000 replications (values $\geq$ 50 are shown). Green, black, red and blue font colours denote species related to the genera *Polerovirus*, *Polemovirus*, *Enamovirus* and *Sobemovirus*, respectively. The underlined label in bold font represents the cacao isolate discovered in this study.

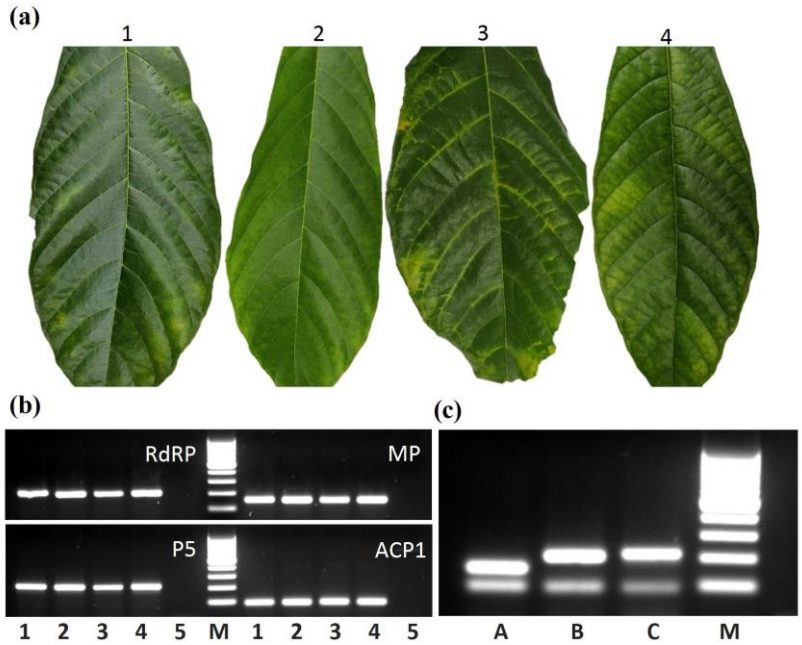

Figure 3. Detection of cacao polerovirus (CaPV). (**a**) Pictures of leaves collected from accessions ARF 2 (1), GU 139/A (2), RB 37 [BRA] (3) and TSH 1188 (4) for testing for CaPV. (**b**) RT-PCR-based detection of CaPV in cacao accessions 1–4, with primers designed from RNA-directed RNA polymerase (RdRP), movement protein (MP) and PLRV readthrough domain protein (P5). *T. cacao* acyl carrier protein 1 (ACP1) was used as internal control. Lane 5 was loaded with no-RT reaction. (**c**) Optimization of multiplex RT-PCR. Lanes A, B and C contained RT-PCR products amplified with a combination of ACP1 with RdRP, MP or P5 primers, respectively. A 100 bp increment ladder is shown in lanes M for size reference.

With regard to the amplification of the complete genome, primers designed from both ends of the NGS-derived CaPV genome failed to amplify the targeted sequence. A second set of primers designed from 5′ and 3′ UTRs, for amplification of the coding regions of CaPV, amplified a band of expected size (Figure 4a), which was cloned and sequenced by Sanger sequencing. Alignment of the Sanger contig of 5794 nt with the NGS-derived CaPV genome revealed > 99% identity with 98% coverage (nt 33–5826). No polymorphisms were found between both sequences in 5′, 3′ and intergenic UTRs. The ORF0, which encodes P0, contained six single nucleotide polymorphisms (SNPs). The P1 region encoded by ORF1 was found to be the most variable region with 26 SNPs. ORF2, which encodes RdRP, was found to be more conserved with seven SNPs. ORF3 and ORF4, which encode CP and MP, had one SNP each. Five SNPs were found in ORF5, which encodes the readthrough domain protein.

3.4. Screening of Germplasms

Singleplex diagnostic PCR assays are prone to false negative results, mainly due to the presence of PCR inhibitors in the reaction. We have developed multiplex RT-PCR assays that include a CaPV specific primer set and a cacao ACP1 primer set as ICs. The amplification of the ACP1 target would validate the negative results for CaPV. The accession RB 37 was used for the optimization of the multiplex assay. Combinations of ICs with each of the three CaPV primer sets successfully amplified the primary and secondary targets (Figure 3c).

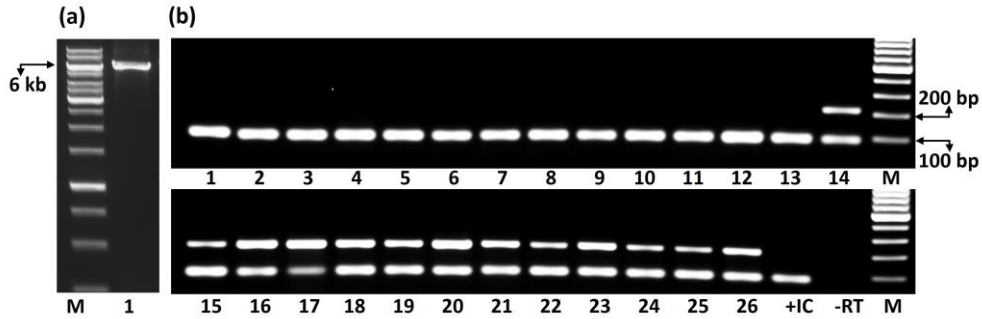

Figure 4. Agarose gel electrophoresis. (**a**) RT-PCR amplification of the complete genome of cacao Polerovirus virus (CaPV) from RB37. (**b**) Multiplex RT-PCR-based assay for testing of CaPV in the International Cocoa Quarantine Centre, Reading (ICQC-R) germplasm. A 212 bp fragment of CaPV and a 90 bp fragment of internal control [cacao ACP1 gene (LOC18599903)] were amplified from accessions RB 46 [BRA] (1), RB 49 [BRA] (2), RIM 189 [MEX] (3), EET 95 [ECU] (4), PMCT 93 (5), ARF 12 (6), PA 169 [PER] (7), CC 252 (8), CC 137 (9), EET 183 [ECU] (10), UF 712 (11), CRIOLLO 21 [CRI] (12), APA 5 (13), BE 8 (14), C SUL 3 (15), EET 387 [ECU] (16), BE 5 (17), RB 37 [BRA] (18), RB 43 [BRA] (19), RB 48 [BRA] (20), BE 2 (21), CAS 3 (22), CRIOLLO 17 [CRI] (23), GU 139/A (24), ARF 2 (25) and TSH 1188 (26). Letter M denotes the lanes loaded with 1 kb (**a**) or 100 bp size marker (**b**).

Twenty-six cacao accessions, received from CATIE between 1998 and 2021, were tested with multiplex RT-PCR. Each accession was screened with three multiplex primer combinations. Figure 4b represents the RT-PCR products amplified with P5 and ACP1 primer combination and loaded on the gel in chronological order of accession date. Gel images of screening with two other primer combinations are presented in Figure S2. Thirteen accessions, imported from 1998 to 2011, tested negative, whereas the remaining 13 accessions, imported from 2011 to 2021, tested positive for CaPV (Figure 3b; Table S4). Pictures of the leaves of 13 accessions that tested positive did not show a consistent pattern of symptoms. (Figure S3). The accessions RB 37 [BRA], RB 43 [BRA] and RB 48 [BRA] showed yellow vein banding, whereas ARF 2 and TSH 1188 had interveinal yellowing. The remaining leaves did not show visible symptoms. Yellow vein banding was also observed in the leaf of RB 46, which tested negative for CaPV (Figure S4). It is noteworthy to mention that the set of 26 accessions tested for polerovirus tested negative for the presence of CaMMV and CYVBV.

4. Discussion

This study has generated several important findings. It has utilised the SRA transcript database to identify, for the first time, the widespread presence of a polerovirus, now named CaPV, in a range of cacao germplasms, and to reconstruct a near-complete genome with all the genome features known for poleroviruses from other species [14]; previously, only a fragment of the genome had been reported in cacao [16]. The full-length sequence was then used to design primers that enabled the cloning of a genomic sequence and then the development of diagnostic primers, which were used to screen a set of germplasms in the ICQC-R collection. This analysis identified virus-positive plants from a range of diverse germplasms and confirmed that these plants all originated from material transferred from CATIE to the ICQC-R between 2011 and 2021. The CATIE germplasm was also the source of the leaves collected in 2018 by PSU and found to give the positive results in the SRA database (Table 2). It is important to note that not all the leaves from the PSU study tested positive (Figure 5), and that there is no suggestion that all cacao material at CATIE is infected. In addition, the data show infected trees from several genetic groups (Table 2), a finding that suggests there is no obvious genetic basis for the presence or absence of infection.

Figure 5. Location of the selected 31 cacao trees (BioProject PRJNA558793) held at Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Costa Rica. White and yellow font colour represent the trees tested positive and negative for Cacao polerovirus, respectively.

These results can now be used to extend the understanding of viruses known to infect cacao and, also, to inform policy relating to the international transfer of cacao germplasms. As regards the diversity of the cacao virome, previous studies have focused almost exclusively on badnaviruses, particularly on those known to cause CSSVD, but also on the mild strains of CaMMV and CYVBV. The latter viruses are known to occur in asymptomatic plants [7]. In this context, it is especially relevant now to consider the evidence for any link between the presence of the CaPV and symptomology. On one hand, the previous studies [16,17] to identify the presence of a polerovirus in cacao were conducted on plants that showed symptoms but did not find the virus in all symptomatic plants [17]. On the other hand, some asymptomatic plants tested positive in the present study (Figure 3a; Figure S3). Moreover, there is a broad range of reported symptoms related to infections caused by cacao mild badnaviruses [4–6,8], and, in some cases, these symptoms are similar to those reported for polerovirus infection [17]. Taken together, these results provide no certainty about any potential impact of CaPV on cacao production, and no information on any possible synergy between CaPV and other viruses, nor on whether the symptoms may be exacerbated by nutrient deficiency, known to cause similar symptoms in leaves [32]. These issues can only be resolved by a much more thorough large-scale analysis of plants deliberately co-infected with CaPV and mild strains of badnaviruses, and their possible interactions with abiotic stresses.

The present findings also provide some limited circumstantial evidence concerning the origin of the CaPV in the CATIE collection. As shown in Table S4, virus-positive plants were limited to those received in the ICQC-R after 2010; those received prior to that date tested negative. Possible explanations for this variation include latent infection in germplasms introduced into the CATIE gene bank in the 2000s, cross infection during

repropagation, or a sudden infection of the collection, from an unknown source, prior to 2010. Interestingly, there was no evidence of any spatial clustering in the infected clones in the field (Figure 5), a finding that might argue against the infection having entered the field from any specific direction. It would be informative to conduct a study of plants growing close to the collection, which might act as alternative host for CaPV and as a reservoir of infection. For example, it has been reported that another polerovirus, cotton leafroll dwarf virus, was found in 23 species of weeds growing adjacent to cotton fields in Georgia, USA [33]. A study of insects in the vicinity of the CATIE collection would also be valuable. It is reported that poleroviruses, such as CLRDV [34], are only, or mostly, transmitted by the cotton aphid *Aphis gossypii*, though there is no available experimental evidence of this from cacao. However, there is some evidence that biolistic methods can be used for the mechanical transmission of other species of poleroviruses [34], and it would be helpful to test the possibility of mechanical transmission of CaPV.

In summary, this study has extended the diversity of viruses known to infect cacao and provided robust diagnostic technique for CaPV. It should be noted that such discovery is an ongoing process and, as shown in other crops, it is likely that the virome will be extended, potentially to many more species. For example, five new poleroviruses were recently identified by detailed analysis of transcriptome data from six plant species [35]. At a practical level, the data provided here confirm the value of molecular diagnostic testing in complementing virus-indexing by visual observation, to ensure the safe international movement of cacao germplasms.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/pathogens12111284/s1. Figure S1: Multiple sequence alignment of the de novo assembled cacao polerovirus contig (CaPV) discovered in this study and the previously reported five partial sequences of cacao leafroll virus; Figure S2: agarose gel electrophoresis; Figure S3: pictures of sampled leaves from the accessions that tested positive for cacao polerovirus; Figure S4: pictures of sampled leaves from the accessions that tested negative for cacao polerovirus; Table S1: summary of cacao (*Theobroma cacao* L.) RNA sequencing datasets used in the study; Table S2: sequence identity of the cacao polerovirus isolate discovered in this study with 43 selected species from four genera of the *Solemoviridae* family using amino acid sequences of the coat protein; Table S3: sequence identity of the cacao polerovirus isolate discovered in this study with 43 selected species from four genera of *Solemoviridae* family using amino acid sequences of RNA-directed RNA polymerase; Table S4: summary of multiplex RT-PCR test for the detection of cacao polerovirus in selected germplasms held at International Cocoa Quarantine Centre, Reading, UK (ICQC-R).

Author Contributions: Conceptualization, I.U. and J.M.D.; methodology, I.U.; software, I.U.; validation, I.U. and J.M.D.; formal analysis, I.U. and M.K.; investigation, I.U.; resources, I.U. and J.M.D.; data curation, I.U.; writing—original draft preparation, I.U.; writing—review and editing, I.U. and J.M.D.; funding acquisition, J.M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Cocoa Research Association Ltd. (CRA) (Funding number CRA-2021), Cocoa Research UK Ltd. (Funding number 2022), and the United States Department of Agriculture (USDA) (Funding number 58-0206-1-186-F).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequencing data generated in this study were submitted to NCBI GenBank and have accession numbers OR605721 and Third Party Annotation (TPA) database accession number BK064704.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Statista. Revenue of the Confectionery Market Worldwide from 2016 to 2027, by Segment. Available online: https://www.statista.com/forecasts/1309970/worldwide-confectionery-market-size-by-category (accessed on 20 September 2023).
2. Muller, E.; Ravel, S.; Agret, C.; Abrokwah, F.; Dzahini-Obiatey, H.; Galyuon, I.; Kouakou, K.; Jeyaseelan, E.C.; Allainguillaume, J.; Wetten, A. Next Generation Sequencing Elucidates Cacao Badnavirus Diversity and Reveals the Existence of More than Ten Viral Species. *Virus Res.* **2018**, *244*, 235–251. [CrossRef]
3. ICCO. Quarterly Bulletin of Cocoa Statistics, Vol. XLIX, No.2, Cocoa Year 2022/23. Available online: https://www.icco.org/wp-content/uploads/Production_QBCS-XLIX-No.-2.pdf (accessed on 20 September 2023).
4. Chingandu, N.; Zia-Ur-Rehman, M.; Sreenivasan, T.N.; Surujdeo-Maharaj, S.; Umaharan, P.; Gutierrez, O.A.; Brown, J.K. Molecular Characterization of Previously Elusive Badnaviruses Associated with Symptomatic Cacao in the New World. *Arch. Virol.* **2017**, *162*, 1363–1371. [CrossRef] [PubMed]
5. Puig, A.S.; Ramos-Sobrinho, R.; Keith, C.; Kitchen, N.; Gutierrez, O.A.; Goenaga, R.J. First Report of Cacao Mild Mosaic Virus (CaMMV) Associated with Symptomatic Commercial Cacao (*Theobroma cacao* L.) trees in Puerto Rico. *Plant Dis.* **2020**, *104*, 3089. [CrossRef]
6. Puig, A.S. Detection of Cacao Mild Mosaic Virus (CaMMV) Using Nested PCR and Evidence of Uneven Distribution in Leaf Tissue. *Agronomy* **2021**, *11*, 1842. [CrossRef]
7. Ullah, I.; Daymond, A.J.; Hadley, P.; End, M.J.; Umaharan, P.; Dunwell, J.M. Identification of Cacao Mild Mosaic Virus (CaMMV) and Cacao Yellow Vein-Banding Virus (CYVBV) in Cocoa (*Theobroma cacao*) Germplasm. *Viruses* **2021**, *13*, 2152. [CrossRef] [PubMed]
8. Kandito, A.; Hartono, S.; Trisyono, Y.A.; Somowiyarjo, S. First Report of Cacao Mild Mosaic Virus Associated with Cacao Mosaic Disease in Indonesia. *New Dis. Rep.* **2022**, *45*, e12071. [CrossRef]
9. Muller, E.; Ullah, I.; Dunwell, J.M.; Daymond, A.J.; Richardson, M.; Allainguillaume, J.; Wetten, A. Identification and Distribution of Novel Badnaviral Sequences Integrated in the Genome of Cacao (*Theobroma cacao*). *Sci. Rep.* **2021**, *11*, 8270. [CrossRef]
10. Ullah, I.; Dunwell, J.M. Bioinformatic, Genetic and Molecular Analysis of Several Badnavirus Sequences Integrated in the Genomes of Diverse Cocoa (*Theobroma cacao* L.) Germplasm. *Saudi J. Biol. Sci.* **2023**, *30*, 103648. [CrossRef]
11. Edula, S.R.; Bag, S.; Milner, H.; Kumar, M.; Suassuna, N.D.; Chee, P.W.; Kemerait, R.C.; Hand, L.C.; Snider, J.L.; Srinivasan, R.; et al. Cotton Leafroll Dwarf Disease: An Enigmatic Viral Disease in Cotton. *Mol. Plant Pathol.* **2023**, *24*, 513–526. [CrossRef]
12. Walker, P.J.; Siddell, S.G.; Lefkowitz, E.J.; Mushegian, A.R.; Adriaenssens, E.M.; Alfenas-Zerbini, P.; Davison, A.J.; Dempsey, D.M.; Dutilh, B.E.; García, M.L.; et al. Changes to Virus Taxonomy and to the International Code of Virus Classification and Nomenclature Ratified by the International Committee on Taxonomy of Viruses (2021). *Arch. Virol.* **2021**, *166*, 2633–2648. [CrossRef]
13. Osman, T.A.; Coutts, R.H.; Buck, K.W. In Vitro Synthesis of Minus-Strand RNA by an Isolated *Cereal Yellow Dwarf Virus* RNA-Dependent RNA Polymerase Requires VPg and a Stem-Loop Structure at the 3′ End of the Virus RNA. *J. Virol.* **2006**, *80*, 10743–10751. [CrossRef] [PubMed]
14. Delfosse, V.C.; Barrios-Barón, M.P.; Distéfano, A.J. What we Know about Poleroviruses: Advances in Understanding the Functions of Polerovirus Proteins. *Plant Pathol.* **2021**, *70*, 1047–1061. [CrossRef]
15. Silva, J.M.F.; Al Rwahnih, M.; Blawid, R.; Nagata, T.; Fajardo, T.V.M. Discovery and Molecular Characterization of a Novel Enamovirus, Grapevine Enamovirus-1. *Virus Genes* **2017**, *53*, 667–671. [CrossRef] [PubMed]
16. Adegbola, R.O.; Keith, C.V.; Gutierrez, O.A.; Goenaga, R.J.; Brown, J.K. A Previously Undescribed Polerovirus (*Solemoviridae*) Infecting *Theobroma cacao* Germplasm. *Plant Dis.* **2023**, *107*, 975. [CrossRef]
17. Ramos-Sobrinho, R.; Ferro, M.M.M.; Lima, G.S.A.; Nagata, T. First Report of Cotton Leafroll Dwarf Virus Infecting Cacao (*Theobroma cacao* L.) Trees in Brazil. *Plant Dis.* **2022**, *107*, 1251. [CrossRef]
18. Csorba, T.; Kontra, L.; Burgyán, J. Viral Silencing Suppressors: Tools Forged to Fine-Tune Host-Pathogen Coexistence. *Virology* **2015**, *479–480*, 85–103. [CrossRef]
19. Sivakumaran, K.; Fowler, B.C.; Hacker, D.L. Identification of Viral Genes Required for Cell-to-Cell Movement of Southern Bean Mosaic Virus. *Virology* **1998**, *252*, 376–386. [CrossRef]
20. Prüfer, D.; Tacke, E.; Schmitz, J.; Kull, B.; Kaufmann, A.; Rohde, W. Ribosomal Frameshifting in Plants: A Novel Signal Directs the −1 Frameshift in the Synthesis of the Putative Viral Replicase of Potato Leafroll Luteovirus. *EMBO J.* **1992**, *11*, 11–17. [CrossRef]
21. Brierley, I.; Gilbert, R.J.C.; Pennell, S. Pseudoknot-Dependent Programmed -1 Ribosomal Frameshifting: Structures, Mechanisms and Models. In *Recoding: Expansion of Decoding Rules Enriches Gene Expression*; Atkins, J., Gesteland, R., Eds.; Springer: New York, NY, USA, 2010; Volume 24, pp. 149–174. [CrossRef]
22. Veidt, I.; Lot, H.; Leiser, M.; Scheidecker, D.; Guilley, H.; Richards, K.; Jonard, G. Nucleotide Sequence of Beet Western Yellows Virus RNA. *Nucleic Acids Res.* **1988**, *16*, 9917–9932. [CrossRef]
23. Smirnova, E.; Firth, A.E.; Miller, W.A.; Scheidecker, D.; Brault, V.; Reinbold, C.; Rakotondrafara, A.M.; Chung, B.Y.W.; Ziegler-Graff, V. Discovery of a Small non-AUG-initiated ORF in Poleroviruses and Luteoviruses that is Required for Long-distance Movement. *PLoS Pathog.* **2015**, *11*, e1004868. [CrossRef]
24. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* **2014**, *9*, e108277. [CrossRef] [PubMed]

25. Stamatakis, A. RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef] [PubMed]
26. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef] [PubMed]
27. Ullah, I.; Wang, Y.; Eide, D.J.; Dunwell, J.M. Evolution, and Functional Analysis of Natural Resistance-Associated Macrophage Proteins (NRAMPs) from *Theobroma cacao* and their Role in Cadmium Accumulation. *Sci. Rep.* **2018**, *8*, 14412. [CrossRef] [PubMed]
28. Hämälä, T.; Guiltinan, M.J.; Marden, J.H.; Maximova, S.N.; Depamphilis, C.W.; Tiffin, P. Gene Expression Modularity Reveals Footprints of Polygenic Adaptation in *Theobroma cacao*. *Mol. Biol. Evol.* **2020**, *37*, 110–123. [CrossRef]
29. Ullah, I.; (University of Reading, Reading, UK); Guiltinan, M.; (The Pennsylvania State University, Pennsylvania, PA, USA). Personal communication, 1 September 2023.
30. aus dem Siepen, M.; Pohl, J.O.; Koo, B.J.; Wege, C.; Jeske, H. Poinsettia Latent Virus is not a Cryptic Virus, but a Natural Polerovirus–Sobemovirus Hybrid. *Virology* **2005**, *336*, 240–250. [CrossRef]
31. Sõmera, M.; Fargette, D.; Hébrard, E.; Sarmiento, C. ICTV Report Consortium, 2021. ICTV Virus Taxonomy Profile: *Solemoviridae* 2021. *J. Gen. Virol.* **2021**, *102*, 001707. [CrossRef]
32. Muller, A. Deficiency Symptoms in Cacao Seedlings Observed in Pot Experiments with the Bouma-Janssen Method. *Neth. J. Agri. Sci.* **1979**, *27*, 211–220. [CrossRef]
33. Sedhain, N.P.; Bag, S.; Morgan, K.; Carter, R.; Triana, P.; Whitaker, J.; Kemerait, R.C.; Roberts, P.M. Natural Host Range, Incidence on Overwintering Cotton and Diversity of Cotton Leafroll Dwarf Virus in Georgia USA. *Crop Prot.* **2021**, *144*, 105604. [CrossRef]
34. Hoffmann, K.; Verbeek, M.; Romano, A.; Dullemans, A.M.; Van den Heuvel, J.F.J.M.; Van der Wilk, F. Mechanical Transmission of Poleroviruses. *J. Virol. Methods* **2001**, *91*, 197–201. [CrossRef]
35. Kavi Sidharthan, V.; Nagendran, K.; Baranwal, V.K. Exploration of Plant Transcriptomes Reveals Five Putative Novel Poleroviruses and an Enamovirus. *Virus Genes* **2022**, *58*, 244–253. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.