

# Article The Secret Life of the Inhibitor of Virus Replication

Peter Palukaitis <sup>1,\*</sup>, Masoud Akbarimotlagh <sup>2</sup>, Eseul Baek <sup>1</sup> and Ju-Yeon Yoon <sup>3,4,\*</sup>

- <sup>1</sup> Department of Horticulture Sciences, Seoul Women's University, Seoul 01797, Republic of Korea
- <sup>2</sup> Plant Pathology Department, Faculty of Agriculture, Tarbiat Modares University, Tehran 14115-111, Iran
- <sup>3</sup> Department of Plant Protection and Quarantine, Jeonbuk National University, Jeonju 54896, Republic of Korea
- <sup>4</sup> Department of Agricultural Convergence Technology, Jeonbuk National University,
  - Jeonju 54896, Republic of Korea
- Correspondence: scripath1@yahoo.co.uk (P.P.); juyeon74@gmail.com (J.-Y.Y.); Tel.: +82-63-270-4188 (J.-Y.Y.)

**Abstract:** The inhibitor of virus replication (IVR) is an inducible protein that is not virus-target-specific and can be induced by several viruses. The GenBank was interrogated for sequences closely related to the tobacco IVR. Various RNA fragments from tobacco, tomato, and potato and their genomic DNA contained IVR-like sequences. However, IVRs were part of larger proteins encoded by these genomic DNA sequences, which were identified in Arabidopsis as being related to the cyclosome protein designated anaphase-promoting complex 7 (APC7). Sequence analysis of the putative APC7s of nine plant species showed proteins of 558-561 amino acids highly conserved in sequence containing at least six protein-binding elements of 34 amino acids called tetratricopeptide repeats (TPRs), which form helix-turn-helix structures. The structures of Arabidopsis APC7 and the tobacco IVR proteins were modeled using the AlphaFold program and superimposed, showing that IVR had the same structure as the C-terminal 34% of APC7, indicating that IVR was a product of the *APC7* gene. Based on the presence of various transcription factor binding sites in the APC7 sequences upstream of the IVR coding sequences, we propose that IVR could be expressed by these *APC7* gene sequences involving the transcription factor SHE1.

**Keywords:** defense response factors; solanaceous; tobacco; Arabidopsis; inhibitor of virus replication; cyclosome; anaphase-promoting complex 7; protein structures

# 1. Introduction

The recruiting or pirating of host proteins to support the replication or movement of plant viruses is a well-characterized form of parasitism, as mutant recessive genes encoding many of these factors fail to function in enabling virus infection, thus contributing to virus resistance [1,2]. However, another form of recruitment that occurs is the ability of plants to adapt host proteins with roles in normal growth and development to serve novel roles in resistance to plant pathogens. One example of this is the use of kinase signaling pathways that serve in growth, development, or ameliorating abiotic stress in mediating resistance responses [3]. Another example is the generation of novel plant effector molecules, which can either act directly on pathogens or affect their gene expression or translation products. Better-known plant effectors that are specifically associated with defense systems include proteins such as pathogenesis-related (PR) proteins [4], particular translation factors [5], and various components of the RNA silencing system, viz., Dicer-like endonucleases, Argonaute nucleases ("slicers"), and RNA-dependent RNA polymerases (RDRs) [6,7]. Lesser-known plant effectors include those induced during infection under specific circumstances, such as ribosome-inactivating proteins or various proteinaceous inhibitors of virus infection, initially by tobacco mosaic virus (TMV) in tobacco (Nicotiana tabacum) but often expanded to other viruses (rev. in [8]).

Among those proteins induced by TMV infection that were shown to have an antiviral role, the antiviral factor (AVF) and the inhibitor of virus replication (IVR) are the



Citation: Palukaitis, P.; Akbarimotlagh, M.; Baek, E.; Yoon, J.-Y. The Secret Life of the Inhibitor of Virus Replication. *Viruses* **2022**, *14*, 2782. https://doi.org/10.3390/ v14122782

Academic Editors: Michael Taliansky, Andrew J. Love and Alex M. Murphy

Received: 15 November 2022 Accepted: 12 December 2022 Published: 14 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/). best-characterized. AVF was described by Sela and Appelbaum [9,10] as a family of phosphorylated glycoproteins stimulated by TMV infection in tobacco, one of which, gp35, is a  $\beta$ -1,3-glucanase, whereas another, gp22, is an isoform of the PR-5 protein (a thaumatinlike protein, also known as PR-S in earlier papers [4]) [11,12]; however, PR-S was shown not to have any antiviral activity when overexpressed in transgenic tobacco plants [13], whereas AVF was shown to inhibit TMV infection of tobacco if previously mixed with virus [9]. AVF was reported to function by stimulating plants to produce nucleotides with antiviral activities [14]. The regulation of the expression of the tobacco *AVF* gene has not been examined.

IVR was first described by Loebenstein and Gera [15], who infected tobacco protoplasts made from the TMV-resistant cultivar Samsun NN (SNN tobacco hereafter) with TMV. Then, they showed that culture fluid contained a substance that could inhibit TMV infection of protoplasts from either SNN tobacco or the susceptible tobacco without the N gene, cultivar Snn. The substance was not produced from either healthy SNN tobacco protoplasts or healthy or TMV-infected Snn tobacco protoplasts and was therefore specifically produced by TMV infection in protoplasts containing the N gene. The substance could be applied up to 18 h post inoculation (hpi) with TMV, showing that it inhibited replication of the virus and not uptake. In addition, IVR could be concentrated and fractionated by gel filtration into two active peaks of 26 kDa and 57 kDa, suggesting that it may form dimers [15]. Subsequent work with IVR showed that it could be recovered from protoplasts directly; could inhibit virus synthesis in leaf disks, as well as on sprayed leaves; could inhibit infection by cucumber mosaic virus (CMV) (in SNN tobacco and cucumber) and potato virus X (in N. glutinosa); and could not be induced by CMV in SNN tobacco [16]. In addition, IVR was not able to affect TMV directly and was insensitive to treatment with RNase but was sensitive to treatment with proteases and heating ( $60^{\circ}C/10$  min) and was therefore probably proteinaceous [16]. The TMV-induced synthesis of IVR in tobacco NN protoplasts was sensitive to treatment with the antimetabolites actinomycin D and chloramphenicol applied at either 5 or 24 hpi but insensitive to the same treatment applied at 48 hpi [17], indicating that IVR required RNA and protein synthesis for its production. The discovery that IVR could be collected from the intercellular fluid of non-inoculated parts of TMVinoculated leaves, where TMV was applied either in parallel trips or at the base of the leaves, and could be collected (albeit in lesser quantities) from systemically-infected leaves made the procurement of larger amounts of enriched IVR easier [18]. Antisera prepared with the two IVR fractions (isolated from protoplast culture medium) reacted identically, suggesting that the 57 kDa protein was a dimer of the 26 kDa protein [18,19]. Antisera against IVR could neutralize the inhibitory activity of IVR but did not react with either AVF or human interferon, indicating that these all were different antiviral factors [19]. Polyacrylamide gel electrophoresis fractionation of the precipitated culture medium from TMV-infected SNN tobacco protoplasts produced a 23 kDa protein that exhibited IVR activity, which also could be neutralized by antiserum prepared against this protein [20].

Reciprocal interspecific crosses made between *N. glutinosa* and *N. debneyi* produced amphidiploid hybrids that showed strong resistance to infection by TMV, producing much smaller local lesions than those produced on *N. glutinosa* (~93% reduction in size), from which little if any infectious virus could be recovered [21]. These plants also constitutively produced a PR protein designated as  $b_{1"}$  [21,22], which was produced in both parent plants only after virus infection (TMV in *N. glutinosa* and tobacco necrosis virus in *N. debneyi*). The  $b_{1"}$  protein later was considered a member of the PR-1 class of PR proteins [23]. Thus, these plants appeared to be primed for expression of proteins associated with resistance. In addition, the resistance to TMV was not abrogated by incubation at 30 °C but required 35 °C for the loss of resistance [22]. When hybrid plants of the *N. glutinosa* × *N. debneyi* cross (designated H9) were examined for IVR expression with and without TMV infection, in both cases, IVR was detected by serological assays, showing that the IVR present in H9 plants was indistinguishable from the IVR in SNN tobacco [24]. Extracts from H9 leaves were assayed in SNN tobacco leaf disks and protoplasts and were shown to be inhibitory to infection by TMV using a local lesion assay, demonstrating the presence of functional IVR in the H9 plants [24]. When the H9 plants were incubated at 35 °C, TMV was able to infect those plants systemically, as it did in SNN tobacco plants, whereas IVR could not be detected in either set of plants maintained at 35 °C [25]. Thus, IVR expression is also regulated by the temperature-sensitive expression of other defense genes in *N*-gene *Nicotiana* species [26,27].

A cDNA clone of IVR from SNN tobacco was selected by phagemid cloning and screening for the expressed IVR protein with antisera to the 23 kDa protein. The cDNA clone (designated NC330) was 1016 bp and contained an open reading frame (ORF) encoding a 199-amino-acid (aa) protein, with a mass of 21,651 Da [28]. A probe generated from the cDNA clone showed that RNA of the same size was detected in TMV-infected SNN tobacco but not in non-inoculated SNN tobacco leaves, in mock-inoculated Snn tobacco or in TMV-infected Snn tobacco. The E.coli-expressed protein interacted with the anti-23 kDa serum but migrated farther than the IVR obtained from TMV-infected Samsun NN protoplasts, with an apparent mass of ~21 kDa. Nevertheless, the expressed IVR-like protein also showed the ability to inhibit virus replication in TMV-infected Snn tobacco leaf disks [28]. Transgenic expression of the IVR cDNA clone in Snn tobacco behind a constitutive cauliflower mosaic virus 35S promoter led to some plants in each of four generations examined showing expression of the IVR-like protein and resistance to infection by TMV, although homozygous, resistant, IVR-expressing transgenic plants apparently could not be obtained [29]. Plants showing the presence of IVR mRNA were not necessarily resistant, and highly resistant plants could produce plants that segregated for resistance and susceptibility, even hypersusceptibility (producing more TMV than the controls). In addition, IVR-expressing transgenic plants became susceptible to TMV infection when grown above 30  $^\circ$ C, and IVR transgenic seeds germinated poorly or not at all in the dark or at 34 °C [29], indicating unusual early developmental effects due to constitutive expression of IVR.

Infection of SNN tobacco by TMV or potato virus Y (PVY) induced the expression of several other genes involved in defense responses, including a mitochondrial alternative oxidase, the transcriptional factor (TF) signaling hub effector 1 (SHE1, previously known as ERF5 [30]), the RNA silencing signal amplifying enzyme RDR6, and IVR [31], as well as TF MYB1 [27,32] and peroxidases [32]. The transgenic silencing of the phytohormone-induced *RDR1* gene in SNN tobacco showed inhibitory effects on the PVY-induced expression of those defense response genes. The silencing of these genes resulted in an increase in the accumulation of PVY [31,32]. SHE1 was shown to be involved in resistance to TMV, with transgenic overexpression of SHE1 enhancing resistance in SNN tobacco; however, inducible SHE1 was not activated at 32 °C [30]. Recently, we found that IVR interacted with SHE1 and that silencing of SHE1 in SNN tobacco inhibited the TMV-induced expression of IVR, whereas constitutive overexpression of SHE1 led to constitutive expression of IVR [33], suggesting that SHE1 is a TF in the pathway producing IVR. Here, we further characterize IVR and consider its origin from a plant factor involved in the regulation of mitosis.

#### 2. Materials and Methods

#### 2.1. Yeast Two-Hybrid Assays

The Gateway System (Invitrogen, Carlsbad, CA, USA) was used to introduce RT-PCR products into the expression clones via the Gateway LR reaction (Invitrogen) using pAS-attR(BD) and pACT2-attR(AD), as described previously [34]. The primers used to generate the RT-PCR products are listed in Table S1. Yeast cells were cotransformed and interactive proteins growing in yeast cells were selected as described previously [34], using the YEASTMAKER Yeast Transformation System 2 Kit and User Manual (PT1172-1, Clontech, Mountain View, CA, USA). Transformants were selected on restrictive media plates without X-gal, lacking various amino acids or adenine, to detect transcription of reporter genes for HIS, LEU, TRP, ADE, and MEL1 [34].

#### 2.2. Comparisons of DNA and Protein Sequences

The NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and Dbfetch (http:// www.ebi.ac.uk/Tools/dbfetch/dbfetch/) systems was used to obtain sequences of *IVR*-like genes and anaphase-promoting complex 7 (*APC7*)-like genes, as well as sequences of their encoded proteins, and perform alignments of sequences. Before analysis, the tobacco IVR sequence was edited to correct errors causing a frame shift between nucleotides 491 and 558, corresponding to the C-terminal region amino acids 164–175 and 186 [33].

# 2.3. Protein Structure Modeling by AlphaFold

IVR and APC7 protein were modeled using the AlphaFold program [35] with a userfriendly interface for accessing AlphaFold2 provided online via Github notebooks (https://colab. research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb# scrollTo=XUo6foMQxwS2, accessed on 8 May 2022), the structural prediction of which is supported by AlphaFold2 combined with a fast multiple sequence alignment generation stage using MMseq2 [36,37]. The modeling of homo- and heterocomplexes also was achieved using ColabFold.

The atomic coordinates and per-residue confidence estimates for predicted structures scale from 0 to 100. Higher scores are related to higher confidence. This confidence criterion is termed pLDDT (predicted local distance difference test) [38]. The pLDDT is defined in four levels: the first level includes high model confidence of residues with pLDDT  $\geq$  90; the second level shows the confidence model representing residues with 90 > pLDDT  $\geq$  70; the third level, representing residues with 70 > pLDDT  $\geq$  50, has low confidence; and the final level residues with pLDDT < 50 correspond to very low confidence [39]. The pLDDT is represented by a scale from red (bad) to blue (good) in AlphFold2 plots.

UniProt Blasts (https://www.uniprot.org/uniprotkb/, accessed on 28 July 2022) were performed with the IVR protein sequence; some of these that had 100% similarity with UniProt were used for AlphaFold predictions, which were performed with the APC7 from *Nicotiana sylvestris* (LOC104246274-anaphase-promoting complex subunit 7 isoform X2-Nicotiana sylvestris (Wood tobacco) | UniProtKB | UniProt, https://www.uniprot.org/uniprotkb/A0A1U7YEH5/entry, accessed on 28 July 2022).

#### 2.4. Identifying TPR Elements

Tetratricopeptide repeat (TPR) elements present in IVR-like proteins and TPC7 proteins were identified from information provided in BLAST searches, as well as from data provided by the AlphaFold program [35] and as described in [40–42].

#### 3. Results

# 3.1. IVR1 and MYB1 Interact with the CMV 1a Protein

Previously, we showed that the CMV 1a protein could interact with the TF SHE1 in several systems, including the yeast two-hybrid system (Y2H) [34]. Here, we used the Y2H system to determine whether other early components of the *N*-gene-mediated defense response also interacted with the CMV 1a protein (Figure 1). Interactions between the CMV 1a replication-associated protein plus the CMV 2a replicase protein were used as positive controls, and the absence of an interaction of CMV 1a with the TMV helicase domain [43], which binds to the N protein, was used as a negative control for interaction with CMV 1a. CMV 1a also interacts with itself to form dimers [44]. Of the examined tobacco proteins, the N protein and the two chaperones (RAR1 and HSP90) that interact with each other and with the N protein via HSP90 [45] did not react with the CMV 1a protein, whereas IVR and the TF MYB1 did interact with the CMV 1a protein (Figure 1).

BD	AD	-Trp	-Leu	-Leu/-Trp	-Leu/-Trp/-His	-Leu/-Trp/-His/-Ade
pAS-1a	-	+a	-	_	-	-
pAS-1a	pACT(AD)	+	+	+	-	-
pAS-1a	pACT-1a	+	+	+	+	+
pAS-1a	pACT-2a	+	+	+	+	+
pAS-1a	pACT-N	+	+	+	_	
pAS-1a	pACT-HSP90	+	+	+	-	-
pAS-1a	pACT-RAR1	+	+	+	-	-
pAS-1a	pACT-MYB1	+	+	+	+	+
pAS-1a	pACT-IVR	+	+	+	+	+
pAS-1a	pACT-TMV Hel	+	+	+	_	_

#### Summary of the interactions between CMV 1a and other proteins by Y2H

a + and -: positive and negative reaction.

**Figure 1.** Yeast two-hybrid assay for interactions between the cucumber mosaic virus (CMV) 1a protein and various host defense factor proteins. The CMV 1a protein bound to the binding domain (BD) in the plasmid pAS was propagated in yeast cells, either alone or together with plasmids (pACT), expressing the activation domain (AD) fused to the CMV 1a protein, the CMV 2a protein (both positive controls), the TMV helicase domain of the 126K protein (negative control), the resistance gene protein N from tobacco, the defense chaperone proteins HSP90 and RAR1, the transcription factor MYB1, and the defense factor IVR. The metabolites listed at the top represent amino acids (Trp, Leu, and His) or the base adenine (Ade), which were missing for the plating media in which the yeast cells were grown. The pAS plasmid constitutively expresses Trp, and the pACT plasmid constitutively expresses Leu. Cells containing interacting test proteins induce the expression of both His and Ade. The growth of yeast cells on media-specific plates, indicating an interaction of the test proteins, is shown by "+", whereas the absence of growth is indicated by "-".

MYB1 is a TF that is activated by salicylic acid (SA) and TMV infection of *N*-gene tobacco [27] and is required for the *N*-gene-mediated resistance response to infection against TMV [46]. MYB1 binds to the PR1a promoter region but is a minor factor in the expression of the tobacco *PR-1a* gene, where TFs WRKY12 and TGA1a are the major regulators of PR-1a expression [47]. The consequences of the interactions of the CMV 1a protein with IVR and MYB1 is unknown but could be either part of the defense response or a counter-defense action by the CMV 1a protein. Here, we focused on the nature of the *IVR* gene and its encoded protein.

# 3.2. Sequence Analysis of IVR-like Proteins from Solanaceous Plants

At the time the *IVR* gene was sequenced [28] or expressed in transgenic tobacco SNN plants [29], none of the solanaceous genomes had been completely sequenced. Therefore, there were limited data available (1999–2005) for comparison of IVR sequences. An expressed sequence tag (EST) of 479 bp (GenBank Accession AW932904; 21 May 2001) obtained from tomato (*Solanum lycopersicum* cv. TA496) fruit as part of an unpublished study entitled "Generation of ESTs from tomato fruit tissue" was very similar (in antisense orientation) to part of the tobacco *IVR* gene from ORF nucleotide 454 to the termination codor; the 3' nontranslated regions (NTRs) showed more differences, including insertions (Figure 2). Subsequently, a reference library of tomato cv. Micro-Tom [48] produced a tomato fruit cDNA with sequences very similar the tobacco *IVR*, although the sequence of the ORF was more than twice as long (GenBank Accession AK328373; 3 May 2010). The 3' NTR sequences were identical to those of the shorter tomato EST shown in GenBank Accession AW932904 (Figure 2), but after the end of the shorter tomato EST, the longer tomato cDNA differed considerably in sequence from the tobacco *IVR* sequence to such an extent that it was no longer detectable by alignment tools (Figure 2).

Nt Sl

Nt S1

NE

S1 S1



GATGTCGAGGATGCTGATGCAGACCAGGAAGAGGCTGAGCTT 1154



**Figure 2.** Nucleotide sequences of the *IVR*-like genes and 3' nontranslated regions (3' NTRs) of various cDNAs, expressed sequence tags, or mRNAs deduced from genomic DNA clones from the plant species tobacco (Nt) and tomato (Sl), as well as the two parent species of Nt, Ns, and Nto. Numbers to the right and at the end of the sequences refer to the positions of the sequence in Nt beginning at the start of the IVR open reading frame and, in the other species, the locations of the beginning of either cloned sequence or the beginning of the *IVR*-like gene within that cloned sequence. Differences between the Nt sequence and the other sequence are indicated by turquoise highlight. Differences between Ns and Nto beyond the 3' NTR of the Nt sequence are indicated by yellow and green colors. The starting AUG and the terminating UGA are highlighted, except in Nto (where it is underlined), because the Nto mRNA sequence contains an insert, leading to a frameshift. "poly A" refers to the polyadenylated tail that follows the end of the 3' NTR. Dashes were added to maintain the alignment when some sequences contained insertions relative to others.

Primers for PCR amplification of the tobacco IVR were used for RT-PCR on RNAs isolated from potato (Solanum tuberosum) cv. Phureja, which was cloned. The sequence of the ORF in the cloned PCR product was only 4 nt different from a 600 bp sequence within the genomic scf00035\_44 of the potato cv. DM 1-3 516 R44 (AEWC01007128.1; 24 May 2011) and was very similar to the sequences of both the tobacco IVR (AJ009684) and the tomato Micro-Tom sequence (AK328373) (data not shown), as was the sequence of the encoded proteins (see below). N. benthamiana (Nb) also has sequences very similar the tobacco IVR gene (GIUP01022078.1; 4 October 2021). IVR-like sequences also were available from one of the parents of tobacco, N. sylvestris (Ns) (two identical, partial EST clones: BP744385 and BP745492; both 28 May 2004), which contained sequences of the IVR ORF from nucleotide 190 to the termination codon plus 167 sequences of the 3' NTR. In this case, the limited 3' NTR sequences between the IVR-like gene of Ns and N. tabacum (Nt) were nearly identical, containing one substitution, as well as three nucleotide deletions and five nucleotide insertions in Ns (data not shown). Because the Ns sequences were incomplete, the same analysis was performed again using mRNA sequences derived from the genomic DNA sequences of the NsIVR-like gene (XM\_009802059.1; 21 October 2014) (Figure 2). In this case, there was only one deletion and one substitution in the 3' NTR

of *Ns*IVR-like RNA vs. the 3' NTR of the *Nt*IVR RNA. In contrast, the sequences of the 3' NTR of *Nt*IVR compared with the IVR-like sequences from the other tobacco parent, *N. tomentosiformis* (*Nto*) (XM\_018778567; 20 April 2020), which also was derived from a genomic DNA clone, showed 22 substitutions, 12 insertions, and 60 deletions (52 in 4 blocks of 9-22 nt) in the *Nto*IVR-like 3' NTR (Figure 2). Within the ORF of the IVRs, *Ns* showed no sequence differences from *Nt*, whereas *Nto* showed 15 substitutions and 6 inserts (Figure 2). Overall, the data indicate that the *IVR* gene of *Nt* tobacco is derived from the *Ns* parent.

The genomic sequences of *IVR*-like genes from *Nt*, tomato, potato, and *Nb* were compared and found to share many common features (Figure 3). The ORFs of the potato and tomato IVR-like genes (Figure 3A,B) were split into five exons in which the sizes of corresponding exons between species were almost identical. This was also the case for the four introns, among which the second intron was the largest. The *NtIVR* gene contained a similar arrangement of introns and exons, but the second intron was much smaller (1495 bp for tobacco vs. 2981 bp for potato and 2886 bp for tomato (Figure 3C vs. Figure 3A,B). In contrast, the *IVR*-like genes varied in size, with *Nt* having the shortest length (2837 bp), whereas potato (4210 bp) and tomato (4104 bp) were longer. In the case of Nb, the numbers and lengths of the exons were similar to those in the *NtIVR* gene (Figure 3C vs. Figure 3D), but the entire *Nb* gene was much longer and contained duplications of exons (Figure S1). Specifically, exons 1–3 were duplicated, separated by introns some 4000 bp further along the same BAC fragment after the end of the complete copy of the *IVR*-like gene (Figure S1B), although the second intron of the partially duplicated *IVR*-like gene was much shorter (703 bp). In addition, in the complete copy of the IVR-like gene, the second intron (2350 bp) contained an inverted copy of the last exon (Figure S1A), whereas upstream of the complete copy, there were other exons (or a fragment) inserted in inverted orientations (Figure S1C). It is not known whether the *IVR* gene is expressed in *Nb* during defense responses.



**Figure 3.** Genome organization of *IVR*-like genes in four solanaceous plants: (**A**) potato, (**B**) tomato, (**C**) tobacco, and (**D**) *N. benthamiana*. Based on genomic sequences that contain *IVR*-like genes and mRNA-derived sequences that contain IVR-like sequences, the positions of exons and introns could be determined. In many cases, additional copies were found, some containing minor base changes and others containing deletions or insertions, suggesting that dysfunctional copies are present in these genomes. This was particularly the case for *N. benthamiana* (see Figure S1). Numbers above the "chromosome line" represent the transcripts coordinates, and numbers below the lines or introns represent the genome sequence coordinates from specific genomic scaffolds.

# 3.3. Sequence Analysis of IVR-like Proteins from Various Plant Families

An *Arabidopsis thaliana* protein-coding sequence (At2g39090; 2 February 2001) encoding a 276 aa partial protein was identified, showing high sequence similarity to the 199 aa tobacco IVR sequence. A later sequence analysis identified this Arabidopsis coding sequence in a 558 aa protein (NP\_850309; 16 September 2003) (Figure S2) containing TPRs, which are involved in protein–protein interactions [44], and showing sequence similarity to both human and mouse anaphase-promoting complex subunit 7 (APC7). A comparison of the corrected tobacco IVR protein (from cv. Samsun NN) sequence against the *Nt*APC7 protein sequence (XP\_016480892.1; 3 May 2016) (from cv. TN90) showed that in their overlap regions, they differed by only one amino acid (APC7 amino acid 444 is alanine, whereas the corresponding *Nt*IVR amino acid (82) is serine), indicating that IVR is derived from APC7. A comparison of *Nt*APC7 against those of the APC7 of the parental species from which *Nt* is derived (*Ns* and *Nto*) showed that there were minor differences in sequence between all three proteins (Figure 4), with few differences in sequence between *Nt*APC7 (positions 150, 210, 331, and 444) and between *Nt*APC7 and *Nto*APC7 (positions 278, 339, and 342/343).



**Figure 4.** Amino acid sequence comparisons of anaphase-promoting complex (APC7) proteins from six solanaceous species, Arabidopsis, and two monocot grass species. The species examined were pepper (Ca), tomato (Sl), potato (St), *N. sylvestris* (Ns), *N. tabacum* (Nt), *N. tomentosiformis* (Nto), Arabidopsis (At), rice (Os), and wheat (Ta). Sequences were aligned with triangles inserted to maintain the alignments, where one or more amino acids were absent relative to the same position in other species. Similar sequences were grouped together; the similarity was based on type of amino acid: basic (H, K, and R), acidic (D and E), amide (N and Q), hydroxylated aliphatic (S and T), small aliphatic (A, L, I, M, and V), or phenyl group (F and Y). Not within groups are C, G, P, and W, as they all affect the folding of the protein. Highlight key: yellow—all sequences within a column were similar; light green—all but one of the amino acids in a column were similar; turquoise—the majority

or a plurality of the amino acids in a column were similar; light or dark grey—two to four sequences in a column were similar; red—similar sequences found only in the *Nicotiana* species; and white—a novel amino acid at this position. Note: the sequence of the *Nto*APC7 appears to be lacking either the N-terminal 216 aa with no start codon (XP\_009629102.1, derived from mRNA XM\_009630807.1; 15 October 2014) or the N-terminal 254 aa beginning with the amino acid sequence "MEKVEA" (XP\_018634083.1, derived from mRNA XM\_018778567.2; 20 April 2020).

A further comparison of the *Nicotiana* spp. APC7 sequences with other solanaceous species, namely tomato, potato, and pepper (*Capsicum annuum*), as well as with *At*APC7 and the sequences of APC7 from two monocotyledonous species, rice (*Oryza sativa*) and wheat (*Triticum aestivum*) (Figure 4), showed a remarkable degree of conservation of sequence. A total of 67% of the amino acid sequences at any given position were similar (yellow highlight) among the nine plant species, with 12% of the sequences at any given position among the nine plant species containing only one non-conserved amino acid (light green) and 20% of the amino acids located at specific positions among the nine plant species similar either in a majority or plurality of occasions (turquoise). The methionine corresponding to the N terminus of the IVR is located at amino acid position 363 of most of the sequences; however, neither Arabidopsis nor the two monocot species or only *Ns* and *Nt* contained novel amino acid selections at 12 positions (red), whereas the amino acid sequences at 59 positions were unique to the two monocot species (Figure 4).

# 3.4. Analysis of the TPR Units in Tobacco IVRs vs. AtAPC7

TPR units contain ~34 aa and form a helix–turn–helix structure [41]. The structure of AtAPC7 was predicted to contain 10 TPR units. Of these, the NsAPC7 was predicted to contain six TPR units (Table 1 and Figure 5; XP\_009800361.1; 21 October 2014). However, the human APC7 was shown to contain noncanonical TPR units, in which some of the conserved sequences were not always present at the designated locations in the TPR units [42]; in particular, in the TPR motif, x<sub>3</sub>Wx<sub>2</sub>LGx<sub>2</sub>Yx<sub>8</sub>Ax<sub>3</sub>Fx<sub>2</sub>Ax<sub>4</sub>P, W-4 was always absent, and L-7 and G-8 were usually absent. It is difficult to rationalize why some of the TPR units are thus identified, given that many did not have most of the canonical motif sequences present. This applies to TPR units identified in both tobacco and Arabidopsis, as well as such sequences identified as TPR units in Arabidopsis but not identified as such in the tobacco species. Specifically, the first AtAPC7 TPR unit (Table 1) had only one conserved amino acid of this motif, which was also the case for the similar Nicotiana APC7 sequences (Figures 4 and 5). In the case of the second AtAPC7 TPR unit (Table 1), the corresponding sequences of the three Nicotiana APC7 proteins contained five of the eight canonical TPR motif sequences, whereas the AtAPC7 sequence contained only three of the eight canonical TPR motif sequences (Figures 4 and 5). In the third AtAPC7 TPR unit, equivalent to the first NsAPC7 unit, the AtPC7 corresponding sequences only contained three of the eight canonical TPR motif sequences, whereas the three Nicotiana APC7 proteins only contained two of the eight canonical TPR motif sequences (Figures 4 and 5). The fourth AtAPC7TPR unit and the equivalent second *Nicotiana* APC7 units, as well as the fifth *At*APC7 TPR unit, all contained only three of the eight canonical TPR motif sequences. The third Nicotiana APC7 TPR units, which overlapped with part of the fifth and all of the sixth AtAPC7 TPR units, contained six of the eight canonical TPR motif sequences, whereas the sixth AtAPC7 TPR unit only contained four of the eight canonical TPR motif sequences, including the terminal proline. This is also the motif in which the Nicotiana IVR sequences began (Figures 4 and 5). The seventh AtAPC7 TPR unit and the equivalent fourth Nicotiana APC7 TPR units both contained six of the eight canonical TPR motif sequences. The eighth AtAPC7 TPR unit and the equivalent fifth *Nicotiana* APC7 TPR unit both contained only three of the eight canonical TPR motif sequences (Figures 4 and 5). The ninth AtAPC7TPR unit with no Nicotiana equivalent contained no prolines in this region or adjacent to this region and only had one or two of the other TPR sequences aligned (Figures 4 and 5). Finally, the tenth AtAPC7 TPR unit and the equivalent sixth *Nicotiana* APC7 TPR unit had five and six of the eight canonical TPR motif sequences, respectively (Figures 4 and 5).

Table 1.	Comparison o	f the predicted	tetratricopeptide	e repeats (TI	PR) in AtAP	C7 and NsAPC7.
----------	--------------	-----------------	-------------------	---------------	-------------	----------------

A	tAPC7	NsAPC7			
TPR No.	Borders (aa No.)	TPR No.	Borders (aa No.)		
1	43–76				
2	138–171				
3	212–245	1	219–243		
4	246–279	2	248–278		
5	314–346				
6	348–380	3	339–379		
7	381–413	4	383–411		
8	414-448	5	417–446		
9	450-482				
10	483–515	6	485–514		

1 MDVPKDQMST LLDHGLFPSA QMLGCFLVSS SSVSIETSPH LKAENLVLLG DALFREKEYK

61	RAIHAYKQAL	HCHRIIPKQN	ATPIRSALSA	STRSSSPNSF	NISAINENEV	KFKIASSHCA		
	[Hypothetical TPR]							
121	LNENRAALLE	MEGIPSKARN	LOMNLMMGKL	YRSSRHTRAA	VTCYKECLRH	CPYI IEAIIA		
	TPR no. 1							
181	LAELGVAAKD	IISLFPQTPT	RGARSTFDHF	DSSRWLQR <mark>YV</mark>	EAQCCIASND	YKGGLEIFAE		
	TPR no. 2							
241	LLRRFPCNIH	ILLEMAKVEA	IIGKTEEAIT	DFEKVRSTDP	YVITYMDEYA	MLLKQKSDYS		
	TPR no 3							
301	KINKIVHDII	NIDPTRPEVF	VALSVLWERK	dergalsy <mark>ae</mark>	KSIRIDERHI	PGYIMKGNLF		
	TPR no. 4							
361	FS <mark>M</mark> NRPEAAV	VAFRGAQELR	PDLRSYQGLV	RSYLAISKIK	EALHAAREAM	KAMPQSAKAL		
	TPR no. 5							
421	KLVGDVHASN	TSGREKAKKF	YESSLRLEPG	YLGAALALAE	LHVIEGRTVD	AVTLLERYLK		
TPR no. 6								
481	DWAD <mark>DSLHVK</mark>	LAQVFAATNM	PQDALSHYQA	ALRINPQNEA	AKKGLERLEK	QLKGVDPDAP		

541 EEDEENDVED ADADQEEAEL L -561

**Figure 5.** Mapping of tetratricopeptide repeat (TPR) units and dimerization domain of anaphasepromoting complex (APC7) on the amino acid sequences of ACP7 from *Nicotiana sylvestris* (*Ns*). Highlight key: light green—dimerization domain predicted from human APC7 [42]; dark blue hypothetical TPR before TRP no. 1; turquoise—predicted *Ns*TPR unit no. 1 (amino acids 219–243), no. 2 (amino acids 248–278), no. 3, (amino acids 339–379), no. 4 (amino acids 383–411), no. 5 (amino acids 417–446), and no. 6 (amino acids 485–514); and yellow—methionine start codon for IVR. The double-underlined sequences represent a double TPR unit designated TPR11.

# 3.5. Structural Analysis of the NtIVR

The AlphaFold program [35] was used to model the three-dimensional (3D) structure of IVR. The information explaining the levels of confidence in this structure are given in Section 2.3 and shown in Figure 6B,C. This structure contains 10 helical regions (Figure 6A) from amino acid 5 to amino acid 171 (bordered by prolines and at the edges of TPR units), including two  $\alpha$  helices (nos. VI and VII) between TPR units 3 and 4 and another  $\alpha$  helix (no. X) after TPR unit 4, followed by a C-terminal 28-amino-acid random structure (Figure 6A,B). These helical regions include two each in three TPR units (nos. 2, 3, and 4 in Figure 6D, equivalent to TPR nos. 4, 5, and 6, respectively in Figure 5) and one in the part of the split APC7 TPR no. 3 (Figure 5) located in the N-terminal 17 as of IVR (Figure 6D).



**Figure 6.** Three-dimensional structure and protein sequence of IVR. (**A**) (Left) 3D structure for IVR modeled by AlphaFold, showing 10 helices (from the N terminus, identified by Roman numerals) followed by a disordered random structure for the C-terminal 28 aa. The C terminus is at the far right of the molecule. (Right) Table showing the locations of the 10 helices in the IVR sequence, as well as their positions with regard to the IVR TPRs. (**B**) The predicted local distance difference test (pLDDT) and (**C**) the predicted aligned error (PAE) scores are shown. The pLDDT is a per-residue confidence metric (scale of 1–100). Based on the pLDDT and PAE indices, the predicted 3D structure derived from the protein sequence of IVR has high confidence scores for the domains at residues 1-174 (green in C) but not with the region between 175 and 199 (the C terminus; white in C). The predicted IVR structure in which IVR amino acids 1 and 2 (M and R), 19(P), 33 (I) to 34 (S), 52 (P) to 53 (Q), 67 (S) to 70 (S), 87 (P) to 88 (G), 119 (D) to 121 (A), 154 (Pro) to 155 (Q), and 172 (G) to 199 (L) are confident (pLDDT 70-90), but other amino acids show very high pLDDT indices (pLDDT > 90). (**D**) Amino acid sequence of IVR, with TPR units highlighted in turquoise. The methionine initiation codon is at position 1. Sequences within underlined C-terminal 99 aa were found to contain a binding domain to the SHE1 transcription factor [33].

# 3.6. Structural Analysis of the NtIVR vs. AtAPC7

The 3D structure of *At*APC7 was modeled (AlphaFold DB Q8VY89), as was the structure of the *Ns*APC7 (LOC104246274-anaphase-promoting complex subunit 7 isoform X2-Nicotiana sylvestris (Wood tobacco) | UniProtKB | UniProt, https://www.uniprot.org/uniprotkb/A0A1U7YEH5/entry, accessed on 28 July 2022). The structure of the 199 aa *Nt*IVR (blue) of similar sequence was modeled together with the C-terminal 201 aa of

*At*APC7, i.e., amino acids 361-561 (red) (Figure 7). In four different views of the superimposed structures, the two molecules were identical, except for the non-structured C-terminal 28 aa (Figure 7). Nine of the ten helices were readily discernable, but due to overlapping by the helical bundles, a tenth helix is difficult to discern in various views. The superimposition showed that there were no differences in the number, size, or position of helices, indicating that the 85-122 and 152-171 regions of IVR amino acids must also form helices. This supports the conclusion that IVR is derived from APC7.



**Figure 7.** Superimposed 3D structures of NtIVR with the C-terminal 201 aa of Arabidopsis anaphasepromoting complex (*At*APC7). The structure of the 199 aa tobacco IVR sequence (blue) was modeled together with the structure of *At*APC7 amino acids 361-561 (red). Four views for each superimposition of the two structures are shown, displaying the alignment of various helical units, as well as positions of the C-terminal 28 aa random strands.

When the complete *At*APC7 sequence was superimposed on the *Nt*IVR sequence in the model (Figure 8), it was clear that the *At*IVR molecule largely formed two discrete domains. The N-terminal 360 aa formed a cluster of helical bundles generated from various TPR units plus other helices and large looped regions between several of the helical bundles, as described for human APC7 [42]; beyond this region of the molecule were helical bundles in the superimposed regions, corresponding to the alignment of structures between *At*APC7 and *Nt*IVR, as shown in Figure 7.



**Figure 8.** Superimposed 3D structures of NtIVR with that of the complete Arabidopsis anaphasepromoting complex (AtAPC7). The structure of the 199 aa tobacco IVR sequence (blue) was modeled together with the structure of the 561 aa AtAPC7 (red). Four views for each superimposition of the two structures are shown, displaying the alignment of various helical units (as shown in Figure 7) and the separate N-terminal 360 aa domain of AtAPC7 (red), as well as the positions of the C-terminal 28 aa random strands.

# 4. Discussion

The data presented here showed that IVR, a long-studied antiviral factor from tobacco [49], was able to interact with the CMV 1a protein, just as IVR interacted with the TF SHE1 [33] and the SHE1 and CMV 1a proteins interacted with each other [34]. To gain a better understanding of the nature of the IVR, we examined the sequences of IVR-like genes in a number of plant species and identified a larger protein, APC7, that contained the IVR-like sequences in the C-terminal 36% of the APC7 protein (Figures 4 and 5). In addition, the protein sequences encoded by the APC7 gene are highly conserved in many species ([42]; Figure 4). APC7 contains numerous TPR units, which are involved in interactions with other proteins, usually with the C-terminal region of those proteins, in one or more of several motifs, such as the isoleucine–arginine (IR) motif [42]. APC7 is a component of the cellular cyclosome (aka APC), a E3 ubiquitin ligase controlling the progression of mitotic division [46]. APC is composed of at least 13 subunits, including a cullin homolog (APC2), a ring-H2 finger domain (APC11), a Doc domain protein (APC10) with an IR motif that promotes substrate binding, and four subunits containing TPR elements (APC 3, 6, 7, and 8). APC also requires two adaptor proteins (Cdc20 and Cdh1, both containing WD40 domains). The adaptors and various subunits permit recognition of substrates [50]. APC4 and APC5 are required for the TPR subunits to bind to APC1 [46,47]. The C-terminal halves of APC7 and APC3 contain TPR units needed for the binding of the adaptor Cdc20 and the substrate Nek2A, which interact with the TPR units via their IR peptides [51]. In addition, CP7 subunits form dimers through the N-terminal 38 aa dimerization interface [42].

Modeling of the *Nt*IVR protein structure, both by itself (Figure 6) and superimposed on either the corresponding region of APC7 (Figure 7) or the complete APC7 molecule (Figure 8), showed that except for the C-terminal 28 aa in both proteins, the two molecules were of identical structure in their overlapping regions. IVR lacks the dimerization domain present near the N terminus of APC7 and, except in Arabidopsis, does not contain any cysteine residues in the IVR sequence (Figure 4). Thus, it seems likely that IVR could still form complexes with some of the other components of the APC, as well as with some adaptor and substrate proteins. Whether this would be sufficient to activate the ubiquitin ligase components of the APC is not known. If it could, IVR might have some role in proteasome targeting, albeit with a different specificity than for the APC. Such an altered specificity might be suitable for targeting specific viral encoded proteins.

IVR was considered to form a dimer based on two peaks of antiviral activity detected by gel filtration chromatography, with 74% inhibition activity associated with the monomer unit (ca. 26-27 kDa) and 37% with the dimer unit (56-57 kDa) [15]. This suggests that either the dimer unit was less active than the monomer unit or that most of the protein was in monomer form. We could not detect self-interaction of IVR in the Y2H system (data not shown). Thus, we suggest that the original dimer-sized activity detected by Loebenstein and Gera [15] may have been caused by the presence of the full-length 62.8 kDa APC7 rather than a dimer of the 21.6 kDa IVR monomer.

Because IVR has the same sequence and structure as the C-terminal 201 aa of APC7 (Figures 4–8), it seems likely that IVR was generated via the *APC7* gene. The question remains as to how IVR is formed. Is the APC7 protein processed to form IVR by some specific protease cleavage, or is IVR mRNA transcribed from transcription start sites present upstream of the IVR coding sequence in the *APC7* gene? The third option—that IVR is the result of alternative splicing of the APC7 transcript—is unlikely, as there are no additional sequences at the 5' NTR of the IVR mRNA that are derived from sequences further upstream in the *APC7* gene. Although we cannot rule out protein processing of APC7 to produce IVR, we note that IVR mRNA of ca. 1000 nt was detected by northern blot hybridization of gel-fractionated RNAs extracted from TMV-induced SNN tobacco but not RNAs of uninduced SNN tobacco [28]. This suggests that the most likely origin of IVR mRNA is by transcription from transcription starts sites upstream of the IVR coding sequence, at least in solanaceous plants.

Previously, we showed that IVR and the TF SHE1 were both induced during TMV infection in NahG tobacco plants, expressing the enzyme salicylate dehydroxylase [33], which inhibits the accumulation of SA [52]. In addition, silencing or overexpressing the SHE1 gene caused a parallel change in the expression of IVR [33]. This led us to suggest that SHE1 was a TF involved in expression of IVR, either by binding to a promoter region upstream of the *IVR* gene or further upstream. SHE1 was found to bind weakly to the GCC sequence, an element of the ethylene-responsive element-binding protein (EREBP) site [30]. However, unlike other ethylene-responsive TFs, SHE1 (formerly *Nt*ERF5) was not induced by ethylene, SA, or jasmonic acid [30]. Therefore, we examined the sequences upstream of the IVR coding sequence in the NtAPC7 gene for possible EREBP sites. Four such sites were identified: 28–30 bp, 306–308 bp, 558–560 bp, and 857–859 bp upstream of the initiation codon of IVR (Figure 9). We also examined this region of the NtAPC7 gene of ca. 1500 bp for other TF binding sites. We found one binding site for the <u>YACGT</u>GG/TClike ABA-responsive element-binding protein site (ABRE) 576-580 bp upstream of the IVR initiation codon and five binding sites for WRKY TFs 284-289 bp, 801-806 bp, 826-831 bp, 1482–1487 bp, and 1496–1501 bp upstream of the IVR initiation codon (Figure 9). As we stated at the end of Section 3.1, it is not unusual for two or more TFs to work together (synergistically or additively [47]). Therefore, it is possible that more than one TF may be involved in the expression of IVR from the upstream sequences within the APC7 gene. Future experiments will examine whether SHE1 can bind to sequences within this region of the *APC7* gene and whether other TFs are also involved in the expression of IVR.

8401	TTGCGATATA	TGAAGATAAA	TGTTATGGCG	GTTATAAAAA	ATGTAAGGGC	AGGAGTCTTG
8461	CTAATCTTCA	CACCTGAATA	GTCTAACTTC	AGGGGATACC	CTCCTGCTCA	ТТGАСТ <mark>САСТ</mark>
8521	GTAA <mark>TTGA</mark> TT	AGCAAATGAA	AACAGCTTCA	TTGGGTTCTC	TGTCTCCTAA	ATTTTGTATA
8581	TCTGCTGTTG	CTTTTCTTTT	TTCTAATAGC	ACAAAAAGAA	ACACATAACA	CATATACCAC
8641	TTGTATACTG	GAAGATTTAT	АТТАТСАААА	AAATCTGTTT	GTCTTTGAAC	TGGGAGAACA
8701	GGGGCAAAGT	GCTATCAAAA	TGTCCTCTAG	CCAAAAACTC	GTTTCTACTT	GGTCTTAGTT
8761	TTTGTTTGAT	TGTTCTTATT	CTGCAGGTGG	TCTGGAAATT	TTTGCTGAGC	TTCTGCGACG
8821	ATTCCCTTGT	AACATACATA	TACTGCTTGA	AATGGCTAAG	GTTAGGGAAG	AATGACTATA
8881	TGGAAAATTA	TTGTAAATAT	TCATTGAATA	GTACAATGTT	CCTTTTATTC	ATAGTTCCTT
8941	TTTACATGTG	TTTGGCAGTA	GGTCAAATGG	GCAATACAAT	GATCATGATT	AACATGTCAG
9001	AGGCTCATGG	TTGTATCATT	TGAAATATAA	CAGGGTTATG	TGAAAGTAGT	TGTTCCTTTT
9061	CTTTCTTGTG	GATCAAATAT	TCTCTATACT	GGCTATCTCC	TTTACTGGTG	TCAGATTCAC
9121	GTAGACAGGT	CGATATAGTG	AAGTATCACA	T <mark>GCC</mark> TAGCAT	TATCTTTCCA	TGTTTTAGAT
9181	TTGTCTGGTC	TAATCATTTA	CATCTTTGCC	GTGCTTTCAA	GGGAATTCCT	TGATTGTAAT
9241	GATTGTTCTA	AATTACCCAG	ATTTAGAGAA	CTACTCTTCG	TATTTGTTTC	TTTTCACGGA
9301	TGGTCGAATG	CTCATTTCTT	AGGATGACAG	TTTTTACCCT	AAACACTTAA	AGGACACAGA
9361	GTCTCAGCTT	TTTCTCTGGC	AACTTTGTTC	TGAAAATAAG	CAAAATGTTA	TGCGTATCCC
9421	TGTCTCATAT	Т <mark>САССЛ</mark> ТТСТ	GAAG <mark>GTTGA</mark> A	GCC <mark>ATCATTG</mark>	GAAAAACTGA	AGAAGCTATA
9481	ACAGATTTTG	AGAAAGGTAA	AACGAGTGCT	ACATCTTGAT	GTGATATTTT	TATCATTTCT
9541	AAATGTGTAC	TACATATAAC	ATTGTATGGT	CTCTACAG <mark>GT</mark>	TCGATCCACT	GATCCGTATG
9601	TTATAACTTA	TATGGACGAG	TATGCTATGC	TCTTAAAGCA	GAAATCTGAT	TATTCGAAGT
9661	TGAATAAATT	AGTACATGAT	CTACTAAACA	TAGACCCGAC	CAGGCC TGAG	GTTTTTGTGG
9721	CCTTG <mark>T</mark> CTGT	CTTATGGGAA	AGGAAAGACG	AGAGAGGGGC	TTTATCTTAT	GCTGAGAAGG
9781	TTAGTTTTGT	TTAGTT <mark>TAC</mark>	TTT TTGAGAT	TGTTGTTGAT	GAATCAAGCA	TAAACTAGTG
9841	ATAATGTTGC	TTTTTTGCAA	CCAGAGCATC	CGCATTGATG	AGAGACATAT	ACCAGGTTAT
9901	<mark>ATAATGAAG</mark> G	TGCGCTTAGA	TTTTGTTCTTG	TGTGTAGTGA	ATACTTTTCA	ATGCCACATA
9961	G <mark>TTAGA</mark> TTAT	TCTTTGTTTT	G <mark>GCC</mark> AACATC	TAGGGAAATT	TATTCTTCTC	GATGAATCGA
10021	COTGAAGOTG	CTGTGGTTGC	CTTTAGAGGT	GCTCAAGAGT	TGAGACCTGA	TOTTCOTTOT

**Figure 9.** Putative transcription start sites and promoter sites for transcription of the *Nt*IVR mRNA within the *NtAPC7* gene (cv. K326) from accession AWOJ01567524 (28 April 2015). Yellow highlighted regions constitute exons from upstream APC7 sequences, with the bottom one shown containing the IVR initiation start codons ATG (in turquoise highlight); introns are not highlighted. Putative promoter sites: WRKY TF binding sites [TTGACC/T] (teal highlight); GCC element of ethylene-responsive element-binding protein (EREBP) TF binding sites [AGCCGCC] (violet highlight); and <u>YACGT</u>GG/TC-like ABA-responsive element-binding protein site (ABRE) sequence (blue). Putative transcription start-site sequences [YY<u>A</u>N(T/A)YY], where transcription starts with the underlined A (in red or light grey highlight, with dark grey highlight for shared nucleotides in overlapping start sites). Note that some promoters and transcription start sites are located within introns, which are less conserved between species.

To allow for expression of IVR from APC7, transcription start sites need to be present upstream of the translation initiation site. In the sequence shown in Figure 9, there are five potential transcription start sites located between 39 bp and 221 bp upstream of the translation initiation codon (shown in red and grey highlights). The first two and last two overlap in the canonical initiation site sequence, YYAN(T/A)YY.

If our conclusions are correct that IVR was generated by the repurposing a functional domain of a larger protein to provide additional defense functions, then this may not be the only such situation. There are many genes that are active only at specific times in the cell cycle, and they may be available for other functions during other times that would not interfere with their primary function.

**Supplementary Materials:** The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/v14122782/s1, Table S1: Oligonucleotide primers used in reverse transcription-polymerase chain reaction amplification for the yeast two-hybrid assay; Figure S1: Genome organization of IVR-like genes in *Nicotiana benthamiana*; Figure S2: Comparison of an unknown protein-encoding gene fragment of *Arabidopsis thaliana* with the sequence encoding the full-length Arabidopsis anaphase-promoting complex subunit 7 (APC7). References [53–57] are cited in the supplementary materials.

**Author Contributions:** Conceptualization, J.-Y.Y. and P.P.; formal analysis, M.A., J.-Y.Y. and P.P.; investigation, J.-Y.Y., M.A., E.B. and P.P.; figure preparation, J.-Y.Y., E.B., M.A. and P.P.; writing and editing, P.P. and J.-Y.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported in part by grant number 32003533HD020 (to P.P.) from the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET); and in part by a grant from IPET (to J.-Y.Y.) through the Agriculture, Food and Rural Affairs Convergence Technologies Program for Educating Creative Global Leader Program, funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA) (321001–03). M.A. acknowledges financial support from Tarbiat Modares University. E.B. was supported by a postdoctoral fellowship from the NRF of the Republic of Korea (NRF-2019R1A6A3A03032528).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The data presented in this study are available in the article and Supplementary Materials, as well as in online databases.

Acknowledgments: This article is dedicated to the memory of Gad Loebenstein (1929–2018) for his pioneering work on IVR, together with his long-time collaborator, Abed Gera.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Wang, A. Dissecting the molecular network of virus-plant interactions: The complex roles of host factors. *Annu. Rev. Phytopathol.* 2015, 53, 45–66. [CrossRef] [PubMed]
- 2. Garcia-Ruiz, H. Susceptibility genes to plant viruses. *Viruses* 2018, 10, 484. [CrossRef] [PubMed]
- 3. Sun, T.; Zhang, Y. MAP kinase cascades in plant development and immune signaling. *EMBO Rep.* **2022**, 23, e53817. [CrossRef] [PubMed]
- 4. Van Loon, L.C.; van Strien, E.A. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.* **1999**, *55*, 85–97. [CrossRef]
- 5. Sanfaçon, H. Plant translation factors and virus resistance. *Viruses* **2015**, *7*, 3392–3410. [CrossRef]
- 6. Carr, J.P.; Lewsey, M.G.; Palukaitis, P. Signaling in induced resistance. Adv. Virus Res. 2010, 76, 57–121.
- Zvereva, A.S.; Pooggin, M.M. Silencing and innate immunity in plant defense against viral and non-viral pathogens. *Viruses* 2012, 4, 2578–2597. [CrossRef]
- 8. Cillo, F.; Palukaitis, P. Transgenic resistance. Adv. Virus Res. 2014, 90, 35–146.
- 9. Sela, I.; Appelbaum, S.W. Occurrence of antiviral factors in virus-infected plants. Virology 1962, 17, 543–548. [CrossRef]
- 10. Sela, I. Plant virus interactions related to resistance and localization of viral infections. Adv. Virus Res. 1981, 26, 301–337.
- Edelbaum, O.; Ilan, N.; Grafi, G.; Sher, N.; Stram, Y.; Novick, D.; Tan, N.; Sela, I.; Rubenstein, M. Two antiviral proteins from tobacco: Purification and characterization by monoclonal antibodies to human β-interferon. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 588–592. [CrossRef] [PubMed]

- 12. Edelbaum, O.; Sher, N.; Rubenstein, M.; Novick, D.; Tan, N.; Moyer, M.; Ward, E.; Ryals, J.; Sela, I. Two antiviral proteins, gp35 and gp22, correspond to β-1,3-glucanase and an isoform of PR-5. *Plant Mol. Biol.* **1991**, *17*, 171–173. [CrossRef] [PubMed]
- 13. Linthorst, H.J.M.; Meuwissem, R.L.J.; Kauffmann, S.; Bol, J.F. Constitutive expression of pathogenesis-related proteins PR-1, GRP, and PR-S in tobacco has no effect on virus infection. *Plant Cell* **1989**, *1*, 285–291. [PubMed]
- 14. Reichman, M.; Devash, Y.; Suhadolnik, R.J.; Sela, I. Human leukocyte interferon and antiviral factor (AVF) from virus-infected plants stimulate plant-tissues to produce nucleotides with antiviral activity. *Virology* **1983**, *128*, 240–244. [CrossRef] [PubMed]
- 15. Loebenstein, G.; Gera, A. Inhibitor of virus replication released from tobacco mosaic virus-infected protoplasts of a local lesion-responding tobacco cultivar. *Virology* **1981**, *114*, 132–139. [CrossRef]
- 16. Gera, A.; Loebenstein, G. Further studies of an inhibitor of virus replication from tobacco mosaic virus-infected protoplasts of a local lesion-responding tobacco cultivar. *Phytopathology* **1983**, *73*, 111–115. [CrossRef]
- Gera, A.; Loebenstein, G.; Shabtai, S. Enhanced tobacco mosaic virus production and supressed synthesis of a virus inhibitor in protoplasts exposed to antibiotics. *Virology* 1983, 127, 475–478. [CrossRef]
- Spiegel, S.; Gera, A.; Salomon, R.; Ahl, P.; Harlap, S.; Loebenstein, G. Recovery of an inhibitor of virus replication from the intercellular fluid of hypersensitive tobacco infected with tobacco mosaic virus and from uninfected induced-resistant tissue. *Phytopathology* 1989, 79, 258–262. [CrossRef]
- 19. Gera, A.; Loebenstein, G. Evaluation of antisera to an inhibitor of virus replication (IVR). *J. Phytopathol.* **1989**, *124*, 366–371. [CrossRef]
- 20. Gera, A.; Loebenstein, G.; Salamon, R.; Franck, A. Inhibitor of virus replication from protoplasts of a hypersensitive tobacco cultivar infected with tobacco mosaic virus is associated with a 23-K protein species. *Phytopathology* **1990**, *80*, 78–81. [CrossRef]
- 21. Ahl, P.; Gianinazzi, S. *b*-proteins as a constitutive component in highly (TMV) resistant interspecific hybrids of *Nicotiana glutinosa* × *Nicotiana debneyi*. *Plant Sci. Lett.* **1982**, *26*, 173–181. [CrossRef]
- 22. Gianinazzi, S.; Ahl, P. The genetic and molecular basis of b-proteins in genus Nicotiana. *Neth. J. Plant Pathol.* **1983**, *89*, 275–281. [CrossRef]
- Tahiri-Alaloui, A.; Dumas-Gaudot, E.; Gianinazzi, S.; Antoniw, J.F. Expression of the PR-b<sub>1"</sub> gene in roots of two *Nicotiana* species and their amphidiploid hybrid infected with virulent and avirulent races of *Chalara elegans*. *Plant Pathol.* **1993**, 42, 728–736. [CrossRef]
- 24. Loebenstein, G.; Gera, A.; Gianinazzi, S. Constitutive production of an inhibitor of virus replication in the interspecific hybrid of *Nicotiana glutinosa* × *Nicotiana debneyi. Physiol. Mol. Plant Pathol.* **1990**, *37*, 145–151. [CrossRef]
- 25. Gera, A.; Tam, Y.; Teverovsky, E.; Loebenstein, G. Enhanced tobacco mosaic virus production and suppressed synthesis of the inhibitor of virus replication in protoplasts and plants of local lesion responding cultivars exposed to 35 °C. *Physiol. Mol. Plant Pathol.* **1993**, *43*, 299–306. [CrossRef]
- 26. Malamy, J.; Hennig, J.; Klessig, D.F. Temperature-dependent induction of salicylic acid and its conjugates during the resistance response to tobacco mosaic virus. *Plant Cell* **1992**, *4*, 359–366. [CrossRef] [PubMed]
- 27. Yang, Y.; Klessig, D.F. Isolation and characterization of a tobacco mosaic virus-inducible *myb* oncogene homolog from tobacco. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14972–14977. [CrossRef] [PubMed]
- 28. Akad, F.; Teverovsky, E.; David, A.; Czosnek, H.; Gidoni, D.; Gera, A.; Loebenstein, G. A cDNA from tobacco codes for an inhibitor of virus replication (IVR)-like protein. *Plant Mol. Biol.* **1999**, *40*, 969–976. [CrossRef] [PubMed]
- Akad, F.; Teverovsky, E.; Gidoni, D.; Elad, Y.; Kirschner, B.; Ray-David, D.; Czosnek, H.; Loebenstein, G. Resistance to *Tobacco mosaic virus* and *Botrytis cinerea* in tobacco transformed with complementary DNA encoding an inhibitor of viral replication-like protein. *Ann. Appl. Biol.* 2005, 147, 89–100. [CrossRef]
- 30. Fischer, U.; Dröge-Laser, W. Overexpression of *Nt*ERF5, a new member of the tobacco ethylene response transcription factor family enhances resistance to *Tobacco mosaic virus*. *Mol. Plant-Microbe Interact*. **2004**, *17*, 1162–1171. [CrossRef]
- 31. Rakhshandehroo, F.; Takeshita, M.; Squires, J.; Palukaitis, P. The influence of RNA-dependent RNA polymerase 1 on *Potato virus Y* infection and on other antiviral response genes. *Mol. Plant-Microbe Interact.* **2009**, *22*, 1312–1318. [CrossRef] [PubMed]
- 32. Rakhshandehroo, F.; Behboodi, B.S.; Mohammadi, M. Changes in peroxidase activity and transcript level of the *MYB1* gene in transgenic tobacco plants silenced for the *RDR-1* gene after systemic infection with *Potato virus* Y<sup>0</sup>. J. Phytopathol. **2012**, 160, 187–194. [CrossRef]
- 33. Yoon, J.-Y.; Baek, E.; Kim, M.; Palukaitis, P. The Virus-Induced Transcription Factor SHE1 Interacts with and Regulates the Expression of the Inhibitor of Virus Replication (IVR) in N Gene Tobacco; Department of Agricultural Convergence Technology, Jeonbuk National University: Jeonju, Republic of Korea, 2022; to be submitted.
- 34. Yoon, J.-Y.; Palukaitis, P. Cucumber mosaic virus 1a protein interacts with the tobacco SHE1 transcription factor and partitions between the nucleus and the tonoplast membrane. *Plant Pathol. J.* **2021**, *37*, 182–193. [CrossRef] [PubMed]
- Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A. Highly accurate protein structures prediction with AlphaFold. *Nature* 2021, 596, 583–589. [CrossRef]
- ColabFold. Available online: https://conlab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2 .ipynb#scrollT0=iceGdbe\_Pmt9 (accessed on 10 August 2021).
- 37. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Lim, H.; Ovchinnikov, S.; Steinegger, M. ColabFold:making protein folding accessible to all. *Nature Methods* **2022**, *19*, 679–682. [CrossRef]
- Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. LDDT: A local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* 2013, 29, 2722–2728. [CrossRef]

- Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Židek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; et al. Highly accurate protein structure prediction for the human proteome. *Nature* 2021, 596, 590–596. [CrossRef]
- Wilson, C.G.; Kajander, T.; Regan, L. The crystal structure of NlpI. A prokaryotic teratricopetide repeat protein with a globular fold. FEBS J. 2005, 272, 166–179. [CrossRef]
- 41. Thornton, B.R.; Toczyski, D.P. Precise destruction: An emerging picture of the APC. Genes Dev. 2006, 20, 3069–3078. [CrossRef]
- 42. Han, D.; Kim, K.; Kim, Y.; Kang, Y.; Lee, J.Y.; Kim, Y. Crystal structure of the N-terminal domain of anaphase-promoting complex subunit 7. *J. Biol. Chem.* **2009**, *284*, 15137–15146. [CrossRef]
- Ueda, H.; Yamaguchi, Y.; Sano, H. Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants. *Plant Mol. Biol.* 2006, *61*, 31–45. [CrossRef] [PubMed]
- O'Reilly, E.K.; Wang, Z.; French, R.; Kao, C.C. Interactions between structural domains of the RNA replication proteins of plant-infecting RNA viruses. J. Virol. 1998, 72, 7160–7169. [CrossRef] [PubMed]
- Liu, Y.; Burch-Smith, T.; Schiff, M.; Feng, S.; Dinesh-Kumar, S.P. Molecular chaperone Hsp90 associates with resistance protein N and its signaling proteins SGT1 and Rar1 to modulate an immune response in plants. *J. Biol. Chem.* 2004, 279, 2101–2108. [CrossRef] [PubMed]
- Liu, Y.; Schiff, M.; Dinesh-Kumar, S.P. Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, *COI1* and *CTR1* in *N*-mediated resistance to tobacco mosaic virus. *Plant J.* 2004, *38*, 800–809. [CrossRef] [PubMed]
- 47. van Verk, M.C.; Pappaioannou, D.; Neeleman, L.; Bol, J.F.; Linthorst, H.J.M. A novel WRKY transcription factor is required for induction of *PR-1a* gene expression by salicylic acid and bacterial elicitors. *Plant Physiol.* **2008**, *146*, 1983–1995. [CrossRef]
- 48. Aoki, K.; Yano, K.; Suzuki, A.; Kawamura, S.; Sakurai, N.; Suda, K.; Kurabayashi, A.; Suzuki, T.; Tsugane, T.; Watanabe, M.; et al. Large-scale analysis of full-length cDNAs from the tomato (*Solanum lycopersicum*) cultivar Micro-Tom, a reference system for the Solanaceae genomics. *BMC Genom.* 2010, *11*, 210. [CrossRef]
- 49. Loebenstein, G. Local lesions and induced resistance. Adv. Virus Res. 2009, 75, 73–117.
- 50. Visintin, R.; Prinz, S.; Amon, A. *CDC20* and *CDH1*: A family of substrate-specific activators of APC-dependent proteolysis. *Science* **1997**, 278, 460–463. [CrossRef]
- 51. Vodermaier, H.C.; Gieffers, C.; Maurer-Stroh, S.; Eisenhaber, F.; Peters, J.-M. TRP subunits of the anaphase-promoting complex mediate binding to the activator protein CDH1. *Curr. Biol.* **2003**, *13*, 1459–1468. [CrossRef]
- 52. Mur, L.A.J.; Bi, Y.-M.; Darby, R.M.; Firek, S.; Draper, J. Compromising early salicylic acid accumulation delays the hypersensitive response and increases viral dispersal during lesion establishment in TMV-infected tobacco. *Plant J.* **1997**, *12*, 1113–1126. [CrossRef]
- Rizzo, T.R.; Palukaitis, P. Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 1. J. Gen. Virol. 1989, 70, 1–11. [CrossRef] [PubMed]
- 54. Rizzo, T.R.; Palukaitis, P. Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 2. J. Gen. Virol. 1988, 69, 1777–1787. [CrossRef] [PubMed]
- 55. Whitham, S.; Dinesh-Kumar, S.P.; Choi, D.; Hehl, R.; Corr, C.; Baker, B. The product of the tobacco mosaic virus resistance gene *N*: Similarity to Toll and the Interleukin-1 receptor. *Cell* **1994**, *78*, 1101–1115. [CrossRef] [PubMed]
- 56. Liu, Y.; Schiff, M.; Marathe, R.; Dinesh-Kumar, S.P. Tobacco Rar1, EDS1, and NPR/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. *Plant J.* **2002**, *30*, 415–429. [CrossRef]
- 57. Goelet, P.; Lomonossoff, G.P.; Butler, P.J.; Akam, M.E.; Gait, M.J.; Karn, J. Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5818–5822. [CrossRef] [PubMed]