Journal of Agricultural Science and Technology (JAST), 28(6)
In Press, Pre-Proof Version

Design of degenerate primers for PCR amplification of genomic DNA encoding eIF4G in *Verticillium dahliae* using local alignment search

Abbas Jolodar[1]

Abstract

Verticillium dahliae is a soilborne fungal pathogen that causes a vascular wilt disease in numerous economically significant crops and ornamentals worldwide. In this study, a partial genomic DNA fragment encoding eukaryotic translation initiation factor 4 gamma was amplified from *V. dahliae* using degenerate primers design based on the local alignment search method. To identify homologous proteins in closely related species, the NCBI GenBank database was queried using the HomoloGene program and a target keyword. The aligned part of the selected sequences revealed two conserve motifs, which facilitated the design of a pair of degenerate primers. Using Touchdown PCR, a 503-bp genomic DNA from *V. dahliae* was successfully amplified and sequenced, designated as *Vd*eIF4G. This fragment contained two open reading frames (ORFs), encoding a total of 135 amino acids, separated by a 97-nucleotide intron. The intron-exon junction adhered to the GU-AG rule. At the nucleotide level, taxonomic analysis of *Vd*eIF4G revealed three matches within the genus *Verticillium,* covering 81% of overlapping regions with sequence similarity ranging from 97.99 to 100%. Phylogenetic analysis at the protein level showed that *Vd*eIF4G was identical to *V. dahliae* but exhibited divergence from *V. nonalfalfae* and *V. alfalfae* with the lowest genetic distance between these sequences being 1.8%. Furthermore, the presence of small and polar amino acids, along with a glycine-rich domain suggests potential involvement in host interactions and may contribute to fungal virulence. Degenerate primers designed using a local alignment search method effectively amplified homologous sequences of the corresponding eIF4G domain-containing protein from *V. dahliae.*

Keywords: Degenerate primers, eIF4G, Fungi, Intron, Touchdown PCR, *Verticillium dahlia.*

Introduction

Verticillium dahliae Kleb is a well-known soilborne fungal pathogen that causes a vascular wilt disease in plants, leading to significant economic losses across a broad range of crops worldwide. Annual losses caused by this pathogen are estimated in the billions of dollars

[1]Biochemistry and Molecular Biology Section, Department of Basic Sciences, Faculty of Veterinary Medicine, Shahid Chamran University of Ahvaz, Ahvaz, Islamic Republic of Iran. e-mail: jolodara@scu.ac.ir or jolodara@yahoo.com

(Inderbitzin and Subbarao, 2014). *V. dahliae* has an exceptionally wide host range, infecting trees, herbaceous plants, plantation crops, and mushrooms, thriving in subtropical, tropical regions to cold and warm climates.

To date, *Verticillium* wilt has been reported in nearly 80 plant genera, affecting more than 410 plant species globally (Pegg and Brady, 2002). The pathogen primarily invades plants through the roots, leading to a range of symptoms, including external manifestations such as leaf chlorosis and necrosis, as well as internal vascular discoloration (Chen *et al.,* 2021). The characteristics of the genus *Verticillium*, its role in wilt disease, and methods for its control have been extensively studied (Barbara and Clowes 2003; Faradin and Touma 2006; Klosterman *et al.* 2009). Continued research on its pathogenic mechanisms and resistance strategies remains crucial for managing its widespread impact on agricultural productivity.

In eukaryotic cells, eIF4E plays a pivotal role in the initiation and regulation of translation. By interacting with the 5'-cap structure of mRNA and its translation partner, eIF4G, eIF4E facilitates the recruitment of mRNAs to the ribosome (Prevot *et al.,* 2003). The presence of eIF4E, eIF4A, and eIF4G is essential for cap recognition and subsequent RNA helicase activities, which drive efficient protein synthesis (Von der Haar *et al.,* 2004). While the function of eIF4G as a structural scaffold in translation initiation is well-documented, the broader biological significance of eIF4G-related proteins remains largely unexplored. Further evidence from deletion and mutagenesis studies has provided insight into the functional role of the NIC domain within eIF4G. In both yeast (Dominguez et al. 1999; Neff and Sachs 1999) and human (Imataka and Sonenberg 1997; Morino *et al.* 2000) research has demonstrated that the NIC region is essential for interaction with the helicase eIF4A, underscoring its significance in translation initiation and RNA unwinding mechanisms. To date, the eIF4G family is among the least characterized within eIF4 protein families, warranting further investigation into its molecular mechanisms and functional diversity.

Genome-wide sequencing projects have identified the gene families corresponding to each eIF4 factor across various eukaryotic species. To date, two eIF4G proteins have been characterized in humans, namely eIF4G-I (Bradley *et al.,* 2002), and one has been identified in rabbit, eIF4G-I (Lamphear *et al.,* 1993). In plants, wheat possesses two eIF4G proteins (Lax et al., 1985, 1986) along with eIF(iso)4G found in other plants species (Browning *et al.,* 1992; Kim *et al.,* 1999). In fungi, *Saccharomyces cerevisiae* contains two elF4G homologs, eIF4631 and TIF4632 (Goyer *et al.,* 1993), while *Schizosaccharomyces pombe* has a single

characterized eIF4G (Hashemzadeh-Bonehi *et al.,* 2003). Additionally, *Drosophila melanogaster*, possesses a distinct homolog, Dm-eIF4G (Hernandez *et al.,* 1998).

In this study, a partial genomic DNA (gDNA) fragment encoding the N-terminal region of the eIF4G protein from the soilborne fungal pathogen *V. dahliae* was successfully amplified using primers designed through a local alignment search method.

Materials and Methods

Design of degenerate primers

Using the HomoloGene program, protein sequences containing "eukaryotic initiation factor 4G" related to the target gene were retrieved. An extensive across multiple organisms identified four fungal sequences from different species in the *Verticillium* genus. These sequences were retrieved from the GenBank database to construct a specialized protein collection database for *V. dahliae* relatives. The maximum e-value accepted in this study was 8e-07. Sequences with no homology below this threshold were excluded. Based on the conserved regions identified in the retrieved sequences, two degenerate primers were designed using local alignment search method (Jolodar, 2019).

Genomic DNA extraction

V. dahliae isolates were cultured on potato dextrose agar (PDA) to promote mycelial growth, following the method described by Yan *et al.* (2022). One hundred mg of mycelium with 300 mg of the sterilized sand (white quartz, Sigma) were ground to powder for 30 sec with a pestle in the presence of liquid nitrogen. 600 µl of extraction buffer [100 mM Tris-HCl (pH 8.0), 20 mM EDTA, 0.5 M NaCl and 1% sodium dodecylsulfate] was added to the soaked mycelium. The mixture was extracted with equal volumes of phenol/chloroform/isoamyl alcohol (25:24:1) and centrifuged at 12,000 rpm for 5 min. The supernatant was transferred to a new microtube and re-extracted with 400 µl of chloroform. The supernatant was precipitated with 1 volume of ice-cold isopropanol and centrifuged at 12,000 rpm for 10 min. The precipitate was washed with 70% ice-ethanol and was resuspended in 400 µl of TE buffer (10 mM Tris-HCl, 1 mM Na2EDTA, pH 8.0) containing 20 µg/ml ribonuclease A. The samples were incubated at 37°C for 20 min and then extracted with equal volumes of phenol/ chloroform /isoamyl alcohol. The aqueous phase was precipitated by adding 2.5 volumes of ice-cold absolute ethanol and 1/10 volume of 3 M sodium acetate. The samples were then centrifuged at 12,000 rpm for 15 min. The DNA precipitate was washed with ice-cold 70% ethanol, dried and suspended in 100 µl of TE buffer (10mM Tris- HCl, 1mM EDTA). The

DNA fragments were visualized on a 1% agarose gel in TAE buffer (40 mM Tris/Acetate, 1 mM EDTA) (pH 8.0) at 5 V/cm between positive and negative poles. The gels were stained with 50 µg/ml ethidium bromide and photographed under ultraviolet light. Sterile distilled water was used instead of DNA as a negative control.

TD-PCR Amplification

Degenerate primers were designed on two highly conserved regions identified through a local alignment search method. Using these conserved sequences, two degenerate oligonucleotide primers were designed and synthesized. The forward primer, EIF4-F (5'-GCGgaattcAAYAAYWSAAYAAYATG), corresponding to the protein-coding sequence (NNSNNM), and the reverse primer, EIF4-R (5'-GCGaagcttATRTGNGCNCCNGG), aligns with the protein sequence QPGAHI. In these sequences N, R, W and Y represent nucleotide variations: N (A/G/C/T), R (A/G), W (A/T) and Y (C/T), ensuring primer compatibility across multiple homologous targets. To enhance PCR efficiency and facilitate directional cloning, *Eco*RI and *Hind*III restriction sites were incorporated at the 5' ends of the forward and reverse primers, respectively. Because cleavage sites positioned near the primer ends can result in inefficient digestion by certain restriction enzymes, additional bases (GCG) were appended to the 5' ends. This modification ensures optimal enzyme recognition and effective digestion of the amplified fragments. The restriction enzyme recognition sequences are shown in lowercase letters. TD-PCR reactions were performed under the following conditions: 94°C × 3 min for one cycle. 94°C × 30 s, 42°C × 50 s, 72°C × 1 min for 15 cycles with a 1°C increment per cycle. 94°C × 30 s, 57°C × 50 s, 72°C × 40 s for 20 cycles. Each PCR reaction was performed in a total volume of 25 µl containing 20 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl2, 0.2 mM dNTPs of each, 0.5 units of *Taq* DNA polymerase, 0.4 µM of each primer, and 100 ng of genomic DNA. PCR products were loaded on 1% agarose gel containing 50 µg/mL safe stain (Cinagen, Iran) electrophoresis and visualized by UV light.

DNA sequencing and analysis

To determine the sequence of the amplified gene fragments, single bands were purified using an agarose purification kit (Cinagen, Iran) and then sequenced using an Applied Biosystems 373 automated sequencer. Nucleotide and amino acid sequences were edited and analyzed using BioEdit software. Primers were designed using Primer3 (Untergasser *et al.*, 2012). The sequences were compared using the BLASTn algorithm from the NCBI Genbank website (ncbi.nlm.nih.gov). The translation of the sequences into amino acids was analyzed using the

EXPASY tool program (Gasteiger *et al.,* 2005) (ca.expasy.org/tools/dna.html). Multiple alignments were performed using the CLUSTAL_W program (genome.jp/tools-bin/clustalw) (Thompson *et al.,* 1994) and edited with the BOXSHADE software (Stothard *et al.,* 2000) (bioinformatics.org/SMS/multi_align.html). Phylogeny and genetic distance analysis were performed using the 1000-joining-neighborhood bootstrap test using the MEGA7 software (Tamura *et al.,* 2007). Intron splice sites were determined using NetAspGene -1.0 program (services.healthtech.dtu.dk/services/NetAspGene-1.0/). For obtaining statistical data of proteins, the Gene infinity program was used (org/sms/sms_proteinstats.html).

Results

Design of degenerate primers

To identify homologous transcripts in species closely related to *V. dahliae*, the HomoloGene program was utilized within the NCBI GenBank database. The search was performed using the keyword "eukaryotic initiation factor 4G" in the NCBI gene database through the using the BLASTp algorithm. The query was restricted to the *Plectosphaerellaceae* family, which comprises various plant-pathogenic genera and soilborne fungal species. By using this technique, a total of four homologous sequences were identified and retrieved for local alignment, including *V. dahliae* (PNH37575.1), *V. longisporum* (CRK25417.1), *V. alfalfae* (XP_003002440.1) and *V. nonalfalfae* (XP_028490779.1). Degenerate primers were designed based on the sequence numbering of *V. dahliae* (PNH37575.1) targeting amino acid positions between 88-93 for the sense primer and 221-226 for the antisense primers (Figure 1). Two conserved regions, each containing at least six amino acids, were selected as primer-banding sites. These primers were optimized to meet several critical parameters, including appropriate binding temperature, a balanced GC content, prevention of dimenization through self-complementary ends, and minimal secondary structure formation. Since some amino acids are encoded by multiple triplet codons, the primer sequences were designed using IUPAC ambiguity codes to accommodate codon variability. The degeneracy of each primer was determined by multiplying the degeneracy values of the respective amino acids according to the IUPAC coding system. The degeneracy was calculated to be 256-fold for the both primer.

Fig. 1. Alignments of representative fungus EIF4G protein sequences for designing degenerate primers. Shading indicates identity (black) or conservative substitutions (grey). The conserved amino acids for designing degenerate primers are boxed and primer sites indicated by arrows.

TD-PCR amplification of *Vd*EIF4G

A total of 4 µg of genomic DNA was extracted from 0.5 g of fungal tissue and subsequently used in polymerase chain reactions. Initially, PCR optimization was conducted using 50 ng of genomic DNA as the template. A pair of primers was tested under conventional PCR conditions with an annealing temperature of 58°C; however, no amplification was observed. To improve results, the genomic DNA concentration was increased, but this adjustment also failed to produce a positive outcome. Since higher DNA concentration did not enhance amplification, the TD-PCR technique was implemented. In this method, the PCR reaction was initiated at an annealing temperature of 42°C for 15 cycles, with a gradual increase of 1°C per cycle, followed by an additional 20 cycles at a final annealing temperature of 57°C. Using degenerate primers in TD-PCR, a 503 bp fragment, designated as *Vd*eIF4G was successfully amplified.

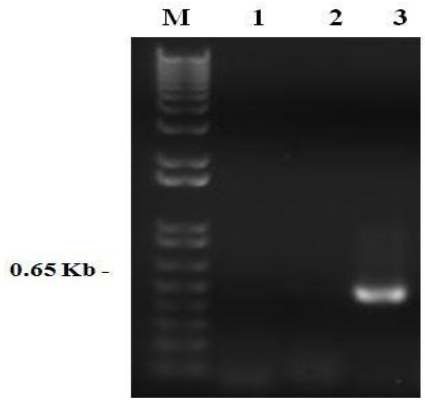

Fig. 2. Agarose gel electrophoresis of PCR products amplified from *V. dahliae* genomic DNA. Lane M represents the DNA size marker. Lanes 1, 2, and 3 correspond to the following controls: (1) a negative control using sterile water in place of DNA, (2) a negative control without *Taq* DNA polymerase, and (3) the PCR-amplified product.

Characterization of *Vd*EIF4G

The amplified *VdEIF4G* gene from *Verticillium dahliae* was confirmed through sequence analysis. The sequence exhibited a GC content of 62.6%, with nucleotide composition distributed as 19.3% A, 37.6% C, 25.0% G, and 18.1% T. The amplified fragment was compared with the NCBI GenBank database using the BLASTn program with the 'highly similar sequences (megablast)' algorithm. The analysis revealed sequence similarity ranging from 97.88% to 100%, with matches to the only three incomplete fungal mRNA sequences from the *Plectosphaerellaceae* family. These included the eukaryotic initiation factor 4F subunit p130 from *V. dahliae* VdLs.17 (VDAG_06653, XM_009659357.1), an uncharacterized protein from *V. nonalfalfae*, D7B24_003156 (XM_028637349.1), and the eukaryotic initiation factor 4F subunit p130 from *V. alfalfae* VaMs.102 (XM_003002394.1). No genomic sequence similarity was observed. As shown in Figure 3, the amplified genomic fragment of *VdEIF4G* contained two open reading frames (ORFs), encoding a total of 135 amino acids, separated by a 97-bp intron. The coding sequence produces a 135-amino-acid protein interrupted by a single intron, with a composition of aliphatic (35.56%), aromatic (2.96%), basic (8.89%), and acidic (17.78%) side groups. Among the amino acids, serine (16.30%), glycine (14.81%), and proline (13.33%) were the most abundant, with 22, 20, and 18 residues, respectively. Further analysis using the Gene Infinity program identified a glycine-rich domain within the amino acids positioned between residues 30 and 40 (Figure 3), potentially contributing to protein flexibility and functional interactions.

Fig. 3. Displays the nucleotide and predicted amino acid sequences of *VdEIF4G*. Exonic regions are represented in uppercase letters, while intronic sequences are shown in lowercase. The donor (GT) and acceptor (AG) sequences at the intron-exon boundaries are highlighted in bold. Additionally, the putative glycine-rich domain is emphasized using both bold and underlined formatting to distinguish its position within the sequence**.**

In the taxonomic report of *Vd*EIF4G at the nucleotide level only showed 3 hits from the genus *Verticillium* in 81% of overlapping regions with a similarity between 97.99 and 100%. At the protein level, this report showed a total of 155 hits in the sub-class *Hypocreomycetidae*, of which 146, 25, and 23 were from the order *Glomerellales*, the family *Plectosphaerellaceae*, and the genus *Verticillium*, respectively.

Phylogenetic analysis of *Vd*EIF4G

To analyze *VdEIF4G* within a broader evolutionary context, phylogenetic analysis was conducted based on deduced amino acid sequences from *V. dahliae*, along with 15 sequences from the *Plectosphaerellaceae* family, which includes numerous plant-pathogenic genera and soilborne fungal species. The constructed phylogram revealed two distinct clusters (A and B). In Cluster B, *VdEIF4G* was grouped with other *Verticillium* species, supported by a strong bootstrap score of 99**.** As expected, *VdEIF4G* exhibited high genetic similarity with various *eIF4G* homologs within the genus. This evolutionary relationship was confirmed by comparison with the four available homolog sequences in GenBank from *Verticillium*, demonstrating that *VdEIF4G* is closely related to *V. longisporum* (CRK25417.1) while showing slight divergence from *V. alfalfa* (XP_003002440.1) and *V. nonalfalfae* (XP_028490779.1). Additionally, the sequence of *Colletotrichum trifolii* (TDZ41177.1) was positioned within a separate sub-cluster, distinct from other *Verticillium* species, supported by a moderate bootstrap score of 43. Meanwhile, all non-*Verticillium* fungal sequences belonging to Cluster A were grouped together, with a bootstrap score of 94, indicating a clear

phylogenetic separation from *Verticillium* species. *Xylaria hypoxylon* (TGJ85310.1) used as an out group.

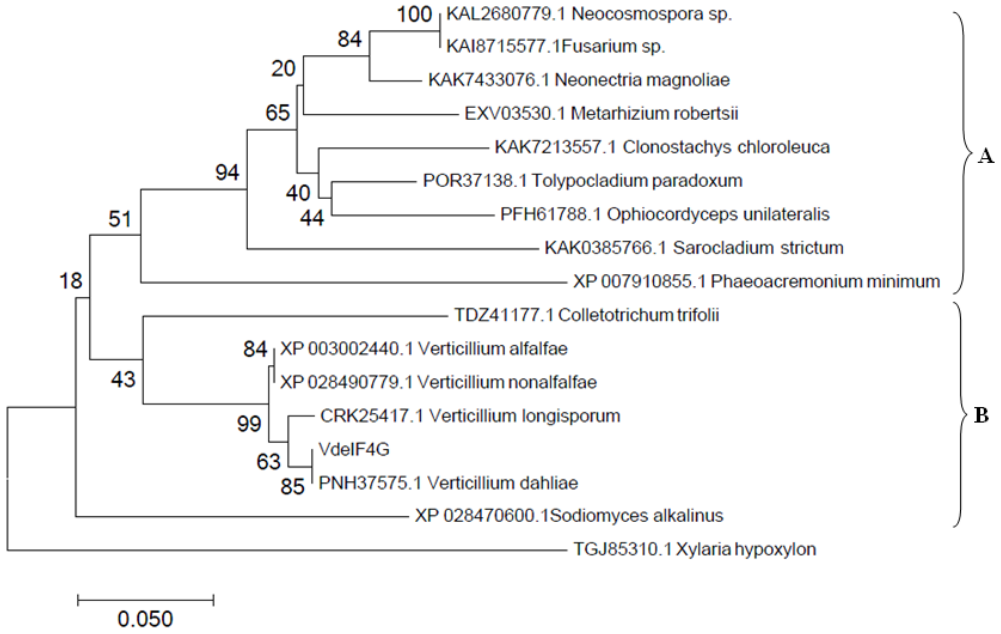

Fig. 4. Phylogenetic tree constructed from the amino acid sequences of *VdEIF4G* and related sequences using Neighbor-Joining analysis. Bootstrap values are based on 1,000 replicates, indicating the statistical support for each branch. Accession numbers corresponding to the species are listed before their names, representing related gene sequences retrieved from the GenBank database. The numbers positioned above the branches illustrate the evolutionary relationships between the grouped sequences.

The genetic distance of *VdEIF4G* was calculated using MEGA7 software (Tamura *et al.*, 2007). Among the 13 related protein sequences available in GenBank, *VdEIF4G* exhibited the lowest genetic distance (1.8%) compared to *V. longisporum* (CRK25417.1) and *V. alfalfa* (XP_003002440.1).

Table 1. Genetic pairwise distances of *Vd*EIF4G in comparison to related fungal proteins.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1-VdeIF4													
2-CRK25417.1 V. longisporum	0.018												
3-XP_003002440.1 V. alfalfae	0.018	0.018											
4-TDZ41177.1 C. trifolii	0.175	0.175	0.167										
5-POR37138.1_T. paradoxum	0.202	0.202	0.184	0.272									
6-KAL2680779.1 Neocosmospora sp.	0.202	0.202	0.184	0.263	0.114								
7-KAI8715577.1 Fusarium sp.	0.202	0.202	0.184	0.263	0.114	0.000							
8-XP_028470600.1 Salkalinus	0.184	0.202	0.184	0.263	0.272	0.246	0.246						
9-KAK0385766.1 S. strictum	0.254	0.202	0.237	0.298	0.158	0.175	0.175	0.307					
10-EXV03530.1 M. robertsii	0.202	0.202	0.184	0.246	0.096	0.123	0.123	0.263	0.193				
11-XP_007910855.1 P. minimum	0.281	0.263	0.263	0.289	0.254	0.272	0.272	0.307	0.298	0.246			
12-KAK7213557.1 C. chloroleuca	0.211	0.202	0.202	0.272	0.096	0.132	0.132	0.254	0.184	0.132	0.198		
13-KAK7433076.1 N. magnoliae	0.202	0.202	0.184	0.272	0.079	0.061	0.061	0.246	0.167	0.096	0.254	0.114	
14-PFH61788.1 O. unilateralis	0.228	0.202	0.211	0.263	0.096	0.158	0.158	0.289	0.211	0.123	0.281	0.132	0.123

Discussion

V. dahliae is a soilborne fungal pathogen that infiltrates host plants through the roots and spreads within the water-conducting tissues, leading to Verticillium wilt. Verticillium wilt symptoms stem from vascular blockage by fungal growth and plant aggregates, with pathogen toxins further driving disease and tissue damage (Fradin et al., 2006). Historically, the symptoms of *Verticillium* wilt have been attributed to vascular occlusion, caused by the accumulation of fungal mycelia and aggregated plant biomacromolecules. Additionally, phytotoxicity resulting from pathogen-secreted toxins further contributes to disease progression and host tissue damage (Fradin *et al.*, 2006).

In this study, the design of degenerate primers enabled the isolation of a partial genomic DNA (gDNA) fragment of the EIF4G gene from *V. dahliae*, utilizing closely related fungal species. To achieve this, degenerate primers were designed based on several known sequences of EIF4G proteins in different fungal species to identify homologous target genes in *V. dahliae*. Generally, designing primers becomes more challenging when the target organism is more distantly related. The primers were constructed based on two blocks of

conserved protein sequences that were aligned. For optimal primer design, the conserved regions had to be located in close proximity. The designed primers were required to meet several key criteria, including an appropriate annealing temperature, a suitable GC content range, and non-self-complementary ends to prevent self-complementarity (Kampke *et al.,* 2001). Traditional methods for designing degenerate primers utilize software programs such as Gene Fisher (Giegerich *et al.,* 1996), CODEHOP (Rose *et al.,* 2003), and PrimaClade (Gadberry *et al.,* 2005). However, these programs rely on a high degree of sequence conservation across all aligned sequences, which can limit their applicability. In this study, we adopted an alternative approach to identify sufficiently large conserved regions, enabling the design of an effective degenerate primer pair.

The careful optimization of the TD-PCR technique was a critical factor in achieving successful amplification using degenerate primers. In this method, the annealing temperature is initially set 5°C below the calculated melting point of the primers to enhance primer-template hybrid formation. Studies have shown that TD-PCR can significantly improve the success rate of degenerate PCR. At lower stringency, annealing facilitates the formation of a greater number of primer-template hybrids. As the PCR cycles progress, the annealing temperature is gradually increased, enhancing specificity and ensuring the amplification of complete primer-template hybrids. This stepwise temperature adjustment improves both efficiency and accuracy in sequence amplification (Scalon *et al.,* 2014)

Recent studies have revealed that the physical presence of introns in fungal genomes plays a crucial role in regulating cell survival under starvation conditions (Jo and Choi, 2015; Parenteau et al., 2019). In addition to their regulatory functions, introns serve as hosts for noncoding RNA (ncRNA) genes, including microRNAs and snoRNAs. A genomic survey of *Verticillium dahliae* indicated that 79.3% of its protein-coding genes contain at least one intron, with an average intron length of 100 bp (Muzafar *et al.,* 2021; Jin *et al.,* 2017). In fungal genomes, introns are generally shorter compared to other eukaryotes, with 88–99.8% exhibiting canonical splice sites **(**GU-AG**).** In contrast, non-canonical splice sites **(**GC-AG**)** are found in 1–2% of introns, while AU-AC splice sites occur in 0.09% of cases (Sieber *et al.,* 2018). The annotation of eukaryotic genes, particularly the identification of correct intron-exon structures, remains one of the most complex challenges in genome research (Sharp, 1994). Two primary obstacles in predicting splice sites accurately include: (1) the presence of numerous GT and AG dinucleotides that do not correspond to true splice junctions, leading to false positives, and (2) the occurrence of atypical splice sites, which can

result in false negatives (Pucker *et al.,* 2017). Analysis of the *Vd*EIF4G genomic sequence confirmed that splicing predominantly follows the GU-AG rule, with the gene interrupted by a 97-nucleotide intron. Intron-exon boundaries were identified through consensus donor-acceptor sequences. Key signals used for this determination include junction sensors, which typically feature a GU dinucleotide sequence signal (C/A)AG/GT(A/G)AGT at the +1/+2 position at the 5′ end of the intron, and an AG sequence signal (T/C)nN(C/T)AG/G at the 3′ splice site (Mount, 1989).

Certain amino acids, including glycine, serine, and proline, along with polar and charged residues, are commonly associated with intrinsically disordered proteins (IDPs). These amino acids contribute to structural flexibility and prevent stable folding, allowing proteins to adopt dynamic conformations. Analysis of *Vd*EIF4G revealed a high proportion of glycine (14.81%), serine (16.30%), and proline (13.33%), strongly suggesting that it exhibits intrinsic disorder. These residues enhance molecular flexibility; promote structural flexibility, and influence protein folding and adaptability. Additionally, we hypothesize that a glycine-rich domain, may be associated with RNA-binding, stress response, and a protein-protein interaction especially in fungal pathogens was identified in *Vd*eIF4G. Given its glycine-rich nature and predicted disorder, this protein may play a role in fungal virulence (Nishimura *et al.,* 2016), host interactions, or RNA metabolism and signal transduction pathways (Mangeon *et al.,* 2010). Further experimental studies, including structural characterization and functional assays, are required to validate its biological significance in *V. dahliae* pathogenicity.

In conclusion, degenerate primers designed using a local alignment search method successfully amplified homologous sequences of the eIF4G domain-containing protein from *V. dahliae*.

Acknowledgement

This work was financially supported by research grant from the Vice President of the Research Affairs Office at the Shahid Chamran University of Ahvaz, Ahvaz- Iran.

References

1. Barbara, D. J and Clewes, E. 2003. Plant pathogenic Verticillium species: how many of them are there? *Mol Plant Pathol.,* **4:** 297-305.

2. Bradley, C. A., Padovan, J. C., Thompson, J. L., Benoit, C. A., Chait, B. T. and Rhoads, R. E. 2002. Mass spectrometric analysis of the N terminus of translational initiation factor eIF4G-1 reveals novel isoforms. *J. Biol. Chem.,* **277:** 12559-12571.
3. Browning, K. S., Webster, C., Roberts, J. K. and Ravel, J. M., 1992. Identification of an isozyme from protein synthesis initiation factor 4F in plants. *J. Biol. Chem.*, **267**: 10096-10100.
4. Chen, J. Y., Klosterman, S. J., Hu, X. P., Dai, X. F. and Subbarao, K. V. 2021. Key insights and research prospects at the dawn of the population genomics era for Verticillium dahliae. *Annual Review of Phytopathology*, **59**: 31-51.
5. Dominguez, D., Altmann, M., Benz, J., Baumann, U., Trachsel, H. 1999. Interaction of translation initiation factor eIF4G with eIF4A in the yeast *Saccharomyces cerevisiae.* J. Biol. Chem., 274: 26720–26726.
6. Fradin, E. F. and Thomma, B. P. H. J. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by V. dahlia and V. albo-atrum. *Mol Plant Pathol.,* **7:** 71-86.
7. Gadberry, M. D., Malcomber, S. T., Doust, A. N. and Kellogg E. A. 2005. Primaclade-a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics*, **21(7):** 1263-4. doi: 10.1093/bioinformatics/bti134. PMID: 15539448.
8. Gasteiger, E., Hoogland, C., Gattiker, A, Duvaud, S., Wilkins, M. R. and Appel, R. D.2005. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa (NJ): Humana Press; 2005. p. 571-607.
9. Giegerich, R., Meyer, F. and Schleiermacher C. 1996. GeneFisher-software support for the detection of postulated genes. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **4:** 68-77(1996). PMID: 8877506.
10. Goyer, C., Altmann, M., Lee, H. S., Blanc, A., Deshmukh, M. and Woolford J. R. 1993. TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function. *Mol. Cell. Biol.,* **13:** 4860-4874.
11. Inderbitzin, P., Subbarao, K. V. 2014. Verticillium systematics and evolution: how confusion impedes Verticillium wilt management and how to resolve it. *Phytopathology,* **104**: 564-574.

12. Hashemzadeh-Bonehi, L., Curtis, P. S., Morley, S. J., Thorpe, J. R. and Pain, V. M. 2003. Overproduction of a conserved domain of fission yeast and mammalian translation initiation factor eIF4FG causes aberrant cell morphology and results in disruption of the localization of F-actin and the organization of microtubules. *Genes Cells,* **8:** 163-178.
13. Hernandez, G., Castellano, M. M., Agudo, M. and Sierra, J. M. 1998. Isolation and characterization of the cDNA and the gene for eukaryotic translation initiation factor 4G from Drosophila melanogaster. *Eur. J. Biochem.,* **253:** 27-35.
14. Imataka, H. and Sonenberg, N. 1997. Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol. Cell Biol.,* **17:** 6940-6947.
15. Jo, B.S. and Choi, S. S. 2015. Introns: the functional benefits of introns in genomes. *Genomics Inform.,* **13:** 112-8.
16. Jolodar, A. 2019. Designing of Degenerate Primers-Based Polymerase Chain Reaction (PCR) for Amplification of WD40 Repeat-Containing Proteins Using Local Alignment Search Method. *Journal of Sciences,* **30(2):**105-111. Doi.10.22059/jsciences.2019.271719.1007349.
17. Kampke, T., Kieninger, M. and Mecklenburg, M. 2001. Efficient primer design algorithms. *Bioinformatics.* **17(3):** 214-25. Doi.10.1093/bioinformatics/17.3.214. PMID: 11294787.
18. Kim, C. Y., Takahashi, K., Nguyen, T. B., Roberts, J. K., and Webster, C. 1999. Identification of a nucleic acid binding domain in eukaryotic initiation factor eIF*iso*4G from wheat. *J. Biol. Chem.,* 274: 10603-10608.
19. Klosterman, S. J., Atallah, Z. K., Vallad, G. E., Subbarao, K. V. 2009. Diversity, pathogenicity and management of Verticillium species. *Annu Rev Phytopathol.,* **47:** 39-62.
20. Lamphear, B. J., Yan, R., Yang, F., Waters, D., Liebig, H. D. and Klump, H. 1993. Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of the 2A proteases from human Coxsackievirus and rhinovirus. *J. Biol. Chem.,* **268:** 19200-19203.
21. Mangeon, A., Junqueira, R.M. and Sachetto-Martins, G. 2010. Functional diversity of the Plant glycine-rich proteins superfamily. *Plant Signal. Behav.,* **5:** 99-104.

22. Morino, S., Imataka, H., Svitkin, Y. V., Pestova, T. V. and Sonenberg, N. 2000. Eukaryotic translation initiation factor 4E (eIF4E) binding site and the middle one-third of eIF4GI constitute the core domain for cap-dependent translation, and the C-terminal one-third functions as a modulatory region. *Mol. Cell Biol.,* **20:** 468-477.
23. Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.,* **22:** 459-472. Doi:10.1093/nar/10.2.459.
24. Muzafar, S., Sharma, R. D., Chauhan, N. and Prasad, R. 2021. Intron distribution and emerging role of alternative splicing in fungi. *FEMS Microbiol Lett.,* **26:** 368(19):fnab135. Doi: 10.1093/femsle/fnab135. PMID: 34718529.
25. Neff, C. L. and Sachs, A. B. 1999. Eukaryotic translation initiation factors 4G and 4A from *Saccharomyces cerevisiae* interact physically and functionally. *Mol. Cell Biol.,* **19:** 5557-5564.
26. Nishimura, T., Mochizuki, S., Ishii-Minami, N., Fujisawa, Y., Kawahara, Y., Yoshida, Y., Okada, K., Ando, S., Matsumura, H., Terauchi, R., Minami, E. and Nishizawa, Y. 2016. Magnaporthe oryzae glycine-rich secretion protein, Rbf1 critically participates in pathogenicity through the focal formation of the biotrophic interfacial complex*, PLoS Pathog,* **12:** e1005921.
27. Pegg, G. F. and Brady, B. L. 2002. Verticillium wilts. CAB International, Wallingford potato early dying. *Plant Dis.,* **84:** 1241-1245.
28. Parenteau, J., Maignon, L. and Berthoumieux, M. 2019. Introns are mediators of cell response to starvation. *Nature,* **565:** 612-7.
29. Rose T. M., Henikoff J. G. and Henikoff S. 2003. CODEHOP (Consensus-Degenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res.,* **31(13):** 3763-6. Doi: 10.1093/nar/gkg524. PMID: 12824413.
30. Scalon M. C, da Silva T. F, Aquino L. C, Carneiro F. T., Lima M. G., Lemos M. D. and Paludo G. R. 2014. Touchdown polymerase chain reaction detection of polycystic kidney disease and laboratory findings in different cat populations. *J Vet Diagn Invest.,* **26(4):** 542-546. Doi: 10.1177/1040638714536561.
31. Sharp, P. A. 1994. Split genes and RNA splicing. *Cell,* **77(6):** 805-15. Doi:10.1016/0092-8674(94)90130-9.
32. Sieber, P., Voigt, K. and Kammer, P. 2018. Comparative study on alternative splicing in human fungal pathogens suggests its involvement during host invasion. *Front Microbiol.,* **9**: 2313.

33. Stothard, P. 2000. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques,* **28(6):**1102-4.
34. Tamura, K., Dudley, J, Nei, M. and Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution,* **24:**1596-1599. Doi.10.1093/molbev/msm092.
35. Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research,* **22:** 4673-4680. Doi.10.1093/nar/22.22.4673.
36. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C. and Remm M. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res.,* **40(15):**e115.
37. Von der Haar, T., Gross, J. D., Wagner, G. and McCarthy, J. E. G. 2004. The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat. Struct. Mol. Biol.,* **11:** 503-511.
38. Yan, N., Addrah, M. E., Zhang, Y., Jia, R., Kang, L. and Zhao J. 2022. Identification and comparison of biological characteristics and pathogenicity of different mating types of *Verticillium dahliae* isolated from potato and sunflower. *Sci Rep.*, **12:**14164.

Journal of Agricultural Science and Technology (JAST), 28(6)
In Press, Pre-Proof Version

طراحی آغازگرهای دجنره برای افزوده سازی DNA ژنومی رمز گزاری کننده فاکتور شروع ترجمه یوکاریوتی4 جی *Verticillium dahliae* با استفاده از روش جستجوی همتراز موضعی

عباس جلودار

چکیده

Verticillium dahlia یک پاتوژن قارچی خاکی است که باعث بیماری پژمردگی عروقی در بسیاری از گیاهان زراعی و زینتی مهم اقتصادی در سراسر جهان می شود. در این مطالعه، یک قطعه DNA ژنومی که فاکتور شروع ترجمه یوکاریوتی گاما 4 را کد می‌کند از *V. dahliae* با استفاده از طراحی پرایمرهای دجنره بر اساس روش جستجوی همتراز موضعی افزوده سازی شد. برای شناسایی پروتئین‌های همولوگ در گونه‌های نزدیک، پایگاه داده های ژن NCBI با استفاده از برنامه HomoloGene و با استفاده از کلمه کلیدی هدف مورد بررسی قرار گرفت. بخش همتراز توالی های انتخاب شده دو موتیف محافظت را نشان داد که طراحی یک جفت آغازگر دجنره را تسهیل کرد. با استفاده از Touchdown PCR، یک قطعه DNA ژنومی با اندازه 503 نوکلئوتید از *V. dahliae* با موفقیت افزوده سازی و توالی یابی گردید که به عنوان *Vd*EIF4G تعیین شد. این قطعه حاوی دو قاب خوانش باز ترجمه (ORFs) بود که در مجموع 135 اسید آمینه را رمزگذاری می‌کرد، که توسط یک اینترون 97 نوکلئوتیدی از هم جدا می شدند. اتصال اینترون- اگزون در این ژن از قانون GU-AG پیروی می کرد. تجزیه و تحلیل طبقه‌بندی در سطح نوکلئوتیدی *Vd*EIF4G مشابهت با سه توالی موجود در جنس *Verticillium* نشان داد که 81٪ از مناطق همپوشانی را با شباهت توالی بین 97.99 تا 100٪ را پوشش می‌داد. تجزیه و تحلیل فیلوژنیک در سطح پروتئینی نشان داد که *Vd*EIF4G با *V. dahliae* یکسان بود اما، واگرایی از *V. nonalfalfae* و *V. alfalfae* را با کمترین فاصله ژنتیکی بین این دو توالی‌ها (1.8٪) نشان داد. علاوه بر این، حضور آمینو اسیدهای کوچک و قطبی، همراه با یک دامنه غنی از گلایسین، دخالت بالقوه این پروتئین را در تعاملات بین پاتوژن و میزبان نشان می دهد که ممکن است با حدت بیماریزایی این قارچ در ارتباط باشد. در این مطالعه، آغازگرهای دجنره طراحی شده با استفاده از روش جستجوی همتراز موضعی انجام شدند که به طور موثری توالی های همولوگ پروتئین حاوی دامنه eILF4G مربوطه را از *V. dahlia* افزوده سازی نمودند.