Check for updates

Genome-wide identification of the opsin protein in *Leptosphaeria maculans* and comparison with other fungi (pathogens of *Brassica napus*)

Marzieh Mohri[1]*, Ali Moghadam[2]*, Lenka Burketova[3] and Pavel Ryšánek[1]

[1]Department of Plant Protection, Faculty of Agrobiology, Food, and Natural Resources, Czech University of Life Sciences, Prague, Czechia, [2]Institute of Biotechnology, Shiraz University, Shiraz, Iran, [3]Institute of Experimental Botany, Czech Academy of Sciences, Prague, Czechia

The largest family of transmembrane receptors are G-protein-coupled receptors (GPCRs). These receptors respond to perceived environmental signals and infect their host plants. Family A of the GPCR includes opsin. However, there is little known about the roles of GPCRs in phytopathogenic fungi. We studied opsin in *Leptosphaeria maculans,* an important pathogen of oilseed rape (*Brassica napus*) that causes blackleg disease, and compared it with six other fungal pathogens of oilseed rape. A phylogenetic tree analysis of 31 isoforms of the opsin protein showed six major groups and six subgroups. All three opsin isoforms of *L. maculans* are grouped in the same clade in the phylogenetic tree. Physicochemical analysis revealed that all studied opsin proteins are stable and hydrophobic. Subcellular localization revealed that most isoforms were localized in the endoplasmic reticulum membrane except for several isoforms in *Verticillium* species, which were localized in the mitochondrial membrane. Most isoforms comprise two conserved domains. One conserved motif was observed across all isoforms, consisting of the BACTERIAL_OPSIN_1 domain, which has been hypothesized to have an identical sensory function. Most studied isoforms showed seven transmembrane helices, except for one isoform of *V. longisporum* and four isoforms of *Fusarium oxysporum*. Tertiary structure prediction displayed a conformational change in four isoforms of *F. oxysporum* that presumed differences in binding to other proteins and sensing signals, thereby resulting in various pathogenicity strategies. Protein−protein interactions and binding site analyses demonstrated a variety of numbers of ligands and pockets across all isoforms, ranging between 0 and 13 ligands and 4 and 10 pockets. According to the phylogenetic analysis in this study and considerable physiochemically and structurally differences of opsin proteins among all studied fungi hypothesized that this protein acts in the pathogenicity, growth, sporulation, and mating of these fungi differently.

1. Introduction

Fungi must recognize signals from various plant cells and tissues at very early and later stages in order to respond to these signals and infect the host plant. Guanine nucleotide-binding proteins (G-proteins) convey the received information cell surface receptors. G-proteins are composed of three subunits: (α (38–52 kDa), β (35 kDa), and γ (8–10 kDa)) that come together to create heterotrimeric proteins. Intracellular signaling linked to G-proteins is triggered when an external ligand binds to the receptor. G-protein-coupled receptors (GPCRs) are considered the largest family of cell surface receptors in eukaryotes and prokaryotes and possess seven transmembrane (TM) domains. GPCRs are associated with intracellular G-proteins. The recognition of various signal molecules or ligands from host plants is carried out through the signaling from the fungal G-protein coupled receptor (GPCR). After perception of the signal by GPCRs, GDP changes to GTP at the α subunit of the G protein. Therefore, it results in the reduction/activation of target effectors (e.g., adenylate cyclase, protein kinases, ion channels, and phospholipases) by a conformational change of receptors (Wall et al., 1995; Kulkarni et al., 2005; Terakita, 2005; Brown et al., 2018; El-Defrawy and Hesham, 2020; Gao et al., 2021). GPCRs have been identified bioinformatically in various fungi, and they are divided into six families: the rhodopsin family (A), the secretin-receptor family (B), the metabotropic glutamate receptor family (C), fungal pheromone P- and α-factor receptors (D), fungal pheromone A- and M-factor receptors (E), and cyclic-AMP receptors from *Dictyostelium* (F). Opsin belongs to family A of the GPCR (rhodopsin family) (Nadarajah et al., 2018; El-Defrawy and Hesham, 2020).

Signal transduction by GPCRs enables fungi to coordinate cell transport, metabolism, and growth, which in turn promotes their survival, reproduction, and virulence. In fungi, several GPCR-regulated signaling pathways, such as the cAMP-activated Protein Kinase A (PKA) pathway, the mitogen-activated protein kinases (MAPK) cascades pathway, and the phospholipase C (PLC) pathway, affect cell growth, morphogenesis, mating, stress responses, metabolism, and pathogenicity (Gao et al., 2021). For example, one of the important plant hormones is Indole-3-acetic acid (IAA), involved in different responses such as growth and effects on Car0-like rhodopsins (they were early called Opsin-related proteins (OPRs)-like rhodopsins). It indicates the association of Car0-like rhodopsins with plant-fungus interactions (Lyu et al., 2016).

Rhodopsins have been found in a wide range of organisms, such as Eukarya, Bacteria, and Archaea. They consist of opsin apoproteins and a covalently linked retinal that absorbs photons to convert energy or stimulate intracellular/intercellular signaling. The functions of microbial rhodopsin are the conversion of the light to the electrochemical potential for the energy of the cell which is called light-driven ion pumps and also to regulate the cell processes by using light. The light-driven proton pumps are presumed to function in the generation of pH gradients or pH homeostasis. They might act as the acidification of some cell compartments that could be involved in the activation of certain biochemical responses (Waschuk et al., 2005; Idnurm et al., 2010). Animal rhodopsins play

a variety of roles pertaining to vision and the sensation of light for nonvisual reasons such as circadian rhythms, sensing dawn/dusk, and body color change (Shichida and Matsuyama, 2009; Schmidt et al., 2011). Rhodopsins are divided into two groups: the first group is type II rhodopsins (animal rhodopsins), which consist of G-protein-coupled photoreceptors binding 11-*cis*-retinal and the auxiliary photoisomerases binding all-*trans*-retinal. The second group is type I rhodopsins (microbial rhodopsins), consisting of all-*trans*-retinal-binding transport and photosensory proteins of several organisms such as archaea, bacteria, and lower eukaryotes. Both type I and type II rhodopsins have the structure of seven TM α- helices consisting of an N-terminus outside the cell and a C-terminus inside the cell (Brown, 2004; Klare et al., 2008; Ernst et al., 2014).

Microbial opsins are transmembrane helix proteins in animals, archaea, fungi, and other eukaryotic microorganisms. These proteins belong to a class of retinal binding and play a role as light-responsive ion pumps or sensory receptors (Xue et al., 2008). Although the GPCRs play a role in several processes, such as pathogenicity, the genes have not been well studied in phytopathogenic fungi (Idnurm and Howlett, 2001; Terakita, 2005; Meena et al., 2018; Nadarajah et al., 2018). For example, there is little known about rhodopsins in one of the largest divisions of fungi, basidiomycetes. There are few studies regarding light responses in *Ustilago maydis* (a maize pathogen) belonging to basidiomycetes. There is a hypothesis that *Umops1* and *Umops2* are induced by white light in axenic cultures, and thus, the fungus synthesizes β-carotene (Estrada et al., 2009). Moreover, a few fungal opsins have been identified in *Neurospora crassa*, *Leptosphaeria maculans*, *Fusarium fujikuroi*, *Cryptococcus neoformans*, and *Botrytis cinerea,* but their functions are not yet known (Bieszke et al., 1999a; Idnurm and Howlett, 2001; Brown, 2004; Estrada and Avalos, 2009; Heller et al., 2012). Two GPCRs from class III in *Aspergillus* (GprC and GprD) are analogs to Gpr1 in *Saccharomyces cerevisiae* and have functions in sugar and oxylipin sensing. The deletion of these two genes (*gprC* or *gprD*) in *A. fumigatus* highlighted the fungus's inability to produce several toxic secondary metabolites. This deletion in turn affected the growth and pathogenicity in murine infections (Gehrke et al., 2010; de Souza et al., 2013). Another GPCR from class V (a common glucose and tryptophan sensor) is GprH, which has been found in several filamentous fungi and has a role in hyphal growth and sexual reproduction under carbon starvation in *A. nidulans* (Brown et al., 2015). The first opsin protein in fungi has been identified in *N. crassa* as Nop-1, an archaeal opsin homolog. The function of Nop-1 as rhodopsin in *N. crassa* has been revealed by gene and heterologous expression of this gene in *Pichia pastoris* (Bieszke et al., 1999a,b; Li et al., 2007). Transcriptomic analysis of *NOP-1* in wild-type and mutants has revealed that this gene acts in the late stage of conidiation. Its act is associated with the regulation of the expression of two genes involved in conidiation (*Al-2, NOC-10*, and *NOC-13*). Therefore, *NOP-1* is introduced as being involved in the modulation of carotenogenesis and repression of conidiation-specific gene expression in *N. crassa* (Bieszke et al., 2007). The study of the homolog gene *Sop1* in *Sclerotinia sclerotiorum* has revealed different functions of this gene, which is involved in the growth and virulence of *S. sclerotiorum* (Lyu et al., 2016).

The second most important oilseed crop is oilseed rape in the world. The global oilseed rape production is 31.8 million metric tons in 2022/2023.[1] Many pests and diseases associated with *B. napus* include the crucifer flea beetle, cabbage stem flea beetle, clubroot, sclerotinia stem rot, and blackleg (Walker and Booth, 2001). Fungi represent the most serious pathogens in oilseed rape. Ascomycetes, *Leptosphaeria* spp., cause significant yield losses in all cultivation areas worldwide. *L. maculans* causes so-called blackleg disease, which is demonstrated by leaf and stem lesions (Fitt et al., 2006).

Few studies have comprehensively analyzed the structure and characteristics of microbial opsins in fungi. Microbial opsins are involved in responses to different environmental signals, such as sensory receptors. For example, light sensing influences fungal development, especially sporulation and mating (Idnurm and Heitman, 2005). Therefore, the identification of these proteins and understanding their functions in fungi could be useful in different areas of plant pathology and plant breeding to provide new resistance genotypes against these proteins. Therefore, this study aimed to identify the potential opsins in *L. maculans* and six other fungal pathogens of oilseed rape, as well as analyze their characters at the bioinformatics level that might have similar functions.

Our study provides a comprehensive analysis of opsin in *L. maculans* and six other fungi: *Alternaria alternata*, *S. sclerotiorum*, *B. cinerea*, *Verticillium dahliae*, *V. longisporum*, and *F. oxysporum,* which cause diseases in oilseed rape. To achieve this goal, we used bioinformatics pipelines for sequence alignment, domain structure classification, transmembrane helices analysis, subcellular location prediction, conserved domains and motifs, the prediction of secondary and tertiary structures, and protein-protein interactions. These results provide a crucial understanding of the opsin protein in seven fungal pathogens of oilseed rape and identify candidate genes for fungal pathogenicity, growth, and sporulation.

2. Results

2.1. Domain identification of opsin proteins from different fungi

The proteins obtained from BLASTP were analyzed for the presence of conserved domains using the PROSITE database (https://prosite.expasy.org/scanprosite/). According to this database, bacterial rhodopsins were specifically detected, and they showed two patterns: the first pattern illustrated the isoforms containing BACTERIAL_OPSIN_1, and the second pattern indicated BACTERIAL_OPSIN_RET. Two of the above-mentioned patterns were found in two isoforms of *L. maculans* (LMO-Q9HGT7 and LMSO-E4ZUL6), three isoforms of *A. alternata* (AAPOL-OWY42043.1, AAPOL-A0A177E1U0, and AAPOL-XP_018391147.1), *S. sclerotiorum*, and *B. cinerea*. The BACTERIAL_OPSIN_1 domain was found in one isoform of *L. maculans* (LMCAO-7BMH), one isoform of *A. alternata* (AAO1-A0A177E306), all isoforms of *V. dahliae* and *V. longisporum*, and nine isoforms of *F. oxysporum* (FOO1-RKK65588.1, FOCO1-KAG7002927.1, FOHP-RKK90771.1,

FOHP-RKL21281.1, FOHPC-EGU75234.1, FOHPC-EGU78064.1, FOHPC-EGU79527.1, FOHPC-KAF6515179.1, and FOHPC-KAF6524808.1). Two isoforms of *F. oxysporum* contain the BACTERIAL_OPSIN_RET domain (FOHPF-QKD57451.1 and FOHPF-EWZ29335.1). Two other isoforms of *F. oxysporum* (FOHP-RKK62641.1 and FOHP-RKK62984.1) showed that Bac rhodopsin conserved domain in the Pfam database (http://pfam.xfam.org/) (Figure 1).

2.2. Characterization of the opsin protein

Several physicochemical and biochemical parameters of the opsin protein in *L. maculans* and six other fungal species were calculated using the ProtParam tool of the ExPASy. They are listed in Table 1. These parameters indicated the variation of the length of the opsin protein from 190 aa (VLHP-CRK17520.1) to 334 aa (FOHPF-EWZ29335.1). The lowest molecular weight was also found in *V. longisporum* species (VLHP-CRK17520.1), at 21166.72 *kDa,* and the highest one in *F. oxysporum* (FOHPF-EWZ29335.1), at 36625.13 *kDa*. The protein charge stability was investigated using isoelectric points (pI) (Gasteiger et al., 2005). The range of isoelectric points varies from 5.54 in *B. cinerea* to 9.08 in *F. oxysporum*. The instability index value was calculated for the experimental evaluation of a protein's stability. When the instability index values are lower than 40, the protein is predicted to be stable (Guruprasad et al., 1990). The calculation of the instability index revealed that all proteins were stable experimentally. The aliphatic index of a protein is the relative volume that is occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine) (Ikai, 1980; Gasteiger et al., 2005). Three isoforms of *A. alternata* (AAPOL-OWY42043.1, AAPOL-A0A177E1U0, and AAPOL-XP_018391147.1) have a lower aliphatic index, whereas other isoforms have the highest thermostability, ranging from 100.81 to 125.47. To calculate the grand average of hydropathicity (GRAVY) of a peptide or protein, the sum of hydropathy values of all of the amino acids was divided by the number of residues in the sequence (Kyte and Doolittle, 1982; Gasteiger et al., 2005). The GRAVY analysis revealed that all isoforms of the opsin protein were hydrophobic, with a GRAVY value above 0.

The amino acid compositions of opsin proteins across species were compared using the heatmap package of the R program. This analysis shows that leucine, glycine, valine, isoleucine, threonine, and alanine have a greater portion of this protein than other amino acids (Figure 2). A comparison of the percentage amino acid composition of opsin proteins across all isoforms in terms of polar, nonpolar, positively and negatively charged amino acids showed that only one isoform of *A. alternata*, and all isoforms of *S. sclerotiorum, B. cinerea*, and *L. maculans* were enriched in nonpolar amino acids such as, alanine, valine and leucine, and polar amino acids such as, threonine. Three other isoforms of *A. alternata* have a high percentage of alanine, valine, and leucine from the non-polar amino acid group, along with aspartic acid from the negatively charged amino acid group.

The complex of two species of *Verticillium* (*V. dahliae* and *V. longisporum*) indicated the highest percentage of non-polar amino acids, including alanine, valine, leucine, glycine, and serine from the polar amino acid group. The percentage of amino

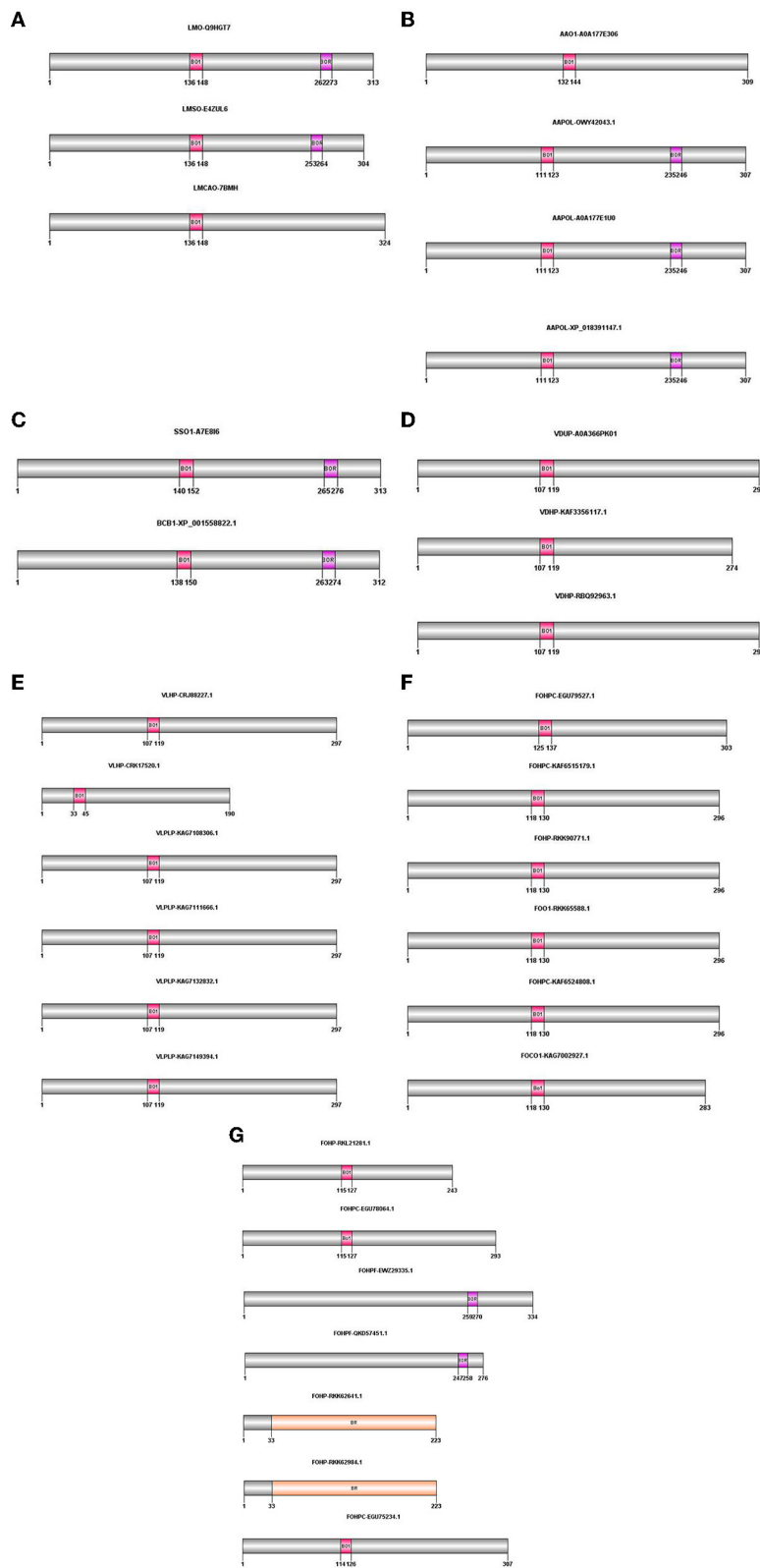---

FIGURE 1

Domain architecture of opsin protein in all species using PROSITE, Pfam; The positions of these domains were drawn using Domain Graph (DOG) software. **(A)** *L. maculans*, **(B)** *A. alternata*, **(C)** *S. sclerotiorum, B. cinerea*, **(D)** *V. dahliae*, **(E)** *V. longisporum*, **(F, G)** *F. oxysporum*. BO1, BACTERIAL_OPSIN_1; BOR, BACTERIAL_OPSIN_RET; BR, Bac rhodopsin.

TABLE 1 Physicochemical parameters of opsin proteins of all isoforms using the ProtParam tool of the ExPASy.

Name	Number of aa	Molecular weight	pI	Instability index	Aliphatic index	Gravy index
LMO-Q9HGT7	313	34,273.93	5.88	22.37	114.38	0.506
LMSO-E4ZUL6	304	33,365.87	5.71	23.7	113.59	0.507
LMCAO-7BMH	324	35,649.51	6.25	21.89	110.49	0.408
AAO1-A0A177E306	309	33,753.27	5.62	26.28	105.79	0.442
AAPOL-OWY42043.1	307	34,417.76	6.22	25.58	94.72	0.25
AAPOL-A0A177E1U0	307	34,371.73	6.04	26.8	95.37	0.27
AAPOL-XP_018391147.1	307	34,371.73	6.04	26.8	95.37	0.27
SSO1-A7E8I6	313	34,485.83	5.74	29.86	105.69	0.438
BCB1-XP_001558822.1	312	34,323.01	5.54	23.97	114.39	0.548
VDUP-A0A366PK01	297	32,637.57	6.33	35.24	100.81	0.354
VDHP-KAF3356117.1	274	30,341.16	8.33	39.2	101.46	0.438
VDHP-RBQ92963.1	297	32,637.57	6.33	35.24	100.81	0.354
VLHP-CRJ88227.1	297	32,664.64	6.33	33.53	102.12	0.373
VLHP-CRK17520.1	190	21,166.72	5.56	34.22	117.47	0.675
VLPLP-KAG7108306.1	297	32,593.56	6.48	34.95	101.14	0.372
VLPLP-KAG7111666.1	297	32,621.57	6.33	35.24	101.14	0.363
VLPLP-KAG7132832.1	297	32,562.49	6.48	32.89	102.12	0.375
VLPLP-KAG7149394.1	297	32,531.48	6.48	32.89	102.12	0.381
FOO1-RKK65588.1	296	32,354.52	8.83	16.65	111.35	0.494
FOCO1-KAG7002927.1	283	30,868.78	9.08	17.06	108.55	0.456
FOHP-RKK62641.1	223	24,313.78	8.33	27.27	120.72	0.915
FOHP-RKK62984.1	223	24,353.07	9.06	30.68	122.91	0.987
FOHP-RKK90771.1	296	32,349.51	8.83	20.74	110.71	0.486
FOHP-RKL21281.1	243	26,344.3	6.53	15.55	101.11	0.4
FOHPF-EWZ29335.1	334	36,625.13	8.19	41.64	117.69	0.647
FOHPF-QKD57451.1	276	30,123.73	8.58	28.22	125.47	0.905
FOHPC-EGU75234.1	307	34,037.45	7.71	21.74	102.35	0.314
FOHPC-EGU78064.1	293	31,886.93	8.76	19.89	107.51	0.511
FOHPC-EGU79527.1	303	33,144.41	8.82	13.91	111.02	0.491
FOHPC-KAF6515179.1	296	32,384.59	8.83	15.12	112.67	0.512
FOHPC-KAF6524808.1	296	32,328.44	8.83	16.07	109.7	0.475

acids in *F. oxysporum* is divided into three groups: the first group includes FOHPC-EGU79527.1, FOHPC-KAF6515179.1, FOHP-RKK90771.1, FOO1-RKK65588.1, FOHPC-KAF6524808.1, FOCO1-KAG7002927.1, FOHP-RKL21281.1, and FOHPC-EGU78064.1 and is rich in non-polar amino acids, including alanine, valine, isoleucine, glycine, and threonine from a polar group of amino acids. The second group of *F. oxysporum* (FOHPF-EWZ29335.1, FOHPF-QKD57451.1, FOHP-RKK62641.1, and FOHP-RKK62984.1) has a high percentage of non-polar amino acids, including alanine, valine, leucine, and isoleucine, and polar amino acids, including threonine and serine. The last isoform of *F. oxysporum* (FOHPC-EGU75234.1) presents a high percentage of alanine, valine, and leucine from non-polar

amino acids, and threonine from polar amino acids, arginine as a positive amino acid, and aspartic acid as a negative amino acid (Supplementary Figure S1). The coevolutionary amino acid analysis revealed that tryptophan had been conserved among all species (Supplementary Figure S2).

2.3. Subcellular localization of the opsin protein in the different fungal pathogens of *B. napus*

Subcellular localization of the opsin protein of all isoforms was conducted using Loctree3 and illustrated that the
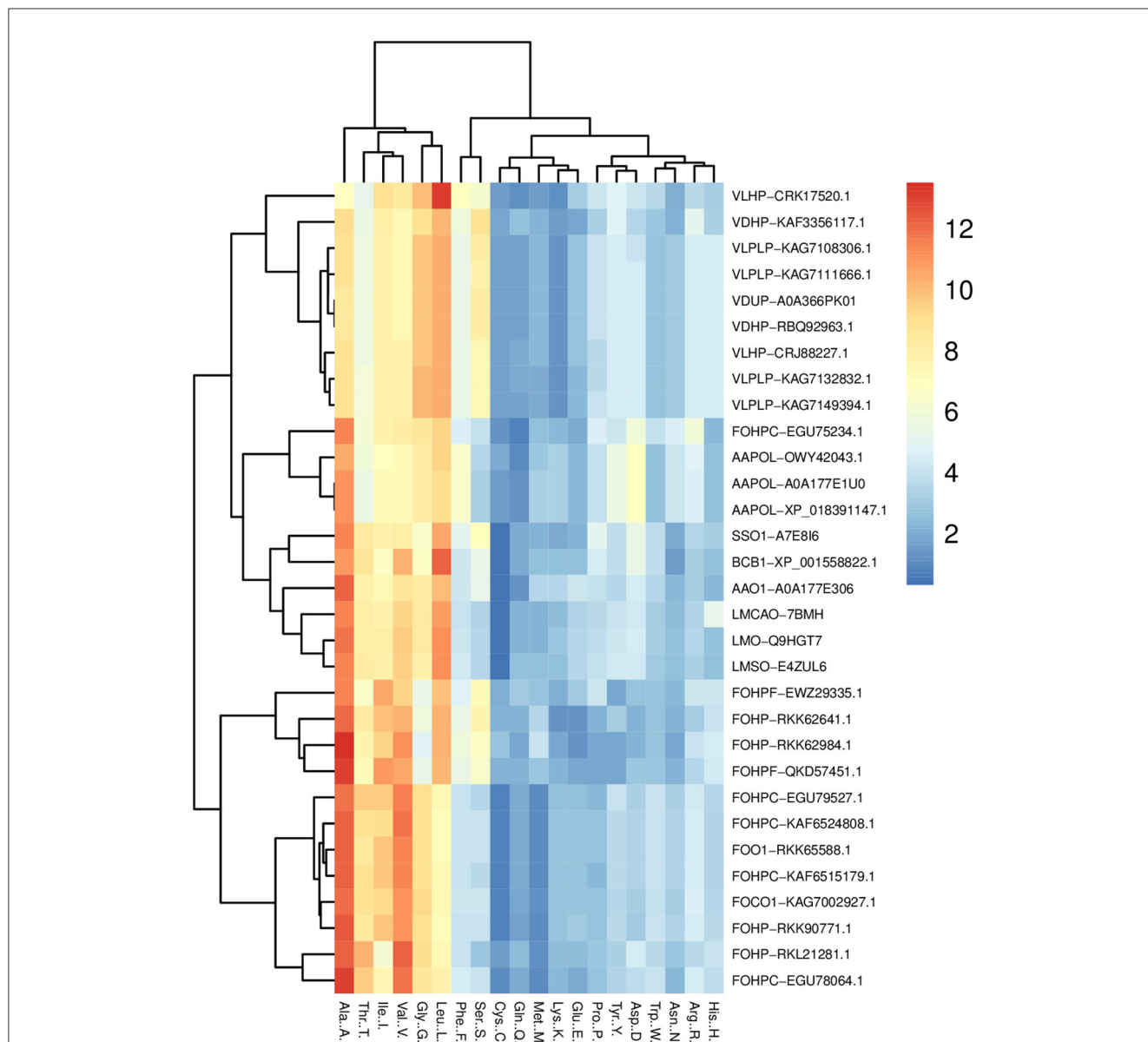
FIGURE 2
Heat map of the percentage comparison of the amino acids of opsin protein across all species (*L. maculans*, *A. alternata*, *S. sclerotiorum*, *B. cinerea*, *V. dahliae*, *V. longisporum*, and *F. oxysporum*) using R program. It illustrates the high percentage of six amino acids (leucine, glycine, valine, isoleucine, threonine, and alanine) in the opsin protein of these fungi.

majority of these proteins were localized in the endoplasmic reticulum membrane except for the two isoforms of *V. dahliae* (VDUP-A0A366PK01 and VDHP-RBQ92963.1) and five isoforms of *V. longisporum* (VLHP-CRJ88227.1, VLPLP-KAG7108306.1, VLPLP-KAG7111666.1, VLPLP-KAG7132832.1, and VLPLP-KAG7149394.1), which were localized in the mitochondrial membrane.

The other subcellular compartment is the nucleus. Proteins are synthesized in the cytoplasm, and they may be transported to the nucleus if they are needed there. There are binding sites on the protein sequence, known as nuclear-localized signals (NLSs), that mediate the protein's transportation to the nucleus (Brameier et al., 2007; Lu et al., 2021). Nuclear-localized signal analysis was conducted using the NucPred program. This analysis

revealed that three isoforms of *A. alternata* (AAPOL-OWY42043.1, AAPOL-A0A177E1U0, and AAPOL-XP_018391147.1) and all isoforms of *Verticillium*, except for isoform VLHP-CRK17520.1, possessed NLS (Table 2). Transmembrane domain analysis was conducted using the PolyPhobius program. This analysis illustrates that most isoforms possess seven transmembrane domains, except for VLHP-CRK17520.1 with five TMs, FOHP-RKK62641.1, FOHP-RKK62984.1, and FOHP-RKL21281.1 that possess six TMs, and FOHPF-QKD57451.1 with eight TMs (Table 2; Supplementary Figure S3). This analysis shows that one sole isoform possesses a signal peptide (*B. cinerea*). Three isoforms of *A. alternata* (AAPOL-OWY42043.1, AAPOL-A0A177E1U0, and AAPOL-XP_018391147.1), *S. sclerotiorum*, and all isoforms of *Verticillium*, except for two isoforms, VDHP-KAF3356117.1

and VLHP-CRK17520.1, possess N-glycan. Four isoforms of *F. oxysporum* (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKL21281.1, and FOHPC-EGU78064.1) possess N-glycan, and FOHPC-EGU75234.1 shows two sites of N-glycan in its sequence.

2.4. Phylogenetic tree

An alignment file generated from multiple sequence alignment using the ClustalW tool in MEGAX, besides two outgroups, was used to construct a phylogenetic tree and analyze the evolutionary relationships of the opsin protein among different species (Figure 3). In the phylogenetic tree, 31 isoforms of the opsin protein from different fungal species were clustered into six major groups and six subgroups. Clades 1 and 2 included *F. oxysporum* and were divided into two subgroups. Two different species (*S. sclerotiorum* and *B. cinerea*) were grouped in the third clade. Clade 4 was divided into two subgroups: one subgroup consisted of one isoform of *A. alternata* (AAO1-A0A177E306), and three isoforms of *L. maculans* were grouped in the second subgroup of this clade. Clade 5 was also divided into two subgroups: the other isoforms of *A. alternata* were grouped in one subgroup of this clade, and one isoform of *F. oxysporum* (FOHPC-EGU75234.1) was in the other subgroup of this clade. Clade 6 consisted of a complex of two species of *Verticillium*: *V. dahliae* and *V. longisporum*.

A maximum parsimony tree was also created using the above-mentioned alignment file. The results showed identical main clades as the above-mentioned tree (Supplementary Figure S6).

2.5. Conserved motifs of opsin homologs

The conserved motifs present in fungal species were determined by subjecting the amino acid sequence of the opsin protein to the MEME program. This analysis represented the fluctuation of motifs in the opsin protein during evolution. One highly conserved motif (motif 1) was found in all species. Further investigation also revealed that the conserved domain BACTERIAL_OPSIN_1 was located within motif 1. The second conserved domain, BACTERIAL_OPSIN_RET, which was found in two isoforms of *L. maculans*, three isoforms of *A. alternata*, *S. sclerotiorum*, and *B. cinerea,* was located within motif 4, whereas this domain of two isoforms of *F. oxysporum* was not detected within these motifs (Figure 4).

2.6. Consensus sequences of conserved motifs

According to the conserved motifs and conserved domains analyses, two motifs (1 and 4) were shared in almost all isoforms, which illustrated the importance of these motifs in the opsin protein across all fungal species. Then, Weblogo 3 was used to construct the consensus sequences of these two motifs. Motif 1 was 50 aa long and was present across all species. Amino acid residue number 12 (tryptophan) was conserved in all species except

for four species of *F. oxysporum* (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKK62641.1, and FOHP-RKK62984.1). In contrast, amino acid residue number 24 (cysteine) was more conserved in these four isoforms of *F. oxysporum* as well as two other isoforms of *F. oxysporum* (FOHP-RKL21281.1, FOHPC-EGU78064.1) and two species of *Verticillium* (*V. dahliae* and *V. longisporum*) (Supplementary Figure S4).

Motif 4 was 40 aa long and was present in all isoforms, except for the two isoforms of *F. oxysporum* (FOHP-RKK62641.1 and FOHP-RKK62984.1) (Supplementary Figure S5). The amino acid residue number 13 (tyrosine) was entirely conserved in all isoforms.

2.7. Features of the secondary and tertiary structures and gene structure

The prediction of the secondary structure features was performed using Jpred4 software and Proteus2 software. According to this prediction, the lowest and highest percentages of helix content were detected in FOHP-RKL21281.1 (35 %) and FOHPF-QKD57451.1 (74 %), respectively. The percentage of helix content of other isoforms ranged from 43 % to 64 %. Most varieties of beta-sheet content and coil content were observed in *F. oxysporum*: isoform FOHP-RKK62641.1 did not possess any beta-sheets, whereas the highest percentage of beta-sheet belonged to isoform FOHP-RKL21281.1 (18 %); the highest content of coil also belonged to isoform FOHP-RKL21281.1 (48 %), and isoform FOHPF-QKD57451.1 had the lowest content of coil (25 %) (Table 3).

The tertiary structure of all isoforms was predicted using SWISS-MODEL software. The conformational changes were observed in four isoforms of *F. oxysporum* (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKK62641.1, and FOHP-RKK62984.1) (Supplementary Figure S3).

The analysis of exon-intron structure led to insight into the evolution of this gene. Introns were not identified across all isoforms.

2.8. The topography of the PPI networks, GO, and binding site analysis of the opsin protein

Biomolecular networks provide information regarding the molecular functions and interpretation of genomic variation. There are several biomolecular networks according to different purposes and scopes, including (1) networks of gene-gene regulatory events in transcription, networks of kinases/phosphatases, networks of metabolites and (2) networks of protein-protein interactions. Protein–protein interactions can occur under various physiological conditions and in different cell types. Proteins can also interact with each other indirectly, such as through a transcription factor that regulates gene expression and the production of another protein (Szklarczyk et al., 2019, 2021). The network of protein–protein interactions of the opsin protein of all isoforms was constructed using STRING software (version 11.5) (Figure 5).

TABLE 2 Subcellular localization (using Loctree3), nuclear-localized signal prediction (using NucPred), and transmembrane helices analysis (using the PolyPhobius) of opsin protein of all species.

Name	Subcellular localization	NucPred	Number of transmembrane helices
LMO-Q9HGT7	Endoplasmic reticulum membrane	-	7
LMSO-E4ZUL6	Endoplasmic reticulum membrane	-	7
LMCAO-7BMH	Endoplasmic reticulum membrane	-	7
AAO1-A0A177E306	Endoplasmic reticulum membrane	-	7
AAPOL-OWY42043.1	Endoplasmic reticulum membrane	**NLS**	7
AAPOL-A0A177E1U0	Endoplasmic reticulum membrane	**NLS**	7
AAPOL-XP_018391147.1	Endoplasmic reticulum membrane	**NLS**	7
SSO1-A7E8I6	Endoplasmic reticulum membrane	-	7
BCB1-XP_001558822.1	Endoplasmic reticulum membrane	-	7
VDUP-A0A366PK01	Mitochondrion membrane	**NLS**	7
VDHP-KAF3356117.1	Endoplasmic reticulum membrane	**NLS**	7
VDHP-RBQ92963.1	Mitochondrion membrane	**NLS**	7
VLHP-CRJ88227.1	Mitochondrion membrane	**NLS**	7
VLHP-CRK17520.1	Endoplasmic reticulum membrane	-	5
VLPLP-KAG7108306.1	Mitochondrion membrane	**NLS**	7
VLPLP-KAG7111666.1	Mitochondrion membrane	**NLS**	7
VLPLP-KAG7132832.1	Mitochondrion membrane	**NLS**	7
VLPLP-KAG7149394.1	Mitochondrion membrane	**NLS**	7
FOO1-RKK65588.1	Endoplasmic reticulum membrane	-	7
FOCO1-KAG7002927.1	Endoplasmic reticulum membrane	-	7
FOHP-RKK62641.1	Endoplasmic reticulum membrane	-	6
FOHP-RKK62984.1	Endoplasmic reticulum membrane	-	6
FOHP-RKK90771.1	Endoplasmic reticulum membrane	-	7
FOHP-RKL21281.1	Endoplasmic reticulum membrane	-	6
FOHPF-EWZ29335.1	Endoplasmic reticulum membrane	-	7
FOHPF-QKD57451.1	Endoplasmic reticulum membrane	-	8
FOHPC-EGU75234.1	Endoplasmic reticulum membrane	-	7
FOHPC-EGU78064.1	Endoplasmic reticulum membrane	-	7
FOHPC-EGU79527.1	Endoplasmic reticulum membrane	-	7
FOHPC-KAF6515179.1	Endoplasmic reticulum membrane	-	7
FOHPC-KAF6524808.1	Endoplasmic reticulum membrane	-	7

Gene ontology (GO) is introduced as a common language in many organisms. GO is used to investigate and describe the function of proteins. The combination of gene ontology assessment and protein–protein interaction analysis enables the prediction of functions for unknown proteins. The three categories of GO are as follows: (1) biological process, which demonstrates the contribution of genes or proteins to specific biological events; (2) molecular function, which highlights the biochemical activities of the protein, such as ligand binding ligands, and (3) cellular component, which identifies the active location of a protein within cells. There is still a lack of knowledge regarding each category of GO for several proteins (Ashburner et al., 2000; Deng et al., 2004; Jain and Bader, 2010). An

investigation into GO was conducted across all isoforms using STRING software (version 11.5) (Supplementary Table S1). Gene set enrichment analysis using STRING software showed no significant results.

The binding sites of the opsin protein were investigated using the P2Rank program. The total number of ligands in *L. maculans* ranged between 12 and 13, along with four pockets. Two ligands and six pockets were observed in three isoforms of *A. alternata*, and isoform AAO1-A0A177E306 of *A. alternata* possessed five ligands and six pockets. One identical ligand was observed in two species, *S. sclerotiorum,* and *B. cinerea*; they showed four and six pockets, respectively. *V. dahliae* and *V. longisporum* also possess a similar ligand with five pockets. The number of ligands and
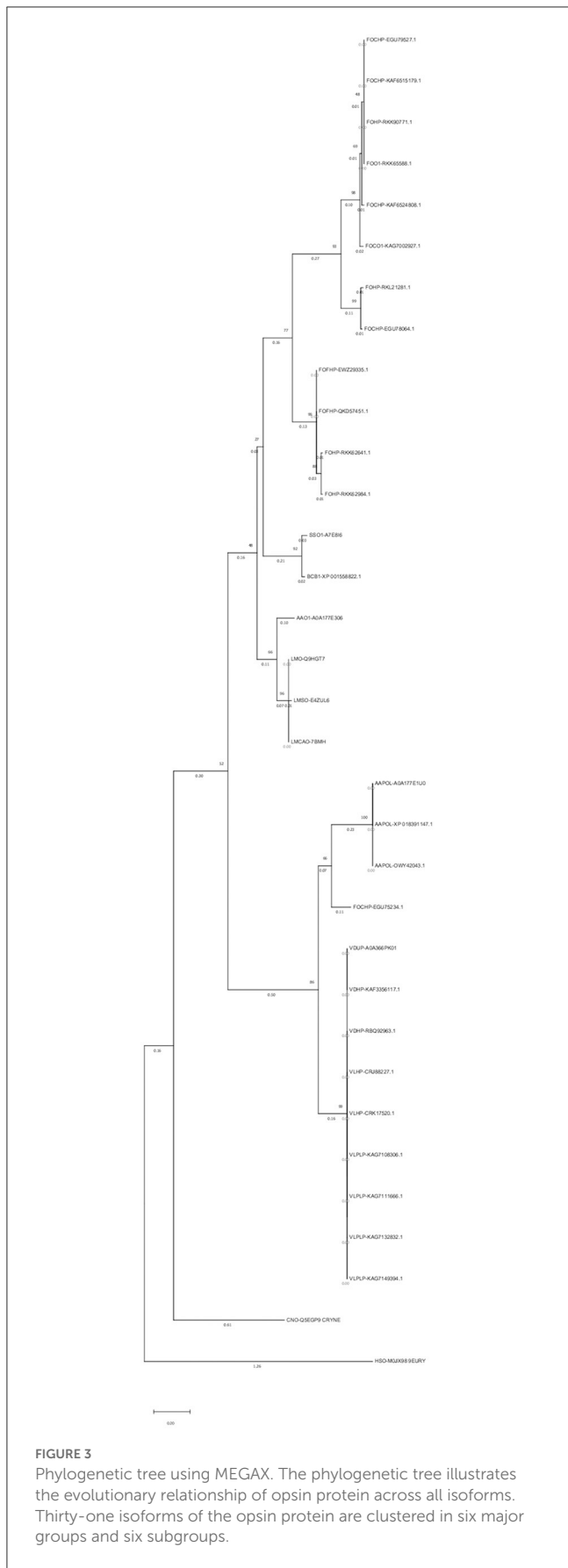
FIGURE 3
Phylogenetic tree using MEGAX. The phylogenetic tree illustrates the evolutionary relationship of opsin protein across all isoforms. Thirty-one isoforms of the opsin protein are clustered in six major groups and six subgroups.

pockets in *F. oxysporum* ranged between 0 and 3 and 4 and 10, respectively (Table 3).
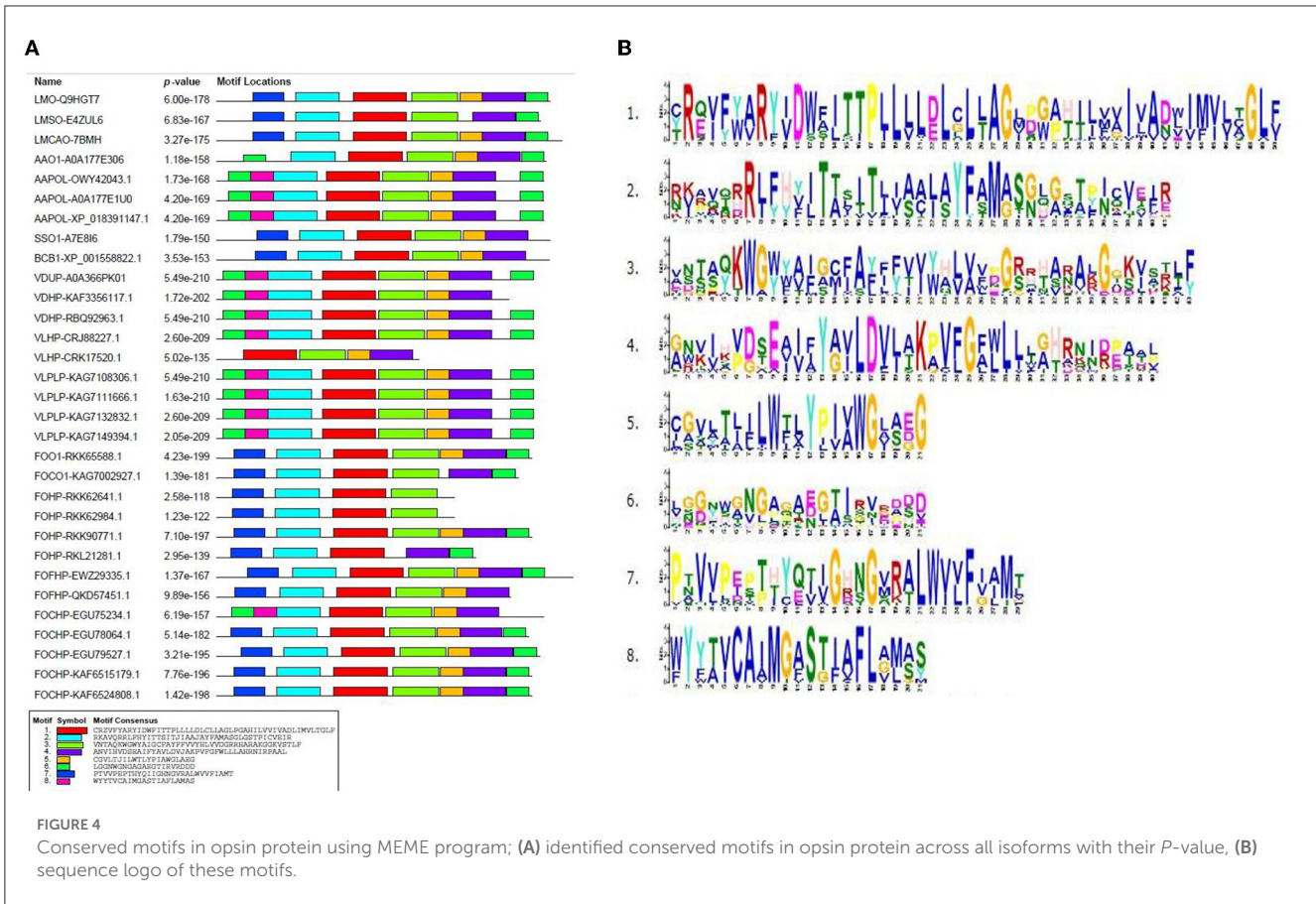
3. Discussion

3.1. Opsin characterization

The initial step in fungal-plant interactions is the recognition of the appropriate plant host. Various signals (chemical or physical signals) are perceived by fungi and in turn, they respond by differentiation, movement to an appropriate infection site, and formation of invasion-related structures (Kumamoto, 2008; Bonfante and Genre, 2010). For example, *Magnaporthe grisea*, the causative agent of rice blast, responds to contact with an appropriate surface by producing structures named appressoria. GPCR which is encoded by *PTH11* in this fungus, interacts with G proteins. Three subunits of G proteins (3 Gα subunits, 2 Gβ subunits, and one Gγ subunit) are encoded by the genome of *M. grisea*. And it has been reported that the formation of appressoria will be incomplete under the condition of deletion of each of these subunits (Nishimura et al., 2003; Kulkarni et al., 2005). In this research, opsin, a member of the rhodopsin family (the first family of GPCRs), of six phytopathogenic ascomycetous fungi (*L. maculans*, *A. alternata*, *S. sclerotiorum*, *B. cinerea*, *V. dahliae*, *V. longisporum*, and *F. oxysporum*) was studied.

Protein domains are evolutionarily conserved sequences with the basic components of protein structure and function (Murzin et al., 1995; Buljan et al., 2010). Evolutionary processes can lead to the emergence of new domains within proteins. Consequently, these new domains may contribute to the development of novel functions in organisms, potentially expanding their functional repertoire (Itoh et al., 2007). In comparison, fungi possess different numbers of domains than other organisms. Domain duplication during recombination causes the expansion of the domain in the fungal genome (Björklund et al., 2006). A study identified the bacteriorhodopsin domain in the microbial opsin homolog Sop1 in *S. sclerotiorum* (Lyu et al., 2016). Seven putative transmembrane helix domains have been reported in an opsin gene, *nop-1,* in *N. crassa* (Bieszke et al., 1999a). BACTERIAL_OPSIN_1 and BACTERIAL_OPSIN_RET are the two identified domains in this study. Most isoforms of studied fungi with the same domain/s were grouped in the same clade in the phylogenetic tree, which is hypothesized to have an identical biological function. Further analysis revealed that BACTERIAL_OPSIN_1 was located within the conserved motif of all species, indicating the importance of this conserved domain during evolutionary events. Bacterial opsins play a role in Light-dependent ion transport and sensory in a family of halophilic bacteria. These proteins are integral membrane proteins that contain seven transmembrane domains. The two classes of these proteins are as follows: (1) light-driven proton pumps, bacteriorhodopsin and archaerhodopsin and (2) a light-driven chloride pump, halorhodopsin (Oesterhelt, 1989; Soppa et al., 1993; Purschwitz et al., 2006).

The membrane composition influences the membrane protein's function and stability (Lee, 2004; Hunte and Richers, 2008;

FIGURE 4
Conserved motifs in opsin protein using MEME program; **(A)** identified conserved motifs in opsin protein across all isoforms with their *P*-value, **(B)** sequence logo of these motifs.

Phillips et al., 2009). Tryptophan is an amino acid that highly affects protein folding (Sanchez et al., 2008; Fiedler et al., 2010). Tryptophan is also involved in anchoring the hydrophobic transmembrane part across the membrane and causing hydrophobic mismatch, which highly affects the change of the transmembrane segment and the membrane structure (Sun et al., 2008). Transmembrane proteins also possess tryptophan amino acid near the end of the transmembrane helices. The above-mentioned functions of tryptophan are also reported in transmembrane proteins. Therefore, it is presumed that this amino acid could influence the function of opsin proteins as one of the transmembrane proteins (Von Heijne, 2007; De Jesus and Allen, 2013). Our study revealed that tryptophan was highly conserved across all isoforms (Supplementary Figure S2). Consensus sequence analysis of motif 1 demonstrated that tryptophan amino acid (residue 12) was highly conserved among all clades of the phylogenetic tree except for clade 2 (Supplementary Figure S4). According to the secondary and tertiary structure analysis of isoforms in clade 2 (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKK62641.1, and FOHP-RKK62984.1), the conformation of these isoforms has changed, confirming the importance of tryptophan regarding its effect on the structure of the protein membrane. Hence, it is hypothesized that these isoforms also have different functions.

3.1.1. Subcellular localization

Proteins are transported to the appropriate location to initiate their functions. Therefore, understanding protein localization provides important information concerning the function of the protein and its interaction with other proteins (Sadowski et al., 2008; Cosson et al., 2013; Peng and Gao, 2014). Subcellular localization prediction demonstrated that all isoforms were located in the endoplasmic reticulum membrane, except for seven isoforms of *Verticillium,* which were located in the mitochondrial membrane. The endoplasmic reticulum is a part of the secretory pathway while mitochondrion depends on another pathway (routes of lipid distribution). Therefore, it could be suggested that isoforms that possess the same subcellular localization are involved in the same cellular pathway (Jan et al., 2014; Acoba et al., 2020; Tamura et al., 2020).

The GPCRs are seven-transmembrane proteins. To initiate the intracellular protein transport of GPCRs, their insertion into the endoplasmic reticulum is required. For this purpose, two various pathways are used in membrane proteins with an extracellular N-terminal tail, such as GPCRs; the first group uses the uncleaved signal anchor sequence, which is the first transmembrane domain of the receptor. Moreover, the second subset possesses cleavable signal peptides. According to the

TABLE 3 The prediction of the secondary structure (helix content, beta-sheet content, coil content) (using Jpred4) and tertiary structure (ligands and pockets) (using SWISS-MODEL and P2Rank) of opsin protein of all isoforms.

Name	Helix content %	Beta sheet content %	Coil content %	Number of ligands	Number of pockets
LMO-Q9HGT7	55	6	39	11+ 2OLA	4
LMSO-E4ZUL6	54	8	38	10 LFA + 2 OLA	4
LMCAO-7BMH	56	5	40	-	-
AAO1-A0A177E306	54	10	36	5	6
AAPOL-OWY42043.1	51	11	38	2 (LFA7,14)	-
AAPOL-A0A177E1U0	49	10	41	2 (LFA7,14)	6
AAPOL-XP_018391147.1	50	10	40	2 (LFA7,14)	6
SSO1-A7E8I6	47	6	47	1 (LFA6)	4
BCB1-XP_001558822.1	56	11	34	1 (LFA6)	6
VDUP-A0A366PK01	51	10	39	1 (LFA7)	5
VDHP-KAF3356117.1	64	7	29	-	-
VDHP-RBQ92963.1	50	11	39	1 (LFA7)	5
VLHP-CRJ88227.1	51	10	38	1 (LFA7)	5
VLHP-CRK17520.1	59	15	26	1 (LFA7)	5
VLPLP-KAG7108306.1	51	10	39	1 (LFA7)	5
VLPLP-KAG7111666.1	51	11	39	1 (LFA7)	5
VLPLP-KAG7132832.1	51	10	39	1 (LFA7)	5
VLPLP-KAG7149394.1	51	10	38	1 (LFA7)	5
FOO1-RKK65588.1	55	11	33	2	10
FOCO1-KAG7002927.1	57	11	32	3 (LFA1,4,21)	10
FOHP-RKK62641.1	64	0	36	0	7
FOHP-RKK62984.1	60	4	36	0	9
FOHP-RKK90771.1	55	12	33	2 (LFA4, 21)	10
FOHP-RKL21281.1	35	18	48	0	9
FOHPF-EWZ29335.1	43	10	46	0	7
FOHPF-QKD57451.1	74	2	25	0	7
FOHPC-EGU75234.1	45	16	39	1 (LFA7)	4
FOHPC-EGU78064.1	44	12	44	0	8
FOHPC-EGU79527.1	52	10	38	2 (LFA4,21)	10
FOHPC-KAF6515179.1	56	7	36	2 (LFA4,21)	10
FOHPC-KAF6524808.1	55	11	34	2 (LFA4,21)	10

statistical analysis, the nature of the N-terminus of the GPCRs may correlate with the presence of the signal peptide. Possessing the signal peptides is required for the GPCRs, for which the post-translational translocation of their N-terminus is either impaired or impossible. Signal peptide prediction in this study shows that only one isoform (*B. cinerea*) possesses the signal peptide in its N-terminus, which may impact the function of this isoform. Moreover, the transmembrane topology of these isoforms displayed N-terminus tail differences among these isoforms, such as VLHP-CRK17520.1. Various N-terminus tails and numbers of the transmembrane have been found among the isoforms of *F. oxysporum*, and also the conformational change within

four isoforms of this species (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKK62641.1, and FOHP-RKK62984.1) has been found. It has been hypothesized that these varieties of the N-terminus, the number of transmembranes, and the conformational change might affect the function of these isoforms (Supplementary Figure S3).

3.1.2. Phylogenetic tree

Two of the most important pieces of information can be provided by reconstructing a phylogenetic tree: first, the evolutionary history of living organisms on earth; and second, the

evidence history of the climate and geological development of earth (Uncu et al., 2015). The evolutionary relationship among these fungi was investigated by creating a phylogenetic tree.

3.1.2.1. Maximum likelihood tree

ClustalW and bootstrap values of 1,000 iterations were used. Six major groups and six subgroups in the phylogenetic tree indicated divergent evolution among these isoforms. The phylogenetic tree also showed that most isoforms with identical domains were grouped in the same clade. However, proteins with the same domain can function differently due to mutations in their sequences (Cannon et al., 2004). This method helps determine the tree topology and branch lengths with the greatest likelihood. The likelihood is calculated for several models of nucleotide or amino acid substitution processes for each residue in an alignment (Felsenstein, 1981).

Most isoforms of *F. oxysporum* were grouped in the first and second clades; *S. sclerotiorum* and *B. cinerea* were grouped in the third clade; the fourth clade included *L. maculans* and one isoform of *A. alternata* (AAO1-A0A177E306); three other isoforms of *A. alternata* and one isoform of *F. oxysporum* (FOHPC-EGU75234.1) belonged to the fifth clade; the last clade included two species of *Verticillium* (*V. dahliae* and *V. longisporum*).

3.1.2.2. Maximum parsimony tree

A phylogenetic tree with the maximum parsimony method was created to precisely analyze these isoforms. The above-mentioned alignment file was also used for creating this phylogenetic tree. In comparison with the first phylogenetic tree, main clades were also observed in the second phylogenetic tree. Each clade was divided into several sub-groups in the second phylogenetic tree. It could be presumed that isoforms in these subgroups were evolutionary close to each other (Supplementary Figure S6).

3.1.3. Post-translational modifications

One of the post-translational modifications of proteins in the cell is glycosylation. Asparagine (N)-linked glycosylation of proteins is formed by the covalent binding of oligosaccharides onto asparagine residues of polypeptide chains (Koch, 1992; Landolt-Marticorena and Reithmeier, 1994; Kukuruzinska and Lennon, 1998; Shental-Bechor and Levy, 2008; Aebi et al., 2010; Schwarz and Aebi, 2011; Aebi, 2013). Translational N-glycosylation of membrane proteins occurs in the endoplasmic membrane and Golgi, where the processes of folding and maturation of membrane proteins take place and influences protein quality control (Chandler and Costello, 2016). Although glycans influence the above-mentioned functions, the role of glycans in these functions, such as folding proteins, is ambiguous. It has been reported that glycans function as chaperones and impact protein folding (Mitra et al., 2006; Shental-Bechor and Levy, 2008). The glycosylation assessment in this study illustrated that the isoforms with N-glycan sites were grouped in the same clade in the phylogenetic tree, suggesting that these isoforms might be involved in the same functions. There are several exceptions regarding N-glycosylation in the phylogenetic tree. *B. cinerea* possesses a peptide signal instead of N-glycosylation that might act differently. Two isoforms
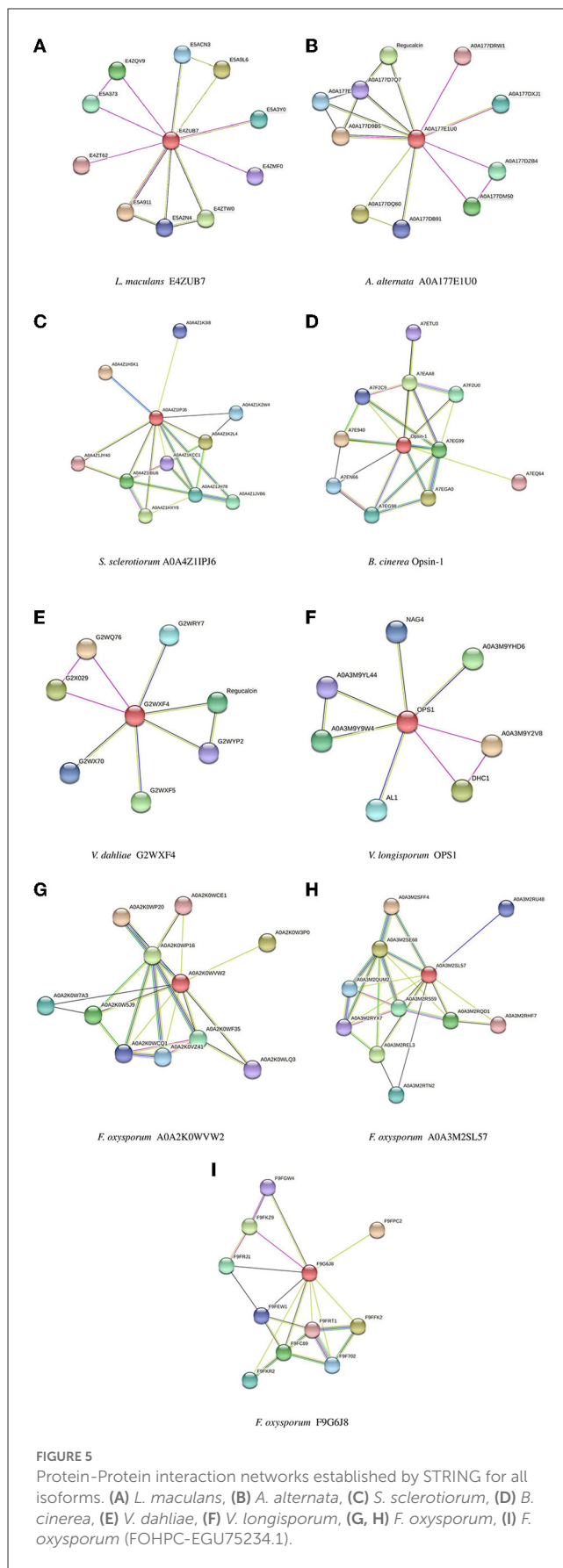

FIGURE 5

Protein-Protein interaction networks established by STRING for all isoforms. **(A)** *L. maculans*, **(B)** *A. alternata*, **(C)** *S. sclerotiorum*, **(D)** *B. cinerea*, **(E)** *V. dahliae*, **(F)** *V. longisporum*, **(G, H)** *F. oxysporum*, **(I)** *F. oxysporum* (FOHPC-EGU75234.1).

of *Verticillium* (VDHP-KAF3356117.1 and VLHP-CRK17520.1) do not possess any N-glycan sites while they are grouped with other isoforms of this fungus. Isoform FOHPC-EGU75234.1 is a subgroup of *A. alternata* and possesses two N-glycan sites. The location of the glycans affects the stabilization of proteins (Shental-Bechor and Levy, 2008). There is a hypothesis that these differences in N-glycosylation might influence the function, stabilization, and other characteristics of these fungi.

3.1.4. Protein structures

The primary structure, the linear amino acid sequence, and the description of the amino acid sequence in the polypeptide chain are the four main stages of protein structure characterization. The side chains of these amino acids can be polar, non-polar, or positive or negative charged, for example. These variations are crucial to the structure of proteins because they help retain a length of protein in a certain shape or conformation and connect side chains together. A polypeptide chain's secondary structure is its local spatial shape. The protein chain's adjacent amino groups and carboxyl groups form hydrogen bonds to form this structure. It is known as alpha helices and beta sheets. Tertiary structure is the three-dimensional arrangement of monomeric and multimeric protein molecules. It is also called a polypeptide. A quaternary structure is the three-dimensional structure of a macromolecule protein (Sun et al., 2004; Hartl and Hayer-Hartl, 2009). The secondary structure is made up of β-sheets and α-helices. The tertiary structure of a protein is created through interactions between the side chains of the amino acids. The tertiary structure of a protein is influenced by the polarity and charge of amino acid constituents. Every modification to a protein's fundamental structure and every instance of protein misfolding have an impact on the protein's structure. The functions of proteins are impacted by their structure. (Sun et al., 2004; Moghadam et al., 2016). The secondary and tertiary structures of all isoforms were predicted in this study. The percentage of α-helix content varies between 35% and 74 % among all isoforms, and the β-sheet content ranges between 0 and 18 % across all isoforms. A conformational change of the tertiary structure was observed in four isoforms of *F. oxysporum* (FOHPF-QKD57451.1, FOHPF-EWZ29335.1, FOHP-RKK62641.1, and FOHP-RKK62984.1).

Fungi have short introns and long exons compared to their mammalian counterparts (Kupfer et al., 2004). Different selections and mutations on introns affect the intron densities. Losing the amount of intron over a specific time could reflect general evolution (Lynch, 2002; Denoeud et al., 2010; Venkatesh et al., 2014). Gene structure analysis in this study did not show any introns in any of the studied isoforms. It could be due to an evolutionary event. A lower number of introns in a gene is interpreted as genetic evolution (Jacob and Smith, 2017; Naro and Sette, 2017).

3.1.5. Protein-protein networks

The most crucial information regarding the biological and cellular functions of proteins is provided by the prediction of protein-protein interactions. In a protein–protein interaction network, two or more proteins are bound together to accomplish their functions. It means that the function and activity of proteins

are often modulated by other connected proteins in the network of their protein-protein interactions (Phizicky and Fields, 1995). Protein E4ZUB7 from *L. maculans* and protein A0A177E1U0 from *A. alternata* showed an identical connection with RRM domain-containing protein. These proteins of *L. maculans* and *A. alternata,* along with proteins G2WXF4 and OPS1 from *V. dahliae* and *V. longisporum,* respectively, displayed similar connections with Dynein heavy chain, cytoplasmic. Protein E4ZUB7 from *L. maculans* and OPS1 from *V. longisporum* illustrated another identical connection with SGL domain-containing protein. Protein A0A177E1U0 from *A. alternata* and protein G2WXF4 from *V. dahliae* demonstrated a similar connection with Regucalcin. Proteins of two species of *Verticillium* (*V. dahliae* and *V. longisporum*) exhibit a similar connection with SBDS domain-containing protein in the network of their protein–protein interactions. Protein A0A4Z1IPJ6 from *S. sclerotiorum* and protein Opsin-1 from *B. cinerea,* and these two proteins (A0A2K0WVW2, A0A3M2SL57) from different isoforms of *F. oxysporum* exhibited two identical connections with two proteins, Amino_oxidase domain-containing protein and PAS domain-containing protein. These proteins, A0A4Z1IPJ6 from *S. sclerotiorum* and protein Opsin-1 from *B. cinerea,* and A0A2K0WVW2 from *F. oxysporum,* also demonstrated a connection with Photolyase/cryptochrome alpha/beta domain-containing protein. The protein of *S. sclerotiorum* (A0A4Z1IPJ6) illustrated one more similar connection, the same as two proteins (A0A2K0WVW2, A0A3M2SL57) from different isoforms of *F. oxysporum* with Bifunctional lycopene cyclase/phytoene synthase. These two proteins of *F. oxysporum* had the connection with Cryptochrome DASH. GO analysis of all isoforms illustrates that there is still a challenge in several studied isoforms whose functions are unknown (Supplementary Table S1). The gene set enrichment analysis did not show any significant results among all isoforms.

Endogenous peptides or protein ligands activate most GPCRs in humans. The different structures of the ligand-binding pocket between GPCRs and rhodopsin have illustrated that rhodopsins are not only bound to the ligands but also held in the pocket, which is required for the activation of rhodopsins by the photon-triggered isomerization of retinal, while the other GPCRs can be activated by binding to the ligands. The modes of peptide ligands vary in the receptors even within the same phylogenetic family, which indicates the specificity of their physiological function (Zhou et al., 2012; Kaiser and Coin, 2020). Our study reveal that most isoforms that are grouped in the same clade in the phylogenetic tree possess the same ligand and pocket.

3.2. Fungi characterization

L. maculans belongs to the Dothideomycetes class and the Pleosporales order. Several other important pathogens are in this order, such as those from the genus *Alternaria*. Pseudothecia production has been observed among most fungi of the Dothideomycetes class, forming either a single or a group of fruiting bodies. Fungi in the Dothideomycetes class have varieties of host plants, but their life cycles and pathogenicity strategies are often similar (Eyal, 1999; Nolan et al., 1999;

Kaczmarek and Jedryczka, 2012). *L. maculans* is introduced as a hemibiotrophic fungus. Several plants were suggested as hosts for *L. maculans* in the past, such as *Phaseolus* sp., *Humulus* sp., *Swertia perennis*, *Teucrium* sp., and *Artemisia campestris*. Afterward, reports highlighted crucifers and mainly *Brassica* crops as the hosts of *L. maculans* (Mendes-Pereira et al., 2003). Fungi belonging to the genus *Leptosphaeria* can cause allergic responses of types I and II in humans (Hasnain, 1993). The genus *Alternaria* from the Dothideomycetes class mainly includes saprophytic fungi; some species are plant pathogens. There is a broad host range, such as cereals, ornamentals, oil crops, vegetables (cauliflower, broccoli, carrot, potato, tomato), citrus, and apple. Several species of *Alternaria* cause cancer in mammals. Moreover, spores of this genus are also introduced as airborne allergens (Meena et al., 2010). According to the similarity among fungi of the Dothideomycetes class, it has been hypothesized that isoforms AAO1-A0A177E306 and isoforms of *L. maculans,* which are grouped in one clade in the phylogenetic tree, can play a similar role in sporulation and have the identical growth pathway. They could also be considered allergenic pathogens for humans. Although the other isoforms of *A. alternata* may have different sporulation pathways, they may also have a less allergic effect on humans.

Two necrotrophs and soilborne fungi, *S. sclerotiorum* (white mold fungus) and *B. cinerea* (gray mold fungus) are very close taxonomically. Their host range is broader than that of most other plant pathogens (Bolton et al., 2006; Williamson et al., 2007). Both fungi produce sclerotia that survive in the soil. They also produce melanized sclerotium, which is involved in their lifecycle by germinating vegetatively or sexually. *Botrytis* spp. also belongs to the Sclerotiniaceae family. Proteins encoded by the genomes of *B. cinerea* and *S. sclerotiorum* are 83% identical, which points to a very close relationship between these two genera (Groves and Loveland, 1953; Nees, 1973; Amselem et al., 2011). Despite many similarities in developmental and physiological features between these two fungi, they vary in their regulation and potential for sporulation. *S. sclerotiorum* produces ascospores and is dispersed via airborne spores, whereas *B. cinerea* produces ascospores, but its dispersal is predominantly via conidia. Furthermore, they are different in the regulation of sexual sporulation, with *S. sclerotiorum* being homothallic while *B. cinerea* is heterothallic (Bolton et al., 2006; Williamson et al., 2007). Most similarities between these two fungi confirm the phylogenetic analysis of this study, and the above-mentioned differences that have been found in this study suggest that the opsin protein may be involved in the melanized sclerotium pathway and can have an identical function in their growth and sporulation.

The widespread distribution of *Fusarium* species has been reported on plants, in soil, and in water as parasites, endophytes, or saprophytes. *Fusarium* causes several varieties of diseases in on-field, horticultural, ornamental, and forest crops, such as wilts, blights, rots, and cankers. Varieties of toxic secondary metabolites are produced by fusaria, such as trichothecenes and fumonisins, that are involved in the contamination of agricultural products; trichothecenes can also be a virulence factor in plant diseases. A few *Fusarium* species have been introduced as opportunistic human pathogens that cause corneal infections (O'Donnell et al., 2004). Life cycle, specialization, host adaptation, and specificity differ among *Fusarium* species. *F. oxysporum* has a broad host

range, including monocotyledonous and dicotyledonous plants. Moreover, it can also infect immunocompromised patients and other mammals (O'Donnell et al., 2004; Ortoneda et al., 2004). *F. oxysporum* is an anamorphic species. It is a heterogeneous and polytypic morphospecies that is introduced as the most abundant soilborne fungus (Waalwijk et al., 1996; O'Donnell and Cigelnik, 1997; Snyder and Hansen, 2013). Most of the *F. oxysporum* isoforms were grouped in the same clade in the phylogenetic analysis, which displays similarity within this forma. Despite these similarities, several subgroups have been observed in this phylogenetic tree analysis of the opsin protein, which points to the different performances of the opsin protein within this species, demonstrating different recognition and function strategies of this protein on different hosts. It has been hypothesized that one isoform of *F. oxysporum* (FOHPC-EGU75234.1), which was grouped with three isoforms of *A. alternata,* can not be an allergenic pathogen to humans.

The genus *Verticillium* involves soilborne plant pathogenic species that cause vascular wilt diseases on a wide range of plants, especially in the temperate and subtropical regions of the world (Fradin and Thomma, 2006). The *Verticillium* genus comprises two plant pathogenic species: *V. dahliae,* which forms microsclerotia. *V. albo-atrum,* which produces dark resting mycelium, and the other four species are saprophytes and non-pathogenic species (*V. nigrescens*, *V. nubilum*, *V. tricorpus*, and *V. theobromae*) (Barbara and Clewes, 2003). *V. dahliae* causes verticillium wilt in a wide range of host plants, including woody, dicotyledonous species of herbaceous annuals and perennials (Bhat and Subbarao, 1999). Subsequent studies have revealed the occurrence of parasexual hybridization between *V. albo-atrum* and *V. dahliae* to evolve *V. longisporum* (Karapapa et al., 1997; Zeise and Von Tiedemann, 2001). Several studies have illustrated the host specificity of *V. longisporum* in brassicaceous hosts, whereas *V. dahliae* is restricted to non-brassicaceous hosts (Zeise and Von Tiedemann, 2001, 2002a,b). The phylogenetic analysis of this study separated *Verticillium* species from other fungi, demonstrating higher similarity between these two species. These two species were separated within their clade in this phylogenetic analysis of the opsin protein, suggesting that this protein acts differently among these two species of the genus *Verticillium* concerning plant–pathogen interaction, pathogenicity, and sporulation.

4. Material and methods

4.1. Collection of opsin protein sequences

Protein sequences of opsins from *L. maculans*, *A. alternata*, *S. sclerotiorum*, *B. cinerea*, *V. dahliae*, *V. longisporum*, and *F. oxysporum* were retrieved from the NCBI and UniProt databases. Their gene IDs are listed in Supplementary Table S2. Then, BLASTP was conducted across the proteomes of the above-mentioned fungi using the opsin protein sequences. The identified opsin proteins were analyzed for the presence of conserved domains using the Pfam [a protein families database that identifies proteins by using two alignments and a profile hidden Markov model (HMM)] (Finn et al., 2014), SMART (a database that is used to identify and annotate the genetically mobile domains and

analysis of their architectures) (Schultz et al., 2000), ENSEMBLE (a comprehensive database that provides information on the annotation of individual genomes of many species such as plants and microorganisms) (Birney et al., 2004), Conserved Domain Database (CDD) (the annotation of protein sequences is provided by this resource along with the location of conserved domain footprints) (Marchler-Bauer et al., 2011), and PROSITE (two signatures (patterns and generalized profiles) are used in this database to identify conserved regions) (Hulo et al., 2006) databases. Therefore, these databases provided comprehensively conserved domain analyses by applying different alignment and annotation methods. The position and scale of protein domains were drawn using Domain Graph (DOG) software (http://dog.biocuckoo.org/).

4.2. Physicochemical properties of opsin homologs

Physicochemical properties, including isoelectric point (pI), molecular weight (Mw), instability index (II), grand average of hydropathicity (GRAVY), and aliphatic index (AI), were investigated for identified opsin proteins using the ProtParam tool available on the ExPASy website (https://web.expasy.org/protparam/) and EMBOSS Pepstats (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/). The coevolution of amino acids was investigated using MISTIC (Mutual Information Server to Infer Coevolution) (http://mistic.leloir.org.ar/index.php).

4.3. Subcellular localization and transmembrane domains

Subcellular localization of opsin proteins was predicted using Loctree3 (https://rostlab.org/services/loctree3/). Nuclear localization signal prediction was conducted using NucPred (https://nucpred.bioinfo.se/cgi-bin/single.cgi). The integral transmembrane (TM) domains were investigated using PolyPhobius (https://phobius.sbc.su.se/poly.html) and Protter (http://wlab.ethz.ch/protter/start/). The prediction of signal peptide cleavage sites was performed using TargetP-2.0 (https://services.healthtech.dtu.dk/services/TargetP-2.0/).

4.4. Phylogenetic analysis

Multiple sequence alignment of the opsin sequences of these fungal species was performed using the ClustalW tool in MEGAX. Then, the alignment file was used to create a phylogenetic tree alongside with Maximum Likelihood method (1000 repeats) with the parameters of the Jones-Taylor-Thornton (JTT) matrix-based model, uniform rates among sites, and partial deletion of gaps in MEGAX software. Opsin sequences of one species of yeast (*Cryptococcus neoformans*) and one species of archaea (*Haloarcula sinaiiensis*) were downloaded from the UniProt database to be used as outgroup in the phylogenetic tree. The alignment file was also used to generate a phylogenetic tree, alongside with Maximum

parsimony method in MEGAX. In the second phylogenetic tree the calculation of bootstrap values was also at 1000 iterations with the parameters of the Subtree-Pruning-Regrafting (SPR) method and partial deletion of gaps in MEGAX software. Two above-mentioned outgroups were also used in the second phylogenetic tree.

4.5. Identification of conserved motifs and sequence logo

Conserved motifs of opsin proteins across all fungal species were predicted using the MEME (https://meme-suite.org/meme/) tool with the maximum number of eight motifs, and the other parameters were set to default. The sequence logos of conserved motifs were generated using Weblogo 3 (http://weblogo.threeplusone.com/).

4.6. Protein and gene structures of opsin homologs

The features of the secondary structure were predicted using Jpred4 software (https://www.compbio.dundee.ac.uk/jpred/) and Proteus2 software (http://www.proteus2.ca/proteus2/), and the SWISS-MODEL software (https://swissmodel.expasy.org/) was used for the prediction of the fatures of the tertiary structure of opsin proteins in all fungal species. To investigate the exon and intron structures, blastx was performed at NCBI for each isoform.

4.7. Protein-protein interactions and gene ontology (GO) analyses

Protein–protein interactions (PPI) of the opsin protein in all fungal species were constructed using String software (https://string-db.org/). Afterward, constructed networks were investigated considering the GO annotation, including Biological Processes (BP), Molecular Functions (MF), and Cellular Component (CC). String software was also used for Gene set enrichment analysis across all isoforms. Binding sites, including ligand-binding pockets of the opsin protein, were predicted using P2Rank (https://prankweb.cz/).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

AM and MM: conception, design, and data curation of the study. AM: methodology, investigation, validation, software, supervision, review, and editing. MM: formal analysis, methodology, software, investigation, and writing. All authors

contributed to the manuscript revision and read and approved the submitted version.

Funding

This study was supported by a grant QK1710397 from the Ministry of Agriculture of the Czech Republic: Characterization of the compatibility of relations between agents that cause blackleg and oilseed rape varieties as a basis for increasing the growing rentability of this crop in the Czech Republic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1193892/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Comparison of the amino acid composition of opsin protein in all isoforms; (A) *L. maculans*, (B) *A. alternata*, (C) *S. sclerotiorum*, (D) *B. cinerea*, (E) *V. dahliae*, (F) *V. longisporum*, (G–I) *F. oxysporum*. The y-axis shows the percentage of amino acids of opsin protein in each isoform; the X-axis presents the name of all amino acids.

SUPPLEMENTARY FIGURE 2
The coevolutionary analysis of the amino acids of opsin protein across all isoforms using Mistic. Labels in the outer circle illustrate the amino acid codes of the opsin protein of all isoforms. Red boxes show the highly conserved positions and blue boxes indicate less conserved positions. Red arrows show highly conserved amino acid (Tryptophan).

SUPPLEMENTARY FIGURE 3
Transmembrane domains analysis of the opsin protein across all isoforms using Protter; (A) *L. maculans*, (B) *S. sclerotiorum*, (C) *B. cinerea*, (D) *A. alternata*, (E) *V. dahliae*, (F) *V. longisporum*, and (G) *F. oxysporum*. Red circles show N-glycan sites.

SUPPLEMENTARY FIGURE 4
Consensus sequences of motif 1 in opsin protein across all isoforms using Weblogo 3; (A) isoforms of *F. oxysporum* (FOHPC-EGU79527.1, FOHPC-KAF6515179.1, FOHP-RKK90771.1, FOO1-RKK65588.1, FOHPC-KAF6524808.1, FOCO1-KAG7002927.1); (B) isoforms of *F. oxysporum* (FOHP-RKL21281.1, FOHPC-EGU78064.1); (C) isoforms of *F. oxysporum* (FOHPF-QKD57451.1, FOHP-EWZ29335.1); (D) isoforms of *F. oxysporum* (FOHP-RKK62641.1 and FOHP-RKK62984.1); (E) *S. sclerotiorum*, and *B. cinerea*; (F) *L. maculans* and one isoform of *A. alternata* (AAO1-A0A177E306); (G) *A. alternata* and one isoform of *F. oxysporum* (FOHPC-EGU75234.1); (H) *V. dahliae* and *V. longisporum*.

SUPPLEMENTARY FIGURE 5
Consensus sequences of motif 4 in opsin protein across all isoforms using Weblogo 3; (A) isoforms of *F. oxysporum* (FOHPC-EGU79527.1, FOHPC-KAF6515179.1, FOHP-RKK90771.1, FOO1-RKK65588.1, FOHPC-KAF6524808.1, FOCO1-KAG7002927.1); (B) isoforms of *F. oxysporum* (FOHP-RKL21281.1, FOHPC-EGU78064.1); (C) isoforms of *F. oxysporum* (FOHPF-QKD57451.1, FOHPF-EWZ29335.1); (D) *S. sclerotiorum*, and *B. cinerea*; (E) *L. maculans* and one isoform of *A. alternata* (AAO1-A0A177E306); (F) *A. alternata* and one isoform of *F. oxysporum* (FOHPC-EGU75234.1); (G) *V. dahliae* and *V. longisporum*.

SUPPLEMENTARY FIGURE 6
Phylogenetic tree using MEGAX (Maximum Parsimony method). The maximum parsimony method presents a close evolutionary relationship of opsin protein across all isoforms.

References

Acoba, M. G., Senoo, N., and Claypool, S. M. (2020). Phospholipid ebb and flow makes mitochondria go. *J. Cell Biol.* 219, 1–21. doi: 10.1083/jcb.202003131

Aebi, M. (2013). N-linked protein glycosylation in the ER. *Mol. Cell Res.* 1833, 2430–2437. doi: 10.1016/j.bbamcr.2013.04.001

Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010). N-glycan structures: recognition and processing in the ER. *Trends Biochem.* Sci. 35, 74–82. doi: 10.1016/j.tibs.2009.10.001

Amselem, J., Cuomo, C. A., van Kan, J. A. L., Viaud, M., Benito, E. P., Couloux, A., et al. (2011). Genomic analysis of the necrotrophic fungal pathogens sclerotinia sclerotiorum and botrytis cinerea. *PLoS Genet.* 7, 8. doi: 10.1371/journal.pgen.1002230

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Mol. Breed.* 35, 25–29. doi: 10.1038/75556

Barbara, D. J., and Clewes, E. (2003). Plant pathogenic verticillium species: how many of them are there? *Mol. Plant Pathol.* 4, 297–305. doi: 10.1046/j.1364-3703.2003.00172.x

Bhat, R. G., and Subbarao, K. V. (1999). Host range specificity in Verticillium dahliae. *Phytopathology* 89, 1218–1225. doi: 10.1094/PHYTO.1999.89.12.1218

Bieszke, J. A., Braun, E. L., Bean, L. E., Kang, S., Natvig, D. O., and Borkovich, K. A. (1999a). The nop-1 gene of Neurospora crassa encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. *Proc. Natl. Acad. Sci. U.S.A.* 96, 8034–8039. doi: 10.1073/pnas.96.14.8034

Bieszke, J. A., Li, L., and Borkovich, K. A. (2007). The fungal opsin gene nop-1 is negatively-regulated by a component of the blue light sensing pathway and influences conidiation-specific gene expression in Neurospora crassa. *Curr. Genetics* 52,: 149–157. doi: 10.1007/s00294-007-0148-8

Bieszke, J. A., Spudich, E. N., and Spudich, J. L. (1999b). A eukaryotic protein, NOP-1, binds retinal to form an archaeal. *Biochemistry* 38, 14138. doi: 10.1021/bi9916170

Birney, E., Andrews, T. D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., et al. (2004). An overview of Ensembl. *Genome Res.* 14, 925–928. doi: 10.1101/gr.1860604

Björklund, Å. K., Ekman, D., and Elofsson, A. (2006). Expansion of protein domain repeats, *PLoS Comp. Biol.* 2, 0959–0970. doi: 10.1371/journal.pcbi.0020114

Bolton, M. D., Thomma, B. P., and Nelson, B. D. (2006). Sclerotinia sclerotiorum (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Mol. Plant Pathol.* 7, 1–16. doi: 10.1111/j.1364-3703.2005.00316.x

Bonfante, P., and Genre, A. (2010). Mechanisms underlying beneficial plant - Fungus interactions in mycorrhizal symbiosis. *Nat. Communicat.* 1, 4. doi: 10.1038/ncomms1046

Brameier, M., Krings, A., and MacCallum, R. M. (2007). NucPred– predicting nuclear localization of proteins. *Bioinformatics* 23, 1159–1160. doi: 10.1093/bioinformatics/btm066

Brown, L. S. (2004). Fungal rhodopsins and opsin-related proteins: eukaryotic homologues of bacteriorhodopsin with unknown

functions. *Photochem. Photobiol. Sci.* 3, 555–565. doi: 10.1039/b315527g

Brown, N. A., Dos Reis, T. F., Ries, L. N. A., Caldana, C., Mah, J-H., Yu, J-H., et al. (2015). G-protein coupled receptor-mediated nutrient sensing and developmental control in *Aspergillus nidulans. Mol. Microbio.* 98, 420–439. doi: 10.1111/mmi.13135

Brown, N. A., Schrevens, S., van Dijck, P., and Goldman, G. H. (2018). Fungal G-protein-coupled receptors: Mediators of pathogenesis and targets for disease control. Nat. *Microbiol.* 3, 402–414. doi: 10.1038/s41564-018-0127-5

Buljan, M., Frankish, A., and Bateman, A. (2010). Quantifying the mechanisms of domain gain in animal proteins. *Genome Biol.* 11, 7. doi: 10.1186/gb-2010-11-7-r74

Cannon, S. B., Mitra, A., Baumgarten, A., Young, N., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biol.* 4, 1–21. doi: 10.1186/1471-2229-4-10

Chandler, K. B., and Costello, C. E. (2016). Glycomics and glycoproteomics of membrane proteins and cell-surface receptors: Present trends and future opportunities. Electrophoresis 37, 1407–1419. doi: 10.1002/elps.201500552

Cosson, P., Perrin, J., and Bonifacino, J. S. (2013). Anchors aweigh: Protein localization and transport mediated by transmembrane domains, Trends in Cell *Biology.* 23: 511–517. doi: 10.1016/j.tcb.2013.05.005

De Jesus, A. J., and Allen, T. W. (2013). The role of tryptophan side chains in membrane protein anchoring and hydrophobic mismatch. *Biomembranes.* 1828, 864–876. doi: 10.1016/j.bbamem.2012.09.009

de Souza, W. R., Morais, E. R., Krohn, N. G., Savoldi, M., Goldman, M. H. S., Rodrigues, F., et al. (2013). Identification of metabolic pathways influenced by the g-protein coupled receptors GprB and GprD in *Aspergillus nidulans. PLoS ONE* 8, 1–13. doi: 10.1371/journal.pone.0062088

Deng, M., Tu, Z., Sun, F., and Chen, T. (2004). Mapping gene ontology to proteins based on protein-protein interaction data. *Bioinformatics* 20, 895–902. doi: 10.1093/bioinformatics/btg500

Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J-M., Da Silva, C., Brinkmann, H., et al. (2010). Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science* 330, 1381–1385. doi: 10.1126/science.1194167

El-Defrawy, M. M. H., and Hesham, A. E.-L. (2020). G-protein-coupled receptors in fungi. *Fungal Biol.* 2020, 37–126. doi: 10.1007/978-3-030-41870-0_3

Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S., and Kandori, H. (2014). Microbial and animal rhodopsins. *Chem. Rev.* 114, 126–163. doi: 10.1021/cr4003769

Estrada, A. F., and Avalos, J. (2009). Regulation and targeted mutation of opsA, coding for the NOP-1 Opsin orthologue in Fusarium fujikuroi. *J. Mol. Biol.* 387, 59–73. doi: 10.1016/j.jmb.2009.01.057

Estrada, A. F., Brefort, T., Mengel, C., Díaz-Sánchez, V., Alder, A., Al-Babili, S., et al. (2009). Ustilago maydis accumulates β-carotene at levels determined by a retinal-forming carotenoid oxygenase, Fungal Genetics and *Biology.* 46: 803–813. doi: 10.1016/j.fgb.2009.06.011

Eyal, Z. (1999). The septoria tritici and stagonospora nodorum blotch diseases of wheat. *Eur. J. Plant Pathol.* 105, 629–641. doi: 10.1023/A:1008716812259

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evolution.* 17, 368–376. doi: 10.1007/BF01734359

Fiedler, S., Broecker, J., and Keller, S. (2010). Protein folding in membranes. *Cell. Mol. Life Sci.* 67, 1779–1798. doi: 10.1007/s00018-010-0259-0

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: The protein families database. *Nucleic Acids Res.* 42, 222–230. doi: 10.1093/nar/gkt1223

Fitt, B. D. L., Brun, H., Barbetti, M., and Rimmer, S. (2006). World-wide importance of phoma stem canker (Leptosphaeria maculans and L. *biglobosa*) on oilseed rape (Brassica napus), *Eur. J. Plant Pathol.* 114, 3–15. doi: 10.1007/s10658-005-2233-5

Fradin, E. F., and Thomma, B. P. H. J. (2006). Physiology and molecular aspects of Verticillium wilt diseases caused by V. *dahliae and V. albo-atrum. Mol. Plant Pathol.* 7, 71–86. doi: 10.1111/j.1364-3703.2006.00323.x

Gao, J., Xu, X., Huang, K., and Liang, Z. (2021). Fungal G-protein-coupled receptors: a promising mediator of the impact of extracellular signals on biosynthesis of ochratoxin A. *Front. Microbiol.* 12, 1–15. doi: 10.3389/fmicb.2021.631392

Gasteiger, E., Hoogland, C., Gattiker, A., and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *Humana Press.* 112, 531–552. doi: 10.1385/1-59259-890-0:571

Gehrke, A., Heinekamp, T., Jacobsen, I. D., and Brakhage, A. A. (2010). Heptahelical receptors GprC and GprD of aspergillus fumigatus are essential regulators of colony growth, hyphal morphogenesis, and virulence. *Appl. Environm. Microbiol.* 76, 3989–3998. doi: 10.1128/AEM.00052-10

Groves, J. W., and Loveland, C. A. (1953). The connection between Botryotinia Fuckeliana and Botrytis Cinerea. *Mycologia* 45, 415–425. doi: 10.1080/00275514.1953.12024279

Guruprasad, K., Reddy, B. V. B., and Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng. Des. Select.* 4, 155–161. doi: 10.1093/protein/4.2.155

Hartl, F. U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and *in vivo. Nat. Stru. Mol. Biol.* 16, 574–581. doi: 10.1038/nsmb.1591

Hasnain, S. M. (1993). Allergenic implications of airborne leptosphaeria ascospores. *Grana* 32, 315–318. doi: 10.1080/00173139309429996

Heller, J., Ruhnke, N., Espino, J. J., Massaroli, M., Collado, I. G., and Tudzynski, P. (2012). The mitogen-activated protein kinase BcSak1 of botrytis cinerea is required for pathogenic development and has broad regulatory functions beyond stress response. *Mol. Plant-Microbe Interact.* 25, 802–816. doi: 10.1094/MPMI-11-11-0299

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., et al. (2006). The PROSITE database. *Nucleic Acids Res.* 34, 227–230. doi: 10.1093/nar/gkj063

Hunte, C., and Richers, S. (2008). Lipids and membrane protein structures. *Curr. Struct. Biol.* 18, 406–411. doi: 10.1016/j.sbi.2008.03.008

Idnurm, A., and Heitman, J. (2005). Light controls growth and development via a conserved pathway in the fungal kingdom. *PLoS Biol.* 3: 0615–0626. doi: 10.1371/journal.pbio.0030095

Idnurm, A., and Howlett, B. J. (2001). Characterization of an opsin gene from the ascomycete Leptosphaeria maculans. *Genome* 44, 167–171. doi: 10.1139/g00-113

Idnurm, A., Verma, S., and Corrochano, L. M. (2010). A glimpse into the basis of vision in the kingdom Mycota. *Fungal Genet. Biol.* 47, 881–892. doi: 10.1016/j.fgb.2010.04.009

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J. Biochem.* 88: 1895–1898.

Itoh, M., Nacher, J. C., Kuma, K., Goto, S., and Kanehisa, M. (2007). Evolutionary history and functional implications of protein domains and their combinations in eukaryotes. Genome *Biol.* 8, 1–15. doi: 10.1186/gb-2007-8-6-r121

Jacob, A. G., and Smith, C. W. J. (2017). Intron retention as a component of regulated gene expression programs. *Human Genetics.* 136, 1043–1057. doi: 10.1007/s00439-017-1791-x

Jain, S., and Bader, G. D. (2010). An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology. *BMC Bioinformat.* 11, 562. doi: 10.1186/1471-2105-11-562

Jan, C. H., Williams, C. C., and Weissman, J. S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science* 346, 748–751. doi: 10.1126/science.1257522

Kaczmarek, J., and Jedryczka, M. (2012). Characterization of two coexisting pathogen populations of Leptosphaeria spp., the cause of stem canker of brassicas. *Acta Agrobotanica* 64, 3–14. doi: 10.5586/aa.2011.012

Kaiser, A., and Coin, I. (2020). Capturing peptide–GPCR interactions and their dynamics, Molecules. *MDPI AG.* 25, 20. doi: 10.3390/molecules25204724

Karapapa, V. K., Bainbridge, B. W., and Heale, J. B. (1997). Morphological and molecular characterization of Verticillium longisporum comb. nov., pathogenic to oilseed rape, *Mycolog. Res.* 101, 1281–1294. doi: 10.1017/S0953756297003985

Klare, J. P., Chizhov, I., and Engelhard, M. (2008). Microbial rhodopsins: scaffolds for ion pumps, channels, and sensors. *Res. Probl. Cell Different.* 45, 73–122. doi: 10.1007/400_2007_041

Koch, G. L. E. (1992). The endoplasmic reticulum. *Adv. Mol. Cell Biology* 4, 161–182. doi: 10.1016/S1569-2558(08)60177-0

Kukuruzinska, M. A., and Lennon, K. (1998). Protein N-glycosylation: molecular genetics and functional significance. *Critical Rev. Oral Biol. Med.* 9, 415–448. doi: 10.1177/10454411980090040301

Kulkarni, R. D., Thon, M. R., Pan, H., and Dean, R. A. (2005). Novel G-protein-coupled receptor-like proteins in the plant pathogenic fungus Magnaporthe grisea. *Genome Biol.* 6, 1–14. doi: 10.1186/gb-2005-6-3-r24

Kumamoto, C. A. (2008). Fungal contact sensing. *Nat. Rev. Microbiol.* 6, 667–673. doi: 10.1038/nrmicro1960

Kupfer, D. M., Drabenstot, S. D., Buchanan, K. L., Lai, H., Zhu, H., Dyer, D. W., et al. (2004). Introns and splicing elements of five diverse fungi. *Eukaryotic Cell* 3, 1088–1100. doi: 10.1128/EC.3.5.1088-1100.2004

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105–132. doi: 10.1016/0022-2836(82)90515-0

Landolt-Marticorena, C., and Reithmeier, R. A. F. (1994). Asparagine-linked oligosaccharides are localized to single extracytosolic segments in multi-span membrane glycoprotiens. *Biochem. J.* 302, 253–260. doi: 10.1042/bj3020253

Lee, A. G. (2004). How lipids affect the activities of integral membrane proteins. *Biomembranes* 1666, 62–87. doi: 10.1016/j.bbamem.2004.05.012

Li, L., Wright, S. J., Krystofova, S., Park, G., and Borkovich, K. A. (2007). Heterotrimeric G protein signaling in filamentous fungi. *Annual Rev. Microbiol.* 61, 423–452. doi: 10.1146/annurev.micro.61.080706.093432

Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., et al. (2021). Types of nuclear localization signals and mechanisms of protein import into the nucleus. *BioMed Central.* 19, 1–10. doi: 10.1186/s12964-021-00741-y

Lynch, M. (2002). Intron evolution as a population-genetic process. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6118–6123. doi: 10.1073/pnas.092595699

Lyu, X., Shen, C., Fu, Y., Xie, J., Jiang, D., Li, G., et al. (2016). The microbial opsin homolog sop1 is involved in Sclerotinia sclerotiorum development and environmental stress response. *Front. Microbiol.* 6, 1–14. doi: 10.3389/fmicb.2015.01504

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., et al. (2011). CDD: A Conserved Domain Database for the functional annotation of proteins, *Nucleic Acids Res,* 39, 225–229. doi: 10.1093/nar/gkq1189

Meena, M., Swapnil, P., Zehra, A., Amir, M., Dubey, M., Patel, C. P., et al. (2018). Virulence factors and their associated genes in microbes. *Microb. Biotechnol.* 2018, 181–205. doi: 10.1016/B978-0-444-63503-7.00011-5

Meena, P. D., Awasti, R., Chattopadhyay, C., and Kolte, S. (2010). Alternaria blight : a chronic disease in rapeseed-mustard. *J. Oilseed Brassica* 1, 1–11.

Mendes-Pereira, E., Balesdent, M-H., Brun, H., and Rouxel, T. (2003). Molecular phylogeny of the Leptosphaeria maculans-L. *biglobosa* species complex. *Mycological Res.* 107, 1287–1304. doi: 10.1017/S0953756203008554

Mitra, N., Sinha, S., Ramya, T. N. C., and Surolia, A. (2006). N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 31, 156–163. doi: 10.1016/j.tibs.2006.01.003

Moghadam, A., Niazi, A., Afsharifar, A., and Mohsen Taghavi, S. (2016). Expression of a recombinant Anti-HIV and Anti-Tumor protein, MAP30, in nicotiana tobacum hairy roots: A pH-stable and thermophilic antimicrobial protein. 11:e0159653. doi: 10.1371/journal.pone.0159653

Murzin, A. G., Brenner, S., Hubbard, T., and Chothial, C. (1995). SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540. doi: 10.1016/S0022-2836(05)80134-2

Nadarajah, K., Kumar, S., and Omar, N. (2018). Recent advances in understanding the fungal G-protein-coupled receptors. *Malaysian J. Microbiol.* 14, 611–623. doi: 10.21161/mjm.1461823

Naro, C., and Sette, C. (2017). Timely-regulated intron retention as device to fine-tune protein expression. *Cell Cycle* 16, 1321–1322. doi: 10.1080/15384101.2017.1337983

Nees, E. (1973). "Mycologie Systématique d Arenberg, Appliquée Belgium Catholique (With Text-figures) (for Trichophaea)," in *Botrytis-like of conidiophore belong Botrytis Pers., lectotypified increasing by (Clements fungi, exhibiting by Hughes (1953).* Available online at: urlhttps://typeset.io/pdf/botrytis-and-botrytis-like-genera-55ud0bczdi.pdf

Nishimura, M., Park, G., and Xu, J. R. (2003). The G-beta subunit MGB1 is involved in regulating multiple steps of infection-related morphogenesis in Magnaporthe grisea. *Mol. Microbiol.* 6, 231–243. doi: 10.1046/j.1365-2958.2003.03676.x

Nolan, S., Cooke, B. M., and Monahan, F. J. (1999). Studies on the interaction between Septoria tritici and Stagonospora nodorum in wheat. *Eur. J. Plant Pathol.* 105, 917–925. doi: 10.1023/A:1008733227236

O'Donnell, K., and Cigelnik, E. (1997). Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus Fusarium are nonorthologous. *Mol. Phylogenet. Evolut.* 7, 103–116. doi: 10.1006/mpev.1996.0376

O'Donnell, K., Sutton, D. A., Rinaldi, M. G., Magnon, K. C., Cox, P. A., Revankar, S. G., et al. (2004). Genetic diversity of human pathogenic members of the Fusarium oxysporum complex inferred from multilocus DNA sequence data and amplified fragment length polymorphism analyses: Evidence for the recent dispersion of a geographically widespread clonal lineag. *J. Clin. Microbiol.* 42, 5109–5120. doi: 10.1128/JCM.42.11.5109-5120.2004

Oesterhelt, D. (1989). Two pumps, one principle: light-driven ion transport in halobacteria. *Trends Biochem. Sci.* 2, 57–61. doi: 10.1016/0968-0004(89)90044-3

Ortoneda, M., Guarro, J., Madrid, M. P., Caracuel, Z., Roncero, M. I. G., Mayayo, E., et al. (2004). *Fusarium oxysporum* as a multihost model for the genetic dissection of fungal virulence in plants and mammals. *Infect. Immunity* 72, 1760–1766. doi: 10.1128/IAI.72.3.1760-1766.2004

Peng, C., and Gao, F. (2014). Protein localization analysis of essential genes in prokaryotes, *Scient. Reports* 4, 1–7. doi: 10.1038/srep06001

Phillips, R., Ursell, T., Wiggins, P., and Sens, P. (2009). Emerging roles for lipids in shaping membrane-protein function. *Nature* 459, 379–385. doi: 10.1038/nature08147

Phizicky, E. M., and Fields, S. (1995). Protein-protein interactions: Methods for detection and analysis. *Microbiol. Rev.* 59, 94–123. doi: 10.1128/mr.59.1.94-123.1995

Purschwitz, J., Müller, S., Kastner, C., and Fischer, R. (2006). Seeing the rainbow: light sensing in fungi. *Curr. Microbiol.* 9, 566–571. doi: 10.1016/j.mib.2006.10.011

Sadowski, P. G., Groen, A. J., Dupree, P., and Lilley, K. S. (2008). Sub-cellular localization of membrane proteins. *Proteomics* 8, 3991–4011. doi: 10.1002/pmic.200800217

Sanchez, K. M., Gable, J. E., Schlamadinger, D. E., and Kim, J. E. (2008). Effects of tryptophan microenvironment, soluble domain, and vesicle size on the

thermodynamics of membrane protein folding: Lessons from the transmembrane protein OmpA. *Biochemistry* 47, 12844–12852. doi: 10.1021/bi800860k

Schmidt, T. M., Chen, S. K., and Hattar, S. (2011). Intrinsically photosensitive retinal ganglion cells: Many subtypes, diverse functions. *Trends Neurosci.* 34, 572–580. doi: 10.1016/j.tins.2011.07.001

Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000). SMART: A web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 28, 231–234. doi: 10.1093/nar/28.1.231

Schwarz, F., and Aebi, M. (2011). Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Structural Biol.* 21, 576–582. doi: 10.1016/j.sbi.2011.08.005

Shental-Bechor, D., and Levy, Y. (2008). Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8256–8261. doi: 10.1073/pnas.0801340105

Shichida, Y., and Matsuyama, T. (2009). Evolution of opsins and phototransduction. *Philos. Trans. R. Soc. Lond., B.* 364, 2881–2895. doi: 10.1098/rstb.2009.0051

Snyder, W. C., and Hansen, H. N. (2013). The species concept in fusarium. *Botanical Society Am. Stable.* 27, 64–67. doi: 10.1002/j.1537-2197.1940.tb14217.x

Soppa, J., Duschl, J., and Oesterhelt, D. (1993). Bacterioopsin, haloopsin, and sensory opsin I of the halobacterial isolate Halobacterium sp. strain SG1: three new members of a growing family. *J. Bacteriol.* 175, 2720–2726. doi: 10.1128/jb.175.9.2720-2726.1993

Sun, H., Greathouse, D. V., and Andersen, O. S. (2008). The preference of tryptophan for membrane interfaces: Insights from N-methylation of tryptophans in gramicidin channels, Journal of Biological Chemistry. Â© 2008 ASBMB. *Am. Soc. Biochem. Mol. Biol.* 283, 22233–22243. doi: 10.1074/jbc.M802074200

Sun, P. D., Foster, C. E., and Boyington, J. C. (2004). Overview of protein structural and functional folds. *Curr. Protoc. Protein Sci.* 17, 1–189. doi: 10.1002/0471140864.ps1701s35

Szklarczyk, D., Gable, A., Nastou, K., Lyon, D., Krisch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. doi: 10.1093/nar/gkaa1074

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Tamura, Y., Kawano, S., and Endo, T. (2020). Lipid homeostasis in mitochondria. *Biol. Chem.* 401, 821–833. doi: 10.1515/hsz-2020-0121

Terakita, A. (2005). Protein family review the opsins. *Genome Biol.* 6, 213. doi: 10.1186/gb-2005-6-3-213

Uncu, A. O., Uncu, A., Celik, I., Doganlar, S., and Frary, A. (2015). A primer to molecular phylogenetic analysis in plants. *Crit. Rev. Plant Sci.* 34, 454–468. doi: 10.1080/07352689.2015.1047712

Venkatesh, B., Lee, A. P., Ravi, V., Maurya, A. K., Lian, M. M., Swann, J. B., et al. (2014). Elephant shark genome provides unique insights into gnathostome evolution. *Nature* 505, 174–179. doi: 10.1038/nature12826

Von Heijne, G. (2007). The membrane protein universe: whats out there and why bother? *J. Internal Med.* 261, 543–557. doi: 10.1111/j.1365-2796.2007.01792.x

Waalwijk, C., de Koning, J., Baayen, R., and Gams, W. (1996). Discordant groupings of Fusarium sfrom sections Elegans, Liseola and Dlaminia based on ribosomal ITS1 and ITS2 sequences. *Mycologia* 88, 361–368. doi: 10.1080/00275514.1996.12026663

Walker, K. C., and Booth, E. J. (2001). Agricultural aspects of rape and other Brassica products. *Eur. J. Lipid Science and Technol.* 103, 441–446. doi: 10.1002/1438-9312(200107)103:7<441::AID-EJLT441>3.0.CO;2-D

Wall, M. A., Coleman, D. E., Lee, E., Iñiguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., et al. (1995). The structure of the G protein heterotrimer Giα1β1γ2ʹ. *Cell* 83, 1047–1058. doi: 10.1016/0092-8674(95)90220-1

Waschuk, S. A., Shi, L., and Brown, L. S. (2005). Leptosphaeria rhodopsin: Bacteriorhodopsin-like proton pump from a eukaryote, *Proc. Natl. Acad. Sci. U.S.A.* 102, 6879–6883. doi: 10.1073/pnas.0409659102

Williamson, B., Tudzynski, B., Tudzynski, P., and van Kan, J. A. L. (2007). Botrytis cinerea: The cause of grey mould disease. *Mol. Plant Pathol.* 8, 561–580. doi: 10.1111/j.1364-3703.2007.00417.x

Xue, C., Hsueh, Y. P., and Heitman, J. (2008). Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS Microbiol. Rev.* 32, 1010–1032. doi: 10.1111/j.1574-6976.2008.00131.x

Zeise, K., and Von Tiedemann, A. (2001). Morphological and physiological differentiation among vegetative compatibility groups of *Verticillium dahliae* in relation to V. *longisporum, J. Phytopathol.* 149, 469–475. doi: 10.1046/j.1439-0434.2001.00659.x

Zeise, K., and Von Tiedemann, A. (2002a). Application of RAPD-PCR for virulence type analysis within *Verticillium dahliae* and V. *longisporum. J. Phytopathol.* 150, 557–563. doi: 10.1046/j.1439-0434.2002.00799.x

Zeise, K., and Von Tiedemann, A. (2002b). Host specialization among vegetative compatibility groups of *Verticillium dahliae* in relation to *Verticillium longisporum. J. Phytopathol.* 150, 112–119. doi: 10.1046/j.1439-0434.2002.00730.x

Zhou, X. E., Melcher, K., and Xu, H. E. (2012). Structure and activation of rhodopsin. *Acta Pharmacologica Sinica.* 33, 291–299. doi: 10.1038/aps.2011.171