MALAYSIAN JOURNAL OF BIOCHEMISTRY & MOLECULAR BIOLOGY

The Official Publication of The Malaysian Society For Biochemistry & Molecular Biology (MSBMB)
http://mjbmb.org

IN SILICO CHARACTERISATION OF *A Polygonum minus* SEQUENCE AS PUTATIVE BEACH DOMAIN PROTEIN

Nor Azlan Nor Muhammad[1], Chong Sheng Cheah[2], Nor Syafinaz Yaakob[2*]

[1]*Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[2]*Centre for Drug and Herbal Development, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia*

*Corresponding Author: nsy@ukm.edu.my

History
Received: 11 November 2022
Accepted: 8 July 2023

Keywords:

Bioinformatics; Hypothetical protein; Polygonum minus; BEACH domain-containing protein; Plant autophagy

Abstract

Polygonum minus has been widely studied due to its significant content of flavonoid, phenolic and terpenoid compounds which are well-associated with medicinal properties. However, most of the pathways that regulate the development and synthesis of secondary metabolites of *P. minus* remained unknown. Identifying candidate genes and proteins that are involved in the biosynthetic pathways would contribute to a better understanding of the bioactive compound synthesis of *P. minus.* This study analysed and characterised a large 2151 amino acid hypothetical protein in *P. minus* using bioinformatic tools and databases. Sequence homology search, conserved domain prediction, protein hierarchical clustering, structure prediction and sub-cellular localisation prediction were done. Results from sequence homology search showed that it is a BEACH domain-containing protein. Analysis of its conserved domain revealed the presence of Concanavalin A-like lectin domain (ConA), Pleckstrin homology-like domain (PH), BEACH domain and WD40 domain repeats. Hierarchical clustering of protein supported that it has a close relationship with BEACH protein family members. The structure of the hypothetical protein was proposed and revealed the presence of a small pocket that functions for the binding of dsRNA. Sub-cellular localisation analysis suggested that the hypothetical protein is localized in the endoplasmic reticulum. It is proposed that the hypothetical protein is involved in the autophagy process in *P. minus.* This study serves as one of the keys to understanding the role of the hypothetical protein in the plant autophagy process which is important in immunity defence and stress tolerance of *P. minus.* Manipulation of the autophagy process may result in a better yield of secondary metabolites and promotes the survival rate of *P. minus*.

INTRODUCTION

Natural resources, especially plants are an important source of new drugs, new drug leads and new chemical entities for management of diseases [1]. *Polygonum minus* (syn. *Persicaria minor*), is known as 'kesum' in Malaysia. It is commonly found in Southeast Asian countries such as Malaysia, Thailand, Vietnam and Indonesia. It produces a sweet and pleasant aroma and is commonly used as a flavouring agent in Malaysian cuisines [2]. *P. minus* is classified in *Magnoliophyta* division, *Polygonales* order, *Polygonaceae* family and genus of Polygonum [3].

P. minus is a promising plant for drug discovery and development as secondary metabolites of *P. minus* exert several pharmacological activities, such as antioxidant, antiviral, antimicrobial and antifungal activities [4]. Apart from that, secondary metabolites produced by *P. minus* can potentially act as protection against insects [5]. Amino acid and fatty acid composition studies revealed that *P. minus* extracts consisted largely of glutamic and aspartic acids, alanine, proline, and leucine while hydroxyproline,

ornithine, cystine, methionine, and tryptophan were found in smaller amounts. Fatty acids including oleic, linoleic, and palmitic dominated the fatty acid content in *P. minus* extracts [6]. *P. minus* extracts also contain a significant amount of flavonoid, phenolic and terpenoid compounds that can act as antioxidants [7]. Thus, the identification of candidate genes and proteins involved in the biosynthetic pathway of these molecules would contribute to a better understanding and allow manipulations on metabolite synthesis in P. minus. [8].

Transcriptome data generated from RNA-seq analysis revealed that there are three unusually large hypothetical proteins (size of more than 2000 amino acids) that are present in *P. minus* [9]. Proteins with unknown functions may be termed hypothetical proteins (HPs) or putative proteins because these proteins have not been functionally understood, characterised and described [10]. One of the large hypothetical proteins in *P. minus* is analysed using bioinformatic analyses in this study. The identity of the protein and its characteristics are described herein.

MATERIALS AND METHODS

Transcriptome Data of the Hypothetical Protein

The hypothetical protein originated from transcriptome data of *P. minus* generated from RNA-seq analysis done by [9]. The hypothetical protein has 2151 amino acids, depicted as the following:

MRKDQLPFFDLIGNDSGIVIRTPMHWPISKGFTFSCWLRIENFPGVGAMGLFSFLSEHGKGCFALLTKDKLIYESIS
QKRQCVSLNVNIVTKKWHFLCITHSIGRAFSGGSILRCYVNGALVSSERIRYPRINELLTSCSIGCKIHLPRGDEESP
SHSIKEASSFLGQIGPIYMFNDVITPEQVLGISSLGPSYMYSFLDNENAHEPLPGGVLDAKDGLASKMIFGLNAQA
SNGRALYNVAPLLESAPDKSSFEATVMTGTQLCSRRLLQQIIYVVGGVSVFFPLFTQSQWYENEDDRDEHSLLIPI
TKERLTAEVIELIAFVLDENLANQQQMHLISGFSILGFLLQSVPPHQLNLETLSALKHLFNVVVANSGLSEFLVKEA
VSFIFLNPLVWVHTGFKIQRELCMFLSQQFDNDPRLLSLCRLPRIIDLIRQYYWDNAKCCSFAGGKRTLPSMPETI
VGERPSREEIQKIRLLMLSLGEMSLRQNISSADVKALVAFFETSQDMACIEDVLHMVIRAISLKSLLASFVEQVNSI
GGCQVFINLLQREYEPIRLLSLQFLGRVLIGLPSEKKGLKFFSLAVGRSKSISESTRRTSFRIQQLLFAAISDRLFRFP
VTEDLCAALFDALLGGASPKQVLQKQSQIEKQRSKRSGSQFYLPQILPLIFKFLSECENVVHRSKILADLLDLLDS
NPLNIESLMENGWNAWLTASMELAVVKNYKMEPSVKCDNETSEQHSVKILYSTVLGHYILSVKGGWQFLEETA
NFILMKFEQGSVSYHYFLRDLYEDLIWRLWKISSEENILALQPCRDNTLYLMKLIDELLLSEIDHKLPFPARSSELT
PDIFDIQSSRDIGSALSEALVGGSNEQISSNPPTVKQSVTTEDAVRDEEWWTLYDLLWIIISEMHGKGQRKMLMK
SASSAGPSLGQRARGLVESLNIPAAEMAAVVVSGGISNALGGKSSKNTDKAMLLRGEKCPRIVFRLLVLYLCRAS
LPRVSRCVQQIIPLLPGLLVGDDENSKNRLQLFIWSLLSVRSEYGMLDDGSRFHVMSHLIRETVNCGKSMLATSIA
GREDPPETGSSNKDTGALQNLIQKDRILAAVSEESTYIKTLKTDRNKQLHEWRVKVDESFSGESSQKKNFEDEVQ
SCLSLILSSDDNRRSAFQLAYDEERQIVAEKWIHVFRSLIDERGPWSANPFPNCDVRRWKLDKTEDSWRRRSKLR
QNYRFDEKLCQPSASASGKEVVVPSGEIKLGSGLLPEQMKQFLLKGIRRITDEVSSEPNESDTESSSQKPLNSEEIS
DHQSDLVKDVALSKDPLQERNESSSTEVEGGEALMSVPCVLVTPKRRLFGQLAVMHNALHFFGEFLVEGTAGFS
VFKNYDSISNPESGKQEHLAGIEKQKFLKLPAHLTSHSEKQNGIDSLDKQPKNVKRHRRWNLSKVKAVHWTRY
LLRYTAIEVFFSNSVAPVFFNFASQQLAKDFATLVVNTRNEFLFPKGSNRDRAGAISLVDRRIAQDLAETARESW
RRREMTNFEYLMILNTLAGRSYNDLTQYPIFPWVIADYSSETLDFNKSSTFRDLSKPVGALDSKRFEVFEDRYRN
FDDPDIPSFYYGSHYSSGGIVLYYLIRLEPFSTLHRNLQGGKFDHADRLFQSIEGTYKNCLSNTSDVKELIPEFFYM
PEILVNSNSYHLGVRQDGEPIADVCLPPWAKGSPEEFINKNREALESEYVSSNLHNWIDLIFGYKQRGKPAVEAA
NIFYYVTYEGAVDLENMEDELQRSAIEDQIANFGQTPIQIYRKKHPRRGPPIPIAHPLRFAPGSINLTSVVSSTRYPA
SSLLYVNVFDSNVVLVSQGLTMSVKMWLTTQLQSGGNFTFSSSQDPFFGIGSEVLSYRKIGSPLAEGFELGAQCF
ATLQTLSENFLISCGNWENSFQVVSLSDGRVVQSIRAHKDVVSAVAVTSDGSILATGSYDTTVMVWEVLPSRGH
EKRGRNVQSELPRKDYVISETPFHILCGHDDVITCLYVSVELDIVISGSKDGTCVFHTLREGKYIRSIRHPCGSPLS
KLVVSRHGWLVLYSDDDLSLHLYSINGKHIATSESNGRLNCVELSVCGEFLACAGDQGQITVRSIRSLELIMKYA
GAGKVITSLAVTPEECFLAGTKEGILLVY

Sequence Homology Search

Sequence homology search was done to look for the hypothetical protein's homologues. BLASTp [11] and HHpred [12] were used for the homology search. BLASTp was set to search against the UniProt/SwissProt database, a manually curated protein database [13]. Since the hypothetical protein originated from *P. minus*, the 'organism' column was set to 'plants (taxid 3193)' to narrow down the search results by searching only against plant proteins in the database. The 'E-value threshold' column was set to '1E-6'. Other parameters remained at default values. Matches with less than 30% query coverage or sequence identity of less than 20% were excluded. The hypothetical protein was searched against other protein databases as well, including the non-redundant protein dataset (NR) [14] and UniProt/TrEMBL using BLASTp [15].

HHpred was also used to conduct a sequence homology search. The database of *Arabidopsis thaliana* proteomes was selected as the database of choice. The alignment mode was

set to be local alignment because it is more sensitive in detecting remote relationships between a hypothetical protein and known proteins in the selected database. Other search options remained default.

Sub-cellular Localisation

The sub-cellular localisation of the hypothetical protein was investigated using bioinformatic tools including TargetP version 1.1 [16], WoLF PSORT [17], BaCelLo [18], MultiLoc2 [19], PSLpred [20] and LocTree3 [21].

TMHMM 2.0 [22,23] was used for predicting the propensity of a protein to be a membrane protein. SignalP 3.0 [24,25] and SecretomeP 1.0 [26] were used for the discrimination of secretory and non-secretory proteins.

Conserved Domain Prediction

Several tools were used to predict the conserved domains of the hypothetical protein. Publicly available bioinformatics protein family databases such as Pfam [27], SUPERFAMILY[28], CATH [29], CDD [30], CDART [31] and SMART [32,33] were used to analyse the hypothetical protein as well. CDART and SMART were used for similarity search based on domain architecture and profiles rather than by direct sequence similarity.

Hierarchical Clustering Protein

ProtoNet Version 6.1 [34] provided an automatic hierarchical clustering of the protein sequences. The ‘‘Classify your protein’’ option in ProtoNet was used for assignment of a biological function to the hypothetical protein. All settings remained default.

Structure Prediction

Three-dimensional structures of the hypothetical protein were determined using SWISS-MODEL and Phyre2. SWISS-MODEL is an automated system for 3D structure modelling using homology modelling method from the amino acid sequence of the hypothetical protein. SWISS-MODEL works by searching against the database to obtain a similar protein as a suitable template for modelling [35]. Phyre2 models the 3D structure of a hypothetical protein using homologous sequences and known structures [36].

RESULTS

Sequence Homology Search

Using the BLASTp tool, the hypothetical protein was searched against non-redundant protein dataset (NR). Results revealed that the hypothetical protein shows similarity with uncharacterised protein BVRB and BEACH domain-containing protein LvsC with unknown function in *Beta vulgaris*. The search against UniProt/TrEMBL database showed a match with an uncharacterised protein in *B. vulgaris* as well. However, since both NR and UniProt/TrEMBL contain proteins which are automatically annotated and are not reviewed [37], therefore the hypothetical protein was further searched against UniProt Swiss-Prot, a manually annotated and reviewed protein database for a more accurate and reliable result.

Further sequence homology search using BLASTp tool against UniProt SwissProt protein database revealed four possible homologs of the hypothetical protein in *A. thaliana* which were within the selected criteria (more than 30% query cover or sequence identity more than 20%). The homologs are BEACH domain homolog B (E-value = 0.0), BEACH domain homolog C1 (E-value = 2E-59), BEACH domain homolog C2 (E-value = 5E-133) and SPIRRIG-BEACH domain homolog A2 (E-value = 1E-133). HHpred analysis towards *A. thaliana* proteome produced three hits, NP_182087.1 (E-value = 7E-271), NP_171805.1 (E-value = 3E-266) and NP_192175.2 (E-value = 4E-257) which are all BEACH domain homologs.

Alignment of the hypothetical protein against curated BEACH domain-containing proteins namely the LvsA, SPIRRIG, ALFY (Human WD repeat and FYVE domain-containing protein), LYST (Lysosomal-trafficking regulator), Neurobeachin, LRBA (Lipopolysaccharide-responsive and beige like anchor protein) and FAN using BLASTp tool showed that the hypothetical protein had significantly high similarity with LvsA protein (56%), ALFY protein (47%) and LYST protein (40%). Pooled results of sequence homology searches suggested that the hypothetical protein was closely related to the BEACH family protein.

Conserved Domain Prediction

Pfam and SUPERFAMILY analyses showed six domain matches where Concanavalin A-like lectin domain, PH domain, BEACH domain and WD40 repeats were detected from N to C terminal. The domain assignation was further validated by CATH and SMART analysis which revealed the presence of BEACH domain protein and WD repeats in the hypothetical protein. Analysis using CDD (Conserved Domain Database) and CDART (Conserved Domain Architecture Retrieval Tool) revealed four known conserved domains, which were Concanavalin A-like lectin (Laminin G domain, pfam02138), PH-like superfamily (Cl17171), BEACH domain (pfam02138) and WD40 Superfamily (Cl02567). Two domains of unknown function were revealed as well, which were DUF4704 (pfam15787) and DUF4800 (Cl24581) (Figure 1).

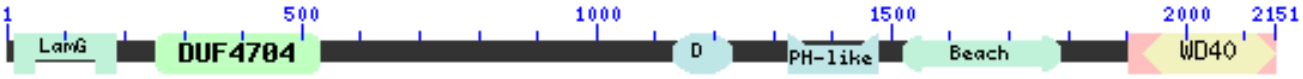

Figure 1. Domains in *P. minus* BEACH containing hypothetical protein sequence revealed by CDART tool.

Hierarchical Clustering of Protein

The hypothetical protein was further analysed using ProtoNet, where the relationship between proteins is presented in a hierarchical tree (Figure 2). ProtoNet revealed that the hypothetical protein belongs to WD40 Superfamily Cluster (Cluster 3933940), where the cluster had eight proteins. The protein cluster consisted of BEACH domain proteins such as LvsC-like protein and the rest were uncharacterised BEACH domain and WD40 repeats proteins.

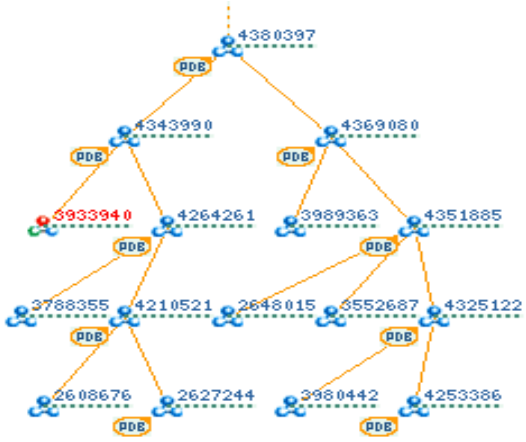

Figure 2. Hierarchical clustering of proteins generated from ProtoNet showed that the hypothetical protein belongs to WD40 Cluster (Cluster 3933940 indicated in red). Cluster WD40 393340 has cluster 4343990 (WD 40 cluster) as parents and cluster 438039 (BEACH Neurobeachin as grandparent). It has close relationship with Concanavalin A-like lectin/glucanase (cluster 4264261), BEACH cluster (Cluster 3989363) and Neurobeachin cluster (cluster 4351885).

Cluster WD40 393340 had cluster 4343990 (WD 40 cluster) as parents and cluster 4380397 (BEACH Neurobeachin) as grandparent. It has a close relationship with the Concanavalin A-like lectin/glucanase cluster (cluster 4264261), BEACH cluster (Cluster 3989363) and Neurobeachin cluster (cluster 4351885). Hierarchical clustering from ProtoNet further validated previous information provided from sequence homology search and conserved domain detection tools, including BLASTp, HHpred, Pfam, SUPERFAMILY, CDD and CDART. The hypothetical protein is a member of WD40 cluster which includes BEACH protein domains. Thus, the hypothetical protein will show a high degree of similarity with BEACH and WD40 family proteins such as neurobeachin and concanavalin A protein when analysed using the tools mentioned.

Structure Prediction

The structure was predicted using Phyre2. There were twenty models generated using different templates. The top three protein templates selected for protein modelling were C1mi1A (neurobeachin), d1mi1a1 (BEACH domain), and d1tt77a1 (BEACH domain). The top two template hits from the SWISS-MODEL analysis were 1t771A (Lipopolysaccharide-responsive and beige-like anchor protein) and 1mi1A (neurobeachin). The top-ranking model from Phyre2 (neurobeachin) was selected for further analysis as this was also the template found by SWISS-MODEL.

The structure was modelled by Phyre2 using template c1mi1A, a neurobeachin protein with a similarity of 45% with the hypothetical protein. The confidence of the model generated is 100% with a resolution of 2.90Å. The structure predicted was sent to ProFunc for further analysis. ProFunc revealed the presence of BEACH motif in the hypothetical protein. Matching folds are the PH-BEACH domain of human neurobeachin (PDB: 1mi1) and PH-BEACH domain of human LRBA/BGL (PDB: 1t77).

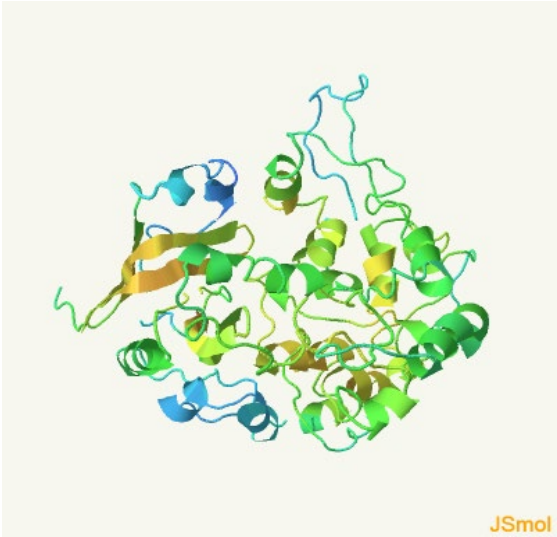

Figure 3. Overall structure predicted using Phyre2. Using ProQ2 quality assessment tool attached in Phyre2, the model is coloured according to the quality of the protein model. From the model, it can be seen that the majority of the chains are coloured green and light green, indicating that the model is modelled well.

Nest analysis of ProFunc revealed a significant hit of nest. Nests are structural motifs that are often found in functionally important regions of protein structures. This significant nest hit was labelled in red. The score of 2.97 indicated that the nest is functionally significant (significance threshold >2.0). The nest is made up of 3 highly conserved residues of histidine, leucine and glycine. The nest started with histidine residue at position 1679 and ended with glycine at position 1681.

3D functional template searches in ProFunc revealed that the hypothetical protein has no hits in enzyme active site templates and ligand binding templates. However, there was a dsRNA-binding domain (PDB:1di2) revealed in DNA-binding templates. According to PDB, it is composed of two chains of interferon-inducible double-stranded RNA-dependent protein kinase activator A homolog B which is involved in the biochemical function of double-stranded RNA binding (Figure 4).

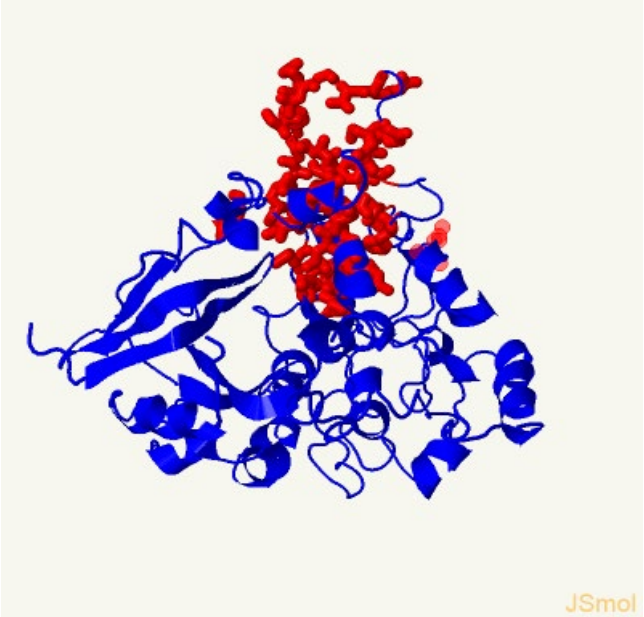

Figure 4. ProFunc 3D functional template search results revealed the presence of a DNA-binding template for double-stranded RNA protein.

Sub-cellular Localisation

Variable results were obtained for sub-cellular localisation analysis of the hypothetical protein. Several tools were employed for cross-checking to ensure the accuracy of the results. TargetP 1.1 predicted that the protein is localised in places other than chloroplast, mitochondria and signal peptide with high probability. Thus, more tools were employed to further confirm the possible location of the hypothetical protein.

The result from Secretome 1.0 showed that the hypothetical protein is not a secretory protein because of low NN and odd score. TMHMM 2.0 prediction revealed that the hypothetical protein is not located at the transmembrane with high probability. This showed that the hypothetical protein does not have a transmembrane sequence.

WoLF PSORT revealed 14 nearest neighbours that showed similarity with the hypothetical protein and their site of localisation, where five of them were localised in the nucleus, four localised in the cytoplasm, two in chloroplast and two in cytoskeleton; consequently, the possible site of localisation could be narrowed down to either cytoplasm or nucleus.

Subsequent analysis with PSLpred and MultiLoc2 revealed that it is a cytoplasmic protein with a probability of 68.1% and 58% respectively. However, BaCelLo suggested that the hypothetical protein is localised in the nucleus. LocTree3 was employed to further confirm the sub-cellular localisation of the hypothetical protein. LocTree3 was able to predict and categorise eighteen localisation classes for eukaryotic proteins. The result revealed that the hypothetical protein is localised in the endoplasmic reticulum near the nucleus. This is consistent with results from previous sub-cellular localisation tools.

DISCUSSION

In a previous study, a set of transcriptome data of *P. minus* was generated in order to understand its secondary metabolite synthesis pathways. This hypothetical protein which served as the subject of this study was one of the large proteins that are present in the transcriptome data of *P. minus* [9]. The protein in *P. minus* hypothetical protein was analysed using sequence homology search, conserved domain prediction, hierarchical classification of protein, structural prediction, and subcellular localisation prediction.

Concanavalin A (ConA)-like lectin domain exists in many BEACH family proteins, such as SPIRRIG, human lysosomal trafficking regulator (LYST) and Neurobeachin (NBEA). ConA domain can co-exist with PH domain, BEACH domain and WD40 repeats in BEACH family proteins [38]. ConA resides in the endoplasmic reticulum and Golgi complex [39]. It could be involved in oligosaccharide bindings, and also associated with glycoprotein trafficking and sorting in the secretory pathway of eukaryote cells from endoplasmic reticulum to the Golgi complex, related to vesicle fusion mechanism [40]. A study found that plant lectins ConA is capable of inducing autophagic pathways in cells and its role was tested in cancerous hepatoma cells [41]. In the experiment, ConA protein was tested in ML-1, Huh-7, HepG2 and CT-26 hepatoma cell lines. The result showed that the ConA induced autophagy in hepatoma cells and their growth were inhibited. ConA was then tested on liver tumour nodules in severe immune deficiency (SCID) mice and the result showed that the liver tumour nodule formation was inhibited. ConA initiated autophagic cell death via BNIP3-mediated mitochondria autophagy pathway [42]. ConA binds and

becomes internalised to the mitochondria surface via clathrin-mediated endocytosis. This results in a decrease of mitochondrial membrane potential, initiating the autophagic pathway. With BNIP3 and LC3-II, autophagosome will form and sequestrates cytoplasm, and fuse with lysosome [43]. Thus, it is implicated that the ConA lectin domain in the hypothetical protein could be involved in autophagy cell cycle for *P. minus* plant.

PH domain is known as Pleckstrin homology-like domain and it exists in plant proteins as well, such as phosphatidylinositol 4-kinase in *Arabidopsis* [44]. This motif exists in proteins that bind to ligands, such as G-protein βγ subunits and polyphosphorylated inositol lipids [45]. PH domain is generally involved in recruiting proteins to the appropriate cellular location through the binding of phosphatidylinositols and is essential for phosphatidylinositol 4-kinase activity [46]. Phosphatidylinositol 4-kinase is involved in phosphoinositide-dependent phospholipase C (PI-PLC) signalling in plant cells [47]. PI-PLC can be found in cytosol and membrane and it is involved in PLC signalling pathways which plays role in stresses, deficiency and growth and development processes, such as tip growth, stomatal function, CO_2 fixation and gravitropism [48,49].

BEACH (Beige and Chediak-Higashi) domain is a family of conserved proteins in eukaryotes. The BEACH domain is important for the function of several proteins and is implicated in membrane trafficking [50]. It is reported that BEACH domain-containing proteins play a role in endosome and lysosome organelle biogenesis [51]. Examples of BEACH domain-containing proteins are human lysosomal trafficking regulator (LYST), LPS-responsive and beige-like anchor (LRBA), neurobeachin and SPIRRIG. It was reported that disruption of LYST leads to Chediak-Higashi syndrome, characterised by dysfunction of lysosomes and melanosomes that lead to severe immunodeficiency, albinism, poor blood coagulation and neurologic problems. LYST defect mice showed significant elevation of endoplasmic reticulum proteins in lysosomal membrane of liver cells of mice. LYST protein plays a role in organised incorporation of proteins into lysosomal membrane during recycling or maturation processes of lysosomal membrane [52]. Literature suggested that BEACH domain proteins in plants are essential in vacuole trafficking activities, plant immunity and salt stress regulations [53]. SPIRRIG protein is the only plant BEACH family protein that has been well characterised. SPIRRIG is a large protein which is preceded by PH domain, then followed by BEACH domain, ending with WD40 repeats [54]. It plays a role in salt stress tolerance and vesicle trafficking processes in *Arabidopsis* [53].

WD40 domain consists of 44-60 repeating units that ended with tryptophan (W) and aspartate (D) and is widely distributed in eukaryotic proteins. WD40 domain constantly exists with PH in the same protein [55]. It is involved in a variety of functions, such as vesicle trafficking, cell cycle control, transcription regulation, signal transduction, autophagy and apoptosis [50]. The function of WD40 repeats can act as a platform for interaction with other proteins. The WD repeats in *Arabidopsis* can provide a binding site for other proteins and act as integral components of protein complexes [56]. It is also worth noting that the present study also discovered the presence of two domains which are yet to be characterised. DUF4704 is a domain of unknown function which is found in eukaryotes on neurobeachin proteins. DUF4800 is functionally uncharacterised, which is found in eukaryotes as well.

The sub-cellular localisation of the BEACH domain hypothetical protein was supported by literature where it has been shown that all BEACH domain proteins are localised in cytoplasm areas rather than the nucleus. SPIRRIG, a BEACH domain protein in *A. thaliana* plant which is responsible for salt stress tolerance is localised in the cytoplasm [54]. LvsA is a BEACH domain protein whereby in *Dictyostelium* amoeba is responsible for cytokinesis and osmoregulation, is localised in the contractile vacuole [57]. Thus, it is evident that BEACH domain proteins are localised in the cytoplasm, thus implying that the hypothetical protein is localized in the cytoplasm area, specifically in the endoplasmic reticulum of cells in *P. minus*.

The protein was clustered in cluster 3933940 and was named as WD40 superfamily by SCOP. WD40 is short and consists of approximately 40 amino acids. It is commonly found in eukaryotes and involved in cellular functions such as signal transduction and transcription regulation to cell cycle controls [58]. It was proposed that repeats of WD40 carried out their function by acting as a platform for assembling of protein complexes [59]. This suggests that the presence of WD40 repeats in the hypothetical protein is important for protein binding, and subsequently contributes to signal transduction and regulation of cell cycle.

There was a hit of dsRNA-binding domain in ProFunc tool, which was PDB:1di2. According to literature, this domain presents in Protein Kinase R (PKR). PKR is the only member of eukaryotic initiation factor 2a (eIF2a) phosphorylating family of kinases that is induced by interferon. PKR is activated by binding of dsRNA onto the dsRNA-binding domain of PKR. PKR plays role in immune response mediation, signal transductions and apoptosis [60,61].

Information gathered about the domains present in the hypothetical protein implicates that these domains in the BEACH domain family hypothetical protein are working together to carry out a function involving vesicle trafficking and signal transduction of a biological process. Summing up all of the findings in sequence homology search, conserved domain prediction, hierarchical classification of protein, structural prediction, and subcellular localisation prediction, it is proposed that the hypothetical protein is involved in autophagic pathway in *P. minus*, which is similar to a BEACH domain-containing protein ALFY (autophagy-linked FYVE). The annotation of the hypothetical protein as a protein related to autophagy is consistent with all findings

from the analysis mentioned. ALFY protein shows high degree of similarity with the hypothetical protein. Although LvsA has the highest similarity with the hypothetical protein (56%), LvsA protein is considered having different function from the hypothetical protein. This is because LvsA that exists in contractile vacuole of *Dictyostelium* is responsible for osmoregulation of the amoeba [62], which is not the same with *P. minus* plant cell. Furthermore, the hypothetical protein has PH-BEACH-WD40 domain at the end of the protein, which resembles ALFY protein [63] and is localised in cytoplasmic area. Presence of Concanavalin A lectin domain may support the role of these domains to carry out autophagy cell cycle because Concanavalin A protein is reported to play a role in autophagy cell cycle as well.

During the autophagy process, WD40 repeats are essential to interact with the autophagy protein Atg5. PH-BEACH domains are involved in phosphoinositides binding and mediate protein-protein interaction. The PH-BEACH domain will interact with autophagy receptor p62 [64]. The involvement of the BEACH domain protein in autophagy was further described by [63]: first, the misfolded proteins that are selected for autophagy will aggregate and become ubiquitinated, and then be recognised by p62. The ALFY protein then is recruited and interacts with the autophagy receptor p62. This interaction will facilitate the sequestration of protein into larger aggregates packaged into autophagosomes to permit degradation upon fusion to the lysosome [65].

Autophagy is important for housekeeping in plants and adaptation to drastic changing environmental stresses such as oxidative stress, starvation, pathogen invasion and salt tolerance [66]. This process is regulated by phosphatidylinositol-3 kinase complex, ubiquitination of protein and formation of ATG complexes, which lead to membrane recruitment in autophagosome formation [67]. A study showed that the absence of autophagy gene RNA AtATG18a increases sensitivity to oxidative stress, resulting in a higher amount of oxidised protein accumulated in the *Arabidopsis thalia* model. This suggests autophagy plays a role in degrading oxidised protein in plant cells [68]. Plants can defend against microbial pathogens via immunity related programmed cell death, which plays a critical role in resistance towards pathogens. Suppression of the ATG gene in *A. thalia* model will results in widespread infection [69]. Autophagy genes were also shown to a play role in osmotic stress, in a study where an autophagy gene GFP–AtAtg8f-HA was expressed to adjust the root architecture in response to salt stress [70]. During nutrient starvation, Atg proteins are upregulated to facilitate the recycling process of nutrients in plants. Models with knockout of autophagy genes are more sensitive towards starvation of nitrogen and sugar [71]. Autophagy promotes the longevity of *A. thaliana* during caloric restriction condition. Knockout of autophagy related genes causes shorter lifespan of *Arabidopsis.* In short, regulation of autophagy process in plants is critical for their development including dealing with stresses, plant immunity and starvation [72].

It was shown that the content of secondary metabolite is affected by autophagy process as well [73]. In the study, comparison of the content of secondary metabolites between wild-type *Arabidopsis* and ATG mutants was done using liquid chromatography-mass spectrometry (LC-MS). Results showed that ATG mutants had a significantly lesser of some secondary metabolites compared to wild type, such as flavonols, indole and aliphatic glucosinolates. Autophagy-related protein has also been reported to implicate the metabolism of phytohormones such as jasmonates and salicylic acid [74] which have been shown to possess medicinal relevance. In *P.minus*, methyl jasmonate was shown to influence the levels of aldehydes and terpenes [75], which among others have been described with pharmacological properties of *P.minus* [76].

It is noted that the findings of this study are limited by the predictive nature of computational analyses and further experiments on the protein characterisation are required to support the findings reported. However, the present findings about the hypothetical protein which is involved in autophagy may be a key to a better understanding of the regulation and development processes of *P. minus*. These findings contribute to proteomics profiling of *P.minus* where in a holistic view, this is important for sustainable cultivation of a healthy plant, for instance, cultivating a *P. minus* plant that is more resistant to starvation, stresses and pathogen invasion by manipulation of autophagy processes in *P. minus*, thus improving the cultivation yield. Moreover, this profiling may also open up avenues towards manipulation of its metabolites production as *P. minus* is now increasingly reported for its medicinal properties where alteration of autophagy in plants may result in better yield of secondary metabolites which can contribute to a better pharmacological effect of *P. minus*. Future work shall embark on empirically validating the autophagy function of this hypothetical protein and subsequently characterising its role in regulating the levels of important metabolites, especially those with medicinal properties.

ACKNOWLEDGEMENTS

We thank Prof. Dr. Zeti Azura Mohamed Hussein from INBIOSIS, UKM for contributing the sequence data used for this study.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

REFERENCES

1. Bhanot, A., Sharma, R., Singh, S., Noolvi, M. N., and Singh, S. (2013) In vitro anti-cancer activity of ethanol extract fractions of *Aerva lanata* L. *Pak. J. Biol. Sci.* **16**(22), 1612.

2. Bunawan, H., Choong, C. Y., Md-Zain, B. M., Baharum, S. N., and Noor, N. M. (2011) Molecular systematics of *Polygonum minus* Huds. based on ITS sequences. *Int. J. Mol. Sci.* **12**(11), 7626-7634.

3. Qader, S. W., Abdulla, M. A., Chua, L. S., and Hamdan, S. (2012) Potential bioactive property of *Polygonum minus* Huds (kesum) review. *Sci. Res. Essays* **7**(2), 90-93.

4. Ee, S.-F., Mohamed-Hussein, Z.-A., Othman, R., Shaharuddin, N. A., Ismail, I., and Zainal, Z. (2014) Functional characterization of Sesquiterpene synthase from *Polygonum minus*. *ScientificWorldJournal* **2014.**

5. Seman-Kamarulzaman, A.-F., Mohamed-Hussein, Z.-A., Ng, C. L., and Hassan, M. (2016) Novel NAD+-Farnesal Dehydrogenase from *Polygonum minus* Leaves. Purification and Characterization of Enzyme in Juvenile Hormone III Biosynthetic Pathway in Plant. *PLoS One* **11**(8), e0161707.

6. Shevchenko, A., Muzychkina, R., and Korul, D. Y. (2016) Amino-and Fatty-Acid Compositions of the Kazakhstan Plant *Polygonum minus*. *Chem. Nat. Compd.* **52**(4), 771-772.

7. Rusdi, N. A., Goh, H.-H., and Baharum, S. N. (2016) GC-MS/Olfactometric characterisation and aroma extraction dilution analysis of aroma active compounds in *Polygonum minus* essential oil. *Plant Omics* **9**(4), 289-294.

8. Roslan, N. D., Yusop, J. M., Baharum, S. N., Othman, R., Mohamed-Hussein, Z.-A., Ismail, I., Noor, N. M., and Zainal, Z. (2012) Flavonoid biosynthesis genes putatively identified in the aromatic plant *Polygonum minus* via expressed sequences tag (EST) analysis. *Int. J. Mol. Sci.* **13**(3), 2692-2706.

9. Loke, K.-K., Rahnamaie-Tajadod, R., Yeoh, C.-C., Goh, H.-H., Mohamed-Hussein, Z.-A., Noor, N. M., Zainal, Z., and Ismail, I. (2016) RNA-seq analysis for secondary metabolite pathway gene discovery in *Polygonum minus*. *Genomics Data* **7**12-13.

10. Shahbaaz, M., ImtaiyazHassan, M., and Ahmad, F. (2013) Functional annotation of conserved hypothetical proteins from *Haemophilus influenzae* Rd KW20. *PloS one* **8**(12), e84263.

11. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**(17), 3389-3402.

12. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**(suppl 2), W244-W248.

13. Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., and Bairoch, A. (2007) Uniprotkb/swiss-prot. *Plant Bioinformatics: Methods and Protocols* 89-112.

14. Li, W., Jaroszewski, L., and Godzik, A. (2001) Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**(3), 282-283.

15. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., and Phan, I. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**(1), 365-370.

16. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2**(4), 953-971.

17. Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C., and Nakai, K. (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**(suppl 2), W585-W587.

18. Pierleoni, A., Martelli, P. L., Fariselli, P., and Casadio, R. (2006) BaCelLo: a balanced subcellular localization predictor. *Bioinformatics* **22**(14), e408-e416.

19. Blum, T., Briesemeister, S., and Kohlbacher, O. (2009) MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinf.* **10**(1), 1

20. Bhasin, M., Garg, A., and Raghava, G. (2005) PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics* **21**(10), 2522-2524.

21. Goldberg, T., Hecht, M., Hamp, T., Karl, T., Yachdav, G., Ahmed, N., Altermann, U., Angerer, P., Ansorge, S., and Balasz, K. (2014) LocTree3 prediction of localization. *Nucleic Acids Res.* **42**(W1), W350-W355.

22. Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**(3), 567-580.

23. Sonnhammer, E. L., Von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. in *Ismb* .

24. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**(4), 783-795.

25. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**(1), 1-6.

26. Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004) Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng., Des. Sel.* **17**(4), 349-356.

27. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., and Sangrador-Vegas, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**(D1), D279-D285.

28. Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* **313**(4), 903-919.

29. Sillitoe, I., Lewis, T. E., Cuff, A., Das, S., Ashford, P., Dawson, N. L., Furnham, N., Laskowski, R. A., Lee, D., and Lees, J. G. (2015) CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res.* **43**(D1), D376-D381.

30. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., and Hurwitz, D. I. (2014) CDD: NCBI's conserved domain database. *Nucleic Acids Res.* gku1221.

31. Marchler-Bauer, A., Anderson, J. B., DeWeese-Scott, C., Fedorova, N. D., Geer, L. Y., He, S., Hurwitz, D. I., Jackson, J. D., Jacobs, A. R., and Lanczycki, C. J. (2003) CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **31**(1), 383-387.

32. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A.* **95**(11), 5857-5864.

33. Letunic, I., Doerks, T., and Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* **43**(D1), D257-D260.

34. Rappoport, N., Linial, N., and Linial, M. (2013) ProtoNet: charting the expanding universe of protein sequences. *Nat. Biotechnol.* **31**(4), 290-292.

35. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., and Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**(Web Server issue), W252-258.

36. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**(6), 845-858.

37. Consortium, U. (2011) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* gkr981.

38. Kahr, W. H., Hinckley, J., Li, L., Schwertz, H., Christensen, H., Rowley, J. W., Pluthero, F. G., Urban, D., Fabbro, S., and Nixon, B. (2011) Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome. *Nat. Genet.* **43**(8), 738-740.

39. Schrag, J. D., Procopio, D. O., Cygler, M., Thomas, D. Y., and Bergeron, J. J. (2003) Lectin control of protein folding and sorting in the secretory pathway. *Trends Biochem. Sci* **28**(1), 49-57.

40. Burgess, A., Mornon, J.-P., de Saint-Basile, G., and Callebaut, I. (2009) A concanavalin A-like lectin domain in the CHS1/LYST protein, shared by members of the BEACH family. *Bioinformatics* **25**(10), 1219-1222.

41. Chang, C. P., Yang, M. C., Liu, H. S., Lin, Y. S., and Lei, H. Y. (2007) Concanavalin A induces autophagy in hepatoma cells and has a therapeutic effect in a murine in situ hepatoma model. *Hepatology* **45**(2), 286-296.

42. Lei, H.-Y., and Chang, C.-P. (2007) Induction of autophagy by concanavalin A and its application in anti-tumor therapy. *Autophagy* **3**(4), 402-404.

43. Li, W.-w., Yu, J.-y., Xu, H.-l., and Bao, J.-k. (2011) Concanavalin A: a potential anti-neoplastic agent targeting apoptosis, autophagy and anti-angiogenesis for cancer therapeutics. *Biochem. Biophys. Res. Commun.* **414**(2), 282-286.

44. Mikami, K., Iuchi, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2000) A novel *Arabidopsis thaliana* dynamin-like protein containing the pleckstrin homology domain. *J. Exp. Bot.* **51**(343), 317-318.

45. Stevenson, J. M., Perera, I. Y., and Boss, W. F. (1998) A phosphatidylinositol 4-kinase pleckstrin homology domain that binds phosphatidylinositol 4-monophosphate. *J. Biol. Chem.* **273**(35), 22761-22767.

46. Stevenson-Paulik, J., Love, J., and Boss, W. F. (2003) Differential regulation of two Arabidopsis type III phosphatidylinositol 4-kinase isoforms. A regulatory role for the pleckstrin homology domain. *Plant Physiol.* **132**(2), 1053-1064.

47. Delage, E., Ruelland, E., Zachowski, A., and Puyaubert, J. (2012) Eat in or take away? How phosphatidylinositol 4-kinases feed the phospholipase C pathway with substrate. *Plant Signaling Behav.* **7**(9), 1197-1199.

48. Singh, A., Bhatnagar, N., Pandey, A., and Pandey, G. K. (2015) Plant phospholipase C family: Regulation and functional role in lipid signaling. *Cell Calcium* **58**(2), 139-146.

49. Dowd, P. E., and Gilroy, S. (2010) The emerging roles of phospholipase C in plant growth and development. in *Lipid Signaling in Plants*, Springer. pp 23-37.

50. Cullinane, A. R., Schäffer, A. A., and Huizing, M. (2013) The BEACH Is Hot: A LYST of Emerging Roles for BEACH-Domain Containing Proteins in Human Disease. *Traffic* **14**(7), 749-766.

51. Barrett, A., and Hermann, G. J. (2016) A Caenorhabditis elegans Homologue of LYST Functions in Endosome and Lysosome-Related Organelle Biogenesis. *Traffic* **17**(5), 515-535.

52. Zhang, H., Fan, X., Bagshaw, R. D., Zhang, L., Mahuran, D. J., and Callahan, J. W. (2007) Lysosomal membranes from beige mice contain higher than normal levels of endoplasmic reticulum proteins. *J. Proteome Res.* **6**(1), 240-249.

53. Teh, O.-k., Hatsugai, N., Tamura, K., Fuji, K., Tabata, R., Yamaguchi, K., Shingenobu, S., Yamada, M., Hasebe, M., and Sawa, S. (2015) BEACH-domain proteins act together in a cascade to mediate vacuolar protein trafficking and disease resistance in Arabidopsis. *Mol. Plant* **8**(3), 389-398.

54. Steffens, A., Bräutigam, A., Jakoby, M., and Hülskamp, M. (2015) The BEACH domain protein SPIRRIG is essential for Arabidopsis salt stress tolerance and functions as a regulator of transcript stabilization and localization. *PLoS Biol* **13**(7), e1002188.

55. Rebecchi, M., and Scarlata, S. (1998) Pleckstrin homology domains: a common fold with diverse functions. *Annu. Rev. Biophys. Biomol. Struct.* **27**(1), 503-528.

56. Van Nocker, S., and Ludwig, P. (2003) The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. *BMC Genomics* **4**(1), 1.

57. Kwak, E., Gerald, N., Larochelle, D. A., Vithalani, K. K., Niswonger, M. L., Maready, M., and De Lozanne, A. (1999) LvsA, a Protein Related to the Mouse Beige Protein, Is Required for Cytokinesis in Dictyostelium. *Molecular Biology of the Cell* **10**(12), 4429-4439.

58. Li, D., and Roberts, R. (2001) Human Genome and Diseases: WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci.* **58**(14), 2085-2097.

59. Smith, T. F., Gaitatzes, C., Saxena, K., and Neer, E. J. (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem. Sci* **24**(5), 181-185.

60. Clemens, M. J., and Elia, A. (1997) The double-stranded RNA-dependent protein kinase PKR: structure and function. *J Interferon Cytokine Res* **17**(9), 503-524.

61. Pindel, A., and Sadler, A. (2011) The role of protein kinase R in the interferon response. *J. Interferon Cytokine Res.* **31**(1), 59-70.

62. Wu, W. I., Yajnik, J., Siano, M., and De Lozanne, A. (2004) Structure-Function Analysis of the BEACH Protein LvsA. *Traffic* **5**(5), 346-355.

63. Isakson, P., Holland, P., and Simonsen, A. (2013) The role of ALFY in selective autophagy. *Cell Death Differ.* **20**(1), 12-20.

64. Clausen, T. H., Lamark, T., Isakson, P., Finley, K. D., Larsen, K. B., Brech, A., Øvervatn, A., Stenmark, H., Bjørkøy, G., and Simonsen, A. (2010) p62/SQSTM1 and ALFY interact to facilitate the formation of p62 bodies/ALIS and their degradation by autophagy. *Autophagy* **6**(3), 330-344.

65. Filimonenko, M., Isakson, P., Finley, K. D., Anderson, M., Jeong, H., Melia, T. J., Bartlett, B. J., Myers, K. M., Birkeland, H. C., and Lamark, T. (2010) The selective macroautophagic degradation of aggregated proteins requires the PI3P-binding protein Alfy. *Mol. Cell* **38**(2), 265-279.

66. Han, S., Yu, B., Wang, Y., and Liu, Y. (2011) Role of plant autophagy in stress response. *Protein Cell* **2**(10), 784-791.

67. Bassham, D. C. (2007) Plant autophagy—more than a starvation response. *Curr. Opin. Plant Biol.* **10**(6), 587-593.

68. Liu, Y., and Bassham, D. C. (2012) Autophagy: pathways for self-eating in plant cells. *Annu. Rev. Plant Biol.* **63**215-237.

69. Lai, Z., Wang, F., Zheng, Z., Fan, B., and Chen, Z. (2011) A critical role of autophagy in plant resistance to necrotrophic fungal pathogens. *Plant J.* **66**(6), 953-968.

70. Slavikova, S., Ufaz, S., Avin-Wittenberg, T., Levanony, H., and Galili, G. (2008) An autophagy-associated Atg8 protein is involved in the responses of Arabidopsis seedlings to hormonal controls and abiotic stresses. *J. Exp. Bot.* **59**(14), 4029-4043.

71. Thompson, A. R., and Vierstra, R. D. (2005) Autophagic recycling: lessons from yeast help define the process in plants. *Curr. Opin. Plant Biol.* **8**(2), 165-173.

72. Minina, E. A., Sanchez-Vera, V., Moschou, P. N., Suarez, M. F., Sundberg, E., Weih, M., and Bozhkov, P. V. (2013) Autophagy mediates caloric restriction-induced lifespan extension in Arabidopsis. *Aging Cell* **12**(2), 327-329.

73. Avin-Wittenberg, T., Bajdzienko, K., Wittenberg, G., Alseekh, S., Tohge, T., Bock, R., Giavalisco, P., and Fernie, A. R. (2015) Global analysis of the role of autophagy in cellular metabolism and energy homeostasis in Arabidopsis seedlings under carbon starvation. *Plant Cell* **27**(2), 306-322.

74. Gou, W., Li, X., Guo, S., Liu, Y., Li, F., Xie, Q. (2019) Autophagy in Plant: A New Orchestrator in the Regulation of the Phytohormones Homeostasis. *Int J Mol Sci.* **20**(12):2900.

75. Rahnamaie-Tajadod, R., Goh, H. H., & Mohd Noor, N. (2019). Methyl jasmonate-induced compositional changes of volatile organic compounds in *Polygonum minus* leaves. *J Plant Physiol.* 240, 152994.

76. Christapher, P.V., Parasuraman, S., Christina, J.M., Asmawi, M.Z., Vikneswaran, M. (2015) Review on *Polygonum minus*. Huds, a commonly used food additive in Southeast Asia. *Pharmacognosy Res.* **7**(1):1-6.