МАТЕМАТИЧЕСКАЯ БИОЛОГИЯ, БИОИНФОРМАТИКА / MATHEMATICAL BIOLOGY, BIOINFORMATICS

DOI: https://doi.org/10.60797/jbg.2024.25.1

ASSESSMENT OF THE HOMOLOGY OF HOG1, 14-3-3 AND STE11 IN REPRESENTATIVES OF SYMBIOTIC MYCORRHIZAL FUNGI

Research article

Kuzmin P.A.[1], Malov V.O.[2], Krylov P.A.[3,] *
[1] ORCID : 0000-0002-1303-765X;
[2] ORCID : 0000-0003-2766-0124;
[1, 2, 3] Federal Scientific Center for Ecology, Integrated Land Reclamation and Protective Aforestation of the Russian Academy of Sciences, Volgograd, Russian Federation

* Corresponding author (krylov-p[at]vfanc.ru)

Abstract

Adaptation to drought of woody and shrubby plants is associated with species features/characteristics and the presence of symbiotic relationships with arbuscular mycorrhiza. There is a wide variety of arbuscular mycorrhiza in nature, the most widespread genera are: Glomus, Diversispora, Paraglomus, Acaulospora, Entrophospora, Gigaspora, Scutellospora. The signaling proteins Hog1, 14-3-3, Ste11 in arbuscular mycorrhiza fungi play an important role in triggering mechanisms of adaptation to stress in the symbiotic relationship "fungus-plant". In this regard, the goal of the work was to evaluate the homology of amino acid sequences of the signaling proteins Hog1, 14-3-3 and Ste11 in different symbiotic mycorrhizal fungi. For the study, their amino acid sequences were taken from the Protein database (NCBI) using the BLAST program, a search for protein homologues was performed, subsequently multiple alignment was performed by the MEGA11 program using the MUSCLE algorithm. Phylogenetic analysis was carried out in the MEGA11 program using the maximum likelihood method. As a result, high conservativity for the Hog1, 14-3-3 proteins and high variability for Ste11 were revealed. The divergence of a pair of homologues of the protein 14-3-3 in Glomeromycotina before the separation of the subphylum and Ste11 in Glomeraceae after the formation of the family was also shown. As a result, we determined that Ste11 is a promising protein of interest for searching for intraspecific variations associated with resistance to drought and other abiotic stress factors in the mushroom-plant system.

Keywords: Hog 1, 14-3-3, Ste11, arbuscular mycorrhiza, multiple alignment, phylogenetics, desertification.

ОЦЕНКА ГОМОЛОГИИ HOG1, 14-3-3 И STE11 У ПРЕДСТАВИТЕЛЕЙ СИМБИОТИЧЕСКИХ МИКОРИЗНЫХ ГРИБОВ

Научная статья

Кузьмин П.А.[1], Малов В.О.[2], Крылов П.А.[3,] *
[1] ORCID : 0000-0002-1303-765X;
[2] ORCID : 0000-0003-2766-0124;
[1, 2, 3] Федеральный научный центр агроэкологии, комплексных мелиораций и защитного лесоразведения Российской академии наук, Волгоград, Российская Федерация

* Корреспондирующий автор (krylov-p[at]vfanc.ru)

Аннотация

Адаптация к засухе древесно-кустарниковых растений связана с видовой спецификой и наличием симбиотических связей с арбускулярной микоризой. В природе отмечается широкое разнообразие арбускулярной микоризы, самыми распространенными таксономическими группами являются: *Glomus*, *Diversispora*, *Paraglomus*, *Acaulospora*, *Entrophospora*, *Gigaspora*, *Scutellospora*. В запуске механизмов адаптации к стрессу при симбиотической связи «гриб-растение» важную роль играют сигнальные белки Hog1, 14-3-3, Ste11 у представителей арбускулярной микоризы. В связи с этим целью работы стало оценить гомологию аминокислотных последовательностей сигнальных белков Hog1, 14-3-3 и Ste11 у разных симбиотических микоризных грибов. Для исследования были взяты их аминокислотные последовательности из базы данных Protein (NCBI) и при помощи программы BLAST был выполнен поиск гомологов белков в последующем с которыми проводилось множественное выравнивание программой MEGA11 с использованием алгоритма MUSCLE. Филогенетический анализ проводился с помощью программы MEGA11 методом максимального правдоподобия. В результате для белков Hog1, 14-3-3 была выявлена высокая консервативность и высокая вариабельность для Ste11. Также было показано расхождение пары гомологов белка 14-3-3 у *Glomeromycotina* до отделения подтипа и Ste11 у *Glomeraceae* после формирования семейства. В итоге мы определили, что Ste11 является перспективным белком интереса для поиска внутривидовых вариаций, ассоциированных с устойчивостью к засухе и прочим абиотическим факторам стресса в системе «гриб-растение».

Ключевые слова: Hog1, 14-3-3, Ste11, арбускулярная микориза, множественное выравнивание, филогенетика, опустынивание.

Introduction

Climate change leads to desertification and land degradation, which leads to in biodiversity of woody and shrubby plants decrease [1]. The problem of reducing desertification in southern regions and preventing degradation of agricultural lands stands before Russian science more and more sharply at present times. One of the methods to stop desertification, particularly, sands fixation is planting of several woody and shrubby plants, such as *Calligonum aphyllum* (Pall.) Gürke, *Tamarix ramosissima* Ledeb., *Haloxylon aphyllum* (Minkw.) Iljin, *Krascheninnikovia ceratoides* L. и *Ulmus pumila* L. Species features influences crucially on plants adaptation to drought, common in territories with arid climate, but important, sometimes, a key factor is symbiotic mycorrizal fingi presense, which provide nutrition and protection from unfavoravle factors, favor addditional protective mechanisms formation and, so, increasing plant sustainability to stress factors [2]. Mycorrizal arbuscular fungi in *C. aphyllum*, *T. ramosissima*, *H. aphyllum*, *K. ceratoides* and *U. pumila* are not yet studied [2], so their particular role in drought adaptation processes are not yet described. However, we can assume, that in the arid climate those genie ill be amongst mycorrizal species: *Glomus*, *Diversispora*, *Paraglomus*, *Acaulospora*, *Entrophospora*, *Gigaspora*, *Scutellospora* or other genie of symbiotic fungi. Now there are evidences of arbuscular mycorriza positive influence on plant's drought tolerance [2].

Drought tolerance increase was shown for plants *Glycine max* L., but species diversity of mycorrizal species was not studied in that research [3]. *Ipomoea batatas* (L.) Lam endured drought easier with inoculation of fungi genie *Glomus sp*.and *Acaulospora sp.* [4]. Majority of such resources were carried out on agricultural plants [5], but endomycorrizal fungi species diversity forming symbiosis with desert and semi-desert plants is not studied alike protective mechanisms formation in plants forming a symbiosis.

Proteins Hog1, 14-3-3, Ste11 are playing important role in drought adaptation mechanisms launching in arbuscular mycorriza fungi in symbiotic bond "fungus-plant". Hog1 and Ste11 are part of kinase cascade inhibiting large amount of genes, associated with drought tolerance [6], like genes of aquaporins, regulatory proteins 14-3-3. 14-3-3 is a large grope of proteins, genomes of many organisms contain several homologues of genes of this group. These genes are present in genomes of fungi, plants and animals. In fungi, these genes take part in many processes like neutral thregalase Nth1 activation [7] and change in expression of 220 genes in *Saccharomyces cerevisiae* genome [8]. Ste 11 is a MAPKKK kinase, signal carrier, participant of several signal pathways. Its involvement in fungi pheromones response [9], plant infection by pathogenic fungus [10] and drought tolerance [11] was discovered. These proteins are of interest in drought tolerance mechanisms formation research and possibilities of its improvement for plant and fungus system.

Regarding this, the goal of our research was to estimate amino-acid sequence homology of signal proteins Hog1, 14-3-3 and Ste11 amongst different mycorrizal fungi species for further study of symbiotic bond "fungus-plant" formation features in conditions of drought.

Research methods and principles

Signal proteins Hog1, 14-3-3, Ste11 presented among diverse species of arbuscular mycorriza were selected for research. Amino-acid sequences for alignment building were taken from Protein (NCBI) database and search for homologous of Hog1, 14-3-3 и Ste11 was carried out using program BLASAT. Protein sequences of Rhizophaus irregular, which genes are annotated and localised in chromosome, were selected as initial query. Homologous annotated as Hog1 were selected using instrument ProteinBLAST [11], in case of absence of annotated homologous – closest uncharacterized homology. Those species were selected to build an alignment:

1) *Rhizophaus irregularis* (XP_025185210.1);
2) *Rhizophaus diaphanus* (RGB42528.1);
3) *Rhizophaus clarus* (GBB87021.1);
4) *Funneliformis geosporum* (CAI2181536.1);
5) *Glomus cerebriforme* (RIA92778.1);
6) *Geosyphon pyriformis* (KAG9291825.1);
7) *Ambispora gerdemannii* (CAG8436886.1);
8) *Diversispora epigaea* (RHZ50009.1);
9) *Acaulospora morrowiae* (CAG8446886.1);
10) *Racocetra persica* (CAG8495386.1);
11) *Cetraspora pellucida* (CAG8501949.1);
12) *Dentiscutata erythropus* (CAG8739878.1);
13) *Umbelopsis sp.* PIM 123. (KAH8549606.1).

Proteins 14-3-3 and Ste11 were selected according to the same principle, but with slight changes in the species of the presented proteins, which are associated with the presence or absence of characterized amino acid sequences in the NCBI Protein databases [12]. 14-3-3 proteins of the species were taken:

1) *Rhizophagus irregularis* DAOM (XP 025173882.1);
2) *Rhizophaus irregularis* (CAJ16742.1, PKY38380.1);
3) *Glomus cerebriforme* (RIA97653.1);
4) *Gigaspora rosea* (RIB16586.1);
5) *Gigaspora margarita* (KAF0484528.1);
6) *Gigaspora rosea* (RIB2229.1);
7) *Glomus cerebriforme* (RIA79980.1);
8) *Gigaspora margarita* (KAF0420307.1);
9) *Funneliformis caledonium* (CAG8442663.1);
10) *Funneliformis geosporum* (CAI2173046.1);
11) *Acaulospora morrowiae* (CAG8439107.1);
12) *Cetraspora pellucida* (CAG8612119.1);
13) *Diversispora Epigaea* (RHZ47341.1);
14) *Ambispora gerdemannii* (CAG8474580.1);
15) *Geosiphon pyriformis* (KAG9289207.1);

16) *Dentiscutata erythropus* (CAG8685133.1);
17) *Racocetra persica* (CAG8610165.1);
18) *Umbelopsis sp* PMI 123 (KAH8548527.1).

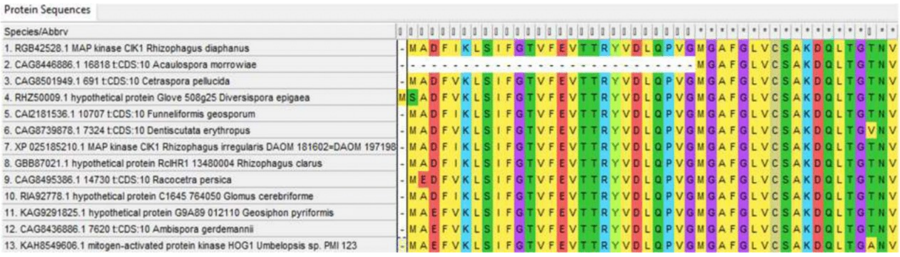
To search for the closest Ste11 homologues, the following amino acid sequences were taken from mycorrhizal symbiotic fungi:

1) *Rhizophagus irregularis* DAOM (XP 025186107.1);
2) *Rhizophaus irregularis* (XP 025172837.1);
3) *Rhizophagus diaphanus* (RGB34169.1, RGB38268.1);
4) *Rhizophagus clarus* (GES82803, GBB91729.1);
5) *Funneliformis caledonium* (CAG8474800.1, CAG8456822.1);
6) *Funneliformis geosporum* (CAI2186005.1);
7) *Acaulospora morrowiae* (CAG8664532.1, CAG8489781.1);
8) *Glomus cerebriforme* (RIA89482.1);
9) *Dentiscutata erythropus* (CAG8619186.1);
10) *Gigaspora rosea* (RIB01303.1);
11) *Gigaspora margarita* (CAG8472639.1);
12) *Geosiphon pyriformis* (KAG9304061.1);
13) *Ambispora leptoticha* (CAG8599465.1);
14) *Diversispora Epigaea* (RHZ44604.1);
15) *Racocetra fulgida* (CAG8484893.1);
16) *Umbelopsis sp* PMI 123 (CAG8583455.1, KAH8553767.1).

To identify differences between signaling proteins, bioinformatic analysis was performed. It included multiple alignment of the amino acid sequences of the studied proteins using the Mega 11 program (MEGA, Japan) [13] according to the MUSCLE algorithm [14] with default settings: penalty for gap opening -2.90, penalty for gap expansion 0, hydrophobicity multiplier 1.20. Alignment filtering was done manually with minimal changes: only the terminal unaligned regions were removed. Also, using MEGA 11, the most suitable model of amino acid substitutions was selected: LG + G for Hog 1 and 14-3-3, and JTT + G + F for Ste 11. Based on the amino acid sequences, a phylogenetic tree was constructed using the maximum likelihood method with bootstrap estimation -support, in the MEGA11 program (MEGA, Japan).

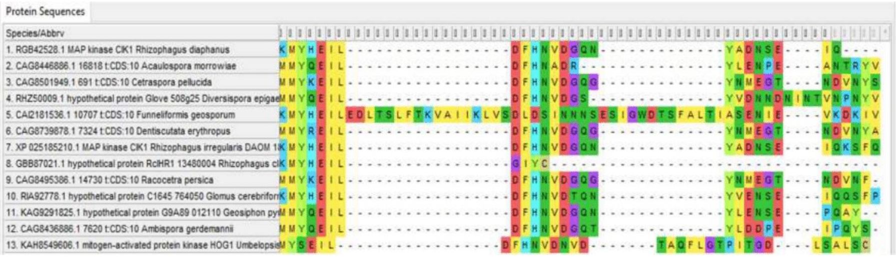
Main results and Discussion

The Hog1 alignment showed that this protein is highly conserved. The greatest distance was between *Umbelopsis sp.* PMI 123 and *Funneliformis geosporum*: 0.18947. *Umbelopsis* is the most distant taxon from the rest, it belongs to the subphylum *Mucoromycotina*, while all other species belong to the subphylum *Glomeromycotina*. Among the subphylum *Glomeromycotina*, the greatest distance was found between *Acaulospora morrowiae* and *Dentiscutata erythropus*: 0.11892 No significantly variable regions were found throughout the amino acid sequence (Fig. 1), with the exception of the C-terminal region, where the length of the variable sequence is up to 50 amino acids (Fig. 2).



Species/Abbrv	
1. RGB42528.1 MAP kinase CIK1 Rhizophagus diaphanus	-MADFIKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
2. CAG8446886.1 16818 t:CDS:10 Acaulospora morrowiae	----------------------------MGAFGLVCSAKDQLTGTNV
3. CAG8501949.1 691 t:CDS:10 Cetraspora pellucida	-MADFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
4. RHZ50009.1 hypothetical protein Glove 508g25 Diversispora epigaea	MSADFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
5. CAI2181536.1 10707 t:CDS:10 Funneliformis geosporum	-MADFIKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
6. CAG8739878.1 7324 t:CDS:10 Dentiscutata erythropus	-MADFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGVNV
7. XP 025185210.1 MAP kinase CIK1 Rhizophagus irregularis DAOM 181602=DAOM 197198	-MADFIKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
8. GBB87021.1 hypothetical protein RclHR1 13480004 Rhizophagus clarus	-MADFIKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
9. CAG8495386.1 14730 t:CDS:10 Racocetra persica	-MEDFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
10. RIA92778.1 hypothetical protein C1645 764050 Glomus cerebriforme	-MADFIKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
11. KAG9291825.1 hypothetical protein G9A89 012110 Geosiphon pyriformis	-MAEFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
12. CAG8436886.1 7620 t:CDS:10 Ambispora gerdemannii	-MAEFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGTNV
13. KAH8549606.1 mitogen-activated protein kinase HOG1 Umbelopsis sp. PMI 123	-MAEFVKLSIFGTVFEVTTRYVDLQPVGMGAFGLVCSAKDQLTGANV

Figure 1 - N-terminus of Hog1 alignment
DOI: https://doi.org/10.60797/jbg.2024.25.1.1



Species/Abbrv	
1. RGB42528.1 MAP kinase CIK1 Rhizophagus diaphanus	KMYHEIL-----------------DFHNVDGQN-------------YADNSE----IQ----
2. CAG8446886.1 16818 t:CDS:10 Acaulospora morrowiae	MMYQEIL-----------------DFHNADR---------------YLENPE----ANTRYV
3. CAG8501949.1 691 t:CDS:10 Cetraspora pellucida	MMYKEIL-----------------DFHNVDGQG-------------YNMEGT----NDVNYS
4. RHZ50009.1 hypothetical protein Glove 508g25 Diversispora epigae	MMYQEIL-----------------DFHNVDGS--------------YVDNNDNINTVNPNYV
5. CAI2181536.1 10707 t:CDS:10 Funneliformis geosporum	KMYHEILEDLTSLFTKVAIIKLVSDLDSINNNSESIGWDTSFALTIASENIE----VKDKIV
6. CAG8739878.1 7324 t:CDS:10 Dentiscutata erythropus	MMYREIL-----------------DFHNVDGQG-------------YNMEGT----NDVNYA
7. XP 025185210.1 MAP kinase CIK1 Rhizophagus irregularis DAOM 1	KMYHEIL-----------------DFHNVDGQN-------------YADNSE----IQKSFQ
8. GBB87021.1 hypothetical protein RclHR1 13480004 Rhizophagus cl	KMYHEIL-----------------GIYC----------------------------------
9. CAG8495386.1 14730 t:CDS:10 Racocetra persica	MMYKEIL-----------------DFHNVDGQG-------------YNMEGT----NDVNF-
10. RIA92778.1 hypothetical protein C1645 764050 Glomus cerebrifor	KMYHEIL-----------------DFHNVDTQN-------------YVENSE----IQQSFP
11. KAG9291825.1 hypothetical protein G9A89 012110 Geosiphon py	MMYQEIL-----------------DFHNVDGQN-------------YLENSE----PQAY--
12. CAG8436886.1 7620 t:CDS:10 Ambispora gerdemannii	MMYQEIL-----------------DFHNVDGQT-------------YLDDPE----IPQYS-
13. KAH8549606.1 mitogen-activated protein kinase HOG1 Umbelopsis	MYSEIL-----------------DFHNVDNVD-------TAQFLGTPITGD----LSALSC

Figure 2 - C-terminal variable regions of Hog1 proteins
DOI: https://doi.org/10.60797/jbg.2024.25.1.2

No major insertions or deletions were noted in the central region, with the exception of the missing first 27 amino acids in *Acaulospora morrowiae* and insertions at the C-terminus in *Funneliformis geosporum*. Hog1 phylogenetic tree, in general, repeats species phylogenetic structure. (Fig. 3).

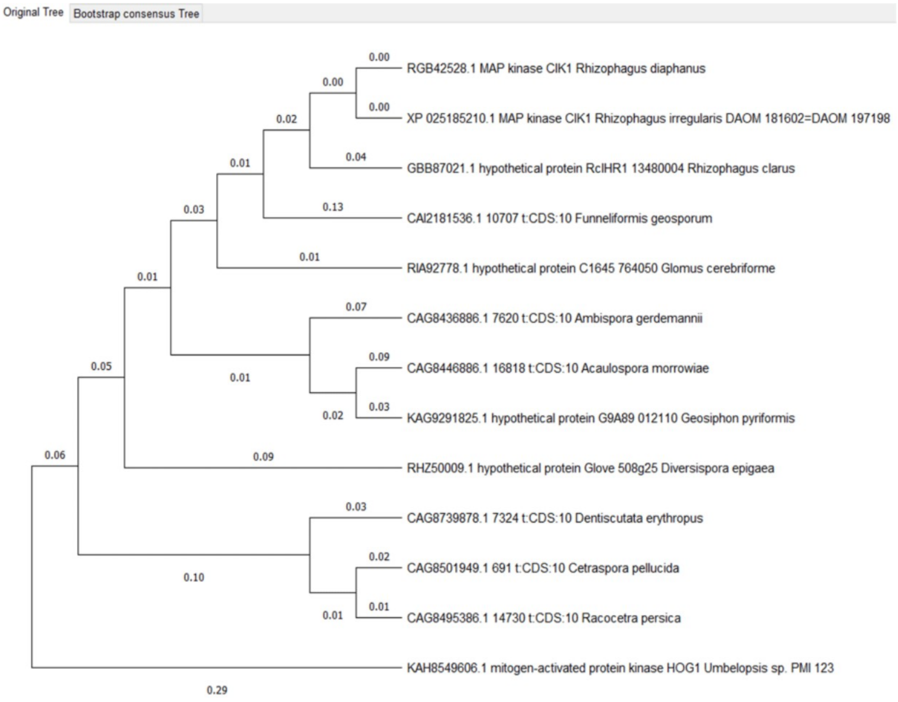
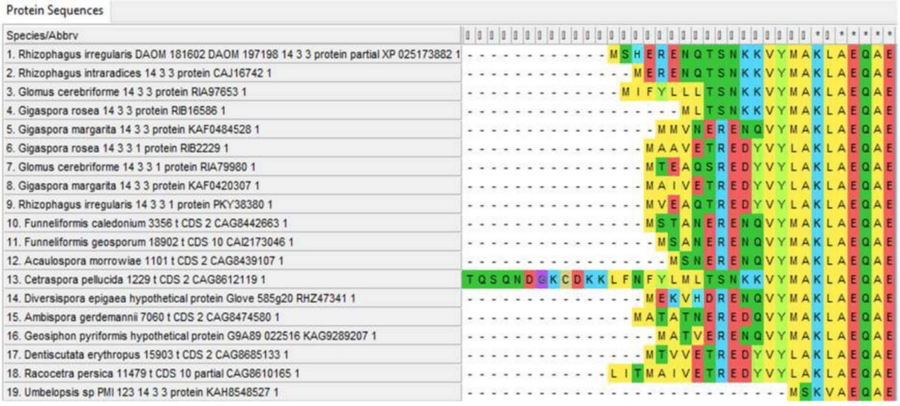


Figure 3 - Hog 1 proteins phylogenetic tree
DOI: https://doi.org/10.60797/jbg.2024.25.1.3

The most distant protein from the rest is *Umbelopsis sp*. PIM 123, which belongs to a different subphylum. Its branch is much longer than all the others. The proteins *Dentisculata Cetraspora* and *Racocetra*, belonging to the family *Gigasporaceae* of the order *Diversisporales*, are classified into a separate clade. All proteins of species of the order Glomerales also form a separate clade. The branch lengths of *Rhizophaus irregularis* and *Rhizophaus diaphanus* are zero, and indeed these proteins have no substitutions, but *Rhizophaus irregularis* has several additional amino acids at the C-terminus of the chain.

The clade consisting of the proteins *Geosyphon pyriformis*, *Ambispora gerdemannii* and *Acaulospora morrowiae* differs from the phylogenetic tree of organisms, *Ambispora* and *Geosyphon* belong to the order *Archaeosporales*, and *Acaulospora* belongs to the order *Diversisporales* and the family *Acaulosporaceae*.

14-3-3 protein is also conserved, featured by a small variable region at the N-terminus, as well as a longer sequence in *Cetraspora pellucida*, far beyond the general conserved part of the protein at the N-terminus. (Fig. 4)

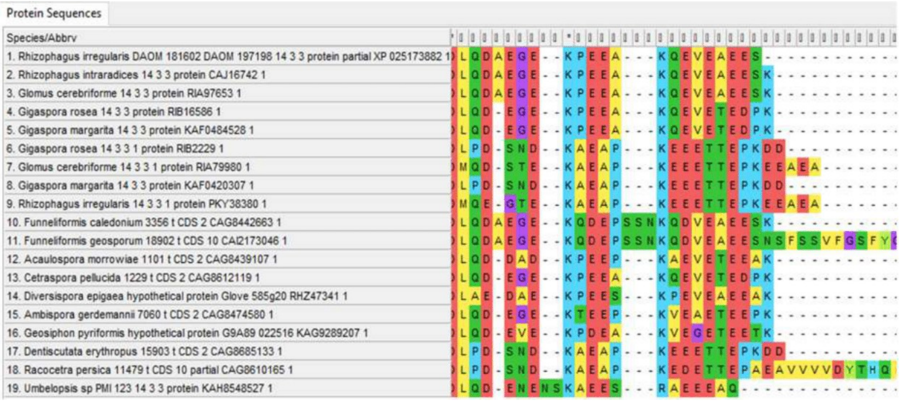


Protein Sequences

Species/Abbrv	
1. Rhizophagus irregularis DAOM 181602 DAOM 197198 14 3 3 protein partial XP 025173882 1	- - - - - - - - - - - - M S H E R E N Q T S N K K V Y M A K L A E Q A E
2. Rhizophagus intraradices 14 3 3 protein CAJ16742 1	- - - - - - - - - - - - - - M E R E N Q T S N K K V Y M A K L A E Q A E
3. Glomus cerebriforme 14 3 3 protein RIA97653 1	- - - - - - - - - - - - - M I F Y L L L T S N K K V Y M A K L A E Q A E
4. Gigaspora rosea 14 3 3 protein RIB16586 1	- - - - - - - - - - - - - - - - - - M L T S N K K V Y M A K L A E Q A E
5. Gigaspora margarita 14 3 3 protein KAF0484528 1	- - - - - - - - - - - - - - - - M M V N E R E N Q V Y M A K L A E Q A E
6. Gigaspora rosea 14 3 3 1 protein RIB2229 1	- - - - - - - - - - - - - - - M A A V E T R E D Y V Y L A K L A E Q A E
7. Glomus cerebriforme 14 3 3 1 protein RIA79980 1	- - - - - - - - - - - - - - - M T E A Q S R E D Y V Y L A K L A E Q A E
8. Gigaspora margarita 14 3 3 protein KAF0420307 1	- - - - - - - - - - - - - - - M A I V E T R E D Y V Y L A K L A E Q A E
9. Rhizophagus irregularis 14 3 3 1 protein PKY38380 1	- - - - - - - - - - - - - - - M V E A Q T R E D Y V Y L A K L A E Q A E
10. Funneliformis caledonium 3356 t CDS 2 CAG8442663 1	- - - - - - - - - - - - - - - M S T A N E R E N Q V Y M A K L A E Q A E
11. Funneliformis geosporum 18902 t CDS 10 CAI2173046 1	- - - - - - - - - - - - - - - - M S A N E R E N Q V Y M A K L A E Q A E
12. Acaulospora morrowiae 1101 t CDS 2 CAG8439107 1	- - - - - - - - - - - - - - - - - M S N E R E N Q V Y M A K L A E Q A E
13. Cetraspora pellucida 1229 t CDS 2 CAG8612119 1	T Q S Q N D G K C D K K L F N F Y L M L T S N K K V Y M A K L A E Q A E
14. Diversispora epigaea hypothetical protein Glove 585g20 RHZ47341 1	- - - - - - - - - - - - - - - M E K V H D R E N Q V Y M A K L A E Q A E
15. Ambispora gerdemannii 7060 t CDS 2 CAG8474580 1	- - - - - - - - - - - - - - M A T A T N E R E D Q V Y M A K L A E Q A E
16. Geosiphon pyriformis hypothetical protein G9A89 022516 KAG9289207 1	- - - - - - - - - - - - - - - - M A T V E R E N Q V Y M A K L A E Q A E
17. Dentiscutata erythropus 15903 t CDS 2 CAG8685133 1	- - - - - - - - - - - - - - - M T V V E T R E D Y V Y L A K L A E Q A E
18. Racocetra persica 11479 t CDS 10 partial CAG8610165 1	- - - - - - - - - - - - L I T M A I V E T R E D Y V Y L A K L A E Q A E
19. Umbelopsis sp PMI 123 14 3 3 protein KAH8548527 1	- M S K V A E Q A E

Figure 4 - N-terminal regions of 14-3-3 proteins
DOI: https://doi.org/10.60797/jbg.2024.25.1.4

No insertions or deletions in the internal part of the sequence were found in any species, but on the variable C-terminus additional amino acid sequences are present in *Funneliformis geosporum* and *Racocetra persica*. (Fig. 5)



Species/Abbrv	
1. Rhizophagus irregularis DAOM 181602 DAOM 197198 14 3 3 protein partial XP 025173882 1	LQDAEGE--KPEEA---KQEVEAEES-----------
2. Rhizophagus intraradices 14 3 3 protein CAJ16742 1	LQDAEGE--KPEEA---KQEVEAEESK----------
3. Glomus cerebriforme 14 3 3 protein RIA97653 1	LQDAEGE--KPEEA---KQEVEAEESK----------
4. Gigaspora rosea 14 3 3 protein RIB16586 1	LQD-EGE--KPEEA---KQEVETEDPK----------
5. Gigaspora margarita 14 3 3 protein KAF0484528 1	LQD-EGE--KPEEA---KQEVETEDPK----------
6. Gigaspora rosea 14 3 3 1 protein RIB2229 1	LPD-SND--KAEAP---KEEETTEPKDD---------
7. Glomus cerebriforme 14 3 3 1 protein RIA79980 1	MQD-STE--KAEAP---KEEETTEPKEEAEA------
8. Gigaspora margarita 14 3 3 protein KAF0420307 1	LPD-SND--KAEAP---KEEETTEPKDD---------
9. Rhizophagus irregularis 14 3 3 1 protein PKY38380 1	MQE-GTE--KAEAP---KEEETTEPKEEAEA------
10. Funneliformis caledonium 3356 t CDS 2 CAG8442663 1	LQDAEGE--KQDEPSSNKQDVEAEESK----------
11. Funneliformis geosporum 18902 t CDS 10 CAI2173046 1	LQDAEGE--KQDEPSSNKQDVEAEESNSFSSVFGSFY
12. Acaulospora morrowiae 1101 t CDS 2 CAG8439107 1	LQD-DAD--KPEEP---KAEVETEEAK----------
13. Cetraspora pellucida 1229 t CDS 2 CAG8612119 1	LQD-EGE--KPEEA---KQEVETEDPK----------
14. Diversispora epigaea hypothetical protein Glove 585g20 RHZ47341 1	LAE-DAE--KPEES---KPEVEAEEAK----------
15. Ambispora gerdemannii 7060 t CDS 2 CAG8474580 1	LQD-EGE--KTEEP---KVEAETEEPK----------
16. Geosiphon pyriformis hypothetical protein G9A89 022516 KAG9289207 1	LQD-EVE--KPDEA---KVEGETEETK----------
17. Dentiscutata erythropus 15903 t CDS 2 CAG8685133 1	LPD-SND--KAEAP---KEEETTEPKDD---------
18. Racocetra persica 11479 t CDS 10 partial CAG8610165 1	LPD-SND--KAEAP---KEDETTEPAEAVVVVDYTHQ
19. Umbelopsis sp PMI 123 14 3 3 protein KAH8548527 1	LQD-ENENSKAEES---RAEEEAQ-------------

Figure 5 - Variable C-terminal regions of 14-3-3 protein
DOI: https://doi.org/10.60797/jbg.2024.25.1.5

In the case of 14-3-3, two proteins were taken from some species, characterized as 14-3-3, but, nevertheless, the distances between these proteins are greater than between homologues from other species. For example, for *Rhizophaus irregularis* the distance between two 14-3-3 is 0.286386, and between 14-3-3 *Rhizophaus irregularis* (XP_025173882.1) and its closest homologue in *Gigaspora rosea* is 0.01988. Accordingly, the proteins present in two copies form two different clades, and the clade including 14-3-3 *Rhizophaus irregularis* PKY38380.1 diverges even earlier than the *Umbelopsis sp* PIM 123, indicating a very early divergence of 14-3-3 homologs, earlier than the divergence of subphylums. *Glomeromycotina* and *Mucoromycotina* (Fig. 6).

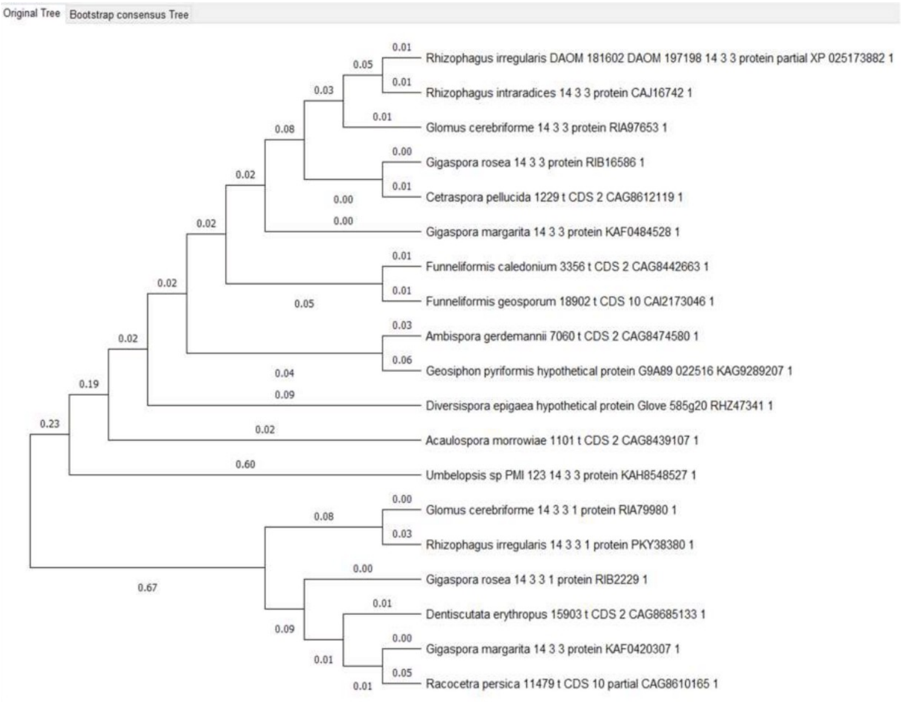
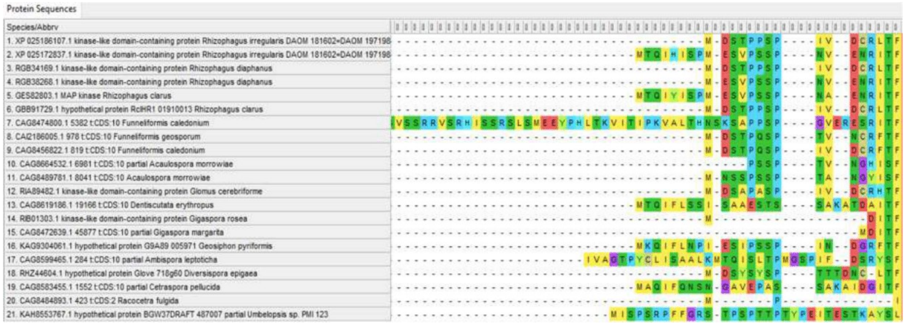


Figure 6 - 14-3-3 proteins phylogenetic tree
DOI: https://doi.org/10.60797/jbg.2024.25.1.6

The Ste11 protein is the least conserved of those presented in this article. It has variable regions not only at the ends (Fig. 7), but also in the central part of the sequence (Fig. 8).

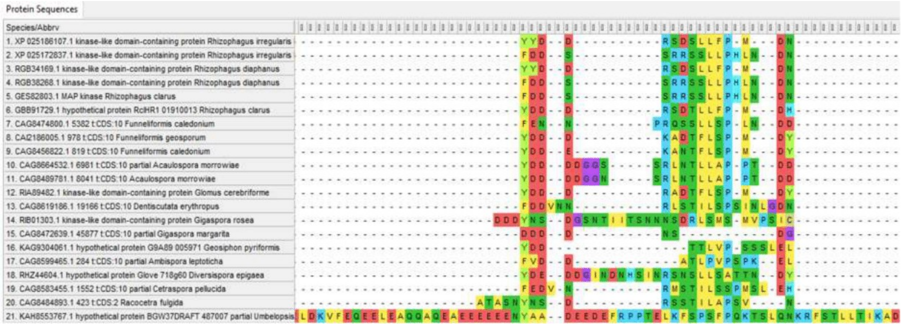


Protein Sequences

Species/Abbrv

1. XP 025186107.1 kinase-like domain-containing protein Rhizophagus irregularis DAOM 181602=DAOM 197198
2. XP 025172837.1 kinase-like domain-containing protein Rhizophagus irregularis DAOM 181602=DAOM 197198
3. RGB34169.1 kinase-like domain-containing protein Rhizophagus diaphanus
4. RGB38268.1 kinase-like domain-containing protein Rhizophagus diaphanus
5. GES82803.1 MAP kinase Rhizophagus clarus
6. GBB91729.1 hypothetical protein RcHR1 01910013 Rhizophagus clarus
7. CAG8474800.1 5382 t:CDS:10 Funneliformis caledonium
8. CAI2186005.1 978 t:CDS:10 Funneliformis geosporum
9. CAG8456822.1 819 t:CDS:10 Funneliformis caledonium
10. CAG8664532.1 6981 t:CDS:10 partial Acaulospora morrowiae
11. CAG8489781.1 8041 t:CDS:10 Acaulospora morrowiae
12. RIA89482.1 kinase-like domain-containing protein Glomus cerebriforme
13. CAG8619186.1 19166 t:CDS:10 Dentiscutata erythropus
14. RIB01303.1 kinase-like domain-containing protein Gigaspora rosea
15. CAG8472639.1 45877 t:CDS:10 partial Gigaspora margarita
16. KAG9304061.1 hypothetical protein G9A89 005971 Geosiphon pyriformis
17. CAG8599465.1 284 t:CDS:10 partial Ambispora leptoticha
18. RHZ44604.1 hypothetical protein Glove 718g60 Diversispora epigaea
19. CAG8583455.1 1552 t:CDS:10 partial Cetraspora pellucida
20. CAG8484893.1 423 t:CDS:2 Racocetra fulgida
21. KAH8553767.1 hypothetical protein BGW37DRAFT 487007 partial Umbelopsis sp. PMI 123

Figure 7 - Variable N-terminus of Ste11 proteins
DOI: https://doi.org/10.60797/jbg.2024.25.1.7



Protein Sequences

Species/Abbrv

1. XP 025186107.1 kinase-like domain-containing protein Rhizophagus irregularis
2. XP 025172837.1 kinase-like domain-containing protein Rhizophagus irregularis
3. RGB34169.1 kinase-like domain-containing protein Rhizophagus diaphanus
4. RGB38268.1 kinase-like domain-containing protein Rhizophagus diaphanus
5. GES82803.1 MAP kinase Rhizophagus clarus
6. GBB91729.1 hypothetical protein RcIHR1 01910013 Rhizophagus clarus
7. CAG8474800.1 5382 t:CDS:10 Funneliformis caledonium
8. CAI2186005.1 978 t:CDS:10 Funneliformis geosporum
9. CAG8456822.1 819 t:CDS:10 Funneliformis caledonium
10. CAG8664532.1 6981 t:CDS:10 partial Acaulospora morrowiae
11. CAG8489781.1 8041 t:CDS:10 Acaulospora morrowiae
12. RIA89482.1 kinase-like domain-containing protein Glomus cerebriforme
13. CAG8619186.1 19166 t:CDS:10 Dentiscutata erythropus
14. RIB01303.1 kinase-like domain-containing protein Gigaspora rosea
15. CAG8472639.1 45877 t:CDS:10 partial Gigaspora margarita
16. KAG9304061.1 hypothetical protein G9A89 005971 Geosiphon pyriformis
17. CAG8599465.1 284 t:CDS:10 partial Ambispora leptoticha
18. RHZ44604.1 hypothetical protein Glove 718g60 Diversispora epigaea
19. CAG8583455.1 1552 t:CDS:10 partial Cetraspora pellucida
20. CAG8484893.1 423 t:CDS:2 Racocetra fulgida
21. KAH8553767.1 hypothetical protein BGW37DRAFT 487007 partial Umbelopsis

Figure 8 - Insertions and variable regions in the internal part of Ste11 proteins amino-acid sequence
DOI: https://doi.org/10.60797/jbg.2024.25.1.8

Also in the internal part of the sequence there are long insertions in the protein sequences of *Umbelopsis sp*. PMI 123, *Diversispora epigaea*, *Acaulospora morrowiae*, *Funneliformis caledonium*, *Ambispora leptoticha*. Many species also have two Ste11 genes in their genomes, and, as in the case of 14-3-3, the distance between them is greater than between the closest homologues from other species, for example, the distance between *Rhizophagus irregularis* XP_025186107.1 and *Rhizophagus irregularis* XP_025172837.1 is 0.30817, while the distance between *Rhizophagus irregularis* XP_025186107.1 and *Funneliformis caledonium* CAG8456822.1 is 0.15729. But, unlike 14-3-3, all proteins of species of the family Glomeraceae converge into one clade, which indicates the divergence of these proteins after the formation of the family (Fig. 9).

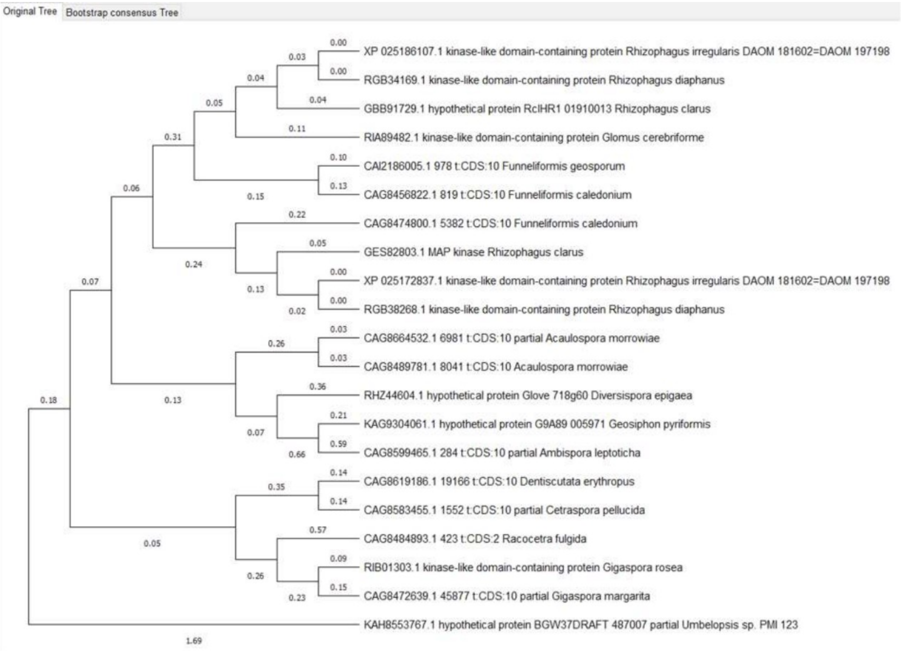


Figure 9 - Ste11 proteins phylogenetic tree
DOI: https://doi.org/10.60797/jbg.2024.25.1.9

Both 14-3-3 and Hog1 proteins are highly conserved, making them unsuitable for use in phylogenetics. Phylogenetic analysis is also complicated by the presence in the genomes of fungi of several homologues of proteins of these groups, the differences in functions of which have not yet been studied.

However, variability in the sequences of these proteins is present not only at the ends of the chain, which is typical for most proteins, but also in the central part. And although at the moment there is no data on the variability of the amino acid sequences of these proteins within species and the differences between different types of endomycorrhizal fungi in relation to the drought resistance factor, such data can be obtained in the framework of further studies, which leaves hope for linking variations in these proteins with the resistance of the “fungus-plan” pair to drought and its mechanisms.

Promoter regions of the genes of the studied proteins also require attention. Since the proteins themselves are highly conserved, markers for selecting symbiotic fungi as inocula should be sought in the promoter regions of their genes, attention to upstream and downstream regulators should be paid as well.

The Ste11 protein is much more variable; the distance to other proteins can reach 0.80988 (*Umbelopsis* and *Diversispora epigaea*); within *Glomeromycotina*, distances often exceed 0.4. This variability of the protein gives hope for detecting variations in its sequence associated with drought resistance of the “fungus-plant” system, which makes it a priority object for further research.

Conclusion

Alignments and phylogenetic trees were constructed for the Hog1, 14-3-3 and Ste11 proteins, which showed high conservativity of the Hog1 14-3-3 proteins and variability of the Ste11 protein. The divergence of the homologue pair 14-3-3 in Glomeromycotina before subphylum separation and Ste11 in Glomeraceae after the formation of the family was shown. Ste11 has been identified as a promising protein of interest for the search for intraspecific variations associated with resistance to drought and other abiotic stress factors.

Финансирование

24-26-00174 Видовое разнообразие арбускулярной эндомикоризы и ее влияние на засухоустойчивость ксерофитных растений аридных территорий Юга России.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Funding

24-26-00174 Species diversity of arbuscular endomycorrhizae and its effect on drought resistance of xerophytic plants in arid territories of Southern Russia.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Куликов А.И. О глобальном изменении климата и его экосистемных следствиях / А.И. Куликов, Л.Л. Убугунов // Аридные экосистемы. — 2014. — 3(60). — с. 5-13.
2. Wang Q. The role of arbuscular mycorrhizal symbiosis in plant abiotic stress / Q. Wang, M. Liu, Z. Wang et al. // Front Microbiol. — 2024. — 14. — p. 1323881. — DOI: 10.3389/fmicb.2023.1323881.
3. Pavithra D. Arbuscular mycorrhizal fungi inoculation enhances drought stress tolerance of plants / D. Pavithra, N. Yapa // Groundw. Sustain. Develop. — 2018. — 7. — p. 490–494. — DOI: 10.1016/j.gsd.2018.03.005.
4. Yooyongwech S. Arbuscular mycorrhizal fungi (AMF) improved water deficit tolerance in two different sweet potato genotypes involves osmotic adjustments via soluble sugar and free proline / S. Yooyongwech, T. Samphumphuang, R. Tisarum et al. // Sci. Hortic. — 2016. — 198. — p. 107-117. — DOI: 10.1016/j.scienta.2015.11.002.
5. Bahadur A. Mechanistic Insights into Arbuscular Mycorrhizal Fungi-Mediated Drought Stress Tolerance in Plants / A. Bahadur, A. Batool, F. Nasir et al. // Int J Mol Sci. — 2019. — 20(17). — p. 4199. — DOI: 10.3390/ijms20174199.
6. Wang S. Host- and virus-induced gene silencing of HOG1-MAPK cascade genes in Rhizophagus irregularis inhibit arbuscule development and reduce resistance of plants to drought stress / S. Wang, X. Xie, X. Che et al. // Plant Biotechnol J. — 2023. — 21(4). — p. 866-883. — DOI: 10.1111/pbi.14006.
7. Kopecka M. Role of the EF-hand-like motif in the 14-3-3 protein-mediated activation of yeast neutral trehalase Nth1 / M. Kopecka, D. Kosek, Z. Kukacka et al. // J. Biol. Chem. — 2014. — 289(20). — p. 13948-13961. — DOI: 10.1074/jbc.M113.544551.
8. Ichimura T. Transcriptomic and proteomic analysis of a 14-3-3 gene-deficient yeast / T. Ichimura, H. Kubota, T. Goma et al. // Biochemistry. — 2004. — 43(20). — p. 6149-6158. — DOI: 10.1021/bi035421i.
9. Costa A.C.B.P. Signal-mediated localization of Candida albicans pheromone response pathway components / A.C.B.P. Costa, R.P. Omran, C. Law et al. // G3 (Bethesda). — 2021. — 11(3). — p. jkaa033. — DOI: 10.1093/g3journal/jkaa033.
10. Wang X. Mitogen-activated protein kinase cascade CgSte50-Ste11-Ste7-Mk1 regulates infection-related morphogenesis in the poplar anthracnose fungus Colletotrichum gloeosporioides / X. Wang, D. Lu, C. Tian // Microbiol. Res. — 2021. — 248. — p. 126748. — DOI: 10.1016/j.micres.2021.126748.
11. Altschul S.F. Basic local alignment search tool / S.F. Altschul, W. Gish, W. Miller et al. // J. Mol. Biol. — 1990. — 215. — p. 403-410.
12. NCBI Protein. — 2024 — URL: https://www.ncbi.nlm.nih.gov/protein (accessed: 10.04.2024)

13. Tamura K. MEGA11: Molecular Evolutionary Genetics Analysis version 11 / K. Tamura, G. Stecher, S. Kumar // Molecular Biology and Evolution. — 2021. — 38. — p. 3022-3027.

14. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput / R.C. Edgar // Nucleic Acids Res. — 2004. — 32(5). — p. 1792-1797. — DOI: 10.1093/nar/gkh340.

Список литературы на английском языке / References in English

1. Kulikov A.I. O global'nom izmenenii klimata i ego ekosistemnyh sledstvijah [Global Climate Change And Its Impact On Ecosystems] / A.I. Kulikov, L.L. Ubugunov // Arid Ecosystems. — 2014. — 3(60). — p. 5-13. [in Russian]

2. Wang Q. The role of arbuscular mycorrhizal symbiosis in plant abiotic stress / Q. Wang, M. Liu, Z. Wang et al. // Front Microbiol. — 2024. — 14. — p. 1323881. — DOI: 10.3389/fmicb.2023.1323881.

3. Pavithra D. Arbuscular mycorrhizal fungi inoculation enhances drought stress tolerance of plants / D. Pavithra, N. Yapa // Groundw. Sustain. Develop. — 2018. — 7. — p. 490–494. — DOI: 10.1016/j.gsd.2018.03.005.

4. Yooyongwech S. Arbuscular mycorrhizal fungi (AMF) improved water deficit tolerance in two different sweet potato genotypes involves osmotic adjustments via soluble sugar and free proline / S. Yooyongwech, T. Samphumphuang, R. Tisarum et al. // Sci. Hortic. — 2016. — 198. — p. 107-117. — DOI: 10.1016/j.scienta.2015.11.002.

5. Bahadur A. Mechanistic Insights into Arbuscular Mycorrhizal Fungi-Mediated Drought Stress Tolerance in Plants / A. Bahadur, A. Batool, F. Nasir et al. // Int J Mol Sci. — 2019. — 20(17). — p. 4199. — DOI: 10.3390/ijms20174199.

6. Wang S. Host- and virus-induced gene silencing of HOG1-MAPK cascade genes in Rhizophagus irregularis inhibit arbuscule development and reduce resistance of plants to drought stress / S. Wang, X. Xie, X. Che et al. // Plant Biotechnol J. — 2023. — 21(4). — p. 866-883. — DOI: 10.1111/pbi.14006.

7. Kopecka M. Role of the EF-hand-like motif in the 14-3-3 protein-mediated activation of yeast neutral trehalase Nth1 / M. Kopecka, D. Kosek, Z. Kukacka et al. // J. Biol. Chem. — 2014. — 289(20). — p. 13948-13961. — DOI: 10.1074/jbc.M113.544551.

8. Ichimura T. Transcriptomic and proteomic analysis of a 14-3-3 gene-deficient yeast / T. Ichimura, H. Kubota, T. Goma et al. // Biochemistry. — 2004. — 43(20). — p. 6149-6158. — DOI: 10.1021/bi035421i.

9. Costa A.C.B.P. Signal-mediated localization of Candida albicans pheromone response pathway components / A.C.B.P. Costa, R.P. Omran, C. Law et al. // G3 (Bethesda). — 2021. — 11(3). — p. jkaa033. — DOI: 10.1093/g3journal/jkaa033.

10. Wang X. Mitogen-activated protein kinase cascade CgSte50-Ste11-Ste7-Mk1 regulates infection-related morphogenesis in the poplar anthracnose fungus Colletotrichum gloeosporioides / X. Wang, D. Lu, C. Tian // Microbiol. Res. — 2021. — 248. — p. 126748. — DOI: 10.1016/j.micres.2021.126748.

11. Altschul S.F. Basic local alignment search tool / S.F. Altschul, W. Gish, W. Miller et al. // J. Mol. Biol. — 1990. — 215. — p. 403-410.

12. NCBI Protein. — 2024 — URL: https://www.ncbi.nlm.nih.gov/protein (accessed: 10.04.2024)

13. Tamura K. MEGA11: Molecular Evolutionary Genetics Analysis version 11 / K. Tamura, G. Stecher, S. Kumar // Molecular Biology and Evolution. — 2021. — 38. — p. 3022-3027.

14. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput / R.C. Edgar // Nucleic Acids Res. — 2004. — 32(5). — p. 1792-1797. — DOI: 10.1093/nar/gkh340.