Spectrum of proline-specific peptidases in dikarya fungi

N.A. Alkin[1], Y.E. Dunaevsky[2], M.A.Belozersky[2], G.A.Belyakova[1], V.F. Tereshchenkova[3], I.Y.Filippova[3], E.N. Elpidina[2]

[1]Department of Mycology and Algology, Faculty of Biology, Moscow State University, Moscow, Russia. nikita9801@mail.ru; *[2]Department of Plant Proteins, A.N.Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia.* dun@genebee.msu.ru, *[3]Division of Natural Compounds, Chemical Faculty, Moscow State University, Moscow, Russia.*

Introduction. In constrast to 19 proteinogenic amino acids, proline is an imino acid widely distributed in the majority of proteins, and the bonds formed by this residue are resistant to peptidases with broad specificity, but can be hydrolyzed by proline-specific peptidases (PSP). The most common needs of proline-created peptide bond hydrolysis are digestion of proline-rich substrates (for instance, plant prolamins) and degradation of bioactive peptides [1-2]. PSP are found both in prokaryotic and eukaryotic organisms, still up-to-date the information about the distribution of these enzymes among fungal species is limited. The hydrolysis of proteins, which contain proline residue, is further valuable for food industry since it was demonstrated that proline-rich peptides may contribute a bitter taste to food products [3].

The objective of this research is a search for genes of PSP in genomes of different species of dikaryotic fungi (Ascomycota and Basidiomycota) and subsequent analysis of predicted PSP sequences.

Materials and methods. The search of sequenced and annotated fungal genomes was accomplished in NCBI public database (https://www.ncbi.nlm.nih.gov/). PSP sequences in the genomes of chosen fungal species were obtained by BLAST cluster services (https://blast.ncbi.nlm.nih.gov/Blast.cgi, blastx for acid prolyl endopeptidase (PEP), blastp in other cases) using as a query 13 human PSP sequences, 2 nonspecific human exopeptidases, capable of hydrolyzing the bonds formed by the carboxy group of any amino acids including proline: leucyl aminopeptidase (LAP) and cytosolic non-specific dipeptidase (CND), and well characterized *Aspergillus niger* acid PEP. Amino acid sequences of human PSP and information about their active sites was achieved in UniProt database (http://www.uniprot.org/), nucleotide sequence of acid PEP from *A. niger* was obtained from StrainInfo database (http://www.straininfo.net/). Multiple alignment of amino acid sequences and phylogenetic tree building were performed via Clustal Omega service (http://www.ebi.ac.uk/Tools/msa/clustalo/), mutations in the active sites were identified in comparison to the amino acid sequences of human PSP. Potential transmembrane domains of the detected proteins were analyzed using TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) and potential signal peptides were

predicted through SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP).

Results and discussion. We have analyzed 33 fungal species with sequenced genomes, which have diverse systematic position (25 species from 3 subphyla of Ascomycota, 8 species from 3 subphyla of Basidiomycota), life forms (mycelial and yeast) and ecological niches (13 biotrophic and necrotrophic plant pathogens, 11 xylotrophic and saprotrophic free-living fungi, 7 animal parasites and nematophagous fungi, etc.).

Fungal genomes contained homologs for 10 from 15 examined human peptidases, some of which were presented in several copies. Fungal homologs to dipeptidyl peptidases (DPP) 6, 8, 9, 10 and aminopeptidase 2 (APP 2) were absent in all species. In contrast, some PSP, such as APP 1 and APP 3 were present in all investigated fungal species. It is interesting that there may be a taxon-specific pattern of the PSP distribution among the fungal species. Indeed, prolyl oligopeptidase (POP) and LAP were present in the genomes of every checked Basidiomycota species, and were absent in all studied species from phylum Ascomycota, with the only exception of *Beauveria bassiana* from the Cordicipitaceae family, whereas the other examined species of this family (*Cordyceps confragosa* and *Metarhizium anisopliae*) lacked homologs for POP and LAP. It should be noted that the evidence of POP activity in Basidiomycota was achieved previously in *Agaricus bisporus* only [4].

Homologs of *A. niger* acid PEP were detected in all examined species excluding those, which are members of Saccharomycotina subphylum, characterized by yeast or yeast-mycelial life forms. All analyzed fungal species had homologs for human DPP 4, except the members of Magnaporthales order (*Magnaporthe oryzae* and *Gaeumannomyces tritici*).

Fibroblast activation protein (FAP) was detected in seven fungal species, all of which are members of Sordariomycetes class. However, *B. bassiana* from this clade lacked homologs of FAP. Still it should be stressed out that FAP belongs to the same peptidase family as widespread DPP 4 does, therefore the lack of FAP gene could be connected with the presence of two DPP 4 homologs in *B. bassiana* genome.

Prolyl carboxypeptidase (PRCP) as well as *A. niger* acid PEP belong to the S28 family of MEROPS classification, and close human PRCP homolog was found in nematophagous fungus *Pochonia chlamydosporia*, while genomes of the other examined species lacked close PRCP homologs.

Prolidase (XPD) homologs were detected in the majority of the examined fungal species, however the distribution of this enzyme among taxa is rather mosaic: homologs of XPD were found in genomes of ascomycetes and basidiomycetes, still both taxa have exceptions which lack

this enzyme (for example, *Taphrina deformans* from Ascomycota, *Ustilago maydis* and *Tilletia indica* from Basidiomycota).

Human dipeptidyl peptidase 2 (DPP 2) reveals low level of homology with fungal proteins, only 2 possible fungal homologs were detected in *B. bassiana* and *Colletotrichum tofieldiae*.

Most of the detected PSP sequences have the same active site composition as human PSP. The conservative amino acid triad of serine peptidases (serine, histidine, and aspartic acid) is present in each fungal homologs of DPP 4, FAP and POP, and this is the evidence in favor of enzyme activity of these peptidases. A wide dissemination and distinct conservatism of DPP 4 and similar PSPs among fungal species may point to a high biological prominence of these PSP.

Analysis of the fungal PSPs on the presence of potential transmembrane domains demonstrates that fungal species, which have one homolog of human transmembrane PSP, also contain predicted transmembrane region. In case of several homologs of one transmembrane human PSP in the fungal genome, at least one of them contains transmembrane domain while other (others) include potential signal peptide, i.e. it is a secreted protein. Thus, it may be assumed that these PSPs are paralogs and perform different functions.

A phylogenetic tree of the examined fungal species based on amino acid sequence of DPP 4 proteins was built. It consists of three clades: DPP 4 homologs in Basidiomycota, transmembrane DPP 4 homologs of Ascomycota and secreted DPP 4 homologs of Ascomycota. When considering each of these clusters separately, remarkable resemblance with the results gained in major reviews on fungal systematics is achieved [5-6]. It is revealing that the matching between the phylogenetic tree of DPP 4 sequences and the independent data from aforementioned works is evident both at high-level taxa (phyla) and at much smaller taxa like families and genera. This result may support homology of the detected polypeptides and demonstrates a potential opportunity of using widespread PSPs as markers in phylogenetic analysis.

Conclusions.

1. For the first time, a systematic analysis of PSP spectrum in fungal genomes was performed.
2. Analysis of 33 fungal genomes revealed homologs for 10 from 15 examined human peptidases, capable of hydrolyzing the bonds formed by proline residue, and one specific for fungi acid PEP.
3. APP1, APP3, DPP4, CND and PEP were present in the majority of studied species; POP and LAP were found mainly in Basidiomycetes, FAP was specific to Sordariomycetes, XPD was distributed mosaically, and PRCP was found in single species of fungi.

4. Presence or lack of the examined PSP in fungal species is mostly determined by taxonomical position.

The work was supported by the RFBR grants № 15-03-06675, 15-04-08523, 15-04-08689, 16-34-01012_мол-а.

References.

1. V. F. Tereshenkova et al. Dipeptidyl peptidase 4 – an important digestive peptidase in *Tenebrio molitor* larvae, *Insect Biochem Mol Biol*, **76**:38-48.

2. D. F. Cunningham, B. O'Connor (1997) Proline specific peptidases, *BBA*, **1343:**160–186.

3. L. Lemieux, Re Simard (1992) Bitter flavour in dairy products. II. A review of bitter peptides from caseins: their formation, isolation and identification, structure masking and inhibition, *Lait*, **72**:335-382.

4. A. K. Sattar et al. (1990) Purification and characterization of an extracellular prolyl endopeptidase from *Agaricus bisporus*, *J Biochem*, **107**:256-261.

5. T. Y. James et al. (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny, *Nature*, **443**:818-882.

6. P. B. Matheny et al. (2006) Major clades of Agaricales: a multilocus phylogenetic overview, *Mycologia*, **98**:982-995.