

The glyphosate target enzyme 5-enolpyruvyl shikimate 3-phosphate synthase (EPSPS) contains several EPSPS-associated domains in fungi

► **Tuomas Tall** ^{1,*}, **Pere Puigbò**, PhD ^{1,2,3}

1. University of Turku; 2. Eurecat Technology Center of Catalonia; 3. Rovira i Virgili University; * tall@utu.fi

Introduction

Glyphosate is one of the most used herbicides against weeds that targets the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). EPSPS is the central enzyme in the shikimate pathway to synthesize 3 essential amino acids in plants, fungi and prokaryotes. Although this pathway is not found in animals, herbicide may affect the biodiversity of free-living and host-associated microorganisms. Thus, the study of the glyphosate target enzyme will provide clues about the potential effect of the herbicide in organisms. In recent studies, we analyzed the distribution of the EPSPS protein in organisms and we estimated the differential sensitivity to the herbicide.

Aims

In this study, we analyze the distribution and evolution of the EPSPS and EPSPS-associated domains. In plants and prokaryotes, the EPSPS is a single-domain, thus the potential sensitivity to the herbicide glyphosate depends on the type of amino acids in the EPSPS proteins. However, the effect of the herbicide on the multi-domain structure of the EPSPS fungi is still unclear.

Methods and Materials

We obtained protein data of the EPSPS and EPSPS-associated domains from the Pfam database (<http://pfam.xfam.org>). First, we analyzed frequencies of EPSPS-associated domains across 390 fungal species, and then we summarized the main domain function in a Venn's diagram. We also analyzed the distribution of the associated domains by taxonomical groups in fungi. Cytoscape was used to reconstruct a bipartite network of protein domains and fungal species. The analysis of the evolution of the EPSPS-associated domains in fungi was performed by Dolan parsimony with the program Count.

Results

The EPSPS-associated domains can be classified in four partially overlapping groups (figure 2): shikimate (shikimate pathway proteins), enzymes (proteins with catalytic function), expression (domains whose products are needed in controlling gene expression) and structural function (proteins that do not have a catalytic).

A total of 22 domains EPSPS-associated domains can be found in EPSPS proteins (in diverse domain architectures) across 390 fungal species (Table 1). Ascomycota is the most variable phylum, in terms of EPSPS-associated domains, and contains several rare domains. However, it is the phylum that contains least presence of the Shikimate DH domain (figure 1A).

Majority of multi-domain architectures (220 out of 390) contain five domains and approximately 1/3 of the sequences (134 out of 390) have all six domains (Table 1). The most common multi domain architecture is composed by a group of five enzymes (HQ synthase, EPSPS, SKI, DHquinase I and Shikimate dh N). The second most common architecture, includes Shikimate DH as sixth domain (Figure 1b and table 1).

The Dolan parsimony analysis (figure 3) shows that the domain Shikimate DH domain is likely present in the common ancestor of fungi and lost in some lineages. Moreover, rarer domains are late inclusions to the multi-domain structure.

References

- Leino L, Tall T, Helander M, Saloniemi I, Saikonen K, Ruuskanen S, Puigbò P. (2020) Classification of the glyphosate target enzyme (5-enolpyruvylshikimate-3-phosphate synthase). *Biorxiv* (pre-print). <https://doi.org/10.1101/2020.05.27.118265>
- Ronrio MJ, Ruuskanen S, Helander M, Saikonen K, Saloniemi I, Puigbò P. (2020) Adaptation of bacteria to glyphosate: a microevolutionary perspective of the enzyme 5-enolpyruvylshikimate 3-phosphate (EPSP) synthase. *Biorxiv* (pre-print). <https://doi.org/10.1101/2020.06.16.154065>

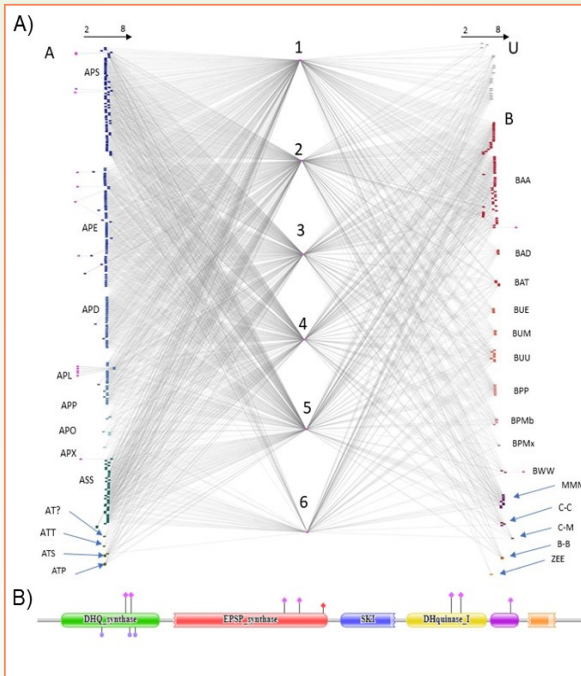


Figure 1. A bipartite network (made with cytoscape) of fungal species and EPSPS-associated domains. A) A bipartite network of fungal samples. Letters stands for different phyla and classes of fungi. On the left side there are fungi of the phylum ascomycota (A) and on the right side there are fungi of phylum Basidiomycota (B) Mucoromycota (M), Chytridiomycota (C), Blastocladiomycota (Bl) and Zoopagomycota (Z), Unknown taxa (U). Numbers correspond to fungal multidomain: 1) DH synthase 2) EPSP synthase 3) SKI 4) DH quinase type 1 5) Shikimate dh N and 6) Shikimate DH. Arrows on the top of the figure note the number of domains in multidomain, which ranges between 2 to 8 domains. Subphylum and class starting from upper left are Pezizomycotina: Sordariomycetes (APS), Eurotiomycetes (APE), Dothideomycetes (APD), Leotiomycetes (APL), Pezizomycetes (APP), Orbiliomycetes (APO) and Xylonomycetes (APX); Saccharomycotina: Saccharomycetes (ASS); Taphrinomycotina: incertae sedis (AT?), Taphrinomycetes (ATT), Schizosaccharomycetes (ATS) and Pneumocystidomycetes (ATP); Agaricomycotina: Agaricomycetes (BAA), Dacrymycetes (BAD) and Tremellomycetes (BAT); Ustilaginomycotina: Exobasidiomycetes (BUE), Malasseziomycetes (BUM) and Ustilaginomycetes (BUU); Pucciniomycotina: Pucciniomycetes (BPP), Microbotryomycetes (BPMb) and Mixiomycetes (BPMx); Wallemiomycotina: Wallemiomycetes (BWW); Mucoromycotina: Mucoromycetes (MMM); Chytridiomycota: Chytridiomycetes (C-C) and Monoblepharidomycetes (C-M); Blastocladiomycota: Blastocladiomycetes (B-B); Entomophthoromycotina: Entomophthoromycetes (ZEE). **B)** Example of a multi-domain EPSPS protein of 1616 amino acids (protein A0A060S9A7_PYCCL from *Pycnoporus cinnabarinus*). In this example, the SKI and Shikimate DH are fractured.

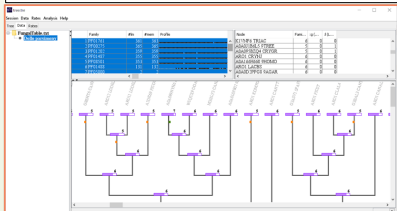


Figure 3. Screenshot of the output of the program COUNT. The image shows a part of phylogenetic tree (centre): domain family profile (top left) and node information (top right). The multi domain structure of the EPSPS in fungi is ancestral (6 domains) and there are recent domain gains and losses.

1	2	3	4	5	6	7	8	N
EPSPS	EPSPS							1
EPSPS	SKI							2
DHG 5	EPSPS							4
EPSPS	DHG							8
EPSPS	DHG							1
EPSPS	Mutlpha							1
EPSPS	mt11							1
EPSPS	FAD							1
EPSPS	binding 5							1
EPSPS	SDN/N	SDH						1
DHG 5	EPSPS	SKI						1
DHG 5	DHG 5	EPSPS	EPSPS					1
DHG 5	EPSPS	SKI	DHG					2
EPSPS	EPSPS	SKI	DHG/N					1
EPSPS	SKI	DHG	DHG/N					5
DHG 5	EPSPS	DHG	DHG/N					1
DHG 5	EPSPS	SKI	DHG	SDN/N				215
EPSPS	SKI	DHG	SDN/N	SDH				3
DHG 5	EPSPS	SKI	DHG	SDN/N	SDH			1
DHG 5	DHG 5	EPSPS	UCH	SKI				1
DHG 5	DHG 5	EPSPS	SKI	DHG	SDN/N			1
DHG 5	DHG 5	EPSPS	SKI	DHG	SDN/N			2
SKN/N	DHG 5	EPSPS	SKI	DHG	SDN/N			2
Gp10	DHG 5	EPSPS	SKI	DHG	SDN/N			1
hyrd 43								
DHG 5	EPSPS	SKI	DHG	SDN/N	SDH			128
DHG 5	EPSPS	SKI	DHG	SDN/N	SDH			1
DHG 5	EPSPS	SKI	DHG	SDN/N	SDH			1
ReuGAP	CRAL	DHG 5	EPSPS	SKI	DHG	SDN/N		1
TRD2								
GDA	GDA	DHG 5	EPSPS	SKI	DHG	SDN/N		1
asac								
DUF1977	DHG 5	EPSPS	SKI	DHG	SDN/N	SDH		1
DHG 5	EPSPS	EPSPS	SKI	DHG	DHG	SDN/N	SDH	1
RNA_JSL	RNA_JSL	RNA_JSL	RNA_JSL	RNA_JSL	RNA_JSL	RNA_JSL	RNA_JSL	2
Rib1_C1	Rib1_C2	Rib1_C3	Rib1_C4	Rib1_C5	Rib1_C6	Rib1_C7	Rib1_C8	

Table 1. Multidomain architectures and their frequencies. Sorted from lowest to highest domain number. The six most important domains have been colour coded according to the colour scheme from Pfam. Novel domains are in white.

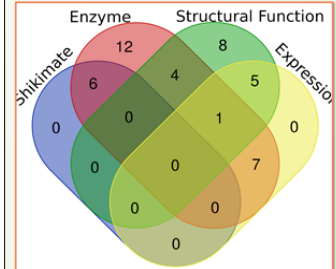


Figure 2. A Venn's diagram of the EPSPS-associated domains. Blue is Shikimate pathway, red is enzymes, green is structural function and yellow is gene expression. The diagram was made with the following tool: <http://bioinformatics.psb.ugent.be/webtools/Venn/>.

Conclusion

The EPSPS multi-domain structure in fungi ranges between two to eight domains. An evolutionary analysis shows that the ancestral state of the EPSPS protein in included six domains (DHQ synthase, EPSPS, SKI, DHquinase I, Shikimate DH N and Shikimate DH). Thus, there have been multiple independent domain gains/losses throughout the evolution of the EPSPS in fungi. Further analyses should determine the importance of the EPSPS-associated domains on the EPSPS sensitivity to glyphosate.

