Brief Report

Adaptation of bacteria to glyphosate: a microevolutionary perspective of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase

Miia J. Rainio,1 Suvi Ruuskanen,1 Marjo Helander,1 Kari Saikkonen,2 Irma Saloniemi1 and Pere Puigbó3,4*
1Department of Biology, University of Turku, Turku, Finland.
2Biodiversity Unit, University of Turku, Turku, Finland.
3Nutrition and Health Unit, Eurecat Technology Centre of Catalonia, Reus, Catalonia, Spain.
4Department of Biochemistry and Biotechnology, Rovira i Virgili University, Tarragona, Catalonia, Spain.

Abstract

Glyphosate is the leading herbicide worldwide, but it also affects prokaryotes because it targets the central enzyme (5-enolpyruvylshikimate-3-phosphate, EPSP) of the shikimate pathway in the synthesis of the three essential aromatic amino acids in bacteria, fungi and plants. Our results reveal that bacteria may easily become resistant to glyphosate through changes in the 5-enolpyruvylshikimate-3-phosphate synthase active site. This indicates the importance of examining how glyphosate affects microbe-mediated ecosystem functions and human microbiomes.

Background

Evolutionarily conserved 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) is the key enzyme of the shikimate pathway in the synthesis of three essential aromatic amino acids (phenylalanine, tyrosine and tryptophan) across taxa, including bacteria, fungi and plants (Bentley, 1990; Richards et al., 2006). The EPSPS enzyme is also involved in the production of the para-aminobenzoate (McConkey, 1999). Thus, glyphosate may have an effect on the synthesis of folate even in environments rich in phenylalanine, tyrosine and tryptophan. The EPSPS enzyme is also known as AroA (EC:2.5.1.19) and catalyses the reaction from shikimate-3-phosphate to chorismate. Because the shikimate pathway does not occur in vertebrates, the use of glyphosate, targeting the EPSPS enzyme by competing for the binding site with the second substrate of the EPSPS (phosphoenol pyruvate, PEP) (Schönbrunn et al., 2001), is regarded safe for use in food production.

Due to its affordable price, effectiveness and broad-spectrum ability to kill weeds, glyphosate has become the most commonly used herbicide worldwide (Woodburn, 2000; Helander et al., 2012; Myers et al., 2016). Although glyphosate antibiotic properties are known (Kurenbach et al., 2015; Kurenbach et al., 2018), its possible effects on microbiomes (Funke et al., 2006; Motta et al., 2018; Gómez-Gallego et al., 2020; Ruuskanen et al., 2020) have largely been neglected until recently (Motta et al., 2018; Mesnage et al., 2019). As microbes have driven eco-evolutionary processes since the origin of life, we note the importance of thoroughly understanding the possible undesirable effects of glyphosate on ecosystem structures, functions and services (Bote et al., 2019). Here, we propose that the wide use of glyphosate may potentially affect microbial communities by (i) eradicating susceptible microbes and (ii) leading to evolution of resistance.

Results and discussion

We analysed 663 microbial genomes distributed across 32 alignable tight genomic clusters (ATGCs) (Novichkov et al., 2009; Puigbó et al., 2014), with the EPSPSclass server (http://puigbo.me/programs/EPSPSclass) (Leino et al., 2020) to identify changes in sensitivity to glyphosate. The algorithm classifies EPSPS proteins based on the presence and absence of amino acid markers in the active site that determine whether the enzyme is putatively sensitive (class I) or resistant (class II, III and IV) to the glyphosate (see methods). Currently, the majority of species of prokaryotes belong to class I, thus are putatively affected by the herbicide (Leino et al., 2020). We analysed 32 ATGCs that amply represented bacteria and one archaeal group. The EPSPS sequence was not found in ATGC033 (Mycoplasma), small-genome single-
membrane bacteria (Meseguer et al., 2003); ATGC046 (Rickettsia), obligate intracellular parasites (Zomorodipour and Andersson, 1999) and ATGC056 (Lactobacillus). Even though Lactobacillus strains in the ATGC dataset do not have the EPSPS sequence, other strains have the enzyme and present differential sensitivity to glyphosate. Previous studies have shown that some Lactobacillus species are susceptible to glyphosate (Shehata et al., 2013).

A microevolutionary analysis of the EPSPS enzyme shows that phylogenetics best explains bacterial sensitivity to glyphosate (Fig. 1, Table 1 and Supplementary Table 1). In general, Firmicutes are significantly more (putatively) resistant to glyphosate than Proteobacteria, whereas Actinobacteria is the most sensitive group to glyphosate. Furthermore, bacterial lifestyle is also associated with sensitivity, i.e. facultative host-associated (FHA) bacteria are more sensitive to glyphosate than free-living (FL) bacteria (Supplementary Tables 2 and 3; Supplementary Fig. 1). However, an analysis with a greater number of samples would be necessary to test whether this association with lifestyle is independent of phylogeny. Intracellular parasites seem to be the most putatively sensitive group; however, due to the limited amount of data, we cannot determine statistical significance for this group’s sensitivity. We hypothesize that sensitivity mirrors the strength of selective force of glyphosate on bacteria. FL bacteria are directly exposed to glyphosate spray at high concentrations, whereas host-associated microbes can be exposed to lower concentrations when the glyphosate is moving through the plant from leaves to roots, and intracellular parasites might be protected by the cell membrane. Genomes with %GC below 40% tend to be more resistant to glyphosate (Fig. 2A). Spearman correlation coefficients between sensitivity and genome parameters from the original ATGC dataset (Puigbò et al., 2014) were not statistically significant (see methods and Supplementary Table 4). To statistically decouple the contribution of %GC, ATGC clusters

Fig. 1. Distribution of the predicted sensitivity to glyphosate across the species tree. Pie charts indicate the proportion of species that are putatively sensitive (green) or resistant (red) to glyphosate, and unclassified (black). See full description in Table 1.
Adaptation of bacteria to glyphosate

(closely related phylogenetic species), number of genes, taxonomic groups (e.g., Actinobacteria, Proteobacteria, Firmicutes) and lifestyle (FL, FHA and intracellular parasites) to the bacterial sensitivity to glyphosate (sensitive, resistant, unknown), we carried out a multiple correspondence analysis (MCA) and random forest (RF) classification model. The MCA analysis revealed that Firmicutes and FL bacteria are mostly resistant, whereas Actinobacteria, Proteobacteria and FHA appeared to be sensitive to glyphosate (Fig. 2B and V). The RF suggests that ATGC clusters (i.e. phylogenetics) are the most important definer of sensitivity to glyphosate (Fig. 2D).

EPSPS is the central enzyme in the metabolic pathway for the synthesis of three essential amino acids in bacteria (Steinrücken and Amrhein, 1980); thus, its protein sequence is presumably evolutionarily conserved (Novick and Doolittle, 2020). Indeed, the EPSPS sequence is highly conserved within 17 ATGCs that preserve sensitivity status across species. However, 12 ATGCs contained species with different sensitivities to glyphosate (Fig. 3, Table 1 and Supplementary Figs 2–13). Although the active site of the EPSPS is highly conserved in prokaryotes, in some species, the sensitivity to glyphosate may be altered with few changes in the active site (Fig. 3).

### Table 1. Summary table of the percentage of species sensitive/resistant to glyphosate.

| ATGC     | Genera       | Taxonomy   | Lifestyle | N     | CDS   | GC    | P (%) | Cl  | ClII | ClIII | ClIV | U     | %S (%) |
|----------|--------------|------------|-----------|-------|-------|-------|-------|-----|------|-------|------|-------|--------|
| ATGC001  | Enterobacteria| Proteobacteria| FHA      | 109   | 4652  | 0.528 | 8.24  | 109 | 1    | 0     | 0    | 0     | 99     |
| ATGC002  | Enterobacter–Klebsiella| Proteobacteria| FHA      | 11    | 5375  | 0.583 | 8.16  | 10  | 0    | 0     | 0    | 1     | 91     |
| ATGC003  | Streptococcus| Firmicutes  | FL       | 22    | 2114  | 0.408 | 9.37  | 22  | 0    | 0     | 0    | 0     | 0      |
| ATGC004  | Streptococcus| Firmicutes  | FL       | 22    | 1865  | 0.385 | 8.97  | 22  | 0    | 0     | 0    | 0     | 0      |
| ATGC005  | Streptococcus| Firmicutes  | FL       | 16    | 1977  | 0.422 | 9.21  | 15  | 0    | 0     | 0    | 0     | 4      |
| ATGC014  | Bacillus     | Firmicutes  | FL       | 31    | 5704  | 0.361 | 9.42  | 30  | 0    | 0     | 0    | 0     | 2      |
| ATGC015  | Bacillus     | Firmicutes  | FL       | 24    | 4089  | 0.459 | 8.91  | 24  | 0    | 0     | 0    | 0     | 0      |
| ATGC021  | Chlamydia    | Chlamydia   | P        | 44    | 898   | 0.417 | 8.87  | 44  | 0    | 0     | 0    | 0     | 100    |
| ATGC022  | Chlamydia–Chlamydia| Chlamydia| P        | 19    | 1126  | 0.397 | 8.93  | 3   | 0    | 0     | 0    | 0     | 17.15  |

ATGCs containing species with potential differential sensitivity to glyphosate are highlighted in bold. FHA: facultative host-associated; FL: free-living; P: intracellular parasite. CDS: number of protein coding genes; N: number of genomes; GC: median GC; P: percentage of aromatic amino acids. Cl: class I; ClII: class II; ClIII: class III; ClIV: class IV; %S: percentages of sensitive species.
potential purifying selection effect against sensitive bacteria but also because (i) some bacterial groups may adapt rapidly to become resistant to glyphosate and (ii) glyphosate-based herbicides may enhance multidrug resistance in bacteria (Liu et al., 2020). Previous studies have shown that glyphosate can affect the microbiome composition of insects, including bees (Motta et al., 2018), potato beetles (Gómez-Gallego et al., 2020) and birds (Ruuskanen et al., 2020). Moreover, preliminary results suggest a putative effect in the human (Leino et al., 2020) and rat (Mesnage et al., 2019) gut microbiomes. In addition, glyphosate may affect ecosystem functions including nutrient cycle and carbon balance, plant growth–promoting microbes (e.g. mycorrhizae and nitrogen-fixing bacteria), microbially mediated plant-herbivore and pathogen interactions, and pollination biology depending on the related microbial community composition and its sensitivity to glyphosate (Helander et al., 2018; Gómez-Gallego et al., 2020).

An analysis of variability in the amino acid landscape (see methods) shows that the amino acids in the active site of the EPSPS that bind PEP are 23% more conserved than the rest of the amino acid residues (Supplementary Fig. 14), in agreement with results from previous studies (Schönbrunn et al., 2001). Moreover, amino acids flanking binding sites (±5 residues) also show higher levels of conservation than average, possibly due to selection preserving the function of the EPSPS. Thus, additional residues within the active site may determine sensitivity to the herbicide. Bacteria potentially more resistant to glyphosate tend to have fewer amino acid substitutions than sensitive ones, e.g. the binding amino acids of actinobacteria and FL bacteria (more resistant to glyphosate) tend to be more conserved (Fig. 5 and Supplementary Tables 5 and 6). Although there are homogeneous and heterogeneous clusters in relation to glyphosate sensitivity, the active site is equally conserved (Supplementary Table 7). Further analyses of amino acid variabilities in EPSPS (Supplementary Table 8) should help elucidate the effects of glyphosate on microbial communities. Moreover, predictions based on amino acid markers may partially explain resistance to glyphosate. However, glyphosate resistance may also occur due to alternative mechanisms, such as efflux pumps or elevated expression of the EPSPS (Staub et al., 2012).
Conclusions
In conclusion, this study shows that phylogenetic groups and bacterial lifestyles are key factors determining sensitivity to glyphosate. Our results suggest that the sensitivity of bacterial species to glyphosate varies and can change in the short evolutionary time of ATGC. Moreover, microbes more exposed to the herbicide (FL bacteria) are putatively more resistant. Nevertheless, because single mutations in the EPSPS active site and rapid HGT events may change the status of sensitivity, heavy use of glyphosate may impact microbial biodiversity independently of taxonomy and lifestyle. The central and unanswered questions are (i) what are the effects of glyphosate-modulated microbiomes on ecosystem functions and services and (ii) on human wellbeing.

Methods
Protein sequences
Protein data were obtained from the ATGCs database, which contains data on >4.5 million proteins encoded in >1500 genomes of prokaryotes (approximately 60% of proteins and 62% of genomes from RefSeq as of June 2013) that met the same criteria as the original ATGCs (Novichkov et al., 2009), as described in Puigbò et al. (2014). The ATGC database provides an ideal environment for a microevolutionary analysis, as it provides a widespread representation of genes from a representative set of prokaryotes. However, other databases (e.g. the Clusters of Orthologous Groups (Galperin et al., 2019)) would be more suitable for a comprehensive macroevolutionary analysis of the EPSPS. EPSPS protein sequences were annotated through BLAST mapping onto the Clusters of Orthologous Genes (COG) database (Galperin et al., 2015). EPSPS belongs to COG0128 (category E, amino acid transport and metabolism). We analysed 32 ATGCs that amply represented bacteria and one archaeal group. Each ATGC contains at least 10 species by the definition in the original study (Puigbò et al., 2014), and up to 109 species are present in the enterobacterial cluster (ATGC001). The final ATGC dataset used in this study includes 332 FHA species (13 ATGCs), 312 FL species (17 ATGCs) and...
19 intracellular parasites (two ATGCs). The EPSPS sequence was not found in ATGC033 (*Mycoplasma*), ATGC046 (*Rickettsia*) and ATGC056 (*Lactobacillus*). Even though *Lactobacillus* strains in the ATGC dataset do not have the EPSPS sequence, the enzyme is present in other strains (Supplementary Table 9).

**Classification of the EPSPS**

Glyphosate inhibits the shikimate pathway by inhibiting the enzyme EPSPS (Schönbrunn *et al*., 2001; Abraham, 2020). The EPSPS class has been classified with the EPSPSClass server (http://ppuigbo.me/programs/EPSPSClass/).
EPSPS Class), which assesses the type of potential sensitivity of the EPSPS enzyme to the herbicide glyphosate (Leino et al., 2020). The amino acid markers to determine potential glyphosate sensitivity to glyphosate (Supplementary Table 10) are also available at Leino et al. (2020) and in the web server. EPSPS enzymes can be classified as class I (alpha or beta) (Light et al., 2016; Firdous et al., 2018), class II (Barry et al., 1997; Priestman et al., 2005), III (Carozzi et al., 2006) or IV (Lira et al., 2013) based on the presence of known amino acid markers (class I, class II and class IV) and motifs (class III) (Barry et al., 1997; Priestman et al., 2005; Carozzi et al., 2006; Lira et al., 2013; Light et al., 2016; Firdous et al., 2018). In addition, a large portion of microbial species are yet unclassified, and their potential sensitivities to glyphosate are unknown (Leino et al., 2020). Different uses of amino acid markers distinguish the EPSPS enzymes into four classes using the reference EPSPS sequences from the ATGCs and the reference gene families, median gene cluster content, median guanine+cytosine (median GC), median dN/dS, synteny distance (dY) and percentage of aromatic amino acids were not statistically significant (p-value >0.05) (Supplementary Table 4). We performed a chi-squared test to determine associations between sensitivity to glyphosate and lifestyle (FL vs FHA) and taxonomic group (Actinobacteria, Firmicutes and Proteobacteria). A Kruskal–Wallis test was used to analyse conservation of the amino acids in the EPSPS. The importance of each variable (ATGC, taxonomic groups, %GC, genome size and lifestyle) to determine potential sensitivity to glyphosate was performed with a MCA and a RF classification model. Input variables are available in the Supplementary Table 11.

Acknowledgements
Pere Puigbó’s research is supported by funds from the Turku Collegium for Science and Medicine (Turku, Finland). Miia Rainio was funded by the Academy of Finland (grant no. 311077 to Marjo Helander).

References
Abraham W. (2020) US7771736B2 - Glyphosate formulations and their use for the inhibition of 5-enolpyruvylshikimate-3-phosphate synthase - Google Patents [Internet], [cited 2020 Apr 24]. URL https://patents.google.com/patent/US7771736B2/en
Barry GF, Kishore GM, Padgette SR, Stallings WC. (1997) US5627061A - Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases, USA.
Bentley, R. (1990). The shikimate pathway - a metabolic tree with many branches. Crit Rev Biochem Mol Biol 25: 307–384.
Bote, K., Pöppe, J., Merle, R., Makarova, O., and Roesler, U. (2019) Minimum inhibitory concentration of glyphosate and of a glyphosate-containing herbicide formulation for Escherichia coli isolates - differences between Pathogenicand non-pathogenic isolates and between host species. Front Microbiol 10: 932.
Carozzi N, Carr B, Hammer P. (2006) US20060253921A1 - Identification of a new class of EPSP synthases, USA.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797.
Firdous, S., Iqbal, S., Anwar, S., and Jabeen, H. (2018) Identification and analysis of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene from glyphosate-resistant Ochrobactrum intermedium Sq20. Pest Manag Sci 74: 1184–1196.
Funke, T., Han, H., Healy-Fried, M.L., Fischer, M., and Schönbärrn, E. (2006) Molecular basis for the herbicide resistance of roundup ready crops. Proc Natl Acad Sci U S A 103: 13010–13015.
Galperin, M.Y., Kristensen, D.M., Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2019) Microbial genome analysis: the COG approach. Brief Bioinformatics 20: 1063–1070.
Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43: D261–D269.
Gómez-Gallego, C., Rainio, M.J., Collado, M.C., Mantziari, A., Salminen, S., Saikkonen, K., et al. (2020) Glyphosate-based herbicide affects the composition of microbes associated with Colorado potato beetle (Leptinotarsa decemlineata). FEMS Microbiol Lett 367: Inaa050.
Healy-Fried, M.L., Funke, T., Priestman, M.A., Han, H., and Schönbärrn, E. (2007) Structural basis of glyphosate...
tolerance resulting from mutations of Pro101 in Escherichia coli 5-enolpyruvylshikimate-3-phosphate synthase. J Biol Chem 282: 32949–32955.

Helander, M., Saloniemi, I., Oncini, M., Drulle, M., Salminen, J.-P., and Saikkonen, K. (2013) Glycophosphate decreases mycorrhizal colonization and affects plant-soil feedback. Sci Total Environ 642: 285–291.

Helander, M., Saloniemi, I., and Saikkonen, K. (2012) Glycophosate in northern ecosystems. Trends Plant Sci 17: 569–574.

Kurenbach, B., Hill, A.M., Godsoe, W., van Hamelsveld, S., and Heinemann, J.A. (2018) Agrichemicals and antibiotics in combination increase antibiotic resistance evolution. PeerJ 6: e5801.

Kurenbach, B., Marjoshi, D., Amábile-Cuevas, C.F., Ferguson, G.C., Godsoe, W., Gibson, P., et al. (2015) Sublethal exposure to commercial formulations of the herbicides dicamba, 2,4-dichlorophenoxyacetic acid, and glyphosate cause changes in antibiotic susceptibility in Escherichia coli and Salmonella enterica serovar typhimurium. MBio 6: e00009–15.

Leino, L., Tall, T., Helander, M., Saloniemi, I., Saikkonen, K., Ruuskanen, S., et al. (2020) Classification of the gylphosate target enzyme (5-enolpyruvylshikimate-3-phosphate synthase) for assessing sensitivity of organisms to the herbicide. J Hazard Mater 124556.

Light, S.H., Krishna, S.N., Minasov, G., and Anderson, W.F. (2016) An unusual cation-binding site and distinct domain interactions distinguish class II Enolpyruvylshikimate-3-phosphate synthases. Biochemistry 55: 1239–1245.

Lira JM, Cicchillo RM, Nair SK. (2013) US20130217577A1 - 5-Enolpyruvylshikimate 3-phosphate synthase from Staphylococcus aureus is insensitive to glyphosate. FEBS Lett 579: 728–732.

Puigbó, P., Lobkovsky, A.E., Kristensen, D.M., Wolf, Y.I., and Koonin, E.V. (2014) Genomes in turmoil: quantification of genome dynamics in prokaryote supergenomes. BMC Biol 12: 66.

Richards, T.A., Dacks, J.B., Campbell, S.A., Blanchard, J.L., Foster, P.G., McLeod, R., and Roberts, C.W. (2006) Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. Eukaryot Cell 5: 1517–1531.

Ruuskanen, S., Rainio, M.J., Gómez-Gallego, C., Selenius, O., Salminen, S., Collado, M.C., et al. (2020) Glyphosate-based herbicides influence antioxidants, reproductive hormones and gut microbiome but not reproduction: a long-term experiment in an avian model. Environ Pollut 266: 115108.

Schönbrunn, E., Eschenburg, S., Shuttleworth, W.A., Schloss, J.V., Armrhein, N., Evans, J.N., et al. (2001) Inter-action of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. Proc Natl Acad Sci U S A 98: 1376–1380.

Shehata, A.A., Schrödl, W., Aldin, A.A., Hafez, H.M., and Krüger, M. (2013) The effect of glyphosate on potential pathogens and beneficial members of poultry microbiota in vitro. Curr Microbiol 66: 350–358.

Staub, J.M., Brand, L., Tran, M., Kong, Y., and Rogers, S.G. (2012) Bacterial glyphosate resistance conferred by over-expression of an E. coli membrane efflux transporter. J Ind Microbiol Biotechnol 39: 641–647.

Steinrücken, H.C., and Armrhein, N. (1980) The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimate acid-3-phosphate synthase. Biochem Biophys Res Commun 94: 1207–1212.

Woodburn, A.T. (2000) Glyphosate: production, pricing and use worldwide. Pest Manag Sci 56: 309–312.

Zomorodipour, A., and Andersson, S.G. (1999) Obligate intracellular parasites: Rickettsia prowazekii and Chlamydia trachomatis. FEBS Lett 452: 11–15.

**Supporting Information**

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

**Appendix S1**: Supplementary Information