Computational and Structural Biotechnology Journal 19 (2021) 2160–2169

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Widespread distribution of *hmf* genes in Proteobacteria reveals key enzymes for 5-hydroxymethylfurfural conversion

Raúl A. Donoso [a,b], Fabián González-Toro [a], Danilo Pérez-Pantoja [a,*]

[a] Programa Institucional de Fomento a la Investigación, Desarrollo e Innovación (PIDi), Universidad Tecnológica Metropolitana, Santiago, Chile
[b] Center of Applied Ecology and Sustainability (CAPES), Santiago, Chile

ARTICLE INFO

Article history:
Received 7 November 2020
Received in revised form 5 April 2021
Accepted 5 April 2021
Available online 16 April 2021

Keywords:
HMF
Furans
Biodegradation
Proteobacteria
Metabolic engineering

ABSTRACT

Furans represent a class of promising chemicals, since they constitute valuable intermediates in conversion of biomass into sustainable products intended to replace petroleum-derivatives. Conversely, generation of furfural and 5-hydroxymethylfurfural (HMF) as by-products in lignocellulosic hydrolysates is undesirable due its inhibitory effect over fermentative microorganisms. Therefore, the search for furans-metabolizing bacteria has gained increasing attention since they are valuable tools to solve these challenging issues. A few bacterial species have been described at genetic level, leading to a proposed HMF pathway encoded by a set of genes termed *hmf*/*psf*, although some enzymatic functions are still elusive. In this work we performed a genomic analysis of major subunits of furoyl-CoA dehydrogenase orthologues, revealing that the furoic acid catabolic route, key intermediate in HMF biodegradation, is widespread in proteobacterial species. Additionally, presence/absence profiles of *hmf*/*psf* genes in selected proteobacterial strains suggest parallel and/or complementary roles of enzymes with previously unclear function that could be key in HMF conversion. The furans utilization pattern of selected strains harboring different *hmf*/*psf* gene sets provided additional support for bioinformatic predictions of the relevance of some enzymes. On the other hand, at least three different types of transporter systems are clustered with *hmf*/*psf* genes, whose presence is mutually exclusive, suggesting a core and parallel role in furans transport in Proteobacteria. This study expands the number of bacteria that could be recruited in biotechnological processes for furans biodetoxification and predicts a core set of genes required to establish a functional HMF pathway in heterologous hosts for metabolic engineering endeavors.

1. Introduction

The sustainable production of bio-based chemicals is greatly reliant on the development of viable biorefineries to use different biomass resources for their conversion into diversified products [1]. These bioproducts have the potential to replace those produced by petroleum refineries [1]. Furans represent one of the most promising classes of intermediates in the conversion of biomass into bio-based products since they constitute sustainable alternatives to various petroleum-derived chemicals, having properties analogous to and even exceeding those derived from fossil hydrocarbons [2,3]. Particularly, 5-hydroxymethylfurfural (HMF) has been termed as the "Sleeping Giant" of sustainable chemistry because of its vast synthetic potential, and it has been broadly accepted by the scientific community as a key element to bridge the gap from a fossil-based economy to a sustainable manufacturing system [2,4]; being the focus of nearly 1000 papers published per year [2] and reaching a global market of 56 million US$ in 2019 [5]. HMF can be produced from the dehydration of hexoses (mainly fructose) or via acid-catalyzed hydrolysis/dehydration of cellulose [6]. The latter synthetic route is the second aspect of HMF chemistry that is relevant for biotechnology industry, this is due its production as by-product during the acid-based pretreatment of lignocellulose to allow the subsequent digestion of cellulose for biofuels fermentation [6]. The generation of furans in lignocellulosic hydrolysates is unwanted due its highly inhibitory effect over fermentative microorganisms, and consequently HMF is included along with furfural in the so-called "lignocellulose-derived inhibitors" (LDIs) group of noxious compounds whose adverse effects include decreased viability, ethanol yield, and productivity [7,8]. In particular, at levels found in hydrolysates, HMF can produce a reduction of approximately 33% in growth rate of an ethanologenic bacterium such as *Zymomonas mobilis* [9], and

* Corresponding author.
E-mail address: danilo.perez@utem.cl (D. Pérez-Pantoja).

https://doi.org/10.1016/j.csbj.2021.04.017

in mixtures with other LDIs is able to severely affect ethanol yield and productivity by 90 and 85% respectively, when tested in the fermentative yeast *Scheffersomyces stipitis* [10].

Several approaches have been explored to overcome the negative effect of furan inhibitors, including decreasing their generation by adjusting pretreatment conditions, developing more tolerant ethanologenic strains, and detoxifying the hydrolysates [7,8,11]. Regarding this latter approach, biological detoxification, using a microorganism to remove inhibitors from the hydrolysate, is an attractive alternative due minimal generation of waste streams, no need for chemical inputs, and the feasibility to be performed directly in the fermentation vessel avoiding an extra process step [7,8,11]. The cornerstone of this approach is to use a very proficient microorganism having abilities to catabolize multiple furans without metabolization of sugars intended for fermentation [12–15]. Altogether, the requirement of furans-metabolizing strains both for biological detoxification and for biocatalysis aimed to obtain fine products have promoted the study of furans biodegradation over the last years [16–18].

A limited number of bacterial species owning the ability to use furfural or HMF as sole carbon sources have been characterized at genetic level, with *Pseudomonas putida* Fu1 [19], *Cupriavidus basilensis* HMF14 [20], and *Pseudomonas putida* ALS1267 [21] being the best characterized examples. The pioneering work of Nichols and Mertens provided the first pieces of evidence about furans biodegradation in *P. putida* strain Fu1 by identification of catabolic and regulatory *psf* genes [19]. Later, Koopman *et al.* [20] performed the genetic characterization of HMF and furfural degradation in *C. basilensis* HMF14, both converging to 2-furoic acid (FA) to be converted to 2-oxoglutarate using coenzyme A intermediates and enzymes encoded by *hmfABCDE* genes (Fig. 1). In the case of HMF catabolism, the presence of *hmfFGH* genes is also required since they are involved in the upper part of the degradative route, supporting the conversion of HMF into FA (Fig. 1) [18,20]. Specifically, *hmfFG* genes encode a decarboxylase, able to catalyze conversion of key intermediate 2,5-furandicarboxylic acid (FDCA) to FA, whilst *hmfH* encodes a FAD-dependent oxidoreductase putatively involved in two consecutive oxidations of 5-hydroxymethyl-2-furancarboxylic acid (HMFCA) to generate FDCA [18,20]. No specific aldehyde dehydrogenases have been identified for initial HMF oxidation into HMFCA, suggesting that such activity is related to broad-specificity dehydrogenases protecting cells against toxic aldehydes [18,20,22]. Recently, the genes involved in HMF catabolism have been characterized in *P. putida* ALS1267, which carries all *hmf* genes previously described in strain HMF14 except for *hmfH*, suggesting that another gene would replace this FAD-dependent oxidoreductase [21]. Interestingly, two additional genes are present in *hmf* operon from *P. putida* ALS1267, which were also previously found in *P. putida* Fu1 and termed *psfG* and *psfA,* encoding a putative short-chain alcohol dehydrogenase and a putative NAD-dependent aldehyde dehydrogenase, respectively [19]. These enzymatic functions have been suggested to be related to the oxidation of furfuryl alcohol/furfural and HMF alcohol/HMF [19,21], but additional support has not been provided.

The interest to establish a furans detoxification pathway in fermentative microorganisms has driven metabolic engineering endeavors for functional expression of HMF route in genetically tractable strains [23–26]. Two studies focused on transferring the entire HMF pathway from *P. putida* ALS1267 [21] and from a previously unrecognized furans-degrading *Paraburkholderia phytofirmans* PsJN strain [24], to industrially relevant *P. putida* strains. These efforts have not been completely successful since the furans-degrading phenotype was only partially achieved as revealed by lack of growth on HMF or a very slow proliferation rate by the engineered strains [21,24]; most probably because the selection of catabolic genes was incomplete, or their expression was inadequate.

Considering that metabolic engineering has as a central goal to identify gene targets for the optimization of a metabolic phenotype [27], we focus on genome analysis of furans biodegradation to pro-

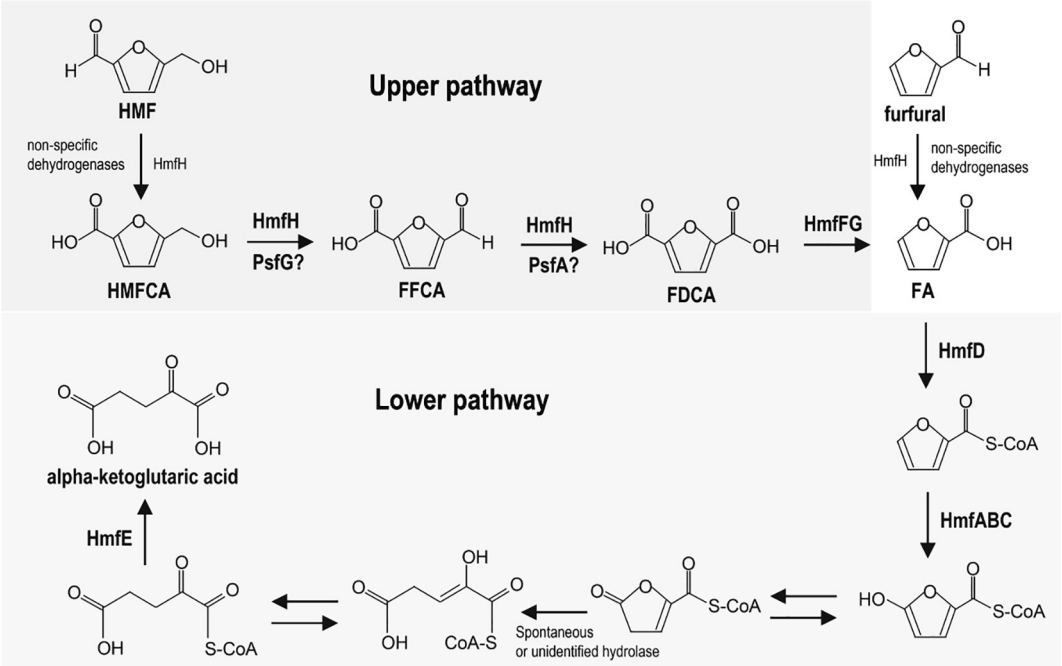

Fig. 1. Predicted metabolic pathways of HMF and furfural in *Cupriavidus* and *Pseudomonas* species converged toward FA as central intermediate. The diagram of reconstructed pathways has been adapted and integrated from previous reports [19–21]. FA conversion into α-ketoglutaric acid is accomplished by *hmfABCDE*-encoded enzymes constituting the lower pathway. HMFCA and FFCA dehydrogenase activities would be encoded by *psfG* and *psfA* genes, respectively. HmfH, HMF/furfural oxidoreductase; HmfFG, FDCA decarboxylase; HmfD, furoyl-CoA synthetase; HmfABC, furoyl-CoA dehydrogenase; HmfE, 2-oxoglutaryl-CoA hydrolase. HMF, 5-hydroxymethylfurfural; HMFCA, 5-hydroxymethyl-2-furancarboxylic acid; FFCA, 5-formylfuran carboxylic acid; FDCA, 2,5-furandicarboxylic acid; FA, furoic acid.

vide additional insights about genetic requirements for HMF conversion. In this work we performed an extensive survey on the presence of *hmf/psf* genes in proteobacterial genomes revealing a previously unrecognized distribution of the HMF pathway, broadly expanding the diversity of species suitable for furans conversion. In addition, the presence-absence pattern of some genes suggests redundant or complementary roles that should be considered in metabolic engineering endeavors towards biological detoxification of furans and/or biocatalytic production of high-value intermediates.

2. Material and methods

2.1. Database survey

Protein similarity searches were performed using the blastP algorithm from NCBI and IMG/M databases using default parameters [28,29]. The HmfA sequences were retrieved from non-redundant protein database of GenBank [30]. Only proteins displaying at least 45% amino acid identity with the *hmfA*-encoded protein of *C. basilensis* HMF14 [20] were considered for analysis. Proteobacterial genomes that harbor *hmfA* homologues were downloaded from NCBI database for identification of additional catabolic and complementary functions. Gene products of lower (*hmfBCDE*) and upper (*hmfFGH* and *psfAG*) pathways, transcriptional regulators (AraC- and LysR-type), and transporters (*hmfT1*) with verified or putative functions located in the *hmf* gene clusters of *C. basilensis* HMF14, *P. putida* ALS1267, and *P. putida* Fu1 [19–21] were used as bait for homologue searches in genomes previously selected using BLAST software [28] on a local computer. Identification of conserved domains within a protein was performed using the Conserved Domains Database from the NCBI website [31].

2.2. Phylogenetic reconstructions

Evolutionary relationships from the concatenated alignment of the single-copy *hmfABCDE* genes were inferred by IQ-TREE web server tools (http://iqtree.cibiv.univie.ac.at/) designed for estimate maximum-likelihood phylogenies [32] employing ModelFinder as model-selection method [33] and Ultrafast Bootstrap Approximation as bootstrap approach [34] with the -m TEST, -bb 1000 and -alrt 1000 options. Sequence alignments for phylogenetic reconstruction were calculated with MAFFT online service (https://mafft.cbrc.jp/alignment/server/) using Auto (FFT-NS-1, FFT-NS-2, FFT-NS-i or L-INS-i; depends on data size) strategy [35]. Visualization and edition of phylogenetic trees were performed by the Interactive Tree Of Life (iTOL) online tool (https://itol.embl.de/) [36].

2.3. Bacterial strains and growth conditions

The bacterial strains assayed in this study were *Azospirillum oryzae* COC8 [37], *Cupriavidus basilensis* DSM 11853 [38], *Cupriavidus pinatubonensis* JMP134 [39], *Paraburkholderia phytofirmans* PsJN [40], *Paraburkholderia xenovorans* LB400 [41], *Pseudomonas mendocina* VN230 [42], *Pseudomonas sp*. ALS1131 [43], and *Pseudomonas sp*. ALS1279 [43]. All strains were grown in mineral salts medium (14 g/L $Na_2HPO_4{\cdot}12H_2O$, 2 g/L KH_2PO_4, 50 mg/L $Ca(NO_3)_2{\cdot}4H_2O$, 1 g/L $(NH_4)_2SO_4$, 200 mg/L $MgSO_4{\cdot}7H_2O$, 278 μg/L $FeSO_4{\cdot}7H_2O$, 70 μg/L $ZnCl_2$, 100 μg/L $MnCl_2{\cdot}4H_2O$, 62 μg/L H_3BO_3, 190 μg/L $CoCl_2{\cdot}6H_2O$, 17 μg/L $CuCl_2{\cdot}2H_2O$, 24 μg/L $NiCl_2{\cdot}6H_2O$, 36 μg/L $Na_2MoO_4{\cdot}2H_2O$), supplemented with 2.5 mM HMF, HMFCA, 5-formylfuran carboxylic acid (FFCA), FDCA or FA as sole carbon and energy source. The cultures were incubated in a 96-well microplate (Thermo Fisher Scientific, Rochester, NY, USA) at 30 °C, and the optical density at 600 nm (OD_{600}) was measured with a Synergy HTX multimode plate reader (BioTek, Winooski, VT, USA). Cell cultures were inoculated with 100-fold dilutions of overnight cultures grown on R2A broth. At least two biological replicates were performed for each growth measurement.

2.4. Chemicals

HMF, FFCA and FA were purchased from Sigma-Aldrich (Steinheim, Germany). HMFCA was purchased from Cayman Chemical Company (Michigan, USA). FDCA was purchased from Alfa Aesar (Massachusetts, USA).

3. Results and discussion

3.1. Furans biodegradation genes are widespread in α-, β- and γ-lineages of the phylum Proteobacteria

In order to evaluate the presence of furans biodegradative genes in proteobacterial genomes we chose as gene marker the HmfA product corresponding to the major subunit of furoyl-CoA dehydrogenase, a key enzyme in the lower route responsible for 2-furoyl-CoA metabolization (Fig. 1). Using HmfA from *C. basilensis* HMF14 [20] as query (1015 aa), we conducted a restricted search for Proteobacteria in the non-redundant protein sequences database of GenBank as of September 2020, identifying nearly ~900 HmfA-like sequences displaying at least 45% amino acid identity in strains that belong to α-, β- and γ-Proteobacteria lineages. We selected 98 bacterial species with genome sequences available to include at least one member of each genus harboring HmfA homologues in order to examine the presence of additional *hmf* genes (Fig. 1). We also include the genomes of three additional *Pseudomonas* species (ALS1131, ALS1267 and ALS1279) with verified furans-degrading abilities [43,44] collecting a total of 101 strains. Nevertheless, two bacterial species harboring HmfA, representing *Inquilinus* (GCA_002195995.1) and *Acidisphaera* (GCA_000964365.1) genus of the α-Proteobacteria lineage (order *Rhodospirillales*) were subsequently discarded due to low quality genome assembly. From the 99 Proteobacterial genomes analyzed (Table S1), we found that 50 belong to α-Proteobacteria class subdivided in order *Rhizobiales* (30), *Rhodobacterales* (14), *Rhodospirillales* (5) and *Sphingomonadales* (1), representing 14 different families (Fig. S1); 35 belong to β-Proteobacteria class subdivided in order *Burkholderiales* (33) and *Rhodocyclales* (2), representing 6 different families (Fig. S1); and 14 belong to γ-Proteobacteria class subdivided in order *Alteromonadales* (1), *Chromatiales* (1), *Oceanospirillales* (5), *Pseudomonadales* (4) (including three *bona-fide* degrading strains added), *Salinisphaerales* (1) and *Xanthomonadales* (2), representing 8 different families (Fig. S1). These strains were isolated from diverse locations such as oceans, plants, freshwater, groundwater, sediments, soils, activated sludge, wastewater, saline springs and human host, among other environments (Table S1). Additionally, we detected the *hmfA* gene in metagenome-assembled genomes belonging to *Candidimonas* (TAM85542.1), *Comamonas* (MPT12907.1), *Confluentimicrobium* (MAQ45687.1), *Filomicrobium* (MAI44436.1), *Kaistia* (ODT20617.1), *Lautropia* (ODS98322.1), *Methylibium* (KNZ31457.1) and *Verminephrobacter* (MBA4060343.1) genera, obtained from a wide diversity of environmental sources. These genomes not represented by isolated bacterial species expand the range of proteobacterial genera putatively harboring the furans-degrading pathway. Then, we analyzed the presence of remaining lower pathway genes (*hmfBCDE*) in selected strains, observing that all bacterial species have the full set of genes required for metabolization of FA, the central intermediate in fur-

ans catabolism (Fig. 1; Table S1). These results would indicate a ubiquitous presence of furans in the environment and their relevance as natural substrates.

In order to establish phylogenetic relationship among *hmf* genes of selected strains we constructed a dendrogram of concatenated amino acid sequences of lower pathway genes. The resulting HmfABCDE tree showed a strict partition between Proteobacteria lineages with few exceptions (Fig. 2). Interestingly, γ-Proteobacteria *Pseudoxanthomonas composti* GSS15 isolated from compost [45] and *Luteimonas huabeiensis* HB-2 isolated from stratum water [46], both belonging to the family *Xanthomonadaceae*, are inserted in the clade of sequences belonging to α-Proteobacteria (Fig. 2), suggesting lateral gene transfer of lower pathway genes. Furthermore, vast majority of strains (85) harbor the canonical *hmfABCDE* gene order (Fig. 3; Table S1), indicating that this organization is highly conserved to that reported by Wierckx *et al.* [18]. The second most common organization was *hmfABCDFGE*, which is restricted to β-Proteobacteria (11) and includes the insertion of *hmfFG* genes of upper pathway (Fig. 3; Table S1). Only *Rhodoplanes sp.* Z2-YC6860 and *Noviherbaspirillum denitrificans* TSA40 have lower pathway genes in different gene contexts (Fig. 3), raising questions about its functionality in these strains. Remarkably, β-Proteobacteria strains *Acidovorax sp.* KKS102, *Bordetella flabilis* AU10664 and *Paracandidimonas soli* DSM 100048 harbor two copies of *hmfABCDE* gene cluster (Fig. 3), suggesting functional redundancy of lower pathway in these bacteria. In the case of *P. soli* DSM 100048 both copies are very close related indicating a recent event of gene duplication. From the total number of analyzed strains, 44 species harbor exclusively lower pathway genes and not those related to upper route for HMF, suggesting that FA is *per se* a relevant growth substrate or it is metabolized as central intermediate for channeling of furfural by non-specific dehydrogenases. On the other hand, 55 selected strains carry extra genes related to upper pathway

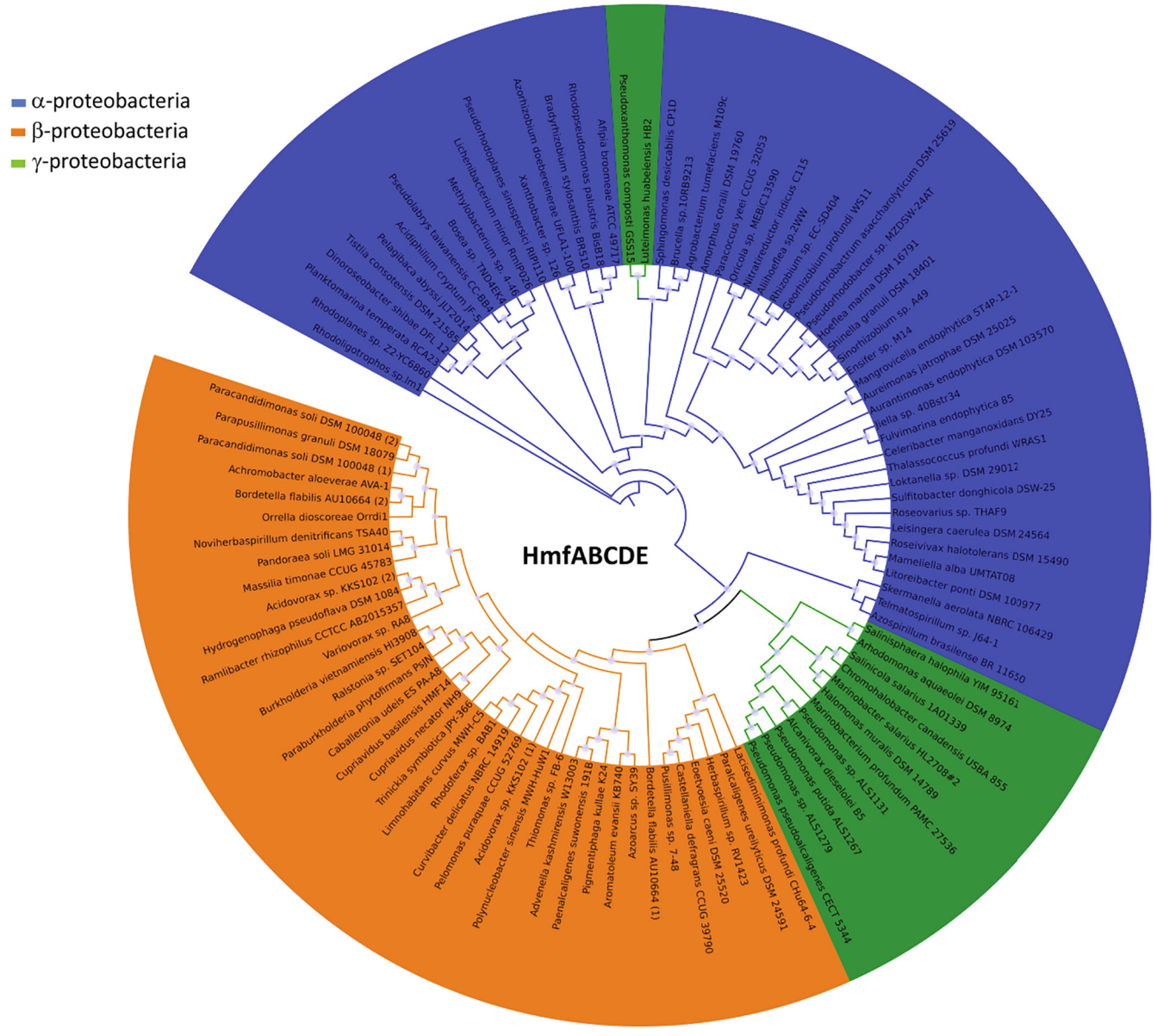

Fig. 2. Evolutionary relationships among concatenated HmfABCDE protein homologues identified in proteobacterial genomes analyzed in this work. Maximum likelihood topology provided by IQ-TREE [32] and based on sequence alignments calculated using MAFFT [35] is shown with SH-like approximate likelihood ratio support values (n = 1000) given at each node (values > 50% are shown). The substitution model selected was LG + F + I + G4 as determined by ModelFinder [33]. Blue, α-proteobacteria; orange, β-proteobacteria; green, γ-proteobacteria. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

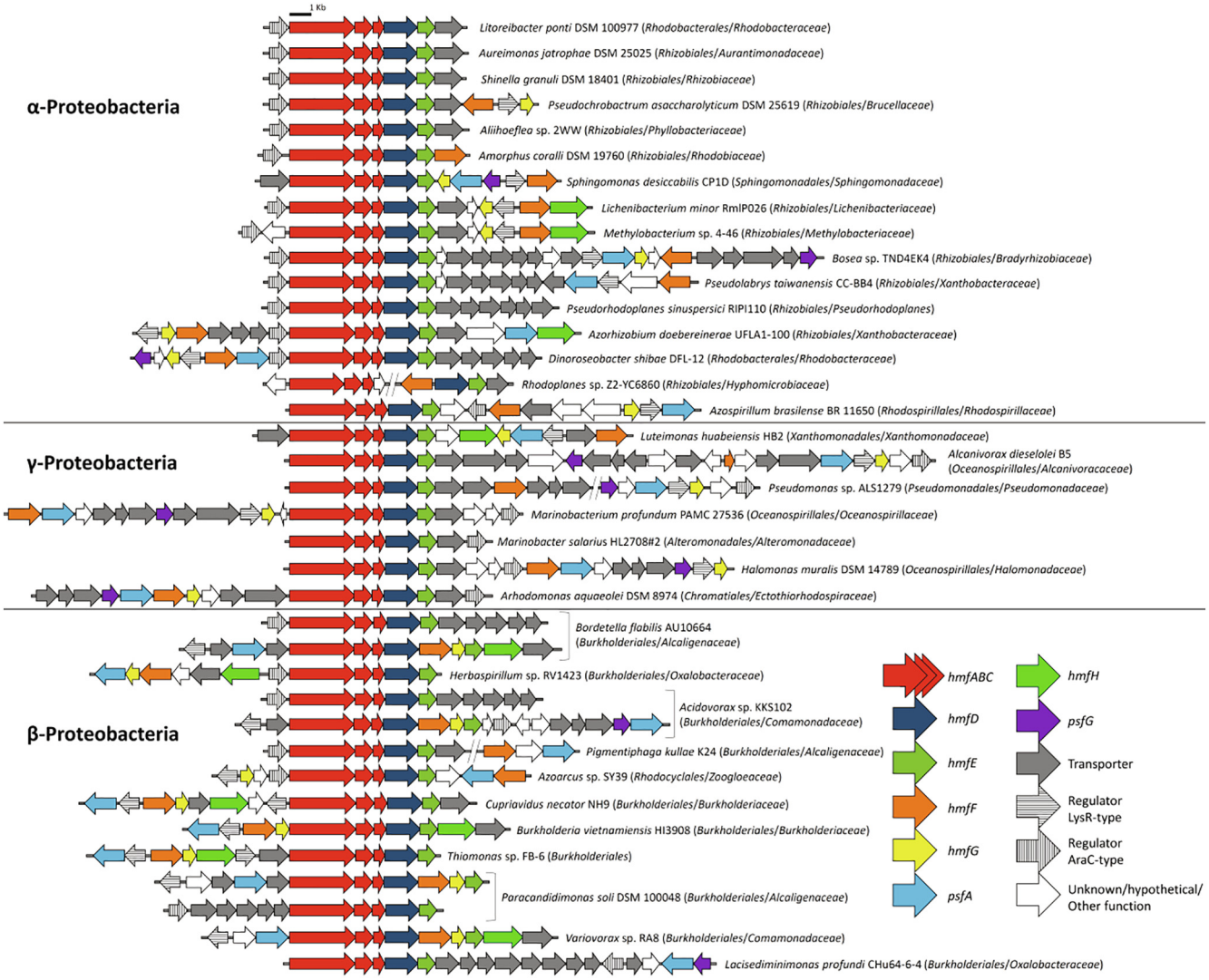

Fig. 3. Clusters of *hmf* genes in representative proteobacterial species. The highly conserved organization of *hmfABC* genes was used to align the *hmf* gene clusters from selected species of α-, β- and γ- lineages of Proteobacteria. Genes putatively encoding enzymes, transporters and regulators of the biodegradative pathway for HMF were colored or shaded according to their function.

(Fig. 4), suggesting broader spectrum of furan compounds degraded by these species.

3.2. Additional catabolic enzymes could be part of the HMF pathway in Proteobacteria

A key step in HMF degradation it is played by *hmfFG* genes, which encoded an enzymatic activity able to catalyze FDCA decarboxylation to FA (Fig. 1). Since HmfG is homologous to UbiX, which synthesizes the prenylated flavin mononucleotide cofactor required by UbiD family of enzymes [47,48], apparently the key catalytic function in FDCA decarboxylation is performed by HmfF, which belongs to UbiD family [49]. We detected 55 bacterial genera out of 99 selected representative strains that harbor *hmfFG* genes, in which 18 belong to α-Proteobacteria (36% of selected strains belonging to this lineage), 28 are classified as β-Proteobacteria (80%), and 9 are included in the γ-Proteobacteria lineage (64%) (Fig. 4), indicating that more than half of the whole set of bacterial genera harboring a FA-metabolizing route could also catalyze conversion of FDCA into FA. The presence of *hmfFG* genes is particularly predominant in β-Proteobacteria, suggesting that channeling of additional furans has a higher relevance for this lineage. Analogous to HmfABCDE tree (Fig. 2), a HmfF dendrogram showed a clear separation between α- and β-Proteobacteria lineages; but on the contrary, HmfF homologues belonging to γ-Proteobacteria showed a more complex distribution, not been clearly distinguished as a specific clade (Fig. S2). This would be suggestive that in γ-Proteobacteria this enzymatic function may have been acquired by lateral gene transfer from α- and β- lineages.

An enzyme playing a crucial role in the upper pathway of furans biodegradation is the HMF oxidoreductase encoded by *hmfH*, previously described in *C. basilensis* HMF14, which has been extensively studied by its ability to convert HMF into FDCA in heterologous strains *P. putida* S12 and *Raoultella ornithinolytica* BF60, turning a valuable tool for production of this high-value precursor [22,26,50–52]. HmfH is a member of the glucose-methanol-choline (GMC) oxidoreductase family and mainly converts HMFCA into FFCA, and secondarily it would be also involved in the subsequent oxidation into FDCA [18,26]. However, a role for unspecified aldehyde dehydrogenases catalyzing oxidation of FFCA into FDCA during heterologous expression of *hmf* genes in *P. putida* S12 has been also proposed [22]. In the selected strains harboring the lower route for furans we found overall 23 species that harbor *hmfH* gene,

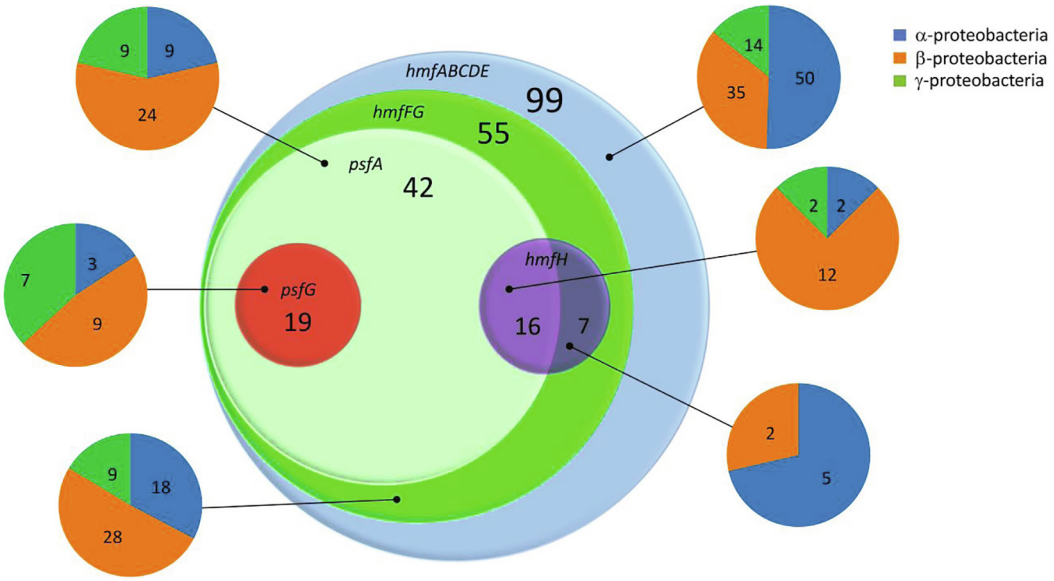

Fig. 4. Presence and distribution of *hmfFG, hmfH, psfG* and *psfA* genes among proteobacterial genomes analyzed in this work. The presence of genes putatively encoding different enzymatic functions of the upper HMF pathway was analyzed in genomes of 99 proteobacterial species containing the *hmfABCDE* cluster responsible for the lower part of the route. The number of species harboring each gene is included in the circle representing it. External pie charts represent the distribution of each gene denoted in Venn diagram among different proteobacterial lineages. Blue, α-proteobacteria; orange, β-proteobacteria; green, γ-proteobacteria. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

mainly α- (7) and β-Proteobacteria (14) (Fig, 4). We only detected this oxidoreductase-encoding gene in two γ-Proteobacteria, *P. composti* GSS15 and *L. huabeiensis* HB-2 (Table S1). Notably, the *hmfABCDE* genes of both strains were previously clustered into the α-Proteobacteria clade in *hmfABCDE* tree (Fig. 2), strongly suggesting that these clusters have been acquired by lateral gene transfers, and consequently pinpointing that *hmfH* gene would not be genuinely associated to *hmf* clusters originated in γ-Proteobacteria. This indicates that in most HMF-degrading γ-Proteobacteria the consecutive oxidations of HMFCA to generate FDCA should be performed by enzymes other than HmfH homologues. Notably, in the search for alternative enzymes to be recruited for biocatalytic FDCA production a remote HmfH homologue has been identified in *Methylovorus sp*. MP688 (46% aa identity) and termed HmfO [53]. Due to its biocatalytic potential, the properties of HmfO have been comprehensively studied [52,54–57]. It should be noted that most probably HMFCA or FFCA are not physiological substrates for this enzyme since no *hmfABCDE* genes could be identified by us in the genome of *Methylovorus sp*. MP688 or related *Methylovorus* species, excluding a role for HmfO in a catabolic route for HMF.

Interestingly, the *psfA* gene encoding a putative aldehyde dehydrogenase previously identified in *P. putida* strains Fu1 and ALS1267 [19,21] was additionally detected in 42 genomes out of 55 strains harboring *hmfFG* genes, including α- (9), β- (24) and γ-Proteobacteria (9) (Fig. 4; Fig. S3), suggesting a relevant role of this enzyme in the upper pathway for furans. A homologue of this putative aldehyde dehydrogenase (63% aa identity with *psfA*) was previously detected in close proximity to *hmf* genes of *C. basilensis* HMF14 and termed Adh [18]. Notably, during the establishment of a whole cell biocatalyst system by heterologous expression of *hmfH* gene in *P. putida* S12 [26], it was demonstrated that FDCA production was improved by achieving a high and constant conversion rate by co-expression of *adh* and a transporter gene termed *hmfT1* (see 3.4 section), avoiding transient accumulation of FFCA detected in presence of *hmfH* and *hmfT1* [22]. These results suggest that a PsfA-like aldehyde dehydrogenase would be involved in conversion of FFCA into FDCA, complementing the activity of HmfH over HMFCA. Accordingly, we detected that 16 out of 23 strains containing the HmfH-encoding gene also own a *psfA* gene in its genome (Fig. 4), supporting importance of aldehyde dehydrogenase activity of PsfA (most probably as FFCA dehydrogenase) in HMF pathway (Fig. 1). It should be noted that in 7 strains harboring a *hmfH* gene (5 α- and 2 β-Proteobacteria) a *psfA* gene could not be found (Fig. 4; Table S1), suggesting that if these strains are able to grow in HMF as carbon source they should recruit a different dehydrogenase activity to replace for PsfA. Most probably, the PsfA aldehyde dehydrogenase has an underestimated role in the HMF degradation pathway (Fig. 1), masked by the pool of endogenous aldehyde dehydrogenases that protects cells against aldehydes toxicity.

Another gene originally identified in *Pseudomonas* strains Fu1 and ALS1267 is *psfG*, encoding a putative short-chain alcohol dehydrogenase putatively involved in furfuryl/HMF alcohol oxidation [19,21]. The *psfG* gene was found in 19 species comprising α- (3), β- (9) and γ-Proteobacteria (7), all of them simultaneously harboring *hmfFG* and *psfA* genes (Fig. 4). A comparison of the amino acid sequence of *psfG* (249 aa) from ALS1267 strain [21] in the UniProtKB/Swiss-Prot database showed a 39% sequence identity with (*S*)-1-Phenylethanol dehydrogenase from *Aromatoleum aromaticum* EbN1, which is a typical secondary alcohol dehydrogenase that catalyzes the NAD-dependent stereospecific oxidation of (*S*)-1-phenylethanol to acetophenone in the degradation of ethylbenzene [58]. Both enzymes are classical members of the short-chain dehydrogenase/reductase (SDR) family with a reaction spectrum comprising the NAD(P)(H)-dependent oxidoreduction of hydroxy/keto groups [59]. It is worth mentioning that strains containing *psfG* on no occasion also carry the *hmfH* gene (Fig. S3), suggesting that both enzymes could have analogous functions not previously considered, catalyzing HMFCA conversion to FFCA (Fig. 1), which is the unidentified activity in strain ALS1267 that lacks a *hmfH*-encoding gene [21]. Nevertheless, entire *hmf* cluster, including *hmfABCDEFG* and *psfAG* genes, was cloned and transformed in strain KT2440, allowing it to grow on furfural as a sole carbon source but not on HMF, suggesting that an additional gene is still lacking to support growth on HMF [21], or not enough

expression of genes encoding the upper part of the route was achieved. A null mutant of *psfG* gene in ALS1267, or in related strains harboring a similar gene set, would help to clarify its role in HMF degradation pathway, which may lead to its use as a biocatalyst in the production of furan derivatives.

3.3. Furans utilization pattern of selected strains harboring different hmf/psf gene sets provides support for key role of some enzymes

Based on the analysis of presence/absence profiles of *hmf*/*psf* genes in proteobacterial genomes we selected representative strains harboring distinct gene sets to provide support for key role of some enzymes. Eight strains belonging to α- (1), β- (4), and γ- (3) Proteobacteria were selected to test their ability to grow on different intermediates of HMF pathway as sole carbon and energy source. The selected strains display clusters ranging from full *hmfABCDE-hmfFG-psfA-hmfH* or *hmfABCDE-hmfFG-psfA-psfG* gene sets to solely *hmfABCDE* core genes. As shown in Fig. 5 the strains *Cupriavidus basilensis* DSM 11853 and *Pseudomonas sp*. ALS1279 harboring the *hmfABCDE-hmfFG-psfA-hmfH* or *hmfABCDE-hmfFG-psfA-psfG* genes respectively were able to grow on all intermediates tested with exception of FDCA, suggesting that both gene sets may provide a complete HMF pathway, and that *psfG* or *hmfH* genes could encode analogous functions. The inability to grow using FDCA as substrate is most probably related with the absence of a specific transporter system for this dicarboxylic acid. On the contrary, the strains *Cupriavidus pinatubonensis* JMP134 and *Pseudomonas sp*. ALS1131 harboring only *hmf*ABCDE genes are unable to grow in any intermediate except FA. In the case of *Paraburkholderia xenovorans* LB400 the presence of *hmfFG* genes along with *hmfABCDE* could allow to this strain to proliferate on FDCA as sole carbon source, in addition to FA. Most probably this strain harbors a specific transport system for FDCA uptake that is not found in *C. basilensis* DSM 11853 or *Pseudomonas sp*. ALS1279, whose *hmf* clusters would be devoted to HMF consumption instead FDCA metabolization. Noteworthy, the presence of *psfA* in addition to *hmfFG* and *hmfABCDE* genes in *Pseudomonas mendocina* VN230 and *Azospirillum oryzae* COC8 enable these strains to grow on FFCA besides FDCA and FA as substrates, supporting its role as a FFCA dehydrogenase. Finally, a remarkable case in HMF metabolization is represented by *Paraburkholderia phytofirmans* PsJN. It has been previously reported that this strain is able to grow on HMF although with very low growth rates, even 30 times slower than other strains harboring *hmf* gene clusters [21]. In the culture conditions tested by us, no significant growth in presence of HMF was assessed after 72 h of incubation. However, it was possible to measure significant growth in HMFCA and FFCA, supporting the notion that *hmfH* is able to supply the absence of *psfA* catalyzing consecutive oxidations of HMFCA to generate FDCA, but that the growth on HMF is limited by inefficient conversion into HMFCA either by HmfH or housekeeping aldehyde dehydrogenases. Altogether the results of *in vivo* growth tests support the requirement of a *hmfABCDE-hmfFG-psfA-psfG/hmfH* gene set and most probably a proficient aldehyde dehydrogenase acting on HMF for the complete metabolization of this substrate.

3.4. Transcriptional regulators associated to hmf genes are related to AraC- and LysR-type families.

On the contrary to the metabolic features of the HMF biodegradative route that has been roughly established, the molecular bases of transcriptional regulation of the pathway are mostly unknown. As previously reported, *C. basilensis* HMF14 carries *hmf* genes in two different gene clusters, both preceded by a LysR-type transcriptional regulator-encoding gene located in opposite orientation, termed *hmfR1* and *hmfR2,* which could master the expression of HMF pathway by sensing substrates or intermediates of the route [20], but experimental evidence supporting this role has not been provided yet. In the case of *P. putida* Fu1, the inactivation of a homologous LysR-type regulatory protein (termed PsfB) abolished growth on

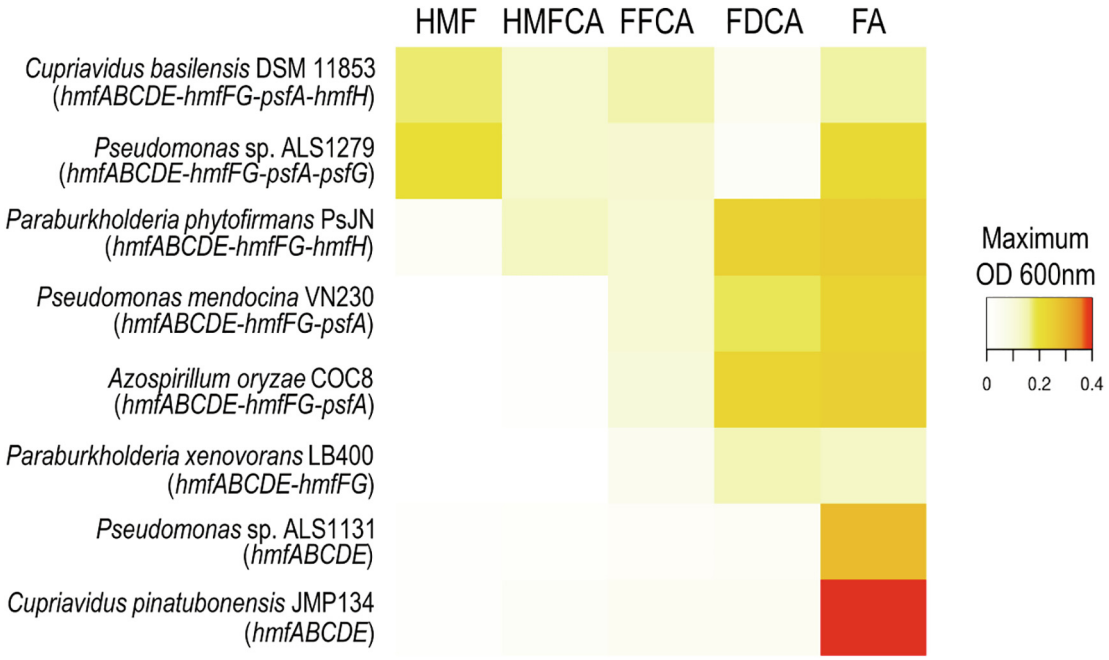

Fig. 5. Growth on furan intermediates of several Proteobacterial species owing different *hmf* and *psf* gene sets. Strains belonging to *Azospirillum*, *Cupriavidus*, *Paraburkholderia* and *Pseudomonas* genus were grown in mineral salt medium with 2.5 mM HMF, HMFCA, FFCA, FDCA or FA as a sole carbon and energy sources. The presence of *hmf* and *psf* genes in each strain is depicted under species name. Shading indicates maximum optical density (OD) at 600 nm measured during 48 h of incubation (average of two biological replicates). HMF, 5-hydroxymethylfurfural; HMFCA, 5-hydroxymethyl-2-furancarboxylic acid; FFCA, 5-formylfuran carboxylic acid; FDCA, 2,5-furandicarboxylic acid; FA, furoic acid.

FA in this strain and their chemotactic responses towards furans [19,60]. Alternatively, it was reported the selection of an adaptive mutation in a regulatory gene coding for an activator of the AraC/XylS family in the furfural-degrading *Pseudomonas pseudoalcaligenes* CECT 5344 strain [61]. This point mutation affected the HTH region of the protein allowing efficient growth of the strain on FA and furfural; and on the contrary, the whole inactivation of the AraC-type regulator-encoding gene reversed the phenotype [61]. Both, LysR- and AraC-type regulators are commonly associated with carbon source assimilation, stress response and pathogenesis in bacteria [62,63]. Their involvement in transcriptional regulation of furan catabolism requires additional work to elucidate their specific roles. Accordingly, we found 82 bacterial genera from the total of selected strains harboring a homologous activator protein belonging to AraC-family in the context of *hmf* genes, whilst a transcriptional regulator belonging to LysR-family was identified in 49 species (Fig. S4). Notably, all bacterial strains that carry exclusively lower pathway-encoding genes (44) only own an AraC-type regulator (Fig. S4; Table S1), suggesting that *hmfABCDE* genes supporting FA metabolism are specifically mastered by this transcriptional regulator, comparable to the scenario reported by *P. pseudoalcaligenes* CECT 5344 [61]. Interestingly, the point mutation (L261R) that turns functional that AraC regulator is in fact a conserved residue (Arg) in the remaining homologues putatively involved in FA catabolism, as revealed by alignment of all AraC proteins identified in this study (data not shown), indicating that the inactive version of the regulator in wild type strain CECT 5344 is exceptional. On the other hand, all bacterial species that contain LysR-type regulators harbor upper and lower pathway genes (Fig. S4), indicating that this transcriptional regulator could be mainly associated to the presence of a full catabolic pathway for furans. As well, 33 bacterial species possessed both AraC- and LysR-type transcriptional regulators associated to *hmf* contexts (Fig. S4), analogous to the scenario described in clusters of *P. putida* Fu1 and ALS1267 [19,21]; suggesting that the upper pathway would be controlled by a LysR-type regulator and the lower route by the previously described AraC-type activator. The identity of the substrates/intermediates acting as effectors of transcriptional activation in strains harboring this putative regulatory architecture deserves further research. Finally, 11 selected strains mostly belonging to β-proteobacteria (7) harboring upper and lower pathway genes, contain solely a LysR-type regulator associated to *hmf* cluster without an accompanying AraC-type regulator (Table S1), therefore these genes could be regulated by a single activator/effector pair. Nevertheless, 6 species belonging to *Burkholderiaceae* family of β-Proteobacteria contain a second copy of LysR-type regulator, analogous to reported by strain HMF14 [20], suggesting that specifically in these strains this extra copy could be adapted to recognize an additional substrate/intermediate.

The identification of substrates/intermediates acting as effectors of the different transcriptional regulators putatively involved in furans catabolism is relevant not only to understand the molecular basis of pathway activation, but also from a biotechnological perspective on their potential for biosensing. Since cytotoxic, genotoxic, and carcinogenic effects of HMF and furans derivatives have been demonstrated in several animal models and mammalian cell cultures, its quantification becomes relevant in alimentary industry because this compound is produced during processing of several foodstuffs [64–67]. As an alternative to sophisticated chromatographic methods for HMF quantification, sensor protein-based biosensors would allow a fast and low-cost procedure to detect this compound and/or their transformation products. Transcriptional activators that sense HMF and/or intermediates of the catabolic route could be a valuable resource to construct biosensors able to detect these molecules, becoming highly desirable to describe the profile of effectors recognized by the proposed regulators.

3.5. Transporters-encoding genes associated to furans biodegradation

The role of transport in biodegradation of furans is largely unknown. The only piece of evidence about an active transport of furans comes from the establishment of a whole cell biocatalyst system in *P. putida* S12, in which the heterologous expression of a transporter termed HmfT1 had a considerable effect on the improvement in FDCA production from HMF mediated by HmfH activity of *C. basilensis* HMF14 [22]. This effect was attributed to a strong decrease in transient HMFCA accumulation, suggesting that the main function of *hmfT1* gene product was uptake of this metabolite and therefore indicating the importance of transporter-encoding genes in furan metabolism [22,52]. HmfT1 belongs to Major Facilitator Superfamily, an ancient, extensive and ubiquitous family able to transport a broad spectrum of ions and solutes across membranes via facilitated diffusion, symport, or antiport [68]; and been assigned to the metabolite:H + symport subfamily [69]. We identified 29 strains that harbor HmfT1 transporter-encoding genes associated to *hmf* clusters, mostly belonging to α- (10) and β-proteobacteria (16) (Table S2). Most of these strains (24) possess upper pathway genes (Fig. S5), supporting that its main function is likely to be the transport of upper pathway substrates or intermediates as proposed in *C. basilensis* HMF14 [20,22]. In addition, we detected 39 species harboring a transporter-encoding gene associated to lower pathway genes (Table S2), which is homologous to a gene described in *P. pseudoalcaligenes* CECT 5344 and termed *benE* despite its linkage to *hmf* genes [61]. Similar to HmfT1, this transporter belongs to Major Facilitator Superfamily, although it was included in a different branch named Benzoate:H + symport subfamily, which first characterized member was the benzoate transporter encoded by *benE* from *Acinetobacter calcoaceticus* [69]. Interestingly, *benE*-like transporters were found primarily in α- (26) and γ-proteobacteria (11), and most strains (32) harbor only lower pathway genes (Fig. S5), suggesting that its main function could be FA transport.

Remarkably, we found that *hmf* clusters from 24 species belonging to α- (9) and β-proteobacteria (15) carry a putative transporter gene closely related to high affinity branched-chain amino acid transport system (LivKHMGF) from *Escherichia coli* (Fig. S5; Table S2), that relies in coupling of ATP-hydrolysis for uptake of substrates [70]; suggesting alternative mechanisms for furans transport. Remarkably, almost all strains (91) harbor one homologue of HmfT1-, BenE- or LivKHMGF-like transporters associated to *hmf* genes and its presence exclude the others (Fig. S5), except for *Bordetella flabilis* AU10664 that harbors two copies of *hmfABCDE* cluster owing either HmfT1- or a LivKHMGF-like transporter associated to each copy. These observations suggest that these three different transporter types could accomplish analogous functions in furans transport. Finally, a diverse array of other transporter types could be also found associated to *hmf* genes but each to a much lower extent (Fig. S5), including homologues related to ABC-type nitrate/sulfonate/bicarbonate transport system (*tauABC*), TRAP-type C4-dicarboxylate transport system (*dctPQM*), TRAP-type uncharacterized transport system (Imp), tripartite tricarboxylate transporter substrate binding protein (*tctC*), and Na^+:solute symporter family protein (SSS). This broad diversity of transporters putatively involved in furans catabolism would reflect that different substrates or intermediates in the route can be subject to active transport in degradative bacteria.

4. Conclusion

A systematic analysis of *hmf* genes in proteobacterial genomes has allowed establishment of their phylogenetic relationships, clusters organization and presence-absence patterns of catabolic

and accessory functions. From this study a previously unrecognized distribution of the HMF catabolic pathway could be observed, expanding the array of potential degradative microorganisms to be recruited in biodetoxifying processes to remove furan inhibitors from biomass hydrolysates. Most importantly, the persistent presence of some genes, or excluding patterns among them, allow us to predict the full gene set required to establish a functional HMF pathway in a different host. In addition to the FA-metabolizing lower route encoded by *hmfABCDE* genes, the presence of *hmfFG – psfA - hmfH*/*psfG* genes it seems to be a key for proper HMF funneling into central metabolism. Furthermore, the recruitment of at least one of the main transporter types described as associated to *hmf* genes would be significant for a proficient metabolic engineering endeavor.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by FONDECYT 1201741, ANID PIA/Anillo ACT172128, ANID PIA/BASAL FB0002 grants.

Author contributions

RD performed most of bioinformatics studies, analyzed and interpreted all the achieved results, and wrote the original draft. FG-T performed preliminary bioinformatic analyses. DP-P conceived and designed throughout the studies, supervised the follow-up, and reviewed and edited the final manuscript. All authors read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.04.017.

References

[1] Kawaguchi H, Hasunuma T, Ogino C, Kondo A. Bioprocessing of bio-based chemicals produced from lignocellulosic feedstocks. Curr Opin Biotechnol 2016;42:30–9.
[2] Galkin KI, Ananikov VP. When Will 5-Hydroxymethylfurfural, the "Sleeping Giant" of Sustainable Chemistry, Awaken?. ChemSusChem 2019;12 (13):2976–82.
[3] Lalanne L, Nyanhongo GS, Guebitz GM, Pellis A. Biotechnological production and high potential of furan-based renewable monomers and polymers. Biotechnol Adv 2021;48:107707. https://doi.org/10.1016/j.biotechadv.2021.107707.
[4] Hu L, He A, Liu X, Xia J, Xu J, Zhou S, et al. Biocatalytic Transformation of 5-Hydroxymethylfurfural into High-Value Derivatives: Recent Advances and Future Aspects. ACS Sustainable Chem Eng 2018;6(12):15915–35.
[5] Global 5-hydroxymethylfurfural (5-HMF) (CAS 67-47-0) Market 2019 by Manufacturers, Regions, Type and Application, Forecast to 2024. Report ID: MSR1550260, Published Date: February 03, 2019.
[6] Rasmussen H, Sørensen HR, Meyer AS. Formation of degradation compounds from lignocellulosic biomass in the biorefinery: sugar reaction mechanisms. Carbohydr Res 2014;385:45–57.
[7] Jönsson LJ, Martín C. Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. Bioresour Technol 2016;199:103–12.
[8] Jönsson LJ, Alriksson B, Nilvebrant N-O. Bioconversion of lignocellulose: inhibitors and detoxification. Biotechnol Biofuels 2013;6(1):16. https://doi.org/10.1186/1754-6834-6-16.
[9] Franden MA, Pienkos PT, Zhang M. Development of a high-throughput method to evaluate the impact of inhibitory compounds from lignocellulosic hydrolysates on the growth of *Zymomonas mobilis*. J Biotechnol 2009;144 (4):259–67.
[10] Nigam JN. Ethanol production from wheat straw hemicellulose hydrolysate by *Pichia stipitis*. J Biotechnol 2001;87(1):17–27.
[11] Almeida JRM, Bertilsson M, Gorwa-Grauslund MF, Gorsich S, Lidén G. Metabolic effects of furaldehydes and impacts on biotechnological processes. Appl Microbiol Biotechnol 2009;82(4):625–38.
[12] Dong H, Bao J. Metabolism: biofuel via biodetoxification. Nat Chem Biol 2010;6 (5):316–8.
[13] Parawira W, Tekere M. Biotechnological strategies to overcome inhibitors in lignocellulose hydrolysates for ethanol production: review. Crit Rev Biotechnol 2011;31(1):20–31.
[14] Nichols NN, Dien BS, Cotta MA. Fermentation of bioenergy crops into ethanol using biological abatement for removal of inhibitors. Bioresour Technol 2010;101(19):7545–50.
[15] Kannisto MS, Mangayil RK, Shrivastava-Bhattacharya A, Pletschke BI, Karp MT, Santala VP. Metabolic engineering of *Acinetobacter baylyi* ADP1 for removal of *Clostridium butyricum* growth inhibitors produced from lignocellulosic hydrolysates. Biotechnol Biofuels 2015;8:198.
[16] Wang Y, Brown CA, Chen R. Industrial production, application, microbial biosynthesis and degradation of furanic compound, hydroxymethylfurfural (HMF). AIMS Microbiol 2018;4(2):261–73.
[17] Igeño MI, Sánchez-Clemente R, Población AG, Guijo MI, Merchán F, Blasco R. Biodegradation of 5-(hydroxymethyl)-furfural and furan derivatives. Proceedings 2018;2(20):1283.
[18] Wierckx N, Koopman F, Ruijssenaars HJ, de Winde JH. Microbial degradation of furanic compounds: biochemistry, genetics, and impact. Appl Microbiol Biotechnol 2011;92(6):1095–105.
[19] Nichols NN, Mertens JA. Identification and transcriptional profiling of *Pseudomonas putida* genes involved in furoic acid metabolism. FEMS Microbiol Lett 2008;284(1):52–7.
[20] Koopman F, Wierckx N, de Winde JH, Ruijssenaars HJ. Identification and characterization of the furfural and 5-(hydroxymethyl)furfural degradation pathways of *Cupriavidus basilensis* HMF14. Proc Natl Acad Sci U S A 2010;107 (11):4919–24.
[21] Crigler J, Eiteman MA, Altman E. Characterization of the Furfural and 5-Hydroxymethylfurfural (HMF) Metabolic Pathway in the Novel Isolate *Pseudomonas putida* ALS1267. Appl Biochem Biotechnol 2020;190(3):918–30.
[22] Wierckx N, Elink Schuurman TD, Blank LM, Ruijssenaars HJ (2015) Whole-Cell Biocatalytic Production of 2,5-Furandicarboxylic Acid. In: Kamm B. (eds) Microorganisms in Biorefineries. Microbiology Monographs, vol 26. Springer, Berlin, Heidelberg.
[23] Hossain GS, Yuan H, Li J, Shin H-D, Wang M, Du G, et al. Metabolic engineering of *Raoultella ornithinolytica* BF60 for production of 2,5-furandicarboxylic acid from 5-hydroxymethylfurfural. Appl Environ Microbiol 2017;83(1). https://doi.org/10.1128/AEM.02312-16.
[24] Guarnieri MT, Ann Franden M, Johnson CW, Beckham GT. Conversion and assimilation of furfural and 5-(hydroxymethyl)furfural by *Pseudomonas putida* KT2440. Metab Eng Commun 2017;4:22–8.
[25] Baptista SL, Costa CE, Cunha JT, Soares PO, Domingues L. Metabolic engineering of *Saccharomyces cerevisiae* for the production of top value chemicals from biorefinery carbohydrates. Biotechnol Adv 2021;47:107697. https://doi.org/10.1016/j.biotechadv.2021.107697.
[26] Koopman F, Wierckx N, de Winde JH, Ruijssenaars HJ. Efficient whole-cell biotransformation of 5-(hydroxymethyl)furfural into FDCA, 2,5-furandicarboxylic acid. Bioresour Technol 2010;101(16):6291–6.
[27] Herrgård M, Panagiotou G. Analyzing the genomic variation of microbial cell factories in the era of "New Biotechnology". Comput Struct Biotechnol J 2012;3 (4):e201210012. https://doi.org/10.5936/csbj.201210012.
[28] Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nuc Acids Res 2008;36(Web Server issue): W5–9.
[29] Chen IA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. NuclAcids Res. 2019; 47(D1):D666-D677.
[30] NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucl Acids Res 2018;46(D1):D8–D13.
[31] Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A. CDD/SPARCLE: the conserved domain database in 2020. Nucl Acids Res. 2020; 48(D1):D265-D268.
[32] Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32(1):268–74.
[33] Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 2017;14(6):587–9.
[34] Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 2018;35 (2):518–22.
[35] Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20(4):1160–6.
[36] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucl Acids Res. 2019;47(W1):W256–9.

[37] Xie CH, Yokota A. *Azospirillum oryzae* sp. nov., a nitrogen-fixing bacterium isolated from the roots of the rice plant *Oryza sativa*. Int J Syst Evol Microbiol 2005;55(Pt 4):1435–8.
[38] Vandamme P, Coenye T. Taxonomy of the genus *Cupriavidus*: a tale of lost and found. Int J Syst Evol Microbiol 2004;54(Pt 6):2285–9.
[39] Lykidis A, Pérez-Pantoja D, Ledger T, Mavromatis K, Anderson IJ, Ivanova NN, et al. The complete multipartite genome sequence of *Cupriavidus necator* JMP134, a versatile pollutant degrader. PLoS ONE 2010;5(3):e9729.
[40] Weilharter A, Mitter B, Shin MV, Chain PSG, Nowak J, Sessitsch A. Complete genome sequence of the plant growth-promoting endophyte *Burkholderia phytofirmans* strain PsJN. J Bacteriol 2011;193(13):3383–4.
[41] Chain PSG, Denef VJ, Konstantinidis KT, Vergez LM, Agullo L, Reyes VL, et al. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. Proc Natl Acad Sci U S A 2006;103(42):15280–7.
[42] Donoso R, González-Toro F, Guajardo-Parra M, Araya-Nail M, Farkas C, Maldonado-Agurto R, et al. Genome sequencing and characterization of novel proteobacterial strains having the ability to use furan aldehydes as only carbon source. N Biotechnol 2018;44:S84–5.
[43] Farkas C, Donoso RA, Gárate-Castro C, Villegas P, Durán RE, Seeger M, et al. Draft genome sequences of two *Pseudomonas* Strains that are able to use furan derivatives as their sole carbon source. Microbiol Resour Announc. 2020;9(2). https://doi.org/10.1128/MRA.01131-19.
[44] Lee SA, Wrona LJ, Cahoon AB, Crigler J, Eiteman MA, Altman E. Isolation and characterization of bacteria that use furans as the sole carbon source. Appl Biochem Biotechnol 2016;178(1):76–90.
[45] Lin J, Yang G, Tang J, Li Z, Yu Z, Zhuang Li. *Pseudoxanthomonas composti* sp. nov., isolated from compost. Antonie Van Leeuwenhoek 2019;112(8):1213–9.
[46] Wu G, Liu Y, Li Q, Du H, You J, Li H, Ke C, Zhang X, Yu J, Zhao T. Luteimonas huabeiensis sp. nov., isolated from stratum water. Int J Syst Evol Microbiol. 2013; 63(Pt 9):3352-3357.
[47] Marshall SA, Payne KAP, Leys D. The UbiX-UbiD system: The biosynthesis and use of prenylated flavin (prFMN). Arch Biochem Biophys 2017;632:209–21.
[48] Saaret A, Balaikaite A, Leys D. Biochemistry of prenylated-FMN enzymes. Enzymes 2020;47:517–49.
[49] Payne KAP, Marshall SA, Fisher K, Cliff MJ, Cannas DM, Yan C, et al. Enzymatic carboxylation of 2-furoic acid yields 2,5-furandicarboxylic acid (FDCA). ACS Catal 2019;9(4):2854–65.
[50] Yuan H, Li J, Shin H-D, Du G, Chen J, Shi Z, et al. Improved production of 2,5-furandicarboxylic acid by overexpression of 5-hydroxymethylfurfural oxidase and 5-hydroxymethylfurfural/furfural oxidoreductase in *Raoultella ornithinolytica* BF60. Bioresour Technol 2018;247:1184–8.
[51] Yuan H, Liu Y, Li J, Shin H-D, Du G, Shi Z, et al. Combinatorial synthetic pathway fine-tuning and comparative transcriptomics for metabolic engineering of *Raoultella ornithinolytica* BF60 to efficiently synthesize 2,5-furandicarboxylic acid. Biotechnol Bioeng 2018;115(9):2148–55.
[52] Pham NN, Chen C-Y, Li H, Nguyen MTT, Nguyen PKP, Tsai S-L, et al. Engineering Stable *Pseudomonas putida* S12 by CRISPR for 2,5-Furandicarboxylic Acid (FDCA) Production. ACS Synth Biol 2020;9(5):1138–49.
[53] Dijkman WP, Fraaije MW. Discovery and characterization of a 5-hydroxymethylfurfural oxidase from *Methylovorus* sp. strain MP688. Appl Environ Microbiol 2014;80(3):1082–90.
[54] Hsu C-T, Kuo Y-C, Liu Y-C, Tsai S-L. Green conversion of 5-hydroxymethylfurfural to furan-2,5-dicarboxylic acid by heterogeneous expression of 5-hydroxymethylfurfural oxidase in *Pseudomonas putida* S12. Microb Biotechnol 2020;13(4):1094–102.
[55] Viña-Gonzalez J, Martinez AT, Guallar V, Alcalde M. Sequential oxidation of 5-hydroxymethylfurfural to furan-2,5-dicarboxylic acid by an evolved aryl-alcohol oxidase. Biochim Biophys Acta Proteins Proteom 2020;1868 (1):140293. https://doi.org/10.1016/j.bbapap.2019.140293.
[56] Viñambres M, Espada M, Martínez AT, Serrano A. Screening and Evaluation of New Hydroxymethylfurfural Oxidases for Furandicarboxylic Acid Production. Appl Environ Microbiol 2020;86(16):e00842–e920.
[57] Wu S, Liu Q, Tan H, Zhang F, Yin H. A Novel 2,5-Furandicarboxylic Acid Biosynthesis Route from Biomass-Derived 5-Hydroxymethylfurfural Based on the Consecutive Enzyme Reactions. Appl Biochem Biotechnol 2020;191(4):1470–82.
[58] Kniemeyer O, Heider J. (*S*)-1-phenylethanol dehydrogenase of *Azoarcus* sp. strain EbN1, an enzyme of anaerobic ethylbenzene catabolism. Arch Microbiol 2001;176(1-2):129–35.
[59] Kavanagh KL, Jörnvall H, Persson B, Oppermann U. Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell Mol Life Sci 2008;65(24):3895–906.
[60] Nichols NN, Lunde TA, Graden KC, Hallock KA, Kowalchyk CK, Southern RM, et al. Chemotaxis to furan compounds by furan-degrading *Pseudomonas* strains. Appl Environ Microbiol 2012;78(17):6365–8.
[61] Igeño MI, Macias D, Blasco R. A Case of Adaptive Laboratory Evolution (ALE): Biodegradation of Furfural by *Pseudomonas pseudoalcaligenes* CECT 5344. Genes (Basel). 2019;10(7):499.
[62] Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL. Arac/XylS family of transcriptional regulators. Microbiol Mol Biol Rev 1997;61(4):393–410.
[63] Maddocks SE, Oyston PCF. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology (Reading) 2008;154(Pt 12):3609–23.
[64] Kowalski S, Lukasiewicz M, Duda-Chodak A, Zięć G. 5-Hydroxymethyl-2-furfural (HMF)–heat-induced formation occurrence in food and biotransformation—a review. Pol J Food Nutr Sci. 2013;63(4):207–25.
[65] Islam MN, Khalil MI, Islam MA, Gan SH. Toxic compounds in honey. J Appl Toxicol 2014;34(7):733–42.
[66] Koszucka A, Nowak A. Thermal processing food-related toxicants: a review. Crit Rev Food Sci Nutr 2019;59(22):3579–96.
[67] Farag MR, Alagawany M, Bin-Jumah M, Othman SI, Khafaga AF, Shaheen HM, et al. The toxicological aspects of the heat-borne toxicant 5-hydroxymethylfurfural in animals: a review. Molecules 2020;25(8):1941. https://doi.org/10.3390/molecules25081941.
[68] Yan N. Structural biology of the major facilitator superfamily transporters. Annu Rev Biophys 2015;44(1):257–83.
[69] Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. Microbiol Mol Biol Rev 1998;62(1):1–34.
[70] Adams MD, Wagner LM, Graddis TJ, Landick R, Antonucci TK, Gibson AL, et al. Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli*. J Biol Chem 1990;265 (20):11436–43.