**Supplementary Document**

Title: **Systems Biology-Guided Biodesign of Consolidated Lignin Conversion**

Lu Lin[1,2,3,4], Yanbing Cheng[1,3,4], Yunqiao Pu[5], Su Sun[1,4,6], Xiao Li[1,3,4], Mingjie Jin[7], Elizabeth A. Pierson[8], Dennis C. Gross[3], Bruce E. Dale[7], Susie Y. Dai[1,6,9], Arthur J. Ragauskas[5,10], Joshua S. Yuan[1,3,4,*]

1. Synthetic and Systems Biology Innovation Hub (SSBiH), Texas A&M University, College Station, TX, 77843
2. College of Ocean Sciences, Zhejiang University, Hangzhou, China
3. Department of Plant Pathology and Microbiology, Texas A&M University, College Station, TX 77843
4. Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843
5. Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
6. Department of Veterinary Pathobiology, Texas A&M University, College Station, TX 77843
7. Department of Chemical Engineering and Materials Science, Michigan State University, East Lansing, MI 48824
8. Department of Horticulture Science, Texas A&M University, College Station, TX 77843
9. Office of the Texas State Chemist, Texas A&M University, College Station, TX 77843
10. Department of Chemical and Biomolecular Engineering, Department of Forestry, Wildlife, and Fisheries, University of Tennessee, Knoxville, TN 37996

**Running title:** Biodesign of consolidated lignin conversion

* For correspondence: Joshua S. Yuan, 2123 TAMU, College Station, TX 77843, syuan@tamu.edu

## Part I. Methods

*1.1. Lignin pretreatment experiment*

To optimize the fermentation conditions, lignin was pretreated with sodium hydroxide. Briefly, 1% lignin M9 in medium was adjusted to pH 12.5 with 50% NaOH solution to dissolve the lignin powder. Subsequently, to support strain growth, lignin medium was then adjusted to pH 7.0-7.5 using 2 M HCl.

*1.2. Whole-genome sequencing and sequence analysis*

Shotgun sequencing libraries were constructed [1] and sequenced using two platforms. Illumina sequencing was performed at the Texas A&M Agrilife Genomics Core with a paired-end library of 100-bp reads on an Illumina HiSeq 2500. The PacBio sequencing was performed by the Duke University IGSP Genome Sequencing & Analysis Core Resources with a library of 5 kb insert size. Illumina short reads were first assembled to obtain contigs using SOAPdenovo [2], and the contigs were further analyzed using a hybrid assembly approach to integrate the pre-assembled contigs with PacBio long reads following the RS_AHA_Scaffolding analysis protocol in SMRT Analysis v2.2.0 (Pacific Biosciences of California, Inc). Gene models were predicted using the online server Prodigal v1.20 [3] and were blasted against the Uniprot database for functional annotation. A whole-genome atlas for the A514 strain was generated using GeneWiz browser v0.94 (http://www.cbs.dtu.dk/services/gwBrowser/).

Genome and gene sequence data for twelve *P. putida* strains with publicly available completed genomes were downloaded from NCBI. Protein sequence similarity searches and multiple protein sequence alignments were performed to identify the genes involved

in lignin depolymerization, aromatic compound catabolism and PHA synthesis in the A514 strain and twelve other *P. putida* strains [4-7].

*1.3. Total protein isolation*

For intracellular total protein extraction, cells from 100 mL of culture were harvested in the late-exponential phase by centrifugation (9000 rpm for 10 min at 4 °C) and washed twice with 10 mL PBS buffer (pH 7.4). Cells were then resuspended in 5 mL Alkaline-SDS buffer (5% SDS; 50 mM Tris-HCl, pH 8.5; 150 mM NaCl; 0.1 mM EDTA; 1 mM $MgCl_2$; 50 mM DTT) with 1 mM PMSF. Cell lysis was performed via tip-probe sonication (Branson, Danbury CT) using 4 cycles of 30 s. Lysates were then centrifuged at 12,000 g for 15 min at 4 °C to collect supernatant. For the secretome analysis, extracellular protein was extracted from culture supernatants harvested in the late-exponential phase by centrifugation. The supernatant (from extracellular supernatant or intracellular lysates) was precipitated with 100% trichloroacetic acid (TCA) overnight at -20 °C. The protein pellet was collected by centrifugation at 12,000 g at 4 °C for 30 min and washed three times with chilled acetone. A SpeedVac was used to remove residual acetone, and Reagent Type 4 Working Solution, provided in the Total Protein Extraction Kit (Sigma-Aldrich, St. Louis, MO), was used to incubate the pellet. The pellet was then centrifuged at 13,000 g for 30 min, and supernatant was collected and stored at -80 °C for future proteomics use. All of the samples were prepared and analyzed in duplicate.

*1.4. MudPIT-based shotgun proteomics experiments and data analysis*

Approximately 100 μg of each protein sample was digested with Trypsin Gold, Mass Spectrometry Grade (Promega, WI, USA), and desalted using a Sep-Pak plus C18

column. The desalted sample was then loaded onto a biphasic capillary column. The two-dimensional liquid chromatography separation and tandem mass spectrometry conditions followed the previously described protocol [8, 9]. Tandem mass spectra were extracted from the raw files and converted into an MS2 file. The MS2 file was searched against the A514 protein database generated from gene models created in the genomics analysis. Details of the proteomics data analysis have been previously described[8, 9].

*1.5. Plasmid and strain construction*

For the lignin depolymerization module, putative signal peptide sequences from secreted proteins that were identified in the secretome proteomics analysis were predicted using SignalP 4.0 [10]. The DNA fragments encoding signal peptides from phosphate-binding protein (Pbp, PputA514_0110), outer membrane lipoprotein (OprI, PputA514_0982), outer membrane porin F (OprF, PputA514_3701) and *E. carotovora* PelB [11] were amplified from the A514 genome and were then fused to the multifunctional dye peroxidase (DyP2) coding sequence from *Amycolatopsis* sp. 75iv2 [12] to produce 1.4-kb PCR fragments. These fragments were sub-cloned into plasmid pGP1099 to generate pGPbpDyP2, pGOprIDyP2, pGOprFDyP2 and pGPelBDyP2, respectively (Table S2). In addition, a 2.3-kb PCR fragment encoding the 2-kb vanillate demethylase A (*vanA*) and vanillate O-demethylase oxidoreductase (*vanB*) genes (PputA514_0644-0645) flanked by their 281-bp promoter ($P_{van}$) was amplified from the A514 genome and ligated into the plasmid pPROBE-GT to generate pGVAN. Similarly, a PCR fragment containing the $P_{van}$ promoter, *phaJ4*, and either the *phaC1* or the *phaC2* gene was amplified and sub-cloned into plasmid pPROBE-GT to construct pGJ4C1 and pGJ4C2 (Table S2). The resultant

plasmids were transformed into A514 cells through chemical transformation [13], followed by selection with 30 μg/mL gentamicin.

To integrate all three functional modules, the lignin depolymerization module and aromatic compound utilization module were integrated into one plasmid. $P_{1099}$-*pelb*/*dyp2* (amplified from pGPelbDyP2) and *vanAB* (amplified from pGVAN) were combined into a 4-kb fragment and subcloned into the plasmid pPROBE-TT to derive the construct pTDV (Fig. S2) [14]. The plasmids pTDV and pGJ4C1 were co-transformed into A514 cells, and the transformants were subjected to selection on LB agar supplemented with 30 μg/mL gentamicin and 15 μg/mL tetracycline.

*1.6. The heterologous DyP2 enzyme activity assay*

The DyP2 activity assay was performed at 25 °C using ABTS (2, 2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid, Sigma-Aldrich, St. Louis, MO) as the substrate with a BioTek spectrophotometer. The assay reactions were performed in a 300-μl mixture containing 50 mM sodium acetate buffer at pH 4.5, 0.5 mM $H_2O_2$, 2.5 mM ABTS, and 150 μl extracellular supernatants from A514 engineering strains. Activity was monitored based on an increase in absorption at 420 nm ($\varepsilon 420 = 36\ mM^{-1}\ cm^{-1}$) [15]. All experiments were prepared and analyzed in triplicate.

*1.7. Vanillic acid degradation assay*

Cells were grown in M9 minimal medium containing 50 mM vanillic acid. At the indicated time points, the 1 mL cell cultures were centrifuged for 5 min at 16,000 g, and the concentration of vanillic acid in the supernatant was measured via absorbance at 289 nm, as previously described [16].

*1.8. PHA extraction*

Liquid cultures (100 mL) were harvested in the stationary phase and were centrifuged at 9,000 g for 15 min at 4 °C, washed twice with 15 mL Nanopure water, and lyophilized (Lyophilizer alpha 1–4 LSC, Christ, Germany) at -59 °C and 0.140 mbar for a minimum of 24 h. The lyophilized cells were weighed and then supplemented with chloroform at a ratio 1 : 15 (w ⁄ v) and shaken at 60 °C and 120 rpm overnight to extract polymer produced in the cells. The chloroform solution was filtered to remove any cell debris and then concentrated by rotary evaporation. PHA was precipitated by adding a 10-fold volume of pre-chilled methanol. The methanol–chloroform mixture was decanted, and the pure polymer was washed with fresh iced methanol. To obtain the pure product, the precipitant was then re-dissolved in chloroform, and the process was repeated for clean-up. The collected PHA was dried at 25 °C in a vacuum dryer to remove all residual solvent and weighed. Note: chloroform is a hazardous solvent[17]. The use of chloroform was followed the laboratory biosafety guidelines in this study, including the storage of hazardous chemicals, hazardous waste operations and emergency response (http://rmtd.mt.gov/safetylosscontrol/library/labchem).

*1.9. GC/MS (Gas Chromatograph/Mass Spectrometry) analysis of PHA composition*

For GC/MS analysis of the PHA polymer, wild-type A514 and recombinant A514 strains were prepared and lyophilized as described above. The dry cells (~15 mg) were dissolved in 2 mL of methanol-sulfuric acid (85:15) solution and 2 mL chloroform containing 0.01 mg/mL 3-methylbenzoic acid (internal standard) and then incubated at 100 °C for 4 h. After cooling, 1 mL of demineralized water was added to and the organic phase, which contains the resulting methyl esters of monomers. The organic phase was

filtered and analyzed using a GC/MS 2010 SE plus Gas Chromatograph (Shimadzu, Japan) as previously described [18]. All experiments were performed in triplicate.

*1.10. Lignin Analysis*

All of the NMR experiments were performed on a Bruker Avance 400-MHz spectrometer[19, 20]. For quantitative $^{31}P$ NMR analysis, lignin (ca. 20 mg) was dissolved in a solvent of pyridine/CDCl3 (1.6/1.0 v/v, 500 μL) and derivatized with 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane [21]. The spectrum was acquired using an inverse-gated decoupling pulse sequence (WaltFz-16), a 90° pulse, and a 25-s pulse delay. N-hydroxy-5 norbornene-2, 3-dicarboximide (4.0 mg/mL) was used as the internal standard. A total of 128 scans were collected for each sample. NMR data were processed using TopSpin 2.1 (Bruker BioSpin) and MestreNova (Mestre Labs) software packages.

**Part II. Additional details for the lignin degradation mechanisms of the A514 strain, as revealed in the genomics study.**

The draft genome of the A514 strain was annotated and compared with twelve other sequenced *P. putida* strains (Table S3). In addition to enzymes involved in lignin depolymerization, the comparative genomics analysis identified genes encoding enzymes potentially involved in peripheral reactions to lignin degradation in the A514 genome. These enzymes included oxidases (glyoxal oxidase, glucose oxidase and methanol oxidase) to generate $H_2O_2$ production for sustainable peroxidase function, quinone reductases to reduce iron for the Fenton reaction, and cytochrome P450 monooxygenases for the downstream events of the lignin degradation process [4] (Fig. 2B and Table S3). Overall, the genome analysis revealed that the A514 genome contains genes that potentially encode enzymes required for lignin depolymerization.

The A514 strain also contains catabolic pathways suitable for degrading lignin-derived aromatic oligomers into aromatic monomers, as discussed in the main text [5, 22]. Three major pathways were identified based on gene models derived from the A514 genome [23, 24]. First, a β-aryl ether degradation pathway was identified. β-aryl ether is the most abundant linkage in lignin (50-70%) [6] and can potentially be catabolized by a cascade of reactions in the A514 strain. The chemical linkage can first be oxidized by NAD-dependent dehydrogenase LigD (PputA514_6255) and then converted to ketone aromatic monomers by a glutathione-dependent β-etherase (LigEFG, PputA514_4050 and PputA514_5894). The ketone product can be subsequently catabolized to vanillic acid, a key central metabolite for the β-ketoadipate pathway (Fig. S3 and Fig. S4). Second, a biphenyl degradation pathway was identified. Depending on the lignin source,

the biphenyl component could account for up to 10% of lignin. The *bphABCD* genes (PputA514_5045, PputA514_3697 and PputA514_6181) encode enzymes to transform the biphenyl component into chloro-benzoate, which can be further catabolized into catechol, another key intermediate of the β-ketoadipate pathway (Fig. S3 and S4). Third, a diarylpropane degradation pathway was observed in the A514 strain. In fungi, the $C_\alpha$-$C_\beta$ bond is often cleaved by a lignin peroxidase to produce aromatic aldehyde products [5]. The A514 genome contains a gene potentially encoding lignostilbene dioxygenase (PputA514_2615) to catalyze the oxidative cleavage of lignostilbene to produce two molecules of vanillic acid (Fig. S4). The lignostilbene dioxygenase is expected to play an important role in catalyzing the cleavage of $C_\alpha$-$C_\beta$ linkage in relevant oligomers.

**Part III. Mechanisms for lignin utilization and conversion, as revealed by proteomics analysis of the A514 strain under different carbon sources.**

Systems biology analysis of lignin utilization mechanisms in the A514 strain is essential for the design of a lignin bioconversion strategy. A proteomics-based systems biology approach revealed that the efficient conversion of lignin to PHA required synergy at multiple steps, including lignin depolymerization, aromatic compound degradation, and PHA biosynthesis. The proteomics analysis identified 1500, 1303 and 1760 proteins produced by A514 cells grown on glucose, VA and lignin substrates, respectively. The proteome covered 23-27% of the 6491 predicted proteins of *P. putida* A514. The proteome for each carbon source included approximately 250 to 500 uniquely expressed proteins. Differential protein expression patterns were analyzed using glucose substrate as the control. The study revealed that lignin and vanillic acid induced a broad spectrum of aromatic compound degradation pathways (Fig. 1C). The Gene Ontology (GO) analysis of differentially expressed proteins indicated that vanillic acid treatment induced the following functional categories: cellular amino acid metabolism, generation of precursor metabolites, energy and catabolic metabolism, catabolic process, and lipid metabolic process; while the cells grown in lignin substrate had the following GO terms enriched: cellular amino acid metabolism, catabolic metabolism, and response to stress. The detailed pathways for lignin utilization are as follows.

*Lignin depolymerization*

Proteomics analysis suggested that two complementary systems for lignin depolymerization were induced when A514 cells were grown on lignin as the sole carbon source. First, as discussed in the main text, a dye peroxidase-based system was induced

for lignin depolymerization. A dye peroxidase (DyP, PputA514_2985) was coordinately up-regulated with oxidases to potentially achieve sustainable and synergistic lignin depolymerization (Fig. 2A and Fig. 2B). In addition to the DyP, an intracellular generic peroxidase, peroxidasin (PputA514_3972), was also over-expressed (Fig. 2A), but its relevance to lignin depolymerization could not be established. The limited number of peroxidases induced by lignin treatment highlighted the limitations of bacterial systems in lignin degradation compared to fungal systems[4]. For this reason, we designed the functional module to enhance the peroxidase-based lignin depolymerization system. Second, quinone reductases involved in iron reduction and hydroxyl radical recycling were up-regulated (Fig. 2A and 2B). The result suggested an active Fenton reaction system for lignin depolymerization in A514 cells. In addition, cytochrome P450 monooxygenases and other enzymes related to downstream lignin degradation were also up-regulated [4] (Fig. 2A). Overall, the proteomics analysis validated the idea that a lignin depolymerization capacity existed in the A514 strain, even though the enzyme system was much less sophisticated and efficient than those of white rot fungus. Therefore, enhancing the lignin depolymerization capacity of the A514 strain is warranted.

*Catabolic pathways for lignin-derived aromatic compounds*

*Pseudomonas* strains have been well established for their capacity to mineralize and degrade a variety of aromatic compounds, including lignin derivatives and toxic aromatic pollutants [25]. As discussed in the main text, the comparative genomics and proteomics analyses both indicated that the A514 strain degrades lignin-derived aromatics through a catabolic network, where different peripheral pathways funnel structurally diverse substrates into a few key intermediates, such as vanillic acid. These intermediates can be

subsequently ring-cleaved and converted in a few central pathways to enter primary metabolic pathways through acetyl-CoA, succinate and other TCA cycle compounds [7, 25]. First, as discussed in the main text, the protocatechuate branch of the β-ketoadipate pathway was activated to catabolize the most common phenylpropenoids derived from lignin [25]. Enzymes in both the protocatechuate branch of the β-Ketoadipate pathway and the relevant peripheral pathways were over-expressed under lignin and vanillic acid treatments (Fig. 1C and Table S1). Second, the catechol branch of the β-ketoadipate pathway was also induced when culturing A514 cells on lignin (Fig. 1C and Table S1). Complimentary to the protocatechuate branch of the β-ketoadipate pathway, the catechol branch of the β-ketoadipate pathway is able to catabolize aromatic compounds, e.g., benzene and salicylate. For example, salicylate hydroxylase (PputA514_4661) was up-regulated on lignin substrate to catalyze salicylate to catechol, which could, in turn, be transformed to acetyl-CoA via the catechol branch of the β-ketoadipate pathway. Third, besides the β-ketoadipate pathway for ortho cleavage of aromatic compounds, alternative *meta*-cleavage or *para*-cleavage routes were also induced in A514 cells under lignin, highlighting the versatile metabolism in the A514 strain to degrade various lignin-derived aromatic compounds (Fig. 1C). Fourth, the homogentisate pathway and its corresponding peripheral pathways were up-regulated when A514 cells utilized lignin as the sole carbon source. In particular, TyrB (aromatic amino acid aminotransferase, PputA514_6118) and Hpd (4-hydroxyphenylpyruvate dioxygenase, PputA514_2311) were both induced on lignin substrate compared to glucose. These enzymes are involved in degrading aromatic amino acids, e.g., phenylalanine and tyrosine, into homogentisate, which, in turn, served as a central intermediate to be further degraded into acetoacetate and fumarate via the

homogentisate pathway. The induction of aromatic amino acid degradation pathways might be part of the overall augmentation of aromatic compound degradation.

Proteomics analysis revealed a coordinative up-regulation of cofactor-related proteins along with the enzymes for aromatic compound catabolism under lignin treatment. Most of the catabolic enzymes require cofactors, such as manganese, heme and copper, for electron transfer in redox reactions. For example, ferric iron uptake regulator, heme transporter, and TonB-dependent siderophore receptor (PputA514_4501, PputA514_2027 and PputA514_2407) were all induced when A514 cells were grown on lignin. The coordinative expression of these proteins provided iron and other cofactors for enzymes (e.g., alcohol dehydrogenase and aldehyde dehydrogenase) to perform redox and other reactions for lignin degradation.

**References**


1. L. Camarena, V. Bruno, G. Euskirchen, S. Poggio and M. Snyder, *PLoS Pathog.*, 2010, **6**, e1000834.
2. R. Li, H. Zhu, J. Ruan, W. Qian, X. Fang, Z. Shi, Y. Li, S. Li, G. Shan, K. Kristiansen, H. Yang and J. Wang, *Genome Res.*, 2010, **20**, 265-272.
3. D. Hyatt, G. L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer and L. J. Hauser, *BMC bioinformatics*, 2010, **11**, 119.
4. D. Floudas, M. Binder, R. Riley, K. Barry, R. A. Blanchette, B. Henrissat, A. T. Martinez, R. Otillar, J. W. Spatafora, J. S. Yadav, A. Aerts, I. Benoit, A. Boyd, A. Carlson, A. Copeland, P. M. Coutinho, R. P. de Vries, P. Ferreira, K. Findley, B. Foster, J. Gaskell, D. Glotzer, P. Gorecki, J. Heitman, C. Hesse, C. Hori, K. Igarashi, J. A. Jurgens, N. Kallen, P. Kersten, A. Kohler, U. Kues, T. K. Kumar, A.

Kuo, K. LaButti, L. F. Larrondo, E. Lindquist, A. Ling, V. Lombard, S. Lucas, T. Lundell, R. Martin, D. J. McLaughlin, I. Morgenstern, E. Morin, C. Murat, L. G. Nagy, M. Nolan, R. A. Ohm, A. Patyshakuliyeva, A. Rokas, F. J. Ruiz-Duenas, G. Sabat, A. Salamov, M. Samejima, J. Schmutz, J. C. Slot, F. St John, J. Stenlid, H. Sun, S. Sun, K. Syed, A. Tsang, A. Wiebenga, D. Young, A. Pisabarro, D. C. Eastwood, F. Martin, D. Cullen, I. V. Grigoriev and D. S. Hibbett, *Science*, 2012, **336**, 1715-1719.

5. T. D. Bugg, M. Ahmad, E. M. Hardiman and R. Rahmanpour, *Nat. Prod. Rep*, 2011, **28**, 1883-1896.
6. E. Masai, Y. Katayama and M. Fukuda, *Biosci. Biotechnol. Biochem.*, 2007, **71**, 1-15.
7. J. I. Jimenez, B. Minambres, J. L. Garcia and E. Diaz, *Environ. Microbiol*, 2002, **4**, 824-841.
8. Y. Zhang, S. Liu, S. Y. Dai and J. S. Yuan, *BMC bioinformatics*, 2012, **13 Suppl 15**, S8.
9. Y. Zhang, P. Gao, Z. Xing, S. Jin, Z. Chen, L. Liu, N. Constantino, X. Wang, W. Shi, J. S. Yuan and S. Y. Dai, *Mol. Cell Proteomics*, 2013, **12**, 3431-3442.
10. T. N. Petersen, S. Brunak, G. von Heijne and H. Nielsen, *Nat Methods*, 2011, **8**, 785-786.
11. T. Dammeyer, M. Steinwand, S. C. Kruger, S. Dubel, M. Hust and K. N. Timmis, *Microb. Cell. Fact*, 2011, **10**, 11.
12. M. E. Brown, T. Barros and M. C. Chang, *ACS Chem. Biol.*, 2012, **7**, 2074-2081.
13. A. A. Mercer and J. S. Loutit, *J. Bacteriol.*, 1979, **140**, 37-42.

14. W. G. Miller, J. H. Leveau and S. E. Lindow, *Mol. Plant Microbe. Interact.*, 2000, **13**, 1243-1250.

15. Y. Shi, L. Chai, C. Tang, Z. Yang, H. Zhang, R. Chen, Y. Chen and Y. Zheng, *Biotechnol. Biofuels*, 2013, **6**, 1.

16. M. Thanbichler, A. A. Iniesta and L. Shapiro, *Nucleic Acids Res.*, 2007, **35**, e137.

17. D. Prat, A. Wells, J. Hayler, H. Sneddon, C. R. McElroy, S. Abou-Shehada and P. J. Dunn, *Green Chemistry*, 2016, **18**, 288-296.

18. Q. Wang and C. T. Nomura, *J. Biosci. Bioeng.*, 2010, **110**, 653-659.

19. B. B. Hallac, Y. Pu and A. J. Ragauskas, *Energy & Fuels*, 2010, **24**, 2723-2732.

20. Y. Q. Pu, F. Chen, A. Ziebell, B. H. Davison and A. J. Ragauskas, *Bioenerg. Res.*, 2009, **2**, 198-208.

21. Y. Pu, S. Cao and A. J. Ragauskas, *Energy Environ. Sci.*, 2011, **4**, 3154-3166.

22. T. D. Bugg, M. Ahmad, E. M. Hardiman and R. Singh, *Curr. Opin. Biotechnol*, 2011, **22**, 394-400.

23. X. Wu, S. Monchy, S. Taghavi, W. Zhu, J. Ramos and D. van der Lelie, *FEMS microbiology reviews*, 2011, **35**, 299-323.

24. A. J. Ragauskas, G. T. Beckham, M. J. Biddy, R. Chandra, F. Chen, M. F. Davis, B. H. Davison, R. A. Dixon, P. Gilna, M. Keller, P. Langan, A. K. Naskar, J. N. Saddler, T. J. Tschaplinski, G. A. Tuskan and C. E. Wyman, *Science*, 2014, **344**, 1246843.

25. J. I. Jiménez, J. Nogales, J. L. García and E. Díaz, in *Handbook of Hydrocarbon and Lipid Microbiology*, ed. K. Timmis, Springer Berlin Heidelberg, 2010, ch. 91, pp. 1297-1325.