



**HAL**  
open science

# Engineering microbial pathways for production of bio-based chemicals from lignocellulosic sugars: current status and perspectives

Jean François, Ceren Alkim, Nicolas Morin

## ► To cite this version:

Jean François, Ceren Alkim, Nicolas Morin. Engineering microbial pathways for production of bio-based chemicals from lignocellulosic sugars: current status and perspectives. *Biotechnology for Biofuels*, 2020, 13, pp.118. 10.1186/s13068-020-01744-6 . hal-02912905

**HAL Id: hal-02912905**

**<https://hal.inrae.fr/hal-02912905v1>**

Submitted on 7 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

REVIEW

Open Access



# Engineering microbial pathways for production of bio-based chemicals from lignocellulosic sugars: current status and perspectives

Jean Marie Francois<sup>1,2\*</sup> , Ceren Alkim<sup>1,2</sup> and Nicolas Morin<sup>1,2</sup>

## Abstract

Lignocellulose is the most abundant biomass on earth with an annual production of about  $2 \times 10^{11}$  tons. It is an inedible renewable carbonaceous resource that is very rich in pentose and hexose sugars. The ability of microorganisms to use lignocellulosic sugars can be exploited for the production of biofuels and chemicals, and their concurrent biotechnological processes could advantageously replace petrochemicals' processes in a medium to long term, sustaining the emerging of a new economy based on bio-based products from renewable carbon sources. One of the major issues to reach this objective is to rewire the microbial metabolism to optimally configure conversion of these lignocellulosic-derived sugars into bio-based products in a sustainable and competitive manner. Systems' metabolic engineering encompassing synthetic biology and evolutionary engineering appears to be the most promising scientific and technological approaches to meet this challenge. In this review, we examine the most recent advances and strategies to redesign natural and to implement non-natural pathways in microbial metabolic framework for the assimilation and conversion of pentose and hexose sugars derived from lignocellulosic material into industrial relevant chemical compounds leading to maximal yield, titer and productivity. These include glycolic, glutaric, mesaconic and 3,4-dihydroxybutyric acid as organic acids, monoethylene glycol, 1,4-butanediol and 1,2,4-butanetriol, as alcohols. We also discuss the big challenges that still remain to enable microbial processes to become industrially attractive and economically profitable.

**Keywords:** White biotechnology, Lignocellulose, Microbial physiology, Metabolic engineering, Synthetic biology, Bio-based products, Chemicals

## Background

The 21st century is challenging the human society with major social and economic issues owing to the irreversible global warming and the steady rise of environmental pollution, both events mainly resulting from our slave reliance on fossil resources. The expected rarefaction of these resources prompted us to innovate in alternative

and sustainable solutions, which in particular must revive the natural cycles of carbon and nitrogen. One of the cornerstones of this new societal and economical deal coined 'Bioeconomy' [1, 2] stands on our ability to exploit renewable carbon resources to produce environmentally gentle fuels and chemicals that shall profitably replace those derived at present from fossil resources. Among the potential renewable resources, lignocellulosic biomass is the most abundant feedstock on earth with an annual output of around  $2 \times 10^{11}$  metric tons [3]. This feedstock is a non-edible carbon source, which enables to circumvent the resource contention for 'food' versus 'chemicals

\*Correspondence: fran\_jm@insa-toulouse.fr; jean.marie-francois@inra.fr

<sup>1</sup> Toulouse Biotechnology Institute, CNRS, INRA, LISBP INSA, 135 Avenue de Rangueil, Toulouse Cedex 04 31077, France

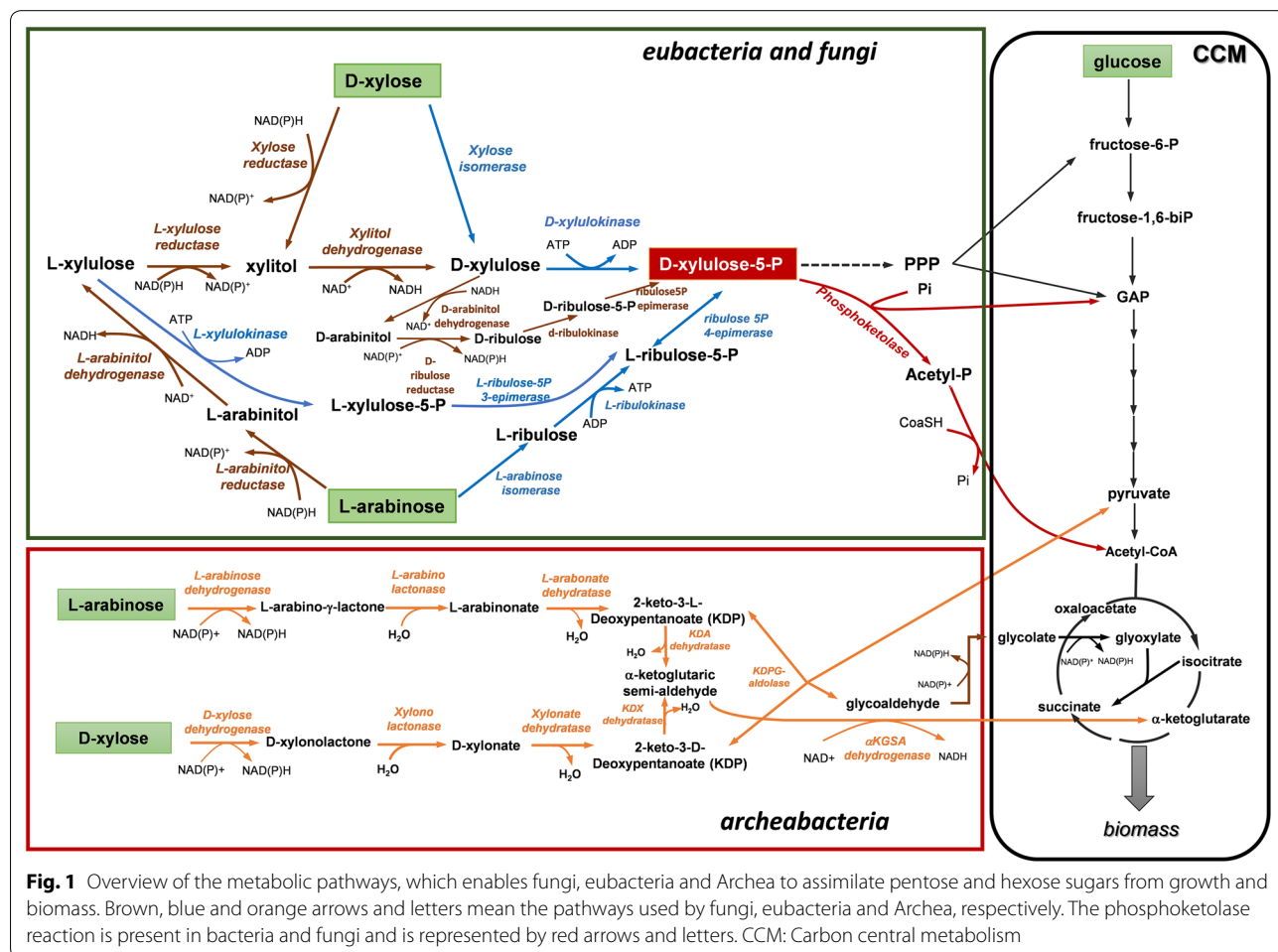
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

and fuels' purposes [4]. In addition to cellulose, which is a polymer of  $\beta$ -1, 4 linked glucosyl units, lignocellulose is composed of a hemicellulosic fraction that can reach 75% of total biomass whose major sugars are D-xylose and L-arabinose [5]. Microorganisms are naturally able to grow on pentose sugars either through the oxidative pentose phosphate pathway (PPP) for Eubacteria [6] and fungi [7] or via the pentose oxidation pathway in Archea [8–10] (Fig. 1). However, with the desire to fully convert lignocellulosic sugars into bio-based natural or non-natural products, the classical metabolic network harbored by these microorganisms is not optimally shaped for this purpose. Synthetic biology and evolutionary engineering emerge as very promising scientific and technological strategies that can meet the biotechnological objective without disrupting cellular homeostasis and keeping at minimum growth of microbial system. Applied to microbial organism, synthetic biology consists in rewiring existing pathways with novel enzymatic functions and/or to implement new non-natural-metabolic pathways in a microbial chassis with the ensuing goal of maximizing

product yield, titer and productivity [11, 12]. In this review, we will first recall the existing natural pathways that have been elaborated by microbial organisms to assimilate pentose and hexose (coined C5 and C6 sugars). We will then report on the exploitation of the non-phosphorylating pentose pathway, the reconstruction and implementation of natural and non-natural pathways into industrially tractable microorganisms, enabling efficient conversion of lignocellulosic sugars into a wide range of economically attractive bio-based chemicals. Biofuels and bioethanol are not considered in this review as these bio-based products have received a lot of attention and reviews over the last 10 years [13–17]. In addition, this review complements and extends that of [18], which was focused on metabolic engineering strategies for the production of C2–C4 diols from lignocellulosic sugars and that of [19], which described specific advancements on engineering strategies for utilization and conversion of xylose into chemicals. We will conclude about the challenges that remain to incorporate these microbial cell factories as key players in the emerging Bio-based Economy.



**Fig. 1** Overview of the metabolic pathways, which enables fungi, eubacteria and Archea to assimilate pentose and hexose sugars from growth and biomass. Brown, blue and orange arrows and letters mean the pathways used by fungi, eubacteria and Archea, respectively. The phosphoketolase reaction is present in bacteria and fungi and is represented by red arrows and letters. CCM: Carbon central metabolism

### Overview of the natural metabolic pathways for microbial assimilation of lignocellulosic sugars

Pentose and hexose sugars derived from lignocellulosic biomass are ideal carbon source for the production of bio-platform molecules by industrially tractable microorganisms such as *E. coli*, *Corynebacterium glutamicum* and yeasts. However, a major limiting factor in the use of these sugars is the difficulty to extract them from lignocellulosic material due to the extreme structural complexity of this feedstock. Though it is not the purpose here to discuss about lignocellulose features, which has been treated in many excellent reviews [20–22], it is worth to notice that deconstruction and solubilization of lignocellulose are strategic issues for commercialization of biomass-based technologies. Relevant progress in this field has been recently reviewed, notably in three directions which are actually complementary. A first one is the recent discovery of deep eutectic solvents (DESs), also termed green solvents showing high-efficient biomass deconstruction and being more compatible for future treatment with enzymes and microorganisms [23]. A second direction is focusing on engineering the composition and structure of the cell wall with the ultimate goal to create tailor-made biomass with a cell wall optimized for particular industrial or agricultural applications, without affecting plant growth and development [24, 25]. Innovation in this research field would also come from engineering native biomass biogradors, such as *Clostridium thermocellum* or *Caldicellulosiruptor* [26, 27] that are able to convert directly lignocellulose into monomers and thus could fulfill the consolidated bioprocessing concept originally set up by Lynd et al. in which biomass conversion into products can be carried out by a single fermenting microbe [28].

Prokaryotic and eukaryotic microorganisms have elaborated different pathways to assimilate pentose (D-xylose and L-arabinose) and hexose (D-glucose) sugars as depicted in Fig. 1. Table 1 provides a summary of genes and their corresponding encoded enzymes involved in pentose assimilation pathways by bacteria and fungi. When data are available, catalytic efficiencies of these enzymes on their specific substrates are also reported. In filamentous fungi and xylose-fermenting yeasts, D-xylose and L-arabinose are metabolized by an oxidoreductive pathway (Fig. 1, arrows and symbols in brown) into xylulose-5-P (Xu5P), which thereafter enters the central carbon metabolism (CCM) via the non-oxidative branch of the pentose phosphate pathway (PPP). The assimilation of these pentoses begins with their uptake that is carried out by a relatively large, yet poorly characterized sugar transporter family [29–31] (not illustrated in the picture). They are reduced into their corresponding alcohol by an

NADPH-dependent D-xylose (or L-arabinose) reductase, followed by a dehydrogenation catalyzed by an NAD<sup>+</sup>-dependent xylitol (or L-arabitol) dehydrogenase to yield D-xylulose and L-xylulose, respectively. D-Xylulose is readily phosphorylated on carbon 5 by a D-xylulokinase to D-xylulose-5-P, which fuels the CCM at the level of fructose-6-P and glyceraldehyde-3-phosphate (GAP) through transketolase/transaldolase-catalyzed reactions. Prior to its incorporation into CCM, L-xylulose derived from L-arabinose must be reduced into xylitol by an NADPH-dependent L-xylulose reductase followed by an oxidation catalyzed by an NAD<sup>+</sup>-xylitol dehydrogenase [32]. In some fungi, such as *Pichia stipitis*, D-xylulose can enter the non-oxidative pentose phosphate pathway via another route, which involves its reduction into D-arabinitol, followed by oxidation to D-ribulose and then phosphorylation to ribulose-5-P [33]. Overall, the assimilation of D-xylose and L-arabinose by filamentous fungi is characterized by a dissimilarity in the use of redox cofactors while the process is redox neutral. This dissimilarity of redox cofactors is still considered as a major problem for efficient fermentation of pentose sugars into ethanol by metabolically engineered yeasts [34].

In eubacteria (e.g., *Escherichia coli*, *Corynebacterium*, *Lactococcus*, *Streptomyces* species, etc.), the assimilation of pentose sugars also starts with their uptake mediated by active transporters [29]. Once they get into the cell, they are isomerized by a corresponding isomerase into D-xylulose or L-ribulose. D-xylulose further enters the CCM via Xu5P, by direct phosphorylation catalyzed by a D-xylulokinase. Alternatively Xu5P can be obtained from L-ribulose through a phosphorylation and epimerization catalyzed by a L-ribulokinase and a L-ribulose 5-P 4-epimerase, respectively (Fig. 1, arrows in blue). The isomerase pathway is thus making the distinction between eubacteria and fungi for pentose assimilation as the latter relies on an oxido-reductive pathway. Nevertheless, exceptions exist such as for the anaerobic fungi *Pyromyces* sp. that can metabolize D-xylose via the isomerase route [35], whereas the bacteria *Mycobacterium smegmatis* utilizes the oxido-reductive pathway for assimilation of D-arabinose [36]. Another deviation to this general rule of pentose assimilation is the phosphorylytic cleavage of Xu5P into glyceraldehyde-3P (GAP) and acetyl-Pi catalyzed by a xylulose-5-P/fructose-6-P phosphoketolase (PK) (Fig. 1, arrows in red). This PK route has been originality identified in lactic bacteria and bifidobacteria [37–40], and turns out to be widely present in filamentous fungi [41] and unicellular algae [42, 43]. In most cases, PPP and PK routes coexist in a same organism, raising the question about which is the most useful for pentose assimilation. In *Clostridium acetobutylicum*

**Table 1 Description of enzymes and corresponding genes implicated in microbial assimilation of lignocellulosic sugars**

Pathway	Enzyme name	Enzyme abbreviation	Gene name	Specific constant kcat/Km (s <sup>-1</sup> M <sup>-1</sup> ) [substrat]	Microbial origin	Reference
Pentose oxido-reductive pathway	D-Xylose reductase	XR	<i>XYL1</i> <i>xyrA</i>	na na	<i>P. stipitis</i> <i>A. niger</i>	[32]
	Xylitol dehydrogenase	XDH	<i>xdh1</i> <i>XYL2</i>	na na	<i>Filamentous fungi</i> <i>S. cerevisiae</i>	[136] [137]
	D-Xylulokinase	XK	<i>xylB</i> <i>XKS1</i> <i>xkiA</i> <i>XYL3</i>	879 10 <sup>3</sup> [xylulose] na	<i>E. coli</i> <i>S. cerevisiae</i> <i>Aspergillus niger</i> <i>Pichia stipitis</i>	[138] [139] [140] [141]
	L-Arabinose reductase	AR	<i>XYL1</i> <i>larA</i>	na na	<i>T. reesei</i> <i>Aspergillus niger</i>	[32] [142]
	L-Arabitol 4-dehydrogenase	ARDH	<i>ladA</i>	na	<i>Filamentous fungi*</i>	[33, 143]
	L-Xylulose reductase (NADPH)	IXR	<i>lrx3</i> <i>lrxA</i>	712 [L-xylulose] 47,10 <sup>3</sup> [L-xylulose]	<i>T. reesei</i> <i>Aspergillus niger</i>	[144] [145]
	L-Xylulose reductase (NADH)		<i>ALX1</i>	na	<i>Ambrosiozyma monospora</i>	[146]
	D-Arabinitol 2-dehydrogenase/ribulose reductase	RIR	<i>ARD</i>	na	<i>Pichia stipitis</i> <i>Candida tropicalis</i>	[33] [147]
	D-Ribulokinase	RIK	<i>YDR109c</i>	80 10 <sup>3</sup> [D-ribulose]	<i>S. cerevisiae</i> ,	[148]
	Isomerase pathway	Xylose isomerase	XI	<i>xylA</i>	na	<i>E. coli</i>
D-Xylulokinase		XK	<i>xylB</i>	879 10 <sup>3</sup> [xylulose]	<i>E. coli</i>	[138]
L-Arabinose isomerasz		AI	<i>araA</i>	na	<i>E. coli</i>	[150]
L-Ribulokinase		RKI	<i>araB</i>	na	<i>E. coli</i>	[151, 152]
L-Xylulokinase		XKI	<i>xylK</i>	na	<i>E. coli</i>	[108]
L-Ribulose-5-P 3-epimerase		UlaE	<i>ulaE</i>	na	<i>E. coli</i>	[153]
L-Ribulose 5-P 4-epimerase		RPE	<i>araD</i>	na	<i>E. coli</i>	[154]
Non-phosphorylative pathway	D-Xylose dehydrogenase	XylDH	<i>xylB</i> <i>xylB</i> <i>HVO_B0028</i>	10 <sup>4</sup> [D-Xylose] 157,10 <sup>3</sup> [xylose]	<i>Caulobacter crescenti</i> <i>S. solfataricus</i> <i>H. vulcanii</i>	[46, 48, 155, 156]
	D-Xylonate lactonase	XAL	<i>xylC</i>	na na	<i>C. crescenti</i> <i>H. vulcanii</i>	[48, 81]
	D-Xylonate dehydratase	XAD	<i>xylD</i> <i>HVO B0038A</i> <i>yagE/yjhG</i>	40 10 <sup>3</sup> [D-xylonate] 4 10 <sup>3</sup> (D-Xylonate)	<i>C. crescenti</i> <i>H. vulcanii</i> <i>Escherichia coli</i>	[48, 70, 155]
	L-Arabinose 1-dehydrogenase	AraDH	<i>araA</i>	32.3 10 <sup>3</sup> [L-arabinose]	<i>B. multivorans</i>	[81]
	L-Arabinolactonase	AraL	<i>araL</i>	na	<i>B. multivorans</i> , <i>S. solfataricus</i>	[81]
	L-Arabinonate dehydratase	AraD	<i>araC</i>	83 [L-arabinonate]	<i>B. multivorans</i> <i>S. solfataricus</i>	[81] [156]
	2-Keto-3-deoxy-D-xylo-nate dehydratase	KdxD	<i>xylX</i>	260 [2-keto-3-deoxy-D-pentanoate] 530 [2-keto-3-deoxy-D-pentanoate]	<i>C. crescenti</i> <i>B. xenovorans</i>	[80, 81]
	2-Keto-3-deoxy-L-arabinonate dehydratase	KdaD	<i>araD</i>	23 [2-keto-3-deoxy-L-arabinonate]	<i>B. xenovorans</i>	[81]
	2-Ketoglutarate semialdehyde dehydrogenase	KGSADH	<i>xylA</i> <i>ycbD</i> <i>HVO_B0039</i>	4.6 10 <sup>5</sup> [α-ketoglutaric semialdehyde]	<i>A. brasilense</i> <i>C. crescenti</i> <i>S. solfataricus</i> <i>H. valcanii</i>	[48, 155–157]
	2-Keto-3-deoxy-D/L-pentanoate aldolase	KDPA	<i>kdxA</i> <i>yagE/yjhH</i>	450 [2-keto-3-deoxy-D-pentanoate] 7000 [pyruvate]	<i>S. solfataricus</i> <i>Escherichia coli</i>	[47, 62]

na: Not accessible

for which this question has been recently addressed it, was found that the PK route is more useful than PPP and that L-arabinose is assimilated before D-xylose [40].

In a third pathway, which is the privilege of archaeobacteria (Fig. 1, arrows and symbols in orange), D-xylose and L-arabinose (as well as D-ribose or D-arabinose) are catabolized by a so-called non-phosphorylating pentose (NP) pathway to either  $\alpha$ -2-ketoglutarate ( $\alpha$ KG) or pyruvate and glycolaldehyde. The conversion of pentose into  $\alpha$ KG takes place by the Weimberg pathway [8]. This pathway proceeds in five consecutive enzymatic steps. The first step is the oxidation of the pentose (D-xylose or L-arabinose) into pentono-(D/L)-lactone (D-xylonolactone or L-arabinolactone) by a specific pentose dehydrogenase (i.e., D-xylose dehydrogenase or L-arabinose dehydrogenase). These intermediates are then hydrated by a specific lactonase into their corresponding pentonic acid, followed by a dehydration into a 2-keto-3-deoxy L or D-pentanoate (KDP) catalyzed by specific D-xylonate or L-arabinonate dehydratase. A subsequent dehydration catalyzed by a 2-keto-deoxy L- or D-pentanoate dehydratase (KDPH) gives rise to  $\alpha$ -ketoglutaric semialdehyde also termed 2,5-dioxopentanoate (DOP). The last reaction of this pathway is catalyzed by an  $\alpha$ -ketoglutaric semialdehyde dehydrogenase that yields  $\alpha$ -ketoglutarate, which fuels CCM at the level of the tricarboxylic acid cycle (TCA) [9, 44]. An alternative route to this catabolic pathway, which has been discovered by Dahms [45] and is therefore termed the 'Dahms pathway', is the aldolytic cleavage of KDP into pyruvate and glycolaldehyde by a KDP aldolase. While pyruvate directly fuels the CCM, glycolaldehyde is incorporated into CCM at the level of glyoxylate, which requires two reactions generating two reduced equivalents. This was demonstrated in *Sulfolobus* sp., engaging a highly oxygen-sensitive glycolaldehyde oxidoreductase and a soluble NADH-dependent glyoxylate reductase [46]. Although the NP pathway has been known for over 50 years, there are still many issues to be clarified regarding the promiscuity of enzymes in this pathway, the dependence to redox cofactors  $\text{NAD}^+$  or  $\text{NADP}^+$  as well as the co-existence of the Weimberg and Dahms route in a same microorganism. In this respect, Nunn et al. [46] showed that the hyperthermophilic archaea *Sulfolobus solfataricus* and *acidocaldarius* employ the same enzymes for the oxidation of D-xylose and L-arabinose to generate 2-keto-3-deoxy-D-xylonate (KDX) and 2-keto-3-deoxy-L-arabinonate (kDa), which thereafter are cleaved by a same KDP aldolase into pyruvate and glycolaldehyde [47]. In these archaeobacteria, the Weimberg and Dahms route coexists at a ratio of about 50/50 [46], whereas in other bacteria such as *Haloferax volcanii*, D-xylose is mainly metabolized by the Weimberg pathway

and the xylose dehydrogenase has almost no activity on other C5 sugars [48].

### **The non-phosphorylating pentose pathway combined with the central carbon metabolism for conversion of lignocellulosic sugars into bio-based chemicals**

#### **Main reason to refactor carbon metabolism for biotechnological objectives**

As stated above, many microorganisms are able to assimilate pentose and hexose sugars. However, this capability is not economically viable with the biotech industry sector aiming at producing bio-based chemicals from these lignocellulosic sugars, due to several constraints as follows. First, pentose assimilation is often repressed in the presence of glucose [49, 50]. This can strongly hamper the fermentation process, and different strategies have been elaborated to alleviate this glucose repression effect. These include engineering hexose transporter to release xylose uptake from glucose inhibition in yeast [51], abolition of the cAMP-CRP-dependent glucose repression in bacteria [52], adapted laboratory evolution [53–55], use of a microconsortium of different *E. coli* strains, each engineered to preferentially catabolize a different hexose—glucose, galactose, or mannose [56], as well as fermentation strategy such as continuous culture [57]. Second, the common metabolite intermediate from C5 and C6 metabolism is a C3 compound, such as GAP or PEP, which can be problematic if the desired chemical is made with C4 or C5 carbon skeleton. Yet importantly, traditional pathways comprise lengthy enzymatic reactions that are subject to transcriptional and posttranscriptional regulation, whose complexity hampers performance in terms of yield and productivity.

These limitations could be also overcome by a rewiring of the existing carbon metabolism pathways to reach the targets in less reaction steps, or by implementing non-natural or artificial pathways that are orthogonal to the classical pathways. The implementation of the nonphosphorylating pentose (NP) pathway present in archaeobacteria [44] in industrially relevant microorganisms such as *E. coli*, yeast or *C. glutamicum* [58–60] is an interesting solution, as it can equip the microbial cells with new functionalities to co-assimilate C5 with C6 sugars. The key intermediate of this pathway is 2-keto-3-deoxy-D-xylonate (KDX) or 2-keto-3-L-arabinonate (KDA) obtained from D-xylose or L-arabinose through the three reaction steps involving a dehydrogenase, a lactonase and a dehydratase (Figs. 2 and 3). It should be noted that lactone intermediate can be spontaneously hydrolyzed into D-xylonate and L-arabinonate. This may explain that the genes *xylC* or *araB* encoding these reactions are often not considered in the pathway construction. The genes

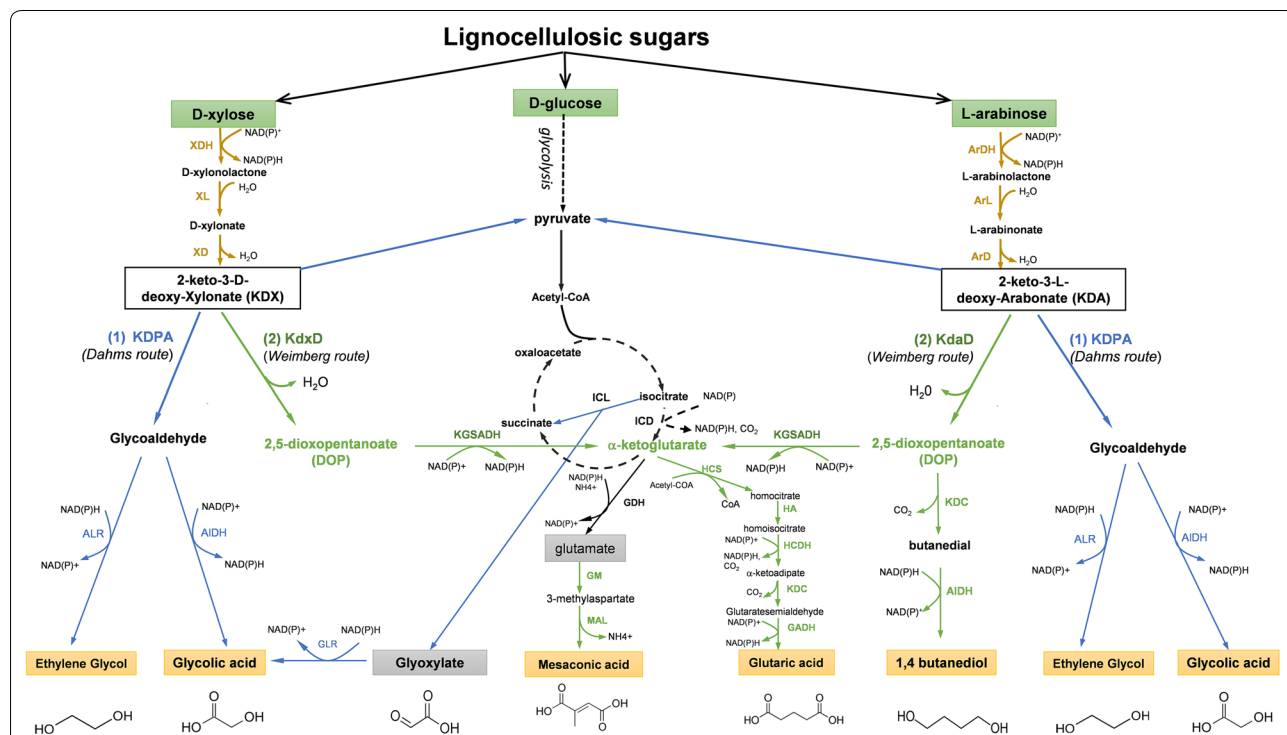
encoding these enzymes have been isolated from various archaeobacteria such as *Caulobacter crescentus* or *Burkholderia multivorans* and heterologously expressed in *E. coli*, yeast or *C. glutamicum* (see Additional file 1: Table S1 for a description of these enzymes, their catalytic efficiency and corresponding genes used in heterologous expression). Also, *E. coli* possesses two genes *yagF* and *yjhG* encoding a dehydratase that catalyzes the dehydration of D-xylonate into KDX [58, 61]. As indicated in Figs. 2 and 3, the produced KDX or KDA can be taken in charge by three potential downstream pathways, to yield a wide range of bio-based products. For simplification, these three pathways have been termed NP\_Dahms, NP\_Weimberg and NP\_KDC. The orthogonality of these pathways enable them to function alone or in connection with the CCM, with the latter being able to provide either growth support or additional product yield, as it will be discussed in the next paragraphs.

The bio-based chemicals that are considered at length in this review as they were reported to be produced through these NP pathways are listed in Table 2. These products are nowadays exclusively manufactured by chemical means but their bioproduction from

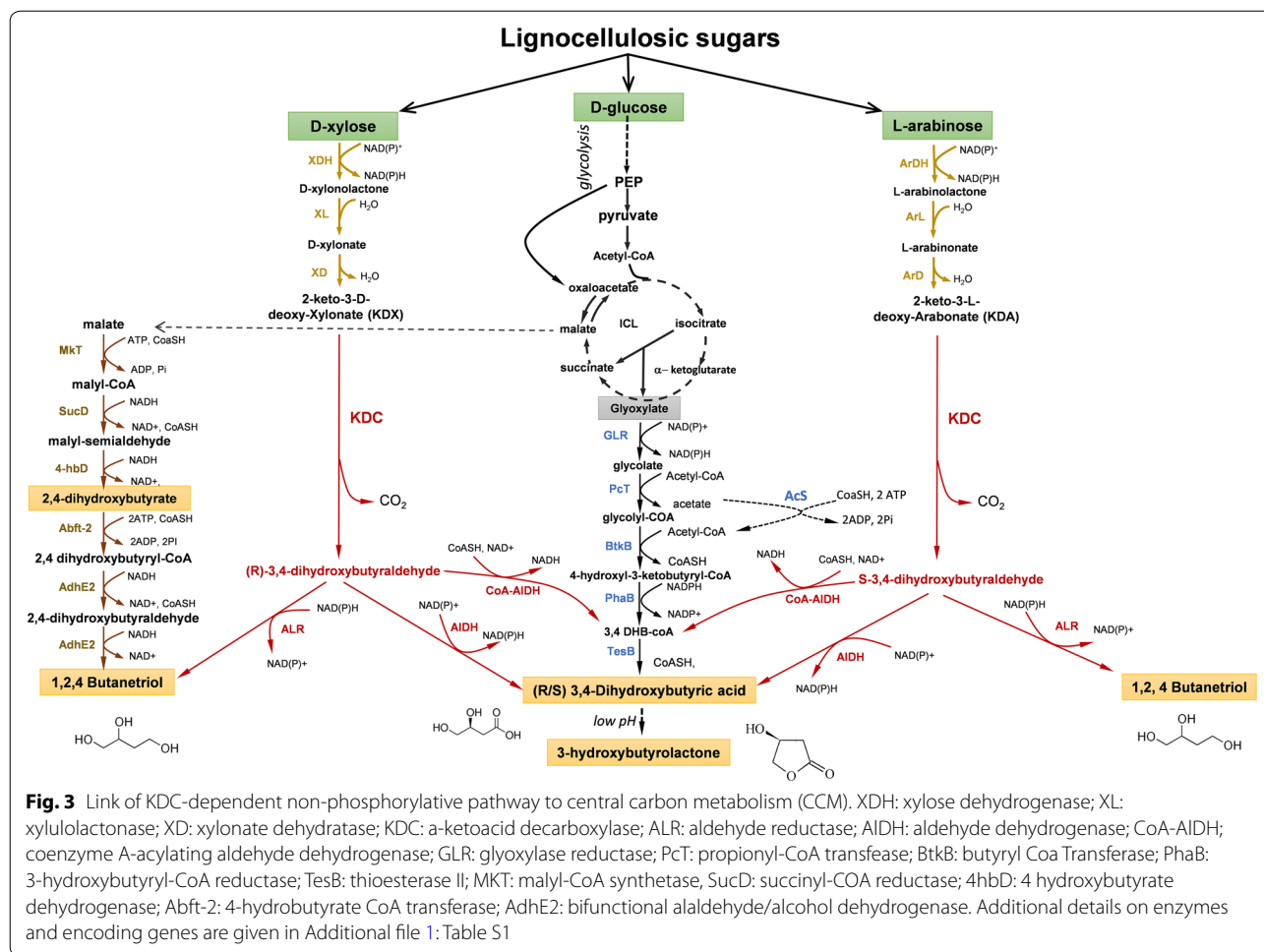
lignocellulose is both scientifically sound and industrially feasible. These chemicals furthermore cover a large spectrum of application in chemical, pharmaceutical and food industries, with a global market size of USD 8.8 billions in 2017 and an expected compound annual growth rate (CAGR) of 12.6% from 2018 to 2025 (see <https://www.grandviewresearch.com/industry-analysis/bio-based-platform-chemicals-market>).

### The NP\_Dahms metabolic route for ethylene glycol and glycolic acid

The NP\_Dahms pathway starts by the aldolytic cleavage of KDX or KDA into glycolaldehyde and pyruvate catalyzed by a 2-keto-3-deoxy-pentanoate aldolase (KDPA). While pyruvate can be used as a carbon source for growth, glycolaldehyde is either oxidized into glycolic acid (GA) or reduced into ethylene glycol (EG) (Fig. 2). In *E. coli*, two genes *yjhH* and *yagE* encode a functional KDPA [62], as evidenced by the finding that a double *yjhH yagE* mutant was unable to grow on D-xylonate [63]. In addition, KDPA encoded by *yagE* is likely more active than that encoded by *yjhH* since levels of GA were higher in engineered strains with *yagE* [63]. The fate of



**Fig. 2** Link between the Dahms (pathway [1]) or Weimberg-dependent (Pathway [2]) non-phosphorylative pathway to central carbon metabolism (CCM). XDH: xylose dehydrogenase; XL: xylulolactonase; XD: xylonate dehydratase; KDPA: 2-keto-3-deoxy-pentanoate aldolase; KdxD: 2-keto-3-deoxy-D-xylonate dehydratase; KdaD: 2-keto-3-deoxy-L-arabinonate dehydratase; KGSADH: a ketoglutarate semialdehyde dehydrogenase; ALR: aldehyde reductase; AIDH: aldehyde dehydrogenase; GLR: glyoxylase reductase; GDH: glutamate dehydrogenase; GM: glutamate mutase; MAL: methylaspartate lyase; HCS: homocitrate synthase; HA: homocitrate aconitase; HCDH: homocitrate dehydrogenase; KDC: a-ketoacid decarboxylase; GADH: glutaric acid dehydrogenase. Further details of the enzymes and their encoding genes are given in Additional file 1: Table S1



glycolaldehyde to either EG or GA will depend on the expression levels of genes encoding aldehyde reductase or aldehyde dehydrogenase. Oxidation of glycolaldehyde to GA is mainly ensured by a glycolaldehyde dehydrogenase encoded by *aldA* [64], the overexpression of which promoted increased GA production [63, 65]. However, not a single enzyme among the various aldehyde reductases that possesses *E. coli* [66] is devoted to the reduction of glycolaldehyde to EG. While Chae et al. [67] showed that the overexpression of *yqhD* encoding a broad-substrate NADPH-aldehyde reductase [68] was the most effective in the production yield of EG from xylose, Cabulong et al. [69] reported a better performance with *yjgB* also termed *ahr* encoding another broad-spectrum NADPH-aldehyde reductase. The discrepancy between these two reports is unclear. At variance to these studies, Wang et al. [70] showed that *fucO*, which encodes an NADH-aldehyde reductase, was the most efficient in the production of EG. These authors explained their results by considering a higher NADH availability in their engineered strain due to the deletion of *aldA* and *arcA*, a gene encoding a redox

regulator under microaerobic condition and whose deletion results in derepression of oxygen-repressed genes [71]. Altogether, the NP\_Dahms route to GA or EG comprises only five reaction steps from C5 sugars (Fig. 2), and only the last step of the pathway makes a distinction between the engineered *E. coli* strains for EG and for GA production.

A major problem observed upon implementation of this metabolic route in *E. coli* was a toxic accumulation of xylonic acid that strongly impairs growth. This also clearly identifies the xylonic dehydratase (XD) as rate-limiting enzyme of this pathway. Cabulong et al. [69] partially circumvented this problem by a combination of a weak expression of *xylB* encoding XD together with an overexpression of *yjhG* encoding the endogenous XD. Reduction of *xylB* expression was also carried out by Chae et al. [67] using a synthetic small regulatory RNA inhibiting expression of this gene. Altogether, this five-step reaction of the NP\_Dahms route to EG is redox neutral and can reach 50% of the yield efficiency ( $Y^E$ ) (Table 3). This  $Y^E$  of the pathway is calculated



**Table 2 Relevant commodity and fine chemicals made from petrochemistry, which are foreseen to be produced by economically viable fermentation process economically from lignocellulosic sugars**

Chemicals name	Chemical formula	Molecular mass	Degree of reduction	Application field	Market size (CAGR*) (US dollars)	Reference
Glycolic acid	$C_2H_4O_3$	76	6	Dyeing and tanning agent in textile industry, flavor and preservative in food industry, skin care agent in cosmetics	280 million by 2019 CAGR 11.3%	<a href="https://www.marketsandmarkets.com/Market-Reports/glycolic-polyglycolic-acid-market-1090.html">https://www.marketsandmarkets.com/Market-Reports/glycolic-polyglycolic-acid-market-1090.html</a>
Mono ethylene glycol	$C_2H_6O_2$	62	10	Antifreeze agent, precursor of polyethylene terephthalate (PET) for plastic bottle, use in paints and resins in chemical industry,	26 billion by 2018 CAGR 5%	<a href="https://www.marketresearchfuture.com/reports/mono-ethylene-glycol-market-2709">https://www.marketresearchfuture.com/reports/mono-ethylene-glycol-market-2709</a>
Glyoxylic acid	$C_2H_2O_3$	74	6	Cosmetic, personal care and dermatology, household adhesive, additive for ink and painting, component in emulsion polymers	76million in 2016 CAGR 5.27%	<a href="https://www.marketresearchfuture.com/reports/glyoxylic-acid-market-3158">https://www.marketresearchfuture.com/reports/glyoxylic-acid-market-3158</a>
1,4 Butanediol	$C_4H_{10}O_2$	90	22	Fibers, plastics, medicines, cosmfibretics, artificial leather, pesticides, plasticizers, hardener, solvent, rust remover, etc.	7.75 Billion by 2017 CAGR 6.6%	<a href="https://www.marketsandmarkets.com/Market-Reports/1-4-butanediol-market-685.html">https://www.marketsandmarkets.com/Market-Reports/1-4-butanediol-market-685.html</a>
1,2,4 Butanetriol	$C_4H_{10}O_3$	106	20	Building block for energetic plasticizer, polymers, for electronics materials	8.9 billion by 2025 CAGR 8.2%	<a href="https://www.grandviewresearch.com/press-release/global-1-4-butanediol-market">https://www.grandviewresearch.com/press-release/global-1-4-butanediol-market</a>
3,4 dihydroxybutyric acid hydroxybutyrolactone	$C_4H_6O_4$	120	16	Intermediates for high-value pharmaceutical compounds such as Crestor, Lipitor, Zyxov, Zetia	54 million (2003) CAGR 5.5%	<a href="http://www.mynewsdesk.com/us/tag/3-hydroxybutyrolactone-market">http://www.mynewsdesk.com/us/tag/3-hydroxybutyrolactone-market</a>
Glutaric acid	$C_5H_8O_4$	132	20	Building block for polyesters and polyamides like nylon, ingredient in cleaning products, painting and coating formulation	Data not found	NA
Mesaconic acid	$C_5H_8O_4$	130	18	Precursor for pentadiene (isoprene) by thermocatalysis	Data not found	NA

\*CAGR: compound annual growth rate; NA: not available

according to [72] as the ratio between the stoichiometric efficiency (or pathway yield)  $Y^P$  and the thermodynamic yield  $Y^{th}$ , with the latter obtained as the ratio between the degree of reduction of the substrate to that of the final product. The fact that only 1 mol of EG is produced per mole of C5 sugar is due to the metabolic inability to recycle pyruvate into EG. This may also explain that this chemical cannot be produced from glucose at least using conventional pathways. Nonetheless, fermentation process has been established with *E. coli* equipped with the NP\_Dahms pathway to reach titer of 108 g/L and productivity in the range of 2.25 g/L/h from D-xylose [67] (summarized in Additional file 1: Table S2).

The yeast *S. cerevisiae* was also used as host to express the NP-Dahms pathway for EG production. To this end, the producer strain was constructed by genome integration of *C. crescentus xylB* under the strong *PGK* promoter while the other genes of the Dahms pathway,

namely *xylD*, *yagE* and *aldA*, were expressed on a high-copy plasmid under the strong *TDH* promoter [73]. In addition, the strain was also deleted for competing pathways to avoid loss of xylose into xylitol. Upon incubation with D-xylose, the engineered strain only produced a tiny amount of EG, while exhibiting huge accumulation of D-xylonate as well as of 3-deoxypentanoic acid (3-DPA) and 3,4-dihydroxybutyric acid (3,4-3,4-DHBA). These two metabolites actually revealed the presence of endogenous aldo-keto reductase and 2-ketoacid decarboxylase that compete with KDPA for KDX, respectively (see below for further details). Their accumulation was mainly ascribed to the inefficient catalytic activity of xylonate dehydratase due to low availability of iron in yeast cytosol. Indeed, this enzyme belongs to LlvD/EDD protein family that requires iron for activity [74]. Support to this hypothesis was reported by Salusjarvi et al. [73] who found that the deletion of *FRA2* encoding a cytosolic protein implicated in the regulation of iron, resulted in

**Table 3 Yield and energy cost of bio-based chemicals according to production pathways expressed in *E. coli***

Bio-based products	Production pathway	Substrat	Nbr reaction step	NAD(P)H yield (mole/mole)	ATP yield (mole/mole)	Pathway yield ( $Y^P$ ) <sup>§</sup>	Yield efficiency (%) <sup>*</sup>
Ethylene glycol (EG)	NP-Dahms	Xylose	5	0	0	1	50
	Xylulose-1-P	Xylose	4	-1	-1	0.87	50
	Ribulose-1-P	Xylose	5	-1	-1	0.87	50
	Serine pathway	Glucose	13	+2	0	2	83.3
	X1P (R1P) + serine pathway	Xylose	7	+1	0	2	100
	X1P (R1P) + serine pathway	Xylose + glucose	11	+2	0	4	91
Glycolic acid (GA)	CCM- (GS) <sup>§</sup>	Glucose	18	+6	+2	2	66
	NP-Dahms	Xylose	5	+2	0	1	40
	Xylulose-1-P	Xylose	4	+1	-1	0.96	40
	Ribulose-1-P	Xylose	5	+1	-1	0.96	40
	NP-Dahms + GS	Xylose	10	+2 (+1QH2)	0	2	80
	X1P (R1P) + GS	Xylose	9	+3 (+1QH2)	+1	2	80
	X1P (R1P) + GS	Xylose + glucose	15	+2 + (3 QH2)	+3	4	73
1,4 Butanediol (BDO)	CCM <sup>§</sup>	Glucose	21	0	0	1	92
	NP-Weimberg	Xylose	6	0	0	1	110
1,2,4 Butanetriol (BTO)	CCM <sup>§</sup>	Glucose	17	0	-1	1	83
	NP-KDC	Xylose	5	0	0	1	100
3,4 dihydroxybutyric acid (3,4-DHBA)	CCM	Glucose	22	+3	0	1	66
	NP-KDC	Xylose	5	+2	0	1	80
Mesaconic acid (MSA)	CCM <sup>§</sup>	Glucose	18	+3	+1	1	75
	NP-Weimberg	Xylose	8	+1	0	1	91
Glutaric acid (GTA)	CCM <sup>§</sup>	Glucose	17	+5.66	+1.33	1	
	NP-Weimberg + XI	Xylose	14	+4	0	0.66	62

XI means xylose isomerase pathway

<sup>§</sup>  $Y^P$  is the stoichiometry pathway efficiency strictly calculated from the pathway without balancing ATP and cofactors

<sup>\*</sup>Yield efficiency is the ratio between the thermodynamic or energy yield and the pathway yield ( $Y^P$ ) as defined by [72]. The energy yield of these chemicals as produced from pentose or glucose can be found in Additional file 1: Table S4

<sup>§</sup> CCM-GS means Carbon central metabolism—glyoxylate shunt

a 13-fold increase of the XD activity, which was accompanied by a still modest increase of EG (14 mg/L) from D-xylose.

As already described above, GA is the other relevant bio-based product that can be obtained from pentose through the NP\_Dahms. At first glance, it only requires the rerouting of glycolaldehyde by the overexpression of *aldA* encoding an NAD<sup>+</sup>-dependent glycolaldehyde reductase [64]. However, this pathway generates 2 mol of reduced cofactors (Table 3) that must be regenerated through the respiratory chain coupled with ATP synthesis. This energy can be spent for the pentose uptake and for cellular maintenance. Contrary to EG, GA can also be naturally produced in *E. coli* through the glyoxylate shunt (GS). Therefore, a linkage between GS and the NP\_Dahms pathway at the level of pyruvate might be envisioned, resulting in the production of 2 mol GA per mole of pentose (Table 3). This linkage was attempted in two recent papers. In a first report [63], the connection of the two pathways was created from two different plasmids, termed module 1 and module 2, in an *E. coli* strain deleted for competing pathways, namely xylose isomerase (*xylA*), xylulose-5-P kinase (*xylB*) and glyoxylate oxidase (*glcDEF*). Module 1 was a low-copy plasmid carrying genes encoding the NP-Dahms pathway, i.e., *xdh-yagF-yagE-aldA* as well as *grhA* encoding a glyoxylate reductase (GLR). Module 2 was a high-copy plasmid carrying genes favoring GS, namely *aceA* encoding isocitrate lyase and *aceK*, which codes for the bifunctional isocitrate dehydrogenase kinase/phosphatase to reduce isocitrate dehydrogenase activity. A production of GA at 0.46 g/g D-xylose was obtained with this engineered *E. coli* strain. Same yield on D-xylose was obtained by Liu et al. [65] following a similar strategy except that *ackA* encoding acetate kinase was deleted instead of overexpressing *aceK*.

Glycolic acid can be also produced from glucose, by natural pathways that comprise no less than 13 steps, with a maximal stoichiometric yield of 2 (Table 3). Through conventional metabolic engineering, the carbon flux from glucose to glycolic acid can be favored, which merely implicates activation of the glyoxylate shunt and inactivation of competing pathways at the level of pyruvate and isocitrate. Hence, Deng et al. [75] engineered a GA producer *E. coli* strain from glucose that initially grew poorly on glucose, likely due to a high expression of *aceK* causing a blockage of TCA at the isocitrate dehydrogenase level. Following adaptive evolution by repeating transfer to mineral glucose (M9) medium, a producer strain with improved growth rate was isolated and shown to produce 54 g/L GA at 63% of the pathway yield (Additional file 1: Table S2). In a more recent report, the same authors reported a

production of 65 g/L GA at a yield approaching 92% of the maximal pathway yield with an improved producer strain, which was obtained according to the following subtle genomic modifications. Expression levels of *aceA*, *ghrA* and *aceK* were optimized by placing these genes on a low-copy plasmid under the strong *trc* promoter. Flow of acetyl-COA into the initial TCA reaction was favored by overexpression of the citrate synthase encoding *gltA*. Competing pathway for pyruvate and GA consumption were deleted such as *ldhA* encoding the lactate dehydrogenase and *aldA* encoding the glycolaldehyde dehydrogenase. Finally, the fermentation process was optimized by adjusting the C/N ratio at 15:1 (summarized in Additional file 1: Table S2).

Production of GA from glucose by *Corynebacterium glutamicum* and yeasts has been also attempted [76, 77]. However, use of these microorganisms for this production presented several difficulties. For *C. glutamicum*, the maximal pathway yield ( $Y^P$ ) is 1.7 mol glycolic acid/mol glucose instead of 2 since not all NADPH is fulfilled by glycolysis. On the other hand, *C. glutamicum* does not possess a glyoxylate reductase. On top of that, glyoxylate shunt is inactive under glucose growth condition. Therefore, GA was produced using an engineered *C. glutamicum* deleted for *aceB* encoding malate synthase, attenuated for expression of *icd* encoding isocitrate dehydrogenase and overexpressing the *E. coli* *grhA* encoding glyoxylate reductase. Due to this attenuation of *icd*, which impairs TCA cycle, acetate had to be added in culture medium. A production of 5.3 g/L GA was obtained according to a 1:1 C-mole ratio between glucose and acetate [76]. The group of Pentillä et al. [77] reported the engineering of the glyoxylate shunt pathway in the yeast *S. cerevisiae* and *K. lactis* for GA production. The metabolic engineering in both yeasts requested to remove genes encoding malate synthase and cytosolic isocitrate dehydrogenase and to express the *A. thaliana* *GLYR1* encoding the glyoxylate reductase. D-Xylose was used as the carbon source to avoid glucose repression of the glyoxylate shunt [78]. The titer obtained by these engineered yeast was only 1 g/L with *S. cerevisiae* while it was 15 times higher with the engineered *K. lactis*. The observed difference was explained by the fact that *K. lactis* exhibits neither glucose repression of respiration nor a Crabtree effect, which is known to divert carbon and reducing power into ethanol [79].

#### The NP\_Weimberg metabolic route to yield BDO and other chemicals

In a second route, which we called NP\_Weimberg pathway, the 2-keto-3-deoxy-(D/L)-pentanoate is dehydrated by a specific 2-keto-3-deoxy-D-xylonate dehydratase (KdxD) or a 2-keto-3-deoxy-L-arabinonate dehydratase

(KdaD) to yield 2,5-dioxopentanoate (DOP). DOP is a strategic platform because it can be oxidized into  $\alpha$ -ketoglutarate, from which a panel of other bio-based chemicals can be produced (Fig. 2). To find out an efficient NP\_Weimberg pathway that could generate these chemicals, Zhang et al. [80, 81] devised an original screening strategy that employed an *E. coli* mutant defective in isocitrate dehydrogenase ( $\Delta icd$  mutant) unable to produce  $\alpha$ -ketoglutarate from glucose by TCA. Therefore, growth rescue of this mutant on D-xylose or L-arabinose exclusively relied on the expression of genes clusters isolated from various Archea that harbor the NP\_Weimberg pathway. An interesting result that came out from this screening was to find that the dehydratase enzyme encoded by *xylX* (for D-xylose degradation pathway) or *araE* (for L-arabinose degradation pathway) was clearly the bottleneck enzyme in this pathway.

As indicated in Table 3, mesaconic acid (MSA) can be produced from C5 sugars at a molar yield of 1 mol per mole of pentose. To achieve this production, Bai et al. [80] equipped *E. coli* strain with three compatible plasmids. Plasmid 1 (medium-copy) carried NP\_Weimberg pathway genes to produce DOP from D-xylose, genes encoding glutamate mutase (GlmE) and methyl aspartate lyase (Mal) to generate mesaconate from DOP were cloned in a plasmid 2 (high-copy), and plasmid 3 (low-copy) carried four genes necessary for the regeneration and reactivation enzymatic system for coenzyme 12 that is essential for GlmE activity. This strain was moreover deleted for *yagE* and *yjhH* encoding KDPA to eliminate Dahms route and for *sucA* to avoid loss of  $\alpha$ -ketoglutarate downstream in the TCA cycle. Altogether, the engineered strain was able to produce 14.7 g/L of MSA after 2 days at a yield of 0.74 g/g D-xylose, which corresponds to 85% of the pathway yield  $Y^P$  (Table 3). Although this pathway is ATP neutral, the extra NAD(P)H generated can be reoxidized through respiratory chain coupled with ATP synthesis to yield energy needed for sugar uptake and maintenance. Since the production of MSA by NP\_Weimberg route is restricted to C5 sugars, Wang et al. [82] recently developed an MSA producer strain from glucose by rerouting carbon flux to  $\alpha$ -KG via overexpression of *ppc*, *gltA* and *icd* encoding PEP carboxylase, citrate synthase and isocitrate dehydrogenase, respectively. In addition, PEP was made more available for PPC-dependent anaplerotic reactions by deletion of *pstG* and overexpression of *galP* and *ppsA* encoding sugar importer and PEP synthase. This engineered strain was able to produce 23.1 g/L MSA at a yield of 0.46 g/g glucose, which correspond to 64% of the pathway yield (Table 3 and Additional file 1: Table S2). The construction of an *E. coli* strain that harbors these two engineered pathways has not been

reported yet, which would enable an MSA production from pentose and glucose at the same time.

Contrary to MSA, glutaric acid cannot be produced from pentose sugars by the NP-Weimberg pathway alone because this route cannot provide one of its key intermediates, i.e., acetyl-CoA. Indeed, de novo biosynthesis of glutaric acid from  $\alpha$ -KG requires five steps (Fig. 2) and engages an acetyl-CoA at the first step catalyzed by homocitrate synthase. This first step as well as the two following ones have been implemented in *E. coli* from the leucine biosynthesis genes of the yeast *S. cerevisiae* [83], which enables  $\alpha$  +1 carbon chain extension of  $\alpha$ -ketoglutarate to yield  $\alpha$ -adipate. This intermediate is further decarboxylated by a promiscuous  $\alpha$ -keto acid decarboxylase encoded by *L. lactis kivD* to glutarate semi-aldehyde, and then oxidized into glutarate by a succinate semi-aldehyde dehydrogenase encoded by *gadB* from *P. putida* [84]. Since NP-Weimberg alone cannot produce glutaric acid, Wang et al. [85] conceived a strategy that assessed the synergistic effect of installing either one of combination of the three xylose pathways, namely xylose isomerase (XI), NP\_Dahms and NP\_Weimberg pathways in the production yield and titer of this organic acid. While the XI pathway has the potential to produce 0.56 mol glutaric acid per mol D-xylose, the NP\_Weimberg pathway alone was unproductive because it cannot generate acetyl-CoA, whereas it can be provided by the NP\_Dahms pathway via pyruvate (Fig. 2). However, the drawback of this branching pathway is an equimolar production of a C-2 compound glycolaldehyde that is converted into GA. Consequently, the maximal yield obtained by combining NP\_Dahms and Weimberg is 0.5 mol of glutaric acid and 1 mol GA per mole of xylose. Hence, to increase the yield (thus to reduce loss of carbon as CO<sub>2</sub>), a combination of XI with NP\_Weimberg should be the best since the calculated pathway yield of this arrangement is 0.625 mol glutaric acid per mole xylose. Wang et al. [85] tested this hypothesis by engineering *E. coli* with either one or a combination of two or of the three pathways. They found that the combination of NP\_Weimberg and XI yielded the best producer strain, as expected from the model. Even though the titer was only 602 mg/L, it was 5 times and 2 times higher than using XI alone or NP\_Weimberg + Dahms, respectively. Additional improvements were attempted by restricting XI pathway to the production of acetyl-CoA while NP\_Weimberg was dedicated to production of  $\alpha$ -ketoglutarate by deleting *gltA* encoding citrate synthase. Unexpectedly, these genetic interventions did not improve yield and titer, probably because they affected cell growth. In addition, a likely reason for low yield and titer of glutaric acid may be found in the promiscuity activity of the  $\alpha$ -ketoacid decarboxylase encoded by

*Lactococcus lactis kivD* as this enzyme can readily decarboxylate upstream intermediates of the pathway, namely 2-keto-3-deoxy-D-xylonate and DOP (see below). Thus, in addition to the problem of providing enough acetyl-CoA in the pathway [85], provision of  $\alpha$ -ketoglutarate by the NP\_Weimberg route could be another limitation that explains the low production yield of glutaric acid from D-xylose. Pathways engineering for de novo biosynthesis of glutaric acid from glucose via  $\alpha$ -ketoglutarate by *E. coli* have been also reported [83]. A modest production of 0.3 g/L glutaric acid from 20 g/L glucose in 48 h was obtained, which corresponded to 10% of the pathway yield. Since this yield is 3.5-fold lower than that obtained from xylose using XI+NP\_Weimberg, this argues that a potentially viable process for glutaric acid should request co-feeding strategy of C5 and C6 sugars combining CCM with NP-Weimberg pathways. Nevertheless, these engineered pathways are still far less efficient than the glutaric acid production build on L-lysine overproducing *C. glutamicum* strain. This lysine-based pathway offers a theoretical yield of 0.75 mol/mol glucose, which is higher than the  $\alpha$ -ketoglutarate pathway (0.66 mol/mol glucose, Table 3). From lysine, only 4 additional enzymatic steps are needed which are a conversion of lysine to 5-aminovalerate by a lysine monooxygenase (DavB) and  $\delta$ -aminovaleramidase (DavA), a transaminase to produce glutarate semialdehyde (GabT) and finally a glutarate semi-aldehyde dehydrogenase (gadB) to yield glutarate. While genes encoding these enzymes were overexpressed, high yield (75% of the maximal) and titer of glutaric acid (90 g/L, see Additional file 1: Table S2) in a fed-batch process were achieved thank to the identification of a 5-aminovalerate importer encoding gene whose overexpression resulted in a reimport of 5-aminovalerate [86].

The production of 1,4-butanediol (BDO) from C5 sugars has been also attempted through the NP-Weimberg pathway [81], and presents the advantage of requiring 3 times less enzymatic steps than that from glucose [87, 88]. From DOP, only two enzymatic steps catalyzed by  $\alpha$ -keto acid decarboxylase and an aldehyde reductase, respectively, are requested (Fig. 2). Tai et al. [81] identified *Lactococcus lactis* KivD as the most efficient  $\alpha$ -ketoacid decarboxylase (KDC) from an enzymatic screening of various bacterial 2-keto acid decarboxylases to convert DOP into butanedial. Assuming YqhD as the most active NADPH-dependent aldehyde reductase in *E. coli* [68], these authors reported the production of a few g/L of BDO from xylose in an *E. coli* strain equipped with the NP\_Weimberg pathway genes, that also overexpressed *kivD* and was deleted for competing pathways (i.e., xylose isomerase, Dahms pathway). However, this engineered strain produced high amount of

1,2,4-butanetriol (BTO), whereas no BTO was produced using L-arabinose, as the carbon source. These results were explained by the fact that KDC from *Lactococcus lactis* was even more active on KDX than on DOP and was inactive on KDA [81]. Therefore, protein engineering of KivD was realized, leading to a KivD<sup>V411I</sup> variant exhibiting catalytic efficiency 15-fold reduced on KDX and twofold increased on DOP [81]. Replacement of the wild-type KivD by KivD<sup>V411I</sup> variant in the engineered *E. coli* equipped with the NP\_Weimberg pathway resulted in a production of 0.37 g BDO per g D-xylose. Interestingly, this yield was in the same range as that obtained with a metabolically engineered *E. coli* strain for BDO production from glucose [88] (see Additional file 1: Table S2). This may indicate that there is plenty room for optimization of BDO production from pentose sugars by the NP\_Weimberg pathway. However, this process is not expected to be economically viable due to the high cost of xylose used as the sole carbon source. On the other hand, the introduction of the NP\_Weimberg pathway in an already good BDO producer strain from glucose [88] could increase the marketability of this organism for BDO production from lignocellulosic sugars.

#### The NP\_KDC-dependent metabolic route yielding BTO and 3,4-DHBA

The decarboxylation of the NP pathway intermediate KDX or KDA into (*R,S*) 3,4-dehydroxybutyraldehyde constitutes a third non-natural route that can lead to the production of 1,2,4-butanetriol (BTO) or 3,4-dihydrobutyric acid (3,4-DHBA). This metabolite can be further cyclized at acidic pH to give rise to 3-hydroxybutyrolactone ( $\gamma$ HBL) (Fig. 3). These chemicals are widely exploited as precursors for polymers, plasticizer and pharmaceutical drugs (Table 2). The NP\_KDC pathway has been initially established by Frost et al. [89]. These authors identified benzoylformate decarboxylase encoded by *mdlC* from *Pseudomonas putida* as able to decarboxylate KDX. They reported a production of 1.6 g/L BTO from 10 g/L xylonic acid by an *E. coli* DH5 $\alpha$  strain carrying *mdlC* on a high-copy plasmid, while the two other reactions in the pathway were catalyzed by the endogenous xylonate dehydratase and aldehyde reductase. Subsequent works were dedicated to install an optimized NP\_KDC pathway for pentose conversion into BTO or 3,4-DHBA in *E. coli*. This comprised the additional expression of a xylose (XDH) or arabinose dehydrogenase (ArDH) from *Caulobacter crescenti* (Table 1), selection of an appropriate aldehyde reductase or aldehyde dehydrogenase to produce BTO or 3,4-DHBA, or alternatively, expression of a CoA-acylating aldehyde dehydrogenase for the CoA-branch pathway of 3,4-DHBA (Fig. 3). This optimization also included the elimination of the Dahms route and

D-xylose pathways into the pentose phosphate pathway [83, 90].

The expression of the NP\_KDC pathway in *E. coli* led to the identification of bottlenecks at three enzymatic steps in the pathway and which may explain that titer and productivity of BTO and 3,4-DHBA are still far from any industrial exploitation (Additional file 1: Table S2). A first limitation was to find a huge accumulation and excretion of xylonic acid in *E. coli* equipped with the NP\_KDC pathway. Only partial reduction of this co-product was achieved by overexpression of endogenous *E. coli yjhG* or *Caulobacter crescentus xylD* encoding xylonate dehydratase [90–94]. Since XD belongs to the IlvD/EDD protein family that requires iron–sulfur (Fe–S) cluster in its active center [74], availability of iron is essential to warrant efficient activity of XD, which likely becomes limiting when XD is highly overexpressed in the cell. A second issue dealt with the last step in the pathway to produce BTO or 3,4-DHBA, which engaged either an aldehyde reductase or an aldehyde dehydrogenase. This problem turned out to be very critical because all *E. coli* strains equipped with the NP\_KDC pathway were found to produce both chemicals at once [83, 92, 95]. A study of Valdehuesa et al. [95] showed that the loss of function of *yqhD* almost completely abolished BTO production from D-xylose, but did not increase the other chemical. On the other hand, Wang et al. [93] screened six different *E. coli* aldehyde reductases and found that overexpression of an NADH-dependent aldehyde reductase encoded by *adhP* resulted in higher BTO titer. A possible explanation for these different results is that the two authors used different *E. coli* background strains to engineer the NP\_KDC pathway. While Valdehuesa et al. [90] used an MG1655 derivative strain for which the overexpression of *yqhD* is detrimental for growth, Wang et al. [93] expressed the NP\_KDC pathway in an *E. coli* BL21 strain that exhibits for unexplained reasons a higher capacity for BTO synthesis than MG1655. Nevertheless, the titer of BTO obtained with these various engineered *E. coli* expressing the NP-KDC pathway was in the range of 1 to 5 g/L with a yield that was at best 27% of the pathway yield (see Additional file 1: Table S2). To favor 3,4-DHBA at the expense of BTO, Wang et al. [96] overexpressed the endogenous *ynel* encoding a succinate semialdehyde dehydrogenase, which resulted in a twofold increase of 3,4-DHBA while levels of BTO were not reduced. An alternative strategy was to convert 3,4-dihydroxybutanal in its acyl-CoA derivative by a CoA-acylating aldehyde dehydrogenase followed by a thioesterase that releases CoA and 3,4-DHBA. This possibility was found to function in vivo but it did not improve titer and yield of the

engineered strain [94] (see Additional file 1: Table S2). The high promiscuity and the low catalytic efficiency of the 2-keto-acid decarboxylase on KDX (and very likely on KDA) are the third issue that makes NP-KDC pathway still poorly effective for converting pentose into BTO and 3,4-DHBA. As indicated in Table 1, the catalytic efficiency of KDC on KDX is at least one order of magnitude lower than that of the other enzymes in the pathway. Until now, only few 2-keto-acid decarboxylases have been screened for their activity on KDX. According to Wang et al. [93], KdcA from *Lactococcus lactis* was found to be more active than KivD or MdlC in the production of BTO, whereas Gao et al. [94] found that production of 3,4-DHBA was equivalent whether the KDC was encoded by *L. lactis kivD* or *P. putida mdlC*.

The NP\_KDC pathway has been recently expressed in the yeast *S. cerevisiae*. A production of 1.7 g/L of BTO from D-xylose at a yield of 25% (mol/mol) and 1.1 g/L from direct fermentation of rice straw hydrolysate have been reported [97]. The motivation to use yeast as a host organism for the production of this chemical resides in part in its capacity to ferment inexpensive complex hydrolysates obtained from mild acidic treatment of biomass, which are not tolerated by *E. coli* [98, 99]. A major issue in this work was to express an active xylonate dehydratase since this enzyme requires iron–sulfur (Fe–S) for activity [74], while cytosolic iron is rather low in yeast. Bamba et al. [97] showed that deletion of *BOL2* encoding a cytosolic protein involved in the repression of iron regulon [100] combined with the overexpression of *tTYW1*, whose effect was reported to increase cytosolic iron, resulted in an improved production of BTO from D-xylose with a 24% molar yield (see Additional file 1: Table S2).

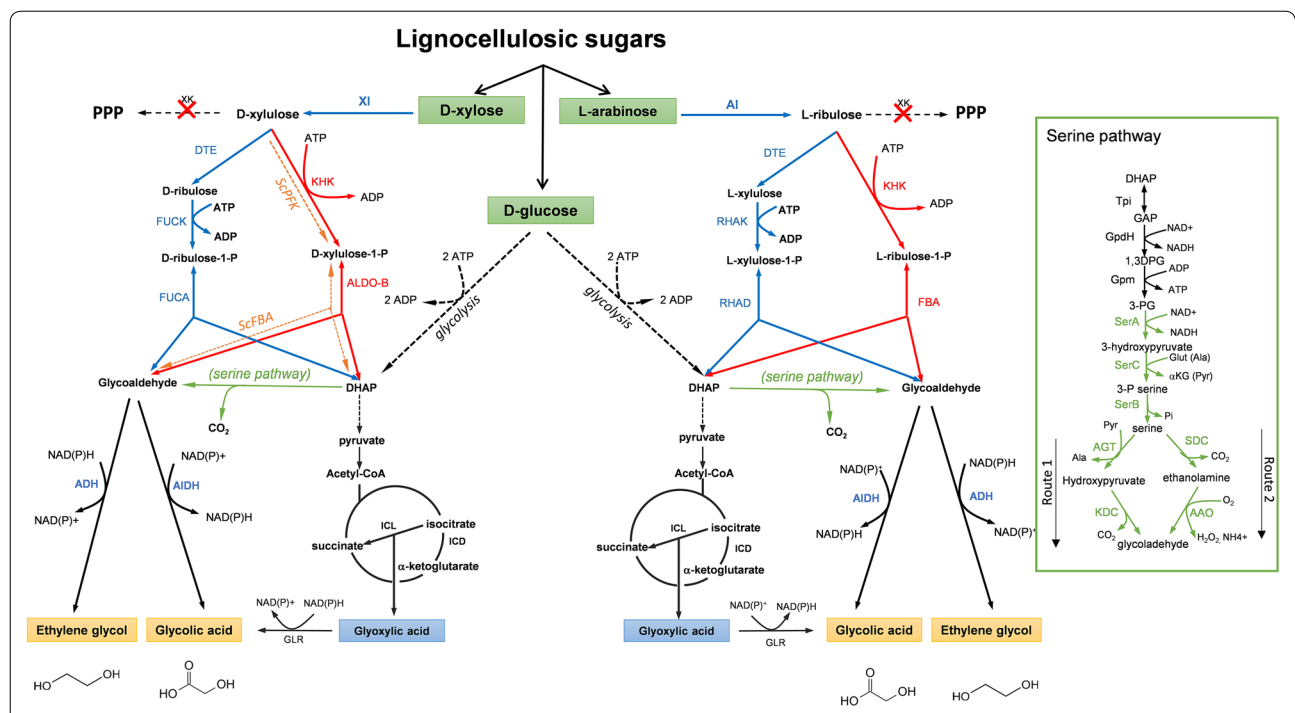
Metabolic pathway has also been engineered in *E. coli* for the production of BTO and 3,4-DHBA or its lactone derivative from glucose [101–103]. As indicated in Fig. 3, the biosynthetic pathway of BTO from glucose comprises a natural part that corresponds to the metabolism of glucose into malate in 9 steps, and a non-natural route involving 6 reaction steps from malate to BTO. The non-natural pathway that converts malate to BTO was assembled using two compatible plasmids, each carrying a specific genes module. Genes of module 1 encode enzymes that convert malate into 2,4-dihydroxybutyric acid in 3 steps. The first step is catalyzed by a malate thiokinase (MTK) whose *mtkAB* gene was isolated from the archaeobacterium *Methylobium petroleiphilum*. The produced malyl-CoA is then reduced into malate semialdehyde (MSA) by a succinate semialdehyde dehydrogenase encoded by *sucD*

from *Porphyromonas gingivalis*. The last step is the reduction of MSA into 2,4-dihydroxybutyric acid that was found to be catalyzed by 4-hydroxybutyrate dehydrogenase encoded by *4-hbB* from *Porphyromonas gingivalis*. Module 2 was constructed with *abfT-2* of *Porphyromonas gingivalis* encoding 4-hydroxybutyrate COA transferase and *adhE2* of *Clostridium acetobutylicum* that encodes the bifunctional aldehyde/alcohol dehydrogenase. Upon transformation of *E. coli* strain BL21 with these two compatible plasmids, a tiny amount (120 ng/L) of BTO was detected. Although the feasibility of this non-natural pathway was demonstrated, many side products like BDO and 4-hydroxybutyrate were also accumulated, indicating that most of the requested enzymes for this non-natural pathway showed poor specificity for the intermediates of this pathway [101]. In addition, the requirement for reduced cofactors and ATP makes this pathway not beneficial with a maximal yield of 0.65 mol BTO/mol glucose (Table 3). Figure 3 also schematically illustrates the non-natural enzymatic steps leading to the synthesis of 3,4-DHBA from glucose. The whole pathway encompasses 15 enzymatic reactions, including a path of 5 non-natural reactions that converts glyoxylate to 3,4-DHBA. The construction of this pathway has been made using a transacetylase encoded by *pct* from *Megasphaera elsdenii*, a thiolase and acyl-CoA reductase encoded by *btkB* and *phaB*, respectively, from *Cupriavidus necator* and finally a thioesterase encoded by the endogenous *E. coli tesB* gene. In addition, overexpression of acetyl-CoA synthetase could be required to recover acetate into acetyl-CoA. Overall, this pathway could optimally produce 1 mol 3,4-DHBA from 1 mol glucose with the generation of 2 mol reduced equivalents and 1 mol ATP (Table 3). While this positive energetic balance is an advantage for strains' stability and viability, major problem to be solved is many co-products generated due to the low specificity of the involved enzymes for their non-natural substrates. This is mainly the case for BtkB, which can condense two acetyl-CoA into 3-hydroxybutyryl CoA, or can produce equally 2 hydroxy-3-ketobutyryl-CoA and 4-hydroxy-3-ketobutyryl-CoA from glycolyl-CoA and acetyl-coA, which thereafter generates equal amount of 2,3-3,4-DHBA and 3,4-3,4-DHBA [103]. Metabolic burden caused by the number of genes to assemble this pathway adds another difficulty that affects product titers. Altogether, the complexity of this glucose-depending pathway for the synthesis of 3,4-DHBA and  $\gamma$ HBL raises question about its feasibility toward industrial development.

### Orthogonality of non-natural pentose-1-phosphate pathway to central carbon metabolism for optimal conversion of lignocellulosic sugars into bio-based products

The implementation of non-phosphorylating (NP) pathway in microbial systems for conversion of lignocellulosic sugars into bio-based products presents two main difficulties that may hinder use of these microbial strains for industrial exploitation. A first limitation is the strong dependency of the xylonate/arabinonate dehydratase to iron, which strongly affects in vivo activity, and explains high and eventually toxic accumulation of xylonic acid in engineered strains harboring the NP pathway. The second limitation relates to the fact that only pentose sugars can be converted into bio-based product through the NP. To circumvent the first limitation and to partially solve the second one, non-natural pathways have been recently designed that enable the production of bio-based products from lignocellulosic sugars at higher production yield and productivity. The example presented in this review will focus on the bioproduction of EG and GA.

As depicted in Fig. 4, three novel non-natural pathways have been designed, assembled and tested in vivo in *E. coli*, *C. glutamicum* or yeast for the production of EG and GA from pentose and glucose (also reviewed in [104]). The first two pathways draw their originality on the phosphorylation of the pentose sugar D-xylose or D-arabinose on its C-1 carbon instead of C-5 carbon as it occurs in the pentose phosphate pathway. Cam et al. [105] identified the mammalian ketohexokinase encoded by *khkC* [106] as very effective to phosphorylate D-xylose or L-arabinose into xylulose-1P or arabinose-1P, whereas Stephanopoulos' group [107] trusted on the native fuculo-1-kinase and rhamnulokinase encoded by *fucK* and *araB* implicated in D-arabinose and L-lyxose metabolism [108] to yield similar pentose phosphorylation. However, the latter strategy required a step prior to phosphorylation that is catalyzed by a tagatose epimerase encoded by *dtE* from *Pseudomonas cichorii*, which converts D-xylulose into D-ribulose and L-ribulose into L-xylulose. The pentose-1P generated can be aldolytically cleaved into glycolaldehyde (C2) and DHAP (C3), which requires an aldolase that turns out to be rather specific for the pentose-1P produced. The mammalian aldolase B (AldoB) was found to be very efficient to cleave xylulose-1P, but ineffective on L-ribulose-1P [109], whereas endogenous aldolases encoded by *fucA* and *rhaD* were used to cleave D-ribulose-1P and L-xylulose-1P [107] (see Additional file 1: Table S3 for a summary of enzymes used in these non-natural pathways). Production of EG was found to be dependent on the aldehyde reductase encoded by *yqhD* but its maximal conversion from glycolaldehyde also needed overexpression of *fucO* and deletion of *aldA*



**Fig. 4** Xylulose 1-Phosphate, Ribulose-1-Phosphate and serine pathway connected to central carbon metabolism for conversion of pentose and hexose into GA and EG. XI: xylose isomerase; DTE: D-tagatose 3-epimerase; FUKC: fuculokinase; FUCA: fuculo-phosphate aldolase; KHKC: ketohexokinase C; Aldo-B: mammalian aldolase B; ScFBA: *S. cerevisiae* FBP aldolase; RHAK: L-rhamno-1-P kinase; RHAD: Rhamno-1-P aldolase; ICL: isocitrate lyase; AIDH: aldehyde dehydrogenase; ALR: aldehyde reductase. Inset of Fig. 4 is shown the two serine pathways. TPI: triose phosphate Isomerase; GpdH: glyceraldehyde-3-phosphate dehydrogenase; SerA: 3-hydroxypyruvate dehydrogenase; SerC: serine:  $\alpha$ -KG transaminase; SerB: phosphoserine phosphatase; AGT: alanine-glyoxylate transaminase; KDC:  $\alpha$ -keto-acid decarboxylase; SDC: serine adecarboxylase; AAO: amino acid oxidase. Additional details on enzyme and encoding genes are given in Additional file 1: Table S3

[107, 110]. Conversely, maximal production of GA from glycolaldehyde [105] requires not only solely overexpression of *aldA* but also the deletion of *glcD* encoding the membrane-associated glycolate oxidase [111]. Expression of these non-natural pathways in *E. coli*, in which competing pathways of xylose and arabinose into PPP were deleted, resulted in the production of EG and GA that reached 85 to 90% of the maximal yield expected (1 mol/mol sugar) (see Additional file 1: Table S2). This 10 to 15% missing carbon raises the question of its fate from xylose since, at first glance, the issue of the two-carbon fraction of this sugar should be EG and/or GA.

Xylulose-1-Phosphate (X1P) and Ribulose-1-Phosphate (R1P) pathways have been tentatively expressed in the yeast *S. cerevisiae*. In a first report by [112], a weak production of 0.25 g/L EG was obtained with a  $\Delta xks1$  yeast mutant defective in xylulose-5-kinase that overexpressed XI, *khkC* and *FBA1* encoding respectively xylose isomerase, mammalian ketohexokinase and the endogenous FBP aldolase. However, the production of EG was accompanied by a significant accumulation of xylitol, glycerol and ethanol. This finding indicated the presence of a very active aldo-reductase/alcohol dehydrogenase in

spite of the fact that this engineered strain was deleted for *GRE3* encoding a major aldo-reductase implicated in the reduction of xylose into xylitol [113]. Another report revealed that *S. cerevisiae* may have a native capability to produce EG from glucose, given that it harbors a high xylose isomerase (XI) activity that has been obtained by in vivo evolutionary engineering. Because of this high XI activity, xylulose accumulates in the cells and can be illegitimately phosphorylated into its C1 carbon by phosphofructo-6-kinase encoded by *PFK1/2* [114]. In addition, the loss of function of the *XKS1*-encoded xylulose-5-P kinase favored GE production. A production of 4 g/L EG from 20 g/L xylose was obtained with such an XI-evolved  $\Delta xks1$  strain that also overexpressed an *E. coli* aldehyde reductase encoded by *fucO* [114].

As stated above, a limitation of these non-natural pathways is that only the two-carbon fraction is diverted into the products, leaving the C3 fraction (DHAP) into the biomass. Thus, to increase the yield, it might be worth to recapture this C3 fraction to glycolaldehyde. A possibility is through the serine pathway, which can lead to glycolaldehyde from DHAP via 8 reaction steps from which the first five enzymatic reactions correspond to the natural



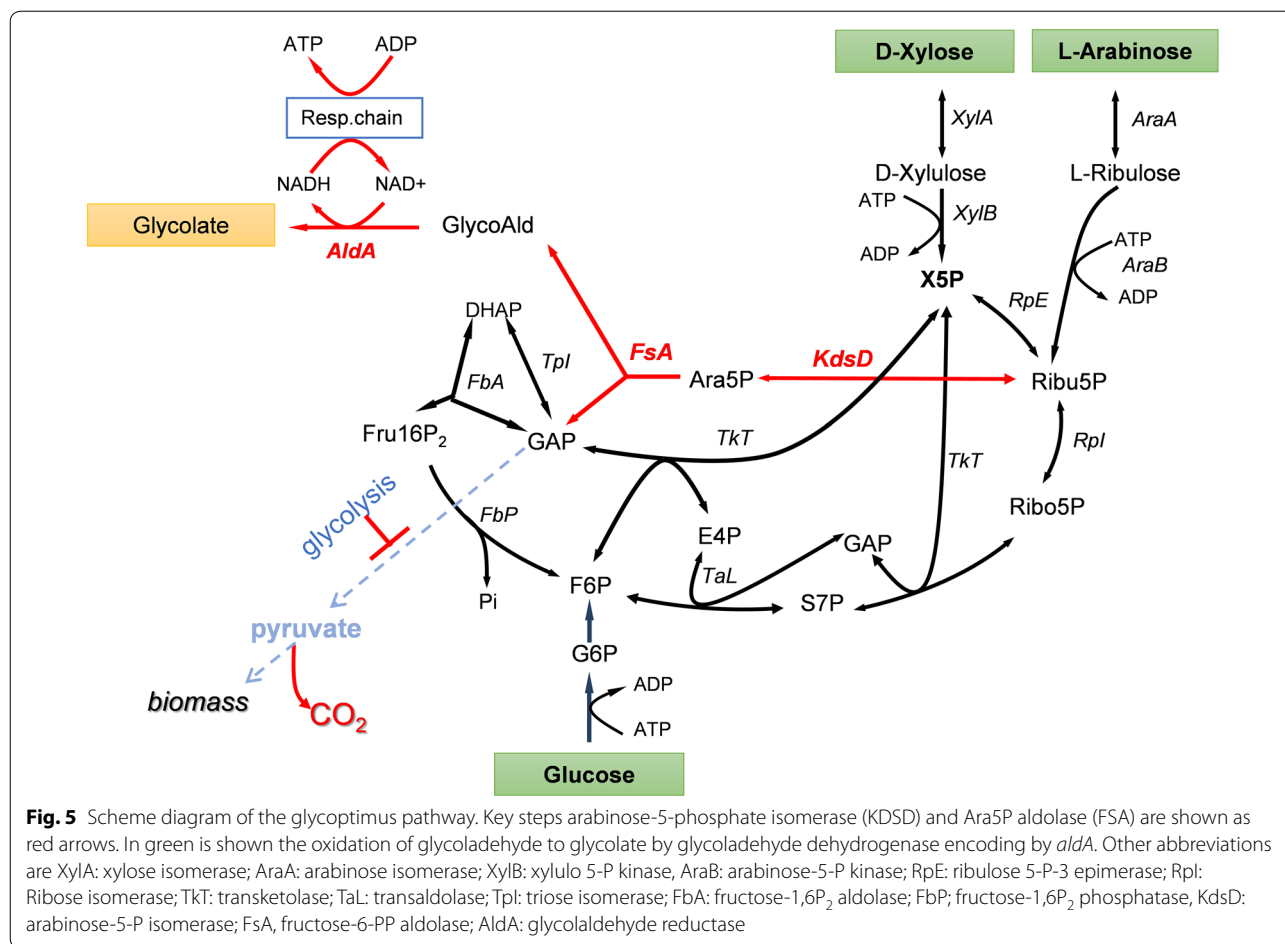
pathway for serine biosynthesis. From serine, there are actually two different routes to yield glycolaldehyde (see inset Fig. 4). A first pathway (route 1) involves a deamination of serine into hydroxypyruvate by an aminotransferase (TA) or amino acid dehydrogenase followed by decarboxylation into glycolaldehyde by a  $\alpha$ -ketoacid decarboxylase (KDC). This route will generate two extra NADH per DHAP produced. The second route (route 2) involves a decarboxylation of serine into ethanolamine by a serine decarboxylase (SDC) followed by a reduction into glycolaldehyde by a monoamineoxidase (AAO). This route will generate hydrogen peroxide in addition to NADH. In theory, the serine pathway shall yield 2 mol EG/mole glucose with 2 extra NADH (Table 3) that can be reoxidized through respiratory chain coupled to ATP synthesis. Pereira et al. [115] implemented the latter route into an *E. coli* strain and after some genetic modifications that was intended to favor the carbon flux in the serine pathway, the engineered strain produced 4.1 g/L EG from 20 g/L glucose which corresponded to 30% of the pathway yield (see Additional file 1: Table S2). Main reasons invoked for this relatively low yield were that serine is a major amino acid implicated in several essential metabolic pathways [116] and that C3 intermediate cannot be easily channeled into the serine pathway due to its central position in the carbon metabolism for growth [115]. In addition, this pathway generates  $H_2O_2$  that can be toxic for the cell [117, 118]. As *Corynebacterium glutamicum* is a good producer of serine [119], Chen et al. [120] evaluated which one of the two serine pathways was the most efficient in the production of EG from glucose in this bacteria. This strategy required to identify an efficient  $\alpha$ -ketoacid decarboxylase and a transaminase for route 1. The decarboxylase was screened by incubating a *C. glutamicum* strain that overexpressed the *E. coli* aldehyde reductase encoded by *yqhD* with hydroxypyruvate as the substrate to produce EG, whereas serine was used as the substrate for the transaminase. The benzoylformate decarboxylase encoded by *mdlC* from *P. putida* and alanine-glyoxylate transaminase from *Arabidopsis thaliana* (AtAGT) were retained by this screening test. Similarly, an active amino acid oxidase (AAO) and serine decarboxylase (SDC) were screened based on the production of EG using the same *C. glutamicum* strain as above but incubated with ethanolamine and serine, respectively. This screening identified the monoamineoxidase from *Arthrobacter* sp. and serine decarboxylase from *Arabidopsis thaliana* as the best candidates in route 2. Subsequently, these pathways were introduced in a *C. glutamicum* strain that had been engineered from hyperproduction of serine, notably by deletion of *sda* encoding serine deaminase and overexpressing the *serABC* operon. These authors showed that a strain that expressed route

2 was roughly 3 times more efficient than that carrying route 1 for EG production from glucose. However, a higher production of EG was obtained in a strain that contained both route [120], even though the titer was still low (3.5 g/L) and the yield was only 12% of the maximal pathway yield (Additional file 1: Table S2), which is surprisingly lower than with the engineered *E. coli* constructed by Gu et al. [119].

While the X1P and R1P pathways were originally designed to convert C5 into bio-based products, the co-functioning of these synthetic pathways with the glyoxylate shunt (GS) has been investigated for the production of GA from both xylose and glucose. In theory, this combination shall produce 2 mol GA/mol of sugar with a positive redox and ATP balance (see Table 3). Preliminary attempts of GA production by the co-function of both pathways on D-xylose were awkward since Alkim et al. [109] did not find further increase of GA as compared to X1P pathway alone. On the contrary, Pereira et al. [107] reported a significant increased yield from 0.44 to 0.63 g/g xylose by combining R1P with GS pathway. A plausible explanation for this discrepancy is that the engineered strain made by Alkim et al. [109] was defective in *icd* encoding  $NADP^+$ -dependent isocitrate dehydrogenase, the deletion of which is known to strongly impair TCA cycle [121]. This problem was overridden by a co-feeding of glucose and xylose which resulted in an improved GA yield up to 0.63 g/g total sugar with a yield on the xylose fraction that reached 75% of the maximal theoretical yield [109] (Table 3). Although an engineering strategy with serine pathway instead of GS has not been experimentally verified yet, one can anticipate that the combination of the X1P (or R1P) with serine pathway should also improve the EG production yield in co-feeding with xylose and glucose.

### Refactoring metabolic pathways for conversion of lignocellulosic sugars into bio-based products aimed at optimal carbon conservation

A general feature of metabolic pathways is their inherent carbon loss as  $CO_2$ . There has been several attempts to refactor carbon metabolism in order to either overcome when possible or minimize this carbon lost and hence to optimize carbon conservation, as demonstrated by Liao and coworkers for the non-oxidative glycolysis (NOG). This reengineering of glycolytic pathway can lead to the maximal conversion 1 mol of glucose into 3 mol of acetate [122]. According to this objective, we recently designed a cyclic non-natural pathway—termed the glycoptimus pathway—enabling conversion of C5 and C6 sugar into GA with optimal carbon conservation (Fig. 5). The proposed pathway theoretically results in a yield of 3 mol GA per mole glucose or 2.5 mol GA



per mole pentose, which is 50% higher than obtained by the natural-glyoxylate-dependent pathway. The key feature of glycoptimus relies on the use of arabinose-5P (Ara5P) isomerase encoded by *kdsD* and a class I aldolase encoded by *fsaA/fsaB* to generate glycolaldehyde and GAP from Xu5P. While glycolaldehyde is oxidized into GA, GAP can be shuttled back into the non-oxidative pentose phosphate pathway via transketolase and transaldolase reaction. This cycling scheme also generates 3 mol of NADH that needs to be reoxidized to allow a continuous running of the pathway. The reoxidation of NADH requires an active respiratory chain, which is coupled to the synthesis of ATP. This ATP provision is useful for the uptake and the phosphorylation of sugars, as well as for strain stability and maintenance. The in vitro and in vivo functions of the glycoptimus pathway have been validated although the yield of GA from glucose and xylose did not exceed 30% of the maximal yield expected. This poor efficiency of the pathway could be due to several problems and notably that the blockage of the lower part of glycolysis at the level of GAPDH may be detrimental for cell viability. In addition, many competitive

pathways siphoning intermediates of the cycle have not been dismantled yet in the engineered strains and carbon flux in the *kdsD-fsaA* pathway needs to be optimised [123]. Finally, the reoxidation of NADH is likely to generate more ATP than needed for the sole functioning of the glycoptimus cycle, which could result in feedback inhibition due to accumulation of intermediates and eventually depletion of cofactors. Solving these issues together with the development of an appropriate fermentation process is critically important to develop an economically viable process of glycolic acid production from lignocellulosic sugars.

**Conclusion and prospective**

The sustainability and competitiveness of biotechnological processes for the production of bio-based chemicals from renewable carbon sources critically depend on the need to equip the microbial platform with efficient metabolic pathways able to readily transform these carbon resources into bio-based products at titer, rate and yield (TRY) that are economically viable. For high-volume, low-priced chemicals such as those considered here, the

higher these three performance indices are, the more competitive biotechnological processes become compared to petrochemical processes. Even though many microorganisms on earth are able to assimilate pentose and hexose sugars, their carbon metabolism is not optimally fashioned to achieve these performances. This necessitates to rewire their actual metabolic network, to plug (non-natural) pathways that exist in other organisms or to create new ones (artificial or synthetic), while ensuring cellular homeostasis. Metabolic diversity found in Nature, combined with enzymes' promiscuity, provides inspiration for implementing non-natural or constructing synthetic metabolic pathways that could meet these goals.

In this review, we showed that plugging the non-phosphorylating pentose pathway or implementing synthetic pentose-1-P pathway in the industrially tractable organism *E. coli*, *Corynebacterium glutamicum* or *S. cerevisiae* empower these microbial cells with strong capability to convert lignocellulosic sugars into a wide range of added-value chemicals. Although these microbial processes are far from industrial performances, there are strong indications that this objective could be soon reached. First, these pathways are orthogonal to the natural metabolic networks, which in theory make them insensitive to endogenous metabolic and genetic regulation, while being deliberately controllable by mean of biosensors/biocontrollers [124]. This orthogonality may allow to partition carbon utilization into the production and growth purpose, enabling tunable dynamic control to favor one purpose over the other, as it was demonstrated for the production of glucaric acid from glucose [125]. Second, the number of reaction steps from the substrate to the product in the constructed pathway has to be as short as possible. The synthetic pathway for BDO from xylose is a best example since it only encompasses 6 steps, instead of the 21 required from glucose. Less reaction steps go in hand with less optimization workload and less intermediates to be siphoned off in branching pathways. Consequently, higher yield and higher productivity should be expected for production of bio-based chemicals.

There still remain many challenges ahead to better improve microbial cell factories and concurrent biotechnology processes to make them industrially more appealing. We anticipate at least three major actions that must be actively worked on. The first one dealt with further rewiring of carbon metabolism for the conversion of all available renewable carbon resources, which include not solely lignocellulosic biomass, but CO<sub>2</sub> and methane from landfills. Although the conversion of these carbon sources into bio-based chemicals still largely depends on traditional metabolic paths, new routes must be designed and constructed that should be orthogonal to

the central carbon metabolism without perturbing the cellular homeostasis. As for instance, new route for methane and CO<sub>2</sub> assimilation into ethanol has been proposed by Liao and coworkers [126]. The same authors also provided various pathways enabling rewiring carbon metabolism to achieve maximal carbon conservation (reviewed in [127]). Innovation in yet inexistent pathways can be guided from exploration of the huge metagenomics data using chemoinformatics tools. This strategy has been recently applied to explore novel biosynthetic pathways for EG synthesis from C1 carbon source [128]. A second action that is critically determinant for high productivity and yield will be to increase catalytic efficiency and substrate specificity of pathway enzymes. However, getting these criteria can be a daunting task especially for enzymes in non-natural or synthetic pathways for which the substrates are not natural. This problem was clearly illustrated with  $\alpha$ -ketoacid decarboxylase of *Lactococcus lactis*, whose promiscuity has been exploited to decarboxylate KdxD, DOP and  $\alpha$ -adipate, but concomitantly generated metabolic side products due to this promiscuous activity [93]. Nevertheless, enzyme promiscuity provides an evolution starting point to yield specialized enzymes that catalyze the desired reactions on non-natural substrate. Rational enzyme engineering is a first approach to investigate the possibility to create enzyme with new function or with higher affinity to non-natural substrate than on its natural substrate. This strategy has been successful to completely change the specificity of an *E. coli* aspartate kinase encoded by *lysC* from its natural substrate—aspartate—to malate [129]. However, this rational approach is insufficient in many cases, asking for enzyme evolution either by random mutagenesis or by in vivo evolutionary engineering. Direct enzyme evolution can be very powerful if an appropriate high-throughput assay can be set up. Droplet-based microfluidic is thus well appropriate for this evolution, and its combination with IVTT (in vitro transcription–translation system) enables to screen for even larger libraries as it avoids transformation step in *E. coli* prior to the screen [130, 131]. Alternative to this method is the direct in vivo enzyme evolution using automated micro-reactors as developed by Altar company (<http://www.altar.bio/about-us/>). The use of this technique requires a phenotypic screening assay that exclusively relies on the vital function of the enzyme of interest. As an example, this strategy has been used to evolve an amino acid transaminase that can convert L-homoserine into the non-natural intermediate 2-keto-4-hydroxybutyrate (HOB), which thereafter could transfer its one carbon unit (CHO) to tetrahydrofolate (H4F), replacing the natural transfer C1-moieties from serine and glycine to H4F in an *E. coli* [132]. Fermentation process is likely a third target to be

addressed early in the development of the microbial platforms as nicely reviewed by [133]. This aspect is really of the outmost importance when it deals with lignocellulosic hydrolysates or any raw materials that are on the one hand very cheap but on the other hand ill characterized, causing harsh fermentation conditions. Hence, key for industrial success, these new engineered microbial systems must be readily challenged with these harsh conditions, such as increasing their tolerance to toxic compounds and their global robustness by adapted laboratory evolution techniques [134, 135]. Besides these aspects, the choice between production phases coupled or decoupled to growth can be decisive in setting the fermentation process as it should be mostly determined by the capacity to reach maximal production yield, while assuring redox and energy balance.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s13068-020-01744-6>.

**Additional file 1: Table S1.** Enzymes and corresponding genes that has been used in the expression of the that has been used in the Dahms, Weimberg and KDC-dependent non phosphorylating (NP) pathway for the production of bio-based commodity chemicals from hexose and pentose sugars by *E. coli* as described in Figs. 3 and 4. **Table S2.** Production titer and yield of bio-based chemicals from lignocellulosic sugars by engineered microorganisms. **Table S3.** Enzymes and corresponding genes that have been used in the xylulose-1P and ribulose-1P synthetic pathways for the production of bio-based commodity chemicals from pentose sugars. **Table S4.** Thermodynamic yield and pathway yield of bio-based chemicals from sugars.

## Abbreviations

DHBA: 3,4-Dihydroxybutyric acid; 3-DPA: 3-Deoxypentanoic acid;  $\alpha$ KG:  $\alpha$ -2-Ketoglutarate; AAO: Amino acid oxidase; Ara5P: Arabinose-5-phosphate; BDO: 1,4-Butanediol; BTO: 1,2,4-Butanetriol; CAGR: Compound annual growth rate; CCM: Central carbon metabolism; DHB: 2,4-Dihydroxybutyric acid; DOP: 2,5-Dioxopentanoate; EG: Ethylene glycol; GA: Glycolic acid; GAP: Glyceraldehyde-3-P; GLR: Glyoxylate reductase; H4F: Tetrahydrofolate; HOB: 2-Keto-4-Hydroxybutyrate; IVTT: In vitro transcription-translation system; kDa: 2-Keto-3-deoxy-L-arabinonate; KDC:  $\alpha$ -Ketoacid decarboxylase (KDC); KDP: D-2-Keto-3-deoxy-pentanoate; KDPA: 2-Keto-3-deoxy-pentanoate Aldolase; KDPH: 2-Keto-3-deoxy-pentanoate Hydratase; KDX: 2-Keto-3-deoxy-D-xylonate; KDA: 2-Keto-3-deoxy-L-arabinonate; MSA: mesaconic acid; MTK: Malate thiokinase; NOG: Non-oxidative glycolysis; NP: Non phosphorylating pentose pathway; PEP: Phosphoenolpyruvate; PK: Phosphoketolase; PPP: Pentose phosphate pathway; R1P: Ribulose-1-Phosphate; SDC: Serine decarboxylase; TA: Transaminase; TCA: Tricarboxylic acid; TRY: Titer, rate and yield; X1P: Xylulose-1-Phosphate; XD: Xylonic dehydratase; XI: Xylose isomerase; Xu5P: Xylulose-5-Phosphate.

## Acknowledgements

The authors wish to acknowledge Agence Nationale de la Recherche (ANR), Ministry of Research and Education (MENRT) and TWB for technical and financial support.

## Authors' contributions

All authors contributed to the writing of this review. All authors read and approved the final manuscript.

## Funding

Researches in the authors' laboratory were supported by French Research National Agency, grant ANR-10-BTBR-05, Grant ANR-CE-0024, by Agence De l'environnement et Maîtrise de l'Energie (ADEME, convention no 1782C0056), and by Toulouse White Biotechnology (Exploratory Project Pentosys Grant no 1337504).

## Availability of supporting data

Not applicable.

## Ethical approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interest.

## Author details

<sup>1</sup>Toulouse Biotechnology Institute, CNRS, INRA, LISBP INSA, 135 Avenue de Rangueil, Toulouse Cedex 04 31077, France. <sup>2</sup>Toulouse White Biotechnology (TWB, UMS INRA/INSA/CNRS), NAPA CENTER Bât B, 3 Rue Ariane 31520, Ramonville Saint-Agnes, France.

Received: 11 December 2019 Accepted: 1 June 2020

Published online: 08 July 2020

## References

- Flores BY, Tangney M. Synthetic biology in the driving seat of the bioeconomy. *Trends Biotechnol.* 2017;35:373–8.
- Aguilar A, Magnien E, Thomas D. Thirty years of European biotechnology programmes: from biomolecular engineering to the bioeconomy. *N Biotechnol.* 2013;30:410–25.
- Paul S, Dutta B. Challenges and opportunities of lignocellulosic biomass for anaerobic digestion. *Resour Conserv Recycl.* 2018;130:174.
- Graham-Rowe D. Agriculture: beyond food versus fuel. *Nature.* 2011;474:56–8.
- Venkateswar RL, Goli JK, Gentela J, Koti S. Bioconversion of lignocellulosic biomass to xylitol: an overview. *Bioresour Technol.* 2016;213:299–310.
- Papagianni M. Recent advances in engineering the central carbon metabolism of industrially important bacteria. *Microb Cell Fact.* 2012;11:50.
- Hahn-Hagerdal B, Karhumaa K, Jeppsson M, Gorwa-Grauslund MF. Metabolic engineering for pentose utilization in *Saccharomyces cerevisiae*. *Adv Biochem Eng Biotechnol.* 2007;108:147–77.
- Weimberg R. Pentose oxidation by *Pseudomonas fragi*. *J Biol Chem.* 1961;236:629–35.
- Sato T, Atomi H. Novel metabolic pathways in Archaea. *Curr Opin Microbiol.* 2011;14:307–14.
- Kohler KA, Blank LM, Frick O, Schmid A. D-Xylose assimilation via the Weimberg pathway by solvent-tolerant *Pseudomonas taiwanensis* VLB120. *Environ Microbiol.* 2015;17:156–70.
- Yadav VG, De Mey M, Lim CG, Ajikumar PK, Stephanopoulos G. The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab Eng.* 2012;14:233–41.
- Ceroni F, Carbonell P, Francois JM, Haynes KA. Editorial—synthetic biology: engineering complexity and refactoring cell capabilities. *Front Bioeng Biotechnol.* 2015;3:120.
- Duwe A, Tippkötter N, Ulber R. Lignocellulose-biorefinery: ethanol-focused. *Adv Biochem Eng Biotechnol.* 2019;166:177–215.
- Paulino de Souza J, Dias do Prado C, Eleutherio ECA, Bonatto D, Malavazi I, Ferreira da Cunha A. Improvement of Brazilian bioethanol production—challenges and perspectives on the identification and genetic modification of new strains of *Saccharomyces cerevisiae* yeasts isolated during ethanol process. *Fungal Biol.* 2018;122:583–91.
- Liu H, Sun J, Chang JS, Shukla P. Engineering microbes for direct fermentation of cellulose to bioethanol. *Crit Rev Biotechnol.* 2018;38:1–17.

16. Zhao X, Xiong L, Zhang M, Bai F. Towards efficient bioethanol production from agricultural and forestry residues: exploration of unique natural microorganisms in combination with advanced strain engineering. *Bioresour Technol*. 2016;215:84–91.
17. Ko JK, Lee SM. Advances in cellulosic conversion to fuels: engineering yeasts for cellulosic bioethanol and biodiesel production. *Curr Opin Biotechnol*. 2018;50:72–80.
18. Zhang Y, Liu D, Chen Z. Production of C2–C4 diols from renewable bioresources: new metabolic pathways and metabolic engineering strategies. *Biotechnol Biofuels*. 2017;10:299.
19. Zhao Z, Xian M, Liu M, Zhao G. Biochemical routes for uptake and conversion of xylose by microorganisms. *Biotechnol Biofuels*. 2020;13:21.
20. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*. 2007;315:804–7.
21. Li M, Pu Y, Ragauskas AJ. Current understanding of the correlation of lignin structure with biomass recalcitrance. *Front Chem*. 2016;4:45.
22. Chundawat SP, Beckham GT, Himmel ME, Dale BE. Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu Rev Chem Biomol Eng*. 2011;2:121–45.
23. Satlewal A, Agrawal R, Bhagia S, Sangoro J, Ragauskas AJ. Natural deep eutectic solvents for lignocellulosic biomass pretreatment: recent developments, challenges and novel opportunities. *Biotechnol Adv*. 2018;36:2032–50.
24. Van de Wouwer D, Boerjan W, Vanholme B. Plant cell wall sugars: sweeteners for a bio-based economy. *Physiol Plant*. 2018;164:27–44.
25. Brandon AG, Scheller HV. Engineering of bioenergy crops: dominant genetic approaches to improve polysaccharide properties and composition in biomass. *Front Plant Sci*. 2020;11:282.
26. Chung D, Cha M, Guss AM, Westpheling J. Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. *Proc Natl Acad Sci USA*. 2014;111:8931–6.
27. Holwerda EK, Worthen RS, Kothari N, Lasky RC, Davison BH, Fu C, Wang ZY, Dixon RA, Biswal AK, Mohnen D, et al. Multiple levers for overcoming the recalcitrance of lignocellulosic biomass. *Biotechnol Biofuels*. 2019;12:15.
28. Lynd LR, van Zyl WH, McBride JE, Laser M. Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol*. 2005;16:577–83.
29. Jojima T, Omumasaba CA, Inui M, Yukawa H. Sugar transporters in efficient utilization of mixed sugar substrates: current knowledge and outlook. *Appl Microbiol Biotechnol*. 2010;85:471–80.
30. Colabardini AC, Ries LN, Brown NA, Dos Reis TF, Savoldi M, Goldman MH, Menino JF, Rodrigues F, Goldman GH. Functional characterization of a xylose transporter in *Aspergillus nidulans*. *Biotechnol Biofuels*. 2014;7:46.
31. Li J, Xu J, Cai P, Wang B, Ma Y, Benz JP, Tian C. Functional analysis of two L-arabinose transporters from filamentous fungi reveals promising characteristics for improved pentose utilization in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2015;81:4062–70.
32. Seiboth B, Metz B. Fungal arabinan and L-arabinose metabolism. *Appl Microbiol Biotechnol*. 2011;89:1665–73.
33. Jin YS, Cruz J, Jeffries TW. Xylitol production by a *Pichia stipitis* D-xylulokinase mutant. *Appl Microbiol Biotechnol*. 2005;68:42–5.
34. Young E, Lee SM, Alper H. Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol Biofuels*. 2010;3:24.
35. Harhangi HR, Akhmanova AS, Emmens R, van der Drift C, de Laat WT, van Dijken JP, Jetten MS, Pronk JT. Op den Camp HJ. xylose metabolism in the anaerobic fungus *Piromyces* sp. strain E2 follows the bacterial pathway. *Arch Microbiol*. 2003;180:134–41.
36. Wojtkiewicz B, Szmidszinski R, Jezierska A, Cocito C. Identification of a salvage pathway for D-arabinose in *Mycobacterium smegmatis*. *Eur J Biochem*. 1988;172:197–203.
37. Heath EC, Hurwitz J, Horecker BL, Ginsburg A. Pentose fermentation by *Lactobacillus plantarum*. I. The cleavage of xylulose 5-phosphate by phosphoketolase. *J Biol Chem*. 1958;231:1009–29.
38. Meile L, Rohr LM, Geissmann TA, Herensperger M, Teuber M. Characterization of the D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase gene (xfp) from *Bifidobacterium lactis*. *J Bacteriol*. 2001;183:2929–36.
39. Liu L, Zhang L, Tang W, Gu Y, Hua Q, Yang S, Jiang W, Yang C. Phosphoketolase pathway for xylose catabolism in *Clostridium acetobutylicum* revealed by 13C metabolic flux analysis. *J Bacteriol*. 2012;194:5413–22.
40. Aristilde L, Lewis IA, Park JO, Rabinowitz JD. Hierarchy in pentose sugar metabolism in *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 2015;81:1452–62.
41. Panagiotou G, Andersen MR, Grotkjaer T, Regueira TB, Hofmann G, Nielsen J, Olsson L. Systems analysis unfolds the relationship between the phosphoketolase pathway and growth in *Aspergillus nidulans*. *PLoS ONE*. 2008;3:e3847.
42. Fabris M, Matthijs M, Rombauts S, Vyverman W, Goossens A, Baart GJ. The metabolic blueprint of *Phaeodactylum tricornutum* reveals a eukaryotic Entner–Doudoroff glycolytic pathway. *Plant J*. 2012;70:1004–14.
43. Xiong W, Lee TC, Rommelfanger S, Gjersing E, Cano M, Maness PC, Ghirardi M, Yu J. Phosphoketolase pathway contributes to carbon metabolism in cyanobacteria. *Nat Plants*. 2015;2:15187.
44. McClintock MK, Wang J, Zhang K. Application of nonphosphorylative metabolism as an alternative for utilization of Lignocellulosic Biomass. *Front Microbiol*. 2017;8:2310.
45. Dahms AS. 3-Deoxy-D-pentulose aldolase and its role in a new pathway of D-xylose degradation. *Biochem Biophys Res Commun*. 1974;60:1433–9.
46. Nunn CE, Johnsen U, Schonheit P, Fuhrer T, Sauer U, Hough DW, Danson MJ. Metabolism of pentose sugars in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. *J Biol Chem*. 2010;285:33701–9.
47. Archer RM, Royer SF, Mahy W, Winn CL, Danson MJ, Bull SD. Syntheses of 2-Keto-3-deoxy-D-xylonate and 2-Keto-3-deoxy-L-arabinonate as stereochemical probes for demonstrating the metabolic promiscuity of *Sulfolobus solfataricus* towards D-Xylose and L-Arabinose. *Chemistry*. 2013;19:2895–902.
48. Johnsen U, Dambeck M, Zaiss H, Fuhrer T, Soppa J, Sauer U, Schonheit P. D-xylose degradation pathway in the halophilic archaeon *Haloferax volcanii*. *J Biol Chem*. 2009;284:27290–303.
49. Oreb M, Dietz H, Farwick A, Boles E. Novel strategies to improve co-fermentation of pentoses with D-glucose by recombinant yeast strains in lignocellulosic hydrolysates. *Bioengineered*. 2012;3:347–51.
50. Kim JH, Block DE, Mills DA. Simultaneous consumption of pentose and hexose sugars: an optimal microbial phenotype for efficient fermentation of lignocellulosic biomass. *Appl Microbiol Biotechnol*. 2010;88:1077–85.
51. Farwick A, Bruder S, Schadeweg V, Oreb M, Boles E. Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci USA*. 2014;111:5159–64.
52. Yao R, Hirose Y, Sarkar D, Nakahigashi K, Ye Q, Shimizu K. Catabolic regulation analysis of *Escherichia coli* and its crp, mlc, mgsA, pgi and ptsG mutants. *Microb Cell Fact*. 2011;10:67.
53. Kim SM, Choi BY, Ryu YS, Jung SH, Park JM, Kim GH, Lee SK. Simultaneous utilization of glucose and xylose via novel mechanisms in engineered *Escherichia coli*. *Metab Eng*. 2015;30:141–8.
54. Mohamed ET, Mundhada H, Landberg J, Cann I, Mackie RI, Nielsen AT, Herrgard MJ, Feist AM. Generation of an *E. coli* platform strain for improved sucrose utilization using adaptive laboratory evolution. *Microb Cell Fact*. 2019;18:116.
55. Lane S, Xu H, Oh EJ, Kim H, Lesmana A, Jeong D, Zhang G, Tsai CS, Jin YS, Kim SR. Glucose repression can be alleviated by reducing glucose phosphorylation rate in *Saccharomyces cerevisiae*. *Sci Rep*. 2018;8:2613.
56. Chappell TC, Nair NU. Co-utilization of hexoses by a microconsortium of sugar-specific *E. coli* strains. *Biotechnol Bioeng*. 2017;114:2309–18.
57. Borirak O, Bekker M, Hellingwerf KJ. Molecular physiology of the dynamic regulation of carbon catabolite repression in *Escherichia coli*. *Microbiology*. 2014;160:1214–23.
58. Liu H, Valdehuesa KN, Nisola GM, Ramos KR, Chung WJ. High yield production of D-xylonic acid from D-xylose using engineered *Escherichia coli*. *Bioresour Technol*. 2012;115:244–8.
59. Toivari M, Nygard Y, Kumpulainen EP, Vehkomaki ML, Bencina M, Valkonen M, Maaheimo H, Andberg M, Koivuola A, Ruohonen L, et al. Metabolic engineering of *Saccharomyces cerevisiae* for bioconversion of D-xylose to D-xylonate. *Metab Eng*. 2012;14:427–36.
60. Radek A, Krumbach K, Gatgens J, Wendisch VF, Wiechert W, Bott M, Noack S, Marienhagen J. Engineering of *Corynebacterium glutamicum*

- for minimized carbon loss during utilization of D-xylose containing substrates. *J Biotechnol*. 2014;192:156–60.
61. Jiang Y, Liu W, Cheng T, Cao Y, Zhang R, Xian M. Characterization of D-xylonate dehydratase YjhG from *Escherichia coli*. *Bioengineered*. 2015;6:227–32.
  62. Bhaskar V, Kumar M, Manicka S, Tripathi S, Venkatraman A, Krishnaswamy S. Identification of biochemical and putative biological role of a xenolog from *Escherichia coli* using structural analysis. *Proteins*. 2011;79:1132–42.
  63. Cabulong RB, Lee WK, Banares AB, Ramos KRM, Nisola GM, Valdehuesa KNG, Chung WJ. Engineering *Escherichia coli* for glycolic acid production from D-xylose through the Dahms pathway and glyoxylate bypass. *Appl Microbiol Biotechnol*. 2018;102:2179–89.
  64. Baldoma L, Aguilar J. Involvement of lactaldehyde dehydrogenase in several metabolic pathways of *Escherichia coli* K12. *J Biol Chem*. 1987;262:13991–6.
  65. Liu M, Ding Y, Xian M, Zhao G. Metabolic engineering of a xylose pathway for biotechnological production of glycolate in *Escherichia coli*. *Microb Cell Fact*. 2018;17:51.
  66. Rodriguez GM, Atsumi S. Isobutyraldehyde production from *Escherichia coli* by removing aldehyde reductase activity. *Microb Cell Fact*. 2012;11:90.
  67. Chae TU, Choi SY, Ryu JY, Lee SY. Production of ethylene glycol from xylose by metabolically engineered *Escherichia coli*. *AIChE J*. 2018;64:4193–200.
  68. Jarboe LR. YqhD: a broad-substrate range aldehyde reductase with various applications in production of biorenewable fuels and chemicals. *Appl Microbiol Biotechnol*. 2011;89:249–57.
  69. Cabulong RB, Valdehuesa KN, Ramos KR, Nisola GM, Lee WK, Lee CR, Chung WJ. Enhanced yield of ethylene glycol production from D-xylose by pathway optimization in *Escherichia coli*. *Enzyme Microb Technol*. 2017;97:11–20.
  70. Wang Y, Xian M, Feng X, Liu M, Zhao G. Biosynthesis of ethylene glycol from D-xylose in recombinant *Escherichia coli*. *Bioengineered*. 2018;9:233–41.
  71. Alexeeva S, Hellingwerf KJ, Teixeira de Mattos MJ: requirement of ArcA for redox regulation in *Escherichia coli* under microaerobic but not anaerobic or aerobic conditions. *J Bacteriol*. 2003;185:204–9.
  72. Dugar D, Stephanopoulos G. Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nat Biotechnol*. 2011;29:1074–8.
  73. Salusjarvi L, Toivari M, Vehkomaki ML, Koivistoinen O, Mojzita D, Niemela K, Penttila M, Ruohonen L. Production of ethylene glycol or glycolic acid from D-xylose in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*. 2017;101:8151–63.
  74. Rahman MM, Andberg M, Koivula A, Rouvinen J, Hakulinen N. The crystal structure of D-xylonate dehydratase reveals functional features of enzymes from the Ilv/ED dehydratase family. *Sci Rep*. 2018;8:865.
  75. Deng Y, Mao Y, Zhang X. Metabolic engineering of *E. coli* for efficient production of glycolic acid from glucose. *Biochem Eng J*. 2015;103:256–62.
  76. Zahoor A, Otten A, Wendisch VF. Metabolic engineering of *Corynebacterium glutamicum* for glycolate production. *J Biotechnol*. 2014;192:366–75.
  77. Koivistoinen OM, Kuivanen J, Barth D, Turkia H, Pitkanen JP, Penttila M, Richard P. Glycolic acid production in the engineered yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. *Microb Cell Fact*. 2013;12:82.
  78. Gancedo JM. Yeast carbon catabolite repression. *Microbiol Mol Biol Rev*. 1998;62:334–61.
  79. Hagman A, Piskur J. A study on the fundamental mechanism and the evolutionary driving forces behind aerobic fermentation in yeast. *PLoS ONE*. 2015;10:e0116942.
  80. Bai W, Tai YS, Wang J, Wang J, Jambunathan P, Fox KJ, Zhang K. Engineering nonphosphorylative metabolism to synthesize mesaconate from lignocellulosic sugars in *Escherichia coli*. *Metab Eng*. 2016;38:285–92.
  81. Tai YS, Xiong M, Jambunathan P, Wang J, Wang J, Stapleton C, Zhang K. Engineering nonphosphorylative metabolism to generate lignocellulose-derived products. *Nat Chem Biol*. 2016;12:247–53.
  82. Wang J, Wang J, Tai YS, Zhang Q, Bai W, Zhang K. Rerouting carbon flux for optimized biosynthesis of mesaconate in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2018;102:7377–88.
  83. Wang J, Wu Y, Sun X, Yuan Q, Yan Y. De novo biosynthesis of glutarate via alpha-keto acid carbon chain extension and decarboxylation pathway in *Escherichia coli*. *ACS Synth Biol*. 2017;6:1922–30.
  84. Park SJ, Kim EY, Noh W, Park HM, Oh YH, Lee SH, Song BK, Jegal J, Lee SY. Metabolic engineering of *Escherichia coli* for the production of 5-aminovalerate and glutarate as C5 platform chemicals. *Metab Eng*. 2013;16:42–7.
  85. Wang J, Shen X, Lin Y, Chen Z, Yang Y, Yuan Q, Yan Y. Investigation of the synergetic effect of xylose metabolic pathways on the production of glutaric acid. *ACS Synth Biol*. 2018;7:24–9.
  86. Rohles CM, Gläser L, Kohstedt M, Giebelmann G, Pearson S, Del campo A, Becker J, Wittmann C. A Bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6,5 using metabolically engineered *Corynebacterium glutamicum*. *Green Chem*. 2018;20:4462.
  87. Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khan-durina J, Trawick JD, Osterhout RE, Stephen R, et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol*. 2011;7:445–52.
  88. Burgard A, Burk MJ, Osterhout R, Van Dien S, Yim H. Development of a commercial scale process for production of 1,4-butanediol from sugar. *Curr Opin Biotechnol*. 2016;42:118–25.
  89. Niu W, Molefe MN, Frost JW. Microbial synthesis of the energetic material precursor 1,2,4-butanetriol. *J Am Chem Soc*. 2003;125:12998–9.
  90. Valdehuesa KN, Liu H, Ramos KR, Park SJ, Nisola GM, Lee WK, Chung WJ. Direct biocconversion of D-xylose to 1,2,4-butanetriol in an engineered *Escherichia coli*. *Process Biochem*. 2014;49:25.
  91. Cao Y, Niu W, Guo J, Xian M, Liu H. Biotechnological production of 1,2,4-butanetriol: an efficient process to synthesize energetic material precursor from renewable biomass. *Sci Rep*. 2015;5:18149.
  92. Sun L, Yang F, Sun H, Zhu T, Li X, Li Y, Xu Z, Zhang Y. Synthetic pathway optimization for improved 1,2,4-butanetriol production. *J Ind Microbiol Biotechnol*. 2016;43:67–78.
  93. Wang X, Xu N, Hu S, Yang J, Gao Q, Xu S, Chen K, Ouyang P. d-1,2,4-Butanetriol production from renewable biomass with optimization of synthetic pathway in engineered *Escherichia coli*. *Bioresour Technol*. 2018;250:406–12.
  94. Gao H, Gao Y, Dong R. Enhanced biosynthesis of 3,4-dihydroxybutyric acid by engineered *Escherichia coli* in a dual-substrate system. *Bioresour Technol*. 2017;245:794–800.
  95. Valdehuesa KN, Lee WK, Ramos KR, Cabulong RB, Choi J, Liu H, Nisola GM, Chung WJ. Identification of aldehyde reductase catalyzing the terminal step for conversion of xylose to butanetriol in engineered *Escherichia coli*. *Bioprocess Biosyst Eng*. 2015;38:1761–72.
  96. Wang J, Shen X, Jain R, Wang J, Yuan Q, Yan Y. Establishing a novel biosynthetic pathway for the production of 3,4-dihydroxybutyric acid from xylose in *Escherichia coli*. *Metab Eng*. 2017;41:39–45.
  97. Bamba T, Yukawa T, Guirimand G, Inokuma K, Sasaki K, Hasunuma T, Kondo A. Production of 1,2,4-butanetriol from xylose by *Saccharomyces cerevisiae* through Fe metabolic engineering. *Metab Eng*. 2019;56:17–27.
  98. Olsson L, Hahn-Hägerdal B. Fermentative performance of bacteria and yeasts in lignocellulose hydrolysates. *Process Biochem*. 1993;28:249–57.
  99. Parawira W, Tekere M. Biotechnological strategies to overcome inhibitors in lignocellulose hydrolysates for ethanol production: review. *Crit Rev Biotechnol*. 2011;31:20–31.
  100. Uzarska MA, Nasta V, Weiler BD, Spantarg F, Ciofi-Baffoni S, Saviello MR, Gonnelli L, Mühlenhoff U, Banci L, Lill R. Mitochondrial Bol1 and Bol3 function as assembly factors for specific iron-sulfur proteins. *ELife*. 2016;5:e16673.
  101. Li X, Cai Z, Li Y, Zhang Y. Design and construction of a non-natural malate to 1,2,4-butanetriol pathway creates possibility to produce 1,2,4-butanetriol from glucose. *Sci Rep*. 2014;4:5541.
  102. Martin CH, Dhamankar H, Tseng HC, Sheppard MJ, Reisch CR, Prather KL. A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-gamma-butyrolactone. *Nat Commun*. 2013;4:1414.
  103. Dhamankar H, Tarasova Y, Martin CH, Prather KL. Engineering *E. coli* for the biosynthesis of 3-hydroxy-gamma-butyrolactone (3HBL) and 3,4-dihydroxybutyric acid (3,4-DHBA) as value-added chemicals from glucose as a sole carbon source. *Metab Eng*. 2014;25:72–81.

104. Salusjarvi L, Havukainen S, Koivistoinen O, Toivari M. Biotechnological production of glycolic acid and ethylene glycol: current state and perspectives. *Appl Microbiol Biotechnol*. 2019;103:2525–35.
105. Cam Y, Alkim C, Trichez D, Trebosc V, Vax A, Bartolo F, Besse P, Francois JM, Walther T. Engineering of a synthetic metabolic pathway for the assimilation of (d)-Xylose into value-added chemicals. *ACS Synth Biol*. 2016;5:607–18.
106. Asipu A, Hayward BE, O'Reilly J, Bonthron DT. Properties of normal and mutant recombinant human ketohexokinases and implications for the pathogenesis of essential fructosuria. *Diabetes*. 2003;52:2426–32.
107. Pereira B, Li ZJ, De Mey M, Lim CG, Zhang H, Hoeltgen C, Stephanopoulos G. Efficient utilization of pentoses for bioproduction of the renewable two-carbon compounds ethylene glycol and glycolate. *Metab Eng*. 2016;34:80–7.
108. Sanchez JC, Gimenez R, Schneider A, Fessner WD, Baldoma L, Aguilar J, Badia J. Activation of a cryptic gene encoding a kinase for L-xylulose opens a new pathway for the utilization of L-xylose by *Escherichia coli*. *J Biol Chem*. 1994;269:29665–9.
109. Alkim C, Trichez D, Cam Y, Spina L, Francois JM, Walther T. The synthetic xylulose-1 phosphate pathway increases production of glycolic acid from xylose-rich sugar mixtures. *Biotechnol Biofuels*. 2016;9:201.
110. Alkim C, Cam Y, Trichez D, Auriol C, Spina L, Vax A, Bartolo F, Besse P, Francois JM, Walther T. Optimization of ethylene glycol production from (D)-xylose via a synthetic pathway implemented in *Escherichia coli*. *Microb Cell Fact*. 2015;14:127.
111. Pellicer MT, Badia J, Aguilar J, Baldoma L. *glc* locus of *Escherichia coli*: characterization of genes encoding the subunits of glycolate oxidase and the *glc* regulator protein. *J Bacteriol*. 1996;178:2051–9.
112. Chomvong K, Bauer S, Benjamin DI, Li X, Nomura DK, Cate JH. Bypassing the pentose phosphate pathway: towards modular utilization of xylose. *PLoS ONE*. 2016;11:e0158111.
113. Träff KL, Otero Cordero RR, van Zyl WH, Hahn-Hägerdal B. Deletion of the GRE3 aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xylA* and *XKS1* genes. *Appl Environ Microbiol*. 2001;67:5668–74.
114. Uranukul B, Woolston BM, Fink GR, Stephanopoulos G. Biosynthesis of monoethylene glycol in *Saccharomyces cerevisiae* utilizing native glycolytic enzymes. *Metab Eng*. 2019;51:20–31.
115. Pereira B, Zhang H, De MM, Lim CG, Li ZJ, Stephanopoulos G. Engineering a novel biosynthetic pathway in *Escherichia coli* for production of renewable ethylene glycol. *Biotechnol Bioeng*. 2016;113:376–83.
116. Kalhan SC, Hanson RW. Resurgence of serine: an often neglected but indispensable amino acid. *J Biol Chem*. 2012;287:19786–91.
117. Manchado M, Michan C, Pueyo C. Hydrogen peroxide activates the SoxRS regulon in vivo. *J Bacteriol*. 2000;182:6842–4.
118. Uhl L, Dukan S. Hydrogen peroxide induced cell death: the major defences relative roles and consequences in *E. coli*. *PLoS ONE*. 2016;11:e0159706.
119. Gu P, Yang F, Su T, Li F, Li Y, Qi Q. Construction of an L-serine producing *Escherichia coli* via metabolic engineering. *J Ind Microbiol Biotechnol*. 2014;41:1443–50.
120. Chen Z, Huang J, Wu Y, Liu D. Metabolic engineering of *Corynebacterium glutamicum* for the de novo production of ethylene glycol from glucose. *Metab Eng*. 2016;33:12–8.
121. Spaans SK, Weusthuis RA, van der Oost J, Kengen SW. NADPH-generating systems in bacteria and archaea. *Front Microbiol*. 2015;6:742.
122. Bogorad IW, Lin TS, Liao JC. Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature*. 2013;502:693–7.
123. Lachaux C, Frazao CR, Krauser F, Morin N, Walther T, Francois J. A new synthetic pathway for the bioproduction of glycolic acid from lignocellulosic sugars aimed at maximal carbon conservation. *Front Bioeng. Biotechnol*. 2019;7:359.
124. Liu D, Evans T, Zhang F. Applications and advances of metabolite biosensors for metabolic engineering. *Metab Eng*. 2015;31:35–43.
125. Gupta A, Reizman IM, Reisch CR, Prather KL. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol*. 2017;35:273–9.
126. Yu H, Liao JC. A modified serine cycle in *Escherichia coli* converts methanol and CO<sub>2</sub> to two-carbon compounds. *Nat Commun*. 2018;9:3992.
127. Liao JC, Mi L, Pontrelli S, Luo S. Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nat Rev Microbiol*. 2016;14:288–304.
128. Ahsanul IM, Hadadi N, Ataman M, Hatzimanikatis V, Stephanopoulos G. Exploring biochemical pathways for mono-ethylene glycol (MEG) synthesis from synthesis gas. *Metab Eng*. 2017;41:173–81.
129. Walther T, Topham CM, Irague R, Auriol C, Baylac A, Cordier H, Dressaire C, Lozano-Huguet L, Tarrat N, Martineau N, et al. Construction of a synthetic metabolic pathway for biosynthesis of the non-natural methionine precursor 2,4-dihydroxybutyric acid. *Nat Commun*. 2017;8:15828.
130. Muller MM, Allison JR, Hongdilokkul N, Gaillon L, Kast P, van Gunsteren WF, Marliere P, Hilvert D. Directed evolution of a model primordial enzyme provides insights into the development of the genetic code. *PLoS Genet*. 2013;9:e1003187.
131. Chiu FWY, Stavrakis S. High-throughput droplet-based microfluidics for directed evolution of enzymes. *Electrophoresis*. 2019;40(21):2860–72.
132. Bouzou M, Perret A, Loreau O, Delmas V, Perchat N, Weissenbach J, Taran F, Marliere P. A synthetic alternative to canonical one-carbon metabolism. *ACS Synth Biol*. 2017;6:1520–33.
133. Wehrs M, Tanjore D, Eng T, Lievens J, Pray TR, Mukhopadhyay A. Engineering robust production microbes for large-scale cultivation. *Trends Microbiol*. 2019;27:524–37.
134. Deparis Q, Claes A, Foulquie-Moreno MR, Thevelein JM. Engineering tolerance to industrially relevant stress factors in yeast cell factories. *FEMS Yeast Res*. 2017;17:fox036.
135. Gong Z, Nielsen J, Zhou YJ. Engineering robustness of microbial cell factories. *Biotechnol J*. 2017;12:1700014.
136. Matsushika A, Inoue H, Kodaki T, Sawayama S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: current state and perspectives. *Appl Microbiol Biotechnol*. 2009;84:37–53.
137. Toivari MH, Salusjarvi L, Ruohonen L, Penttila M. Endogenous xylose pathway in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2004;70:3681–6.
138. Di Luccio E, Petschacher B, Voegtli J, Chou HT, Stahlberg H, Nidetzky B, Wilson DK. Structural and kinetic studies of induced fit in xylulose kinase from *Escherichia coli*. *J Mol Biol*. 2007;365:783–98.
139. Toivari MH, Aristidou A, Ruohonen L, Penttila M. Conversion of xylose to ethanol by recombinant *Saccharomyces cerevisiae*: importance of xylulokinase (XKS1) and oxygen availability. *Metab Eng*. 2001;3:236–49.
140. vanKuyk PA, de Groot MJ, Ruijter GJ, de Vries RP, Visser J. The *Aspergillus niger* D-xylulose kinase gene is co-expressed with genes encoding arabinan degrading enzymes, and is essential for growth on D-xylose and L-arabinose. *Eur J Biochem*. 2001;268:5414–23.
141. Jin YS, Jones S, Shi NQ, Jeffries TW. Molecular cloning of XYL3 (D-xylulokinase) from *Pichia stipitis* and characterization of its physiological function. *Appl Environ Microbiol*. 2002;68:1232–9.
142. Mojzita D, Penttila M, Richard P. Identification of an L-arabinose reductase gene in *Aspergillus niger* and its role in L-arabinose catabolism. *J Biol Chem*. 2010;285:23622–8.
143. Richard P, Londesborough J, Putkonen M, Kalkkinen N, Penttila M. Cloning and expression of a fungal L-arabinitol 4-dehydrogenase gene. *J Biol Chem*. 2001;276:40631–7.
144. Metz B, Mojzita D, Herold S, Kubicek CP, Richard P, Seiboth B. A novel L-xylulose reductase essential for L-arabinose catabolism in *Trichoderma reesei*. *Biochemistry*. 2013;52:2453–60.
145. Mojzita D, Vuoristo K, Koivistoinen OM, Penttila M, Richard P. The 'true' L-xylulose reductase of filamentous fungi identified in *Aspergillus niger*. *FEBS Lett*. 2010;584:3540–4.
146. Verho R, Putkonen M, Londesborough J, Penttila M, Richard P. A novel NADH-linked L-xylulose reductase in the L-arabinose catabolic pathway of yeast. *J Biol Chem*. 2004;279:14746–51.
147. Quong MW, Miyada CG, Switchenko AC, Goodman TC. Identification, purification, and characterization of a D-arabinitol-specific dehydrogenase from *Candida tropicalis*. *Biochem Biophys Res Commun*. 1993;196:1323–9.
148. Singh C, Glaab E, Linster CL. Molecular identification of D-Ribulokinase in budding yeast and mammals. *J Biol Chem*. 2017;292:1005–28.
149. Rosenfeld SA, Stevis PE, Ho NW. Cloning and characterization of the xyl genes from *Escherichia coli*. *Mol Gen Genet*. 1984;194:410–5.

150. Xu Z, Li S, Feng X, Liang J, Xu H. L-Arabinose isomerase and its use for biotechnological production of rare sugars. *Appl Microbiol Biotechnol*. 2014;98:8869–78.
151. LeBlanc DJ, Mortlock RP. Metabolism of D-arabinose: origin of a D-ribulokinase activity in *Escherichia coli*. *J Bacteriol*. 1971;106:82–9.
152. Elsinghorst EA, Mortlock RP. Molecular cloning of the *Escherichia coli* B L-fucose-D-arabinose gene cluster. *J Bacteriol*. 1994;176:7223–32.
153. Shi R, Pineda M, Ajamian E, Cui Q, Matte A, Cygler M. Structure of L-xylulose-5-Phosphate 3-epimerase (UlaE) from the anaerobic L-ascorbate utilization pathway of *Escherichia coli*: identification of a novel phosphate binding motif within a TIM barrel fold. *J Bacteriol*. 2008;190:8137–44.
154. Luo Y, Samuel J, Mosimann SC, Lee JE, Tanner ME, Strynadka NC. The structure of L-ribulose-5-phosphate 4-epimerase: an aldolase-like platform for epimerization. *Biochemistry*. 2001;40:14763–71.
155. Stephens C, Christen B, Fuchs T, Sundaram V, Watanabe K, Jenal U. Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. *J Bacteriol*. 2007;189:2181–5.
156. Brouns SJ, Walther J, Snijders AP, van de Werken HJ, Willemen HL, Worm P, de Vos MG, Andersson A, Lundgren M, Mazon HF, et al. Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. *J Biol Chem*. 2006;281:27378–88.
157. Watanabe S, Kodaki T, Makino K. A novel alpha-ketoglutaric semialdehyde dehydrogenase: evolutionary insight into an alternative pathway of bacterial L-arabinose metabolism. *J Biol Chem*. 2006;281:28876–88.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

