LETTER

doi:10.1038/nature12536

Microbial production of short-chain alkanes

Yong Jun Choi[1] & Sang Yup Lee[1,2]

Increasing concerns about limited fossil fuels and global environmental problems have focused attention on the need to develop sustainable biofuels from renewable resources. Although microbial production of diesel has been reported, production of another much in demand transport fuel, petrol (gasoline), has not yet been demonstrated. Here we report the development of platform *Escherichia coli* strains that are capable of producing short-chain alkanes (SCAs; petrol), free fatty acids (FFAs), fatty esters and fatty alcohols through the fatty acyl (acyl carrier protein (ACP)) to fatty acid to fatty acyl-CoA pathway. First, the β-oxidation pathway was blocked by deleting the *fadE* gene to prevent the degradation of fatty acyl-CoAs generated *in vivo*. To increase the formation of short-chain fatty acids suitable for subsequent conversion to SCAs *in vivo*, the activity of 3-oxoacyl-ACP synthase (FabH)[1], which is inhibited by unsaturated fatty acyl-ACPs[2], was enhanced to promote the initiation of fatty acid biosynthesis by deleting the *fadR* gene; deletion of the *fadR* gene prevents upregulation of the *fabA* and *fabB* genes responsible for unsaturated fatty acids biosynthesis[3]. A modified thioesterase[4] was used to convert short-chain fatty acyl-ACPs to the corresponding FFAs, which were then converted to SCAs by the sequential reactions of *E. coli* fatty acyl-CoA synthetase, *Clostridium acetobutylicum* fatty acyl-CoA reductase and *Arabidopsis thaliana* fatty aldehyde decarbonylase. The final engineered strain produced up to 580.8 mg l^{-1} of SCAs consisting of nonane (327.8 mg l^{-1}), dodecane (136.5 mg l^{-1}), tridecane (64.8 mg l^{-1}), 2-methyl-dodecane (42.8 mg l^{-1}) and tetradecane (8.9 mg l^{-1}), together with small amounts of other hydrocarbons. Furthermore, this platform strain could produce short-chain FFAs using a *fadD*-deleted strain, and short-chain fatty esters by introducing the *Acinetobacter* sp. ADP1 wax ester synthase (*atfA*)[5] and the *E. coli* mutant alcohol dehydrogenase ($adhE^{mut}$)[6].

Bio-based sustainable production of fuels has been attracting increasing interest for our sustainable future[7]. Hydrocarbon, such as alkane or alkene, is of particular interest owing to its potential to be used as an advanced biofuel that is similar to the petro-based fuels currently in use and superior to other biofuels in many aspects, including its high energy content (for example, it has a 30% higher energy content than ethanol)[8]. There have been a few reports on the bio-based production of C13–C17 long-chain hydrocarbons for substituting for diesel[9]. Microbial production of up to 300 mg l^{-1} of long-chain hydrocarbons, mainly pentadecane and heptadecane, was achieved by using an engineered *E. coli* strain harbouring a cyanobacterial alkane biosynthesis operon encoding acyl-ACP reductase and aldehyde decarbonylase[9]. Another study also reported production of even or odd numbered long-chain alkanes in *E. coli* by the overexpression of the *Bacillus subtilis fabH* gene[10]. In these studies, hydrocarbons were produced by decarbonylation of fatty aldehydes, which are directly generated from fatty acyl-ACPs. More recently, long-chain alkanes were produced from fatty acids by using fatty acid reductase and aldehyde decarbonylase[11].

Petrol, a mixture of C4–C12 short-chain hydrocarbons (SCHCs)[12], is a liquid fuel primarily used in internal combustion engines. Although short-chain alcohols were produced to substitute for petrol[13,14], they are inferior to petrol in their fuel properties (Supplementary Table 1). Thus, it is of great interest to produce SCHCs directly that have the potential to be used directly as petrol[15]. However, there has been no report so far about the production of such SCHCs by microbial fermentation. This seems to be because most of the bacterial fatty acids identified are C14–C18 long-chain ones. Here we report the development of engineered *E. coli* strains capable of producing SCAs suitable for petrol by engineering fatty acid biosynthesis and degradation pathways. This was achieved, in a different way from previous studies on the production of long-chain hydrocarbons, by introducing a new pathway involving a mutant fatty acyl-ACP thioesterase, fatty acyl-CoA synthetase, fatty acyl-CoA reductase and fatty aldehyde decarbonylase into engineered *E. coli* supporting generation of short-chain fatty acyl-ACPs. The detailed strategy for the production of SCAs is described in Fig. 1 and Supplementary Fig. 1. This strategy also allows production of short-chain FFAs, fatty esters and fatty alcohols as described below.

In the production of fatty-acid-based biofuels, FFAs derived from fatty acyl-ACPs by thioesterases are important intermediate metabolites. To examine the performance of different thioesterases, the *fadD* gene was deleted in *E. coli* strain W3110 to prevent conversion of FFAs to fatty acyl-CoAs. Among three thioesterases encoded by the *E. coli tesB* gene[16], *E. coli* '*tesA* (a leaderless version of *tesA*) gene[17], and the *Umbellularia californica fatB* gene[18], the *fadD*-deleted W3110 strain expressing 'TesA was found to be the best, producing 313 mg l^{-1} (Fig. 2a) of mixed FFAs (mainly C16 and small amounts of C8, C10, C12 and C14; Fig. 2b).

Because 'TesA preferentially hydrolyses long-chain fatty acyl-ACPs[17], an engineered thioesterase capable of converting short-chain fatty acyl-ACPs to FFAs was needed. Because TesA with a L109P mutation showed hydrolytic activity on both short- and long-chain fatty acyl-ACPs[4], 'TesA was similarly engineered to make 'TesA(L109P). Recombinant *E. coli fadD*-deleted W3110 expressing 'TesA(L109P) was able to produce a FFA mixture of short carbon lengths; C16 FFAs decreased by 91%, whereas C14, C12 and C10 FFAs increased by 6.8-fold, 12.8-fold and 2.2-fold, respectively (Fig. 2b). The percentages of C12 and C14 FFAs produced by *fadD*-deleted W3110 harbouring 'TesA(L109P) were 19.5% and 69.9%, respectively (Fig. 2b). Many thioesterases having different substrate specificities and activities[19] can be similarly used. Other approaches recently reported can also be taken to produce short-chain FFAs[20,21] (see Supplementary Discussion).

For the production of SCAs, the *fadE* gene needs to be deleted to block β-oxidation (Fig. 1a). Thus, the GAS1 strain was constructed by deleting the *fadE* gene in W3110. As in the *fadD*-deleted W3110 strain, the *fadD* gene was also deleted in the GAS1 strain to allow production of FFAs. The *fadD*-deleted GAS1 strain expressing 'TesA(L109P) was also able to produce short-chain FFAs (Fig. 2c); there were some variations in the composition of FFAs but C14 was the most prevalent one as in *fadD*-deleted W3110 expressing 'TesA(L109P).

Formation of short-chain FFAs can be enhanced by promoting the initiation of fatty acid biosynthesis, that is, the formation of β-ketoacyl-ACP by the condensation of acetyl-CoA and malonyl-ACP by 3-ketoacyl-ACP synthase (FabH)[1,22]. The overexpression of the *fabH* gene indeed increased production of short-chain fatty acids; C14

[1]Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 program), BioProcess Engineering Research Center and Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea.
[2]Bioinformatics and BioProcess Engineering Research Center, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea.

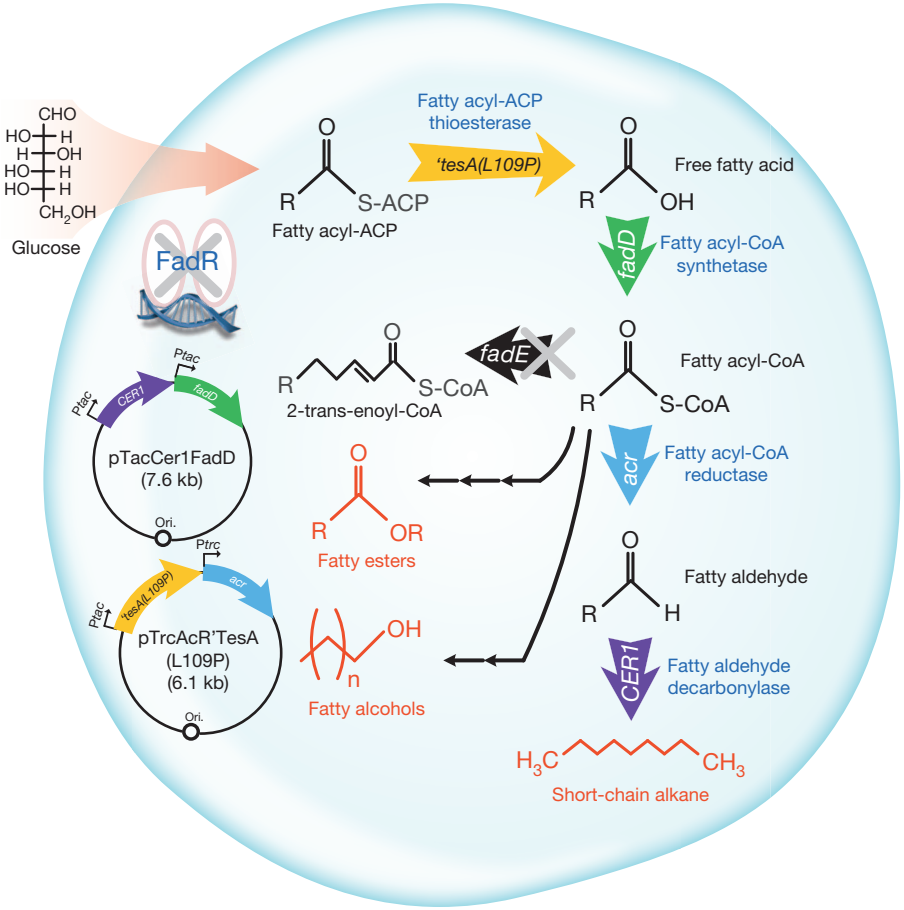

Figure 1 | Metabolic engineering of *E. coli* for the production of short-chain alkanes. The overall strategy for the production of short-chain alkanes (SCAs) is shown. The key enzymes are represented in the same colour. The genes knocked out are indicated by X. Glucose is converted into SCAs, fatty esters and fatty alcohols through the fatty acyl-ACP to fatty acid to fatty acyl-CoA pathway. The detailed metabolic engineering strategies used for the production of SCAs are described in Supplementary Fig. 1. The enzymes encoded by the genes shown are: *acr*, fatty acyl-CoA reductase; *CER1*, fatty aldehyde decarbonylase; *fadD*, fatty acyl-CoA synthetase; *fadE*, acyl-CoA dehydrogenase; *'tesA*(L109P), leaderless mutant thioesterase I. FadR is a global transcriptional regulator protein.

FFAs decreased by 77.4%, whereas C8 and C10 FFAs increased by 65% and 16%, respectively, in the *fadD*-deleted W3110 strain expressing 'TesA(L109P) and FabH. Moreover, the *fadD*-deleted GAS1 strain expressing 'TesA(L109P) and FabH produced C8 FFA by 5.7-fold more than the *fadD*-deleted W3110 strain expressing 'TesA(L109P) (Supplementary Fig. 2). FabH is known to be inhibited by the unsaturated fatty acyl-ACPs[2]. Because FadR positively regulates unsaturated fatty acid biosynthesis by upregulating the β-hydroxyacyl-ACP dehydratase and 3-ketoacyl-ACP synthase I operon (*fabAB*)[3], the *fadR* gene was also deleted in GAS1 to make the GAS2 strain. The *fadD*-deleted GAS2 strain expressing 'TesA(L109P) produced 2.6-fold and 1.6-fold more C10 and C12 FFAs and 64.6% lower C14 FFAs compared with the *fadD*-deleted GAS1 strain expressing 'TesA(L109P) (Fig. 2c, d). Thus, either overexpression of the *fabH* gene or deletion of the *fadR* gene results in the enhanced production of short-chain FFAs. Combining the two approaches of *fadR* deletion and *fabH* amplification did not further improve the production of short-chain FFAs mainly because of growth retardation. The importance of FabH in increasing short-chain fatty acids was also evaluated by knockdown experiments using synthetic RNA (sRNA)[23]. The knockdown of the *fabH* gene decreased the titre of short-chain fatty acids and increased the titre of long-chain fatty acids; C10 FFA decreased by 94%, whereas C14 FFAs increased by twofold in the *fadD*-deleted GAS2 strain (Supplementary Fig. 3). Thus, short-chain FFAs could be successfully produced in *E. coli* by deleting the *fadD, fadE* and *fadR* genes.

Next, conversion of short-chain FFAs to SCAs was performed in the GAS2 strain by amplifying the *fadD* gene and introducing the fatty acyl-CoA reductase and fatty aldehyde decarbonylase reactions (Fig. 1 and Supplementary Fig. 1). Even though *E. coli* is known to use only long-chain FFAs, recombinant *E. coli* overexpressing the *fadD* gene can utilize C8 and C10 FFAs[24]. This suggests that FadD can transfer CoA to both long- and short-chain FFAs. Thus, the GAS3 strain was constructed by replacing the native promoter of the *fadD* gene with the strong *trc* promoter in the chromosome of GAS2. Also, the FadD level was further increased by plasmid-based overexpression of *fadD*. Then the *Clostridium acetobutylicum acr* gene encoding a fatty acyl-CoA reductase for the reduction of fatty acyl-CoAs to fatty aldehydes and the *E. coli* codon-optimized *Arabidopsis thaliana CER1* gene encoding a fatty aldehyde decarbonylase[25] for the decarbonylation of fatty aldehydes to corresponding hydrocarbons were introduced by plasmid-based overexpression under the *trc* and *tac* promoters, respectively.

Fed-batch culture of the GAS3 strain harbouring pTacCer1FadD and pTrcAcR'TesA(L109P) resulted in the production of 396.5 mg l^{-1} of SCHCs, composed of 18.1 mg l^{-1} octene, 34.1 mg l^{-1} 2-octene, 217.0 mg l^{-1} nonane, 100.0 mg l^{-1} dodecane, 24.1 mg l^{-1} tridecane and 3.2 mg l^{-1} tetradecane (Fig. 3a, Supplementary Figs 4 and 5). To increase the titre of SCHCs further, the activity of fatty aldehyde decarbonylase (CER1) needed to be enhanced. The expression level of CER1 was found to be higher at 30 °C, compared with results obtained at other temperatures examined (Supplementary Fig. 10), therefore fed-batch fermentation was performed at 30 °C. Fed-batch culture of the GAS3 strain harbouring pTacCer1FadD and pTrcAcR'TesA(L109P) at 30 °C resulted in the production of 580.8 mg l^{-1} of hydrocarbons composed of 327.8 mg l^{-1} nonane, 136.5 mg l^{-1} dodecane, 64.8 mg l^{-1} tridecane,

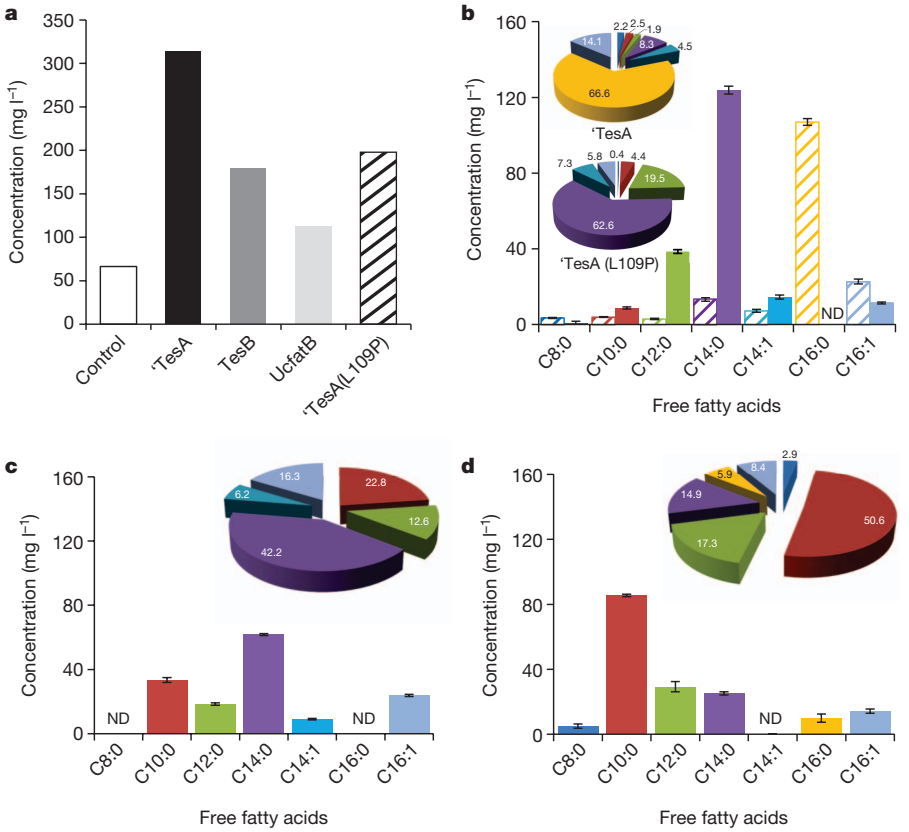

Figure 2 | Effects of three types of acyl-ACP thioesterases and 'TesA(L109P) on free fatty acid production. **a**, The total amounts of FFAs produced by the overexpression of 'TesA, TesB, UcfatB and 'TesA(L109P) in the *fadD*-deleted W3110 strain together with the control strain harbouring an empty vector. **b**, Distribution of FFAs produced in the *fadD*-deleted W3110 strain overexpressing 'TesA (hatched bar) and 'TesA(L109P) (solid bar). The percentage ratios of FFAs produced are also shown. **c**, **d**, The chain length distribution and percentage ratios of FFAs produced in *fadD*-deleted GAS1 (**c**) and *fadD*-deleted GAS2 (**d**) expressing 'TesA(L109P) are shown. Error bars represent the s.d. of experiments conducted in triplicate. Enzymes shown are: UcfatB, thioesterase of *Umbellularia californica*; TesB, thioesterase II of *E. coli*; 'TesA, the leaderless thioesterase I of *E. coli*; and 'TesA(L109P), a mutated leaderless thioesterase I of *E. coli*. ND, not detected.

42.8 mg l^{-1} 2-methyl-dodecane and 8.9 mg l^{-1} tetradecane (Fig. 3b, Supplementary Figs 6 and 7). Because *fadD*-deleted GAS2 harbouring 'TesA(L109P) produced C10 FFA dominantly, decarbonylation yielded mainly nonane (Fig. 3).

The *A. thaliana* fatty aldehyde decarbonylase (CER1) is known to be active towards long-chain fatty aldehydes (>30 carbons)[26]. However, our results suggest that this enzyme is also active towards short-chain fatty aldehydes (Supplementary Table 2). Furthermore, short-chain fatty alcohols as well as trace amounts of fatty aldehydes were also detected at the end of fermentation. Fatty alcohols were probably produced owing to the presence of inherent *E. coli* alcohol dehydrogenase, despite its low activity under aerobic conditions, and also due to the overexpression of fatty acyl-CoA reductase, which has been shown to convert fatty aldehydes to fatty alcohols[9,27]. Thus, the activity of aldehyde decarbonylase needs to be improved to further enhance the production of SCAs (see Supplementary Discussion). One possible way might be the overexpression of CER3, which was recently discovered to enhance the activity of CER1 (refs 28, 29).

Recently an *E. coli* strain integrated with a dynamic sensor–regulator system was developed for the production of C12–C20 long-chain fatty ethyl esters (FAEEs)[5]. Using GAS3 as a platform strain, production of short-chain FAEEs was attempted. For the aerobic production of

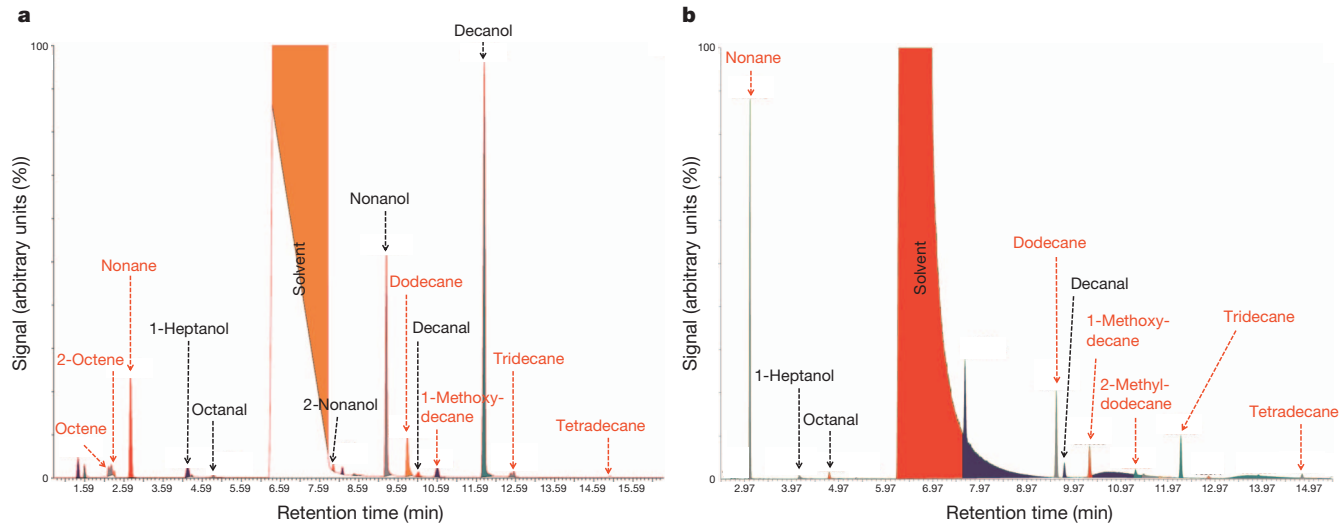

Figure 3 | GC–MS profile of fermentation products. The gas chromatography–mass spectrometry (GC–MS) profile of the hydrocarbon products obtained by fed-batch culture of the engineered GAS3 strain harbouring pTacCer1FadD and pTrcAcR'TesA(L109P) at 31 °C (**a**) and 30 °C (**b**) are shown. The ion spectra of these compounds are shown in Supplementary Figs 4–7.

ethanol, the *E. coli adhE*^{mut} gene[6] was expressed. For the esterification of fatty acyl-CoAs with ethanol, the *Acinetobacter* sp. ADP1 wax ester synthase (*atfA*) gene[5] was expressed. Batch culture of the GAS3 strain harbouring the plasmids pTacAdhE^{mut}FadD and pTrcAtfA*TesA(L109P) allowed production of 477.7 mg l^{-1} of short-chain FAEEs which consisted of 22.4 mg l^{-1} C10, 363.1 mg l^{-1} C12 and 92.2 mg l^{-1} C14 FAEEs (Supplementary Fig. 8). Thus, the platform strain developed here can be used for the production of short-chain FAEEs as well as petrol by using different final metabolic pathways. However, it was interesting to note that the composition of FAEEs produced was somewhat different from what would be expected from the results of SCAs production. As nonane was the major alkane produced, the expectation was to see C10 FAEE as the most abundant one rather than C12 FAEE. This discrepancy seems to be because of the substrate specificity of wax ester synthase, which has higher affinity towards long-chain fatty acyl-CoAs[17].

We have developed new platform *E. coli* strains, the GAS2 strain capable of producing short-chain FFAs, and the GAS3 strain capable of producing SCHCs (petrol), fatty esters and fatty alcohols. This was possible by establishing the corresponding metabolic pathways by engineering *E. coli* fatty acid biosynthesis and degradation pathways and employing an engineered thioesterase. Also, if desired, long-chain alkanes suitable for diesel can be produced by employing the same platform strain together with non-engineered thioesterase (Supplementary Fig. 9). This work will serve as a stepping stone for establishing bioprocesses for the production of short-chain fatty acid derived chemicals and fuels from renewable resources.

METHODS SUMMARY

Bacterial strains and plasmids. *E. coli* strains, plasmids and oligonucleotides used are listed in Supplementary Tables 3 and 4. Detailed procedures for the construction of strains are described in the Methods. All DNA manipulations were performed according to standard procedures[30]. All oligonucleotides were synthesized at GenoTech or Bioneer (Daejeon). Preparation of plasmids and DNA fragments was performed with Qiagen kits. All other chemicals used were of analytical grade and purchased from Sigma-Aldrich.

Culture condition and analysis. Recombinant *E. coli* strains were grown in MR medium (pH 6.8; see Methods) containing 10 g l^{-1} glucose and 3 g l^{-1} yeast extract at 31 °C and shaking at 220 r.p.m. Hydrocarbons were identified and quantified by gas chromatography–mass spectrometry (GC–MS) (Perkin Elmer Turbo Mass Clarus 600 coupled with a quadrupole mass selective detector on EI operated at 70 eV; see Methods for details). Retention times and fragmentation patterns were compared with GC–MS library database (NIST MS Search 2.0).

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 March; accepted 8 August 2013.
Published online 29 September 2013.

1. Han, L., Lobo, S. & Reynolds, K. A. Characterization of β-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J. Bacteriol.* **180,** 4481–4486 (1998).
2. Heath, R. J. & Rock, C. O. Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. *J. Biol. Chem.* **271,** 1833–1836 (1996).
3. Nunn, W. D., Giffin, K., Clark, D. & Cronan, J. E. Jr. Role for *fadR* in unsaturated fatty acid biosynthesis in *Escherichia coli*. *J. Bacteriol.* **154,** 554–560 (1983).
4. Lo, Y. C., Lin, S. C., Shaw, J. F. & Liaw, Y. C. Substrate specificities of *Escherichia coli* thioesterase I/protease I/lysophospholipase L1 are governed by its switch loop movement. *Biochemistry* **44,** 1971–1979 (2005).
5. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnol.* **30,** 354–359 (2012).
6. Holland-Staley, C. A., Lee, K., Clark, D. P. & Cunningham, P. R. Aerobic activity of *Escherichia coli* alcohol dehydrogenase is determined by a single amino acid. *J. Bacteriol.* **182,** 6049–6054 (2000).
7. Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* **488,** 320–328 (2012).
8. Lennen, R. M., Braden, D. J., West, R. A., Dumesic, J. A. & Pfleger, B. F. A process for microbial hydrocarbon synthesis: overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnol. Bioeng.* **106,** 193–202 (2010).
9. Schirmer, A., Rude, M. A., Li, X., Popova, E. & del Cardayre, S. B. Microbial biosynthesis of alkanes. *Science* **329,** 559–562 (2010).
10. Harger, M. *et al.* Expanding the product profile of a microbial alkane biosynthetic pathway. *ACS Synthet. Biol.* **2,** 59–62 (2013).
11. Howard, T. P. *et al.* Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **110,** 7636–7641 (2013).
12. Altin, O. & Eser, S. Carbon deposit formation from thermal stressing of petroleum fuels. *Am. Chem. Soc. Div. Fuel Chem.* **49,** 764–766 (2004).
13. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451,** 86–89 (2008).
14. Choi, Y. J., Park, J. H., Kim, T. Y. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for the production of 1-propanol. *Metab. Eng.* **14,** 477–486 (2012).
15. Gary, J. H. & Handwerk, G. E. *Petroleum Refining: Technology and Economics* 4th edn (Marcel Dekker, 2001).
16. Naggert, J. *et al.* Cloning, sequencing, and characterization of *Escherichia coli* thioesterase II. *J. Biol. Chem.* **266,** 11044–11050 (1991).
17. Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463,** 559–562 (2010).
18. Pollard, M. R., Anderson, L., Fan, C., Hawkins, D. J. & Davies, H. M. A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of *Umbellularia californica*. *Arch. Biochem. Biophys.* **284,** 306–312 (1991).
19. Jing, F. *et al.* Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochem.* **12,** 44 (2011).
20. Zheng, Y. *et al.* Boosting the free fatty acid synthesis of *Escherichia coli* by expression of a cytosolic *Acinetobacter baylyi* thioesterase. *Biotechnol. Biofuels* **5,** 76 (2012).
21. Torella, J. P. Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proc. Natl Acad. Sci. USA* **110,** 11290–11295 (2013).
22. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. Isolation and characterization of the β-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J. Biol. Chem.* **267,** 6807–6814 (1992).
23. Na, D. *et al.* Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nature Biotechnol.* **31,** 170–174 (2013).
24. Zhang, H., Wang, P. & Qi, Q. Molecular effect of FadD on the regulation and metabolism of fatty acid in *Escherichia coli*. *FEMS Microbiol. Lett.* **259,** 249–253 (2006).
25. Aarts, M. G., Keijzer, C. J., Stiekema, W. J. & Pereira, A. Molecular characterization of the *CER1* gene of *Arabidopsis* involved in epicuticular wax biosynthesis and pollen fertility. *Plant Cell* **7,** 2115–2127 (1995).
26. McNevin, J. P., Woodward, W., Hannoufa, A., Feldmann, K. A. & Lemieux, B. Isolation and characterization of eceriferum (*cer*) mutants induced by T-DNA insertions in *Arabidopsis thaliana*. *Genome* **36,** 610–618 (1993).
27. Reiser, S. & Somerville, C. Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase. *J. Bacteriol.* **179,** 2969–2975 (1997).
28. Bernard, A. *et al.* Reconstitution of plant alkane biosynhtesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell* **24,** 3106–3118 (2012).
29. Bourdenx, B. Overexpression of *Arabidopsis ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stress. *Plant Physiol.* **156,** 29–45 (2011).
30. Sambrook, J. R. D. *Molecular Cloning*: *A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank Y. H. Lee for her assistance in cloning work and S. J. Choi for performing the fermentation experiments for checking reproducibility. This work was supported by the Advanced Biomass Research and Development Center of Korea (ABC-2010-0029799) through the Global Frontier Research Program of the Ministry of Science, ICT and Future Planning (MSIP) through the National Research Foundation (NRF). Systems metabolic engineering work was supported by the Technology Development Program to Solve Climate Changes on Systems Metabolic Engineering for Biorefineries (NRF-2012-C1AAA001-2012M1A2A2026556) by MSIP through NRF.

Author Contributions S.Y.L. conceived and supervised the project. Y.J.C. performed all experiments and analysed the data. Y.J.C. and S.Y.L. wrote the manuscript together. Both authors approved the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.Y.L. (leesy@kaist.ac.kr).

METHODS

Construction of plasmids. For the construction of pTacCer1, the *A. thaliana CER1* gene was used. For better expression in *E. coli*, its codon usage was optimized to *E. coli* (synthesized at Bioneer and cloned into a vector to make plasmid pGEM-T-easy-optiCER1). The codon-optimized *CER1* gene was amplified with the primers CER1f and CER1r using pGEM-T-easy-optiCER1 as a template. The PCR product was digested with SacI and XbaI, and ligated with SacI/XbaI digested pTac15K, creating pTacCer1. For the construction of pTacCer1FadD, pTacFadD was first constructed to clone the *tac* promoter in front of the *fadD* gene. The *fadD* gene was amplified with the primers Fadf and Fadr using the genomic DNA of *E. coli* as a template. The PCR product was digested with EcoRI and KpnI, and ligated with EcoRI/KpnI digested pTac15K. The *E. coli fadD* gene tagged with *tac* promoter was amplified with the primers tacfadDf and tacfadDr using pTacFadD as a template and cloned into pTacCer1 digested with SphI to construct pTacCer1FadD.

The pTrcAcR'TesA was constructed in a similar manner as above. For the construction of pTrcAcR, the fatty acyl-CoA reductase gene was amplified by PCR with the primers AcRf and AcRr using the genomic DNA of *C. acetobutylicum* as a template. The PCR product was digested with KpnI and BamHI, and ligated with KpnI/BamHI digested pTrc99a. The resulting plasmid, pTrcAcR, contains the fatty acyl-CoA reductase gene under the control of the strong *trc* promoter. To make pTrcAcR'TesA, the leaderless version of the *E. coli tesA* gene ('*tesA*) was first cloned in pTac15K. The '*tesA* gene was amplified with the primers 'tesAf and 'tesAr using the genomic DNA of *E. coli* as a template. The PCR product was digested with EcoRI and SacI, and ligated with EcoRI/SacI digested pTac15K, resulting in pTac'TesA. The '*tesA* gene attached to the *tac* promoter was amplified by PCR with the primers tac'tesAf and tac'tesAr using pTac'tesA as a template. The PCR product was digested with BamHI and ligated with BamHI-digested pTrcAcR, creating pTrcAcR'TesA. In the case of 'tesA(L109P), the primers 'tesAf and mt'tesA2 were used to amplify a 357 base pair (bp) DNA fragment. The primer mt'tesA2 contains a single mutated base (T329C). The second 245-bp DNA fragment was amplified using the primers mt'tesA3 and and 'tesAr. The primer mt'tesA3 contains a single mutated base (A329G). The two DNA fragments were purified and mixed, and the complete 572-bp fragment was amplified by overlapping PCR using the primers 'tesAf and 'tesAr. The PCR product was digested with EcoRI and SacI, and ligated with EcoRI/SacI-digested pTac15K, resulting in pTac'TesA(L109P). The '*tesA*(L109P) gene attached to the *tac* promoter was amplified by PCR with the primers tac'tesAf and tac'tesAr using the pTac'TesA(L109P) as a template. The PCR product was digested with BamHI, and ligated with BamHI digested DNA fragment of pTrcAcR resulting in pTrcAcR'TesA(L109P). For the construction of pTac'TesA(L109P)FabH, the *fabH* gene was amplified by PCR with the primers fabHf and fabHr using the genomic DNA of *E. coli* as a template. In the same manner, sRNA specific to the *fabH* gene was amplified with primers sfabHf and sfabHr using anti-fabH-sRNA[23] as a template to construct the pTac'TesA(L109P)sFabH.

To make pTacTesB and pTacUcfatB, the *tesB* gene of *E. coli* and the *fatB* gene of *Umbellularia californica* were amplified with the primer pairs tesBf-tesBr and UcfatBf-UcfatBr, respectively. The PCR products were digested with EcoRI and SacI, and ligated with EcoRI/SacI-digested pTac15K to construct pTacTesB and pTacUcfatB, respectively. For the construction of pTrcAtfA, the *Acinetobacter* sp. ADP1 wax ester synthase gene (*atfA*) was amplified by PCR with the primers atfAf and atfAr using the genomic DNA of *Acinetobacter* sp. ADP1 as a template. The PCR product was digested with KpnI and BamHI, and ligated with KpnI/BamHI digested pTrc99a. The resulting plasmid, pTrcAtfA, contains the wax ester synthase gene under the control of the strong *trc* promoter. To make pTrcAtfA'TesA(L109P), the 'tesA(L109P) gene attached to the *tac* promoter was cloned into pTrcAtfA in the same way used for constructing pTrcAcR'TesA(L109P).

For the construction of pTacAdhE^{mut}, the mutant *E. coli* alcohol dehydrogenase gene was amplified by PCR with the primers adhEf and adhEr using the pTacDA_tac_adhE^{mut} plasmid[14] as a template. The PCR product was digested with EcoRI and SacI, and ligated with EcoRI/SacI digested pTac15K. The resulting plasmid, pTacAdhE^{mut}, contains the mutant alcohol dehydrogenase gene under the control of the strong *tac* promoter. To make pTacAdhE^{mut}FadD, the *fadD* gene attached to the *tac* promoter was amplified with the primers tacfadDf and tacfadDr using pTacFadD as a template, and was cloned into pTacAdhE^{mut} in the same way used for constructing pTacCer1FadD.

Genome engineering. Deletion of the *fadE* (acyl-CoA dehydrogenase), *fadR* (a global transcriptional regulator) and *fadD* (acyl-CoA synthetase) genes were performed using one-step inactivation method[31]. Replacement of the native promoters of the *fadD* gene with the *trc* promoter was performed by PCR-mediated λ-Red recombination method[32]. Detailed procedures for genome engineering are as follows. Recombinant *E. coli* harbouring pKD46 was cultivated at 30 °C, and the expression of λ-Red recombinase was induced by adding 10 mM L-arabinose. Then electrocompetent cells were prepared by the standard protocol[30]. For the deletion of the *fadE, fadR* and *fadD* genes, the DNA fragment containing the lox71 site, the chloramphenicol resistance gene and the lox66 site fused together was obtained by PCR with the primer pairs fadEKOf-fadEKOr, fadRKOf-fadRKOr and fadDKOf-fadDKOr, respectively, and using pMloxC, which contains the chloramphenicol resistance gene flanked by lox sequences (Cre recognition target), as a template. The PCR product was introduced by electroporation into *E. coli* expressing the λ-Red recombinase. The mutants, in which gene inactivation occurred by homologous recombination, were selected on the LB agar plate containing 34 μg ml^{-1} chloramphenicol and subsequently screened by direct colony PCR. To construct marker-free mutant strains, the chloramphenicol resistance gene was eliminated by using a helper plasmid, pJW168 (ref. 33), harbouring the Cre recombinase, ampicillin resistance gene and a temperature-sensitive replication origin. The expression of Cre recombinase was induced by isopropyl-β-D-thiogalactoside (IPTG). The chloramphenicol resistant mutants were transformed with pJW168, and ampicillin-resistant transformants were selected on LB agar plates containing 100 μg ml^{-1} ampicillin and 1 mM IPTG at 30 °C. Those colonies that lost chloramphenicol resistance were selected. Among them, positive colonies were cultivated in LB medium without antibiotic at 42 °C. Then they were examined for the loss of all antibiotic resistance markers by colony PCR and antibiotic resistance.

For the strong expression of the *fadD* gene, the native promoter of the *fadD* gene was replaced with the *trc* promoter. Substitution of the native promoter of the *fadD* gene by the *trc* promoter was performed by PCR mediated λ-Red recombination[32]. A 1,189 base pair DNA fragment containing the strong *trc* promoter, the lox71 site, the chloramphenicol resistance gene and the lox66 site fused together was obtained by PCR using the primers fadDpchf and fadDpchr. Plasmid pMtrc9 was used as a template. The PCR product was introduced by electroporation into *E. coli* harbouring pKD46 (ref. 31), expressing the λ-Red recombinase. The mutants in which promoter replacement occurred by homologous recombination were selected on the LB agar plate containing 34 μg ml^{-1} chloramphenicol, and subsequently screened by direct colony PCR. Replacement of the promoter of the *fadD* gene with the *trc* promoter was confirmed by direct colony PCR followed by sequencing. To construct marker-free mutant strains, removal of antibiotic marker was performed in the same manner as described for the chromosomal gene inactivation described above.

Culture condition. The MR medium (pH 6.8) contains per litre: 6.67 g KH_2PO_4, 4 g $(NH_4)_2HPO_4$, 0.8 g $MgSO_4{\cdot}7H_2O$, 0.8 g citric acid and 5 ml trace metal solution. The trace metal solution contains per litre of 0.5 M HCl: 10 g $FeSO_4{\cdot}7H_2O$, 2 g $CaCl_2$, 2.2 g $ZnSO_4{\cdot}7H_2O$, 0.5 g $MnSO_4{\cdot}4H_2O$, 1 g $CuSO_4{\cdot}5H_2O$, 0.1 g $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$ and 0.02 g $Na_2B_4O_7{\cdot}10H_2O$. Glucose and $MgSO_4{\cdot}7H_2O$ were sterilized separately. Seed cultures were prepared in 15 ml test tubes containing 5 ml LB medium at 31 °C overnight in a rotary shaker at 220 r.p.m. One ml of overnight culture was used to inoculate the 250 ml flask containing 100 ml of LB medium. Cells cultured for 12 h were used to inoculate a 6.6 l bioreactor (Bioflo 3000; New Brunswick Scientific) containing 2 l of MR medium containing 3 g l^{-1} of yeast extract and 20 g l^{-1} of glucose. Fermentation was carried out at 30 °C. For fed-batch fermentation, 100 ml of feeding solution (consisting of 400 g l^{-1} glucose and 14 g l^{-1} $MgSO_4{\cdot}7H_2O$) was added when the glucose concentration in the culture broth decreased below 1 g l^{-1}. The pH was controlled at 6.8 by the automatic feeding of 25% (v/v) NH_4OH. To fully recover volatile hydrocarbon products, octanol was used as a trapping solvent (see analytical methods section below). The dissolved oxygen concentration was maintained above 40% of air saturation by supplying air at 1 vvm (air volume/working volume/min) and by automatically controlling the agitation speed. For the expression of genes under *trc* and *tac* promoter, 1 mM IPTG (Sigma-Aldrich) was added at the attenuance of 600 nm ($D_{600\,nm}$; or OD_{600}) of 0.5. When necessary, 100 μg ml^{-1} ampicillin and 50 μg ml^{-1} kanamycin were added to the medium.

Analytical methods. Cell growth was monitored by measuring the absorbance at $D_{600\,nm}$ using an Ultrospec3000 spectrophotometer (Pharmacia Biotech). Glucose concentration was measured using a glucose analyser (model 2700 STAT; Yellow Springs Instrument). For the analysis of FFAs, 5 ml of liquid culture sample was acidified with 500 μl of acetic acid for pH adjustment, followed by the addition of 5 ml of ethylacetate. The culture tubes were vigorously vortexed for 5 min. The organic layer was taken out and ethylacetate was evaporated by blowing pure N_2 gas. The near dry FFAs were subjected to methanolysis with 15% (v/v) sulphuric acid in the presence of benzoic acid as an internal standard. The resulting fatty acid methyl esters (FAMEs) were analysed by gas chromatography (GC) as follows. The oven temperature was initially maintained at 80 °C for 5 min and ramped to 230 °C at 7.5 °C min^{-1}. Temperature was gradually increased at a rate of 10 °C min^{-1} up to 260 °C. One microlitre aliquot of FAMEs solution was injected into the gas chromatograph. The carrier gas flow rate was 2.0 ml min^{-1}, with a split ratio of 1:25. The retention time and fragmentation peaks were compared with standards to confirm the peaks.

Alkanes and hydrocarbons produced by fermentation were analysed as follows. First, it was thought that alkanes and hydrocarbons produced would be evaporated and also found in fermentation broth. Thus, both fermentation broth and octanol trapped products were analysed. The evaporated SCHCs were trapped with 100 ml

octanol using a cooling device (Supplementary Fig. 11). For the analysis of SCHCs trapped in octanol, the octanol solution containing dissolved SCHCs was directly injected into the electro-ionization (EI) GC–MS. For the analysis of SCHCs in the fermentation broth, 1 ml fermentation broth was extracted with 1 ml of ethylacetate and analysed by EI GC–MS. The GC–MS analyses were performed with Perkin Elmer Turbo Mass Clarus 600 coupled with a quadrupole mass selective detector on EI operated at 70 eV. The Perkin Elmer Elite 1 column (30 m length, 0.25 mm inner diameter, 0.25 μm film thickness) was used at 50:1 split ratio. The temperature of ion source was adjusted to 200 °C. The GC analyses were performed as follows. The oven temperature was initially maintained at 80 °C for 3 min. Temperature was gradually increased at a rate of 5 °C min^{-1} up to 270 °C. One μl of sample solution was injected. The retention times and fragmentation patterns were compared with GC–MS library database (NIST MS Search 2.0) for the identification of compounds. The concentrations of hydrocarbon compounds were measured by comparing the peak intensities of the samples and standards of known concentrations during the EI GC–MS analyses.

***In vitro* fatty aldehyde decarbonylase assay.** To test the activity of fatty aldehyde decarbonylase, the recombinant W3110 strain harbouring pTac15K or pTacCer1 was cultured in LB medium and 1 mM of IPTG was added at the $D_{600\,nm}$ of 0.5. Five millilitres of culture sample was collected by centrifugation at 16,600*g* for 10 min. The cell pellet was resuspended in 200 μl of Tris-HCl (pH 7.2) containing 0.2 mM of dithiothreitol (DTT) and centrifuged again. The cell pellet was lysed by sonication (Bioreutor KRB-01, Bio Medical Science) and its supernatant was collected. The assay mixture (1 ml) contained 100 mM Tris-HCl (pH 7.5), 0.1 mM DTT, 0.5 mM NADPH, 100 mM decanaldehyde, 100 μl of cell extract. The change in absorbance was measured at 340 nm every 5 s. The extinction coefficient of NADPH used was 6.23 mM^{-1} cm^{-1}. Enzyme activity was determined by using a spectrophotometer (GeneZuant 1300; GE Healthcare). One unit (U) of enzyme activity was defined as an amount of enzyme needed to catalyse 1 μmol of fatty aldehyde per minute into alkane. The specific activity was defined as unit per mg of protein.

31. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97,** 6640–6645 (2000).
32. Yuan, L. Z., Rouviere, P. E., Larossa, R. A. & Suh, W. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metab. Eng.* **8,** 79–90 (2006).
33. Palmeros, B. *et al.* A family of removable cassettes designed to obtain antibiotic-resistance-free genomic modifications of *Escherichia coli* and other bacteria. *Gene* **247,** 255–264 (2000).