Food Materials Research

ARTICLE

Open Access

https://doi.org/10.48130/FMR-2022-0015
Food Materials Research **2022**, 2:15

A multi-layer genome mining and phylogenomic analysis to construct efficient and autonomous efflux system for medium chain fatty acids

Hu Peng[1], Lin Zhou[1], Xuguo Duan[2], Zhaojun Wang[1], Zhaoshi Wang[1], Mo Xia[1], Mingsheng Dong[1], and Junjun Wu[3,1*]

[1] *College of Food Science and Technology, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China*
[2] *Department of Food Science and Technology, College of Light Industry and Food Engineering, Nanjing Forestry University, Nanjing 210037, China*
[3] *School of Biotechnology, Jiangnan University, 1800 Lihu Road, Wuxi 214122, China*
* Corresponding author, E-mail: wujunjun@jiangnan.edu.cn

Abstract

Medium-chain fatty acids (MCFAs) are important components for food, pharmaceutical and fuel industries. Nevertheless, engineering microorganisms to produce MCFAs often induces toxicity and stresses towards host strains, which could be alleviated via accelerating the export of MCFAs from cells. However, current secretory systems are inefficient and require inducible promoters. Here, a multi-layer genome mining and phylogenomic analysis was developed to identify efficient efflux transporters. Firstly, based on the genomic mining of 397 strains throughout various representative species, the evolutionary history of efflux transporters was recapitulated, and further experimental analysis revealed that *acrE* from *Citrobacter* exhibited the best performance. Secondly, according to the further mining of 797 *Citrobacter* genomes and 1084 *Escherichia* genomes, a detailed phylogenomic analysis of efflux transporter-centric genomic vicinities was performed. This led to the identification of efficient efflux pump combination *acrE* and *acrF*. These efflux pumps were then combined with the quorum-sensing circuit from *Enterococcus faecalis* to regulate MCFA efflux in an autonomous manner, which achieved a 4.9-fold boost in MCFA production and firstly demonstrated the efficient and autonomous efflux pump specially for MCFAs. The integrative omics technologies described here are enabling the utilization of the increasingly large database and the effective mining of target gene diversities.

Citation: Peng H, Zhou L, Duan X, Wang Z, Wang Z, et al. 2022. A multi-layer genome mining and phylogenomic analysis to construct efficient and autonomous efflux system for medium chain fatty acids. *Food Materials Research* 2:15 https://doi.org/10.48130/FMR-2022-0015

INTRODUCTION

Medium chain fatty acids (MCFAs) represent molecules with one carboxylic acid bound to a medium alkyl chain (C6-C10), constituting important food constituents and essential feedstocks of biofuels or oleo-chemical industries. Compared to their long-chain counterparts with a long alkyl chain (C12 or more), the shorter chain lengths confer MCFAs with significant characters such as higher carbon conversion yield and lower freezing/cloudy point, suggesting their potential as substitutes for fossil fuels[1,2]. Furthermore, MCFAs exhibit other unique physicochemical properties, for instance, little tendency to deposit as body fat, weight control benefits, antimicrobial effects, immune-modulating effects, and improving clinical symptoms, constituting their unique advantages as food constituents or even chemotherapeutic agents[3,4].

Currently, natural source extraction or petrol-based synthesis are the main processes by which to obtain MCFAs. In nature, MCFAs present only in coconut and palm kernel with low concentrations, ranging from 7.9% to 15% of total fatty acids. Due to the seasonal/regional limitations, long breeding cycles and low concentrations, plant extraction is not amendable for industrialization[2,5,6]. Besides, the growing scarcity of fossil fuels and environmental anxiety of rising petrol-based manufacturing costs, and owing to food safety considerations, this manner is unfavorable in the food and pharmaceutical industries[1]. Accordingly, efficient, scalable and sustainable procedures to obtain MCFAs from cheap and renewable resources are required as an impetus towards MCFAs more widespread adoption.

Numerous advantages inherent to microbial conversion procedures, for instances, rapid replication speeds, the capability of utilizing renewable feedstocks or acting during mile pressures and temperatures, and easy realization of large-scale fermentation[2,5,7–9], means it is an attractive alternative for fatty acid production. Previous pioneering studies have firstly demonstrated efficient MCFA production at 1.1−1.3 g/L via utilizing reversal of β-oxidation cycle (r-BOX) associate with leveraging thioesterases[10–12]. A series of our studies achieved the highest titer (3.8−15.6 g/L) reported to date through identifying pathway bottlenecks[13], satisfying redox cofactor requirement[14], or constructing artificial micro-aerobic metabolism[15]. All of these results have illustrated that *E. coli*-based bioconversion so far presents a good chassis to produce MCFAs.

Despite the apparent capability for microbial production of MCFAs, product toxicity is a common issue in strain engineering, which would result in physiological perturbations including reducing cell viability and membrane integrity, inducing membrane stress responses and losing proton motive

force[16–18]. One promising strategy to abate this problem is improving the transport speed of MCFAs from cells, and our previous study has demonstrated that expressing transporter from *E. coli* responsible for accelerating MCFA export could improve the production of MCFAs[19]. However, current secreting system is constructed based on the endogenous transporters derived from *E. coli*, which is inefficient and requires inducible promoters for conducting the transport function. This is still incompatible with large-scale production.

The rapid buildup of genomic information has revealed that metabolic abilities of virtually all organisms are vastly underappreciated[20,21], and sequenced microbial genomes may contain numerous efflux pumps and offer a vastly unexplored resource for mining novel pumps. Here, in order to efficiently mine genomes during large genomic datasets, a multi-layer genome mining and phylogenomic analysis was developed to screen a library of uncharacterized heterologous pumps among over 2000 microbial genomes. This led to the identification of efficient efflux pump combination *acrE* and *acrF* from *Citrobacter tructae*. When combining with the quorum-sensing (QS) circuit from *Enterococcus faecalis*, MCFA efflux presented as an autonomous behavior without inducer supplementation or human supervision, and this achieved a 4.9-fold boost in MCFA production.

MATERIALS AND METHODS

General procedures

E. coli JM109 and BL21 (DE3) were used for all molecular experiments and bio-catalysis, respectively. The plasmids of pACYCDuet-1, pCDFDuet-1, and pETDuet-1 (Novagen, Darmstadt, Germany) used in this study required the supplementation of 20 μg/mL of chloramphenicol, 40 μg/mL of streptomycin, 100 μg/mL of ampicillin, respectively, to maintain in the same cell. T4 DNA ligase, FastDigest restriction enzymes, and Phusion DHA polymerase (Novagen, Darmstadt, Germany) were employed to perform standard molecular manipulations. UV/vis spectrophotometer (UVmini-1240, Shimadzu, kyoto, Japan) was utilized to measure cell growth (OD_{600}).

General phylogenomic reconstruction of MCFA transporter families

Genomes for general phylogenomic analysis of MCFA transporter families such as AcrE, MdtE, and MdtC, were selected from 397 representative species of prokaryotic microorganisms. These genome assemblies, which were obtained from NCBI FTP site based on the screening parameters such as completeness (≥ 80%), contig numbers (cut-off ≤ 400), N50 (≥ 20,000 bases)[22], were annotated through Rapid Annotation using Subsystem Technology[23]. The blast database was created based on these annotated genome assemblies via the makeblastdb program in Linux, and the executing parameters were set as dbtype prot, and parse_seqids, respectively. The amino acid sequences of AcrE, MdtE, and MdtC from *E. coli* were utilized as queries for bioinformatics screening to predict target regions responsible for MCFA efflux within the constructed blast database associated with the parameters such as E-value cutoff of 1E-12 and bit score cutoff of 200. MUSCLE v3.8.31 was then used to align, trim and concatenate the obtained homologs[24], and IQ-TREE was utilized for phylogenomic reconstruction based on the resulting matrix[25]. During phylogenomic reconstruction, ModelFinder was used to identify the suitable model of substitution, and ultrafast bootstrap was set as 10,000 replicates.

Analysis of detailed evolutionary divergence in *Citrobacter* and *Escherichia* species

In order to comprehensively analyze transporter-centric phylogenies which contained the genomic context surrounding the target gene *acrE*, genomes deposited as *Citrobacteria* or *Escherichia* were retrieved from the NCBI FTP site with the appropriate filter parameters such as contig number (cut-off ≤ 400), N50 (≥ 20,000 bases), and completeness (≥ 80%), resulting in 797 genomes of *Citrobacteria* and 1,084 genomes of *Escherichia*. Based on this, the evolutionary relationships focusing on the genomic context encompassing *acrE* gene among different organisms were analyzed through CORASON[21,26] via retrieving gene neighborhood of *acrE* up to 20 genes upstream and downstream from genomes.

Construction of MCFA efflux pump library

Primers and plasmids utilized here are shown in Supplemental Tables S1 and S2, respectively. In order to clearly annotate each primer or gene, all the names of these genetic parts contained both abbreviated species and gene names. The plasmid of pCDFD-T7-*bktB*-T7-*fadB*-T7-*ter*-T7-*ydiI*-T7-*acs*, which was used for MCFA production, was derived from our previous study[19]. All the predicted efflux pumps were amplified from the genomic DNA prepared by Ezup Column Bacteria Genomic DNA Purification Kit (Sangon Biotech, Shanghai, China), or synthesized by GenScript (Nanjing, China). Primers Pf_PA-others(*Nde*I) and Pr_PA-others(*Xho*I), Pf_PA-mdtC(*Nde*I) and Pr_PA-mdtC(*Xho*I), Pf_SC-mdtC(*Nde*I) and Pr_SC-mdtC(*Xho*I), Pf_SE-mdtC(*Nde*I) and Pr_SE-mdtC(*Xho*I), Pf_SE-acrE(*Nde*I) and Pr_SE-acrE(*Xho*I), Pf_SE-acrA(*Nde*I) and Pr_SE-acrA(*Xho*I) were used to amplify other efflux RND transporter periplasmic adaptor subunit families of *Pseudomonas aeruginosa*, *mdtC* of *Pseudomonas aeruginosa*, *mdtC* of *Streptomyces coelicolor*, and *mdtC* of *Salmonella enterica*, *acrE* of *Salmonella enterica*, *acrA* of *Salmonella enterica* from corresponding genomic DNA into *Nde*I/*Xho*I site of pETDuet-1 through Gibson assembly kit (New England Biolabs), resulting in plasmids of pETD-PA-others, pETD-PA-*mdtC*, pETD-SC-*mdtC*, pETD-SE-*mdtC*, pETD-SE-*acrE*, pETD-SE-*acrA*, respectively.

Primers Pf_CTR-acrE(*Nde*I) and Pr_CTR-acrE(*Xho*I), Pf_CTR-acrA(*Nde*I) and Pr_CTR-acrA(*Xho*I), Pf_CTR-mdtE(*Nde*I) and Pr_CTR-mdtE(*Xho*I), Pf_CTE-acrE(*Nde*I) and Pr_CTE-acrE(*Xho*I), Pf_CTE-acrA(*Nde*I) and Pr_CTE-acrA(*Xho*I), Pf_ES-acrE(*Nde*I) and Pr_ES-acrE(*Xho*I), Pf_ES-acrA(*Nde*I) and Pr_ES-acrA(*Xho*I), Pf_BA-acrE(*Nde*I) and Pr_BA-acrE(*Xho*I), Pf_BA-acrA(*Nde*I) and Pr_BA-acrA(*Xho*I), Pf_CU-acrE(*Nde*I) and Pr_CU-acrE(*Xho*I), Pf_CU-acrA(*Nde*I) and Pr_CU-acrA(*Xho*I), Pf_KV-acrE(*Nde*I) and Pr_KV-acrE(*Xho*I), Pf_KV-acrA(*Nde*I) and Pr_KV-acrA(*Xho*I), Pf_KV-others(*Nde*I) and Pr_KV-others(*Xho*I), Pf_RT-acrA(*Nde*I) and Pr_RT-acrA(*Xho*I), Pf_RT-others(*Nde*I) and Pr_RT-others(*Xho*I), Pf_AG-others(*Nde*I) and Pr_AG-others(*Xho*I), Pf_SF-others(*Nde*I) and Pr_SF-others(*Xho*I), Pf_CR-others(*Nde*I) and Pr_CR-others(*Xho*I), Pf_MP-others(*Nde*I) and Pr_MP-others(*Xho*I), Pf_ZA-acrA(*Nde*I) and Pr_ZA-acrA(*Xho*I) were

used to amplify *acrE* of *Citrobacter tructae, acrA* of *Citrobacter tructae, mdtE* of *Citrobacter tructae, acrE* of *Citrobacter telavivum, acrA* of *Citrobacter telavivum, acrE* of *Enterobacter soli, acrA* of *Enterobacter soli, acrE* of *Buttiauxella agrestis, acrA* of *Buttiauxella agrestis, acrE* of *Cronobacter universalis, acrA* of *Cronobacter universalis, acrE* of *Klebsiella variicola, acrA* of *Klebsiella variicola,* other efflux RND transporter periplasmic adaptor subunit families of *Klebsiella variicola, acrA* of *Raoultera terrigena,* other efflux RND transporter periplasmic adaptor subunit families of *Raoultera terrigena,* other efflux RND transporter periplasmic adaptor subunit families of *Acetobacter ghanensis,* other efflux RND transporter periplasmic adaptor subunit families of *Solimonas flava,* other efflux RND transporter periplasmic adaptor subunit families of *Caulobacter rhizosphaerae,* other efflux RND transporter periplasmic adaptor subunit families of *Methylibium petroleiphilum, acrA* of *Zavarzinia aquatilis* from corresponding pUC57 derived plasmids (GenScript, Nanjing, China) into *Nde*I/*Xho*I site of pETDuet-1 through Gibson assembly kit (New England Biolabs, Ipswich, UK), resulting in plasmids of pETD-CTR-*acrE*, pETD-CTR-*acrA*, pETD-CTR-*mdtE*, pETD-CTE-*acrE*, pETD-CTE-*acrA*, pETD-ES-*acrE*, pETD-ES-*acrA*, pETD-BA-*acrE*, pETD-BA-*acrA*, pETD-CU-*acrE*, pETD-CU-*acrA*, pETD-KV-*acrE*, pETD-KV-*acrA*, pETD-KV-others, pETD-RT-*acrA*, pETD-RT-others, pETD-AG-others, pETD-SF-others, pETD-CR-others, pETD-MP-others, pETD-ZA-*acrA*, respectively.

To fuse each predicted efflux pump to GFP individually, the stop codon of each predicted efflux pump was removed and two rounds of PCR was used to introduce a Gly-Ser-Gly linker between these two genes[27]. During the first round, two sets of primers such as Pf_PA-others-GSG-GFP(*Eco*NI) and Pr_PA-others-GSG-GFP, Pf_PA-others-GSG-GFP and Pr_PA-others-GSG-GFP(*Xho*I) were used. Secondly, primers Pf_PA-others-GSG-GFP(*Eco*NI)/Pr_fused-GFP(*Xho*I) were used to connect two above PCR products via overlapping extension PCR, resulted in pACYC-PA-others-GSG-GFP harboring fused gene construct encoding PA_others, three amino acid linker, and GFP. Similarly, Pf_PA-mdtC-GSG-GFP(*Eco*NI)/Pr_PA-mdtC-GSG-GFP and Pf_PA-mdtC-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_SC-mdtC-GSG-GFP(*Eco*NI)/Pr_SC-mdtC-GSG-GFP and Pf_SC-mdtC-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_SE-mdtC-GSG-GFP (*Eco*NI)/Pr_SE-mdtC-GSG-GFP and Pf_SE-mdtC-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_SE-acrE-GSG-GFP(*Eco*NI)/Pr_SE-acrE-GSG-GFP and Pf_SE-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_SE-acrA-GSG-GFP(*Eco*NI)/Pr_SE-acrA-GSG-GFP and Pf_SE-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CTR-acrE-GSG-GFP(*Eco*NI)/Pr_CTR-acrE-GSG-GFP and Pf_CTR-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CTR-acrA-GSG-GFP(*Eco*NI)/Pr_CTR-acrA-GSG-GFP and Pf_CTR-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CTR-mdtE-GSG-GFP(*Eco*NI)/Pr_CTR-mdtE-GSG-GFP and Pf_CTR-mdtE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CTE-acrE-GSG-GFP(*Eco*NI)/Pr_CTE-acrE-GSG-GFP and Pf_CTE-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_ES-acrE-GSG-GFP(*Eco*NI)/Pr_ES-acrE-GSG-GFP and Pf_ES-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_ES-acrA-GSG-GFP(*Eco*NI)/Pr_ES-acrA-GSG-GFP and Pf_ES-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_BA-acrE-GSG-GFP(*Eco*NI)/Pr_BA-acrE-GSG-GFP and Pf_BA-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_BA-acrA-GSG-GFP(*Eco*NI)/Pr_BA-acrA-GSG-GFP and Pf_BA-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CU-acrE-GSG-GFP(*Eco*NI)/Pr_CU-acrE-GSG-GFP and Pf_CU-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CU-acrA-GSG-GFP(*Eco*NI)/Pr_CU-acrA-GSG-GFP and Pf_CU-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_KV-acrE-GSG-GFP(*Eco*NI)/Pr_KV-acrE-GSG-GFP and Pf_KV-acrE-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_KV-acrA-GSG-GFP(*Eco*NI)/Pr_KV-acrA-GSG-GFP and Pf_KV-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_KV-others-GSG-GFP(*Eco*NI)/Pr_KV-others-GSG-GFP and Pf_KV-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_RT-acrA-GSG-GFP(*Eco*NI)/Pr_RT-acrA-GSG-GFP and Pf_RT-acrA-GSG-GFP/Pr_fused-GFP (*Xho*I), Pf_RT-others-GSG-GFP(*Eco*NI)/Pr_RT-others-GSG-GFP and Pf_RT-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_AG-others-GSG-GFP(*Eco*NI)/Pr_AG-others-GSG-GFP and Pf_AG-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_SF-others-GSG-GFP(*Eco*NI)/Pr_SF-others-GSG-GFP and Pf_SF-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_CR-others-GSG-GFP(*Eco*NI)/Pr_CR-others-GSG-GFP and Pf_CR-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_MP-others-GSG-GFP(*Eco*NI)/Pr_MP-others-GSG-GFP and Pf_MP-others-GSG-GFP/Pr_fused-GFP(*Xho*I), Pf_ZA-acrA-GSG-GFP (*Eco*NI)/Pr_ZA-acrA-GSG-GFP and Pf_ZA-acrA-GSG-GFP/Pr_fused-GFP(*Xho*I) were used to fuse other predicted efflux pumps to GFP, this resulted in pACYC-PA-*mdtC*-GSG-GFP, pACYC-SC-*mdtC*-GSG-GFP, pACYC-SE-*mdtC*-GSG-GFP, pACYC-SE-*acrE*-GSG-GFP, pACYC-SE-*acrA*-GSG-GFP, pACYC-CTR-*acrE*-GSG-GFP, pACYC-CTR-*acrA*-GSG-GFP, pACYC-CTR-*mdtE*-GSG-GFP, pACYC-CTR-*acrE*-GSG-GFP, pACYC-ES-*acrE*-GSG-GFP, pACYC-ES-*acrA*-GSG-GFP, pACYC-BA-*acrE*-GSG-GFP, pACYC-BA-*acrA*-GSG-GFP, pACYC-CU-*acrE*-GSG-GFP, pACYC-CU-*acrA*-GSG-GFP, pACYC-KV-*acrE*-GSG-GFP, pACYC-KV-*acrA*-GSG-GFP, pACYC-KV-others-GSG-GFP, pACYC-RT-acrA-GSG-GFP, pACYC-RT-others-GSG-GFP, pACYC-AG-others-GSG-GFP, pACYC-SF-others-GSG-GFP, pACYC-CR-others-GSG-GFP, pACYC-MP-others-GSG-GFP, pACYC-ZA-*acrA*-GSG-GFP, respectively.

Primers Pf_CTR-envR(*Nde*I) and Pr_CTR-envR(*Xho*I) were used to amplify *envR* of *Citrobacter tructae* from corresponding pUC57 derived plasmids (GenScript, Nanjing, China) into *Nde*I/*Xho*I site of pETDuet-1 through Gibson assembly kit (New England Biolabs), resulting in plasmids of pETD-CTR-*envR*. Primers Pf_EC-envR(*Nde*I) and Pr_EC-envR(*Xho*I) were used to amplify *envR* of *E. coli* from genomic DNA into *Nde*I/*Xho*I site of pETDuet-1 through Gibson assembly kit (New England Biolabs), resulting in plasmids of pETD-EC-*envR*. The lambda-red recombination-based method[28] was used to construct the EC_*envR* knockout mutant. Briefly, primers Pf_Kan^{FRT}-EC-envR and Pr_ Kan^{FRT}-EC-envR were used to amplify the FRT-flanked kanamycin resistance gene (*Kan*^{FRT}) from the plasmid pKD13[28], which included 40 bp of homology with the ends of EC-*envR* in both sides. This design would facilitate integration of this cassette into the corresponding sites. After transforming these cassettes, proper colonies were verified via colony PCR and following sequencing. The FRT-flanked Kan would be excised by FLP recombinase via pCP20 plasmid[28]. Primers Pf_CTR-acrF(G)/Pr_CTR-acrF(G), and Pf_pETD-CTR-acrE(G)/ Pr_pETD-CTR-acrE(G) were used to amplify CTR-*acrF* of *Citrobacter tructae* from corresponding pUC57 derived plasmids (GenScript, Nanjing, China) into pETD-CTR-*acrE* through Gibson assembly kit (New England Biolabs), resulting in plasmids of pETD-CTR-*acrE*-CTR-*acrF*.

Construction of autonomous MCFA secreting systems

Primers and plasmids utilized here were shown in Supplemental Tables S1 and S2, respectively. Primer sets of

Pf_Ptrc-PrgX(PETD)/Pr_Ptrc-PrgX(Pi), Pf_Pi-ccfA(G)/Pr_ccfA(G), Pf_prgZ(G)/Pr_prgZ(G), Pf_PprgQ-CTR-acrE(G)/Pr_PprgQ-CTR-acrE(G), and Pf_PprgQ-CTR-acrF(G)/Pr_PprgQ-CTR-acrF(PETD) were used to amplify *prgX* under Ptrc promoter, *ccfA* under Pi promoter, *prgZ* under P1 promoter, CTR_*acrE* under PprgQ promoter, and CTR_*acrF* under PprgQ promoter from pACYC-Ptrc-*prgX*, pETD-Ptrc-*ccfA*-Ptrc-*prgZ*, and corresponding pUC57 derived plasmids (GenScript, Nanjing, China) into *Eco*NI/*Xho*I site of pETDuet-1 through Gibson assembly kit (New England Biolabs) (i = 1–6). This would result in the plasmid of pETD-Ptrc-*prgX*-Pi-*ccfA*-PprgQ-CTR-*acrE*-PprgQ-CTR-*acrF* (i = 1–6).

Fluorescence intensity measurement

During the shake flask culture, LB medium associate with corresponding antibiotics was firstly utilized to culture engineered strains overnight (37 °C, 220 rpm orbital shaking). MOPS minimal medium supplemented with 10 g/L D-glucose was then used for re-culture with OD_{600} of 0.1, and the culture condition was then altered to 30 °C when OD_{600} reached 0.6[29]. At this time, 1 mM IPTG was added to induce the expression. Cell fluorescence and cell density were measured after 30 h of culture using Cytation 3 imaging reader system (BioTek, Winooski, USA).

Analytical methods

Each experiment was conducted in triplicate and the deviation was represented by the error bar. The extracellular and intracellular MCFA measurement was conducted based on our previous study[19]. Briefly, the supernatant of 1 mL cell culture was obtained (10,000 g, 5 min) for extracellular MCFA measurement, whereas the cell pellet of 1 mL cell culture was recovered (10,000 g, 5 min) with 1 mL deionized water for intracellular MCFA measurement. Based on our previous studies[13,14], gas chromatograph mass spectrometer (GC-MS) QP2010 Plus (Shimadzu) equipped with GC-MS column (Rtx-5 MS capillary with length of 30 m, film thickness of 0.25 μm, diameter of 0.25 mm) was utilized for the following free fatty acid extraction and quantification.

Culture conditions

MOPS minimal medium supplemented with 15 g/L D-glucose was used to perform the fermentations as demonstrated in our previous studies[13,14]. The overnight incubation in LB medium was firstly conducted to prepare the pre-inocula, which were then diluted into 50 mL MOPS minimal medium with an initial OD_{600} of 0.1 in 500-mL flasks. The parameters of 37 °C and 220 rpm orbital shaking were used to conduct the fermentation. The culture temperature was then altered to 30 °C with the supplementation of 1 mM IPTG when the OD_{600} reached 0.5–0.6. The concentration of MCFAs, including both extracellular and intracellular levels, was measured after a fermentation time of 48 h.

Batch culture

Seed culture, which was performed on rotary shakers overnight (37 °C, 220 rpm), was then diluted into 3-L BioFlo 115 fermentor (New Brunswick Scientific Co, Edison, NJ, USA) as an OD_{600} of 0.1. This fermentor included 1.5 L MOPS minimal medium associate with corresponding antibiotics and 10 g/L D-glucose. During the fermentation, concentrated D-glucose (800 g/L) was used to maintain D-glucose concentration at 5 g/L. The cultivation temperature was changed to 30 °C when OD_{600} reached 0.5–0.6 associoate with the supplementation of 1 mM IPTG. 12.5% NH_4OH solution or phosphoric acid solution was used to keep the pH at 6.5, and the agitation cascade (200–500 rpm) was utilized to keep the dissolved oxygen concentration at 30% saturation. Each MCFA fermentation was conducted in triplicate, and the deviation was represented via the error bar.

RESULTS

General evolutionary recapitulation of MCFA transporter families

Our previous secreting system screened numerous endogenous transporters including famous AcrAB-TolC system and other triphosphate (ATP)-binding cassette superfamily or annotated multidrug efflux superfamily, and found that the overexpression of resistance nodulation cell division family transporter *acrE*, *mdtE* and *mdtC* together with the deletion of multidrug efflux pump *cmr* from *E. coli* achieved the best performance[19]. However, owing to the rapid accumulation of genomic information, other sequenced microbial genomes may contain numerous efflux pumps and present a greatly unexplored resource for mining novel pumps. In order to screen the most favorable candidates during large genomic datasets, a multi-layer genome mining and phylogenomic analysis was developed. Firstly, the general evolutionary recapitulation of MCFA transporter families was investigated by comprehensive and systematic phylogenomics, and the input of the customized blast database for this analysis was constructed with 397 genomes belonging to different representative prokaryotic species.

Our previous study identified that *acrE*, *mdtE* and *mdtC* from *E. coli* were responsible for accelerating MCFA export[19]. Hence, the amino acid sequences of AcrE, MdtE, and MdtC from *E. coli* were utilized as queries for the bioinformatics screen to predict target regions responsible for MCFA efflux within the constructed blast database. This screen was performed under E-value cutoff of 1E-12 and bit score cutoff of 200. The homology hits for AcrE, MdtE, and MdtC were 287, 284, 1446, respectively, among the constructed blast database, and the evolutionary relationships of AcrE, MdtE, and MdtC homology hits are presented in Figs 1, 2 & 3, respectively.

As seen in Fig. 1, the homologues of AcrE were distributed in 134 genomes, and most genomes contained more than one homology hit, indicating the deep genomic mining for the target gene. The information of these homologues was then confirmed via BLASTp. It was found that the homology hits of AcrE mainly included AcrE families, AcrA families, MdtE families, and other efflux RND transporter periplasmic adaptor subunits such as MexX, MexA. Whereas the homologues of MdtE were also distributed in 134 genomes, and most genomes also contained multiple homology hits (Fig. 2). Similarly, the homology hits of MdtE also mainly comprised AcrE families, AcrA families, MdtE families, and other efflux RND transporter periplasmic adaptor subunits, indicating the analogous evolutionary relationships between AcrE and MdtE. MdtC presented totally different evolutionary history

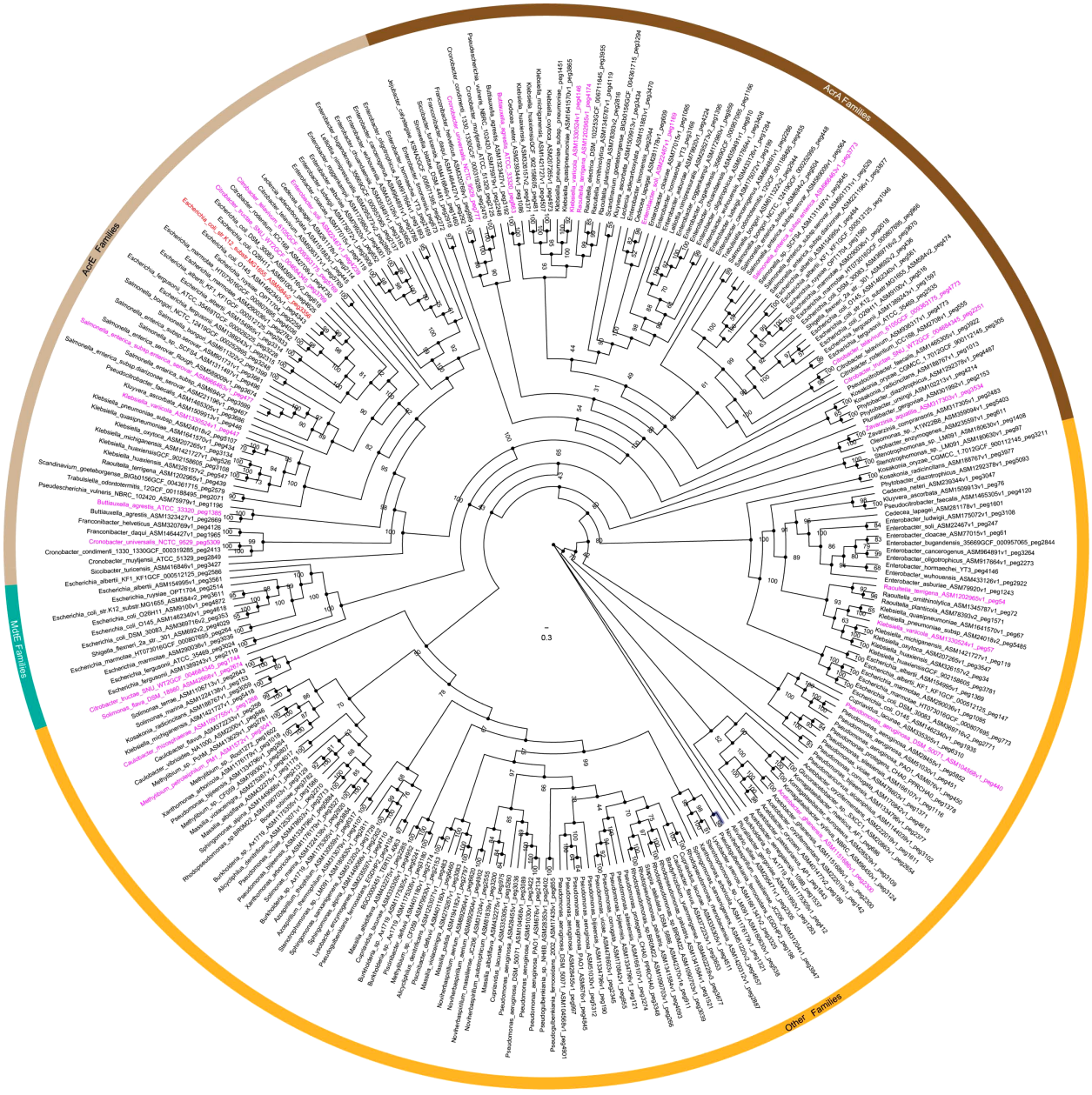

Fig. 1 The evolutionary relationships of AcrE homology hits. When using AcrE as a query, the evolutionary relationships of 287 homology hits were analyzed and each homolog information was confirmed with BLASTp. It was found that the homology hits of AcrE mainly included AcrE families, AcrA families, MdtE families, and other efflux RND transporter periplasmic adaptor subunits. The violet and red indicated the selected predicted efflux pumps for further analysis.

compared with AcrE and MdtE, and the 1446 homology hits were distributed in 236 genomes (Fig. 3a). These homology hits could be divided into seven different enzyme families such as MdtC, MdtB, AcrD, AcrF, MdtF, AcrB, and CusA families, and the evolutionary history of MdtC was further recapitulated (Fig. 3b).

When utilizing AcrE (Fig. 1) or MdtE (Fig. 2) as a query to mining genomes, homologues from *Citrobacteria*, *Salmonella*, and *Enterobacteria* species presented the closest evolutionary relationships with *Escherichia* species among both AcrE and AcrA families; Among the MdtE families, only the homologue from *Citrobacteria tructae* and *Escherichia* species existed; Whereas among efflux RND transporter periplasmic adaptor subunit families, merely homologues from partial *Escherichia* species were existing along with other species such as *Pseudomonas*, *Acetobacteria* species.

When using MdtC as a query to mining genomes, homologues from *Citrobacteria*, *Enterobacteria*, and *Salmonella* species presented the closest evolutionary relationships with *Escherichia* species, and *Enterobacteria* species exhibited closer evolutionary relationships than *Salmonella* species among MdtC families (Fig. 3b), whereas these two species bestowed different evolutionary behaviors when using AcrE or MdtE as queries. Furthermore, the taxonomic relationship of each species was defined via constructing a species tree with the amino acid sequences of their RNA polymerase beta

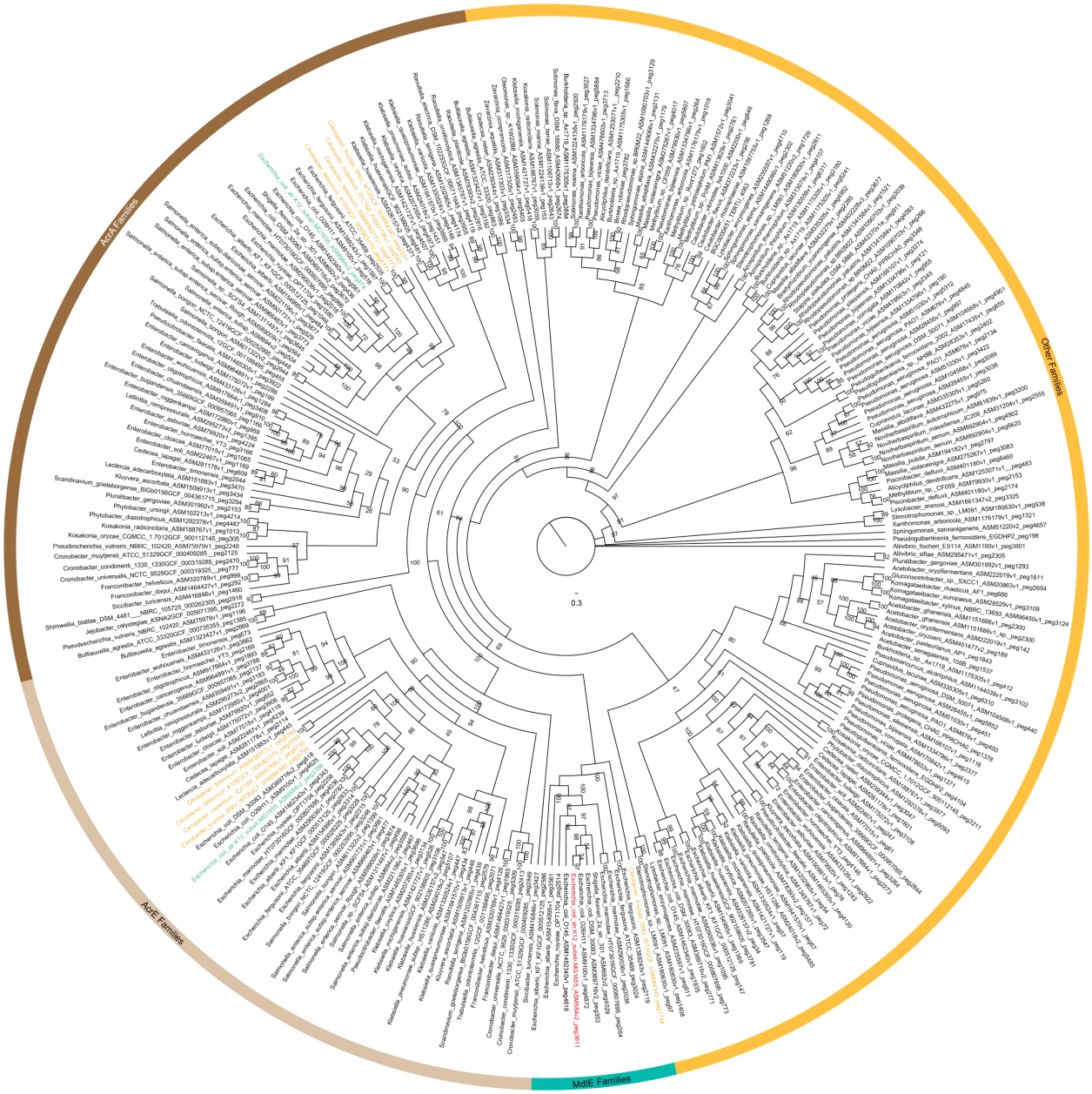

Fig. 2 The evolutionary relationships of MdtE homology hits. When using MdtE as a query, the evolutionary relationships of 284 homology hits were analyzed and each homolog information was confirmed with BLASTp. It was found that the homology hits of MdtE also mainly comprised AcrE families, AcrA families, MdtE families, and other efflux RND transporter periplasmic adaptor subunits. The colored areas indicate the relationships between *Citrobacteria* and *E. coli* species.

subunits (RpoB) (Fig. 4). It was found that *Salmonella* species exhibited closer evolutionary relationship with *Escherichia* species than *Citrobacteria* species, which was different when using AcrE, MdtE, or MdtC as queries, suggesting the interesting engineering targets of homologues from *Citrobacteria* species.

Construction of MCFA efflux pump library

The above bioinformatic metric rendered the ability to rank the entire set of pumps and pick a portion that manifested a uniform distribution of candidates. To construct the library, the predicted efflux pumps were amplified from the genomic DNA or synthesized by GenScript (Nanjing, China), and this library harbored 29 predicted efflux pumps, all of which had not been previously characterized for MCFA transport. This library mainly focused on AcrE or MdtE homologues, as in our previous study[19] demonstrated that these two transporters derived from *E. coli* exhibited better performance than MdtC. Besides, due to the large size of MdtC (> 3,000 bp), it is costly and not convenient to amplify or synthesize numerous MdtC homologues.

AcrE or mdtE homologues from *Citrobacter tructae* and *Citrobacter telavivum* among AcrE families, AcrA families, and MdtE families were selected, as we observed that these

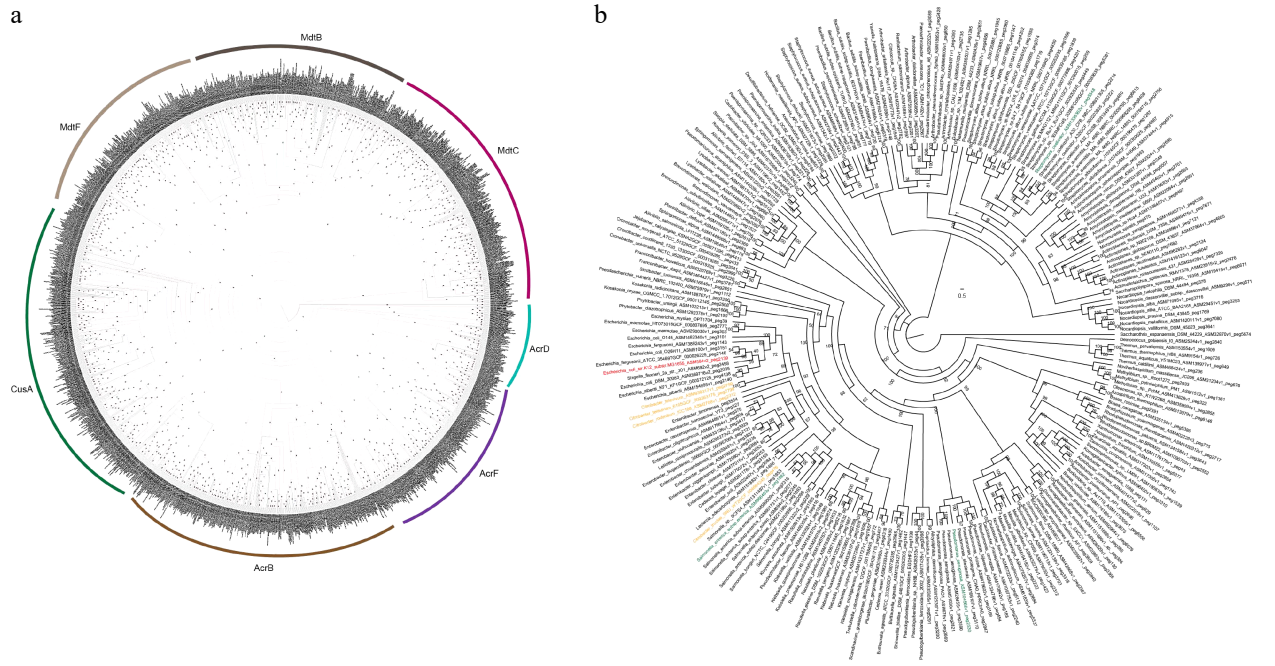

Fig. 3 The evolutionary relationships of MdtC homology hits. (a) When using MdtC as a query, the evolutionary relationships of 1,446 homology hits were analyzed and each homolog information was confirmed with BLASTp. These homology hits could be divided into seven different enzyme families such as MdtC, MdtB, AcrD, AcrF, MdtF, AcrB, and CusA families. (b) The evolutionary history of MdtC families was further recapitulated. The colored areas indicate the selected predicted efflux pumps for further analysis.

species presented different evolutionary trajectories. For instance, under the same search parameters, when using AcrE as a query, suitable hits were obtained and occurred in similar evolutionary positions among the AcrE families (Fig. 1); Whereas only suitable hits from *Citrobacter tructae* were observed when using MdtE as a query among the MdtE family; When using MdtC as a query, suitable hits were obtained in both species, yet they occurred in different evolutionary positions among the MdtC families (Fig. 3). Other AcrE/MdtE homologues were selected from *Salmonella enterica*, *Enterobacter soli*, *Buttiauxella agrestis* and *Cronobacter universalis* among AcrE or AcrA families, *Klebsiella variicola* among AcrE families, AcrA families, or other efflux RND transporter periplasmic adaptor subunit families, *Raoultera terrigena* among AcrA families or other efflux RND transporter periplasmic adaptor subunit families, *Pseudomonas aeruginosa*, *Acetobacter ghanensis*, *Solimonas flava*, *Caulobacter rhizosphaerae*, and *Methylibium petroleiphilum* among other efflux RND transporter periplasmic adaptor subunit families, *Zavarzinia aquatilis* among AcrA families. Several MdtC homologues from *Citrobacter tructae*, *Citrobacter telavivum*, *Pseudomonas aeruginosa*, *Streptomyces coelicolor*, and *Salmonella enterica* were also selected for further investigation.

To efficiently identify suitable transporters with the capability to export MCFAs from cells, a simple test system constructed in our previous study[19], was utilized. This test system consisted of two individual plasmids (Fig. 5b), which could stably maintain in one cell owing to their distinct replication origins and antibiotic resistance markers. The first plasmid pCDFD-T7-*bktB*-T7-*fadB*-T7-*ter*-T7-*ydiI*-t7-*acs* carrying thiolase (BktB) of *Ralstonia eutropha*, 3-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase (FadB) of *E. coli*, transenoyl-CoA reductase (Ter) of *Euglena gracilis*, thioesterase (YdiI) of *E. coli*, and acetyl-CoA synthetase (Acs) of *E. coli*, was responsible for MCFA production (Fig. 5a), whereas the other pETDuet-1 derived plasmid was utilized for the expression of various bacterial transporters.

A set of 29 predicted efflux pumps were then overexpressed individually, and three different measurements including the extracellular MCFA concentration, the intracellular MCFA concentration, and the total MCFA concentration, were used to screen each target pump. Firstly, as these candidates have not been characterized previously, to assure their reliable gene expression, GFP was tagged to each candidate to measure translational output and normalized fluorescence measurements were conducted for each one by dividing measured fluorescence values to the OD_{600} of that well (Supplemental Fig. S1). As seen from Fig. 5c, it was found that homologues among AcrE/MdtE families exhibited better performance than among MdtC families and AcrA families, and the top-performing candidate pumps existed in *Citrobacteria* species.

Detailed evolutionary divergence of MCFA transporter families in *Citrobacter* species

Although the top-performing efflux pumps exist in *Citrobacteria* species, AcrE homologues from different *Citrobacteria* species exhibited dissimilar behaviors. Besides, MCFA transporter homologues from *Citrobacteria* species occurred in divergent evolutionary positions, suggesting the necessity for future engineering efforts. Hence, genomes deposited as *Citrobacteria* were retrieved from the NCBI FTP site with the appropriate filter parameters such as contig number (cut-off ≤ 400), N50 (≥ 20,000 bases), and complete ness (≥ 80%) to remove low-quality genomes and eliminate redundancy at

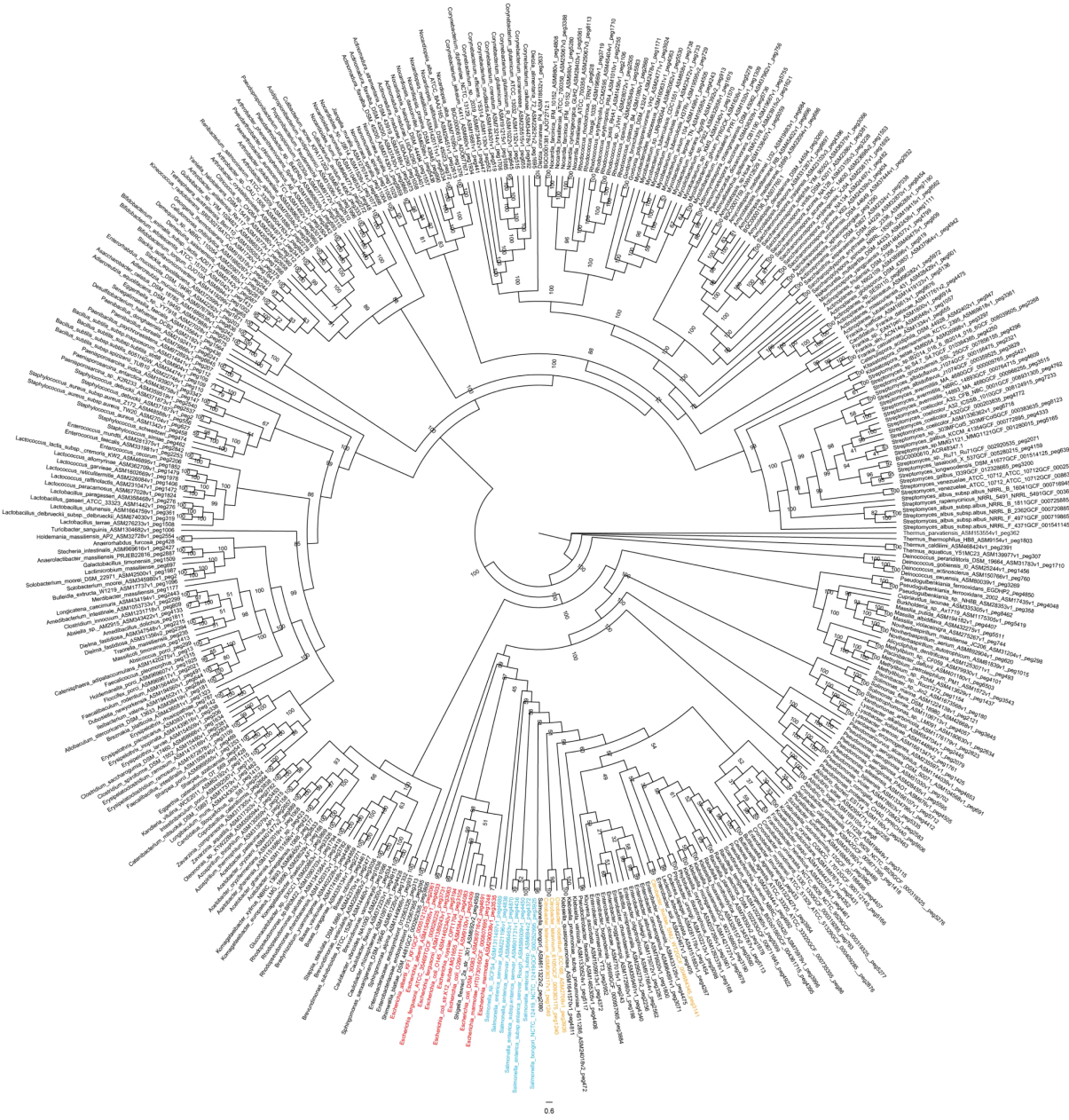

Fig. 4 The taxonomic relationship of each species used for general evolutionary recapitulation. The taxonomic relationship of each species was defined via constructing a species tree with the amino acid sequences of their RNA polymerase beta subunits (RpoB). It was found that *Salmonella* species exhibited closer evolutionary relationship with *Escherichia* species than *Citrobacteria* species, which was different when using AcrE, MdtE, or MdtC as queries.

the strain level. This resulted in a subset of 797 genomes used hereafter, to comprehensively analyze transporter-centric phylogenies which contained the genomic context surrounding target genes.

Analysis of this AcrE-centric phylogenetic tree exhibited in Fig. 6a revealed that EnvR homologues, a predicted AcrEF/EnvCD operon regulator, were present in most *Citrobacteria* species. Hence, we then asked whether this transcriptional regulator could further affect MCFA production. It was found that overexpression of EnvR from *Citrobacter tructae* decreased extracellular MCFA production by 32% (Fig. 6b), suggesting that EnvR might act as a repressor. The AcrE-centric phylogenetic tree based on genomes of *Escherichia* species were then constructed, and 1084 genomes deposited as *Escherichia* were retrieved from the NCBI FTP site. This phylogenetic tree also manifested that EnvR homologues were existing in most *Escherichia* species (Fig. 7a), and it was observed that overexpression of EnvR from *E. coli* decreased extracellular MCFA production by 39%, whereas the deletion of endogenous EnvR further increased extracellular MCFA production by 168% associated with the overexpression of EnvR from *Citrobacter tructae* (Fig. 7b).

Although the deletion of endogenous EnvR rendered the increase of extracellular MCFA production, we also observed

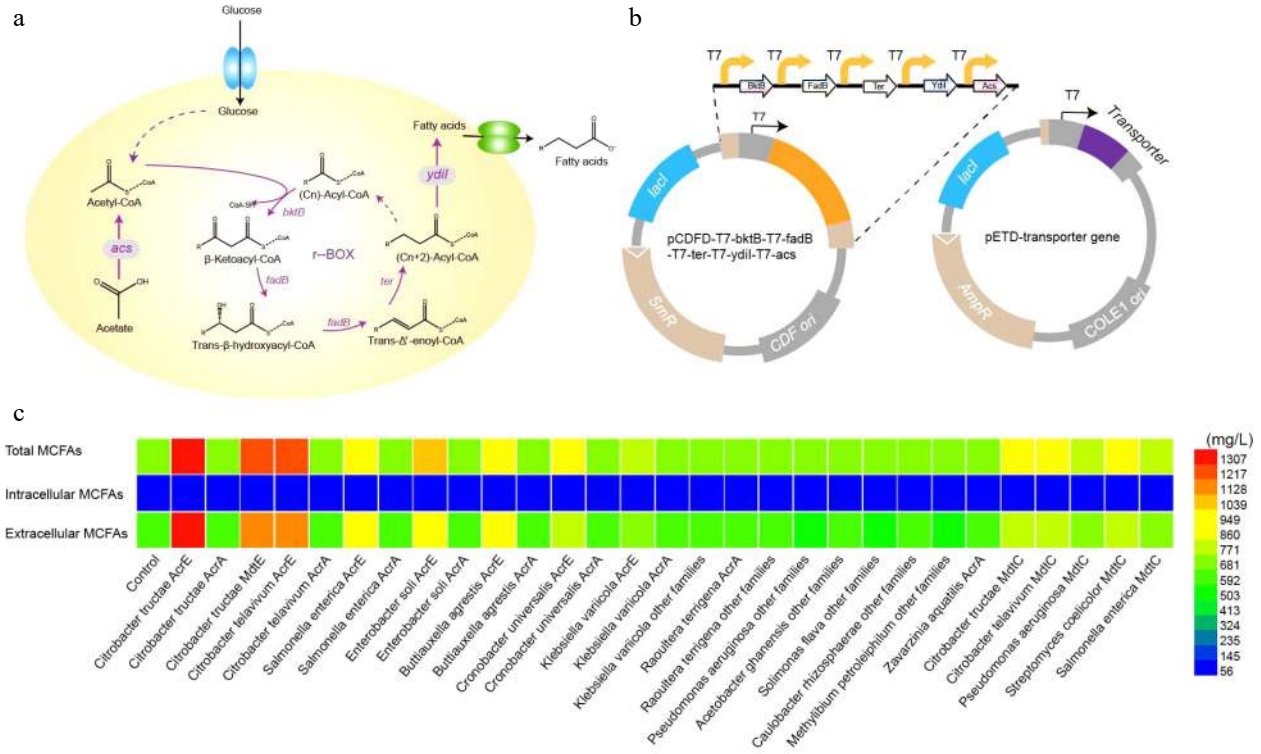

Fig. 5 Construction of MCFA efflux pump library. (a) Microbial production of MCFAs from D-glucose via the reversal of *β*-oxidation cycle and transporter engineering. (b) Illustration of the test system. This test system consisted of two individual plasmids. The first plasmid pCDFD-T7-*bktB*-T7-*fadB*-T7-*ter*-T7-*ydiI*-t7-*acs* was responsible for MCFA production, whereas the other pETDuet-1 derived plasmid was utilized for the expression of various bacterial transporters. (c) Effect of predicted efflux pump engineering on extracellular, intracellular and total MCFA production. Each experiment in this study was conducted in triplicate and error bars signify standard deviation (SD) with 95% confidence interval (CI).

the decrease of the cell growth (Supplemental Fig. S2). This would exert a negative influence on the total MCFA production and indicated that EnvR was not only involved in MCFA export, but also possessed unknown essential functions. In order to prevent the deactivation of the entire regulon by deleting EnvR, we sought to investigate whether there was a new protein potentially involved in MCFA export. As EnvR was a predicted AcrEF operon regulator, AcrF from *Citrobacter tructae* was then overexpressed associated with AcrE. It was found that overexpression of both AcrE and AcrF exhibited the best performance (2.5-fold) among all the candidates (Fig. 7b), demonstrating that AcrE and AcrF were responsible for MCFA export.

Construction of autonomous MCFA secreting systems

In order to convert MCFA efflux to an autonomous behavior without inducer supplementation and human supervision, we turned to combining quorum-sensing (QS) circuitry with the efflux pumps. Our previous studies described two robust and autonomous QS-based circuits deriving from peptide pheromone responsive QS system of *Enterococcus faecalis* (QEX), and optimized acyl-homoserine lactone responsive QS system of *Vibrio fisheri* (QVX) by introducing T7 RNA polymerase as a genetic amplifier[26,30]. As the optimized QVX circuity needs the expression of T7 RNA polymerase, this would affect the utilization of T7 promoter for driving other pathway genes. Hence, in this study, T7 promoter driving the expression of efflux pumps was replaced by QEX circuity.

During the QEX circuity, the operator sequence of the response promoter PprgQ was repressed by the master protein regulator PrgX, and the activation of this response promoter only occurred when heptapeptide cCF10 synthesized by heptapeptide CcfA bound to protein regulator PrgX (Fig. 8a)[30]. Our previous studies demonstrated that the components of functional QEX circuity must contain protein regulator PrgX and surface cCF10-binding protein PrgZ driven by constitutive Ptrc and P1 promoters, respectively, to assure both the low leakiness and robust response behavior of QEX circuity[30], whereas signal synthase CcfA was driven by constitutive promoters with different strength ranging from high strength P1 to low strength P6, to trigger QEX circuity at various times.

As seen in Fig. 8b, it was observed that different triggering times of QEX circuity driving the efflux pumps exerted different impact on extracellular MCFA concentrations and total MCFA concentrations. We found that an early or delayed triggering of efflux pumps led to the decrease of extracellular or total MCFA concentrations compared to the suitable triggering time (i = 2), further demonstrating the importance of examining the impact of different triggering times on efflux efficiency. It was presumed that during the early fermentation time, product toxicity did not present as an issue in strain engineering, and the early expression of efflux pumps would exert an extra metabolic burden on host strains; whereas the delay triggering of efflux pumps would not efficiently alleviate the product toxicity.

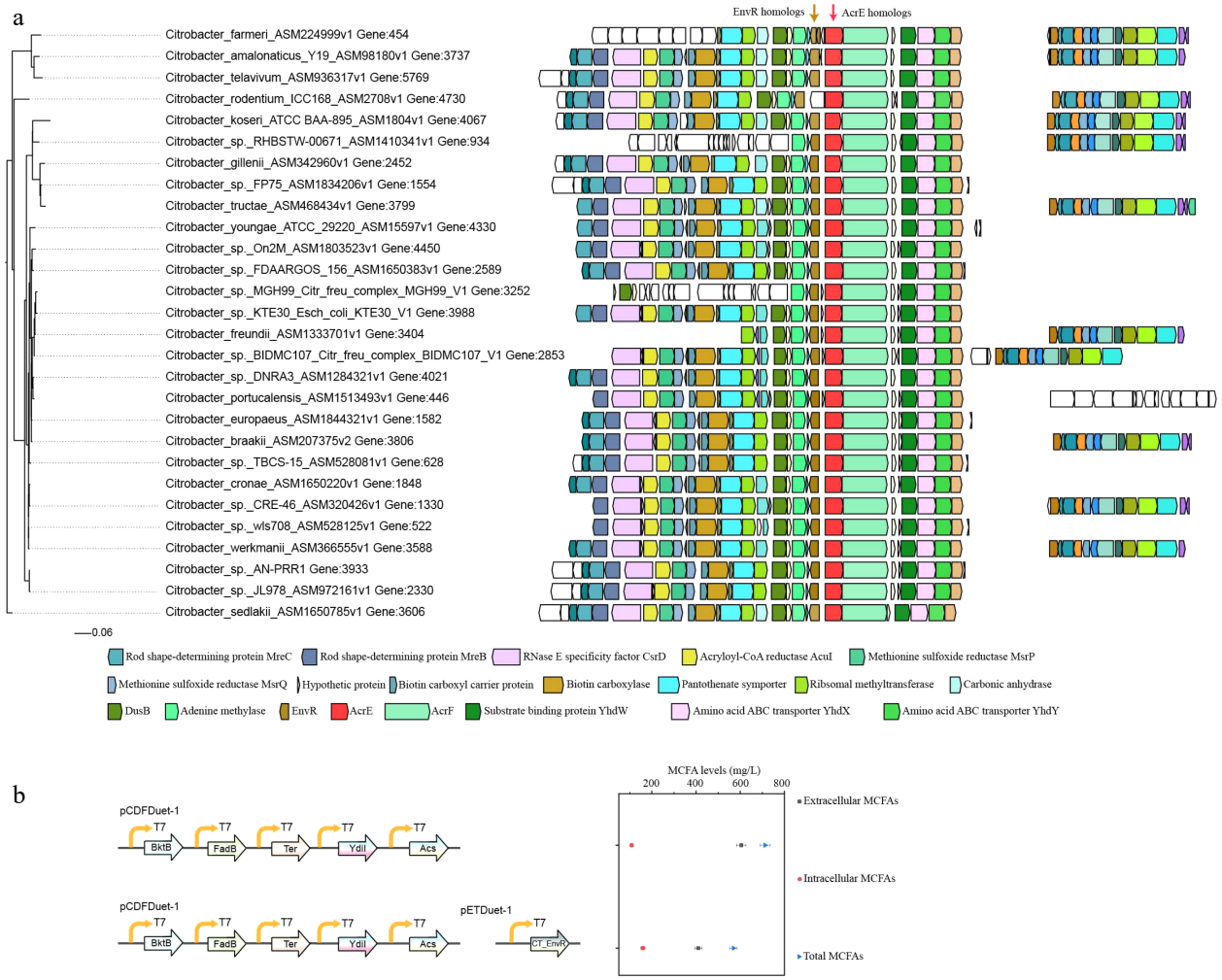

Fig. 6 Detailed evolutionary divergence of MCFA transporter families in *Citrobacter* species. (a) Analysis of this AcrE-centric phylogenetic tree based on genomes from *Citrobacter* species. This revealed that EnvR homologues, a predicted AcrEF/EnvCD operon regulator, were present in most *Citrobacteria* species. (b) Effect of transcriptional regulator EnvR engineering on MCFA production. CT_EnvR indicated *envR* of *Citrobacter tructae*. Experiments in this study were conducted in triplicate and error bars signify SD with 95% CI.

We also evaluated the performance of this autonomous MCFA secreting system in scaled-up bioreactors (Fig. 8c), which presented as more industrially relevant procedures. The autonomous secreting system was then evaluated in a 5-L fermenter with the conduction of dissolved oxygen (30%), glucose (5 g/L) and pH control (6.5). It was observed that engineered strains in bioreactors exhibited better performance than in shake flasks, and a nearly 4.9-fold increase in MCFA titers (6.9 g/L) was observed. It was presumed that engineered strains in bioreactors produced higher concentration of MCFAs than shake flasks, and this would render more product toxicity to host strains, thus limiting their performance in bioreactors, whereas our autonomous secreting system would unleash their potential in target product synthesis.

DISCUSSION

Most bio-chemicals present toxic effects and stresses towards host strains during high concentrations, which are essential for developing an economically viable and scalable bio-process[1,16,31]. Furthermore, extracting MCFAs through harvesting engineered organisms also exhibits energy- and cost-intensive characteristics. Numerous studies found that microbial efflux pumps could provide host strains the ability of resistance to high target product concentrations in fermentation broth via improving the secretion of endogenous compounds. More importantly, expediting product secretion could decrease product inhibition and improve target flux through reversible reactions due to the maintainence of low intracellular target product levels[16,17]. However, the information of efflux pumps specially responsible for MCFA transport is limited. Here, a multi-layer genome mining analysis combining with quorum-sensing circuit was developed to screen a library of uncharacterized heterologous pumps among over 2000 microbial genomes, and these efforts rewired the MCFA efflux to a robust and autonomous behavior without inducer supplementation or human supervision, paving the way to develop economically feasible bioprocesses.

The current MCFA secreting system is built on the basis of endogenous transporters, which require both over-expression of *acrE*, *mdtE*, *mdtC* and deletion of *cmr* from *E. coli*[19].

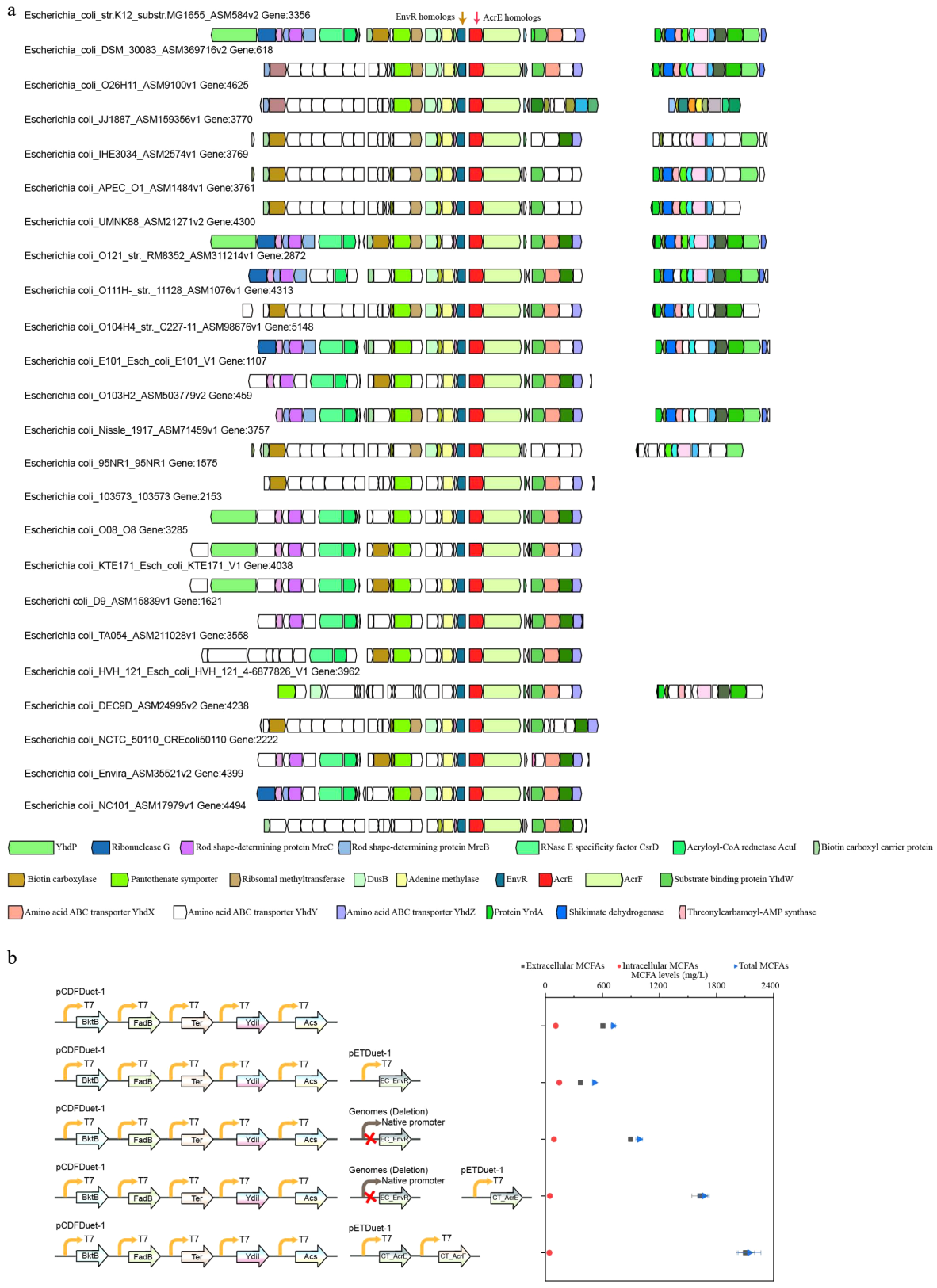

Fig. 7 Detailed evolutionary divergence of MCFA transporter families in *Escherichia* species. (a) Analysis of the AcrE-centric phylogenetic tree based on genomes of *Escherichia* species. This phylogenetic tree also manifested that EnvR homologues were existing in most *Escherichia* species. (b) Effect of transcriptional regulator EnvR engineering on MCFA production. EC_EnvR indicated *envR* of *E. coli*; CT_EnvR indicated *envR* of *Citrobacter tructae*; CT_AcrE indicated *acrE* of *Citrobacter tructae*; CT_AcrF indicated *acrF* of *Citrobacter tructae*. Experiments in this study were conducted in triplicate and error bars signify SD with 95% CI.

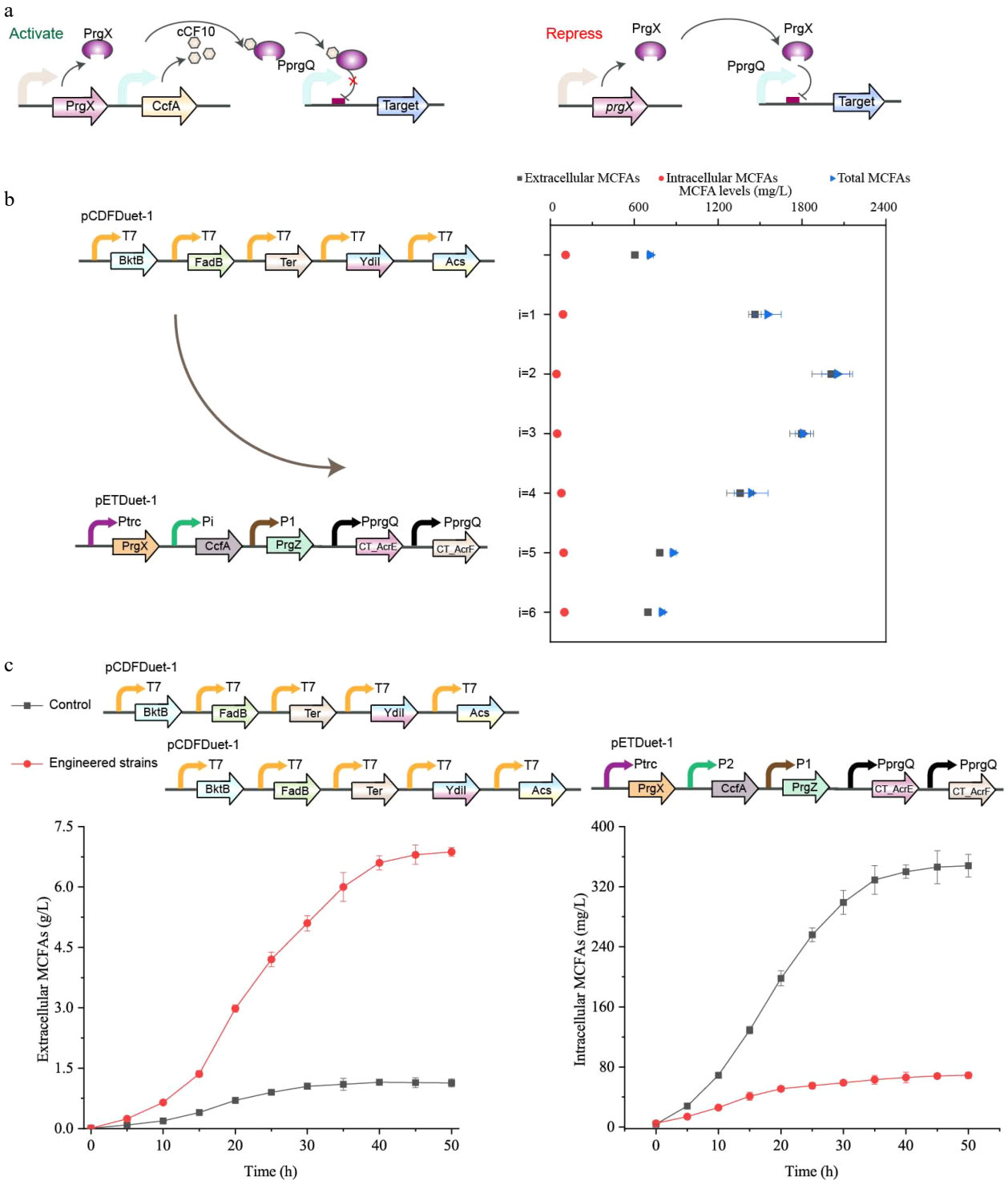

Fig. 8 Construction of autonomous MCFA secreting systems. (a) Schematic of QEX circuity. (b) The effect of replacing T7 promoter with QEX circuity on MCFA production. The signal synthase CcfA was driven by constitutive promoters with different strength ranging from high strength P1 to low strength P6, to trigger QEX circuity at various times. (c) The evaluation of the performance of this autonomous MCFA secreting system in scaled-up bioreactors. Experiments in this study were conducted in triplicate and error bars signify SD with 95% CI.

However, fueled by rapid developments in high-throughput sequencing, numerous other sequenced microbial genomes contain abundant efflux pumps and present a largely unexplored resource for mining novel pumps[20,21]. In order to efficiently mine genomes during large genomic datasets, a multi-layer genome mining and phylogenomic analysis was developed. In the first layer, the general evolutionary recapitulation of target gene families was performed by comprehensive and systematic phylogenomics based on 397 genomes belonging to different representative prokaryotic

species. In the second layer, special species which exhibited great potential after experimental verification were selected for future engineering efforts, and target gene-centric phylogenies, which contained the genomic context surrounding target genes based on all the genomes derived from these species, was conducted. This allowed us to perform detailed analyses of how gene cluster architectures evolved from their constituent independent enzymes or sub-clusters. This multi-layer analysis would enable us to identify hidden regulons related to target genes. Hence, this multi-layer bioinformatic framework could help us to effectively screen uncharacterized heterologous target genes or pathways across large strain collections during genome mining.

MCFA efflux in organisms by nature could sense environmental changes in real time, and self-regulate cellular pathway fluxes, which would maximize product yields and minimize human supervision over the fermentation process control. Whereas current MCFA efflux systems required inducible promoters to conduct the transport function[19], and this was still incompatible with large-scale production[30,32,33]. In order to transform current MCFA efflux systems to an autonomous behavior eliminating inducer supplementation and human supervision, peptide pheromone responsive QS system of *Enterococcus faecalis* was combined with the efflux pumps. It was found that suitable triggering times of QEX circuity driving the efflux pumps yielded the best effect, and an early or delayed triggering of efflux pumps led to the decrease of extracellular or total MCFA concentrations, demonstrating the importance of examining the impact of different triggering times on efflux efficiency (Fig. 8b). This is, to our knowledge, the first report of autonomous and robust MCFA efflux system, and our autonomous secreting system would unleash microbial potential in target product synthesis, providing a valuable tool for advancing the field of high-value oleochemical research.

Supporting information

Detailed information regarding the construction of MCFA efflux pump library and autonomous MCFA secreting systems, experimental details on the quantitation of MCFAs, culture conditions and batch culture are shown. The results regarding the confirmation of expressing each predicted efflux pump, cell growth of engineered strains, DNA sequences of modified genes (Supplemental Table S3) are also presented.

ACKNOWLEDGMENTS

Thanks to Pablo Cruz-Morales (Senior researcher, DTU Biosustain) to help us with the phylogenomics analysis for mining the efflux pumps. This work was financially supported by Natural Science Foundation of Jiangsu Province (BK20202002), Excellent Youth Foundation of Jiangsu Scientific Committee (BK20211526), Jiangsu Agricultural Science and Technology Innovation Fund (SCX(20)3332), National Natural Science Foundation of China (No. 31972060), Fellowship of China Postdoctoral Science Foundation (2020T130305), Fundamental Research Funds for the Central Universities (KYGD202003), China Postdoctoral Science Foundation (2018M640491), Postdoctoral Research Funding of Jiangsu Province (2018K030B), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/FMR-2022-0015)

Dates

Received 8 August 2022; Accepted 7 September 2022; Published online 21 September 2022

REFERENCES

1. Sarria S, Kruyer NS, Peralta-Yahya P. 2017. Microbial synthesis of medium-chain chemicals from renewables. *Nature Biotechnology* 35:1158–66
2. Zhu Z, Hu Y, Teixeira PG, Pereira R, Chen Y, et al. 2020. Multidimensional engineering of *Saccharomyces cerevisiae* for efficient synthesis of medium-chain fatty acids. *Nature Catalysis* 3:64–74
3. Nagao K, Yanagita T. 2010. Medium-chain fatty acids: Functional lipids for the prevention and treatment of the metabolic syndrome. *Pharmacological Research* 61:208–12
4. Zentek J, Buchheitrenko S, Ferrara F, Vahjen W, Van Kessel AG, et al. 2011. Nutritional and physiological role of medium-chain triglycerides and medium-chain fatty acids in piglets. *Animal Health Research Reviews* 12:83–93
5. Keasling J, Garcia Martin H, Lee TS, Mukhopadhyay A, Singer SW, et al. 2021. Microbial production of advanced biofuels. *Nature Reviews Microbiology* 19:701–15
6. Yan Q, Pfleger BF. 2020. Revisiting metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering* 58:35–46
7. Bharathiraja B, Sridharan S, Sowmya V, Yuvaraj D, Praveenkumar R. 2017. Microbial oil - A plausible alternate resource for food and fuel application. *Bioresource Technology* 233:423–32
8. Wu H, Lee J, Karanjikar M, San KY. 2014. Efficient free fatty acid production from woody biomass hydrolysate using metabolically engineered *Escherichia coli*. *Bioresource Technology* 169:119–25
9. Reddy MV, Kumar G, Mohanakrishna G, Shobana S, Al-Raoush RI. 2020. Review on the production of medium and small chain fatty acids through waste valorization and CO_2 fixation. *Bioresource Technology* 309:123400
10. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R. 2011. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* 476:355–59
11. Clomburg JM, Vick JE, Blankschien MD, Rodríguez-Moyá M, Gonzalez R. 2012. A synthetic biology approach to engineer a functional reversal of the β-oxidation cycle. *ACS Synthetic Biology* 1:541–54
12. Clomburg JM, Blankschien MD, Vick JE, Chou A, Kim S, et al. 2015. Integrated engineering of β-oxidation reversal and ω-oxidation pathways for the synthesis of medium chain ω-functionalized carboxylic acids. *Metabolic Engineering* 28:202–12
13. Wu J, Zhang X, Xia X, Dong M. 2017. A systematic optimization of medium chain fatty acid biosynthesis via the reverse beta-oxidation cycle in *Escherichia coli*. *Metabolic Engineering* 41:115–24

14. Wu J, Zhang X, Zhou P, Huang J, Xia X, et al. 2017. Improving metabolic efficiency of the reverse beta-oxidation cycle by balancing redox cofactor requirement. *Metabolic Engineering* 44:313–24
15. Wu J, Wang Z, Duan X, Zhou P, Liu P, et al. 2019. Construction of artificial micro-aerobic metabolism for energy- and carbon-efficient synthesis of medium chain fatty acids in *Escherichia coli*. *Metabolic Engineering* 53:1–13
16. Doshi R, Nguyen T, Chang G. 2013. Transporter-mediated biofuel secretion. *PNAS* 110:7642–47
17. Dunlop MJ, Dossani ZY, Szmidt HL, Chu HC, Lee TS, et al. 2011. Engineering microbial biofuel tolerance and export using efflux pumps. *Molecular Systems Biology* 7:487
18. Lennen RM, Politz MG, Kruziki MA, Pfleger BF. 2013. Identification of transport proteins involved in free fatty acid efflux in *Escherichia coli*. *Journal of Bacteriology* 195:135–44
19. Wu J, Wang Z, Zhang X, Zhou P, Xia X, et al. 2019. Improving medium chain fatty acid production in *Escherichia coli* by multiple transporter engineering. *Food Chemistry* 272:628–34
20. Medema MH, de Rond T, Moore BS. 2021. Mining genomes to illuminate the specialized chemistry of life. *Nature Reviews Genetics* 9:553–71
21. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, et al. 2020. A computational framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology* 16:60–68
22. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, et al. 2018. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution* 35:543–8
23. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al. 2008. The RAST server: Rapid annotations using subsystems technology. *BMC Genomics* 9:75
24. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:1–19
25. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32:268–74
26. Wu J, Zhou L, Duan X, Peng H, Liu S, et al. 2021. Applied evolution: Dual dynamic regulations-based approaches in engineering intracellular malonyl-CoA availability. *Metabolic Engineering* 67:403–16
27. Zhang Y, Li SZ, Li J, Pan X, Cahoon RE, et al. 2006. Using unnatural protein fusions to engineer resveratrol biosynthesis in yeast and mammalian cells. *Journal of the American Chemical Society* 128:13030–31
28. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* 97:6640–45
29. Neidhardt FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. *Journal of Bacteriology* 119:736–47
30. Wu J, Bao M, Duan X, Zhou P, Chen C, et al. 2020. Developing a pathway-independent and full-autonomous global resource allocation strategy to dynamically switching phenotypic states. *Nature Communications* 11:5521
31. Salvador López JM, Van Bogaert INA. 2021. Microbial fatty acid transport proteins and their biotechnological potential. *Biotechnology and Bioengineering* 118:2184–201
32. Kim EM, Woo HM, Tian T, Yilmaz S, Javidpour P, et al. 2017. Autonomous control of metabolic state by a quorum sensing (QS)-mediated regulator for bisabolene production in engineered *E. coli*. *Metabolic Engineering* 44:325–36
33. Gupta A, Reizman IMB, Reisch CR, Prather KLJ. 2017. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology* 35:273–79