Integrated metagenomics identifies a crucial role for trimethylamine-producing *Lachnoclostridium* in promoting atherosclerosis

Yuan-Yuan Cai1,2,5, Feng-Qing Huang2,5, Xingzhen Lao1, Yawen Lu2, Xuejiao Gao3, Raphael N. Alolga2, Kunpeng Yin2, Xingchen Zhou1, Yun Wang1, Baolin Liu2, Jing Shang2,5, Lian-Wen Qi2,5 and Jing Li1,2,4,5

Microbial trimethylamine (TMA)-lyase activity promotes the development of atherosclerosis by generating of TMA, the precursor of TMA N-oxide (TMAO). TMAO is well documented, but some cannot be said about TMA-producing bacteria. This work aimed to identify TMA-producing genera in human intestinal microbiota. We retrieved the genomes of human-associated microorganisms from the Human Microbiome Project database comprising 1751 genomes, Unified Human Gastrointestinal Genome collection consisting 4644 gut prokaryotes, recapitulated 4930 species-level genome bins and public gut metagenomic data of 2134 individuals from 11 populations. By sequence searching, 216 TMA-lyase-containing species from 102 genera were found to contain the homologous sequences of *cntA/B*, *yeaW/X*, and/or *cutC/D*. We identified 13 strains from 5 genera with *cntA* sequences, and 30 strains from 14 genera with *cutC* showing detectable relative abundance in healthy individuals. *Lachnoclostridium* (p = 2.9e–05) and *Clostridium* (p = 5.8e–04), the two most abundant *cutC*-containing genera, were found to be much higher in atherosclerotic patients compared with healthy persons. Upon incubation with choline (substrate), *L. saccharolyticum* effectively transformed it to TMA at a rate higher than 98.7% while that for *C. sporogenes* was 63.8–67.5% as detected by liquid chromatography-triple quadrupole mass spectrometry. In vivo studies further showed that treatment of *L. saccharolyticum* and choline promoted a significant increase in TMAO level in the serum of ApoE−/− mice with obvious accumulation of aortic plaque in same. This study discloses the significance and efficiency of the gut bacterium *L. saccharolyticum* in transforming choline to TMA and consequently promoting the development of atherosclerosis.

npj Biofilms and Microbiomes (2022) 8:11; https://doi.org/10.1038/s41522-022-00273-4

INTRODUCTION

Numerous studies on intestinal microbiota over the past ten years have confirmed their pivotal roles in human health and diseases1,2,3. Cardiovascular diseases remain the leading cause of mortality and morbidity worldwide, accounting for 17.8 million deaths globally4. Accumulation of cholesterol on the arterial wall, immune responses, and chronic inflammation have been reported as major biological events in atherosclerosis5. Recently, the gut microbiota as a whole, has recently been found to be an important contributor to the progression of cardiovascular diseases6,7,8. The gut microbiota can regulate host cholesterol homeostasis aside from the host genetic contribution8. Increasing evidence demonstrates that gut microbiota-derived metabolites, such as short-chain fatty acids9, secondary bile acids10, and lipopolysaccharide11, play crucial roles in the development of atherosclerosis.

Trimethylamine (TMA) is a small molecular weight byproduct of intestinal microbial metabolism of dietary choline, carnitine, and phosphatidylcholines12–14. TMA is converted to trimethylamine-N-oxide (TMAO) in the liver via flavin-containing monoxygenase 315–17. A link between TMAO and cardiovascular diseases first emerged in 2011, when investigators found a dose-dependent association between plasma concentrations of TMAO and risk of cardiovascular diseases among persons with heart diseases15. A growing number of studies have since confirmed plasma TMAO as an independent risk factor for atherosclerosis, thrombus forma-

The probable mechanisms by which TMAO contributes to cardiovascular diseases could involve enhanced foam cell formation15, enhanced activation of platelets with increased calcium release16, and adverse ventricular remodeling17. In addition, TMAO partly accounts for exacerbated atherosclerosis by promoting forward cholesterol transport and by inhibiting the reverse transport of same13,15,20.

TMA, a precursor of TMAO, is produced by gut commensals using three different enzyme complexes. Choline TMA-lyase (cutC/D) which was discovered from the anaerobic sulfate-reducing *Desulfovibrio desulfuricans*, uses choline to produce TMA18. CutC is a specific glycyl radical enzyme, while cutD is its activator19. Two-component Rieske-type oxygenase/reductase (cutA/B), which uses y-butyrobetaine as a substrate to generate TMA12,13. In spite of available credible information on TMAO-producing bacteria24–26 and their importance in human health, little is known about their abundance in different populations, their relationship with diseases, and TMAO conversion abilities.

In this study, using an integrated metagenomic approach, we analyzed reference genomes from the three reference

1School of Life Science and Technology, China Pharmaceutical University, Nanjing 210009, China. 2State Key Laboratory of Natural Medicines, School of traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing 210009, China. 3Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing 210028, China. 4Key Laboratory of Drug Quality Control and Pharmacovigilance (China Pharmaceutical University), Ministry of Education, Nanjing 210009, China. 5These authors contributed equally: Yuan-Yuan Cai, Feng-Qing Huang. 6Email: shangjing21cn@163.com; Qilw@cpu.edu.cn; lj_cpu@126.com
RESULTS
Taxonomic identification of TMA-producing bacteria
The gut microbiota is known to convert carnitine, betaine, and choline to TMA by utilizing three TMA-lyase complexes, cntA/B, yeaW/X, and cutC/D, respectively. CntA, yeaW, and cutC are TMA-lyases, while cntB, yeaX, and cutD are their activators (Fig. 1a). We identified the TMA-lyase sequences of the TMA-producing bacteria based on the three reference resources.

The final 885 TMA-lyase reference sequences were identified from three resources, 451 from the Human Microbiome Project database, 216 from the Unified Human Gastrointestinal Genome, and 218 reference sequences from 4930 species-level genome bins by taking the initial 983 sequences as query (Supplementary Table 1). These sequences were distributed in 216 bacteria species from 102 genera as shown in Supplementary Table 2. Specifically, cntA existed in 15 genera, cntB in 20 genera, yeaW in 15 genera, and yeaX in 17 genera (Fig. 1b). All these genera belong to the phylum of Proteobacteria (Fig. 1b). CutC was distributed in 3 genera from Actinobacteria, 27 from Firmicutes, 18 from Proteobacteria and 4 from other phyla like Bacteroidetes, Desulfovibacterota_A, and Fusobacteria. CutD was distributed in 4 genera from Actinobacteria, 39 from Firmicutes, 30 from Proteobacteria and 10 from other phyla like Bacteroidetes, Desulfovibacterota_A, Dictyoglomi, Fusobacteria, Spirochaetes, Thermotoga, and Verrucomicrobia (Fig. 1b). Furthermore, cntA/B and yeaW/X coexisted in 15 genera, while cutC/D coexisted in 48 genera (Fig. 1c).

It is worth noting that cntA/B and yeaW/X coexisted in 49 species (Supplementary Table 2). For sequence similarity, the average sequence identity was 81.64% for cntA and yeaW, and 65.63% for cntB and yeaX (Fig. 1d). Compared to cntA/B and yeaW/X, the sequence identity of cutC/D was relatively low; 32.27% for cutC and cntA/yeaW, and 33.60% for cutD and cntB/yeaX (Fig. 1d). Since the TMA-lyases cntA/yeaW/cutC usually coexisted with their corresponding activators cntB/yeaX/cutD and there was high homologous similarity between cntA/B and yeaW/X rather than cutC/D, we focused our attention on cntA and cutC in subsequent analysis.

Relative abundance of cntA and cutC in the human intestine
An integrated reference metagenomic sequence dataset was acquired from several public cohorts (Supplementary Table 3). Multivariable-adjusted analysis showed that population played a major role in the relative abundance (R.A.) of cntA (27.15% contribution to total variations) and cutC (21.02% contribution to total variations, Supplementary Fig. 1). A total of 645 healthy individuals worldwide were included to figure out the R.A. of TMA-producing bacteria with sequence alignment. Five genera containing cntA were found in the human intestine, and Escherichia occupied 66.5% followed by Klebsiella with 17.7% and then Shigella with 15.4% (Fig. 2a). The cumulative R.A. of cntA varied widely among populations ($\rho_{\text{org}} = 0.013$). Swedish exhibited the highest cumulative R.A. of cntA ($4.55 \times 10^{-3}$), while Americans exhibited the lowest cumulative R.A. ($2.48 \times 10^{-9}$) (Fig. 2a). For cutC, the top five most prevalent cutC-carrying genera were Lachnoclostridium (22.8%), Desulfovibrio (13.5%), Klebsiella (12.5%), Clostridium (11.2%) and Escherichia (7%) (Fig. 2b). The cumulative R.A. of cutC also varied largely among populations ($\rho_{\text{org}} = 6.1e-06$). Japanese showed the highest cutC cumulative R.A. ($2.66 \times 10^{-5}$), while Swedish showed the lowest cumulative R.A. ($2.02 \times 10^{-5}$).

Subsequence analysis of the TMA-producing bacterial strains was based on the Human Microbiome Project which has defined strain taxonomic information among the three reference sources. We conducted phylogenetic analysis of cntA and cutC in candidate TMA-producing bacterial strains. Totally, 13 strains were found to have detectable R.A. containing cntA, 8 of which belonged to the dominant genera Escherichia (Fig. 2c, left panel). In the phylogenetic classification, the cntA from Shigella was close to Escherichia. Two strains from Klebsiella and one strain from Pseudomonas were farther on the phylogenetic tree, suggesting that their genetic sequences differ widely. The R.A. of cntA in the 13 strains ranged from $1.62 \times 10^{-6}$ to $2.55 \times 10^{-5}$ and the prevalence ranged from 1.90 to 6.72% (Fig. 2c, right panel).

CutC existed in 30 strains with detectable R.A. in three phyla, Actinobacteria, Firmicutes, and Proteobacteria. There were three main branches in the phylogenetic tree of cutC. The first branch was mainly from Clostridium (R.A., $3.71 \times 10^{-6} - 7.54 \times 10^{-6}$ and prevalence, 2.20–7.80%). The next branch was mainly from Lachnoclostridium (R.A., $4.54 \times 10^{-6} - 1.72 \times 10^{-5}$ and prevalence, 3.18–5.53%) and Desulfovibrio (R.A., $3.80 \times 10^{-7} - 4.61 \times 10^{-7}$ and prevalence, 2.24–4.66%). The third branch was mainly from Escherichia (R.A., $1.06 \times 10^{-6} - 3.99 \times 10^{-6}$ and prevalence, 1.30–3.41%) and Klebsiella (R.A., $2.25 \times 10^{-6} - 1.30 \times 10^{-5}$ and prevalence, 1.73–4.97%). Among them, the highest genus containing cutC was Lachnoclostridium (Fig. 2d).

Correlation of cntA and cutC with various diseases
We then explored the correlation of cntA and cutC with various diseases such as, adenoma, colorectal cancer, impaired glucose tolerance, type 2 diabetes, cardiovascular disease, hypertension, and obesity using several public case-control gut metagenomic datasets (Supplementary Table 3). Interestingly, the R.A. of cntA showed no significant differences in the various disease groups compared with healthy individuals. There was an elevated trend but no statistically significant difference between persons with atherosclerosis and the healthy individuals (Supplementary Fig. 2). Surprisingly, cutC showed significantly increased R.A. in the patients with atherosclerosis (Mann–Whitney U test $p^a = 0.033$) and a significant difference in prevalence of same (Chi-squared test $p^b = 0.0091$) compared with healthy individuals. Moreover, cutC showed significantly increased R.A. in the patients with obesity (Mann–Whitney U test $p^a = 0.013$) and a significant difference in prevalence (Chi-squared test $p^b = 0.033$) compared with healthy individuals (Fig. 3a). No significant differences in the R.A. of cutC were observed in other diseases compared with the healthy controls. The prevalence of cutC was significant higher in patients with colorectal cancer compared with healthy individuals (Chi-squared test $p^b = 0.045$).

In further analysis of the correlation between the risk of atherosclerosis and cutC-containing genera, we observed that the R.A. of Lachnoclostridium (Mann–Whitney U test $p^a = 2.9e-05$), Clostridium ($p^a = 5.8e-04$) and Olsenella ($p^a = 0.018$) were significantly increased in patients with atherosclerosis compared with the healthy persons (Fig. 3b). Other genera showed no significant differences in the R.A. between the atherosclerotic patients and healthy individuals (Fig. 3b). Lachnoclostridium and Clostridium showed highest prevalence in this cohort. The prevalence of Escherichia (Chi-squared test $p^b = 1.6e-06$, Klebsiella ($p^b = 0.00046$) and Desulfovibrio ($p^b = 0.024$) were significant different. The genera Anaerococcus, Hungatella, and Proteus were not detected in this dataset.
In silico and in vitro comparisons of candidate TMA-producing genera

Since cutC but not cntA was significantly increased in patients with atherosclerosis, we then compared the TMA-producing ability of 5 candidate cutC-containing genera with high R.A. We performed homologous modeling, molecular docking, and dynamic simulation analysis between choline and cutC protein from different strains (specific sequence is shown in Supplementary Table 4). The homologous modeled structure of cutC from L. saccharolyticum WM1 gave the lowest binding energy (-4.97 kJ/mol) with choline compared with L. asparagiforme DSM 15981 (-4.49 KJ/mol) and L. phytofermentans ISDg (-4.41 KJ/mol). In addition, the binding energy between choline and cutC from L. saccharolyticum WM1 was lower than the other four cutC-containing genera (from -4.26 to -4.82 KJ/mol) (Fig. 4a). The seven active site residues of cutC were all conserved among these cutC-containing genera (Fig. 4b).

Fig. 1  Taxonomic distribution of TMA-producing bacteria. a Schematic diagram of three microbial transformation pathways of TMA-lyases, cntA/B, yeaW/X, and cutC/D. b The phylogeny and genera distribution of cntA/B, cutC/D, and yeaW/X. The plus signs indicate that the genus contains the genes which encoded the corresponding TMA-lyase. c Venn diagram analysis of the coexisting genera distribution of cntA/B, cutC/D, and yeaW/X. d Homology trees of three TMA-lyases and their corresponding activators. The numbers on the corner represent the average sequence identities between the TMA-lyases connected by the branch. TMA: trimethylamine.

In silico and in vitro comparisons of candidate TMA-producing genera

Fig. 4a. The seven active site residues of cutC were all conserved among these cutC-containing genera (Fig. 4b). However, the site residues close to the seven active sites, such as V225, M344, V404, and M406, appeared to differ.

Y.-Y. Cai et al.
Published in partnership with Nanyang Technological University
npj Biofilms and Microbiomes (2022) 11

Published in partnership with Nanyang Technological University

npj Biofilms and Microbiomes (2022) 11
Fig. 2  R.A. distribution and taxonomic identification of cntA and cutC. R.A. distribution of cntA (a) and cutC (b) in 11 populations from six continents. Stack columns display the cumulative R.A. of different genera and pie charts in the corner show the distribution of the abundant genera. The detailed 11 populations from six continents include: Africa: HZ (the Hadza ethnic group of Tanzania); Asia: CN (China), JP (Japan), and KR (South Korea); Europe: DK (Denmark), SE (Sweden), AT (Austria), and FR (France); Oceania: AU (Australia); North America US (the United States); South America: PE (Peru). Significant differences in R.A. were conducted by Kruskal–Wallis rank sum test. Taxonomic identification of cntA (c) and cutC (d). The left panels show the phylogenetic tree and the highlighted branches correspond to the colors of different genera on the middle panel. The bar charts on the right panel indicate the average R.A. and the prevalence of cntA and cutC among different intestinal strains. R.A.: relative abundance. Error bars denote the means ± s.e.m.
among these genera, possibly accounting for the differences in the cutC enzyme activity. The binding conformation between choline and cutC from L. saccharolyticum WM1 was relatively stable within 100 ns by the dynamic simulation (Fig. 4c). In this binding conformation, choline exhibited three hydrogen bonds with two key residues, i.e., Cys497 and Glu499 (Fig. 4d). These findings demonstrated that L. saccharolyticum WM1 exhibited favorable binding affinity to substrate choline. The binding conformation of cutC from other candidate cutC-containing genera are showed in Supplementary Fig. 3.
L. saccharolyticum WM1 was chosen for subsequent in vitro- and in vivo-experiments. The TMA-producing activities of several cutC-containing strains were then investigated by incubating each strain with the substrate choline. The concentration of TMA was accurately quantified using an isotope-labeled internal standard method by liquid chromatography-triple quadrupole mass spectrometry. L. saccharolyticum WM1, effectively converted choline to TMA in a time-dependent manner at 12, 24, 36, 48, and 96 h (Fig. 4e). The conversion rates of L. saccharolyticum WM1 for various concentrations of choline (2 mM, 4 mM and 6 mM) were higher than 98.68% (Fig. 4f). C. sporogenes BNCC 104015 showed moderate conversion rates for choline (63.79−67.51%). B. fragilis BNCC 352061 and P. intermedia BNCC336948, two negative control strains without cutC enzymes, showed no capability in transforming choline to TMA (less than 0.1%) (Fig. 4f).

L. saccharolyticum promotes the development of atherosclerosis in ApoE−/− mice

After showing that L. saccharolyticum effectively converted choline to TMA by in vitro incubation, we then investigated the possible implication of L. saccharolyticum in atherosclerosis in ApoE−/− mice. We compared the atherosclerotic phenotypes of mice in the following four groups: normal group, normal + L. saccharolyticum group, choline group, and choline + L. saccharolyticum group (Fig. 5a). L. saccharolyticum administration notably elevated the level of L. saccharolyticum in fecal samples (Fig. 5b). The serum levels of TMAO, the downstream metabolite of TMA, significantly increased in the choline + L. saccharolyticum group compared with the normal group (p = 9.9e−04), normal + L. saccharolyticum group (p = 3.4e−05) and choline group (p = 5.3e−04) (Fig. 5c). Compared with the normal control, the plaques and the stenosis of the blood vessels showed an increased trend in the L. saccharolyticum group and the choline group. The plaques and the stenosis of aortic arch were greatly pronounced in the choline + L. saccharolyticum group (Fig. 5d). More importantly, the level of L. saccharolyticum in the fecal samples was positively correlated with TMAO level in serum (R = 0.31, p = 0.0052), atherosclerosis area (R = 0.54, p = 0.00242) and vessel lesion area (R = 0.48, p = 0.0085, Supplementary Fig. 4a). Administration of choline increased the thickness of the vessel wall and inflammatory cells infiltration and lipid deposition (Fig. 5e). These changes were more obvious in the choline + L. saccharolyticum-treated mice (Fig. 5e). L. saccharolyticum and choline administration triggered inflammation in the vessel, as indicated by gene inductions of Il-1β, Tnf-α, Icam-1 and Vcam-1, Mcp-1, Cdf68, and F480 (Fig. 5f). Similarly, L. saccharolyticum administration increased circulating contents of FFAs, while other metabolic parameters remained unaffected (Supplementary Fig. 4b). These results indicated that L. saccharolyticum abundance impaired the aorta and favored the formation of atherosclerosis, probably due to the increased production of TMAO.

DISCUSSION

In this work, we implemented an integrated metagenomics approach combined with bioinformatics and in vitro/in vivo validation to characterize a key role for TMA-lyase bacteria in the promotion of atherosclerotic lesion formation. The major findings of this study include the following: (1) We conducted a taxonomic analysis based on three reference sources and identified 216 TMA-producing species from 102 genera. (2) By analyzing 2134 individuals from 11 populations using public metagenomics data, we identified 13 cntA-containing and 30 cutC-containing bacterial strains. Correlation analysis of TMA-lyase bacteria and various diseases showed that Lachnoclostridium, the most abundant cutC-containing genus in healthy individuals, was significantly increased in atherosclerotic patients. (3) L. saccharolyticum WM1, a representative strain of Lachnoclostridium, effectively converted choline to TMA at a transformation rate near 100%, much higher than 63% for C. sporogenes. This strain was found to elevate serum TMAO and promote the formation of atherosclerosis in ApoE−/− mice when co-administered with choline. Uncovering the TMA-producing bacteria and the conversion capabilities of the human gut microbiota may predict risk of developing atherosclerosis1. Falony et al. identified 102 genomes showing TMA-producing potential by mining public genomic databases25. With emerging updated data on human metagenomics, we performed a more comprehensive taxonomic analysis and identified 216 TMA-producing species from 102 genera. Romano and colleagues identified 8 strains from 79 human intestinal isolates showing significant transformation of choline to TMA26. In line with this observation, 7 of the 8 strains were found in the 30 cutC-containing strains identified in the present study. The remaining one isolate, thus E. tarda ATCC 23685 that was neglected in this study, as it does not appear to contain cutC genes.

It is well established that the risk of developing cardiovascular diseases is associated with genetic variations29, dietary factors31, and TMAO concentration in blood15,16. Factors that determine the TMA-lyase abundance in the human gut metagenome are still not fully characterized. We observed that population type played a major role in the R.A. of cntA and cutC, age had a slight impact, and BMI and gender had no observable effect. Population structure is also associated with several confounding factors such as environment, dietary factors, and host genetic structures. Because of the significance of population structure in the R.A. of TMA-lyase, it was crucial to employ the worldwide open public datasets of the 11 populations captured in this work in order to avoid biases.

In the correlation analysis, we found that the R.A. of cntA showed no significant differences between atherosclerotic patients and healthy individuals, an observation that is consistent with a previous study3. This finding may be attributed to the limited depth of sequencing and query set. In the prevalence analysis of cutC, we found a significant increase in the colorectal cancer patients which is consistent with the findings of Thomas et al. who found cutC to be overabundant in colorectal cancer32. It is worth noting that the R.A. of cutC was markedly elevated in atherosclerotic patients but not in other diseases. Our result is contrary to that of Jie et al. who reported no significant difference in the R.A. of cutC between atherosclerotic patients and healthy persons using the same cohort3. This divergence in final outcome
could be attributed to the 575 query sequences of the cutC enzyme retrieved in this work juxtaposed with the 13 sequences employed by Jie et al. Microbial transplantation of cutC-containing commensals could promote atherosclerosis development and thrombosis. Skye et al. showed that transplantation of Clostridium sporogenes, a cutC-expressing human commensal, was sufficient to transmit thrombosis potential in a host. Consistently, we confirmed that Clostridium sporogenes transformed choline to TMA in vitro at a moderate conversion rate of 63.8%–67.5%. More importantly, this work is the first to identify Lachnoclostridium as the most abundant cutC-containing genus, and show that L. saccharolyticum converts choline to TMA at a rate that is almost 100%.

Fig. 4  In silico and in vitro comparisons of candidate TMA-producing genera. a The binding energies of cutC-containing strains and choline based on homologous modeling. b The 7 site residues of cutC in TMA-producing genera. Conserved amino acid sites are marked in red boxes with amino acid locations indicated above. The secondary structure is represented by background color, yellow represents beta folding, and purple represents alpha spiral. The asterisks indicate hydrogen bond interactions with choline, and the paper clip symbol in the middle indicate the presence of a card-issue structure. c The binding conformation complex of cutC from L. saccharolyticum and choline was stable at 100 ns dynamic equilibrium. d The AutoDock predicted binding conformation complex (from L. saccharolyticum) within which the protein is shown as cartoon and choline as sticks. The residues that have important hydrogen-bond interactions with choline are labeled and these hydrogen bonds are represented as dashed lines. e The typical chromatogram of TMA produced by L. saccharolyticum incubated with choline at different times are shown in the left panel, and the corresponding conversion rate curve is shown in the right panel. f Barplots show the conversion rates of strains from TMA-producing genera based on various choline concentrations, 2, 4, and 6 mM (n = 3). TMA production was assessed using LC-MS. Error bars denote the means ± s.e.m. *p < 0.05, **p < 0.01, ***p < 0.001, differences were conducted by Mann–Whitney–Wilcoxon test.
in the development of atherosclerosis. Our avoidance of high-fat diet in our animal model may have been responsible for the observed non-significant changes in the serum cholesterol levels.

This study has some limitations. (1) The integrated metagenomic analysis based on public metagenome datasets constitutes the primary limitation. All datasets were subjected to the same pipeline in conducting the raw sequencing data, but there might be some biases due to the different quality of each dataset, such as sequencing depth and sample sizes. (2) The effect of Lachnoclostridium on the progress of atherosclerosis was verified using a mouse model, thus a prospective cohort study is necessary in the further studies.

In conclusion, our work provides an integrated metagenomic analysis of the TMA-producing bacteria in the human gut and
reports for the first time a TMA-producing genus, Lachnoclostridium, which was significantly abundant in atherosclerotic patients. Specifically, we verified that L. saccharolyticum WM1 could produce TMA in vitro, increase the TMAO level in serum, and accelerate plaque formation in vivo. Hence, targeting Lachnoclostridium might serve as a potential therapeutic target for the treatment of atherosclerosis.

METHODS
Bioinformatic identification of the TMA-lyase sequences
The first reference genomes resource was from the Human Microbiome Project in September 2014 which contained 1751 bacterial strains covering 1253 species. The second reference genomes resource, thus, the unified Human Gastrointestinal Genome, comprised of 204,938 nonredundant genomes from 4644 gut prokaryotes, and the third reference genomes resource contained recapitulated 4930 species-level genome bins (Supplementary Fig. 5). The genes and related proteins of these bacterial genomes were predicted by MetaGeneMark (v2.8)38, and the taxonomic information of each protein sequence of each individual was based on two steps: (i) Calculation of the entire copies of each gene with correct insert-size; (ii) Calculation of the R.A. of each gene in each individual, only if the other reads mapped outside the genomic region, as previously described in detail42. The cumulative R.A. was calculated by the sum of R.A. of each genus in each population. TMA-lyase was identified in the above population datasets using BLASTP with the parameters: -e-value 1e−5 -outfmt 0 -max_target_seqs 3 and 885 query sequences of e-value 1e−5 and 80% sequence identity as the cutoff values. Persons without zero R.A. were employed for the comparisons within the various cohorts.

Phylogenetic tree
The protein sequences of cutA and cutC were aligned using mafft v7.4552, and the resultant multiple sequences were trimmed for poorly aligned positions with Gblock 9.1b53. RAxML v8.2.1254 was used to build the most likely phylogenetic tree of genomes with the parameters -m GTRCAT -n PROTGAMMAILGX. R package “ggtree”55 was used to construct the phylogenetic tree.

Homology modeling, molecular docking, and molecular dynamics
Protein Homology/analogy Recognition Engine V 2.0456 based on intensive mode was used to predict the homologous structures of cutC in strains from Lachnoclostridium, Clostridium, and other abundant cutC-containing genera. The details of cutC sequence used for modeling are provided in Supplementary Table 4. Choline ligand was downloaded from ZINC database57. The AutoDock (v.4.2.6)58 was employed to generate an ensemble of docked conformations for each ligand bound to its target. We used the genetic algorithm for conformational search and performed 100 individual GA runs to generate docked conformations for each ligand. For each docked conformation, the top-ranked docking pose was optimized in the binding pocket and used as the initial geometry in molecular dynamics. We performed 100 ns molecular dynamics simulation for choline-bound states. All of the molecular dynamics simulations were performed with Amber1659.

Bacterial strains and growth conditions
To verify the in vitro activity of TMA-producing bacteria, Lachnoclostridium saccharolyticum WM1 (ATCC 35040) was cultured anaerobically on ATCC medium 1118 at 37 °C. Clostridium sporogenes (BNCC 104015) was cultured anaerobically in Trypticase soy agar/broth with defibrinated sheep blood at 37 °C. Two representative intestinal strains, Prevotella intermedia (BNCC 352061) and Bacteroides fragilis (BNCC336948), were selected as negative controls and were cultured in anaerobiologically sterilized Trypticase soy agar/broth with defibrinated sheep blood at 37 °C.

For TMA assessment, bacterial cultures were inoculated with choline chloride (Sigma-Aldrich, USA) in sterile tubes and incubated anaerobically (80% N2 and 20% CO2) at 37 °C (typically OD600nm ~ 1.0). The cultures were centrifuged at 16,200 × g for 10 min at 4 °C.

Quantitative assessment of TMA in bacterial culture medium
Quantification of TMA was performed on a liquid chromatography-2A system coupled to a triple quadrupole mass spectrometer (Shimadzu 8050, Japan) using an internal standard method16. Briefly, 10 μL of bacterial culture medium was first mixed with 10 μL of isotope-labeled d9-TMA internal standard (100 μg/mL). Then, 960 μL of n-hexane/n-butyl alcohol (2:1, v/v) and 20 μL of 1 M NaOH were added and vortexed for 3 min to fully extract TMA. After centrifugation (16200 × g, 4 °C, 1 min), 500 μL of the organic layer was transferred to a sealed container and acidified with 200 μL of 0.2 N formic acid. Finally, the mixture was vortexed for 3 min and centrifuged at 4 °C for 1 min. An aliquot of 1 μL of the aqueous layer was injected for LC-MS/MS analysis.

LC separation was carried out on a HILIC column (100 × 2.1 mm, 1.7 μm; WATERS). Pure water and acetonitrile both containing 0.1% formic acid were used as mobile phases A and B, respectively, delivered at a flow rate of 0.4 mL/min. An isocratic elution with phase B/phase A (80:20, v/v) was employed and the detection was operated in the positive ion mode. Multiple reaction monitoring (MRM) transitions were performed at m/z 60.08 → 44.04 for quantitative detection of TMA, and m/z 69 → 49 for d9-TMA. The ESI source parameters were as follows: nebulizing gas flow, 3 L/min; heating gas flow, 10 L/min; drying gas flow, 10 L/min; interface temperature, 300 °C; DL temperature, 250 °C; and
heat block temperature, 400 °C. Multiple reaction monitoring transitions were performed at m/z 60.08 → 44.04 for quantitative detection of TMA, and m/z 69 → 49 for d9-TMA.

**Animal study**

The care and treatment of mice were performed in accordance with the Provisions and General Recommendation of Chinese Experimental Animals Administration Legislation and the study was approved by the Animal Ethics Committee of China Pharmaceutical University (No. 2021-04-002). All procedures conformed to the European Parliament directive on the protection of animals used for scientific purposes (Directive 2010/63/EU) or the NIAID Guide for the Care and Use of Laboratory Animals (NIH Guide for the Care and Use of Laboratory Animals, 8th Edition, National Institutes of Health, 2011). Five to six mice were housed in a pathogen-free environment at a temperature of 22–24 °C, humidity 40–60%, and a 12 h light cycle. After acclimatization for a week, both female and male ApoE −/− mice were randomized into 4 groups: normal group, mice fed standard chow and gavaged with the sterile medium; normal + L. saccharolyticum group, mice fed standard chow and gavaged live L. saccharolyticum at a dose of 5 × 1010 CFUs/100 μl; choline group, mice fed chow supplemented with 1.0% choline (Sigma-Aldrich, USA) and gavaged with the sterile medium; and normal + L. saccharolyticum group, mice fed chow supplemented with 1.0% choline and gavaged with live L. saccharolyticum at a dose of 5 × 1010 CFUs/100 μl. The mice number was 20 (half male and half female) for each group. The group sizes were selected according to the minimum experimental requirements and natural factors such as flight-related injury. After 20 weeks, overnight-fasted mice were anesthetized with 4% isoflurane and their blood was collected from the ventriculus dexter for the assay of glucose (Solarbio, BC2505), non-esterified free fatty acids (Solarbio, BC0595), triglycerides, total cholesterol, low-density lipoprotein, and high-density lipoprotein using commercial Kits (Jiancheng Biotechnology Co., Ltd., Nanjing, China). Their stool samples were collected and immediately stored at −80 °C until further analysis. After that, the mice were euthanized by exsanguination following the guidelines of the Institutional Animal Care, the cardiac and aortic tissues were immediately removed, rinsed with ice-cold physiological saline, and stored at −80 °C or fixed in 4% paraformaldehyde until further analysis.

**Histological analysis**

Cardiac and aortic tissues were fixed with 4% (w/v) paraformaldehyde overnight. Then the tissues were embedded in paraffin, cut into 5 μm slices, and stained with hematoxylin and eosin or oil red O. Images were captured using NanoZoomer 2.0 (Hamamatsu, Japan).

**Assessment of atherosclerotic plaque size**

Perivascular adipose tissue and adventitious blood vessels were carefully removed and the prepared aorta was stained with Sudan IV solution (0.1% Sudan in 50% aceton) for 6 min, followed by de-staining in 80% ethanol for 5 min. The images were captured with a Canon EOS 80D Digital Camera. Areas stained red were considered atherosclerotic lesions and were quantified using ImageJ software.

**Quantitation of L. saccharolyticum abundance in feces**

Fecal samples of each mouse were used for metagenomic DNA extraction using TIANamp Stool DNA Kit according to the manufacturer’s instructions. DNA concentration was measured using NanoDrop 2000 spectrophotometer (Biotek, Germany). Quantitative PCR was carried out using the LC480 detection system (Roche Diagnostics, Basel, Switzerland) and Hiscript Q RT SuperMix (Vazyme biotech, Nanjing, China). Primers used in this work are listed in Supplementary Table S5.

**Quantification of TMAO in serum**

An isotope-labeled internal standard method was employed for the quantitative analysis of TMAO in serum. Targeted detection of TMAO was performed on an Agilent 1290 infinity liquid chromatography system coupled to a triple quadrupole mass spectrometer (Agilent 6470, USA) operated in the positive ion mode. An aliquot of 10 μl serum was precipitated by adding 980 μl of acetonitrile/ water/formic acid (94:5:1, v/v/v) solution and 10 μl of internal standard working solution (2 μg/ml of d9-TMAO), followed by vortex-mixing for 30 s. Precipitated protein was subsequently removed by centrifugation at a speed of 16200 × g at 4 °C for 10 min. Then, 5 μl of supernatant was analyzed by ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS).

Chromatographic evaluation was achieved on an ACQUITY UPLC BEH HILIC column (2.1 × 100 mm, 1.7 μm) maintained at 40 °C. The mobile phase consisted of (A) 10 mM ammonium acetate aqueous solution and (B) 10 mM ammonium acetate/water/acetonitrile (1:9) solution delivered at a flow rate of 0.4 mL/min. The gradient elution program was 5–80% B at 0–7 min, 80–100% B at 7–12 min, 100% B at 12–13 min, and then back to initial conditions, with 2 min for equilibration. Targeted detection of TMAO was performed on an Agilent 1290 infinity LC system coupled to a triple quadrupole mass spectrometer (Agilent, LC-MS/MS 6470) operating in the positive ion mode. Multiple reaction monitoring (MRM) mode was employed and the ion transitions monitored were m/z 76 → 58 for TMAO and 85 → 66 for d9-TMAO. The detailed MS parameters were set as follows: fragmental voltage, 100 V; capillary voltage, 3500 V; nebulizer gas, 35 psi; drying gas flow rate, 10 L/min; drying gas temperature, 300 °C.

**Statistical analysis**

Statistical differences between the two groups were calculated using the Wilcoxon rank-sum test. The multivariable-adjusted analysis using linear regression model was performed to evaluate the R.A. of TMA-lyase against other confounding factors. The Kruskal–Wallis test was for multi-group comparison followed by Holm–Bonferroni correction. The differences among groups in animal experiments were analyzed by single-factor ANOVA with Tukey HSD test. All analyses were performed using R 3.6.1, and p < 0.05 were considered statistically significant.

**DATA AVAILABILITY**

All public metagenomic data used in this manuscript were provided their web links or the accession codes in “Methods” section.

**CODE AVAILABILITY**

All code used for statistical analysis is available from the corresponding author on reasonable request.

Received: 27 June 2021; Accepted: 31 January 2022; Published online: 10 March 2022

**REFERENCES**

1. Jonsson, A. L. & Backhed, F. Role of gut microbiota in atherosclerosis. Nat. Rev. Cardiol. 14, 79–87 (2017).
2. Cho, I. & Blaser, M. J. The human microbiome: At the interface of health and disease. Nat. Rev. Genet. 13, 260–270 (2012).
3. Rath, S., Rud, T., Karch, A., Pieper, D. H. & Vital, M. Pathogenic functions of host microbiota. Microbiome 6, 174 (2018).
4. Virani, S. S. et al. Heart disease and stroke statistics-2020 update: A report from the American Heart Association. Circulation 141, e139–e596 (2020).
5. Hansson, G. K. Inflammation, atherosclerosis, and coronary artery disease. N. Engl. J. Med. 352, 1685–1695 (2005).
6. Jie, Z. et al. The gut microbiome in atherosclerotic cardiovascular disease. Nat. Commun. 8, 945 (2017).
7. Tang, W. H. & Hazen, S. L. The contributory role of gut microbiota in cardiovascular disease. J. Clin. Invest. 124, 4204–4211 (2014).
8. Le Roy, T. et al. The intestinal microbiota regulates host cholesterol homeostasis. BMC Biol. 17, 94 (2019).
9. Kasahara, K. et al. Interactions between Roseburia intestinalis and diet modulate atherogenesis in a murine model. Nat. Microbiol. 3, 1461–1471 (2018).
10. Chen, M. L. et al. Resveratrol attenuates trimethylamine-N-oxide (TMAO)-induced atherosclerosis by regulating TMAO synthesis and bile acid metabolism via remodeling of the gut microbiota. mbio 7, e02210–e02215 (2016).
11. Yoshida, N. et al. Bacteroides vulgatus and bacteroides dorei reduce gut microbial lipopolysaccharide production and inhibit atherosclerosis. Circulation 141, 2486–2498 (2018).
12. Zhu, Y. et al. Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota. Proc. Natl Acad. Sci. USA 111, 4268–4273 (2014).
13. Koeth, R. A. et al. gamma-Butyrobetaine is a proatherogenic intermediate in gut microbiota. Cell Metab. 20, 799–812 (2014).

Y.-Y. Cai et al.
14. Craciun, S. & Balskus, E. P. Microbial conversion of choline to trimethylamine requires a glycolyl radical enzyme. *Proc. Natl Acad. Sci. USA* 109, 21307–21312 (2012).

15. Wang, Z. et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472, 57–63 (2011).

16. Bennett, B. J. et al. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab.* 17, 252–262 (2013).

17. Serino, M., Blasco-Baque, V., Nicolas, S. & Burcelin, R. Far from the eyes, close to the heart: dysbiosis of gut microbiota and cardiovascular consequences. *Curr. Cardiol. Rep.* 16, 540 (2014).

18. Zhu, W. et al. Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk. *Cell* 165, 111–124 (2016).

19. Organy, C. et al. Choline diet and its gut microbe-derived metabolite, trimethylamine-N-oxide, exacerbate pressure overload-induced heart failure. *Circ. Heart Fail* 9, e003144 (2016).

20. Koeth, R. A. et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat. Med.* 19, 576–585 (2013).

21. Thibodeaux, C. J. & van der Donk, W. A. Converging on a mechanism for choline degradation. *Proc. Natl Acad. Sci. USA* 109, 21184–21185 (2012).

22. Seim, H., Löster, H., Claus, R., Kleber, H. P. & Strack, E. Splitting of the C–N bond in carnitine by an enzyme (trimethylamine forming) from membranes of *Acinetobacter calcoaceticus*. *Fems Microbiol. Lett.* 15, 165–167 (1982).

23. Unemoto, T., Hayashi, M., Miyaï, K. & Hayashi, M. Formation of trimethylamine from di-carnitine by *Serratia marcescens*. *Biochem. Biophys. Acta* 121, 220–222 (1996).

24. Rath, S., Heidrich, B., Pieper, D. H. & Vital, M. Uncovering the trimethylamine-producing bacteria of the human gut microbiome. *Nat. Microbiol.* 5, 54 (2017).

25. Falony, G., Vieira-Silva, S. & Raes, J. Microbiome meets big data: The case of gut microbiota-derived trimethylamine. *Annu. Rev. Microbiol.* 69, 305–321 (2015).

26. Romano, K. A., Vivas, E. I., Amador-Noguez, D. & Rey, F. E. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *MBio* 6, e02481 (2015).

27. Human Microbiome Project, C. Structure, function, and diversity of the healthy human microbiome. *Nature* 486, 207–214 (2012).

28. Human Microbiome Project, C. A framework for human microbiome research. *Nature* 486, 215–221 (2012).

29. Craciun, S., Marks, J. A. & Balskus, E. P. Characterization of choline trimethylamine-lyase expands the chemistry of glycolyl radical enzymes. *ACS Chem. Biol.* 9, 1408–1413 (2014).

30. Yamada, Y. Identification of genetic factors and development of genetic risk diagnosis systems for cardiovascular diseases and stroke. *Circ. J. Off. J. Jpn. Circ. Soc.* 70, 1240 (2006).

31. Rodriguezmoneforte, M., Floresmateo, G. & Sánchez, E. Dietary patterns and CVD: A systematic review and meta-analysis of observational studies. *Br. J. Nutr.* 114, 1341–1359 (2015).

32. Thomas, A. M. et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* 25, 667–678 (2019).

33. Gregory, J. C. et al. Transmission of atherosclerosis susceptibility with gut microbial transplantation. *J. Biol. Chem.* 290, 5647–5660 (2015).

34. Skye, S. M. et al. Microbial transplantation with human gut commensals containing CutC is sufficient to transmit enhanced platelet reactivity and thrombosis potential. *Circ. Res.* 122, 1164–1176 (2018).

35. Landskog Jonsson, A. et al. Impact of gut microbiota and diet on the development of atherosclerosis in Apo(−/−) mice. *Arterioscler. Thromb. Vasc. Biol.* 38, 2318–2326 (2018).

36. Almeida, A. et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* 39, 105–114 (2021).

37. Pasolli, E. et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* 176, 649–662 e620 (2019).

38. Zhu, W., Lomsadze, A. & Borodovsky, M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.* 38, e132 (2010).

39. Eichler, K., Bourgis, F., Buchet, A., Kleber, H. P. & Mandrand-Berthelot, M. A. Moleculare characterization of the cai operon necessary for carnitine metabolism in *Escherichia coli*. *Mol. Microbiol.* 13, 775–786 (1994).

40. Rampelli, S. et al. Metagenome sequencing of the Hadza Hunter–Gatherer gut microbiota. *Curr. Biol.* 25, 1682–1693 (2015).

41. Schnorr, S. L. et al. Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* 5, 3654 (2014).

42. Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60 (2012).

43. Mi, Y. L. et al. Stability of gut enterotypes in Korean monozygotic twins and their association with biomarkers and diet. *Sci. Rep.* 4, 7348 (2014).

44. Nishijima, S. et al. The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res.* 23, 125–133 (2016).

45. Feng, Q. et al. Gut microbiome development along the colorectal adenoma–carcinoma sequence. *Nat. Commun.* 6, 6528 (2015).

46. Le Chatelier, E. et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546 (2013).

47. Zeller, G. et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* 10, 766 (2014).

48. Karlsson, F. H. et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103 (2013).

49. Obregón-Tito, A. J. et al. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* 6, 6505 (2015).

50. Luo, R. et al. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1, 6–1 (2012).

51. Kent, W. J. BLAT-the BLAST-like alignment tool. *Genome Res.* 12, 656–664 (2002).

52. Katoh, K. & Standley, D. M. MAFFT: Iterative refinement and additional methods. *Methods Mol. Biol.* 1079, 131–146 (2014).

53. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577 (2007).

54. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313 (2014).

55. Yu, G. C., Smith, D. K., Zhu, H. C., Guan, Y. & Lam, T. T. Y. GGTR3E: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36 (2017).

56. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction, and analysis. *Nat. Protoc.* 10, 845–858 (2015).

57. Irwin, J. J., Sterling, T., Mislinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model* 52, 1757–1768 (2012).

58. Morris, G. M. et al. AutoDock and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791 (2009).

59. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688 (2005).
