A novel cholesterol dehydrogenase from Limosilactobacillus fermentum catalyzes the direct conversion of cholesterol to coprostanol

Urmila Netter[a], Vishakha Bisht[a], Amit Gaurav[a], Rekha Sharma[a], Avik Ghosh[a], Vinod Singh Bisht[b], Kiran Ambatipudi[b], Kapil Tahlan[c], Naveen K Navani[a]*

[a]Chemical Biology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology, Roorkee, Roorkee- 247667, Uttarakhand, India.

[b]Proteomics and Lipidomics Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology, Roorkee, Roorkee- 247667, Uttarakhand, India.

[c]Department of Biology, Memorial University of Newfoundland and Labrador, Livingston, St. John's, Newfoundland and Labrador, Canada.

*Address corresponding to: - Naveen K. Navani, Department of Biosciences and Bioengineering, Indian Institute of Technology, Roorkee, Roorkee- 247667, Uttarakhand, India. Telephone no- +91 1332285677

E-mail address: naveen.navani@bt.iitr.ac.in

Abstract

The gut microbiome harbors enzymes that can transform dietary cholesterol. Understanding this interaction can help tailor the diet to modulate host lipid homeostasis. Despite being exploited commercially as a probiotic, including a role in cholesterol reduction, the molecular mechanism of cholesterol transformation by lactobacilli still needs to be discovered. Herein, we elucidate the role of a novel microbial 3β-OH-Δ^{5-6}-cholesterol-5β-reductase from *Limosilactobacillus fermentum* NKN51, which directly converts cholesterol to coprostanol. Protein engineering provides insights into the catalytic mechanism of 5βChR. Phylogenetic studies indicate an abundance of 5βChR in gut commensal lactobacilli, which shares a common ancestor with plant 5β reductases. Meta-analysis of healthy participants microbiomes highlights the significance of the 5βChR homologs, and shotgun data analysis establishes an association between higher 5βChR abundance in diabetic patients (*p*-0.0213). The discovery and elucidation of the role of lactobacillus 5βChR in cholesterol metabolism may lead to designing functional foods tailored to ameliorate dyslipidemia.

Keywords: Cholesterol, $NADP^+$, *Limosilactobacillus fermentum* NKN51, Coprostanol MALDI-TOF.

Introduction

Since Eli Metchnikoff proposed the mutualistic connection between the gut, prolongation of life status, and gut inhabitants, lactobacilli have maintained a crucial status in human health and nutrition[1]. Lactobacilli help digest complex dietary substrates and protect the host from pathogens, thus touting them as commercially desirable health supplements and aids for food technology. Lactobacilli are given ‘generally recognized as safe’ (GRAS) status by the U.S. Food and Drug

40 Administration and a 'qualified presumption of safety' status from the European Food
41 Safety Authority[2]. Gut microbiota can perform the biotransformation of dietary fats and
42 lipids, including cholesterol. The cholesterol levels are modulated by gut microbiota
43 using mechanisms like bile salt hydrolysis, short-chain fatty acid-mediated cholesterol
44 uptake, and biotransformation of cholesterol[3]. However, elevated cholesterol levels
45 become the critical risk factor for cardiovascular diseases (CVDs) and atherosclerosis,
46 leading to heart failure, obesity, diabetes, fatty liver, and Alzheimer's disease[4-6]. The
47 hydroxysteroid dehydrogenase (HSDH) enzyme family has recently been shown to
48 contribute to gut cholesterol transformation[7]. The commensal microflora harbors
49 pyridine-dependent stereo- and regio-specific HSDHs, which can oxidize, reduce, and
50 epimerize the keto/hydroxyl groups at C-3, C-7, C-12, and C-20 of substrates[8,9].
51 HSHDs primarily include aldo-keto reductase (AKR), long/medium chain
52 dehydrogenase (MDR), and the Rossman fold containing short-chain dehydrogenase
53 (SDR), which catalyze conversion between enol and keto form of hydroxyl groups on
54 alcohol, carbohydrates, and xenobiotics[10-12]. SDRs are an extended family of over
55 2000 tyrosine-dependent oxidoreductases, with tyrosine as a catalytic residue[13]. The
56 SDRs contain conserved motifs spanning the Rossmann fold for $NADP^+/NAD^+$
57 cofactor binding and a catalytic tetrad comprising serine, tyrosine, lysine, and
58 asparagine amino acid residues[14,15]. Plant 5β reductases belong to the SDR class,
59 which contain Rossmann folds with a conserved catalytic motif NFYYYED[16]. The
60 function of 5β reductases in plants has undergone extensive investigations; however,
61 their role in the bacteria domain has been overlooked[17]. Only recently, the mechanism
62 of cholesterol metabolism has been deciphered in obligatory gut anaerobes
63 *Eubacterium coprostanoligenes, Bacteroides thetaiotaomicron,* and *Oscillibacter*
64 species[7,18,19]. The genome of *E. coprostanoligenes* encodes a cholesterol

65 dehydrogenase enzyme, 'Intestinal steroid metabolism A' (IsmA) that converts
66 cholesterol into cholestenone. Further, IsmA was found to convert coprostanone into
67 coprostanol[7]. However, information on the steps in transforming cholestenone into
68 coprostanone is lacking. On the other hand, *B. thetaiotaomicron* exhibits a unique
69 sulfotransferase activity that converts cholesterol into cholesterol sulfate[18]. Further,
70 *Oscillibacter* converts cholesterol to cholestenone, cholesterol glucoside, and
71 hydroxycholesterol[19]. Overall, coprostanol, cholesterol sulfate, cholesterol glucoside,
72 and hydroxycholesterol are excreted, thereby playing a role in cholesterol
73 homeostasis[7,18,19]. Even though the direct conversion of cholesterol to coprostanol in
74 the gut milieu was proposed as early as 1954[20], the mechanism of this transformation
75 has remained an enigma[3,21]. Interestingly, the lactobacilli have been implicated in the
76 modulation of host cholesterol levels for a long time, yet the mechanistic details of this
77 observation have also remained elusive[22-27].

78 Here, we report a novel cholesterol-specific microbial 5β reductase (5βChR) from
79 *Limosilactobacillus fermentum* NKN51, isolated from fermented Himalayan yak milk.
80 Our study presents mechanistic evidence of the direct conversion of cholesterol to
81 coprostanol by 5βChR through its 3β-OH-Δ^{5-6}-cholesterol-5β-reductase activity. We
82 also validate catalytic residues involved in cholesterol and $NADP^+$ binding to 5βChR.
83 Phylogenetic analysis identified 5βChR in lactobacilli, which shared a common
84 ancestor with plant 5β reductases. Metagenomic analysis of healthy individual
85 samples identified a high prevalence of 5βChR. Further, a comparative meta-analysis
86 of a cohort study indicated the correlation of 5βChR with diabetes. Our findings
87 addressed a primarily unexplored frontier of microbial cholesterol metabolism in *L.*
88 *fermentum* NKN51 by identifying a novel enzyme class from the SDR family,

89 establishing a path for unveiling the metabolic capabilities of one of the most popular
90 probiotic bacterium.

91 **Results**

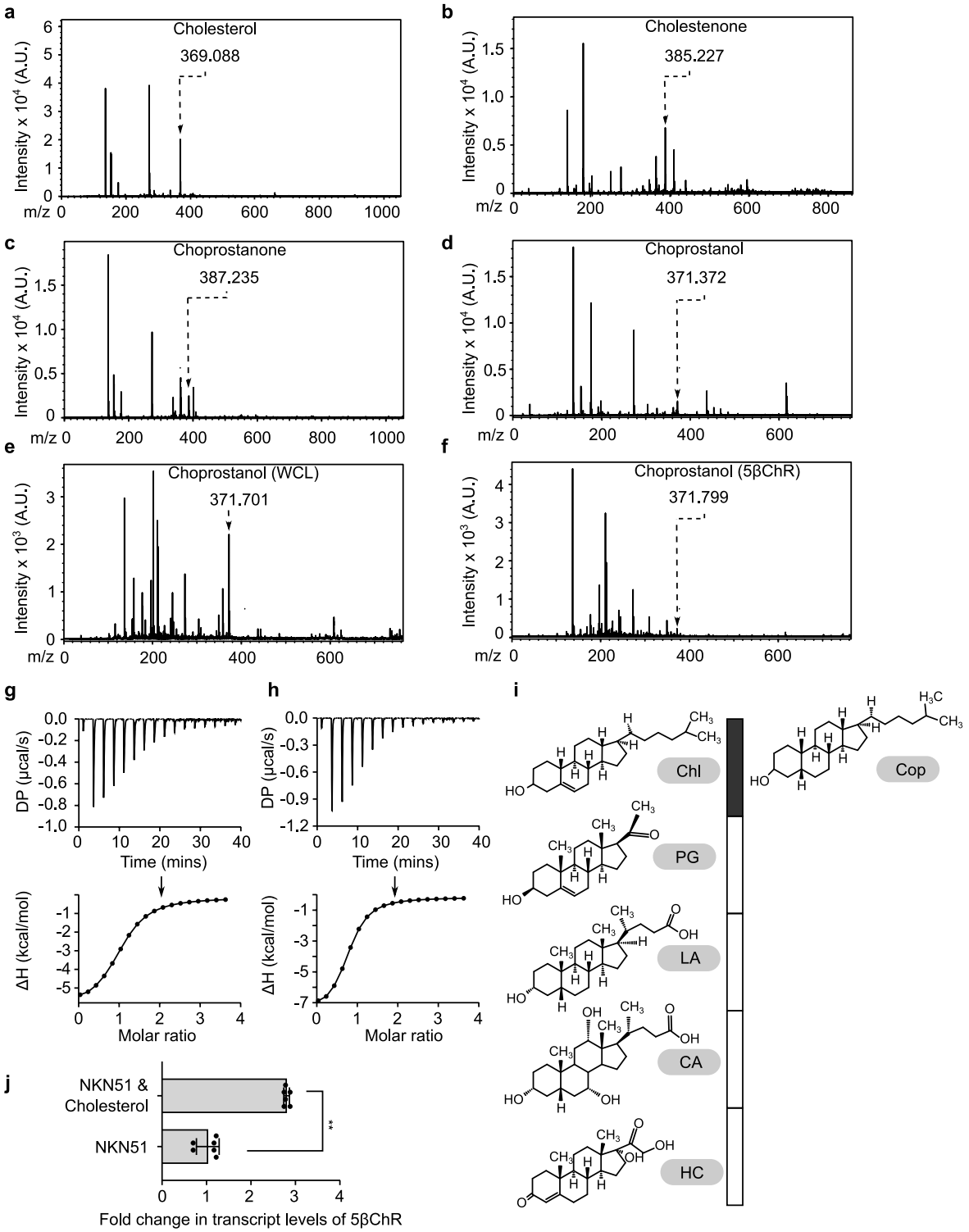
92 ***L. fermentum* NKN51 shows *in-vitro* cholesterol dehydrogenase activity**

93 Lactobacilli are touted as probiotic strains for functional foods that lower host
94 cholesterol levels[28]. To understand the mechanism of cholesterol metabolism in
95 lactobacilli, we performed cholesterol depletion assays with a battery of lactobacillus
96 spp isolated from an ethnic fermented yak milk cheese (churppi) prepared by the
97 native population in the Western Indian Himalayas. *L. fermentum* NKN51 depleted the
98 maximum (~64%) cholesterol from the media compared to other strains
99 **(Supplementary Fig. 1)**, and thus, we proceed with *L. fermentum* NKN51 for further
100 investigations. Since the cholesterol is known to be metabolized to coprostanol
101 through direct or indirect pathways[20], we checked the possible cholesterol metabolism
102 pathways in our bacterial strains by whole-cell lysate assays using thin-layer
103 chromatography (TLC) **(Supplementary Fig. 2)**. These results were further validated
104 by Matrix-assisted laser desorption/ionization- Time of flight (MALDI-TOF) analysis,
105 demonstrating that *L. fermentum* NKN51 strain converted cholesterol to coprostanol
106 **(Fig. 1a-1e)**. We could not detect cholestenone or coprostanone in assays. These data
107 pointed towards the possibility that *L. fermentum* NKN51 might directly convert
108 cholesterol to coprostanol. These analytical assays also highlighted that the formation
109 of coprostanol by NKN51 was $NADP^+$-dependent but oxygen-independent
110 **(Supplementary Fig. 3-4).**

111 **Biochemical characterization of putative microbial cholesterol dehydrogenase**

112 To provide mechanistic insights into the putative cholesterol dehydrogenase enzyme
113 for cholesterol to coprostanol transformation in *L. fermentum* NKN51, the recombinant
114 protein was overexpressed and purified for analysis. The amino acid sequence of a
115 known cholesterol SDR from *E. coprostanoligenes* HL ATCC51222[22]
116 (WP_078769004.1) was used as a query against *L. fermentum* NKN51 draft genome,
117 which selected a sequence with ~30% identity in NKN51. The putative cholesterol
118 dehydrogenase was cloned, and the overexpressed protein of 29.5 kDa was purified
119 by immobilized affinity chromatography (IMAC) **(Supplementary Fig. 5)**. The purified
120 recombinant protein converted cholesterol to coprostanol using $NADP^+$ as a cofactor
121 instead of NAD^+, validating our previous observation. **(Fig. 1f)**. Protein-cofactor-
122 substrate isothermal titration calorimetry (ITC) studies corroborate that cholesterol
123 dehydrogenase had an affinity for $NADP^+$ with K_d 6.88 μM, and the protein-$NADP^+$
124 complex displayed interaction for cholesterol with a K_d 4.66 μM **(Fig. 1g-1h)**. However,
125 protein-$NADP^+$ did not show affinity towards cholestenone or coprostanone,
126 substantiating that both of them are not the cognate substrates of the enzyme
127 **(Supplementary Fig. 6)**. Further, the gel filtration chromatography depicted that the
128 recombinant protein predominantly remains as a dimer in solution **(Supplementary**
129 **Fig 7)**. The substrate specificity assay confirms that the enzyme can only perform
130 saturation of the Δ5-6-oxo bond of the substrate with the OH group at the 3β position.
131 Hence, the enzyme can specifically metabolize cholesterol to coprostanol **(Fig.1i)**
132 **(Supplementary Fig. 8)**. These observations allude to the fact that the cholesterol
133 dehydrogenase belongs to a novel class of SDR, i.e., 3β-OH- Δ^{5-6}- cholesterol-5β-
134 reductase (5βChR). Kinetic studies on protein-substrate confirmed that 5βChR obeys
135 the Michaelis- Menton equation with K_m 19.09 μM and V_{max} 15.35×10^{-6} μmol sec^{-1} $nmol^{-}$
136 ^{1} **(Supplementary Fig. 9)**. Furthermore, 5βChR transcript expression levels were ~1.5

137 fold higher when NKN51 was grown in the presence of cholesterol for 24h, verifying
138 5βChR genetic locus was induced upon cholesterol-supplementation (*p*-0.0022) in the
139 native host **(Fig. 1j)**.


a
Cholesterol
369.088
Intensity x 10^4 (A.U.)
m/z
b
Cholestenone
385.227
c
Choprostanone
387.235
d
Choprostanol
371.372
e
Choprostanol (WCL)
371.701
Intensity x 10^3 (A.U.)
f
Choprostanol (5βChR)
371.799
g
DP (µcal/s)
Time (mins)
ΔH (kcal/mol)
Molar ratio
h
i
Chl
PG
LA
CA
HC
Cop
j
NKN51 &
Cholesterol
NKN51
**
Fold change in transcript levels of 5βChR

140

141 **Figure 1: *L. fermentum* NKN51 harbors a cholesterol dehydrogenase that can**
142 **convert cholesterol to coprostanol**

143 **(a-d).** Matrix-assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF)
144 chromatograms show the peaks of sterol standards, cholesterol, cholestenone,
145 coprostanone, and coprostanol, which are represented by their respective mass-to-
146 charge (m/z) ratios. **(e).** *L. fermentum* NKN51 enzymatic activity is assessed by its
147 whole-cell lysate (WCL) assay in the presence of cholesterol (100 μM) and $NADP^+$
148 (100 μM) for 12 h under anaerobic conditions. The reaction mixture is evaluated by
149 MALDI-TOF, illustrating the presence of coprostanol with an m/z of 371.701. **(f).** The
150 MALDI-TOF spectra graph shows the result of an *in-vitro* biochemical assay of purified
151 cholesterol-5β-reductase (5βChR) incubated with $NADP^+$ (100 μM) and cholesterol
152 (100 μM) in the 20 mM Tris-cl (pH-7.8) for 12 h. It depicts the peak of coprostanol (m/z-
153 371) as a product of cholesterol metabolism. **(g).** Isothermal Titration Calorimetry (ITC)
154 data displaying binding affinity of the 5βChR (20 μM) for $NADP^+$ (1 mM). **(h).** The
155 titration curve provides the interaction between the 5βChR + $NADP^+$ complex and
156 cholesterol (500 μM), where DP stands for differential power, and ΔH represents the
157 enthalpies (normalized with the control). **(i).** The bar graph indicates the conversion of
158 cholesterol to coprostanol (grey block) by 5βChR under the same conditions. The
159 other structurally similar compounds are not metabolized into their respective
160 metabolites. Abbreviations: PG stands for pregnenolone, LA-lithocholic acid, CA-
161 cholic acid, and HC-hydrocortisone. Spectra analysis of this substrate specificity assay
162 is given in Supplementary Figure 8. **(j).** The transcript levels of 5βChR in a cholesterol-
163 rich medium compared to a medium without cholesterol at 24 h under anaerobic
164 conditions show ~1.5-fold change. In the graph, each point represents a biological
165 replicate, with the centre bar representing the mean and the error bars representing

166 the standard deviation. ** denotes the *p* value- 0.0022 calculated by the Mann-Whitney
167 test.

168 **Structural analysis of cholesterol dehydrogenase-5βChR**

169 The 5βChR structure was predicted using AlphaFold2 with high confidence (>90%),
170 high pre-residue confidence score (>90%), and a low domain position confidence.
171 These scores highlight the high local confidence around each amino acid and
172 position/orientation confidence between individual domains. The 5βChR tertiary
173 structure depicts conserved motifs of the SDR superfamily, including the N-terminus
174 dinucleotide binding Rossmann fold consisting of seven central β-strands flanked by
175 six α-helices and a catalytic tetrad (NSYK) **(Fig 2a-2b)**. Molecular docking indicates
176 that cholesterol was bound to the C-terminus of the protein where cholesterol steroid
177 ring-1 interacts with 5βChR through hydrophobic forces by methionine (Met189) and
178 phenylalanine (Phe190). An electrostatic bond was observed between 5βChR
179 threonine (Thr187) and the hydroxyl group of cholesterol. Protein-cholesterol
180 interaction was also accompanied by hydrophobic interactions with proline (Pro182 &
181 Pro184), phenylalanine (Phe190, Phe198 & Phe239), and glycine (Gly183). Hence,
182 the positioning of the sterol in 5βChR was complemented by Van der Waals repulsion
183 energies. $NADP^+$ was bound to the N-terminus of the 5βChR via the conserved
184 glycine-rich (Gly 8, 9, 12, and 14) motif (GXXGXXG) of the SDR enzyme family.
185 Tyrosine (Tyr152), lysine (Lys156), serine (Ser11), and threonine (Thr10 and Thr187)
186 residues of the protein contributed to a total of eight polar interactions with $NADP^+$.
187 Further, Tyr152 and Lys156 amino acid residues of 5βChR were likely to interact with
188 the nicotinamide oxygen (O=C-NH). The hydroxyl chain of Ser11 and Thr10 had an
189 affinity with the pyrophosphate group's oxygen and phosphate group of ribose,

respectively. Van der Waals forces were involved between $NADP^+$ and Ala32, Arg33, Gly89, Ile89, Ser139, Asp146, Met149, Ser11, Ile90, and Ala9 of 5βChR **(Fig. 2c)**. Structural homologs of 5βChR were retrieved from DALI server, which was identified as SDR reductase (PDB ID: 5X8H), Sorbitol dehydrogenase (PDB ID: 4NPC), and Putative oxidoreductase RV2002 (PDB ID: 1NFR). The superimposed images of these enzymes onto 5βChR displayed root square mean value (r.m.s) deviation of 1.3, 1.3, and 1.2 Å, respectively. These proteins share a topological arrangement with 5βChR. However, 5βChR had an extended C-terminus, which is not present in either of the homologs. These observations confirm the previous sequence alignment-based prediction that the cholesterol dehydrogenase, 5βChR, belongs to an enzyme of the SDR category **(Supplementary Fig. 10)**.

Identification of catalytic residues for $NADP^+$ and cholesterol binding in 5βChR

Using site-directed mutagenesis, we elucidated the role of catalytic tetrad (NSYK) and other conserved residues of 5βChR. Since the C-terminus of SDRs was reported for substrate binding[10], we mutated the C-terminus of the protein to investigate its role in cholesterol binding **(Supplementary Fig. 11)**. The TLC and MALDI experiments with the purified mutants revealed that the cholesterol transformation ability of serine (Ser139A), tyrosine (Tyr152A), threonine (Thr187A), methionine (Met189A), and phenylalanine (Phe190A) 5βChR mutants was abrogated **(Supplementary Fig. 12-13**). Further, ITC data revealed that S139A and Y152A mutants lost the $NADP^+$ binding ability **(Supplementary Fig. 14)** yet retained the binding ability for cholesterol with a K_d, 24.8 μM and 4.44 μM, respectively **(Fig 2d-2e).** On the other hand, T187A, M189A, and F190A retained the $NADP^+$ binding ability (K_d 8.77 μM, 7.49 μM, and 5.2 μM respectively) **(Fig 2f-2h) (Supplementary Fig. 14)** and lost the ability to bind to

214 cholesterol, thus revealing that indeed the C-terminus of the 5βChR is important for
215 the substrate interaction. Other mutations, like N136A, K156A, H131A, Y170A, E172A,
216 and T187A, did not affect the ability for conversion of cholesterol to coprostanol, thus
217 indicating their non-significant role in the enzyme activity **(Supplementary Fig 13)**.
218 We compared the $NADP^+$-binding pocket of 5βChR with the 5β-progesterone
219 reductase (5β-POR) complexed with $NADP^+$ (PDB: 2V6G)[16]. The Rossmann fold was
220 responsible for $NADP^+$ binding in plant and microbial 5β reductase **(Fig. 2i-2j)**.
221 However, the lysine residue of a conserved catalytic tetrad of SDRs does not interact
222 with the O-2' and O-3' of nicotinamide ribose in both cases[14,29]. In the 5β-POR, two
223 arginine residues (Arg63 & Arg64), and in 5βChR a serine (Ser139) and tyrosine
224 (Tyr152) were responsible for cofactor binding **(Fig. 2i-2j)**. This reinforces that 5βChR
225 retains the structural features of both SDRs and 5β-POR. The circular dichroism
226 spectra analysis of wild-type and mutant 5βChR enzymes did not reveal any change
227 in the protein structure, indicating that the observations about cofactor and substrate
228 binding were not due to perturbations in the secondary structure of the 5βChR
229 enzymes. The observed changes were likely due to the ability of substrate and cofactor
230 binding ascribed to the amino acids mutated. **(Supplementary Fig. 15)**.

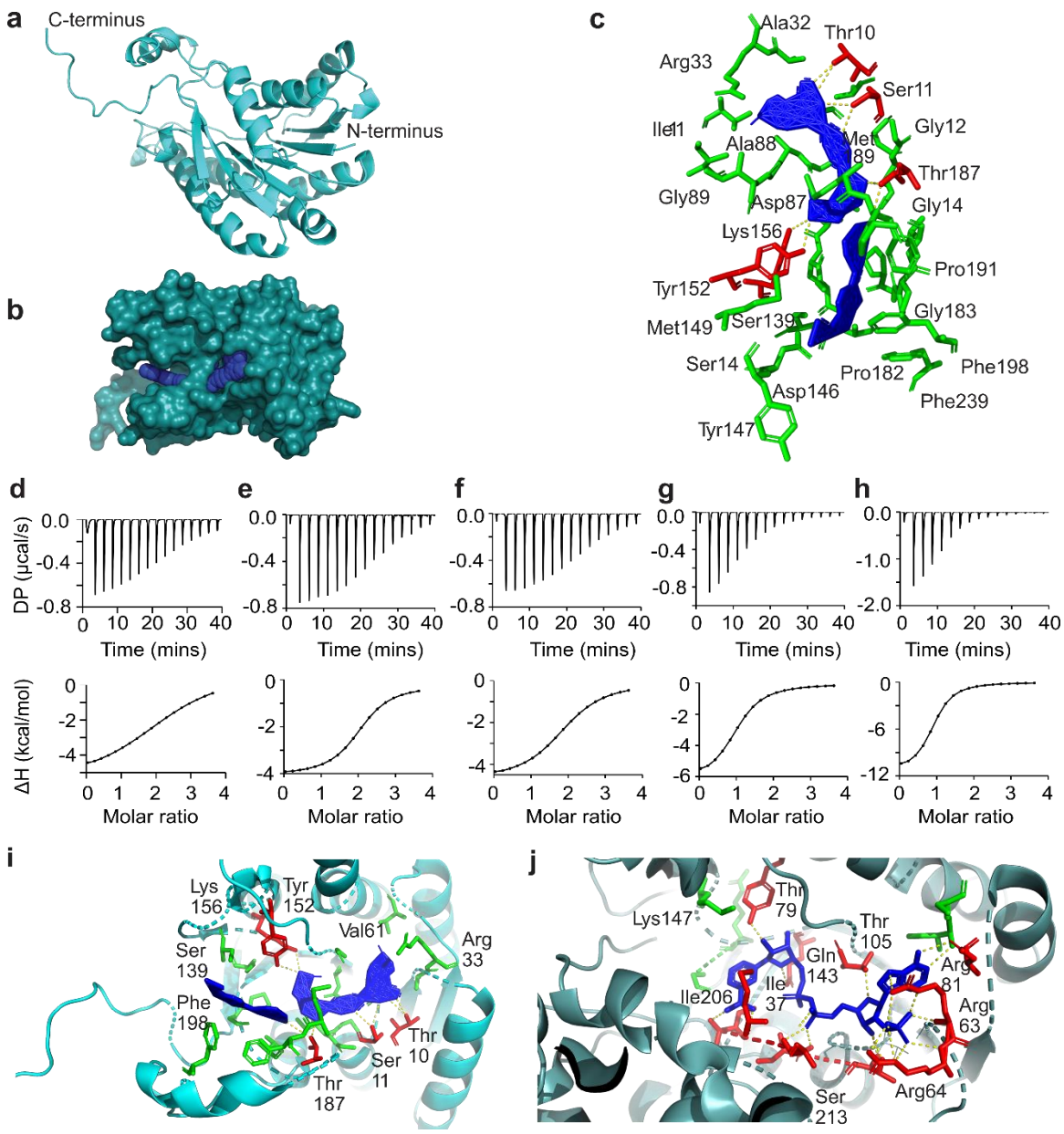


Figure 2: Structure and mutation analysis of 5βChR for ligand binding and comparison study with the nicotinamide binding pocket of plant 5β reductase.

(a). Apo structure of 5βChR deciphered by Alphafold2 and visualized by PyMOL. The structure exhibits the SDR-characteristic N-terminus Rossmann fold consisting of seven central β-strands (β3- β2- β1- β4- β5- β6- β7) flanked by six α-helices. **(b).** The surface representation of 5βChR shows the active binding pocket of cholesterol and $NADP^+$ (blue). **(c).** The molecular docking results of 5βChR-$NADP^+$-cholesterol depict

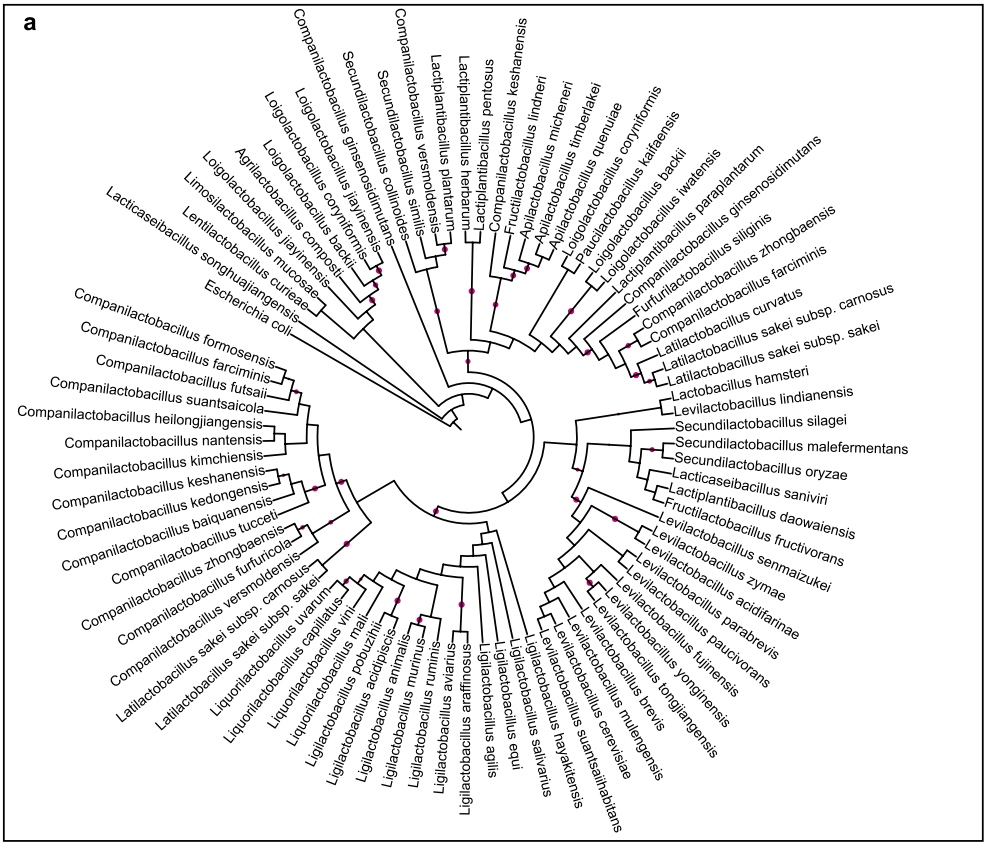
239 the interacting amino acids. The Cofactor ($NADP^+$) binding pocket highlights the ionic
240 interaction (red) (indicated by dashed lines). The amino acids involved in hydrogen
241 bonding are Thr10, Ser11, and Tyr152, while Lys156, Ala32, Arg33, Gly89, Ile89,
242 Ser139, Asp146, Met149, Ser11, Ile90, and Ala9 showcase involvement in van der
243 Waals bonds (green). Protein-substrate (5βChR-cholesterol) complex shows
244 hydrophilic interaction with Thr187, whereas the hydrophobic forces are present
245 between Pro182, Pro184, Met189, Phe190, Phe198, and Phe238. **(d-h).** The
246 normalized enthalpy curves represent the interaction profile of 5βChR mutants with
247 their ligands. S139A (d) and Y152A (e) mutants have an affinity for cholesterol and lost
248 binding for $NADP^+$ (Supplementary Figure 13). T187A, M189A, and F190A mutants
249 do not bind to cholesterol (Supplementary Figure 13). In contrast, they retained affinity
250 for $NADP^+$ (f-h). **(i-j).** The comparison study of the active site of 5βChR and plant 5β
251 reductase (5β-POR) with their respective substrates. The stereo view of 5βChR
252 complexed with $NADP^+$ is mediated by tyrosine (Tyr152) and serine (Ser139) of the
253 conserved tetrad. In 5β-POR - two arginine residues (Arg63 & Arg64) stabilise $NADP^+$
254 interaction. 5β-POR has two Rossmann folds (glycine residues) (GXXGXXG),
255 whereas 5βChR harbors a single N-terminus fold, contributing to $NADP^+$ binding.

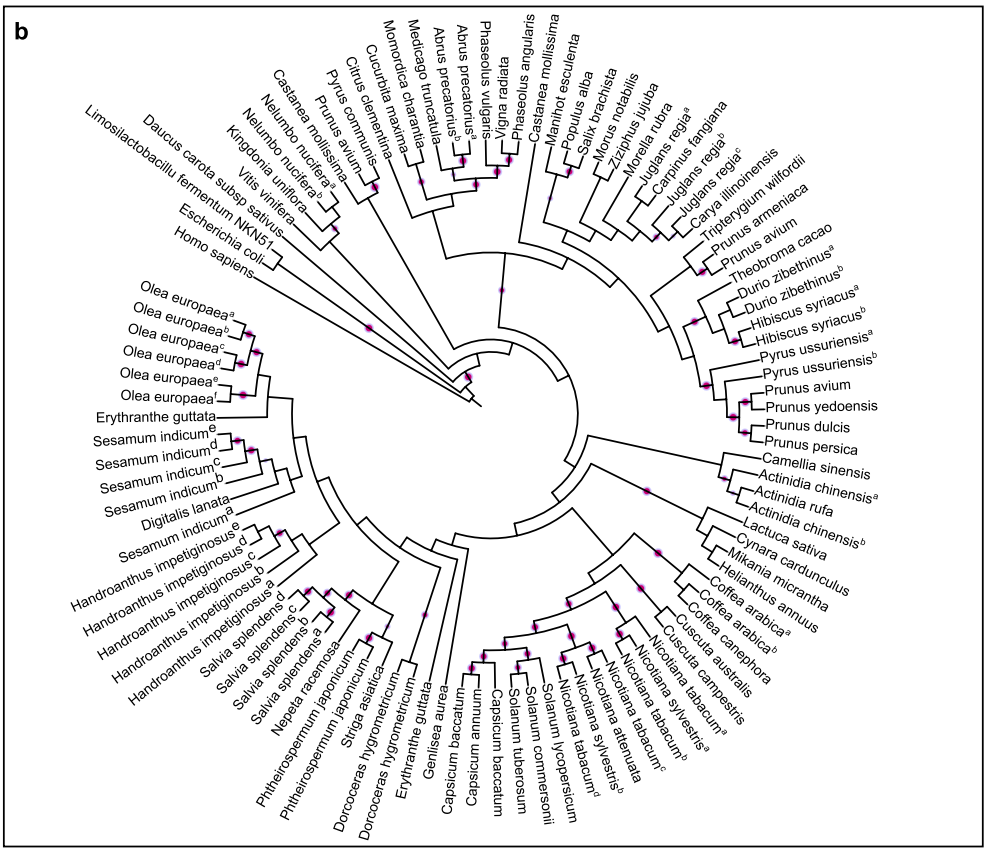
256 **Distribution and abundance analysis of 5βChR**

257 Since 5βChR turned out to be a novel class of enzyme discovered in bacteria, we
258 focused on assessing the diversity and abundance within Lactobacillaceae and overall
259 in eubacteria. A local database of 256 lactobacilli was screened to assess the
260 prevalence of 5βChR in their genomes **(Supplementary File. 1)**. A maximum
261 likelihood method confirmed the presence of 5βChR sequence in 84 lactobacilli
262 genomes, which were further divided into six major clades. Most of these

263 representatives (*Lactiplantibacillus plantarum, Limosilactobacillus mucosae,*
264 *Levilactobacillus parabrevis*, *Latilactobacillus sakei*, *Latilactobacillus curvatus*,
265 *Ligilactobacillus salivarius,* and *Levilactobacillus brevis*) were associated with a
266 healthy commensal gut community[30,31] **(Fig. 3a)**. We traced the evolutionary
267 relationship of microbial 5βChR with plant 5β reductases using the human homolog
268 as an outgroup **(Supplementary File. 1)**. The phylogeny analysis revealed that
269 human, plant, and microbial 5β reductases originated from a common ancient
270 ancestor. **(Fig. 3b)**. With an aim to assess the distribution of 5βChR in different
271 microbial taxa, we retrieved 5000 5βChR homologs and interpreted the data using
272 Cytoscape. 5βChR homologs were grouped into 33, 138, and 269 clusters with
273 threshold values for sequence identity (e-value) of 10^{-60}, 10^{-70}, and 10^{-80}, respectively
274 **(Supplementary Fig. 16) (Supplementary File. 1)**. Finally, a threshold value of 10^{-80}
275 was chosen to map the 5βChR homologs to ensure similar functional and structural
276 aspects, which led to the grouping of proteins into 269 distinct clusters **(Fig. 4a).** The
277 phylum-level taxonomic classification of respective clusters were taken into
278 consideration to assess the evolutionary origin these sequences. Cluster 1 had 934
279 homologs belonging to Pseudomonadota (Proteobacteria) (87.04%), Actinomycetota
280 (Actinobacteria) (11.56%), and Planctomycetota (0.74%). Cluster 2 had 26 phyla,
281 predominately from Chloroflexota (21.9%), Cyanobacteria (12.44%), Acidobacteriota
282 (10.18%), Bacillota (Firmicutes) (8.59%), Actinomycetota (6.7%), Candidatus
283 omnitrophota (5.4%) and an uncharacterized class (4.29%). Cluster 3 was the most
284 diversified SSN cluster, accounting for 36 phyla, out of which Bacillota (54.2%) was
285 the most prevalent phylum, comprising half of the homologs. Interestingly, Bacillota
286 was also a dominant phylum in 5, 6, 8, 9, 10, 11, 16, 52, 56, 83, 128, 139 and 249
287 clusters **(Fig. 4b) (Supplementary Fig. 17).** Pseudomondata was present in clusters

288 4, 7, 12, and 15, while cluster 13 was abundant with Bacteroidata. These observations
289 highlight the presence of the 5βChR in *Clostridia, Eubacteria, Oscillopsia,*
290 *Peptostreptococcus, Fusobacteria Bifidobacteria, and Lactobacillus* **(Fig. 4b-4c)**
291 **(Supplementary Fig. 17)**. The SSN analysis indicates that inhabiting different niches
292 belonging to different phyla harbor 5βChR homologs, possibly providing some benefit
293 to their respective hosts **(Fig. 4c)**.


a
Companilactobacillus ginsenosidimutans
Secundilactobacillus collinoides
Companilactobacillus versmoldensis
Secundilactobacillus similis
Lactiplantibacillus plantarum
Lactiplantibacillus herbarum
Lactiplantibacillus pentosus
Companilactobacillus keshanensis
Fructilactobacillus lindneri
Apilactobacillus micheneri
Apilactobacillus timberlakei
Apilactobacillus quenuiae
Loigolactobacillus coryniformis
Paucilactobacillus kaifaensis
Loigolactobacillus backii
Loigolactobacillus coryniformis
Loigolactobacillus backii
Loigolactobacillus compostii
Loigolactobacillus jiayinensis
Agrilactobacillus iwatensis
Loigolactobacillus mucosae
Limosilactobacillus curieae
Lentilactobacillus parafarraginis
Lacticaseibacillus songhuajiangensis
Escherichia coli
Companilactobacillus formosensis
Companilactobacillus farciminis
Companilactobacillus futsaii
Companilactobacillus suantsaicola
Companilactobacillus heilongjiangensis
Companilactobacillus nantensis
Companilactobacillus kimchiensis
Companilactobacillus keshanensis
Companilactobacillus kedongensis
Companilactobacillus baiquanensis
Companilactobacillus tucceti
Companilactobacillus zhongbaensis
Companilactobacillus furfuricola
Companilactobacillus versmoldensis
Companilactobacillus farciminis
Companilactobacillus curvatus
Companilactobacillus zhongbaensis
Companilactobacillus siliginis
Companilactobacillus ginsenosidimutans
Lactiplantibacillus paraplantarum
Furfurilactobacillus curvatus
Companilactobacillus farciminis
Latilactobacillus sakei subsp. carnosus
Latilactobacillus sakei subsp. sakei
Latilactobacillus hamsteri
Lactobacillus lindianensis
Levilactobacillus silagei
Secundilactobacillus malefermentans
Secundilactobacillus oryzae
Lacticaseibacillus saniviri
Lactiplantibacillus daowaiensis
Fructilactobacillus fructivorans
Levilactobacillus senmaizukei
Levilactobacillus zymae
Levilactobacillus acidifarinae
Levilactobacillus parabrevis
Levilactobacillus paucivorans
Levilactobacillus fujinensis
Levilactobacillus yongjiangensis
Levilactobacillus mulengensis
Levilactobacillus brevis
Levilactobacillus cerevisiae
Levilactobacillus suantsaihabitans
Levilactobacillus hayakitensis
Ligilactobacillus salivarius
Ligilactobacillus equi
Ligilactobacillus agilis
Ligilactobacillus araffinosus
Ligilactobacillus aviarius
Ligilactobacillus ruminis
Ligilactobacillus murinus
Ligilactobacillus animalis
Ligilactobacillus acidipiscis
Ligilactobacillus pobuzihii
Ligilactobacillus mali
Liquorilactobacillus vini
Liquorilactobacillus capillatus
Liquorilactobacillus uvarum
Latilactobacillus sakei subsp. sakei
Latilactobacillus sakei subsp. carnosus
Companilactobacillus versmoldensis
Companilactobacillus furfuricola
Companilactobacillus zhongbaensis
Companilactobacillus tucceti
Companilactobacillus baiquanensis
Companilactobacillus kedongensis
Companilactobacillus keshanensis
Companilactobacillus kimchiensis
Companilactobacillus nantensis
Companilactobacillus heilongjiangensis
Companilactobacillus suantsaicola
Companilactobacillus futsaii
Companilactobacillus farciminis
Companilactobacillus formosensis
Escherichia coli
Lacticaseibacillus songhuajiangensis
Lentilactobacillus curieae
Limosilactobacillus mucosae
Agrilactobacillus jiayinensis
Loigolactobacillus compostii
Loigolactobacillus backii
Loigolactobacillus coryniformis
Loigolactobacillus ginsenosidimutans
Secundilactobacillus collinoides
Companilactobacillus versmoldensis
Lactiplantibacillus plantarum
Lactiplantibacillus herbarum
Lactiplantibacillus pentosus
Companilactobacillus keshanensis
Fructilactobacillus lindneri
Apilactobacillus micheneri
Apilactobacillus timberlakei
Apilactobacillus quenuiae
Loigolactobacillus coryniformis
Paucilactobacillus kaifaensis
Loigolactobacillus backii
Lactiplantibacillus iwatensis
Lactiplantibacillus paraplantarum
Furfurilactobacillus curvatus
Companilactobacillus farciminis
Companilactobacillus zhongbaensis
Companilactobacillus siliginis
Companilactobacillus ginsenosidimutans
Latilactobacillus sakei subsp. carnosus
Latilactobacillus sakei subsp. sakei
Latilactobacillus hamsteri
Lactobacillus lindianensis
Levilactobacillus silagei
Secundilactobacillus malefermentans
Secundilactobacillus oryzae
Lacticaseibacillus saniviri
Lactiplantibacillus daowaiensis
Fructilactobacillus fructivorans
Levilactobacillus senmaizukei
Levilactobacillus zymae
Levilactobacillus acidifarinae
Levilactobacillus parabrevis
Levilactobacillus paucivorans
Levilactobacillus fujinensis
Levilactobacillus yongjiangensis
Levilactobacillus mulengensis
Levilactobacillus brevis
Levilactobacillus cerevisiae
Levilactobacillus suantsaihabitans
Levilactobacillus hayakitensis
Ligilactobacillus salivarius
Ligilactobacillus equi
Ligilactobacillus agilis
Ligilactobacillus araffinosus
Ligilactobacillus aviarius
Ligilactobacillus ruminis
Ligilactobacillus murinus
Ligilactobacillus animalis
Ligilactobacillus acidipiscis
Ligilactobacillus pobuzihii
Ligilactobacillus mali
Liquorilactobacillus vini
Liquorilactobacillus capillatus
Liquorilactobacillus uvarum
Latilactobacillus sakei subsp. sakei
Latilactobacillus sakei subsp. carnosus
Companilactobacillus versmoldensis
Companilactobacillus furfuricola
Companilactobacillus zhongbaensis
Companilactobacillus tucceti
Companilactobacillus baiquanensis
Companilactobacillus kedongensis
Companilactobacillus keshanensis
Companilactobacillus kimchiensis
Companilactobacillus nantensis
Companilactobacillus heilongjiangensis
Companilactobacillus suantsaicola
Companilactobacillus futsaii
Companilactobacillus farciminis
Companilactobacillus formosensis
Escherichia coli
Lacticaseibacillus songhuajiangensis
Lentilactobacillus curieae
Limosilactobacillus mucosae
Agrilactobacillus jiayinensis
Loigolactobacillus compostii
Loigolactobacillus backii
Loigolactobacillus coryniformis
Loigolactobacillus ginsenosidimutans
Secundilactobacillus collinoides
Companilactobacillus versmoldensis
Lactiplantibacillus plantarum
Lactiplantibacillus herbarum
Lactiplantibacillus pentosus
Companilactobacillus keshanensis
Fructilactobacillus lindneri
Apilactobacillus micheneri
Apilactobacillus timberlakei
Apilactobacillus quenuiae
Loigolactobacillus coryniformis
Paucilactobacillus kaifaensis
Loigolactobacillus backii
Lactiplantibacillus iwatensis
Lactiplantibacillus paraplantarum
Furfurilactobacillus curvatus
Companilactobacillus farciminis
Companilactobacillus zhongbaensis
Companilactobacillus siliginis
Companilactobacillus ginsenosidimutans
Latilactobacillus sakei subsp. carnosus
Latilactobacillus sakei subsp. sakei
Latilactobacillus hamsteri
Lactobacillus lindianensis
Levilactobacillus silagei
Secundilactobacillus malefermentans
Secundilactobacillus oryzae
Lacticaseibacillus saniviri
Lactiplantibacillus daowaiensis
Fructilactobacillus fructivorans
Levilactobacillus senmaizukei
Levilactobacillus zymae
Levilactobacillus acidifarinae
Levilactobacillus parabrevis
Levilactobacillus paucivorans
Levilactobacillus fujinensis
Levilactobacillus yongjiangensis
Levilactobacillus mulengensis
Levilactobacillus brevis
Levilactobacillus cerevisiae
Levilactobacillus suantsaihabitans
Levilactobacillus hayakitensis
Ligilactobacillus salivarius
Ligilactobacillus equi
Ligilactobacillus agilis
Ligilactobacillus araffinosus
Ligilactobacillus aviarius
Ligilactobacillus ruminis
Ligilactobacillus murinus
Ligilactobacillus animalis
Ligilactobacillus acidipiscis
Ligilactobacillus pobuzihii
Ligilactobacillus mali
Liquorilactobacillus vini
Liquorilactobacillus capillatus
Liquorilactobacillus uvarum
Latilactobacillus sakei subsp. sakei
Latilactobacillus sakei subsp. carnosus
Companilactobacillus versmoldensis
Companilactobacillus furfuricola
Companilactobacillus zhongbaensis
Companilactobacillus tucceti
Companilactobacillus baiquanensis
Companilactobacillus kedongensis
Companilactobacillus keshanensis
Companilactobacillus kimchiensis
Companilactobacillus nantensis
Companilactobacillus heilongjiangensis
Companilactobacillus suantsaicola
Companilactobacillus futsaii
Companilactobacillus farciminis
Companilactobacillus formosensis
Escherichia coli


b
Phaseolus vulgaris[a]
Vigna radiata
Phaseolus angularis
Abrus precatorius[a]
Abrus precatorius[b]
Medicago truncatula
Momordica charantia
Cucurbita maxima
Citrus clementina
Pyrus communis
Prunus avium
Castanea mollissima
Nelumbo nucifera[a]
Nelumbo nucifera[b]
Kingdonia uniflora
Vitis vinifera
Daucus carota subsp sativus
Limosilactobacillu fermentum NKN51
Escherichia coli
Homo sapiens
Castanea mollissima
Manihot esculenta
Populus alba
Salix brachista
Morus notabilis
Ziziphus jujuba
Morella rubra
Juglans regia[a]
Carpinus fangiana
Juglans regia[b]
Juglans regia[c]
Carya illinoinensis
Pterocarya stenoptera
Prunus armeniaca
Prunus avium
Prunus avium
Theobroma cacao
Durio zibethinus[a]
Durio zibethinus[b]
Hibiscus syriacus[a]
Hibiscus syriacus[b]
Pyrus ussuriensis[a]
Pyrus ussuriensis[b]
Prunus avium
Prunus yedoensis
Prunus dulcis
Prunus persica
Camellia sinensis
Actinidia chinensis[a]
Actinidia rufa
Actinidia chinensis[b]
Lactuca sativa
Cynara cardunculus
Mikania micrantha
Helianthus annuus
Coffea arabica[a]
Coffea arabica[b]
Coffea canephora
Cuscuta australis
Cuscuta campestris
Nicotiana tabacum[a]
Nicotiana sylvestris[a]
Nicotiana tabacum[b]
Nicotiana tabacum[c]
Nicotiana sylvestris[b]
Nicotiana tabacum[d]
Nicotiana attenuata
Solanum lycopersicum
Solanum commersonii
Solanum tuberosum
Capsicum baccatum
Capsicum annuum
Capsicum baccatum
Genlisea aurea
Erythranthe guttata
Dorcoceras hygrometricum
Dorcoceras hygrometricum
Striga asiatica
Phtheirospermum japonicum
Phtheirospermum japonicum
Nepeta racemosa
Salvia splendens[a]
Salvia splendens[b]
Salvia splendens[c]
Salvia splendens[d]
Handroanthus impetiginosus[a]
Handroanthus impetiginosus[b]
Handroanthus impetiginosus[c]
Handroanthus impetiginosus[d]
Handroanthus impetiginosus[e]
Digitalis lanata
Sesamum indicum[a]
Sesamum indicum[b]
Sesamum indicum[c]
Sesamum indicum[d]
Sesamum indicum[e]
Erythranthe guttata
Olea europaea[f]
Olea europaea[e]
Olea europaea[d]
Olea europaea[c]
Olea europaea[b]
Olea europaea[a]

294

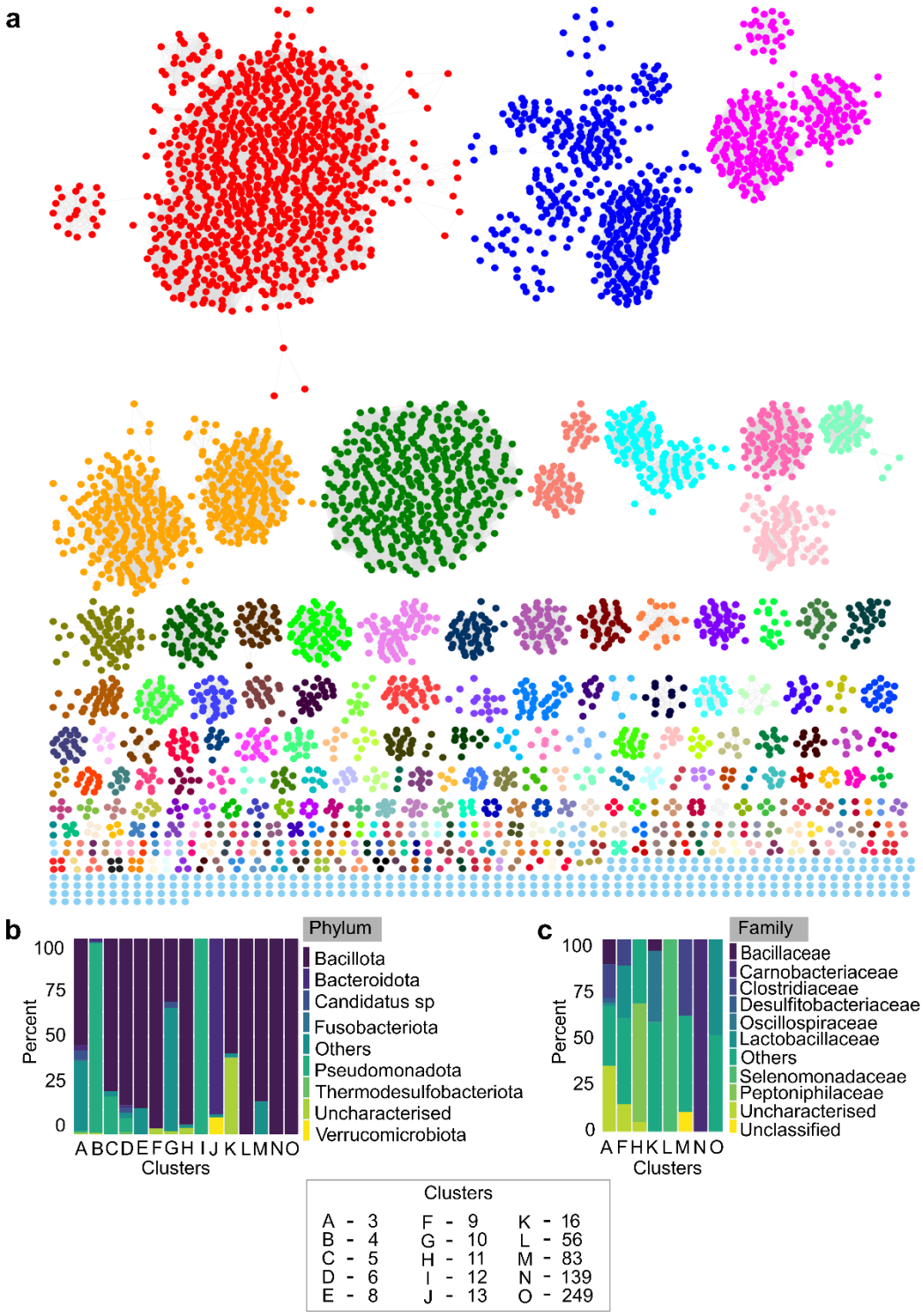
295 **Figure 3**: **Mapping the prevalence of 5βChR within (a) lactobacilli and (b) plants**
296 **by phylogenetic profiling.**

297 **(a).** The cladogram of 5βChR protein across 256 lactobacilli depicts its abundance in
298 84 strains. This study used the PROTCATWAG model in RAxML and rooted using the
299 branch leading to *Lacticaseibacillus songhuajiangensis* and *Escherichia coli* as the
300 outgroups. **(b).** The evolution analysis of 5βChR homologs in plants (~100) is acquired
301 from UniProt in RAxML using the PROTCATWAG model with human 5β reductase as
302 an outgroup. 5βChR and plant 5β reductase evolved from a common ancestor. A pink
303 circle indicates a bootstrap value of more than 70.

304 **Meta-analysis of 5βChR abundance in microbiomes of healthy microbiomes**

305 With an aim to assess the importance of aforementioned 5βChR clusters, we
306 assessed the data through ShortBRED consisting of protein class in 380 high-quality
307 healthy human metagenome sequences (from the human microbiome project),
308 covering samples from stool, buccal mucosa, supragingival plaque, tongue dorsum,
309 anterior nares, and vaginal fornix **(Supplementary File. 1)**. Through EFI-
310 Computationally-Guided Functional Profiling (EFI-CGFP) tool we identified 20 clusters
311 of 5βChR homologs in Human Microbiome Project data **(Fig. 5a)**. Clusters 3, 9, and
312 11 were abundant in the oral microbiome (supragingival plaque, buccal cavity, and
313 tongue dorsum), whereas 2, 3, 6, 9, 11, 52, 83, 104, 119, 122, 128, 140, 156, 172,
314 185, and 24 were dominant clusters in the fecal samples, indicating the presence of
315 5βChR homologs in anaerobes. Bacillota-dominated clusters, 16 and 139, showed
316 higher distribution in the posterior fornix (vaginal microbiome) and anterior nares **(Fig.**
317 **5b) (Supplementary Fig. 18)**. The varying abundance of 5βChR unveiled the
318 enzyme's ecological diversity in humans and reinforced its potential contribution to

maintaining human physiological homeostasis. No clusters were dominated by phyla-Acidobacteria, Nitrospirota, and Chloroflexota, likely indicating their functional insignificance to humans.

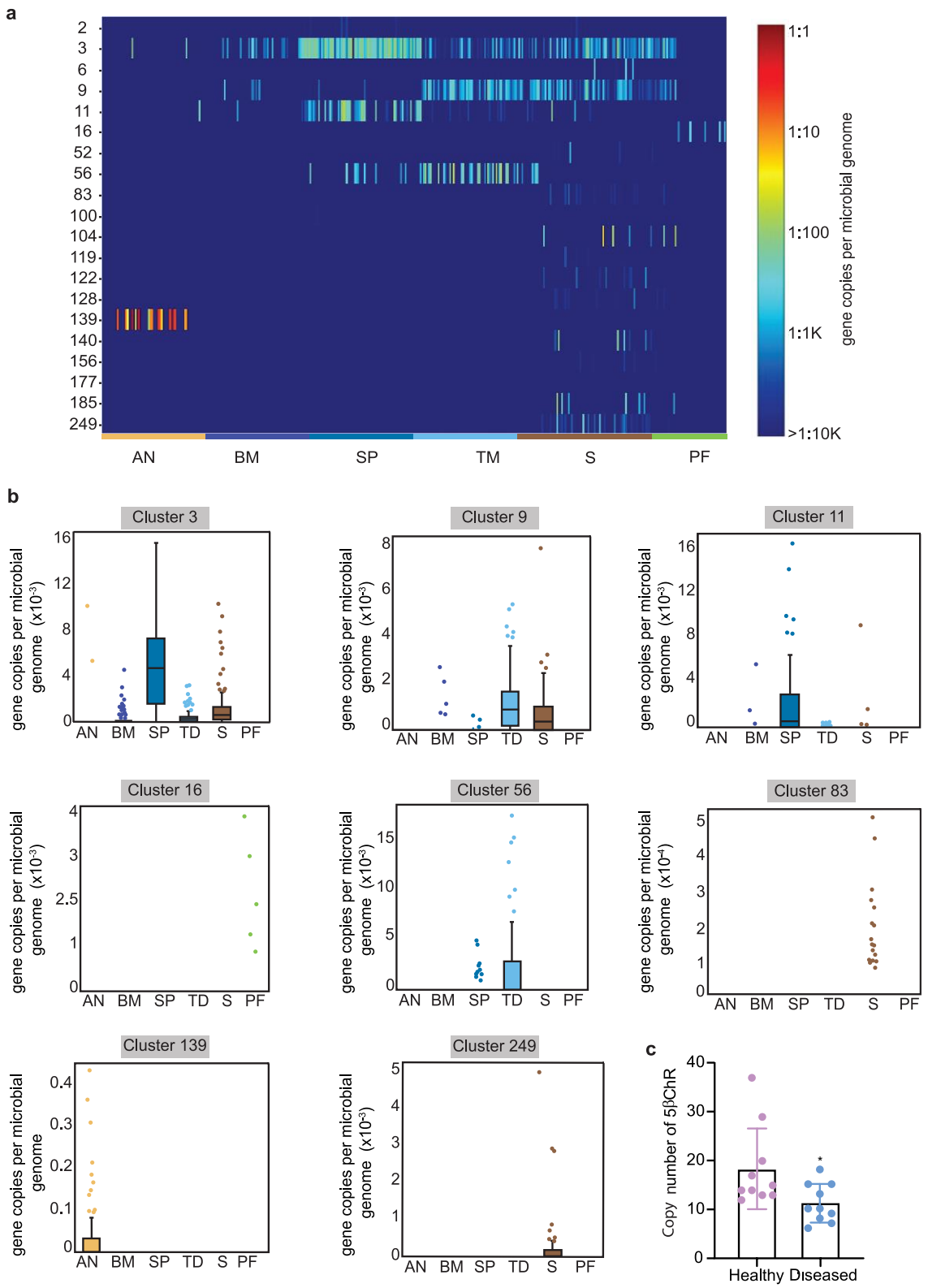


323 **Figure 4: Network maps illustrating the abundance and distribution of the**
324 **5βChR gene family across the different microbes.**

325 **(a)**. Microbial origin 5βChR protein homologs are retrieved from UniProt for sequence
326 similarity network (SSN) analysis. These networks are visualized by the Enzyme
327 Function Initiative-Enzyme Similarity Tool (EFI-EST) with threshold values of 10^{-80}. The
328 proteins are divided into 269 clusters based on sequence identity, structural
329 similarities, and functional profile of 5βChR proteins, highlighting their presence in
330 various microbes. **(b)** The graphs depict the taxonomy of microbial 5βChR homologs
331 from the selected clusters, which are prevalent in human microbiome data. The
332 representative sequences of clusters belong to different bacteria belonging to diverse
333 phyla and families.

334 **5βChR has an inverse association with diabetes**

335 Diabetes can induce changes in lipid levels, leading to diabetic dyslipidemia, which
336 escalates risk factors for CVDs[32-34]. We investigated the relationship between 5βChR
337 and diabetes using Chinese type 2 diabetes shotgun data (PRJNA3855128). The two
338 independent metagenomic cohorts were diabetic individuals (n=10) and healthy
339 controls (n=10). The interpretation of data validated that 5βChR exhibits a significantly
340 higher prevalence in healthy individuals than those with diabetes (*p*-0.0213) **(Fig. 5c)**.
341 This corroborates the inverse relationship between 5βChR and diabetes within this
342 cohort study, thereby indicating that 5βChR expression may play a role to lower the
343 chance of diabetic dyslipidemia.



AN- Anterior nares, BM- Buccal mucosa, SP- Supragingival plaque, TD- Tongue dorsum, S- Stool, and PF- Posterior fornix

Figure 5: Meta-analysis depicts the abundance of 5βChR homologs in healthy microbiomes and comparative analysis in a human diabetic cohort study.

347 **(a).** The heat map depicts the prevalence of 5βChR homologs in the human
348 metagenomic data of 380 healthy individuals. Analysis validates the presence of 20
349 clusters in the participants' metagenome data. **(b).** The box-plot graphs highlight the
350 presence of 5βChR homologs in various body sites. Clusters 3, 9, 11, and 56 dominate
351 the oral microbiome (supragingival plaque, buccal cavity, and tongue dorsum) and
352 stool samples. Cluster 16 is predominant in vaginal samples. The enzymes from other
353 clusters 139 showed higher distribution in anterior nares samples of human
354 microbiomes. Further, clusters 83 and 249 are abundant in stool samples, indicating
355 their abundance in the gut. **(c).** 5βChR has a lower abundance among diabetic patients
356 than healthy individuals in the shotgun metagenomic data of a Chinese type 2 diabetes
357 cohort study. Individuals with 5βChR high copy number are predominantly reported as
358 healthy. The points in the graph represent the number of individuals in each category.
359 *p value- 0.0213 is determined by the Mann-Whitney test.

360 **Discussion**

361 The gastrointestinal tract is a site for cholesterol metabolism and regulation of blood
362 cholesterol levels. Although cholesterol metabolism by lactobacilli has been known for
363 almost 50 years, the exact mechanism has remained elusive[3]. Recently, three
364 independent studies have reported the mechanism of cholesterol metabolism by
365 different bacterial gut residents. Still, the role of lactobacilli has not been addressed in
366 either of the reports[7,18,19]. Here, we provide the first mechanistic evidence for the direct
367 conversion of cholesterol to coprostanol by lactobacilli. We demonstrate that *L.*
368 *fermentum* NKN51 harbors 3β-OH-$\Delta^{5\text{-}6}$-cholesterol-5β-reductase(5βChR),
369 representing a novel class within the SDR family that catalysis the transformation of
370 cholesterol to coprostanol **(Fig. 6)**. The 5βChR sequence exhibits the characteristic

371 features of SDRs. Unlike typical microbial SDRs, 5βChR remains primarily a dimeric
372 protein that can catalysis direct coprostanol formation by reducing cholesterol's 5, 6
373 double bond. Hence, the distinctive features of 5βChR ascribed it to a newly identified
374 class within microbial SDRs. The structural and functional analysis of 5βChR
375 deciphered the molecular basis of the catalysis of cholesterol to coprostanol.
376 Molecular docking-inspired observations of the cofactor and substrate with the wild
377 type and mutant 5βChR identified the critical residues for $NADP^+$ binding and revealed
378 the novel residues important for cholesterol binding to the protein. Evolutionary
379 analysis of plant and microbial 5βChR enzymes suggests both are paraphyletic
380 groups. The phylogeny data confirms the presence of 5βChR in gut lactobacilli,
381 signifying its possible importance in human physiology. The shotgun data analysis
382 confirms low levels of 5βChR transcripts in diabetic patients. Thus, 5βChR may exert
383 an influence to lower the chance of premature coronary heart disease and stroke in
384 diabetic dyslipidemia conditions[35]. Metagenomic mining reveals that the
385 representatives of the 5βChR are distributed across various bacterial genera, with a
386 predominant presence in firmicutes. Our findings suggest that 5βChR homologs are
387 also acquired by uncharacterized bacterial classes, which may be explored further.
388 The prevalence of 5βChR in microbiomes of healthy individuals suggests its
389 paramount multifaceted functional aspects in lipid homeostasis. At the forefront of our
390 work, we unfolded the cholesterol transformation through its direct conversion to
391 coprostanol by a potential probiotic, *L. fermentum*. A similar strategy can unravel the
392 intricate interplay between gut microbe-mediated metabolism of diverse biological
393 molecules that influence physiological homeostasis or initiate complex diseases. By
394 delving into this mechanistic biology, we pave the way for developing next-generation
395 probiotics that can be explored for the formulation of functional foods and may also

synergize with host-targeted lipid-lowering medications, offering a promising intervention for combating hypercholesterolemia.

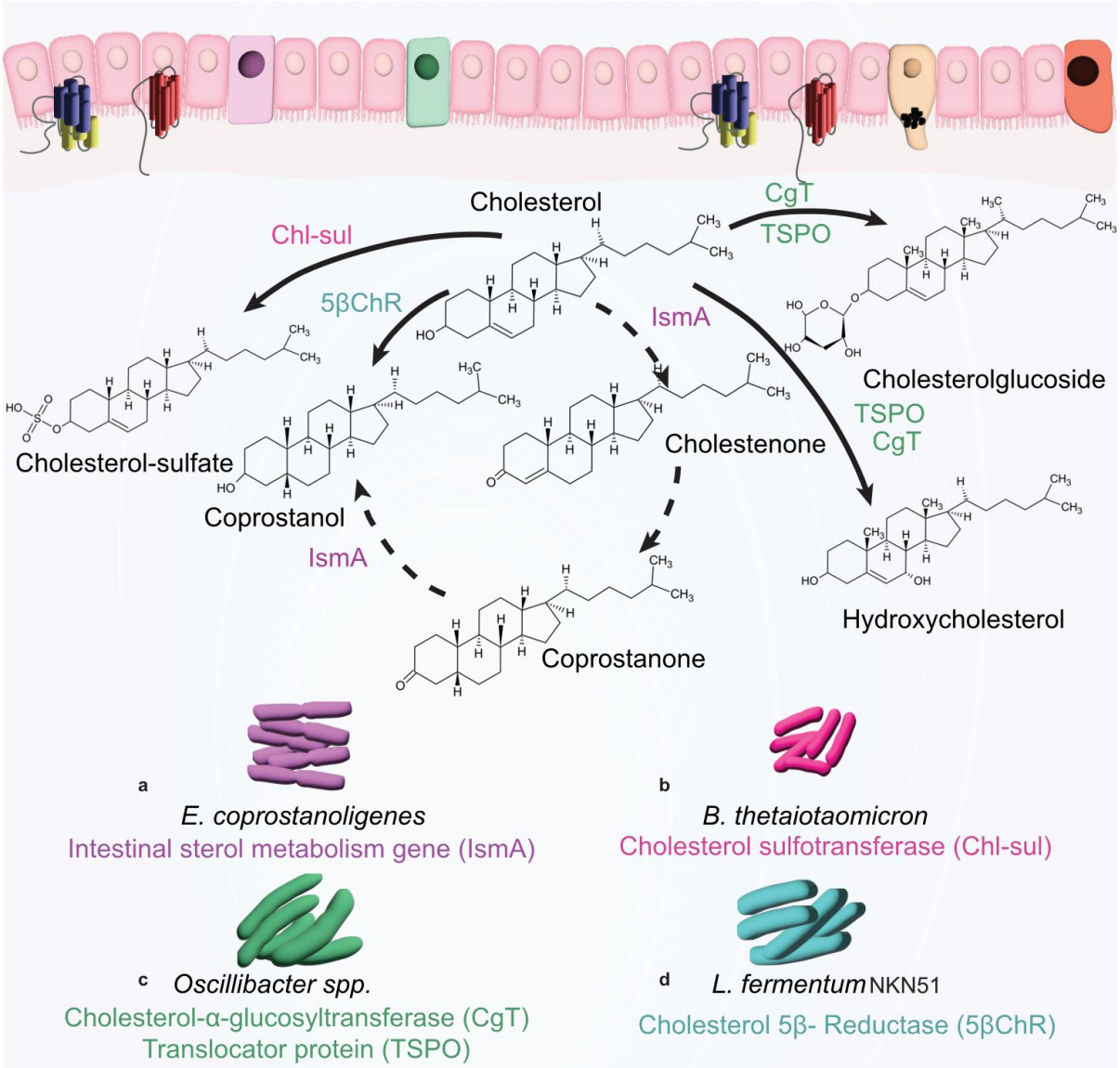


Figure 6: A comprehensive representation of microbially mediated cholesterol transformation in the human gut.

Various gut microbes transform cholesterol in the gut via direct and indirect pathways. **(a).** A cholesterol dehydrogenase (ismA) of *Eubacterium coprostanoligenes* is reported to catalyze the first and last step of the indirect pathway of cholesterol metabolism. **(b).** The gut cholesterol is also converted to cholesterol-sulphate with the help of a

405 cholesterol sulfotransferase encoded by *Bacteroidetes thetaiotaomicron.* **(c).** A group
406 of enzymes, cholesterol-α-glucosyltransferase, and translocator protein expressed by
407 *Oscillibacter* species transformed cholesterol into cholesterolglucoside and
408 hydroxycholesterol while **(d).** a comprehensive study on 5βChR *Limosilactobacillus*
409 *fermentum* NKN51 can directly convert cholesterol to coprostanol by a cholesterol-5β-
410 reductase.

411 **Online Methods**

412 ***In-vitro* cholesterol reduction assay**

413 *Limosilactobacillus fermentum* NKN51[36], *Lacticaseibacillus casei* NKN344, and
414 *Lacticaseibacillus casei* NKN347 were fermented food isolates from our lab. For the
415 cholesterol estimation assay, the bacterial strains were inoculated with cholesterol
416 (Sigma Aldrich, USA) (100 µg/ml) dissolved in methanol (Merck, USA) further, added
417 to sterile anaerobic Man, Rogosa, and Sharpe (MRS) media (Difco Laboratories, USA)
418 supplemented with Sodium thioglycolate (STG) (0.1%) (Tokyo Chemical Industry,
419 China) and incubated at 37°C for 24 h. After incubation, the samples were centrifuged
420 at 13,000 rpm for 5 min. The supernatant was subjected to a cholesterol estimation
421 assay using the Rudel and Morris method[37]. This was done in triplicates, and then
422 average values were plotted. *L. fermentum* NKN51 culture of 12 h was pelleted by
423 centrifugation at 4500 rpm for 20 min at 4°C and resuspended in phosphate buffer
424 saline (PBS) (1x) (Sigma Aldrich, USA) and lysed with Branson probe sonicator 102C
425 (CE) (Thermo Fisher Scientific, USA). Disrupted cells were centrifuged at 13,000 rpm
426 for 30 min at 4°C, and then the supernatant was collected separately. Cholesterol (100
427 µM) dissolved in methanol and $NADP^+$ (100 µM) (Sigma Aldrich, USA) were added to
428 the 500 µL supernatant and incubated at 37°C for 12 h. As described above, sterols

429 were extracted from these reaction mixtures and analyzed by TLC (explained later).
430 Reaction products were also checked with MALDI-TOF (details are in the MALDI-TOF
431 instrument section).

432 **Cloning of putative cholesterol dehydrogenase gene**

433 The genomic DNA of *L. fermentum* NKN51 was isolated using the HiPurA® Bacterial
434 Genomic DNA Purification Kit (Himedia, India). The gene encoding 5βChR was
435 amplified using Phusion High Fidelity DNA polymerase (Thermo Fisher Scientific,
436 USA) with the primers listed in **Supplementary File. 1**. Further, the PCR product and
437 pET28(a) were double digested with NdeI (Thermo Fisher Scientific, USA) and XhoI
438 (Thermo Fisher Scientific, USA). The digested products were ligated using T4 DNA
439 ligase (Thermo Fisher Scientific, USA) and transformed into chemically competent
440 cells of *E. coli*. Transformants were checked with double digestion of the construct,
441 followed by sequencing.

442 **Site-directed mutagenesis of putative Cholesterol dehydrogenase gene**

443 5βChR mutants were constructed with the Q5® Site-Directed Mutagenesis Kit (New
444 England Biolabs, USA). Primers containing the desired mutation were designed with
445 the NEBase Changer webpage tool (https://nebasechanger.neb.com/). The primers
446 were mentioned in the **Supplementary file. 1**, and transformants were confirmed by
447 sequencing.

448 **Overexpression and Purification of N-His Terminal Tagged WT and mutants**
449 **5βChR proteins**

450 Plasmid clones containing 5βChR were transformed into ultra-competent BL21AI
451 (Arabinose Induction) (Thermo Fisher Scientific, USA) cells for protein

overexpression. 5βChR protein (29.5 kDa) overexpression was optimized with IPTG (0.05 mM) (Tokyo Chemical Industry, China) and Arabinose (0.2%) (Himedia, India). For protein purification, the cells were grown in Luria-Bertani broth (Difco Laboratories, USA) supplemented with kanamycin (50 µg/ml) (Sigma Aldrich, USA) overnight at 37°C with shaking at 180 rpm. The secondary culture was inoculated with a primary culture of WT and mutant 5βChR (1%), incubated at 37°C with shaking at 180 rpm until optical density (OD_{600}) reached ~0.8. The bacterial cultures were induced with IPTG (0.05 mM) and 0.2% Arabinose (0.2%) and incubated at 18°C with shaking at 180 rpm for 18 h. Cells were harvested at 4500 rpm for 20 min at 4°C and resuspended into 20mL buffer A (50 mM Tris-Cl, 500 mM NaCl 5% glycerol) (pH-7.8). The resuspended cells were lysed with a Branson probe sonicator 102C (CE) and then centrifuged at 4500 rpm for 30 min and 4°C. The clarified supernatant was loaded into the buffer A saturated His Trap column of AKTA Pure-50L (GE Healthcare). Purified protein fractions were eluted with Buffer B (50 mM Tris-Cl, 500 mM NaCl, 500 mM Imidazole 5% glycerol) (pH-7.8) and analyzed using SDS-PAGE. Further, the protein was dialyzed in dialysis buffer (20 mM Tris-Cl, 300 mM NaCl 5% glycerol) with a 10 kDa membrane (Pall Corporation Centrifugal Devices, USA) for 8 h intervals. The protein was checked on SDS-PAGE and stored in aliquots at -80°C. Purified protein fractions were analyzed for protein-substrate binding and enzymology assays using a method previously explained.

Instrumentation conditions of MALDI-TOF for measurement of sterols

A Bruker ultrafleXtrem mass spectrometer equipped with a smart-beam-II laser was used to identify the cholesterol and related metabolites. A lipid sample mixed with 2,5-Dihydroxybenzoic acid (DHB) matrix (1:1) and 1 µl mixture of sample and matrix were

spotted on the MTP AnchorChip target plate (Bruker Daltonics GmbH). Nd: YAG laser (355 nm/2000Hz/500 shots/50-60% intensity) was used for the ionization of the sample, and mass spectra were acquired in the positive reflectron mode at 100-800 Da mass range using FLEXCONTROL 3.0, (Bruker). Spectra were processed by centroid peak picking, and spectrum analysis was done in LEXANALYSIS 3.0 (Bruker).

qPCR of 5βChR transcript

The chloroform-phenol (VWR, USA) extraction method was used for total RNA purification from triplicate *L. fermentum* NKN51 cell pellets from cultures grown with or without cholesterol for 24 h. DNase-treated RNA samples were used to prepare cDNA using PrimeScript 1st strand cDNA Synthesis Kit (Takara, Japan). 5βChR transcript was quantified by real-time PCR with PowerTrack SYBR Green Master Mix (Applied Biosystems, USA). It was normalized with 16S rRNA gene expressions. Primers used in the analysis are listed in the **Supplementary File. 1**.

Thin Layer Chromatography of sterols

The reaction set-up contains an enzyme (10 nm), $NADP^+$ (100 μM), and substrate (100 μM) (cholesterol, cholestenone, and coprostanone) in the dialysis buffer, incubated at 37°C for 12 h. These reaction mixtures were used for sterol extraction using the previously explained method and analyzed on TLC. The organic layer was evaporated, resuspended in 10 μl of methanol, and spotted on a pre-treated TLC plate (silica gel IB2-F flexible 20 x 20 cm) with methanol at 100°C for 30 min. The sterols were separated with mobile phase (Benzene: Chloroform) (90:10) (Merk, USA). The Plate was visualized after being treated with 15% sulphuric acid in methanol and incubated at 110°C for 10 min.

500 **Gel Filtration chromatography of 5βChR**

501 5βChR gel filtration chromatography (column: Hiload™ 16/600 Superdex™ 200pg)
502 was performed on an AKTA pure-50L chromatography system (GE Healthcare) at 4°C.
503 The concentrated, purified protein (10 mg/mL) was loaded onto the analytical column
504 equilibrated with Tris-Cl (20 mM) and NaCl buffer (200 mM) at pH 7.8 and eluted at a
505 flow rate of 1ml/min. A protein standard consisting of thyroglobulin bovine, γ-globulins,
506 Ovalbumin, Ribonuclease A type I, and p-aminobenzoic acid (pABA) (Merck, India)
507 was used to determine the native molecular mass of 5βChR.

508 **Enzyme kinetic assay**

509 The enzyme activity of the purified protein was measured spectrophotometrically by
510 monitoring $NADP^+$ levels at 340nm. The standard reaction mixture comprised sodium
511 acetate trihydrate buffer (20 mM) (pH 7.8), $NADP^+$ (100 μM), enzyme (20 nm), and
512 substrate concentration varied from 10 to 300 μM. GraphPad Prism was used to plot
513 the data using the Michaelis-Menton saturation curve and Lineweaver-Burk plots to
514 calculate kinetic parameters.

515 **Substrate-specificity assay**

516 Cholic acid, lithocholic acid, pregnenolone, hydrocortisone (100 μM) (Tokyo Chemical
517 Industry, China), or different substrate analogs of cholesterol were separately added
518 with the same cholesterol concentration as a control. The reaction was incubated at
519 37°C for 12 h after the sterols were extracted and run on the MALDI to check the
520 enzyme activity 5βChR for structure analogs of cholesterol.

521 **Circular Dichroism spectroscopy of WT 5βChR and mutants.**

522 CD spectroscopy of 5βChR (WT) or mutants was conducted using a J-815 CD
523 spectropolarimeter (Jasco, Tokyo, Japan). The samples of the protein (20 μM) were
524 prepared in KH_2PO_4 buffer (10 mM) (pH 7.8). For the CD scan, a quartz cell with a 0.1
525 cm was used, and scans were carried out at 25°C from 190 to 260 nm at a speed of
526 50 nm/min with wavelength pitch with three accumulation data. Data were analyzed
527 using GraphPad Prism.

528 **Isothermal Titration Calorimetry assay for binding analysis**

529 ITC analysis was performed using a MicroCal PEAQ-ITC (Malvern Panalytical, United
530 Kingdom) with 300 cell volumes. $NADP^+$ cofactor was dissolved in the dialysis buffer,
531 and the substrate was dissolved in methanol. Purified 5βChR (20 μM) was injected
532 with 18-successive 2 μl aliquots of $NADP^+$ (1 mM) at 150 s intervals. Further, protein-
533 $NADP^+$ was injected similarly with a substrate (500 μM). The data were fitted to a
534 nonlinear regression model using a single binding site with the MicroCal origin
535 software. The thermodynamic parameters were calculated using the Gibbs free energy
536 equation ($\Delta G= \Delta H-T\Delta S$ and $\Delta G= -RT\ln (K_a)$).

537 **5βChR structure determination**

538 5βChR structural modelling was done using ColabFold (v.1.5.5) implementation of
539 Alphafold (https://github.com/sokrypton/ColabFold)[38,39]. The multiple sequence
540 alignment was done using the MMseq2 algorithm[40]. The models were ranked through
541 the pTM score, and the highest-scoring model was used for analysis. PyMOL[41] (v2.5.7
542 Schrödinger, LLC) was used for protein visualization and structural analysis of 5βChR.
543 DALI server[42] was used for structural similarity search.

544 **Molecular docking of multiple ligands docking with 5βChR**

545 The three-dimensional structure of 5βChR was predicted using ColabFold (v.1.5.5),
546 which uses MMseqs2 along with AlphaFold2 in its background[38,39]. The potential
547 ligands of 5βChR-$NADP^+$-cholesterol datasets were downloaded from the ZINC drug
548 database in SD format (https://zinc.docking.org/). AutoDockTools-1.5.7 was used to
549 convert the 5βChR PDB files and ligand structures into PDBQT format[43]. Autodock
550 Vina (v.1.2.3) performed molecular docking with multiple ligands[44]. The protein-ligand
551 interactions were computed and were subsequently visualized by PyMOL[41] (v2.5.7
552 Schrödinger, LLC).

553 **Phylogenetic analysis of 5βChR within lactobacilli and plants**

554 The prevalence of 5βChR was examined by creating a protein-based custom-made
555 database for available lactobacilli from RefSeq using BLAST (v2.15.0)
556 **(Supplementary File. 1)**. 5βChR was used as a query search in the database, and
557 the homologous proteins were retrieved for evolutionary studies. To trace the ancestral
558 relationship of 5βChR in plants, 100 related sequences were taken out from UniProt[45]
559 (https://www.uniprot.org/). MUSCLE (v3.8.31)[46] aligned all respective sequences, and
560 phylogenetic analysis was done using RAxML (v8) by the maximum likelihood Protcat
561 blossom62 model[47].

562 **5βChR sequence similarity network data analysis**

563 For 5βChR sequence similarity network (SSN)[48] analysis, its amino acid sequence
564 was used as a query in the UniProt database using BLASTp with a series of cutoff e-
565 values- 10^{-60}, 10^{-70}, and 10^{-8} and protein identity values of >50%, >60% and >70%
566 respectively. The EFI-EST tool was used to construct the protein SSNs of these
567 sequences with various e-values and visualized by Cytoscape 3.3[49].

568 **Determination of 5βChR abundance using healthy human metagenome data**

569 5βChR abundance was checked in the HMP[50,51] as previously described using the
570 ShortBRED[52,53] **(Supplementary File. 1)**. This was done using the EFI-chemically
571 guided functional profiling (EFI-CGFP) platform[54-56]. Using UniProt, 5βChR sequence
572 representatives were filtered at 80% amino acid similarity thresholds to identify non-
573 redundant peptide markers. ShortBRED-quantify was used with default parameters to
574 quantify the marker abundances in the 380 metagenomic database. The relative
575 abundance is based on reads per kilobase million (RPKM) values calculated and
576 converted into 'copies per microbial genome' as explained earlier[56,57].

577 **Abundance analysis of 5βChR gene in human gut metagenomes under disease**
578 **conditions**

579 The shotgun metagenomic sequencing datasets of the human gut microbiomes
580 describing diabetes were downloaded from the Sequence Read Archive of the NCBI
581 (https://www.ncbi.nlm.nih.gov/sra). The dataset was divided into healthy and diseased
582 based on the metadata provided with the original data. Read quality was checked
583 using FastQC (v0.11.9) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
584 followed by MultiQC (v1.11)[58]. Only reads with a phread score higher than 30 were
585 included in the study. The high-quality sequencing reads were assembled using
586 MEGAHIT (v1.2.9) with a minimum contig length of 500 bp[59]. Assembled sequences
587 were matched to the nucleotide sequence of 5βChR using DIAMOND (v2.0.15.1531)
588 with the following parameters: e-value – 0.00001, query coverage 30, percent identity
589 30[60].

590 **Data availability**

591 The data is provided in the manuscript as supplementary information.

592 **Code availability**

593 No custom-made code was used in this study.

594 **References**

595 1. Underhill, D., Gordon, S., Imhof, B. *et al.* Élie Metchnikoff (1845–1916):
596 celebrating 100 years of cellular immunology and beyond. *Nat Rev*
597 *Immunol* **16**, 651–656 (2016).
598 2. Dempsey E and Corr SC Lactobacillus spp. for Gastrointestinal Health: Current
599 and Future Perspectives. *Front. Immunol*. **13**:840245 (2022).
600 doi:10.3389/fimmu.2022.840245
601 3. Jia, B., Zou, Y., Han, X., Bae, J. W., & Jeon, C. O. Gut microbiome-mediated
602 mechanisms for reducing cholesterol levels: implications for ameliorating
603 cardiovascular disease. *Trends Microbiol.*, **31(1)**, 76–91 (2023).
604 4. Puglielli, L., Tanzi, R. & Kovacs, D. Alzheimer's disease: the cholesterol
605 connection. *Nat Neurosci* **6**, 345–351 (2003).
606 5. Berger, S., Raman, G., Vishwanathan, R., Jacques, P. F., & Johnson, E. J.
607 Dietary cholesterol and cardiovascular disease: a systematic review and meta-
608 analysis. *The Am. J. Clin. Nutr.*, **102(2)**, 276–294 (2015).
609 6. Song, Y., Liu, J., Zhao, K., Gao, L., & Zhao, J. Cholesterol-induced toxicity: An
610 integrated view of the role of cholesterol in multiple diseases. *Cell Metab.*,
611 **33(10)**, 1911–1925 (2021).
612 7. Kenny, D. J., Plichta, D. R., Shungin, D., Koppel, N., Hall, A. B., Fu, B., Vasan,
613 R. S., Shaw, S. Y., Vlamakis, H., Balskus, E. P., & Xavier, R. J. Cholesterol

614 Metabolism by Uncultured Human Gut Bacteria Influences Host Cholesterol
615 Level. *Cell host & microbe*, **28(2)**, 245–257.e6 (2020).
616 8. Mythen SMDevendran SMéndez-García C, Cann I, Ridlon JM. Targeted
617 Synthesis and Characterization of a Gene Cluster Encoding NAD(P)H-
618 Dependent 3α-, 3β-, and 12α-Hydroxysteroid Dehydrogenases from
619 Eggerthella CAG:298, a Gut Metagenomic Sequence. Appl Environ Microbiol.,
620 **84** : e02475-17 (2018).
621 9. Doden, H. L., Pollet, R. M., Mythen, S. M., Wawrzak, Z., Devendran, S., Cann,
622 I., Koropatkin, N. M., & Ridlon, J. M. Structural and biochemical characterization
623 of 20β-hydroxysteroid dehydrogenase from Bifidobacterium adolescentis strain
624 L2-32. *J. Biol Chem.* **294(32)**, 12040–12053 (2019).
625 10. Kallberg, Y., Oppermann, U., Jörnvall, H., & Persson, B. Short-chain
626 dehydrogenases/reductases (SDRs). *Eur J Biochem*., **269(18)**, 4409–4417
627 (2002).
628 11. Penning T.M. Molecular endocrinology of hydroxysteroid dehydrogenases.
629 *Endocr. Rev.*; **18 (9183566)**: 281-305 (1997).
630 12. Ridlon J.M., Ikegawa S., Alves J.M., Zhou B., Kobayashi A., Iida T., Mitamura
631 K., Tanabe G., Serrano M., De Guzman A., Cooper P., Buck G.A., Hylemon P.B.
632 Clostridium scindens: a human gut microbe with a high potential to convert
633 glucocorticoids into androgens. *J. Lipid Res*. **54 (23772041)**: 2437-2449 (2013).
634 13. Kallberg, Y., Oppermann, U., Jörnvall, H., & Persson, B. Short-chain
635 dehydrogenase/reductase (SDR) relationships: a large family with eight
636 clusters common to human, animal, and plant genomes. *Protein Sci*, **11(3)**, 636-
637 641 (2002).

638 14. Filling C., Berndt K.D., Benach J., Knapp S., Prozorovski T., Nordling E.,
639 Ladenstein R., Jörnvall H., Oppermann U. Critical residues for structure and
640 catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.* **277**
641 **(11976334)**: 25677-2568410.1074/jbc.M202160200, (2002)
642 15. Rossmann, M. G., Ford, G. C., Watson, H. C., & Banaszak, L. J. Molecular
643 symmetry of glyceraldehyde-3-phosphate dehydrogenase. *J Mol Biol.* **64(1)**,
644 237–245 (1972).
645 16. Thorn, A., Egerer-Sieber, C., Jäger, C. M., Herl, V., Müller-Uri, F., Kreis, W., &
646 Muller, Y. A. The crystal structure of progesterone 5β-reductase from Digitalis
647 lanata defines a novel class of short-chain dehydrogenases/reductases. *J Biol*
648 *Chem*., **283(25)**, 17260-17269 (2008).
649 17. Chen, M., & Penning, T. M. 5β-Reduced steroids and human Δ4-3-ketosteroid
650 5β-reductase (AKR1D1). *Steroids*, **83**, 17-26 (2014).
651 18. Le, H.H., Lee, MT., Besler, K.R. et al. Characterization of interactions of dietary
652 cholesterol with the murine and human gut microbiome. *Nat Microbiol.* **7**, 1390–
653 1403 (2022).
654 19. Li, C., Stražar, M., Mohamed, A.M.T., Pacheco, J. A., Walker, R. L., … & Xavier,
655 R. J. Gut microbiome and metabolome profiling in Framingham heart study
656 reveals cholesterol-metabolizing bacteria. *Cell*, **187**, 1-19. (2024).
657 20. Rosenfeld, R., Fukushima, D. K., Hellman, L., & Gallagher, T. F. The
658 transformation of cholesterol to coprostanol. *J Biol Chem*, **211(1)**, 301-311
659 (1954).
660 21. Björkhem, I., & Gustafsson, J. A. Mechanism of microbial transformation of
661 cholesterol into coprostanol. *Eur J Biochem.,* **21(3)**, 428–432 (1971).

662 22. Harrison, V. C., & Peat, G. Serum cholesterol and bowel flora in the newborn.
663 *Am. J. Clin. Nutr.*, **28(12)**, 1351-1355 (1975).
664 23. Marshall, V. M., & Taylor, E. Ability of neonatal human Lactobacillus isolates to
665 remove cholesterol from liquid media. *Int. J. Food Sci. Technol.*, **30(5)**, 571-577
666 (1995).
667 24. Lye, H. S., Rusul, G., & Liong, M. T. Removal of cholesterol by lactobacilli via
668 incorporation and conversion to coprostanol. *J. Dairy Sci.*, **93(4)**, 1383-1392
669 (2010).
670 25. Fuentes, M. C., Lajo, T., Carrión, J. M., & Cuné, J. Cholesterol-lowering efficacy
671 of Lactobacillus plantarum CECT 7527, 7528 and 7529 in
672 hypercholesterolaemic adults. *Br. J. Nutr.*, **109(10)**, 1866-1872 (2013).
673 26. Costabile, A., Buttarazzi, I., Kolida, S., Quercia, S., Baldini, J., Swann, J. R., ...
674 & Gibson, G. R. An in vivo assessment of the cholesterol-lowering efficacy of
675 Lactobacillus plantarum ECGC 13110402 in normal to mildly
676 hypercholesterolaemic adults. *PLoS One*, **12(12)**, e0187964 (2017).
677 27. Tjandrawinata, R. R., Kartawijaya, M., & Hartanti, A. W. In vitro evaluation of
678 the anti-hypercholesterolemic effect of Lactobacillus Isolates from various
679 sources. *Front. Microbiol.*, **13**, 825251 (2022).
680 28. Cunningham, M., Azcarate-Peril, M. A., Barnard, A., Benoit, V., Grimaldi, R.,
681 Guyonnet, D., ... & Gibson, G. R. Shaping the future of probiotics and prebiotics.
682 *Trends Microbiol.*, **29(8)**, 667-685 (2021).
683 29. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J.,
684 & Ghosh, D. Short-chain dehydrogenases/reductases (SDR). *Biochemistry*,
685 **34(18)**, 6003-6013 (1995).

686 30. Walter, J. Ecological role of lactobacilli in the gastrointestinal tract: implications
687 for fundamental and biomedical research. *Appl. Environ. Microbiol.*, **74(16)**,
688 4985-4996 (2008).
689 31. Dang, D., Liu, X., Chen, H., Zhao, J., Chen, W., & Yang, B. In silico, in vitro and
690 in vivo safety assessment of three Limosilactobacillus mucosae strains for
691 probiotic candidates. *Food Biosci.*, **57**, 103462 (2024).
692 32. Mihaylova, B., Emberson, J., Blackwell, L., Keech, A., Simes, J., Barnes, E. H.,
693 ... & Baigent, C. The effects of lowering LDL cholesterol with statin therapy in
694 people at low risk of vascular disease: a meta-analysis of individual data from
695 27 randomised trials. *Lancet*, **380(9841)**, 581-590 (2012).
696 33. Bitzur, R., Cohen, H., Kamari, Y., Shaish, A., & Harats, D. Triglycerides and HDL
697 cholesterol: stars or second leads in diabetes? *Diabetes Care*, **32(Suppl 2)**,
698 S373 (2009).
699 34. Fernandez, M. L., & Andersen, C. J. Effects of dietary cholesterol in diabetes
700 and cardiovascular disease. *Clin Lipidol*, **9(6)**, 607-616 (2014).
701 35. https://www.heart.org/en/health-topics/diabetes/diabetes-complications-and-
702 risks/cholesterol-abnormalities--diabetes
703 36. Beniwal, A., Gaurav, A., Sharma, R., Ghosh, T., Kumar, P., Netter, U., ... &
704 Navani, N. K. Draft Genome Sequence of Limosilactobacillus fermentum Strain
705 NKN-51, Isolated from Fermented Yak Milk in the Western Himalayas of India.
706 *Microbiol. Resour. Announc.*, **10(16)**, 10-1128 (2021).
707 37. Rudel, L. L., & Morris, M. D. Determination of cholesterol using o-
708 phthalaldehyde. *J. Lipid Res.*, **14(3)**, 364-366 (1973).

38. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature*, **596(7873)**, 583-589 (2021).
39. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods*, **19(6)**, 679-682 (2022).
40. Steinegger, M., & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.*, **35(11)**, 1026-1028 (2017).
41. Schrödinger, L., & DeLano, W. (2020). *PyMOL*. Retrieved from http://www.pymol.org/pymol
42. Holm, L., & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.*, **38**(Web Server issue), W545–W549 (2010).
43. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.*, **30(16)**, 2785-2791 (2009).
44. Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. *J Chem Inf Model.*, **61(8)**, 3891-3898 (2021).
45. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.*, **43(D1)**, D204-D212 (2015).
46. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797 (2004).
47. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30(9)**, 1312-1313 (2014).

734 *48.* Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D.
735 R., & Whalen, K. L. Enzyme function initiative-enzyme similarity tool (EFI-EST):
736 a web tool for generating protein sequence similarity networks. *Biochim*
737 *Biophys Acta.*, **1854(8)**, 1019-1037 (2015).
738 49. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ...
739 & Ideker, T. Cytoscape: a software environment for integrated models. *Genome*
740 *Res.*, **13(22)**, 426 (1971).
741 50. Human Microbiome Project Consortium. Structure, function and diversity of the
742 healthy human microbiome. *Nature* **486**: 207–214, 2012.
743 doi:10.1038/nature11234.
744 51. Methé, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower,
745 C., ... & Mannon, P. J. A framework for human microbiome research. *Nature*,
746 **486(7402)**, 215 (2012).
747 52. Jonas Coelho Kasmanas, Alexander Bartholomäus, Felipe Borim Corrêa,
748 Tamara Tal, Nico Jehmlich, Gunda Herberth, Martin von Bergen, Peter F
749 Stadler, André Carlos Ponce de Leon Ferreira de Carvalho, Ulisses Nunes da
750 Rocha, HumanMetagenomeDB: a public repository of curated and
751 standardized metadata for human metagenomes, *Nucleic Acids Res.*, Volume
752 49, Issue D1, 8 January 2021, Pages D743–D750
753 53. Kaminski, J., Gibson, M. K., Franzosa, E. A., Segata, N., Dantas, G., &
754 Huttenhower, C. High-specificity targeted functional profiling in microbial
755 communities with ShortBRED. *PLoS Comput. Biol.*, **11(12)**, e1004557 (2015).
756 54. Zallot, R., Oberg, N., & Gerlt, J. A. The EFI web resource for genomic
757 enzymology tools: leveraging protein, genome, and metagenome databases to

758 discover novel enzymes and metabolic pathways. *Biochemistry*, **58(41)**, 4169-
759 4182 (2019).
760 55. Oberg, N., Zallot, R., & Gerlt, J. A. EFI-EST, EFI-GNT, and EFI-CGFP: enzyme
761 function initiative (EFI) web resource for genomic enzymology tools. *J. Mol.*
762 *Biol.*, **435(14)**, 168018 (2023).
763 56. Levin, B. J., Huang, Y. Y., Peck, S. C., Wei, Y., Martinez-Del Campo, A., Marks,
764 J. A., ... & Balskus, E. P. A prominent glycyl radical enzyme in human gut
765 microbiomes metabolizes trans-4-hydroxy-l-proline. *Science*, **355(6325)**,
766 eaai8386 (2017).
767 57. Jia, B., Park, D., Hahn, Y., & Jeon, C. O. Metagenomic analysis of the human
768 microbiome reveals the association between the abundance of gut bile salt
769 hydrolases and host health. *Gut Microbes*, **11(5)**, 1300-1313 (2020).
770 58. Ewels, P., Magnusson, M., Lundin, S., & Käller, M. MultiQC: summarize
771 analysis results for multiple tools and samples in a single report. *Bioinformatics*,
772 **32(19)**, 3047-3048 (2016).
773 59. Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. MEGAHIT: an ultra-fast
774 single-node solution for large and complex metagenomics assembly via
775 succinct de Bruijn graph. *Bioinformatics*, **31(10)**, 1674-1676 (2015).
776 60. Buchfink, B., Xie, C., & Huson, D. H. Fast and sensitive protein alignment using
777 DIAMOND. *Nat. Methods*, **12(1)**, 59-60 (2015).

778 **Acknowledgments**

779 Urmila Netter thanks the Council of Scientific and Industrial Research (CSIR), India,
780 for providing a research fellowship. Vishakha Bisht and Amit Gaurav thank the Ministry
781 of Human Resource and Development and the Department of Biotechnology (DBT)

Govt. of India for the financial support. We thank the Department of Biotechnology (DBT), Govt. of India, for providing the funding to carry out the research work.

Contributions

U.N. designed, data curation, methodology, conceptualization of the project, and writing the original draft. V.B. methodology and writing. A.G. Methodology, Data curation, visualization, and writing. R.S. methodology. A.G. methodology. V.S.B. methodology and data curation. KA. Formal analysis. K.T. Formal analysis and writing. NKN Formal analysis, conceptualization, writing, resources, and project administration.

Ethics declarations

Competing interests

The authors declare no competing interests.