Discovery and mechanistic characterization of a probiotic-origin 3β-OH-Δ^{5-6}-cholesterol-5β-reductase directly converting 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], Hanuman Prasad Sharma[c], Sujata Mohanty[d], Shubham Loat[e], Mihir Sarkar[e], Kapil Tahlan[f], 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]Centralized Core Research Facility, All India Institute of Medical Sciences Ansari Nagar, New Delhi – 110029, India

[d]Stem Cell Facility, (DBT-Centre of Excellence for Stem Cell Research), All India Institute of Medical Sciences, Ansari Nagar, New Delhi– 110029, India.

[e]ICAR-National Research Centre on Yak, Dirang, West Kameng district, Arunachal Pradesh-790101, India

[f]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

Keywords: Cholesterol, $NADP^+$, *Limosilactobacillus fermentum* NKN51, Coprostanol LC-MS.

Conflict of Interest

The authors declare no competing interests.

Abstract

On the one hand, cholesterol is a foundational molecule for various structural and biochemical pathways, while elevated cholesterol levels are associated with cardiovascular diseases. Some selected strains of Lactobacilli are also known for modulating cholesterol levels. However, the molecular mechanism of cholesterol transformation by lactobacilli has remained challenging to reveal. This study reports the discovery and role of a microbial 3β-OH-Δ^{5-6}-cholesterol-5β-reductase from *Limosilactobacillus fermentum* NKN51, which directly converts cholesterol to coprostanol, thereby resolving the enigma. Protein engineering of the reductase enzyme identified the cholesterol and $NADP^+$ interacting amino acids, detailing the catalytic mechanism of 5βChR. Phylogenetic studies emphasize the abundance of 5βChRs in gut commensal lactobacilli, which shares a common ancestor with plant 5β reductases. Meta-analysis results of healthy participant microbiomes underline the significance of 5βChR homologs, while a cohort study reveals an association between higher 5βChR abundance and diabetes. The discovery of the 5βChR enzyme and its molecular mechanism in cholesterol metabolism pave the way for a better understanding of the gut-associated microbiome and the design of practical applications to ameliorate dyslipidemia.

Introduction

46 Cholesterol is a vital animal sterol, serving essential structural and metabolic functions.
47 However, high cholesterol levels are a critical risk factor for cardiovascular diseases
48 (CVDs) and atherosclerosis, which can lead to heart failure, obesity, diabetes, fatty
49 liver, and Alzheimer's disease [1-3]. Gut microbes can transform dietary cholesterol
50 into coprostanol by direct stereospecific reduction of the Δ5 double bond or via forming
51 intermediates such as cholestenone and coprostanone **(Fig. S1)** [4]. The short-chain
52 dehydrogenase (SDRs) enzyme family has recently been shown to contribute to gut
53 cholesterol transformation [5]. The commensal gut microflora possesses SDRs that
54 facilitate the conversion of hydroxyl groups in alcohols, carbohydrates, and xenobiotics
55 from their enol to keto forms [6-8]. The SDRs contain conserved motifs spanning the
56 Rossmann fold for $NADP^+/NAD^+$ cofactor binding and a catalytic tetrad comprising
57 serine, tyrosine, lysine, and asparagine residues [9, 10]. Eight 5β reductases classified
58 as SDRs have been thoroughly studied in plants; however, their role in bacteria has
59 received considerably less attention[11, 12]. Only recently has the mechanism of
60 cholesterol metabolism been described for the obligatory gut anaerobes *Eubacterium*
61 *coprostanoligenes, Bacteroides thetaiotaomicron,* and *Oscillibacter* species [5, 13,
62 14]. The genome of *E. coprostanoligenes* encodes the cholesterol dehydrogenase
63 enzyme 'Intestinal steroid metabolism A' (IsmA), which converts cholesterol into
64 cholestenone and coprostanone into coprostanol [5]. However, information on the
65 steps involved in transforming cholestenone into coprostanone is lacking. On the other
66 hand, *B. thetaiotaomicron* exhibits unique sulfotransferase activity that converts
67 cholesterol into cholesterol sulfate [13]. Furthermore, *Oscillibacter* species convert
68 cholesterol to cholestenone, cholesterol glucoside, and hydroxycholesterol [14].
69 Overall, coprostanol, cholesterol sulfate, cholesterol glucoside, and
70 hydroxycholesterol are excreted, thereby regulating cholesterol homeostasis [5, 13,

14]. These anaerobes do not metabolize cholesterol directly. Although the direct conversion of cholesterol to coprostanol in the gut was suggested as early as 1954, the underlying mechanism of this transformation remains a mystery [4, 15, 16]. Selected strains of probiotic lactobacilli display potential for cholesterol metabolism. However, the specific mechanisms involved in cholesterol reduction have not yet been elucidated [17-22].

We report a novel cholesterol-specific microbial 5β reductase (5βChR) from *Limosilactobacillus fermentum* NKN51 isolated from fermented Himalayan yak milk. Our study provides mechanistic evidence of the direct conversion of cholesterol to coprostanol by 5βChR through its 3β-OH-Δ^{5-6}-cholesterol-5β-reductase activity. The biochemical and structural analysis of 5βChR identify and characterize cholesterol metabolism in lactobacilli. Furthermore, phylogenetic and metagenomic analysis suggests the prevalence of 5βChR in microbes and its association with human physiology. Our findings addressed a primarily unexplored frontier of microbial cholesterol metabolism in *L. fermentum* NKN51 by identifying a novel enzyme class from the SDR family, establishing a path for unveiling the metabolic capabilities of one of the most popular probiotic bacteria.

Results

L. fermentum NKN51 exhibit 3β-OH-Δ^{5-6}-cholesterol-5β-reductase activity

To understand the mechanism of cholesterol metabolism in lactobacilli, we performed cholesterol depletion assays with a battery of Lactobacillus spp. isolated from an ethnic fermented yak milk cheese (churppi) prepared by the native population in the Western Indian Himalayas. Among other strains, *L. fermentum* NKN51 depleted the maximum (~64%) cholesterol in the media **(Fig. S2)**. Thin layer chromatography (TLC)

95 and Liquid chromatography-mass spectrophotometry (LC-MS) analysis of *L.*
96 *fermentum* NKN51 whole-cell lysate confirmed the conversion of cholesterol to
97 coprostanol **(Fig. S3, 1*a*)**. We could not detect cholestenone or coprostanone in these
98 assays, suggesting that *L. fermentum* NKN 51 might directly convert cholesterol to
99 coprostanol in $NADP^+$-dependent but oxygen-independent conditions **(Fig. S4, S5)**.
100 Furthermore, a known cholesterol SDR from *E. coprostanoligenes* HL ATCC51222[5]
101 (WP_078769004.1) was used as a query against the *L. fermentum* NKN51 draft
102 genome, which highlights the putative cholesterol dehydrogenase protein with ~30%
103 identity in NKN51. The putative cholesterol dehydrogenase was cloned, and the 29.5
104 kDa overexpressed protein was purified by immobilized affinity chromatography
105 (IMAC) **(Fig. S6)**. The recombinant purified protein converted cholesterol to
106 coprostanol using $NADP^+$ as a cofactor instead of NAD^+, validating the activity we
107 initially observed with *L. fermentum* whole-cell lysate assays **(Fig. 1*a*)**. To determine
108 the binding affinities of protein with its cofactor and substrates, we performed
109 isothermal titration calorimetry (ITC) analysis. $NADP^+$ had an affinity for protein with
110 K_d 6.88 μM and the protein-$NADP^+$ complex bound with cholesterol with K_d of 4.66 μM
111 **(Fig. 1*b*, *c*)**. However, protein-$NADP^+$ did not display an affinity for cholestenone or
112 coprostanone, substantiating that neither of these compounds is cognate substrates
113 of the enzyme **(Fig. S7)**. Unlike typical microbial SDRs, the cholesterol dehydrogenase
114 remains primarily a dimeric protein in solution, confirmed by size exclusion
115 chromatography **(Fig. S8)**. The substrate specificity experiments validated that protein
116 can specifically catalyze the direct conversion of cholesterol to coprostanol by
117 saturating Δ^5 double bond with the OH group at the 3β position **(Fig. S9, 1*d*)**. Hence,
118 these distinctive catalytic features of the protein ascribed it to a newly identified class
119 within microbial SDRs, i.e., 3β-OH- $\Delta^{5\text{-}6}$- cholesterol-5β-reductase (5βChR). Enzyme-

120 substrate kinetic studies confirmed that 5βChR obeys the Michaelis–Menten equation,
121 with a K_m of 19.09 μM and V_{max} of 15.35x10^{-6} μmol sec^{-1} $nmol^{-1}$ **(Fig. S10)**. At the
122 genetic level, 5βChR transcript expression levels were ~1.5-fold increased when the
123 NKN51 strain was grown in the presence of cholesterol, signifying that the 5βChR
124 genetic locus was induced upon cholesterol supplementation (p=0.0022) in the native
125 host **(Fig. 1*e*)**. The biophysical characterization and *in-vitro* assay of 5βChR map the
126 molecular basis of cholesterol metabolism in *L. fermentum* NKN51.

127 **5βChR interacts with cholesterol via C-terminus and $NADP^+$ by N-terminus**

128 The 5βChR AlphaFold tertiary structure depicts conserved motifs of the SDR
129 superfamily, including the N-terminus dinucleotide binding Rossmann fold consisting
130 of seven central β-strands flanked by six α-helices and a catalytic tetrad (NSYK) **(Fig.**
131 **1*a-b*)**. Molecular docking results predicted that the cholesterol was bound in the C-
132 terminus pocket of the protein. Cholesterol steroid ring-1 interacts with methionine
133 (Met189) and phenylalanine (Phe190) residues of 5βChR through hydrophobic forces.
134 An electrostatic interaction was seen between the threonine (Thr187) and the hydroxyl
135 group of cholesterol. $NADP^+$ was linked to the N-terminus of 5βChR via the conserved
136 glycine-rich (Gly 8, 9, 12, and 14) motif. The tyrosine (Tyr152), lysine (Lys156), serine
137 (Ser11), and threonine (Thr10 and Thr187) amino acid residues contributed to a total
138 of eight polar interactions with $NADP^+$. Furthermore, 5βChR is directed towards the
139 nicotinamide oxygen (O=C-NH) by Tyr152 and Lys156 amino acid residues. The N-
140 terminus Ser11 amino acid residue interacts with the diphosphate group of $NADP^+$. On
141 the other hand, the ribose phosphate group of the $NADP^+$ coordinates with the Thr10
142 amino acid residue of 5βChR **(Fig. 1*c*)**. The superimposition of 5βChR and its
143 structural homologs (SDR reductase, sorbitol dehydrogenase, and putative
144 oxidoreductase RV2002) revealed a shared topological arrangement with root square

145 mean (r.m.s) deviations of 1.3, 1.3, and 1.2 Å, respectively and confirmed that 5βChR
146 belongs to SDRs **(Fig. S11)**.

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

148 Using site-directed mutagenesis, we elucidated the role of the catalytic tetrad (NSYK)
149 and other conserved residues of 5βChR. Since the C-terminus of SDRs was reportedly
150 involved in substrate binding,[6] we mutated the C-terminus of the protein to
151 investigate its role in cholesterol binding **(Fig. S12)**. TLC and LC-MS experiments of
152 5βChR mutants with $NADP^+$ and cholesterol revealed the importance of Ser139,
153 Tyr152, Thr187, Met189, and Phe190 amino acid residues in cholesterol conversion
154 to coprostanol **(Fig. S13, 3*a*)**. Furthermore, ITC data revealed that the S139A and
155 Y152A mutants lost $NADP^+$ binding but retained the cholesterol binding, with K_d values
156 of 24.8 μM and 4.44 μM, respectively. On the other hand, T187A, M189A, and F190A
157 mutants retained $NADP^+$ binding (K_d 8.77 μM, 7.49 μM, and 5.2 μM, respectively) **(Fig.**
158 **3*b*)** but lost the binding ability to cholesterol. Thus, the C-terminus of 5βChR is
159 important for substrate interaction, and the C-terminus of SDRs is indeed necessary
160 for substrate interaction[6]. Other mutations, such as N136A, K156A, H131A, Y170A,
161 E172A, and T187A, did not affect the ability to convert cholesterol to coprostanol, thus
162 indicating their nonsignificant role in enzyme activity **(Fig. 3*a*)**. Circular dichroism
163 spectral analysis of the wild-type and mutant 5βChR enzymes did not predict any
164 change in protein structure. Thus indicating that the changes in cofactor and substrate
165 binding were not due to perturbations in the secondary structure of the 5βChR
166 enzymes. **(Fig. S14)**. Since plant and microbial 5β reductases belong to the SDR
167 superfamily, we performed an in-depth analysis of the structural attributes of plant and
168 microbial 5β reductases. A systematic comparison of the $NADP^+$ binding pocket of
169 microbial 5βChR with the plant origin 5β-progesterone reductases (5β-POR) (PDB:

170 2V6G) was accomplished [11]. It revealed that the $NADP^+$ was bound to the Rossmann
171 fold in plant and bacterial 5β reductase. However, the lysine residue of a conserved
172 catalytic tetrad of SDRs does not interact with the O-2' or O-3' of nicotinamide ribose
173 in either case **(Fig. 4)** [9, 23].

174 **5βChRs is a prevalent enzyme class in microbes and shows its association with**
175 **human physiology**

176 Since 5βChRs are a novel class of enzymes discovered in bacteria, we focused on
177 assessing the diversity and abundance of these enzymes within Lactobacillaceae and
178 overall in eubacteria. A local database of 256 lactobacilli was screened to determine
179 the prevalence of 5βChRs in the genome. The maximum likelihood method confirmed
180 the presence of the 5βChR sequence in 84 lactobacilli genomes, which were further
181 divided into six major clades. Most of these representatives (*Lactiplantibacillus*
182 *plantarum, Limosilactobacillus mucosae, Levilactobacillus parabrevis*,
183 *Latilactobacillus sakei*, *Latilactobacillus curvatus*, *Ligilactobacillus salivarius,* and
184 *Levilactobacillus brevis*) were associated with a healthy commensal gut community
185 **(Fig. 5)** [24, 25]. We traced the evolutionary relationship of microbial 5βChRs with
186 plant 5β reductases using human homologs as an outgroup, which revealed that plant
187 and microbial 5β reductases originated from a common ancient ancestor **(Fig. S15)**.
188 To assess the distribution of 5βChRs in different microbial taxa, we retrieved 5000
189 5βChRs and interpreted the data using Cytoscape. 5βChR homologs were grouped
190 into 33, 138, and 269 clusters with threshold values for sequence identity (e-value) of
191 10^{-60}, 10^{-70}, and 10^{-80}, respectively **(Fig. S16, S17*a*)**. A threshold value of 10^{-80} was
192 chosen to map the 5βChR homologs to ensure similar functional and structural
193 aspects, which led to the grouping of proteins into 269 distinct clusters. The phylum-
194 level taxonomic classification of the respective clusters was considered to assess the

195 prevalence of microbial 5βChRs. Across the different clusters, 5βChRs were
196 predominantly present in Pseudomonadota, Actinomycetota, Planctomycetota,
197 Chloroflexota, Cyanobacteria, Acidobacteriota, Bacillota, and Candidatus
198 omnitrophota **(Fig. S17*b*, S18).** The SSN analysis indicated that inhabiting different
199 niches belonging to different phyla harbored 5βChR homologs, possibly providing
200 some benefit to their respective hosts. To understand the importance of the
201 aforementioned 5βChR clusters, we assessed the data through ShortBRED, which
202 consists of protein classes in 380 high-quality, healthy human metagenome
203 sequences (from the human microbiome project), covering samples from the different
204 regions of the body. Using the EFI-Computationally Guided Functional Profiling (EFI-
205 CGFP) tool, we identified 20 clusters of 5βChR homologs in the Human Microbiome
206 Project data **(Fig. S19)**. Clusters 3, 9, and 11 were abundant in the oral microbiome
207 (supragingival plaque, buccal cavity, and tongue dorsum). In contrast, clusters 2, 3, 6,
208 9, 11, 52, 83, 104, 119, 122, 128, 140, 156, 172, 185, and 24 were the dominant
209 clusters in the fecal samples, indicating the presence of 5βChR homologs in
210 anaerobes. Among the bacteria-dominated clusters, 16 and 139 had more significant
211 distributions in the posterior fornix (vaginal microbiome) and anterior nares,
212 respectively **(Fig. 7, S20)**. The comprehensive analysis of the 5βChR family in the
213 human gut inhabitants highlights this enzyme's prevalence and potential role in
214 maintaining cholesterol homeostasis.

215 **5βChR has an inverse association with diabetes**

216 We explore the correlation of 5βChR with diabetes because diabetes can induce
217 changes in lipid levels, leading to diabetic dyslipidemia, which exacerbates risk factors
218 for CVDs [26-28]. Analyzing the relationship between 5βChR and diabetes using
219 Chinese-type 2 diabetes shotgun data (PRJNA855128). The two independent

220 metagenomic cohorts included diabetic individuals (n=10) and healthy controls (n=10).
221 The interpretation of the data confirmed that 5βChRs exhibited a significantly greater
222 prevalence in healthy individuals than in those with diabetes (*p*=0.0213) **(Fig. 6*b*)**.
223 These findings corroborate the inverse relationship between 5βChR and diabetes in
224 this cohort study, indicating that 5βChR expression might reduce the risk of premature
225 coronary heart disease and stroke in patients with diabetic dyslipidemia [29].

226 **Discussion**

227 The gastrointestinal tract is a site for cholesterol metabolism and regulation of blood
228 cholesterol levels. Although cholesterol metabolism by lactobacilli has been known for
229 almost 50 years, the exact mechanism has remained elusive [16]. Recently, three
230 independent studies have reported the mechanism of cholesterol metabolism by
231 different bacterial gut residents. Still, the role of lactobacilli has not been addressed in
232 either of the reports[5, 13, 14]. Here, we discovered and characterized a novel
233 microbial 5β-reductase under the SDR superfamily, which catalyzes the direct
234 transformation of cholesterol to coprostanol **(Fig. 7)**. Molecular docking observations
235 of the cofactor and substrate with the wild-type and mutant 5βChRs identified the
236 critical residues for $NADP^+$ binding and revealed novel residues important for
237 cholesterol binding to the protein. The structural and mutational analysis of 5βChR
238 represents its similarity with microbial SDRs and plant 5β reductase, categorizing it
239 into a new class of SDRs. Furthermore, evolutionary analysis of plant and microbial
240 5βChR enzymes suggested that both are paraphyletic groups and confirmed the
241 presence of 5βChR in gut lactobacilli. Shotgun data analysis highlights the low levels
242 of 5βChR transcripts in diabetic patients. Metagenomic mining revealed that the
243 representatives of 5βChRs are present in *Clostridia, Eubacteria, Oscillopsia,*
244 *Peptostreptococcus, Fusobacteria Bifidobacteria, and Lactobacillus*. Our findings

245 suggest that 5βChR homologs are also acquired by uncharacterized bacterial classes,
246 which may be explored further. The prevalence of 5βChR in the microbiomes of
247 healthy individuals indicates the critical role this enzyme can play in designing
248 strategies to regulate lipid homeostasis. We believe this study can pave the way for
249 developing next-generation probiotics that can be explored to formulate functional
250 foods to ameliorate dyslipidemia.

251 **Materials and methods**

252 **In vitro cholesterol reduction assay**

253 *Limosilactobacillus fermentum* NKN51,[30] *Lacticaseibacillus casei* NKN344, and
254 *Lacticaseibacillus casei* NKN347 were fermented food isolates from our laboratory.
255 For the cholesterol estimation assay, the secondary culture was inoculated in sterile
256 anaerobic Man, Rogosa, and Sharpe (MRS) media (Difco Laboratories, USA)
257 supplemented with cholesterol (Sigma Aldrich, USA) (100 µg/ml) and sodium
258 thioglycolate (STG) (0.1%) (Tokyo Chemical Industry, China) and incubated at 37°C
259 for 24 h. After incubation, the samples were centrifuged at 13,000 rpm for 5 min. The
260 supernatant was subjected to a cholesterol estimation assay using the Rudel and
261 Morris methods [31]. This procedure was performed in triplicate, and the average
262 values were plotted. For whole-cell lysate assay, a 12 h *L. fermentum* NKN51 culture
263 was pelleted by centrifugation at 4500 rpm for 20 min at 4°C, resuspended in
264 phosphate-buffered saline (PBS) (1x) (Sigma Aldrich, USA), and lysed with a Branson
265 probe sonicator 102C (CE) (Thermo Fisher Scientific, USA) and centrifuged at 13,000
266 rpm for 30 min at 4°C. The 500 µL supernatant (whole cell lysate) with cholesterol (100
267 µM) and $NADP^+$ (100 µM) (Sigma Aldrich, USA) was incubated at 37°C for 12 h.
268 Further, the sterols were extracted by adding double the volume of hexane (Merck,

269 USA), followed by vigorous mixing. The microcentrifuge tube was left for phase
270 separation for an hour at room temperature. After the phase separation, the upper
271 organic layer was transferred into a fresh microcentrifuge tube and evaporated by a
272 refrigerated centriVape concentrator (Labconco, USA). The sterols were resuspended
273 in 10μl of methanol (Merck, USA) and separated by TLC (explained later). The reaction
274 products were also checked via LC-MS (details are provided in the LC-MS instrument
275 section).

276 **Cloning of the putative cholesterol dehydrogenase gene**

277 The genomic DNA of *L. fermentum* NKN51 was isolated using the HiPurA® Bacterial
278 Genomic DNA Purification Kit (Himedia, India). The 5βChR gene was amplified using
279 Phusion High Fidelity DNA polymerase (Thermo Fisher Scientific, USA) with the
280 primers listed in **Table S1**. Furthermore, the PCR product and pET28(a) were double
281 digested with NdeI (Thermo Fisher Scientific, USA) and XhoI (Thermo Fisher
282 Scientific, USA). The digested products were ligated using T4 DNA ligase (Thermo
283 Fisher Scientific, USA) and transformed into chemically competent *E. coli* cells. The
284 transformants were checked via double digestion of the plasmid, followed by
285 sequencing.

286 **Site-directed mutagenesis of the putative cholesterol dehydrogenase gene**

287 WT5βChR carrying plasmid was used to construct its mutants with the Q5® Site-
288 Directed Mutagenesis Kit (New England Biolabs, USA). Primers containing the desired
289 mutation were designed with the NEBase Changer webpage tool
290 (https://nebasechanger.neb.com/). The primers used are listed in **Table S1**, and the
291 transformants were confirmed by sequencing.

292 **Overexpression and purification of the N-His terminal tagged WT and mutant**
293 **5βChR proteins**

294 Confirmed clones containing WT and mutant 5βChR were transformed into
295 ultracompetent BL21AI (arabinose induction) (Thermo Fisher Scientific, USA) cells for
296 protein overexpression. 5βChR protein (29.5 kDa) overexpression was optimized with
297 IPTG (0.05 mM) (Tokyo Chemical Industry, China) and arabinose (0.2%) (Himedia,
298 India). For protein purification, the cells were grown in Luria–Bertani broth (Difco
299 Laboratories, USA) supplemented with kanamycin (50 µg/ml) (Sigma Aldrich, USA)
300 overnight at 37°C with shaking at 180 rpm. The secondary culture was inoculated with
301 a primary culture of WT or the 5βChR mutant (1%) and incubated at 37°C with shaking
302 at 180 rpm until the optical density (OD_{600}) reached ~0.8. The bacterial cultures were
303 induced with IPTG (0.05 mM) and 0.2% arabinose (0.2%) and incubated at 18°C with
304 shaking at 180 rpm for 18 h. The cells were harvested at 4500 rpm for 20 min at 4°C
305 and resuspended in 20 mL of buffer A (50 mM Tris-Cl, 500 mM NaCl 5% glycerol) (pH
306 7.8). The resuspended cells were lysed with a Branson probe sonicator 102C (CE)
307 and then centrifuged at 4500 rpm for 30 min at 4°C. The clarified supernatant was
308 loaded into the buffer A saturated HisTrap column (AKTA Pure-50 L; GE Healthcare).
309 Purified protein fractions were eluted with Buffer B (50 mM Tris-Cl, 500 mM NaCl, 500
310 mM imidazole 5% glycerol) (pH 7.8) and analyzed via SDS–PAGE. Furthermore, the
311 protein was dialyzed in dialysis buffer (20 mM Tris-Cl, 300 mM NaCl 5% glycerol) with
312 a 10 kDa membrane (Pall Corporation Centrifugal Devices, USA) for 8h intervals. The
313 protein concentration was checked via SDS–PAGE, and the proteins were stored in
314 aliquots at -80°C. Purified protein fractions were subjected to further assays, which
315 were discussed later.

316 **Thin-Layer Chromatography of Sterols**

We checked the enzyme activity by TLC in the presence of cofactor and substrate. The reaction mixture was prepared with an enzyme (WT and mutants) (10 nm), $NADP^+$ (100 μM), and substrate (100 μM) (cholesterol, cholestenone, and coprostanone) in dialysis buffer was incubated at 37°C for 12 h. These reaction mixtures were subjected to sterol extraction (as explained above) and analyzed via TLC. The resuspended sterol was spotted on a pretreated TLC plate (silica gel IB2-F flexible 20 × 20 cm) with methanol at 100°C for 30 min. The sterols were separated with a mobile phase (benzene: chloroform) (90:10) (Merck, USA). After being treated with 15% sulfuric acid in methanol, the plate was visualized and incubated at 110°C for 10 min.

LC-MS instrumentation conditions for the measurement of sterols

The enzyme-cofactor-substrate analysis was conducted on ultra-high performance liquid chromatography (UHPLC) coupled with tandem mass spectrometry. The Exion LC UHPLC was equipped with an online degasser, binary pump, computer-controlled sample cooler, and column heater. Mass spectrometric analysis was on QTRAP 6500+ (Sciex, MA). UHPLC and mass spectrometer parameters were controlled using Analyst (ver 1.7.1; Sciex, MA, USA). The chromatographic separation was achieved on the Discovery C8 column (150X4.6 mm, 5μm; Supelco, Sigma-Aldrich, USA). An isocratic mobile phase consisting of ultra-pure water (A) and acetonitrile (B) at a ratio of 10:90 (A: B) was used for the separation. The flow rate was maintained at 1 mL/min, and the column oven was kept at ambient. A 20 μL of sample was injected into the column. The mass spectrometric analysis was done by an atmospheric pressure chemical ionization (APCI) probe operated at positive ion mode. The multiple reaction monitoring (MRM) mode was used for the quantification of sterols with the following mass transitions: cholesterol (369.3 → 161.2), and coprostanol (371.3 → 95.1).

341 **qPCR of the 5βChR transcript**

342 The chloroform-phenol (VWR, USA) extraction method was used for total RNA
343 purification from triplicate *L. fermentum* NKN51 cell pellets grown with or without
344 cholesterol for 24 h. DNase-treated RNA samples were used to prepare cDNA using
345 the PrimeScript 1st strand cDNA Synthesis Kit (Takara, Japan). The 5βChR transcript
346 was quantified by real-time PCR with PowerTrack SYBR Green Master Mix (Applied
347 Biosystems, USA). The expression was normalized to the 16S rRNA gene expression.
348 The primers used in the analysis are listed in **Table S1**.

349 **Gel filtration chromatography of 5βChR**

350 5βChR gel filtration chromatography (column: Hiload^{TM} 16/600 Superdex^{TM} 200 pg)
351 was performed on an AKTA pure-50 L chromatography system (GE Healthcare) at
352 4°C. The concentrated, purified protein (10 mg/mL) was loaded onto an analytical
353 column equilibrated with Tris-Cl (20 mM) and NaCl buffer (200 mM) at pH 7.8 and
354 eluted at a flow rate of 1 ml/min. A protein standard consisting of thyroglobulin bovine,
355 γ-globulins, ovalbumin, ribonuclease A type I, and p-aminobenzoic acid (pABA)
356 (Merck, India) was used to determine the native molecular mass of 5βChR.

357 **Enzyme kinetic assay**

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

364 **Substrate-specificity assay**

365 Cholic acid, lithocholic acid, and hydrocortisone (100 μM) (Tokyo Chemical Industry,
366 China) substrate analogs of cholesterol were separately added to the reaction, and
367 cholesterol (100 μM) as a control. The reaction mixture was incubated at 37°C for 12
368 h after the sterols were extracted; after that, the samples were subjected to LC-MS to
369 determine the enzymatic activity of 5βChR. The flow injection analysis (FIA) was
370 performed using an electrospray ionization (ESA) source working in positive mode for
371 hydrocortisone and lithocholic acid while operating in negative mode for cholic acid. A
372 gradient mobile phase containing pure acetonitrile was pumped into the mass
373 spectrometer without a column. All injection volumes were kept at 20 μL.

374 **Circular dichroism spectroscopy of WT 5βChR and its mutants.**

375 CD spectroscopy of 5βChR (WT) or 5βChR-mutant strains was conducted using a J-
376 815 CD spectropolarimeter (Jasco, Tokyo, Japan). The protein samples (20 μM) were
377 prepared in KH_2PO_4 buffer (10 mM) (pH 7.8). A quartz cell 0.1 cm in length was used
378 for the CD scan, and scans were carried out at 25°C from 190 to 260 nm at a 50
379 nm/min speed with a wavelength pitch and three accumulation data points. The data
380 were analyzed using GraphPad Prism.

381 **Isothermal titration calorimetry assay for binding analysis**

382 ITC analysis was performed using a MicroCal PEAQ-ITC (Malvern Panalytical, United
383 Kingdom) with 300 cell volumes. The $NADP^+$ cofactor was dissolved in dialysis buffer,
384 and the substrate was dissolved in methanol. Purified 5βChR (20 μM) was injected
385 with 18 successive 2 μl aliquots of $NADP^+$ (1 mM) at 150 s intervals. Furthermore,
386 protein-$NADP^+$ was injected similarly to the substrate (500 μM). The data were fitted
387 to a nonlinear regression model using a single binding site with MicroCal Origin

388 software. The thermodynamic parameters were calculated using the Gibbs free energy
389 equation ($\Delta G = \Delta H - T\Delta S$ and $\Delta G = -RT\ln(K_a)$).

390 **5βChR structure determination**

391 5βChR structural modeling was performed using the ColabFold (v.1.5.5)
392 implementation of Alphafold (https://github.com/sokrypton/ColabFold) [32, 33].
393 Multiple sequence alignment was performed using the MMseq2 algorithm [34]. The
394 models were ranked through the pTM score, and the highest-scoring model was used
395 for analysis. PyMOL [35] (v2.5.7 Schrödinger, LLC) was used for protein visualization
396 and structural analysis of 5βChR. The DALI server [36] was used for structural
397 similarity searches.

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

399 The three-dimensional structure of 5βChR was predicted using ColabFold (v.1.5.5),
400 which uses MMseqs2 along with AlphaFold2 in its background [32, 33]. The potential
401 ligands of the 5βChR-$NADP^+$-cholesterol dataset were downloaded from the ZINC
402 drug database in SD format (https://zinc.docking.org/). AutoDockTools-1.5.7 was used
403 to convert the 5βChR PDB files and ligand structures into PDBQT format [37].
404 Autodock Vina (v.1.2.3) performed molecular docking with multiple ligands [38]. The
405 protein–ligand interactions were computed and visualized with PyMOL [35] (v2.5.7
406 Schrödinger, LLC).

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

408 The prevalence of 5βChR was examined by creating a protein-based custom-made
409 database for available lactobacilli from RefSeq using BLAST (v2.15.0) **(Table S2)**.
410 5βChR was used as a query search in the database, and the homologous proteins
411 were retrieved for evolutionary studies. To determine the ancestral relationship of

412 5βChR in plants, 100 related sequences **(Table S3)** were removed from UniProt
413 (https://www.uniprot.org/) [39]. MUSCLE (v3.8.31)[40] was used to align all the
414 respective sequences, and phylogenetic analysis was performed using RAxML (v8)
415 via the maximum likelihood Protcat blossom62 model [41].

416 **5βChR sequence similarity network data analysis**

417 For 5βChR sequence similarity network (SSN)[42] analysis, the amino acid sequence
418 of 5βChR was used as a query in the UniProt database via BLASTp with a series of
419 cutoff e-values of 10^{-60}, 10^{-70}, and 10^{-8} and protein identity values >50%, >60% and
420 >70%, respectively. The EFI-EST tool was used to construct the protein SSNs of these
421 sequences with various e-values, which were visualized by Cytoscape 3.3 [43].

422 **Determination of 5βChR abundance using healthy human metagenomic data**

423 The abundance of 5βChR was checked in HMP[44, 45] as previously described using
424 the ShortBRED **(Table S4)** [46]-[47]. This was performed using the EFI-chemically
425 guided functional profiling (EFI-CGFP) platform [48-50]. Using UniProt, 5βChR
426 sequence representatives were filtered at 80% amino acid similarity thresholds to
427 identify nonredundant peptide markers. ShortBRED-quantify was used with default
428 parameters to quantify the marker abundances in the 380 metagenomic database.
429 The relative abundance was calculated based on reads per kilobase million (RPKM)
430 values and converted into 'copies per microbial genome', as explained earlier [50, 51].

431 **Abundance analysis of the 5βChR gene in human gut metagenomes under**
432 **disease conditions**

433 The shotgun metagenomic sequencing datasets of the human gut microbiomes
434 describing diabetes were downloaded from the Sequence Read Archive of the NCBI
435 (https://www.ncbi.nlm.nih.gov/sra). The dataset was divided into healthy and diseased

436 based on the metadata provided with the original data. The read quality was checked
437 using FastQC (v0.11.9) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/),
438 followed by MultiQC (v1.11) [52]. Only reads with a phread score higher than 30 were
439 included in the study. The high-quality sequencing reads were assembled using
440 MEGAHIT (v1.2.9) with a minimum contig length of 500 bp [53]. The assembled
441 sequences were matched to the nucleotide sequence of 5βChR using DIAMOND
442 (v2.0.15.1531) with the following parameters: e value – 0.00001, query coverage 30,
443 and percent identity 30 [54].

444 **Authors contribution**

445 U.N. designed the study, curated the data, developed the methodology,
446 conceptualized the project, and wrote the original draft. V.B. methodology and writing.
447 A.G. Methodology, Data curation, Visualization, and writing. RS methodology. A.G.
448 methodology. V.S.B. methodology and data curation. KA. Formal analysis. HPS. LC-
449 MS methodology and data analysis. SM. LC-MS methodology and data analysis. SL.
450 Methodology. MS. Methodology. K.T. Formal analysis and writing. NKN Formal
451 analysis, conceptualization, writing, resources, and project administration.

452 **Acknowledgments**

453 This work was supported by the National Priority Project (NPP) under the Corpus Fund
454 Head (CFH) of the Indian Council of Agricultural Research. Urmila Netter thanks the
455 Council of Scientific and Industrial Research (CSIR), India, for providing a research
456 fellowship. For financial support, Vishakha Bisht and Amit Gaurav thank the Ministry
457 of Human Resource and Development and the Department of Biotechnology (DBT)
458 Govt. of India.

459 **Supporting information**

460 Additional experimental data are provided in the supplementary information section.

461 **References**

462 1. Puglielli, L., Tanzi, R. E. & Kovacs, D. M. (2003) Alzheimer's disease: the
463 cholesterol connection, *Nat Neurosci.* **6**, 345-51.

464 2. Berger, S., Raman, G., Vishwanathan, R., Jacques, P. F. & Johnson, E. J. (2015)
465 Dietary cholesterol and cardiovascular disease: a systematic review and meta-
466 analysis, *Am J Clin Nutr.* **102**, 276-94.

467 3. Song, Y., Liu, J., Zhao, K., Gao, L. & Zhao, J. (2021) Cholesterol-induced toxicity:
468 An integrated view of the role of cholesterol in multiple diseases, *Cell Metab.* **33**, 1911-
469 1925.

470 4. Rosenfeld, R. S., Fukushima, D. K., Hellman, L. & Gallagher, T. F. (1954) The
471 transformation of cholesterol to coprostanol, *J Biol Chem.* **211**, 301-11.

472 5. Kenny, D. J., Plichta, D. R., Shungin, D., Koppel, N., Hall, A. B., Fu, B., Vasan, R.
473 S., Shaw, S. Y., Vlamakis, H., Balskus, E. P. & Xavier, R. J. (2020) Cholesterol
474 Metabolism by Uncultured Human Gut Bacteria Influences Host Cholesterol Level,
475 *Cell Host Microbe.* **28**, 245-257.e6.

476 6. Kallberg, Y., Oppermann, U., Jörnvall, H. & Persson, B. (2002) Short-chain
477 dehydrogenases/reductases (SDRs), *Eur J Biochem.* **269**, 4409-17.

478 7. Penning, T. M. (1997) Molecular endocrinology of hydroxysteroid dehydrogenases,
479 *Endocr Rev.* **18**, 281-305.

480 8. Ridlon, J. M., Ikegawa, S., Alves, J. M., Zhou, B., Kobayashi, A., Iida, T., Mitamura,
481 K., Tanabe, G., Serrano, M., De Guzman, A., Cooper, P., Buck, G. A. & Hylemon, P. B.
482 (2013) Clostridium scindens: a human gut microbe with a high potential to convert
483 glucocorticoids into androgens, *J Lipid Res.* **54**, 2437-49.

484 9. Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E.,
485 Ladenstein, R., Jörnvall, H. & Oppermann, U. (2002) Critical residues for structure and
486 catalysis in short-chain dehydrogenases/reductases, *J Biol Chem.* **277**, 25677-84.
487 10. Rossmann, M. G., Ford, G. C., Watson, H. C. & Banaszak, L. J. (1972) Molecular
488 symmetry of glyceraldehyde-3-phosphate dehydrogenase, *J Mol Biol.* **64**, 237-45.
489 11. Thorn, A., Egerer-Sieber, C., Jäger, C. M., Herl, V., Müller-Uri, F., Kreis, W. &
490 Muller, Y. A. (2008) The crystal structure of progesterone 5beta-reductase from
491 Digitalis lanata defines a novel class of short chain dehydrogenases/reductases, *J Biol*
492 *Chem.* **283**, 17260-9.
493 12. Chen, M. & Penning, T. M. (2014) 5β-Reduced steroids and human Δ(4)-3-
494 ketosteroid 5β-reductase (AKR1D1), *Steroids.* **83**, 17-26.
495 13. Le, H. H., Lee, M. T., Besler, K. R., Comrie, J. M. C. & Johnson, E. L. (2022)
496 Characterization of interactions of dietary cholesterol with the murine and human gut
497 microbiome, *Nat Microbiol.* **7**, 1390-1403.
498 14. Li, C., Stražar, M., Mohamed, A. M. T., Pacheco, J. A., Walker, R. L., Lebar, T.,
499 Zhao, S., Lockart, J., Dame, A., Thurimella, K., Jeanfavre, S., Brown, E. M., Ang, Q.
500 Y., Berdy, B., Sergio, D., Invernizzi, R., Tinoco, A., Pishchany, G., Vasan, R. S.,
501 Balskus, E., Huttenhower, C., Vlamakis, H., Clish, C., Shaw, S. Y., Plichta, D. R. &
502 Xavier, R. J. (2024) Gut microbiome and metabolome profiling in Framingham heart
503 study reveals cholesterol-metabolizing bacteria, *Cell.* **187**, 1834-1852.e19.
504 15. Björkhem, I. & Gustafsson, J. A. (1971) Mechanism of microbial transformation of
505 cholesterol into coprostanol, *Eur J Biochem.* **21**, 428-32.
506 16. Jia, B., Zou, Y., Han, X., Bae, J. W. & Jeon, C. O. (2023) Gut microbiome-mediated
507 mechanisms for reducing cholesterol levels: implications for ameliorating
508 cardiovascular disease, *Trends Microbiol.* **31**, 76-91.

509 17. Harrison, V. C. & Peat, G. (1975) Serum cholesterol and bowel flora in the
510 newborn, *Am J Clin Nutr.* **28**, 1351-5.
511 18. Marshall, V. M. & Taylor, E. (1995) Ability of neonatal human Lactobacillus isolates
512 to remove cholesterol from liquid media, *International journal of food science &*
513 *technology.* **30**, 571-577.
514 19. Lye, H. S., Rusul, G. & Liong, M. T. (2010) Removal of cholesterol by lactobacilli
515 via incorporation and conversion to coprostanol, *J Dairy Sci.* **93**, 1383-92.
516 20. Fuentes, M. C., Lajo, T., Carrión, J. M. & Cuñé, J. (2013) Cholesterol-lowering
517 efficacy of Lactobacillus plantarum CECT 7527, 7528 and 7529 in
518 hypercholesterolaemic adults, *Br J Nutr.* **109**, 1866-72.
519 21. Costabile, A., Buttarazzi, I., Kolida, S., Quercia, S., Baldini, J., Swann, J. R.,
520 Brigidi, P. & Gibson, G. R. (2017) An in vivo assessment of the cholesterol-lowering
521 efficacy of Lactobacillus plantarum ECGC 13110402 in normal to mildly
522 hypercholesterolaemic adults, *PLoS One.* **12**, e0187964.
523 22. Tjandrawinata, R. R., Kartawijaya, M. & Hartanti, A. W. (2022) In vitro Evaluation
524 of the Anti-hypercholesterolemic Effect of Lactobacillus Isolates From Various
525 Sources, *Front Microbiol.* **13**, 825251.
526 23. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J. &
527 Ghosh, D. (1995) Short-chain dehydrogenases/reductases (SDR), *Biochemistry.* **34**,
528 6003-13.
529 24. Walter, J. (2008) Ecological role of lactobacilli in the gastrointestinal tract:
530 implications for fundamental and biomedical research, *Appl Environ Microbiol.* **74**,
531 4985-96.

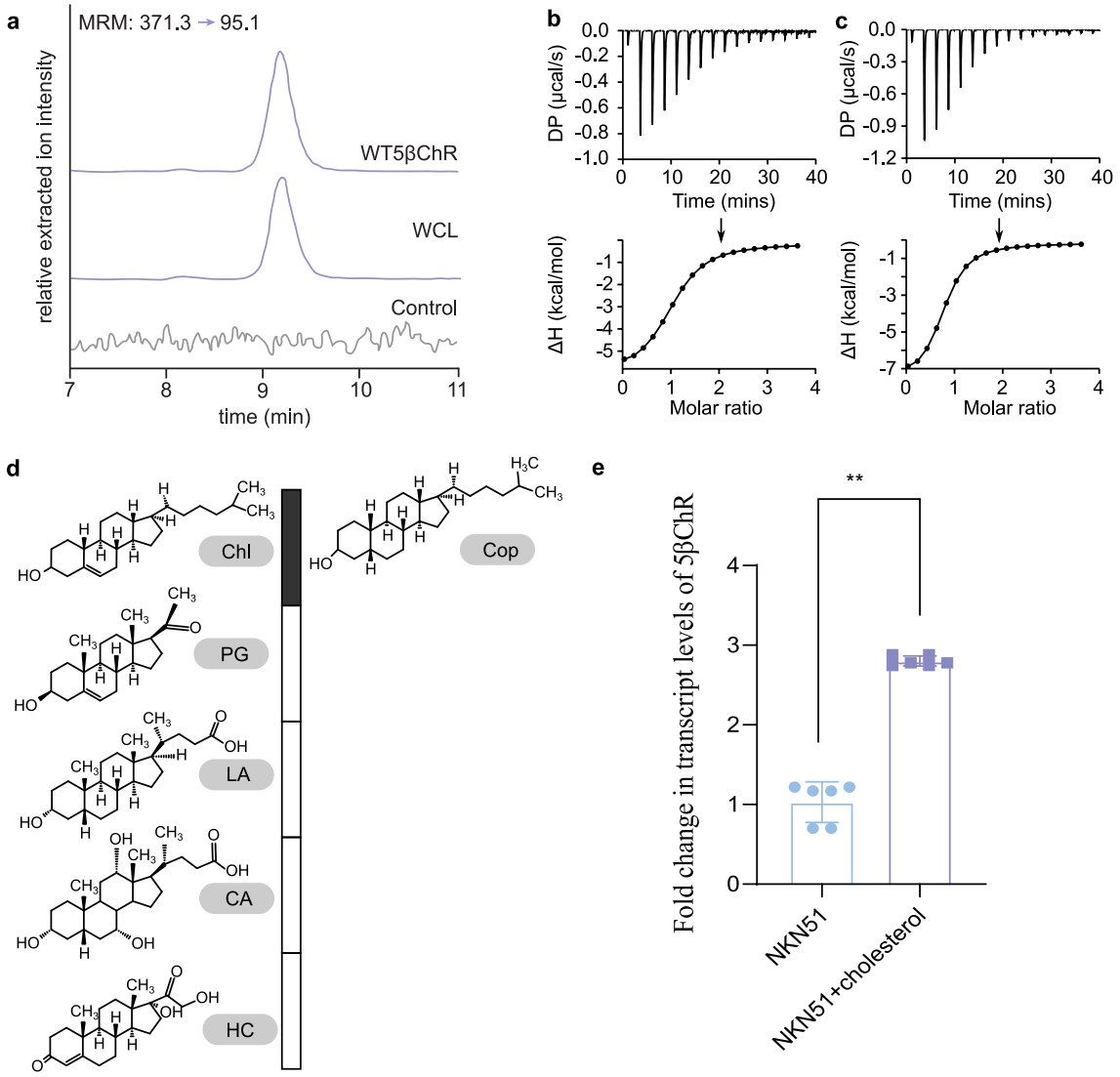
532 25. Dang, D., Liu, X., Chen, H., Zhao, J., Chen, W. & Yang, B. (2024) In silico, in vitro
533 and in vivo safety assessment of three Limosilactobacillus mucosae strains for
534 probiotic candidates, *Food Bioscience.* **57**, 103462.
535 26. Mihaylova, B., Emberson, J., Blackwell, L., Keech, A., Simes, J., Barnes, E. H.,
536 Voysey, M., Gray, A., Collins, R. & Baigent, C. (2012) The effects of lowering LDL
537 cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis
538 of individual data from 27 randomised trials, *Lancet.* **380**, 581-90.
539 27. Bitzur, R., Cohen, H., Kamari, Y., Shaish, A. & Harats, D. (2009) Triglycerides and
540 HDL cholesterol: stars or second leads in diabetes?, *Diabetes Care.* **32 Suppl 2**,
541 S373-7.
542 28. Fernandez, M. L. & Andersen, C. J. (2014) Effects of dietary cholesterol in
543 diabetes and cardiovascular disease, *Clinical Lipidology.* **9**, 607-616.
544 29. Association, A. H. (2024) Cholesterol and Diabetes in, American Heart
545 Association,
546 30. Beniwal, A., Gaurav, A., Sharma, R., Ghosh, T., Kumar, P., Netter, U., Pathania,
547 R. & Navani, N. K. (2021) Draft Genome Sequence of Limosilactobacillus fermentum
548 Strain NKN-51, Isolated from Fermented Yak Milk in the Western Himalayas of India,
549 *Microbiol Resour Announc.* **10**.
550 31. Rudel, L. L. & Morris, M. D. (1973) Determination of cholesterol using o-
551 phthalaldehyde, *J Lipid Res.* **14**, 364-6.
552 32. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
553 Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C.,
554 Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R.,
555 Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M.,
556 Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W.,

557 Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021) Highly accurate protein structure
558 prediction with AlphaFold, *Nature.* **596**, 583-589.
559 33. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M.
560 (2022) ColabFold: making protein folding accessible to all, *Nat Methods.* **19**, 679-682.
561 34. Steinegger, M. & Söding, J. (2017) MMseqs2 enables sensitive protein sequence
562 searching for the analysis of massive data sets, *Nat Biotechnol.* **35**, 1026-1028.
563 35. DeLano, W. L. (2002) Pymol: An open-source molecular graphics tool, *CCP4*
564 *Newsl Protein Crystallogr.* **40**, 82-92.
565 36. Holm, L. & Rosenström, P. (2010) Dali server: conservation mapping in 3D,
566 *Nucleic Acids Res.* **38**, W545-9.
567 37. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D.
568 S. & Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with
569 selective receptor flexibility, *J Comput Chem.* **30**, 2785-2791.
570 38. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. (2021) AutoDock Vina
571 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings, *J Chem*
572 *Inf Model.* **61**, 3891-3898.
573 39. (2015) UniProt: a hub for protein information, *Nucleic Acids Res.* **43**, D204-12.
574 40. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy
575 and high throughput, *Nucleic Acids Res.* **32**, 1792-7.
576 41. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-
577 analysis of large phylogenies, *Bioinformatics.* **30**, 1312-3.
578 42. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R.
579 & Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST):
580 A web tool for generating protein sequence similarity networks, *Biochim Biophys Acta.*
581 **1854**, 1019-37.

582 43. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin,
583 N., Schwikowski, B. & Ideker, T. (2003) Cytoscape: a software environment for
584 integrated models of biomolecular interaction networks, *Genome Res.* **13**, 2498-504.
585 44. (2012) Structure, function and diversity of the healthy human microbiome, *Nature.*
586 **486**, 207-14.
587 45. (2012) A framework for human microbiome research, *Nature.* **486**, 215-21.
588 46. Kasmanas, J. C., Bartholomäus, A., Corrêa, F. B., Tal, T., Jehmlich, N., Herberth,
589 G., von Bergen, M., Stadler, P. F., Carvalho, A. & Nunes da Rocha, U. (2021)
590 HumanMetagenomeDB: a public repository of curated and standardized metadata for
591 human metagenomes, *Nucleic Acids Res.* **49**, D743-d750.
592 47. Kaminski, J., Gibson, M. K., Franzosa, E. A., Segata, N., Dantas, G. &
593 Huttenhower, C. (2015) High-Specificity Targeted Functional Profiling in Microbial
594 Communities with ShortBRED, *PLoS Comput Biol.* **11**, e1004557.
595 48. Zallot, R., Oberg, N. & Gerlt, J. A. (2019) The EFI Web Resource for Genomic
596 Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to
597 Discover Novel Enzymes and Metabolic Pathways, *Biochemistry.* **58**, 4169-4182.
598 49. Oberg, N., Zallot, R. & Gerlt, J. A. (2023) EFI-EST, EFI-GNT, and EFI-CGFP:
599 Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools, *J Mol*
600 *Biol.* **435**, 168018.
601 50. Levin, B. J., Huang, Y. Y., Peck, S. C., Wei, Y., Martínez-Del Campo, A., Marks, J.
602 A., Franzosa, E. A., Huttenhower, C. & Balskus, E. P. (2017) A prominent glycyl radical
603 enzyme in human gut microbiomes metabolizes trans-4-hydroxy-l-proline, *Science.*
604 **355**.

605 51. Jia, B., Park, D., Hahn, Y. & Jeon, C. O. (2020) Metagenomic analysis of the
606 human microbiome reveals the association between the abundance of gut bile salt
607 hydrolases and host health, *Gut Microbes.* **11**, 1300-1313.
608 52. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. (2016) MultiQC: summarize
609 analysis results for multiple tools and samples in a single report, *Bioinformatics.* **32**,
610 3047-8.
611 53. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. (2015) MEGAHIT: an ultra-
612 fast single-node solution for large and complex metagenomics assembly via succinct
613 de Bruijn graph, *Bioinformatics.* **31**, 1674-1676.
614 54. Buchfink, B., Xie, C. & Huson, D. H. (2015) Fast and sensitive protein alignment
615 using DIAMOND, *Nat Methods.* **12**, 59-60.

616 **Figures**



617

618 **Fig.1**

619 *L. fermentum* NKN51 harbors a cholesterol dehydrogenase that can convert
620 cholesterol to coprostanol. **(a).** Extracted ion chromatograph shows the reaction
621 product of *L. fermentum* NKN51 whole-cell lysate (WCL) and purified cholesterol-5β-
622 reductase (5βChR) incubated with $NADP^+$ (100 μM) and cholesterol (100 μM) for 12
623 h. **(b).** Isothermal Titration Calorimetry (ITC) data displaying binding affinity of the
624 5βChR (20 μM) for $NADP^+$ (1 mM). **(c).** The titration curve provides the interaction
625 between the 5βChR + $NADP^+$ complex and cholesterol (500 μM), where DP stands

626 for differential power, and ΔH represents the enthalpies (normalized with the control).
627 **(d).** The bar graph indicates the conversion of cholesterol to coprostanol (grey block)
628 by 5βChR under the same conditions. The other structurally similar compounds are
629 not metabolized into their respective metabolites. Abbreviations: PG stands for
630 pregnenolone, LA-lithocholic acid, CA-cholic acid, and HC-hydrocortisone. Spectra
631 analysis of this substrate specificity assay is given in Supplementary Figure 8. **(e).** The
632 transcript levels of 5βChR in a cholesterol-rich medium compared to a medium without
633 cholesterol at 24 h under anaerobic conditions show ~1.5-fold change. In the graph,
634 each point represents a biological replicate, with the centre bar representing the mean
635 and the error bars representing the standard deviation. ** denotes the *p*-value- 0.0022
636 calculated by the Mann-Whitney test.

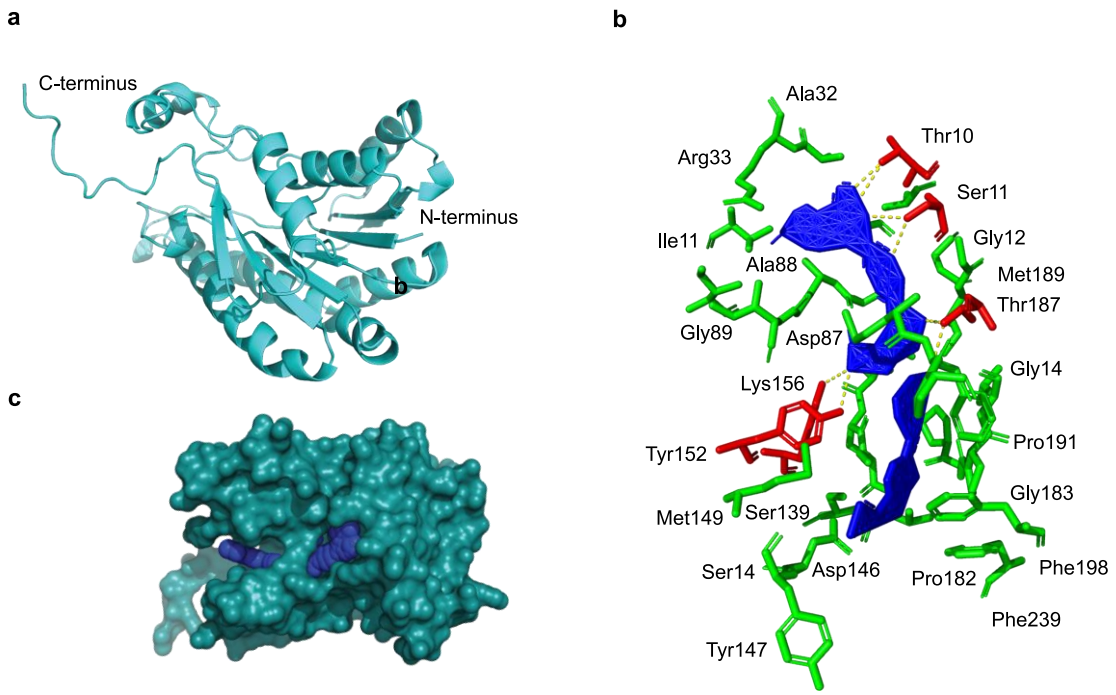


Fig. 2

Structure and mutation analysis of 5βChR for ligand binding and comparison study with the plant 5β reductase. **(a).** Alphafold2 predicted a top-ranked 5βChR structure with high confidence (>90%), a high preresidue confidence score (>90%), and a low domain position confidence, and visualized by PyMOL. **(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 the interacting amino acids. The Cofactor ($NADP^+$) binding pocket highlights the ionic interaction (red) (indicated by dashed lines). The amino acids involved in hydrogen bonding are Thr10, Ser11, and Tyr152, while Lys156, Ala32, Arg33, Gly89, Ile89, Ser139, Asp146, Met149, Ser11, Ile90, and Ala9 showcase involvement in van der Waals bonds (green). Protein-substrate (5βChR-cholesterol) complex shows hydrophilic interaction with Thr187, whereas the hydrophobic forces are present between Pro182, Pro184, Met189, Phe190, Phe198, and Phe238.

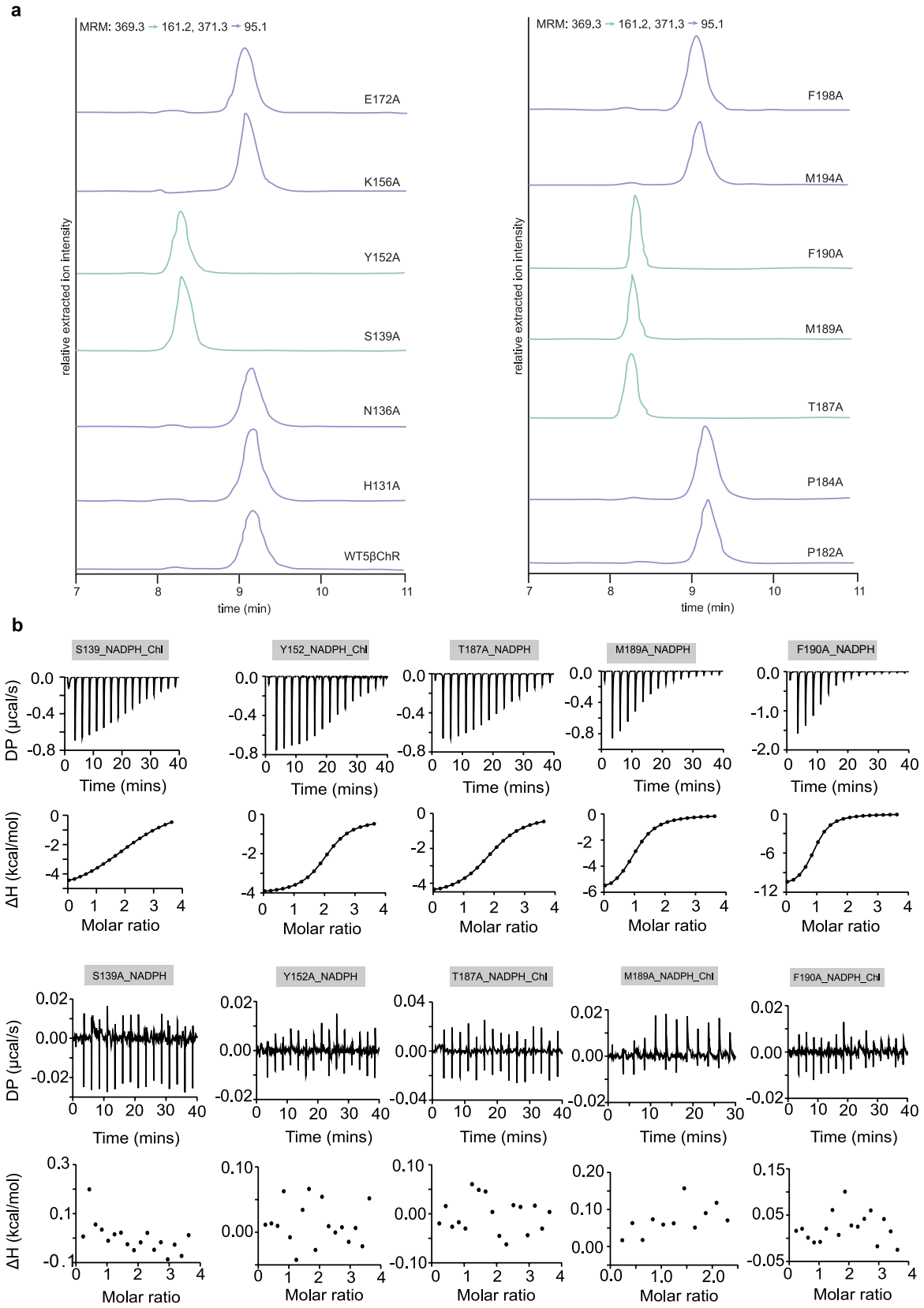
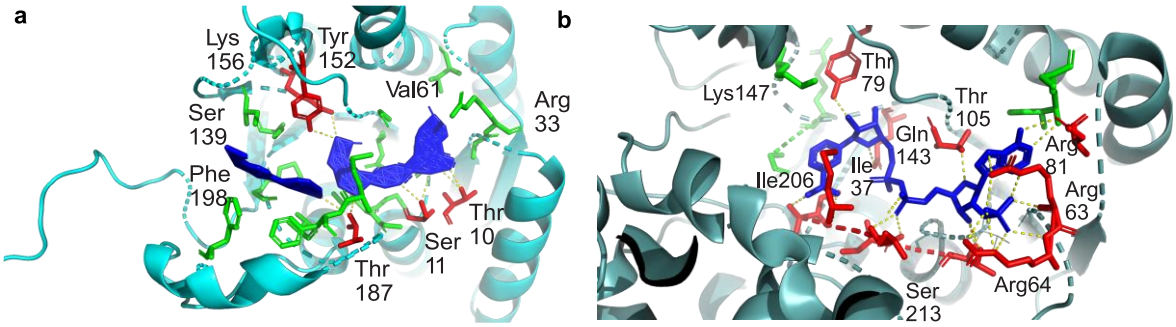

a
MRM: 369.3 → 161.2, 371.3 → 95.1
relative extracted ion intensity
E172A
K156A
Y152A
S139A
N136A
H131A
WT5βChR
time (min)
F198A
M194A
F190A
M189A
T187A
P184A
P182A
b
S139_NADPH_ChI
Y152_NADPH_ChI
T187A_NADPH
M189A_NADPH
F190A_NADPH
DP (µcal/s)
Time (mins)
ΔH (kcal/mol)
Molar ratio
S139A_NADPH
Y152A_NADPH
T187A_NADPH_ChI
M189A_NADPH_ChI
F190A_NADPH_ChI

Fig. 3

654 Characterization of 5βChR mutants. **(a).** The LC-MS data is represented by the
655 extracted ion chromatographs of 5βChR mutants supplemented with $NADP^+$ (100 μM)
656 and cholesterol (100 μM) under similar conditions. **(b).** The normalized enthalpy
657 curves represent the interaction profile of 5βChR mutants with their ligands. S139A
658 and Y152A mutants have an affinity for cholesterol and lost binding for $NADP^+$. T187A,
659 M189A, and F190A mutants do not bind to cholesterol. In contrast, they retained an
660 affinity for $NADP^+$.



661

662 **Fig. 4**

663 Comparison of the active sites of 5βChR and 5β-POR with their respective substrates.
664 The comparison study of the active site of 5βChR and plant 5β reductase (5β-POR)
665 with their respective substrates. The stereo view of 5βChR complexed with $NADP^+$ is
666 mediated by tyrosine (Tyr152) and serine (Ser139) of the conserved tetrad. In 5β-POR
667 - two arginine residues (Arg63 & Arg64) stabilise $NADP^+$ interaction. 5β-POR has two
668 Rossmann folds (glycine residues) (GXXGXXG), whereas 5βChR harbors a single N-
669 terminus fold, contributing to $NADP^+$ binding.

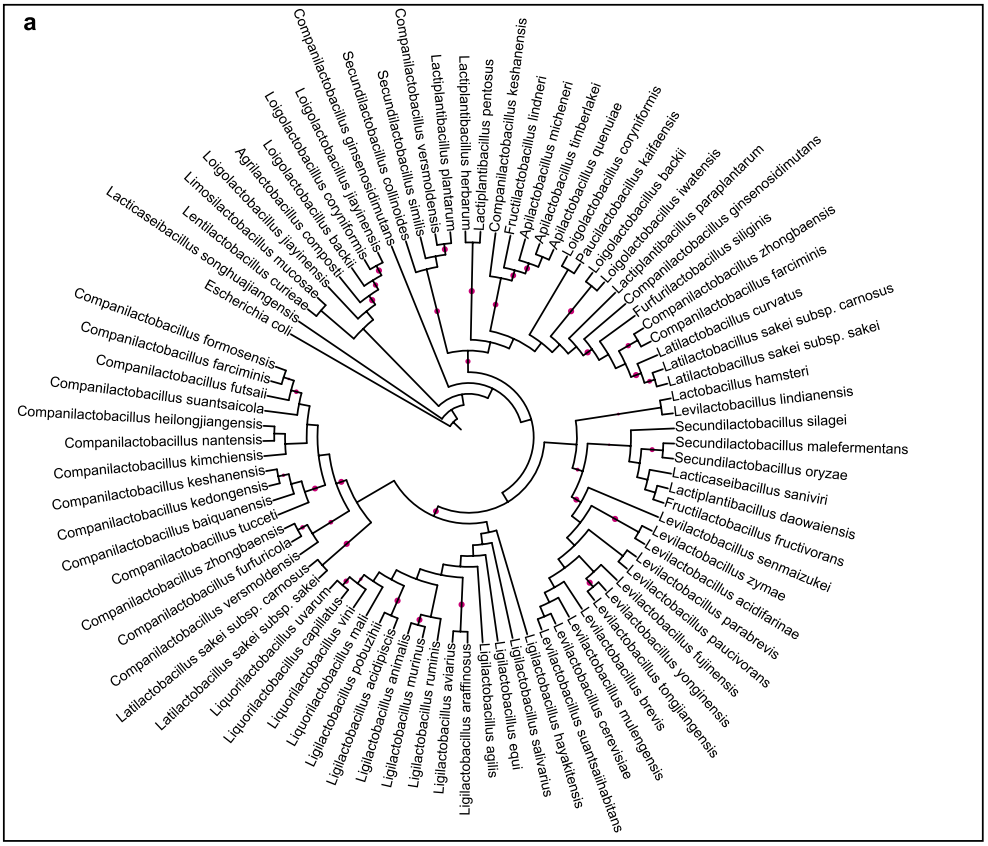
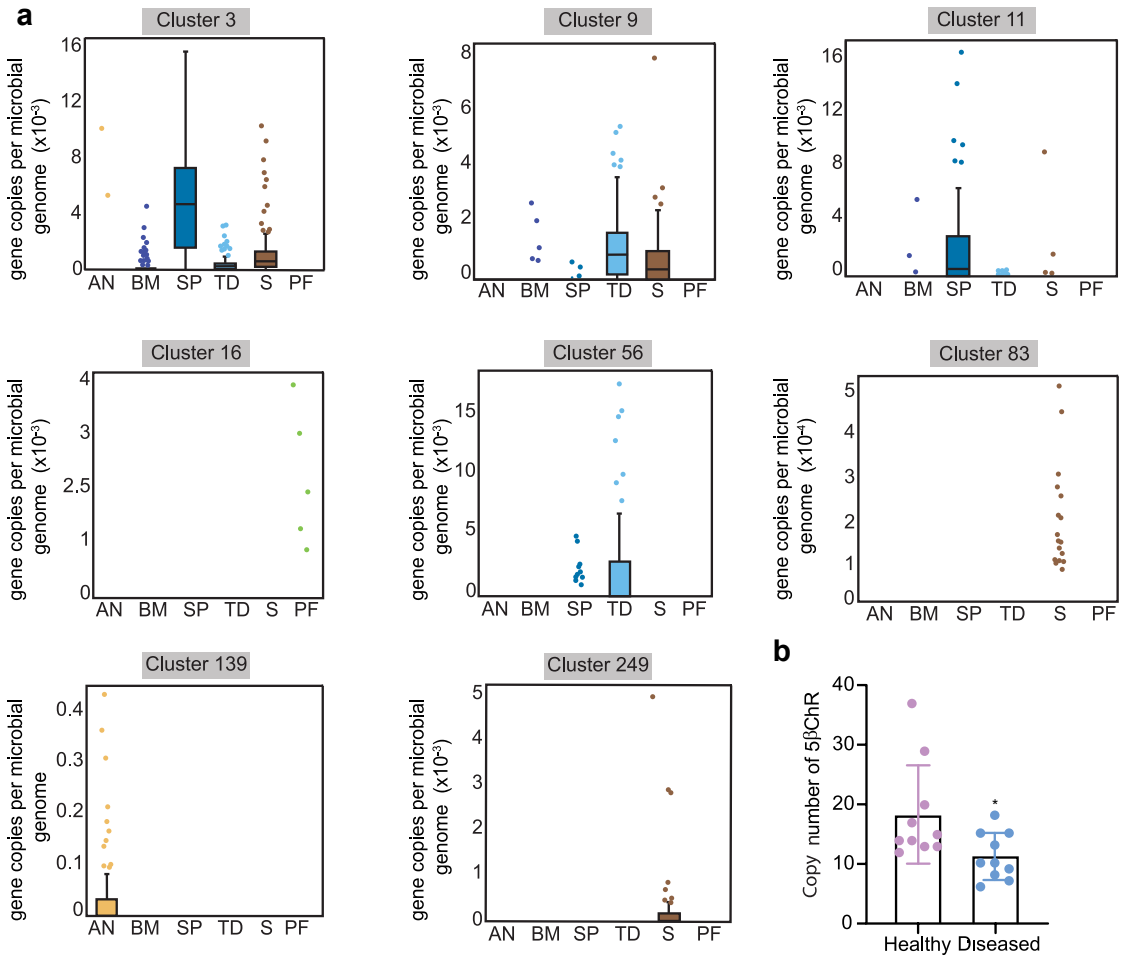


Fig. 5

Mapping the prevalence of 5βChR in lactobacilli in gut lactobacilli. The cladogram of 5βChR protein across 256 lactobacilli depicts its abundance in 84 strains. This study used the PROTCATWAG model in RAxML and rooted using the branch leading to *Lacticaseibacillus songhuajiangensis* and *Escherichia coli* as the outgroups.

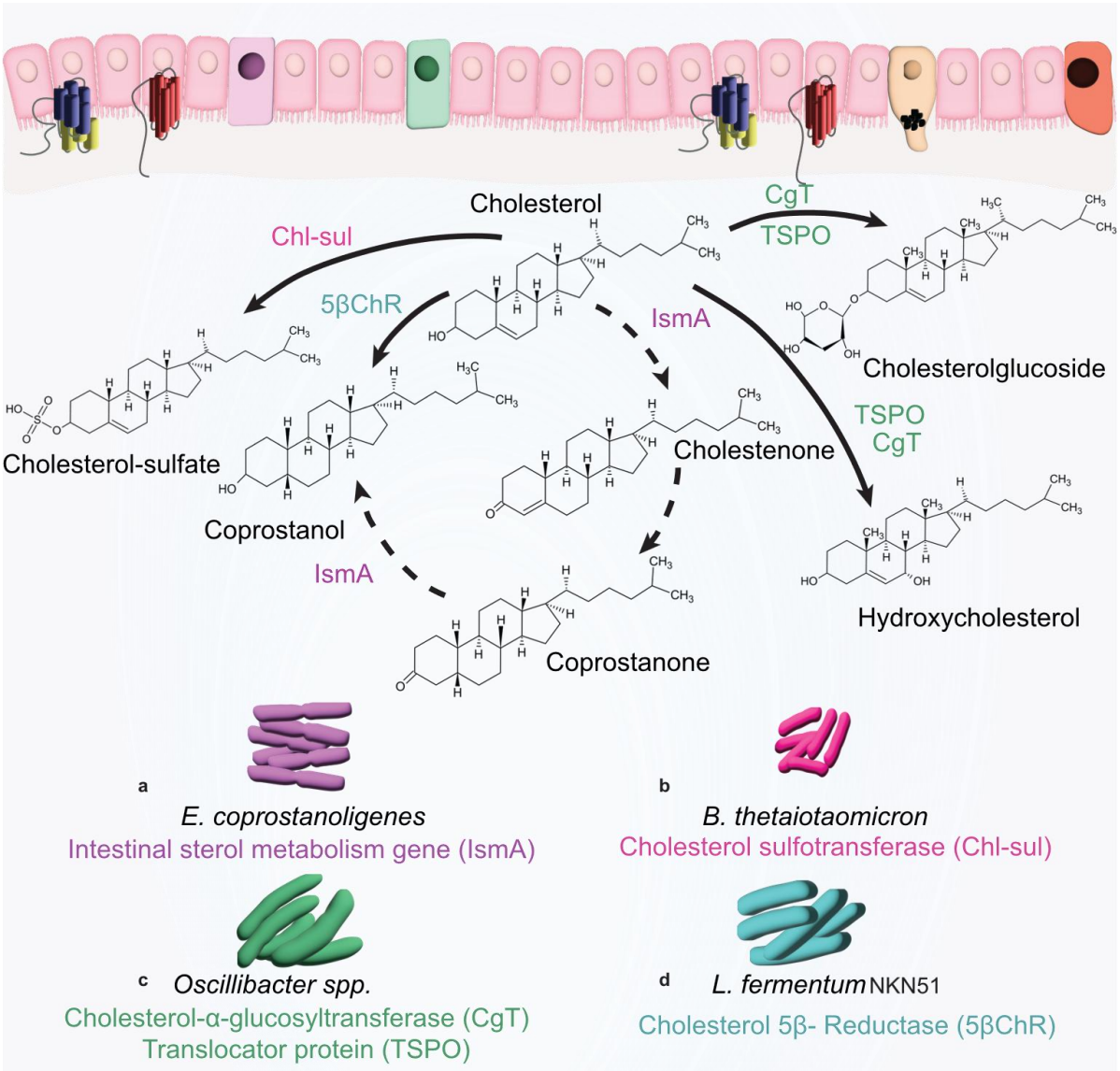


AN- Anterior nares, BM- Buccal mucosa, SP- Supragingival plaque, TD- Tongue dorsum, S- Stool, and PF- Posterior fornix

Fig. 6

Meta-analysis depicting the abundance of 5βChR homologs in healthy microbiomes and comparative analysis in a human diabetic cohort study. **(a).** The box-plot graphs highlight the presence of 5βChR homologs in various body sites. Clusters 3, 9, 11, and 56 dominate the oral microbiome (supragingival plaque, buccal cavity, and tongue dorsum) and stool samples. Cluster 16 is predominant in vaginal samples. The enzymes from other clusters 139 showed higher distribution in anterior nares samples of human microbiomes. Further, clusters 83 and 249 are abundant in stool samples, indicating their abundance in the gut. **(b).** 5βChR has a lower abundance among diabetic patients than healthy individuals in the shotgun metagenomic data of a

687 Chinese type 2 diabetes cohort study. Individuals with 5βChR high copy number are
688 predominantly reported as healthy. The points in the graph represent the number of
689 individuals in each category. *p value- 0.0213 is determined by the Mann-Whitney test.



690

691 **Fig. 7**

692 A comprehensive representation of microbial cholesterol transformation in the human
693 gut. Various gut microbes transform cholesterol in the gut via direct and indirect
694 pathways. **(a).** A cholesterol dehydrogenase (ismA) of *Eubacterium coprostanoligenes*
695 is reported to catalyze the first and last step of the indirect pathway of cholesterol

696 metabolism. **(b).** The gut cholesterol is also converted to cholesterol-sulphate with the
697 help of a cholesterol sulfotransferase encoded by *Bacteroidetes thetaiotaomicron.* **(c).**
698 A group of enzymes, cholesterol-α-glucosyltransferase, and translocator protein
699 expressed by *Oscillibacter* species transformed cholesterol into cholesterolglucoside
700 and hydroxycholesterol while **(d).** A comprehensive study demonstrates that 5βChR
701 expressed by *Limosilactobacillus fermentum* NKN51 can directly convert cholesterol
702 to coprostanol by a cholesterol-5β-reductase.

Supplementary information

Discovery and mechanistic characterization of a probiotic-origin 3β-OH-Δ^{5-6}-cholesterol-5β-reductase directly converting 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], Hanuman Prasad Sharma[c], Sujata Mohanty[d], Shubham Loat[e], Mihir Sarkar[e], Kapil Tahlan[f], 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]Centralized Core Research Facility, All India Institute of Medical Sciences Ansari Nagar, New Delhi – 110029, India

[d]Stem Cell Facility, (DBT-Centre of Excellence for Stem Cell Research), All India Institute of Medical Sciences, Ansari Nagar, New Delhi– 110029, India.

[e]ICAR-National Research Centre on Yak, Dirang, West Kameng district, Arunachal Pradesh-790101, India

[f]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

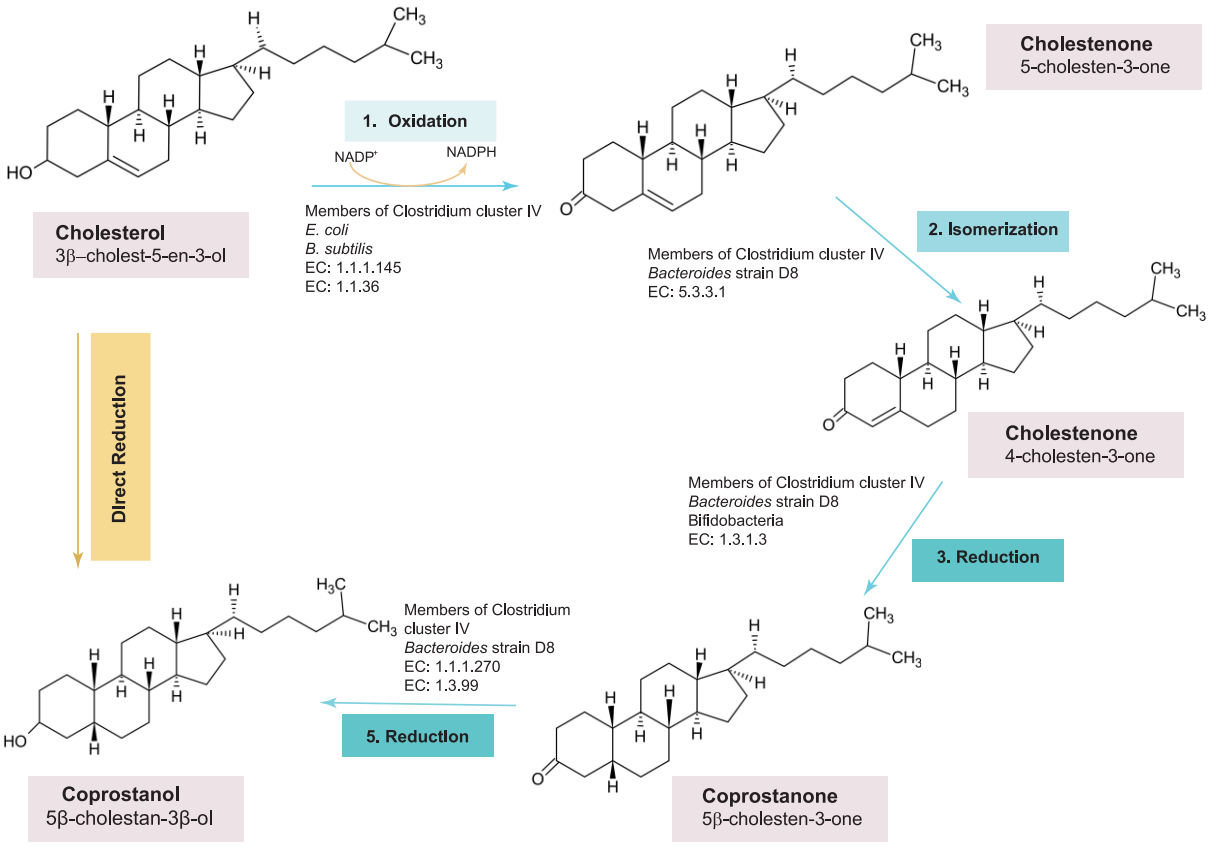
E-mail address: naveen.navani@bt.iitr.ac.in

24 **Table of Contents**

25 **1. Supplementary Figures……...1-14**

26 **2. Supplementary Tables……15-35**

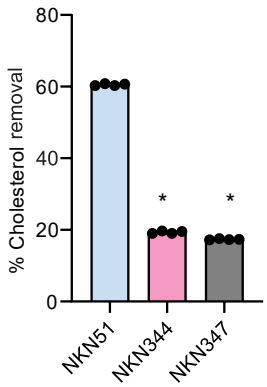
27 **1. Supplementary Figures**



28

29 **Fig. S1** Biotransformation pathways of cholesterol to coprostanol by gut microbes.

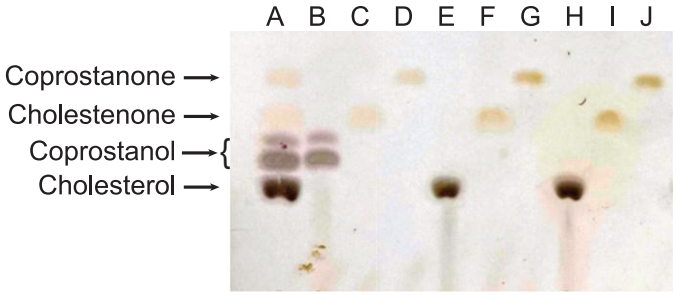
30



31

32 **Fig. S2**

33 Rudel & Morris cholesterol depletion assay representing in-vitro cholesterol removal by different
34 *Lactobacillus* strains**.** Each point represents the biological replicate and *p- value* calculated by the
35 Mann-Whitney test (**p*- 0.0286).



A- Standards of cholesterol, cholestenone, coprostanone, and coprostanol
B- NKN51+cholesterol, C-NKN51+cholestenone, D-NKN51+coprostanone
E- NKN344+cholesterol, F-NKN344+cholestenone, G- NKN344+coprostanone
H- NKN347+cholesterol, I-NKN347+cholestenone, J- NKN347+coprostanone

Fig. S3

Thin-layer chromatography (TLC) screening of cholesterol transformation to coprostanol by lactobacilli. The whole-cell lysate of different lactobacilli strains supplemented with $NADP^+$ (100 μM) and either of substrates, cholesterol (100 μM), cholestenone (100 μM), or coprostanone (100 μM) for 12 h in anaerobic conditions.

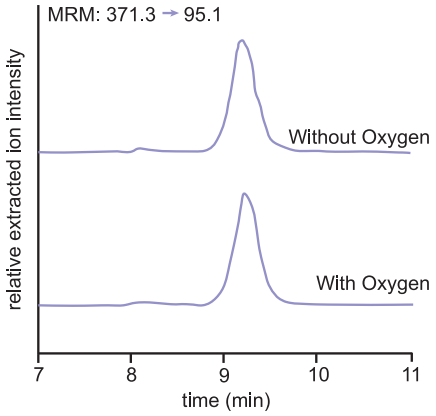


Fig. S4

Analysis of *L. fermentum* NKN51 lysate activity in microaerophilic conditions by Liquid chromatography Mass Spectrometry (LC-MS). The provided extracted ion chromatogram (EIC) represents the product of the whole-cell lysate assay of *L. fermentum* NKN51 supplemented with $NADP^+$ (100 μM) and cholesterol (100 μM) over 12 h in microaerophilic and anaerobic conditions as a control.

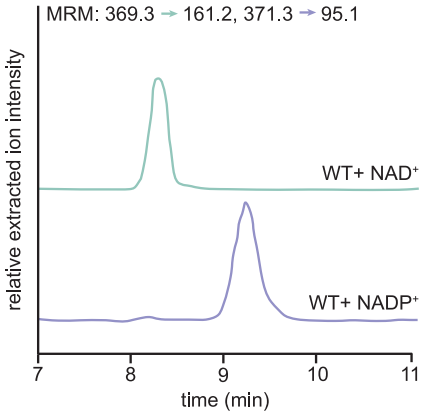


Fig. S5

Investigating the role of NAD^+ as a cofactor for cholesterol metabolism by LC-MS. The reaction mixture of *L. fermentum* NKN51 whole-cell lysate with cholesterol (100 μM) and NAD^+ (100 μM) or $NADP^+$ (100 μM) for 12 h. No reaction was observed with NAD^+ as a cofactor.

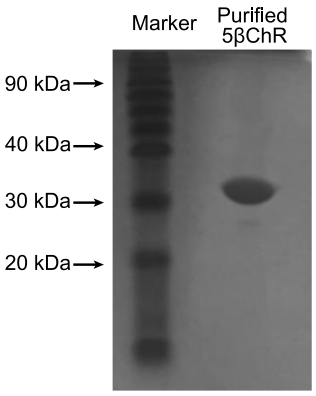


Fig. S6

Recombinant putative cholesterol dehydrogenase protein purification by immobilized affinity chromatography (IMAC). The SDS-PAGE depicts purified 5βChR protein by Ni-NTA-based affinity chromatography, with a theoretical molecular weight of 29.5 kDa.

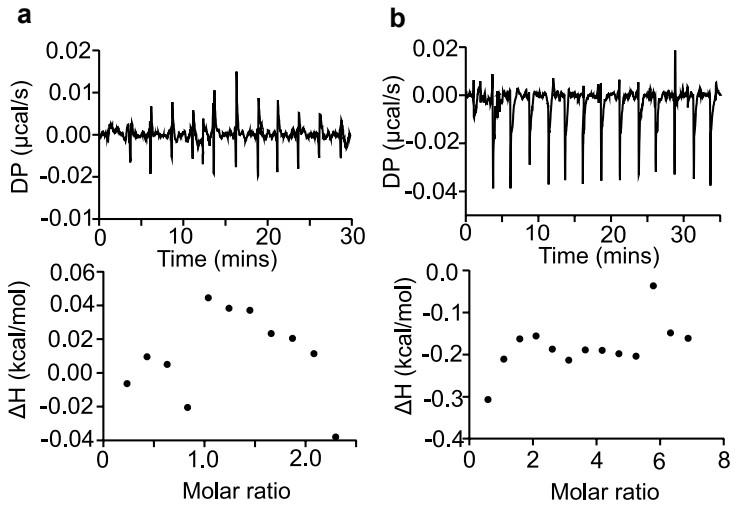


Fig. S7

Binding studies of 5βChR protein with substrates. The Isothermal Titration Calorimetry (ITC) of WT5βChR with (a) cholestenone and (b) coprostanone, wherein the points represent the normalized enthalpy change of the reaction at the different time intervals.

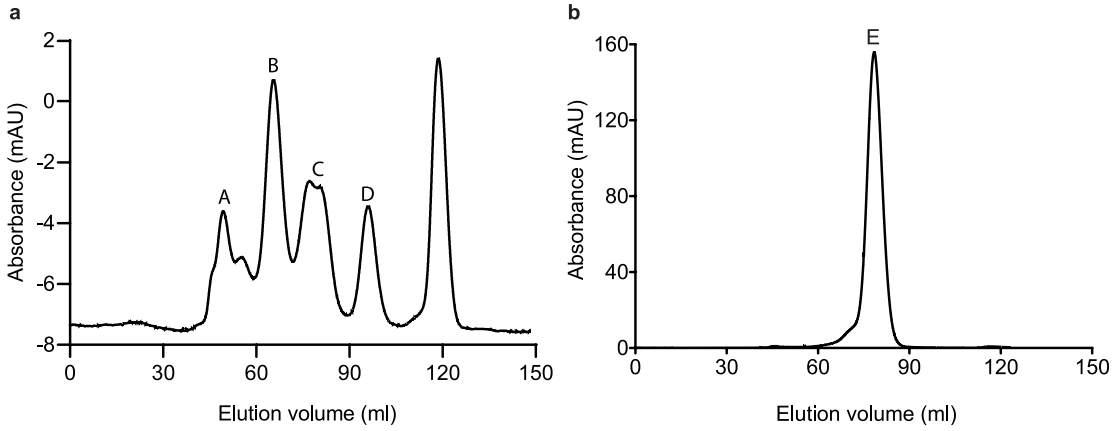


Fig. S8

Calculation of native molecular size of recombinant 5βChR. **(a).** The gel filtration graph of the standard mixture consists of various proteins represented by a distinct peak and their respective molecular weight along with elution volumes, which are as follows: A- Thyroglobulin (670kDa) 49.33ml, B- Gamma globulin (150 kDa) 65.63ml, C- Ovalbumin (44.5 kDa) 80.02 ml and, D- Ribonuclease A (13.7 kDa) 96.97 ml. **(b).** A 29.5 kDa recombinant cholesterol dehydrogenase (E) eluted at 78.43 ml, depicts its dimeric form.

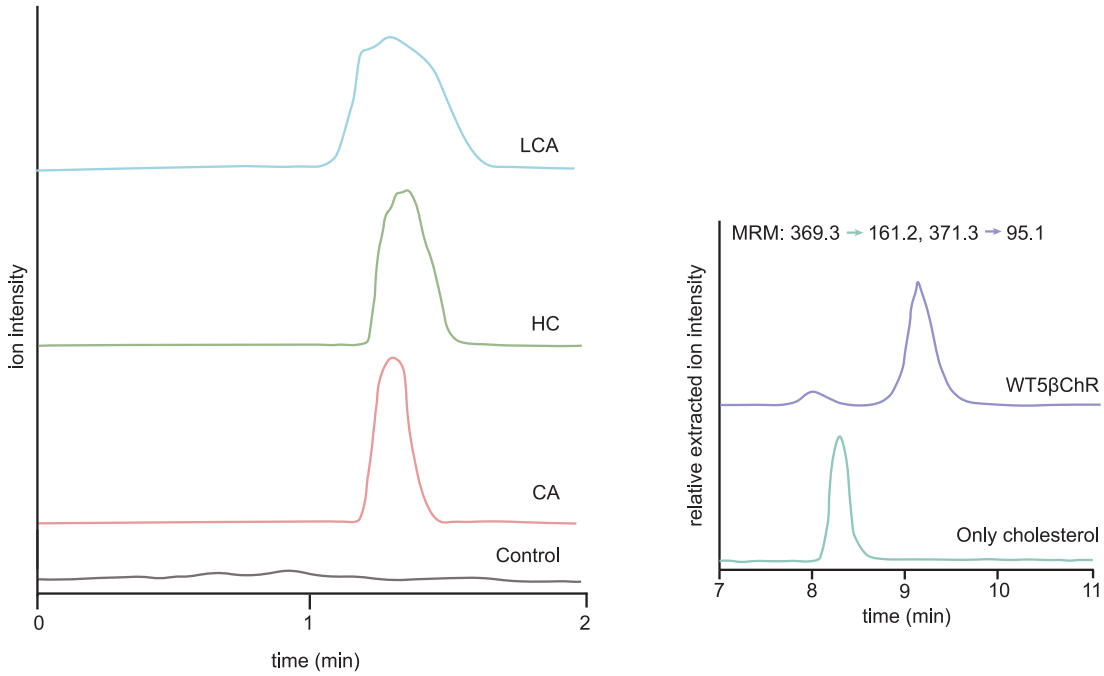


Fig. S9

5βChR enzyme activity with a similar structure compound like cholesterol. The total ion chromatograph demonstrates no reaction observed by 5βChR incubated with either 100 μM cholic acid (CA), hydrocortisone (HC), or lithocholic acid (LCA). EIC shows the coprostanol detection from the reaction of 5βChR incubated with cholesterol (100 μM) and $NADP^+$ (100 μM).

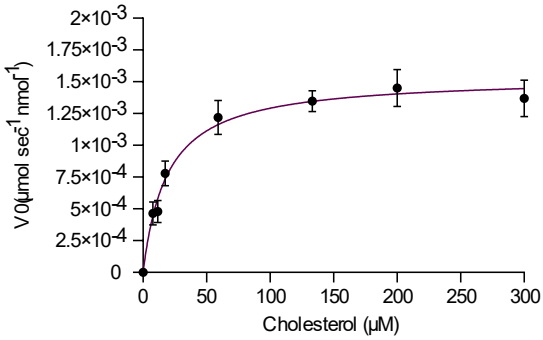


Fig. S10

Enzyme kinetics of WT5βChR. The Michalis-Menton curve is determined from the reaction of cholesterol with 5βChR. The saturation curve predicts the kinetics of cholesterol reduction by recombinant 5βChR enzyme with *Km*- 19.09 μM and V_{max}- 15.35x10^{-6} μmol sec^{-1} $nmol^{-1}$.

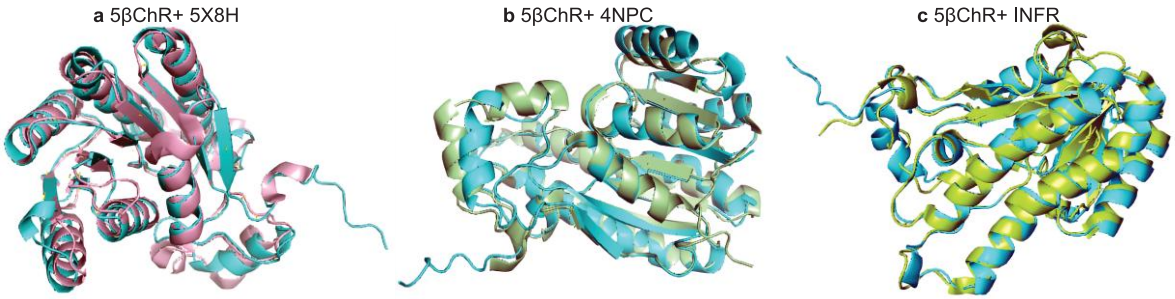


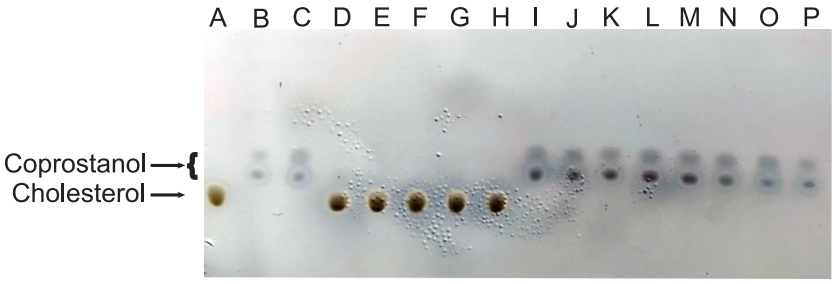
Fig. S11

Structural comparison of 5βChR with its predicted homologs. According to the DALI server, the closest structural homologs of 5βChR are SDR reductase (PDB ID: 5X8H), Sorbitol dehydrogenase (PDB ID: 4NPC), and Putative oxidoreductase RV2002 (PDB ID: 1NFR). These three enzymes are superimposed onto 5βChR with root square mean value (r.m.s) deviation of 1.3, 1.3, and 1.2 Å, respectively.

```
Amino acid number 131     136     139         152       156       172     182184  187 189190  194     198
WT                HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
H131A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
N136A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
S139A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
Y152A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
K156A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
E172A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
P182A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
P184A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
T187A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
M189A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
F190A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
M194A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK......F
F198A             HG..... N........SV...........YN.........KG.........ED.....P.P.......T......M  F......MK.....F
```

Fig. S12

Site-directed mutagenesis of WT5βChR. The highlighted amino acids represent the mutation in conserved residues of WT5βChR. N-terminus residues, H131, N136, S139, Y152, K156 and E172 mutated to alanine. C-terminus residues, P182, P184, T187, M189, F190, M194, and F198, are converted to alanine for cholesterol binding analysis.



A- Cholesterol standarda, B-Coprostanol standard, C-5βChR, D-S139A, E- Y152A, F- T187A, G- M189A, H- F190A, I- H131A, J-N136A, K-156A, L- E172A, M- P182A ,N- P184A, O- M194A, and P- F198A

Fig. S13

Cholesterol transformation profile of WT and mutant 5βChR**.** The reaction mixture of 5βChR WT and mutants with cholesterol (100 μM) and $NADP^+$ (100 μM) are analyzed through TLC. S139A, Y152A, T187A, M189A, and F190A show the spot equivalent to cholesterol. H131A, N136A, 156A, E172A, P182A, P184A, M194A, and F198A highlighting spots corresponding to coprostanol.

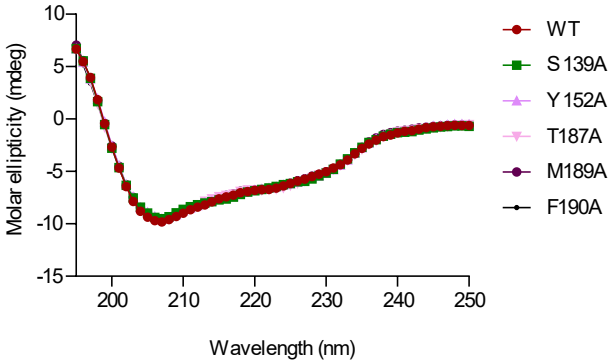


Fig. S14

Circular Dichroism spectra of 5βChR and its mutants. The graph represents the secondary structure profiles of WT5βChR, S139A, Y152A, T187A, M189A, and F190A.

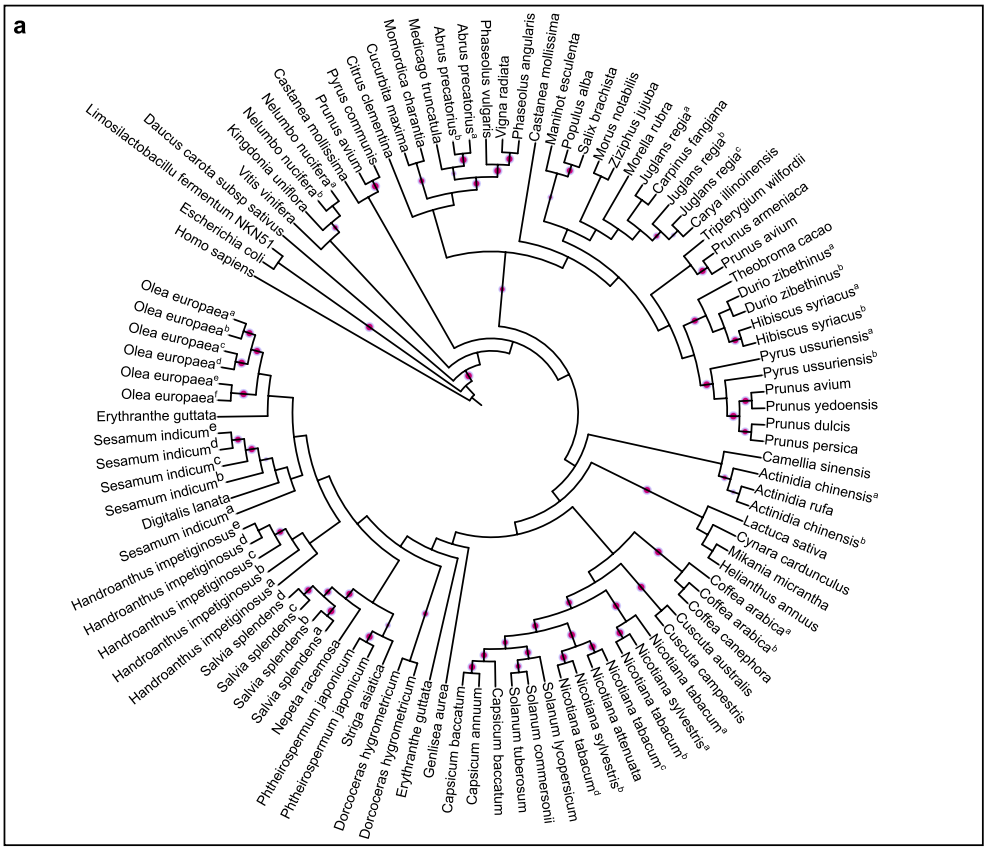
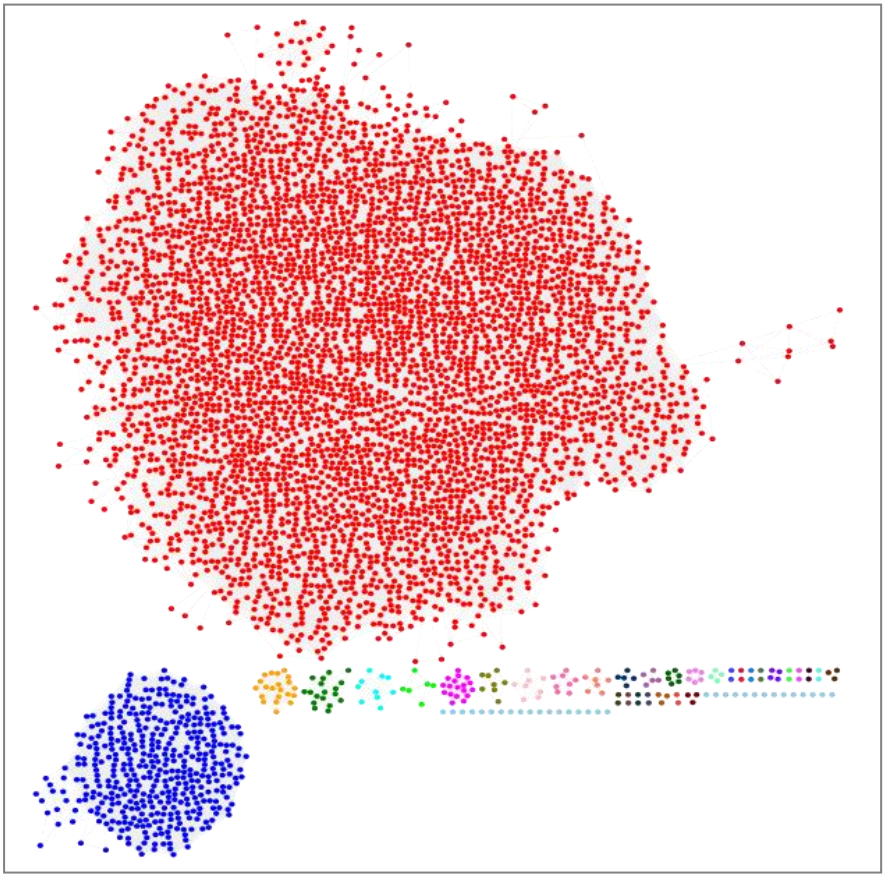


Fig. S15

Mapping the prevalence of 5βChR within plants by phylogenetic profiling. **(a).** The evolution analysis of 5βChR homologs in plants (~100) is acquired from UniProt in RAxML using the PROTCATWAG model with human 5β reductase as an outgroup. 5βChR and plant 5β reductase evolved from a common ancestor. A pink circle indicates a bootstrap value of more than 70.

a



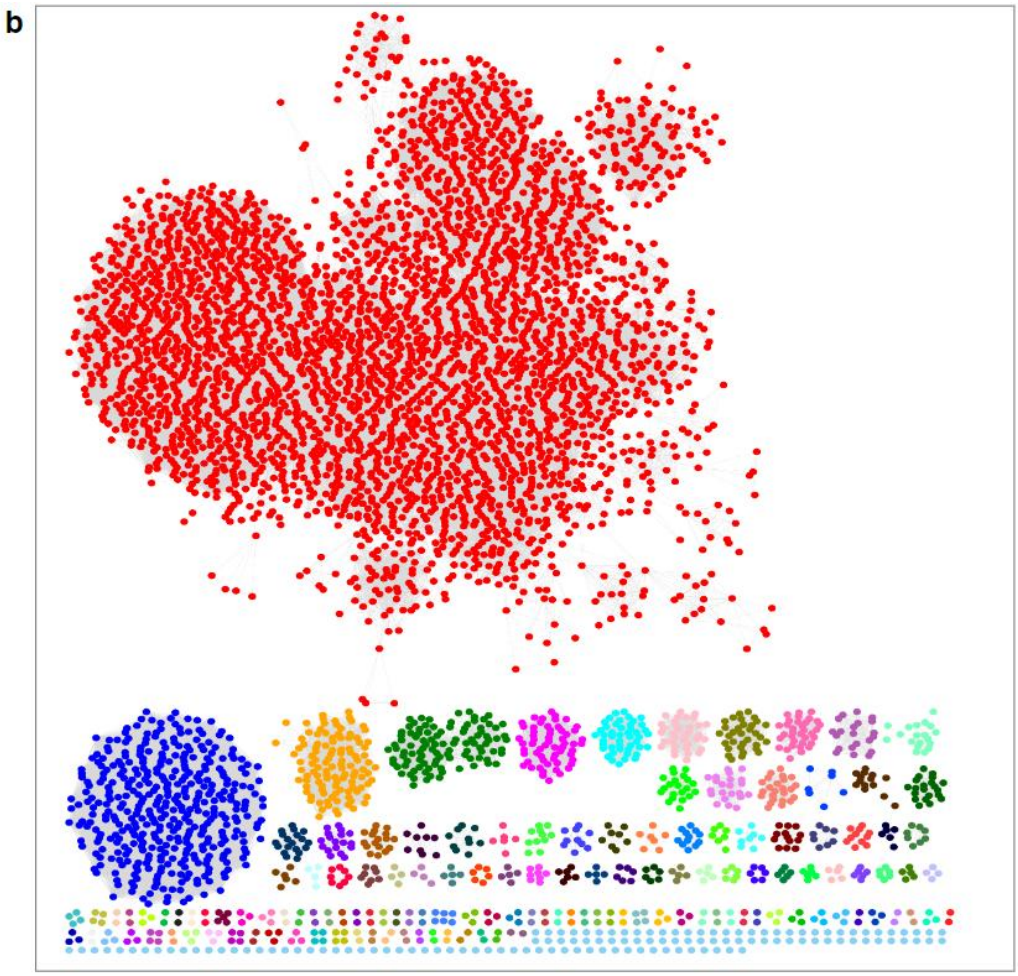


Fig. S16

5βChR sequence similarity network. **(a)** A sequence similarity Network (SSN) is built using 5000 proteins from the UniProt with 60% sequence coverage and >50% sequence identity to 5βChR. 4383 proteins out of 5000 are combined into a single cluster, and the rest of the protein sequences are divided into 32 clusters. **(b)** With a 70% threshold and >60% sequence identity to 5βChR, the 5000 proteins are divided into 138 clusters based on functional similarity and increasing sequence identity.

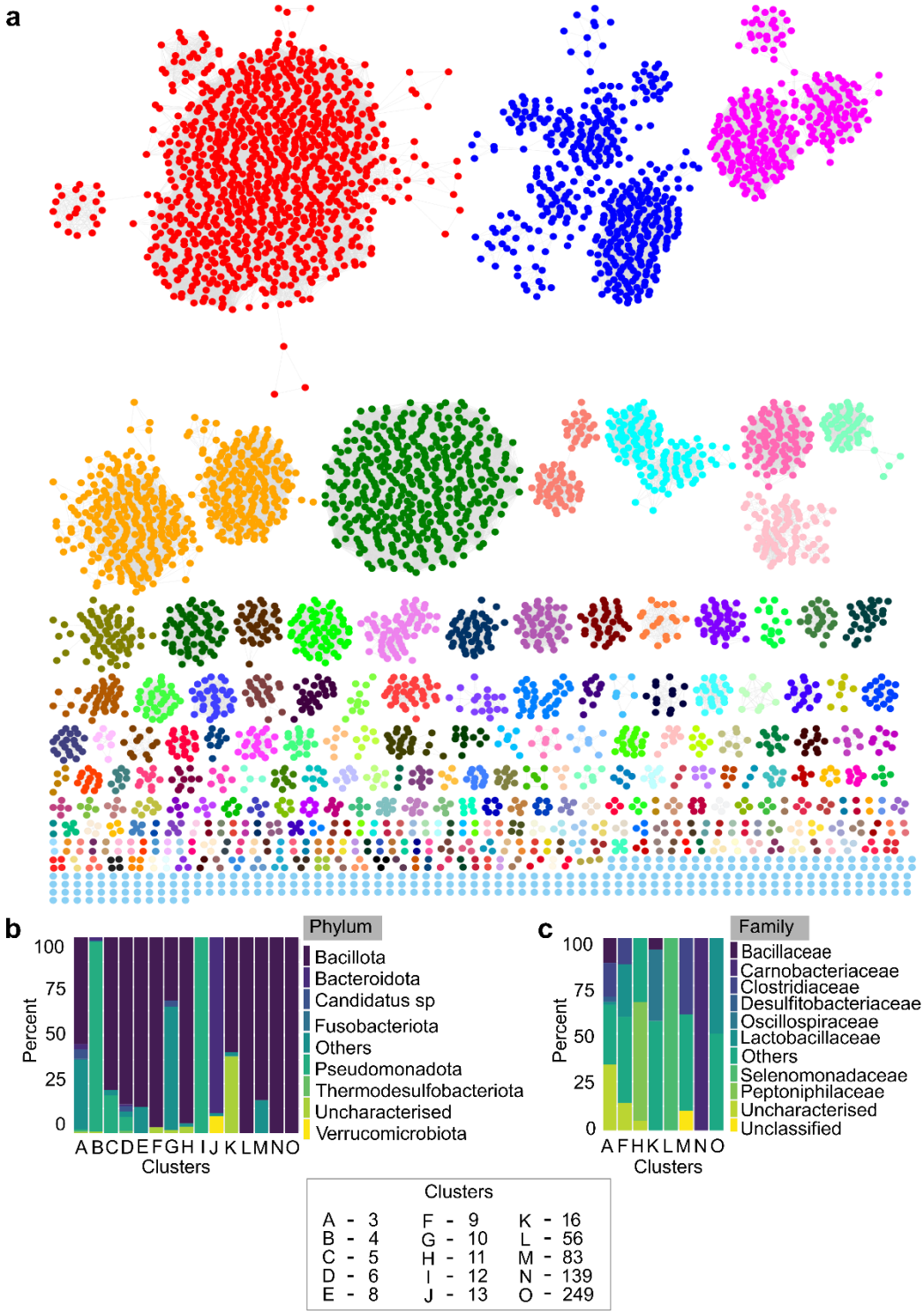


Fig. S17

Network maps illustrating the abundance and distribution of the 5βChR gene family across the different microbes. **(a)**. Microbial origin 5βChR protein homologs are retrieved from UniProt for sequence similarity network (SSN) analysis. These networks are visualized by the Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST) with threshold values of 10^{-80}. The proteins are divided into 269 clusters based on sequence identity, structural similarities, and functional profile of 5βChR proteins, highlighting their presence in various microbes. **(b-c)** The graphs depict the taxonomy of microbial

5βChR homologs from the selected clusters, which are prevalent in human microbiome data. The representative sequences of clusters belong to different bacteria belonging to diverse phyla and families.

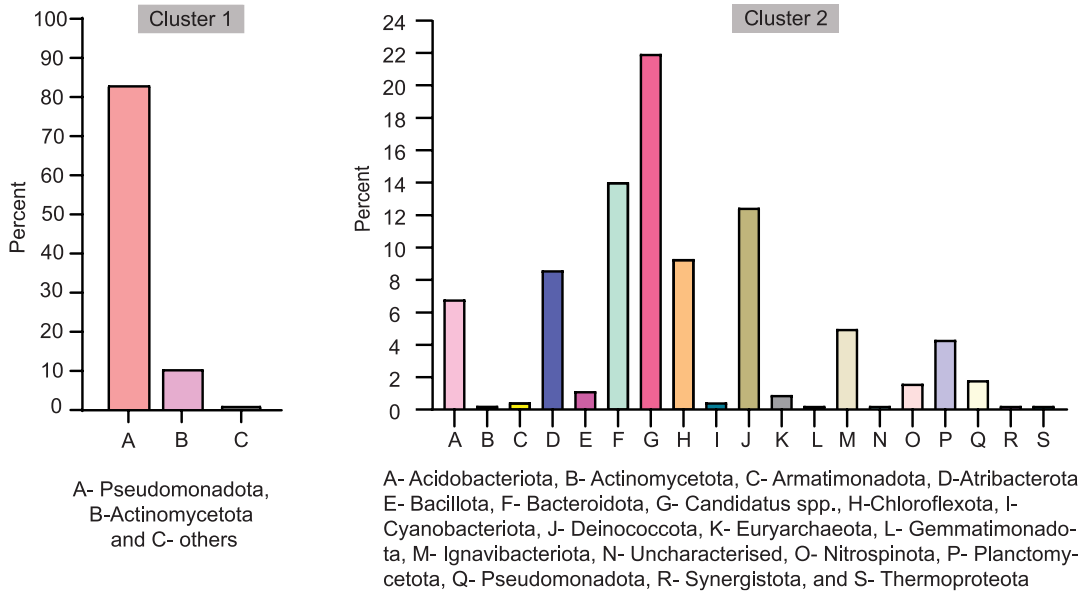


Fig. S18

Assessing the taxonomy of 5βChR representatives. The phyla-based classification of clusters 1 and 2 of 5βChR homologs with >80% sequence coverage and >70% sequence identity. The bar graphs show the prevalence of 5βChR protein in various phyla.

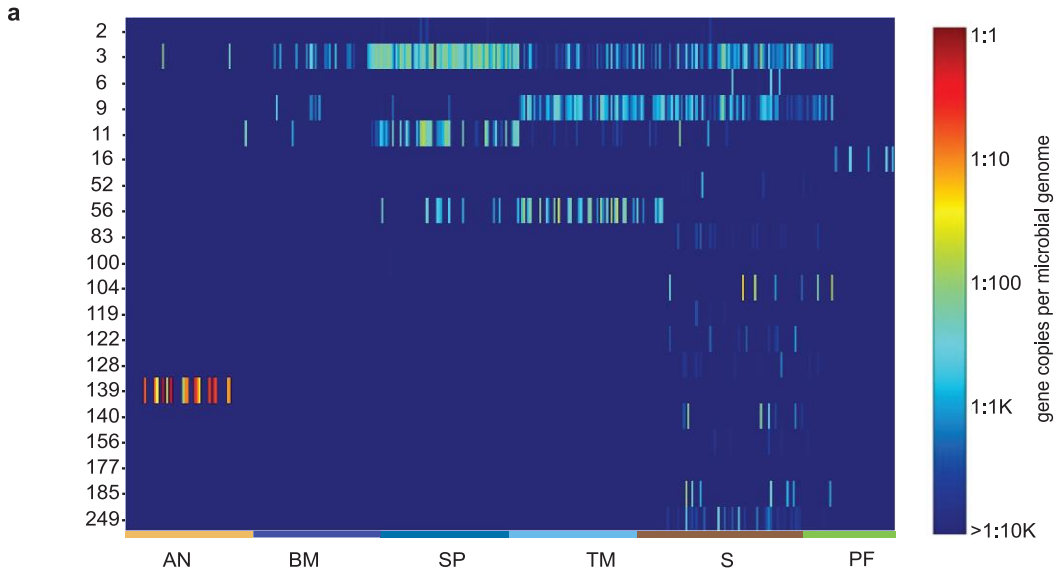
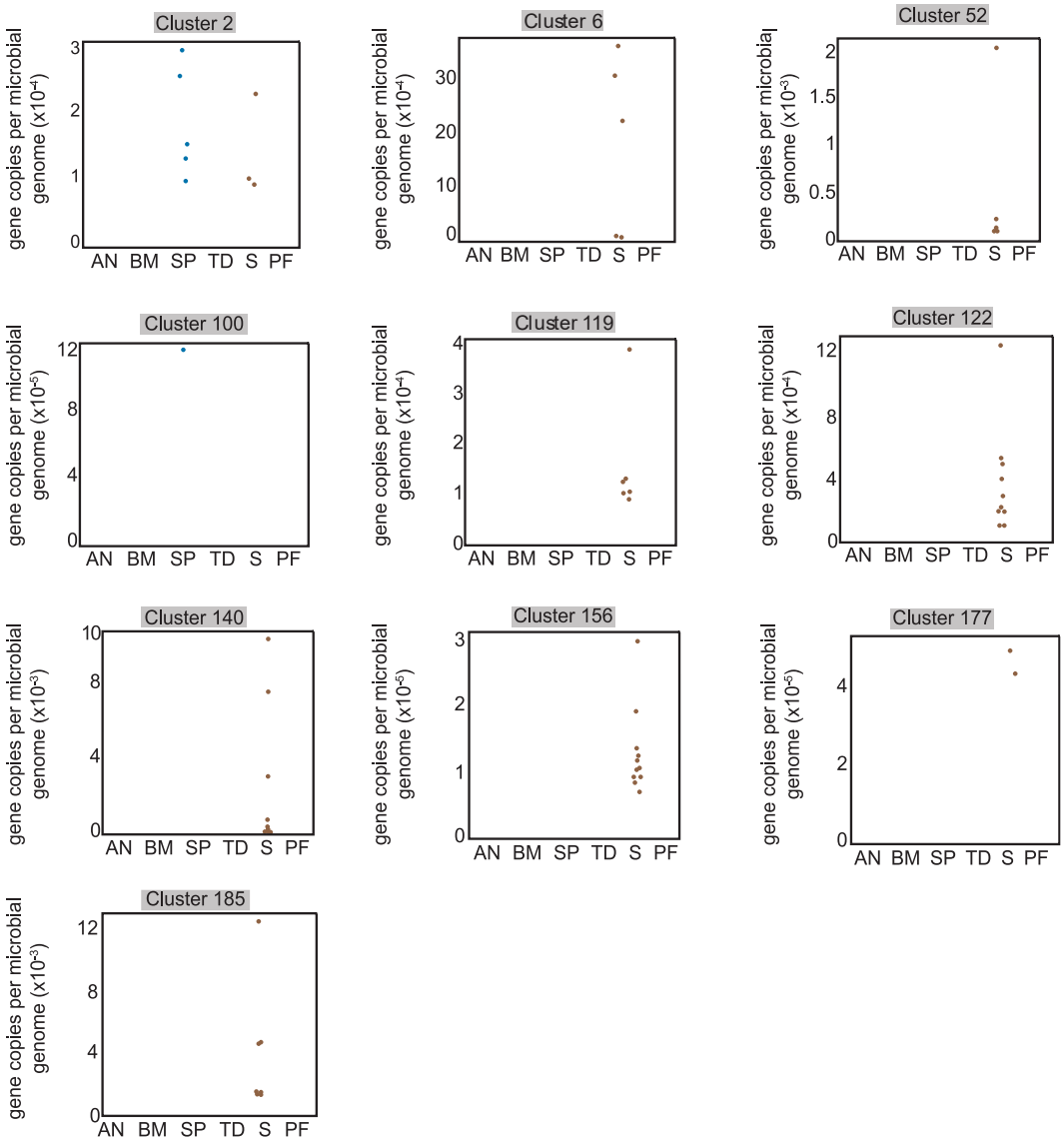


Fig. S19

Meta-analysis depicts the abundance of 5βChR homologs in a healthy microbiome. The heat map illustrates the prevalence of 5βChR homologs in the human metagenomic data of 380 healthy individuals. Analysis validates the presence of 20 clusters in the participants' metagenome data.



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

143 **Fig. S20**

144 The distribution of different 5βChR homolog clusters in the 380-microbiome data of healthy human
145 participants. The box plot displays the various clusters of 5βChR homologs primarily present in stool
146 samples.

147 **2. Supplementary Tables**

148 **Table S1. Primer details used in this study.**

Primer name	Sequence 5'-3'
5βChR FP	TTATCATATGATGCAAACAGTTGTTATCACGGGGGCC
5βChR RP	TTTATCTCGAGTTATACCTGCTTCGGTTGGCCCGTCG
5βChR RT FP	CGTGGAACCAGAGGAAGTGG
5βChR RT RP	TGGCCCGTCGAAATCTCAAA
16S rRNA RT FP	AGTCACGGCTAACTACGTGC
16S rRNA RT RP	ACCGCCTGCACTCTCTTTAC
S139RP	TAACATCTCGgcgGTTTCCGGGATGG
S139FP	ACAATGGCGCCGTGTCGT
Y152RP	GATGGTGGCCgcgAACGCCACCAAG
Y152FP	CCGTAGTCACCACCCATC
K156FP	TAACGCCACCgcgGGGGCCATCATC
K156RP	TAGGCCACCATCCCGTAG
H131 FP	AGACCAACGAgcgGGCGCCATTGTTAACATCTCGTCC
H131 RP	AGCATGGTCGGCAGGACC
N136 FP	CGCCATTGTTgcgATCTCGTCCGTTTC
N136 RP	CCGTGTCGTTGGTCTAGC
E172 FP	TGACTACGGTgcgGACCAGATCCG
E172 RP	ATCGCCATCGCCCGGGTC
P182 FP	CAGCGTGGCCgcgGGTCCAACCA
P182 RP	TTGACGCGGATCTGGTCCTCACC
P184 FP	GGCCCCGGGTgcgACCAACACCC
P184 RP	ACGCTGTTGACGCGGATCTGG
T187 FP	TCCAACCAACgcgCCGATGTTCCCGC
T187 RP	CCCGGGGCCACGCTGTTG
M189 FP	CAACACCCCGgcgTTCCCGCAAGACATGAAGG
M189 RP	GTTGGACCCGGGGCCACG
F190 FP	CACCCCGATGgcgCCGCAAGACATGAAGGGG
F190 RP	TTGGTTGGACCCGGGGCC
M194 FP	CCCGCAAGACgcgAAGGGGACCT
M194 RP	AACATCGGGGTGTTGGTT
F198 FP	GAAGGGGACCgcgGCCAAAAACAG
F198 RP	ATGTCTTGCGGGAACATC

149

150 **Table S2. Lactobacillus strains used for phylogenetic analysis.**

Species
Lactobacillus_delbrueckii_subsp._bulgaricus
Lactobacillus_delbrueckii_subsp._delbrueckii
Lactobacillus_delbrueckii_subsp._indicus
Lactobacillus_delbrueckii_subsp._Jakobsenii
Lactobacillus_delbrueckii_subsp._lactis
Lactobacillus_delbrueckii_subsp._sunkii
Lactobacillus_acetotolerans
Lactobacillus_acidophilus
Lactobacillus_amylolyticus
Lactobacillus_amylovorus
Lactobacillus_apis
Lactobacillus_bombicola
Lactobacillus_crispatus
Lactobacillus_equicursoris
Lactobacillus_gallinarum
Lactobacillus_gasseri
Lactobacillus_gigeriorum
Lactobacillus_jensenii
Lactobacillus_johnsonii
Lactobacillus_hamsteri
Lactobacillus_helsingborgensis
Lactobacillus_helveticus
Lactobacillus_hominis
Lactobacillus_iners
Lactobacillus_intestinalis
Lactobacillus_kalixensis
Lactobacillus_kefiranofaciens_subsp._kefiranofaciens
Lactobacillus_kefiranofaciens_subsp._kefirgranum
Lactobacillus_kimbladii
Lactobacillus_kitasatonis
Lactobacillus_kullabergensis
Lactobacillus_melliventris
Lactobacillus_mulieris
Lactobacillus_panisapium
Lactobacillus_paragasseri
Lactobacillus_pasteurii
Lactobacillus_psittaci
Lactobacillus_rodentium

Lactobacillus_taiwanensis
Lactobacillus_ultunensis
Lactobacillus_xujianguonis
Amylolactobacillus_amylophilus
Holzapfelia_floricola
Bombilactobacillus_mellifer
Bombilactobacillus_bombi
Bombilactobacillus_mellis
Companilactobacillus_alimentarius
Companilactobacillus_allii
Companilactobacillus_baiquanensis
Companilactobacillus_bobalius
Companilactobacillus_crustorum
Companilactobacillus_farciminis
Companilactobacillus_formosensis
Companilactobacillus_furfuricola
Companilactobacillus_futsaii
Companilactobacillus_ginsenosidimutans
Companilactobacillus_halodurans
Companilactobacillus_heilongjiangensis
Companilactobacillus_huachuanensis
Companilactobacillus_hulinensis
Companilactobacillus_insicii
Companilactobacillus_jidongensis
Companilactobacillus_kedongensis
Companilactobacillus_keshanensis
Companilactobacillus_kimchiensis
Companilactobacillus_kimchii
Companilactobacillus_metriopterae
Companilactobacillus_mindensis
Companilactobacillus_mishanensis
Companilactobacillus_musae
Companilactobacillus_nantensis
Companilactobacillus_nodensis
Companilactobacillus_nuruki
Companilactobacillus_paralimentarius
Companilactobacillus_suantsaicola
Companilactobacillus_tucceti
Companilactobacillus_versmoldensis

Companilactobacillus_zhachilii
Companilactobacillus_zhongbaensis
Lapidilactobacillus_concavus
Lapidilactobacillus_bayanensis
Lapidilactobacillus_dextrinicus
Agrilactobacillus_composti
Agrilactobacillus_yilanensis
Schleiferilactobacillus_perolens
Schleiferilactobacillus_harbinensis
Schleiferilactobacillus_shenzhenensis
Lacticaseibacillus_casei
Lacticaseibacillus_baoqingensis
Lacticaseibacillus_brantae
Lacticaseibacillus_camelliae
Lacticaseibacillus_chiayiensis
Lacticaseibacillus_hulanensis
Lacticaseibacillus_jixianensis
Lacticaseibacillus_manihotivorans
Lacticaseibacillus_nasuensis
Lacticaseibacillus_pantheris
Lacticaseibacillus_paracasei_subsp._paracasei
Lacticaseibacillus_paracasei_subsp._tolerans
Lacticaseibacillus_porcinae
Lacticaseibacillus_rhamnosus
Lacticaseibacillus_saniviri
Lacticaseibacillus_sharpeae
Lacticaseibacillus_songhuajiangensis
Lacticaseibacillus_thailandensis
Paralactobacillus_selangorensis
Latilactobacillus_sakei_subsp._carnosus
Latilactobacillus_sakei_subsp._sakei
Latilactobacillus_curvatus
Latilactobacillus_fuchuensis
Latilactobacillus_graminis
Loigolactobacillus_coryniformis_subsp._coryniformis
Loigolactobacillus_coryniformis_subsp._torquens
Loigolactobacillus_backii
Loigolactobacillus_bifermentans
Loigolactobacillus_iwatensis

Loigolactobacillus_jiayinensis
Loigolactobacillus_rennini
Loigolactobacillus_zhaoyuanensis
Dellaglioa_algida
Liquorilactobacillus_mali
Liquorilactobacillus_aquaticus
Liquorilactobacillus_cacaonum
Liquorilactobacillus_capillatus
Liquorilactobacillus_ghanensis
Liquorilactobacillus_hordei
Liquorilactobacillus_nagelii
Liquorilactobacillus_satsumensis
Liquorilactobacillus_oeni
Liquorilactobacillus_sucicola
Liquorilactobacillus_uvarum
Liquorilactobacillus_vini
Ligilactobacillus_salivarius
Ligilactobacillus_acidipiscis
Ligilactobacillus_agilis
Ligilactobacillus_animalis
Ligilactobacillus_apodemi
Ligilactobacillus_araffinosus
Ligilactobacillus_aviarius
Ligilactobacillus_ceti
Ligilactobacillus_equi
Ligilactobacillus_hayakitensis
Ligilactobacillus_murinus
Ligilactobacillus_pobuzihii
Ligilactobacillus_ruminis
Ligilactobacillus_saerimneri
Lactiplantibacillus_plantarum_subsp._plantarum
Lactiplantibacillus_plantarum_subsp._argentoratensis
Lactiplantibacillus_daoliensis
Lactiplantibacillus_daowaiensis
Lactiplantibacillus_dongliensis
Lactiplantibacillus_fabifermentans
Lactiplantibacillus_herbarum
Lactiplantibacillus_modestisalitolerans
Lactiplantibacillus_mudanjiangensis

Lactiplantibacillus_nangangensis
Lactiplantibacillus_paraplantarum
Lactiplantibacillus_pentosus
Lactiplantibacillus_pingfangensis
Lactiplantibacillus_plajomi
Lactiplantibacillus_songbeiensis
Lactiplantibacillus_xiangfangensis
Furfurilactobacillus_rossiae
Furfurilactobacillus_siliginis
Paucilactobacillus_vaccinostercus
Paucilactobacillus_hokkaidonensis
Paucilactobacillus_kaifaensis
Paucilactobacillus_nenjiangensis
Paucilactobacillus_oligofermentans
Paucilactobacillus_suebicus
Paucilactobacillus_wasatchensis
Limosilactobacillus_fermentum
Limosilactobacillus_antri
Limosilactobacillus_coleohominis
Limosilactobacillus_equigenerosi
Limosilactobacillus_frumenti
Limosilactobacillus_gastricus
Limosilactobacillus_gorillae
Limosilactobacillus_ingluviei
Limosilactobacillus_mucosae
Limosilactobacillus_oris
Limosilactobacillus_panis
Limosilactobacillus_pontis
Limosilactobacillus_reuteri
Limosilactobacillus_secaliphilus
Limosilactobacillus_vaginalis
Secundilactobacillus_malefermentans
Secundilactobacillus_collinoides
Secundilactobacillus_kimchicus
Secundilactobacillus_mixtipabuli
Secundilactobacillus_odoratitofui
Secundilactobacillus_oryzae
Secundilactobacillus_paracollinoides
Secundilactobacillus_pentosiphilus

Secundilactobacillus_silagei
Secundilactobacillus_silagincola
Secundilactobacillus_similis
Levilactobacillus_brevis
Levilactobacillus_acidifarinae
Levilactobacillus_bambusae
Levilactobacillus_cerevisiae
Levilactobacillus_fujinensis
Levilactobacillus_fuyuanensis
Levilactobacillus_hammesii
Levilactobacillus_huananensis
Levilactobacillus_koreensis
Levilactobacillus_lindianensis
Levilactobacillus_mulengensis
Levilactobacillus_namurensis
Levilactobacillus_parabrevis
Levilactobacillus_paucivorans
Levilactobacillus_senmaizukei
Levilactobacillus_spicheri
Levilactobacillus_suantsaii
Levilactobacillus_suantsaiihabitans
Levilactobacillus_tangyuanensis
Levilactobacillus_tongjiangensis
Levilactobacillus_yonginensis
Levilactobacillus_zymae
Fructilactobacillus_fructivorans
Fructilactobacillus_florum
Fructilactobacillus_lindneri
Fructilactobacillus_sanfranciscensis
Acetilactobacillus_jinshanensis
Apilactobacillus_kunkeei
Apilactobacillus_apinorum
Apilactobacillus_micheneri
Apilactobacillus_ozensis
Apilactobacillus_quenuiae
Apilactobacillus_timberlakei
Lentilactobacillus_buchneri
Lentilactobacillus_curieae
Lentilactobacillus_diolivorans

Lentilactobacillus_farraginis
Lentilactobacillus_hilgardii
Lentilactobacillus_kefiri
Lentilactobacillus_kisonensis
Lentilactobacillus_otakiensis
Lentilactobacillus_parabuchneri
Lentilactobacillus_parafarraginis
Lentilactobacillus_parakefiri
Lentilactobacillus_raoultii
Lentilactobacillus_rapi
Lentilactobacillus_senioris
Lentilactobacillus_sunkii
Lapidilactobacillus_achengensis
Lapidilactobacillus_gannanensis
Lapidilactobacillus_mulanensis
Lapidilactobacillus_wuchangensis
Lacticaseibacillus_daqingensis
Lacticaseibacillus_hegangensis
Lacticaseibacillus_suibinensis
Lacticaseibacillus_yichunensis
Loigolactobacillus_binensis
Lactiplantibacillus_garii
Levilactobacillus_angrenensis

Table S3. List of 5βChR plants homologs

Name & Uniprot ID
Digitalis_lanata_Q6PQJ
Sesamum_indicum_A0A6I9TR48_SESIN (Sesamum_indicum d)
Erythranthe_guttata_A0A022RBG3_ERYGU
Handroanthus_impetiginosus_A0A2G9HCN6_9LAMI (Handroanthus_impetiginosus c)
Sesamum_indicum_A0A6I9UEB3_SESIN
Handroanthus_impetiginosus_A0A2G9HDB4_9LAMI (Handroanthus_impetiginosus b)
Handroanthus_impetiginosus_A0A2G9FZQ9_9LAMI (Handroanthus_impetiginosus d)
Handroanthus_impetiginosus_A0A2G9G080_9LAMI (Handroanthus_impetiginosus a)
Phtheirospermum_japonicum_A0A830BWA5_9LAMI (Phtheirospermum_japonicum a)
Striga_asiatica_A0A5A7PVK4_STRAF
Olea_europaea_subspe_uropaea_A0A8S0T3M5_OLEEU (Olea_europaea f)
Handroanthus_impetiginosus_A0A2G9HCM6_9LAMI (Handroanthus_impetiginosus e)

Olea_europaea_subsp_europaea_A0A8S0UNL0_OLEEU (Olea_europaea d)
Olea_europaea_subsp_europaea_A0A8S0TDW3_OLEEU (Olea_europaea c)
Olea_europaea_subsp_europaea_A0A8S0T0U6_OLEEU (Olea_europaea e)
Dorcoceras_hygrometricum_A0A2Z7BQP4_9LAMI (Dorcoceras_hygrometricum b)
Nepeta_racemosa_ISY1_NEPRA
Sesamum_indicum_A0A6I9TWM7_SESIN (Sesamum_indicum b)
Genlisea_aurea_S8DQY0_9LAMI
Sesamum_indicum_A0A8M8V613_SESINO (Sesamum_indicum a)
Coffea_arabica_A0A6P6WBJ5_COFAR (Coffea_arabica a)
Coffea_canephora_A0A068TPH4_COFCA (Coffea_arabica b)
Coffea_arabica_A0A6P6VU68_COFAR
Phtheirospermum_japonicum_A0A830BZG9_9LAMI (Phtheirospermum_japonicum b)
Salvia_splendens_A0A8X8Y7X1_SALSN (Salvia_splendens a)
Salvia_splendens_A0A8X8Y1P7_SALSN (Salvia_splendens b)
Nicotiana_attenuata_A0A1J6KD04_NICAT
Nicotiana_tabacum_C0SUC5_TOBAC (Nicotiana_tabacum c)
Solanum_lycopersicum_A0A3Q7IFK3_SOLLCOX
Nicotiana_sylvestris_A0A1U7W281_NICSY (Nicotiana_sylvestris b)
Nicotiana_tabacum_A0A1S4C7K6_TOBAC (Nicotiana_tabacum d)
Salvia_splendens_A0A8X8XA34_SALSN (Salvia_splendens c)
Olea_europaea_subsp_europaea_A0A8S0S199_OLEEU (Olea_europaea a)
Solanum_tuberosum_M1A3I7_SOLTU
Erythranthe_guttata_A0A022RB89_ERYGU
Salvia_splendens_A0A8X8X5Q9_SALSN (Salvia_splendens d)
Solanum_commersonii_A0A9J5WZQ5_SOLCO
Dorcoceras_hygrometricum_A0A2Z7AKE5_9LAMI (Dorcoceras_hygrometricum a)
Capsicum_baccatum_A0A2G2W4K4_CAPBA
Cuscuta_campestris_A0A484KG59_9ASTE
Camellia_sinensis_A0A7J7HHB0_CAMSI
Olea_europaea_subsp_europaea_A0A8S0SEQ7_OLEEU
Cuscuta_australis_ A0A328D8P5_9ASTE
Capsicum_baccatum_A0A2G2XN52_CAPBA
Actinidia_chinensis_var_chinensis_A0A2R6Q5W0_ACTCC (Actinidia_chinensis a)
Juglans_regia_A0A2I4FQH4_JUGRE (Juglans_regia a)
Juglans_regia_A0A2I4FWX2_JUGRE (Juglans_regia c)
Daucus_carota_subsp_sativus_A0A166GIS8_DAUCS
Lactuca_sativa_A0A166GIS8_DAUCS
Actinidia_chinensis_var_chinensis_A0A2R6RGX9_ACTCC (Actinidia_chinensis b)
Carya_illinoinensis_A0A8T1N750_CARIL

Nelumbo_nucifera_A0A1U8 (Nelumbo_nucifera b)
Carpinus_fangiana_A0A5N6RKZ6_9ROSI
Prunus_yedoensis_var_nudiflora_A0A314XHP9_PRUYE
Prunus_dulcis_A0A5E4E997_PRUDU
Cynara_cardunculus_var_scolymus_A0A103XM59_CYNCS
Sesamum_indicum_A0A6I9U2W0_SESIN (Sesamum_indicum c)
Morus_notabilis_W9SGC8_9ROSA
Nelumbo_nucifera_A0A1U8B751_NELNU (Nelumbo_nucifera a)
Prunus_avium_ A0A6P5RCG0_PRUAV
Prunus_persica_A0A251RBQ8_PRUPE
Kingdonia_uniflora_A0A7J7LY63_9MAGN
Durio_zibethinus_A0A6P6B6U9_DURZI (Durio_zibethinus a)
Pyrus_ussuriensis_Pyrus_communis_A0A5N5GW02_9ROSA (Pyrus_ussuriensis a)
Helianthus_annuus_ A0A251TMZ3_HELAN
Abrus_precatorius_A0A8B8LJU1_ABRPR (Abrus_precatorius b)
Pyrus_ussuriensis_Pyrus_communis (Pyrus_ussuriensis b)
Durio_zibethinus_A0A6P5WM40_DURZI (Durio_zibethinus b)
Actinidia_rufa_A0A7J0GN97_9ERIC
Nicotiana_tabacum_A0A1S4B9B5_TOBAC (Nicotiana_tabacum a)
Abrus_precatorius_A0A8B8LHB1_ABRPR (Abrus_precatorius a)
Vigna_radiata_var_radiata_A0A1S3TXR7_VIGRR
Mikania_micrantha_A0A5N6LLS5_9ASTR
Phaseolus_vulgaris_V7AWY5_PHAVU
Theobroma_cacao_A0A061FFF0_THECC
Ziziphus_jujuba_A0A6P3ZUB5_ZIZJJ
Nicotiana_sylvestris_A0A1U7WAQ4_NICSY (Nicotiana_sylvestris a)
Nicotiana_tabacum_A0A1S3XWC9_TOBAC (Nicotiana_tabacum b)
Castanea_mollissima_A0A8J4RLX8_9ROSI
Capsicum_annuum_A0A2G3AN45_CAPAN
Manihot_esculenta_A0A2C9VXF4_MANES
Vitis_vinifera_F6H675_VITVI
Hibiscus_syriacus_A0A6A2WIU1_HIBSY (Hibiscus_syriacus b)
Morella_rubra_A0A6A1W9A5_9ROSI
Pyrus_ussuriensis_Pyrus_communis_A0A5N5I733_9ROSA
Tripterygium_wilfordii_A0A7J7CZC2_TRIWF
Juglans_regia_A0A2I4HD46_JUGRE (Juglans_regia c)
Prunus_avium_A0A6P5RRI7_PRUAV
Cucurbita_maxima_A0A6J1J0N7_CUCMA
Populus_alba_A0A4U5N810_POPAL

Castanea_mollissima_A0A8J4RN14_9ROSI
Phaseolus_angularis_A0A0L9UHR8_PHAAN
Prunus_avium_A0A6P5RFJ5_PRUAV
Salix_brachista_A0A5N5K873_9ROSI
Momordica_charantia_A0A6J1DZW3_MOMCH
Citrus_clementina_V4T9V0_CITCL
Hibiscus_syriacus_A0A6A2WII9_HIBSY (Hibiscus_syriacus a)
Prunus_armeniaca_A0A6J5W5F7_PRUAR
Medicago_truncatula_G8A2B2_MEDTR

153

154 **Table S4. NCBI accession number of metagenome samples.**

Metagenome ID	Body Site
SRS011061	stool
SRS011090	buccal mucosa
SRS011098	supragingival plaque
SRS011126	supragingival plaque
SRS011132	anterior nares
SRS011134	stool
SRS011140	tongue dorsum
SRS011144	buccal mucosa
SRS011152	supragingival plaque
SRS011239	stool
SRS011243	tongue dorsum
SRS011247	buccal mucosa
SRS011255	supragingival plaque
SRS011263	anterior nares
SRS011269	posterior fornix
SRS011271	stool
SRS011302	stool
SRS011306	tongue dorsum
SRS011310	buccal mucosa
SRS011343	supragingival plaque
SRS011355	posterior fornix
SRS011397	anterior nares
SRS011405	stool
SRS011452	stool
SRS011529	stool
SRS011584	posterior fornix

SRS011586	stool
SRS012273	stool
SRS012279	tongue dorsum
SRS012281	buccal mucosa
SRS012285	supragingival plaque
SRS012291	anterior nares
SRS012294	posterior fornix
SRS012663	anterior nares
SRS012902	stool
SRS013155	anterior nares
SRS013158	stool
SRS013164	tongue dorsum
SRS013170	supragingival plaque
SRS013215	stool
SRS013234	tongue dorsum
SRS013239	buccal mucosa
SRS013252	supragingival plaque
SRS013269	anterior nares
SRS013476	stool
SRS013502	tongue dorsum
SRS013506	buccal mucosa
SRS013521	stool
SRS013533	supragingival plaque
SRS013542	posterior fornix
SRS013637	anterior nares
SRS013687	stool
SRS013705	tongue dorsum
SRS013711	buccal mucosa
SRS013723	supragingival plaque
SRS013800	stool
SRS013818	tongue dorsum
SRS013825	buccal mucosa
SRS013836	supragingival plaque
SRS013876	anterior nares
SRS013879	tongue dorsum
SRS013881	buccal mucosa
SRS013945	buccal mucosa
SRS013949	supragingival plaque
SRS013951	stool

SRS013956	anterior nares
SRS014124	tongue dorsum
SRS014126	buccal mucosa
SRS014235	stool
SRS014271	tongue dorsum
SRS014287	stool
SRS014313	stool
SRS014459	stool
SRS014464	anterior nares
SRS014470	tongue dorsum
SRS014472	buccal mucosa
SRS014476	supragingival plaque
SRS014494	posterior fornix
SRS014573	tongue dorsum
SRS014575	buccal mucosa
SRS014578	supragingival plaque
SRS014613	stool
SRS014629	posterior fornix
SRS014682	anterior nares
SRS014683	stool
SRS014684	tongue dorsum
SRS014686	buccal mucosa
SRS014690	supragingival plaque
SRS014888	tongue dorsum
SRS014890	buccal mucosa
SRS014894	supragingival plaque
SRS014901	anterior nares
SRS014923	stool
SRS014979	stool
SRS015038	tongue dorsum
SRS015040	buccal mucosa
SRS015044	supragingival plaque
SRS015051	anterior nares
SRS015054	posterior fornix
SRS015133	stool
SRS015154	buccal mucosa
SRS015158	supragingival plaque
SRS015168	posterior fornix
SRS015190	stool

SRS015209	tongue dorsum
SRS015215	supragingival plaque
SRS015217	stool
SRS015225	posterior fornix
SRS015264	stool
SRS015269	anterior nares
SRS015272	tongue dorsum
SRS015274	buccal mucosa
SRS015278	supragingival plaque
SRS015369	stool
SRS015374	buccal mucosa
SRS015378	supragingival plaque
SRS015395	tongue dorsum
SRS015425	posterior fornix
SRS015430	anterior nares
SRS015434	tongue dorsum
SRS015436	buccal mucosa
SRS015440	supragingival plaque
SRS015450	anterior nares
SRS015470	supragingival plaque
SRS015537	tongue dorsum
SRS015574	supragingival plaque
SRS015578	stool
SRS015640	anterior nares
SRS015644	tongue dorsum
SRS015646	buccal mucosa
SRS015650	supragingival plaque
SRS015663	stool
SRS015745	buccal mucosa
SRS015752	anterior nares
SRS015755	supragingival plaque
SRS015762	tongue dorsum
SRS015782	stool
SRS015893	tongue dorsum
SRS015895	buccal mucosa
SRS015899	supragingival plaque
SRS015921	buccal mucosa
SRS015937	anterior nares
SRS015941	tongue dorsum

SRS015947	supragingival plaque
SRS015960	stool
SRS015989	supragingival plaque
SRS015996	anterior nares
SRS016002	tongue dorsum
SRS016018	stool
SRS016033	anterior nares
SRS016037	tongue dorsum
SRS016039	buccal mucosa
SRS016043	supragingival plaque
SRS016056	stool
SRS016086	tongue dorsum
SRS016088	buccal mucosa
SRS016092	supragingival plaque
SRS016095	stool
SRS016111	posterior fornix
SRS016188	anterior nares
SRS016191	posterior fornix
SRS016196	buccal mucosa
SRS016200	supragingival plaque
SRS016203	stool
SRS016225	tongue dorsum
SRS016267	stool
SRS016292	anterior nares
SRS016297	buccal mucosa
SRS016319	tongue dorsum
SRS016331	supragingival plaque
SRS016335	stool
SRS016342	tongue dorsum
SRS016349	buccal mucosa
SRS016360	supragingival plaque
SRS016434	anterior nares
SRS016495	stool
SRS016501	tongue dorsum
SRS016503	buccal mucosa
SRS016513	anterior nares
SRS016516	posterior fornix
SRS016529	tongue dorsum
SRS016533	buccal mucosa

SRS016553	anterior nares
SRS016559	posterior fornix
SRS016569	tongue dorsum
SRS016575	supragingival plaque
SRS016581	anterior nares
SRS016585	stool
SRS016600	buccal mucosa
SRS016746	supragingival plaque
SRS016752	anterior nares
SRS016753	stool
SRS016954	stool
SRS016989	stool
SRS017013	buccal mucosa
SRS017025	supragingival plaque
SRS017044	anterior nares
SRS017080	buccal mucosa
SRS017103	stool
SRS017120	tongue dorsum
SRS017127	buccal mucosa
SRS017139	supragingival plaque
SRS017156	anterior nares
SRS017191	stool
SRS017209	tongue dorsum
SRS017215	buccal mucosa
SRS017227	supragingival plaque
SRS017244	anterior nares
SRS017247	stool
SRS017304	supragingival plaque
SRS017307	stool
SRS017433	stool
SRS017439	tongue dorsum
SRS017441	buccal mucosa
SRS017445	supragingival plaque
SRS017451	anterior nares
SRS017497	posterior fornix
SRS017511	supragingival plaque
SRS017520	posterior fornix
SRS017521	stool
SRS017533	tongue dorsum

SRS017537	buccal mucosa
SRS017687	buccal mucosa
SRS017691	supragingival plaque
SRS017697	anterior nares
SRS017700	posterior fornix
SRS017701	stool
SRS017713	tongue dorsum
SRS017808	tongue dorsum
SRS017810	buccal mucosa
SRS017814	supragingival plaque
SRS017820	anterior nares
SRS017821	stool
SRS018133	stool
SRS018145	tongue dorsum
SRS018149	buccal mucosa
SRS018157	supragingival plaque
SRS018300	tongue dorsum
SRS018312	anterior nares
SRS018329	buccal mucosa
SRS018337	supragingival plaque
SRS018351	stool
SRS018357	tongue dorsum
SRS018359	buccal mucosa
SRS018369	anterior nares
SRS018394	supragingival plaque
SRS018427	stool
SRS018439	tongue dorsum
SRS018463	anterior nares
SRS018573	supragingival plaque
SRS018575	stool
SRS018585	anterior nares
SRS018591	tongue dorsum
SRS018656	stool
SRS018661	buccal mucosa
SRS018665	supragingival plaque
SRS018671	anterior nares
SRS018739	tongue dorsum
SRS018769	posterior fornix
SRS018778	supragingival plaque

SRS018784	anterior nares
SRS018791	tongue dorsum
SRS018817	stool
SRS019215	anterior nares
SRS019219	tongue dorsum
SRS019221	buccal mucosa
SRS019225	supragingival plaque
SRS019267	stool
SRS019327	tongue dorsum
SRS019329	buccal mucosa
SRS019333	supragingival plaque
SRS019339	anterior nares
SRS019379	posterior fornix
SRS019386	anterior nares
SRS019387	supragingival plaque
SRS019389	tongue dorsum
SRS019391	buccal mucosa
SRS019397	stool
SRS019587	buccal mucosa
SRS019591	supragingival plaque
SRS019597	anterior nares
SRS019600	posterior fornix
SRS019601	stool
SRS019607	tongue dorsum
SRS019968	stool
SRS019974	tongue dorsum
SRS019976	buccal mucosa
SRS019980	supragingival plaque
SRS019986	anterior nares
SRS019989	posterior fornix
SRS020220	tongue dorsum
SRS020226	supragingival plaque
SRS020232	anterior nares
SRS020233	stool
SRS020328	stool
SRS020334	tongue dorsum
SRS020336	buccal mucosa
SRS020340	supragingival plaque
SRS020349	posterior fornix

SRS020386	anterior nares
SRS020856	tongue dorsum
SRS020858	buccal mucosa
SRS020862	supragingival plaque
SRS020868	anterior nares
SRS020869	stool
SRS022137	stool
SRS022143	tongue dorsum
SRS022145	buccal mucosa
SRS022149	supragingival plaque
SRS022158	posterior fornix
SRS022530	tongue dorsum
SRS022532	buccal mucosa
SRS022536	supragingival plaque
SRS022719	tongue dorsum
SRS022721	buccal mucosa
SRS022725	supragingival plaque
SRS022734	posterior fornix
SRS023346	stool
SRS023352	tongue dorsum
SRS023354	buccal mucosa
SRS023358	supragingival plaque
SRS042428	posterior fornix
SRS042457	buccal mucosa
SRS042643	tongue dorsum
SRS043001	stool
SRS043646	buccal mucosa
SRS043663	tongue dorsum
SRS043755	supragingival plaque
SRS044373	tongue dorsum
SRS045004	stool
SRS045049	buccal mucosa
SRS045254	buccal mucosa
SRS045262	buccal mucosa
SRS045313	supragingival plaque
SRS045713	stool
SRS046344	anterior nares
SRS047824	tongue dorsum
SRS048164	stool

SRS048719	buccal mucosa
SRS049389	tongue dorsum
SRS049712	stool
SRS049900	stool
SRS049959	stool
SRS050007	buccal mucosa
SRS050025	anterior nares
SRS050029	buccal mucosa
SRS050184	posterior fornix
SRS050244	tongue dorsum
SRS050628	buccal mucosa
SRS050752	stool
SRS051244	supragingival plaque
SRS051505	posterior fornix
SRS051613	anterior nares
SRS051941	supragingival plaque
SRS052227	tongue dorsum
SRS052330	posterior fornix
SRS052590	anterior nares
SRS052604	supragingival plaque
SRS052697	stool
SRS052876	supragingival plaque
SRS053335	stool
SRS053398	stool
SRS053437	anterior nares
SRS053854	tongue dorsum
SRS054061	anterior nares
SRS054590	stool
SRS054653	supragingival plaque
SRS054687	tongue dorsum
SRS054956	stool
SRS055118	buccal mucosa
SRS055401	supragingival plaque
SRS055426	tongue dorsum
SRS056323	tongue dorsum
SRS056695	posterior fornix
SRS057539	tongue dorsum
SRS057791	tongue dorsum
SRS057807	posterior fornix

SRS058053	supragingival plaque
SRS058213	anterior nares
SRS058808	supragingival plaque

155