

## Chemoproteomics of microbiota metabolites reveals small-molecule agonists for orphan receptor GPRC5A

Xiaohui Zhao<sup>1</sup>, Kathryn R. Stein<sup>1</sup>, Victor Chen<sup>3</sup>, Matthew E. Griffin<sup>1</sup>, Howard C. Hang<sup>1,2\*</sup>

<sup>1</sup> Department of Immunology and Microbiology, Scripps Research, La Jolla, California 92037, United States.

<sup>2</sup> Department of Chemistry, Scripps Research, La Jolla, California 92037, United States.

<sup>3</sup> Laboratory of Chemical Biology and Microbial Pathogenesis, The Rockefeller University, New York, New York 10065, United States.

\*Correspondence: [hhang@scripps.edu](mailto:hhang@scripps.edu)

### ABSTRACT

The microbiota generates diverse metabolites that can engage multiple pathways to modulate host physiology and disease, but their protein targets and mechanism(s) of action have not been fully elucidated. To address this challenge, we focused on indole-3-acetic acid (IAA), a prominent microbiota metabolite, and developed IAA-based chemical reporters for proteomic studies. We discovered that IAA interacts with many proteins in host cells, including small-molecule transporters, receptors and metabolic enzymes. Notably, our functional studies revealed that IAA binds to orphan G protein-coupled receptors such as GPRC5A, but only aromatic monoamines were capable of inducing GPRC5A signaling. Functional profiling of microbiota uncovered specific bacterial species and enzymes that generate GPRC5A agonists. Finally, biochemical characterization of GPRC5A activation identified more potent synthetic agonists as well as key amino acid residues involved in ligand binding. These studies highlight the utility of chemoproteomics to dissect protein targets and mechanisms of action for microbiota metabolites.

## INTRODUCTION

The human microbiota directly modulate host immunity and generate diverse metabolites that engage multiple pathways to regulate host physiology.<sup>1-3</sup> For example, microbe-associated molecular patterns (MAMPs) such as lipopolysaccharides (LPS) can engage toll-like receptors (TLRs) to maintain intestinal epithelial homeostasis<sup>4</sup> and modulate infant autoimmunity<sup>5</sup>. In addition, microbiota-derived cyclic dinucleotides<sup>6</sup> and peptidoglycan fragments<sup>7,8</sup> initiate host immune response and enhance the efficacy of immune checkpoint inhibitor for cancer therapy<sup>9,10</sup>. Beyond generating MAMPs that engage canonical pattern recognition receptors (PRRs), microbiota metabolism of dietary components (e.g., polysaccharides, amino acids and lipids)<sup>11</sup>, host metabolites (e.g. bile acids)<sup>12</sup> and therapeutic drugs<sup>13</sup> can produce diverse molecules in micromolar to millimolar levels, which have local and systemic effects on host health and disease. Notably, aromatic amino acids (aAAs) are transformed by microbiota to diverse metabolites, such as aromatic hydrocarbons, acids and amines.<sup>14,15</sup> These microbiota-derived aromatic compounds have been suggested to engage different classes of protein targets within host cells, including nuclear receptors<sup>16,17</sup> and G protein-coupled receptors (GPCRs)<sup>18,19</sup>, but their other potential protein targets and mechanisms of action have not been fully elucidated.

Advances in chemical proteomics provide an excellent opportunity to characterize protein targets of microbiota metabolites and elucidate their mechanisms of action in host pathways.<sup>20-22</sup> The derivatization of microbiota metabolites with bioorthogonal tags for detection and affinity enrichment as well as photoaffinity moieties for crosslinking enables identification of both covalently and non-covalently labeled proteins for subsequent functional studies.<sup>23,24</sup> To showcase the utility of chemoproteomics for microbiota metabolite studies, we developed metabolic (alk-IAA) and photoaffinity (x-alk-IAA) reporters of indole-3-acetic acid (IAA), a prominent microbiota-generated metabolite<sup>25,26</sup> that exhibits significant effects on host physiology through engaging multiple pathways.<sup>27,28</sup> From chemoproteomic profiling of IAA-interacting proteins in intestinal epithelial cells, we identified many candidate protein targets including small-molecule transporters, immune sensors and GPCRs, highlighting the utility of chemoproteomics to reveal potentially novel mechanisms of action for microbiota metabolites.

Our functional studies revealed that the microbiota generate small-molecule aromatic monoamine agonists for orphan receptor GPRC5A. Specific microbiota species and enzymes, such as *Enterococcus* species that encode tyrosine decarboxylases, are capable of activating GPRC5A. In addition, our chemical and biochemical characterization of GPRC5A activation yielded potent synthetic agonists such as 7-fluorotryptamine (7FTA) as well as key amino acid residues for ligand binding. Given the orphan status of GPRC5A and its association with cancer<sup>29-31</sup> and immunity<sup>32</sup>, these studies showcase how chemoproteomic analysis of microbiota metabolites enables the identification of unpredicted protein targets such as orphan GPCRs and reveals novel mechanism(s) of action between specific microbiota species and metabolic enzymes, microbiota metabolites, and host pathways for potential therapeutic development.

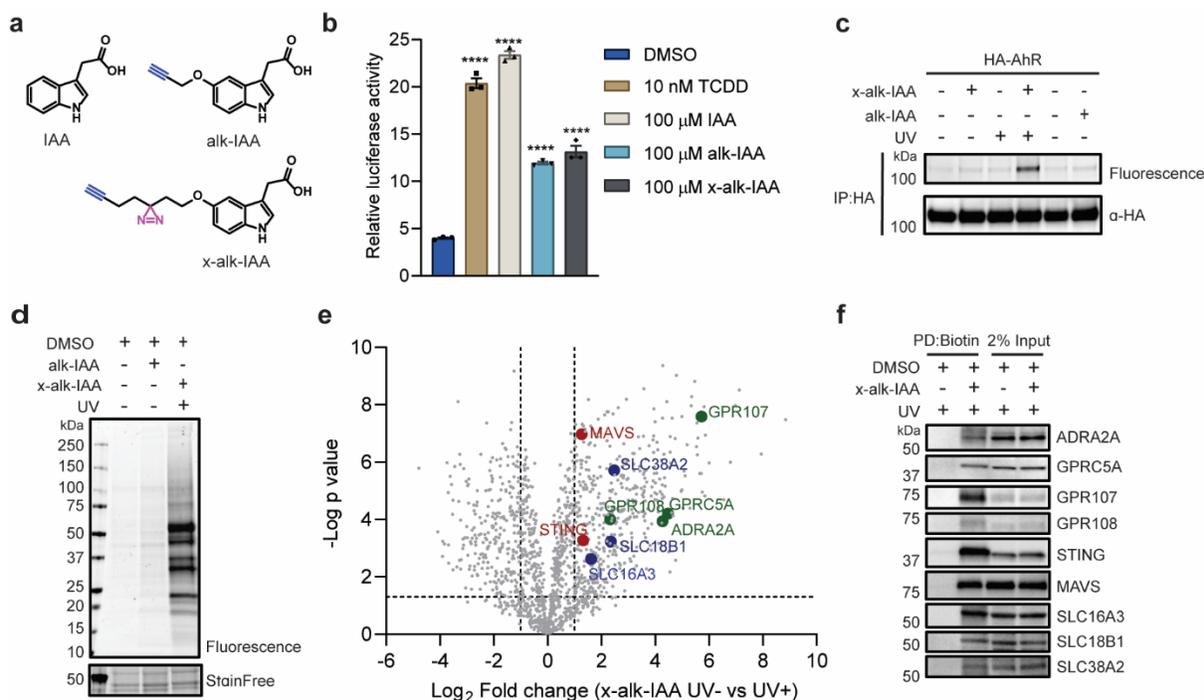
## RESULTS

**Chemoproteomics reveals IAA-interacting proteins.** To investigate the protein targets of microbiota metabolites, we initially focused on IAA and generated a metabolic reporter functionalized with an alkyne chemical tag (alk-IAA) as well as a photoaffinity reporter that also contains a diazirine crosslinker (x-alk-IAA) (**Fig. 1a**). To examine whether these chemical reporters function as active analogs of IAA, we tested them on aryl hydrocarbon receptor (AhR), a previously described IAA protein target,<sup>16,33</sup> using the xenobiotic response element (XRE)-driven luciferase reporter assay in HEK293T cells. Both alk-IAA and x-alk-IAA were able to activate AhR, although less efficient than IAA and a potent ligand 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (**Fig. 1b**). UV irradiation of the chemical reporter treated cells revealed that x-alk-IAA, but not alk-IAA, can photocrosslink HA-tagged AhR, which was immunoprecipitated, reacted with azide-rhodamine and analyzed by in-gel fluorescence profiling (**Fig. 1c**). Cells treated with DMSO, alk-IAA or x-alk-IAA without UV irradiation did not label HA-AhR. Taken together, these experiments indicate that our IAA chemical reporters function as reasonable analogs of IAA in mammalian cells.

We next applied the IAA-based chemical reporters to HT-29 cells, a human colorectal adenocarcinoma epithelial cell line, to identify potential microbiota metabolite protein targets using chemoproteomics (**Supplementary Fig. 1a**). The alk-IAA reporter may be metabolically processed in cells and covalently incorporated onto proteins, while the photoaffinity reporter x-alk-IAA enables the photocrosslinking of non-covalent protein targets with UV. HT-29 cells were thus incubated with these IAA chemical reporters with or without UV treatment and total cell lysates were reacted with azide-modified reagents for in-gel fluorescence profiling or label-free quantitative proteomics. HT-29 cells treated with alk-IAA revealed a few specifically labeled proteins compared to DMSO treated samples (**Fig. 1d**). In contrast, we observed many labeled proteins between 20 to 75 kDa with x-alk-IAA and UV irradiation compared to controls (**Fig. 1d**). Labeling of x-alk-IAA proteins correlated with increasing concentration of x-alk-IAA (**Supplementary Fig. 1b**). We therefore reacted the cell lysates with azide-biotin, performed neutravidin affinity enrichment and analyzed the recovered proteins by label-free quantitative proteomics. A total number of 413 proteins were detected in x-alk-IAA and UV-treated samples in comparison to the without UV control (p value < 0.05, two folds change in difference), visualized by volcano plot (**Fig. 1e, Supplementary Table 1**). The majority of these proteins were additionally found when comparing x-alk-IAA and UV-treated samples versus a DMSO control (**Supplementary Fig. 1c,d, Supplementary Table 1**). Notably, the 18 significant proteins recovered from alk-IAA metabolic labeling were also identified with the x-alk-IAA (**Supplementary Fig. 1d, Supplementary Table 1**). Gene ontology (GO) analysis of the hits revealed many cytosolic metabolic enzymes as well as membrane proteins that localize on organelle membranes including the endoplasmic reticulum (ER), mitochondria, Golgi apparatus and lysosome (**Supplementary Fig. 1e**). KEGG pathway analysis of these hits revealed proteins associated with steroid and *N*-glycan biosynthesis, valine, leucine and isoleucine degradation, lipid (e.g., glycerolipid, glycerophospholipid, sphingolipid, cholesterol) and amino acids (e.g., tryptophan and histidine) metabolism (**Supplementary Fig. 1f**). In addition to these metabolic enzymes,

we identified proteins involved in different functional classes that include small-molecule transporters (SLC16A3, SLC18B1 and SLC38A2), innate immunity effectors (STING and MAVS) and GPCRs (ADRA2A, GPRC5A, GPR107 and GPR108) (**Fig. 1g**). The x-alk-IAA photocrosslinking of these proteins were validated by western blot analysis using specific antibodies (**Fig. 1h**). The recovery of STING with x-alk-IAA is consistent with prior reports, which demonstrate that this innate immune sensor can interact with aromatic compounds containing a free acid moiety.<sup>34,35</sup> Conversely, x-alk-IAA photocrosslinking to GPRC5A, GPR107 and GPR108 was striking, as no small molecule ligands were previously reported for these orphan or putative GPCRs, which highlights the utility of chemoproteomics to identify unpredicted protein targets.

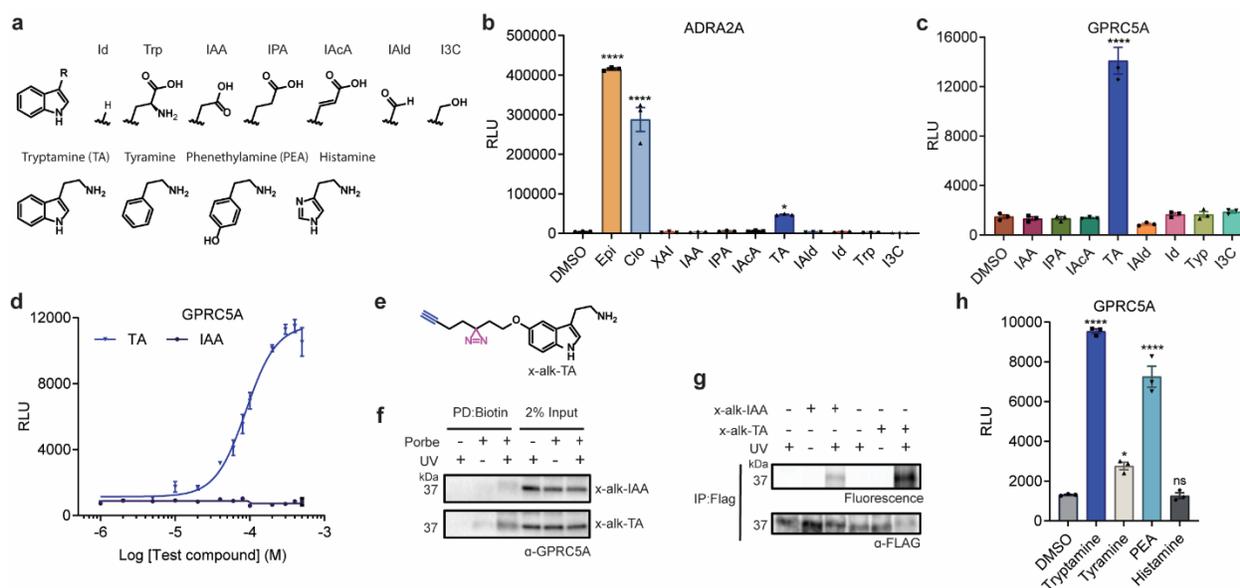
The GPCRs crosslinked by x-alk-IAA are interesting candidate microbiota metabolite protein targets. Alpha-2A adrenergic receptor (ADRA2A) responds to the endogenous ligands including epinephrine and norepinephrine to transduce cellular signaling in the sympathetic nervous system<sup>36</sup> and has been suggested to sense the microbiota-derived metabolite, phenylacetylglutamine<sup>37</sup>. In contrast, GPR107 and GPR108 consist of a 7-transmembrane helix and mainly localize on the *trans*-Golgi network,<sup>38,39</sup> but no ligands have been reported and functions as GPCRs are unclear.<sup>40</sup> GPRC5A is a retinoic acid-induced protein [also termed retinoic acid-induced gene 3 (RAI3) or retinoic acid-induced gene 1 (RAIG1)] that is highly expressed in the lung and gastrointestinal tract and associated with inflammation and cancer progression.<sup>41</sup> GPRC5A mainly localizes to the plasma membrane and endoplasmic reticulum (ER) and is categorized as a class C orphan GPCR. GPRC5A has a short N-terminus that is atypical of class C GPCRs, which usually contain long extracellular N-termini as a ligand-binding domain and a cysteine-rich domain<sup>42</sup>. As GPRC5A is implicated as an immune-associated tumor suppressor with no reported ligands, we investigated whether specific microbiota species and metabolites may function as agonists of this orphan GPCR.



**Fig. 1 Chemoproteomic analysis of indole-3-acetic acid-interacting proteins in mammalian cells. a**, Chemical reporters based on indole-3-acetic acid (IAA) for metabolic labeling (alk-IAA) and photoaffinity profiling (x-alk-IAA) of IAA-interacting proteins. **b**, Activity of TCDD, IAA, alk-IAA and x-alk-IAA on aryl hydrocarbon receptor (AhR) and xenobiotic response element (XRE)-luciferase reporter in HEK293T cells. **c**, Photoaffinity crosslinking of x-alk-IAA (100 μM) to HA-tagged AhR in HEK293T cells revealed by rhodamine labeling. **d**, In-gel fluorescent profiling of IAA-interacting proteins in HT-29 cells using 0.1% DMSO, 100 μM alk-IAA or 100 μM x-alk-IAA. **e**, Volcano plot of the identified proteins from HT-29 cells by chemical proteomics using x-alk-IAA (100 μM) and UV irradiation in comparison with samples without UV. p value is 0.05, two folds change in four replicates. Key hits are highlighted in GPCRs (green), small-molecule transporters (blue) and immunity-associated proteins (red). **f**, Western blot analysis of x-alk-IAA (100 μM) photo-crosslinking proteins. Input indicates 2% of total HT-29 cell lysates. Data indicates mean with SEM from three replicates. One-way ANOVA using Dunnett's multiple comparisons test. \*\*\*\* indicates p-value < 0.0001.

**Aromatic monoamines activate GPRC5A signaling.** To explore whether IAA and indole-derived metabolites modulate GPCR activity of the identified proteins, we utilized PRESTO-Tango<sup>43</sup>, a cell-based reporter assay that relies on β-arrestin 2 recruitment and transcriptional activation. We tested IAA and a collection of indole molecules, including indole (Id), tryptophan (Trp), indole propionic acid (IPA), indole acrylic acid (IAcA), tryptamine (TA), indole aldehyde (IAld), indole-3-carbinol (I3C), and x-alk-IAA photoaffinity reporter (XAI) (**Fig. 2a**). For ADRA2A, TA exhibits weak effects in comparison with its endogenous ligand epinephrine (Epi) and synthetic agonist clonidine (Clo) (**Fig. 2b**). Notably, amongst the collection of indole molecules screened, TA activated GPRC5A most effectively (**Fig. 2c**) with an EC<sub>50</sub> of 88.9 μM (**Fig. 2d**). In addition, TA competes with x-alk-IAA photocrosslinking to GPRC5A and ADRA2A (**Supplementary Fig. 2a,b**). Interestingly, IAA also competed with x-alk-IAA photocrosslinking (**Supplementary Fig. 2a,b**) but did not activate GPRC5A (**Fig. 2c,d**). These results suggest that IAA may bind to GPRC5A with low-affinity, but is incapable of inducing downstream signaling compared to TA. To further investigate TA interaction

with proteins, we synthesized a TA photoaffinity reporter x-alk-TA (**Fig. 2e**) and found that it more effectively photocrosslinked endogenous GPRC5A in HT-29 cells (**Fig. 2f**) and FLAG-tagged GPRC5A expressed in HEK293T cells in comparison to x-alk-IAA (**Fig. 2g**), suggesting TA may bind GPRC5A stronger than IAA. These data suggest that additional microbiota aromatic monoamines may also activate GPRC5A signaling. We therefore evaluated aromatic monoamines such as phenethylamine (PEA), tyramine and histamine. PEA showed comparable activity to TA at 100  $\mu$ M, whereas tyramine and histamine only exhibited minimal or no activity on GPRC5A signaling (**Fig. 2h**). These results demonstrate that aromatic metabolites bind to GPRC5A, but only aromatic amines can effectively induce signaling, which is the first indication for small-molecule agonists of GPRC5A and deorphanizes this class C GPCR.



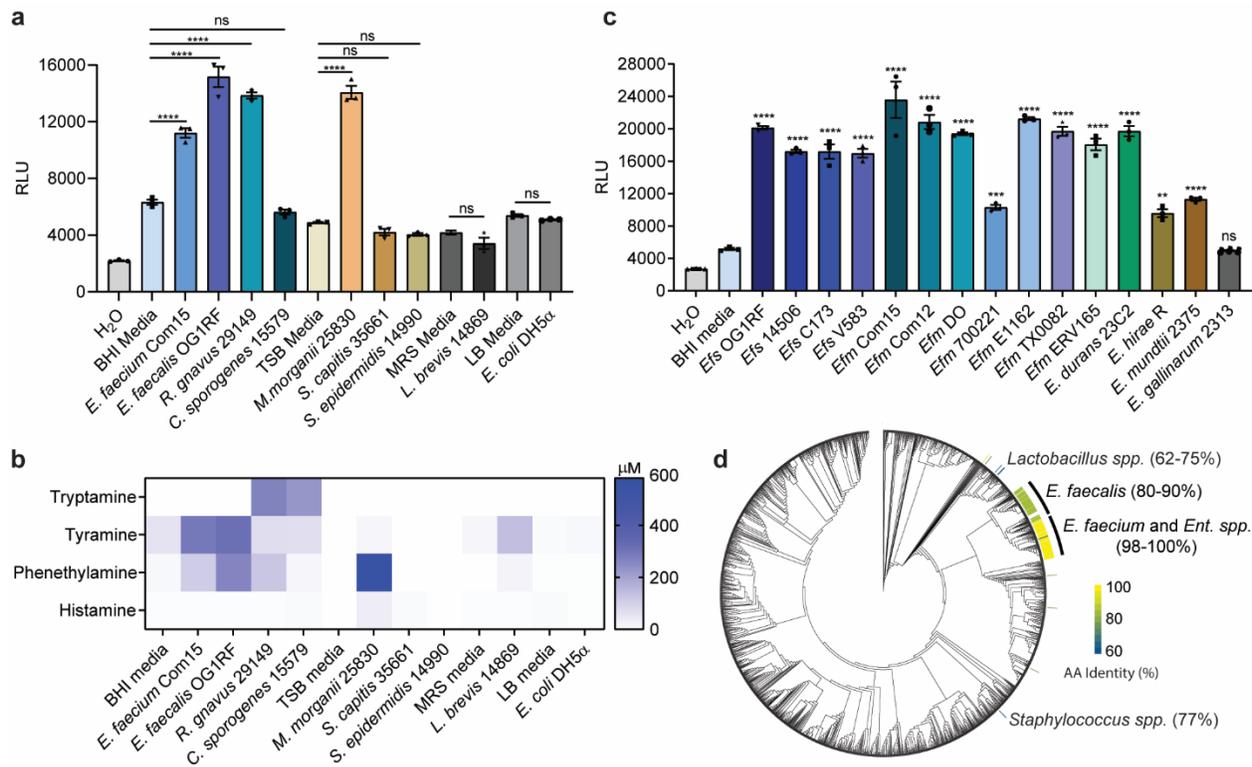
**Fig. 2 Aromatic monoamines activate GPRC5A signaling.** **a**, Chemical structures of indole derivatives and aromatic monoamines. **b,c**, PRESTO-Tango assay for screening a collection of indole derivatives (100  $\mu$ M) on ADRA2A (**b**) and GPRC5A (**c**) activation. 0.1% DMSO, 1  $\mu$ M epinephrine (Epi), 1  $\mu$ M clonidine were used as controls. **d**, Dose-response curves of GPRC5A activation upon IAA and TA treatment. **e**, Chemical structure of TA photoaffinity reporter x-alk-TA. **f**, Pull-down of endogenous GPRC5A from HT-29 cells that were photo-crosslinked by x-alk-IAA or x-alk-TA followed by biotin labeling and neutravidin enrichment. **g**, Photoaffinity crosslinking of x-alk-IAA or x-alk-TA to FLAG-tagged GPRC5A in HEK293T cells identified with rhodamine labeling. **h**, PRESTO-Tango assay for examine aromatic monoamines (100  $\mu$ M) in GPRC5A activation. Data indicates mean with SEM from three replicates. One-way ANOVA using Dunnett's multiple comparisons test. \*\*\*\* $p$ <0.0001, \* $p$ <0.05, ns indicates not significant.

### Microbiota species containing aAAs decarboxylases activate GPRC5A signaling.

We next explored GPRC5A activating microbiota species and enzymes that could generate aromatic monoamines. TA, PEA and tyramine are aromatic monoamines produced by the metabolic processing of aAAs by bacteria and human. For example, two common Firmicutes, *Ruminococcus gnavus* and *Clostridium sporogenes* converts Trp to TA through reductive metabolism by the Trp decarboxylase (TrpDC).<sup>44</sup> Alternatively, PEA is known as a product of bacterial species including *Enterococcus faecalis*<sup>45</sup> and *Morganella morganii*<sup>19</sup>, whereas tyramine has been identified in the

growth media of different *Enterococcus* and *Lactobacillus* species.<sup>45,46</sup> To identify microbiota species that can activate GPRC5A, we examined the growth media of different bacterial species and strains predicted to metabolize aAAs into their corresponding monoamines. The bacterial growth media were subjected to 3 kDa MW-cutoff membrane filtration to remove secreted proteins and retain low molecular weight metabolites. Amongst the bacterial species we initially screened, 3 kDa MW-cutoff membrane filtered growth media from *R. gnavus*, *M. morgani*, *E. faecium* and *E. faecalis* effectively activated GPRC5A (**Fig. 3a**). However, the membrane filtered growth media from *C. sporogenes* and *E. coli* exhibited no GPRC5A activation beyond their control growth media alone (**Fig. 3a**). To determine whether these bacterial species generate aromatic monoamines, we performed targeted LC-MS analysis of the 3 kDa MW-cutoff membrane filtered growth media (**Fig. 3b**). Indeed, *R. gnavus*, *E. faecium* and *E. faecalis* generate high micromolar levels of GPRC5A aromatic monoamine agonists (**Fig. 3b**). Instead of generating different aromatic monoamines, *M. morgani* primarily produced PEAs at higher levels ~500  $\mu$ M (**Fig. 3b**). *C. sporogenes* also generated aromatic monoamines, but at lower micromolar levels (**Fig. 3b**), which may not be sufficient to activate GPRC5A signaling in the PRESTO-Tango assay (**Fig. 3a**). *E. coli* did not produce significant levels of aromatic monoamines that were detectable by our fractionation and LC-MS conditions (**Fig. 3b**). Based on the *E. faecium* and *E. faecalis* results and the importance of *Enterococcus*-host interactions, we evaluated additional *Enterococcus* strains and species, which revealed that *E. faecium* and *E. faecalis* strains as well as *E. durans*, *E. hirae*, and *E. mundtii* membrane filtered growth media activated GPRC5A (**Fig. 3c**). However, the growth media from *E. gallinarum* exhibited no significant effect on GPRC5A activation (**Fig. 3c**).

To further investigate GPRC5A activating microbiota species, we performed a bioinformatic searching and phylogenetic analysis of bacterial amino acid decarboxylases. While TrpDC activity has been reported for *R. gnavus*<sup>44</sup>, no clear orthologs greater than 60% protein sequence similarity of this enzyme are present in *Enterococcus*. However, *E. faecium* contains a predicted Tyr decarboxylase (TyrDC) (**Supplementary Fig. 3**) that is highly conserved in the strains of *E. faecium*, *E. faecalis*, and other *Enterococcus* species identified and sequenced from the Human Microbiome Project (HMP)<sup>47</sup> (**Fig. 3d, Supplementary Table 2**). *E. gallinarum* strain does not appear to have a TyrDC ortholog, but contains a Gln decarboxylase (**Supplementary Fig. 3**) that may explain its inability to activate GPRC5A (**Fig. 3c**). Other microbiota species such as *Staphylococcus* and *Lactobacillus* encode proteins with more than 60% protein similarity to *E. faecium* TyrDC. However, the 3 kDa MW-cutoff membrane filtered growth media from *S. capitis*, *S. epidermidis* and *L. brevis* did not induce GPRC5A activation (**Fig. 3c**) and only produced minimal levels of aromatic monoamines or undetectable levels under the conditions we evaluated (**Fig. 3b**). These results indicated that microbiota species which express aAAs decarboxylases and generate high micromolar levels of aromatic monoamines (*R. gnavus*, *M. morgani* and *Enterococcus* spp.) function as GPRC5A agonists. Different bacterial species that contain decarboxylases may exhibit lower activity, distinct specificity and/or require unique conditions to generate sufficient levels of aromatic monoamines to activate GPRC5A.

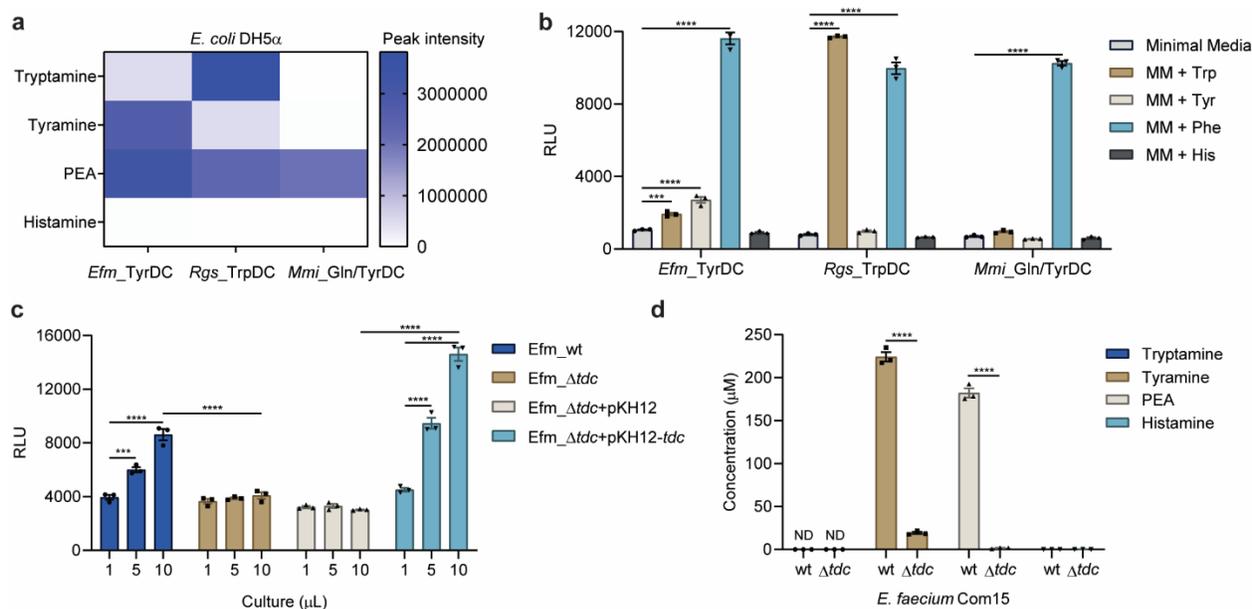


**Fig. 3 Specific microbiota species and strains produce aromatic monoamines that activate GPRC5A signaling.** **a**, GPRC5A activation by PRESTO-Tango analysis of bacterial cultures of representative human microbiota species. **b**, Heat map of aromatic monoamines concentration in different bacterial cultures determined by LC-MS analysis. **c**, GPRC5A activation by PRESTO-Tango analysis of different *Enterococcus* species and strains. **d**, Cladogram of Human Microbiome Project isolates organized by 16S rRNA homology with heatmap indicating amino acid (AA) sequence identity of putative tyrosine decarboxylase orthologs in *Enterococcus*, *Lactobacillus*, and *Staphylococcus* species. *Ent. spp.* indicates the *Enterococcus* isolates without a specific species annotation. Data indicates mean with SEM from three replicates. One-way ANOVA using Tukey's multiple comparisons test in comparison with respective media (a) or Dunnett's multiple comparisons test in comparison with BHI media (c). \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , ns indicates not significant.

**Bacterial aromatic amino acid decarboxylases generate GPRC5A agonists.** To determine the specific aAAs decarboxylases that are involved in the production of aromatic monoamines and activation of GPRC5A, we performed gain- and loss-of-expression in bacteria as well as *in vitro* activity studies. For these studies, we expressed the *E. faecium* tyrosine decarboxylase (*Efm\_TyrDC*), *R. gnavus* tryptophan decarboxylase (*Rgs\_TrpDC*) and *M. morgani* glutamate or tyrosine decarboxylase (*Mmi\_Gln/TyrDC*) in *E. coli* DH5 $\alpha$  (**Supplementary Fig. 4a,b**), examined the 3 kDa MW-cutoff membrane filtered growth media for aromatic monoamines by LC-MS (**Supplementary Fig. 4c-n**) and determined the activation of GPRCA by the PRESTO-Tango assay (**Fig. 4a**). We found that all three enzymes did not generate histamine, but effectively decarboxylated Phe to PEA (**Fig. 4a, Supplementary Fig. 4d,h,i**) and activated GPRCA signaling (**Fig. 4b**). As previously described,<sup>44</sup> *Rgs\_TrpDC* was able to catalyze the decarboxylation of Trp to generate TA (**Fig. 4a, Supplementary Fig. 4i**)

and in our assays activate GPRCA signaling (**Fig. 4b**). Mmi\_Gln/TyrDC on the other hand was specific for Phe (**Fig. 4a, Supplementary Fig. 4l**). Efm\_TyrDC generated high levels of tyramine (**Fig. 4a, Supplementary Fig. 4a**), but has only a mild effect on GPRC5A signaling (**Fig. 4b**), consistent with our data with tyramine alone (**Fig. 2e**). Efm\_TyrDC also generated low levels of TA from Trp (**Fig. 4a, Supplementary Fig. 4e**), which displayed modest activation of GPRC5A (**Fig. 4b**), suggesting this decarboxylase is selective for Tyr and Phe over Trp. Indeed, our analysis of recombinant and purified His6-tagged Efm\_TyrDC (**Supplementary Fig. 5a**) demonstrated that this *Enterococcus* aAAs decarboxylase preferentially utilizes Tyr over Phe, but not Trp, as a substrate under conditions we tested (**Supplementary Fig. 5b,c**). Importantly, the reaction products from Efm\_TyrDC *in vitro* decarboxylation of Phe and Tyr also activated GPRC5A (**Supplementary Fig. 5d**). Kinetic analysis of Efm\_TyrDC activity revealed a  $V_{\max}$  of 0.7483  $\mu\text{M/s}$  and a  $K_M$  value of 280.7  $\mu\text{M}$  with Tyr and  $V_{\max}$  value of 0.04 mM/min and  $K_M$  of 8.5 mM with Phe (**Supplementary Fig. 5e**). These results demonstrate bacterial decarboxylases from diverse microbiota species generate aromatic amines that activate GPRC5A signaling, confirming previous enzymatic activity for Rgs\_TrpDC and predicted activity of Mmi\_Gln/TyrDC and Efm\_TyrDC.

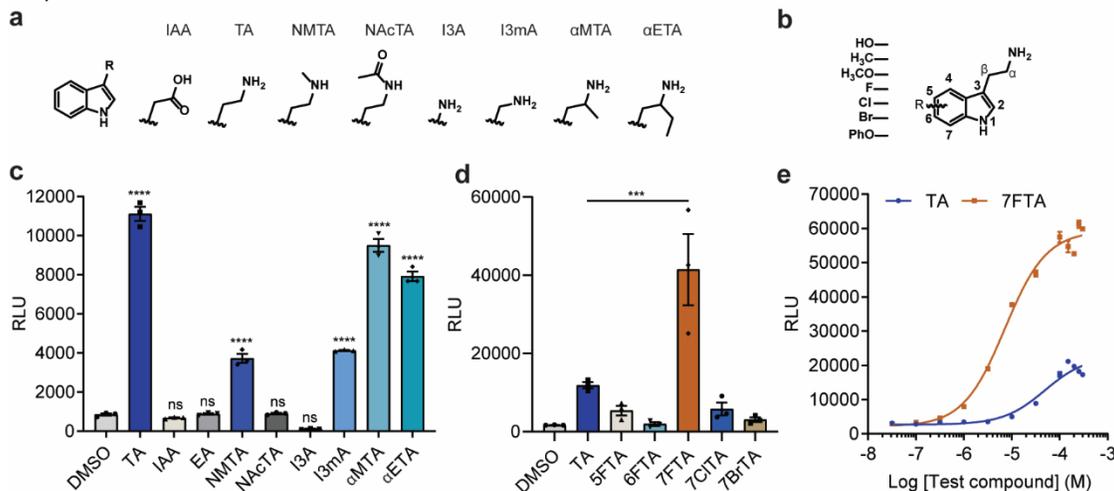
To determine if Efm\_TyrDC is also essential or required to generate aromatic amines in *Enterococcus*, we generated a *E. faecium* Com15 *tyrDC* deletion strain (*Efm\_ΔtyrDC*) using recently developed CRISPR-recombineering methods (**Supplementary Fig. 6a,b**)<sup>48</sup>. Deletion of *TyrDC* did not affect *E. faecium* cell growth (**Supplementary Fig. 5c**), but abolished the effect on GPRC5A activation (**Fig. 4c**) as well as tyramine and PEA levels in the growth media (**Fig. 4d**). Complementation of *Efm\_ΔtyrDC* strain with *TyrDC* by overexpression restored and enhanced GPRC5A activation (**Fig. 4c, Supplementary Fig. 6d**). The results demonstrate *TyrDC* is a key aromatic amino decarboxylase in *E. faecium* Com15, which is highly conserved in many *E. faecium* strains and *Enterococcus* species (**Fig. 3d**) and is capable of generating aromatic monoamines to activate GPRC5A.



**Fig. 4 Bacterial decarboxylases generate aromatic monoamines that activate GPRC5A signaling.** **a**, LC-MS determination of aromatic monoamines produced by overexpression of *E. faecium* Com15 tyrosine decarboxylase (*Efm\_TyrDC*), *R. gnavus* tryptophan decarboxylase (*Rgs\_TrpDC*) and *M. morganii* glutamate/tyrosine decarboxylase (*Mmi\_Gln/TyrDC*) in *E. coli* DH5 $\alpha$ . **b**, PRESTO-Tango assay for comparing bacterial enzyme products on GPRC5A activation. **c**, PRESTO-Tango assay for examining wild-type *E. faecium* Com15 (*Efm\_wt*) and the *E. faecium* Com15 *TyrDC* knock-out (*Efm\_Δtdc*) cultures with *TyrDC* (pKH12-*tdc*) complementation or vector control (pKH12) on GPRC5A activation. **d**, LC-MS determination of aromatic monoamines in *Efm\_wt* and *Efm\_Δtdc* cultures. Data indicates mean with SEM from three replicates. Two-way ANOVA using Tukey's multiple comparisons test. \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$ , ND indicates not detected.

**GPRC5A preferentially senses aromatic amines and is more effectively activated by 7-fluorotryptamine.** To classify the key determinants of aromatic monoamines and identify more potent GPRC5A agonists, we investigated additional natural and synthetic compounds (Fig. 5a,b). We found that the aromatic moiety is essential because ethylamine (EA) was not able to activate GPRC5A (Fig. 5c). We next observed decreased and complete loss of the GPRC5A activation with *N*-methyl tryptamine (NMTA) and *N*-acetyl tryptamine (NAcTA), respectively (Fig. 5c). These data together with the IAA results indicate that aromatic primary amines are the minimal pharmacophore for GPRC5A activation. We therefore tested additional indole-based primary amines, including indole-3-amine (I3A), indole-3-methylamine (I3mA),  $\alpha$ -methyltryptamine ( $\alpha$ MTA) and  $\alpha$ -ethyltryptamine ( $\alpha$ ETA) (Fig. 5a). Three compounds (i.e., I3mA,  $\alpha$ MTA, and  $\alpha$ ETA) retained activity towards GPRC5A (Fig. 5c), demonstrating the tolerance of truncation and branching of the alkyl amine moiety. We then evaluated a collection of tryptamine derivatives bearing various substitutions (methyl, hydroxyl, methoxy, halogen) on distinct positions of the indole ring (Fig. 5b). Notably, 7-fluorotryptamine (7FTA) was ~5-fold more effective at inducing GPRC5A signaling than TA at 100  $\mu$ M (Fig. 5d) and more active than other TA derivatives (Fig. 5d, Supplementary Fig. 7a), with an EC<sub>50</sub> value of 7.2  $\mu$ M (Fig. 5e). Concentrations of

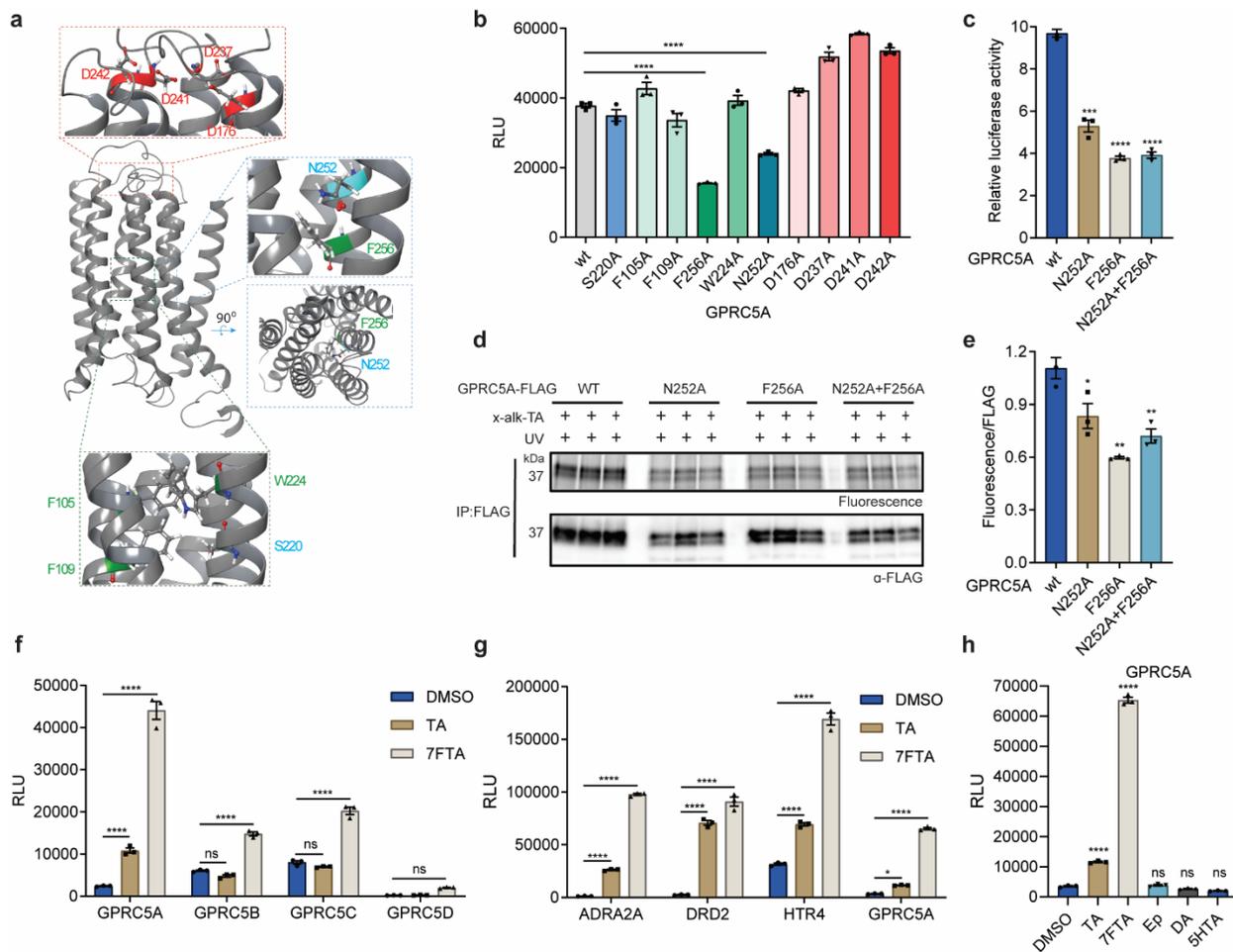
TA and 7FTA below 100  $\mu$ M exhibited no toxicity for mammalian cells (**Supplementary Fig. 7b**).



**Fig. 5 Structure-activity relationship studies of GPRC5A ligands.** **a,b**, Chemical structures of indole and tryptamine derivatives. **c-d**, PRESTO-Tango assay for analysis of various indole and tryptamine derivatives (100  $\mu$ M) on GPRC5A activation. **e**, Dose-response curves of GPRC5A activation upon TA and 7FTA treatment. Data indicates mean with SEM from three replicates. One-way ANOVA using Tukey's multiple comparisons test. \*\*\*\* $p$ <0.0001, \*\*\* $p$ <0.001, ns indicates not significant.

**Aromatic amines bind conserved transmembrane domain of retinoic acid-induced class C orphan GPCRs.** To understand how GPRC5A senses aromatic monoamines, we performed site-directed mutagenesis and photocrosslinking studies to determine the potential ligand binding sites. While the molecular structure of GPRC5A has not been determined, RosettaGPCR<sup>49</sup>, AlphaFold<sup>50</sup> and Robetta<sup>51</sup> provide structural models of its 7-helix transmembrane domain (**Fig. 6a**, **Supplementary Fig. 8a**). Based on the knowledge of ligand binding to aminergic GPCRs and class C GPCRs, we mutated a predicted extracellular patch of acidic amino acid residues as well as a polar and aromatic pocket within the transmembrane domain (**Fig. 6a**). Using the PRESTO-Tango assay upon 7FTA treatment, we found that the extracellular patch of aspartic acids in GPRC5A are not involved in the interaction with aromatic monoamines (**Fig. 6b**). Mutations of N252 and F256 respectively to alanine resulted in decreased GPRC5A activation with 7FTA, which indicates that these two amino acids within a polar and aromatic pocket of transmembrane domain may be crucial for ligand binding (**Fig. 6b**). Alanine mutations of F105, F109, S220 and W224 had no significant effect on 7FTA activation of GPRC5A (**Fig. 6b**). We then generated the N252A and F256A double mutants and found that both single and double mutations decreased the 7FTA mediated GPRC5A activation (**Fig. 6c**) and photoaffinity labeling (**Fig. 6d,e**). These results suggest N252 and F256 are key amino acid residues involved in GPRC5A ligand binding. Nonetheless, these amino acid residues are conserved amongst retinoic acid-induced class C orphan GPCRs (GPRC5A, GPRC5B, GPRC5C and GPRC5D) (**Supplementary Fig. 8b**), which are all activated by 7FTA (**Fig. 6f**). Moreover, both TA and 7FTA function as agonists of GPCRs sensing aromatic monoamines, such as ADRA2A, DRD2 and HTR4 (**Fig. 6g**). However, GPRC5A did not respond to endogenous neurotransmitters like epinephrine, dopamine and serotonin (**Fig. 6h**),

indicating distinct mechanisms and functions of metabolite sensing compared to several canonical GPCRs. Collectively, our results demonstrate that aromatic amines can bind a conserved transmembrane domain of GPRC5A and activate other retinoic acid-induced class C GPCRs.



**Fig. 6 Analysis of aromatic monoamine-GPRC5A interactions.** **a**, Molecular model of GPRC5A made from RosettaGPCR. **b-c**, PRESTO-Tango assay for mutagenesis screening of key residues in GPRC5A-7FTA (100  $\mu$ M) interactions. **d-e**, Photoaffinity crosslinking of x-alk-TA to FLAG-tagged GPRC5A mutants in HEK293T cells revealed by rhodamine labeling. **f-g**, Evaluation of 100  $\mu$ M of TA and 7FTA on class C orphan GPCRs (GPRC5A,B,C,D) and aminergic GPCRs (ADRA2A, DRD2, HTR4) activation. **h**, Examination of endogenous neurotransmitters epinephrine (10  $\mu$ M), dopamine (10  $\mu$ M) and serotonin (100  $\mu$ M) on GPRC5A activation. Data indicates mean with SEM from three replicates. One-way ANOVA using Dunnett's multiple comparisons test (b,c,e,h) or Two-way ANOVA using Tukey's multiple comparisons test (f,g). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , ns indicates not significant.

## DISCUSSION

The human microbiota and their metabolites are important modulators of host physiology. Understanding the molecular basis by which specific microbiota species and metabolites exert their mechanisms of action is therefore crucial for dissecting host-microbe interactions. For example, aAAs are metabolized by specific microbiota into

diverse molecules that play pivotal roles in host health and disease.<sup>14,52</sup> Desaminotyrosine (DAT) produced by *Clostridium orbiscindens* protects mice from influenza virus challenge through augmentation of type I IFN signaling and suppression of lung immunopathology.<sup>53</sup> Alternatively, *Lactobacillus reuteri* generates indole-3-lactic acid (ILA) and IAld that activates AhR and induces intraepithelial CD4<sup>+</sup> CD8 $\alpha$ <sup>+</sup> T cells to maintain intestinal homeostasis.<sup>54</sup> Moreover, indole-3-ethanol, indole-3-pyruvate and IAld can regulate gut barrier function through AhR.<sup>55</sup> Indole metabolites such as IAA and IPA have also been reported to activate pregnane X receptor (PXR) and increases tight junction protein expression and downregulate the expression of pro-inflammatory cytokines.<sup>17</sup> Beyond modulating intestinal homeostasis, *C. sporogenes* has been reported to metabolize tryptophan and generate high levels of indole metabolites that can be detected in serum and modulate systemic immune responses.<sup>15,56</sup> While certain protein targets and receptors have been reported for these indole metabolites, the relatively high levels (micro-millimolar levels) of these metabolites in the intestinal lumen and circulation, suggests they may engage multiple protein targets in host cells.

To address this general challenge of investigating microbiota metabolite mechanisms, we employed chemical proteomics and explored the protein targets of IAA, a prominent microbiota metabolite<sup>14,15</sup> with potentially multiple mechanisms of action<sup>16,17</sup>. Through the generation of metabolic and photoaffinity IAA-based chemical reporters, we identified many indole-metabolite interacting proteins. We focused on GPRC5A and discovered that this previously annotated class C orphan GPCR is able to bind indole metabolites, but is only activated by aromatic monoamines such as TA, PEA or tyramine. These exciting discoveries reveal the first small molecule agonists of GPRC5A and its related class C GPCR homologs (GPRC5B, GPRC5C and GPRC5D). Moreover, this led us to identify and characterize aromatic monoamine-producing microbiota species (*R. gnavus*, *M. morgani* and *Enterococcus* species) and their associated decarboxylases that can generate aromatic amines and activate GPRC5A signaling (**Supplementary Fig. 9**). In addition, our analysis of synthetic aromatic amines revealed 7FTA as a more potent agonist of GPRC5A and its homologs as well as potential sites of ligand binding within their 7-transmembrane domains. Of note, our proteomic studies of IAA-interacting proteins also identified additional candidate GPCRs (GPR107 and GPR108) that have no reported small molecule ligands and numerous other proteins, which may reveal additional mechanisms of action for IAA. Chemoproteomics therefore provides a complementary approach to targeted screening of individual microbiota species<sup>19,57,58</sup> or biosynthetic gene clusters<sup>19,57,58</sup> for novel GPCR ligands.

Our discovery of microbiota metabolites and synthetic ligands of GPRC5A is important, as GPRC5A expression is correlated with lung, pancreatic and colon cancers<sup>30,31,59-61</sup>. Notably, *Gprc5a*<sup>-/-</sup> mice exhibit higher inflammation signatures and develop lung cancer over time<sup>29,32</sup>. In both human and mouse studies, GPRC5A has been implicated in epidermal growth factor receptor (EGFR)<sup>62</sup> and NF- $\kappa$ B<sup>32</sup> signaling, which may account for the modulation of inflammation and cancer incidence. However, no small-molecule agonists and antagonists have been reported for GPRC5A or its homologs. Our results suggest the generation of aromatic monoamines by specific microbiota species are

sensed by GPRC5A, which may crosstalk with cell signaling of other monoamine-sensing GPCRs (ADRA2A, DRD2 and HTR4) and/or key pathways associated with tissue homeostasis or immunity (**Supplementary Fig. 9**). Therefore, the future development of more potent and specific GPRC5A agonists and antagonists based on our studies may provide important pharmacological tools and therapeutic leads to investigate the roles of GPRC5A in physiology and human disease. Lastly, our studies highlight the utility of chemoproteomics to undercover novel mechanisms of action for indole metabolites that could be applied to other classes of microbiota metabolites.

**Acknowledgements** We thank H. Molina from Proteomics Resource Center at The Rockefeller University for proteomic analysis. We also thank G. Barnea (Brown University) for providing HTLA cells, B. Roth (UNC) for PRESTO-Tango plasmids, M. Gilmore (Harvard Medical School) for *E. faecium* Com15, S. F. Brady (The Rockefeller University) for *R. gnavus* ATCC 29149, and C-J. Guo (Weill Cornell Medical School) for *C. sporogenes* ATCC 15579. K.S. acknowledges support from Scripps Research Department of Immunology and Microbiology T32AI007244. H.C.H. acknowledges grant support from NIH-NIGMS R01GM087544.

**Author contributions** X.Z. and H.C.H. conceived of the project. K.R.S. contributed to the analysis of GPRC5A signaling, V.C. prepared TyrDC-KO *E. faecium* Com15 and M.E.G. performed phylogenetic analysis of bacterial decarboxylases. X.Z. performed the experiments and interpreted the data. X.Z. and H.C.H. wrote the manuscript with input of editing from K.R.S., V.C. and M.E.G.

**Competing interests** The authors declare no competing interests.

**Additional information** Supplementary information is available for this paper.

## REFERENCES

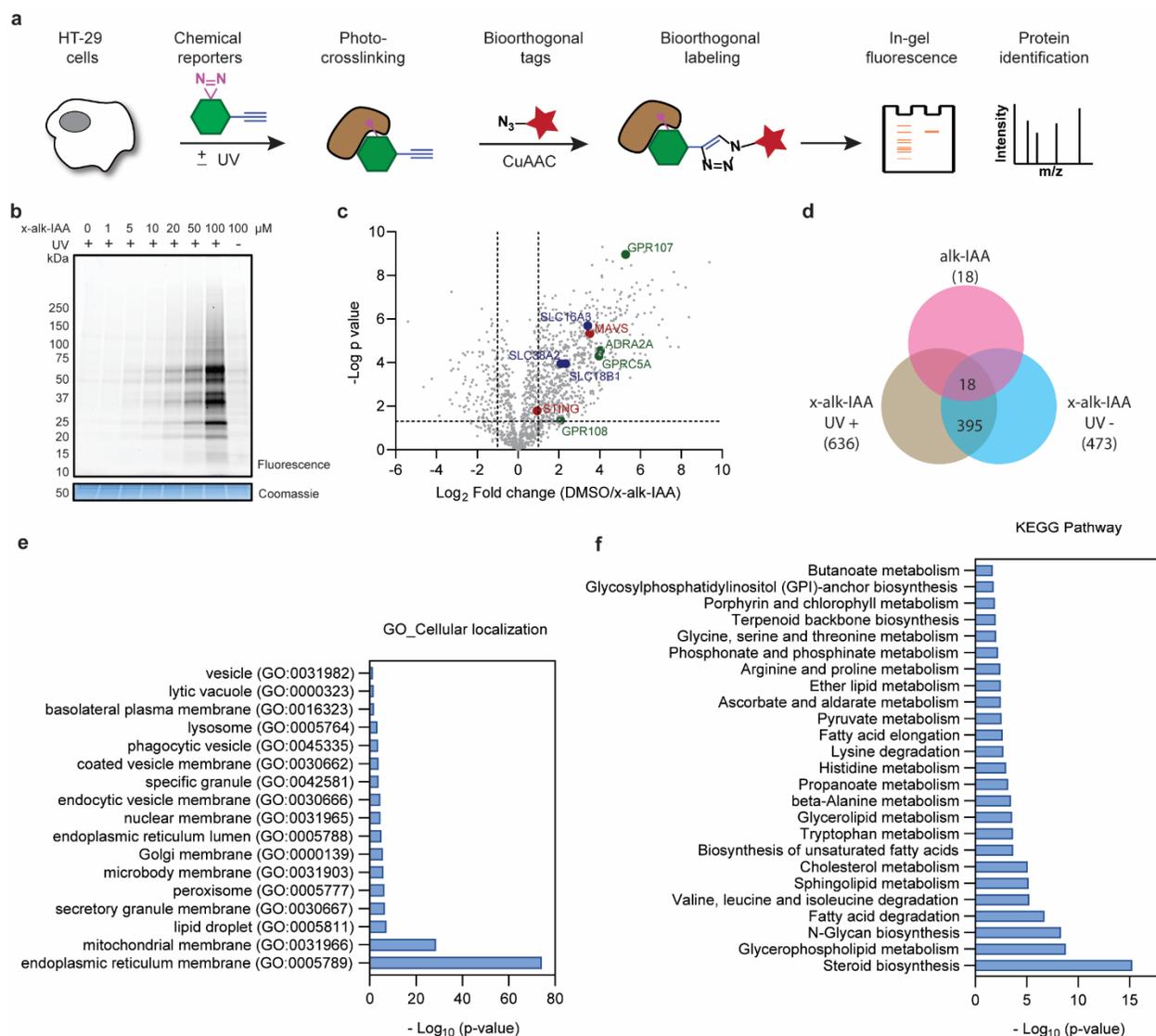
- 1 Koppel, N., Maini Rekdal, V. & Balskus, E. P. Chemical transformation of xenobiotics by the human gut microbiota. *Science (New York, N.Y.)* **356**, doi:10.1126/science.aag2770 (2017).
- 2 Donia, M. S. & Fischbach, M. A. HUMAN MICROBIOTA. Small molecules from the human microbiota. *Science (New York, N.Y.)* **349**, 1254766-1254766, doi:10.1126/science.1254766 (2015).
- 3 McCarville, J. L., Chen, G. Y., Cuevas, V. D., Troha, K. & Ayres, J. S. Microbiota Metabolites in Health and Disease. **38**, 147-170, doi:10.1146/annurev-immunol-071219-125715 (2020).
- 4 Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* **118**, 229-241, doi:10.1016/j.cell.2004.07.002 (2004).
- 5 Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* **165**, 842-853, doi:10.1016/j.cell.2016.04.007 (2016).
- 6 Danilchanka, O. & Mekalanos, J. J. Cyclic dinucleotides and the innate immune response. *Cell* **154**, 962-970, doi:10.1016/j.cell.2013.08.014 (2013).
- 7 Rangan, K. J. *et al.* A secreted bacterial peptidoglycan hydrolase enhances tolerance to enteric pathogens. *Science (New York, N.Y.)* **353**, 1434-1437, doi:10.1126/science.aaf3552 (2016).

- 8 Kim, B. *et al.* Enterococcus faecium secreted antigen A generates mucopeptides to enhance host immunity and limit bacterial pathogenesis. *eLife* **8**, doi:10.7554/eLife.45343 (2019).
- 9 Griffin, M. E. *et al.* Enterococcus peptidoglycan remodeling promotes checkpoint inhibitor cancer immunotherapy. *Science (New York, N.Y.)* **373**, 1040-1046, doi:10.1126/science.abc9113 (2021).
- 10 Lam, K. C. *et al.* Microbiota triggers STING-type I IFN-dependent monocyte reprogramming of the tumor microenvironment. *Cell* **184**, 5338-5356.e5321, doi:10.1016/j.cell.2021.09.019 (2021).
- 11 Nicolas, G. R. & Chang, P. V. Deciphering the Chemical Lexicon of Host-Gut Microbiota Interactions. *Trends in Pharmacological Sciences* **40**, 430-445, doi:10.1016/j.tips.2019.04.006 (2019).
- 12 Chaudhari, S. N. *et al.* Bariatric surgery reveals a gut-restricted TGR5 agonist with anti-diabetic effects. *Nature Chemical Biology* **17**, 20-29, doi:10.1038/s41589-020-0604-z (2021).
- 13 Maini Rekdal, V., Bess, E. N., Bisanz, J. E., Turnbaugh, P. J. & Balskus, E. P. Discovery and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism. *Science (New York, N.Y.)* **364**, doi:10.1126/science.aau6323 (2019).
- 14 Liu, Y., Hou, Y., Wang, G., Zheng, X. & Hao, H. Gut Microbial Metabolites of Aromatic Amino Acids as Signals in Host-Microbe Interplay. *Trends in endocrinology and metabolism: TEM* **31**, 818-834, doi:10.1016/j.tem.2020.02.012 (2020).
- 15 Dodd, D. *et al.* A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* **551**, 648-652, doi:10.1038/nature24661 (2017).
- 16 Lamas, B. *et al.* CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nature medicine* **22**, 598-605, doi:10.1038/nm.4102 (2016).
- 17 Venkatesh, M. *et al.* Symbiotic bacterial metabolites regulate gastrointestinal barrier function via the xenobiotic sensor PXR and Toll-like receptor 4. *Immunity* **41**, 296-310, doi:10.1016/j.immuni.2014.06.014 (2014).
- 18 Bhattarai, Y. *et al.* Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell host & microbe* **23**, 775-785.e775, doi:10.1016/j.chom.2018.05.004 (2018).
- 19 Chen, H. *et al.* A Forward Chemical Genetic Screen Reveals Gut Microbiota Metabolites That Modulate Host Physiology. *Cell* **177**, 1217-1231.e1218, doi:10.1016/j.cell.2019.03.036 (2019).
- 20 Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annual review of biochemistry* **77**, 383-414, doi:10.1146/annurev.biochem.75.101304.124125 (2008).
- 21 Wright, M. H. Chemical Proteomics of Host-Microbe Interactions. *Proteomics* **18**, e1700333, doi:10.1002/pmic.201700333 (2018).
- 22 Parasar, B. *et al.* Chemoproteomic Profiling of Gut Microbiota-Associated Bile Salt Hydrolase Activity. *ACS Central Science* **5**, 867-873, doi:10.1021/acscentsci.9b00147 (2019).
- 23 Parker, C. G. & Pratt, M. R. Click Chemistry in Proteomic Investigations. *Cell* **180**, 605-632, doi:10.1016/j.cell.2020.01.025 (2020).
- 24 Tsukidate, T., Li, Q. & Hang, H. C. Targeted and proteome-wide analysis of metabolite-protein interactions. *Current opinion in chemical biology* **54**, 19-27, doi:10.1016/j.cbpa.2019.10.008 (2020).
- 25 Agus, A., Planchais, J. & Sokol, H. Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. *Cell host & microbe* **23**, 716-724, doi:10.1016/j.chom.2018.05.003 (2018).
- 26 Roager, H. M. & Licht, T. R. Microbial tryptophan catabolites in health and disease. *Nature communications* **9**, 3294, doi:10.1038/s41467-018-05470-4 (2018).

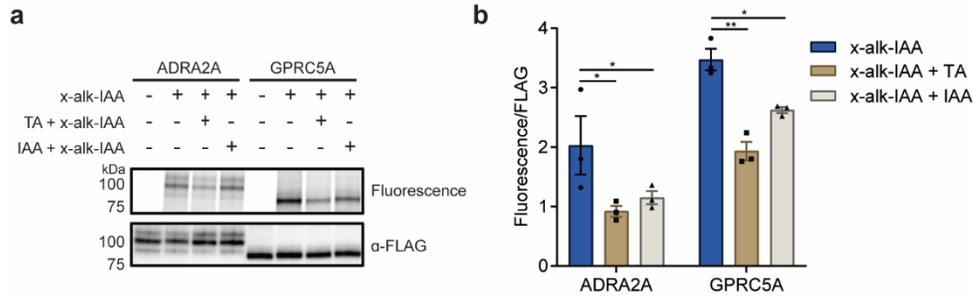
- 27 Krishnan, S. *et al.* Gut Microbiota-Derived Tryptophan Metabolites Modulate Inflammatory Response in Hepatocytes and Macrophages. *Cell reports* **23**, 1099-1111, doi:10.1016/j.celrep.2018.03.109 (2018).
- 28 Liu, Y. *et al.* Gut microbiome alterations in high-fat-diet-fed mice are associated with antibiotic tolerance. *Nature microbiology* **6**, 874-884, doi:10.1038/s41564-021-00912-0 (2021).
- 29 Tao, Q. *et al.* Identification of the retinoic acid-inducible Gprc5a as a new lung tumor suppressor gene. *Journal of the National Cancer Institute* **99**, 1668-1682, doi:10.1093/jnci/djm208 (2007).
- 30 Zhou, H. & Rigoutsos, I. The emerging roles of GPRC5A in diseases. *Oncoscience* **1**, 765-776, doi:10.18632/oncoscience.104 (2014).
- 31 Fujimoto, J. *et al.* Development of Kras mutant lung adenocarcinoma in mice with knockout of the airway lineage-specific gene Gprc5a. *International journal of cancer* **141**, 1589-1599, doi:10.1002/ijc.30851 (2017).
- 32 Deng, J. *et al.* Knockout of the tumor suppressor gene Gprc5a in mice leads to NF-kappaB activation in airway epithelium and promotes lung inflammation and tumorigenesis. *Cancer prevention research (Philadelphia, Pa.)* **3**, 424-437, doi:10.1158/1940-6207.Capr-10-0032 (2010).
- 33 Jin, U. H. *et al.* Microbiome-derived tryptophan metabolites and their aryl hydrocarbon receptor-dependent agonist and antagonist activities. *Molecular pharmacology* **85**, 777-788, doi:10.1124/mol.113.091165 (2014).
- 34 Chin, E. N. *et al.* Antitumor activity of a systemic STING-activating non-nucleotide cGAMP mimetic. *Science (New York, N.Y.)* **369**, 993-999, doi:10.1126/science.abb4255 (2020).
- 35 Pan, B. S. *et al.* An orally available non-nucleotide STING agonist with antitumor activity. *Science (New York, N.Y.)* **369**, doi:10.1126/science.aba6098 (2020).
- 36 Hein, L., Altman, J. D. & Kobilka, B. K. Two functionally distinct alpha2-adrenergic receptors regulate sympathetic neurotransmission. *Nature* **402**, 181-184, doi:10.1038/46040 (1999).
- 37 Nemet, I. *et al.* A Cardiovascular Disease-Linked Gut Microbial Metabolite Acts via Adrenergic Receptors. *Cell* **180**, 862-877.e822, doi:10.1016/j.cell.2020.02.016 (2020).
- 38 Tafesse, F. G. *et al.* GPR107, a G-protein-coupled receptor essential for intoxication by *Pseudomonas aeruginosa* exotoxin A, localizes to the Golgi and is cleaved by furin. *The Journal of biological chemistry* **289**, 24005-24018, doi:10.1074/jbc.M114.589275 (2014).
- 39 Dudek, A. M. *et al.* GPR108 Is a Highly Conserved AAV Entry Factor. *Molecular Therapy* **28**, 367-381, doi:10.1016/j.ymthe.2019.11.005 (2020).
- 40 Elrick, M. M. *et al.* Neuronostatin acts via GPR107 to increase cAMP-independent PKA phosphorylation and proglucagon mRNA accumulation in pancreatic  $\alpha$ -cells. *American journal of physiology. Regulatory, integrative and comparative physiology* **310**, R143-155, doi:10.1152/ajpregu.00369.2014 (2016).
- 41 Cheng, Y. & Lotan, R. Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor. *The Journal of biological chemistry* **273**, 35008-35015, doi:10.1074/jbc.273.52.35008 (1998).
- 42 Ellaithy, A., Gonzalez-Maeso, J., Logothetis, D. A. & Levitz, J. Structural and Biophysical Mechanisms of Class C G Protein-Coupled Receptor Function. *Trends in Biochemical Sciences* **45**, 1049-1064, doi:10.1016/j.tibs.2020.07.008 (2020).
- 43 Kroeze, W. K. *et al.* PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature structural & molecular biology* **22**, 362-369, doi:10.1038/nsmb.3014 (2015).
- 44 Williams, B. B. *et al.* Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell host & microbe* **16**, 495-503, doi:10.1016/j.chom.2014.09.001 (2014).

- 45 Pessione, E. *et al.* First evidence of a membrane-bound, tyramine and beta-phenylethylamine producing, tyrosine decarboxylase in *Enterococcus faecalis*: a two-dimensional electrophoresis proteomic study. *Proteomics* **9**, 2695-2710, doi:10.1002/pmic.200800780 (2009).
- 46 Moreno-Arribas, V., Torlois, S., Joyeux, A., Bertrand, A. & Lonvaud-Funel, A. Isolation, properties and behaviour of tyramine-producing lactic acid bacteria from wine. *Journal of applied microbiology* **88**, 584-593, doi:10.1046/j.1365-2672.2000.00997.x (2000).
- 47 Nelson, K. E. *et al.* A catalog of reference genomes from the human microbiome. *Science (New York, N.Y.)* **328**, 994-999, doi:10.1126/science.1183605 (2010).
- 48 Chen, V., Griffin, M. E., Maguin, P., Varble, A. & Hang, H. C. RecT Recombinase Expression Enables Efficient Gene Editing in *Enterococcus* spp. *Applied and environmental microbiology* **87**, e0084421, doi:10.1128/aem.00844-21 (2021).
- 49 Bender, B. J., Marlow, B. & Meiler, J. Improving homology modeling from low-sequence identity templates in Rosetta: A case study in GPCRs. *PLoS computational biology* **16**, e1007597, doi:10.1371/journal.pcbi.1007597 (2020).
- 50 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
- 51 Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32**, W526-W531, doi:10.1093/nar/gkh468 (2004).
- 52 Zheng, Y. *et al.* The Role of Bacterial-Derived Aromatic Amino Acids Metabolites Relevant in Autism Spectrum Disorders: A Comprehensive Review. **15**, doi:10.3389/fnins.2021.738220 (2021).
- 53 Steed, A. L. *et al.* The microbial metabolite desaminotyrosine protects from influenza through type I interferon. *Science (New York, N.Y.)* **357**, 498-502, doi:10.1126/science.aam5336 (2017).
- 54 Cervantes-Barragan, L. *et al.* *Lactobacillus reuteri* induces gut intraepithelial CD4(+)CD8α(+) T cells. *Science (New York, N.Y.)* **357**, 806-810, doi:10.1126/science.aah5825 (2017).
- 55 Scott, S. A., Fu, J. & Chang, P. V. Microbial tryptophan metabolites regulate gut barrier function via the aryl hydrocarbon receptor. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 19376-19387, doi:10.1073/pnas.2000047117 (2020).
- 56 Guo, C. J. *et al.* Depletion of microbiome-derived molecules in the host using *Clostridium* genetics. *Science (New York, N.Y.)* **366**, doi:10.1126/science.aav1282 (2019).
- 57 Colosimo, D. A. *et al.* Mapping Interactions of Microbial Metabolites with Human G-Protein-Coupled Receptors. *Cell host & microbe* **26**, 273-282.e277, doi:10.1016/j.chom.2019.07.002 (2019).
- 58 Chang, F. Y. *et al.* Gut-inhabiting *Clostridia* build human GPCR ligands by conjugating neurotransmitters with diet- and human-derived fatty acids. *Nature microbiology* **6**, 792-805, doi:10.1038/s41564-021-00887-y (2021).
- 59 Insel, P. A. *et al.* GPCRomics: GPCR Expression in Cancer Cells and Tumors Identifies New, Potential Biomarkers and Therapeutic Targets. *Frontiers in pharmacology* **9**, 431, doi:10.3389/fphar.2018.00431 (2018).
- 60 Jahny, E. *et al.* The G Protein-Coupled Receptor RA13 Is an Independent Prognostic Factor for Pancreatic Cancer Survival and Regulates Proliferation via STAT3 Phosphorylation. *PLoS one* **12**, e0170390, doi:10.1371/journal.pone.0170390 (2017).
- 61 Zougman, A. *et al.* Retinoic acid-induced protein 3: identification and characterisation of a novel prognostic colon cancer biomarker. *European journal of cancer (Oxford, England : 1990)* **49**, 531-539, doi:10.1016/j.ejca.2012.07.031 (2013).
- 62 Lin, X. *et al.* EGFR phosphorylates and inhibits lung tumor suppressor GPRC5A in lung cancer. *Molecular cancer* **13**, 233, doi:10.1186/1476-4598-13-233 (2014).

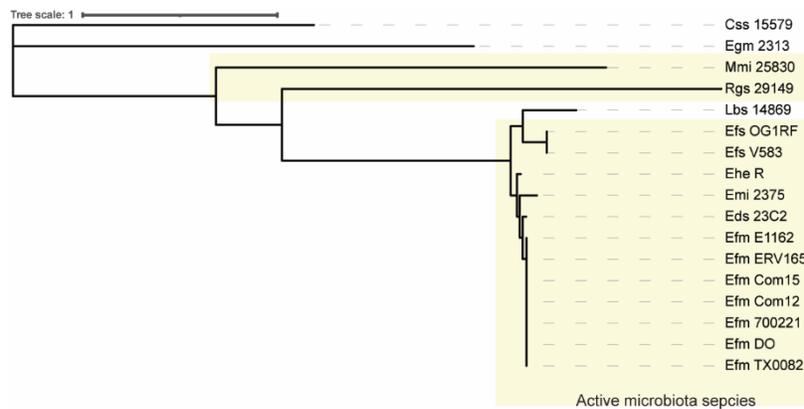
- 63 Hayashi, K.-i. *et al.* Auxin transport sites are visualized in planta using fluorescent auxin analogs. **111**, 11557-11562, doi:10.1073/pnas.1408960111 %J Proceedings of the National Academy of Sciences (2014).
- 64 Lin, J. C. *et al.* An in vivo tagging method reveals that Ras undergoes sustained activation upon transglutaminase-mediated protein serotonylation. *Chembiochem : a European journal of chemical biology* **14**, 813-817, doi:10.1002/cbic.201300050 (2013).
- 65 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).
- 66 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).
- 67 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-w641, doi:10.1093/nar/gkz268 (2019).
- 68 Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)* **25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).
- 69 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PloS one* **5**, e9490, doi:10.1371/journal.pone.0009490 (2010).
- 70 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256-w259, doi:10.1093/nar/gkz239 (2019).
- 71 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).
- 72 Hullahalli, K., Rodrigues, M., Nguyen, U. T. & Palmer, K. An Attenuated CRISPR-Cas System in *Enterococcus faecalis* Permits DNA Acquisition. *mBio* **9**, doi:10.1128/mBio.00414-18 (2018).



**Supplementary Fig. 1 Chemoproteomic analysis of IAA-interacting proteins in HT-29 cells.** **a**, Scheme for in-gel fluorescence profiling and chemoproteomic analysis of microbiota metabolite protein targets in HT-29 cells using photoaffinity chemical reporters. **b**, In-gel fluorescence profiling of IAA-interacting proteins in HT-29 cells using indicated concentration of x-alk-IAA with or without UV irradiation. **c**, Volcano plot of the identified proteins from HT-29 cells by chemical proteomics using x-alk-IAA with UV irradiation compared to DMSO control ( $p$  value is 0.05, two folds change in four replicates). Some of key hits are highlighted in GPCRs (green), transporters (blue) and immunity-associated proteins (red). **d**, Venn diagram for the statistically significant hits identified by chemical proteomics using two chemical reporters with or without UV irradiation. **e**, Gene ontology analysis of the major subcellular localization of x-alk-IAA interacting proteins. **f**, Major KEGG pathways involve x-alk-IAA interacting proteins.



**Supplementary Fig. 2 ADRA2A and GPCR5A are crosslinked by IAA photoaffinity chemical reporters.** **a,b**, Competitive labeling assay using 200  $\mu$ M IAA or TA that compete with 20  $\mu$ M x-alk-IAA in binding on ADRA2A or GPCR5A. Two-way ANOVA using Holm-Sidak's multiple comparisons test. \* $P < 0.05$ , \*\* $P \leq 0.01$ , ns indicates not significant.



**Supplementary Fig. 3 a**, Unrooted phylogenetic clustering of protein sequences of tyrosine decarboxylases of *Enterococcus* species and strains, as well as *L. brevis* 14869, glutamate decarboxylase of *E. gallinarum* 2313, tryptophan decarboxylases of *R. gnavus* 29149 and *C. sporogenes* 15579, glutamate or tyrosine decarboxylases of *M. morgani* 25830. Active strains are indicated by the yellow box. Scale bar represents sequence distance.

**Supplementary Fig. 3 b**, Sequence alignment of amino acid decarboxylases of different microbiota species.



**Supplementary Fig. 3 c**, Sequence alignment of amino acid decarboxylases of different *Enterococcus* species.

```
1      10      20      30      40      50      60      70      80      90
Efs_OG1RF  MNARSKNSLKFVRRNFRMKNEKLAKGEMNLNALFIGDKAENGQLYKDLLIDLVDLHGLWRQNYMFDMPVISSQERTSEYKTVNHMKDVLNE
Efs_V583   MKNEKLAKEGEMNLNALFIGDKAENGQLYKDLLIDLVDLHGLWRQNYMFDMPVISSQERTSEYKTVNHMKDVLNEISSRMRTSHSPVHWTAG
Efm_Com15  MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_Com12  MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_DO     MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_702221 MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_E1162  MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_TX0082 MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Efm_ERV165 MSLSLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Eds_23C2   .MESLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Ehe_R      .MESLSDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Emi_2375   .MESKVNLDLNALNLFIDGDKAENGQIYKALLNELVDLHGLWRQNYMFPQDMPITPEEKSSASFHTVNTKTKDVLSEISARMRTHSPVHWTAGRY
Egm_2313   MLYGKRENREEDYLEPIFGAVCEGEDLPKYKLNQPSVPRMAYQLVQQLDLEGNARLNLAIFCQTYMEPEAVKLTMTQTLKNAIDKSEYPRR

100     110     120     130     140     150     160     170     180
Efs_OG1RF  ISSRMRTSHSPVHWTAGRYWGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYAR
Efs_V583   RYWGHNSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_Com15  WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_Com12  WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_DO     WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_702221 WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_E1162  WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_TX0082 WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Efm_ERV165 WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Eds_23C2   WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Ehe_R      WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Emi_2375   WGHMNSSETLMPSELLAYNFAMLWNGNNVAYESSPATSQMEEVGVGHEFAHLSYKNGWGHIVADGSLANLEGLWYARNIKSLPFAMKEVVKPEL
Egm_2313   AEIENRCVNIADLWHASKEEFMGTSTIGSEACMLGGMAKFAWRNRAEKLGLDLNLAHKPNLVISSGYQVCWEKFCVYWDIMREVFMDEN

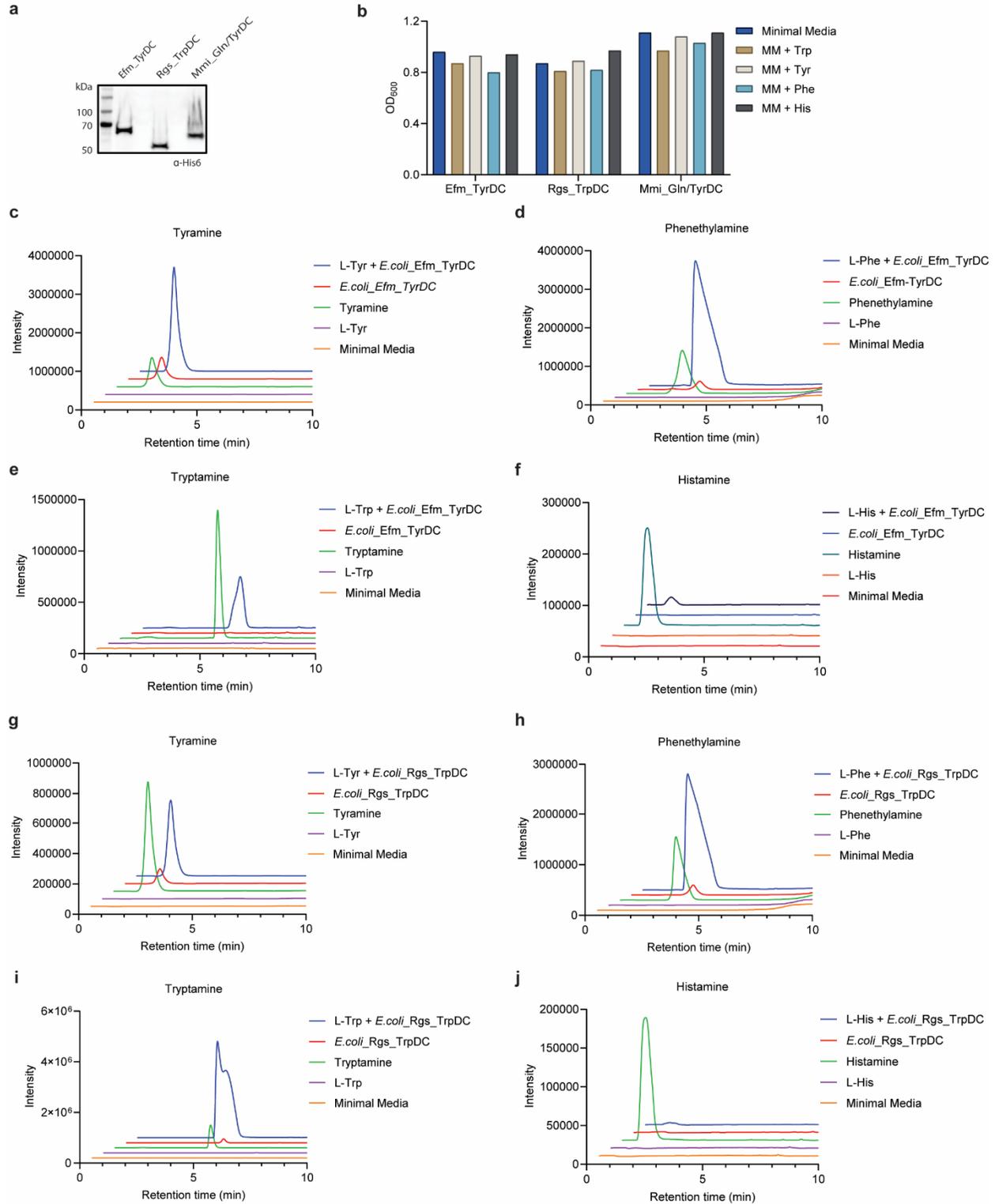
190     200     210     220     230     240     250     260     270
Efs_OG1RF  NIKSLPFAMKEVVKPELVAGKSDWELLNMPKTEIMDLLESDEIDETIKAHASARSQKHLQAIQAGKWLVPQTKHYSWLKAADIIGIGLDQVIVPVP
Efs_V583   GKSDWELLNMPKTEIMDLLESAEDEIKAHASARSQKHLQAIQAGKWLVPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVR
Efm_Com15  GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_Com12  GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_DO     GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_702221 GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_E1162  GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_TX0082 GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Efm_ERV165 GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Eds_23C2   GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Ehe_R      GKSDWELMNLSTTEIMNLLDSVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Emi_2375   QTDWELMNLSTTEIIDLQVPEKIDEIKAHASARSQKHLQKGLWLPQTKHYSWLKAADIIGIGLDQVIVPVPVHDHNYRMDINELEKIVRGL
Egm_2313   HLSINLDKVMDDVDEYTLGIVGILGITYTGRTDIDKGLNLELVEQYNTIDYKVIHVDAASGGLYTPTPEPLVDFQLPNVISTINTSGHKYV

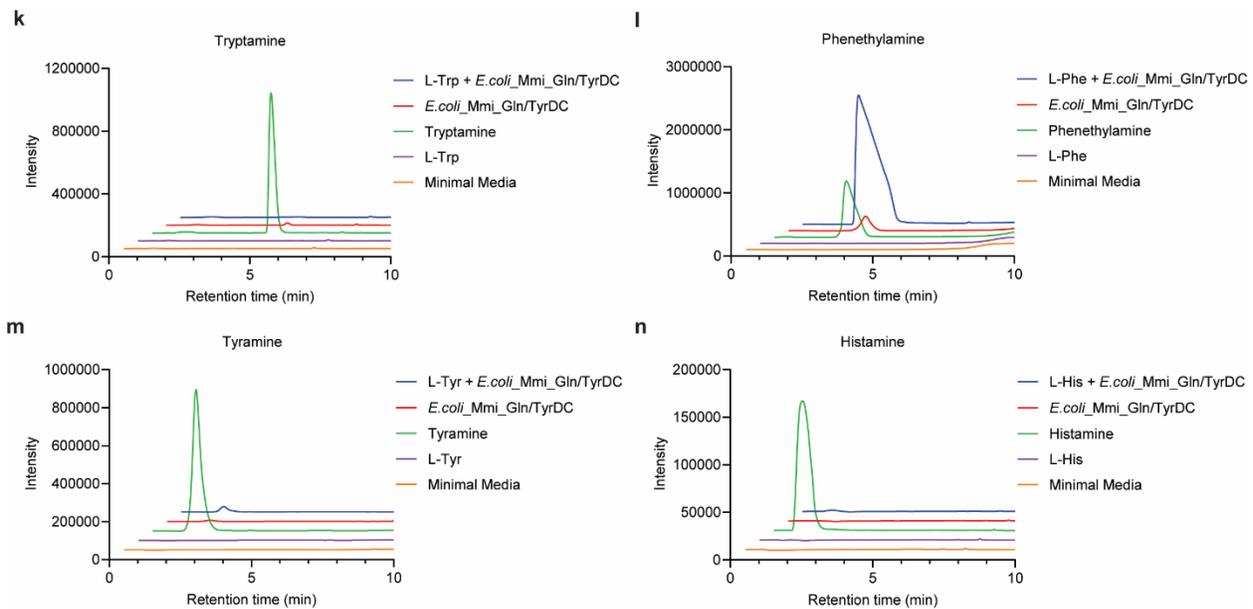
280     290     300     310     320     330     340     350     360     370
Efs_OG1RF  DHNYRMDINELEKIVRGLAEQIPVLGVVGVVGSVEEGAVDSIDKIALRDELMDKGIYFVYVHDAAYGGYGRAIFLDEDNFIPFELKDVHVKY
Efs_V583   GLAEQIPVLGVVGVVGSVEEGAVDSIDKIALRDELMDKGIYFVYVHDAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_Com15  AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_Com12  AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_DO     AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_702221 AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_E1162  AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_TX0082 AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Efm_ERV165 AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Eds_23C2   AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Ehe_R      AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Emi_2375   AAETPILGVVGVVGSVEEGAIDGIDKIVARRVLEKDGIFYFLHVDAAAYGGYGRAIFLDEDNFIPFELKDVHVKYNNVFTENKDYILEEVH
Egm_2313   LVVPGVGVVLRDQSYLPQELVFKVSYLGCLELTAINFSSHAAQLICQYVNFVYGFDCYKATHDKTHKVALYLLAEIKETDLFEMINDGAA

380     390     400     410     420     430     440     450     460
Efs_OG1RF  EYGVFKKKEHISREYVDYKAIELAESVTDPHKMGYIPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efs_V583   VYDAYKAIELAESVTDPHKMGYIPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_Com15  SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_Com12  SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_DO     SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_702221 SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_E1162  SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_TX0082 SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Efm_ERV165 SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Eds_23C2   SAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Ehe_R      RAFKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Emi_2375   RAYKAIIEEAESVTDPHKMGYVPPYSAGGIVIQDIRMRDVISYFATYVFEKADIPALLGAYILEGSKAGATAA
Egm_2313   LPIVCYKLDKDDPKREWTLYDLADRLMKGWQVPAYPELANLDDQIYQRLVIRADFGNMMAHDYIQDMKEALLALDQAHVYHKKQEIKEYGFI

470     480     490     500     510     520     530     540     550
Efs_OG1RF  WAAHHVPLPLNVAAGYKLGASIEGSHHFYNFLNDLTFKVGDKIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efs_V583   LIGASIEGSHHFYNFLNDLTFKVGDKIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_Com15  GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_Com12  GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_DO     GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_702221 GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_E1162  GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_TX0082 GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Efm_ERV165 GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Eds_23C2   GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Ehe_R      GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Emi_2375   GASIEGAHRFFYNFLNLSFKVGDKEIEVHFLTPDFNMVDYVFEKGNDDLVAAMKLNHDVYDSSYVKGSIYGNF
Egm_2313   H...

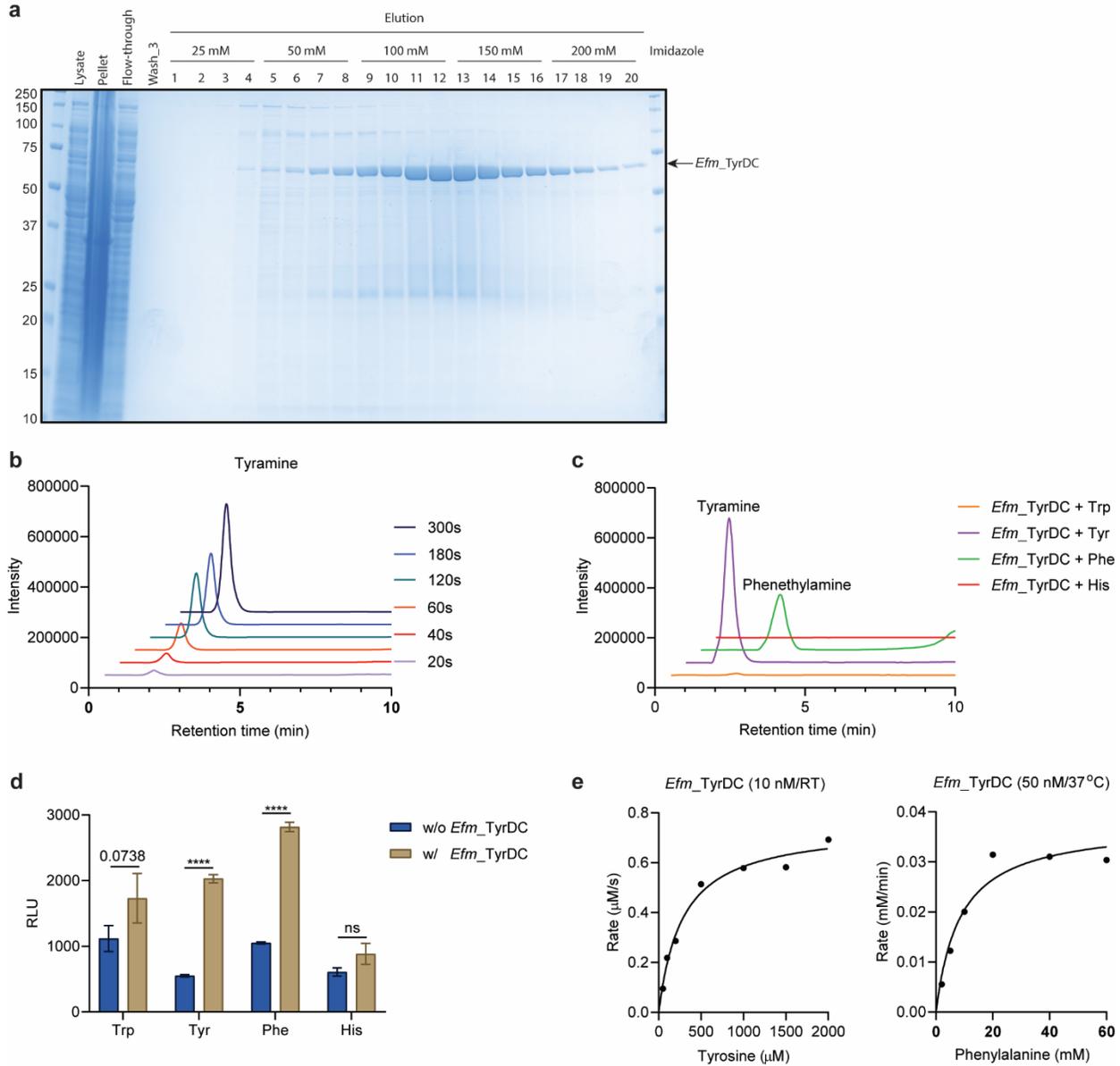
560     570     580     590     600     610     620     630
Efs_OG1RF  EFITSHDFAIPDYGNSPLKRVNSLGFSDDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efs_V583   SPLKRVNSLGFSDENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_Com15  LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_Com12  LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_DO     LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_702221 LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_E1162  LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_TX0082 LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Efm_ERV165 LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Eds_23C2   LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Ehe_R      LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Emi_2375   LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
Egm_2313   LQFVNQLGFSDEENRAGKVTVLRASVMTTPYMNKEHFEEYAEKIKAAALQEKLEKIYADQLLASEAK
```



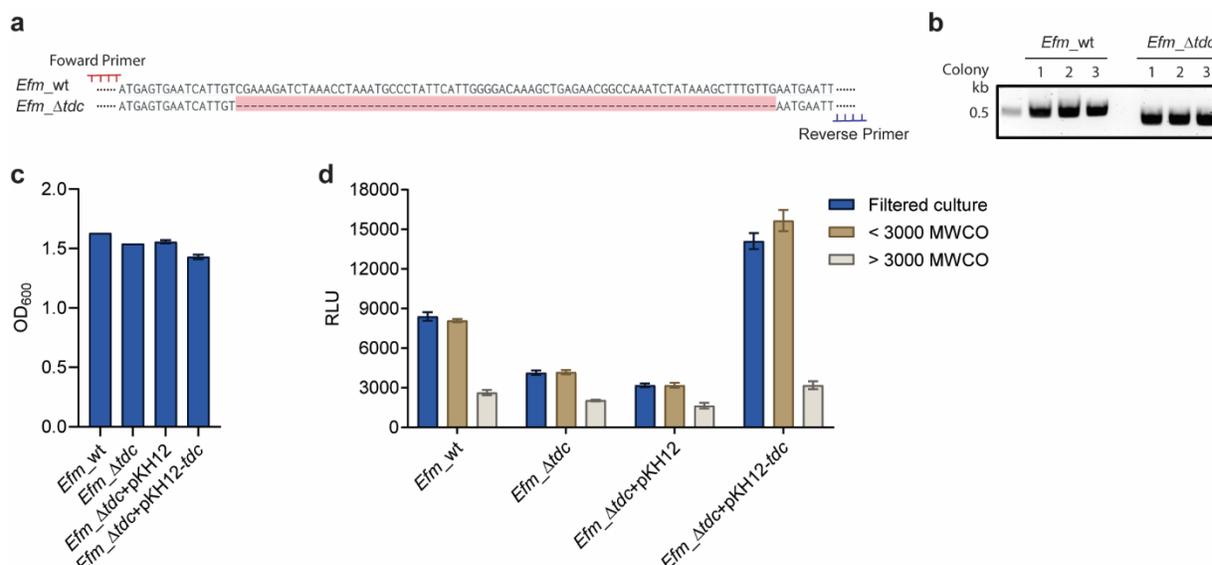


**Supplementary Fig. 4 Bacterial decarboxylases transform specific aAAs to aromatic monoamines.**

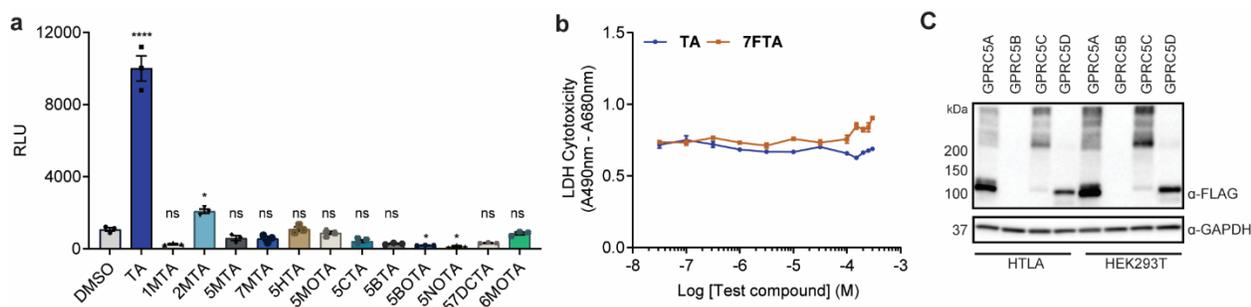
**a**, Western blot analysis of the overexpression of His6-tagged *E. faecium* Com15 tyrosine decarboxylase (Efm\_TyrDC), *R. gnavus* tryptophan decarboxylase (Rgs\_TrpDC) and *M. organii* glutamate/tyrosine decarboxylase (Mmi\_Gln/TyrDC) in *E. coli* DH5a. **b**, Bacterial growth in minimal media supplemented with individual aAAs at 37 °C for 16 h. **c-n**, LC-MS analysis of respective aromatic monoamines produced by *E. coli*\_Efm\_TyrDC (c-f), *E. coli*\_Rgs\_TrpDC (g-j) and *E. coli*\_Mmi\_Gln/TyrDC (k-n).



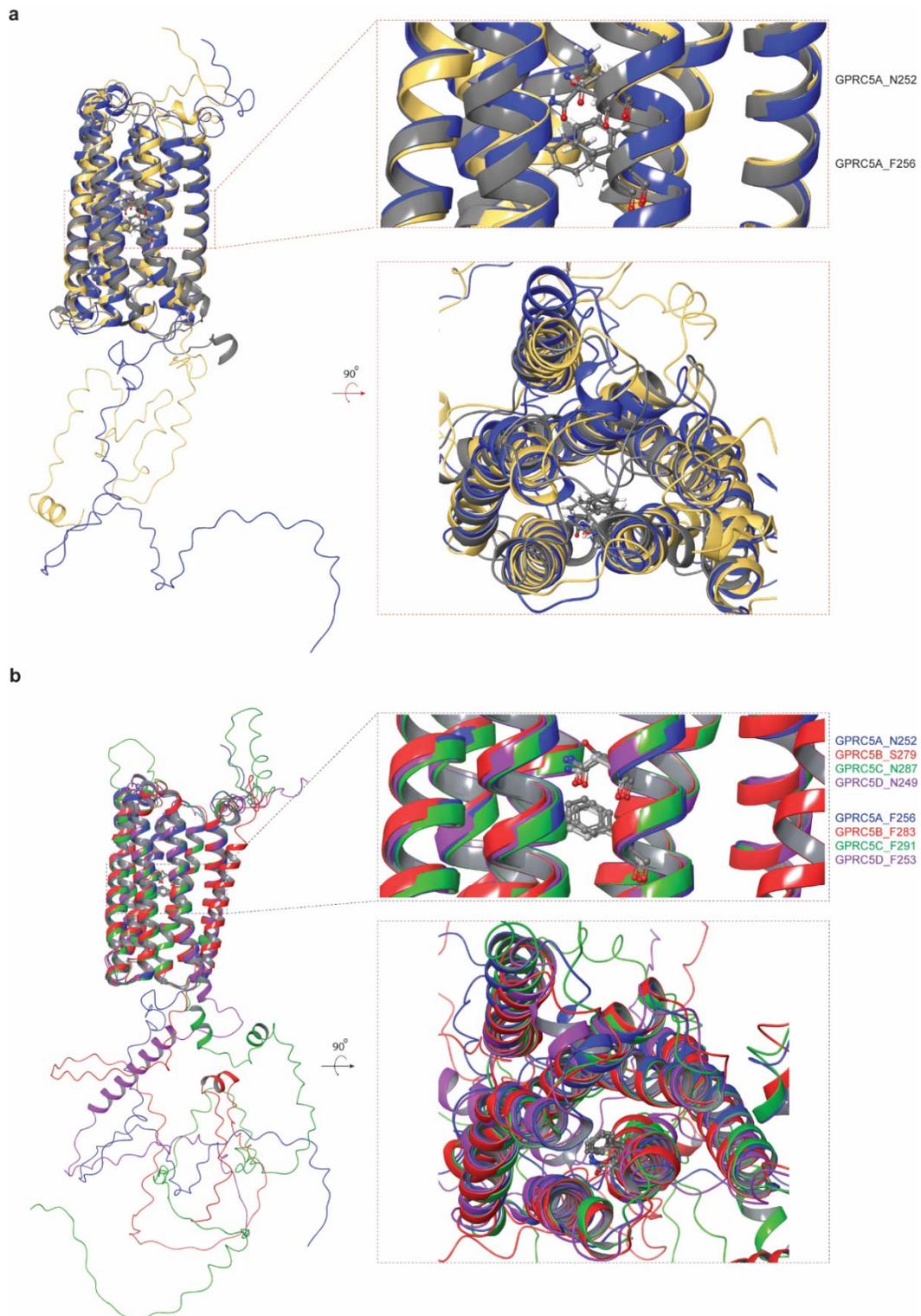
**Supplementary Fig. 5 *In vitro* enzymatic transformation of aAAs to aromatic monoamines.** **a**, SDS-PAGE analysis of the purification of His6-tagged *E. faecium* Com15 tyrosine decarboxylase (*Efm\_TyrDC*). **b**, LC-MS analysis of the time-dependent formation of tyramine from enzymatic transformation of tyrosine by 100 nM *Efm\_TyrDC* at room temperature. **c**, LC-MS determination of tryptamine, tyramine, phenethylamine and histamine from enzymatic transformation of tryptamine, tyrosine, phenylalanine and histidine by 100 nM *Efm\_TyrDC* at 37 °C for 10 min. **d**, PRESTO-Tango assay for *Efm\_TyrDC* enzymatic products for GPRC5A activation. **e**, Michaelis-Menten kinetics study of *Efm\_TyrDC* catalyzed transformation of tyrosine and phenylalanine. Data indicates mean with SEM from three replicates. Two-way ANOVA using Sidak's multiple comparisons test. \*\*\*\* $p < 0.0001$ , ns indicates not significant.



**Supplementary Fig. 6 Knock-out of tyrosine decarboxylase in *E. faecium* Com15.** **a**, Genetic deletion of segment of DNA disrupts *TyrDC* gene expression in *E. faecium* Com15. **b**, Agarose gel electrophoresis of the PCR products from wild-type and *TyrDC* gene-disrupted *E. faecium* Com15. **c**, Optical density measurement of bacterial growth including wild-type *E. faecium* Com15 (*Efm\_wt*), *E. faecium* Com15 *TyrDC* knock-out (*Efm\_Δtdc*), complementation of vector control (pKH12) or *TyrDC* (pKH12-tdc) in *Efm\_Δtdc* at 37 °C for 16 h. **d**, PRESTO-Tango assay for examining bacterial cultures for GPRC5A activation. Bacterial growth media were filtered through 0.2  $\mu\text{M}$  membrane followed by 3,000 MWKO membrane.

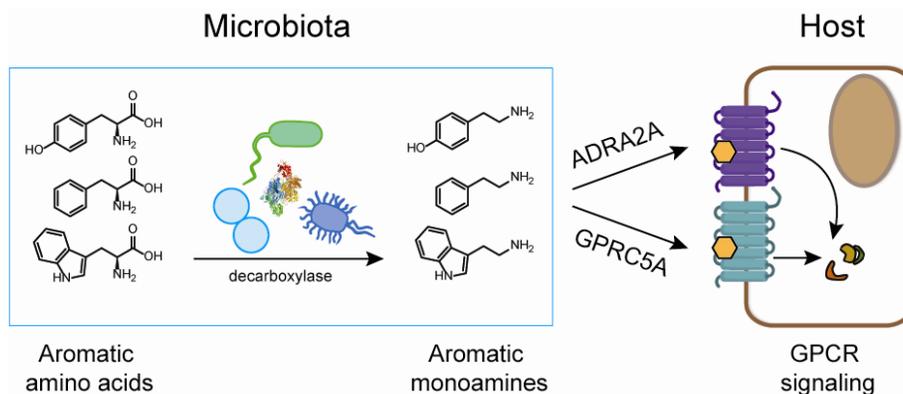
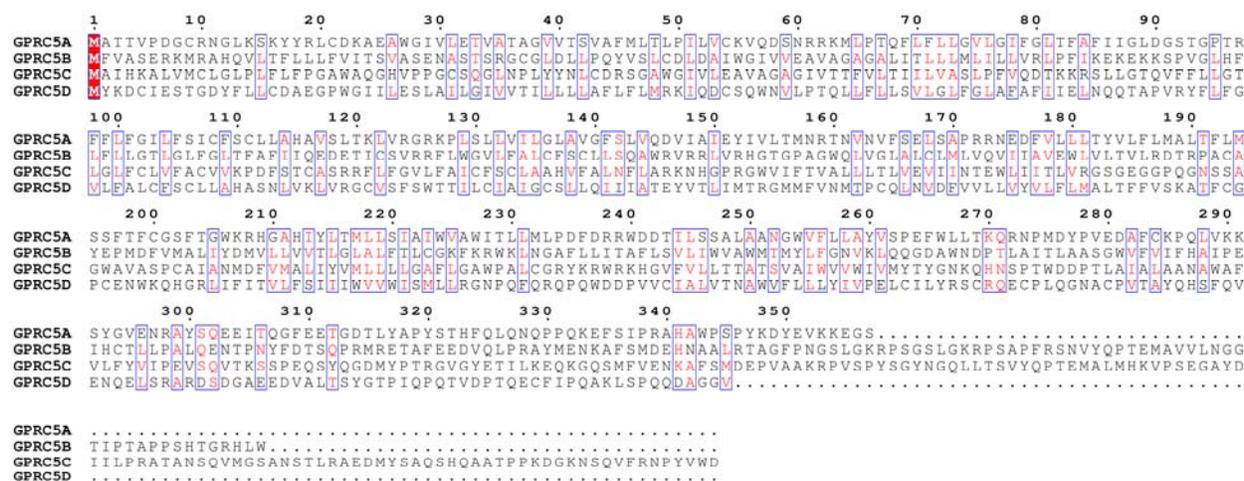


**Supplementary Fig. 7 Structure-activity relationship studies of GPRC5A ligands.** **a**, PRESTO-Tango assay for screening a collection of tryptamine derivatives (100  $\mu\text{M}$ ) for GPRC5A activation. **b**, LDH cytotoxicity assay to examine TA and 7FTA in HEK293T cells. **c**, Western blot analysis of the expression of GPRC5A-, GPRC5B-, GPRC5C-, GPRC5D-Tango proteins in HTLA or HEK293T cells, respectively.



**Supplementary Fig. 8 a**, Structural models of GPCR5A from RosettaGPCR (gray), AlphaFold (blue) and Robetta (yellow). N252 and F256 are indicated. **b**, Structural models of GPCR5A (blue), GPCR5B (red), GPCR5C (green) and GPCR5D (purple) from AlphaFold.

**Supplementary Fig. 8 c**, Sequence alignment of GPRC5A, GPRC5B, GPRC5C, GPRC5D.



**Supplementary Fig. 9** Schematic of specific microbiota metabolism of aromatic amino acids by decarboxylases and activation of GPRC5A and other functionally-related GPCRs within host cells.

**Supplementary Table 1** Primers.

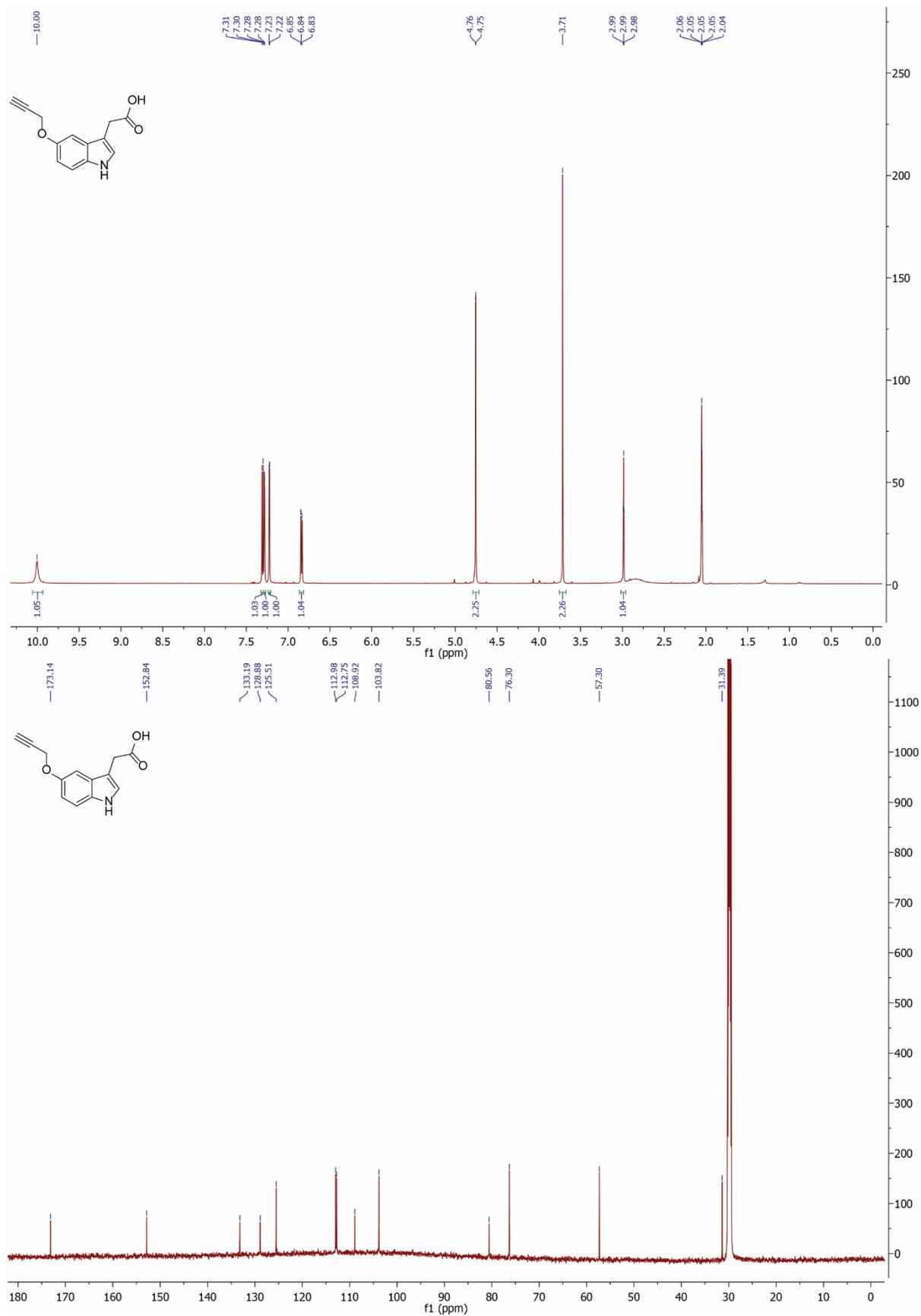
Primer	Sequence (5'-3')
oVC215	aaactcagctttgtccccaatgaatagggcatttg
oVC216	aaaacaaatgcctattcattggggacaaagctga
oVC217	ttacagaataggagaacagggtg
oVC218	cccagccatctttatagctc
oVC219	catgtagtttgacgccagcctaaatgttcgtccaccaattcattacaatgattcactcatagtgaaccctcctaaaaaatattttaa
pKH12_fwd	cgaagcaaaacaccatcaccatcaccattaacattaggatccagatcttc
pKH12_rev	ttcgacaatgattcactcatgatactcctcctcaaaatatttttg
Efm_TyrDC_fwd	atgagtgaatcattgtcgaaag
Efm_TyrDC_rev	cctaattgtaatggatgatggatggtgttttgcttcgcttgccaataattg
Rgs_TrpDC_fwd	tttgaggaggagatcatgatcggaaccgaatatattttaa
Rgs_TrpDC_rev	ggtgatggtgatggtgagctttttcatttccgtataaaatttc
Mmi_TyrDC_fwd	tttgaggaggagatcatgaatcagtcagaactatc
Mmi_TyrDC_rev	ggtgatggtgatggtgttctgactcatggttttaag
PTGPRC5A_D176A_R	gagcagcacgaaggcctcgttccgccg
PTGPRC5A_D176A_F	cggcgaacgaggcctcgtgctgctc
PTGPRC5A_D241A_R	cagtatggtgtcggcccagcggcggtc
PTGPRC5A_D241A_F	gaccgccgctgggccgacaccatactg
PTGPRC5A_D242A_R	gctcagtatggtggcgtcccagcggcg
PTGPRC5A_D242A_F	cgccgctgggacgccaccatactgagc
PTGPRC5A_D237A_R	cccagcggcgggcaaagtctggcaac
PTGPRC5A_D237A_F	gttgccagactttgcccgcgctggg
PTGPRC5A_F105A_R	gaggcaggagaaacaaatagaagccagatcccaaacagaaagaac
PTGPRC5A_F105A_F	gttcttctgtttgggatactggcttctattgtttctcctgcctc
PTGPRC5A_F109A_R	agccaggaggcaggaggcacaatagaaaacagtatcccaaa
PTGPRC5A_F109A_F	ttgggatactgtttctattgtgctcctcctcctggct
PTGPRC5A_F256A_R	aaacataagcaggagagccaccaaccgttggcgg

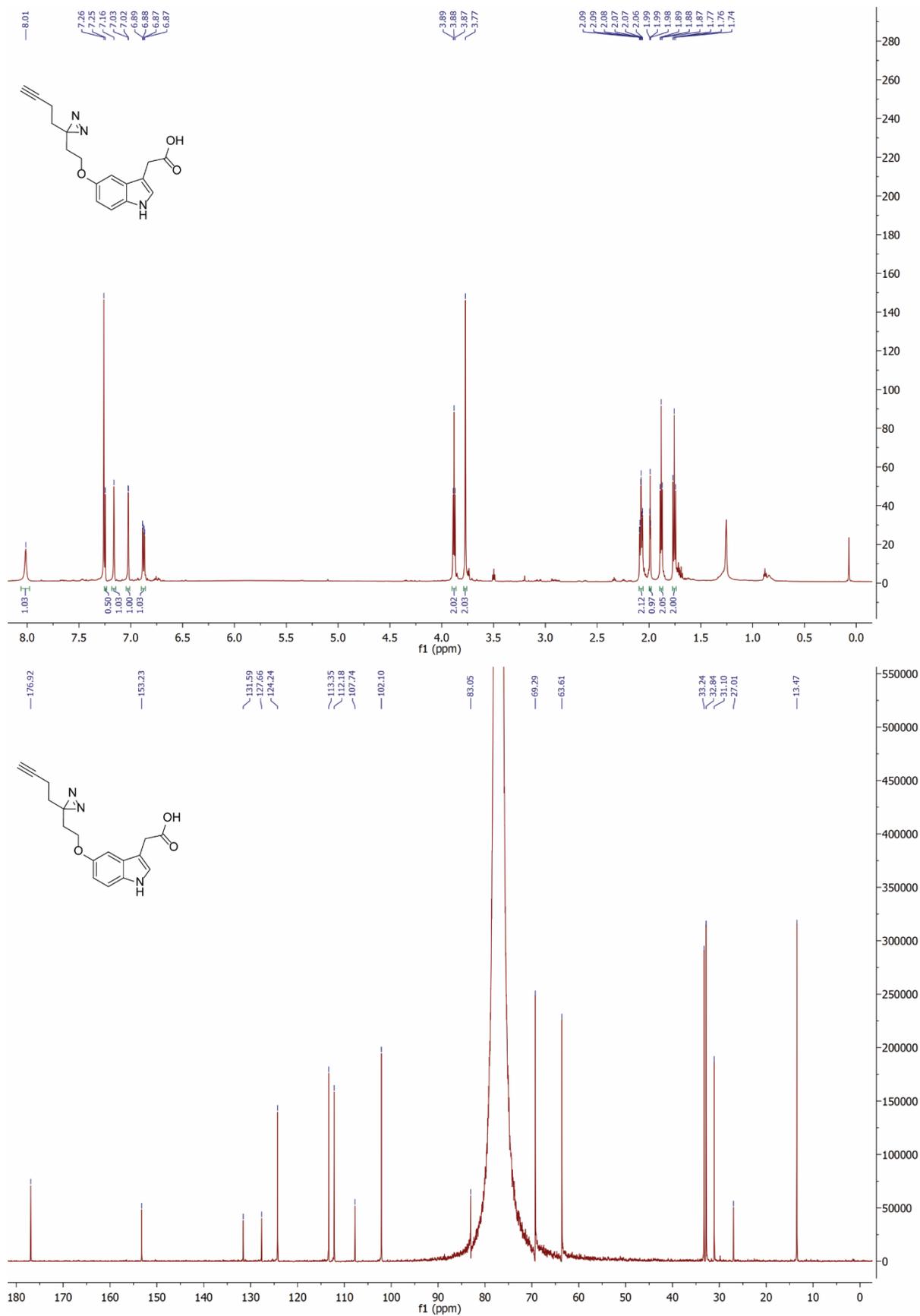
PTGPRC5A_F256A_F	ccgccaacgggtgggtggctctcctcgcttatgttt
PTGPRC5A_W224A_R	aaagtgatccaggccaccgcatagcaatactcagaagc
PTGPRC5A_W224A_F	gcttctgagtattgctatcgcggtggcctggatcacttt
PTGPRC5A_S220A_R	ggccaccagatagcaatagccagaagcatcgtagataa
PTGPRC5A_S220A_F	ttatctgacgatgcttctggctattgctatctgggtggcc
PTGPRC5A_N252A_R	gaaacaccaaccggcggcgagtgac
PTGPRC5A_N252A_F	gtgcactcgccgccgggtgggtgtttc
PTGPRC5A_N252A_F256A_R	cataagcgaggagagccaccaaccggcggcgagtgac
PTGPRC5A_N252A_F256A_F	gtgcactcgccgccgggtgggtggctctcctcgcttatg
GPRC5A_N252A_R	ggaacaccagccagcggcagccaaggcgg
GPRC5A_N252A_F	ccgcctggctgccgctgggtgggtgttc
GPRC5A_F256A_R	ctaacataagccaacagggccaccagccattggcagc
GPRC5A_F256A_F	gctgccaatggctgggtggcctgttgcttatgttag
GPRC5A_N252A_F256A_R	cataagccaacagggccaccagccagcggcagccaaggcgg
GPRC5A_N252A_F256A_F	ccgcctggctgccgctgggtgggtggcctgttgcttatg

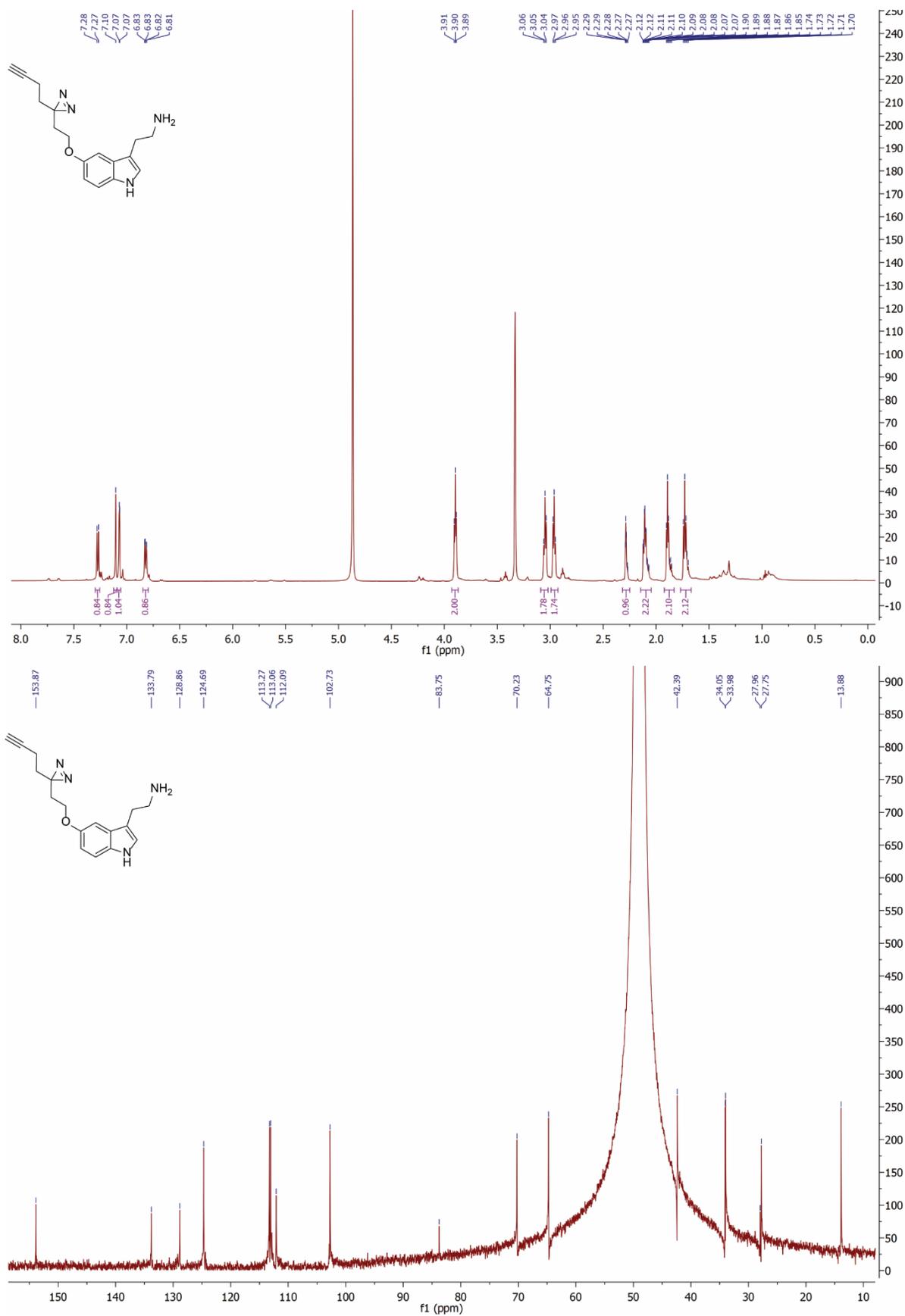
**Supplementary Table 2** Plasmids.

Plamid	Gene	Antibiotic resistance	Origin
pRecT	recT	Erythromycin	Our lab
pCas9-Efm_TyrDC	Cas9/sgRNA	Chloramphenicol	This study
pKH12-Efm_TyrDC	Efm_TyrDC-His6	Chloramphenicol	This study
pKH12-Rgs_TrpDC	Rgs_TrpDC-His6	Chloramphenicol	This study
pKH12-Mmi_Gln/TyrDC	Mmi_Gln/TyrDC-His6	Chloramphenicol	This study
pGL4.43_Luc2P/XRE/Hygro	3x XRE-Luc	Ampicillin	Promega E4121
pRL-TK	Renilla-Luc	Ampicillin	Promega E2241
pCMV3-N-HA-AhR	HA-AhR	Kanamycin	SinoBio HG10456-NY

pcDNA3.1_ADRA2A-Tango	Flag-ADRA2A-tTA	Ampicillin	Addgene 66216
pcDNA3.1_DRD2-Tango	Flag-DRD2-tTA	Ampicillin	Addgene 66269
pcDNA3.1_HTR4-Tango	Flag-HTR4-tTA	Ampicillin	Addgene 66412
pcDNA3.1_GPRC5D-Tango	Flag-GPRC5D-tTA	Ampicillin	Addgene 66385
pcDNA3.1_GPRC5C-Tango	Flag-GPRC5C-tTA	Ampicillin	Addgene 66384
pcDNA3.1_GPRC5B-Tango	Flag-GPRC5B-tTA	Ampicillin	Addgene 66383
pcDNA3.1_GPRC5A-Tango	Flag-GPRC5A-tTA	Ampicillin	Addgene 66382
pcDNA3.1_GPRC5A_D176A-Tango	Flag-GPRC5A_D176A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_D241A-Tango	Flag-GPRC5A_D241A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_D242A-Tango	Flag-GPRC5A_D242A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_D237A-Tango	Flag-GPRC5A_D237A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_F105A-Tango	Flag-GPRC5A_F105A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_F109A-Tango	Flag-GPRC5A_F109A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_F256A-Tango	Flag-GPRC5A_F256A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_W224A-Tango	Flag-GPRC5A_W224A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_S220A-Tango	Flag-GPRC5A_S220A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_N252A-Tango	Flag-GPRC5A_N252A-tTA	Ampicillin	This study
pcDNA3.1_GPRC5A_N252A_F256A-Tango	Flag-GPRC5A_N252A_F256A-tTA	Ampicillin	This study
pCMV3-GPRC5A	GPRC5A	Ampicillin	SinoBio HG12833-UT
pcDNA3_c-Flag	Flag	Ampicillin	Addgene 20011
pcDNA3_GPRC5A-Flag	GPRC5A-Flag	Ampicillin	This study
pcDNA3_GPRC5A_N252A-Flag	GPRC5A_N252A-Flag	Ampicillin	This study
pcDNA3_GPRC5A_F256A-Flag	GPRC5A_F256A-Flag	Ampicillin	This study
pcDNA3_GPRC5A_N252A_F256A-Flag	GPRC5A_N252A_F256A-Flag	Ampicillin	This study







## MATERIALS AND METHODS

**Chemicals.** Commonly used reagents were purchased from Sigma-Aldrich and Fisher Scientific. Specific chemicals were originated from following sources: 2,3,7,8-tetrachlorodibenzo-p-dioxin solution (TCDD, Sigma-Aldrich, Catalog #48599), indole-3-acetic acid (Fisher Scientific, Catalog #11453194), tryptamine (Sigma-Aldrich, Catalog #76706), 2-phenethylamine (Sigma-Aldrich, Catalog #42346), tyramine (Sigma-Aldrich, Catalog #80345), histamine (Sigma-Aldrich, Catalog #59964), 5-fluorotryptamine hydrochloride (AstaTech, Catalog #52030), 6-fluorotryptamine (AstaTech, Catalog #W10003), 7-fluorotryptamine hydrochloride (Cayman Chemical, Catalog #17323), 7-chlorotryptamine hydrochloride (Santa Cruz Biotechnology, Catalog #sc-268384), 7-bromotryptamine hydrochloride (Santa Cruz Biotechnology, Catalog #sc-268370), *N*-methyltryptamine (Cayman Chemical, Catalog #12067), (1*H*-Indol-3-yl)methanamine (Sigma-Aldrich, Catalog # CDS020852),  $\alpha$ -methyltryptamine hydrochloride (Cayman Chemical, Catalog #15692),  $\alpha$ -ethyltryptamine (Cayman Chemical, Catalog #15671), L(-)-epinephrine (Fisher Scientific, Catalog #AC204400010), clonidine hydrochloride (Tocris, Catalog #0690), pyridoxal-5'-phosphate (Oakwood Chemical, Catalog #044568), azido-PEG3-biotin (Click Chemistry Tools, Catalog #AZ104).

The synthesis of alk-IAA, x-alk-IAA and x-alk-TA was according to reported synthetic procedures.<sup>63,64</sup> The product was purified by silica gel flash column chromatography, determined by NMR (Bruker 600 MHz) and LC-HRMS (Orbi XL). alk-IAA: <sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 10.00 (s, 1H), 7.31-7.30 (d, 1H), 7.28 (s, 1H), 7.23-7.22 (d, 1H), 6.85-6.83 (dd, 1H), 4.76-4.75 (d, 2H), 3.71 (s, 2H), 2.99-2.98 (t, 1H); <sup>13</sup>C-NMR (150 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 173.14, 152.84, 133.19, 128.88, 125.51, 112.98, 112.75, 108.92, 103.82, 80.56, 76.30, 57.30, 31.39. HRMS (ESI) calculated for C<sub>13</sub>H<sub>11</sub>NO<sub>3</sub> [M+H]<sup>+</sup>: 230.0817, found 230.0816.

x-alk-IAA: <sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 8.31 (s, 1H), 7.26-7.25 (d, 1H), 7.16 (s, 1H), 7.03-7.02 (d, 1H), 6.89-6.87 (dd, 1H), 3.89-3.87 (t, 2H), 3.77 (s, 2H), 2.09-2.06 (m, 2H), 1.99-1.98 (t, 1H), 1.89-1.87 (t, 2H), 1.77-1.74 (t, 2H); <sup>13</sup>C-NMR (150 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 176.92, 153.23, 131.59, 127.66, 124.24, 113.35, 112.18, 107.74, 102.10, 83.05, 69.29, 63.61, 33.24, 32.84, 31.10, 27.01, 13.47. HRMS (ESI) calculated for C<sub>17</sub>H<sub>17</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 312.1348, found 312.1341.

x-alk-TA: <sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 7.28-7.27 (d, 1H), 7.10 (s, 1H), 7.07 (s, 1H), 6.83-6.81 (dd, 1H), 3.91-3.89 (t, 2H), 3.06-3.04 (t, 2H), 2.97-2.95 (t, 2H), 2.29-2.27 (m, 1H), 2.12-2.07 (m, 2H), 1.90-1.85 (m, 2H), 1.74-1.70 (m, 2H); <sup>13</sup>C-NMR (150 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 153.87, 133.79, 128.86, 124.69, 113.27, 113.06, 112.09, 102.73, 83.75, 70.23, 64.75, 42.39, 34.05, 33.98, 27.96, 27.75, 13.88. HRMS (ESI) calculated for C<sub>17</sub>H<sub>20</sub>N<sub>4</sub>O [M+H]<sup>+</sup>: 297.1715, found 297.1717.

**Cell Culture.** HEK293T and HT-29 cells were obtained from ATCC. HTLA cells (a HEK293 cell line stably expressing a tTA-dependent luciferase reporter and a  $\beta$ -arrestin2-TEV fusion gene) were a gift from the laboratory of Gilad Barnea at Brown University. HEK293T and HT-29 cells were cultured in DMEM (Gibco) supplemented with 10% FBS (GE Healthcare HyClone) and 100 U/mL penicillin-streptomycin

(ThermoFisher). HTLA cells were cultured in DMEM supplemented with 10% FBS and 100 U/mL penicillin-streptomycin, 2 µg/mL puromycin and 100 µg/mL hygromycin B. All cells were cultured in a 37 °C incubator with 5% CO<sub>2</sub>.

**Luciferase reporter assay.** HEK293T cells were seeded in CoStar white 96-well plate (Corning) at 20,000 cells/well for 16 h. The cells were transiently transfected with three plasmids including pCMV3-N-HA-AhR (Sinobiological, HG10456-NY), pGL4.43[luc2P/XRE/Hygro] (Promega, E4121) and pRL-TK (Promega, E2241) using Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000-015). After 24 hours, cells were stimulated with 0.1% DMSO, 10 nM TCDD, 100 µM IAA, alk-IAA or x-alk-IAA for 16 hours. Luminescence was detected after addition of dual-luciferase reporter assay reagent (Promega, E1960), using a BioTek Synergy Neo2 Multi-Mode Microplate Reader with a 0.5 second integration time. Firefly luciferase luminescence normalized to the Renilla luciferase control is shown, with error bars indicating the S.E.M. for three replicates.

**In-gel fluorescence and Western blot.** For in-gel fluorescent labeling, HEK293T cells were grown in 6-well plate at 500,000 cells/well to above 90% confluence before transfection with respective plasmids pCMV3-N-HA-AhR, pcDNA3.1\_ADRA2A-Tango (Addgene 66216), pcDNA3.1\_GPRC5A-Tango (Addgene 66382), pcDNA3\_GPRC5A-Flag, pcDNA3\_GPRC5A\_N252A-Flag, pcDNA3\_GPRC5A\_F256A-Flag, or pcDNA3\_GPRC5A\_N252A\_F256A-Flag. After 24 h, cells were change to OMEM medium containing 0.1% DMSO, 100 µM x-alk-IAA, or 100 µM x-alk-TA, and incubated for 1 h at 37 °C. Then the cells were washed and incubated in cold DPBS on ice, exposed to 365 nm UV light for 10 min. Cells were lysed in 500 µL of lysis buffer including 20 mM HEPES buffer, pH 7.5, 150 mM NaCl, 1% NP-40, 1x cOMplete protease inhibitor cocktail (Sigma-Aldrich) for 15 min. Cell lysates were cleared by centrifugation at 20,000g for 20 min, and the soluble proteins were incubated with EZview red anti-HA or anti-Flag affinity gel (Sigma) for 3 h at 4 °C. The immunoprecipitated proteins were labeled with azido-rhodamine via click chemistry, resolved by SDS-PAGE, and visualized on a Bio-Rad ChemiDoc MP imaging system. For in-gel fluorescent profiling, HT-29 cells were cultured in 6-well plate at 500,000 cells/well to above 90% confluence, treated with 0.1% DMSO, 100 µM alk-IAA, 100 µM x-alk-IAA, or 100 µM x-alk-TA with or without UV irradiation. Cells were lysed in 1% SDS lysis buffer, labelled by azido-rhodamine, resolved by SDS-PAGE and visualized. For Western blot analysis, cell lysates prepared as above were labeled with azido-PEG3-biotin (Click Chemistry Tools, AZ104), incubated with high-capacity neutravidin beads (Thermo Scientific Pierce), resolved by SDS-PAGE and analyzed by Western blot. Antibodies for ADRA2A (14266-1-AP), GPRC5A (10309-1-AP), GPR107 (25076-1-AP), GPR108 (24009-1-AP), STING/TMEM173 (19851-1-AP), MAVS (14341-1-AP), SLC16A3/MCT4 (22787-1-AP) were purchased from Proteintech, SCL18B1 (bs-9530) and SLC38A2 (bs-12125R) from Bioss.

**Chemical proteomics.** HT-29 cells were grown in 10-cm dish to above 90% confluence, treated with DMSO, alk-IAA or x-alk-IAA with or without UV-irradiation, lysed and labeled with azido-PEG3-biotin as shown above. The proteins enriched by neutravidin beads were digested with MS-grade trypsin (Promega). The on-beads digested peptides were collected and desalted by the in-house prepared reversed phase C18 Stage-tip, dried by Speed-vac, redissolved in 0.1% TFA/H<sub>2</sub>O. Tryptic peptides from each of the 4 replicates were separated by the reversed phase C18 column on an EASY-nLC 1000 (Thermo Scientific) system at a flow rate of 200 nL/min using a 70 min gradient from 5% B to 45% B (A: 0.1% formic acid, B: acetonitrile/0.1% formic acid) followed by a washing step at 100% B. Mass spectra were collected on an LTQ-Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Full scan MS spectra were acquired in the Orbitrap at a resolution of 60,000 (at m/z 300). The raw data were analyzed using MaxQuant (version 1.6.0.16)<sup>65</sup>. Parameters were used as default settings for label-free quantification (LFQ). For protein identification, the false discovery rate was set to 0.01, the minimum peptide length allowed was seven amino acids, and the minimum number of unique peptides allowed was set to one. Further data processing was performed using Perseus (version 1.6.0.7).<sup>66</sup> LFQ intensities were log<sub>2</sub> transformed. The data were filtered to remove reverse contaminants, only identified by sites. Proteins were filtered to be detected in at least 3 valid values out of 8 experiments. The replicates were grouped together, and empty values were imputed with random numbers from a normal distribution (width = 0.3 and shift = 1.8). Significantly enriched proteins were identified as having a p-value below 0.05 and at least two folds change when comparing LFQ intensities of chemical reporter-treated group with DMSO vehicle group. The data was visualized in a volcano plot of log<sub>2</sub> ratio change versus the negative logarithmic p values. Gene ontology analysis of the protein hits was performed by Enrichr.

**PRESTO-Tango assay.** The PRESTO-Tango plasmids were a gift from Bryan Roth (Addgene 1000000068).<sup>43</sup> HTLA cells were seeded in a CoStar white 96-well plate (Corning) at 20,000 cells/well for 16 h. The cells were transiently transfected with individual GPCR-Tango plasmid using Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000-015). After 24 h, cells were stimulated with various molecules of different concentration for 16 h. Cell culture medium were changed with 50  $\mu$ L Bright-Glo luciferase assay system reagent (Promega) that was firstly diluted 20 folds with 1x DPBS buffer. After incubation for 20 min at room temperature in dark, luminescence was detected using a BioTek Synergy Neo2 Multi-Mode Microplate Reader with a 0.5 second integration time.

**Bacterial culture.** Some commercially available bacteria species and strains are sourced as follows: *E. coli* DH5a (NEB), *E. faecalis* OG1RF (ATCC 47077), *E. faecalis* PCI 1325 (ATCC 14506), *E. faecalis* V583 (ATCC 700802), *E. faecium* DO/TX0016 (ATCC BAA-472), *E. faecium* (ATCC 700221), *E. faecium* TX0082 (BEI Resources), *E. faecium* ERV165 (BEI Resources), *E. durans* 23C2 (ATCC 6056), *E. hirae* R (ATCC 8043), *E. mundtii* NCDO 2375 (ATCC 43186), *E. gallinarum* NCDO 2313 (ATCC

51559), *M. morgani* ATCC 25830 (Microbiologics), *S. capitis* ATCC 35661 (Microbiologics), *S. epidermidis* ATCC 14990 (Microbiologics), *L. brevis* ATCC 14869 (Microbiologics). *E. faecium* Com15 was provided as a gift by Michael Gilmore (Harvard Medical School), *R. gnavus* ATCC 29149 by Sean F. Brady (The Rockefeller University), *C. sporogenes* ATCC 15579 by Chun-Jun Guo (Weill Cornell Medical School). Bacteria were cultured in the anaerobic chamber (Coy Laboratory Products) at 37 °C for 48 h or bacterial shaker incubator of 220 rpm at 37 °C for 16 h in LB, BHI, TSB or MRS media respectively. Bacterial growth was determined by optical density measurement. Bacterial cultures were pelleted by centrifugation at 5,000 g, supernatants were filtered through 0.2 µm membrane (Millipore) followed by 3,000 Da MWKO membrane (Amicon).

**Phylogenetic analysis.** The analysis of TyrDC orthologs across different species in the human microbiome was conducted as previously described.<sup>9</sup> Briefly, GenBank coding sequences from genome FASTA assemblies within the Human Microbiome Project<sup>47</sup> were downloaded via BioProject Accession number PRJNA28331 and queried by BLASTP using TyrDC *E. faecium* Com15 (AZV35689.1). For genomes with multiple hits, the hit sequence with the highest percent identity with the query sequence was chosen as the TyrDC ortholog. Separately, GenBank genomic assemblies were queried for 16S rRNA sequences using Barrnap (GitHub.com/tseemann/barrnap). The longest predicted 16S rRNA sequence for each strain was used for alignment by Clustal Omega<sup>67</sup> using default settings. The multiple sequence alignment was trimmed by TrimAl v1.3 (automated method)<sup>68</sup>, and an approximate maximum-likelihood phylogenetic tree was constructed using FastTree v2.1.10<sup>69</sup>. Data were visualized as a cladogram by iTOL.<sup>70</sup> For the unrooted phylogenetic clustering, protein sequences were aligned by Clustal Omega with output in FASTA format. The FASTA alignment was then uploaded to the IQ-TREE tool using default settings with the auto substitution model, and 10,000 bootstrap alignments were calculated using the ultrafast setting<sup>71</sup>. The resulting phylogenetic tree file was visualized using iTOL.

**LC-MS analysis.** Chemicals and membrane-filtered bacterial cultures were analyzed by 1290 Infinity II LC/MSD system (Agilent technologies) using Poroshell 120 EC-C18 column (2.1 x 50 mm, 4 µm). The flow rate was 0.5 mL/min using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B at room temperature. The gradient was applied as below: 0-2 min: 5% B, 2-5 min: 5-95% B, 5-7 min: 95% B, 7-10 min: 95-5% B. MSD API-ES SIM mode on sample target masses of tryptamine (m/z = 161.1), tyramine (m/z = 138.1), phenethylamine (m/z = 122.1), histamine (m/z 112.1). Extracted intensity and retention time were analyzed by GraphPad Prism 8.

**Generation of TyrDC-KO *E. faecium* Com15 strains.** TyrDC knockout of *E. faecium* Com15 was generated using CRISPR-Cas9 mediated recombineering as previously described.<sup>48</sup> In brief, pRecT containing *E. faecium* Com15 was co-electroporated with

ssDNA (Integrated DNA technologies) encoding homology arms targeting *TyrDC*, and pCas9 harboring gRNA targeting the *TyrDC* deletion region. The resulting knockout strain carries a 74 base-pair deletion in the *TyrDC* gene, verified via PCR and sanger sequencing (Genewiz). The strain was then cured of pRecT and pCas9 through passaging to generate isogenic mutants. Single colony was inoculated for preparation of glycerol stock. pKH12-*TyrDC* was generated by cloning his-tagged *E. faecium TyrDC* into psrtA1 with *srtA-1* omitted from the plasmid. The resulting plasmid contains a *TyrDC* under the constitutive *bacA* promoter in a pKH12 backbone<sup>72</sup>. pKH12-*TyrDC* or vector pKH-12 was transformed into *E. faecium* Com15 through electroporation. Single colony was inoculated and cultured in BHI media for 16 h. Cleared bacterial cultures were filtered, analyzed by LC-MS and PRESTO-Tango assay.

**Cloning and analysis of aAAs decarboxylase.** *TyrDC* was amplified from *E. faecium* Com15 genomic DNA, *TrpDC* from *R. gnavus* ATCC 29149, *Gln/TyrDC* from *M. morgani* ATCC 25830, and inserted to pKH12 vector by Gibson assembly (NEB), respectively. The plasmids were transformed into chemical competent *E. coli* DH5 $\alpha$  (NEB). Single colony was inoculated in 5 mL of LB medium containing 25  $\mu$ g/mL chloramphenicol for 16 h, and was then transferred to into Minimal media supplemented with 10 mM Trp, 2 mM Tyr, 10 mM Phe or 10 mM His of a OD<sub>600</sub> value of 0.1. After 16 h, bacteria were pelleted, the supernatants were filtered through 0.2  $\mu$ m membrane followed by 3,000 Da MWKO membrane. The filtered cultures were analyzed by LC-MS and PRESTO-Tango assay.

**Analysis of *E. faecium* Com15 TyrDC enzymatic activity.** For preparation of TyrDC, *E. coli* DH5 $\alpha$  transformed with pKH12-*TyrDC* were grown in 500 mL LB medium containing 25  $\mu$ g/mL chloramphenicol starting from OD<sub>600</sub> of 0.1 for 24 h at 37 °C with shaking at 220 rpm. Cells were pelleted, resuspended in 40 mL 50 mM Tris pH 8 containing 0.25 M NaCl, lysed by sonication. The lysate was clarified by centrifugation, incubated with 2 mL of Ni-NTA resin at 4 °C for 16 h. The resins were loaded on a column, washed three times with lysis buffer, eluted using a gradient of 25 mM to 200 mM imidazole in lysis buffer. Fractions containing pure protein were combined and dialyzed over two rounds into in 50 mM Tris pH 8 containing 0.20 M NaCl and 10% w/v glycerol. The dialyzed protein was concentrated using spin columns, concentration was determined by BSA assay, protein aliquots were frozen in liquid nitrogen and stored at – 80 °C.

Purified TyrDC was incubated with pyridoxal-5'-phosphate (PLP) in the reaction buffer (200 mM sodium acetate buffer, pH 5.5) for five minutes. The enzymatic reaction was initiated by adding the preincubated TyrDC-PLP mix with substrates dissolved in the reaction buffer to yield the final concentrations for 100  $\mu$ M PLP, 100 nM enzyme and 10  $\mu$ M-2000  $\mu$ M tyrosine or 1 mM-60 mM phenylalanine at room temperature for 1 min or 37 °C for 10 min, respectively. Aliquots of the reaction were quenched by diluting 10-fold into ice cold methanol, centrifugated and 2  $\mu$ l of the supernatant was analyzed. Rates were calculated and fit to a standard Michaelis–Menten kinetics curve in Graphpad prism 8.

