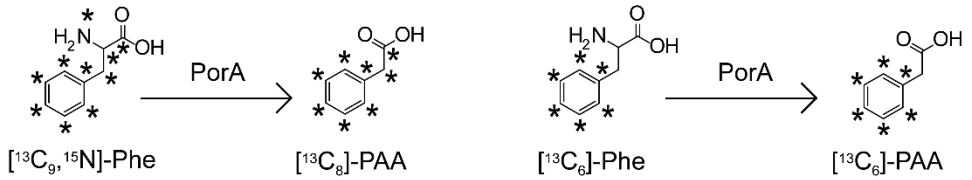
Cell Host & Microbe, Volume *31*

Supplemental information

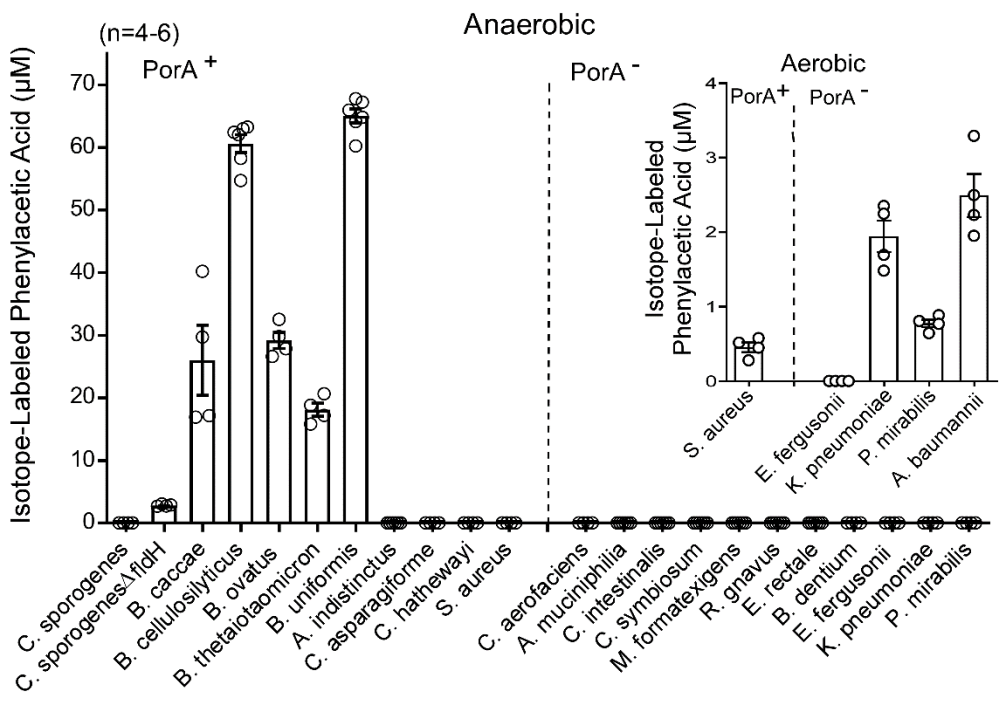
Two distinct gut microbial pathways contribute to meta-organismal production of phenylacetylglutamine with links to cardiovascular disease

Yijun Zhu, Mohammed Dwidar, Ina Nemet, Jennifer A. Buffa, Naseer Sangwan, Xinmin S. Li, James T. Anderson, Kymberleigh A. Romano, Xiaoming Fu, Masanori Funabashi, Zeneng Wang, Pooja Keranahalli, Shawna Battle, Aaron N. Tittle, Adeline M. Hajjar, Valentin Gogonea, Michael A. Fischbach, Joseph A. DiDonato, and Stanley L. Hazen

A


PorA
PorA
[13C9,15N]-Phe
[13C8]-PAA
[13C6]-Phe
[13C6]-PAA

B


(n=4-6)
Anaerobic
PorA +
PorA -
Isotope-Labeled Phenylacetic Acid (μM)
Aerobic
PorA+
PorA -
Isotope-Labeled Phenylacetic Acid (μM)
C. sporogenes
C. sporogenesΔfldH
B. caccae
B. cellulosilyticus
B. ovatus
B. thetaiotaomicron
B. uniformis
A. indistinctus
C. asparagiforme
C. hathewayi
S. aureus
C. aerofaciens
A. muciniphilia
C. intestinalis
C. symbiosum
M. formatexigens
R. gnavus
E. rectale
B. dentium
E. fergusonii
K. pneumoniae
P. mirabilis
S. aureus
E. fergusonii
K. pneumoniae
P. mirabilis
A. baumannii

C

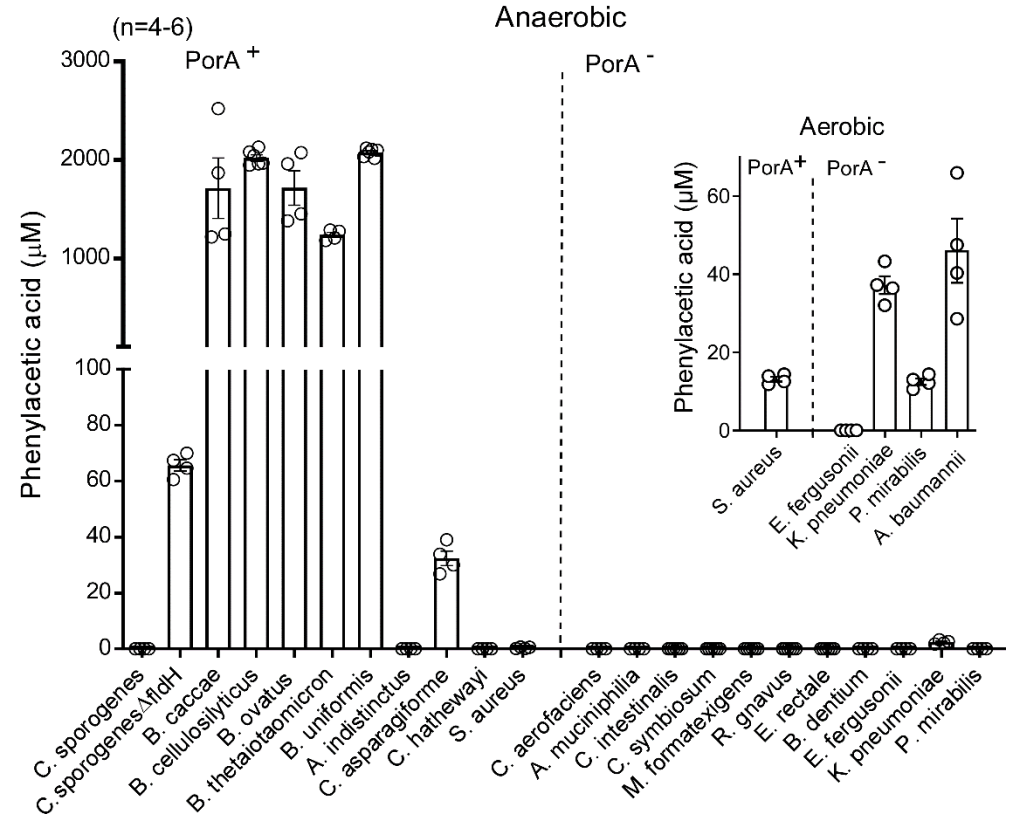

(n=4-6)
Anaerobic
PorA +
PorA -
Phenylacetic acid (μM)
Aerobic
PorA+
PorA -
Phenylacetic acid (μM)
C. sporogenes
C.sporogenesΔfldH
B. caccae
B. cellulosilyticus
B. ovatus
B. thetaiotaomicron
B. uniformis
A. indistinctus
C. asparagiforme
C. hathewayi
S. aureus
C. aerofaciens
A. muciniphilia
C. intestinalis
C. symbiosum
M. formatexigens
R. gnavus
E. rectale
B. dentium
E. fergusonii
K. pneumoniae
P. mirabilis
S. aureus
E. fergusonii
K. pneumoniae
P. mirabilis
A. baumannii

Fig. S1. Screening for isotope-labeled and unlabeled phenylacetic acid (PAA) production among diverse panel of human commensal bacteria, related to main Figure 1.

PAA and [$^{13}C_8$]- or [$^{13}C_6$]-PAA production from phenylalanine (Phe) and [$^{13}C_9$,^{15}N]-Phe or [$^{13}C_6$]-Phe by human commensals grouped by the presence of PorA (ClOSPO_00147, or related 2-oxoacid:ferredoxin oxidoreductases) protein homologues. All cultures were incubated under anaerobic conditions for 48 h before the supernatants were harvested and processed for LC-MS/MS analyses of the respective isotope-labelled PAA. For the facultative anaerobes, parallel experiments were performed under aerobic conditions. For *A. baumannii*, incubation was done under aerobic conditions only. (**A**) Structure of [$^{13}C_9$,^{15}N]-Phe and [$^{13}C_6$]-Phe and the corresponding PAA isotopologues. (**B** and **C**) The production of both isotope-labeled and unlabeled PAA, respectively. n= 4~6 for each strain.

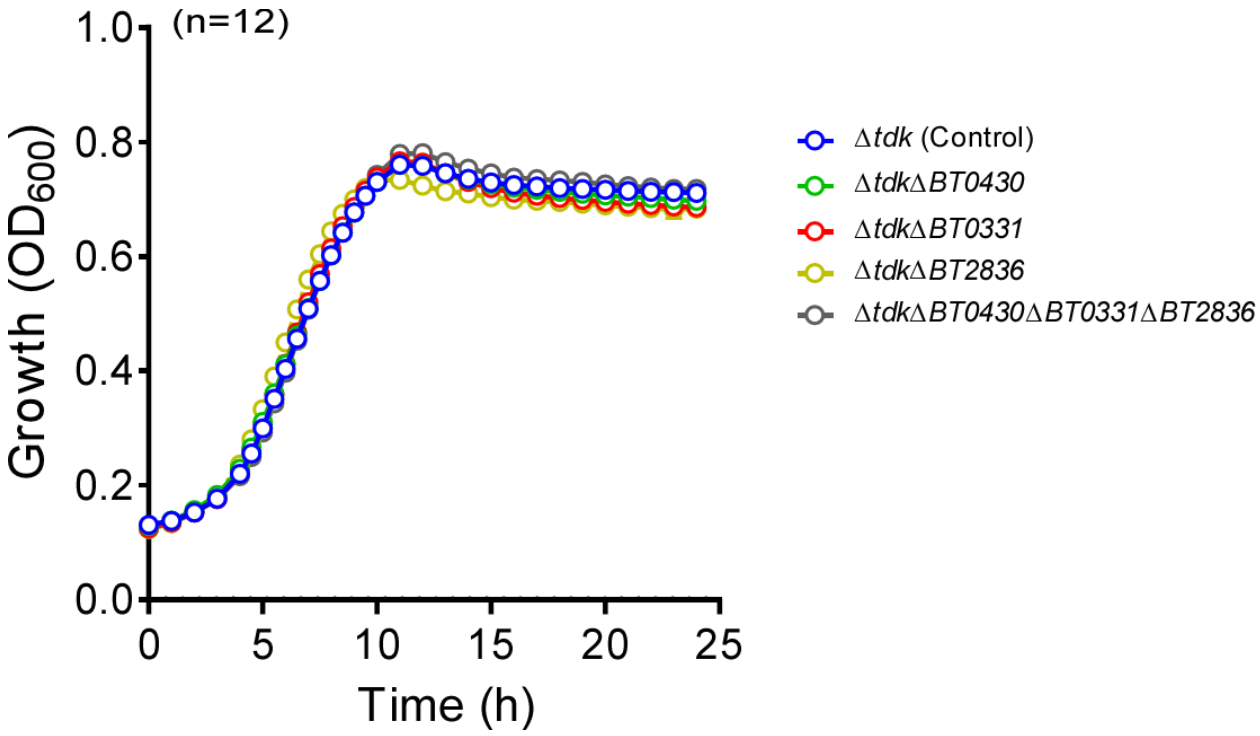
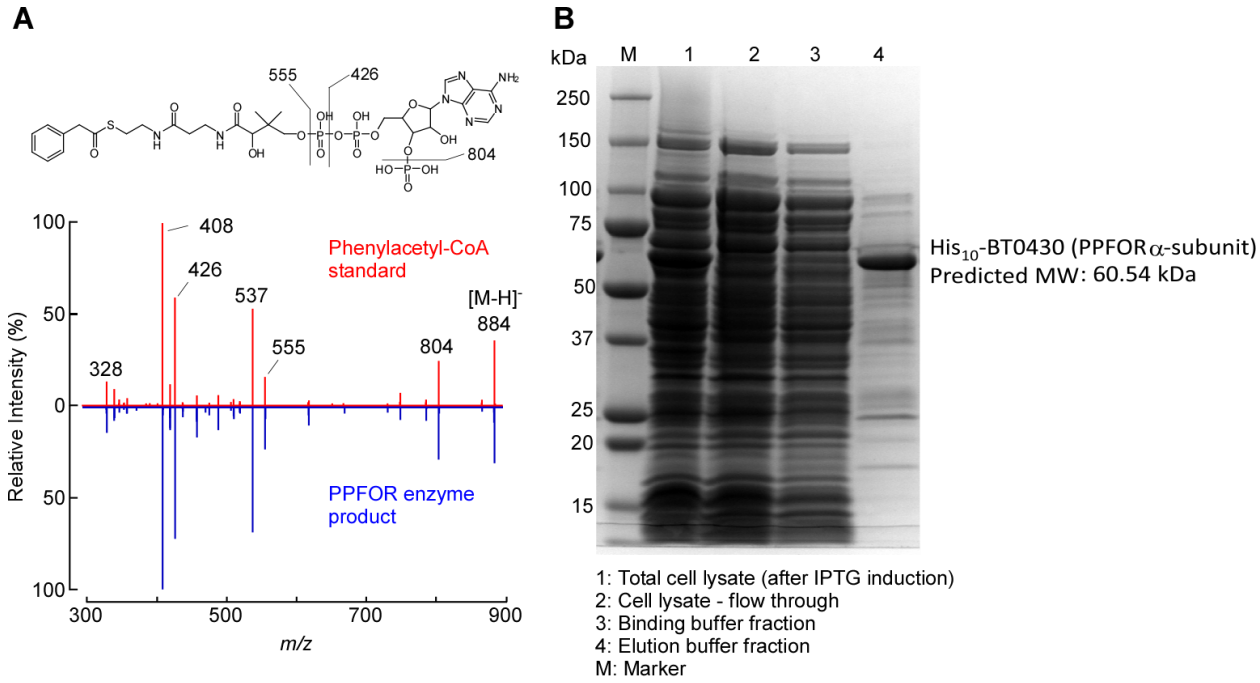


Fig. S2. Growth curve of *B. thetaiotaomicron* mutants, related to main Figure 1.

B. thetaiotaomicron Δ*tdk*, Δ*tdk*Δ*BT0430,* Δ*tdk*Δ*BT0331,* Δ*tdk*Δ*BT2836* and Δ*tdk*Δ*BT0430*Δ*BT0331*Δ*BT2836* knockout mutants were cultured in BHI media supplemented with 100 μM [$^{13}C_9$,^{15}N]-Phe at 37 °C under anaerobic condition in 96-well plate. For each mutant, 4 independent cultures were prepared, and each was subcultured in 3 wells. Data points represent the mean±SE from the 12 wells for each mutant.



C

In-solution digestion of *E. coli* cell lysate containing recombinant PPFOR (BT0430/BT0429)

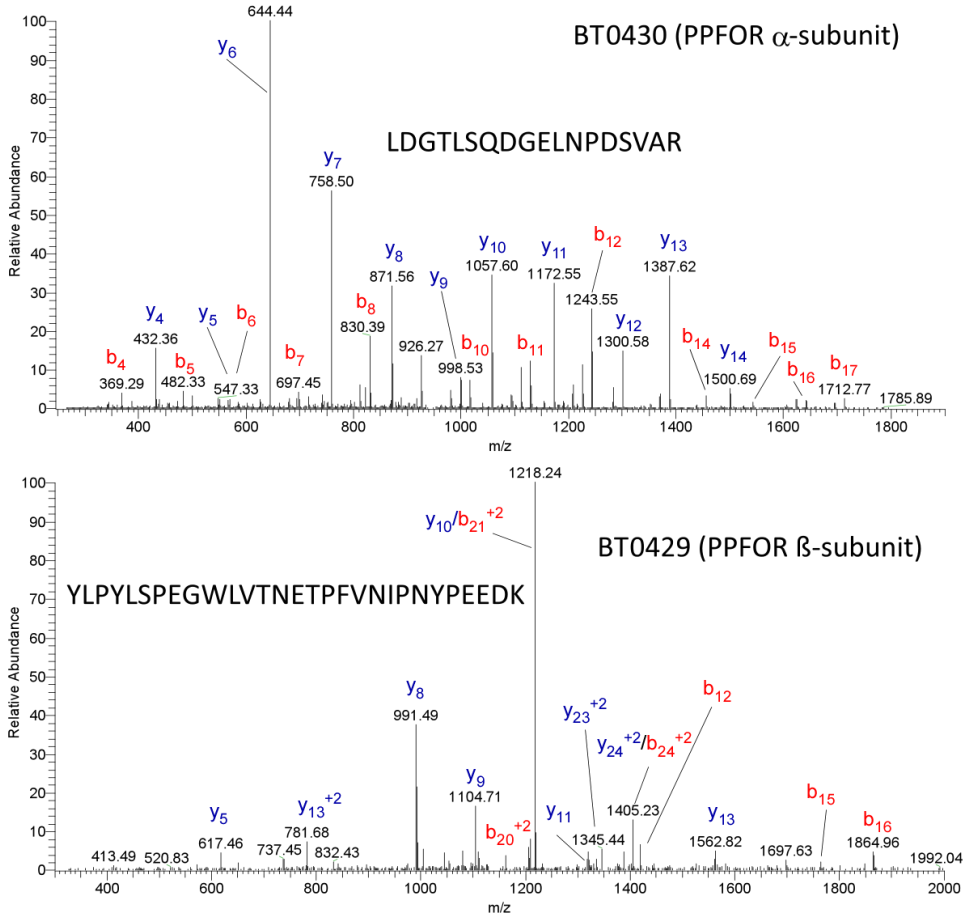


Fig. S3. Phenylacetyl-CoA is the product of phenylpyruvate:ferredoxin oxidoreductase (PPFOR) on phenylpyruvate, related to main Figure 1.

(A) Collision-induced dissociation (CID) spectra of phenylacetyl-CoA standard (red) and the enzymatic product of the recombinant PPFOR (blue).

(B) Coomassi-blue staining and imaging of denaturing SDS-PAGE gel showing the induced *E. coli* cell lysate containing the recombinant His_{10}-tagged α-subunit of PPFOR before and after purification using Ni-NTA column. Both α and β subunits of *B. thetaiotaomicron* PPFOR (BT0430/BT0429) were expressed in *E. coli* BL21 (DE3), but only BT0430 (α-subunit) was labelled on its *N*-terminus with His_{10}-tag and therefore can be observed on the gel after Ni-NTA column purification.

(C) Trypsin digestion of the cut BT0430 gel band and in-solution trypsin digestion of the induced *E. coli* cell lysate followed by LC-MS/MS proteomics analysis confirmed the expression of both α and β subunits in the induced *E. coli*. A total of 24 and 2 peptides corresponding to recombinant BT0430 and BT0429, respectively, were observed in the trypsin digestion of the cut BT0430 gel band, and the induced *E. coli* cell lysate (see also Table S4). The MS spectrum of one representitive peptide from each subunit is shown.

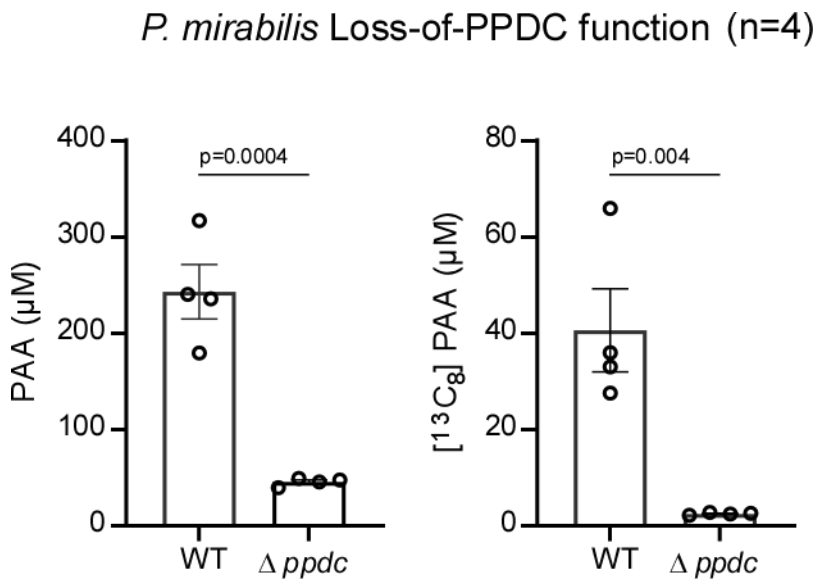
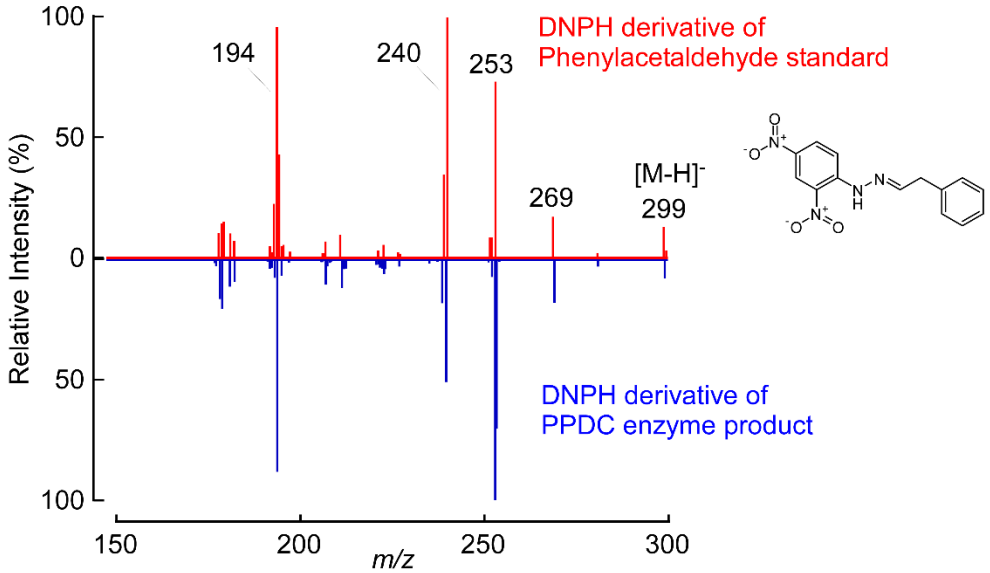


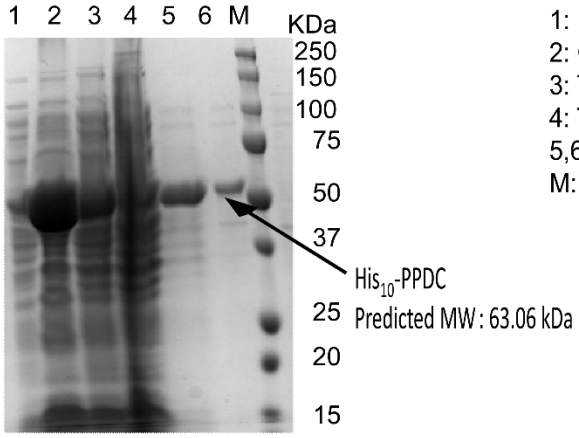
Fig. S4. Phenylacetic acid (PAA) and [$^{13}C_8$]-PAA production by *P. mirabilis* wild-type and Δ*HMPREF0693_2975* mutant (Δ*ppdc*) in LB media supplemented with [$^{13}C_9$,^{15}N]-phenylalanine under aerobic condition, related to main Figure 3.

Supernatant was sampled after 48 hours of incubation. PAA and [$^{13}C_8$]-PAA were quantified using LC-MS/MS. Data points represent the mean±SE from four independent replicates. Significance was determined using Student's *t*-test.

A



B



1: Binding buffer fraction
2: Clear cell lysate flow through
3: Total cell protein, induced by IPTG
4: Total cell protein, not induced
5,6: Eluted fraction
M: Marker

C

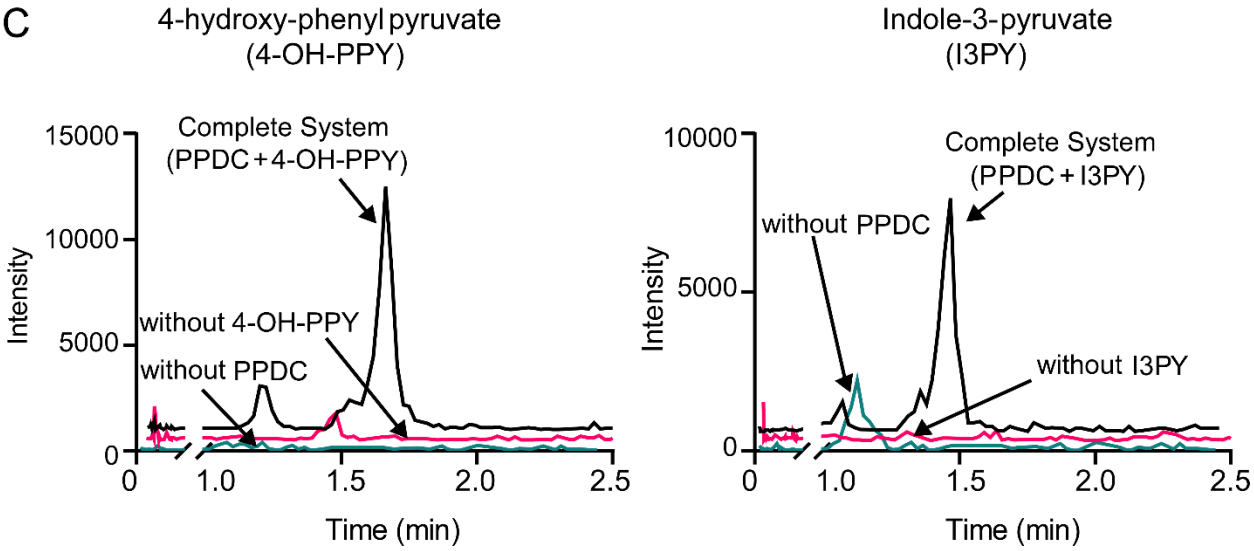


Fig. S5. Phenylacetaldehyde is the product of phenylpyruvate decarboxylase (PPDC) on phenylpyruvate while 4-hydroxyphenylacetaldehyde and indoleacetaldehyde are the products of 4-hydroxyphenylpyruvate (4-OH-PPY) and indole-3-pyruvate (I3PY), respectively, related to main Figure 3.

(A) CID fragmentation spectra of the 2,4-dinitrophenylhydrazone (DNPH) derivative of phenylacetaldehyde standard (red) and the enzymatic product of the recombinant PPDC (blue).

(B) Coomassi blue staining and imaging of denaturing SDS-PAGE gel showing the induced *E. coli* BL21 (DE3) cell lysate containing the recombinant His_{10}-tagged PPDC before and after purification using Ni-NTA column.

(C) 4-Hydroxyphenylacetaldehyde and indoleacetaldehyde were derivatized with DNPH and the signature parent-daughter ion transitions for their hydrazones were monitored by LC-MS/MS: *m/z* 315.2→149.1 for 4-OH-phenylacetaldehyde-DNPH, and *m/z* 338.3→172.2 for indoleacetaldehyde-DNPH derivatives.

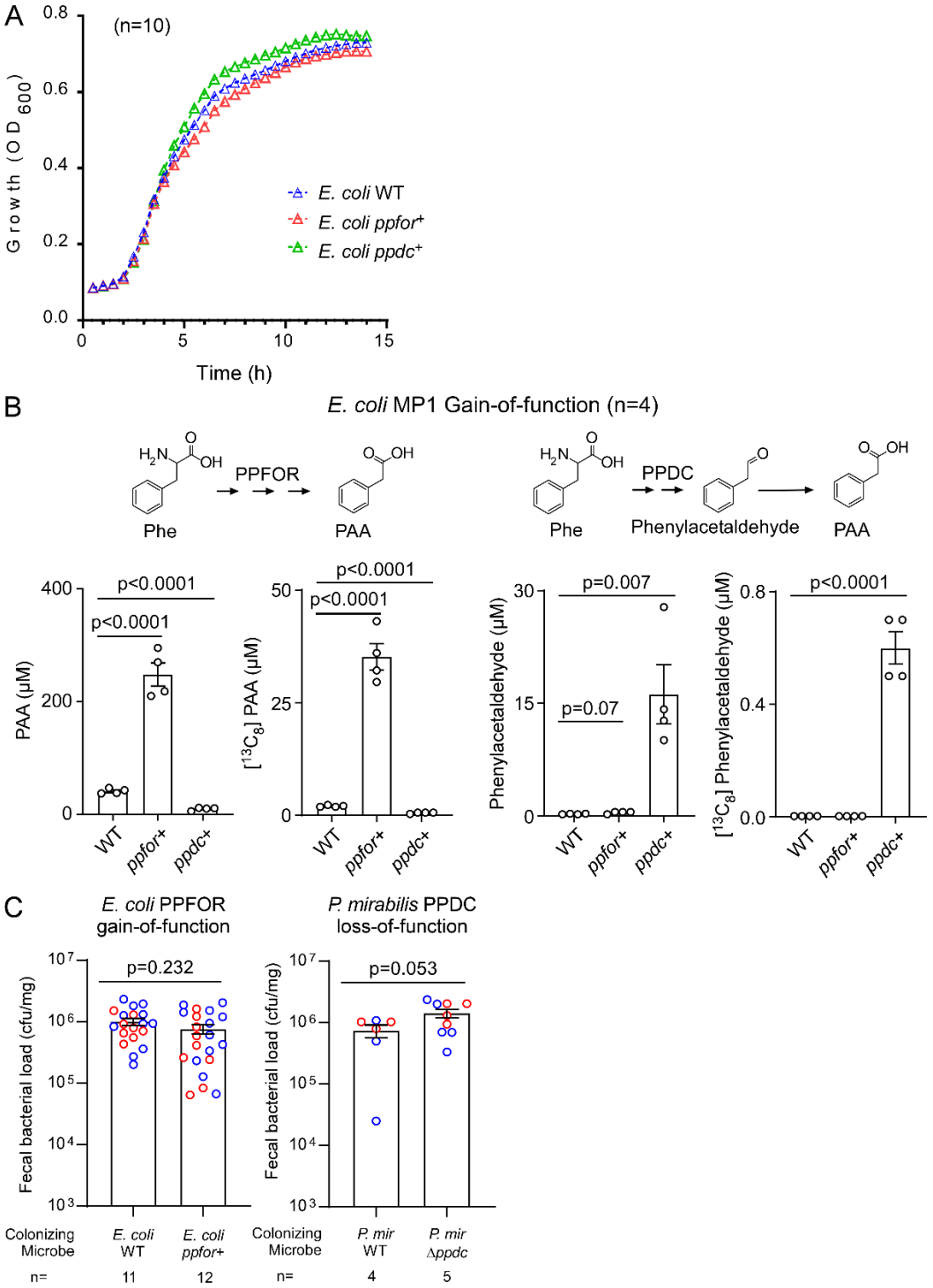


Fig. S6. PPFOR and PPDC gain-of-function in *E. coli* MP1 and colonization efficiency of wild-type strains and mutants of *E. coli* MP1, and *P. mirabilis* in gnotobiotic mice studies, related to main Figure 5.

(A) The two knock-in mutants *E. coli ppfor*$^{+}$, and *E. coli ppdc*$^{+}$ showed similar growth rate as the wild-type *E. coli* MP1. All three strains were cultured in LB media aerobically in a 96 well-plate and their growth was monitored over time. Data points represent the mean±SE from 10 wells for each.

(B) Chromosomal expression of *BT0429/BT0430* genes (The two subunits for *B. thetaiotaomicron* PPFOR enzyme) in *E. coli* MP1 but not *HMPREF0693_2975* (coding for *P. mirabilis* PPDC enzyme) enabled the full conversion of phenylalanine (Phe) to phenylacetic acid (PAA). Similar results were obtained when following isotopically labelled phenylalanine. PPDC knock-in, however, enabled the conversion of phenylalanine to phenylacetaldehyde. Data is shown as individual data points and mean±SE. Significance was determined using Student's *t*-test.

(C) Germ-free mice were randomized and inoculated with either wild-type *E. coli* MP1 versus its gain-of-function mutant *E. coli ppfor*$^{+}$, while in an independent experiment, germ-free mice were inoculated with wild-type *P. mirabilis* or its loss-of-function mutant *P. mirabilis* Δ*HMPREF0693_2975* (Δ*ppdc*) as explained in the methods. Fecal samples were taken at day 4 and day 7 after colonization, weighed, diluted then plated on LB agar plates for counting. Day 4 and day 7 samples are shown as blue and red open circles, respectively. "n" represents the number of mice in each group. For some mice, fecal samples were only able to be collected at one time point (either 4 or 7 days). Bars represent mean±SE. Significance was determined using Student's *t*-test.

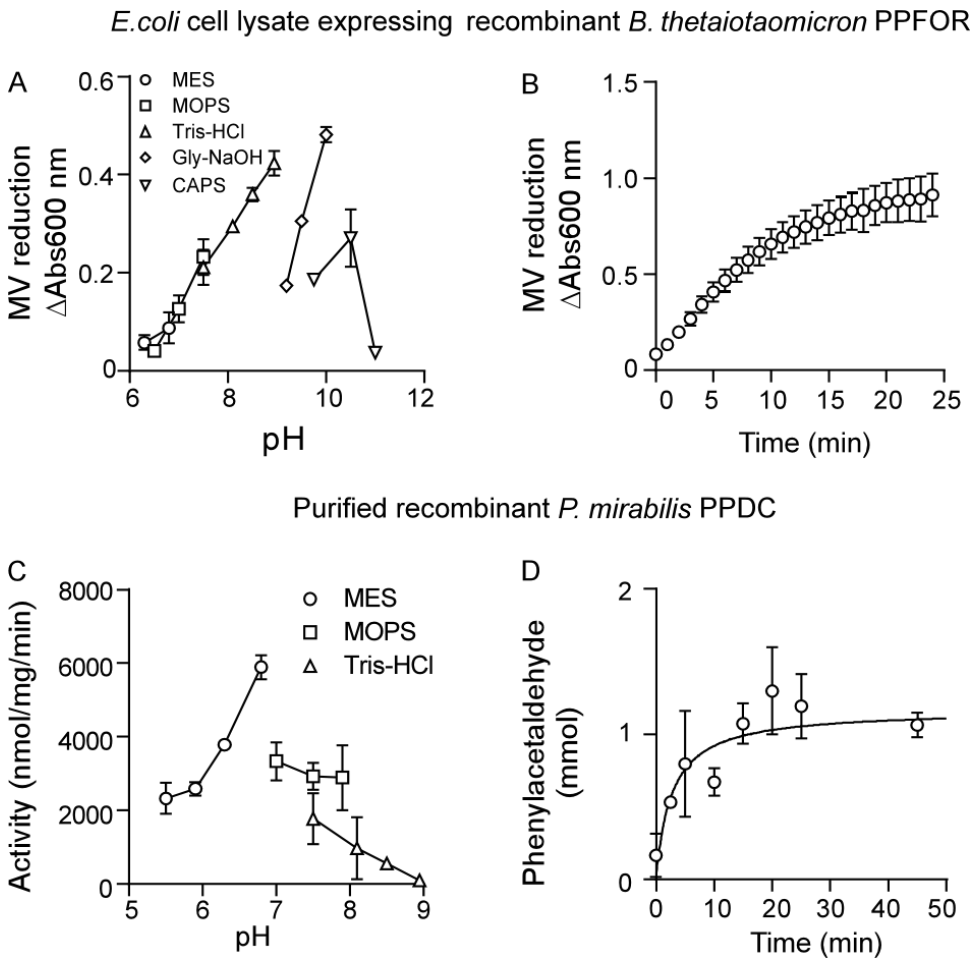


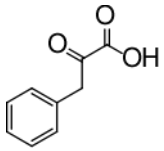
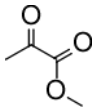
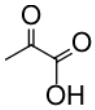
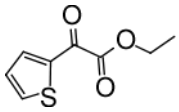
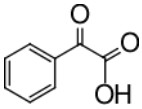
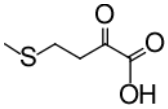
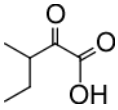
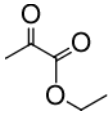
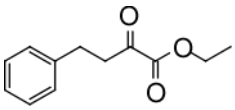
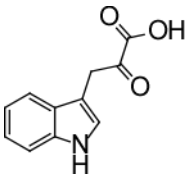
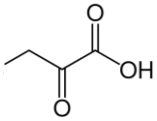
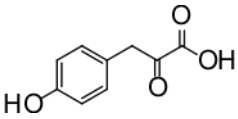
Fig. S7. Optimization of recombinant PPFOR and PPDC enzymatic reaction condition, related to main Figures 1 and 3.

(A, B) Optimization of reaction conditions for induced *E. coli* BL21 (DE3)/ pET16b-PPFOR cell lysate harboring the recombinant PPFOR. (**A**) Enzyme activity at different pH and buffer system. (**B**) Reaction time course for ~50 µg of induced *E. coli* BL21 (DE3)/ pET16b-PPFOR cell lysate in 100 µl of 100 mM Glycine-NaOH buffer, pH 10.5.

(C, D) Optimization of reaction conditions for the recombinant *P. mirabilis* PPDC after expressing in and purifying from *E. coli* BL21 (DE3). (**C**) Enzyme activity at different pH and buffer system. (D) Reaction time course for ~2 µg of PPDC in 100 µl of MES buffer, pH 6.8.

All buffers were used at 100 mM final concentration. MOPS, 3-morpholin-4-ylpropane-1-sulfonic acid; CAPS, N-cyclohexyl-3-aminopropanesulfonic acid; MES, 2-(N-morpholino) ethanesulfonic acid.

Table S2. Substrate specificity of recombinant PPFOR and PPDC, related to main Figures 1, and 3.

Structure/name	Relative activity (%)		Structure	Relative activity (%)	
	Rec. PPFOR (*B.thetaiotaomicron*)	Rec. PPDC (*P. mirabilis*)		Rec. PPFOR (*B.thetaiotaomicron*)	Rec. PPDC (*P. mirabilis*)
 Phenylpyruvate	100.00±8.66	100.00±11.03	 Methyl 2-oxopropanoate	10.51±7.75	20.29±2.15
 Pyruvic acid	2.12±3.68	91.94±3.69	 Ethyl thiophene-2-glyoxylate	2.87±1.46	19.04±16.56
 Benzoylformic acid	13.06±8.28	73.78±15.52	 4-methylthio-2-oxobutanoic acid	7.22±1.21	16.31±14.12
 3-methyl-2-oxovaleric acid	7.54±2.26	57.91±9.39	 Ethyl pyruvate	3.18±0.55	4.75±0.73
 Ethyl 2-oxo-4-phenylbutanoate	11.78±4.49	43.85±6.60	 Indolepyruvate	24.31±9.04	Not determined*
 2-oxobutanoic acid	3.93±0.66	34.43±1.52	 4-hydroxyphenyl-pyruvate	24.18±9.09	Not determined*

Note: Both *B. thetaiotaomicron* PPFOR (BT0430/0429) and *P. mirabilis* PPDC (HMPREF0693_2975) were cloned and expressed from pET16b plasmid in *E. coli* BL21 (DE3). For the recombinant PPFOR, the reaction was performed using the induced *E. coli* cell lysate and the listed substrates and the activity on various substrates were determined by the reduction of methyl viologen monitored by the change in

absorbance at 600 nm anaerobically as described in the methods. For the recombinant PPDC, the reaction was done using the purified enzyme, and the activity was estimated by the reduction of NAD^+ monitored by the change in the absorbance at 340 nm in the presence of aldehyde dehydrogenase under aerobic condition. The relative activity is expressed as a percentage of the highest activity in the presence of phenylpyruvate.

* PPDC showed activity on indolepyruvate and 4-hydroxyphenylpyruvate, but relative activity was not determined as indolepyruvate and 4-hydroxyphenylpyruvate have a high background absorbance at 340 nm. The putative products, indoleacetaldehyde and 4-hydroxyphenylacetaldehyde, were therefore examined using LC-MS/MS after derivatization with 2,4-dinitrophenylhydrazine (see Methods). The putative products, indoleacetaldehyde-2,4-dinitrophenylhydrazone and 4-hydroxyphenylacetaldehyde-2,4-dinitrophenylhydrazone were detected with their predicted signature parent to daughter transition (Fig. S5). Because the authentic indoleacetaldehyde or 4-hydroxyphenylacetaldehyde to build standard curve were not commericaly available, PPDC relative activity on Indolepyruvate or 4-hydroxyphenylpyruvate was not determined.

Table S4. Proteomics analyses of the recombinant PPFOR (BT0430/BT0429) expressed in *E. coli* BL21 (DE3), related to main Figure 1.

Protein accesion	[M+H]	m/z	z	Δm (ppm)	Ret Time (Min)	Residues	Sequence	Xcorr Scores	
								Digestion of the BT0430 gel band	In-solutin digestion of crude *E. coli* lysate
Q8AAN4	779.39	779.3933	1	-0.08	21.1738	213-218	QDEFIK	1.59	-
Q8AAN4	820.49	820.4923	1	-0.51	28.9885	299-306	GYLGIGIK	1.57	-
Q8AAN4	862.55	431.7791	2	-0.07	22.2586	206-212	YKVLLAR	2.07	-
Q8AAN4	915.53	458.2688	2	0.52	24.5511	263-270	IGQYPLPK	2.41	-
Q8AAN4	932.49	466.7473	2	0.26	22.1575	156-163	TGEPILMoR	2.51	-
Q8AAN4	934.50	467.7534	2	0.32	16.2782	228-238	YIDGPNKK	2.17	-
Q8AAN4	1024.46	512.7332	2	0.82	17.1684	219-227	ASEESPYNK	2.97	-
Q8AAN4	1043.62	348.5464	3	-0.03	20.7019	263-271	IGQYPLPKK	2.82	-
Q8AAN4	1146.60	573.8061	2	0.14	20.6321	517-525	RECIQTLAR	3.48	-
Q8AAN4	1198.73	400.2485	3	0.23	27.1559	296-306	HLKGYLGIGIK	3.58	-
Q8AAN4	1325.70	663.3547	2	0.06	41.0946	359-369	DMYITLTEVLK	-	2.8
Q8AAN4	1341.70	671.3524	2	0.34	37.6739	359-369	DMoYITLTEVLK	2.67	-
Q8AAN4	1428.69	714.8466	2	-0.15	31.4391	60-73	TAMoEAALGMSFVGK	4.52	-
Q8AAN4	1444.68	722.8447	2	0.76	28.3369	60-73	TAMoEAALGMoSFVGK	4.87	-
Q8AAN4	1617.85	809.4274	2	-0.52	35.522	503-516	EEIEYNGVSVIIPR	4.67	4.1
Q8AAN4	1749.93	583.982	3	-0.05	33.8578	470-485	IESICIGLGVDPAHIR	3.25	-
Q8AAN4	1761.84	881.4236	2	-0.23	22.9306	179-193	EQKPQNSISFSEDPR	3.28	2.99
Q8AAN4	1784.83	595.6164	3	0.47	24.8847	213-227	QDEFIKASEESPYNK	4.07	-
Q8AAN4	1886.91	943.9578	2	-0.35	29.5298	311-328	LDGTLSQDGELNPDSVAR	5.69	5.2
Q8AAN4	2100.03	1050.519	2	-0.14	27.9465	309-328	GRLDGTLSQDGELNPDSVAR	6.26	-
Q8AAN4	2106.01	702.6752	3	0.11	30.9603	81-100	HVGMoNVAADCFVNSAITGVK	4.96	-
Q8AAN4	2427.12	1214.064	2	0.07	28.1527	101-123	GGLIVIAADDPSMHSSQNEQDSR	-	5.96
Q8AAN4	2443.12	815.0434	3	0.09	27.0526	101-123	GGLIVIAADDPSMoHSSQNEQDSR	5.4	-
Q8AAN4	2801.43	934.4811	3	0.4	45.0584	270-295	QLLQLVESCDEILVLEDGQPFVEK	-	5.35
Q8AAN5	1540.86	514.2911	3	0.49	27.3779	54-68	ISDKPIASDLIPSGK	-	2.85
Q8AAN5	3424.66	1142.226	3	0.59	46.4015	86-111	YLPYLSPEGWLVTNETPFVNIPNYPEEDK	-	2.89

Peptides in blue correspond to BT0430 (PPFOR α-subunit)

Peptides in red correspond to BT0429 (PPFOR β-subunit)