



# Developing a highly efficient hydroxytyrosol whole-cell catalyst by de-bottlenecking rate-limiting steps

Jun Yao<sup>1,2,5</sup>, Yang He<sup>3,5</sup>, Nannan Su <sup>3</sup>, Sakshibeedu R. Bharath<sup>3</sup>, Yong Tao<sup>1</sup>, Jian-Ming Jin<sup>4</sup>✉, Wei Chen<sup>1</sup>✉, Haiwei Song <sup>3</sup>✉ & Shuang-Yan Tang<sup>1</sup>✉

Hydroxytyrosol is an antioxidant free radical scavenger that is biosynthesized from tyrosine. In metabolic engineering efforts, the use of the mouse tyrosine hydroxylase limits its production. Here, we design an efficient whole-cell catalyst of hydroxytyrosol in *Escherichia coli* by de-bottlenecking two rate-limiting enzymatic steps. First, we replace the mouse tyrosine hydroxylase by an engineered two-component flavin-dependent monooxygenase HpaBC of *E. coli* through structure-guided modeling and directed evolution. Next, we elucidate the structure of the *Corynebacterium glutamicum* VanR regulatory protein complexed with its inducer vanillic acid. By switching its induction specificity from vanillic acid to hydroxytyrosol, VanR is engineered into a hydroxytyrosol biosensor. Then, with this biosensor, we use in vivo-directed evolution to optimize the activity of tyramine oxidase (TYO), the second rate-limiting enzyme in hydroxytyrosol biosynthesis. The final strain reaches a 95% conversion rate of tyrosine. This study demonstrates the effectiveness of sequentially de-bottlenecking rate-limiting steps for whole-cell catalyst development.

<sup>1</sup>CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>3</sup>Institute of Molecular and Cell Biology, 61 Biopolis Drive, Singapore, Singapore. <sup>4</sup>Beijing Key Laboratory of Plant Resources Research and Development, Beijing Technology and Business University, Beijing, China. <sup>5</sup>These authors contributed equally: Jun Yao, Yang He. ✉email: [jinjianming@btbu.edu.cn](mailto:jinjianming@btbu.edu.cn); [chenw@im.ac.cn](mailto:chenw@im.ac.cn); [haiwei@imcb.a-star.edu.sg](mailto:haiwei@imcb.a-star.edu.sg); [tangsy@im.ac.cn](mailto:tangsy@im.ac.cn)

Whole-cell catalysts have been frequently used for natural product biosynthesis owing to their stability, recyclability and environmental friendliness. Whole-cell catalysts harboring multiple-step biosynthetic pathways avoid the time-consuming catalyst preparation process in *in vitro* catalysis<sup>1</sup>. For combinatorial whole-cell biocatalysts, the biosynthetic pathway efficiency is the most important factor influencing catalyst productivity. Besides the metabolic engineering strategies used for directing the carbon flux into a targeted pathway, the pathway enzymes are crucial for driving the pathway rates, selectivity, and overall productivity<sup>2,3</sup>. High-efficient pathway enzymes, especially those catalyzing rate-limiting steps, are critical for the end product productivity<sup>4</sup>. Rate-limiting steps of the biosynthetic pathways in combinatorial whole-cell catalysts might not be the same as those in the original hosts, due to the different protein expression levels, enzymatic activities and substrate or cofactor levels in the engineered host cells. When one rate-limiting step is unblocked, most probably another rate-limiting step will appear. A high-efficient whole-cell catalyst can be developed when all rate-limiting steps are sequentially solved.

Besides the conventional method of overexpressing enzymes catalyzing the rate-limiting steps, protein engineering is a frequently used strategy. The activity, specificity, selectivity, and allosteric property of the pathway enzymes have been engineered to improve pathway performance<sup>4–6</sup>. An alternative solution is to redesign the pathways to bypass rate-limiting steps or allow shortcuts for synthesis. To facilitate pathway optimization in whole-cell catalysts, the end-product biosensors are powerful tools, allowing rapid selection of hyper-producing cells<sup>7</sup>. The strategy of *in vivo*-directed evolution combined with biosensor screening is more advantageous than *in vitro* engineering by better adapting to the intracellular environment, including the concentration of substrates, intermediates, and cofactors, enabling the selection of hyper-producing whole-cell catalysts under real *in vivo* environments<sup>8–10</sup>. Development of customized biosensors for targeted products is challenging.

Hydroxytyrosol is a powerful antioxidant scavenger of free radicals that confers cell protection<sup>11,12</sup>. The hydroxytyrosol biosynthetic pathway, which uses tyrosine as a substrate, was first reported by Satoh<sup>13</sup>. In this pathway, the tyrosine hydroxylation step catalyzed by mouse tyrosine hydroxylase using the cofactor tetrahydropterin (MH4) severely limits biosynthetic efficiency, resulting in <20% hydroxytyrosol yield. Chung et al.<sup>14</sup> used the aromatic aldehyde synthase AAS and the 4-hydroxyphenylacetic acid 3-hydroxylase HpaBC to turn tyrosine to hydroxytyrosol in *Escherichia coli*, however the titer of hydroxytyrosol is only 208 mg L<sup>-1</sup>. Li et al.<sup>15</sup> established an artificial pathway for hydroxytyrosol biosynthesis from tyrosine by using the ketoacid decarboxylase KDC, the alcohol dehydrogenase ADH and also the 4-hydroxyphenylacetic acid 3-hydroxylase HpaBC, and finally obtained a hydroxytyrosol titer of ~1243 mg L<sup>-1</sup> in fed-batch experiment.

In this study, we aim to demonstrate the feasibility of developing a powerful hydroxytyrosol whole-cell catalyst using *E. coli*, by sequentially solving the activities of enzymes catalyzing the rate-limiting steps of the pathway. Strategies of structure-based enzyme redesign and *in vivo*-directed evolution are adopted, until all the rate-limiting steps are solved and the conversion rate of end product reaches 95%.

## Results

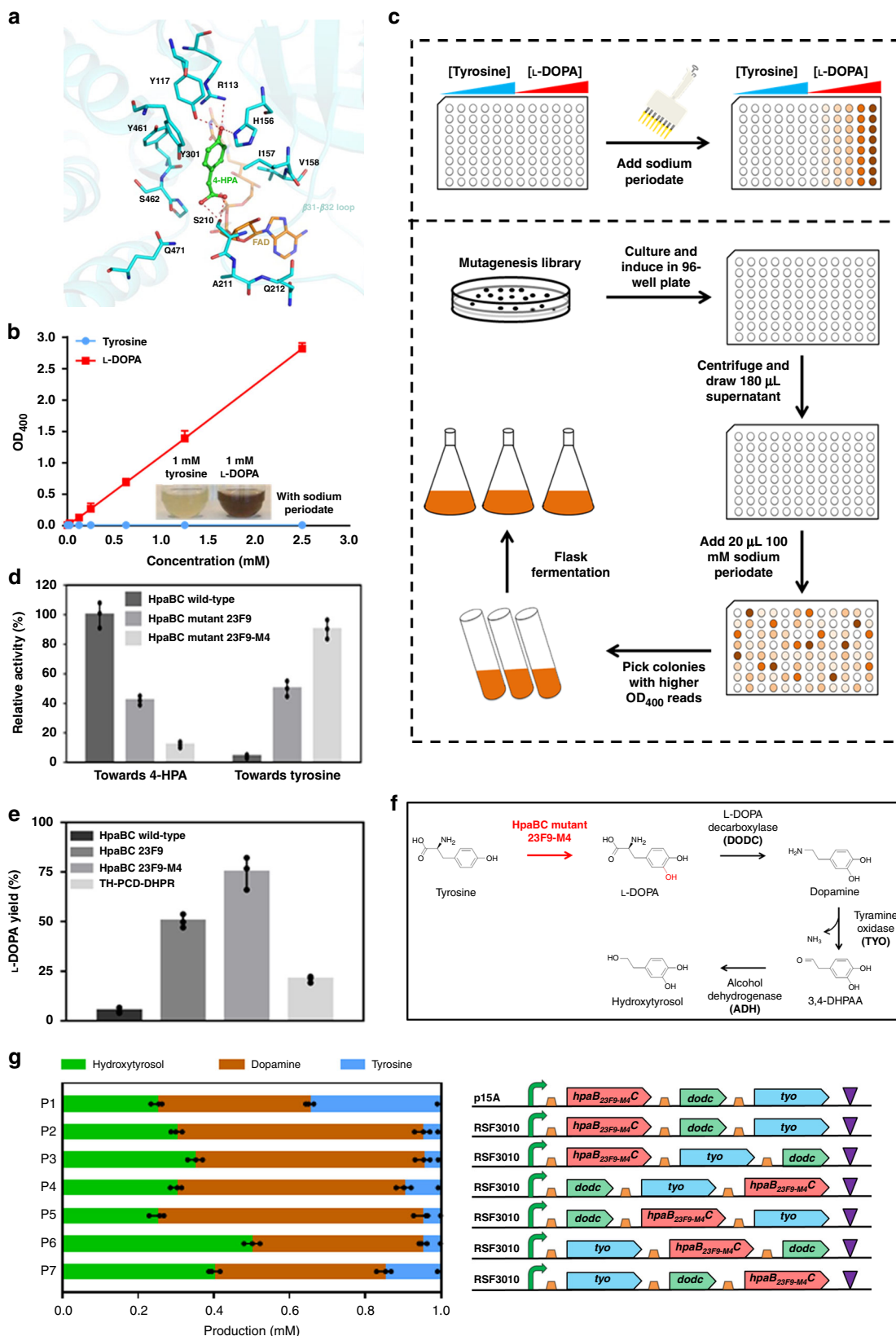
### Development of a high-active microbial tyrosine hydroxylase.

To improve the low efficiency of the tyrosine hydroxylation step catalyzed by mouse tyrosine hydroxylase<sup>13</sup>, a microbial tyrosine

hydroxylase was redesigned by engineering HpaBC, a two-component flavin-dependent monooxygenase from *E. coli*. HpaBC catalyzes the first step in the 4-hydroxyphenylacetate (4-HPA) degradation pathway, producing 3,4-dihydroxyphenylacetate (HPC). In spite of a broad substrate spectrum, HpaBC exhibits extremely low activity on tyrosine (approximately 5% of its activity on 4-HPA)<sup>16</sup>.

To explore the mechanism of the catalytic versatility of HpaBC, we modeled an active HpaBC bound with FAD and 4-HPA based on the reported structures of EcHpaBC (*E. coli*)<sup>17</sup> and ThHpaBC (*Thermus thermophilus*)<sup>18</sup>. We hypothesized that the binding of 4-HPA to HpaBC, like its homolog in ThHpaBC, would induce a conformational change in loop  $\beta$ 32–33 (Supplementary Fig. 1). As shown in Fig. 1a, the hydroxyl group of 4-HPA would form hydrogen bonds with the side chains of R113, Y117, and H155, which are highly conserved and involved in the catalytic reaction. The carboxyl group of 4-HPA is predicted to be coordinated by the side chain of S210 through two hydrogen bonds. The flexibility of loop  $\beta$ 32–33 is likely to define the substrate selectivity of HpaBC. Therefore, we chose residues S210, A211 and Q212 for saturation mutagenesis in order to evolve HpaBC into a highly active tyrosine hydroxylase which can efficiently convert tyrosine into L-DOPA.

A high-throughput screening method was developed according to the hydroxylase activity of HpaBC<sup>19</sup>. As shown in Fig. 1b, when reacted with sodium periodate, L-DOPA turned brown, whereas tyrosine exhibited no color change. OD<sub>400</sub> can be used to quantify this colorimetric difference, and a positive linear correlation was observed between OD<sub>400</sub> and L-DOPA concentrations. Therefore, the sodium periodate assay was used for high-throughput screening of the L-DOPA formation activity of HpaBC in 96-well plates (Fig. 1c). The saturation mutagenesis library was transformed into strain BW25113 and subjected to high-throughput screening. From ~50,000 clones, 12 mutant strains that exhibited improved tyrosine hydroxylation capability compared with that of the strain expressing the wild-type enzyme were selected (Supplementary Fig. 2). The mutant HpaBC expressed by mutant strain exhibiting the highest hydroxylation activity, 23F9, was purified for activity assay, along with the wild-type enzyme. As shown in Fig. 1d, the relative hydroxylase activity of the 23F9 mutant on tyrosine was ~12-fold higher than that of the wild-type enzyme. To further improve the tyrosine hydroxylation activity, a random mutagenesis library was constructed based on mutant 23F9. High-throughput screening of this library resulted in six mutant strains showing increased absorbance at 400 nm (Supplementary Fig. 3a) and produced high titers of L-DOPA (Supplementary Fig. 3b), compared with those of strain expressing the 23F9 mutant. The HpaBC mutant 23F9-M4 exhibited a 1.8-fold higher hydroxylation activity on tyrosine than that of the 23F9 mutant (Fig. 1d). Furthermore, the hydroxylation activity of these HpaBC mutants on the original substrate 4-HPA reduced considerably (Fig. 1d). The plasmids expressing these HpaBC mutants were purified from the selected mutant strains and used to retransform strain BW25113, and enhanced L-DOPA yields was confirmed. The yield of L-DOPA from strains expressing mutant 23F9 and 23F9-M4 was found to increase by ~10- and 15-fold, respectively, compared with that of the strain expressing wild-type HpaBC (Fig. 1e). The L-DOPA yield from strain expressing the mouse tyrosine hydroxylase and the MH4 regeneration cycle constructed according to Satoh et al.<sup>13</sup> was ~21%, around 4-fold of that of the strain expressing wild-type HpaBC (Fig. 1e). The sequencing results revealed amino acid substitutions in the selected HpaB mutants (Table 1). Therefore, a highly active microbial tyrosine hydroxylase was successfully redesigned to replace mouse tyrosine hydroxylase.



**Redesign of the hydroxytyrosol biosynthetic pathway.** The HpaBC mutant 23F9-M4 was used to redesign the hydroxytyrosol biosynthetic pathway, with the removal of the regeneration system of cofactor MH4<sup>13</sup> (Fig. 1f). The redesigned pathway was

constructed via a plasmid expressing an operon consisting of genes encoding mutant 23F9-M4, L-DOPA decarboxylase (DODC), and tyramine oxidase (TYO), as well as the *E. coli* endogenous alcohol dehydrogenase (ADH). As the knockout of

**Fig. 1 Development of a microbial tyrosine hydroxylase and redesign of the hydroxytyrosol biosynthetic pathway.** **a** Model of HpaBC bound with 4-HPA and FAD. 4-HPA and FAD are shown as stick and colored in green and brown, respectively. Hydrogen bonds are shown as red dash. **b** The standard curve of the periodate assay. **c** Flowchart of the high-throughput screening of HpaBC. **d** Specific activities of the wild-type and mutant HpaBCs on 4-HPA and tyrosine. **e** L-DOPA yields of strain BW25113 expressing wild-type, HpaBC mutant 23F9, HpaBC mutant 23F9-M4 and the reported pathway enzymes TH-PCD-DHPR<sup>13</sup>. **f** Redesign of the hydroxytyrosol biosynthetic pathway using HpaBC mutant 23F9-M4 as a highly efficient tyrosine hydroxylase. **g** Optimization of plasmid copy number and gene orders in the operon encoding the biosynthetic pathway. The concentration of the indicated compounds was determined, and no L-DOPA was accumulated (The green, brown and blue bars indicate the concentrations of hydroxytyrosol, dopamine and tyrosine, respectively). 4-HPA, 4-hydroxyphenylacetate; 3,4-DHPAA, 3,4-dihydroxyphenyl acetaldehyde. The data shown in **b**, **d**, **e** and **g** are from three replicate experiments and are expressed as the mean±SD. Source data underlying Figs. 1b, d, e, and g are provided as a Source Data file.

**Table 1 Amino acid substitutions in variants.**

Protein	Amino acid substitutions				
HpaBC	15	210	211	212	284
Wild-type	T	S	A	Q	D
23F9	T	F	K	F	D
23F9-M4	P	F	K	F	E
VanR	103	148	168		
Wild-type	Y	R	Y		
HyT5	K	I	L		
HyT12	L	L	M		
HyT19	N	V	P		
HyT31	P	V	N		
HyT55	G	R	P		
HyT134	I	S	P		
HyT172	T	M	W		
TYO	33	60	103	149	238
Wild-type	E	H	D	D	V
YM9	E	Q	Y	E	V
YM9-2	D	Q	Y	E	E

gene *feaB* has been reported to be essential to achieve high yield and high purity of hydroxytyrosol<sup>13</sup>, we knocked out the *feaB* gene in strain BW25113 to obtain strain BHYT<sup>19</sup>. The plasmid copy number and the order of the three genes in the operon were optimized, in order to improve the hydroxytyrosol biosynthesis efficiency (Fig. 1g). Plasmid P2 carrying the RSF3010 replication origin resulted in high hydroxytyrosol production compared with that of the plasmid carrying the p15A replication origin. The gene order in plasmid P6 presented the highest hydroxytyrosol production (Fig. 1g). It was found that with the redesigned tyrosine hydroxylation step with plasmid P6, the tyrosine conversion rate reached 95%, which indicated that it was no longer a rate-limiting step. Although this rate-limiting step was solved, the hydroxytyrosol yield of the whole pathway was only around 50%. By analyzing the intracellular concentrations of biosynthetic intermediates, it was found that dopamine accumulated in large amounts in strain BHYT harboring plasmid P6 (Fig. 1g), indicating that the oxidation of dopamine to 3,4-dihydroxyphenyl acetaldehyde (3,4-DHPAA) was insufficient. This reaction was catalyzed by TYO, and it has become the rate-limiting step of the redesigned pathway. As this reaction generated NH<sub>3</sub> and H<sub>2</sub>O<sub>2</sub> which are toxic to the cell in high amounts, to avoid toxic intermediate accumulation, balanced activities between TYO and downstream pathway enzymes are critical for efficient hydroxytyrosol biosynthesis. Therefore, in vivo-directed evolution was performed to optimize the TYO activity. To minimize the negative effect of the toxic intermediates on biosynthesis, a high-throughput screening tool of the end product would help. Thus, a hydroxytyrosol biosensor was developed by engineering the regulatory protein VanR.

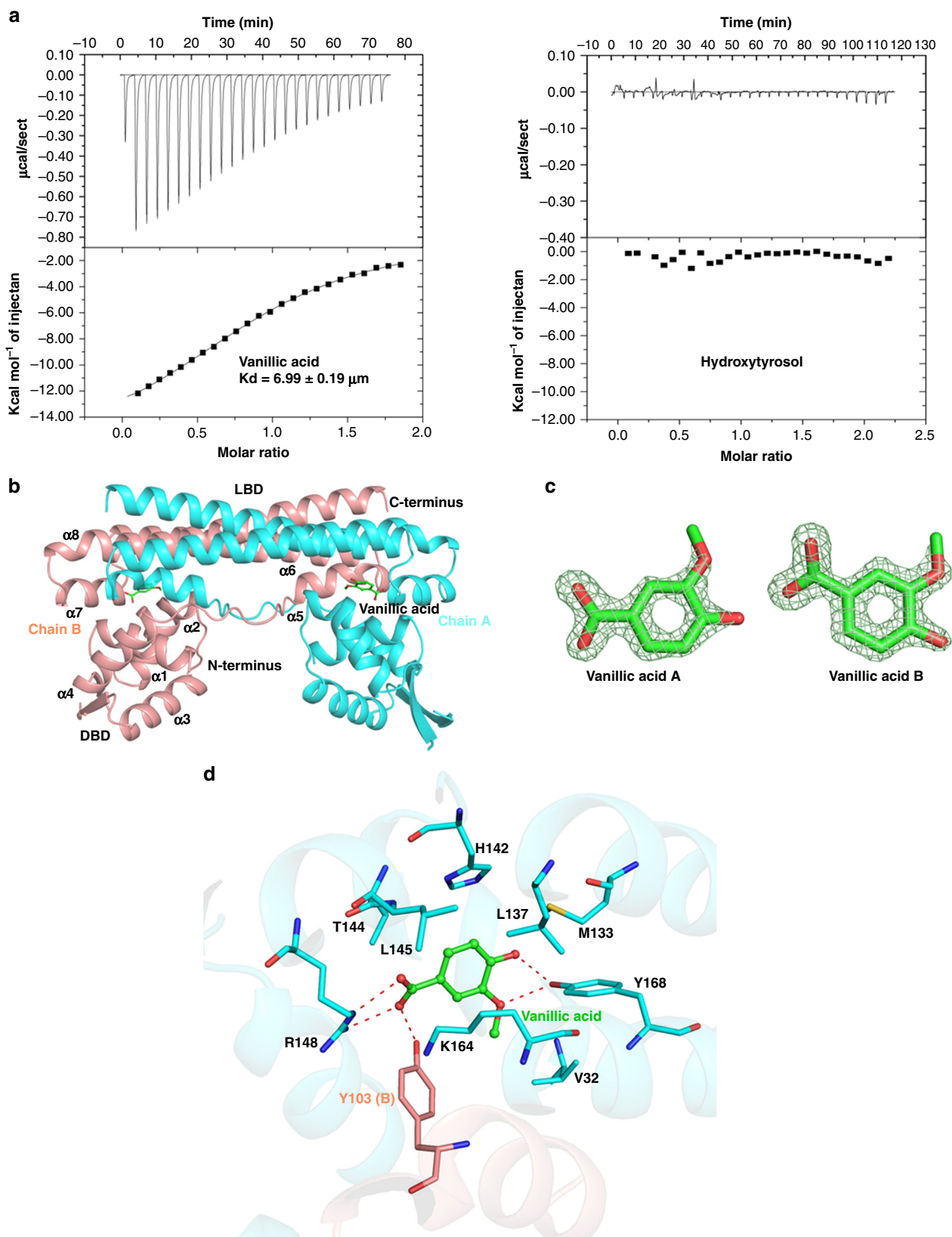
**Crystal structure of VanR with bound vanillic acid.** A regulatory protein from *C. glutamicum*, VanR, was chosen for

effector specificity engineering to develop a hydroxytyrosol-responsive regulator. VanR belongs to the PadR family of regulatory proteins and has been proved to specifically respond to its effector vanillic acid, a hydroxytyrosol analog<sup>20</sup>. As shown in Fig. 2a, vanillic acid binds to VanR with moderate affinity ( $K_d = 6.99 \mu\text{M}$ ), but no binding was observed between hydroxytyrosol and VanR.

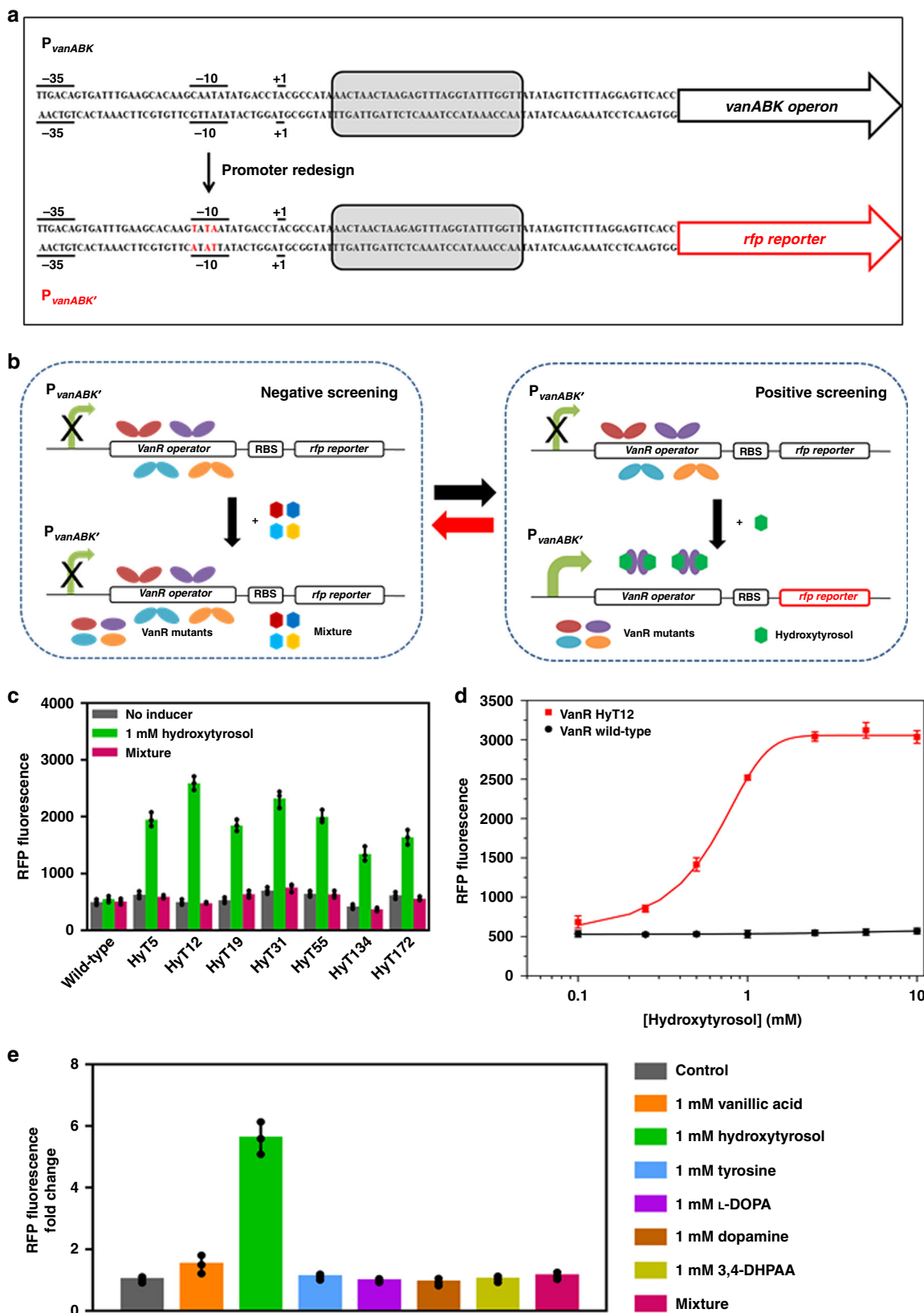
To facilitate VanR engineering, we determined the crystal structure of VanR complexed with vanillic acid (Supplementary Table 1). VanR functions as a homodimer and each monomer consists of an N-terminal winged helix-turn-helix (wHTH) domain which interacts with operator DNA, and a C-terminal homodimerization domain (Fig. 2b). Electronic densities of vanillic acids were found in the interface of two subunits (Fig. 2c, Supplementary Fig. 4a, b, Supplementary Fig. 5). Vanillic acid binds to a hydrophobic interdomain pocket formed by V29, Y168, M133, L145, and H142 from subunit A and Y103 of subunit B (Fig. 2d). Besides hydrophobic interaction, polar contacts largely contribute to the recognition of vanillic acid by VanR. Specifically, the carboxyl group of vanillic acid is electrostatically stabilized by salt bridges with R148 of subunit A and by hydrogen bond with the side chain of Y103 of subunit B. Furthermore, the hydroxyl and methoxyl groups on the benzene ring of vanillic acid form two hydrogen bonds with the side chains from Y168 of subunit A (Fig. 2d). Those three residues were highly conserved in different bacteria (Supplementary Fig. 4c), suggesting they are essential for ligand recognition.

**Development of a hydroxytyrosol-responsive VanR mutant.** In *C. glutamicum*, VanR regulates the expression of the *vanABK* operon, which is responsible for the catabolism of vanillic acid. In the absence of vanillic acid, the VanR protein binds to a 28-bp operator sequence downstream of +1 position of the  $P_{vanABK}$  promoter and represses the transcription of the *vanABK* operon. When vanillic acid is present, it binds to the VanR protein, inducing a conformational change which releases the promoter DNA, thus activating the operon transcription<sup>20</sup>. Directly introducing the whole promoter region (95-bp) into *E. coli* cells could not initiate downstream gene transcription (data not shown). To adapt a functional VanR regulatory system in *E. coli*, after sequence analysis, the -10 region of promoter  $P_{vanABK}$  (CAATAT) was replaced with the -10 region of the *tac* promoter ( $P_{tac}$ ) (TATAAT) frequently used in *E. coli* (Fig. 3a) to generate promoter  $P_{vanABK}$ . The resultant promoter successfully regulated the downstream *rfp* gene expression in response to vanillic acid in the presence of the constitutively expressed VanR protein in *E. coli*, with plasmid pVanR (Supplementary Figs. 6, 7).

To alter the effector specificity of VanR, amino acid residues Y103, R148, and Y168, were chosen for simultaneous site-saturation mutagenesis (Fig. 2d). The library was constructed using plasmid pVanR and used to transforming strain *E. coli* BW25113 ( $3.5 \times 10^5$  transformants). The transformants were screened by fluorescence-activated cell sorting (FACS). One round of negative screening was performed by sorting the least



**Fig. 2** Crystal structure elucidation of VanR complexed with vanillic acid. **a** ITC titrations of vanillic acid and hydroxytyrosol to wild-type VanR. The upper panels show the binding isotherms and the lower panels show the integrated heat for each injection fitted to a single-site model. **b** Crystal structure of the VanR with bound vanillic acid. Subunits A and B are colored in cyan and light red respectively. **c** Fo-Fc electron density map of vanillic acid at the contour of 2.0 colored in light green; **d** The interaction of VanR with vanillic acid. Vanillic acid is shown in green stick and polar contacts are shown as red dash lines.



fluorescent cells ( $2.8 \times 10^5$ ) in the presence of 1 mM each of tyrosine, L-DOPA, dopamine and 3,4-DHPAA, in order to remove the VanR variants responsive to the substrate and intermediates in hydroxytyrosol biosynthesis. One round of

positive screening was then performed by sorting the most fluorescent cells ( $3.5 \times 10^3$ ) in the presence of 1 mM hydroxytyrosol (Fig. 3b). After seven rounds of dual screening, seven variants responsive to hydroxytyrosol but not the substrate and

**Fig. 3 Adaptation of the VanR regulatory system in *E. coli* and its induction specificity alteration.** **a** Adaptation of the VanR regulatory system in *E. coli* by redesigning the  $-10$  region of  $P_{vanABK}$ . The substituted base pairs are highlighted in red. **b** High-throughput screening strategy for engineering the induction specificity of VanR. **c** RFP fluorescence of strain BW25113 harboring plasmid pVanR expressing wild-type or mutant VanRs in the absence of any inducer (black bar) or in the presence of 1 mM hydroxytyrosol (green bar) or in the presence of the mixture (Magenta bar). **d** RFP fluorescence from strain BW25113 harboring plasmid pVanR expressing VanR wild-type (black circle) or the HyT12 mutant (red square) as a function of the hydroxytyrosol concentrations. **e** RFP fluorescence of strain BW25113 harboring plasmid pVanR expressing the HyT12 mutant in response to the indicated compounds. Mixture represents a mixture of 1 mM each of tyrosine, L-DOPA, dopamine and 3,4-DHPAA. The data shown in **c**, **d** and **e** are from three replicate experiments and are expressed as the mean $\pm$ SD. Source data underlying Fig. 3c-e are provided as a Source Data file.

intermediates were selected. Among them variant HyT12 exhibited the highest induction fold towards hydroxytyrosol and was used for downstream characterizations (Fig. 3c). The sequencing results revealed amino acid substitutions (Y103L, R148L and Y168M) in HyT12. Figure 3d depicts RFP fluorescence of strain BW25113 expressing variant HyT12, as a function of hydroxytyrosol concentration. Approximately 6-fold induction of RFP expression was achieved over a hydroxytyrosol concentration range of 0–10 mM, with half-maximal induction occurring at  $\sim$ 0.5 mM. Wild-type VanR exhibited no response towards hydroxytyrosol. The induction effects of the biosynthetic substrate tyrosine and the intermediates L-DOPA, 3,4-DHPAA and dopamine, as well as the original inducer vanillic acid, were determined individually, by comparing with those of hydroxytyrosol. As shown in Fig. 3e, the HyT12 variant showed a good specific response towards hydroxytyrosol, compared with that of the other tested compounds, indicating that negative screening using a mixture of the biosynthetic substrate and intermediates was definitely effective. This induction specificity is critical for the application of HyT12 variant as a hydroxytyrosol biosensor for high-throughput screening. Additionally, we modeled the binding of hydroxytyrosol to the VanR mutants. As shown in the modeling data in Supplementary Fig. 8a, we found a significantly higher Rosseta function score of HyT12 than that of the other mutants, indicating a considerable conformational change in HyT12 to accommodate the ligand. As depicted in Supplementary Fig. 8b, hydroxytyrosol would form hydrogen bond with the side chain of T144 and H142, and the hydroxyl tail was accommodated by a hydrophobic cavity formed by V32, K164, A167, M168 from subunit A and A102 and L103 from subunit B. The amino acid substitutions in various VanR variants are shown in Table 1.

**Application of a whole-cell hydroxytyrosol biosensor.** As depicted in Fig. 4a, strain H2 was constructed by integrating the constitutively expressed HyT12 and the  $P_{vanABK}$ - $lacZ$  construct into the chromosome of strain BHYT. We used  $\beta$ -galactosidase encoded by  $lacZ$  as a reporter to facilitate the subsequent screening on agar plates. Strain H2 was then transformed with plasmid P6 or the control plasmid without the *tyo* gene absent. As shown in Fig. 4b, the results indicated that the developed whole-cell hydroxytyrosol biosensor could effectively report intracellular hydroxytyrosol production.

In vivo-directed evolution was used to optimize the activity of TYO in the biosynthetic pathway and the hydroxytyrosol whole-cell biosensor was used for high-throughput screening. A random mutagenesis library of TYO was constructed with plasmid P6 and was used to transform strain H2. The deep blue colonies, which were supposed to be hydroxytyrosol hyper-producers, were selected on agar plates supplemented with 3 mM tyrosine and  $40 \mu\text{g mL}^{-1}$  5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactopyranoside (X-GAL). Twelve clones were selected from  $1.0 \times 10^6$  transformants screened, and they were re-screened by quantifying hydroxytyrosol production from liquid culture. Variant YM9

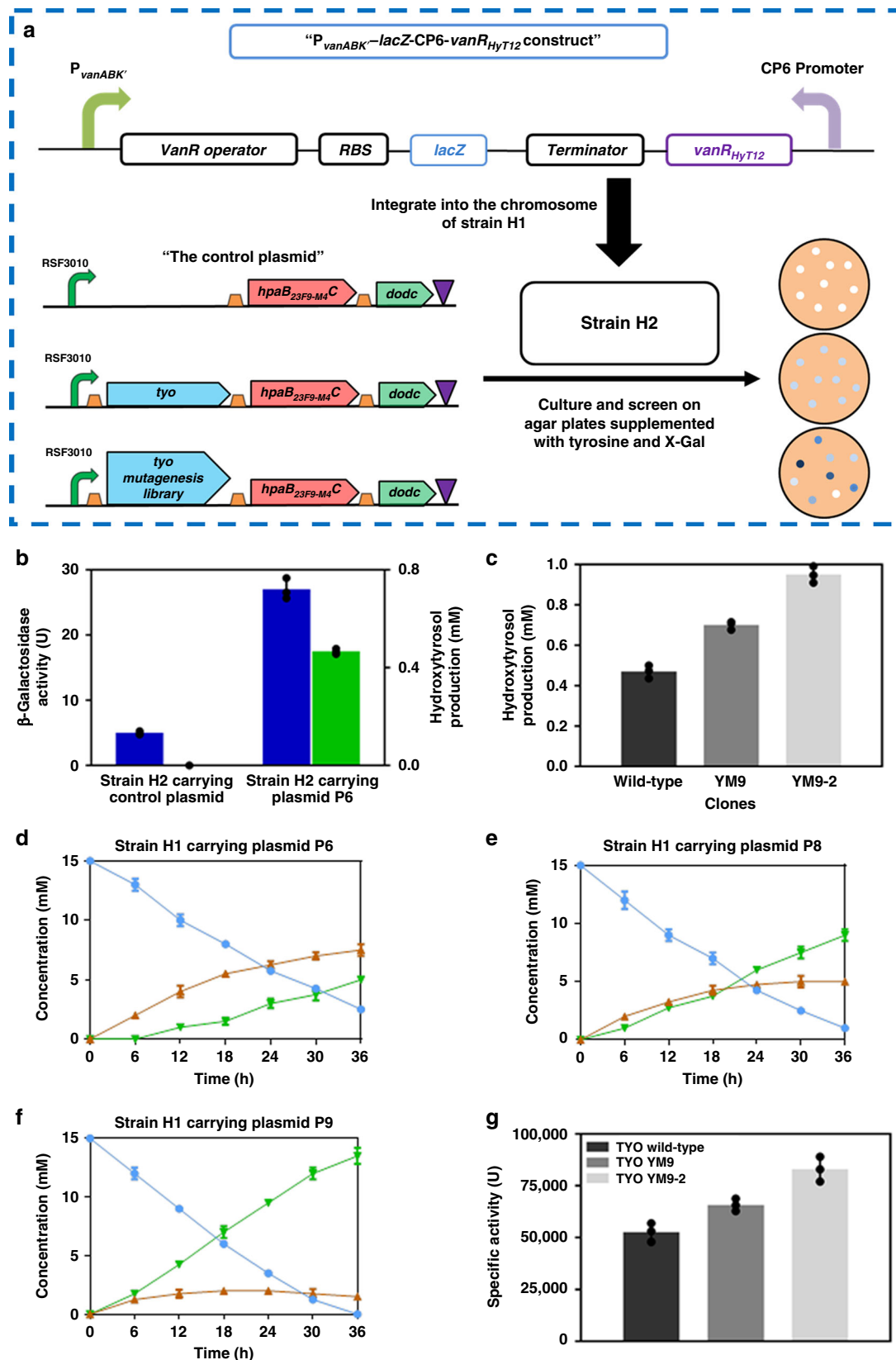
was found to produce the highest amount of hydroxytyrosol (Supplementary Fig. 9a). To further optimize the TYO activity, a random mutagenesis library was constructed using variant YM9 as a template, and the library was subjected to high-throughput screening as described above except with 6 mM tyrosine supplementation. Six clones exhibiting high hydroxytyrosol production were selected, and variant YM9-2 produced the highest amount of hydroxytyrosol (Supplementary Fig. 9b). As shown in Fig. 4c, when cultured in the presence of 1 mM tyrosine, clones YM9 and YM9-2 displayed a  $\sim$ 1.49- and 2.02-fold increase in hydroxytyrosol production, respectively, compared with that of the strain harboring pathway expressing wild-type TYO. The hydroxytyrosol yield of clone YM9-2 reached 95%. Sequencing results revealed amino acid substitutions in the mutated TYOs (Table 1).

As shown in Fig. 4d-f, the time course of the concentration of hydroxytyrosol and intermediates, as well as the consumption of tyrosine, was determined, with strain BHYT harboring plasmids P6, P8 (plasmid P6 expressing TYO YM9), and P9 (plasmid P6 expressing TYO YM9-2) cultured in a 5-L fermentor. The strains were cultured in the presence of 15 mM tyrosine. For strain harboring plasmid P6, the hydroxytyrosol yield was 50% and a large amount of dopamine was accumulated. In strain harboring plasmid P8 (pathway expressing the TYO mutant YM9), dopamine accumulation was alleviated and hydroxytyrosol production was improved. As for strain harboring plasmid P9 (pathway expressing the TYO mutant YM9-2), the accumulation of dopamine was considerably reduced and the hydroxytyrosol yield reached 90% (Fig. 4f), comparing with 33.3% in the presence of wild-type TYO (Fig. 4d).

Fed-batch fermentation and whole-cell biotransformation were both performed to produce hydroxytyrosol from tyrosine. In fed-batch fermentation, strain harboring plasmid P9 was cultured in a 5-L fermentor in the presence of 1 mM L-arabinose, and 7.5 mM of tyrosine was fed with at 0, 12, 24 and 36 h. The time course of hydroxytyrosol production was analyzed (Supplementary Fig. 10). It was found that after 48 h at 37 °C, the hydroxytyrosol production was  $\sim$ 3,373  $\text{mg L}^{-1}$ . In whole-cell biotransformation, cells harboring plasmid P9 was resuspended in 1 L ( $\text{OD}_{600} = 50$ ) of 100 mM Tris-HCl buffer (pH 7.5) with 15 mM tyrosine. An additional 30 mM of tyrosine was supplemented at 12 h and the biotransformation was extended to 48 h at 37 °C. A rotating speed of 100 rpm and air flow rate of 1.0 vvm were used, with pH maintained at 7.5. The final hydroxytyrosol production reached 4,690  $\text{mg L}^{-1}$ , which was higher than reported so far.

#### Analysis and characterization of wild-type and mutant TYOs.

The genes encoding wild-type or mutant TYOs were cloned into the pET28a vector and expressed in *E. coli* BL21(DE3) as N-terminal His-tagged fusion proteins. The enzymes were then purified for activity assay and kinetic parameter determination. As shown in Fig. 4g, the TYO variants YM9 and YM9-2 exhibited  $\sim$ 1.3- and  $\sim$ 1.6-fold increase in specific activity,



respectively, relative to that of the wild-type enzyme. This corresponded well with hydroxytyrosol productions (Fig. 4c). The results indicated that the improved activity of the rate-limiting enzyme, TYO, underlies enhanced hydroxytyrosol productions in the selected hyper-producing strains. The kinetic parameters of the mutant TYOs were determined, along

with the wild-type enzyme (Table 2). Our data revealed that the affinity of the mutant enzymes for substrate dopamine increased. The catalytic efficiency ( $k_{cat}/K_m$ ) of the mutant enzymes YM9 and YM9-2 showed ~1.7- and ~3.3-fold improvements compared with that of the wild-type enzyme, respectively (Table 2).



**Fig. 4 Design of a hydroxytyrosol biosensor and its application in TYO activity optimization.** **a** Design of a hydroxytyrosol biosensor and the strategy of applying it in *in vivo*-directed evolution of TYO for improved hydroxytyrosol biosynthesis. **b**  $\beta$ -Galactosidase activity (blue) and hydroxytyrosol productions (green) in strain H2 harboring control plasmid and plasmid P6, after culturing at 37 °C for 36 h. **c** Hydroxytyrosol productions in the wild-type and selected clones in the first and second rounds of *in vivo*-directed evolution, after culturing at 37 °C for 36 h, in the presence of 1 mM tyrosine; Time courses of productions of hydroxytyrosol and dopamine, as well as tyrosine consumption, from strain BHYT carrying plasmid P6 (**d**), P8 (**e**), or P9 (**f**) in a 5-L fermentor. Blue circle, brown triangle and green triangle indicate the concentration of tyrosine, dopamine and hydroxytyrosol, respectively. **g** Specific activities of wild-type and mutant TYOs. The data shown in **b**, **c**, **d**, **e**, **f** and **g** are from three replicate experiments and are expressed as the mean  $\pm$  SD. Source data underlying Fig. 4b-g are provided as a Source Data file.

**Table 2 Kinetic parameters of wild-type and mutant TYOs.**

Enzymes	$K_m$	$k_{cat}$	$k_{cat}/K_m$
TYO wild-type	8.48 $\pm$ 1.06	115.11 $\pm$ 21.95	13.52 $\pm$ 1.22
TYO YM9	6.67 $\pm$ 0.72	156.71 $\pm$ 16.99	23.48 $\pm$ 0.27
TYO YM9-2	4.30 $\pm$ 0.37	194.42 $\pm$ 16.43	45.20 $\pm$ 0.45

$\pm$  values represent standard deviations from three independent data points. Source data are provided as a Source Data file.

## Discussion

Tyrosine hydroxylase is the key enzyme in the biosynthesis of catecholamines dopamine, epinephrine and norepinephrine in animals, which are hormones and neurotransmitters in both the central and peripheral nervous systems<sup>21</sup>. Tyrosine hydroxylase uses tetrahydrobiopterin as a cofactor, and its activity is regulated by phosphorylation and also feedback inhibited by the catecholamine neurotransmitters<sup>22</sup>. Due to its complicated regulation properties and the necessity for cofactor regeneration cycle, acting as the first catalytic step, it severely reduces the biosynthesis efficiency of the hydroxytyrosol pathway<sup>13</sup>. To redesign this pathway, a non-natural microbial tyrosine hydroxylase was developed by engineering substrate specificity of the two-component flavin-dependent monooxygenase, HpaBC, from *E. coli* strain BL21(DE3). The wild-type HpaBC showed negligible activity on tyrosine<sup>16</sup>. By engineering its substrate-binding pocket, a highly active tyrosine hydroxylase was obtained, which possessed no allosteric inhibitory property and was expressed very well in *E. coli* cells. Thus, it was suitable for replacing the animal-sourced tyrosine hydroxylase for hydroxytyrosol biosynthesis in *E. coli*. The conversion rate of tyrosine significantly increased in the presence of this microbial tyrosine hydroxylase.

Innate transcriptional regulatory protein (TRP)-based biosensors are widely-used in natural product high-throughput screening<sup>23–26</sup>. However, under most circumstances, innate transcription factors responsive to targeted compounds are not available. Studies on development and application of non-natural transcription factors with altered specificity are limited<sup>8–10</sup>. Since there is no regulatory protein responsive to hydroxytyrosol reported so far, a non-natural hydroxytyrosol biosensor has been customized by altering the induction specificity of VanR regulatory protein. The specificity of a biosensor used as a high-throughput screening tool to optimize the pathway efficiency is critical, because it should not respond to the biosynthetic substrate or intermediates. However, performing negative screenings for each of substrate and intermediates in hydroxytyrosol biosynthetic pathway is time-consuming. Thus, a mixture of these compounds was used in a single round of negative screening, while hydroxytyrosol was used in the positive screening. A VanR mutant specifically responsive to hydroxytyrosol was successfully selected, indicating the effectiveness of this screening strategy. Thus the strategy for customizing biosensors for small molecules synthesized through multiple-step pathways has been advanced.

Structural information provides key clues for redesigning proteins including enzymes, sensors, and probes. Polar contacts between ligand and protein, including hydrogen bond and salt bridge, are very critical for redesigning. The crystal structures of both HpaBC and VanR are essential for the development of this redesigned synthetic pathway of hydroxytyrosol. Binding of substrate to HpaBC induced a conformational transition of loop from an open state to a closed state. The flexibility of this loop probably allows different analogs to enter the catalytic pocket. Changes in amino acids of this loop probably result in a different selectivity of substrates. In our model, S210 had a polar contact with 4-HPA, however, A211 and Q212 could also be involved in the recognition of ligands due to flexibility. T196 and A197 of ThHpaBC, located at similar positions as S210 and A211 of EcHpaBC would form hydrogen bond with 4-HPA. As for VanR, all three residues had polar contacts with vanillic acid and were chosen to engineer in our screening. Although HyT12 showed the highest response to the ligand, the other six variants also demonstrate comparable effects. Interestingly, we found the highest Rosetta function score in HyT12 that showed the highest response to ligand. This indicates that changes in the overall conformation of VanR might be important for its function, as the highest response of HyT12 might come from the release of DNA in the DBD domain upon ligand binding. Compared to rational design, the semi-rational design might be more efficient in our case.

Balancing pathway enzyme activities has been proved critical for pathway efficiency<sup>23,24</sup>. The reaction catalyzed by TYO, the rate-limiting enzyme of the redesigned hydroxytyrosol pathway, generated ammonia and hydrogen peroxide<sup>25,26</sup>, thus, an optimal activity instead of extremely high activity would benefit the whole pathway efficiency. *In vivo*-directed evolution combined with high-throughput screening of the end product is a good strategy for optimizing this enzyme to achieve improved hydroxytyrosol biosynthetic efficiency, by avoiding the toxicity caused by intermediate accumulation. *In vivo*-directed evolution is advantageous for pathway engineering because it better optimizes the metabolic fluxes by adapting to the intracellular environment, which could not be well managed by *in vitro* protein engineering strategies. On the contrary, for the rate-limiting enzymes with extremely low activity, it is more rapidly to improve the pathway efficiency by simply replacing them with redesigned enzymes, for example, the redesigned microbial tyrosine hydroxylase described above.

Engineering pathway rate-limiting enzyme to improve biosynthetic efficiency has been adopted in combinatorial biosynthesis. However, unblocking one rate-limiting step does not necessarily solve the overall efficiency of the whole-cell catalysts, especially when the catalytic steps are more than three. Subsequent rate-limiting steps usually come into being. Sequentially identifying and solving these steps until the overall conversion rate reaches close to 100% is the final goal and crucial task of whole-cell catalyst development. The strategies for de-bottlenecking rate-limiting steps can be diverse, depending on the features of the steps. Apart from the two strategies used in this

study, other strategies include cofactor regeneration, branch pathway blocking, and so on.

As a value-added compound with a variety of biological activities, hydroxytyrosol biosynthesis has been reported through various pathways<sup>13–15,19</sup>. Previously, with tyrosine as substrate, hydroxytyrosol productions of 29.29 mg L<sup>-1</sup> (~19% yield)<sup>13</sup>, 208 mg L<sup>-1</sup><sup>14</sup>, 1243 mg L<sup>-1</sup> (~48% yield)<sup>15</sup> and 1890 mg L<sup>-1</sup> (82% yield)<sup>19</sup> have been achieved. L-DOPA is used for treating Parkinson's disease and also a precursor for synthesizing valued benzyloisoquinoline alkaloids<sup>27,28</sup>. Various biotechnological approaches have been applied for L-DOPA production, including microbial fermentation, immobilized tyrosinase as well as electroenzymatic system<sup>29</sup>. Previous studies have revealed an over 3.5 mg L<sup>-1</sup> of L-DOPA production from glucose in yeast, using a tyrosine hydroxylase developed from a cytochrome P450 L-DOPA oxidase<sup>30,31</sup>. In our study, hydroxytyrosol production from fed-batch fermentation has reached ~3373 mg L<sup>-1</sup>. Whole-cell transformation further improved the production to ~4690 mg L<sup>-1</sup>. Furthermore, the HpaBC mutant 23F9-M4, which synthesized 443.7 mg L<sup>-1</sup> L-DOPA from 3 mM tyrosine, is also potentially applicable for L-DOPA production.

In conclusion, in this study, the first rate-limiting enzyme in the hydroxytyrosol biosynthetic pathway, mouse tyrosine hydroxylase, was replaced with an artificially designed high-efficient microbial tyrosine hydroxylase. This redesigned step converted 95% of tyrosine to L-DOPA, and the rate-limiting step of the pathway moved to dopamine oxidation catalyzed by TYO. The second rate-limiting step was optimized by in vivo-directed evolution, with the aid of a hydroxytyrosol biosensor. The crystal structure of the regulatory protein VanR from *C. glutamicum* complexed with its inducer was elucidated based on which a hydroxytyrosol biosensor was developed by altering the induction specificity of VanR. With this biosensor as an end product high-throughput screening tool, the TYO activity was optimized, further improving the hydroxytyrosol yield to 95%, the highest as reported to date<sup>13–15,19</sup>. This study has demonstrated that a powerful whole-cell biocatalyst can be developed by sequentially de-bottlenecking the rate-limiting steps, which should be generally applied in other combinatorial biosynthesis studies.

## Methods

**General.** The biosynthetic substrate, product and intermediates used in this study were obtained from Sigma-Aldrich (St. Louis, USA). The T4 polynucleotide kinase, T4 DNA ligase, restriction enzymes and DNA polymerase used in this study were obtained from Takara Bio (Dalian, China). Hieff Clone Plus Multi One Step Cloning Kit was obtained from YEASEN Biotechnology (Shanghai, China). Oligonucleotides were synthesized by Life Technologies (Shanghai, China).

*E. coli* strains MC1061 and BL21(DE3) were used for cloning and protein overexpression, respectively. Strains were grown at 37 °C in the indicated medium. Ampicillin (100 µg mL<sup>-1</sup>) and kanamycin (50 µg mL<sup>-1</sup>) were supplemented when required. Hydroxytyrosol biosynthesis was performed in yeast extract M9Y medium (M9 minimal salts, 1% (w/v) glucose, 5 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub> supplemented with 0.025% (w/v) of yeast extract)<sup>13</sup>, while other cultures were done in Luria-Bertani (LB) medium if not specially indicated.

**Plasmid construction.** All plasmids and primers used in this study are listed in Supplementary Data 1 and 2, respectively. The *dodc* gene encoding L-DOPA decarboxylase (Genbank accession No. BK006920.1) was amplified from the genome of strain *Pseudomonas putida* KT2440. The *tyo* gene from *Micrococcus luteus* encoding tyramine oxidase (Genbank accession No. AB010716.1) was synthesized after codon optimization<sup>19</sup>. The purified gene fragments *hpaB*<sub>23F9-M4C</sub>, *dodc*, *tyo*, *tyo*<sub>MUT1-tyoMUT1-22</sub> were assembled with the vector fragment of pFA or pRSF<sup>19</sup>, using Hieff Clone Plus Multi One Step Cloning Kit (YEASEN, Shanghai, China), resulting in plasmids pFA-*hpaB*<sub>23F9-M4C-dodc-tyo</sub> (P1), pRSF-*hpaB*<sub>23F9-M4C-dodc-tyo</sub> (P2), pRSF-*hpaB*<sub>23F9-M4C-tyo-dodc</sub> (P3), pRSF-*dodc-tyo-hpaB*<sub>23F9-M4C</sub> (P4), pRSF-*dodc-hpaB*<sub>23F9-M4C-tyo</sub> (P5), pRSF-*tyo-hpaB*<sub>23F9-M4C-dodc</sub> (P6), and pRSF-*tyo-dodc-hpaB*<sub>23F9-M4C</sub> (P7), pRSF-*tyoMUT1-hpaB*<sub>23F9-M4C-dodc</sub> (P8), pRSF-*tyoMUT1-22-hpaB*<sub>23F9-M4C-dodc</sub> (P9). The genes were cloned in different orders in the operons.

The *th* gene encoding tyrosine hydroxylase (TH) from mouse (Genbank accession No. NP\_033403), the *pcd* gene encoding pterin-4a-carbinolamine dehydratase (PCD) from human (Genbank accession No. NP\_000272) and the

*dhpr* gene encoding dihydropteridine reductase (DHPR) from human (Genbank accession NO.P09417) were synthesized by General Biosystems (Chuzhou, China) after codon optimization (Supplementary Data 3). The DNA fragments containing genes *th*, *pcd*, *dhpr* and vector pRSF<sup>19</sup> were amplified using primer pairs *th*-Gibson-1/*th*-Gibson-2, *pcd*-Gibson-1/*pcd*-Gibson-2, *dhpr*-Gibson-1/*dhpr*-Gibson-2 and pRSF-Gibson-1/pRSF-Gibson-2, respectively, and then were assembled using Hieff Clone Plus Multi One Step Cloning Kit (YEASEN, Shanghai, China), resulting in plasmid pRSF-*th-pcd-dhpr*.

The fragment containing promoter region of *vanABK* operon (*P*<sub>*vanABK*</sub>) was amplified using the genome of strain *C. glutamicum* ATCC 13032 as template with primers *P*<sub>*vanABK*</sub>-Gibson-1 and *P*<sub>*vanABK*</sub>-Gibson-2. The vector fragment was amplified using plasmid pSHYA as template with primers pSHYA-*backbone*-Gibson-1 and pSHYA-*backbone*-Gibson-2. The two purified fragments were assembled using Hieff Clone Plus Multi One Step Cloning Kit (YEASEN, Shanghai, China), resulting in plasmid pVan. The mutations in the promoter region were introduced with Stratagene QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies Inc, Santa Clara, USA).

The fragment containing the *vanR* gene encoding transcriptional regulatory protein (Genbank accession No. 1020332) was amplified using the genome of strain *C. glutamicum* ATCC 13032 as template with primers *vanR*-Gibson-1 and *vanR*-Gibson-2. Using plasmid pVan as template, amplification was performed with primers pVan-*backbone*-Gibson-1 and pVan-*backbone*-Gibson-2. The two purified fragments were assembled using Hieff Clone Plus Multi One Step Cloning Kit, resulting in plasmid pVanR.

The wild-type or mutant *vanR* genes were amplified using plasmid pVanR as template with primers *VanR*-for-*EcoRI* and *VanR*-rev-*XhoI*, and ligated into plasmid pGEX-6p-1 after digestion with *EcoRI* and *XhoI*, resulting in plasmid pGEX-6p-1-*vanR* carrying wild-type or mutant *vanR* gene.

**Strain construction and culture conditions.** The *lacZ* gene was amplified from the genomic DNA of strain MG1655 with primers *lacZ*-for-*KpnI* and *lacZ*-rev-*BglII* and ligated into plasmid pVanR carrying gene encoding the VanR mutant HyT12, after digestion with *KpnI* and *BglII*. Then the fragment containing *P*<sub>*vanABK*</sub> controlled *lacZ* (*P*<sub>*vanABK*</sub>-*lacZ*) and gene encoding the VanR mutant HyT12 was amplified with primers pAH156-*P*<sub>*vanABK*</sub>-for and pAH156-*P*<sub>*vanABK*</sub>-rev, and assembled with the PCR fragment amplified with primers *vanR*-pAH156-for and pAH156-*P*<sub>*vanABK*</sub>-rev using the CRIM plasmid pAH156<sup>32</sup> as template, by Gibson Assembly. The construct was then integrated into the chromosome of strain BHYT using helper plasmid pAH69<sup>32</sup>, resulting in strain H2. The integration was verified by PCR.

**Library construction.** A site-saturation mutagenesis library of HpaBC with amino acid residues S210, A211 and Q212 in the HpaB component randomized was constructed on plasmid pFA-*hpaBC*<sup>19</sup> with primers *hpaBC*-Saturated-for and *hpaBC*-Saturated-rev<sup>19</sup>. Around 10<sup>5</sup> transformants were collected for downstream screening.

To construct a random mutagenesis library based on the HpaB mutant 23F9, Error-prone PCR was performed using plasmid pFA-*hpaB*<sub>23F9C</sub> as template with primers *hpaB*-ErrorProne-for and *hpaB*-ErrorProne-rev. The PCR reaction was done as described<sup>19</sup>. The PCR products obtained, containing randomly mutated *hpaB*<sub>23F9C</sub> gene, were used as the megaprimer to perform the MEGAWHOP PCR<sup>33</sup> using plasmid pFA-*hpaB*<sub>23F9C</sub> as template. *DpnI* (20 U) digestion was performed at 37 °C for 2 h and then inactivated at 80 °C for 20 min. Then the PCR products were used to transform strain *E. coli* MC1061 and around 10<sup>5</sup> transformants were recovered. Randomly picked clones from the libraries were sequenced, and an average of 2.5 nucleotide mutations per clone was confirmed. All colonies from the agar plates were used for plasmid isolation to prepare the plasmid library.

To construct the site-saturation mutagenesis library of VanR, two parallel PCR reactions were performed with plasmid pVanR as template using the following two sets of primers: *vanR1*-for/*vanR1*-rev and *vanR2*-for/*vanR2*-rev. The two amplified fragments, *vanR*-F1 and *vanR*-F2, contained the site-saturation mutagenesis at positions 103, 143 and 168. Equimolar aliquots of the two fragments (1.6 nmole each) were PCR-assembled without primers, and then primers *vanR*-F0-for and *vanR*-F0-rev were added to amplify the PCR products. The PCR products were used as megaprimer to perform MEGAWHOP PCR using pVanR as template as described above. Around 5 × 10<sup>4</sup> transformants were recovered. Randomly picked clones from the library were sequenced, and the mutations at the indicated positions were confirmed with no additional point mutations.

The construction of the random mutagenesis library of TYO was similar to the HpaBC random mutagenesis library as described above, except that primer pairs *tyo*-ErrorProne-for/*tyo*-ErrorProne-rev and template plasmid pRSF-*tyo-hpaB*<sub>23F9-M4C-dodc</sub> carrying the gene encoding TYO wild-type or mutant MUT1 were used. A total of 10<sup>4</sup> transformants were recovered and randomly picked clones were sequenced and contained an average of 3 (library of wild-type *tyo*) or 2.5 (library of *tyo*-mut1) nucleotide mutations per clone.

**High-throughput screening of the mutagenesis libraries.** The HpaBC library was screened with the periodate assay method<sup>34</sup>. Single colonies of strain BHYT harboring mutants from the site-saturation mutagenesis library of HpaB or random mutagenesis library of HpaB mutant 23F9 were grown in 0.6 mL M9Y medium in 96-well plates at 37 °C for 24 h, induced with 1 mM L-arabinose and

supplemented with 3 mM tyrosine as substrate. Cells were harvested by centrifugation (1840 × g, 10 min, 4 °C). 180 μL of the supernatant was added into 20 μL of 100 mM sodium periodate, and OD<sub>400</sub> of the reaction mixture was determined by a SynergyMx Multi-Mode Microplate Reader (BioTek, Vermont, USA). Mutants showing higher absorbance than the parent enzymes were selected.

The VanR plasmid library was used to transform strain BW25113 (3.5 × 10<sup>5</sup> transformants were recovered). FACS screening was performed on a FACSAria II sorter (BD, San Jose, USA). Fluorescence was excited at 561 nm and the fluorescence emission was detected using a 610/20 nm band-pass filter. In the first round of negative screening, cells were precultured overnight in LB medium and then diluted to OD<sub>600</sub> = 0.2 in the same medium in the presence of 1 mM tyrosine, 1 mM L-DOPA, 1 mM dopamine and 1 mM 3,4-DHPAA. The cells were then grown for 12 h and the least fluorescent 2.8 × 10<sup>5</sup> cells were collected from a total of 3.5 × 10<sup>5</sup> cells, to eliminate clones that were induced by substrate and biosynthetic intermediate compounds. The collected cells were then induced with 1 mM hydroxytyrosol and grown for 12 h to enable positive screening, in which the most fluorescent 3.5 × 10<sup>3</sup> cells were sorted from a total of 3.5 × 10<sup>5</sup> cells (Supplementary Fig. 11). These itinerary screens were repeated for seven times, and seven clones were selected for rescreening in test tubes.

The TYO random mutagenesis libraries were screening on agar plates. The libraries were used to transform strain H2. After incubation at 37 °C for 24 h on YM9 agar supplemented with 3–6 mM tyrosine, 1 mM L-arabinose, 20 μg mL<sup>-1</sup> X-GAL and 50 μg mL<sup>-1</sup> kanamycin, the darkest blue colonies (by the eye) were selected for quantification of hydroxytyrosol productions in liquid culture.

**HPLC quantification.** A colony of strain BW25113 harboring plasmid pFA-*hpaBC*<sup>19</sup> carrying wild-type or mutant *hpaBC* gene, or strain BHYT harboring the hydroxytyrosol biosynthetic pathway, was grown in M9Y medium at 37 °C till OD<sub>600</sub> = 0.6, then the cells were induced with 1 mM L-arabinose and 3 mM tyrosine was supplemented. After further grown at 37 °C for 24 h, the cell culture was centrifuged at 12,000 × g for 5 min, and the supernatant was filtrated through a 0.22 μm filter and analyzed with HPLC. Concentrations of L-DOPA, tyrosine and hydroxytyrosol were determined by HPLC<sup>19</sup>. The concentrations of dopamine were determined after derivatization with dansylchloride, detected at 254 nm, by HPLC<sup>19</sup>.

**Protein expression and purification.** A single colony of strain BL21(DE3) harboring plasmid pGEX-6p-1-*vanR* carrying wild-type or mutant *vanR* (Y103L, R148L and Y168M), was grown in LB medium at 37 °C till OD<sub>600</sub> = 0.6, then induced with 0.4 mM IPTG. Full length of wild-type and mutant VanR were expressed as GST fusion protein in strain BL21(DE3). Proteins were purified using a glutathione Sepharose 4B column. After the GST-tag was cleaved by PreScission protease (GE Healthcare), the proteins were further purified by a second glutathione Sepharose 4B column, the collected fractions containing target proteins were then loaded into MonoQ column for further purification with a linear gradient elution, and fractions containing VanR or mutant VanR from MonoQ were finally purified through Superdex 75 (GE Healthcare) column equilibrated with buffer (20 mM Tris, pH 8.0, 200 mM NaCl, 2 mM DTT). The wild-type and mutant VanR were concentrated to 10 mg mL<sup>-1</sup> and stored at -80 °C for use.

The purification of HpaB and HpaC were performed using nickel-nitrilotriacetic acid (Ni-NTA) column (Qiagen, Valencia, USA)<sup>19</sup>. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to assess the protein purity. Protein concentration was assayed with Bradford method<sup>35</sup>.

**Crystallization and structure determination.** Purified VanR was concentrated to 7 mg mL<sup>-1</sup>. VanR and vanillic acid were incubated for on ice 2 h before crystallization. Crystal was obtained in 0.2 M magnesium chloride, 0.1 M Tris pH 8.0, 12% PEG4000 by hanging drop method at 18 °C. Then the crystal was flash-cooled in liquid nitrogen to -170 °C. The diffraction data were collected at 100 K at the synchrotron radiation at beam line of Shanghai Synchrotron Radiation Facility. The diffraction data sets were integrated and scaled with XDS package<sup>36,37</sup>. The structure of VanR complexed with vanillic acid was solved by molecular replacement using the structure of smu.1604c from *Streptococcus* mutants UA159 (<https://www.rcsb.org/structure/3L9F>) as the initial model with Phaser<sup>38</sup>. Refmac 5 and Coot9 in CCP4 suite were used for the refinement and model building<sup>39,40</sup>.

**Isothermal titration calorimetry.** For isothermal titration calorimetry (ITC measurement), the purified wild-type or mutant VanR was dialysis overnight at 4 °C against 20 mM phosphate buffer containing 150 mM sodium chloride, pH 7.4. Vanillic acid was dissolved in the same buffer. Isothermal titration calorimetry was carried out at 25 °C on an ITC200 system (MicroCal). 200 μM vanillic acid was titrated into 20 μM VanR. A titration of vanillic acid into buffer was carried out as a control. Data were fitted and analyzed using Origin 7 (MicroCal).

**Molecular modeling.** To generate the active HpaBC complexes with FAD and 4-HPA, we modeled an active structure using ThHpaBC (<https://www.rcsb.org/structure/2YYJ>) as a template by Modoller 9.20<sup>18,41</sup>. We replace the loop of apo HpaBC (<https://www.rcsb.org/structure/6QW0>) with the previously modeled loop (residues 206–218) by YASARA. Then the model was further minimized using

GROMACS under force field gromacs 96<sup>42</sup>. Finally, FAD and 4-HPA were docked to model by Autodock Vina<sup>43</sup>.

The structures of screened VanR mutants were manually generated in Coot<sup>44</sup>. Then Rosseta ligand docking<sup>45,46</sup> was used to calculate the binding model. The conformation with lowest docking score were choose as final model, and Rosseta functions were used to evaluate the change of mutant proteins.

**Activity assay of HpaBC and TYO.** The activity assay of TYO was performed at 37 °C by monitoring the absorbance increase at 240 nm which indicates the formation of H<sub>2</sub>O<sub>2</sub><sup>47</sup>. A reaction mixture (200 μL) for the assay contained 100 mM phosphate buffer (pH 7.0), 0.75 mM dopamine and 0.002 mg mL<sup>-1</sup> purified TYO protein. Reaction was performed at 37 °C for 10 min. The H<sub>2</sub>O<sub>2</sub> formation was monitored at 240 nm, with a SynergyMx Multi-Mode Microplate Reader (BioTek, USA). One unit of activity was defined as the amount of enzyme that catalyzes the production of 1 μM of H<sub>2</sub>O<sub>2</sub> per min.

The HpaBC activity was assayed by quantifying the product formation with HPLC. One unit of activity was defined as the amount of enzyme catalyzing the conversion of 1 mM tyrosine per minute under the assay conditions<sup>19</sup>.

**Kinetic parameters determination.** The kinetic parameters of TYO enzyme were determined by Lineweaver-Burk plots. TYO activity assay was performed as described above at different dopamine concentrations (0.75–1.50 mM).

**Measurement of β-galactosidase activities.** The β-galactosidase activity was measured using the substrate o-nitrophenyl-β-D-galactopyranoside (ONPG)<sup>48</sup>. Crude enzyme extracts were prepared by ultrasonication, and 10 μL of the crude enzyme extract was incubated with 140 μL of McIlvaine buffer (200 mM Na<sub>2</sub>HPO<sub>4</sub>, 100 mM citric acid, pH 6.0) containing 1 mM ONPG at 37 °C for 10 min. The reaction was stopped by adding 150 μL of 200 mM Na<sub>2</sub>CO<sub>3</sub>. The absorbance increase at 420 nm resulted from the release of o-nitrophenol (ONP) was measured. One unit (U) of enzyme activity was defined as the amount of enzyme required to liberate the equivalent of 1 μmol of ONP per minute under the assay conditions.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

A reporting summary for this article is available as a Supplementary Information file. Data supporting the findings of this work are available within the paper and its Supplementary Information files. The datasets generated and analyzed during the current study are available from the corresponding author upon request. Structure of VanR complexed with vanillic acid has been deposited in the Protein Data Bank under accession code of 6LG2 [<https://www.rcsb.org/structure/6LG2>]. The source data underlying Figs. 1b, d, e, g, 3c–e, and 4b–g, Table 2, as well as Supplementary Figs. 2, 3, 7, 8a, 9, and 10 are provided as a Source Data file.

Received: 30 August 2019; Accepted: 11 February 2020;

Published online: 23 March 2020

## References

1. Wachtmeister, J. & Rother, D. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Curr. Opin. Biotech.* **42**, 169–177 (2016).
2. Dhamankar, H. & Prather, K. L. J. Microbial chemical factories: recent advances in pathway engineering for synthesis of value added chemicals. *Curr. Opin. Struct. Biol.* **21**, 488–494 (2011).
3. Mir, R., Jallu, S. & Singh, T. P. The shikimate pathway: review of amino acid sequence, function and three-dimensional structures of the enzymes. *Crit. Rev. Microbiol.* **41**, 172–189 (2015).
4. Erb, T. J., Jones, P. R. & Bar-Even, A. Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.* **37**, 56–62 (2017).
5. Lv, X. M. et al. Combinatorial pathway optimization in *Escherichia coli* by directed co-evolution of rate-limiting enzymes and modular pathway engineering. *Bioengineer.* **113**, 2661–2669 (2016).
6. Machovina, M. M. et al. Enabling microbial syringol conversion through structure-guided protein engineering. *Proc. Natl. Acad. Sci. USA* **116**, 13970–13976 (2019).
7. Tang, S. Y. et al. Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. *J. Am. Chem. Soc.* **135**, 10099–10103 (2013).
8. Xiong, D. D. et al. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metab. Eng.* **40**, 115–123 (2017).
9. Li, H. et al. Monitoring in vivo metabolic flux with a designed whole-cell metabolite biosensor of shikimic acid. *Biosens. Bioelectron.* **98**, 457–465 (2017).

10. Chen, W. et al. Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis. *Metab. Eng.* **30**, 149–155 (2015).
11. Granados-Principal, S., Quiles, J. L., Ramirez-Tortosa, C. L., Sanchez-Rovira, P. & Ramirez-Tortosa, M. C. Hydroxytyrosol: from laboratory investigations to future clinical trials. *Nutr. Rev.* **68**, 191–206 (2010).
12. Bernini, R., MSG, Montani, Merendino, N., Romani, A. & Velotti, F. Hydroxytyrosol-derived compounds: a basis for the creation of new pharmacological agents for cancer prevention and therapy. *J. Med. Chem.* **58**, 9089–9107 (2015).
13. Satoh, Y., Tajima, K., Munekata, M., Keasling, J. D. & Lee, T. S. Engineering of L-tyrosine oxidation in *Escherichia coli* and microbial production of hydroxytyrosol. *Metab. Eng.* **14**, 603–610 (2012).
14. Chung, D., Kim, S. Y. & Ahn, J. H. Production of three phenylethanoids, tyrosol, hydroxytyrosol, and salidroside, using plant genes expressing in *Escherichia coli*. *Sci. Rep.* **7**, 2578 (2017).
15. Li, X. L. et al. Establishing an artificial pathway for efficient biosynthesis of hydroxytyrosol. *ACS Synth. Biol.* **7**, 647–654 (2018).
16. Prieto, M. A., Perezaranda, A. & Garcia, J. L. Characterization of an *Escherichia coli* aromatic hydroxylase with a broad substrate range. *J. Bacteriol.* **175**, 2162–2167 (1993).
17. Shen, X. L. et al. Structural insights into catalytic versatility of the flavin-dependent hydroxylase (HpaB) from *Escherichia coli*. *Sci. Rep.* **9**, 7087 (2019).
18. Kim, S. H. et al. Crystal structure of the oxygenase component (HpaB) of the 4-hydroxyphenylacetate 3-monooxygenase from *Thermus thermophilus* HB8. *J. Biol. Chem.* **282**, 33107–33117 (2007).
19. Chen, W. et al. Promiscuous enzymatic activity-aided multiple-pathway network design for metabolic flux rearrangement in hydroxytyrosol biosynthesis. *Nat. Commun.* **10**, 960 (2019).
20. Heravi, K. M., Lange, J., Watzlawick, H., Kalinowski, J. & Altenbuchner, J. Transcriptional regulation of the vanillate utilization genes (vanABK operon) of *Corynebacterium glutamicum* by VanR, a PadR-like repressor. *J. Bacteriol.* **197**, 959–972 (2015).
21. Daubner, S. C., Le, T. & Wang, S. Z. Tyrosine hydroxylase and regulation of dopamine synthesis. *Arch. Biochem. Biophys.* **508**, 1–12 (2011).
22. Maass, A., Scholz, J. & Moser, A. Modeled ligand-protein complexes elucidate the origin of substrate specificity and provide insight into catalytic mechanisms of phenylalanine hydroxylase and tyrosine hydroxylase. *Eur. J. Biochem.* **270**, 1065–1075 (2003).
23. Rogers, J. K. et al. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res.* **43**, 7648–7660 (2015).
24. Skjoedt, M. L. et al. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol.* **12**, 951–958 (2016).
25. Liu, Y. F. et al. Biosensor-based evolution and elucidation of a biosynthetic pathway in *Escherichia coli*. *ACS Synth. Biol.* **6**, 837–848 (2017).
26. Yu, H. et al. Engineering transcription factor BmoR for screening butanol overproducers. *Metab. Eng.* **56**, 28–38 (2019).
27. Barbeau, A. L-Dopa therapy in Parkinson's disease: a critical review of nine years experience. *Can. Med. Assoc. J.* **101**, 59–68 (1970).
28. Hagel, J. M. & Facchini, P. J. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant Cell Physiol.* **54**, 647–672 (2013).
29. Min, K., Park, K., Park, D. H. & Yoo, Y. J. Overview on the biotechnological production of L-DOPA. *Appl. Microbiol. Biot.* **99**, 575–584 (2015).
30. DeLoache, W. C. et al. An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat. Chem. Biol.* **11**, 465–471 (2015).
31. Savitskaya, J., Protzko, R. J., Li, F. Z., Arkin, A. P. & Dueber, J. E. Iterative screening methodology enables isolation of strains with improved properties for a FACS-based screen and increased L-DOPA production. *Sci. Rep.* **9**, 5815 (2019).
32. Haldimann, A. & Wanner, B. L. Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.* **183**, 6384–6393 (2001).
33. Miyazaki, K. Megawhop cloning: A method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Method Enzymol.* **498**, 399–406 (2011).
34. Orenes-Pinero, E., Garcia-Carmona, F. & Sanchez-Ferrer, A. A new process for obtaining hydroxytyrosol using transformed *Escherichia coli* whole cells with phenol hydroxylase gene from *Geobacillus thermoglucosidasius*. *Food Chem.* **139**, 377–383 (2013).
35. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
36. Kabsch, W. XDS. *Acta Crystallogr. D.* **66**, 125–132 (2010).
37. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D.* **62**, 72–82 (2006).
38. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* **60**, 2126–2132 (2004).
40. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D.* **67**, 355–367 (2011).
41. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **47**, 5.6.1–32 (2014).
42. Pronk, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
43. Trott, O. & Olson, A. J. Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
44. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D.* **66**, 486–501 (2010).
45. Lyskov, S. et al. Serverification of molecular modeling applications: the Rosetta online server that includes everyone (ROSIE). *PLoS ONE* **8**, e63906 (2013).
46. DeLuca, S., Khar, K. & Meiler, J. Fully flexible docking of medium sized ligand libraries with Rosetta ligand. *PLoS ONE* **10**, e0132508 (2015).
47. Yumoto, I., Fukumori, Y. & Yamanaka, T. Purification and characterization of catalase from a facultative alkalophilic *Bacillus*. *J. Biochem.* **108**, 583–587 (1990).
48. Liang, C. N., Xiong, D. D., Zhang, Y., Mu, S. S. & Tang, S. Y. Development of a novel uric-acid-responsive regulatory system in *Escherichia coli*. *Appl. Microbiol. Biot.* **99**, 2267–2275 (2015).

### Acknowledgements

We thank Dr. Junying Jia and Dr. Shuang Sun from Institute of Biophysics, Chinese Academy of Sciences for FACS screening. Financial support was provided by the National Key Research and Development Program of China (Grant 2018YFA0900701) and the National Natural Science Foundation of China (Grant No. 31961133016, 31870072, 31971382 and 31971337).

### Author contributions

S.-Y.T., H.S., W.C., and J.-M.J conceived the idea and supervised the research. J.Y. and W.C. did the construction of strains, plasmids and libraries. Y.H. did molecular modeling. J.Y. did the screening of HpaBC variants and TYO variants. Y.H., N.S., and S.B. did protein expression, purification, crystallization and structure determination. C.W. did the screening of VanR variants and the construction of biosensor. J.Y. and W.C. did the HPLC analysis. J.Y. did enzyme activity assay and kinetic parameters determination. Y.H. did isothermal titration calorimetry (ITC). W.C. did flux analysis. Y.T. contributed reagents/materials/analysis tools for fed-batch fermentation as well as whole-cell bio-transformation and critically reviewed the manuscript. S.-Y.T., H.S., W.C., and J.-M.J performed data analysis and wrote the paper. All of the authors reviewed, approved and contributed to the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information


Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-14918-5>.

Correspondence and requests for materials should be addressed to J.-M.J., W.C., H.S. or S.-Y.T.

Peer review information *Nature Communications* thanks Michael Jensen and the other, anonymous, reviewer for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020