

ARTICLE

Received 23 Oct 2016 | Accepted 6 Apr 2017 | Published 18 May 2017 | Updated 13 Jul 2018

DOI: 10.1038/ncomms15544

OPEN

# Improving 10-deacetylbaaccatin III-10- $\beta$ -O-acetyltransferase catalytic fitness for Taxol production

Bing-Juan Li<sup>1,2</sup>, Hao Wang<sup>1,2</sup>, Ting Gong<sup>1,2</sup>, Jing-Jing Chen<sup>1,2</sup>, Tian-Jiao Chen<sup>1,2</sup>, Jin-Ling Yang<sup>1,2</sup> & Ping Zhu<sup>1,2</sup>

The natural concentration of the anticancer drug Taxol is about 0.02% in yew trees, whereas that of its analogue 7- $\beta$ -xylosyl-10-deacetyltaxol is up to 0.5%. While this compound is not an intermediate in Taxol biosynthetic route, it can be converted into Taxol by de-glycosylation and acetylation. Here, we improve the catalytic efficiency of 10-deacetylbaaccatin III-10-O-acetyltransferase (DBAT) of *Taxus* towards 10-deacetyltaxol, a de-glycosylated derivative of 7- $\beta$ -xylosyl-10-deacetyltaxol to generate Taxol using mutagenesis. We generate a three-dimensional structure of DBAT and identify its active site using alanine scanning and design a double DBAT mutant (DBAT<sup>G38R/F301V</sup>) with a catalytic efficiency approximately six times higher than that of the wild-type. We combine this mutant with a  $\beta$ -xylosidase to obtain an *in vitro* one-pot conversion of 7- $\beta$ -xylosyl-10-deacetyltaxol to Taxol yielding 0.64 mg ml<sup>-1</sup> Taxol in 50 ml at 15 h. This approach represents a promising environmentally friendly alternative for Taxol production from an abundant analogue.

<sup>1</sup>State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, 1 Xian Nong Tan Street, Beijing 100050, China. <sup>2</sup>Key Laboratory of Biosynthesis of Natural Products, National Health and Family Planning Commission of PRC, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, 1 Xian Nong Tan Street, Beijing 100050, China. Correspondence and requests for materials should be addressed to P.Z. (email: zhuping@imm.ac.cn).

**T**axol or paclitaxel (generic name) is one of the most well-known anticancer drugs among natural products<sup>1</sup>. It exerts anticancer activity through promoting tubulin assembly into microtubules and preventing their disassembly<sup>2</sup>. Indications include refractory ovarian cancer, breast cancer and squamous cancers<sup>1</sup>. Taxol is also used to treat psoriasis<sup>3</sup>, rheumatoid arthritis<sup>4</sup> and restenosis<sup>5</sup> for which Taxol is used for the coronary stent or balloon catheter coating<sup>6–8</sup>. In the last decade, at least nine new taxoid derivatives or formulations have undergone clinical trials<sup>1</sup>, including Nab-paclitaxel, Cabazitaxel and Larotaxel.

Taxol is mainly produced by yew trees. With an extremely low content, its concentration ranges from 0 to 0.069% (tree material dry weight), with the highest percentage found in the bark<sup>9</sup>. Therefore, at the early stage of drug development, the destructive collection of bark for Taxol threatened wild resources and was ultimately prohibited around the world<sup>10</sup>. For example, all *Taxus* species were under the Grade I national protection in the National Key Protected Wild Plants List (1999) published by the Chinese government ([http://www.gov.cn/gongbao/content/2000/content\\_60072.htm](http://www.gov.cn/gongbao/content/2000/content_60072.htm)). *T. wallichiana* was also listed on Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) in 1995, and afterwards other *Taxus* species, such as *T. chinensis* and *T. cuspidata*, were added to the Appendix II. Consequently, alternative Taxol sources have been explored to meet increasing clinical requirements<sup>11–14</sup>. Currently, Taxol produced from seedling culture, forestation of yew trees and chemical semi-synthesis of Taxol from the precursor 10-deacetylbaccatin III (10-DAB) have become the major Taxol sources for clinical supply<sup>15,16</sup>. 10-DAB is also one of the key intermediates in the biosynthetic pathway of Taxol and is converted into baccatin III by the enzyme 10-deacetylbaccatin III-10-O-acetyltransferase (DBAT)<sup>17</sup>.

7- $\beta$ -Xylosyl-10-deacetyltaxol (XDT) is an analogue of Taxol. This analogue is not involved in the Taxol biosynthetic pathway from 10-DAB to Taxol, but may share the same precursor 10-DAB. Its content in yew trees is as much as 0.5% (dry weight of the bark)<sup>18</sup>. XDT content is even higher than that of Taxol in nursery cultivated yew trees<sup>19,20</sup>. This analogue is usually discarded during Taxol extraction, causing both resource waste and potential environmental pollution. If the 7- $\beta$ -xylosyl group is removed by a specific  $\beta$ -xylosidase, then the resulting 10-deacetyltaxol (DT) can also be used for the chemical semi-synthesis of Taxol by acetylation at the C10 position<sup>21,22</sup>. The corresponding bacteria with the  $\beta$ -xylosidase activity were explored by several other groups<sup>23–27</sup>, in which a nanoscale cellulosome-like multienzyme complex was found from *Cellulimicrobium cellulans*<sup>27</sup>. Recently, we even cloned and characterized a new  $\beta$ -xylosidase (designated as LXYL-P1-2) from *Lentinula edodes*, which showed much lower sequence identities than other known enzymes including the reported  $\beta$ -xylosidases<sup>28,29</sup>. This enzyme can specifically hydrolyse the  $\beta$ -xylosyl group of XDT and is heterologously expressed in *Pichia pastoris*. We also achieved a large-scale fermentation (up to 1,000-l fermenter)<sup>30,31</sup> and bioconversion (up to 10-l reaction)<sup>30,32</sup> of the recombinant yeast, and the DT yield reached 10.58 g l<sup>-1</sup> in one-liter reaction volume, which was much higher than the earlier reports<sup>23–26</sup> (Supplementary Table 1).

These efforts prompted the exploration of an environmentally friendly method to acetylate the C10 position of DT by acetyltransferase, and to even combine the enzyme with LXYL-P1-2 in a one-pot reaction to produce Taxol from XDT. Apparently, DBAT is an excellent candidate, because its natural substrate 10-DAB harbours the same taxane nucleus as that of DT except for their difference at the C13 side-chain. DBAT has

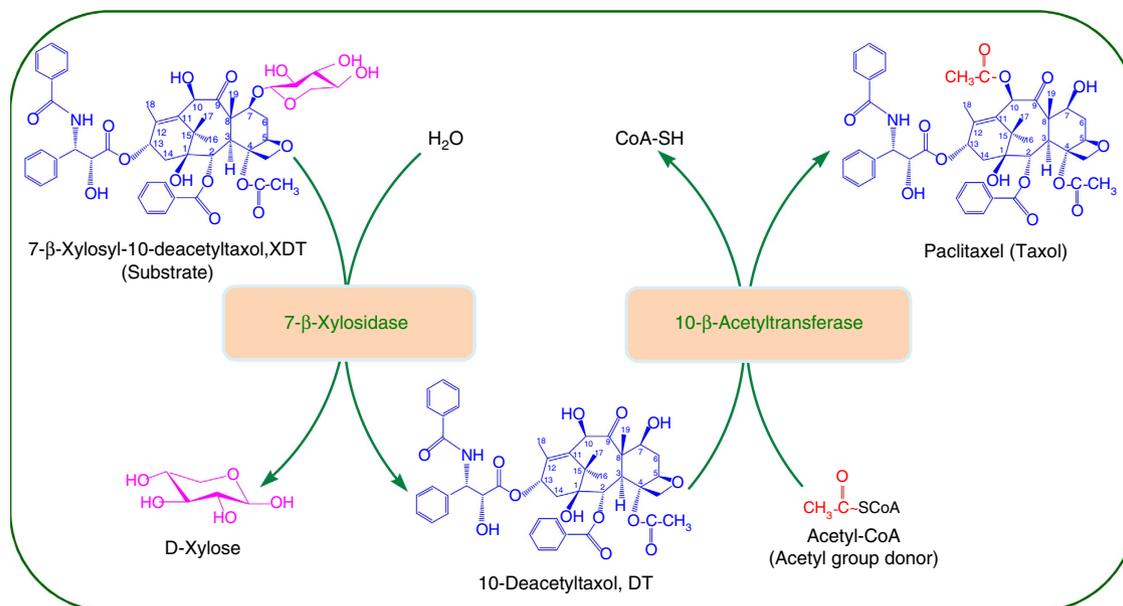
been widely studied by several groups<sup>33–36</sup>. Menhard and Zenk<sup>33</sup> purified the O-acetyltransferase from cell suspension cultures of *T. chinensis* and characterized it. The purified acetyltransferase was found to be a monomeric protein of 71 kDa with an optimum pH of 9 and an optimum temperature of 35 °C. Walker and Croteau<sup>34</sup> cloned the DBAT gene from *T. cuspidata*, heterologously expressed it in *Escherichia coli* and characterized the recombinant enzyme. The recombinant enzyme was also monomeric but with a molecular weight of 49 kDa and an optimum pH of 7.4. It seemed to acetylate the 10-hydroxyl group with a high degree of regioselectivity, as the enzyme did not acetylate the 1 $\beta$ -, 7 $\beta$ - or 13 $\alpha$ -hydroxyl groups of 10-DAB, nor did it acetylate the 5 $\alpha$ -hydroxyl group of taxa-4(20),11(12)-dien-5 $\alpha$ -ol<sup>35</sup>. Ondari and Walker<sup>35</sup> also demonstrated that the recombinant DBAT can acetylate docetaxel, a semi-synthetic taxoid of Taxol, at its natural C10 hydroxyl position, and suggested that acetylation at the C10 hydroxyl can occur after attachment of the phenylpropanoyl side chain to 10-DAB on the biosynthetic pathway. Pennington *et al.*<sup>36</sup> tried to find the acetylation activity of a partially purified DBAT from the leaves and 3-year-old cell suspension cultures of *T. cuspidata* both on 10-DAB and on DT. They observed the formation of baccatin III from 10-DAB in the presence of acetyl CoA, while repeated attempts to detect significant DT acetyltransferase activity were not successful. The inconsistent appearance of Taxol in the assays was thought to be the possible presence of an uncharacterized 10-O-acetyl-transferase as a contaminant in the enzyme preparation. Moreover, all of the products were not subjected to spectroscopic elucidation.

Here, we confirm the enzymatic acetylation of DT at its C10 hydroxyl position via the recombinant DBAT of *T. cuspidata* by spectroscopic detection. However, its catalytic efficiency was much lower than that on the natural substrate 10-DAB. Thus, we use protein engineering of DBAT to improve its enzymatic activity against DT, generating a double mutant DBAT<sup>G38R/F301V</sup> that has a  $k_{cat}/K_m$  of 3.54 M<sup>-1</sup> s<sup>-1</sup> on DT (DBAT with a  $k_{cat}/K_m$  of 0.61 M<sup>-1</sup> s<sup>-1</sup>). Finally, we construct an *in vitro* one-pot reaction system harbouring LXYL-P1-2, and the improved acetyltransferase to enzymatically convert XDT into Taxol (Fig. 1). This approach might be a promising alternative for Taxol production.

## Results

### Activities of the recombinant DBATs against 10-DAB and DT.

Six full-length DBAT cDNAs derived from *Taxus cuspidata* (GenBank accession: Q9M6E2.1), *T. brevifolia* (GenBank accession: EU107143.1), *T. baccata* (GenBank accession: AF456342.1), *T. canadensis* (GenBank accession: EU107134.1), *T. wallichiana* var (GenBank accession: EU107140.1) and *T. x media* (GenBank accession: AY452666.1) were synthesized and expressed in *E. coli*. The purified recombinant DBATs were obtained through Ni-affinity chromatography and molecular size exclusion chromatography, with purities greater than 95%. The catalytic activities of the DBATs against 10-DAB (control) and DT were detected, as shown in Fig. 2. Except for the recombinant DBAT of *T. baccata* origin whose activity was nearly undetectable, all other recombinant enzymes exhibited acetylation activity against 10-DAB and DT, in which the DBAT of *T. canadensis* exhibited the lowest conversion rate (Fig. 2). The optimum temperatures for 10-DAB and DT were 40 °C (Fig. 2a) and 37.5 °C (Fig. 2c), respectively. While the optimum pH values for both of the substrates were 5.5 (Fig. 2b,d), which was not consistent with the result in an earlier report on DBAT<sup>34</sup>. Under the same conditions, the DBATs of *T. wallichiana* var., *T. brevifolia*, *T. cuspidata* and *T. x media* origins showed similar conversion rates for 10-DAB (Fig. 2a,b),



**Figure 1 | One-pot reaction system for the bioconversion of XDT to Taxol.** The *in vitro* one-pot reaction system contained a specific 7- $\beta$ -xylosidase, the improved 10- $\beta$ -acetyltransferase, the substrate 7- $\beta$ -xylosyl-10-deacetyltaxol (XDT) and the acetyl group donor acetyl-CoA, respectively. The 7- $\beta$ -xylosidase deglycosylates XDT into the intermediate product 10-deacetyltaxol (DT) and then 10- $\beta$ -acetyltransferase acetylates 10-deacetyltaxol into the final product Taxol.

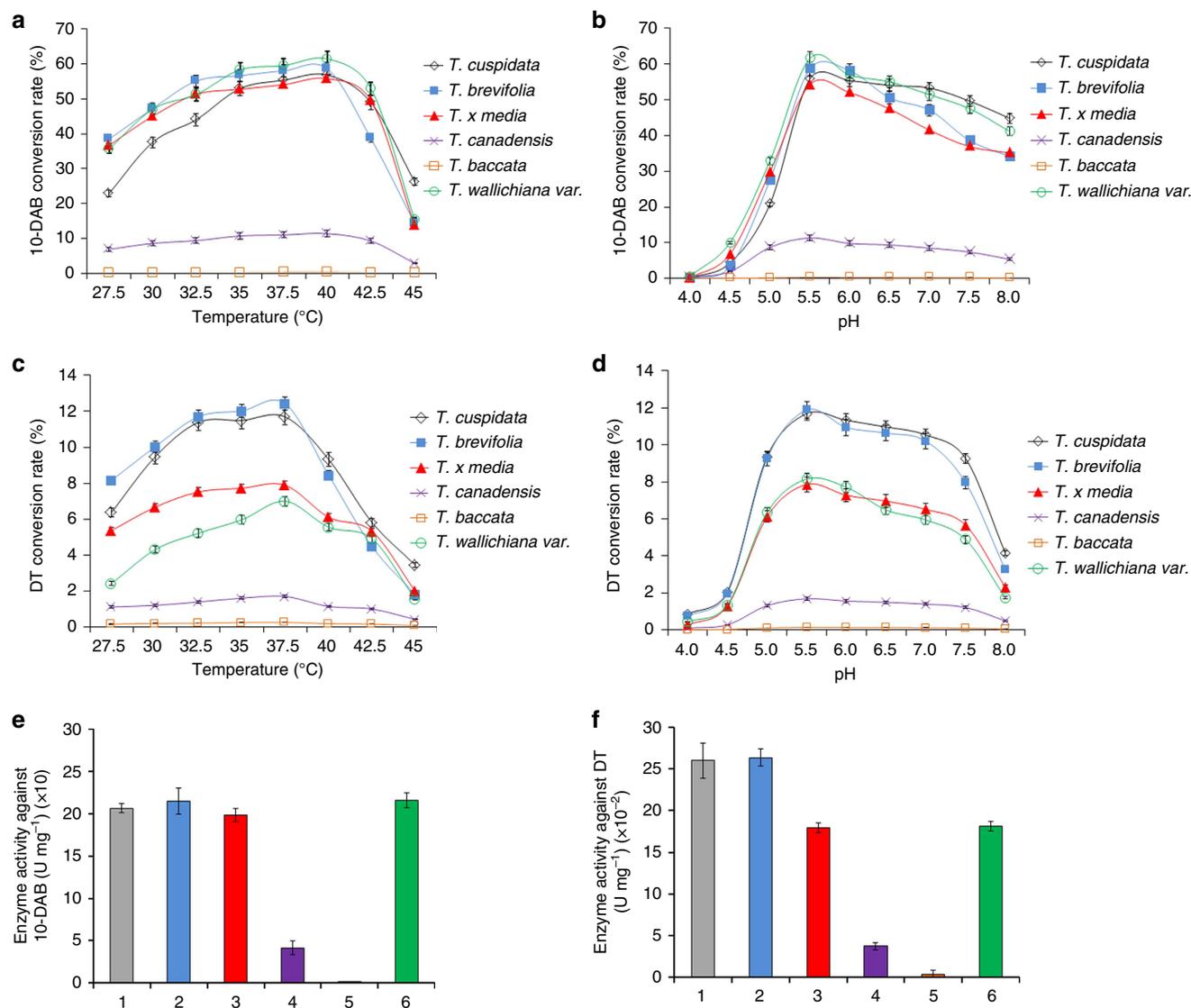
while the DBATs of *T. brevifolia* and *T. cuspidata* were obviously higher than those of *T. x media* and *T. wallichiana var.* regarding conversion rates against DT (Fig. 2c,d). Accordingly, under optimum conditions and among the six recombinant enzymes, the DBATs of *T. cuspidata*, *T. brevifolia*, *T. canadensis* and *T. wallichiana var.* showed similar high specific activities (from 198.7 to 215.7 U mg<sup>-1</sup>) against 10-DAB, almost 200 times higher than that of *T. baccata* (1.1 U mg<sup>-1</sup>) (Fig. 2e, Supplementary Table 2) which may partially explain why the latter accumulated significantly more 10-deacetylbaaccatin III in the twigs of *T. baccata* than other species<sup>37</sup>. Regarding the substrate DT, the DBATs of *T. cuspidata* and *T. brevifolia* showed the highest specific activities ( $\sim 0.26$  U mg<sup>-1</sup>), again over 80 times more than that of *T. baccata* (0.003 U mg<sup>-1</sup>) (Fig. 2f; Supplementary Table 2). However, all of the specific activities against DT dropped nearly three orders of magnitude compared to those of the same enzymes against the natural substrate 10-DAB (Fig. 2e,f; Supplementary Table 2).

The catalytic products of the recombinant DBAT of *T. cuspidata* were chosen for high-performance liquid chromatography/mass spectrometry (HPLC-MS) analysis. The [M + H]<sup>+</sup> and [M + Na]<sup>+</sup> peaks of the product in DBAT + DT group were 854.21 and 876.46, respectively (Fig. 3i), which were consistent with those of the standard (*m/z* 853.75 [M + H]<sup>+</sup>, 876.07 [M + Na]<sup>+</sup>) (Fig. 3g) and previous literature<sup>38</sup>. Furthermore, the nuclear magnetic resonance (NMR) spectroscopy data (Supplementary Fig. 1, Supplementary Table 3) of the purified product were in accordance with those reported in additional literature<sup>39</sup>. Thus, the acetylated DT was demonstrated to be Taxol.

**Three-dimensional structure prediction of DBAT.** To improve the specific activity of DBAT against the unnatural substrate DT, we carried out site-directed mutagenesis of the protein via a ‘semi-rational design’ strategy. Because its three-dimensional (3D) structure was not currently available, DBAT structure and active centre

were predicted by molecular modelling through online prediction software, Swissmodel (<http://swissmodel.expasy.org/>). Other tools included Discovery Studio 4.1, Procheck, Autodock 4.4, Autodock Vina, Chimera and Chemical Office. DBAT belongs to a plant acyl-CoA dependent acyltransferase superfamily, BAHD<sup>40</sup>, and is grouped into clade V in which 3D structures of *Coffea canephora* HCT: 4G22, *Sorghum bicolor* HCT: 4KE4 and *Panicum virgatum* PvHCT2: 5FAL have been reported. Although the sequences of the BAHD superfamily members are divergent, evidence has shown that these proteins possess highly similar 3D structures<sup>41</sup>. On the basis of the sequence alignment between DBAT and several other BAHD members (Supplementary Fig. 2), the *C. canephora* HCT (GenBank accession: ABO47805.1) was selected as the template for homology modelling of the DBAT structure. The sequence alignment between DBAT and the template is shown in Fig. 4a. The two enzymes exhibited 30% sequence identity and 45% sequence similarity, which was most similar among the aligned BAHD members. The predicted model was assessed by the Ramachandran plot<sup>42</sup>, indicating that the model was reasonable with 87.5% of the structure in the most favored regions, and only 0.8% of the residues in the disallowed regions (Supplementary Fig. 3). The predicted DBAT structure consisted of two nearly equal domains connected through a loop and was totally composed of 14  $\beta$ -sheets and 13  $\alpha$ -helices (Fig. 4b). DBAT showed similar hydrophobicity and distribution of electronegativity residues with the template (Supplementary Fig. 4).

It is generally accepted that the highly conserved HXXXD motif is crucial for the activity of BAHD members<sup>43</sup>. The motif has also been found in the DBAT structure, in which His<sup>162</sup> most likely acts as a catalytic site (Fig. 4b). In addition, the structure of the solvent channel of BAHD members is also critical for activity. The acyl donor substrate binding pocket of BAHD members is at the front side of the molecule (the position of the conserved DFGWG motif has been defined at the front side)<sup>43</sup>, which helps position the acceptor substrate pocket of DBAT at the back side (Fig. 4b, Supplementary Fig. 5). To identify the putative residues participating in the binding and/or catalysis of 10-DAB or DT,



**Figure 2 | Optimum temperature and pH and the specific activities of the DBATs. (a,b)** The optimum temperature and pH of the recombinant DBATs against 10-DAB. **(c,d)** The optimum temperature and pH of the recombinant DBATs against DT. **(e)** The specific activities of the six recombinant DBATs against 10-DAB. **(f)** The specific activities of the six recombinant DBATs against DT. (1. *Taxus cuspidata*, 2. *Taxus brevifolia*, 3. *Taxus x media*, 4. *Taxus canadensis*, 5. *Taxus baccata*, 6. *Taxus wallichiana var.*). The data represent the means  $\pm$  s.d.,  $n=3$ .

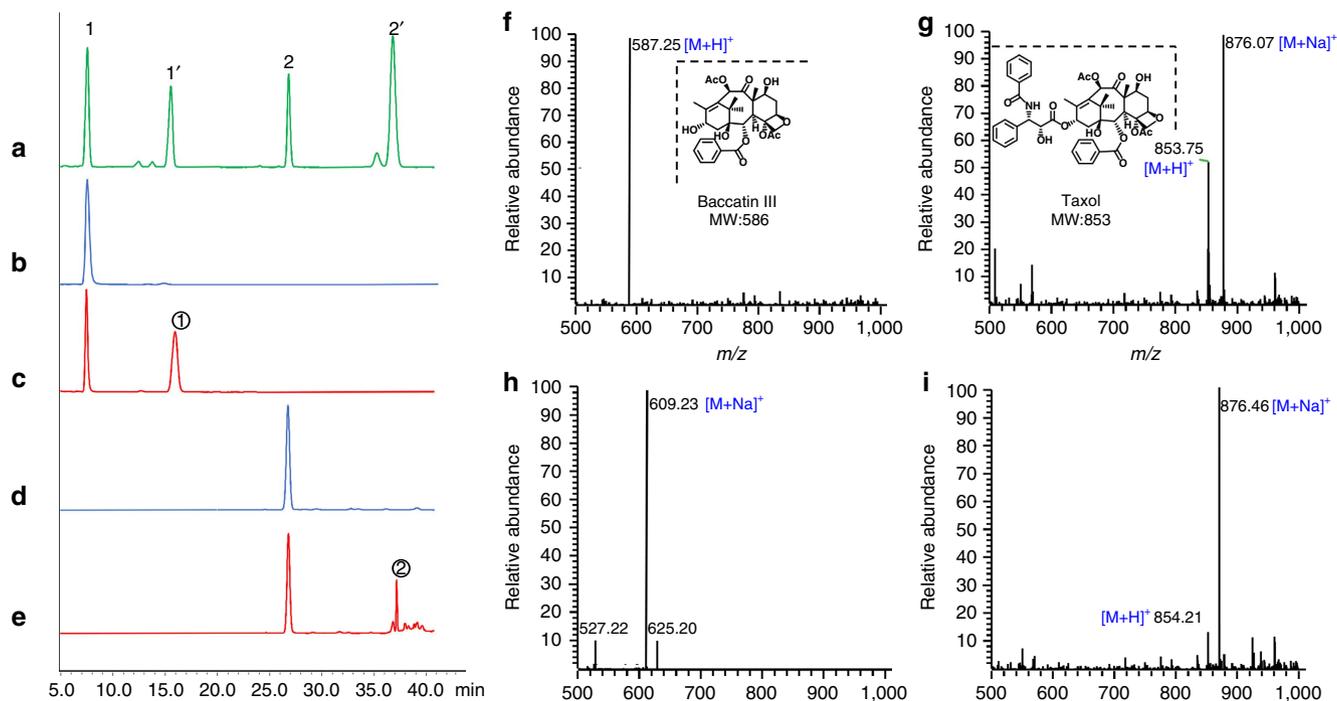
molecular docking was conducted (Fig. 4c,d). The residues within 5 Å distance from DT were selected as ‘hot spots’ for functional analysis.

#### L-alanine scanning mutagenesis of defined amino acids.

Enzyme activity generally declines if the amino acids involved in the substrate binding or catalytic sites are subjected to L-alanine scanning mutagenesis (mutated to Ala). However, this activity may be maintained or even increased if other amino acid residues are mutated to L-alanine<sup>44</sup>. The following amino acid residues within the 5 Å zone around the substrate (Fig. 4d) were chosen for L-alanine scanning mutagenesis: Pro<sup>37</sup>, Gly<sup>38</sup>, Arg<sup>40</sup>, Glu<sup>41</sup>, Phe<sup>44</sup>, Phe<sup>160</sup>, His<sup>162</sup>, Ile<sup>164</sup>, Cys<sup>165</sup>, Phe<sup>301</sup>, Ser<sup>351</sup>, Asn<sup>353</sup>, Gly<sup>359</sup>, Gly<sup>361</sup>, Arg<sup>363</sup>, Ser<sup>396</sup> and Phe<sup>400</sup>. Results (Fig. 5) showed that H162A had no measurable enzyme activity against 10-DAB or DT, which was consistent with the function of this residue as a potential catalytic base. Similarly, each of the R363A, G361A and I164A mutations lost virtually all activity against 10-DAB or DT, which also implied that these residues are potential substrate binding or catalytic sites. F44A activity was nearly undetectable

against DT and significantly declined against 10-DAB, which means that this residue may be involved in catalysis especially to the unnatural substrate DT. Other mutations, such as G359A, F400A, P37A and F160A, also negatively affected enzyme activities, but the impact of each mutation on 10-DAB and DT were quite different. However, F301A and G38A activities against DT significantly increased compared with that of the wild-type control, leading to 1.6 (F301A) and 1.45 (G38A) times more active than the wild-type, respectively. But their activities against 10-DAB were apparently degraded (F301A) or slightly decreased (G38A) (Fig. 5, Supplementary Table 4).

**Saturation mutagenesis of Gly<sup>38</sup> and Phe<sup>301</sup>.** To identify more active mutants against DT, we performed saturation mutagenesis of Gly<sup>38</sup> and Phe<sup>301</sup> using degenerated primers and the whole plasmid PCR amplification technique, changing the codon to NNK (N = A/T/C/G, K = G/T). A total of 38 mutants from Gly<sup>38</sup> and Phe<sup>301</sup> were obtained and activities against DT were detected as shown in Fig. 6a,b (also see Supplementary Table 5 and 6). The mutants DBAT<sup>G38R</sup>, DBAT<sup>G38S</sup>, DBAT<sup>G38D</sup>, DBAT<sup>G38H</sup>,



**Figure 3 | HPLC-MS analysis of the DBAT acetylated products.** HPLC analysis: (a) Standards of 10-DAB (1), baccatin III (1'), DT (2) and Taxol (2'). (b) Control of (c) (only 10-DAB, without the enzyme). (c) Reaction of DBAT on 10-DAB. (d) Control of (e) (only DT, without the enzyme). (e) Reaction of DBAT on DT. ESI-MS analysis: (f,g) Standards of baccatin III (1') and Taxol (2') in a. (h,i) Products ① and ② in c,e.

DBAT<sup>F301V</sup> and DBAT<sup>F301C</sup> surpassed the corresponding DBAT<sup>G38A</sup> or DBAT<sup>F301A</sup> in terms of activity, for which DBAT<sup>G38R</sup> and DBAT<sup>F301V</sup> were 2.19 and 2.85 times more active than the wild-type, respectively.

Two mutants, DBAT<sup>G38R</sup> and DBAT<sup>F301V</sup>, were combined to form the double mutant DBAT<sup>G38R/F301V</sup>. This combination did not change the optimum temperature and optimum pH when compared to those of the wild-type DBAT (Fig. 6c,d), but further increased enzyme activity 3.7 times against DT (Table 1). However, except for DBAT<sup>G38R</sup>, which showed two times greater activity against 10-DAB, both DBAT<sup>G38R/F301V</sup> and DBAT<sup>F301V</sup> exhibited decreased activities against 10-DAB compared with the control (Table 1). Enzyme kinetic analysis showed that catalytic efficiency ( $k_{cat}/K_m$ ) of DBAT<sup>G38R/F301V</sup> against DT surpassed those of DBAT<sup>G38R</sup> and DBAT<sup>F301V</sup>, and was significantly higher than that of the wild-type control ( $P < 0.001$ , Table 2).

#### Selection of DBAT<sup>G38R/F301V</sup> concentration in reaction system.

The substrate DT concentration was assigned to 2 mM and dissolved in the reaction buffer solution with a final concentration of 5% dimethyl sulfoxide (DMSO, v/v). The suitable enzyme concentration was determined under 37.5 °C and pH 5.5. The Taxol yield increased with the increasing amount of DBAT<sup>G38R/F301V</sup>. However, when the enzyme concentration reached 2.0 mg ml<sup>-1</sup>, the increasing amplitude of Taxol yield began to plateau (Fig. 7a). Because Taxol yields were similar between enzyme concentrations of 1.5 and 2.0 mg ml<sup>-1</sup> at the same time points (Fig. 7a), to limit the enzyme consumption, a 1.5 mg ml<sup>-1</sup> enzyme concentration was used to observe changes in the concentrations of DT and Taxol in the reaction system. The Taxol yield reached 452 μg ml<sup>-1</sup> after 12 h (Fig. 7c). As the linear portion of the time course plot was within 3 h (Fig. 7a), the supplementing strategy was utilized beginning at 3 h (near the inflection point), and additional DBAT<sup>G38R/F301V</sup> was repeatedly supplemented in the amount of 1.5 mg ml<sup>-1</sup> in 3 h intervals (Fig. 7d). The Taxol yield reached 641 μg ml<sup>-1</sup> after 15 h (Fig. 7d).

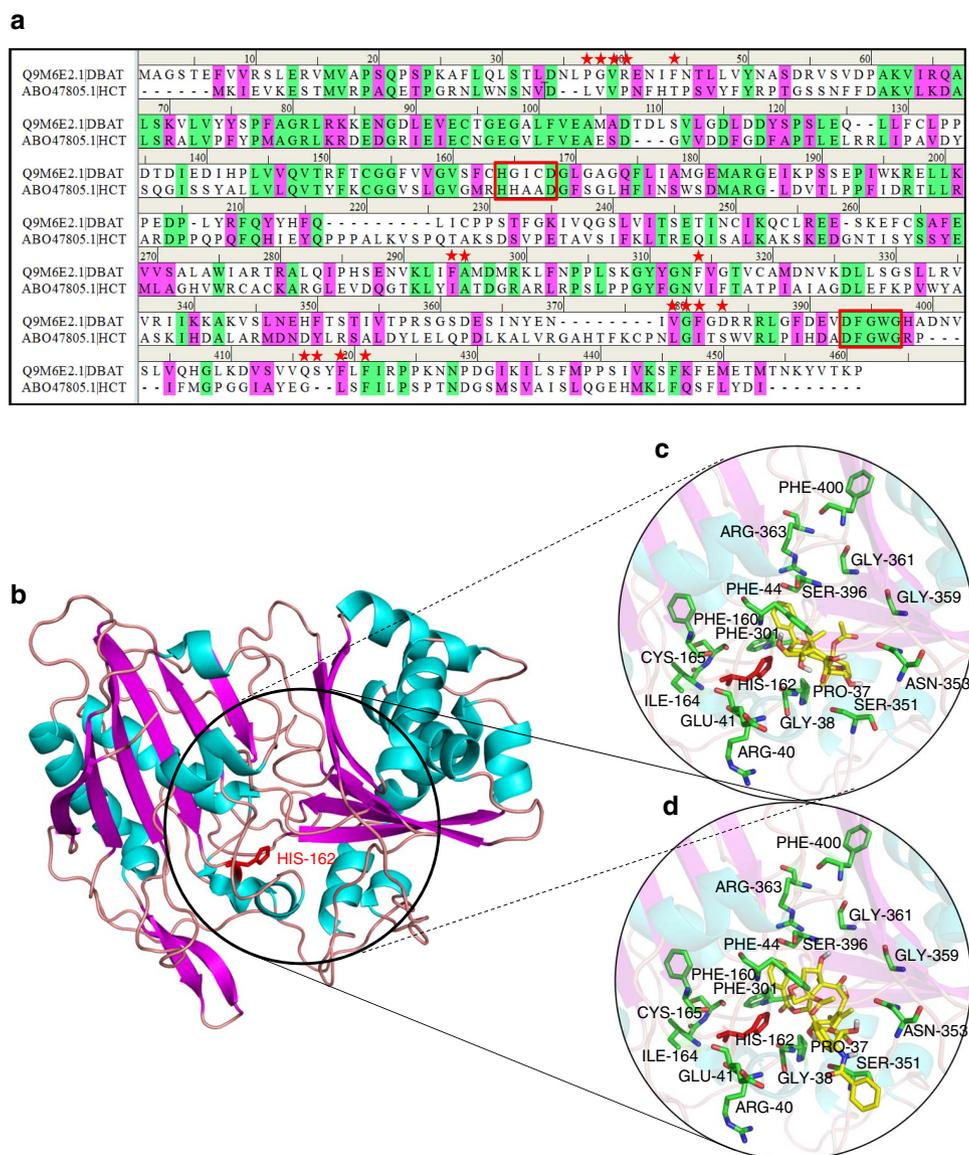
**Selection of LXYL-P1-2 concentration in reaction system.** The optimum temperature and optimum pH of LXYL-P1-2 were 45 °C and 4.5, respectively<sup>28</sup>. Since catalytic efficiency of LXYL-P1-2 against XDT was much higher than that of DBAT<sup>G38R/F301V</sup> against DT ( $9.7 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  versus  $3.5 \text{ M}^{-1} \text{ s}^{-1}$ ), the following one-pot reaction used optimum conditions of DBAT<sup>G38R/F301V</sup>, that is, 37.5 °C and pH 5.5, to maximize Taxol production. The final concentration of XDT assigned was 2 mM, and the influence of LXYL-P1-2 concentration on DT yield was detected (Fig. 7b). DT yields reached 830, 987, 1,054 and 1,093 μg ml<sup>-1</sup> at 12 h, when LXYL-P1-2 concentrations were 0.25, 0.5, 1.0 and 1.5 mg ml<sup>-1</sup>, respectively. All of the yields met the requirement for the next reaction from DT to Taxol and the relatively lower concentration of 0.5 mg ml<sup>-1</sup> LXYL-P1-2 was chosen in the following one-pot reaction system.

**Construction of one-pot reaction system.** In the 1-ml one-pot reaction system, the two enzymes and the substrate XDT were added and the reaction was performed under the aforementioned conditions. Taxol yields reached 657 and 349 μg ml<sup>-1</sup>, respectively, after 15 h with or without the supplementing strategy (Fig. 7e,f).

The one-pot reaction volume was amplified from 1 ml to 50 ml with the supplementing strategy and similar yields of Taxol were obtained, in which the Taxol yield reached 635 μg ml<sup>-1</sup> in the 50-ml reaction volume at 15 h (Fig. 7g, Supplementary Table 7). The product from the entire 50-ml reaction was extracted through ethyl acetate. Once concentrated, the product was purified by HPLC (Fig. 7h) for structure elucidation and was confirmed to be Taxol by MS, <sup>1</sup>H-NMR and <sup>13</sup>C-NMR analysis (Supplementary Fig. 1, Supplementary Table 3).

#### Discussion

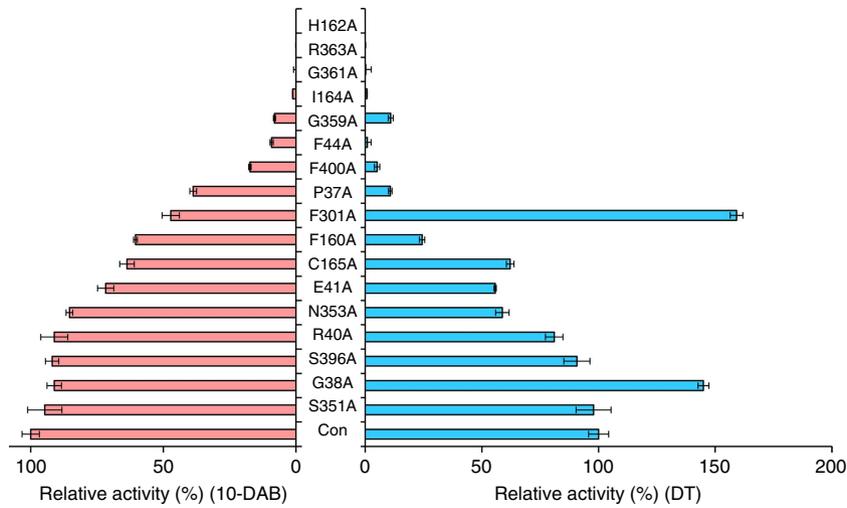
Taxol is a highly-oxygenated natural product with eleven chiral carbons. This diterpenoid compound harbours a main skeleton of A-, B- and C-rings, an oxetane ring (D-ring), the



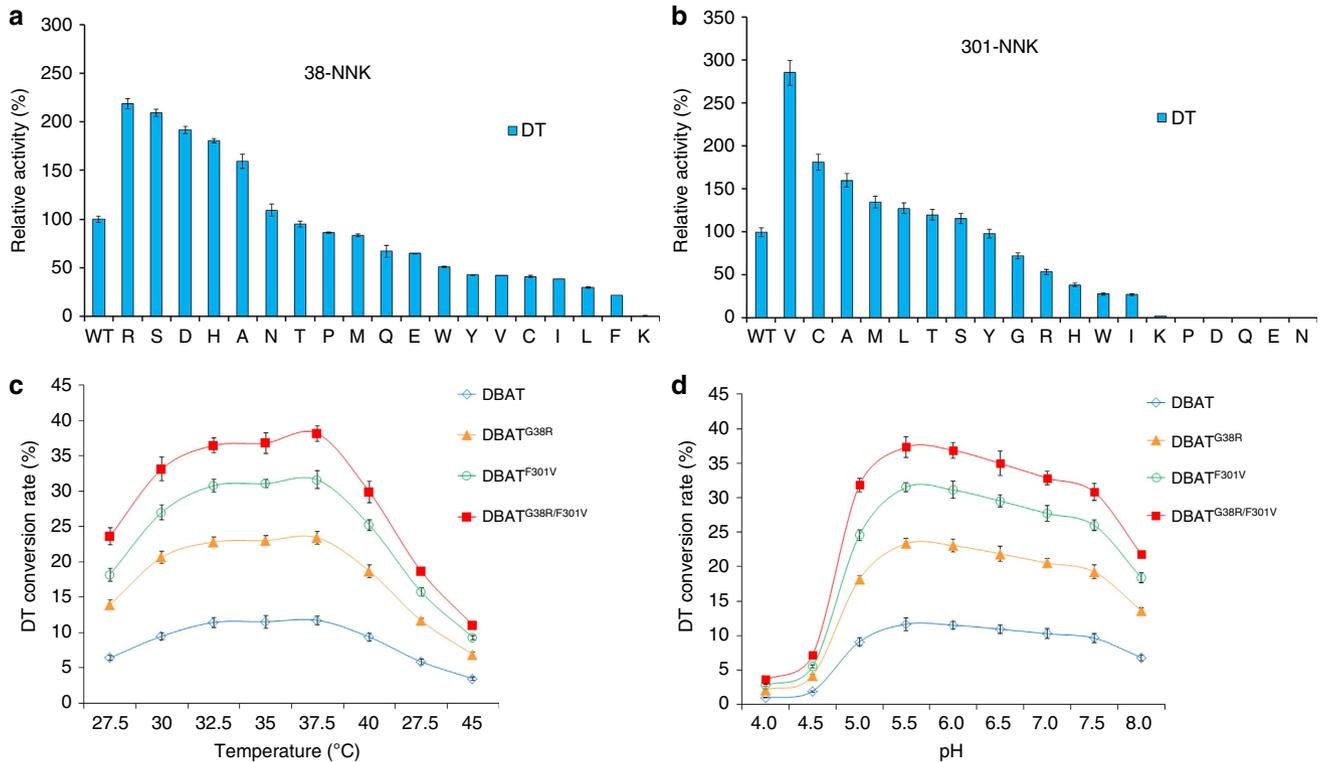
**Figure 4 | The DBAT structure predicted by homology modeling (template: HCT) and molecular docking of DBAT with the substrates. (a)** Sequence alignment between DBAT and the template HCT (ABO47805.1). Identical and similar residues are highlighted in green and pink, respectively. Residues predicted to be involved in acyl acceptor substrate binding are highlighted with red stars, and the conserved HXXXD and DFGWG motifs are marked in red box. **(b)** Homology model of DBAT constructed with Swiss-Model (the distance 5 Å around the substrate position was circled). **(c)** Enlarged view of molecular docking of DBAT with 10-DAB. **(d)** Enlarged view of molecular docking of DBAT with DT.

N-benzoylphenylisoserine side-chain appended to C13 of the A-ring, and the benzoyl group at C2 of the B-ring. Other functional groups include two hydroxyl groups (at C1 and C7) and two acetyl groups (at C4 and C10). It has been postulated that 20 enzymes are involved in 19 steps of the biosynthetic pathway from the universal diterpenoid precursor geranylgeranyl diphosphate (GGPP) to Taxol, of which six corresponding genes remains to be identified, i.e., C1 $\beta$ -hydroxylase, C4,C20-epoxydase, C4b,C20-oxomutase, C9 $\alpha$ -hydroxylase, C9-oxidase and C2' side-chain hydroxylase<sup>45–47</sup>. The first committed step is the cyclization of GGPP to taxa-4(5), 11(12)-diene. This olefin is then structurally modified by eight cytochrome P450-mediated oxygenations, two subsequent acetylations, a benzoylation, oxetane ring formation and oxidation at C9 en route to the late intermediate baccatin III, to which the C13 side-chain is attached to yield the final product Taxol<sup>45–49</sup>. In comparison with Taxol, 7- $\beta$ -xylosyl-10-deacetylaxol (XDT) possesses a  $\beta$ -xylosyl group at the C7 position but lacks an acetyl group at the C10 position.

XDT and Taxol most likely share the same precursor 10-DAB, but the former is synthesized by an unknown branch pathway most likely through the intermediate DT (the compound is also frequently isolated from *Taxus* species in the similar amount of Taxol<sup>37</sup>). The abundant production of XDT is likely a kind of detoxifying mechanism to the plant host, since as the mitotic inhibitor, the *in vitro* cytotoxicity of Taxol was much higher than that of XDT (Supplementary Table 8). Baccatin III is one of the key intermediates in the biosynthetic pathway of Taxol, which is produced by the C10 hydroxyl acetylation of 10-DAB catalysed by DBAT. This enzyme has been systematically investigated by other groups<sup>33–36</sup>. The present study revealed that the DBATs of different *Taxus* species exhibited various activities on 10-DAB. Among the six studied DBATs, higher and similar activities against 10-DAB were found in *T. cuspidata*, *T. brevifolia*, *T. x media* and *T. wallichiana* var. origins, but the activities of *T. canadensis* and *T. baccata* origins were obviously lower than those of the four species, with the activity of *T. baccata* origin



**Figure 5 | Relative activities of L-alanine scanning mutations of DBAT against 10-DAB (left) and DT (right).** Con: DBAT. The data represent the means  $\pm$  s.d.,  $n = 3$ .



**Figure 6 | Activities of Gly<sup>38</sup> and Phe<sup>301</sup> saturation mutagenesis on DT and temperature- and pH- course profiles of the mutants.** (a) Relative activities of Gly<sup>38</sup> saturation mutations on DT. (b) Relative activities of Phe<sup>301</sup> saturation mutations on DT. (c) Temperature-course profiles of the DBAT mutants. (d) pH-course profiles of the DBAT mutants. The data represent the means  $\pm$  s.d.,  $n = 3$ .

**Table 1 | Specific activities of the DBAT mutants against DT and 10-DAB.**

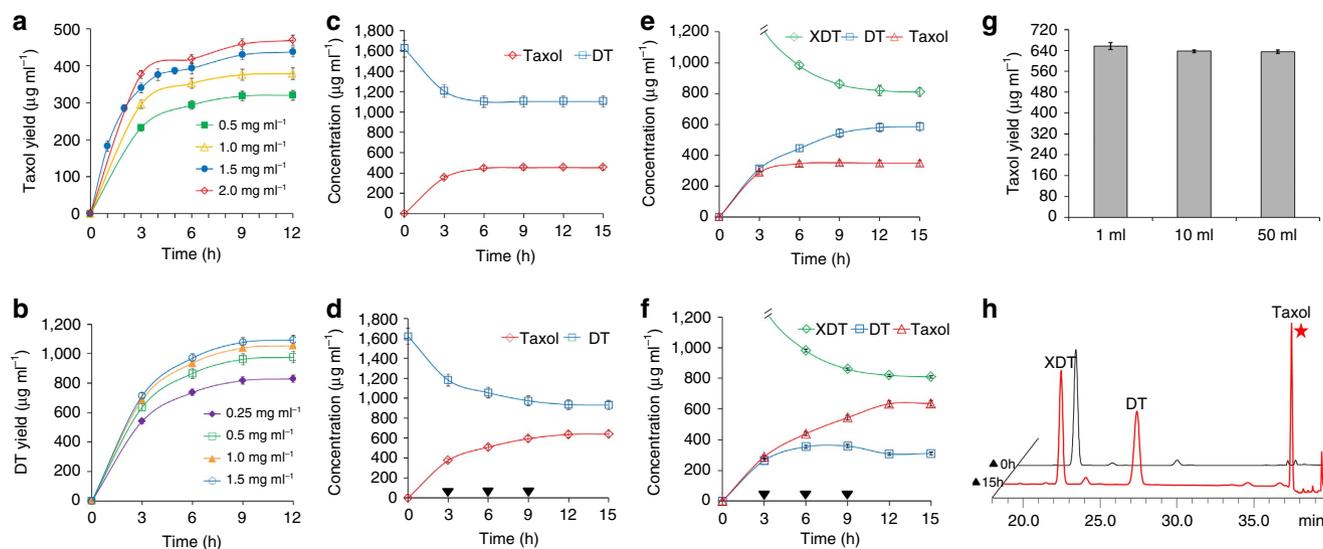
	DT		10-DAB	
	U mg <sup>-1</sup> ( $\times 10^{-2}$ )	Relative activity (%)	U mg <sup>-1</sup> ( $\times 10$ )	Relative activity (%)
DBAT	26.00 ( $\pm 1.12$ )	100.00	20.66 ( $\pm 0.67$ )	100.00
DBAT <sup>G38R</sup>	56.88 ( $\pm 3.11$ )**	218.77	44.34 ( $\pm 3.76$ )**	214.62
DBAT <sup>F301V</sup>	74.10 ( $\pm 8.45$ )**	285.00	10.87 ( $\pm 0.23$ )	52.60
DBAT <sup>G38R/F301V</sup>	96.21 ( $\pm 5.13$ )**	370.03	14.32 ( $\pm 0.71$ )	69.31

The data represent the means  $\pm$  s.d.,  $n = 3$ . \* $P < 0.05$  versus DBAT (Control), \*\* $P < 0.01$  versus DBAT (Control) (Student's *t*-test).

**Table 2 | Kinetic parameters of the DBAT mutants against DT.**

	DBAT	DBAT <sup>G38R</sup>	DBAT <sup>F301V</sup>	DBAT <sup>G38R/F301V</sup>
$V_{max}$ ( $\mu\text{M min}^{-1}$ )	0.204 ( $\pm 0.004$ )	0.499 ( $\pm 0.029$ )***	0.242 ( $\pm 0.008$ )**	0.918 ( $\pm 0.096$ )***
$K_m$ (mM)	0.546 ( $\pm 0.029$ )	0.368 ( $\pm 0.061$ )*	0.178 ( $\pm 0.017$ )***	0.424 ( $\pm 0.133$ )
$k_{cat}$ ( $\text{s}^{-1}$ ) ( $\times 10^{-3}$ )	0.334 ( $\pm 0.001$ )	0.815 ( $\pm 0.005$ )***	0.396 ( $\pm 0.013$ )**	1.499 ( $\pm 0.157$ )***
$k_{cat}/K_m$ ( $\text{M}^{-1} \text{s}^{-1}$ )	0.611 ( $\pm 0.243$ )	2.218 ( $\pm 0.787$ )*	2.22 ( $\pm 0.771$ )*	3.539 ( $\pm 0.118$ )***

The data represent the means  $\pm$  s.d.,  $n = 3$ . \* $P < 0.05$  versus DBAT (Control), \*\* $P < 0.01$  versus DBAT (Control), \*\*\* $P < 0.001$  versus DBAT (Control) (Student's  $t$ -test).



**Figure 7 | Time-concentration curves of XDT and DT as well as Taxol in the reactions and the Taxol yields in the one-pot reaction system.** (a) Time-concentration curves of Taxol with DT as the substrate and catalysed by different concentrations of DBAT<sup>G38R/F301V</sup>. (b) Time-concentration curves of DT with XDT as the substrate and catalysed by different concentrations of LXYL-P1-2. (c) Time-concentration curves of DT and Taxol catalysed by 1.5 mg ml<sup>-1</sup> of DBAT<sup>G38R/F301V</sup> without DBAT<sup>G38R/F301V</sup> supplement. (d) Time-concentration curves of DT and Taxol catalysed by 1.5 mg ml<sup>-1</sup> of DBAT<sup>G38R/F301V</sup> with DBAT<sup>G38R/F301V</sup> supplement (downward black triangle, adding time). (e) Time-concentration curves of XDT, DT and Taxol in the one-pot reaction system (0.5 mg ml<sup>-1</sup> of LXYL-P1-2 plus 1.5 mg ml<sup>-1</sup> of DBAT<sup>G38R/F301V</sup>) without DBAT<sup>G38R/F301V</sup> supplement. (f) Time-concentration curves of XDT, DT and Taxol in the one-pot reaction system with DBAT<sup>G38R/F301V</sup> supplement (downward black triangle, adding time). (g) Scaling-up of the one-pot reaction. (h) HPLC analysis of the one-pot reaction at 15 h (reaction volume: 50 ml). The data represent the means  $\pm$  s.d.,  $n = 3$ .

being the lowest (Fig. 2, Supplementary Table 2). The decreased activity of *T. baccata* origin against 10-DAB may at least partially explain why this species frequently accumulates significantly more 10-DAB in its twigs than other species<sup>37,50</sup>, since apart from the variation of the expression levels, such a low activity may hamper the enzymatic step from 10-DAB to baccatin III, leading to the abundant accumulation of 10-DAB in the plant.

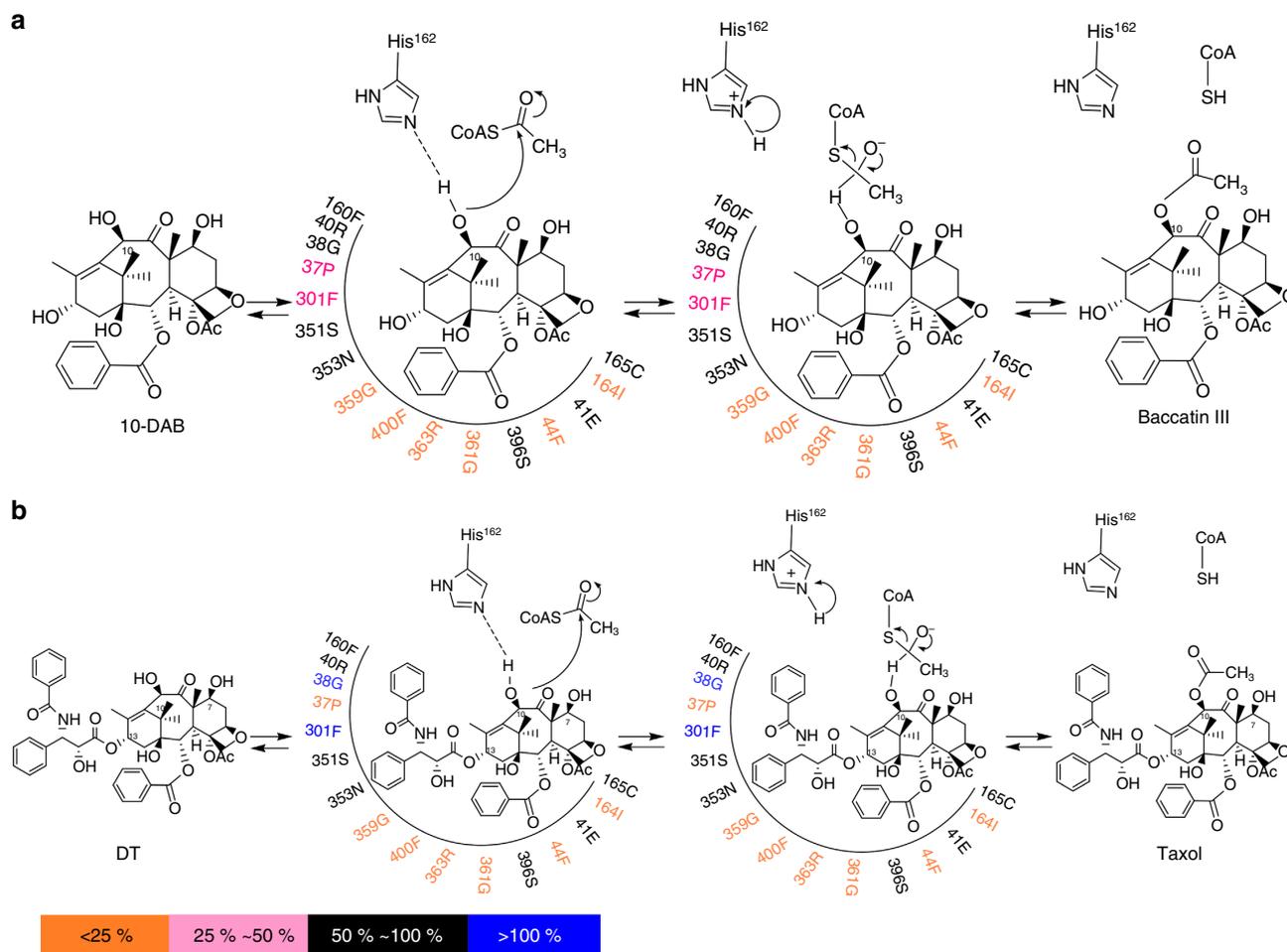
To the unnatural substrate DT, the enzyme activity of each species was only about 0.1% of the same enzyme against the natural substrate 10-DAB, implying that the larger C13 side-chain of DT sharply decreased the fitness of the enzyme to the unnatural substrate. This finding also suggests that under natural conditions acetylation at the C10 position is less likely to occur after attachment of the phenylpropanoyl side chain to 10-DAB on the biosynthetic pathway.

To explore the relationship between the structure and function of DBAT, especially to improve its catalytic efficiency against DT, this study modelled a 3D structure of DBAT on the structure of *C. canephora* HCT (GenBank accession: ABO47805.1), as the template (30% sequence identity and 45% sequence similarity with DBAT) and obtained a reasonable model with 87.5% of the structure in the most favoured regions. The model indicated that the enzyme is composed of two nearly equal domains connected through a loop (Fig. 4b). The highly conserved HXXXD and DFGWG motifs help position the active centre and acetyl

acceptor pocket of DBAT at the back side, in which His<sup>162</sup> probably acts as a catalytic site. After molecular docking, the points within 5 Å distance from DT were selected as 'hot spots' for individual mutations of the residues to L-alanine (L-alanine scanning) to identify putative residues participating in the binding and/or catalysis of 10-DAB or DT (Fig. 4c,d). The strategy of the 'hot spots' selection for semi-rational mutagenesis has been demonstrated to be helpful in determining the active sites of the enzyme and increasing likelihood of finding improved enzymes in protein engineering<sup>44,51,52</sup>.

Results confirmed the residue His<sup>162</sup> as the catalytic base, as the H162A mutation almost fully lost enzyme activity. Similar phenomena were found in R363A, G361A, I164A, F44A, G359A and F400A mutations, implying that Arg<sup>363</sup>, Gly<sup>361</sup>, Ile<sup>164</sup>, Phe<sup>44</sup>, Gly<sup>359</sup> and Phe<sup>400</sup> are also potential substrate binding or catalytic sites (Fig. 5). Furthermore, Phe<sup>44</sup>, Phe<sup>400</sup> and Pro<sup>37</sup> are more crucial to DT than to 10-DAB in enzymatic catalysis. The impact of the amino acid residues around the substrate within the 5 Å region on the acetylation of 10-DAB or DT is respectively summarized in Fig. 8, which exhibits the extent of the activity maintenance after L-alanine scanning mutagenesis.

The residues Gly<sup>38</sup> and Phe<sup>301</sup> may not have participated in the DT binding, and the G38A and F301A mutations even showed significant increase in activities against DT (Fig. 5), which led to performing site-specific saturation mutations to identify more



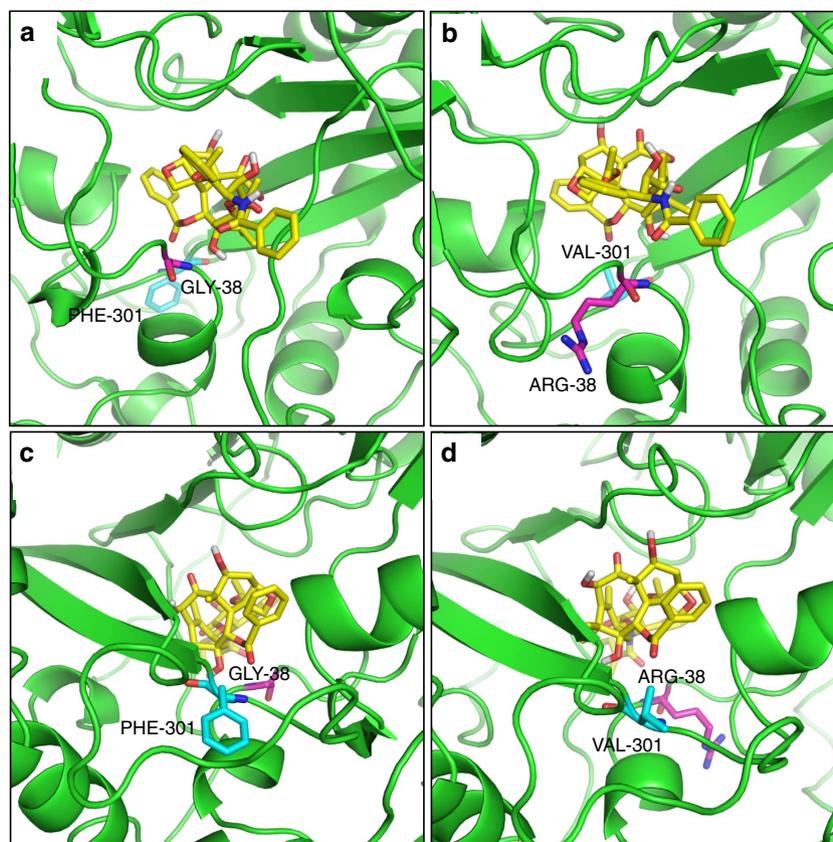
**Figure 8 | The putative catalytic processing of 10-DAB or DT by DBAT supported by DBAT activity assays after site-specific mutagenesis.** In the catalytic process, His<sup>162</sup> of DBAT acts as a general base catalyst and deprotonates the substrate 10-DAB or DT. **(a)** For 10-DAB, Gly<sup>359</sup>, Phe<sup>400</sup>, Arg<sup>363</sup>, Gly<sup>361</sup>, Phe<sup>44</sup> and Ile<sup>164</sup> of DBAT seem to form the substrate binding and/or catalytic sites, as supported by the significant decrease in the activity of the enzyme when these residues are mutated to L-alanine. **(b)** For DT, Pro<sup>37</sup>, Gly<sup>359</sup>, Phe<sup>400</sup>, Arg<sup>363</sup>, Gly<sup>361</sup>, Phe<sup>44</sup> and Ile<sup>164</sup> of DBAT seem to form the substrate binding and/or catalytic sites, due to the significant decrease in the activity after L-alanine scanning mutagenesis. Gly<sup>38</sup> and Phe<sup>301</sup> do not seem to be involved in substrate binding or be part of the catalytic site, and the G38A and F301A mutations increase the activity of DBAT towards DT.

active mutants; thus, the two mutants DBAT<sup>G38R</sup> and DBAT<sup>F301V</sup> were obtained. The two mutants respectively showed 2.19 and 2.85 times more activity than the wild-type on DT (Table 1). Furthermore, a synergistic effect between the two mutations (G38R and F301V) was observed in the double mutant DBAT<sup>G38R/F301V</sup>, which showed approximately six times the catalytic efficiency on DT compared with DBAT (Table 2).

The double mutant DBAT<sup>G38R/F301V</sup> was subjected to molecular docking with DT and the virtual complex was compared with DBAT-DT virtual complex in which the amino acid residues of 38 and 301 sites were highlighted (Fig. 9). As the side-chain of Arg is more hydrophilic than that of Gly, and the 38 site is near the surface of DBAT (Fig. 9a,b), the increased hydrophilic property of G38R mutation is likely one of the reasons that led to the elevated catalytic efficiency ( $k_{cat}/K_m$ ) (Table 2). The G38R mutation also positively influenced the enzyme activity on 10-DAB (Table 1), which suggests that the mutant favours the entry of both substrates. The site Phe<sup>301</sup> was in the interior of the substrate solvent channel of DBAT and its phenyl side-chain may hamper the entry of DT deeper into the binding pocket. In contrast, Val not only maintains the hydrophobicity but is smaller in volume (98.71 Å<sup>3</sup>) than that of Phe (131.1 Å<sup>3</sup>). Thus the F301V mutation may reduce the hindrance of the enzyme against DT (Fig. 9c,d), which likely

explains why this mutation led to increased catalytic efficiencies for DT (Table 2). In addition, the activity of F301V mutation towards its natural substrate 10-DAB was nearly half that of the wide type DBAT (Table 1). Altogether, these results indicate a combined favourable effect of a more opened pocket and a bulkier substrate such as DT, while 10-DAB could go too far due to its smaller size. A similar phenomenon of declined activity towards its natural substrate 10-DAB was also observed in the F301A mutation (Fig. 5). These results suggested that the bulkier Phe<sup>301</sup> may play a role for switching the acyl acceptor substrate specificity of DBAT. Since the possible conformational changes of the amino acid side-chains induced by the substrate and the interactions between the substrate and the related amino acid side-chains cannot be more exactly characterized by the molecular docking and the virtual 3D structure analysis, a future study on the enzyme-substrate complex crystallization and the 3D structure analysis is required, which may decipher the detailed catalytic mechanism of the enzyme on the substrate.

Fully utilizing the abundant analogue XDT for production of Taxol not only reduces resource waste and potential environmental pollution, but also alleviates the shortage for the clinical supply of the drug. Conversion of XDT through DT to Taxol by the enzymatic approach is more environmentally friendly than the chemical approach. The enzymatic deglycosylation of XDT to DT has been



**Figure 9 | Molecular docking of DBAT and DBAT<sup>G38R/F301V</sup> with DT.** DT carbon atoms are coloured in yellow and oxygen atoms in red. The side chains of Gly<sup>38</sup>/Arg<sup>38</sup> and Phe<sup>301</sup>/Val<sup>301</sup> are also shown. Gly<sup>38</sup>/Arg<sup>38</sup> carbon atoms are coloured in magenta and oxygen atoms in red. Phe<sup>301</sup>/Val<sup>301</sup> carbon atoms are coloured in cyan and nitrogen atoms in blue. **(a)** Side view of DBAT with DT, showing Gly<sup>38</sup>; **(b)** Side view of DBAT<sup>G38R/F301V</sup> with DT, showing Arg<sup>38</sup>. The increased hydrophilic properties of Arg<sup>38</sup> compared to that of Gly<sup>38</sup> favors the entry of both DT and 10-DAB substrates and leads to an elevated catalytic efficiency in the G38R mutant. **(c)** Opposite side view of DBAT with DT, showing Phe<sup>301</sup>. **(d)** Opposite side view of DBAT<sup>G38R/F301V</sup> with DT, Showing Val<sup>301</sup>. Val<sup>301</sup> maintains the hydrophobicity but has a smaller volume (98.71 Å<sup>3</sup>) than that of Phe (131.1 Å<sup>3</sup>), which leads to an increased catalytic efficiency in the F301V mutant due to the reduced hindrance.

achieved by several groups<sup>23–29</sup>, in which the specific  $\beta$ -xylosidase (designated as LXYL-P1-2) of the fungus *L. edodes* origin was well characterized<sup>28,29</sup>. The enzymatic acetylation of DT to Taxol can be realized by using the 10- $\beta$ -acetyltransferase double mutant DBAT<sup>G38R/F301V</sup>, which showed the elevated activity against DT.

In the present study, the *in vitro* one-pot reaction system of de-glycosylation/acetylation was constructed with the specific  $\beta$ -xylosidase LXYL-P1-2 and the 10- $\beta$ -acetyltransferase double mutant DBAT<sup>G38R/F301V</sup>, using XDT as the substrate for the synthesis of Taxol. As catalytic efficiency of LXYL-P1-2 on XDT was apparently higher than on 7-xylosyltaxol, likewise, catalytic efficiency of DBAT<sup>G38R/F301V</sup> on DT was much higher than on XDT, thus, the reaction order should be XDT through DT to Taxol in the reaction system (Fig. 1). Additionally, since the catalytic efficiency of LXYL-P1-2 against XDT was obviously higher than that of DBAT<sup>G38R/F301V</sup> against DT, optimum conditions of DBAT<sup>G38R/F301V</sup> and DBAT<sup>G38R/F301V</sup> supplementing strategy were used in the one-pot reaction (2 mM of XDT, 0.5 mg ml<sup>-1</sup> of LXYL-P1-2, 1.5 mg ml<sup>-1</sup> of DBAT<sup>G38R/F301V</sup> at 37.5 °C and pH 5.5; additional 1.5 mg ml<sup>-1</sup> of DBAT<sup>G38R/F301V</sup> was repeatedly supplemented in 3-h intervals) to maximize the production of Taxol. The Taxol yield reached 657  $\mu$ g ml<sup>-1</sup> at 15 h in the 1-ml reaction system (Fig. 7f), which was comparable to Taxol production (641  $\mu$ g ml<sup>-1</sup>) in the single enzyme (DBAT<sup>G38R/F301V</sup>) reaction system that applied DT as the substrate (Fig. 7d). The two-enzyme coupled one-pot reaction was scaled up from 1 ml to 50 ml, and similar yields were

obtained, of which 0.64 mg ml<sup>-1</sup> Taxol was produced in 50-ml reaction (Fig. 7g, Supplementary Table 7).

As clinical demand increases, the long-term supply of Taxol will continue to be an important issue. This study may provide another choice for Taxol production from the ample analogue XDT. As activity of DBAT<sup>G38R/F301V</sup> against DT is not high enough, future work will be focused on the improvement of DBAT by protein engineering to elevate enzyme activity and/or stability. The supply of acetyl CoA to the reaction system is another concern, as it is not a bulk commodity and has become a limiting factor for the large-scale reaction. The practical method for massively producing acetyl CoA is probably the genetic engineering method. The *in vivo* synthetic biology (constructing biosynthetic pathway from XDT to Taxol in a microbial chassis cell) is a promising strategy to produce Taxol from the substrate XDT. In the long term, the *in vivo* and *de novo* synthetic biology for Taxol production through a microbial cell factory should be a favourable approach that does not depend on plant-based processes<sup>14,53</sup>. The latter is based on a thorough understanding of the complex biosynthetic pathway and regulation of Taxol biosynthesis. Considering the complicated structure and certain undefined steps, obtaining the final product Taxol through such strategy is still confronted with many challenges.

## Methods

**Materials and strains.** The *dbat* gene of *Taxus cuspidata* (GenBank accession: Q9M6E2.1) was synthesized by ViewSolid Biotech Crop (Beijing, China). Other

*dbat* genes of *Taxcus brevifolia* (GenBank accession: EU107143.1), *T. baccata* (GenBank accession: AF456342.1), *T. canadensis* (GenBank accession: EU107134.1), *T. wallichiana* var (GenBank accession: EU107140.1), and *T. x media* (GenBank accession: AY452666.1) were synthesized by SynBio Research Platform at Tianjin University (Tianjin, China). Phusion polymerase, restriction enzymes, and T4 ligase were purchased from New England Biolabs (Ipswich, MA). *Escherichia coli* JM109 competent cells and plasmids were purchased from TransGen Biotech (Beijing, China). Taxol was kindly provided by Beijing Union Pharmaceutical Factory, 10-DAB and baccatin III were purchased from J&K Scientific Ltd (Beijing, China). XDT and DT (HPLC purity > 98%) were purified by our laboratory. Acetyl-coenzyme A was purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals were of analytical grade unless otherwise indicated.

**Recombinant protein expression and purification.** The full-length *dbat* genes were amplified using a forward primer (F: 5'-CTTAGGAGTCA-TATGTCATCATCATCATCATG-3') with *Nde* I site (underlined) and His-tag coding sequence and a reverse primer (R: GCAGTCCGACTCTA-GACTAAGGTTTGTAGTTAC) with *Xba* I site (underlined). Each of the PCR products was digested by *Nde* I and *Xba* I for directional ligation into vector pCWoRI. After ligation, the construct was transformed into *E. coli* JM109 competent cells. *E. coli* cells harbouring the recombinant plasmids were grown overnight at 37 °C and 200 r.p.m. in 10 ml Luria-Bertani (LB) medium containing ampicillin (100 µg ml<sup>-1</sup>) in a shake flask. The overnight cultures were suspended in 100 ml fresh Terrific Broth (TB) medium [1.2% (w/v) Tryptone, 2.4% (w/v) Yeast Extract, 0.4% (v/v) Glycerol, 17 mM KH<sub>2</sub>PO<sub>4</sub>, 72 mM K<sub>2</sub>HPO<sub>4</sub>] at a final concentration of 1% (v/v), and grown at 37 °C and 200 r.p.m. for 2–3 h until OD<sub>600</sub> reached 0.8. Then, isopropyl-β-D-thiogalactopyranoside (IPTG) was added at a final concentration of 1 mM and the cell cultures were incubated for an additional 18 h at 20 °C and 200 r.p.m.. Ampicillin at 100 µg ml<sup>-1</sup> was added to the medium when required.

The recombinant DBAT was partially purified by affinity chromatography in Ni-NTA agarose (General Electric Company, USA) according to the manufacturer's instructions. Briefly, the induced cells were resuspended in the sample buffer (20 mM imidazole, 100 mM NaCl in 20 mM Tris-HCl buffer, pH 7.5) with a concentration of 15% (v/v). After high pressure crushing (800 bar, 3 times), the sample was centrifuged with 13,400 g at 4 °C for 30 min and the supernatant was filtered through a 0.45 µm membrane for the further purification. The Ni-NTA agarose was equilibrated by the equilibration buffer (same as sample buffer). The elution buffer contained the same components in addition to 200 mM imidazole. The partially purified sample was concentrated via ultrafiltration (30 kDa, Millipore, Billerica, MA). DBAT was further purified by molecular size chromatography (Agilent Zorbax Bio Series GF-250). The equilibration buffer contained 0.1 M NaCl, 0.1 M K<sub>2</sub>HPO<sub>4</sub>, and 0.1 M KH<sub>2</sub>SO<sub>4</sub>, pH 7.5. Proteins were eluted with the same buffer at a flow rate of 0.2 ml min<sup>-1</sup>. The purified protein was detected by SDS-PAGE gels (12%) followed by Coomassie Blue or silver staining. The concentration of protein was determined with Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

**Characterization and kinetics of the recombinant DBATs.** Determination of the optimal temperature on 10-DAB: A total reaction volume of 100 µl contained 500 µM 10-DAB, 500 µM acetyl-CoA, and 0.02 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer, pH 5.5. The reaction was performed under the following temperatures: 27.5, 30, 32.5, 35, 37.5, 40, 42.5, and 45 °C for 20 min.

Determination of the optimal temperature on DT: A total reaction volume of 100 µl contained 500 µM DT, 500 µM acetyl-CoA, and 0.5 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer, pH 5.5. The reaction was performed under the following temperatures: 27.5, 30, 32.5, 35, 37.5, 40, 42.5, and 45 °C for 180 min.

Determination of the optimal pH on 10-DAB: A total reaction volume of 100 µl contained 500 µM 10-DAB, 500 µM acetyl-CoA, and 0.02 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer with pH 4.0, 4.5, 5.0, 5.5 and 6.0; or in 50 mM sodium phosphate buffer with pH 6.5, 7.0, 7.5 and 8.0. The reaction was performed under 40 °C for 20 min.

Determination of the optimal pH on DT: A total reaction volume of 100 µl contained 500 µM DT, 500 µM acetyl-CoA, and 0.5 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer with pH 4.0, 4.5, 5.0, 5.5, and 6.0; or in 50 mM sodium phosphate buffer with pH 6.5, 7.0, 7.5, and 8.0. The reaction was performed under 40 °C for 180 min.

Reactions were terminated by adding 500 µl methanol and analysed by HPLC [Cosmosil-C18, flow rate: 1 ml min<sup>-1</sup>, column temperature: 28 °C, UV detection wavelength: 230 nm, mobile phase: acetonitrile (A)–water (B). Gradient elution conditions see Supplementary Tables 9 and 10]. The conversion rate was calculated as the product peak area/(substrate peak area + product peak area) × 100%.

Enzymatic activity analysis: The enzymatic activities of DBATs were determined by measuring the amount of baccatin III (substrate 10-DAB) or Taxol (substrate DT) under the optimal conditions. To measure baccatin III production, 100 µl reaction volume contained 500 µM 10-DAB, 500 µM acetyl-coenzyme A, and 0.02 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer, pH 5.5. The reaction was performed under 40 °C for 20 min. To measure Taxol production, 100 µl

reaction volume contained 500 µM DT, 500 µM acetyl-coenzyme A, and 0.5 mg ml<sup>-1</sup> DBAT in 50 mM sodium acetate buffer, pH 5.5. The reaction was performed under 37.5 °C for 180 min. Reactions were terminated as described above. One unit of activity was defined as the amount of enzyme that catalysed the formation of 1 nM baccatin III or Taxol per minute. The kinetic data on DT were processed via a proportional weighted fit using a nonlinear regression analysis program based on Michaelis–Menten enzyme kinetics<sup>28,54</sup>. Briefly, the kinetic parameters were determined against DT in a concentration range of 15.625–2,000 µM, maintaining acetyl-CoA at a saturation concentration of 2,000 µM. Reactions were conducted under the aforementioned conditions. To set up the standard curves and linear regression equations, different concentrations of baccatin III, DT, and Taxol were prepared respectively: 2, 1, 0.2, 0.1, 0.04, and 0.008 (mg ml<sup>-1</sup>). The HPLC conditions were as described above. The linear relationship between the concentration (mg ml<sup>-1</sup>) and peak area of baccatin III, DT and Taxol was respectively analysed and each equation of the best-fit lines ( $R^2 > 0.99$ ) was determined. The product yield was calculated according to the standard curves (Supplementary Table 11). Values were expressed as the means ± s.d. ( $n = 3$ ). *P* values were calculated by using a two-tailed Student's *t*-test.  $P < 0.05$  was considered statistically significant.

**Homology modelling of DBAT three-dimensional structure.** The protein sequences of DBATs from *Taxus* species were retrieved from the NCBI Protein Sequence Database (<http://www.ncbi.nlm.nih.gov/protein>). Protein Structure Database (Protein Data Bank, PDB) (<http://www.rcsb.org/pdb/>) provided the template structure of hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase [*Coffea canephora*] (GenBank accession: ABO47805.1, PDB ID: 4G22), and BAHD family proteins with known structures: hydroxycinnamoyltransferase [*Sorghum bicolor*] (PDB ID: 4KE4), hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase [*Panicum virgatum*] (PDB ID: 5FAL), vinorine synthase [*Rauvolfia serpentina*] (GenBank accession: CAD89104.2, PDB ID: 2BGH), and anthocyanin malonyltransferase homologue [*Chrysanthemum x morifolium*] (GenBank accession: BAF50706.1, PDB ID: 2E1U). The three-dimensional (3D) model of DBAT protein was built by using SWISSMODEL (<http://swissmodel.expasy.org/>)<sup>55</sup>. To evaluate the accuracy of the model, PROCHECK was employed<sup>42</sup>.

**L-alanine scanning and saturation mutagenesis.** The primers with the single site mutation are listed in Supplementary Table 12. The recombinant plasmid pCWoRI-*dbat* was employed as a template. The L-alanine scanning mutations and saturation mutations were all conducted by using whole-plasmid amplification PCR<sup>56</sup>. PCR was performed with Phusion DNA polymerase, and the temperature program consisted of 98 °C for 30 s; 25 cycles of 10 s at 98 °C, 30 s at 55 °C, 3 min 20 s at 72 °C, and a final 10 min extension at 72 °C. The PCR products were purified and digested by *Dpn* I at 37 °C for 4 h and then transformed into JM109 competent cells. After the sequences were confirmed, the corresponding clones were used for the enzymatic activity analysis as mentioned above.

**One-pot reaction system.** To select the proper concentration of the DBAT mutant in the reaction system, 0.5, 1.0, 1.5 and 2.0 mg ml<sup>-1</sup> enzyme solutions, each with 2 mM DT in 5% DMSO (v/v) as the substrate and 2 mM acetyl-CoA as the acetyl donor, in 500 µl reaction volume were respectively tested under the optimum conditions of 37.5 °C and pH 5.5. To compensate for the enzyme activity lost during the reaction, the DBAT mutant supplementing strategy was used at interval of 3 h. To select the proper concentration of the β-xylosidase LXYL-P1-2 in the reaction system, 0.25, 0.5, 1.0 and 1.5 mg ml<sup>-1</sup> enzyme solutions, each with 2 mM XDT in 5% DMSO (v/v) as the substrate, in 500 µl reaction volume were respectively tested under the optimum conditions of the DBAT mutant. The XDT conversion rate was calculated according to the literature<sup>28</sup>. The reactions were conducted for as long as 15 h. An aliquot of 50 µl sample solution was generally taken at 3, 6, 9 and 12 h respectively for HPLC analysis.

Finally, the ideal concentrations of the DBAT mutant and LXYL-P1-2 were combined to construct the one-pot reaction system with 2 mM XDT as the substrate and 2 mM acetyl-CoA as the acetyl donor. The reaction volume was scaled up from 1 ml through 10 to 50 ml.

**Data availability.** The homology model for DBAT was deposited to the Swiss-Model with accession code Q9M6E2. All other data that support the findings of this study are available within the article and its supplementary files and from the corresponding author upon reasonable request.

## References

- Liu, W. C., Gong, T. & Zhu, P. Advances in exploring alternative Taxol sources. *RSC Adv.* **6**, 48800–48809 (2016).
- Horwitz, S. B. Taxol (paclitaxel): mechanisms of action. *Ann. Oncol.* **5**, S3–S6 (1993).
- Ehrlich, A. *et al.* Micellar paclitaxel improves severe psoriasis in a prospective phase II pilot study. *J. Am. Acad. Dermatol.* **50**, 533–540 (2004).

4. Brahn, E. Use of taxol in the treatment of rheumatoid arthritis. US Patent No. 5,583,153 (1996).
5. Bonaventura, K. *et al.* Cost-effectiveness of paclitaxel-coated balloon angioplasty and paclitaxel-eluting stent implantation for treatment of coronary in-stent restenosis in patients with stable coronary artery disease. *Clin. Res. Cardiol.* **101**, 573–584 (2012).
6. Heldman, A. W. *et al.* Paclitaxel stent coating inhibits neointimal hyperplasia at 4 weeks in a porcine model of coronary restenosis. *Circulation* **103**, 2289–2295 (2001).
7. Unverdorben, M. *et al.* Paclitaxel-coated balloon catheter versus paclitaxel-coated stent for the treatment of coronary in-stent restenosis: the three-year results of the PEPCAD II ISR study. *EuroIntervention* **11**, 926–934 (2015).
8. Scheller, B. *et al.* Long-term follow-up after treatment of coronary in-stent restenosis with a paclitaxel-coated balloon catheter. *JACC Cardiovasc. Interv.* **5**, 323–330 (2012).
9. Fang, W., Wu, Y., Zhou, J., Chen, W. & Fang, Q. Qualitative and quantitative determination of taxol and related compounds in *Taxus cuspidata* Sieb et Zucc. *Phytochem. Anal.* **4**, 115–119 (1993).
10. Hamilton, A. & Hamilton, P. in *People and Plants Conservation Series*. (ed. Walters, M.) 264 (2006).
11. Fett-Neto, A. G. *et al.* Cell culture of *Taxus* as a source of the antineoplastic drug taxol and related taxanes. *Biotechnology* **10**, 1572–1575 (1992).
12. Expósito, O. *et al.* Biotechnological production of taxol and related taxoids: current state and prospects. *Anti-Cancer Agents Med. Chem.* **9**, 109–121 (2009).
13. Nicolaou, K. C. *et al.* Total synthesis of taxol. 1. Retrosynthesis, degradation, and reconstitution. *J. Am. Chem. Soc.* **117**, 624–633 (1995).
14. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**, 70–74 (2010).
15. Malika, S. *et al.* Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: A review. *Process Biochem.* **46**, 23–34 (2011).
16. Gueritte-Voegeléin, F. *et al.* Chemical studies of 10-deacetyl baccatin III: hemisynthesis of taxol derivatives. *Tetrahedron* **42**, 4451–4460 (1986).
17. Croteau, R. *et al.* Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97 (2006).
18. Chattopadhyay, S. K. *et al.* A process for the production of taxol. EP Patent No. 0905130 B1 (2002).
19. Fu, Y. J. *et al.* Simultaneous determination of main taxoids in *Taxus* needles extracts by solid-phase extraction-high-performance liquid chromatography with pentafluorophenyl column. *Biomed. Chromatogr.* **23**, 63–70 (2009).
20. Luo, H. *et al.* Determination of main taxoids in *Taxus* species by microwave-assisted extraction combined with LC-MS/MS analysis. *J. Sep. Sci.* **32**, 192–201 (2009).
21. Chattopadhyay, S. K., Sharma, R. P. & Kumar, S. Process for the production of important taxol analogues 10-deacetyl taxol A, B and C. US Patent No. 6,028,206 (2000).
22. Sisti, N. J., Swindell, C. S. & Chander, M. C. Paclitaxel synthesis from precursor compounds and methods of producing the same. US Patent No. 5,948,919 (1999).
23. Hanson, R. L. *et al.* Enzymatic hydrolysis of 7-xylosyltaxanes by xylosidase from *Moraxella* sp. *Biotechnol. Appl. Biochem.* **26**, 153–158 (1997).
24. Yang, L. & Ge, G. B. A microorganism with  $\beta$ -xylosidase activity and its application in taxane xyloside. China Patent No. ZL200610046296.6 (2010).
25. Wang, K. *et al.* Microbial hydrolysis of 7-xylosyl-10-deacetyltaxol to 10-deacetyltaxol. *J. Mol. Catal. B-Enzym.* **68**, 250–255 (2011).
26. Yang, L., Luan, H. & Liu, X. A *Cellulosimicrobium cellulans*, its hydrolase and in the use of transformation of taxanes. International Application Patent, CT/CN2008/000618 (WO2009/030098 A1) (2009).
27. Dou, T. Y. *et al.* Functional and structural properties of a novel cellulosome-like multienzyme complex: efficient glycoside hydrolysis of water-insoluble 7-xylosyl-10-deacetylpaclitaxel. *Sci. Rep.* **5**, 13768 (2015).
28. Cheng, H. L. *et al.* Cloning and characterization of the glycoside hydrolases that remove xylosyl groups from 7- $\beta$ -xylosyl-10-deacetyltaxol and its analogues. *Mol. Cell Proteomics* **12**, 2236–2248 (2013).
29. Zhu, P. *et al.* Glycosyl hydrolase with beta-xylosidase and beta-glucosidase activities and uses thereof. US Patent No. 9,206,405 (2015).
30. Yu, W. B., Liang, X. & Zhu, P. High-cell-density fermentation and pilot-scale biocatalytic studies of an engineered yeast expressing the heterologous glycoside hydrolase of 7- $\beta$ -xylosyltaxanes. *J. Ind. Microbiol. Biotechnol.* **40**, 133–140 (2013).
31. Liu, W. C. *et al.* Scaling-up Fermentation of *Pichia pastoris* to demonstration-scale using new methanol-feeding strategy and increased air pressure instead of pure oxygen supplement. *Sci. Rep.* **6**, 18439 (2016).
32. Liu, W. C. & Zhu, P. Pilot studies on scale-up biocatalysis of 7- $\beta$ -xylosyl-10-deacetyltaxol and its analogues by an engineered yeast. *J. Ind. Microbiol. Biotechnol.* **42**, 867–876 (2015).
33. Menhard, B. & Zenk, M. H. Purification and characterization of acetyl coenzyme A: 10-hydroxytaxane O-acetyltransferase from cell suspension cultures of *Taxus chinensis*. *Phytochemistry* **50**, 763–744 (1999).
34. Walker, K. & Croteau, R. Molecular cloning of a 10-deacetyl baccatin III-10-O-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **97**, 583–587 (2000).
35. Ondari, M. E. & Walker, K. D. The taxol pathway 10-O-acetyltransferase shows regioselective promiscuity with the oxetane hydroxyl of 4-deacetyl taxanes. *J. Am. Chem. Soc.* **130**, 17187–17194 (2008).
36. Pennington, J. J. *et al.* Acetyl CoA:10-deacetyl baccatin-III-10-O-acetyltransferase activity in leaves and cell suspension cultures of *Taxus cuspidata*. *Phytochemistry* **49**, 2261–2266 (1998).
37. Van Rozendaal, E. L. M., Lelyveld, G. P. & van Beek, T. A. Screening of the needles of different yew species and cultivars for paclitaxel and related taxoids. *Phytochemistry* **53**, 383–389 (2000).
38. Kerns, E. H. *et al.* Profiling taxanes in *Taxus* extracts using LC/MS and LC/MS/MS techniques. *J. Nat. Prod.* **57**, 1391–1403 (1994).
39. Falzone, C. J., Benesi, A. J. & Lecomte, J. T. Characterization of taxol in methylene chloride by NMR spectroscopy. *Tetrahedron Lett.* **33**, 1169–1172 (1992).
40. D'Auria, J. C. Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* **9**, 331–340 (2006).
41. Walker, A. M. *et al.* Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme A-dependent transferases and synthases. *Plant Physiol.* **162**, 640–651 (2013).
42. Laskowski, R. A. *et al.* PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
43. Ma, X. *et al.* Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J. Biol. Chem.* **280**, 13576–13583 (2005).
44. Kong, X. D. *et al.* Engineering of an epoxide hydrolase for efficient bioreolution of bulky pharmaco substrates. *Proc. Natl Acad. Sci. USA* **111**, 15717–15722 (2014).
45. Ramirez-Estrada, K. *et al.* Transcript profiling of jasmonate-elicited *Taxus* cells reveals a  $\beta$ -phenylalanine-CoA ligase. *Plant Biotechnol. J.* **14**, 85–96 (2016).
46. Onrubia, M. *et al.* Bioprocessing of plant in vitro systems for the mass production of pharmaceutically important metabolites: paclitaxel and its derivatives. *Curr. Med. Chem.* **20**, 880–891 (2013).
47. Rischer, H. *et al.* Plant cells as pharmaceutical factories. *Curr. Pharm. Des.* **19**, 5640–5660 (2013).
48. Walker, K. Application of plant and bacterial enzymes in the biocatalysis of paclitaxel (110.3). *FASEB J.* **28**(Suppl 1): 110–113 (2014).
49. Malik, S. *et al.* Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: a review. *Process Biochem.* **46**, 23–34 (2011).
50. Bentebibel, S. *et al.* Effects of immobilization by entrapment in alginate and scale-up on paclitaxel and baccatin III production in cell suspension cultures of *Taxus baccata*. *Biotechnol. Bioeng.* **89**, 647–655 (2005).
51. Kozakov, D. *et al.* The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protocols* **10**, 733–755 (2015).
52. Aslan, A. S. *et al.* Semi-rational design of *Geobacillus stearothermophilus*  $\alpha$ -lactate dehydrogenase to access various chiral  $\alpha$ -hydroxy acids. *Appl. Biochem. Biotechnol.* **179**, 474–484 (2016).
53. Biggs, B. W. *et al.* Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **113**, 3209–3214 (2016).
54. Prism, G. P. *Version 5.01* (GraphPad Software Inc., 2007).
55. Schwede, T. *et al.* SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).
56. Kirsch, R. D. & Joly, E. An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. *Nucleic Acids Res.* **26**, 1848–1850 (1998).

## Acknowledgements

We thank D. Lu and Y.S. Guo in our Institute for their help in the protein homology modelling, and Y. Li in our Institute for the *in vitro* cytotoxic detection of Taxol, DT and XDT (with several tumour cell lines). We also thank Y.J. Yuan's lab (Tianjin University, China) for synthesizing the *dbat* genes. This work was supported by the National Natural Science Foundation of China (Grant nos 81573325, 31270796 and 30770229), the National Mega-project for Innovative Drugs (Grant no. 2012ZX09301002-001-005), and the Youth Fund of Peking Union Medical College (Grant no. 3332015137).

## Author contributions

B.-J.L. designed and performed the experiments and drafted the manuscript. H.W. helped to prepare the mutant proteins. T.G. helped to analyse the enzymatic reaction results. T.-J.C. helped to prepare LXYL-P1-2 protein and make molecular docking. J.-J.C. and J.-L.Y. helped to analyse the data. P.Z. conceived, designed and supervised the study and revised the manuscript. All authors approved the final manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Li, B.-J. *et al.* Improving 10-deacetylbaconin III-10- $\beta$ -O-acetyltransferase catalytic fitness for Taxol production. *Nat. Commun.* **8**, 15544 doi: 10.1038/ncomms15544 (2017).

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



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

# Author Correction: Improving 10-deacetylbaocatin III-10- $\beta$ -O-acetyltransferase catalytic fitness for Taxol production

Bing-Juan Li, Hao Wang, Ting Gong, Jing-Jing Chen, Tian-Jiao Chen, Jin-Ling Yang & Ping Zhu

*Nature Communications* 8:15544 doi: 10.1038/ncomms15544 (2017); Published 18 May 2017, Updated 13 Jul 2018

In the originally published version of this Article, financial support was not fully acknowledged. The PDF and HTML versions of the Article have now been corrected to include support from the National Natural Science Foundation of China grant number 81573325.



**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) 2018