

1 **Supplemental Materials**

2 **A Novel Tetrahydrofolate (THF)-dependent Methyltransferase**
3 **Catalyzing the Demethylation of Dicamba in *Sphingomonas* sp. Ndbn-20**

4 Li Yao,^a Lin-Lu Yu,^a Jun-Jie Zhang,^a Xiang-Ting Xie,^c Qin Tao,^c Xin Yan,^a Qing Hong,^a
5 Ji-Guo Qiu,^a Jian He^{*,a,b} and De-Rong Ding^{*,c}

6 ^a Key Laboratory of Agricultural Environmental Microbiology, Ministry of Agriculture, College of
7 Life Sciences, Nanjing Agricultural University, Nanjing, Jiangsu 210095, PR China.

8 ^b Laboratory Centre of Life Science, College of Life Sciences, Nanjing Agricultural University,
9 Nanjing, Jiangsu 210095, PR China.

10 ^c DBN Biotech Center, Beijing DBN Technology Group Co., Ltd. Address: NO. 2, Yuanmingyuan
11 West Road, Haidian District, Beijing 100193, PR China.

12

13 * Author for correspondence: Jian He and De-Rong Ding

14 E-mail address: hejian@njau.edu.cn; Derong.Ding@dbn.com.cn

15

16 Tel: +86-25-84396685; Fax: +86-25-84395326

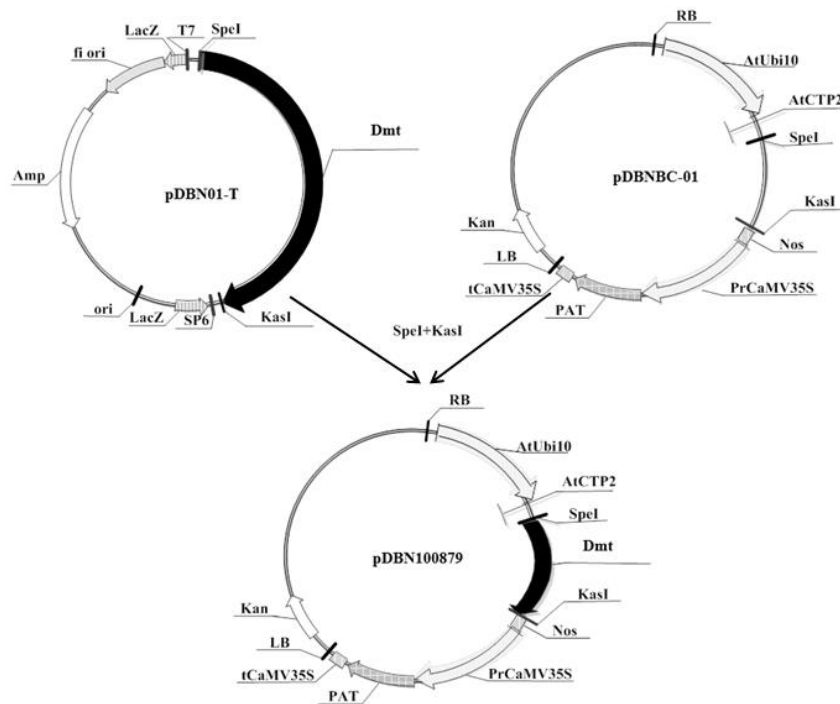
17

18 **Running title:** Tetrahydrofolate (THF)-dependent dicamba methyltransferase

19 **Keywords:** Dicamba; Tetrahydrofolate (THF)-dependent methyltransferase; Dicamba
20 resistance; *Sphingomonas* sp. Ndbn-20

21

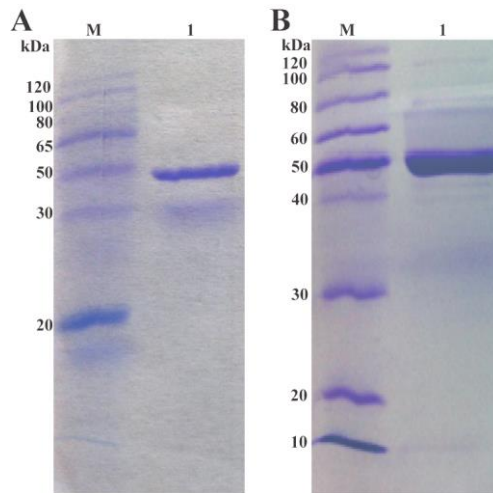
22 **FIG S1** Construction route of the recombinant vector pDBN10879. AtUbi10 represents
23 the promoter of the *Arabidopsis thaliana* polyubiquitin10 gene; AtCTP2 represents the
24 chloroplast transit peptide coding sequence of the *Arabidopsis thaliana* EPSP synthase
25 gene; tCaMV35S represents the terminator of the cauliflower mosaic virus 35S gene;
26 prCaMV35S represents the promoter of the cauliflower mosaic virus 35S gene; PAT
27 represents the glufosinate resistance gene; and Dmt represents the optimized *dmt* gene.



28

29

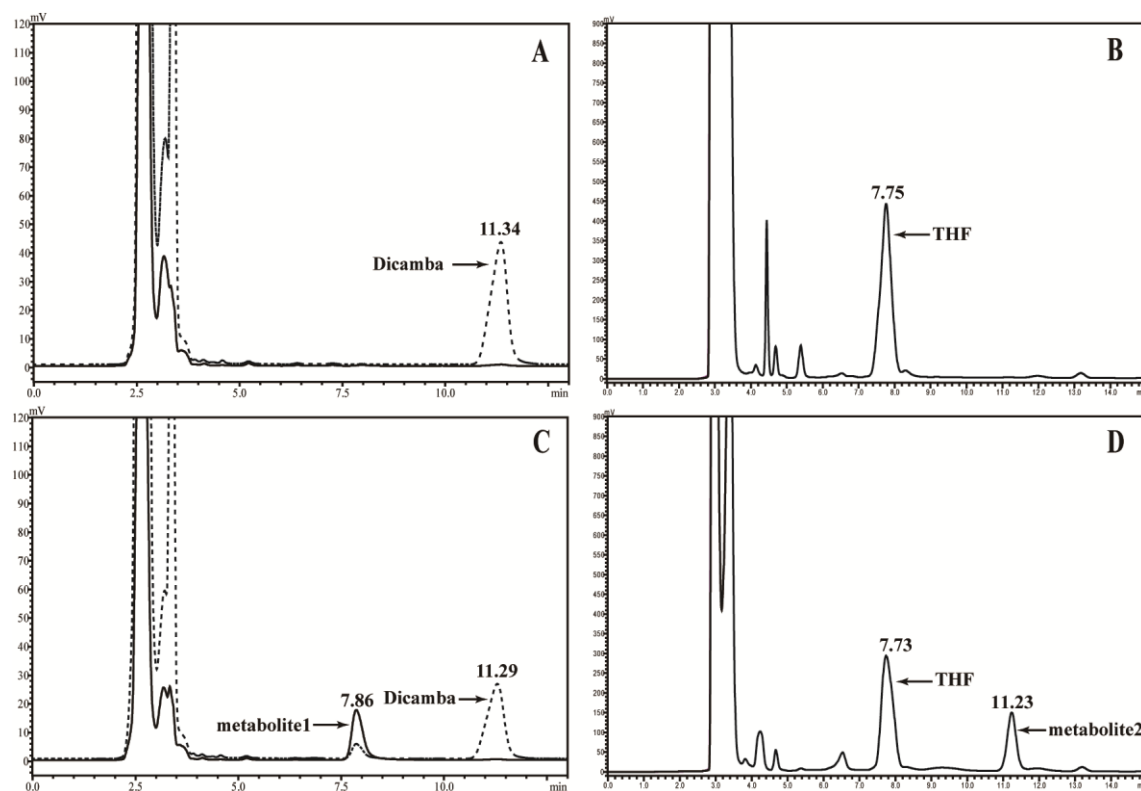
30 **FIG S2** SDS-PAGE analysis of the purified methyltransferases from scaffold 02 (A) and
31 Dmt (B). The proteins were separated on a 12% SDS-polyacrylamide gel and stained with
32 Coomassie brilliant blue. (A) Lane M, molecular mass standards; lane 1, purified
33 His₆-methyltransferase from scaffold 02. (B) Lane M, molecular mass standards; lane 1,
34 purified His₆-Dmt.



35

36

37 **FIG S3** HPLC analysis of the products during dicamba conversion by Dmt
38 (methyltransferase from scaffold 66). A and C, HPLC spectrum of the dicamba
39 conversion without or with Dmt, respectively. The sample was dissolved in methanol, and
40 the mobile phase was a mixture of ultrapure water (58.4%), acetonitrile (31.7%),
41 methanol (7.5%) and acetic acid (2.4%). The dotted line indicates dicamba (275 nm), and
42 the solid line indicates metabolite 1 (319 nm). B and D, HPLC spectrum of the THF
43 conversion without or with Dmt, respectively. The sample was dissolved in 0.1 mol/L
44 KH_2PO_4 buffer, the mobile phase was a mixture of 0.05 mol/L of pH 3.0 KH_2PO_4 buffer
45 (90%) and acetonitrile (10%), and the detection wavelength was 298 nm.



46

47

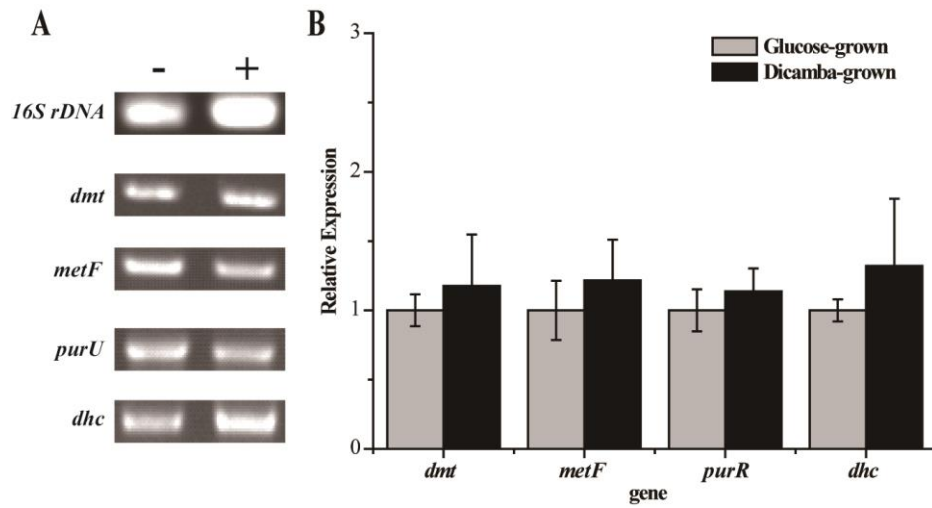
48 **FIG S4** DNA elements in the *dmt* gene cluster promoter. The -35 and -10 boxes are
49 shown in boxes, and the TSS is shown by an arrowhead.

CAA^{-35 box}**TTGGCA**AGGAGATTTTCCCGAA^{-10 box}**CGTCATGAT**GGGAG^{TSS}CGATATAAGCCCTTGCC
CATGTGGAAACGTTCCCATATTGATATGGCAGGGTCGTCGGGTGGTTCGCCGTCTTTCC
CGCTGGCCAATGCGCATCGCCGCCTTGCCCAGCGGGTCAGGGCGCGGCCGCTAGGGTC
GGCTTCGTCTTGCCGGAGTGTCGAAAACGCACCGGCCAGATCCAATTTATATCACGG
AGAGAGGT**ATGGTG**
Dmt →

50

51

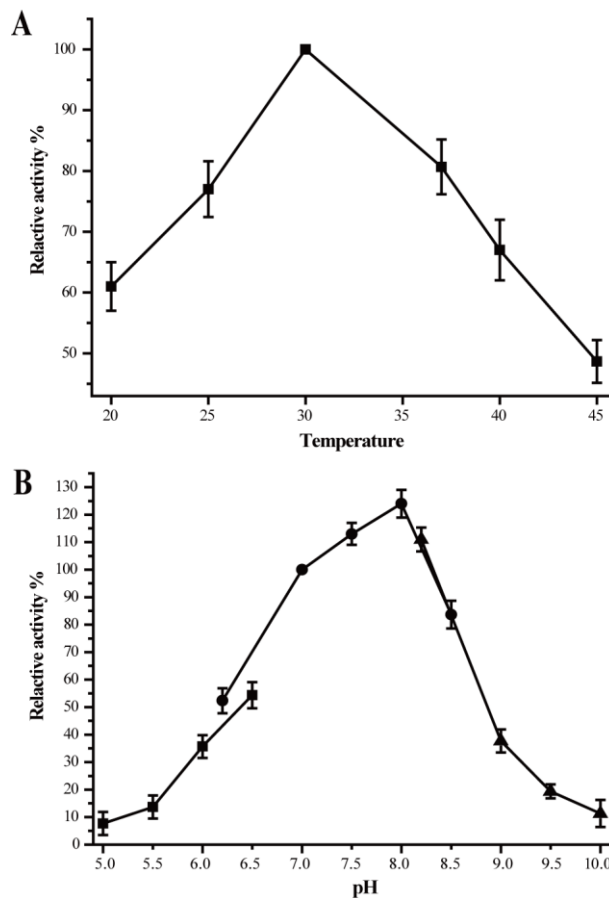
52 **FIG S5** Transcriptional analysis of *dmt*, *metF*, *purU* and *dhc* in scaffold 66 of strain
53 Ndbn-20 grown with glucose or dicamba. A: Agarose gel electrophoresis of the RT-PCR
54 products using the 16S rRNA gene as an internal control. B: Relative expression levels of
55 the four genes, as determined by RT-qPCR. The data were derived from three
56 independent measurements, and the error bars indicate standard deviations.



57

58

59 **FIG S6** Effect of temperature (A) and pH (B) on the activity of Dmt. For the temperature
60 test, Dmt was incubated under standard conditions at pH 8.0 and different temperatures
61 (20-45 °C) for 10 min, and the relative activity was calculated by assuming that the
62 activity observed at 30 °C was 100%. For the pH test, Dmt was incubated in the pH range
63 of 5.0 to 10.0 at 30 °C for 10 min, three different buffering systems were used: ■ indicated
64 100 mM phosphate buffer (pH 5.0 to 6.5), ● indicated 100 mM Tris HCl buffer (pH 6.2 to
65 8.5) and ▲ indicated 100 mM glycine-NaOH buffer (pH 8.2 to 10.0). The relative
66 activity was calculated by assuming that the activity observed at pH 7.0 was 100%.



67

68

69 **FIG S7** Alignment of nucleotide sequences of the *dmt* gene and the codon-optimized *dmt*

70 gene. The nucleotides optimized were marked out in red.

<i>dmt</i>	ATG GTG CGG TCG GTT CAG GAC ATT GTC GAT GAC AAC CCG GAC TTG GTG AAC TAT TTC AAA TAC CAG ACC GGA GGA TAC CCG ACC GCA TCG	90
codon-optimized_ <i>dmt</i>	ATG GTG AGA AGC GTT CAG GAT ATT GTG GAT GAC AAC CCA GAC CTG GTT AAT TAC TTC AAG TAT CAA ACA GGC GGA TAC GCC ACC GCT TCA	90
Consensus	atg gtg g gtt cag ga att gt gat gac aac cc gac tg gt aa ta ttc aa ta ca ac gg gga tac gc age gc tc	
Amino acid	M V R S V Q D I V D D N P D L V N Y F K Y Q T G G Y A S A S	
<i>dmt</i>	CTC ATC GAC ATC CTT CCG CCT TTC ATT CCT GCC GAG TTC ACG AAT TGG CGC GAG GAG CAG CGC GCT TGG CAC GAA CCG CCG GTG ATG TTC	180
codon-optimized_ <i>dmt</i>	CCT ATT GAT ATC TTG CCG CCT TTC ATC CCC GCC GAG TTT ACG AAC TGG AGG GAG GAA CAG CGC GCT TGG CAT GAA GCC GCT GTT ATG TTT	180
Consensus	ct at ga atc t cgg cct ttc at cc gcc gag tt acg aa tgg g gag ga cag cgc gct tgg ca gaa gc gc gt atg tt	
Amino acid	L I D I L R P F I P A E F T N W R E E Q R A W H E A A V M F	
<i>dmt</i>	GAC CAG TCG CAT CAC ATG TCA GTG ACC TAC ATC AAG GGT CCG GAT GCC AAG CCG ATG CTG AAT TAT CTG TCG CCG TCC ACC TTC TCC AAT	270
codon-optimized_ <i>dmt</i>	GAT CAA ACG CAC CAT ATG TCA GTC ACC TAC ATT AAG GGG CCG GAC GCA AAA CCG ATG CTC AAC TAT CTG AGT CCA TCG ACC TTC TCC AAT	270
Consensus	ga ca ca ca atg tca gt acc tac at aag gg ccg ga g ca aa gcc atg ct aa tat ctg cc tc acc ttc tcc aat	
Amino acid	D Q S H H M S V T Y I K G P D A K A M L N Y L S P S T F S N	
<i>dmt</i>	CTG TCC ACC GAT CGT GGC AAA CAA TAT TTC GTC GCC AAT CCC GAC GGC AAG CAT ATC GGC GAC TGT GTC CTT CAT TAT TAT CCG GAT GAT	360
codon-optimized_ <i>dmt</i>	CTC TCT ACT GAT CCC GGC AAG CAG TAC TTT GTC GCA AAT CCT GAT GGG AAA CAC ATT GGC GAC TGC GTG CTG CAT TAC TAT CGA GAT GAC	360
Consensus	l t c ac gat cg ggc aa ca ta tt gtc ga ac cc pa dd g ka ca at ggc gac tg gt ct cat ta tat ggc gat ga	
Amino acid	E S T D R G K Q Y F V A N P D G K H I G D C G V L H Y Y G D D	
<i>dmt</i>	GAG GGG TTC GAG CTG ATC AGC GGC AAT CCG CTG CAA AGC TGG GTG CGG TAC CAT TGC CAG ACG GGC GAA TAT AAC GTC GAG CCG AGG CAC	450
codon-optimized_ <i>dmt</i>	GAG GGT TTC GAA CTC ATC AGT GGG AAC CCC CTG CAG TCG TGG GTG AGA TAC CAC TGT CAA ACA GGC GAG TAT AAT GIT GAA GCC CGG CAC	450
Consensus	gag gg ttc ga ct atc ag ggc aa cc ctg ca tgg gtg g tac ca tg ca ac gcc ga tat aa tv ea ge gcc cac	
Amino acid	E G F E L I S G N P L Q S W V R Y H C Q T G E Y N V E A R H	
<i>dmt</i>	GAT CCA ACG TCA CCC TTG GCC AAA GGG CAC CCG CCG AAG AAT TAT CGC TTT CAG CTT GAA GGC CCC AAC GCC CCG CAG ATC CTC GAT GAT	540
codon-optimized_ <i>dmt</i>	GAC CCG TCC TCT CCA CTC GCC AAG GGA CAT AGG CCG ACG AAA TAT AGG TTC CAG CTC GAG GGC CCT AAC GCT CCG CAA ATC CTG GAT GAC	540
Consensus	ga cc c tc cc t ggc aa gg ca gg eg ac aa tat g tt cag ct ga ggc cc aac gc cg ca atc ct gat ga	
Amino acid	D P S S P L G K G H R R T K Y R F Q L E G P N A R Q I L D D	
<i>dmt</i>	GTC TCC GGT GGA GGA ATG CCA GAG CTC CAG TTT TTC CCG ACC CCA TTC GTC ACC ATC GCG GGC TGC CCG GTC CAC ATC CTG CCG CAT GCC	630
codon-optimized_ <i>dmt</i>	GTC TCC GGT GGT GGG ATG CCC GAG CTT CAA TTT TTC AGG ACA GCA TTT GTC ACG ATC GCG GGA TGT AGG GTG CAC ATA TTG AGG CAC GCC	630
Consensus	gtc tcc gt gg gg atg cc gag ct ca tt tt g ac ge tt gtc ac atc gcc gg tg g gt cac at tg gg ca ggc	
Amino acid	V C V G G M P E L Q F F R T A F V T I A G C R V H I L R H G	
<i>dmt</i>	ATG GCC GCC CAT ACC GGC GCA GAG ATC TCC GGT CCC TGG GAC GAG ATG GAC ACC GGT CGT GAT CCG ATC CTG AAG CCC GCC GAA AAG CTC	720
codon-optimized_ <i>dmt</i>	ATG GCC GGA CAT ACA CGT GCT GAG ATC TCA GGT CCG TGG GAT GAG ATG GAT ACG GTG CCC GAT GCA ATA CTT AAG CCG GGG GAG AAA CTT	720
Consensus	atg gcc gg cat acc ggc gca gag atc tcc ggt cc tgg gac gag atg gat acg gtg ccc gat gca ata ctt aag ccg ggg gag aa ctt	
Amino acid	M A G H T G A E I S G P W D E M T V R D A I L K A G E K L I	
<i>dmt</i>	GGC CTG CCG CCG GCC GGC ACA TGG GCT TAC TAC ACC CCG ACC CTG GAG AAT GGA TGG ATC CCG ACG CCC TTC CCC CCG ATT TAC ACC AGC	810
codon-optimized_ <i>dmt</i>	GGC CTC AGG AGG GCC GGA ACA TGG GCC TAC TAT TCA GCT ACG TTG GAG AAC GGC TGG ATT CCA ACC CCT TTC CCC CCT ATC TAC ACC TCC	810
Consensus	ggc ct g g ggc gg aca tgg gc tac ta gc ac t g gag aa gg tgg at cc ac cc ttc ccc gc at tac acc c	
Amino acid	G L R R G G T W A Y Y S A L T L E N G W I P T P F P A G Y T	
<i>dmt</i>	GAC GCA ATG AAG CCG TAT CCG GAG TGG CTG CCG GCA GCC AGT CTT GAA GGC CAG ATG CAC GTC GCT GGC ACG TAT TAT GGC GAG CAG ATC	900
codon-optimized_ <i>dmt</i>	GAC GCA ATG AAG CCC TAC AGA GAG TGG CTC CCA GCA CCG AGT CTT GAG GGT CAG ATG CAC GTC GCA GGA TGG TAC TAT GGC GAG CAG ATA	900
Consensus	gac gc atg aag cc ta g gag tgg ct cc gca gc agt ct ga gg cag atg cac gtc gc gg ta tat ggc gag ca at	
Amino acid	S D M K A Y R E W A P A A S L E G Q M H V A W S Y Y G E Q I	
<i>dmt</i>	GAG GAT TAT TAT TGG GAC CCG TGG TCG CTT GGC TAC GAC AAG TTC GTC AAG TTT GAT CAC GAC TTC ATC GCG CCG CCG CCG CTG GAG AAG	990
codon-optimized_ <i>dmt</i>	GAA GAT TAC TAT TGG GAC CCG TGG TCC CTC GGG TAC GAT AAG TTC GTT AAA TTT GAT CAC GAC TTC ATC GCG AGG GCC CCT CTG GAG AAG	990
Consensus	ga gat ta tat tgg gac gc tgg tc ct gg tac ga aag ttc gt aa ttt gat cac gac ttc atc ggc g gc gc ctg gag aag	
Amino acid	E D Y Y W D A W S L G Y D K F V K F D H D F I G R A A L E K	
<i>dmt</i>	ATG AAG CCG CAG CCT CAT CGT GTC AAA CCG CTC CTG GTG TGG AAC GAA GAC GAT GTC ATG AAG GTG ATG CCG TCG CAA TTC GCC GAT CCG	1080
codon-optimized_ <i>dmt</i>	ATG AAA CCG CAG CCA CAT AGA GTC AAG CCG CTC CTG GTC TGG AAC GAA GAT GAC GTC ATG AAA GTG ATG GCA TCT CAA TTC GCC GAT CCG	1080
Consensus	atg aa ccg cag cc cat g gtt aa cg etc ctg gt tgg aac gaa ga ga gtc atg aa gtg atg gc tc caa ttc ggc gat gg	
Amino acid	M K P Q P H R V K R L L V W N E D D V M K V M A S Q F G D G	
<i>dmt</i>	CCA ATC TAT AAA CCG ATC GAC ATG CCG CTG GCC TTT TAT GGC CTG CCA CAG GCC GAT GAC GTC CCG TCG ACC GAC GGC CGA CAT GTC GCC	1170
codon-optimized_ <i>dmt</i>	CCT ATA TAC AAG GCA ATT GAC ATG CCT CTT CCG TTT TAT GGA ITG CCC CAG CCG GAT GAC GTG AGA AGC ACC GAT GGC CGG CAC GTT GCC	1170
Consensus	cc at ta aa ge at gac atg cc ct gc ttt tat gg tg cc cag gc gat gac gt g acc ga ggc cg ca gt ggc	
Amino acid	P I Y K A I D M P L A F M G L P Q A D D V R S T D G R H V G	
<i>dmt</i>	GTG TCG CAA TGG TGC GGT TAC ACG GTC AAT GAG CAC CCG ATG ATG TCG CTT TGC GGG ATC GAC GAG GCA TCC GCC CCG CCG GGT ACC GAA	1260
codon-optimized_ <i>dmt</i>	GTG TCC CAA TGG TGC GGA TAC ACT GTT AAT GAG CAT AGG ATG ATG TCC CTT TGT GGT ATA GAC GAA CCG TCT CGA CCG CCG GGA ACA GAG	1260
Consensus	gtg tc caa tgg tgc gg tac ac gt aat gag ca gg atg atg tc ctt tg gg at gac ga ge tc ge gcc ccg gg ac ga	
Amino acid	V S Q W C G Y T V W E H R M M S L C G I D E A S A A P G T E	
<i>dmt</i>	GTC GTC ATC ACC TGG GGC GAG CCG AAC GGC GGC CCG CCA AAT CCG CAA ACC GAA CGT CAT GAG CAG ACG ACG ATC CCC CCG ACG GTG CCG	1350
codon-optimized_ <i>dmt</i>	GTG GTC ATT ACG TGG GGT GAA AGG AAC GGT GGG GCT CCA AAT AGA CAG ACT GAG CCG CAC GAA CAA ACC ACT ATT CCC CCG ACC GTT GCT	1350
Consensus	gt gtt at ac tgg gg ga g aac gg g gc cc aat g ca ac ga cg ca ga ca ac ac at cgc gc ac gt gc	
Amino acid	V V I T W G E R N G G G P A N R Q T E R H E Q T T I R A T V A	
<i>dmt</i>	CIT ACC CCG GCC GAT GCC CTG TCG CAG GCC CCG GTC CCC CCG ACG CTC	1398
codon-optimized_ <i>dmt</i>	CTT ACT CCA GCA GAC GCT TTG TCC CAG GGA AGG GTC CCG CCA ACA TTG	1398
Consensus	ctt ac cc gc ga ge tg tc cag gg gg gtc ccg ccg acc ctg	
Amino acid	L T P A D A L S Q G R V P Q G R	