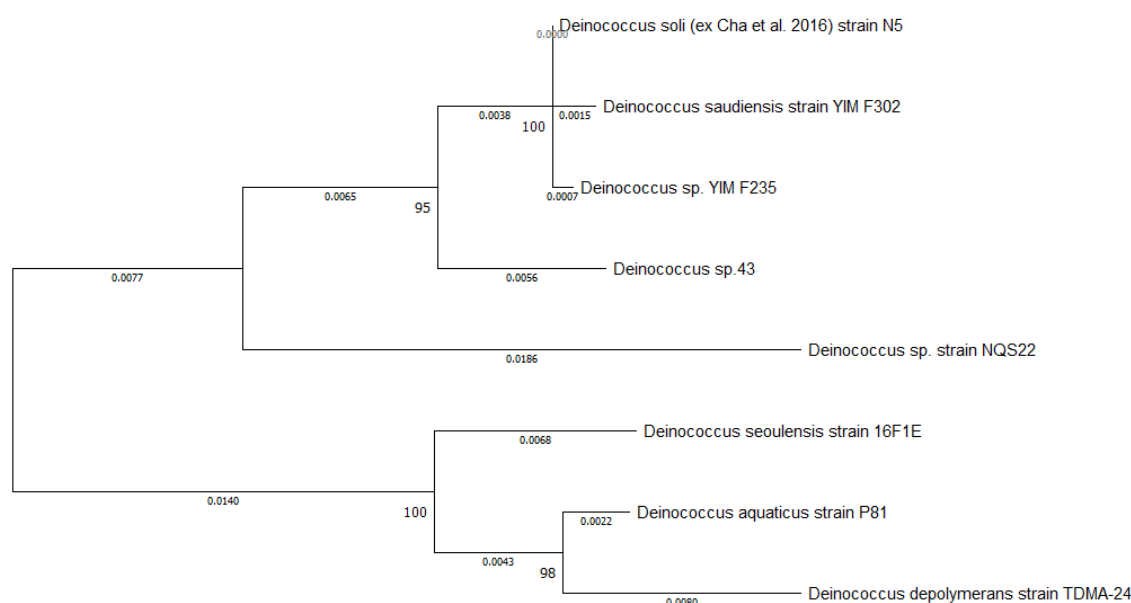




1. The numbers at the nodes indicate the levels of bootstrap support based on Kimura distance and the neighbour-joining method.



**Figure S1.** The numbers at the nodes indicate the levels of bootstrap support based on Kimura distance and the neighbor-joining method.

2. Homologous evolutionary relationship of key genes in the upstream flavonoid synthesis pathway. Evolutionary analysis was performed using MEGA X. Clustal W was used to compare each protein sequence. Based on the Poisson model, the phylogenetic tree was constructed by NJ (Neighbor-joining) method and evaluated with bootstrap value 5000. In order to further confirm the Protein homology, the similarity search of Hidden Markov model (HMM) (E value  $<1e-5$ ) was used to identify the characteristic Protein family (Pfam) domains contained in the two closest phylogenetic protein sequences. Relevant protein sequences were submitted to InterProScan 5 for conserved domain search.

(1).CHI:

> KALODDLK\_00111 tRNA 5-methylaminomethyl-2-thiouridine biosynthesis bi-functional protein MnmC

MIGAGIAGASVAYFLGRAGAQVTVDAGAHRASDVPSALINPVRGQSGGV DAR-  
ALDGM RFT-  
WTLLRDLKAAGHAVPHAQLGVL RPIPD DRARARFERNLPAALNHAWLSSAD APEPL  
APGWGHV LHL PDGGWVDGPAFTRALVQASGAQVVTGRAQDW-  
TARTVTLAGGDTLSGDAV-  
VFCGGSVGV TW RGEAATHRRG TLLTLDRAVTRVPLSFGAYLAPDARGGV LGATFETPS  
PTWTPDGLPLASLGWLLGKGAALTDLRGARVTGHW TGT RL SGLNAG-  
PQGGGTWWLTGLSSKGFL LGPLLAELASELMSAARPG

> WP\_029242752.1 FAD-dependent oxidoreductase [*Pseudomonas viridiflava*]

MSTLFETDALIIGGGIVGASAALALALKGKRVALLERDFCGSHSSGVNY-  
GGVRRQGRPLSQLPLSQRAHHIWGNLRDLIGIDGEYQRSGLKLARS DQDMYALRAY  
AQASQGFG LDLQLLDRAELR ARF-  
PWAGDVAVGASLCADDGHANPRLVSPAFARAAR-  
LAGAQVFEQAQVTQVNHDGQAFIVETANGLKLRAPWLLNCAGAWAGQLAAQFNE  
PVP MYSGHPAMLVTEPLPMFMDVSTGVEGGGIYARQVARGNCV LGGGQGFALD-  
PAR-  
TRPGQAAVLDILRNAVELYPPLKGAQAIR TWSGTEGYLPDREPVLGPSLAQPGL LHGF

GFAGAGFQIGPAAGEALAEWVCAGHSSISLDAFSIGRFRTEQALSITDSTLIVSPRSA-  
PEPVSNVIPLHRGRSPL



```

549
gi|1847250994|dbj|BCA79084.1|  MIGAGIAGASVAYFLGRAGAQTVDAGAHRASDVPSALINPVRQSGGVDRALDGMRF 60
MIV-----RLLL-----SLLLLAPLPVL-----AVEVEGVQL 27
**          :*          *: *:          *: *: *:

549
gi|1847250994|dbj|BCA79084.1|  TWTLRLDLKAAGHAVPHAQLGVLRLPDPDRARARFERNLPAALNHAWLSSADAPEPL--- 117
---PDRLEVQGETLHHNGHGIRKLF-----FD-----IYIGSLYTATPVRS 67
*.: *: *: *: *: *: *: *: *: *: *: *: *:

549
gi|1847250994|dbj|BCA79084.1|  ----APGWGHVHLHPDGGW---VDG---PAFTRALVQASGAQVVTG---RAQDWTART 162
SETLASPGAKLI---LMDFLYKKVEKEKIEAFREGLANQSAA---LASGEAKAFLWF--- 121
:* : *: *: *: *: *: *: *: *: *: *: *:

549
gi|1847250994|dbj|BCA79084.1|  VTLAGGDTLSGDVVFVCGGSGVGTWRGEAATHRRGTLLTLDRAVTRVPLSFGAYLAPDAR 222
---GRDFVRGDRVELELGGDG-----TVTARHNGELLGTLHSTAFEAELGIYLGDK-- 170
* *: *: *: *: *: *: *: *: *: *: *: *:

549
gi|1847250994|dbj|BCA79084.1|  GGVLGATFETPSPTWTPDGLPLASLGWLLGKGAALTDLRGARVTGHWTRLSGLNAGPQ 282
-----PADARLKAGMLQKQKER----- 186
* : *: *: *: *:

549
gi|1847250994|dbj|BCA79084.1|  GGGTWWLTGLSSKGFLGPLLAELASELSAARPG 318
----- 186

```

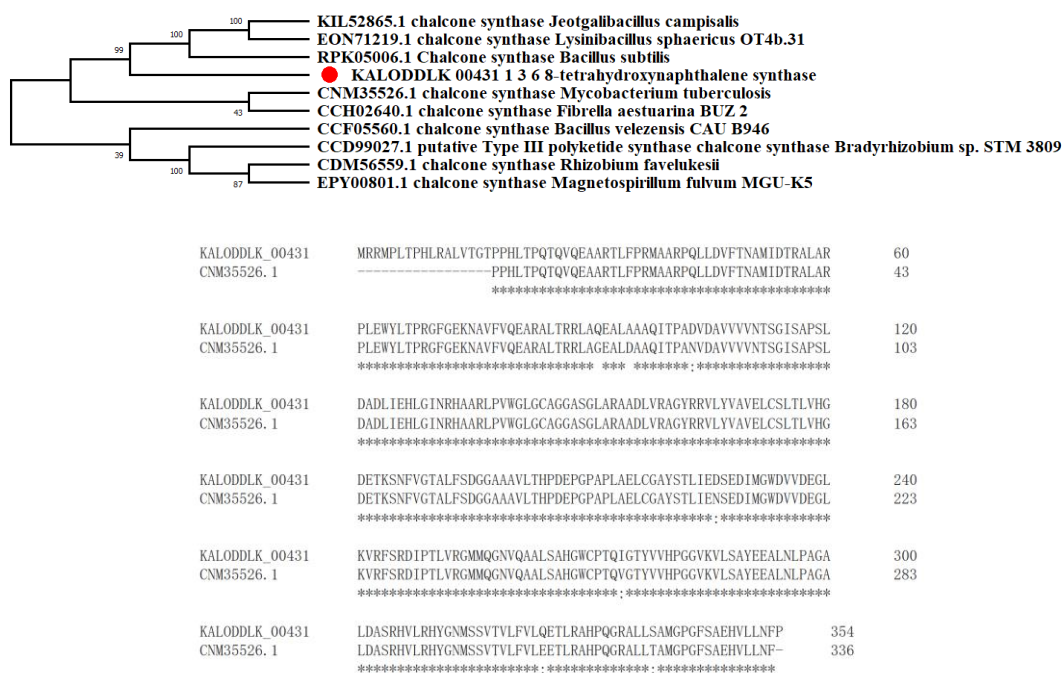
## (2).CHS:

>KALODDLK\_00431 1,3,6,8-tetrahydroxynaphthalene synthase

MRRMPLTPHLRALVTGTPPHLTPQTQVQEAARTLFPMAARPQLLDVFT-  
NAMIDTRALARPLEWYLTPRGFGEKNAVFVQEARALTRRLAQEALAAQITPADVD  
AVVVVNTSGISAPSLDADLIEHLGINRHAARLPVWGLGCAGGASGLARAADLVRA-  
GYR-  
RVLYVAVELCSLTLVHGDETKSNFVGTALFSDGGAAAVLTHPDEPGPAPLAELCGAYST  
LIEDSEDIMGWDVDEGLKVRFSRDIPTLVRG-  
MMQGNVQAALSAHGWCPQTQIGTYVHPGGVKVLSAYEEALNLPAGALDASRHVLR  
HYGNMSSVTVFLVLEETLRAHPQGRALLSAMGPGFSAEHVLLNFP

>CNM35526.1 chalcone synthase [Mycobacterium tuberculosis]

PPHLTPQTQVQEAARTLFPMAARPQLLDVFTNAMIDTRALARPLEWYL-  
PRGFGEKNAVFVQEARALTRRLAGEALDAAQITPANVDVVVNTSGISAPSLDADLI  
EHLGINRHAARLPVWGLGCAGGASGLARAADLVRAGYRRVLYVAVELCSL-  
LVHGDETK-  
SNFVGTALFSDGGAAAVLTHPDEPGPAPLAELCGAYSTLIENSEDIMGWDVDEGLKV  
RFSRDIPTLVRGMMQGNVQAALSAHGWCPQTQVGTYYVHPGGVKVLSAYEEAL-  
NLPAGAL-  
DASRHVLRHYGNMSSVTVFLVLEETLRAHPQGRALLTAMGPGFSAEHVLLNFP



(3).F3H:

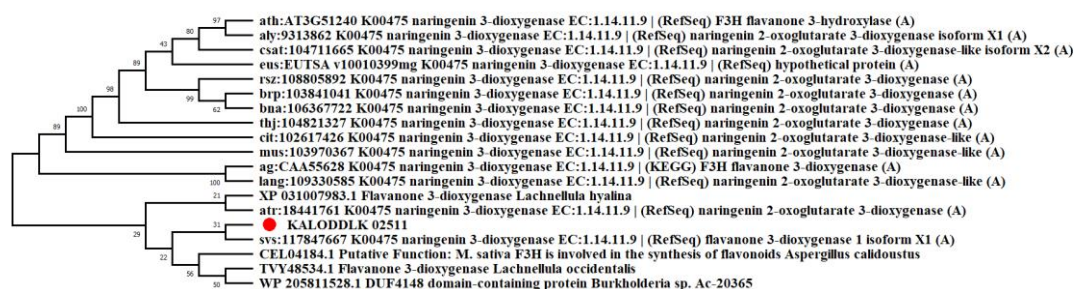
>KALODDLK\_02511 Catechol-2,3-dioxygenase

MTTDSVLPASTHVGAHVLNVRDLNGTAAFYATLLGLSATTLTETDVTLAAHGT-  
PLLHL-

HAAPDLPRAPVSRPGLYHTAFLLPTRAALGRWLAAHARLGHRIKSGDHLVSEAFYLN  
DPEGNGIEVYADRPRDWTWRDQVQMDTKAVDAAAVLASAGIDPATLDGTAP-  
FTAPQGTTVGHVHLKVGSAQAARWYADTLGLDVVADLGSAAFWSGGYHHHVG  
LNEWHSAGQPAPSTPAAGLAGVTFHAPDLGALRAHLAGRADVQNEGHLTLRD-  
PWGNRVTVAQ

>svs:117847667 K00475 naringenin 3-dioxygenase [EC:1.14.11.9] | (RefSeq) flavanone  
3-dioxygenase 1 isoform X1 (A)

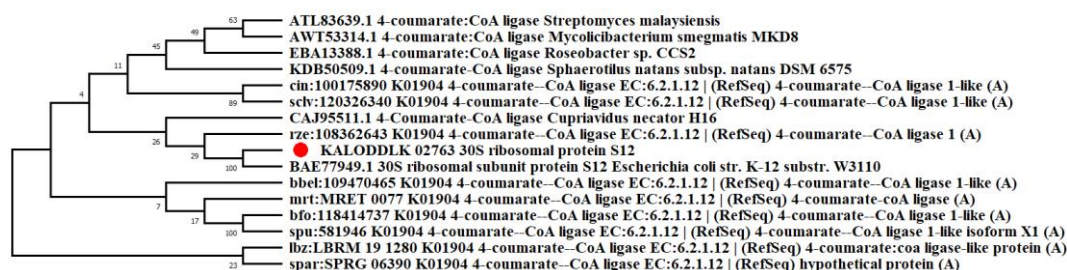
MLLYKAAAQAAYGVRYEQQPSAFRSTIIDREAAARTTTTMAPASSAAPFLPT-  
TASSEAT-  
LPPSFVREEDERPKVPHDSFSDEVPIVSLDGIGGGERRAEIRARVAAACEDWGIFQVVD  
HGVDAAALVAEMARLARDFFALPPQEKLRFDMSGGKKG-  
FIVSSHLQGEAVQDWREIVTY-  
FSYPVKARDYSRWPDKPAAWRAVVEQYSEQLVGLSCKLLGVLSEAMGLEAGALAEA  
CVEMDQKVNVNFYPRCQPPDLTLGLKRHTDPGTITLLQLDLVG-  
GLQATRDGGRTWITVQPVGEAFVNLGDHGHFLSNGRFGKNADHQAVVNSECSRLSI  
ATFQNPAPDATVYPLAVREGEAPILDEPITFAEMYRRKMARDIE-  
VAKLKKQARAAAEKQQVQVQMPQQSASEEFAVPKPESLDEILA



KALODDLK_02511 svs:117847667	MTTDSVLPASTHVGAVHLNVRDLNGTAFFYATLLGLSATTITETDVTLAHGTPLHLHA -----MLLY---KAAAQAAYGVRYEQQPSAFRSTIIDREAAARTTTTMAPASSAAPFLPTTA :* . . . . .** : :** :*: . ,*: : * : : * : :*: * *	60 54
KALODDLK_02511 svs:117847667	-----APDLPRAPVSRPGLYHTA----- SSEATLPPSFVREEDERPKVPHDSFSDSEVPVISLDGIGGGERRAEIRARVAAACEDWGIF * : * : * : * : *	78 114
KALODDLK_02511 svs:117847667	-----FLLPTRAALG-----RWLAHAARLGHRIKSGDHL QVVDHGVDAALVAEMARLARDFFALPPQEKLRFDMSGGKGGFIVSSHLQGEAVQDWREI * * * : * : : : * : . : . : *	107 174
KALODDLK_02511 svs:117847667	VSEAFYLNDEPGNGIEVYADRPDRTWTWRDQVQMDTKAVDAAAVLA-----SAGIDPAT V---TYFSYPV---KARDYSRWPDKPAWRAVVEQYSEQLVGLSCKLLGVLSEAMGLEAGA * * : , * . * : * : : ** * : : * : . : * : * : : *	162 229
KALODDLK_02511 svs:117847667	LDGTAPFTAPQGTTVGHVHLKVGSAQAARWYAD-----TLGLDVADLGSAAFL----- LAEC-----VEMDQ---KVVVNFYPRCQPDLTGLKRHTDPGTITLLQL * : . : : : . : : : * * * * : * : : * *	212 274
KALODDLK_02511 svs:117847667	-----SWGYYHHVGLNEWHSAGQPA---PSTPAAGLAGVT VGGLQATRDGGRTWITVQPEGAFFVNLGDHGHFLSNGRFKNADHQAVNSECSRLSIAT * : * : * : . : : : * : * : : * : * : *	245 334
KALODDLK_02511 svs:117847667	FHAPDLGALRAHLAGADVQNEGHLTLRDPWGNRVTAQ FQNPAPDATVYPLA-----VREGAPILDEP----- * : * : * * : * * : * *	285 360

## (4).4CL:

&gt;KALODDLK\_02763 30S ribosomal protein S12

VCTRVYTTTPKKPNSALRKVCRVRLTNGFEVTSYIGGEGHNLQEHSVILIRG-  
GRVKDLPGVRYHTVRGALDCSGVKDRKQARSKYGVKRPKA>BAE77949.1 30S ribosomal subunit protein S12 [Escherichia coli str. K-12 substr.  
W3110]MCTVVKTTTTPKKPNSALRKIARVRLSSAFEVTAYIPGEGHNLQEHSVVLIRG-  
GRVKDLPGVRYHIVRGSLDTQGVKDRNKSRSKYGTKKPKA

KALODDLK_02763	VCTRVYTTTPKKPNSALRKVCRVRLTNGFEVTSYIGGEGHNLQEHSVILIRGGRVKDLP	60
BAE77949.1	MCTVVKTTTTPKKPNSALRKIARVRLSSAFEVTAYIPGEGHNLQEHSVVLIRGGRVKDLP	60
	:** * *****: . ****: . . ****:* ** *****:*****	

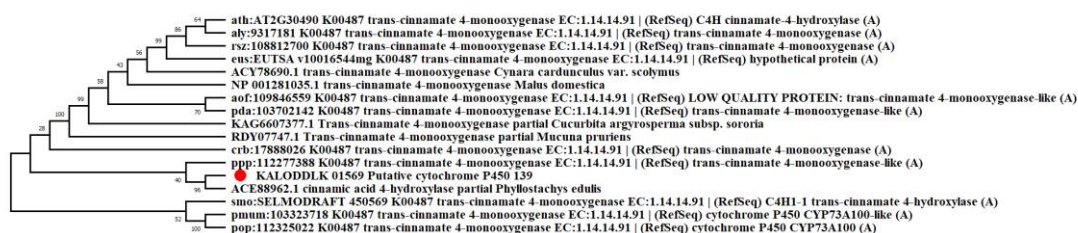
KALODDLK_02763	VRYHTVRGALDCSGVKDRKQARSKYGVKRPKA	92
BAE77949.1	VRYHIVRGSLDTQGVKDRNKSRSKYGTKKPKA	92
	*** **:* * , *****: : *****. * : **	

## (5).C4H:

&gt;KALODDLK\_01569 Putative cytochrome P450 139

GYDIPAESKILVNAWFLANDPKQWVRPDEFRRPERFLGEEKAV-  
KADGNDFRFPFGVGRRSCPGIILALPIIGITLGR>ACE88962.1 cinnamic acid 4-hydroxylase, partial [Phylllostachys edulis]  
GVRLPRGALALYSPYLSGRDPAVWDRPDEFDPGRWAGKPPAWAYLPFGG-  
GERLCLGMHLAQMLIHDALAAL

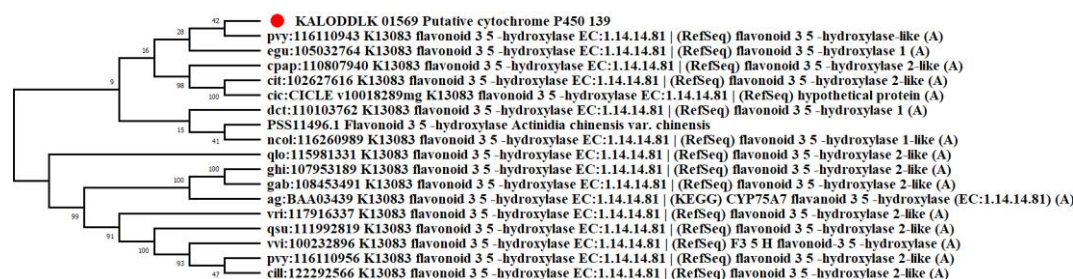




KALODDLK_01569	CPGIILALPIIGITLGR	78
ACE88962.1	CLGMHLAQMILHDAAL	71
	* * * * *	

>KALODDLK 01569 Putative cytochrome P450 139

AVMKEVLRRLHPILPLMVPHCPSESCVVGGYTI PKGARVFVNVAIHRDPSI-  
WENPTEFDPERFVDGKG DYSGKDFTYFPFGSGRRICAGIAMAERMVMFSLASL  
> PSS11496.1 Flavonoid 3',5'-hydroxylase [Actinidia chinensis var. chinensis]  
AALKEVLRLYPPGWMGSRRLSRDLIWGGVRLPRGALALYSPYLSGRD-  
PAVWDRPD EDPGRWAGKPPAWAYI PEGGGERI CI GMHL AQMI IHDALAAI



KALODDLK_01569	ERFVDGKGDYSGDKFTYFPFGSGRRICAGIAMERMVMFSLASL	104
PSS11496.1	GRWAGKP-----PAWAYLPFGGERLCGLMHLAQLIHDAALAAL	98
	*.:..:::*****.*:*:*.:*:.:**:**:	

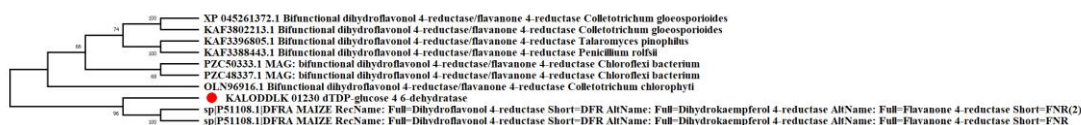
>KALODDLK 01230 dTDP-glucose 4,6-dehydratase

VLVTGASGFVGSWLVMKLLQAGYTVRATVRDPANVGKTKPLMDLP-  
GATERLSIW-  
KADLAEEGSFHDAIRGCTGVFHVATPMDFLSKDPENEVIKPTVEGMISIMRACKEAGT  
VRRIVFTSSAGTVNLEERQRPVYDEES-  
WTDVDFCRRVKMTGWMYFVSKTLAEKAALAYAAEHGLDLVTIIPTLVVGPFISASMP  
SLITALA

>sp|P51108.1|DFRA\_MAIZE RecName: Full=Dihydroflavonol 4-reductase;  
Short=DFR; AltName: Full=Dihydrokaempferol 4-reductase; AltName: Full=Flavanone 4-  
reductase; Short=FNR

VAVTGAEGFIGSHLVETLVRSGARVRAMVLYNSFGSWGWLDDLPEILEH-  
VEVVLGDVRD-  
PVSVREFMRGAEVVYHLAALIAIPYSYAAPHSYVOTNVVGTNLNVLEAARDLOTPLRV

# HTSTSEVYGTARSVPITERHPLQGGSPYSATKI- AADKLVEAYHLSFGLPVVTLRPFNTYGPRQSARAVIPTIISQIA



KALODDLK_01230	VLVTGASGFVGSWLMKLLQAGYTVRATVRDPANVGKTKPLMDLPG-ATERLSIWKADLA	59
sp P51108.1 DFRA_MAIZE	VAVTGAEFGIGSHLVETLVRSGARVRAMVLYNS-FGSWGLDDLPEILEHVEVVLGDVR	59
	* ***, ***, * * *, * : * : * * * * : , * , * * * * * : : , * , *	
KALODDLK_01230	EESGFHDAIRGCTGVFHVATPMDFLSK-DPENEVIKPTVEGMISIMRACKEAGTVRRIVF	118
sp P51108.1 DFRA_MAIZE	DPVSVREFMRGAEEVYHLAALIAIPYSYAAPHSYVQTNVVGTLNVLEAARDLQTPRL-VH	118
	: * , : : : * , * : * : : : : , : : : * , * : : : * , : : * * , *	
KALODDLK_01230	TSSAGTVNLEERQRPVYDEESWTDVDFCRRVKMTGW-MYFVSKTAEKAAALAYAAEHGLD	177
sp P51108.1 DFRA_MAIZE	TSTS-EVYGTARSVPIT-TERHPLQGGSPYSATKIAADKLVEAYHLSFGLP	166
	* * : : * * , * : , * : * * * , * : * , * * , * * , * * , * * , *	
KALODDLK_01230	LVTIIPTLVVGPFISAS-MPPSLITALA	204
sp P51108.1 DFRA_MAIZE	VVTLRPFNTYGPRQSARAVIPTIISQIA	194
	: * * : * , * * * * : * : * : * : *	

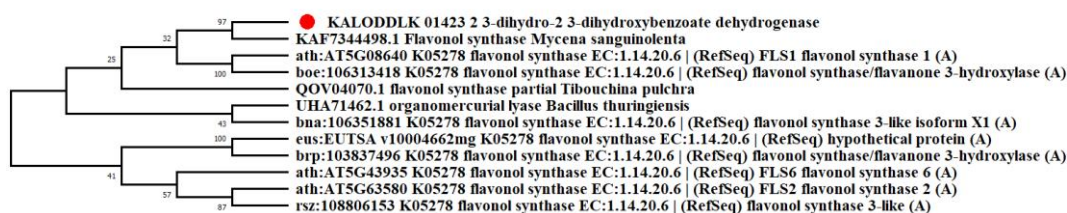
## (8). FLS:

> KALODDLK\_01423 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase

SSKVWLITGTSSGFGRLVSSVLARGDRVIATSRSLPIQLKGTND-  
NLRLLQLDVTAGEEL-  
LQCKMKEAVSSWGRIDVLVNNAGSCHLGILEEGGSALLRRQYEVNVFGLLDVTNACL  
PHLRAQTEATIVVMGSRSAWTCENMGIGPYGSSKAAVHAVAETLSVEVAPFNIR-  
VLIVEPSAFRTRMVR

> KAF7344498.1 Flavonol synthase [Mycena sanguinolenta]

TARVALVTGAAQGIGAASALAAARGARVLATDRQPIPPDLAAQPGVMPY-  
PLDVTDPRAAEQAVAHAEATLGPLTDLNVNAGVLRPGPLTDLSDDEDWHATFAVNTSG  
VFFVSRAATRVMRARGSGSVTVGSNAAHVPRTGMGAYASSKAAAAHL-  
TRTLGLELAGSGVRCNLVSPGSTDTPMQR

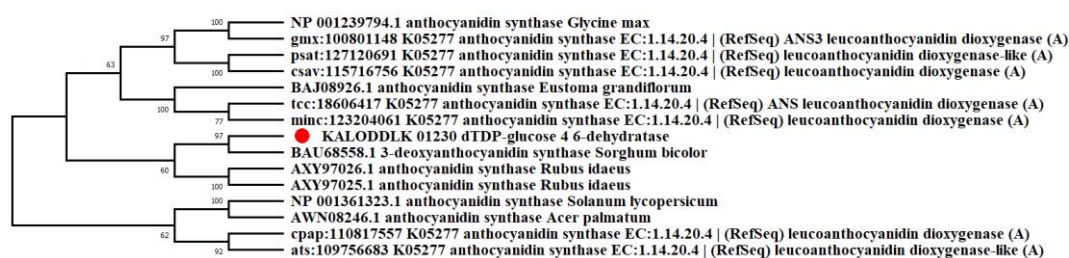


KALODDLK_01423	SSKVWLITGTSSGFGRLVSSVLARGDRVIATSRSLPIQLKGTNDNLRLLQLDVTAGE	60
KAF7344498.1	TARVALVTGAAQGIGAASALAAARGARVLATDRQPIPPDLAA-QPGVMPYPLDVTDP	58
	: : * * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_01423	ELLQCKMKEAVSSWGRIDVLVNNAGSCHLGILEEGGSALLRRQYEVNVFGLLDVTNACL	120
KAF7344498.1	AAE-QAVAHAEATLGPLTDLNVNAGVLRPGPLTDLSDDEDWHATFAVNTSGVFFVSRAATR	117
	: , * : : * : * * * * : * * : , : : * , * : : * , *	
KALODDLK_01423	HLRAQTEATIVVMGSRSAWTCENMGIGPYGSSKAAVHAVAETLSVEVAPFNIRVLIVEPS	180
KAF7344498.1	VMRARGSGSVTVGSNAAHVPRTGMGAYASSKAAAAHLTRTLGLELAGSGVRCNLVSPG	176
	: * : : : : * : * : * , * * * * , : : * , * : * : * : * , *	
KALODDLK_01423	AFRTRMVR	188
KAF7344498.1	STDTPMQR	184
	: * * *	

## (9).ANS:

>KALODDLK\_01230 dTDP-glucose 4,6-dehydratase

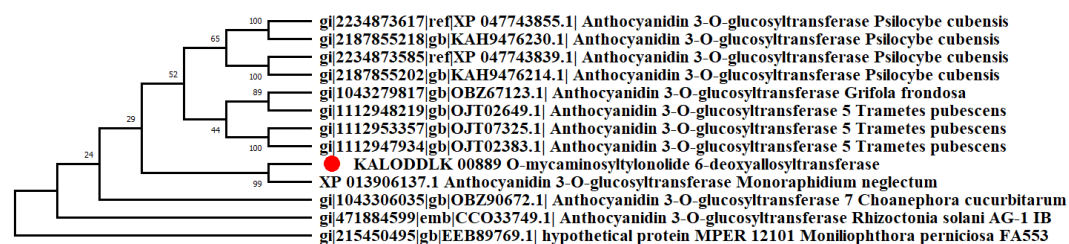
KTACVTGGSGYIGSALIKLLEKGYAVKTTVRNPDDMEKNSHLKDLQKLG-  
PLTVFRADMDEEGSFDDAVAGCDYVFLVAAPLHFEAQDPEKEQIEPAIQGTNLNTRMSC  
VKAGTVRRVILTSSVAAYVFRPDLGDBGHGHVLEDESDVDFLRAHKPPT-  
WSHCVSKVLEKEAGRFAEEHGISLVTILPVIVVGAAPAPKARSSIVDCLSMLS  
> BAU68558.1 3-deoxyanthocyanidin synthase [Sorghum bicolor]  
KLVAVTGAEGFIGSHLVETLVRSGARVRAMVLYNSFGSWGWLDDLPPEILEH-  
VEVVLGDVRD-  
PVSVREFMRGAEVVYHLAALIAIPYSYAAPHSYVQTNVVGTLNVLEAARDLQTPRLV  
HTSTSEVYGTARSVPITERHPLQGQSPYSATKI-  
AADKLVEAYHLSFGLPVVTLRPFNTYGPRQSARAVIPTIISQIA



KALODDLK_01230	KTACVTGGSGYIGSALIKLLEKGYAVKTTVRNPDDMEKNSHLKDLQ—KLGLPLTVFRAD	58
BAU68558.1	KLVAVTGAEGFIGSHLVETLVRSGARVRAMVLYNS—FGSWGWLDDLPPEILEHVEVVLGD	59
	* . . ** . * : *** * : : * . . * : : * . . . * . ** * : * . *	
KALODDLK_01230	MDEEGSFDDAVAGCDYVFLVA—APLHFEAQDPEKEQIEPAIQGTNLNTRMSCVKAGTV	114
BAU68558.1	VRDPVSVREFMRGAEVVYHLAALIAIPYSYAAPH—SYVQTNVVGTLNVLEAARDLQTP	116
	: : * . : : * . : * : * : * . . : : * : : * : : . * . *	
KALODDLK_01230	RRVILTSSVAAYVFRPDLGDBGHGHVLEDESDVDFLRAHKPPTWSHCVSKVLEKEAG	174
BAU68558.1	RLV—HTSTSEVYGTARSVPITERHPL—QGGQSPY—SATKIAADKLVE	159
	* * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_01230	RFAEEHGISLVTILPVIVVGAAPAPKARSSIVDCLSMLS	213
BAU68558.1	AYHLSFGLPVVTLRPFNTYGPRQS—ARAVIPTIISQIA	196
	: . . * : : * : * . * : * : * : * : * : * : * : * : *	

(10).BZ1:

> KALODDLK\_00889 O-mycaminosyltylonolide 6-deoxyallosyltransferase  
QVELFVTHGGLNSTYEGLAAGKPLVVMPPFGDQPVNAQHIVNKGL-  
GAQVDPWTLTAAKLTRAIEAQLADREAA  
> XP\_013906137.1 Anthocyanidin 3-O-glucosyltransferase [Monoraphidium neglec-  
tum]  
RVAAAVHHGGAGTTAAGLAAGVPNVVPPFGDQPFWGDRV-  
QRLGVGPAPVPRRALNERTLADALTRAVTDAGMRDRAAA



KALODDLK_00889	QVELFVTHGGLNSTYEGLAAGKPLVVMPPFGDQPVNAQHIVNKGLGAQVDPWTLTAAKL	59
XP_013906137.1	RVAAAVHHGGAGTTAAGLAAGVPNVVPPFGDQPFWGDRVQRLGVGPAPVPRRALNERTL	60
	: * * * * . : * * * * * * : * * * * * . : : . * : * . *	
KALODDLK_00889	TRAI—EAQLADREAA	73
XP_013906137.1	ADALTRAVTDAGMRDRAAA	79
	: * : * : * * *	



(11).F3'M:

&gt;KALODDLK\_01569 Putative cytochrome P450 139

AVMKESLRLHPPAPLLVPHCPSEDSLLGGFRVPKGSRVFVNVWAIHRDPT-  
 LWENPTEFDP-  
 DRFVNGGVDYSGNDFSYPFGSGRRMCAGMAMAERMVMYSLTLLHSFNWELPKG  
 EKLDLSEQFGIVLKKKLPLI  
 >XP\_031376784.1 flavonoid 3'-monooxygenase CYP75B137-like [Punica granatum]  
 AALKEVLRLYPPGWMGSRRLSRDLIWGGVRLPRGALALYSPYLSGRD-  
 PAVWDRPDEFDPGRWAGKPPAWAYLPFGGGERLCLGMHLAQMLIHDALAALPPLH  
 AVRGDATPLPGLTLGPRGPI

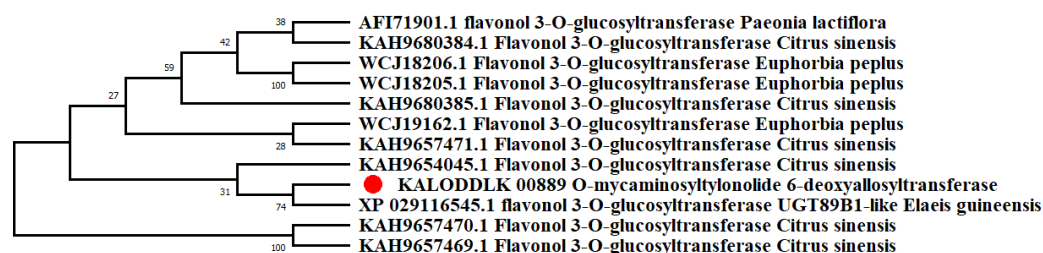


KALODDLK_01569	AVMKESLRLHPPAPLLVPHCPSEDSLLGGFRVPKGSRVFVNVWAIHRDPTLWENPTEFDP	60
XP_031376784.1	AALKEVLRLYPPGWMGSRRLSRDLIWGGVRLPRGALALYSPYLSGRDPAVWDRPDEFDP	59
	*:*:* ***:** : : *:* : ***:** : : : ***:** : ***	
KALODDLK_01569	DRFVNGGVDYSGNDFSYPFGSGRRMCAGMAMAERMVMYSLTLLHSFNWELPKGKLDL	120
XP_031376784.1	GRWAGK-----PPAWAYLPFGGGERLCLGMHLAQMLIHDALAALPPLHA---VRG---DA	108
	.*:.. :*****.*:* ***:** : : :*:** : : *	
KALODDLK_01569	SEQFGIVLKKKLPLI	135
XP_031376784.1	TPLPGLTLGPRGPI	123
	: *:. * : ***	

(12).R06611(Flavonol 3-O-glucosyltransferase):

&gt;KALODDLK\_00889 O-mycaminosyltylonolide 6-deoxyallosyltransferase

PAEEVVAWLDACPPRSVVYVCFGSQYTPTEKQGRA-  
 LAAALERSGVRFVWAIGGGGAV-  
 VPEGFEGRVVEILNHVAVGAFVTHCGWNSVLEAVAAGVVLLAWPMRADQFANARL  
 VVEELGVAVRVGEGEGTPEPEDLARVMAEAEVAEVAWPEMRVRAAALGRK  
 >XP\_029116545.1 flavonol 3-O-glucosyltransferase UGT89B1-like [Elaeis guineensis]  
 PPPALEAFLNSGPPPVSIGFGSMTTPDPQVTTTRAVVAALARSGQRAVLLSGWG-  
 GLSAADVPTDVFVTDSPHDLFPRVAAAVHHGGAGTTAAGLAAGVPNVVVPFFG  
 DQPFWDGRVQRLGVGPAPVPRRALNERTLADALTRAVTDAGMRDRAAALGAR



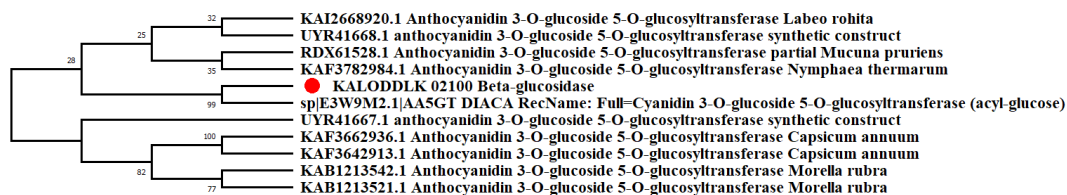


KALODDLK_00889	PAEVVAWLDACPPRSVVVYVCFGSQYTPTE-KQGRALAAALERSGVRVFWAIGGGGAVVP	59
XP_029116545.1	PPPALEAFLNSGPPP--VSIGFGSMITPDPQVTTTRAVVAALARSGQRAVLLSGWGGLSAA	58
	* : *: *: * : * * * * * : * * * * * : * * * * *	
KALODDLK_00889	EGFEGRV--EILN----HVAVGAFVTHCGWNSVLEAVAAGVLLAWPMRADQ-FANARLV	113
XP_029116545.1	DVPDVFVTDSPHDLFPRVAAVHHGGAGTTAAGLAAGVNVVPPFGDQPFWGDRV-	117
	: : : * : : . * * * * * : : : * * * : : * : * * * * :	
KALODDLK_00889	VEELGVAVRVGEGEGTPEPEDLARVMAEAV-AEVAWPEMRVRAAALGRK	162
XP_029116545.1	-----QRLGVGPAPVPRRALNERTLADALTRAVTDAGMRDRAAALGAR	160
	* : * * . * , * : * : * : * : * * * * * :	

(13).UGT75C1(Cyanidin 3-O-glucoside 5-O-glucosyltransferase):.

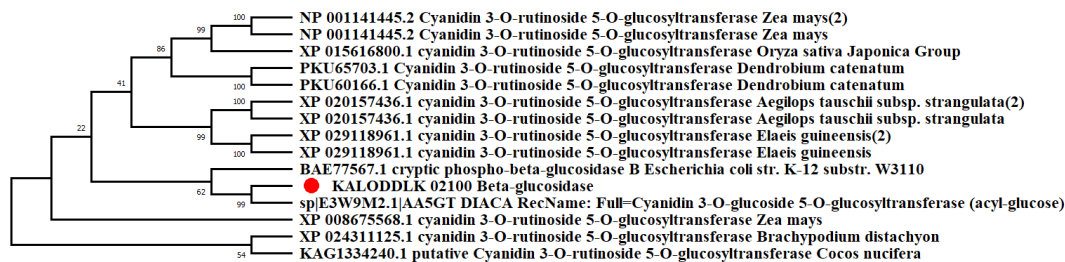
>KALODDLK\_02100 Beta-glucosidase.

DRLDFPKHFIFGASSCAYQVEGAAFEDGRTLSTFDIAAHSGLHPGNG-  
DITSDEYHKYKEDVELMVETGLDAYRFSISWSRLIPNGRGPVNPKGLEYNNLVNALL  
TKGTQPHVTLLHSDLPQALRDEYGGFLFISPKFIDDFVAYADVCFREFGDRVLRHWTT-  
FNE-  
ANFLAFGDENTPASALYLSAHHLLLAHASATRLYRENYQASQRGFIGINVYAYDFIPET  
NTEVDVIAAKRARDFFIGWFVQPLMNGEYPLTMRKNGGPRLPKFTPNETELLTGSYD-  
FIGL-  
NYYTAKTVKDDPVMLTVEPRNYYTDQGLISSYLGNIIDPYQGHPPFNTPWGLHDVLQQ  
FKQVYGNPPVYIHENGEVGDHDADYDKLINDIPRVEYLQGHIRAVLDA-  
VRNGSNVKGyFVWSFLDMYELMYGTFKFTFGLYIDFNDPKLTRHPKLSQKWYSRFL.  
>sp|E3W9M2.1|AA5GT\_DIACA RecName: Full=Cyanidin 3-O-glucoside 5-O-glucosyltransferase (acyl-glucose).  
DPAFFPARFTWGVATSSYQIEGAPREDGKGPSIWDTFCRTPGKVRGGDTGDVAC-  
DH-  
YHRLSDLDLMIRDLGVNAYRFSISWPRILPHGRGAVNQAGLDFYQRLVDGLLTRGITP  
WATLYHWDLPQTLEDEGGWTVRGTAFAFGTYSAVVAAALGDRVKHFITLNEP-  
WCSAYLGYGI-  
GIHAPGRHDLRASFAATHLLVGHGRAMQAIRGQAPGAQAGITLNLHHTYPATDTP  
ADRAAAYRMDGFQNRWYLDVPYVGRGYPDQDMVDLLGDL-  
SPQAQGLVLPDTELMGQPTD-  
FLGVNMYRAVGQDAPGEGFLHARQIRPEGSAYTGFDWEVAPDSLTDLLVRLQEDYA  
PDAIYITENGSTYPDVADEDGNVNDLERTQYLTEHLAATQEALARGAKVAGY-  
FAWSLMDNFEWAEGYDKRFGIVHVDFTQVRTPKLSGRTRYRDL.



[illegible]

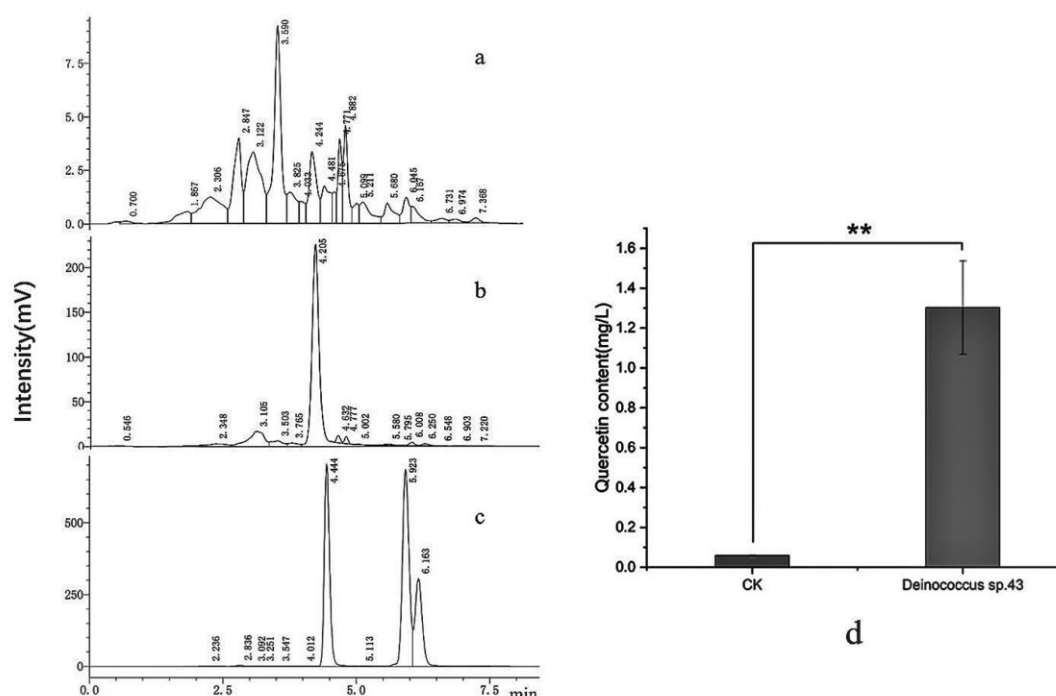
DRLLDFPKHFIFFGASSCAYQVEGAAFEDGRTLSTFDIAAHSGLHPGNG-  
DITSDEYHKYKEDVELMVETGLDAYRFSISWSRLIPNGRGPVNPKGLEYNNLNVNALL  
TKGTQPHVTLLHSDLPQALRDEYGGFLFISPKFIDDFVAYADVCREFGDRVLHWTT-  
FNE-  
ANFLAFGDENTPASALYLSAHHLLLAHASATRLYRENYQASQRGFIGINVYAYDFIPET  
NTEVDVIAAKRARDFFIGWVQPLMNGEYPLTMRKNGGPRLPKFTPNETELLTGSYD-  
FIGL-  
NYYTAKTVKDDPVMLTVEPRNYYTDQGLISSYLGNIIDPYQGHPFFNTPWGLHDVLQQ  
FKQVYGNPPVYIHENGEVGDHDADYDKLINDIPRVEYLQGHIRAVLDA-  
VRNGSNVKG YFVWSFLDMYELMYGKTFGLYIDFNDPKLTRHPKLSQKWYSRFL  
    >sp|E3W9M2.1|AA5GT\_DIACA RecName: Full=Cyanidin 3-O-glucoside 5-O-glu-  
cosyltransferase (acyl-glucose)  
    DPA RF PARFTWGVATSSYQIEGAPREDGKGPSIWDTFCRTPGKVRGGDTGDVAC-  
DH-  
YHRLDSLDLMIRDLGVNAYRFSISWPRILPHGRGAVNQAGLDFYQRLVDGLLTRGITP  
WATLYHWDLPQTLEDEGGWTVRGTA EAFGTYS AVVAAALGDRVKH FITLNEP-  
WCSAYLGYGI-  
GIHAPGRHDLRASFAATHHLLVGHGRAMQAIRGQAPGAQAGITLNLHHTYPATDTP  
ADRAAAYRMDGFGQNRWYLDPVYGRGYPQDMVDLLGDL-  
SPQAQGLVLPGDTELMGQPTD-  
FLGVNMYSRAVGQDAPGEGFLHARQIRPEGSAYTGFDWEVAPDSLTDLLVRLQEDYA  
PDAIYITENGSTYPDVADEDGNVNDLERTQYLTEHLAATQEALARGAKVAGY-  
FAWSLMDNFEWAEGYDKRFGIVHVDFTOVRTPKLSGRTRYRDFL



KALODDLK_02100	DRLDFPKHFIHGASSAYQVEGAFFEDGRTLSTFDIAHS—GHL—PGNGDITSDEYHKY	57
sp E3W9M2.1 AA5GT_DIACA	DPARFPARFTWGVATSSYQIEGAPREDGKGPSIWDTCRTPGKVRGGDTGDVACDHYHRL	60
	* ** : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	KEDVELMVETGLDAYRFSISWSRLIPNGRGPVNPGLLEYNNLVNALLTKGTQPHVTLH	117
sp E3W9M2.1 AA5GT_DIACA	DSLDLMIRD LGVNYRFSISWPRILPHGRGAVNQAGLDFYQRLVDGLLTRGITPWATLYH	120
	..*::: : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	SDLPQALRDEYGGFLISPKFIDDFVAYADVCFREFGDRVLHWTTFNEANFLA—FGDE—	173
sp E3W9M2.1 AA5GT_DIACA	WDLPTILEDEGGWT—VRGTAEAFGTYSAVVAALGDRVKHFITLNEPWCAYLGYGIGI	178
	***: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	—NTPASALYLSAHHLLLAHASATRLYRENYQASQRFIGINVYDFIPETNTEVDV	229
sp E3W9M2.1 AA5GT_DIACA	HAPGRHDLRASFAATHLLVGHGRAMQAIRGQAPGAQA—GITLNLHHTYPATDTPADR	235
	. * : : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	IAAKRARDFFIGWVQPLMNGEYPLTMRKNGGPRLP—KFTPNTELLTGSYDFIGLN	285
sp E3W9M2.1 AA5GT_DIACA	AAAYRMDGFQNRWYLDPVYGRGYPQDMVDLLGDLSPQAQGLVLPGLTLMGQPTDFLGYN	295
	** * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	YTTAKTVKDDPVMLTVEPRNYTDQGLISSVLGNIDPYQGHFFNTPWGLHDVLQQFKQV	345
sp E3W9M2.1 AA5GT_DIACA	MYSRAVGQDAPCEGFLHARQIR—PEGSAYTGDFWEVAPDSLTLVRLQED	345
	* : . * : * : . : * : . : * : * : . : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	YGNPPVYIHENGVEGDHDADYDKLINDIPRVEYLQGHIRAVLDVNRGNSNVKGYFVWSFL	405
sp E3W9M2.1 AA5GT_DIACA	YAPDAIYITENGSTYPDAEDGNVNDLERTQYLTEHLAATQEALARGAKVAGYFAWSLM	405
	* : * : * : * : . * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
KALODDLK_02100	DMYELMYGKFTFGLYYIDFNDPKLTRHPKLSQKWYSRFL	445
sp E3W9M2.1 AA5GT_DIACA	DNFEWAEGYDKRFGIVHVDFTQ—VRTPKLSGRTYRDFL	443
	* : * : * : . * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	

### HPLC results

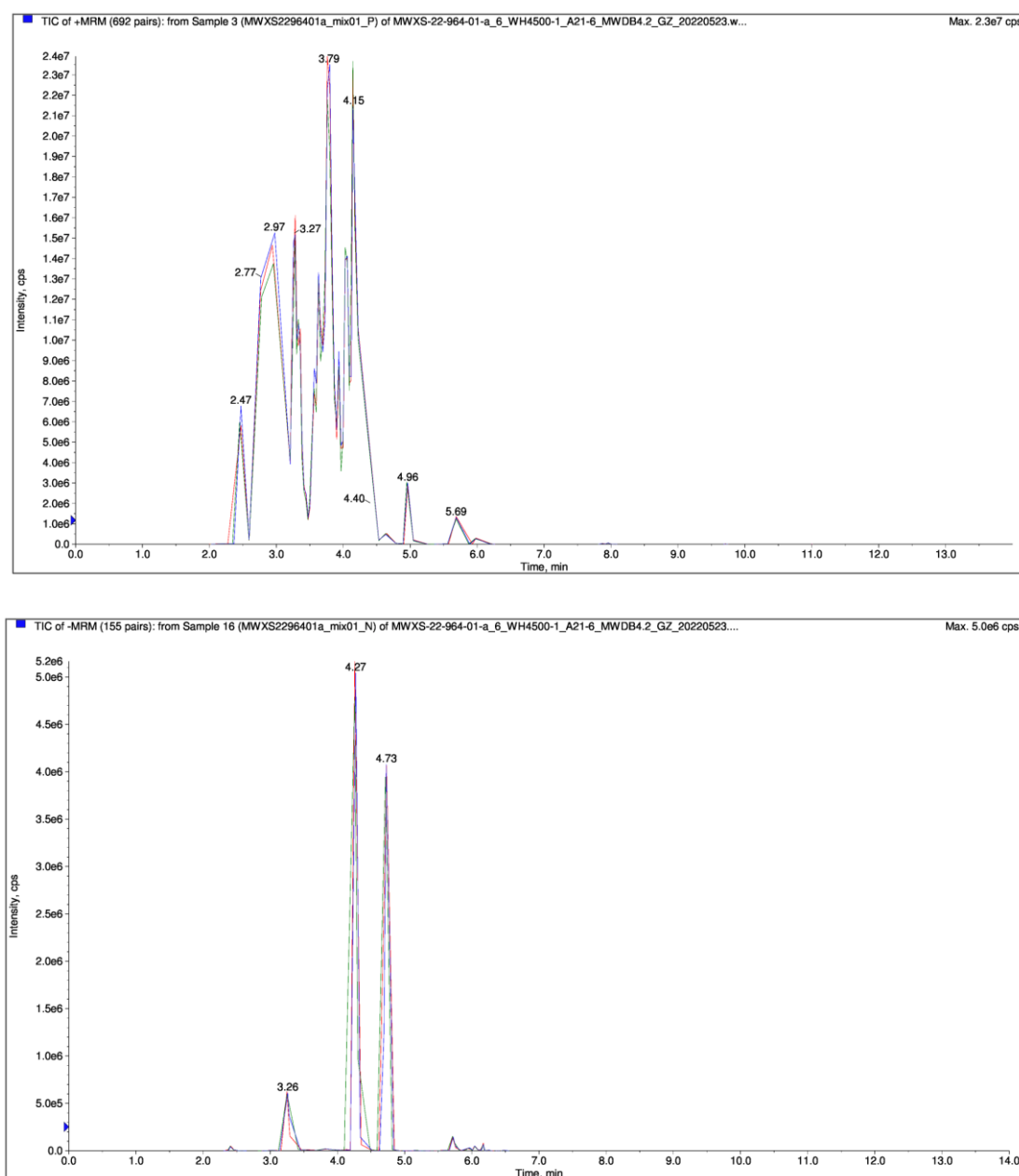
In Figure S2c, Eluting peaks of calibrator were observed at 4.4min, 5.9min and 6.2min, represent quercetin, kaempferol and isopurine respectively. The production of flavonoid during fermentation is determined by its peak area (mAU) and its properties can be determined by the time of peaking. **Figure S2a** shows the content of each substance in the blank medium, three flavonoids were not detected. **Figure S2b** represents bacterial fermentation broth, quercetin was identified.



**Figure S2.** In FS2(c), Eluting peaks of calibrator were observed at 4.4min, 5.9min and 6.2min, represent quercetin, kaempferol and isopurine respectively. The production of flavonoid during fermentation is determined by its peak area (mAU) and its properties can be determined by the time of peaking. FS2(a) shows the content of each substance in the blank medium, three flavonoids were not detected. FS2(b) represents bacterial fermentation broth, quercetin was identified.

### LC-MS detection quality control.

The repeatability of metabolite extraction and instrument detection can be judged by analyzing the overlap of mass spectrometry total ion chromatograms (TIC diagram) of three QC samples. **Figure S3** shown the TIC of three QC samples in positive ion and negative ion modes, which reflected the reliability of sample preparation process and LC-MS detection.



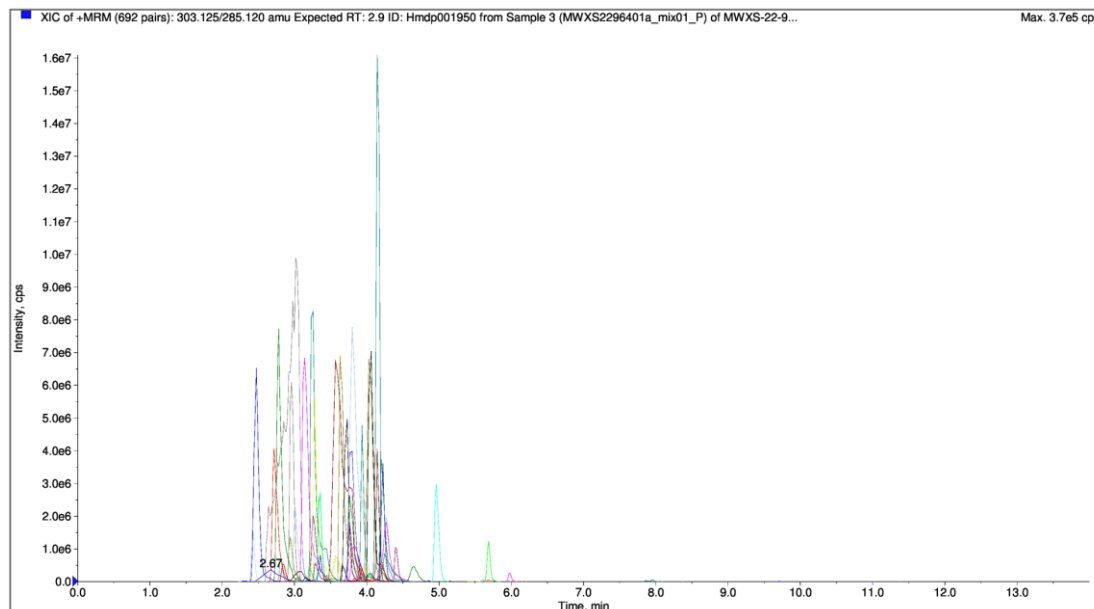
**Figure S3.** The repeatability of metabolite extraction and instrument detection can be judged by analyzing the overlap of mass spectrometry total ion chromatograms (TIC diagram) of three QC samples. FS3 shown the TIC of three QC samples in positive ion and negative ion modes, which reflected the reliability of sample preparation process and LC-MS detection.

Based on metabolite database analysis, we performed qualitative analysis of all detected ion types. In the multi-peak diagram of MRM metabolite detection in multi reaction monitoring mode, the samples that can be detected in positive ion and negative ion modes are displayed, and the peaks of different colors represent different metabolites (**Figure S4**). After filtering the blank control signal, there were relatively many kinds of

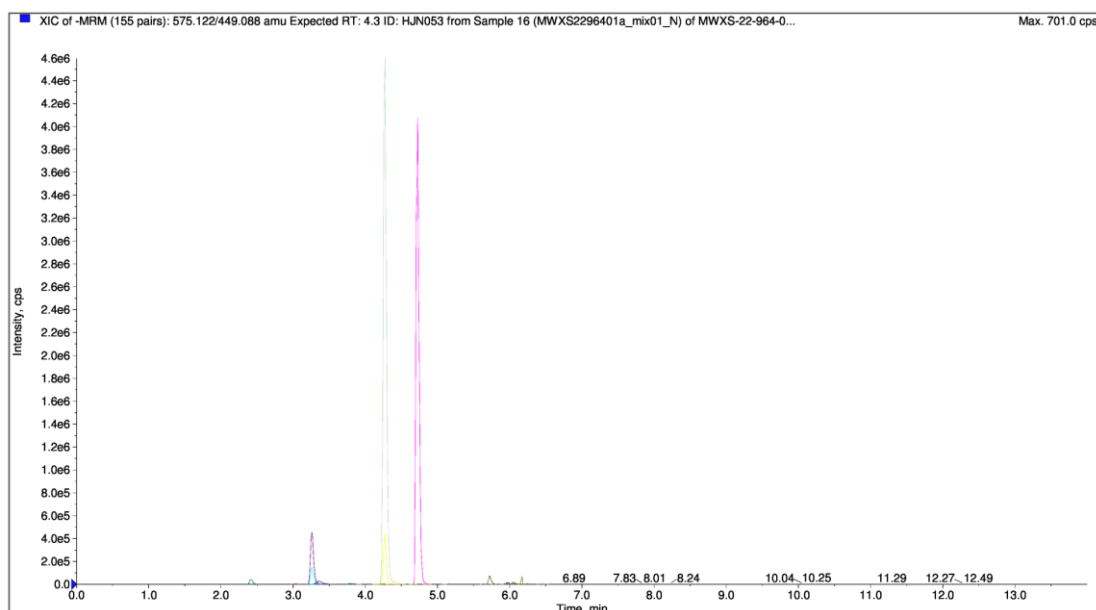


metabolites in the sample, and the number of peaks with strong signal was moderate, which was an ideal microbial experimental model.

positive ion modes



negative ion modes



**Figure S4.** Based on metabolite database analysis, we performed qualitative analysis of all detected ion types. In the multi-peak diagram of MRM metabolite detection in multi reaction monitoring mode, the samples that can be detected in positive ion and negative ion modes are displayed, and the peaks of different colors represent different metabolites (FS4). After filtering the blank control signal, there were relatively many kinds of metabolites in the sample, and the number of peaks with strong signal was moderate, which was an ideal microbial experimental model.