

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## Phylogenetic analysis

### General overview

*Table S1: Detailed information about the enzymes used in the phylogenetic analysis*

Lable	Species	Accession Numbers / Peptide reference	Main Product
<b><i>Antrodia cinnamomea</i> [1]</b>			
AcTPS4		EIW57364	zonarene
AcTPS5		EJF64719	T-cadinol
AcTPS7		EIW53985	nerolidol
AcTPS9		EIW57365	1-epi-cubenol
<b><i>Armillaria gallica</i> [2]</b>			
ArmGa1		PODL13.1	$\Delta$ -6 protoilludene
<b><i>Boreostereum vibrans</i> [3]</b>			
BvCS		AMW90891	$\delta$ -cadinol
<b><i>Coprinus cinereus</i> [4]</b>			
Cop1		XP_001832573	germacrene A
Cop2		XP_001836556	germacrene A
Cop3		XP_001832925	$\alpha$ -muurolene
Cop4		XP_001836356	$\delta$ -cadinene
Cop6		XP_001832549	$\alpha$ -cuprenene
<b><i>Coniophora puteana</i></b>			
Copu1		XP_007772164	---
Copu2		XP_007771895	$\beta$ -copaene
Copu3		XP_007765978	cubebol
<b><i>Fomitopsis pinicola</i> [5]</b>			
Fompi_84944		EPT01290	$\alpha$ -cuprenene
<b><i>Lignosus rhinocerus</i> [6]</b>			
GME3634		KX281943 / ASK39765	$\alpha$ -cadinol
GME3638		KX281944 / ASK39768	$\delta$ -cadinol
GME9210		KX281945	2H-2,4a-ethanonaphthalene
<b><i>Omphalotus olearius</i> [5]</b>			
Omp1		MUStwsD_GLEAN_10001317	$\alpha$ -muurolene
Omp3		MUStwsD_GLEAN_10003938	germacrene A
Omp4		MUStwsD_GLEAN_10005581	$\delta$ -cadinene
Omp5a		MUStwsD_GLEAN_10000810	$\gamma$ -cadinene
Omp5b		MUStwsD_GLEAN_10000811	$\gamma$ -cadinene
Omp6		MUStwsD_GLEAN_10003820	$\Delta$ -6 protoilludene
Omp7		MUStwsD_GLEAN_10000831	$\Delta$ -6 protoilludene
Omp9		MUStwsD_GLEAN_10000543	$\alpha$ -barbatene
Omp10		MUStwsD_GLEAN_10000292	trans-daucadiene
<b><i>Stereum hirsutum</i> [7]</b>			
Stehi_128017		XP_007299839	$\delta$ -cadinene
Stehi_159379		XP_007306912	$\beta$ -barbatene
Stehi2		XP_007301467	$\Delta$ -6 protoilludene
Stehi6		XP_007308318	$\Delta$ -6 protoilludene
Stehi7		XP_007299456	$\Delta$ -6 protoilludene

## Additional file

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## Detailed amino acid sequences

>AcTPS4

MSAQQFTLPDLLAVCPLKDATNPHYAQAAAESTAWVKSYNIFDARKLAFLLQGSSELLVSHAYPYAPY  
EQFRTCCDFVNLLFVVDEVSDDQNGKDARRTGEVYLVNVMRYPDWDDGSALAKMTREFKQRLLAFAAGPN  
SYRRFLMHCCDDYVNAVAREAEYRERGERVLDIDAFQTLRRENSAIRLCFGLFEFALGIDLPDFVFDQDPH  
FMTLYWSAADMVCWSNDVYSYNMEQAKGHTGNNIVTVLMRQKSIQGLQEAADLVGAHFSALMGRFVETK  
KQLPSFGAAALDDAVAKYVAAMEHWWIGNLEWSFESQRYFGAEHTRVKATRNVVLSPADEN

>AcTPS5

MAVTPAPVNGSDSTKEIILKFPDFISPIPYPLRCHPQEREVSRQSEEWLLSMANFSEKQRAKFLTLNA  
GLLSGCYIDCTFDELRVCTDFMNFLLTDDWTDFDFTTGTRGLAECVMNTLYFPDITYKSDTAAHRLT  
KSFWERMRATAGPGCQQRLLSTLDITYFQAIMQQAADRGSRNIPDLEEYILLRRDTSQCKTGFAFIEYA  
ANIDLDPDEVIEHPIKAMSDSTNDLVSWANDVLSYNAEQSRGDTHNLVLCVLMHQNNVDRQEAIEQAGE  
LWRKTLDYFECCHKALPSWGPEIDRAVALYVQGLDDWIIANAESWFETERYFGKEGPTVKKTRQIPLL  
PQRVRA

>AcTPS7

MSAGQNNPPLNVRQILSESASWSEANESALLEPFTHIFSI PGKEIRGQMIAAFNAWLVNPTDKLAIIS  
KVVSMMLHTASLLVDDIEDDAQLRRGVPVAHKIYGI PQTINSANYVYFLAYQELFALRSGVGADDDARE  
RLVPFNELDREVTAELLSLHRGQGLELLWRDSLQCPTEEEEYVSMVNNKTGGLFRVAVKLMMACSTTNI  
TVDYVPLVNLFGVYFQIRDDYMNLQSTQYAKNGFAEDLTEGKFSFPPIVHGVRADTSNRQLLNVLQKR  
PTTPTLKTTHAIGYLRDHTKSFTYTI SVMDDLEVQIREEITRLGGNARLEKIMDTLHVERPPTSQ

>AcTPS9

MSSPSSFVLPDLHAVTPFKGSFNPHYPEAAAESSEWVNSYKVLSDKKRAFFLQGGSELLCAHAYPYAG  
YQQFRTCCDFVNLLFTVDEISDDQNGKRAYETGLTFYNAMSNPAYDDGTVLCKMTKEFTARLLEHCGP  
QTYRRFIKHCKDYIEAVAVEADLRERGERVLDLEAYQTLRRENSAVRFCFGLAGYALGIDLDPDEVVEHP  
AFMAMHLSTVDMVCWSNDLYSYNMEQAMGHTGNNIVTVLMQHKGLDLQGAADYTGTVHFKGLIDTFLDA  
KRSLPSWGPKLDGEVAQYAMAMETWVIGNLNWSFETQRYFGHARHEIKRTRVVQLYPRRIVEESSDEE  
DN

>ArmGal

MSQRIFLPDTLANWQWPRHLNPHYAEVKKASAAWAKSFRAFQTKAQEAFDRCDFNLLASFAYPLADEA  
RLRSGCDLMLNFFVIDEYSDVSTEEVRAQKDIVMDAIRNTEKPRPAGEWIGGEVSRQFWDLAKKTAS  
TQAQKRFIDTFDEYLESVVQQAADRNNSHVGRGIESYLEVRRNTIGAKPSFALLEFDMQLPDESHQSSG  
YQRNLRKSCIDMLCLGNDVVSYNLEQARDDDGHNIVTIAMNELRTDVAGAMIWVDEYHKQLESRFMEN  
FKKVPRWGGPIDLQVARYCDGLGNWVRANDQWSFESERYFGKKGPEIIQRRWITLMPKMVSEELGPQI  
VDGFHL

>BvCS

MSTASSPSLVASEIDSPHHSRTSSPSPTLSPPTSFILPDLVSHCNFPLTYHPAGDEQAAAASLAWMLSF  
VPHFTPKKVAAMNGLQAGELTAYCYHDCPPERLRVVDVDFMNYLFHLDNISDGMMAKNTTQLADWVMNA  
FEWPEKFQPTVNADGEVVEEIAAVKLARDYWSRCIQQAKPGVQQRFKSSMNMFFQAVEQQTNDRDGQV  
VPDLESYIDMRRDTSQCKPVFDLIEYALGFELPEEVVDHPVIKALNQDANDLVTWSNDVFSYNVEQAR  
GDTHNMICIFMEHDGCTLQEAIDRVGGLCKQTIDAFVENKARVPSFAHLGPEVDAWTTGYVQGLQDWI  
VGSLSHWSFMTKRYFQEAGA EVKKTRFVKLLPIEEGRHKHIPP IYASAMVAATA

>Cop1

MVNLSWYWQGGQGNISKSIGPPSQTYTKSVLREQSMTFRMLALQSGLKSAASDHVSTSGSGILRFLSRI  
LPANTTRRCSACCTEMSSLDATIHPVLNFDKKIIVLPDLVSHCNFKLRVSRHRKRITGETKRWLFKGD  
NLVGPARNKYHGLKAGLLTAMTYPDAAYPQLRLCNDFLTYLFHIDNLSDDMDNRGTWSTANEVNLNLY  
HPYTYHGQARVGRMTRDYWRMILTASPGSQQRFIETFDFFFQSVTQQAIDRLTGEIPDLESYIALRR  
DTSGCKPCWALIEYANNLDPDEVMDHPVVRSLGEAANDLVTWSNDIIFSFNVEQSKGDTHNMI PVVMH  
QEGLDLQSAVDFVGEMLDHTSTCCGLESPLWLTPPLRIRRNAPRLYRVFL

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

>Cop2

MPSPAGALPKSFILPDLVNDPCFFPLRVNPLCDEVGRLSEQWFLRHANYSPPRAVAFMALKAGELTAAC  
YPDADAFHLRVSDDFMNFNADDWLDLDDIEDTYGLANCTVRALRDPVNFITDKRAGLMTKSYFSRF  
LKTAGPRTCERFIQTLALYFESVVTQKQARNNGTLPDLESYITIRRNNSGCKPCYALIEFCAGIDLDP  
EVINHPIIQSLEDASNDLIAWSNDIFSFNREQSRHDSFNMVSIVMHQKGFALQEAVNFVGEGLCKKAME  
RFQADKRNLPISWGPEIDGEVAMYVDGLQNWIVGSLNWSIDGTERYFGKDGPGIKKHKRVKLFPKRPLK  
TPAVRVLA

>Cop3

MSTPSSSLTTDESPASFILPDLVSHCPFLRYHPKGDEVAKQTVHWLDSNCPDLTAKERKAMYGLQAG  
ELTGICYPYTTPERLRVADFLNLYFHLNDNISDGMRETAVLADVVMNALWFPEDYRPTKGQAAEEL  
NPGKLARDFWSRCIPDCGPGTQARFKETFGSFFEAVNIQARARDEGVIPDLESYIDVRRDTSGCKPCW  
VLEIYALGIDLDPDFVVEHPVIAALNQGNTNDLVTWSNDIFSYNVEQSKGDTHNMIILMEHHGHTLQSA  
VDYVGSQCQTINTFCENKQQLPSWGPEIDDMVAKYVQGLEDWIVGSLHWSFQTRRYFGDEGQEI KQH  
RLVKLLTVAPPPPPPPPTPPPPQSSDADTKKQKVKAQDGKGPVSDEEVWALVRAEQSKGSILESFLGF  
TTSLSRIFFGYFFAYSH

>Cop4

MRPTARQFTLPDLFSICPLQDATNPWYQAAAESRAWINSYNI FTDRKRAFFIQGSNELLCSHVYAYA  
GYEQFRTCDFVNLLFVDEISDDQNGQDARATGRIFVNAMRDAHWDGSI LAKITHEFRERFVRLAG  
PKTVRRFADLCESYTDVAREAE LRERNQVLGLNDFIALRRQNSAVLLCYSLVEYILGIDLDDDEVYED  
PTFAKAYWAACDFVCWANDVYSYDMEQAKGHTGNNVVTVLMKEKDL SLQEASDYIGRECEKQMRDYLE  
AKSQLLQSTDLPQEAVERYIEALGYWMVGNLVWSFESQRYFGAQHERVKATHVHRLPSSVLEASCDSD  
SDSDC

>Cop6

MPAALPYNVSRDNKWDIKKIIQDFFKRCDVPYQVI PYDTELWNA CLKRAKEKGYPVEPDS PMSLYRSF  
KVGVVITRTSYGHIQDYEILIWVATFTAFVTYADDAFQEDIQHLHSFARTFLQNEKHEHPVLEAFAQF  
LRESSIRFSHFVANTVVSSALRFMMSIALEFEGQNVSVSTEAREYPGYIRILSGLSDIYALFAFPMDL  
PRSTYIQAFPEQIDYINGTNDLLSFYKEELDCE TVNFI SAAATSQQVSKLEVL RNAAEKAAYSYDVVV  
NVLKPYPEALAAWKS FARGFCYFHTSSPRYRLGEMFHDFEHDLVCKCASCTEI

>Copu1

MGLPQATEFVLPDFFAPCPFTLGLTNPHADTVFPEARAWIGKFLPFSGLAWHCDPWAGKEGFRTICDF  
QNLLFLMDELTDMSGEDARVVGESFIRVLKDPVDNESIIAQATREFRTRIADDVATKSVWFGRFVA  
ICKSYADATYIEAEHRENNRVLDDLDFVIARRENSAVRCCFSINEHALGIDLDPDSVFEDPEFLRMYFD  
AVDMIVIVNDVYSYNMEQAKGLAGNNVVTVLEQALGVLDLQAAVNRGGEMFAQKMEGYVRGRGVPPSWG  
AKVDADVEHFFDSVDQWIVGNLEWSTETSRYLGPDHEEIMKTGRVVLRKIESKTK

>Copu2

MSTMDPSEFILPDDFATCPFAFGRTNPHADVVIPEAHAWIVKHVPFVDRKRDEFIQDGFQDLMPHCYP  
WAGKETLRMTCYLNLLFLVDDLTDMMNSDEARGFGESFIRVLNDPAVHDSSQVMQATREFRSRITGT  
GVTESRWFGRFLAIFKLYINAVCAEAEDRENKRILDLDTFTVARRENSAVMVFFAIT EYALGIDLPA  
VYEDPTFLRVYADSADMVILVNDVFSYNREQAKGLDGNNNITVLMQTLDDLQAAVDHVGEMFSQKME  
GCMRGRAMLPISWGKVDADVERFFDALDQWVVGNLEWSSQSPRYLGPEHEEIMRTRRVVLRKVETEIE

>Copu3

MSATPAPTEFILPNLFSVCPLTFGRSNPYDEVIPEARAWIAKYNPFVDSKRAEFVQGCNELLCSRNV  
PYAGREEFRTCDFVNLLFVLDELSDDMGGADARSTCDSFIRVLNDPAPDTSLIAQMTREFRARVAE  
RAKPGCLRRFIALCGTYVEAVCVEAELREQGRVLDLRSFILLRRENSAVRCC LALAEYALGLELPDAV  
FNDPAFQSVYFCAADMVCWSNDVYSYNMEQAKGHTGNNVVTVLMQEHGIDLQAAADRVEVFGQLMEH  
YTSGSRSLPTWGGKVDADAARFLEAAGQWVVGNLEWSFETPRYFGPDHDEVDRDTHRVLK

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

>Fomp1 | 84944

MSRRYQTSRFSFSAKGGKTSLYTSPGDLRALFHRFSLYAGVQLESFTSPKMMQIEVPVMLNLKES SMP  
MNSTVLEDCTAAVAKDAIVQFLRRLGAVLRPSFGNNRDLEERVKEITKTWPF EHRHHPHITTGVMAN  
TTIAYLSDLARAAVAAYTALITALDDPDI FHASGAQNFAQMLCDGSALRDDGVLGQMARVLADMGNH  
FPPFGTSAIIAATLRWCNGELISNPANPFCLRPLSKAFADYQRGLTGVPEAYAAFVGWCKADFPVETDY  
IHVFPDICFFLNHTNDILSFYKETLDGESDSYIHARARLTGKSVTDTLYEVMDEVITTTTERIRKHFGE  
GRMRNAWDREYAGYVWFHTGNPRYRLHELVDTEYMPMY

>GME3634

MAVTPANVASPDSQEIVLKFPDFISPIPYPLRCHVQEREVSRQSEEWLLSMANFSEKQRSKFLTLNGG  
LLSGMCYIDCTFDELRVCTDFMNFLLFTLDDWTFDFDTTGTRGLAECVMNTLYWPHSYQADTAAHRLTK  
SFVWRMKQTAGPGCQQRLMSTLDTYFQAIMQQAADRGSNIPELEEYILLRRDTS GCKIGFAFIEYAA  
NIDL PDDVIEHPIIKAMADATNDLVSWANDVLSYNAEQSRGDTHNLVCV LMAQNGLDRQGAIELAGEL  
WEKTLHLFFECRKNVPSWGSEIDRAVALYIQGLEDWIIANA EWSFETERYFGKDGHLVKKTRQVTL LP  
VRTAA

>GME3638

MRARSFILPDLVSDCPYTLRCNSNCEAVARASEAWMLEDANLSPKRRDAFLRLRGGELTAACY PDTDE  
ACLRVAADFLNFLFLSLDDWSDEFMEDTCGLAQVCV LHDPPDFQTEKAAGKLAKSFFNRFRQTAGP  
RCTRRFIDSMDFFFHAIQAQADRASGSAPSLEEYVALREDTSGCKPCFALIEYAAGMDLPDHVAHP  
TITALE REANACISWSNDLFSYNVEQARGDTHNMI AVIMREDGRSLQEAVEYLGALCKLCIVHFEENR  
AMLPSWGPEIDGEVDRYVLGLQDWMVGALHWSFDTARYFGDEGPAIKKHGVVTL LPRKSSS

>GME9210

MPLSSSVVAFRLPDTLGCWPWRRCLNTHYVEAKQDSASWLESFHPFGPKAQRAF NKCDFNLLASL AYP  
VASKDQLRAGCDLMNVFFVFDEYSDVENEKTVQQLADIIMDALRNPHKPRPAGESL VGEITRQFWART  
IKVASEPSQRRFIETFDDYCQSVVQQAADRSQNHLDVESYLENRRENI GAKPSFALLELD MNLPDEV  
IEHPTIVNLT TWAIDMIILGNDIVSYNVEQARGDDGHNVTIVMHY NVDVQGAMDRIA EWHQRLADQ  
FLTNYNKLPSWGREIDAQVERYIQGIGNWVRANDAWSFESERYFGLNGREIEQSRWVTLL PRVSAEKP  
AVV

>Omp1

MKYTSFALPDLASSCDYNLRFNKYHRVSPETKKWFFRLSPASQADLT TYDAQRF TLLAAVCY PDAGY  
PQLRVCSDFLAYLFYLDNLTDDMDKSTRSVADLV LNSLNEPETFQTQYRIGKMTSDYFKRI IQTSND  
GTKKR FIDTMSFFKSVDDQARDRLAGHIPDLESYIALRRETSGCKTCFSLIEYANNLHIPDEVISHP  
HIEQMETAANDVVSFANDIYSFNIEQSKGDTHNMI PVLMHANPDMDFLEAVSFVRDLTIKAMDRF NEL  
RATLPSWGLDIDKDMKVYVNGLENWVGILFWSFETERYFGKSVRSVKATKTVNLLPSRA

>Omp3

MAIENTIASAPASTPAKQLDTPDHFILPDLVSHCTFPLVYHSNGDAVAAQSVKWLDTNCPDLNDKRRK  
ALYGLQAGELTAYCYNTAPDQRLRVVSDFMNYLFHLDNISDGMMTKDTDALSDAVMNALWFT EYRPT  
KKS DYVQPDEELNAGKLARDFWHRCIQDAGPGCQARFKETLELFFEAVNIQAKARDAGVIPDLESYID  
VRRDTS GCKPCWALIEYGLGIDLDPDYVAEDPIIKSLNQSTNDLVTWSNDIFS YNVEQSRGDTHNMI VI  
LMLYHGHNLQSAIDYVGDLCRQTIIDDFKENRKKIPSWGPEVDDIVKQYVQGLQDWIVGSLHWSFMTTR  
YFGKQGQEVKKNRYVKLLPVGEEANKW

>Omp4

MSSAPTRFLLPDL LLSACPLKGSVNPYYKEAGAESSAWINSYDIFTDRKRAFFVQGCNELLVAHTYPYA  
GYEEFR TCCDFINVL FVLDEVSDEQSGSDARFTGEVFLNALRNPENDDTSKLSKISKEFRARYFKRAG  
PRTAERFLQHCQDYIDCVTREAELRERGEVLDLPSFTALRRENSAIRVCFCLEFAYALGFDLPQEVFDD  
PTFMEMYWAAADLVCWANDVYSYNKEQAQGHGNNIVTVL MKAKDLDLQAACDYIGVYCEELMGRYLS  
AKARLPSWGPEVDAAVAQYVEASGHVWRGNLDWSFETQRYFGAQHAEIKETRLVTLTPAIPEDFSDTG  
SESE

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

>Omp5a

MSPDPTRIVLPDFLAACPFESSTKNPHLKAAGAESSAWVNSHVVFNDRKRAAFMQDIYELLVAYAFPW  
ADYEDFRTMCDFINLLFVLDLSDQNGKDAGYTGKLFMDAMRNI DNGDTSELTELCREFKARYSKRV  
SPQVNERFLQHLQSYTDCVAQEADLRERGEILDLESYVALRRENSAIRPCFDLVEYIIDFDI PQEVID  
HPVFSEMYWASVDLVCWSNDVYSYNVEQAKGHGGSNVVTVLMKEKNL DLQAACDYVGVVYEEELMDRYL  
SAKARLPSWGPEIDAAVGKYILAEAQFVRGNLDWSFDSPRYFGPQHDQVKKTGIVTLTPAPKKFGSDS  
GSESE

>Omp5b

MSPAPSRIVLPDFFASCPFESSTINPHFKAAGAESSAWVNSHVVFNDRKRAALMQNSYELLVAYAFPW  
ASYEDFRTLCDFINLLFVDFDEVSDQNGKDAGYTSKIFMDAMRNI DNGDHSELTELCKEFKARFSRRL  
SPQVNERFLQHLQSYTDCVAQEADLRERGEILDLESYVILRRENSAVRPCFDLVEYIMDFDI PQEVL D  
HPVFSEMYWASVDLVCWSNDVYSYNVEQAKGHRGSNVVTVLMNEKNL DLQAACDYVGVVYQELMDRYL  
SAKARLPSWGPEIDAAVGKYVLAEAQFVRGNLDWSFDTPRYFGPQRDQIKKSRIVTLTPAPKKFGSDS  
GSESE

>Omp6

MIAKNSEIDRFYIPDTLANWPWRHLNPAYPEAKKASAAWLRSFNANERSQKAFDLCDFNLLASLAF  
PLADLYCLRSGCDLMNCFIFIDEYSADVADPQTVRQQADI IMDAIRNPHVPRPRGEFIFGGEAHRQFWER  
AMQGATPTAQRRFIDTYQQYTDAVVQQATDRADNHIRDVEGYFTVRRDTIGAKPSFTLLEFTMDIPDE  
VMGHPVIKDLSLWCIDMLIIGNDLCSYNVEQAHGDDLHNLVTVIMNQYNL DLPGAMEWIGKFHDDIAD  
KFLDTFAKLPSWGPEIDPQIRRYVDGLGNWVRGNDWSFESWRYFRGKGPEIEKTRWVDLMPTEEATI  
TPKYESDSNAAQPAQST

>Omp7

MPETFYLPDCLANWKWKRALNPNYPEVKAASSEWLRSFKAFPKAQEAYDRCDFNLLASLAYPLADKD  
GLRTGCDLMNMFVDFEYSVAHESEVQVQADI IMDALRNPHKPRPVGEWVGGEVTRQFWELAIKTAS  
PQSQKRFIETFDTYTKSVVQQAADRTQHVVTVDEYLEVRRDTIGAKPSFAILELTMDIPDEVIHHT  
IERLAILAIDMILLGNDTASYNVEQARGDDNHNMTIVMHQYKTDIQGALSWEIKYHKELEEEFMQLY  
NSLPKGGQIDVDIARYVDGLGNWVRASDQWGFESERYFGTKAPEIQKTRWVTLMPKKRAEGVGPPEIV  
DISEL

>Omp9

MSQILHLLWSKFSTSLPSTVTIGSDPQTLQLVHSPAPNVNANALEIYKIVDNFLSRCGIRLESTPLDV  
EFYNECKKTLTSHYIGIHSDKVSSEWFKRYLSVGVIIITNAYGHLDNKLTKIYIALYALATCFDDV  
FEKNVDHMSGFNERFMKALPQGDVFLDAFAKVL LDAPKYFGRLASNI IVTSTLDFITSMSVDVLT KGM  
KFNQNLHKFAMACRNMSGIAYTYAPFIFPKEVPFAIYAQCLPDMRIYINHVNDVLSFYKEDKAGETEN  
LASILGQVHPSMTKYQIVQGLADDAEADLRVRTVLSQYQPALDAYNCFRQGYVSFHASSGRYRLDEL  
FSFVEPEPIV

>Omp10

MTLPTEQVELSVCVPESSHTTTRDIMRNFLSQCQIPLQRGVPLDPTFHQECANVLI EDYLKPSAAVTL  
ENLPSLMSSFNPFLLTGVRMASTGYAHLTHTPTRVYVALFTALLVCLDDIFPENVELMCGFNERFIKN  
ETQGEPI LDVAVGLLRSTSKYFSSMLSSNLIVTSALNYVTSLSLDQGLHSIKLAEHSRNFARLCRNMSG  
IPEAFAAFVFPPEVPFTAYIQCFPDLYTYANYVNDVLSFYKEDIAGETENLVSILAQTQPNSSRYQVL  
QQLADEAAAANANIRDILSDQKSILDAYDAFRVGVFVQFHIDSPRYRLAELFPCIDG

>Stehi\_128017

MAAPESSAWVSSYNLFSRKRRTDFITGSNELLVSHTYPHADYDAFRTCCDFVNL LFVIDEISDDQSGK  
AARRTGEVYLNAMRDPEWTDGSDLAKMTQQFRARFLRSVGPQSFRFRFLRHSEYIDCVAKEAEYRERG  
QVLDMDSFKSLRRENSAIRLCFGLFEFTLIGIDL PDSVFEDET FMKMYWASADMVCWANDVYSYNVEQA  
KGHSGNNIVTVLMAARDIDMQAASDYVGEYYAELMEEYMTAKAELASKSFGSRDLDEDVWKYVNAME  
N WPIGNLEWSFKTNRYFGTLHDEVKTRRLVVIKPRKVVV

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

>Stehi\_159379

MAHPTTNPSHRLEQMESVREHIPRLQHFLGEIGYRHTTTPPAPTLDLFLHAHHHWHHVLGPMTSWTVAK  
LNALEDSSSTIFERAYPLSDAEMKFVLAKLTAIAIFLDDSLEDEETYDDIGNFAHRVYLGEAQPTGVL  
TLYHQGIQELSKMHEGDAVFRGLAVAPWITFIDACMLEKRLTLFDSKLRVSPRDLGYQRLRNSTDFTS  
LRAPKATPSEVEVSFPIFLRHKSGIGEAYAAAIKSSRYQELPLSRFVKSMPMIYYIELVNDLMSFY  
KEQLAGETANLIHLQHQSWSKGGQGTGPYGSWTLLDTFSRLCDETRDAAFRVDELLRLDECEKIANGEL  
RGEEVGLSPMDVTMAAQWREFRDGYVSWHLECQRYKLDFIKLSTFE

>Stehi2

PRHINPHYQEVKKASAAWAESFGAFNPKAQHAYNACDFKRLRTGCDLMNMFFVFDEYSDVSSPKDVIQ  
QAAIIMDALRNPYAPRPDDEWVGGEVTRQFWKRAIKTATAGAQRFFIDAFESYTSVQVQAKDRHHGF  
IRDVDSYLEMRRETIGAKPSFVVLQMDMTLPDEVLAHPVIQQLSALSTDMICLGNRRLIQILWTVQDI  
CSYNVEQARGDDLHNIITIAMNQFDIDIAGAMDWVVKYHAKLERKFLYLYNNGLPSWGKELDPQVERY  
VCGL

>Stehi6

MVRSPVSDKFCIPDTLASWPYPRI LNPHYAEEKAASAAWTKGFGAFGPKAQDAFDRCDFKRCRSGCDL  
MNLFFVIDEHS DTHGEETVRKMKDVMDAIRNPHKPRPNDEWIGGEIARQFWERAMCYASEISQRRFI  
DTFDEYLESVVDQAADRDSARIRDIESYINIRNTIGAKPSFVIMEQGMIDPNVFENEVQRLRMAT  
IDMLCLGNDIVSYNIEQARGDDSHNIVRIVMNELDTDVPRAMDWVAQRHTQLEREFFTALSELPTWGE  
PIDGWVKEYVYGLGNWVRANDQWSFESQRYFGTKGMEIMKSRWLSVLPKVRPAEVPQLVDQSL

>Stehi7

MAVATSVATPVPTPAYSAGRAPAKEKKIYLPDTLAEWPWPRAINPHYAEAKEESQAWAASFNAFSPKA  
QHAFNRCDNFLLASLAYPLATKHGCRSGCDLMNLFFVIDEYSDIAPVEEVRQQKDIVMDALRNPHKPR  
PEGEWVGGEVARQFWALTITNASAQSQKHFIETFDEYLDVSVVQQAEDRSESRIQSYIDVRRNTIG  
AKPSFALLELDMDLPDEVLAHPTIQSLSLATIDMLCLGNDIVSYNLEQARGDASHNIITIVMNELNLD  
VNGAMRWVGDFHKQLEKQFFFAFNLPKWGNAELDAQIAVYCDGLGNWVRANDQWSFESERYFGARGL  
EIMETKTLAMPPIQRTEALGPQLVDDSI

## Additional file

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## Sequence analysis

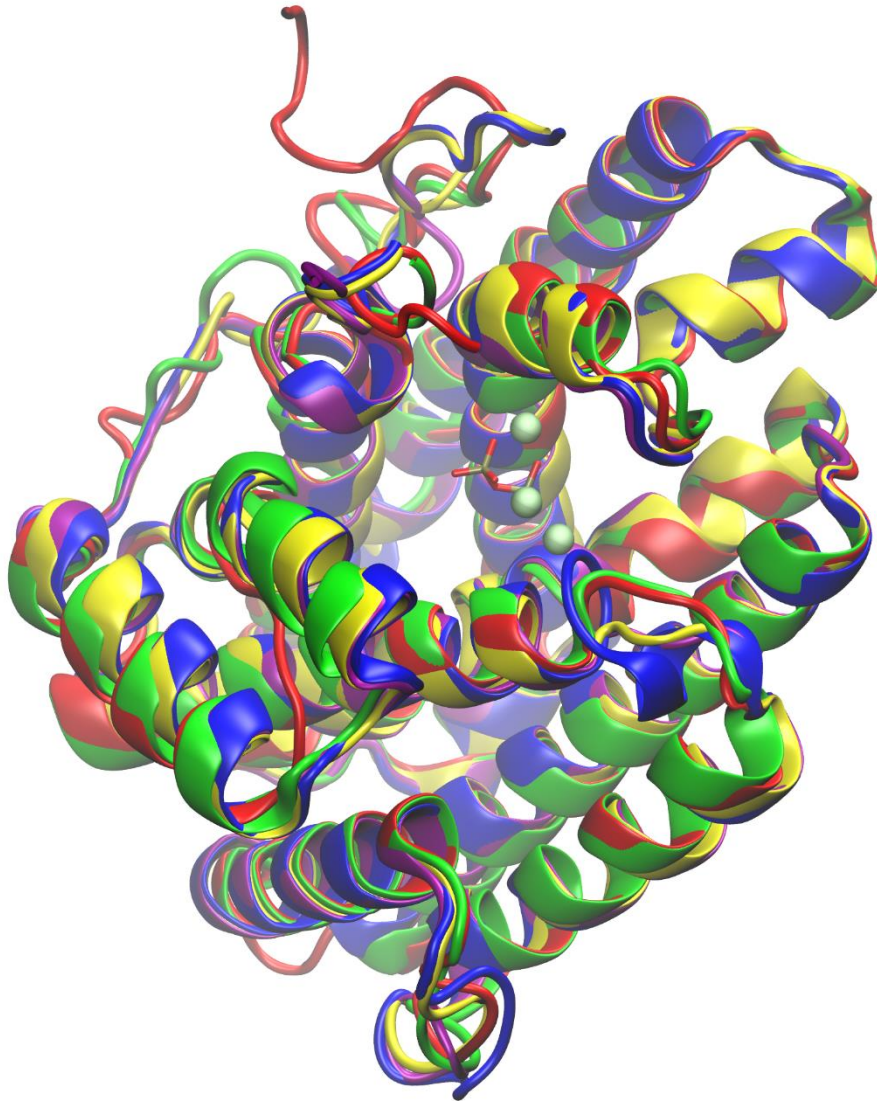
```
Cop4 --MRPTARQFTLPDLFSICPLQ-DATNPWYKQAAAESRAWINSYNIFTDRKRAFFIQGSN 57
Copu1 -MGLPQATEFVLPDFFAPCPFTLGLTNPHADTVFPEARAWIGKFLPFS----- 47
Copu2 -MSTMDPSEFILPDFFATCPFAFGRTNPHADVVIPEAHAWIVKHVPFVDRKRDEFIQDGF 59
Copu3 MSATPAPTEFILPNLFSVCLTFGRSNPYDEVIPEARAWIAKYNPFVDSKRAEFVQGCN 60
      :* **:::* ** . :** . . *::*** .. *
                               D(D/E)xxD
Cop4 ELLCSHVYAYAGYEQFRTCCDFVNLFFVDEISDDQNGQDARATGRIFVNAMRDAHWDG 117
Copu1 -GLAWHCDPWAGKEGFRTICDFQNLFLMDELTDDEMSGEDARVVGESFIRVLKDPVSDNE 106
Copu2 QDLMPHCYPWAGKETLRTMCYLNNLLFLVDDLTDDEMSDEARGFGESFIRVLNDPAVHDS 119
Copu3 ELLCSRVPYAGREEFRTCCDFVNLFFVDELSDDMGGADARSTCDSEIRVLNDPDPDT 120
      * : : ** * : ** * : * ** : * : * : * : * : * : * : * : * : * : * : * :
Pyrophosphate sensor (N/D)Dxx(S/T)xxxE (NSE)
Cop4 RRQNSAVLLCYSLVEYILGIDLDDVEYEDPTFAKAYWAACDFVCWANDVYSYDMEQAKGH 235
Copu1 RRENSAVRCCFSINEHALGIDLPSVFEDEPEFLRMFYDAVDMIVIVNDVYSYNMEQAKGL 226
Copu2 RRENSAVMVFFAITEYALGIDLPAVYEDPTFLRVYADSADMVILVNDVFSYNREQAKGL 239
Copu3 RRENSAVRCCCLALAEYALGELPDVAFNDPAFQSVYFCAADMVCWSNDVYSYNMEQAKGH 238
**::*** : : * : **::* * *::** * * : * : : ***:***: *****
TGNNVVTVLMKEKDLQSLQEAASYIGRECEKQMRDYLEAKSQLLQ-STDLPQEAURYIEAL 294
Copu1 AGNNVVTVLEQALGVDLQAAVNRGGEMFAQKMEGYVRGRGVPPSWGAKVDADVEHFFDSV 286
Copu2 DGNNNITVLMQTLDLDLQAAVDHVGEMFSQKMEGCMRGRAMLPWGVKVDADVERFFDAL 299
Copu3 TGNNVVTVLMQEHGIDLQAAADRVEVFVGQLMEHYTSGRSRSLPTWGGKVDADAARFLEAA 298
*** :*** : :. ** * : * . : * . . . : . : : : :
RY-dimer
Cop4 GYWMVGNLWVSFESQRYFGAQHERVKATHVVHLRPSSVLEASCDSDSDDC 345
Copu1 DQWIVGNLEWSTETSRYLGPDHEEIMKTGRVVLTKIESKTK----- 327
Copu2 DQWVGNLEWSSQSPRYLGPEHEEIMRTRRVVLRKVETEIE----- 340
Copu3 GQWVGNLEWSEFETPRYFGPDHDEVRDTHRVLTK----- 332
. * : *** ** : : * : * : * : * : * : *
```

**Figure S1:** Alignment of Copu1-3 and Cop4 (as reference) amino acid protein sequences. Typical signature motifs for class I terpene cyclases such as the highly conserved NSE-triad and the aspartate-rich D(D/E)xxD motif are highlighted. Additionally, a highly conserved arginine residue 46 AA downstream of the NSE motif, which is indicated as pyrophosphate sensor and a common conserved RY-dimer ~80 AA downstream of the NSE motif, close to the C-terminus is found in all four sequences [8–11].

**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## Enzyme modelling

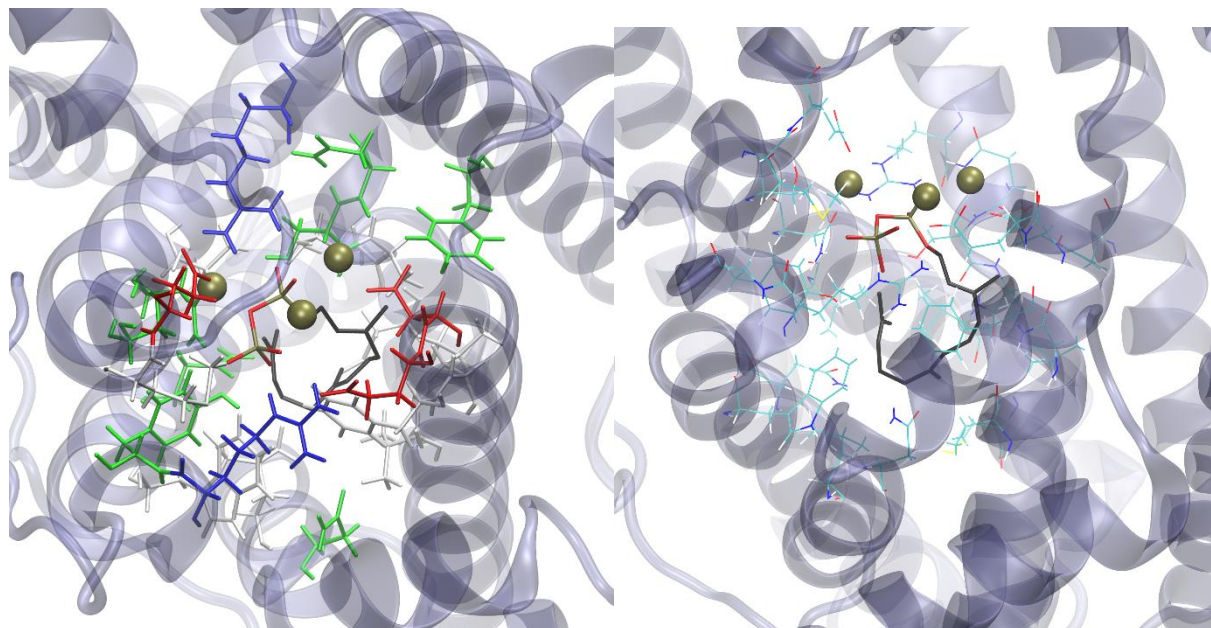


**Figure S2:** Homology modeling and structural alignment of selinadiene cyclase (blue), Copu1 (red), Copu2 (green), Copu3 (yellow) and Cop4 (purple).

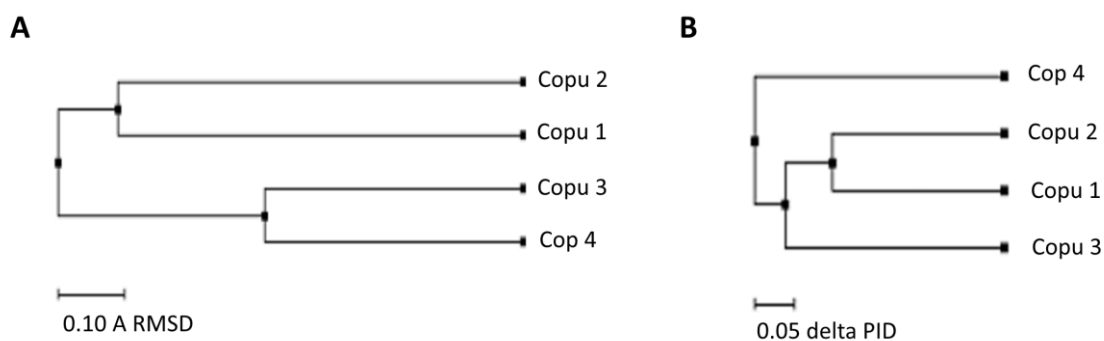


**Additional file**

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018



**Figure S3:** Structure function analyses of *copu3* revealed reliable positioning of FPP in the active sites cleft. Amino acid sidechains in close proximity (5 Angstrom) to the docked FPP molecules (black) are shown. Sidechains are colored based on their physico-chemical characteristics (hydrophobic - white, alkaline - blue, acidic - red, polar - green).



**Figure S4:** Structural identity analysis (RMSD calculation) (A) and sequence identity analyses (B) of the *Copu1-3* and *Cop4* sequences. Sequence identity analysis strictly differentiate sequences from distinct organisms while the structural analysis (RMSD calculation) revealed a close structural relation between *Cop4* and *Copu3*.

### Additional file

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## Structure analysis and quantification

### NMR spectra

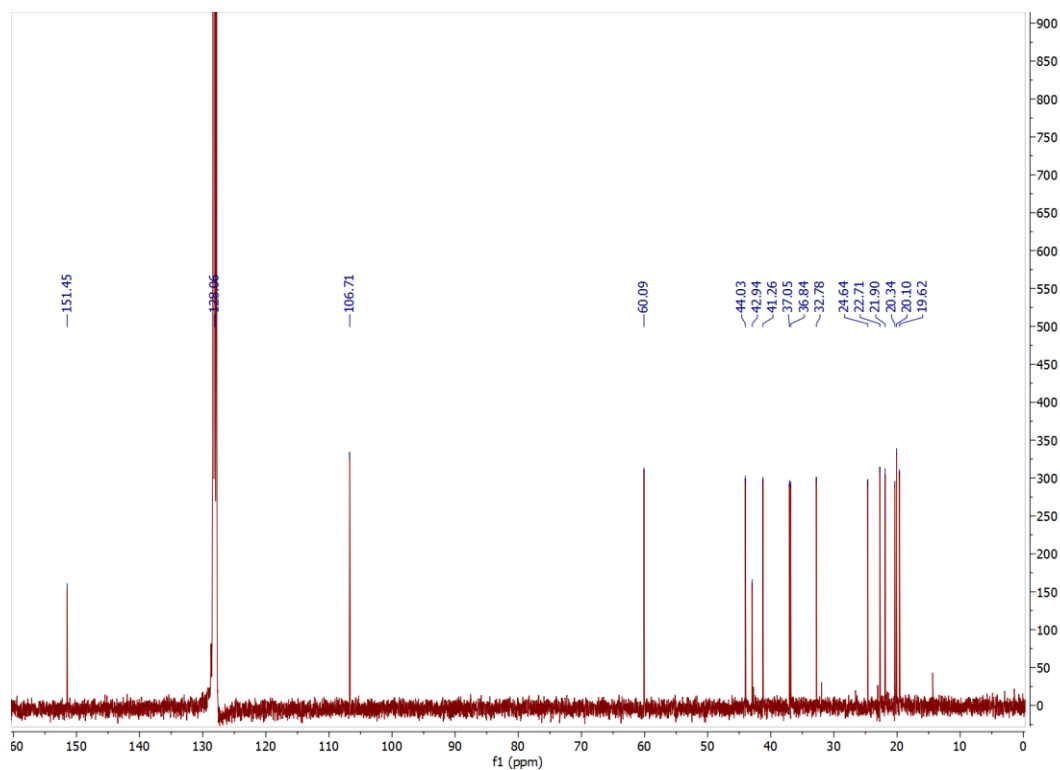


Figure S5:  $\beta$ -Copaene  $^{13}\text{C}$  NMR data recorded in  $\text{C}_6\text{D}_6$ .

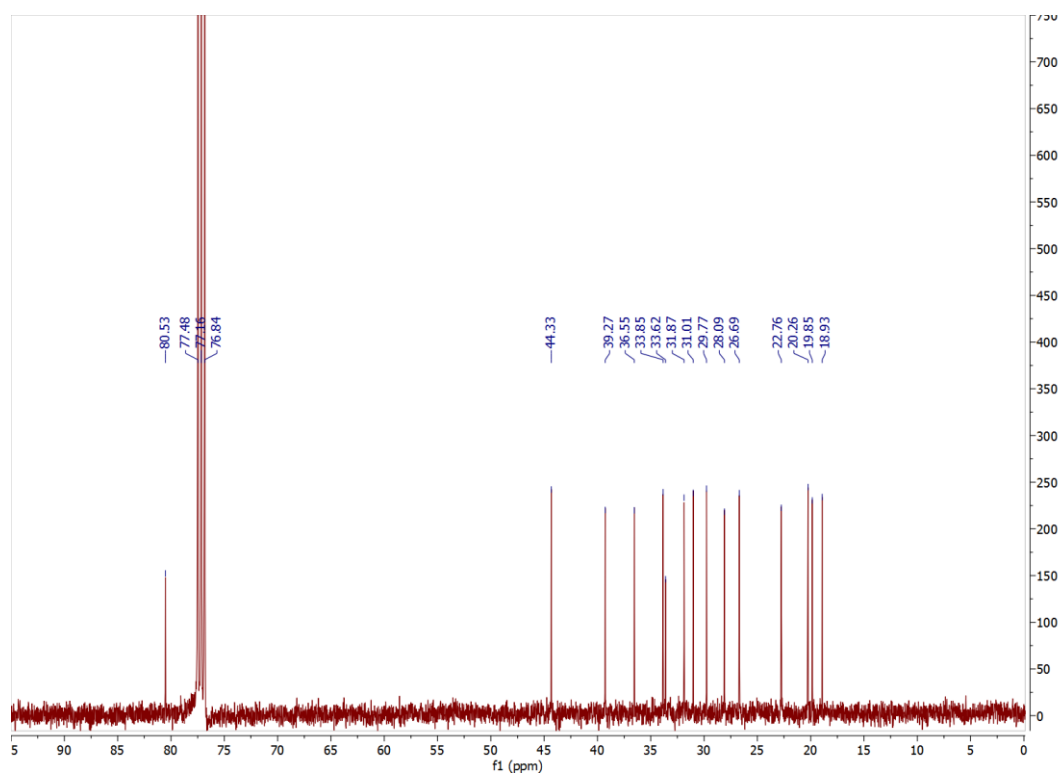


Figure S6: Cubebol  $^{13}\text{C}$  NMR data recorded in  $\text{CDCl}_3$ .

### Additional file

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

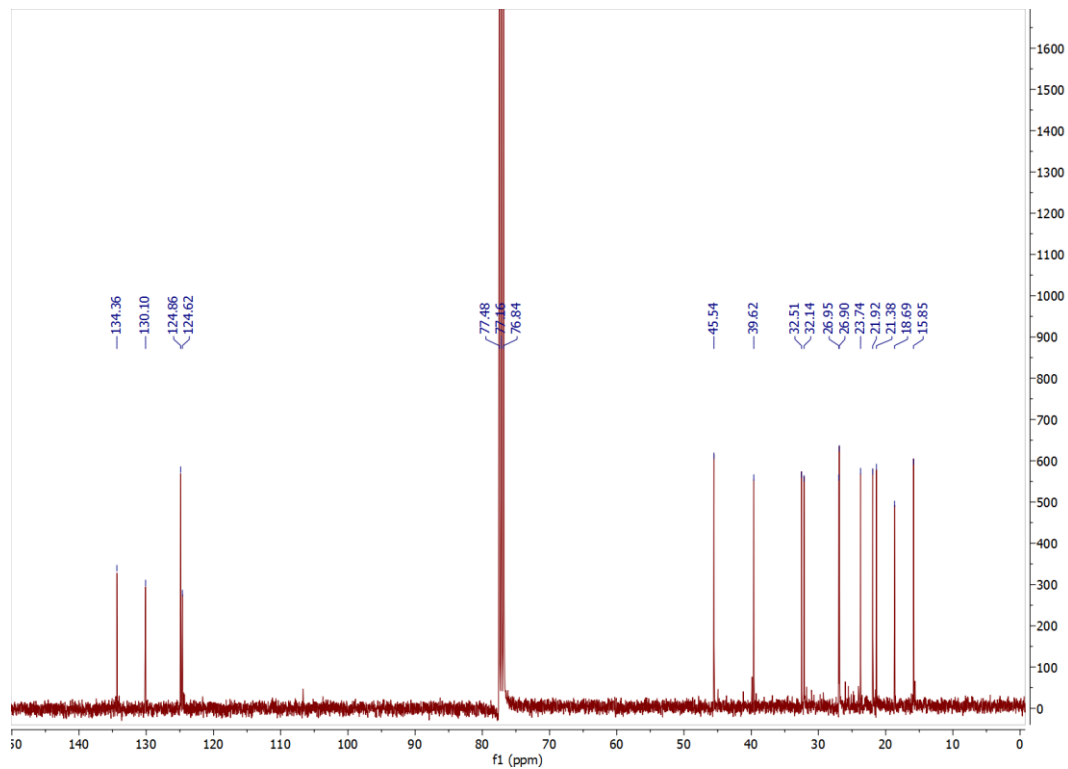


Figure S7:  $\delta$ -Cadinene  $^{13}\text{C}$  NMR data recorded in  $\text{CDCl}_3$ .

### Terpene quantification

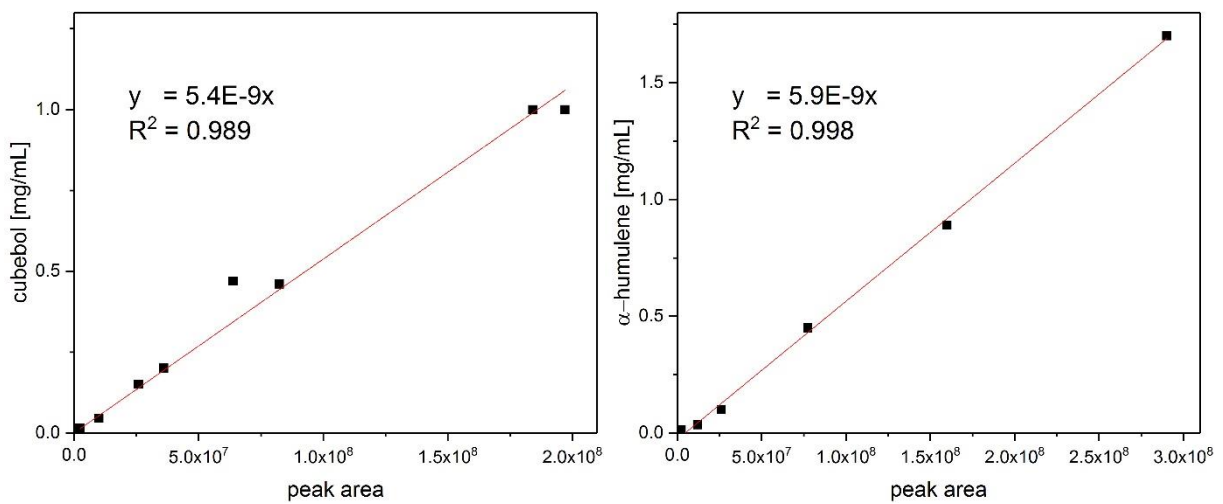


Figure S8: Calibration curves for GC-FID based sesquiterpene quantification. Different concentrations of cubebol and  $\alpha$ -humulene were used for peak-area evaluation via GC-FID.

#### Additional file

Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective  $\beta$ -copaene and cubebol production in *E. coli* – Mischko et al. 2018

## References

1. Lin Y-L, Ma L-T, Lee Y-R, Shaw J-F, Wang S-Y, Chu F-H. Differential Gene Expression Network in Terpenoid Synthesis of *Antrodia cinnamomea* in Mycelia and Fruiting Bodies. *J Agric Food Chem.* 2017;65:1874–86. doi:10.1021/acs.jafc.6b05386.
2. Engels B, Heinig U, Grothe T, Stadler M, Jennewein S. Cloning and characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. *J Biol Chem.* 2011;286:6871–8. doi:10.1074/jbc.M110.165845.
3. Zhou H, Yang Y-L, Zeng J, Zhang L, Ding Z-H, Zeng Y. Identification and Characterization of a  $\delta$ -Cadinol Synthase Potentially Involved in the Formation of Boreovibrins in *Boreostereum vibrans* of Basidiomycota. *Nat Prod Bioprospect.* 2016;6:167–71. doi:10.1007/s13659-016-0096-4.
4. Agger S, Lopez-Gallego F, Schmidt-Dannert C. Diversity of sesquiterpene synthases in the basidiomycete *Coprinus cinereus*. *Mol Microbiol.* 2009;72:1181–95. doi:10.1111/j.1365-2958.2009.06717.x.
5. Wawrzyn GT, Quin MB, Choudhary S, López-Gallego F, Schmidt-Dannert C. Draft genome of *Omphalotus olearius* provides a predictive framework for sesquiterpenoid natural product biosynthesis in Basidiomycota. *Chem Biol.* 2012;19:772–83. doi:10.1016/j.chembiol.2012.05.012.
6. Yap H-YY, Muria-Gonzalez MJ, Kong B-H, Stubbs KA, Tan C-S, Ng S-T, et al. Heterologous expression of cytotoxic sesquiterpenoids from the medicinal mushroom *Lignosus rhinocerotis* in yeast. *Microb Cell Fact.* 2017;16:103. doi:10.1186/s12934-017-0713-x.
7. Quin MB, Flynn CM, Wawrzyn GT, Choudhary S, Schmidt-Dannert C. Mushroom hunting by using bioinformatics: Application of a predictive framework facilitates the selective identification of sesquiterpene synthases in basidiomycota. *Chembiochem.* 2013;14:2480–91. doi:10.1002/cbic.201300349.
8. Dickschat JS. Bacterial terpene cyclases. *Nat Prod Rep.* 2016;33:87–110. doi:10.1039/c5np00102a.
9. Rabe P, Schmitz T, Dickschat JS. Mechanistic investigations on six bacterial terpene cyclases. *Beilstein J Org Chem.* 2016;12:1839–50. doi:10.3762/bjoc.12.173.
10. Baer P, Rabe P, Fischer K, Citron CA, Klapschinski TA, Groll M, Dickschat JS. Induced-fit mechanism in class I terpene cyclases. *Angew Chem Int Ed Engl.* 2014;53:7652–6. doi:10.1002/anie.201403648.
11. Leeper FJ, Vederas JC. *Biosynthesis: Aromatic Polyketides, Isoprenoids, Alkaloids.* Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg; Springer e-books; 2000.