1   # Enhanced production of taxadiene in *Saccharomyces cerevisiae*

2   Behnaz Nowrouzi[a,b+], Rachel Li[c,d+], Laura E. Walls[a,b+], Leopold d'Espaux[c,d], Koray Malci[a,b] , Liang

3   Lungang [a,b], Nestor Jonguitud Borrego[a,b], Albert I. Lerma Escalera[e], Jose R. Morones-Ramirez[e],

4   Jay D. Keasling[c,d,f-h], Leonardo Rios Solis[a,b*]

5

6   *[a]Institute for Bioengineering, School of Engineering, University of Edinburgh, King's Buildings,*

7   *Edinburgh, United Kingdom*

8   *[b]Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, King's Buildings,*

9   *Edinburgh, United Kingdom*

10   *[c]DOE Joint BioEnergy Institute, Emeryville, California, CA 94608, USA*

11   *[d]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley,*

12   *California, CA 94720, USA*

13   *[e]Centro de Investigación en Biotecnología y Nanotecnología, Facultad de Ciencias Químicas,*

14   *Universidad Autónoma de Nuevo León, Apodaca, Mexico*

15   *[f]Departments of Chemical & Biomolecular Engineering and of Bioengineering, University of*

16   *California, Berkeley, Berkeley, CA 94720*

17   *[g]Center for Biosustainability, Danish Technical University, Lyngby, Denmark*

18   *[h]Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes for*

19   *Advanced Technologies, Shenzhen, China*

20

24 ⁺These authors contributed equally

25 *Corresponding author: Dr Leonardo Rios Solis (E -mail: Leo.Rios@ed.ac.uk; Telephone:

26 +441316513561)

Highlights

27

28 - Maximum taxadiene titre of 129 ± 15 mg/L in *Saccharomyces cerevisiae* at 20 °C

29 - Integrating fusion protein tagged- taxadiene synthase improved taxadiene titre.

30 - Consistent taxadiene titres were achieved at the micro-and mini-bioreactor scales.

Abstract

31

32 Cost-effective production of the highly effective anti-cancer drug, paclitaxel (Taxol®), remains

33 limited despite growing global demands. Low yields of the critical taxadiene precursor remains a

34 key bottleneck in microbial production. In this study, the key challenge of poor taxadiene synthase

35 (*TASY*) solubility in *S. cerevisiae* was revealed, and the strains were strategically engineered to

36 relieve this bottleneck. Multi-copy chromosomal integration of *TASY* harbouring a selection of

37 fusion solubility tags improved taxadiene titres 22-fold, up to 57 ± 3 mg/L at 30 °C at shake flask

38 scale. The scalability of the process was highlighted through achieving similar titres during scale

39 up to 25 mL and 250 mL in shake flask and bioreactor cultivations, respectively. Maximum

40 taxadiene titres of 129 ± 15 mg/L and 119 mg/L were achieved through shake flask and bioreactor

41 cultivation, respectively, of the optimal strain at a reduced temperature of 20 °C. The results

42 highlight the positive effect of coupling molecular biology tools with bioprocess variable

43 optimisation on synthetic pathway development.

44

Abbreviations

45

46

47 *BTS1*, Geranylgeranyl diphosphate synthase; *crtE*, Geranylgeranyl diphosphate synthase;

48 DMAPP, Dimethylallyl pyrophosphate; *ERG8*, Phosphomevalonate kinase; *ERG9*, Farnesyl-

49   diphosphate farnesyl transferase (squalene synthase); *ERG10*, 3-hydroxy-3-methylglutaryl-CoA

50   (HMG-CoA) synthase; *ERG12*, Mevalonate kinase; *ERG13*, 3-hydroxy-3-methylglutaryl-CoA

51   (HMG-CoA) synthase; *ERG19*, Mevalonate pyrophosphate decarboxylase; *ERG20*, Farnesyl

52   pyrophosphate synthetase; FPP, Farnesyl diphosphate; GGOH, (E,E,E)-geranylgeraniol; GGPP,

53   (E,E,E)-Geranylgeranyl diphosphate; GPP, Geranyl diphosphate; *HMG1*, 3-hydroxy-3-

54   methylglutaryl-coenzyme   A reductase 1;   *HMG2*,   3-hydroxy-3-methylglutaryl-coenzyme

55   A reductase 2; *IDI*, Isopentenyl-diphosphate delta-isomerase; *IPP*, Isopentenyl pyrophosphate;

56   MBP,   Maltose   binding   protein;   *mvaE*,   Acetyl-CoA   acetyltransferase;   *mvaS*,

57   Hydroxymethylglutaryl-CoA synthase; MVA pathway, Mevalonate pathway; *TASY,* Taxadiene

58   synthase

59

60

61   **1. <u>Introduction</u>**

62   The highly complex diterpenoid drug Paclitaxel (Taxol™) first gained FDA approval in 1992

63   for the treatment of ovarian cancer and has since proven efficacious against a wide range of

64   additional diseases (McElroy and Jennewein, 2018). Direct extraction from its natural source, the

65   bark of Pacific yew (*Taxus brevifolia*), is both destructive and extremely low-yielding. As a result,

66   paclitaxel is currently produced predominantly by semi-synthesis, involving the chemical

67   modification of late precursors extracted from plant cell culture. However, as such methods are

68   high in cost and have limited scalability, the development of a more sustainable source is critical

69   to meet growing global demands (Frense, 2007). One potential solution involves the heterologous

70   expression of the biosynthetic pathway in microbial cell factories. The first committed step in the

71   paclitaxel biosynthetic pathway is the cyclisation of the diterpenoid intermediate, geranylgeranyl

72   diphosphate (*GGPP*) by taxadiene synthase (*TASY*), yielding taxa-4(5),11(12)-diene (taxadiene)

73   as shown in Figure 1.

74   TASY is comprised of three alpha-helical domains; the active site is located within the C-

75   terminal catalytic domain, where GGPP binds and is activated via a cluster of three Mg^{2+} ions

76   (Köksal *et al.*, 2011). TASY enzyme activity has been found to be relatively low compared to

77   other terpene synthases, with a 70-fold lower turnover rate than that of a plant abietadiene synthase

78   with high sequence homology (Jin *et al.*, 2005). In addition, the volume of the TASY active site is

79   larger than that of taxadiene, contributing to previously observed enzyme promiscuity (Soliman
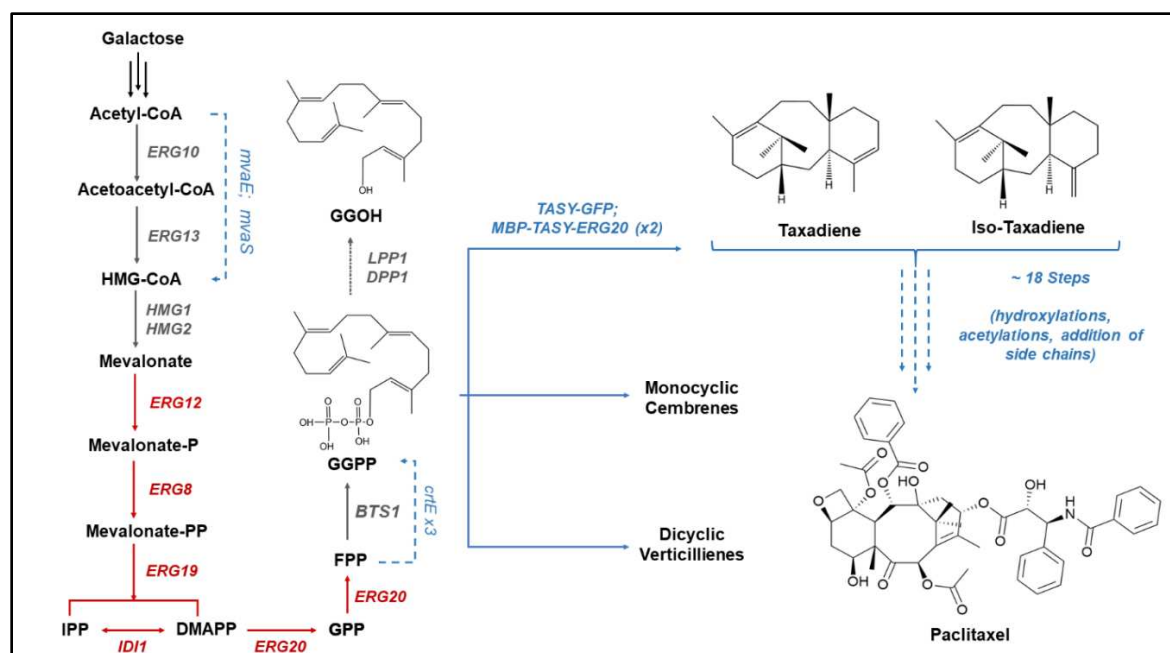
80   and Tang, 2015).

Figure 1: **Engineered taxadiene biosynthetic pathway in *S. cerevisiae*.** Genes highlighted in red represent native genes which have been overexpressed through the integration of one additional copy. Genes highlighted in blue are exogenous genes which were heterologously expressed. Hydroxymethylglutaryl-CoA synthase (*mvaS*) and Acetyl-CoA acetyltransferase (*mvaE*) from *Enterococcus faecalis*. Geranylgeranyl diphosphate synthase (*crtE*) from *Xanthophyllomyces dendrorhous* and Taxadiene synthase (*TASY*) from *Taxus cuspidata*.

The reconstitution of this enzymatic step has been successfully achieved in both *E. coli* (Huang *et al.*, 1998; Ajikumar *et al.*, 2010) and *S. cerevisiae* (Engels, Dahm and Jennewein, 2008; Reider Apel *et al.*, 2017). Through the adoption of a novel multivariate modular approach, Ajikumar *et al.* (2010) improved heterologous taxadiene titres to 300 and 1020 mg/L in *E. coli* shake-flask and fed-batch bioreactor cultivations, respectively. However, when the authors expressed the subsequent enzyme, taxadien-5α-hydroxylase, which is a membrane bound cytochrome P450, a 10-fold reduction in total taxane titre was observed. Membrane bound cytochrome P450s like taxadien-5α-hydroxylase are estimated to comprise around half of the 19 enzymatic steps in the paclitaxel biosynthetic pathway (Kaspera and Croteau, 2006). As the overexpression of such membrane bound enzymes is greatly hindered in *E. coli* (Biggs *et al.*,

92   2016), the construction of the remainder of the pathway is likely to be very challenging in this

93   bacterial host.

94   The eukaryotic host, *S. cerevisiae*, on the other hand, possesses the necessary biosynthetic

95   machinery for the expression of such enzymes, including translocation through the endoplasmic

96   reticulum and a native electron transfer machinery (Delic *et al.*, 2013). As in *E. coli*, early attempts

97   to express *TASY* in *S. cerevisiae* were hindered by GGPP availability with a titre of just 122 μg/L

98   in the wild type strain, which was insufficient for taxadiene synthesis (Engels, Dahm and

99   Jennewein, 2008). DeJong *et al.* (2006) simultaneously incorporated GGPP synthase (*GGPPS*)

100   and *TASY* from *Taxus* sp. into *S. cerevisiae*, leading to a taxadiene titre of 1 mg/L. Through the

101   heterologous expression of a *Sulfolobus acidocaldarius* geranylgeranyl diphosphate synthase

102   (*GDS*), *TASY*, and a truncated HMG-CoA, taxadiene titres were later improved to 8.7 mg/L

103   (Engels, Dahm and Jennewein, 2008). A subsequent study focussed on the careful selection of

104   promoters, integration locus, and solubility tags, leading to the highest reported taxadiene titre of

105   20 mg/L in yeast (Reider Apel *et al.*, 2017).

106   Although taxadiene has been found to be the major product of TASY, with yields over 77%,

107   around 5-13% of the total taxane product has been found to be the isomer taxa-4(20),11(12)-diene

108   (iso-taxadiene) (Williams *et al.*, 2000; Huang *et al.*, 2001; Sagwan-Barkdoll and Anterola, 2018).

109   Small quantities of a product tentatively identified as verticillene and an additional taxadiene

110   isomer (taxa-3(4),11(12)-diene) have also been detected in *E. coli* and *Nicotiana benthamiana*

111   (Sagwan-Barkdoll and Anterola, 2018; Li *et al.*, 2019).

112   The metabolic pathway of paclitaxel is highly complex and development of an alternative

113   recombinant production route remains in the preliminary stages. Despite this, substantial

114   advancements in synthetic biology have been achieved recently (Hsu, Lander and Zhang, 2014).

3

115   Through the application of such tools, there is great potential to accelerate the development of a

116   microbial paclitaxel biosynthetic pathway. This study focussed on the optimisation of TASY

117   enzyme performance in *S. cerevisiae* to alleviate a key early pathway bottleneck. The effect of a

118   number of factors including TASY truncation length, selected promoter and chromosomal gene

119   copy number on pathway expression were evaluated. Cultivation conditions such as culture

120   temperature and exogenous cofactor availability were also considered.

121   **2.**   **<u>Materials and methods</u>**

122   2.1.<u>Yeast strains and media</u>

123   The parent *S. cerevisiae* strain used for episomal expression and integration studies was

124   GTy116   (MATa, *leu2-3,*   *112::HIS3MX6-GAL1p-ERG19/GAL10p-ERG8;*   *ura3-52::URA3-*

125   *GAL1p-MvaSA110G/GAL10p-MvaE* (codon-optimised); *his3Δ1::hphMX4-GAL1p-*

126   *ERG12/GAL10p-IDI1;trp1-289::TRP1_pGAL1-CrtE(X.den)/GAL10p-ERG20;*

127   *YPRCdelta15::NatMX-GAL1p-CrtE(opt)/GAL10p-CrtE*) described previously by Reider Apel *et*

128   *al.* (2017), originating from CEN.PK2-1C (EUROSCARF collection). The *URA3* marker of this

129   strain was further restored to give mGty116. All chemicals and reagents were sourced from Sigma-

130   Aldrich at the highest available purity unless otherwise stated. Episomal expression systems made

131   use of synthetic defined medium minus uracil (CSM-Ura and CSM-Leu, Sunrise Science Products)

132   depending on the selection marker used. These were supplemented with 2% *(w/v)* glucose (SDD-

133   Leu), 2% *(w/v)* galactose (SDG-Leu) or a 1.8% *(w/v)* galactose, 0.2% *(w/v)* glucose mixture

134   (SDGD-Leu). For cultivation of strains harbouring chromosomally integrated genes, a medium

135   containing yeast extract (1%*(w/v)*) and peptone (2 %*(w/v)*), supplemented with 2 % *(w/v)* glucose

136   (YPD), galactose (YPG) or 1.8% *(w/v)* galactose and 0.2% *(w/v)* glucose mixture (YPGD) were

137   used.

138   2.2. Yeast transformation and strain construction

139   Episomal expression was achieved through transforming *S. cerevisiae* with high copy 2-micron

140   plasmids harbouring *TASY* and a *LEU2* selection marker (Supplementary Table 1). The standard

141   LiAc/SS carrier DNA/PEG method was used for all transformations (Gietz and Schiestl, 2007).

142   Chromosomal integration was performed using a cloning free, Cas9-mediated homologous

143   recombination method (Reider Apel *et al.*, 2017). High copy number, Cas9-sgRNA 2-micron

144   plasmids harbouring a *URA3* selection marker derived from pRS426 were used. The gene cassettes

145   were designed using an online tool, CASdesigner. Chromosomal integration at the target site was

146   confirmed by colony PCR and Sanger sequencing (Genewiz Inc., USA; Edinburgh Genomics,

147   UK). Plasmid curing was subsequently performed on successful colonies through sequential

148   culture on YPD agar medium until no growth was observed on concurrent SDD-Leu agar plates.

149   The strains used in this study are summarised in Table 1 and plasmid, primer and tag sequences

150   are tabulated in Supplementary Tables 1-3. All DNA sequences were synthesised by IDT

151   (Integrated DNA Technologies, Inc.). A *TASY* sequence from *Taxus cuspidata* was codon

152   optimised for expression in *S. cerevisiae*. DNA amplification was performed using Phusion Flash

153   High-Fidelity PCR Master Mix (Thermo Fisher Scientific). Promoter and terminator sequences

154   native to the yeast genome were selected (Lee *et al.*, 2015).

155

156   Table 1: **Yeast strains used in this study.**

Strain name	Genotype	Reference
mGTy116	MATa, *leu2-3, 112::HIS3MX6-GAL1p-ERG19/GAL10p-ERG8; ura3-52::URA3-GAL1p-MvaSA110G/GAL10p-MvaE* (codon	(Reider Apel *et al.*, 2017)

5

	optimised); *his3Δ1::hphMX4-GAL1p-ERG12/GAL10p-IDI1; trp1-289::TRP1_GAL1p-CrtE(X.den)/GAL10p-ERG20; YPRCdelta15::NatMX-GAL1p-CrtE(opt)/GAL10p-CrtE*	
LRS1	mGTy116 *Yep_Gal1p-3Xflag- M60TASYBRE-CYC1*	This study
LRS2	mGTy116 *ARS1014a::TDH3p-TASY-ADH1t*	(Reider Apel *et al.*, 2017)
LRS3	mGTy116 *ARS1014::Gal1p-TASY-GFP*	This study
LRS4	LRS3 *ARS1622:: Gal1p-MBP-TASY-ERG20**	This study
LRS5	LRS4 *ARS1114a::Gal1p-MBP-TASY-ERG20**	This study

157

158   2.3. Optimisation of taxadiene production at microscale

159   Preliminary microscale episomal expression studies involved 5 ml cultures of the mGTy116

160   strain harbouring selected high-copy 2-micron plasmids as shown in Supplementary Table 1.

161   Inocula were prepared through transferring a single yeast colony to 5 ml of SDD-Leu medium,

162   followed by incubation overnight at 30 $^\circ$C. Aliquots of the inoculum were diluted with SDGD-Leu

163   medium to give an initial OD_{600} = 0.5. The yeast cultures were incubated at 20 or 30 °C and 200

164   rpm for 96 hours. To minimise air-stripping of the volatile terpene products, a biphasic liquid-

165   liquid system was used through the addition of a 10 % *(v/v)* dodecane overlay after 24 hours of

166   cultivation. At the end of the cultivation, the resulting cultures were centrifuged and the organic

167   dodecane layer was extracted for GC-MS analysis. Microscale culture of the strains harbouring

168   chromosomally integrated genes was done in a similar way as for episomal expression, except that

169   YPD and YPGD media were used for the inoculum and production cultures, respectively.

170

171   2.4.<u>Shake flask cultivations</u>

172   Single LRS5 colonies were transferred from YPD agar to 5 mL YPD medium and incubated

173   at 30 °C and 250 rpm overnight. Aliquots of these cultures were then used to inoculate 20 mL

174   YPG in 250 mL shake flasks to an OD_{600} = 1. A 20% *(v/v)* dodecane overlay was used giving a

175   total working volume of 25 mL. The resulting cultures were incubated at 20 or 30 °C and 250 rpm

176   for 72 hours. At the end of the cultivation, the dodecane overlay was extracted for GC-MS analysis.

177   The final biomass was measured at an optical density of 600 nm.

178   2.5. <u>Batch culture in the bioreactor</u>

179   Cultivations were conducted in MiniBio 500 mL bioreactors

180   (Applikon Biotechnology, The Netherlands) with a working volume of 250 mL. Pre-inocula were

181   prepared by incubating cells in 5 mL of YPD for eight hours. The resulting cultures were used to

182   inoculate secondary 10-mL inocula to an OD_{600} = 1, which were subsequently incubated at 30 °C

183   and 200 rpm overnight. An aliquot of an inoculum culture was then diluted with YPG to give a

184   200 mL culture with an initial OD_{600} = 1. To prevent excess foam production, polypropylene

185   glycol P2000 (Fisher Scientific, UK) was added to a concentration of 0.01 % *(v/v)* and a Rushton

186   turbine was placed at the medium-air interface. A 20 % *(v/v)* dodecane (Fisher Scientific, UK)

187   overlay was also added to minimise product loss due to air stripping. During cultivation, the

188   temperature, dissolved oxygen and pH were monitored online. Biomass was measured through

189   manual sampling twice daily. The adaptive my-Control system

190   (Applikon Biotechnology, The Netherlands) was used to control process parameters. Setpoints of

191   30 % of the saturation and 30 °C were applied for dissolved oxygen and temperature, respectively.

192   The culture pH was controlled to a setpoint of six through automatic addition of

193   1M NaOH. Samples were taken twice daily for analysis of taxane and biomass concentration.

7

194   2.6.Diterpene analysis and quantification

195   The dodecane overlay was analyzed by GC-MS using Trace 1300 GC (ThermoFisher

196   Scientific), equipped with TG-SQC column (15 m × 0.25 mm × 0.25 μm). The mass spectra in the

197   range of 50-650 m/z was recorded on a Thermo Scientific ISQ Series single quadrupole mass

198   spectrometer using EI ionization mode and scan time of 0.204 s. The GC temperature programme

199   began at 120 °C (3 min) and was then raised to 250 °C at a rate of 20 °C/min with 3 min hold time.

200   Xcalibur™ software (ThermoFisher Scientific, USA) was employed for data processing. Pure

201   taxadiene (kindly supplied by the Baran Lab, The Scripps Research Institute) and geranylgeraniol

202   (Sigma Aldrich, UK) were used as standards to identify and quantify taxadiene and GGOH,

203   respectively. Additional taxane products were quantified relative to standard taxadiene

204   concentrations.

205   2.7. Intracellular taxadiene quantification

206   Intracellular taxadiene accumulation was quantified in the 20 °C bioreactor samples. Cell lysis

207   was achieved using a previously described protocol (Westfall *et al.*, 2012) with some

208   modifications. In summary, 1 mL of the aqueous culture phase was centrifuged and the cell pellet

209   was resuspended in a mixture containing 0.4 mL of yeast lysis reagent (2 % (*v/v*) Triton X100,1

210   % *(w/v)* SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA (pH 8.0)) and 0.4 mL

211   of 1M HCl for cell lysis. The resulting suspension was mixed thoroughly prior to the addition of a

212   1-mL dodecane overlay. The samples were subsequently incubated at 30 °C, 200 rpm for 16 hours

213   before separation of the organic phase for GC-MS analysis.

214

215

216

217   ## 3. <u>Results and discussion</u>

218   3.1.   <u>Optimisation of taxadiene titre using episomal *TASY* expression</u>

219   Four key parameters with the potential to affect final taxadiene titre were examined using

220   a high-copy (2-micron) plasmid in the mGty116 *S. cerevisiae* strain. Such parameters included the

221   selected promoter, cultivation temperature, cofactor ($Mg2^+$) concentration, and *TASY* sequence

222   truncation. The results are summarised in Figure 2.
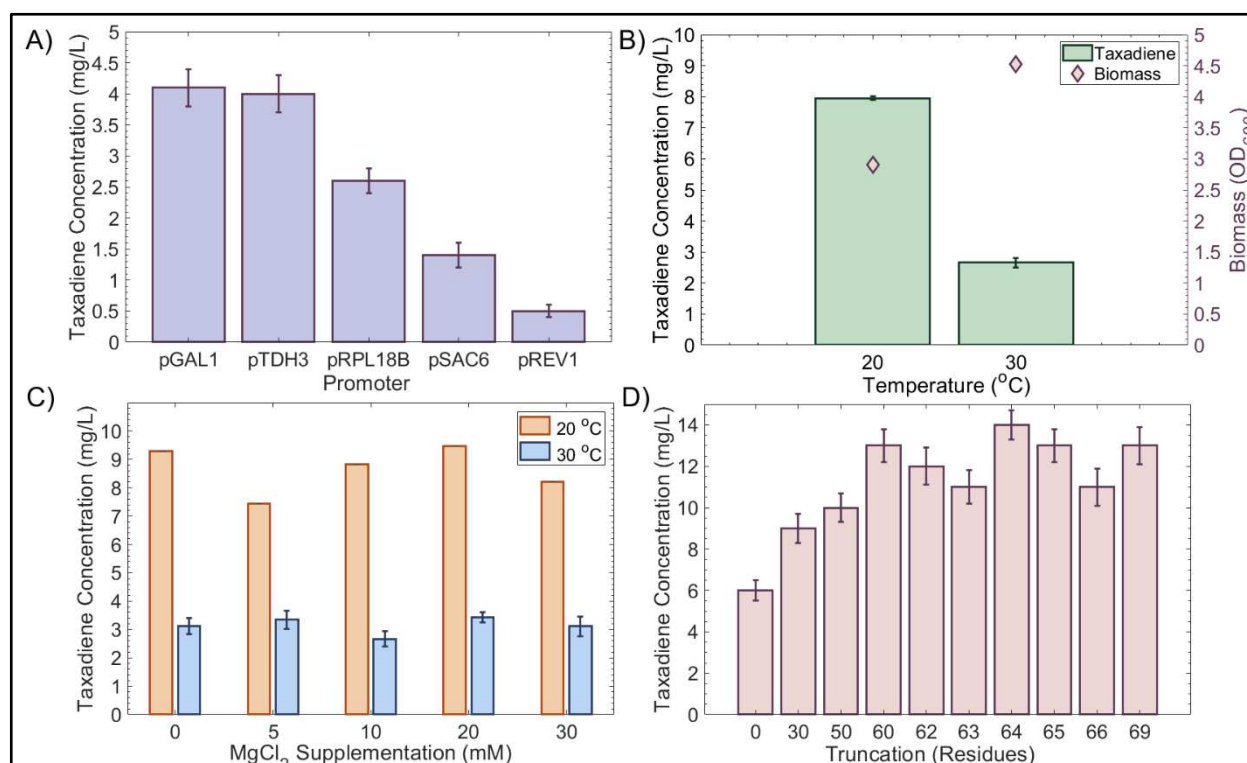
Figure 2: **Optimisation of taxadiene titres using episomal *TASY* expression**. A) Effect of promoter on taxadiene concentration produced by cells incubated at 30 °C. B) Effect of cultivation temperature on taxadiene titre. C) Effect of cofactor (Mg^{2+}) availability at 20 and 30 °C. D) Effect of gene truncation length for taxadiene production at 20 °C. Values shown are mean ± standard deviation for 96-hour triplicate cultivations.

223   In this study, the performance of four constitutive promoters of increasing strength,

224   pREV1, pSAC6, pRP118BP and pTDH3 (Lee *et al.*, 2015) were compared to that of the strong

225   inducible GAL1 promoter in *S. cerevisiae*, grown in SDGD-Leu media (Figure 2A). Performance

226   of the highest strength constitutive TDH3 and inducible GAL1 promoters was highly similar with

227   titres of 4.1 ± 0.3 and 4.0 ± 0.3 mg/L, respectively. Decreasing promoter strength did not improve

228   taxadiene titres, which were 1.6, 2.9 and 8.2-fold lower for the weaker, RPL118, SAC6 and REV1

229   promoters, respectively. As changing the promoter from the inducible GAL1 promoter resulted in

230   no further improvement in taxadiene titre (Figure 2A), pGAL1 was selected for subsequent studies.

231   The growth temperature was found to have an important effect on taxadiene titre as shown in

232   Figure 2B. Reducing the cultivation temperature from 30 °C to 20 °C resulted in a 3.0-fold increase

233   in final taxadiene titre to 8 ± 0.1 mg/L. However, at the lower temperature biomass accumulation

234   was also reduced, with a final OD_{600nm} of 2.90, compared to 4.5 at 30 °C. This was in agreement

235   with previous results where heterologous expression of *TASY* in *E. coli* (Boghigian *et al.*, 2012)

236   and *Bacillus subtilis* (Abdallah *et al.*, 2019) has also been found to be sensitive to culture

237   temperature. Productivity analysis over a broad temperature range (12-37 °C) revealed a

238   temperature optimum of 22 °C for both species (Boghigian *et al.*, 2012). In addition, the highest

239   reported titre of 1020 mg/L was achieved in *E. coli* cultures at a temperature of 22 °C (Ajikumar

240   *et al.*, 2010). Similarly, a recent study by Abdallah *et al.* (2019) showed that *TASY* expression in

241   *B. subtilis* was best at 20 °C, as opposed to 30 °C or 37 °C.

242   TASY relies on the metal ion co-factor, Mg^{2+}, for activation and substrate orientation within

243   the active site (Schrepfer *et al.*, 2016). In a recent study by Tashiro *et al.* (2018), $MgCl_2$

244   supplementation was found to improve activity of an alternative diterpene synthase, pinene

245   synthase, up to 20 mM in *E. coli*. As the concentration of Mg^{2+} in the SDGD-Leu medium was

246   around 0.8 mM, it was hypothesised to be rate limiting. However, supplementation of the

247   cultivation medium with additional $MgCl_2$ did not significantly improve taxadiene titres at 20 or

248   30 °C in this study (Figure 2C). Consequently, additional $MgCl_2$ supplementation was not deemed

249   necessary for subsequent experiments.

10

250   The native *TASY* gene encodes 862 amino acids, including a putative N-terminal sequence of

251   ~137 residues which is cleaved upon maturation in plastids (Köksal *et al.*, 2011). Removal of this

252   sequence was found to reduce inclusion body formation and increase active protein production in

253   *E. coli* (Williams *et al.*, 2000; Köksal *et al.*, 2011). In order to determine the optimal protein length

254   for TASY expression in *S. cerevisiae*, a range of truncation lengths were tested (Figure 2D).

255   Increasing truncation length up to 60 residues improved taxadiene production from 6.0 ± 0.5 to

256   13.0 ± 0.8 mg/L as shown in Figure 2D. In variants harbouring longer truncations, no further

257   improvement in taxadiene titre was observed. The *TASY* variant with a 60-residue truncation was

258   therefore selected for subsequent experiments. These results were in agreement with previous

259   works where truncations of 60 or 79 residues yielded active protein and reduced inclusion body

260   formation, whilst truncations of 93 or more residues produced inactive protein in *E. coli* (Williams

261   *et al.*, 2000).

262   3.2. Chromosomal integration of *TASY*

263   *3.2.1.   Comparing episomal and chromosomal expression*

264   Plasmid-based systems rely on the use of expensive selective media, limiting their industrial

265   relevance. In a previous study, a single copy of the *TASY* gene was chromosomally integrated into

266   the mGty116 strain. Cultivation of the resulting LRS2 (pTDH3-*TASY*-tADH1) strain at 30 °C

267   resulted in a taxadiene titre of 2.6 ± 0.4 mg/L (Reider Apel *et al.*, 2017). Interestingly, this titre

268   was comparable to the 2.7 ± 0.2 mg/L achieved using the high-copy 2-micron plasmid system in

269   this study (Figure 2B). The effect of stable chromosomal *TASY* integration in yeast was therefore

270   investigated in the mGty116 strain of this study. In order to visualise functional TASY expression,

271   *TASY* was tagged with a C-terminus GFP reporter gene (strain LRS3). Cultivation of this strain

272   yielded 12 ± 1 mg/L of taxadiene at 30 °C, 4.7-fold higher than that of the LRS2 strain. This was

11

273   also comparable to the optimal plasmid-based titre of 13 ± 1 mg/L achieved at the lower

274   temperature of 20 °C (Figure 2D). This suggested that although chromosomal integration

275   improved taxadiene titre in comparison to episomal expression, solubility remained a significant

276   bottleneck.

277   3.2.2. *Effect of gene copy number and solubility tags*

278   Fusion tags are often used to improve protein expression and solubility (Kosobokova, Skrypnik

279   and Kosorukov, 2016). A recent study using a range of tags, revealed that maltose binding protein

280   (MBP), had the greatest positive impact on TASY solubility and activity in *S. cerevisiae*

281   (Reider Apel *et al.*, 2017). Prenyltransferase-terpene synthase fusion proteins have also been found

282   to greatly improve terpene production (Ignea *et al.*, 2015; Deng *et al.*, 2016). The

283   prenyltransferase, farnesyl diphosphate synthase, encoded by *ERG20* catalyses the formation of

284   the geranyl diphosphate (GPP) and farnesyl diphosphate (FPP) metabolites, two key enzymatic

285   steps in the mevalonate pathway (Figure 1). ERG20 has been reported to significantly improve

286   isoprenoid concentration when used as a fusion protein with terpene synthases, potentially

287   promoting enzyme solubility as well as increasing the pool of the isoprenoid precursor farnesyl

288   diphosphate (Ignea *et al.*, 2015; Wong *et al.*, 2018). Deng *et al.* (2016) showed that the fusion of

289   a mutant variant of *ERG20* to (*S*)-linalool synthase improved the productivity by around 70%

290   compared to independent integration of the two genes in *S. cerevisia*e. Hence, a cassette containing

291   an N-terminal yeast codon-optimised *MBP* tagged *TASY-ERG20** (F96C; Ignea *et al.*, 2015) fusion

292   gene was developed with the dual aim of increasing precursor availability and improving TASY

293   solubility. This cassette was chromosomally integrated into locus ARS1014a of LRS3 to yield

294   strain LRS4. The integration of a second copy of this dual-tagged *TASY* in locus ARS1622b

295   resulted in strain LRS5 (1x *TASY-GFP*, 2x *MBP-TASY-ERG20**, Figure 3).

12

296

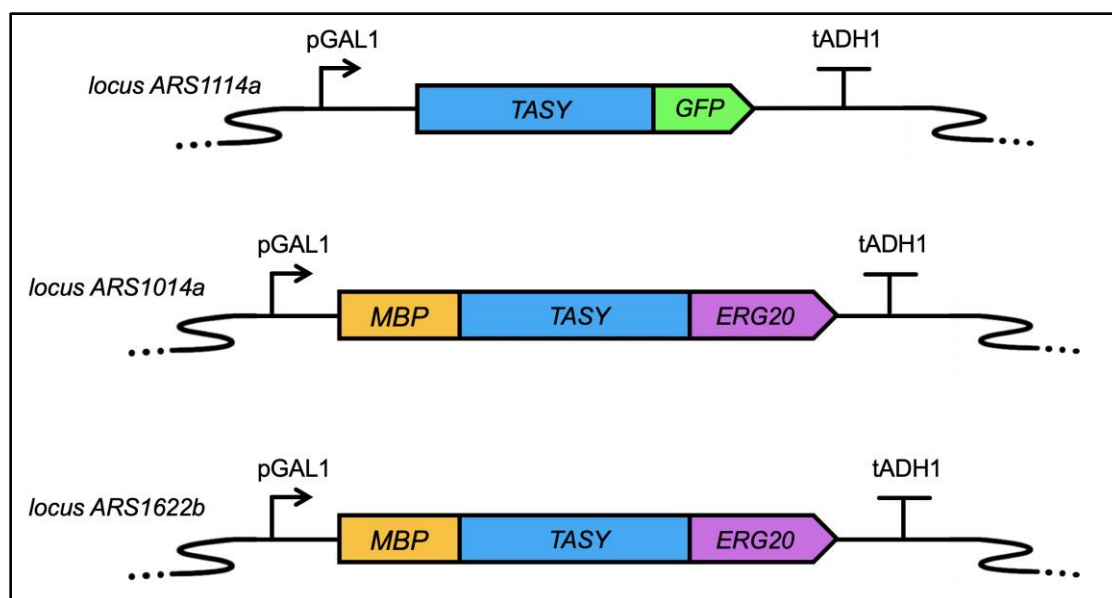

Figure 3: **Chromosomally integrated *TASY* genetic constructs**. From top to bottom are *TASY* cassettes integrated at loci: ARS1114a, ARS1014a, and ARS1622b, respectively. The genetic designs were visualized by SBOLDesigner 2 (Zhang et al., 2017).

297   Cultivation of the newly constructed LRS4 strain expressing two copies of *TASY* resulted in

298   an over three-fold improvement in taxadiene titre to 40 ± 3 mg/L at microscale. Through

299   cultivation of the subsequent LRS5 strain, taxadiene titres were further enhanced to 57 ± 3 mg/L

300   at 30 $^{\circ}$C, the highest titre in yeast reported to date at microscale level (Figure 4).
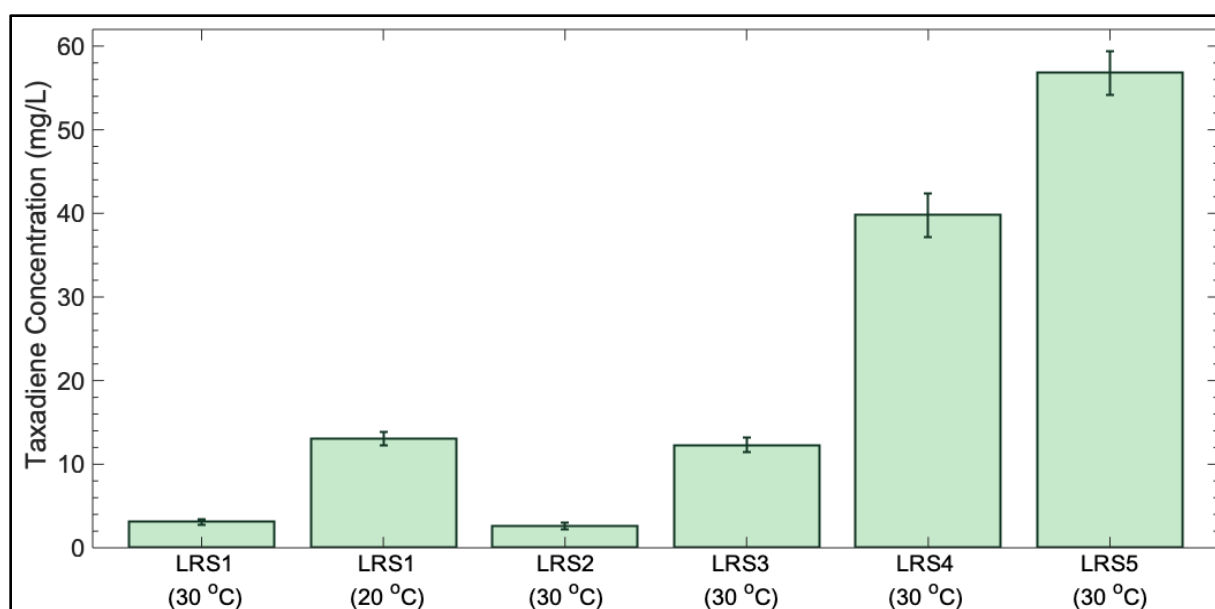
Figure 4: **Summary of taxadiene titers microscale optimisation study**. LRS1 indicates titers resulting from episomal *TASY* expression using the high copy 2-micron plasmid (LRS1) at 30 and 20 °C, respectively. The other titers are from strains with chromosomally integrated *TASY* variants: LRS2 (TDH3p-*TASY*-ADH1t), LRS3 (*TASY-GFP*), LRS4 (*TASY-GFP*; *MBP-TASY-ERG20**) and LRS5 (*TASY-GFP*; *MBP-TASY-ERG20**; *MBP-TASY-ERG20**), respectively, grown at 30 °C.

301   Through effective copy number optimisation and the use of MBP and GFP tags along with an

302   ERG20 fusion, taxadiene production was enhanced up to 22-fold in *S. cerevisiae* (Figure 4). This

303   was likely the result of improved protein expression and solubility, as well as precursor

304   availability. Nevertheless, fluorescent visualisation still showed substantial spotted subcellular

305   localization (Supplementary Figure 1), which was consistent with a previous study highlighting

306   the poor solubility of TASY (Reider Apel *et al.*, 2017).

307   3.3.Taxadiene production scale-up

308   *3.3.1.   Minor product characterisation at increased scale under different temperatures*

309   In order to investigate process scalability, the optimised LRS5 strain was subsequently

310   cultivated in 250-mL shake flasks (25 ml working volume). The effect of temperature on

311   chromosomal *TASY* expression by the optimised strain was also assessed through cultivation at

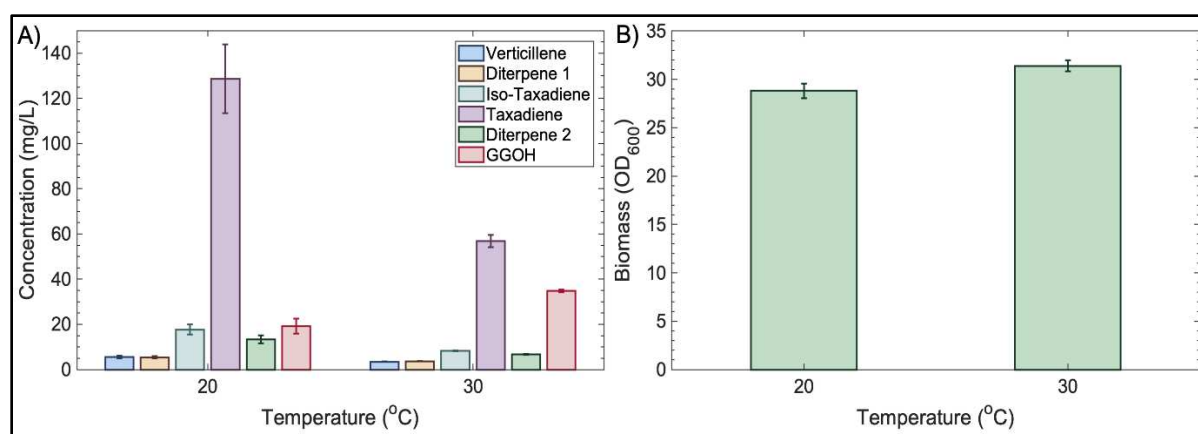312   both 20 °C and 30 °C. The results of this experiment are summarised in Figure 5.

14

313

Figure 5: **Effect of temperature on LRS5 performance in shake flasks.** LRS5 was cultivated in 250 mL shake flasks in YPG media at 20 or 30 °C. Taxane accumulation (A) and yeast growth (B) were evaluated after 72 hours of cultivation. Error bars represent ± standard deviation for triplicate cultivations.

314   The final taxadiene titre at 20 °C was 129 ± 15 mg/L, representing an almost 6.5-fold

315   improvement over the highest reported titre (Reider Apel *et al.*, 2017). The final titre at 30 °C was

316   56.9 ± 2.7 mg/L, which was comparable to the 57 ± 3 mg/L obtained in the smaller scale

317   cultivations (Figure 6). This represented a 2.3-fold lower taxadiene titre at 30 °C compared to 20

318   °C, indicating that despite the addition of solubility tags, further optimisation is likely to be

319   necessary at higher growth temperatures. Interestingly, biomass accumulation was very similar for

320   the strain at the two temperatures, with final OD_{600} values of 29 ± 1 and 31 ± 1 at 20 and 30 °C,

321   respectively (Figure 5B).

322   Although taxadiene was the major product under all conditions studied, additional side

323   products were also generated. At this scale, further characterisation of such side products was

324   performed through a detailed analysis of gas chromatography (Figure 6) and mass spectrometry
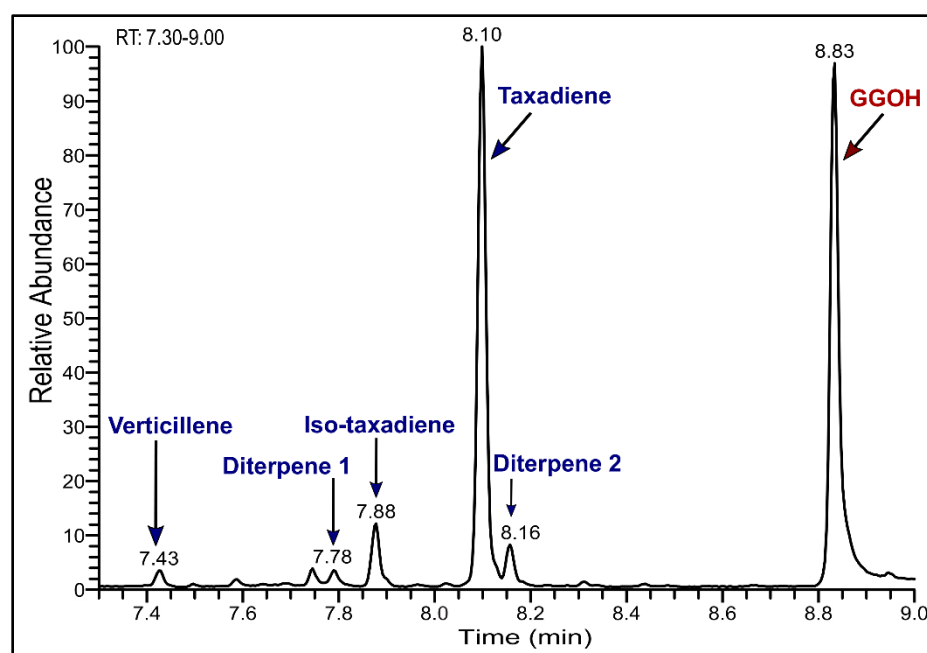
325   data (Supplementary Figure 2).

Figure 6: **LRS5 gas chromatogram.** Results show products produced by LRS5 during the 30 °C shake flask cultivation.

In addition to taxadiene, which was eluted at 8.10 minutes (Supplementary Figure 2D), a further five diterpene products were detected as shown in Figure 6. The mass spectra of the peaks at 7.43 and 7.88 minutes (Supplementary Figures 2A and C) showed a high degree of similarity to compounds previously identified as verticillene and the structural taxadiene isomer, taxa-4(20),11(12)-diene (iso-taxadiene), respectively (Williams *et al.*, 2000; Sagwan-Barkdoll and Anterola, 2018). An additional two diterpene compounds, Diterpene 1 and Diterpene 2 (Supplementary Figures 2B and E) were observed at 7.78 and 8.16 minutes, respectively. The mass spectrum of diterpene 1 closely resembled cembrene-type diterpene (cembrene A) produced by neocembrene synthase in *S. cerevisiae* (Kirby *et al.*, 2010). The cyclisation of GGPP to taxadiene involves production of a cembrenyl intermediate (van Rijn, Escorcia and Thiel, 2019). The relatively high degree of similarity suggests the compound observed in this study could potentially be a side product of this cyclisation step. However, further characterisation is needed to confirm this. The mass spectrum and elution order of diterpene 2 was very similar to that of a novel

16

340   diterpene product of *TASY* expression in *Nicotiana benthamiana* (Li *et al.*, 2019). Production of

341   geranylgeraniol (GGOH) (8.83 minute; Supplementary Figure 2F) was also confirmed using a pure

342   analytical standard. Although GGOH is formed naturally by *S. cerevisiae* through the degradation

343   of GGPP by endogenous phosphatases, titres are typically too low to detect (Song *et al.*, 2017). Its

344   overproduction here indicates that early work to overexpress the mevalonate pathway was

345   successful, however, taxadiene titres could be improved further through pathway optimisation. In

346   addition to the diterpene products, a number of additional potential terpenoid side-products were

347   detected between 4.36 and 6.38 minutes (Supplementary Figures 3 and 4).

348   3.3.2. Scale-up using a mini-bioreactor system

349   A 10-fold scale-up of the LRS5 cultivation was used to assess performance under industrial

350   conditions for taxadiene production. A 500-mL MiniBio bioreactor (Applikon, UK) was employed

351   for this study. This system possesses the online monitoring and control capabilities of larger-scale

352   bioreactors, allowing industrial scale cultivation conditions to be effectively mimicked whilst

353   conserving valuable resources. The results of the bioreactor-scale runs are summarised in Figure
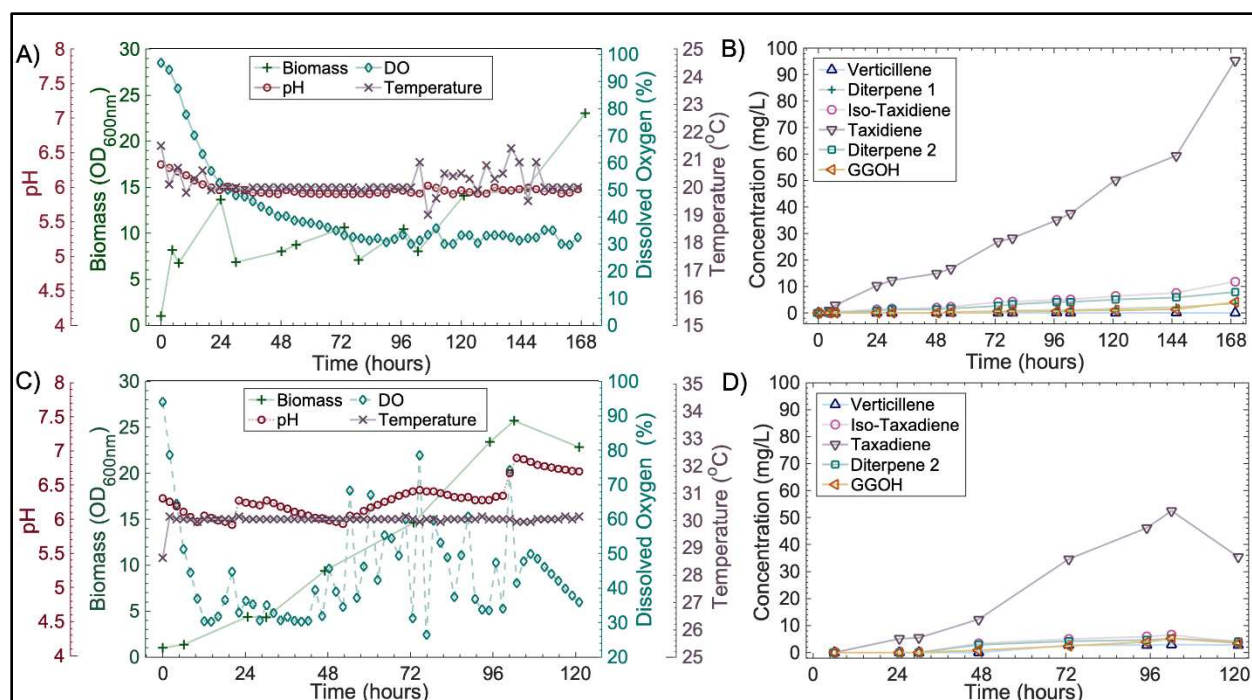
354   7.

355

356

Figure 7: **Bioreactor studies results and respective taxane concentration kinetics at 20 °C (A and B) and at 30 °C (C and D).** LRS5 was cultivated in an Applikon MiniBio 500 bioreactor in yeast extract (1%), peptone (2%), galactose (2%). The pH and dissolved oxygen were monitored online and controlled to set points of 6 and 30%, respectively online.

357   Following 72 hours of cultivation, the OD_{600} values of the 20 and 30 °C cultivations were

358   10.6 and 14.6 respectively as shown in Figures 7A and C. This is significantly lower than the 29

359   ± 1 and 31 ± 1 obtained at shake flask scale (Figure 5B). Taxadiene titres were also lower in the

360   bioreactor studies at 27 and 35 mg/L compared to 129 and 57 mg/L for the equivalent 20 and 30

361   °C shake flask cultures. Despite this, following a further 30 hours of cultivation, the stationary

362   phase of growth was reached and a maximum taxadiene titre of 53 mg/L was attained in the 30 °C

363   reactor. This is highly comparable to those obtained in the 72-hour microscale (Figure 6) and shake

364   flask (Figure 5A) cultivations. A higher taxadiene titre of 109 mg/L was achieved in the 20 °C

365   bioreactor, however, a bioreactor run time of 195 hours was required. Interestingly, production of

366   the endogenous GGOH side product was dramatically reduced at the increased scale with

18

367   maximum titres of 4 and 5 mg/L (Figures 7B and D) compared to 19 ± 3 and 35 ± 1 mg/L for the

368   equivalent 20 and 30 °C shake flask studies (Figure 5A).

369   Unlike prokaryotes, eukaryotic yeast cells have evolved to store hydrophobic compounds

370   within specialised organelles (liquid droplets) (Arhar and Natter, 2019). As a result, it was

371   hypothesised that additional taxadiene may be present within the yeast cells. An investigation into

372   intracellular taxane production was therefore performed on samples taken from the higher-yielding

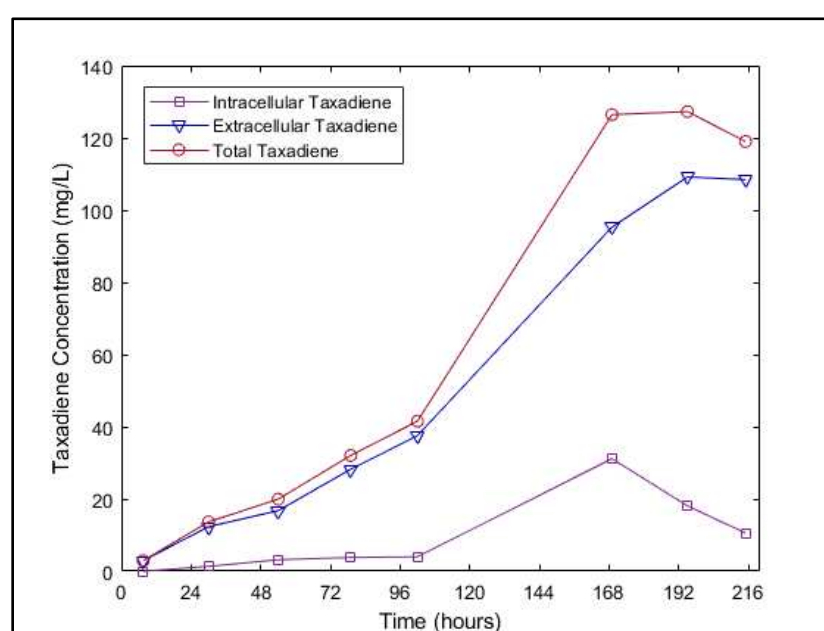373   20 °C bioreactor cultivation as summarised in Figure 8.

374

375   Figure 8: **Kinetics of intracellular and total taxadiene titres in high-yielding 20 °C bioreactor cultivation.**

376   In the first 168 hours of the cultivation, both extracellular and intracellular taxadiene

377   accumulation increased (Figure 8). In the subsequent 48 hours a sharp decline in intracellular

378   taxadiene from 31 to 11 mg/L was observed. This was coupled with an increase in extracellular

379   taxadiene from 95 to 109 mg/L while the total taxadiene concentration remained similar between

380   119 and 127 mg/L (Figure 8). This suggested that the increase in extracellular taxadiene

19

381   concentration observed in the last 48 hr was likely largely due to taxadiene secretion rather than

382   additional taxadiene biosynthesis.

383   A lower cultivation temperature of 20 °C was found to improve TASY performance with

384   higher taxadiene titres being obtained under all the conditions investigated in this study. However,

385   the kinetic analysis revealed a strong positive correlation between biomass and product

386   accumulation at both 20 and 30 °C (Pearson's R = 0.916 and 0.974, respectively). As the growth

387   rate of LRS5 decreased with temperature, longer cultivation time was required at 20 °C. This was

388   exacerbated at larger scale, where cultivation time was tripled (Figure 7A). Although careful

389   optimisation yielded substantial improvements in TASY productivity in this study, the

390   discrepancies between the optimal temperature for the enzyme and host growth remain a challenge.

391   The optimisation of copy number and protein tag combination improved taxadiene titres

392   dramatically in this study. However, the higher titres obtained at reduced temperature indicate

393   further optimisation is needed at the optimal growth temperature of 30 °C.

394   **4. Conclusions**

395   This study reports the successful optimisation of taxadiene biosynthesis in a *Saccharomyces*

396   *cerevisiae* microbial chassis. Low expression and poor solubility of taxadiene synthase (TASY)

397   were identified as a critical bottlenecks. This was alleviated through multi-copy chromosomal

398   integration of *TASY* with a combination of fusion protein tags, improving taxadiene titres 22-fold

399   to 57 ± 3 mg/L at 30 °C. TASY performance was found to be temperature-dependant, with a

400   maximum taxadiene titre of 129 ± 15 mg/L at 20 °C. Similar titres were achieved at larger scale

401   as well, highlighting the scalability of the bioprocess, and representing a 6.5-fold improvement on

402   the highest literature titre.

20

403   This work highlights the benefit of multifactorial approach to biosynthetic pathway

404   optimisation, where a combination of synthetic biology tools and bioprocessing approaches were

405   employed. Although the current study presents significant progress in taxadiene biosynthesis,

406   several tasks remain to improve the expression of taxadiene synthase at 30 °C to allow for further

407   Paclitaxel biosynthetic pathway development in *S. cerevisiae*.

408   **5. Acknowledgements**

417   **6. Declaration of interest**

418   J.D.K. has financial interests in Amyris, Lygos, Demetrix, Napigen, Maple Bio, Apertor Labs,

419   Ansa Biotechnologies, and Berkeley Brewing Sciences.

420

421

422

7. References

423   Abdallah, I. I. *et al.* (2019) 'Metabolic engineering of *Bacillus subtilis* toward taxadiene
424   biosynthesis as the first committed step for taxol production', *Frontiers in Microbiology*,
425   p. 218. doi: 10.3389/fmicb.2019.00218.

426   Ajikumar, P. K. *et al.* (2010) 'Isoprenoid pathway optimization for taxol precursor overproduction
427   in *Escherichia coli*', *Science*, 330(6000), p. 70. doi: 10.1126/science.1191652.

428   Arhar, S. and Natter, K. (2019) 'Common aspects in the engineering of yeasts for fatty acid- and
429   isoprene-based products', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell
430   Biology of Lipids*, 1864(12), p. 158513. doi: 10.1016/j.bbalip.2019.08.009.

431   Biggs, B. W. *et al.* (2016) 'Overcoming heterologous protein interdependency to optimize P450-
432   mediated Taxol precursor synthesis in *Escherichia coli*', *Proceedings of the National
433   Academy of Sciences*, 113(12), pp. 3209–3214. doi: 10.1073/pnas.1515826113.

434   Boghigian, B. A. *et al.* (2012) 'Analysis of heterologous taxadiene production in K- and B-derived
435   *Escherichia coli*', *Applied Microbiology and Biotechnology*, 93(4), pp. 1651–1661. doi:
436   10.1007/s00253-011-3528-4.

437   DeJong, J. M. *et al.* (2006) 'Genetic engineering of taxol biosynthetic genes in *Saccharomyces
438   cerevisiae*', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 93(2), pp. 212–
439   224. doi: 10.1002/bit.20694.

440   Delic, M. *et al.* (2013) 'The secretory pathway: exploring yeast diversity', *FEMS Microbiology
441   Reviews*, pp. 872–914. doi: 10.1111/1574-6976.12020.

442   Deng, Y. *et al.* (2016) 'Enhanced (*S*)-linalool production by fusion expression of farnesyl
443   diphosphate synthase and linalool synthase in *Saccharomyces cerevisiae*', *Journal of
444   Applied Microbiology*, 121(1), pp. 187–195. doi: 10.1111/jam.13105.

446   Engels, B., Dahm, P. and Jennewein, S. (2008) 'Metabolic engineering of taxadiene biosynthesis
447   in yeast as a first step towards Taxol (Paclitaxel) production', *Metabolic Engineering*,
448   10(3), pp. 201–206. doi: 10.1016/j.ymben.2008.03.001.

449   Frense, D. (2007) 'Taxanes: perspectives for biotechnological production', *Applied Microbiology*
450   *and Biotechnology*, 73(6), pp. 1233–1240. doi: 10.1007/s00253-006-0711-0.

451   Gietz, R. D. and Schiestl, R. H. (2007) 'High-efficiency yeast transformation using the LiAc/SS
452   carrier DNA/PEG method', *Nature Protocols*, 2(1), pp. 31–34. doi: 10.1038/nprot.2007.13.

453   Houser, J. R. *et al.* (2012) 'An improved short-lived fluorescent protein transcriptional reporter for
454   Saccharomyces cerevisiae', *Yeast*, 29(12), pp. 519–530. doi: 10.1002/yea.2932.

455   Huang, K. *et al.* (1998) 'Overproduction, in *Escherichia coli* of soluble taxadiene synthase, a key
456   enzyme in the taxol biosynthetic pathway', *Protein Expression and Purification*, 13(1), pp.
457   90–96. doi: 10.1006/prep.1998.0870.

458   Huang, Q. *et al.* (2001) 'Engineering *Escherichia coli* for the synthesis of taxadiene, a key
459   intermediate in the biosynthesis of taxol', *Bioorganic & Medicinal Chemistry*, 9(9), pp.
460   2237–2242. doi: 10.1016/S0968-0896(01)00072-4.

461   Ignea, C. *et al.* (2015) 'Efficient diterpene production in yeast by engineering Erg20p into a
462   geranylgeranyl diphosphate synthase', *Metabolic Engineering*, 27, pp. 65–75. doi:
463   10.1016/j.ymben.2014.10.008.

464   Jin, Y. *et al.* (2005) 'Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl
465   diphosphate to 7-fluoroverticillenes', *Journal of the American Chemical Society*. American
466   Chemical Society, 127(21), pp. 7834–7842. doi: 10.1021/ja050592r.

467   Kaspera, R. and Croteau, R. (2006) 'Cytochrome P450 oxygenases of Taxol biosynthesis',
468   *Phytochemistry Reviews*. Dordrecht, 5(2–3), pp. 433–444. doi: 10.1007/s11101-006-9006-
469   4.

23

470   Kirby, J. *et al.* (2010) 'Cloning of casbene and neocembrene synthases from Euphorbiaceae plants
471   and expression in *Saccharomyces cerevisiae*', *Phytochemistry*, 71(13), pp. 1466–1473. doi:
472   10.1016/j.phytochem.2010.06.001.

473   Köksal, M. *et al.* (2010) 'Taxadiene synthase structure and evolution of modular architecture in
474   terpene biosynthesis', *Nature*. Nature Publishing Group, a division of Macmillan
475   Publishers Limited. All Rights Reserved., 469, p. 116. doi: 10.1038/nature09628.

476   Kosobokova, E. N., Skrypnik, K. A. and Kosorukov, V. S. (2016) 'Overview of fusion tags for
477   recombinant   proteins',   *Biochemistry   (Moscow)*,   81(3),   pp.   187–200.   doi:
478   10.1134/S0006297916030019.

479   Lee, M. E. *et al.* (2015) 'A highly characterized yeast toolkit for modular, multipart assembly',
480   *ACS Synthetic Biology*. American Chemical Society, 4(9), pp. 975–986. doi:
481   10.1021/sb500366v.

482   Li, J. *et al.* (2019) 'Chloroplastic metabolic engineering coupled with isoprenoid pool
483   enhancement for committed taxanes biosynthesis in Nicotiana benthamiana', *Nature
484   Communications*, 10(1), pp. 1–12. doi: 10.1038/s41467-019-12879-y.

485   McElroy, C. and Jennewein, S. (2018) 'Taxol® biosynthesis and production: from forests to
486   fermenters bt - biotechnology of natural products', in Schwab, W., Lange, B. M., and
487   Wüst, M. (eds). Cham: Springer International Publishing, pp. 145–185. doi: 10.1007/978-
488   3-319-67903-7_7.

489   Reider Apel, A. *et al.* (2017) 'A Cas9-based toolkit to program gene expression in *Saccharomyces
490   cerevisiae*', *Nucleic acids research*. 2016/11/24. Oxford University Press, 45(1), pp. 496–
491   508. doi: 10.1093/nar/gkw1023.

492   Sagwan-Barkdoll, L. and Anterola, A. M. (2018) 'Taxadiene-5α-ol is a minor product of
493   CYP725A4 when expressed in *Escherichia coli*', *Biotechnology and Applied
494   Biochemistry*, 65(3), pp. 294–305. doi: 10.1002/bab.1606.

495   Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) 'NIH Image to ImageJ: 25 years of
496   image analysis', *Nature Methods*, 9(7), pp. 671–675. doi: 10.1038/nmeth.2089.

497   Schrepfer, P. *et al.* (2016) 'Identification of amino acid networks governing catalysis in the closed
498   complex of class I terpene synthases', *Proceedings of the National Academy of Sciences*,
499   113(8), p. E958-E967. doi: 10.1073/pnas.1519680113.

500   Song, T.-Q. *et al.* (2017) 'Engineering *Saccharomyces cerevisiae* for geranylgeraniol
501   overproduction by combinatorial design', *Scientific Reports*, 7(1), p. 14991. doi:
502   10.1038/s41598-017-15005-4.

503   Sun, P., Tropea, J. E. and Waugh, D. S. (2011) 'Enhancing the solubility of recombinant proteins
504   in Escherichia coli by using hexahistidine-tagged maltose-binding protein as a fusion
505   partner', *Methods in molecular biology (Clifton, N.J.)*, 705, p. 259. doi: 10.1007/978-1-
506   61737-967-3_16.

507   Tetko, I. V *et al.* (2005) 'Virtual computational chemistry laboratory – design and description',
508   *Journal of Computer-Aided Molecular Design*, 19(6), pp. 453–463. doi: 10.1007/s10822-
509   005-8694-y.

510   van Rijn, J. P. M., Escorcia, A. M. and Thiel, W. (2019) 'QM/MM study of the taxadiene synthase
511   mechanism', *Journal of Computational Chemistry*. Hoboken, USA, 40(21), pp. 1902–
512   1910. doi: 10.1002/jcc.25846.

513   Westfall, P. J. *et al.* (2012) 'Production of amorphadiene in yeast, and its conversion to
514   dihydroartemisinic acid, precursor to the antimalarial agent artemisinin', *Proceedings of*
515   *the National Academy of Sciences of the United States of America*. 2012/01/12. National
516   Academy of Sciences, 109(3), pp. E111–E118. doi: 10.1073/pnas.1110740109.

517   Williams, D. C. *et al.* (2000) 'Heterologous expression and characterization of a "pseudomature"
518   form of taxadiene synthase involved in paclitaxel (taxol) biosynthesis and evaluation of a
519   potential intermediate and inhibitors of the multistep diterpene cyclization reaction',

520   *Archives of Biochemistry and Biophysics*, 379(1), pp. 137–146. doi:
521   10.1006/abbi.2000.1865.

522   Wong, J. *et al.* (2018) 'De novo synthesis of the sedative valerenic acid in *Saccharomyces*
523   *cerevisiae*', *Metabolic Engineering*, 47, pp. 94–101. doi: 10.1016/j.ymben.2018.03.005.

524   Zhang, M. *et al.* (2017) 'SBOLDesigner 2: An intuitive tool for structural genetic design', *ACS*
525   *Synthetic Biology*. American Chemical Society, 6(7), pp. 1150–1160. doi:
526   10.1021/acssynbio.6b00275.