Electronic Supplementary Information

High-throughput screen for sorting cells capable of producing the biofuel feedstock botryococcene

Allen R. Chacko,[a] Daniel E. Amster,[a] Tyler E. Johnson,[a] Sara R. Newman,[a] Angelina V. Gladchuk,[a] Christine J. Sohn,[a] Donna E. Prunkard,[b] Neal A. Yakelis[a] and Jon O. Freeman*[a]

[a.] *Department of Chemistry, Pacific Lutheran University, Tacoma, WA 98447, USA.*

[b.] *Department of Pathology, University of Washington, Seattle, WA 98195, USA*

** Corresponding author. E-mail: freemajo@plu.edu*

Supplementary figures and tables

Fig. S1.....2
Fig. S2.....2
Fig. S3.....3
Fig. S4.....3
Fig. S5.....4
Fig. S6.....5
Fig. S7.....5
Fig. S8.....6
Fig. S9.....6
Fig. S10.....7
Fig. S11.....8
Fig. S12.....9
Fig. S13.....10
Fig. S14.....11
Fig. S15.....12
Fig. S16.....12
References.....12

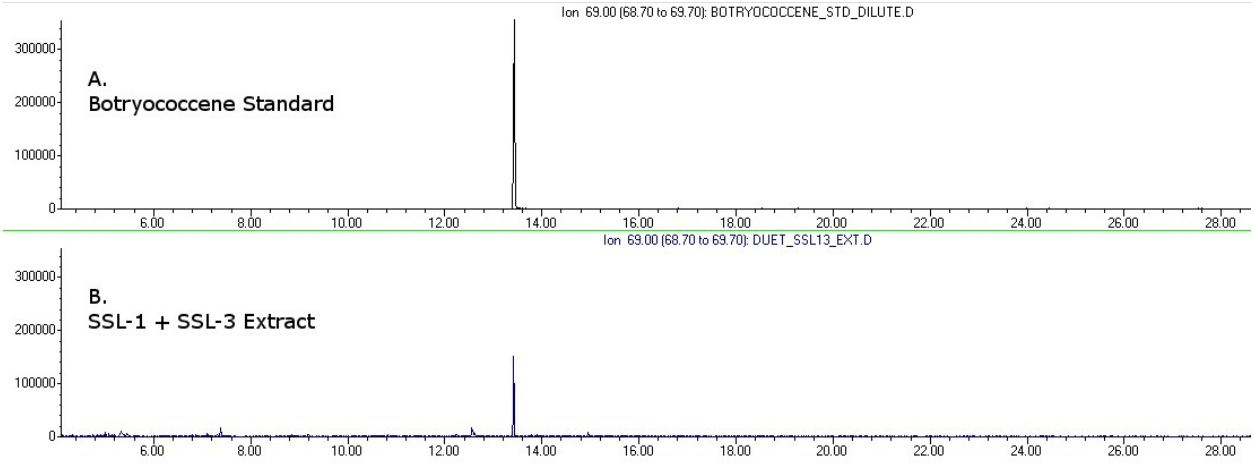


Fig. S1 GC-MS analysis of a botryococcene standard (A) and organic extract from *E. coli* cells harboring the SSL-1 + SSL-3 genes (B). Data presented as extracted-ion chromatograms of *m/z* 69.00.

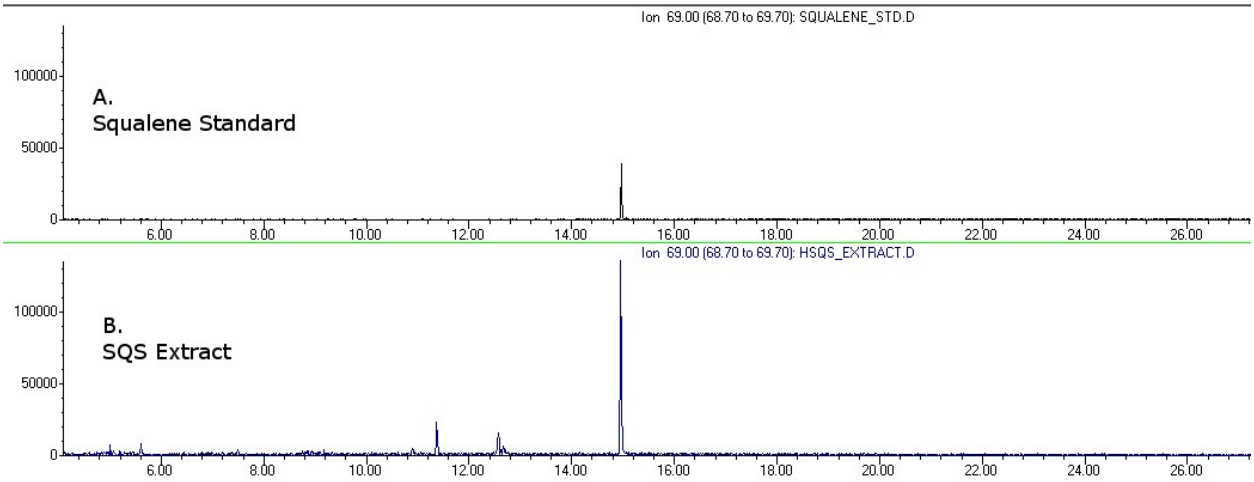


Fig. S2 GC-MS analysis of a squalene standard (A) and organic extract from *E. coli* cells harboring the SQS gene (B). Data presented as extracted-ion chromatograms of *m/z* 69.00.

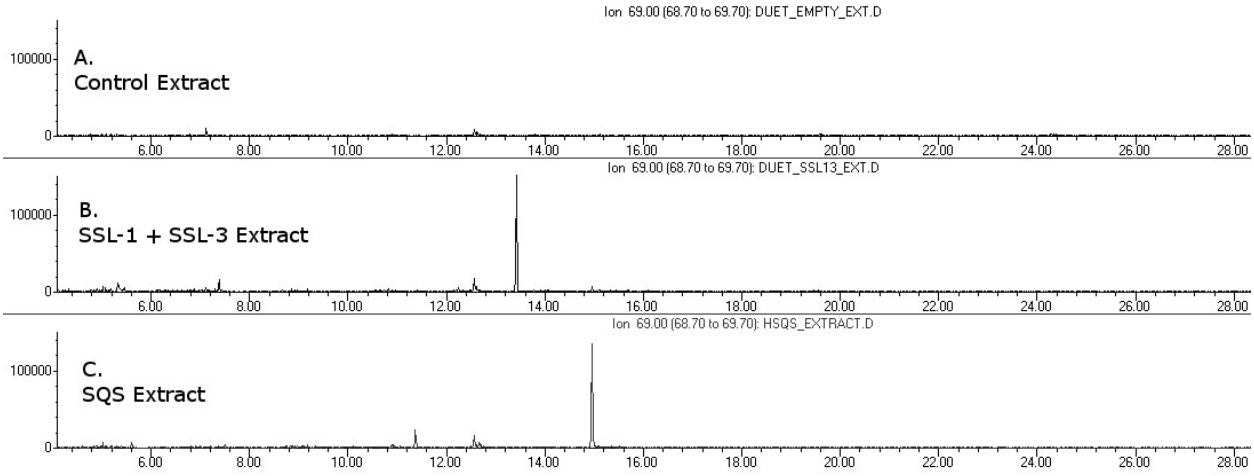


Fig. S3 GC-MS analysis of organic extracts from control *E. coli* cells (A), *E. coli* cells harboring the SSL-1 + SSL-3 genes (B), *E. coli* cells harboring the SQS gene (C). Data presented as extracted-ion chromatograms of m/z 69.00.

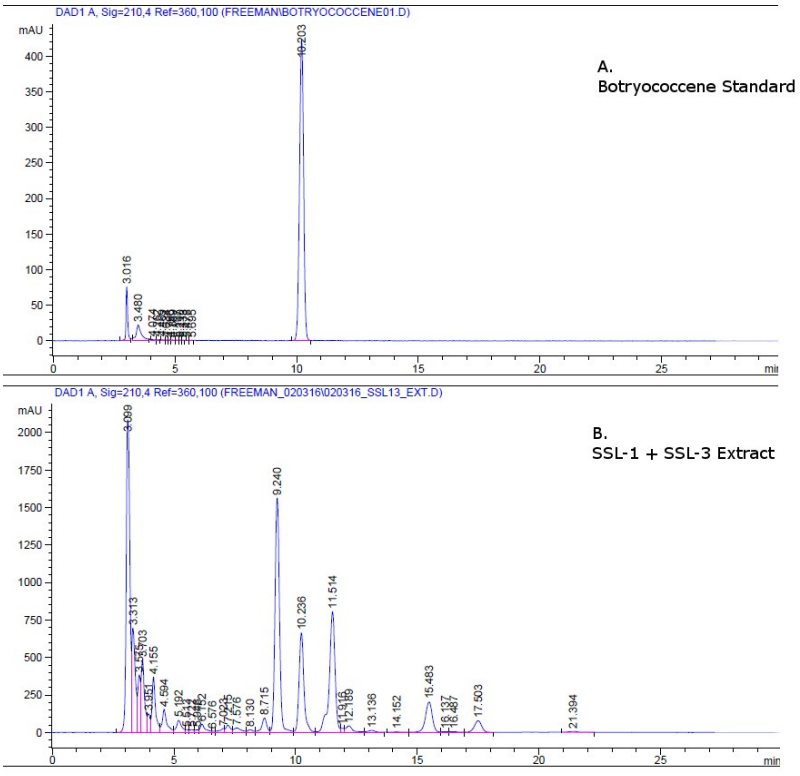


Fig. S4 RP-HPLC chromatogram's of a botryococcene standard (A) and an organic extract from *E. coli* cells harboring the SSL-1 + SSL-3 genes (B). Botryococcene has a retention time = 10.2 min.

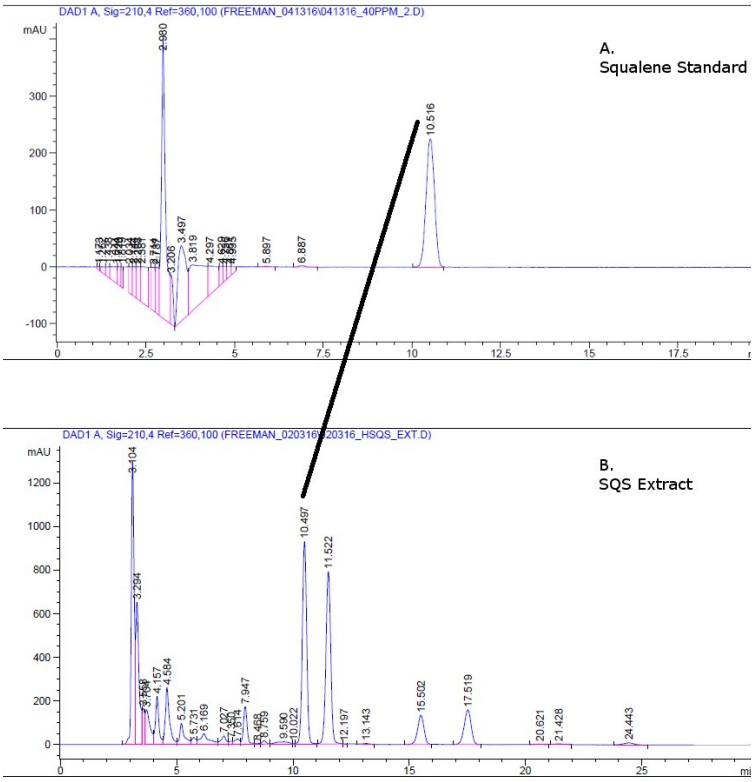


Fig. S5 RP-HPLC chromatogram's of a squalene standard (A) and an organic extract from *E. coli* cells harboring the SQS gene (B). Squalene has a retention time = 10.5 min.

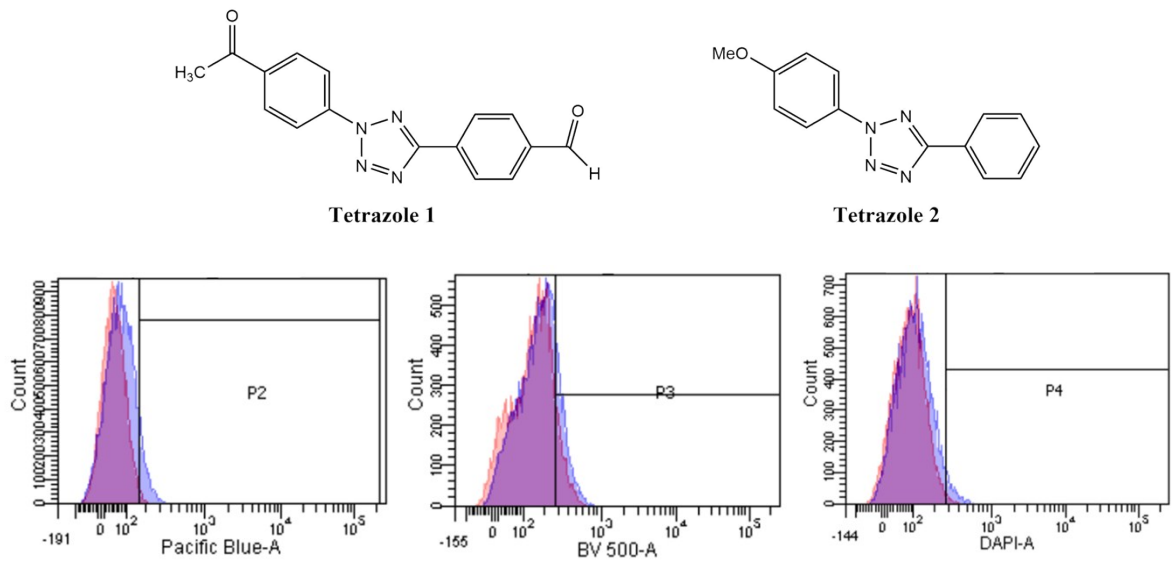


Fig. S6 Structures of tetrazole **1** and tetrazole **2** (top). Tetrazole **2** was previously used by Wang et al. and Song et al. in studies to label *E. coli* and HeLa cells.[1,2] While tetrazole **2** also functioned as desired in this study, tetrazole **1**'s mean fluorescence in the presence of botryococcene was found to be slightly higher when cells were interrogated with a 405 nm laser excitation with emission observed using Pacific Blue (450/50) and BV-500 (525/50) emission filters and with a 355 nm laser excitation with emission observed using a DAPI (450/50) filter (bottom row, Tetrazole **1** = blue, Tetrazole **2** = red). Both tetrazole concentrations were 150 µM and assessed using the same conditions.

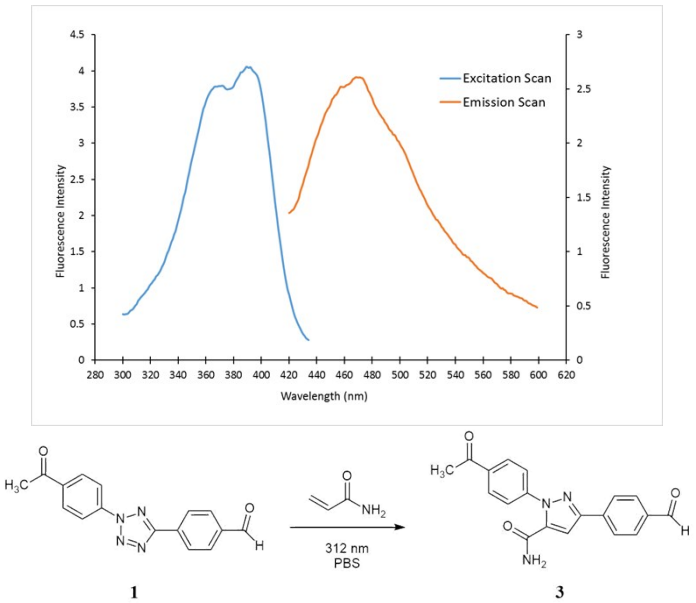


Fig. S7 Excitation and emission scans (top) of pyrazoline **3** generated *in situ* from the photoclick reaction of 150 µM tetrazole **1** with 150 µM acrylamide in PBS buffer. The excitation scan was performed while monitoring emission at 465 nm and the emission scan was carried out using an excitation of 405 nm. The photoclick reaction between tetrazole **1** and acrylamide is depicted on the bottom.

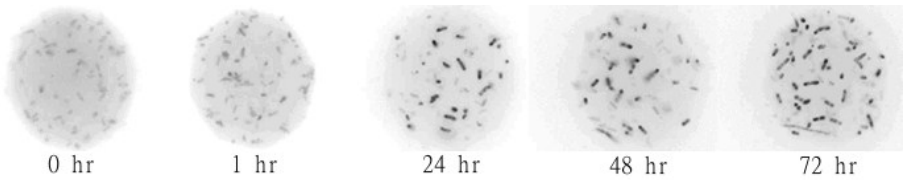


Fig. S8 Fluorescent images of *E. coli* cells co-expressing SSL-1 + SSL-3 enzymes cultured at different time intervals, post photoclick reaction. The increased fluorescence over time suggests an accumulation of botryococcene. Image colors were inverted and grayscaled for printing clarity.

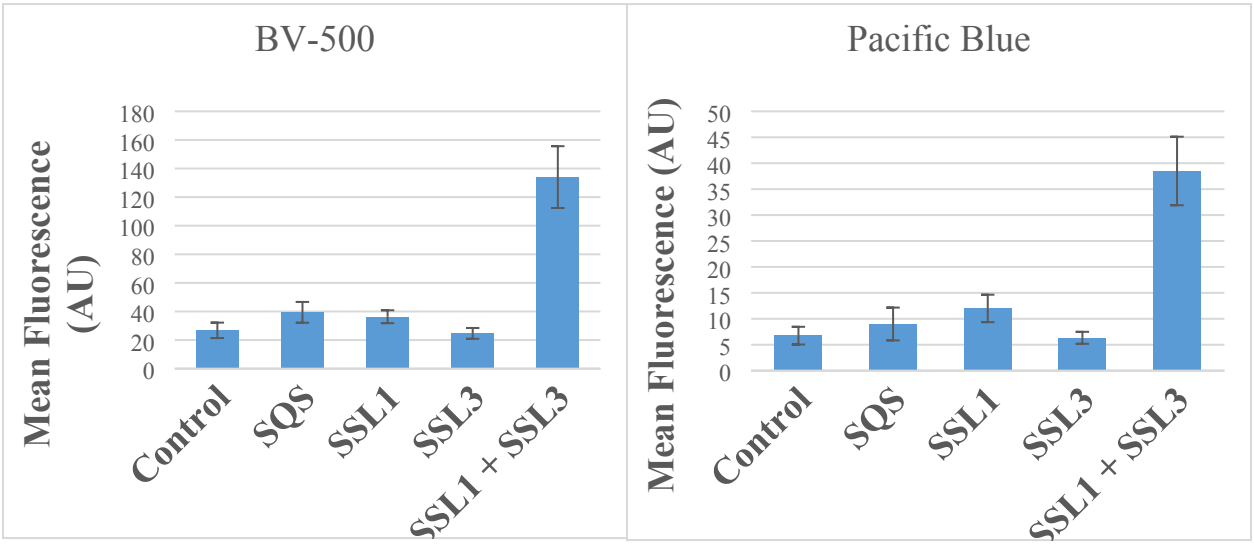


Fig. S9 Mean fluorescence of control, SQS, SSL-1, SSL-3, and SSL-1 + SSL-3 cells as determined by flow cytometry. Cells were interrogated using a 405 nm laser for excitation and emission was monitored using a BV-500 (525/50) or Pacific Blue (450/50) filter.

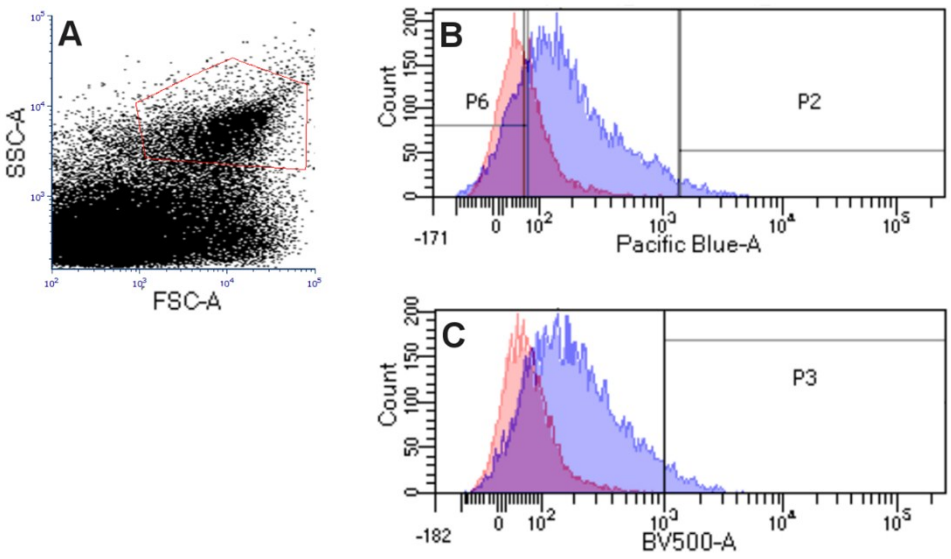


Fig. S10 FACS data depicting a sample of SQS cells (red) and an approximate 50:50 mixture of SQS & SSL-1 + SSL-3 cells (blue). Cells were initially gated to separate from debris, using a forward and side scatter plot (A), and subsequently sorted based on fluorescent intensity. Gates P2 and P3 represent the fluorescent intensity threshold used to sort cells using a 405 nm excitation laser and either Pacific Blue (450/50) or BV500 (525/50) emission filters (B & C, respectively).

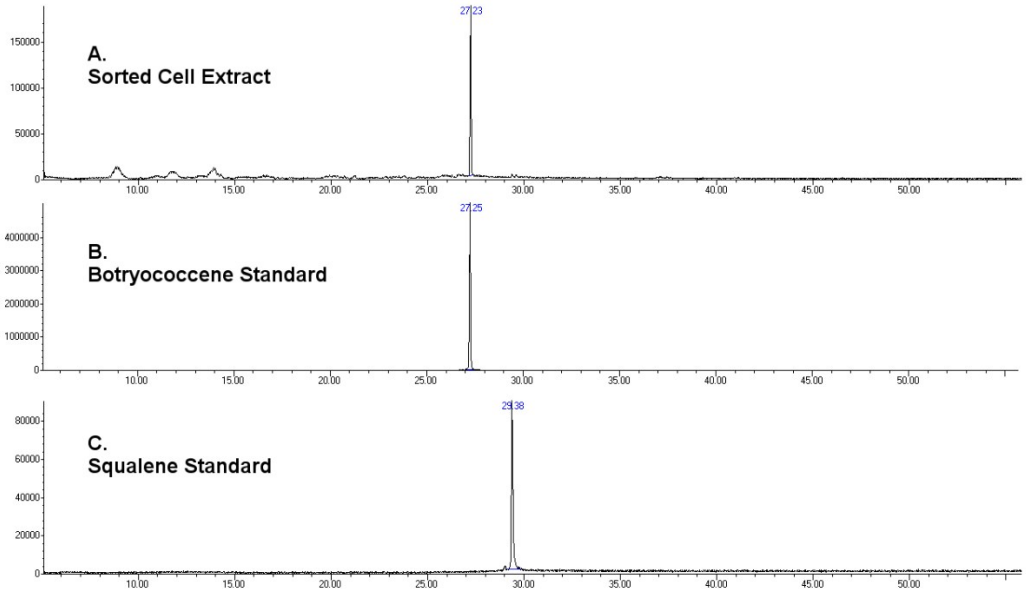
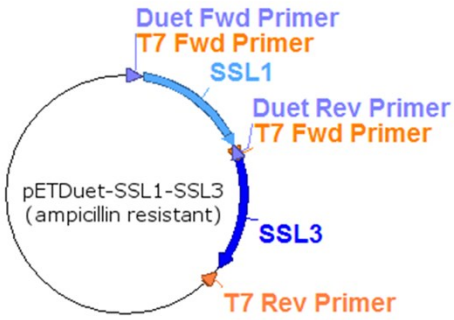
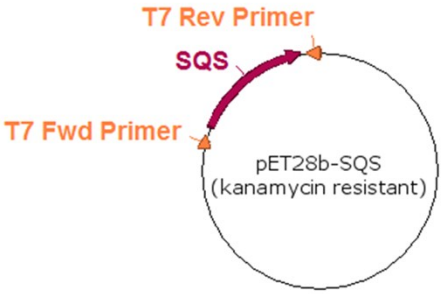


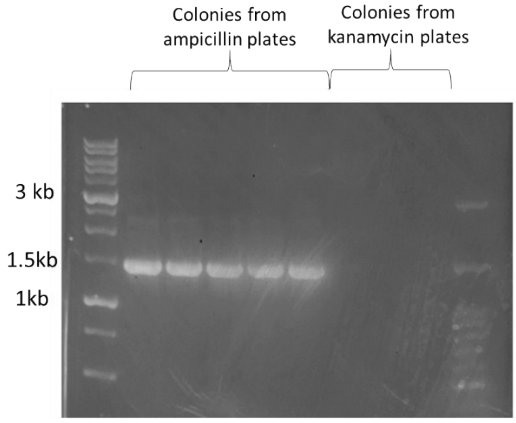
Fig. S11 GC-MS analysis of an organic extract from a culture of sorted E. coli cells (A). Cells were sorted from a 50:50 mixture of E. coli containing SSL-1 + SSL-3 and SQS and cultured in the absence of a selective marker. The chromatogram depicts botryococcene as the primary terpene extract. Botryococcene (B) and squalene (C) standards are shown for reference. Data presented as extracted-ion chromatograms of m/z 69.00.



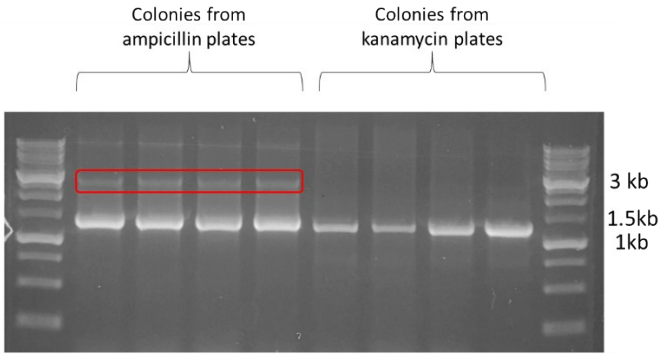


Primers	Predicted Amplicon Length (bp)
Duet Primers	1400
T7 Primers	1400, 2800

Primers	Predicted Amplicon Length (bp)
Duet Primers	N/A
T7 Primers	1300



Colony PCR using Duet Primers



Colony PCR using T7 Primers

Fig. S12 Colony PCR data derived from a FACS experiment that sorted a mixture of *E. coli* cells harboring pETDuet-SSL1-SSL3 and pET28b-SQS. The sorted mixture was cultured in LB media overnight and subsequently plated on LB agar containing either ampicillin or kanamycin. Colony PCR was performed using pETDuet and T7 primers to analyze the plasmids of the resulting *E. coli* colonies. Plasmid maps depicting the location of the primers and anticipated amplicon base pair lengths are shown on top. Gel electrophoresis data of colony PCR products using the Duet and T7 primers from randomly selected *E. coli* colonies are shown below. A total of 60 colonies across twelve plates were investigated; 30 from kanamycin and 30 from ampicillin plates (not all data is shown). PCR products confirmed that colonies from the ampicillin plates contain the pETDuet-SSL1-SSL3 plasmid as predicted, and colonies from the kanamycin plate are a result of pET28b-SQS contamination. Despite the contamination, colony forming units per mL from the overnight LB media was calculated to be $2.73 \pm 0.09 \times 10^9$ and $6 \pm 2 \times 10^2$ for the SSL-1 + SSL-3 and SQS harboring cells, respectively, suggesting trace amounts of SQS contamination.

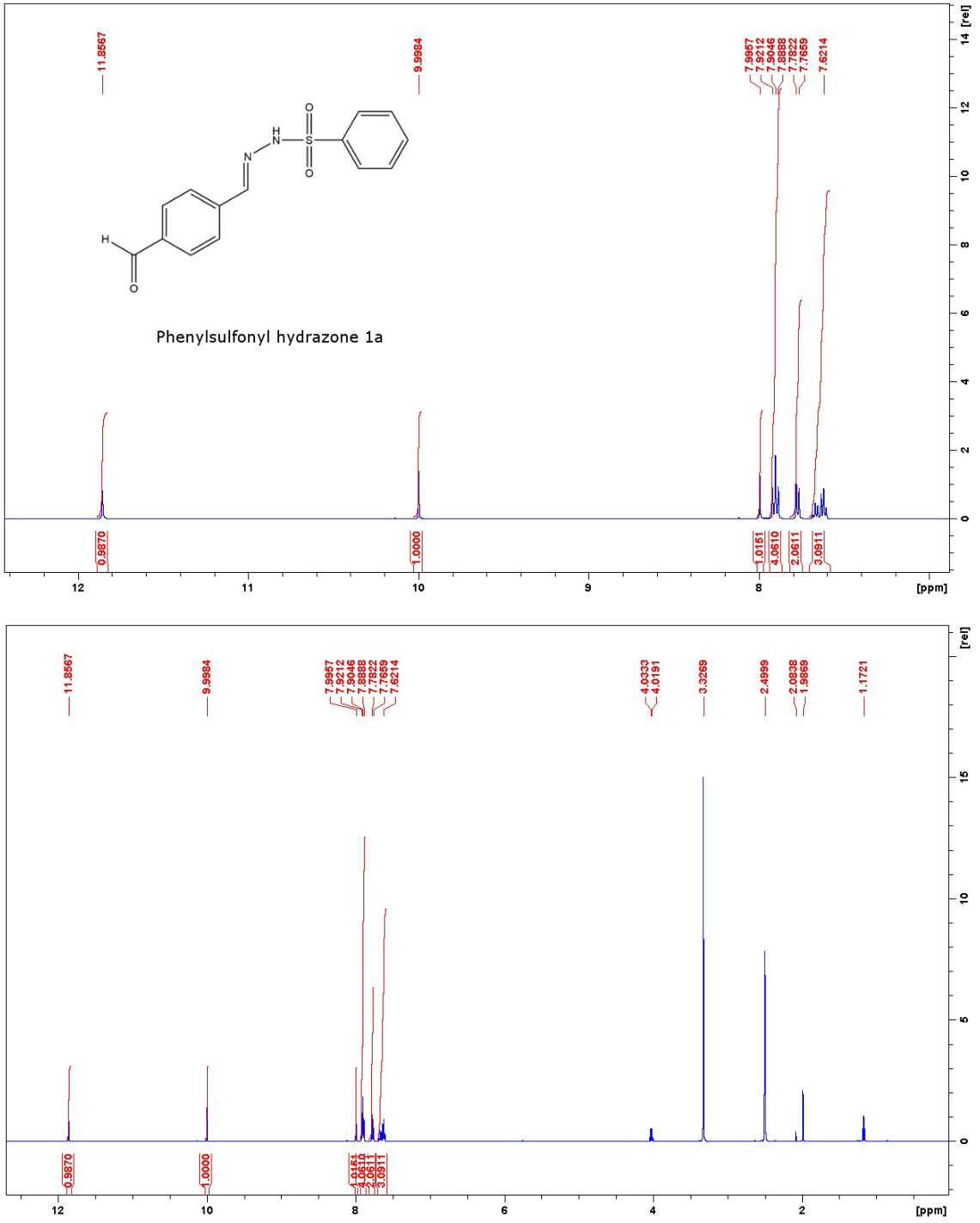


Fig. S13 Terephthalaldehyde (phenylsulfonyl) hydrzaone **1a** 500 MHz ^{1}H NMR spectra in DMSO-d6. The top spectrum depicts the signals corresponding to **1a** and the bottom shows the complete spectrum. Spectra processed and presented with Bruker TopSpin 3.5. Residual solvent signals in the ^{1}H NMR are from ethyl acetate (4.03, 1.99, and 1.17 ppm), DMSO (3.33 ppm), H_2O (2.50 ppm) and acetone (2.08 ppm) from the work-up procedure.

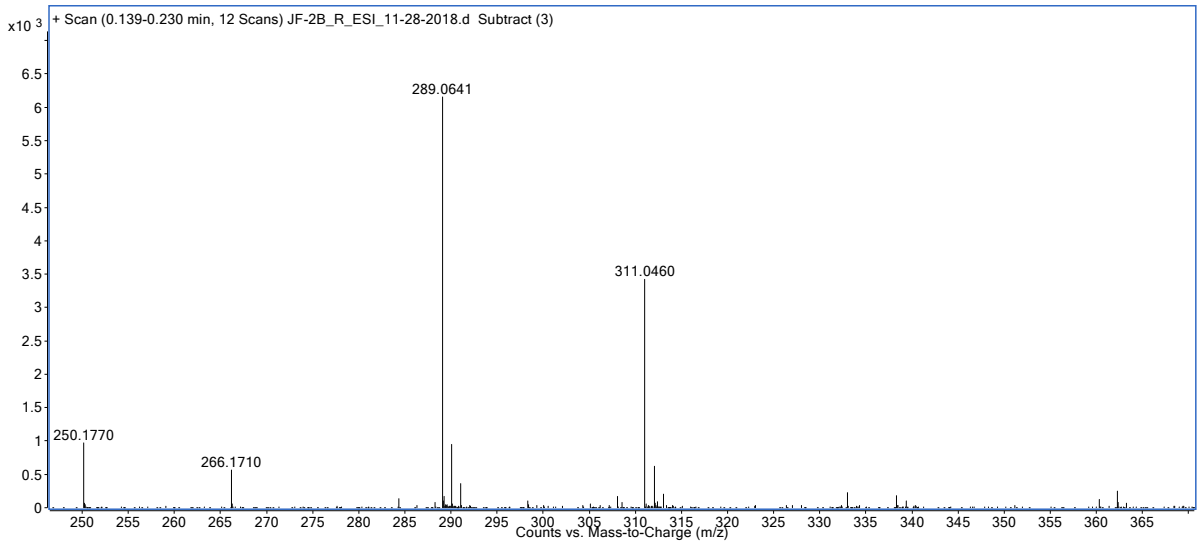


Fig. S14 ESI high resolution mass spectrum of terephthalaldehyde (phenylsulfonyl) hydrazone **1a.** The predicted [M+H]+ m/z value is 289.0641 and the predicted [M+Na]+ m/z value is 311.0461.

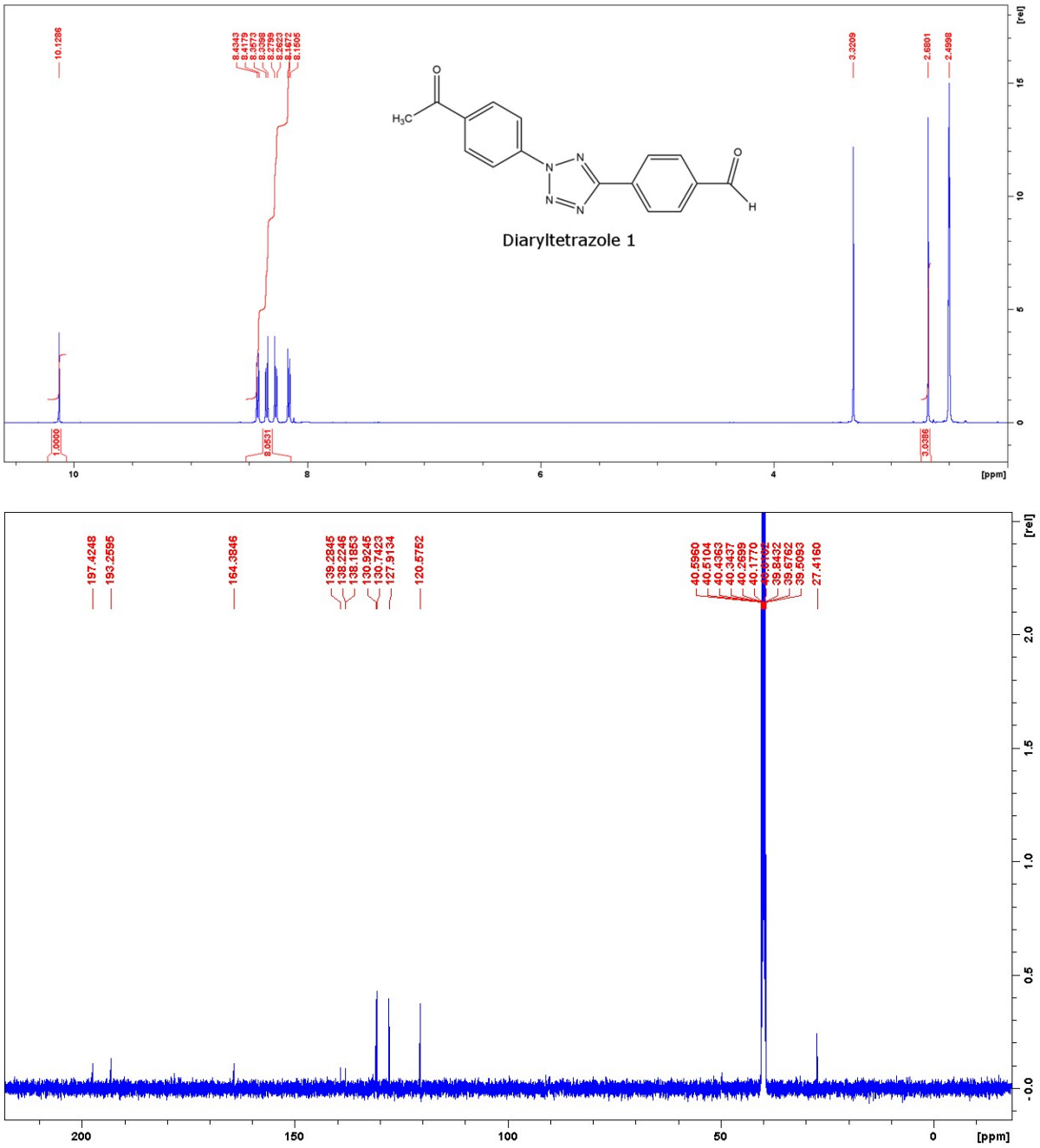


Fig. S15 Diaryltetrazole **1** 500 MHz ^{1}H (top) and ^{13}C (bottom) NMR spectra in DMSO-d6. Spectra processed and presented using Bruker TopSpin 3.5. Residual solvent signals in the ^{1}H NMR are from DMSO (3.33 ppm) and H_2O (2.50 ppm) and in the ^{13}C NMR are from DMSO (39.5 ppm).

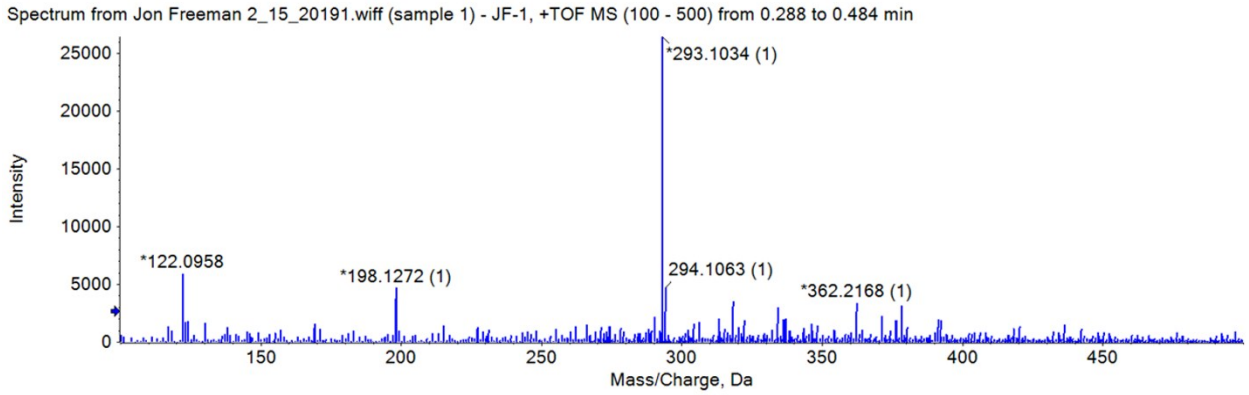


Fig. S16 ESI high resolution mass spectrum of 4-(2-(4-acetylphenyl)-*2H*-tetrazol-5-yl)benzaldehyde **1.** The predicted [M+H]+ m/z value is 293.1033.

References

1 Y. Wang, W. Song, W. J. Hu and Q. Lin, *Angew. Chem. Int. Ed.*, 2009, **48**, 5330–5333.

2 W. Song, Y. Wang, Z. Yu, C. I. R. Vera, J. Qu and Q. Lin, *ACS Chem. Biol.*, 2010, **5**, 875–885.