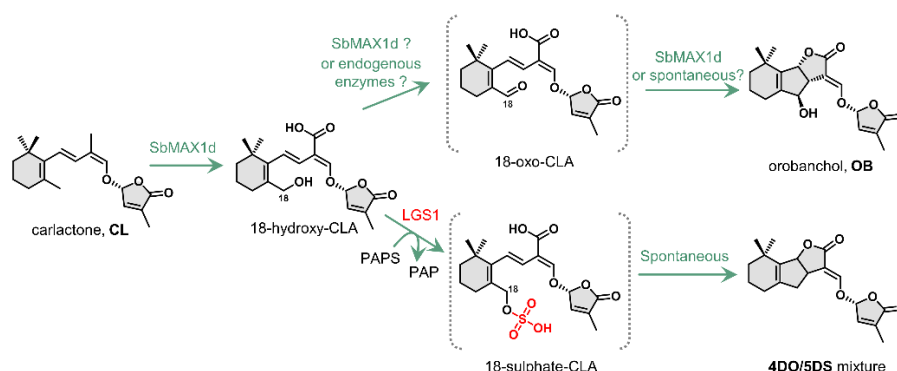


## A unique sulfotransferase-involving strigolactone biosynthetic route in Sorghum

Sheng Wu<sup>1</sup>, Yanran Li<sup>1\*</sup>

<sup>1</sup>Department of Chemical and Environmental Engineering, University of California, Riverside, California 92521, USA

\*Correspondence: Yanran Li, Email: [yanranl@ucr.edu](mailto:yanranl@ucr.edu)



### ABSTRACT

LOW GERMINATION STIMULANT 1 (LGS1) plays an important role in strigolactones (SLs) biosynthesis and *Striga* resistance in sorghum but the catalytic function remains unclear. Using the recently developed SL-producing microbial consortia, we examined the activities of sorghum MAX1 analogs and LGS1. Surprisingly, SbMAX1d (accession # XP\_002458367) synthesized 18-hydroxy-carlactonoic acid (18-hydroxy-CLA) directly from carlactone (CL) through four-step oxidations, and addition of LGS1 led to the synthesis of both 5-deoxystrigol (5DS) and 4-deoxyorobanchol (4DO). Further biochemical characterization found that LGS1 functions after SbMAX1d by converting 18-hydroxy-CLA to 18-sulphate-CLA to provide an easier leaving group to afford a spontaneous formation of 5DS and 4DO. The unique functions of SbMAX1 and LGS1 imply a previously unknown synthetic route towards strigolactones.

## Introduction

Strigolactones (SL) are a group of butanolide-containing molecules originally identified as seed germination stimulants for the parasitic weeds *Striga* and *Orobancha*<sup>1, 2</sup>, and later characterized as phytohormones that play diverse important roles in plant growth and development<sup>3-5</sup>. SLs can be divided into canonical and non-canonical SLs, with canonical SLs further grouped into strigol (*S*)- and orobanchol (*O*)- type SLs according to the stereochemistry of the C-ring<sup>3</sup> (Figure 1). Different SL structures have been reported to exhibit distinct parasitic weed germination activities<sup>6, 7</sup>. For example, SLs exhibiting high germination stimulation activity towards *S. gesnerioides* induced low germination in *S. hermonthica*, while several SLs of high germination stimulation activity to *S. hermonthica* inhibit the germination of *S. gesnerioides*<sup>8</sup>. Recently, LOW GERMINATION STIMULANT 1 (LGS1) has been identified to be responsible for the *Striga* germination stimulant activity in sorghum and missing from the *Striga*-resistant sorghum varieties<sup>9</sup>, which produce distinct SL profiles, i.e., (*S*)-type 5DS and (*O*)-type orobanchol (OB), respectively<sup>9</sup>. LGS1 is a putative sulfotransferase (SOT), which normally catalyzes the transfer of a sulfonate group from 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to a hydroxyl group of acceptor molecules<sup>10</sup>. The mechanism on how LGS1 regulates SL profiles between 5DS and OB in sorghum remains unclear.

SLs are synthesized from carlactone (CL), which is then converted to diverse SL structures by various downstream tailoring enzymes especially cytochrome P450s (CYPs, Figure 1)<sup>4, 11</sup>. The two major groups of CYP that contribute to the structural diversity downstream of CL belong to CYP711A and CYP722C subfamily<sup>12</sup>. The best studied CYP711A is MORE AXILLARY GROWTH1 (MAX1) from *Arabidopsis thaliana* (AtMAX1), which converts CL to carlactonoic acid (CLA) and is functionally conserved in dicots<sup>13</sup>. On the other hand, monocots, especially the

economically significant *Poaceae* family, often encode more than one CYP711As (Table S1, Figure 2a, S1), with diverse functions distinct from AtMAX1<sup>13-16</sup>. For example, rice has five MAX1 homologs, with CYP711A2 catalyzing the conversion of CL to 4-deoxyorobanchol (4DO) and CYP711A3 further oxidizing 4DO to OB<sup>15</sup>. Most CYP711As encoded by monocot plants remain to be characterized. The other major group of SL-synthesizing CYPs, CYP722C subfamily, catalyze the conversion of CLA towards either OB or 5DS<sup>17-19</sup>. Currently, there are two known routes towards the synthesis of (*O*)-type SLs catalyzed by either Group I CYP722C (e.g., VuCYP722C) or OsCYP711A2<sup>15, 18</sup>, while the only known 5DS biosynthetic route is through Group II CYP722C (e.g., GaCYP722C)<sup>17</sup>. However, CYP722Cs are generally missing from the *Poaceae* family including sorghum, which implies that sorghum employs a previously unknown strategy to synthesize (*S*)-type SL.

Here, harnessing the recently developed SL-producing microbial consortia<sup>19</sup> (Figure S2), we investigated SL biosynthesis in *Sorghum bicolor*, which turns out to be distinct from that in rice<sup>15</sup>. We identified SbMAX1d as a unique CYP that catalyzes up to four oxidation steps converting CL to 18-hydroxy-CLA and a small amount of OB. Following this discovery, we found the substrate of LGS1 is likely 18-hydroxy-CLA. The addition of the sulfo group to 18-hydroxy can inhibit further oxidation towards the synthesis of OB, and 18-sulphate-CLA synthesized from LGS1 can spontaneously form comparable amount of 4DO and 5DS with sulphate functioning as an easier leaving group than the original hydroxyl. Our study discovered a second synthetic route towards the synthesis of (*S*)-type SL which employs the unique SOT LGS1. However, the enzyme catalyzing the exclusive conversion of 18-sulphate-CLA to 5DS is still missing and requires further investigation into sorghum (Figure 1).

## Results and Discussion

Same as other *Poaceae* family members, sorghum does not encode CYPs that belong to CYP722C subfamily but encode four MAX1 analogs. To understand the evolutionary relationship of these MAX1 homologues, we conducted a phylogenetic analysis of selected MAX1 analogs from dicotyledons and monocotyledons (Figure 2a, Figure S1). Noticeable, the MAX1 analogs from grasses fall into four different subclades, which are named group a-d here for simplicity (Figure 2a). Sorghum's four MAX1 analogs fall into each of the four groups, while maize and rice only encode MAX1 analogs from group a-c but not group d. To understand the biosynthetic machinery of 5DS and OB in sorghum, MAX1 analogs from sorghum (Table S1) were introduced to the CL-producing microbial consortia (**ECL**, Table S2, Figure 2b). Interestingly, expression of SbMAX1d to CL-producing consortium (**ECL/YSL2d**, Table S2) led to the synthesis of OB and 18-hydroxy-CLA (verified through high-resolution mass spectrometry [HRMS] analysis, Figure S3a). The production of 18-hydroxy-CLA by SbMAX1d is much more efficient than all the SL synthetic CYPs we examined previously (CYP722Cs and OsCYP711A2, resulting in **ECL/YSL3-5** Table S2, Figure 2b, Figure S4)<sup>18</sup>. Likely SbMAX1d first catalyzes three-step oxidation on C19 to synthesize CLA, followed by two-step oxidation on C18 to afford the synthesis of 18-hydroxy-CLA and subsequently 18-oxo-CLA, which can spontaneously convert to OB (Figure 1)<sup>18, 20</sup>. Whether the conversion from 18-hydroxy-CLA to OB is catalyzed by SbMAX1d as shunt product, or by endogenous enzymes in yeast or *E. coli* remains to be investigated.

In addition, SbMAX1c converted CL to CLA and one new peak of molecular weight same as 18-hydroxy-CLA (16Da more than that of CLA) (Figure 2b, S3b). However, due to the low titer of SLs from the microbial consortia and the lack of commercially available standards, we cannot verify

the identities of this compound synthesized by SbMAX1c currently. The other two MAX1 analogs examined simply catalyze the conversion of CL to CLA without further structural modifications (Figure 2b). The MAX1 analogs were also introduced to **ECL/YSL2d** or **ECL/YSL5** that produce 18-hydroxy-CLA and OB or 5DS (resulting strain: **ECL/YSL6-7**, Table S2) but no new conversions were detected (Figure S5). The newly discovered and unique activities of SbMAX1d and SbMAX1c imply the functional diversity of MAX1 analogs encoded by monocot plants, with much remains to be investigated.

To synthesize 5DS by Group II CYP722C (or 4DO by OsCYP711A2), likely C19 functions as the nucleophile to attack C18, which enables C18-hydroxy to recruit one proton and form water as the leaving group (Figure S6)<sup>15, 17</sup>. However, hydroxy group is generally not a favorable leaving group, and it often needs to be activated to trigger the subsequent reactions (e.g., intramolecular cyclization). Common hydroxy activation strategies used in nature include acetylation, phosphorylation, and sulfonation<sup>21-23</sup>. Sulfation/intramolecular cyclization has been reported to be employed in microbial natural product biosynthesis, such as ficellomycin from *Streptomyces ficellus*<sup>23</sup>, but seldom in plant. The discovery of the unique SbMAX1d synthesizing 18-hydroxy-CLA as the major product leads to the hypothesis that LGS1 may modify the 18-hydroxyl group to form 18-sulphate-CLA, which will prohibit further oxidation towards the formation of OB and promote the nucleophilic attack on C18 to form C ring. Introduction of LGS1 to **ECL/YSL2d** (resulting **ECL/YSL8a**, Table S2) resulted in substantial decrease of 18-hydroxy-CLA and the appearance 4DO and 5DS (ratio ~ 1:1, Figure 3a), though the amount is low in comparison to 18-hydroxy-CLA and OB (Figure 3a). Similar to many previous sulfotransferase studies<sup>24</sup>, 18-sulphate-CLA was not detected from *in vivo* assays using SL-producing microbial consortia (Figure S7). 4DO

and 5DS are synthesized in similar levels, which indicates that the conversion from 18-sulphate-CLA to the canonical SL structures are likely spontaneous with 18-sulphate as an easier leaving group than water formed from 18-hydroxy (Figure S8). There is likely other enzyme(s) involved downstream of or simultaneous with LGS1 to guarantee the conversion of 18-sulphate-CLA to 5DS exclusively instead of a 4DO/5DS mixture. We thus examined the function of SbMAX1a, 1b, 1c, SbCYP722B, SbCYP728B35, SbCYP728B1, and ZmCYP728B35 in the 4DO/5DS/18-hydroxy-CLA-producing consortium **ECL/YSL8a** (resulting **ECL/YSL9-10**, Table S2)<sup>25</sup>. Unfortunately, we were unable to see any changes to the ratio between 5DS and 4DO (Figure S9). Further genomics-based analysis on Sorghum is required to identify the missing components that is responsible for the inversion of the stereochemistry on the C ring.

To further validate the proposed mechanism of LGS1 in Sorghum SL biosynthesis (Figure S8), lysates from yeast expressing LGS1 were incubated with spent medium of CL-producing consortia expressing SbMAX1d. When LGS1 was assayed with 18-hydroxy-CLA and PAPS, 18-hydroxy-CLA was nearly completely consumed. 4DO and 5DS were observed, but not 18-sulphate-CLA, which is likely due to the low stability (Figure 4). The addition of PAPS to the lysate assay system results in enhanced consumption of 18-hydroxy-CLA and also synthesis in 4DO/5DS (Figure 4), which indicates that LGS1 is a PAPS-dependent sulfotransferase.

Similar to other plant SOTs, LGS1 is predicted to be localized in cytoplasm. Cytosolic SOTs contain several conserved PAPS-binding motifs, including the one interacts with 5'-phosphate of PAPS (TYPKSGT), 3'-phosphate of PAPS (YxxRNxxDxxVS), and nucleotide of PAPS (GxxGxxK/R)<sup>26</sup>. Multiple sequence alignment indicates that LGS1 contains these motifs, but with some variations (SLPKSGT and YxxRExxDxxVS, respectively) (Figure S10). LGS1 contains the

highly conserved histidine residues (H216)<sup>27</sup> and moderately conserved histidine residues (H317A) (Figure S10) which likely act as a base to remove the proton from the substrate hydroxyl group, thereby forming an oxygen anion, and then attacking the sulfo group of PAPS to complete the transfer of the sulfo group. To determine whether these residues play a key role in catalysis, we conducted site directed mutagenesis on residues likely act as a catalytic base (H216A, H317A) or crucial for PAPS binding (K148A, Y247F)<sup>26</sup>. While LGS1<sup>H216A</sup> (resulting strain: **YSL8f**, Table S2) exhibited same activity as wildtype LGS1, replacing LGS1 with LGS1<sup>K148A</sup>, LGS1<sup>Y247F</sup>, and LGS1<sup>H317A</sup> in **ECL/YSL8a** (resulting strain: **YSL8g-i**, Table S2) completely abolished the synthesis of 4DO and 5DS (Figure S11), implying that these residues are critical to the catalytic activity of LGS1 (Figure S11).

SOTs universally exist in all types of organisms and involve in the modification of both small molecules (e.g., steroids<sup>28</sup>) and macromolecules (e.g., glycosaminoglycans<sup>29</sup>). Among various plant SOTs, the ones from *Arabidopsis thaliana* are the most studied, with 10 out of 21 AtSOTs of known functions or substrates<sup>24,30</sup>. To examine if similar LGS1-involved SL biosynthetic mechanism exists in other plants, likely *Poaceae* plants, we used LGS1 protein sequence as a query to seek for LGS1 analogs (Table S5). We were only able to find one sulfotransferase from *Miscanthus lutarioriparius* (MISOT, 401 a.a., 80% identity) of high similarity to LGS1 (452 a.a.), while the next few on the list are all quite different from LGS1. We selected a few SOTs that exhibit highest similarity to LGS1 including MISOT, SOTs from *Triticum aestivum* (TaSOT, 345 a.a., 55% identity), and *Zea mays* (ZmSOT, 451 a.a., 53% identity), and tested the activity in **ECL/YSL8c-e** (Table S2). As expected, only MISOT was able to synthesize 5DS and 4DO but with a much lower efficiency than LGS1 (Figure S11), while ZmSOT and TaSOT didn't change the SL production profile (Figure 3a). To

further understand the evolutionary relationship between LGS1 and other plant SOTs, we constructed a phylogenetic analysis of various SOTs from plants, animals, bacteria, and fungal (Table S5, Figure 3b). As expected, LGS1 belong to plant SOT family, but are distinct from other characterized plant SOTs<sup>24</sup>. LGS1 and MISOT are located on a unique subbranch that is different from all the other plant SOTs (Figure 3b).

Multiple independent natural LGS1 loss-of-function varieties have been found in *Striga*-prevalent areas in Africa and are rare outside of *Striga*-prone region, which indicates that the lack of *LGS1* gene can adapt to weed parasitism<sup>31</sup>. *M. lutarioriparius* encode four MAX1 analogs, and each exhibits high similarity and corresponds to one of the four SbMAX1s<sup>32</sup>. Because MISOT also exhibits the same activity as LGS1, highly likely *M. lutarioriparius* harness the same LGS1-involving strategy and produces similar SL profiles to sorghum. However, the lack of LGS1 paralogs in other crops (e.g., maize) implies that much remains to be characterized about SL biosynthesis in these economically significant plants. For example, maize has been reported to produce 5DS and non-classical SLs but not (*O*)-type SLs<sup>33-35</sup>. However, same as other members from the *Poaceae* family, maize does not encode CYP722C analogs. The lack of LGS1 functional paralog thus indicates that a different synthetic route towards 5DS remains to be uncovered from maize. The activities of MAX1 analogs from maize (Table S1 and S3) were examined in different microbial consortia, as well (ECL/YSL11, Table S2). ZmMAX1b<sup>36</sup> exhibited similar activity to SbMAX1c: in addition to converting CL to CLA, it produced trace amounts of 18-hydroxy-CLA and an unknown oxidated product as SbMAX1c (Figure S12). ZmMAX1a and c showed no activity towards CL (Figure S12). Our results suggest that the 5DS biosynthesis in maize likely requires unknown types of enzymes yet to be identified.



## Conclusion

Taken together, we clarified the function of LGS1, revealed the cryptic sulfonation in the formation of the canonical SLs skeleton, and explained how the participation of LGS1 redirects the biosynthetic pathway of canonical SLs, thereby switch the product from OB into 5DS.

## Materials and Methods

**Reagents and general procedures.** ( $\pm$ )5-deoxy-strigol (purity >98%) and ( $\pm$ )-orobanchol were purchased from Strigolab (Italy), ( $\pm$ )4-deoxyorobanchol (also named as ( $\pm$ )-2'-epi-5-deoxystrigol) were purchased from Chempep Inc. (USA). Adenosine 3'-phosphate 5'-phosphosulfate lithium salt hydrate (PAPS) and the antibiotics were purchased from Sigma-Aldrich Co. (USA). The BP Clonase II Enzyme Mix, LR Clonase II Enzyme Mix, and Gateway pDONR221 vector were obtained from Invitrogen (USA). The *S. cerevisiae* Advanced Gateway Destination Vector Kit were obtained from Addgene. Expand high fidelity PCR system (Roche Life Science) were used for PCR reactions (Bio-Rad, USA). The *Escherichia coli* TOP 10 competent cells were purchased from Life Technologies. The genes were synthesized by IDT (Coralville, IA) and primers were synthesized by Life Technologies (USA). DNA sequencing were performed at Genewiz (San Diego, CA).

For CL production, XY medium [13.3 g/L  $\text{KH}_2\text{PO}_4$ , 4 g/L  $(\text{NH}_4)_2\text{HPO}_4$ , 1.7 g/L citric acid, 0.0025 g/L  $\text{CoCl}_2$ , 0.015 g/L  $\text{MnCl}_2$ , 0.0015 g/L  $\text{CuCl}_2$ , 0.003 g/L  $\text{H}_3\text{BO}_3$ , 0.0025 g/L  $\text{Na}_2\text{MoO}_4$ , 0.008 g/L  $\text{Zn}(\text{CH}_3\text{COO})_2$ , 0.06 g/L Fe(III) citrate, 0.0045 g/L thiamine, 1.3 g/L  $\text{MgSO}_4$ , 5 g/L yeast extract and 40 g/L xylose, pH 7.0] was prepped and used as previously described<sup>19</sup>. For yeast ectopic expression, synthetic dropout (SD) medium was used [0.425g yeast nitrogen base (YNB, BD Biosciences, USA), 1.25g  $(\text{NH}_4)_2\text{SO}_4$  dissolved in 200 mL distilled water ( $\text{dH}_2\text{O}$ ), autoclave at 121 °C

for 20 min. Add 25 mL 200 g/L glucose and 25 mL 20 g/L amino acid drop-out mix (Clontech. Inc., Japan) solution to prepare the medium]. LC-MS was carried out on a Shimadzu LC-MS 2020 (Kyoto, Japan) with LC-MS grade solvent. High resolution mass spectrometry (HR-MS) analysis was carried on a Synapt G2-Si quadrupole time-of-flight mass spectrometer (Waters) coupled to an I-class UPLC system (Waters).

**Plasmid construction.** All genes were codon optimized for *S. cerevisiae* (Table S6), synthesized, and cloned into the entry vector pDONR221 (Invitrogen) through Gateway® BP reaction. The genes were then introduced to the yeast expression vector through Gateway® LR reaction using destination vectors from Yeast Gateway Kit<sup>37</sup>. LGS1 mutants were constructed through PCR using primers listed in Table S7. PCR was performed using pAG416GPD-LGS1 as the template with Expand high fidelity PCR system. The amplified DNA fragment was purified, recovered using Gibson assembly, and used to transform TOP 10 competent cells. The sequence was verified using *DNA sequencing*.

**Culture conditions for *E. coli*-yeast consortium-based SL production.** The *E. coli* strain ECL for carlactone production (Table S2) were prepared as described previously<sup>19</sup>. Single colony was grown overnight at 37 °C in 1 mL Luria-Bertani (LB) containing 25 µg/mL chloramphenicol, 50 µg/mL spectinomycin and 100 µg/mL ampicillin. 500 µL of the overnight culture was then used to inoculate 5 mL of fresh LB with the corresponding antibiotics and cultured at 37 °C and 220 rpm in the 100 mL Erlenmeyer flask. When OD<sub>600</sub> reached ~0.6, IPTG was added with the final concentration at 0.2 mM, with ferrous sulfate supplemented at the same time (final concentration at 10 mg/L). Then the cultures were incubated at 22°C and 220 rpm for 15 hours. At the same time, single colony of each yeast strain harboring the corresponding cytochrome P450-expression constructs was used to inoculate 1 mL synthetic dropout medium (SDM). The seed culture was incubated at 28 °C and 220 rpm overnight. 100 µL of the overnight grown seed culture was used to inoculate 5 mL of the corresponding SD medium in a 100 mL Erlenmeyer flask and grown at 28 °C

for 15 hours. The *E. coli* and yeast cells were harvested by centrifugation at 3,500 rpm for 5 min. Then they were re-suspended together in 5 mL of XY media and grown at 22°C for 60 hours.

**Metabolite extraction and analysis.** Collect the cell pellet through centrifugation in a 50 mL centrifuge tube at 4,000 rpm for 10 min. The supernatant is transferred to another 50 mL centrifuge tube. The cell pellets were resuspended in 150  $\mu$ L dimethylformamide (DMF) and 850  $\mu$ L acetone, vortexed for 15 min, and centrifuged at 12,000 rpm for 10 min. The medium was extracted with 4 mL of ethyl acetate (EtOAc), vortex for 1 min, and centrifuged at 4,000 rpm for 20 min. The supernatant from both portions were collected and transferred to 1.7 mL microcentrifuge tubes, dried in vacuuge under reduced pressure (Eppendorf concentrator plus), and re-dissolve in 100  $\mu$ L acetone. The extract was centrifuged at 12,000 rpm for 10 min before *LC-MS* analysis. The *LC-MS* analysis were performed on a C<sub>18</sub> column (Kinetex® C18, 100 mm  $\times$  2.1 mm, 100Å, particle size 2.6  $\mu$ m; Phenomex, Torrance, CA, USA). Separation Method I, the parameters were set as follows: column temperature: 40 °C, flow rate: 0.4 mL/min; mobile phase A: water containing 0.1% (v/v) formic acid; mobile phase B: acetonitrile containing 0.1% (v/v) formic acid. The LC program was set as follows: 0-28 min, 5%-100% B; 28-35 min, 100% B; 35-40 min, 5% B.

**Heterologous expression of recombinant LGS1 in *S. cerevisiae*.** We first expressed LGS1 from *E. coli* strain BL21 (DE3) but failed. Then LGS1 was expressed from *S. cerevisiae* CEN.PK2 using low-copy number plasmid (pAG416GPD-LGS1, Table S1). One fresh colony of the LGS1-expressing yeast strain was first cultured in 1mL SDM lacking uracil (SD-Ura) medium, grown at 30°C and 220 rpm for overnight in a shaker incubator. 80  $\mu$ L of the overnight culture was used to inoculate 5 mL SD-Ura medium ( $OD_{600} \approx 0.1$ ), grown at 30°C and 220 rpm for 48 h ( $OD_{600} \approx 20$ ). Cell pellets were then harvested by centrifuging at 3,500 rpm for 2 min, washed with 1 mL of water,

and resuspend in 120  $\mu$ L of 20 mM sodium phosphate buffer (pH=7.4). 50  $\mu$ L of silicon beads [0.5 mm, Research Products International (RPI)] were then added to the cell suspension, which is then chilled on ice, and lysed using cell disruptor (FastPrep<sup>®</sup>-24, MP Biomedicals), The parameters were set as speed: 4.0 m/s, time: 30 s. The homogenate was centrifuge at 13,000 rpm for 2 min, and the supernatant was used for the crude lysate-based enzyme assays.

**Yeast crude lysate-based enzyme assays.** 50  $\mu$ L of crude enzyme extract mentioned above is incubated with 5  $\mu$ L of concentrated metabolic extract dissolved in DMF (extracted from 3 mL co-culture strain), with or without 100  $\mu$ M PAPS, and incubated at 30°C for 1 hour. Enzyme assay using yeast strain expressing an empty vector as the negative control. The reaction mixture was quenched by adding an equal volume of acetonitrile followed by vigorous vortexing to remove the protein. The quenched reaction mixtures were then centrifuged at 13,000 rpm for 2 min. 17  $\mu$ L of supernatant was subjected to LC-MS analysis with the C<sub>18</sub> column (Kinetex<sup>®</sup> C18, 100 mm  $\times$  2.1 mm, 100Å, particle size 2.6  $\mu$ m; Phenomex, Torrance, CA, USA). In order to detect speculative 18-sulphate-CLA, an intermediate with an increased polarity, we use a different separation method: Separation Method II. The parameters were set as follows: column temperature: 40 °C, flow rate: 0.4 mL/min; mobile phase A: water containing 0.1% (v/v) formic acid; mobile phase B: acetonitrile containing 0.1% (v/v) formic acid. The LC program was set as follows: 0-3 min, 5%-11% B; 3-13 min, 11%-19% B; 13-21 min, 19%-27.5% B; 21-24 min, 27.5%-34% B; 24-28 min, 34%-42% B; 28-32min, 42%-90% B; 32-34 min, 90%-100% B; 34-35.5 min, 100%-5% B; 35.5-40 min, 5% B.

## ASSOCIATED CONTENT

### Supporting Information.

Supplementary figures on sorghum SL biosynthesis characterization and LGS1 identification, supplementary tables summarizing plasmids, strains, and gene sequences used in this study (PDF)

## **AUTHOR INFORMATION**

### **Corresponding Author**

\*Yanran Li; E-mail: [yanranl@ucr.edu](mailto:yanranl@ucr.edu).

### **Author Contributions**

S.W and Y.L. conceived the project; S.W. and Y.L. designed the experiments; S.W. performed the experiments and analyzed the results; S.W and Y.L. wrote the manuscript.

### **Funding Sources**

This work is supported by Cancer Research Coordinating Committee Research Award (grant to Y.L., CRN-20-634571).

### **Notes**

Competing financial interests: Y.L. and W.S. filed a provisional patent application on Jan. 28, 2021, "Strigolactone-Producing Microbes and Methods of Making and Using the Same," U.S. Provisional Application No. 63/142,801.

## **ACKNOWLEDGMENT**

We thank the Metabolomics Core Facility at UC Riverside and Dr. Anil Bhatia for instrument access, training, and data analysis; S. Xu for studying protein-protein interaction of SL biosynthetic enzymes identified in this study; A. Zhou for the construction of SYL89; and Dr. K. Zhou for the valuable feedback in the preparation of the manuscript.

## ABBREVIATIONS

5DS, 5-deoxystrigol; 4DO, 4-deoxyorobanchol; OB, orobanchol; SL, Strigolactone; CL, carlactone; CLA, carlactonoic acid; SOT or SULT, sulfotransferase; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; CYP, cytochrome P450; MAX1, MORE AXILLARY GROWTH 1; LGS1, LOW GERMINATION STIMULANT 1; HRMS, high-resolution mass spectrometry.

## REFERENCES

- (1) Cook, C. E., Whichard, L.P., Turner, B., Wall, M.E. Egley, G.H. Germination of witchweed (*Striga lutea* Lour.) : isolation and properties of a potent stimulant. *Science* **1966**, *154*, 1189-1190.
- (2) Samejima, H.; Babiker, A. G.; Mustafa, A.; Sugimoto, Y. Identification of *Striga hermonthica*-resistant upland rice varieties in Sudan and their resistance phenotypes. *Front. Plant Sci.* **2016**, *7*, 634.
- (3) Al-Babili, S.; Bouwmeester, H. J. Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* **2015**, *66*, 161-186.
- (4) Chesterfield, R. J.; Vickers, C. E.; Beveridge, C. A. Translation of strigolactones from plant hormone to agriculture: achievements, future perspectives, and challenges. *Trends Plant Sci.* **2020**, *25*, 1087-1106.
- (5) Zwanenburg, B.; Blanco-Ania, D. Strigolactones: new plant hormones in the spotlight. *J. Exp. Bot.* **2018**, *69*, 2205-2218.
- (6) Yoneyama, K.; Awad, A. A.; Xie, X.; Yoneyama, K.; Takeuchi, Y. Strigolactones as germination stimulants for root parasitic plants. *Plant Cell Physiol.* **2010**, *51*, 1095-1103.
- (7) Zwanenburg, B.; Pospisil, T. Structure and activity of strigolactones: new plant hormones with a rich future. *Mol. Plant* **2013**, *6*, 38-62.
- (8) Nomura, S.; Nakashima, H.; Mizutani, M.; Takikawa, H.; Sugimoto, Y. Structural requirements of strigolactones for germination induction and inhibition of *Striga gesnerioides* seeds. *Plant Cell Rep.* **2013**, *32*, 829-838.
- (9) Gobena, D.; Shimels, M.; Rich, P. J.; Ruyter-Spira, C.; Bouwmeester, H.; Kanuganti, S.; Mengiste, T.; Ejeta, G. Mutation in sorghum *LOW GERMINATION STIMULANT 1* alters strigolactones and causes *Striga* resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 4471-4476.
- (10) Paul, P.; Suwan, J.; Liu, J.; Dordick, J. S.; Linhardt, R. J. Recent advances in sulfotransferase enzyme activity assays. *Anal. Bioanal. Chem.* **2012**, *403*, 1491-1500.
- (11) Wang, Y.; Bouwmeester, H. J. Structural diversity in the strigolactones. *J. Exp. Bot.* **2018**, *69*, 2219-2230.
- (12) Nelson, D. R.; Ming, R.; Alam, M.; Schuler, M. A. Comparison of cytochrome P450 genes from six plant genomes. *Trop. Plant Biol.* **2008**, *1*, 216-235.
- (13) Challis, R. J.; Hepworth, J.; Mouchel, C.; Waites, R.; Leyser, O. A role for *MORE AXILLARY GROWTH1* (*MAX1*) in evolutionary diversity in strigolactone signaling upstream of

MAX2. *Plant Physiol.* **2013**, *161*, 1885-1902.

(14) Marzec, M.; Situmorang, A.; Brewer, P. B.; Braszewska, A. Diverse roles of MAX1 homologues in rice. *Genes* **2020**, *11*, 1348.

(15) Zhang, Y.; van Dijk, A. D.; Scaffidi, A.; Flematti, G. R.; Hofmann, M.; Charnikhova, T.; Verstappen, F.; Hepworth, J.; van der Krol, S.; Leyser, O.; Smith, S. M.; Zwanenburg, B.; Al-Babili, S.; Ruyter-Spira, C.; Bouwmeester, H. J. Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nat. Chem. Biol.* **2014**, *10*, 1028-1033.

(16) Changenet, V.; Macadre, C.; Boutet-Mercey, S.; Magne, K.; Januario, M.; Dalmais, M.; Bendahmane, A.; Mouille, G.; Dufresne, M. Overexpression of a cytochrome P450 monooxygenase involved in orobanchol biosynthesis increases susceptibility to fusarium head blight. *Front. Plant Sci.* **2021**, *12*, 662025.

(17) Wakabayashi, T.; Shida, K.; Kitano, Y.; Takikawa, H.; Mizutani, M.; Sugimoto, Y. CYP722C from *Gossypium arboreum* catalyzes the conversion of carlactonic acid to 5-deoxystrigol. *Planta* **2020**, *251*, 97.

(18) Wakabayashi, T.; Hamana, M.; Mori, A.; Akiyama, R.; Ueno, K.; Osakabe, K.; Osakabe, Y.; Suzuki, H.; Takikawa, H.; Mizutani, M.; Sugimoto, Y. Direct conversion of carlactonic acid to orobanchol by cytochrome P450 CYP722C in strigolactone biosynthesis. *Sci. Adv.* **2019**, *5*, eaax9067.

(19) Wu, S.; Ma, X.; Zhou, A.; Valenzuela, A.; Zhou, K.; Li, Y. Establishment of strigolactone-producing bacterium-yeast consortium. **2021**, doi: <https://doi.org/10.1101/2021.06.29.450423>

(20) Mori, N.; Sado, A.; Xie, X.; Yoneyama, K.; Asami, K.; Seto, Y.; Nomura, T.; Yamaguchi, S.; Yoneyama, K.; Akiyama, K. Chemical identification of 18-hydroxycarlactonic acid as an LjMAX1 product and in planta conversion of its methyl ester to canonical and non-canonical strigolactones in *Lotus japonicus*. *Phytochemistry* **2020**, *174*, 112349.

(21) Chen, X.; Hagel, J. M.; Chang, L.; Tucker, J. E.; Shiigi, S. A.; Yelapaala, Y.; Chen, H. Y.; Estrada, R.; Colbeck, J.; Enquist-Newman, M.; Ibanez, A. B.; Cottarel, G.; Vidanes, G. M.; Facchini, P. J. A pathogenesis-related 10 protein catalyzes the final step in thebaine biosynthesis. *Nat. Chem. Biol.* **2018**, *14*, 738-743.

(22) Muller, W. M.; Schmiederer, T.; Ensle, P.; Sussmuth, R. D. In vitro biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification. *Angew. Chem., Int. Ed.* **2010**, *49*, 2436-2440.

(23) Yue, R.; Li, M.; Wang, Y.; Guan, Y.; Zhang, J.; Yan, Z.; Liu, F.; Lu, F.; Zhang, H. Insight into enzyme-catalyzed aziridine formation mechanism in ficellomycin biosynthesis. *Eur. J. Med. Chem.* **2020**, *204*, 112639.

(24) Hirschmann, F.; Krause, F.; Papenbrock, J. The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions. *Front. Plant Sci.* **2014**, *5*, 556.

(25) Wakabayashi, T.; Ishiwa, S.; Shida, K.; Motonami, N.; Suzuki, H.; Takikawa, H.; Mizutani, M.; Sugimoto, Y. Identification and characterization of sorgomol synthase in sorghum strigolactone biosynthesis. *Plant Physiol.* **2021**, *185*, 902-913.

(26) Xie, L.; Xiao, D.; Wang, X.; Wang, C.; Bai, J.; Yue, Q.; Yue, H.; Li, Y.; Molnár, I.; Xu, Y.; Zhang, L.; Mitchell Aaron, P. Combinatorial biosynthesis of sulfated benzenediol lactones with a phenolic sulfotransferase from *Fusarium graminearum* PH-1. *mSphere* **2020**, *5*, e00949-20.

- (27) Landi, S.; Esposito, S. Bioinformatic characterization of sulfotransferase provides new insights for the exploitation of sulfated polysaccharides in *Caulerpa*. *Int. J. Mol. Sci.* **2020**, *21*, 6681.
- (28) Marsolais, F.; Boyd, J.; Paredes, Y.; Schinas, A. M.; Garcia, M.; Elzein, S.; Varin, L. Molecular and biochemical characterization of two brassinosteroid sulfotransferases from *Arabidopsis*, AtST4a (At2g14920) and AtST1 (At2g03760). *Planta* **2007**, *225*, 1233-1244.
- (29) Kusche-Gullberg, M.; Kjellén, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* **2003**, *13*, 605-611.
- (30) Chan, K. X.; Phua, S. Y.; Van Breusegem, F. Secondary sulfur metabolism in cellular signalling and oxidative stress responses. *J. Exp. Bot.* **2019**, *70*, 4237-4250.
- (31) Bellis, E. S.; Kelly, E. A.; Lorts, C. M.; Gao, H.; DeLeo, V. L.; Rouhan, G.; Budden, A.; Bhaskara, G. B.; Hu, Z.; Muscarella, R.; Timko, M. P.; Nebie, B.; Runo, S. M.; Chilcoat, N. D.; Juenger, T. E.; Morris, G. P.; dePamphilis, C. W.; Lasky, J. R. Genomics of sorghum local adaptation to a parasitic plant. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 4243-4251.
- (32) Miao, J.; Feng, Q.; Li, Y.; Zhao, Q.; Zhou, C.; Lu, H.; Fan, D.; Yan, J.; Lu, Y.; Tian, Q.; Li, W.; Weng, Q.; Zhang, L.; Zhao, Y.; Huang, T.; Li, L.; Huang, X.; Sang, T.; Han, B. Chromosome-scale assembly and analysis of biomass crop *Miscanthus lutarioriparius* genome. *Nat. Commun.* **2021**, *12*, 2458.
- (33) Charnikhova, T. V.; Gaus, K.; Lumbroso, A.; Sanders, M.; Vincken, J.-P.; De Mesmaeker, A.; Ruyter-Spira, C. P.; Screpanti, C.; Bouwmeester, H. J. Zeapyranolactone – A novel strigolactone from maize. *Phytochem. Lett.* **2018**, *24*, 172-178.
- (34) Charnikhova, T. V. G., K. Lumbroso, A. Sanders, M. Vincken, J. P. De Mesmaeker, A. Ruyter-Spira, C. P. Screpanti, C. Bouwmeester, H. J. Zealactones. Novel natural strigolactones from maize. *Phytochemistry* **2017**, *137*, 123-131.
- (35) Awad, A. A.; Sato, D.; Kusumoto, D.; Kamioka, H.; Takeuchi, Y.; Yoneyama, K. Characterization of strigolactones, germination stimulants for the root parasitic plants *Striga* and *Orobanche*, produced by maize, millet and sorghum. *Plant Growth Regul.* **2006**, *48*, 221.
- (36) Yoneyama, K.; Mori, N.; Sato, T.; Yoda, A.; Xie, X.; Okamoto, M.; Iwanaga, M.; Ohnishi, T.; Nishiwaki, H.; Asami, T.; Yokota, T.; Akiyama, K.; Yoneyama, K.; Nomura, T. Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis. *New Phytol.* **2018**, *218*, 1522-1533.
- (37) Alberti, S., Gitler, A.D., Lindquist, S., A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* **2007**, *24* (10), 913-919.



## Figure legends

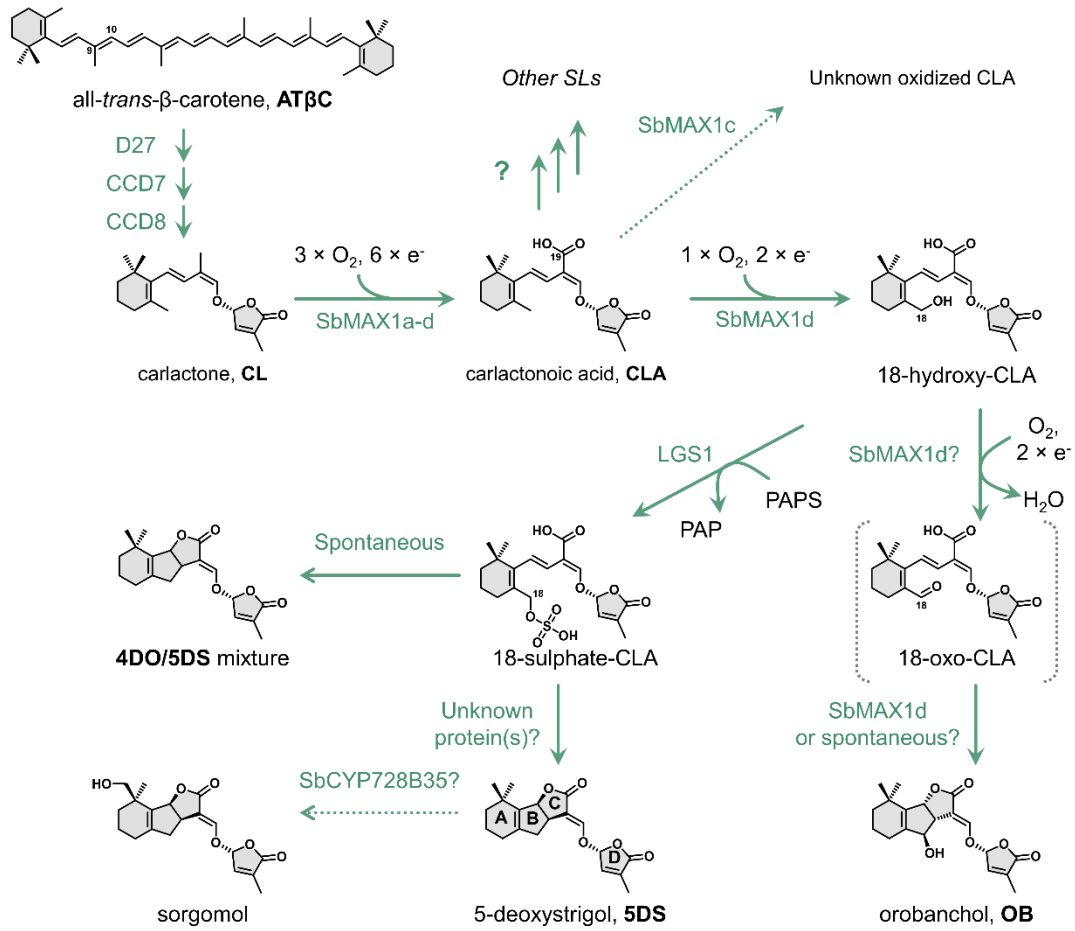
**Figure 1.** The proposed biosynthetic pathway of 5DS and OB in *Sorghum bicolor*. D27, [2Fe-2S]-containing isomerase DWARF27; CCD7, carotenoid cleavage dioxygenase 7; CCD8, carotenoid cleavage dioxygenase 8; SbMAX1d, MAX1 analog d from *S. bicolor*; LGS1, *LOW GERMINATION STIMULANT 1*, a sulfotransferase; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PAP, 3'-phosphoadenosine-5'-phosphate; 4DO, 4-deoxyorobanchol.

**Figure 2.** Functional characterization of MAX1 analogs from *S. bicolor*. (a) Phylogenetic analysis of MAX1 analogs. The Phylogenetic tree was reconstructed in MEGA X using the neighbor-joining method based on amino acid sequence. The MAX1 analogs are from dicotyledons and monocotyledons. Species abbreviations: Sb, *Sorghum bicolor*; Ml, *Miscanthus lutarioriparius*; Zm, *Zea mays*; Bd, *Brachypodium distachyon*; Os, *Oryza sativa*; Amt, *Amborella trichopoda*; At, *Arabidopsis thaliana*; Ac, *Aquilegia coerulea*; Pg, *Picea glauca*; Sm, *Selaginella moellendorffii*. For the accession numbers of proteins, see Table S4. (b) SIM extracted ion chromatogram (EIC) at  $m/z=331.1$  (green), 347.1 (purple), and  $m/z^+=331.1$  (orange), 347.1 (blue) of CL-producing *E. coli* cocultured with *A. thaliana* P450 reductase 1 (ATR1)-expressing yeast i) expressing AtMAX1, ii) – v) expressing SbMAX1a-d, and vi) standards of OB, 4DO, and 5DS. CLA shows characteristic  $m/z=331.1$  (MW=332.40,  $[C_{19}H_{24}O_5-H]^-=[C_{19}H_{23}O_5]^-$ ); 18-OH-CLA shows characteristic  $m/z=347.1$  and  $m/z^+=331.1$  (MW=348.40,  $[C_{19}H_{24}O_6-H]^-=[C_{19}H_{23}O_6]^-$ ); 4DO and 5DS show characteristic  $m/z^+$  signal (MW=330.38,  $[C_{19}H_{22}O_5+H]^+=[C_{19}H_{23}O_5]^+$ ). All traces are representative of at least three biological replicates for each engineered *E. coli-S. cerevisiae* consortium. 18-OH-CLA stands for 18-hydroxy-CLA. MW stands for molecular weight. Strain used for analysis: AtMAX1 (ECL/YSL1, Table S2), SbMAX1a-d (ECL/YSL2a-d, Table S2).

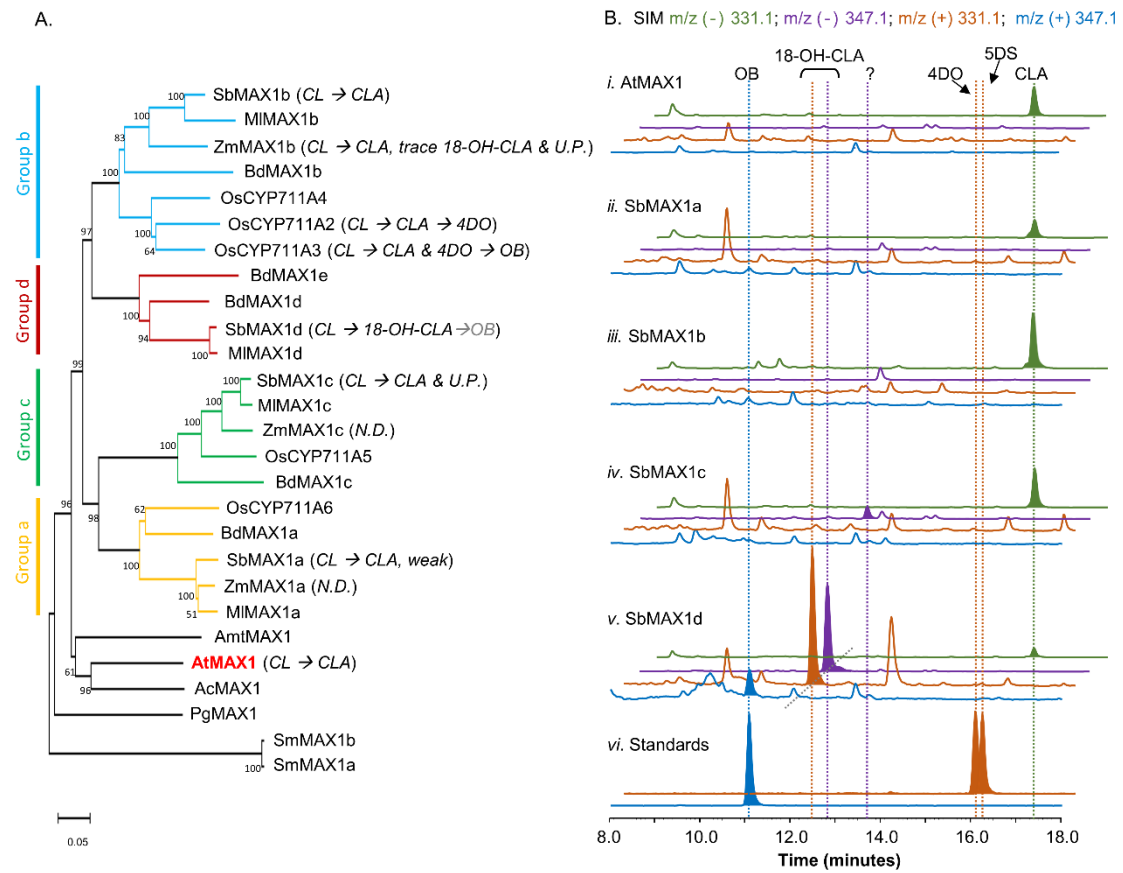
**Figure 3.** Functional characterization of LGS1 and analogs using CL-producing microbial consortium expressing SbMAX1d. (a) SIM EIC at  $m/z=331.1$  (green), 347.1 (purple), and  $m/z^+=331.1$  (orange), 347.1 (blue) of CL-producing *E. coli* cocultured with yeast expressing ATR1, SbMAX1d and i) empty vector (EV), ii) LGS1, iii) LGS1-2, iv) SOT from *Triticum aestivum* (TaSOT), v) SOT from *Zea mays* (ZmSOT), and vi) standards of OB, 4DO and 5DS. All traces are representative of at least three biological replicates for each engineered *E. coli-S. cerevisiae* consortium. (b) Phylogenetic analysis of LGS1. The Phylogenetic tree was reconstructed in MEGA X using the neighbor-joining method based on amino acid sequence. The SOTs are from animals, plants, fungi and cyanobacteria. For the accession numbers of proteins, see Table S5. The sequence of LGS1 is from sorghum WT Shanqui Red, LGS1-2 variation is a reference sequence from NCBI, and is four amino acids (DADD) longer than LGS1, see Table S6.

**Figure 4.** Characterization of LGS1 activity using crude lysate assay. SIM EIC at  $m/z=347.1$  (purple) and  $m/z^+=331.1$  (orange) of crude lysate assay using i) EV-harboring yeast with PAPS, ii) LGS1-expressing yeast without PAPS, iii) LGS1-expressing yeast and PAPS, iv) authentic standard of 4DO and 5DS. The reaction was incubated for 1 hour with extracts of ECL/YSL2d medium, the samples were analyzed using Separation Method II (extraction method see Material and Methods in Supporting Information).

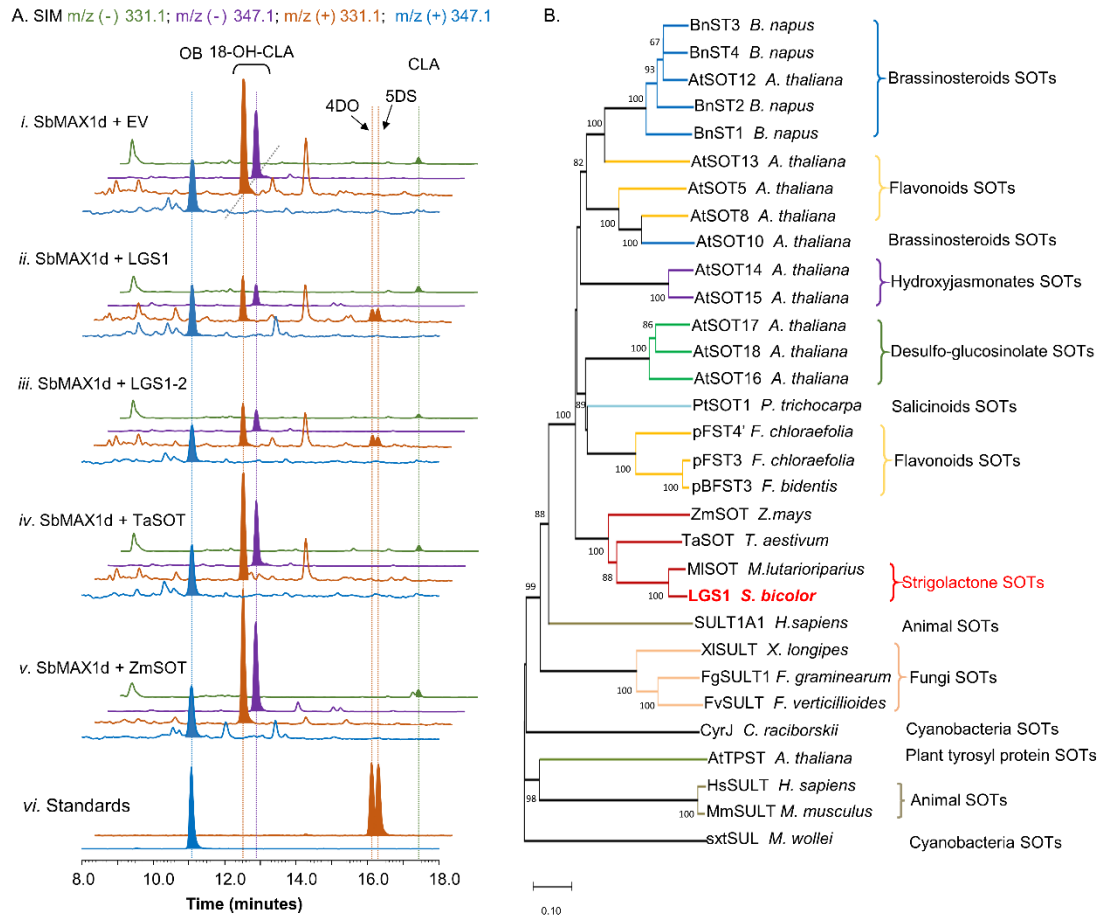
**Figure 1.**



**Figure 2.**



**Figure 3.**



**Figure 4.**

SIM  $m/z$  (-) 347.1;  $m/z$  (+) 331.1

