Science Author Reprints and Color Reimbursement

Administered and Produced by The Sheridan Press
Cindy Eyler, Reprint Customer Service Rep
The Sheridan Press, 450 Fame Avenue, Hanover, PA 17331
(800) 635-7181 ext. # 8008, (717) 632-3535, or fax (717) 633-8929 cindy.eyler@sheridan.com

Order reprints and pay color figure charges online at www.sheridan.com/aaas/eoc

Dear Author:
Science has a combined online form for ordering reprints and paying charges on color figures. You will need to use this form to pay for or receive an invoice for your color figure charges. To start your order, you'll need to enter the last seven digits of the DOI of your paper (this is a 7-digit number at the end of the references of your paper). After filling out the order form, an email will be sent for your records. An invoice will be sent with the reprints. You can pay at the time of your order, indicate that you have a purchase order, or ask to be billed.

Reimbursement for Use of Color in Science

As stated in Information for Contributors and your acceptance letter, authors requesting the use of color are required to pay $650 for the first color figure and $450 for each additional figure to help defray costs related to publishing color in the *Science* issue. **These charges are not related to your reprint order**, but are billed on the same form. Authors of solicited Reviews, Special Issue Perspectives, and Special Issue Reviews are exempt from these charges.

Printed Reprints

Author reprints must be used solely for the author's personal use. If commercial or for-profit use is intended, please contact Rockwater, Inc. at brocheleau@rockwaterinc.com or (803) 359-4578.

Only one invoice will be issued for group orders to multiple locations. Additional order forms may be obtained by contacting The Sheridan Press. **All orders must be received within 60 days of publication date or additional charges will apply.**

Prepayment or an institution purchase order is required to process your order. The online form will provide an invoice.

Delivery
Your order will be shipped within 3 weeks of the *Science* issue date. Allow extra time for delivery. If quicker delivery is necessary, please call for pricing and availability. UPS ground postage and handling are included in the prices (1-5 day delivery). Orders shipped to authors outside the continental US are mailed via an expedited air service at an additional charge. Orders for articles over 1 year past publication will require additional time to produce.

Corrections
If a serious error occurs in the published version of the paper, the error can be corrected in reprints if the editorial office is notified promptly. Please contact the editor or copy editor of your paper with the corrections.

Reprint Order Specifications
All reprints will include either a title page (in black and white or color, depending upon the type of reprint ordered) or the cover of *Science* from the issue in which your article appears. If the cover of *Science* is selected, there will be a $100 additional fee (cover will appear in black & white or color, depending on the type of reprints ordered). This adds one page to the length of your paper.

Pricing
Reprint pricing is shown in the following tables. Orders are limited to 500 copies per author. To convert color articles to black & white reprints, add $200. For articles over 12 months past publication, please contact The Sheridan Press for pricing. **This pricing is valid within 60 days of publication date.**

Black and White Reprints

Quantity	100	200	300	400	500
≤4 pages	300	350	395	435	470
≤8 pages	460	520	575	625	670
≤12 pages	600	665	725	780	830

Color Reprints

Quantity	100	200	300	400	500
≤4 pages	1800	1890	1960	2020	2070
≤8 pages	2140	2235	2310	2375	2430
≤12 pages	2485	2585	2665	2735	2795

Air Shipping Charges
(orders shipped outside the continental US only)

$120 – 8 pages or less and 200 copies or less
$175 – More than 8 pages or more than 200 copies

Instructions for Handling PDF Galley Proofs

It is important that you return galley corrections within 48 hours directly to your copy editor. Please let your copy editor know immediately if there will be any delay.

Dear Author:

Thank you for publishing in *Science*. This letter explains how to mark this PDF file and transmit corrections to your galley proofs. This PDF file includes the following:

1. Instructions for ordering reprints and paying for use of color in figures (p. 1).
2. Detailed instructions for marking the proof (pp. 2-4)
3. Galley proofs of your paper (starting on p. 5).

If your manuscript contains color figures, the colors and resolution in the proofs may appear different from those of the final published figures. Our art department will send separate color figure proofs. Also, although your paper begins at the top of a page in the proofs, it may not when printed in *Science*.

A separate PDF file showing editorial changes to your paper has been, or shortly will be, e-mailed to you by your copy editor. Please use this in checking your proofs, but do not mark corrections on it.

In order to make galley corrections:

1. Please pay color figure charges online at www.sheridan.com/aaas/eoc. The same online form can be used to order reprints.
2. Mark all changes on the galley proofs (this file) directly using Acrobat Reader (free) v. X (available at: http://get.adobe.com/reader/).
3. Additional instructions for marking text are given on the next two pages. To start, select the Comment button on the upper right side of the Adobe Reader screen.
   a. All edits of the text and text corrections should be made with "Text Edits" using insert, replace, or delete/cross out selections. *Please do not use sticky notes, comments, or other tools for actual text edits*.
   b. You can control formatting (e.g., italics or bold) by selecting edited text.
   c. Indicate edits to special or Greek characters with a comment. Use the "sticky note" for comments. Please refrain from using other tools.
   d. Please collect all corrections into one file; import comments provided by multiple authors into one file only (Document menu/Comments/Import Comments).
   e. Be sure to save the marked file and keep a copy.
4. Respond to all of the copy editor's queries listed at the end of the edited manuscript or as embedded PDF notes in the copyedited manuscript, by directly editing the galley text in the PDF with the annotation tools or using sticky notes on the galley PDF.
5. Check reference titles and additional supplementary references on the edited word file sent to you by your copyeditor. Address any edits here in the note back to your copyeditor.
6. Check all equations, special characters, and tables carefully. Check spelling of all author names for accuracy.
7. Make a copy of the corrected galley proofs for yourself. Return the corrected galley proofs to the *Science* copy editor as an attachment to an email.
8. If you cannot mark the proofs electronically, please e-mail a list of corrections to the copy editor.

Thank you for your prompt attention,
The Editors

Headquarters
1200 New York Avenue, NW Washington, DC 20005 USA Telephone: (+1) 202-326-6550 Fax (+1) 202-789-4669
Europe Office
Bateman House, 82-88 Hills road, Cambridge CB2 1LQ, UK Telephone: (+44) 1223-326500 Fax: (+44) 1223 326501
Published by the American Association for the Advancement of Science

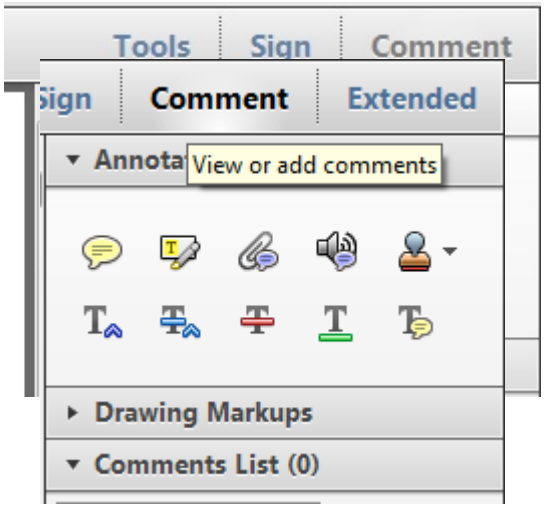
E-Corrections through PDF Annotation

Instructions on how to annotate your galley PDF file using Adobe Acrobat Reader X

To view, annotate and print your galley, you will need Adobe Reader X. This free software can be downloaded from: http://get.adobe.com/reader/. It is available for Windows, Mac, LINUX, SOLARIS, and Android. The system requirements can also be found at this URL.

To make corrections and annotations in your galley PDF with Adobe Reader X, use the commenting tools feature, located by clicking **Comment** at the upper right of your screen. You should then see the Annotations Palette with the following annotation tools. (These tools can also be accessed through View>Comment>Annotations.

Although all the files from SPI will have these commenting tools available, occasionally this feature will not be enabled on a particular PDF. In these cases, you can use the two default commenting tools to annotate your files: Sticky Note and Highlight Text.



To start adding comments, select the appropriate commenting tool from the Annotation Palette.

<u>TO INDICATE INSERT, REPLACE, OR REMOVE TEXTS</u>

- **Insert Text**

  Click the button on the Commenting Palette. Click to set the cursor location in the text and start typing. The text will appear in a commenting box. You may also cut-and-paste text from another file into the commenting box.
- **Replace Text**

  Click the button on the Commenting Palette. To highlight the text to be replaced, click and drag the cursor over the text. Then type in the replacement text. The replacement text will appear in a commenting box. You may also cut-and-paste text from another file into this box.
- **Remove Text**

  Click the button on the Commenting Palette. Click and drag over the text to be deleted. The text to be deleted will then emphasize with a ~~strikethrough~~.

LEAVE A NOTE / COMMENT

- **Add Note to Text**

Click the button on the Commenting Palette. Click to set the location of the note on the document and simply start typing. Kindly refrain from using this feature to make text edits

- **Add Sticky Note**

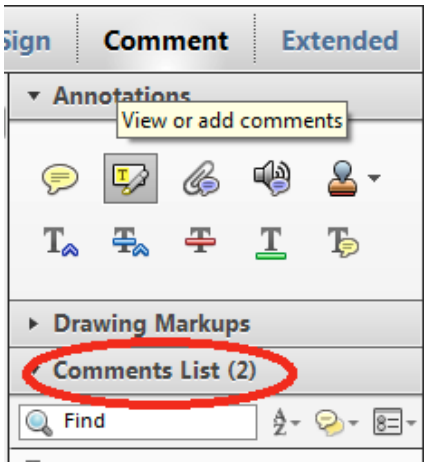
Click the button on the Commenting Palette. Click to set the location of the note on the document and simply start typing. Kindly refrain from using this feature to make text edits

HIGHLIGHT TEXT / MAKE A COMMENT

- Click the button on the Commenting Palette. Click and drag over the text. To make a comment, double click on the highlighted text and simply start typing.

REVIEW

All comments added in the active document are listed in **Comments List** Palette. Navigate by clicking on a correction in the list.



ATTACH A FILE

For equations, tables and figures that need to be added or replaced, or for a large section of text that needs to be inserted, users will find it better to just attach a file.

Click button on the Commenting Palette. And then click on the figure, table or formatted text to be replaced. A window will automatically open allowing you to attach the file.

RESEARCH ARTICLE

PLANT SCIENCE

A specialized metabolic network selectively modulates *Arabidopsis* root microbiota

Ancheng C. Huang[1]*, Ting Jiang[2,3,4]*, Yong-Xin Liu[2,3], Yue-Chen Bai[5,6], James Reed[1], Baoyuan Qu[2,3], Alain Goossens[5,6], Hans-Wilhelm Nützmann[1]†, Yang Bai[2,3,4]‡, Anne Osbourn[1]‡

Plant specialized metabolites have ecological functions, yet the presence of numerous uncharacterized biosynthetic genes in plant genomes suggests that many molecules remain unknown. We discovered a triterpene biosynthetic network in the roots of the small mustard plant *Arabidopsis thaliana.* Collectively, we have elucidated and reconstituted three divergent pathways for the biosynthesis of root triterpenes, namely thalianin (seven steps), thalianyl medium-chain fatty acid esters (three steps), and arabidin (five steps). *A. thaliana* mutants disrupted in the biosynthesis of these compounds have altered root microbiota. In vitro bioassays with purified compounds reveal selective growth modulation activities of pathway metabolites toward root microbiota members and their biochemical transformation and utilization by bacteria, supporting a role for this biosynthetic network in shaping an *Arabidopsis*-specific root microbial community.

Plants have evolved the ability to produce a vast array of specialized metabolites. This metabolic diversification is likely to have been driven by the need to adapt to different environmental niches (*1*). One of the key environmental factors that influences plant health and fitness is the root microbiota (*2*–*4*), yet mechanisms underpinning the establishment of specific root microbiota and how plants direct the microbial communities around their roots remain elusive. Plants are estimated to use ~20% of their photosynthesized carbon to make root-derived organic molecules, stimulating the formation of distinctive root microbiota from surrounding soil (*5*–*8*). However, whether and which specialized metabolites produced by plants can modulate root microbiota is not known. The presence of a substantial number of uncharacterized root-expressed biosynthetic genes in plant genomes also reveals our incomplete understanding of these communication processes.

[1]Department of Metabolic Biology, John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR4 7UH, UK. [2]State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Science, Beijing, China. [3]CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (CAS), Beijing 100101, China. [4]University of Chinese Academy of Sciences, College of Advanced Agricultural Sciences, Beijing 100039, China. [5]Ghent University, Department of Plant Biotechnology and Bioinformatics, Ghent 9052, Belgium. [6]VIB Center for Plant Systems Biology, Ghent 9052, Belgium.
*These authors contributed equally to this work.
†Present address: Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK.
‡Corresponding author. Email: anne.osbourn@jic.ac.uk (A.O.); ybai@genetics.ac.cn (Y.B.)

Triterpenes are specialized plant metabolites that have important functions in plant defense and signaling. They are one of the largest and most structurally diverse families of plant natural products and many have been shown to possess antifungal and antibacterial activities, suggesting potential roles in mediating interactions between the producing plants and microbes (*9*–*11*). Triterpenes are synthesized by means of the mevalonate pathway, the first committed step being carried out by enzymes known as triterpene synthases (TTSs), which collectively are able to make a diverse array of different triterpene scaffolds (*12*). There are 13 predicted TTS genes in the genome of *A. thaliana* accession Col-0, five of which (*MRN*, *THAS*, *ABDS*, *PEN3*, and *BARS*) are located in a total of four plant biosynthetic gene clusters on chromosomes 4 and 5 (Fig. 1A).

Plant biosynthetic gene clusters represent evolutionary genomic hot spots under strong selection pressure (*13*), with the potential to encode metabolites of ecological importance (*14*). Moreover, most genes within these four *Arabidopsis* biosynthetic gene clusters are root-expressed (Fig. 1B), and the arabidiol gene cluster has been implicated in defense against the root-rot pathogen *Pythium irregulare*, suggesting that the derived metabolites might play a role in root-microbe interactions (*15*–*18*). The TTSs of the *A. thaliana* biosynthetic gene clusters appear to share a common ancestor but have functionally diverged and form a monophyletic clade that is distinct from other *A. thaliana* TTSs (fig. S1) (*15*). These clustered TTSs channel primary metabolism into specialized metabolism by converting the ubiquitous precursor 2,3-oxidosqualene into different triterpene scaffolds, specifically marneral/marnerol **M1/M2** (*19*), thalianol **T1** (*20*), arabidiol **A1** (*21*), tirucalladienol **Tr1** (*22*), and baruol **B1** (*23*) by means of the corresponding bi/tri/tetracyclic carbocations, respectively (fig. S1). Four of the 13 cytochrome P450 (CYP) genes in these cluster regions [*At5g48000* (*THAH*), *At4g15330* (*CYP705A1*), *At5g62590* (*MRO*), and *At5g36110* (*CYP716A1*) belonging to the CYP708, CYP705, CYP71, and CYP716 families, respectively] have previously been shown to convert **T1**, **A1**, **M1/M2**, and **Tr1** primarily to 3β,7β-thaliandiol **T2**, 14-apo-arabidiol **A2**, 23-hydroxy-marner-al/ol **M3/M4**, and unknown products, respectively (Fig. 2) (*18*, *24*). However, numerous other uncharacterized genes with predicted functions in specialized metabolism are present in these biosynthetic gene clusters and are coexpressed with the characterized TTS and CYP genes in *A. thaliana* roots (Fig. 1B and fig. S2) (*15*, *18*). Here, we elucidate three biosynthetic pathways derived from the thalianol and arabidiol gene clusters and show that these major root-specialized metabolites play important roles in selectively modulating *Arabidopsis* root bacteria both in vivo and in vitro and contribute substantially to the establishment of an *Arabidopsis*-specific root microbial community.

Elucidation of the thalianol gene cluster–derived pathways

The thalianol biosynthetic gene cluster in *A. thaliana* contains four coexpressed genes, of which the functions of two (the TTS *THAS* and the CYP *THAH*) have been characterized (*15*, *24*). A second CYP (*THAO*, *At5g47990*) that belongs to the CYP705 family mediates a third step in the pathway, but the precise structure of the resulting product has not been determined (*15*). A fourth gene (*THAA1*, *At5g47980*) that is coexpressed with the other pathway genes is predicted to encode an acyltransferase belonging to the BAHD family, but its function is unknown (*15*, *25*). We used *Agrobacterium*-mediated transient expression in *Nicotiana benthamiana* leaves to investigate the functions of the cluster genes *THAO* and *THAA1* (*26*–*28*). We first coexpressed *THAS* with the three other cluster genes encoding the hydroxylase (THAH), CYP705A5 (THAO), and acyltransferase (THAA1) in the leaves of *N. benthamiana* in a combinatorial fashion (table S1). We found that THAO alone could modify thalianol to give two new products—3β,16-thaliandiol **T3** and 16-keto-thalianol **T4**, respectively—when coexpressed with *THAS* (Fig. 2 and fig. S3). Coexpression of *THAO* and *THAH* with *THAS* gave 16-keto-3β,7β-thaliandiol **T5** and 16-keto-3β,7β,15-thaliantriol **T6** as the main products, instead of the desaturated thaliandiol proposed previously (Fig. 2 and fig. S3) (*15*). Furthermore, we found that *THAA1* encodes an acyltransferase that was functional only when coexpressed with all three other cluster genes—*THAS*, *THAH*, and *THAO*—yielding a new product identified as 3β,7β-dihydroxy-16-keto-thalian-15-yl acetate **T7**, along with its C_{17}=C_{18} cis isomer ***cis*-T7** as the minor product (Fig. 2 and fig. S3).

Corrected Text Not Updated

MS no: RAaau6389/CF/BOTANY

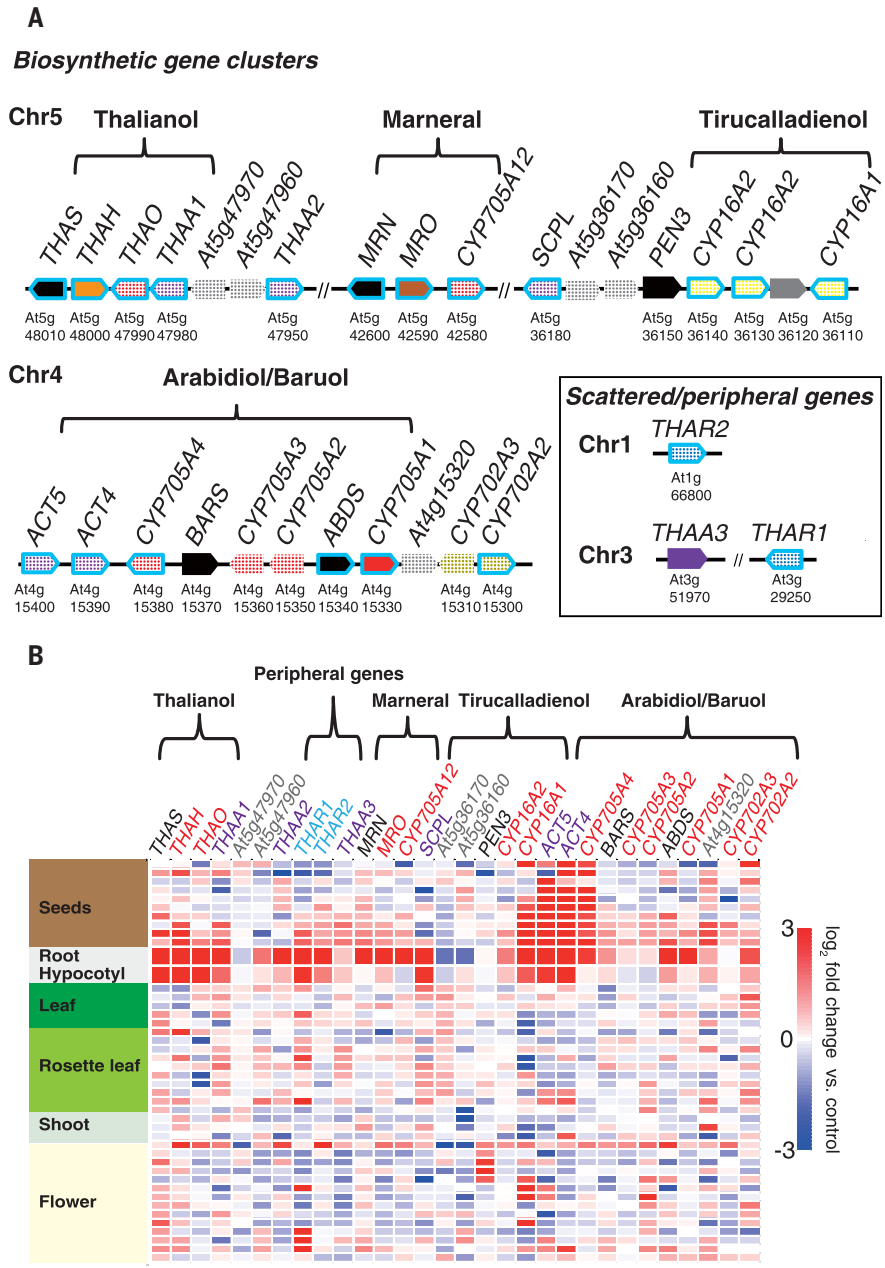


Fig. 1. Identification of a root-expressed triterpene biosynthetic network in *A. thaliana*. (**A**) The four *A. thaliana* triterpene biosynthetic gene clusters and peripheral genes. Genes encoding previously characterized enzymes are shown in solid color, and uncharacterized ones are dotted. Black, TTSs; red, CYP705 family; orange, CYP708 family; brown, CYP71 family; yellow, CYP716 family; dark yellow, CYP702 family; purple, acyltransferases (ACTs); blue, alcohol dehydrogenase (ALDHs); gray, non-biosynthetic genes. The scattered/peripheral genes identified as part of this network are also shown. Coexpressed genes have light blue borders. PEN3 was previously shown to be coexpressed with other cluster genes in roots (*18*). (**B**) Expression profiles of the triterpene cluster genes and additional scattered (peripheral) biosynthetic genes identified as part of this study in different *A. thaliana* tissues. The heatmap was generated by using microarray expression data from the eFP browser (*45*). For clarity, expression profiles for the same tissue across different developmental stages are labeled with the tissue names only. A more detailed heatmap can be found in fig. S1.

This suggests that THAA1 functions after THAS, THAH, and THAO to add an acetate group on the C15 position of the thalianol scaffold after a hydroxy moiety is installed at this position. We also identified another BAHD acyltransferase gene (*THAA2*, *At5g47950*) in close proximity to *THAA1* on chromosome five that was coexpressed with the thalianol cluster genes (Fig. 1). In contrast to THAA1, THAA2 is promiscuous and can act on the C3 hydroxy moiety of different thalianol-derived compounds (**T1** to **T7**) to introduce an acetyl group (**T11** to **T17**) when coexpressed with subsets of the thalianol cluster genes (Fig. 2, fig. S4, and table S1). Our results indicate that the four thalianol cluster genes and the nearby coexpressed gene *THAA2* are functional and yield 7β-hydroxy-16-keto-thalian-3β-15-yl diacetate **T17** when coexpressed in *N. benthamiana*.

Identification of thalianol-derived root metabolites in *Arabidopsis*

We next performed targeted metabolomic analysis to identify thalianol-derived metabolites in *A. thaliana*. Metabolic profiling identified seven major thalianol-derived products (**T1**, **T2**, **T9**, **T10**, and **T18a** to **T18c**) in root extracts from the wild-type Col-0 accession and the *THAS* overexpression line (*thas-oe*) that are absent in those from the *THAS* mutants (*thas-ko1* and *thas-ko2*) (Fig. 3, A to C, and table S2). Mutation of *THAH* (*thah-ko*) led to accumulation of 3β,15-thaliandiol **T3** and 16-keto-thalianol **T4**, whereas mutation of *THAO* (*thao-ko*) resulted in elevated levels of 3β,7β-thaliandiol **T2** and absence of **T3** and **T4** (Fig. 3, A and B, and fig. S5), indicating that THAO functions as a C16 oxidase of thalianol in *A. thaliana*. We were unable to detect **T5** to **T7** or **T17** in either Col-0 wild-type or *thas-oe* root extracts, but we did detect two compounds **T9** and **T10**, which are potential isomers of **T7** and **T17** (fig. S5). Compound **T10** was absent in root extracts from mutants of all thalianol cluster genes and *THAA2* (*thaa2-ko* and *thaa2-crispr*), whereas **T9** was still detected in those of *thaa2-ko* and *thaa2-crispr*, suggesting that **T9** and **T10** are downstream products of the thalianol gene cluster and that THAA2 may act after **T9** to form **T10** (Fig. 3, A to C, and fig. S5). Targeted metabolomics analysis also revealed the presence of thalianol-derived medium-chain saturated triterpene fatty acid esters (TFAEs) in the root extracts of Col-0 and *thas-oe*, including thalianyl palmitate (16:0, **T18a**), myristate (14:0, **T18b**), and laurate (12:0, **T18c**) (Fig. 3, A and B). Compounds **T18a** and **T18b** were also detected in the roots of other mutant lines *thah-ko*, *thao-ko*, *thaa1-crispr*, *thaa2-ko*, and *thaa2-crispr* except *thas-ko1* and *thas-ko2*, suggesting the presence of a branched thalianol-derived biosynthetic pathway. The identities of these TFAEs were confirmed with chemical synthesis.

Identification of missing genes for the biosynthesis of thalianin, TFAEs, and arabidin

To identify the missing genes required for the biosynthesis of metabolites **T9** and **T10**, we carried out a genome-wide search for coexpressed candidate biosynthetic genes in the ATTED-II plant coexpression database (http://atted.jp) using the four thalianol cluster genes as baits (*29*) because no biosynthetic genes in or near the gene cluster region appeared to be coexpressed. We selected eight candidates from the top 20 coexpressed genes for functional analysis in *N. benthamiana* (table S2). We

MS no: RAaau6389/CF/BOTANY

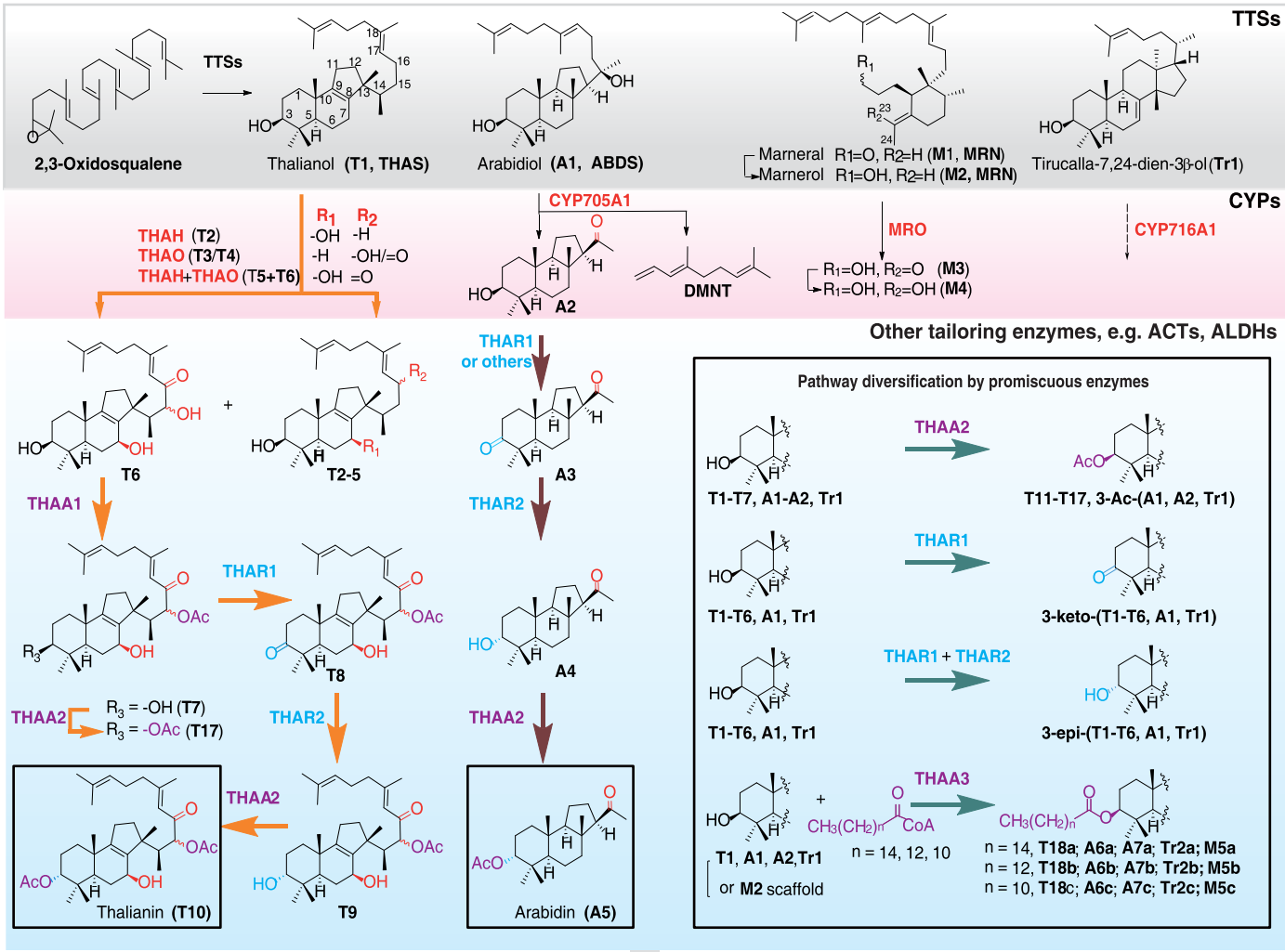


Fig. 2. The triterpene biosynthetic network in *A. thaliana* roots. Transformations catalyzed by TTSs are shaded gray, clustered CYPs are shaded red, and other tailoring enzymes are shaded blue. Chemical diversification of pathway scaffolds and intermediates by promiscuous acyltransferases and alcohol dehydrogenases is highlighted in the black box, bottom right. DMNT, (*E*)-4,8-dimethyl-1,3,7-nonatriene. Black arrows, known transformations; orange and burgundy red arrows, newly characterized steps in the thalianin and arabidin pathways, respectively; cyan arrows, pathway diversification by promiscuous enzymes; dotted black arrow, unknown transformation. The enzymes for each step are color-coded to match the classes of enzyme shown in Fig. 1A, with the exception that all CYPs are in red.

identified two genes, *At3g29250* (*THAR1*) and *At1g66800* (*THAR2*), that encode a pair of promiscuous oxido-reductases capable of epimerizing the C3 hydroxy moiety of **T1** to **T7** (Figs. 1, A and B, and 2 and fig. S6). In this epimerization sequence, THAR1 converted the C3β hydroxy of **T1** to **T7** into the C3 ketones [**3-keto-(T1-T6)** and **T8**], whereas THAR2 reduced the C3-ketones into 3α alcohols [**3-epi-(T1-T6)** and **T9**] sequentially (Fig. 2, fig. S6, and table S9). Coexpression of *THAR1* and *THAR2* with the thalianol cluster genes and *THAA2* completed the biosynthesis of **T10** in *N. benthamiana* (fig. S7). The structure of **T10** was established as 7β-hydroxy-16-keto-thalian-3α-15-yl diacetate by means of nuclear magnetic resonance, and **T10** is named thalianin hereafter (table S10).

To identify genes responsible for the biosynthesis of TFAEs **T18a** to **T18c**, we screened seven *O*-acyltransferase genes (table S4) from *A. thaliana* identified based on annotation (*30*) and found that only one of these [*THAA3*, *At3g51970*; previously reported to function in sterol fatty acid ester biosynthesis (*31*)] could catalyze the formation of thalianyl palmitate **T18a** when coexpressed with *THAS* in *N. benthamiana* (fig. S8, A and B). We also generated a *THAA3* overexpression line (*thaa3-oe*) and found that THAA3 could catalyze the formation of **T18a** in planta because elevated levels of **T18a**, but not **T18b** or **T18c**, were detected in this line. *THAA3* is presumably partially redundant because **T18a** was still detected in the *THAA3* mutant (*thaa3-ko*) (fig. S8, C and D). We further showed that THAA3 could catalyze the formation of **T18b** when coexpressed with *THAS* and a characterized chain-length specific 14:0-acyl carrier protein (ACP) thioesterase gene from *Cuphea palustris*, *CpFatB2* in *N. benthamiana* (fig. S8A) (*32*), suggesting that there are as yet unidentified genes encoding chain-length–specific ACP thioesterases involved in TFAE (**T18a** to **T18c**) biosynthesis. Besides thalianol, THAA3 could also act on other triterpenes—including arabidiol **A1** and its derivative **A2**, the PEN3 product **Tr1**, and marnerol **M2**—to introduce a palmityl or myristyl group depending on the type of fatty acyl CoA available (fig. S9). However, these products were not detected in *A. thaliana* Col-0 or *thaa3-oe* roots, possibly because of their very low abundance.

The promiscuity of enzymes encoded by the nonclustered [referred to as peripheral (*14*)] genes—including *THAR1*, *THAR2*, and *THAA2* and their strong coexpression with other divergent triterpene gene clusters in roots (Figs. 1B and 2)—prompted us to perform combinatorial biosynthesis experiments in *N. benthamiana* using these genes and the arabidiol, marneral, and tirucalladienol cluster TTS and CYP genes.

MS no: RAaau6389/CF/BOTANY

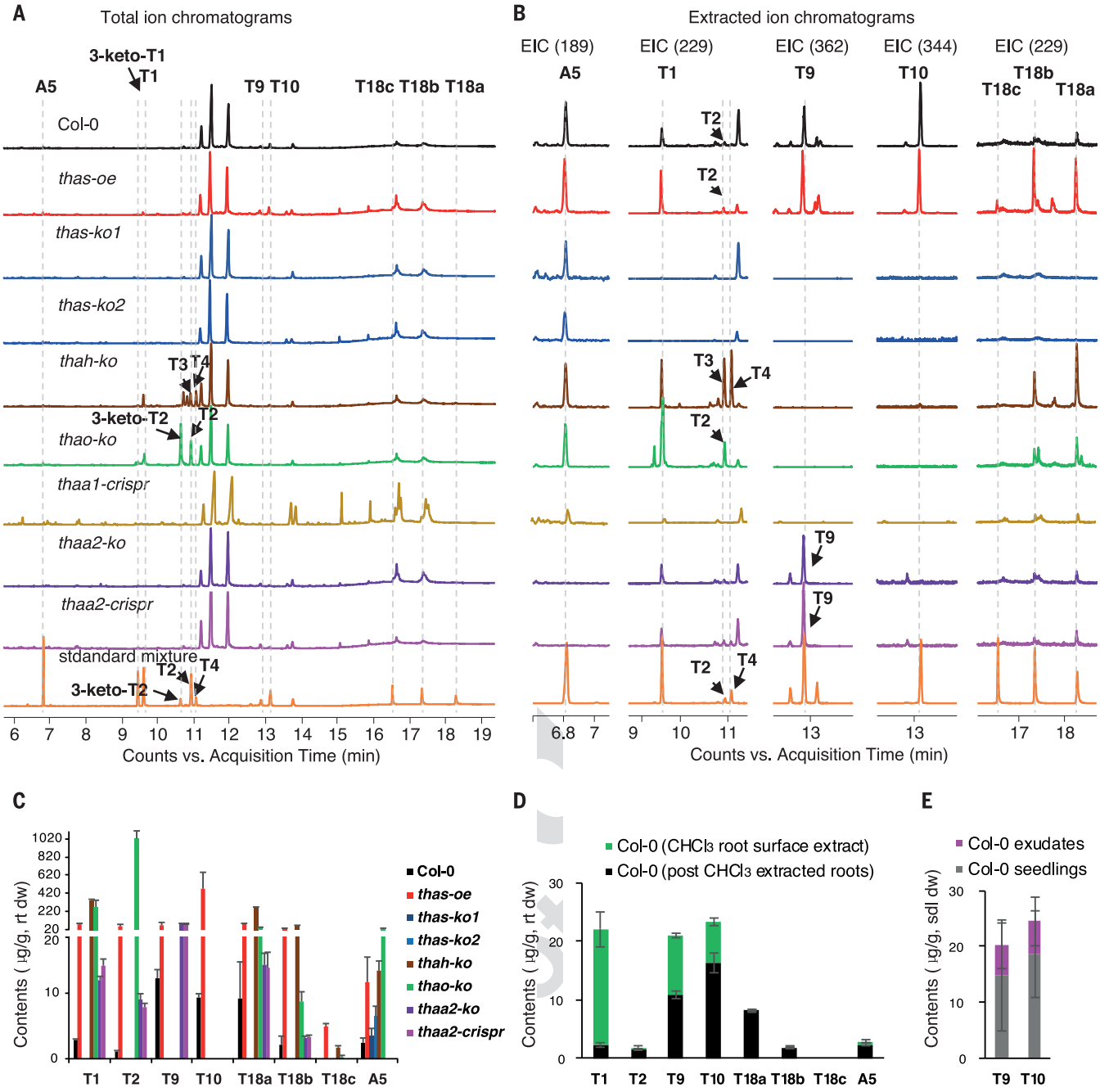


Fig. 3. Metabolite analysis of roots of different *Arabidopsis* lines. (**A**) Comparative GC-MS total ion chromatograms (TICs) of root extracts of the *A. thaliana* wild-type Col-0 (black), *thas-oe* (red), *thas-ko1* (dark blue), *thas-ko2* (blue), *thah-ko* (brown), *thao-ko* (green), *thaa1-crispr* (dark yellow), *thaa2-ko* (purple), *thaa2-crispr* (pink), and an authentic standard mix containing **T1**, **3-keto-T1**, **T2**, **3-keto-T2**, **T4**, **T9**, **T10**, **T18a** to **T18c**, and **A5** (orange). (**B**) Comparative extracted ion chromatograms (EICs) from Fig. 2A for characteristic fragments of the mass spectra of **A5** (189), **T1** to **T4** (229), **T9** (362), **T10** (344), and **T18a** to **T18c** (229). (**C**) Triterpene pathway metabolites detected in whole-root extracts from wild-type and mutant lines. Col-0, wild type. *thas-oe*, *thas-ko1*, *thas-ko2*, *thah-ko*, *thao-ko*, *thaa2-ko*, and *thaa2-crispr* are mutants for *THAS*, *THAH*, *THAO*, and *THAA2* (supplementary materials, materials and methods). (**D**) Triterpenoids detected in chloroform ($CHCl_3$) extracts from the surfaces of fresh roots of *A. thaliana* wild-type Col-0 seedlings and subsequent ethyl acetate extracts of dry and powdered chloroform-extracted roots. All lines used in qualitative and quantitative GC-MS analysis (Fig. 2, A to D) were grown on ¼ mannitol salt (MS) agar plates at 22°C (8 hours dark/16 hours light cycle) for 10 days. (**E**) Determination of compounds **T9** and **T10** by LC-MS in root exudates of Col-0 seedlings grown hydroponically in liquid ¼ MS after 10 days. Error bars represent standard deviation of three biological replicates.

Our results showed that THAR1, THAR2, and THAA2 could also act on arabidiol **A1**, 14-apo-arabidiol **A2**, and tirucalladienol **Tr1** (figs. S10 and S11), but not marnerol **M2**, to give the corresponding C3 ketones, 3α-alcohols, and acetates, respectively. Arabidiol **A1** was fully converted to **A5** when *THAR1*, *THAR2*, and *THAA2* were coexpressed together with *ABDS* and *CYP705A1* (Fig. 2 and fig. S12). **A5** is an *A. thaliana* root metabolite identified through comparative metabolomics analysis of the *A. thaliana* wild type (Col-0) with *ABDS* and *CYP705A1* mutants (*33*). Our results show that *THAR1*, *THAR2*, and *THAA2* encode enzymes that can convert **A2** to **A3**, **A3** to **A4**, and **A4** to **A5**, respectively, reconstituting

MS no: RAaau6389/CF/BOTANY

the complete biosynthesis of **A5** in *N. benthamiana* (Fig. 2, fig. S12, and table S17).

Metabolite analyses of root extracts of *A. thaliana* T-DNA insertion mutants of *THAR1* (*thar1-ko*), *THAR2* (*thar2-ko*), and *THAA2* (*thaa2-ko*) confirm that these three genes are indeed required for the synthesis of thalianin **T10** in *A. thaliana*. **T10** was absent in root extracts from all three mutant genotypes, and the pathway intermediates **T7** and **T17**, **T8**, and **T9**, respectively, accumulated instead (fig. S13). Additionally, *THAR2* and *THAA2* are also responsible for the biosynthesis of 14-apo-arabi-3α-yl acetate **A5** (referred hereafter as arabidin) because mutants *thar2-ko* and *thaa2-ko* lacked **A5** and instead accumulated **A3** and **A4**, respectively (fig. S14). **A5** was still detected in the root extracts from the *thar1-ko* line, indicating that *THAR1* is partially redundant in 14-apo-arabi-3α-yl acetate **A5** biosynthesis. Thus, these enzymes are involved in multiple pathways.

Triterpene biosynthesis affects *Arabidopsis* root microbiota assembly

Having discovered and reconstituted the complete biosynthetic pathways of thalianin **T10**, arabidin **A5**, and TFAEs **T18a** and **T18b** in *N. benthamiana*, we sought to investigate the potential biological role of this metabolic network. The thalianin **T10** and arabidin **A5** pathway genes are coexpressed in *A. thaliana* root epidermis and pericycle/stele, respectively (fig. S15), and thalianol pathway compounds **T1**, **T2**, **T9**, **T10**, and arabidin **A5** could be detected in root surface wax extracts and **T9** and **T10** also in exudates of hydroponically grown seedlings (Fig. 3, D and E, and fig. S16). Moreover, both the thalianin and arabidin pathway genes are up-regulated upon methyl jasmonate treatment

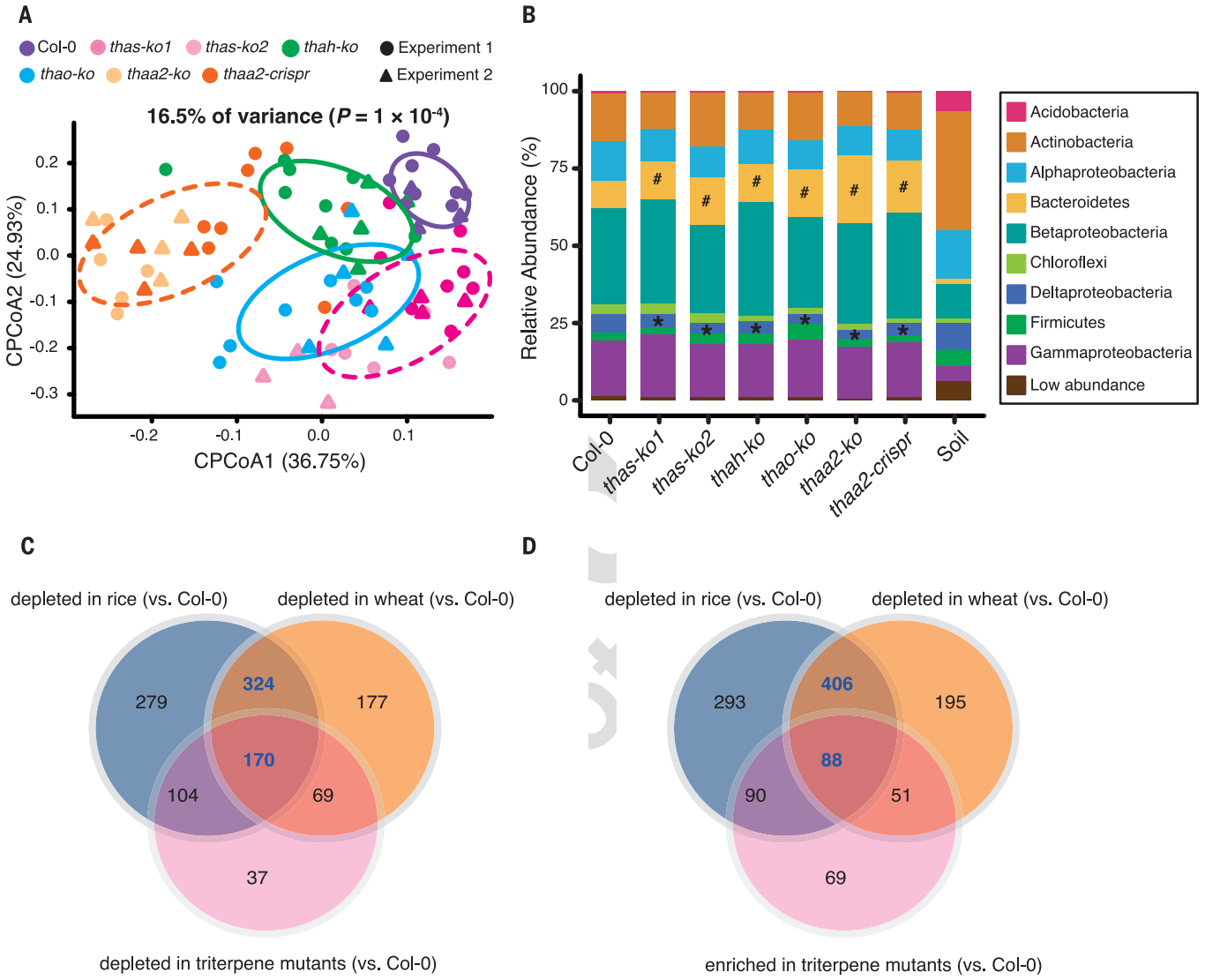


Fig. 4. Modulation of specific root bacterial taxa in triterpene pathway mutants. (**A**) Constrained principal ordination analysis (CPCoA) of Bray-Curtis dissimilarity showing plant genotype effects. Total number of individual plants used for analyses: Col-0 (n = 12), *thas-ko1* (n = 12), *thas-ko2* (n = 9), *thah-ko* (n = 14), *thao-ko* (n = 13), *thaa2-ko* (n = 9), and *thaa2-crispr* (n = 12). Biological replicates (individual plants) from two independent experiments (experiment 1 and 2) are indicated with dots and triangles, respectively. Ellipses include 68% of samples from each genotype. (**B**) Phylum distribution of the root microbiota compositions of the tested *A. thaliana* genotypes. Because the relative abundance of Proteobacteria is greater than 50%, bacteria in this phylum are shown at the class level. Pound sign (#) indicates Bacteroidetes significantly higher than that in Col-0 roots at $P < 0.05$; asterisk indicates Alphaproteobacteria and Deltaproteobacteria significantly lower than that in Col-0 at $P < 0.05$. (**C** and **D**) Venn diagrams showing significant overlap of OTUs (C) depleted or (D) enriched in the root microbiota of *A. thaliana* triterpene mutant lines as compared with the wild type (Col-0) (pink circles), compared with those depleted in the root microbiota of rice (blue circles) and wheat (orange circles) versus the *A. thaliana* Col-0 wild type. The OTU numbers specifically enriched in the root microbiota of *A. thaliana* Col-0 compared with rice and wheat are highlighted in blue and bold in the Venn diagram overlaps.

MS no: RAaau6389/CF/BOTANY

(fig. S17), suggesting that they may have a role in interactions with root microbes (*34*).

We selected mutants disrupted in the thalianin (**T10**), TFAE (**T18a** to **T18c**), and arabidin (**A5**) pathways (*thas-ko1*, *thas-ko2*, *thah-ko*, *thao-ko*, *thaa2-ko*, and *thaa2-crispr*) along with the wild type (Col-0) for root microbiota analyses (table S2). Comprehensive unbiased untargeted metabolomics analysis of whole-root extracts, root-surface extracts, and root exudates from these mutant lines versus Col-0 suggests that the thalianin pathway compounds are the most affected of all root metabolites analyzed, showing the largest fold changes across the respective metabolomes of the mutants (figs. S18 to S22). We grew these lines in natural soil from Changping Farm in Beijing for 6 weeks under controlled experimental conditions. The roots were harvested and washed with phosphate-buffered saline buffer to remove soil particles and loosely attached microbes before carrying out deep 16*S* ribosomal RNA (rRNA) gene sequencing of the root microbiota by using previously established methodology (*3*, *4*, *35*). There were no obvious differences in root phenotypes between the wild-type (Col-0) and mutant lines when the plants were harvested. However, we found that mutants affected in these pathways assembled different root microbiota when compared with the wild type (Col-0). Constrained principal ordination analysis (CPCoA) revealed differences in root microbiota between the wild-type (Col-0) and mutant lines (16.5% of total variance was explained by the plant genotypes, $P < 0.001$, permutational multivariate analysis of variance) (Fig. 4A and tables S18 to S22). Pairs of independent mutant lines for *THAS* (*thas-ko1* and *thas-ko2*) and *THAA2* (*thaa2-ko* and *thaa2-cripsr*) each display similar metabolite defects (figs. S18 to S22) and have similar microbiota profiles and microbial diversity (Fig. 4A, fig. S23, and tables S23 to S26). Furthermore, all the pathway mutants tested showed similar root microbiota modulation patterns for Bacteroidetes (enrichment) and Deltaproteobacteria (depletion) compared with the *A. thaliana* Col-0 wild type at the phylum and operational taxonomic unit (OTU) levels (Fig. 4B, figs. S24 and S25, and tables S27 to S44), which is consistent with these genes operating in the same pathway for thalianin biosynthesis. Of the OTUs that were coenriched/codepleted in the mutants, 93% (28 out of 30) showed higher abundance in Col-0 roots than bulk soil, suggesting active selection of root bacteria by the *A. thaliana* plants (fig. S26 and tables S45 and S46). These data indicate that the triterpene biosynthetic network that we have unveiled contributes to root microbiota assembly and establishment. F4

To understand whether and how this specialized metabolic network might modulate *A. thaliana*–specific root bacteria, we also compared the root bacterial profiles of the *A. thaliana* Col-0 and mutant lines with those from the taxonomically

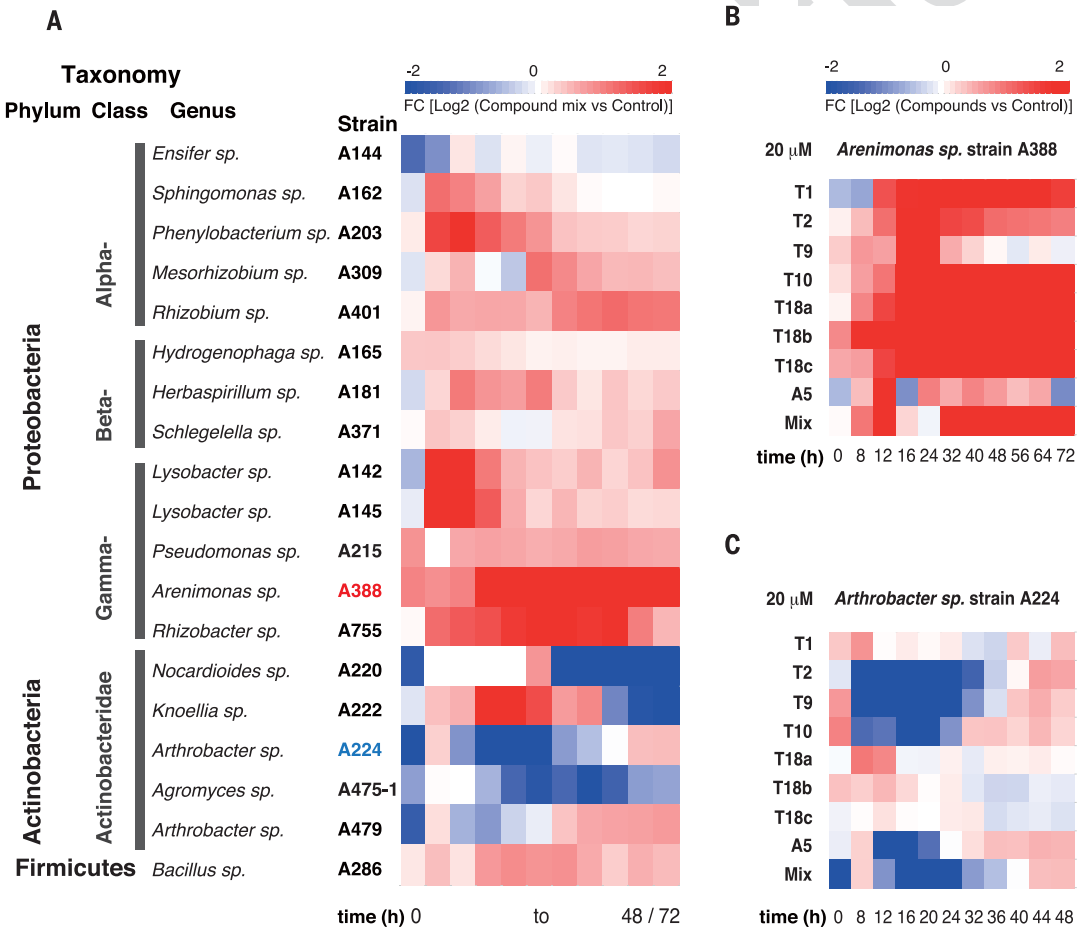


Fig. 5. Effects of pathway metabolites on the growth of isolated *A. thaliana* root–associated bacteria and bacterium-mediated chemical transformations. (**A**) Growth modulation activity of compound mixture **Mix** (10 μM **T1**, 5 μM **T2**, 20 μM **T9**, 20 μM **T10**, 10 μM **T18a**, 5 μM **T18b**, 1 μM **T18c**, and 10 μM **A5**) against the 19 strains of *A. thaliana* root bacteria from different taxa. The heatmap shows $\log_2$-fold change of individual strain treated with **Mix** versus control (ethanol only) over 48 or 72 hours. The corresponding graphical growth curves are depicted in figs. S30 to S34. (**B**) Heatmap showing the $\log_2$-fold change in cell density (OD_{600}) of *Arenimonas* sp. strain A388 treated with eight different purified compounds (**T1**, **T2**, **T9**, **T10**, **T18a**, **T18b**, **T18c**, and **A5**, respectively) at 20 μM and **Mix** over 72 hours. All compounds were dissolved in ethanol. Control, 0 mM, ethanol. (**C**) Heatmap showing the $\log_2$-fold change in cell density (OD_{600}) of *Arthrobacter* sp. strain A224 treated with eight different purified compounds (**T1**, **T2**, **T9**, **T10**, **T18a**, **T18b**, **T18c**, and **A5**, respectively) at 20 μM and **Mix** over 48 hours. All compounds were dissolved in ethanol. Control, 0 mM, ethanol.

MS no: RAaau6389/CF/BOTANY

distant species rice (36) and wheat, also previously grown in the soil from Changping Farm on different occasions. Although the *A. thaliana*, rice, and wheat samples differ from each other in many aspects—including growth conditions, germination periods, and climate conditions (supplementary materials, materials and methods)—the starting inocula (soils) are very similar, sharing substantial overlap in total OTUs (67%, 2377 out of 3531) (fig. S27 and tables S47 to S49). We found that of the 494 OTUs specifically enriched in the root microbiota of *A. thaliana* Col-0 compared with rice and wheat (represented in the overlap between the blue and orange circles in Fig. 4, C and D), 34% (170 out of 494) (Fig. 4C and tables S50 to S55) were depleted, and 18% (88 out of 494) (Fig. 4D and tables S50 to S55) enriched in the root microbiota of the triterpene mutant lines compared with that of *A. thaliana* Col-0. Unlike *A. thaliana*, rice and wheat do not make thalianin, TFAEs, or arabidin (fig. S28). Our results suggest that the specialized triterpene biosynthetic network that we have uncovered may contribute to enrichment of around a third of the *A. thaliana*–specific root bacteria present in the roots of the wild-type line (Fig. 4C) while deterring another 18% of this bacterial population (Fig. 4D). The relatively larger number of OTUs depleted in the triterpene mutants (versus Col-0) (a total of 380) (Fig. 4C, pink circle) in comparison with the total number of OTUs that are enriched in the mutant lines versus Col-0 (298) (Fig. 4D, pink circle) further suggests that this triterpene biosynthetic network may play a more important role in enriching *Arabidopsis*–specific root bacteria rather than repelling other bacteria.

Purified triterpenes selectively modulate root bacteria

To test whether triterpene pathway metabolites directly regulate root microbiota members, we isolated and identified bacteria from the roots of *A. thaliana* Col-0 plants grown in soil from the aforementioned Changping Farm by limiting dilution and barcoded sequencing (35). We selected a total of 19 bacterial strains that belong to 17 genera [within three major bacterial phyla (Proteobacteria, Actinobacteria, and Firmicutes) that shared >97% 16*S* rRNA gene similarity to the OTUs] and that showed differential abundances in the microbiota of *A. thaliana* wild type (Col-0), mutants, and soil (fig. S29 and table S56). These bacterial strains were grown in liquid culture with a formulated cocktail of purified compounds that reflected the content and composition of the pathway metabolites in the roots of *A. thaliana* Col-0 (figs. S30 to S34). We found that most of the Proteobacteria strains tested proliferated faster in the presence of the *A. thaliana* triterpene mixture, whereas all five Actinobacteria strains were inhibited (Fig. 5A and figs. S30 to S34). The OTUs corresponding to these bacterial isolates (16 out of 19) showed consistent enrichment or depletion patterns in plant roots versus soil that corresponded with the growth promotion and inhibitory effects of the triterpene mixture, suggesting that the compounds tested contribute to the active selection of plant root bacteria (figs. S30 to S34). Moreover, the corresponding enrichment or depletion patterns for 10 bacterial genera (59% of the 17 tested) to which the metabolite-sensitive bacterial isolates belong were detected in the microbiota of at least one pathway mutant included in our microbiota analysis; the genera of bacteria for which growth is either promoted or inhibited by the triterpene cocktail mixture are depleted or enriched, respectively, in the pathway mutants compared with the wild type (Col-0) (fig. S35 and tables S57 to S63). Further tests of the sensitive strains with purified individual compounds revealed that pathway metabolites can selectively modulate the growth of bacteria and that small structural differences between compounds can affect activities (fig. S36). For example, we found that Actinobacteria *Arthrobacter* sp. strain A224 was inhibited by compounds **T2**, **T9**, **T10**, and **A5** at 20 μM but not by the other triterpenes tested, whereas all compounds tested showed growth promoting effects on Gammaproteobacteira *Arenimonas* sp. strain A388 (Fig. 5, B and C). We also found that strain A475-1 (*Agromyces sp.*) has alcohol dehydrogenase activity and could selectively convert **T2** into **3-keto-T2** but not **T9**/**T10** (Fig. 6, A and B, and fig. S37A), whereas strain A215 (*Pseudomonas sp.*) has lipase activity and is able to cleave the TFAEs **T18a** to **T18c** to give **T1** and the corresponding fatty acids but not the acetate **T11** (Fig. 6, C and D, and fig. S37, B to D). Moreover, strain A215 could use the cleavage product palmitic acid as a carbon source for proliferation (Fig. 6E). Such diverse and substantial

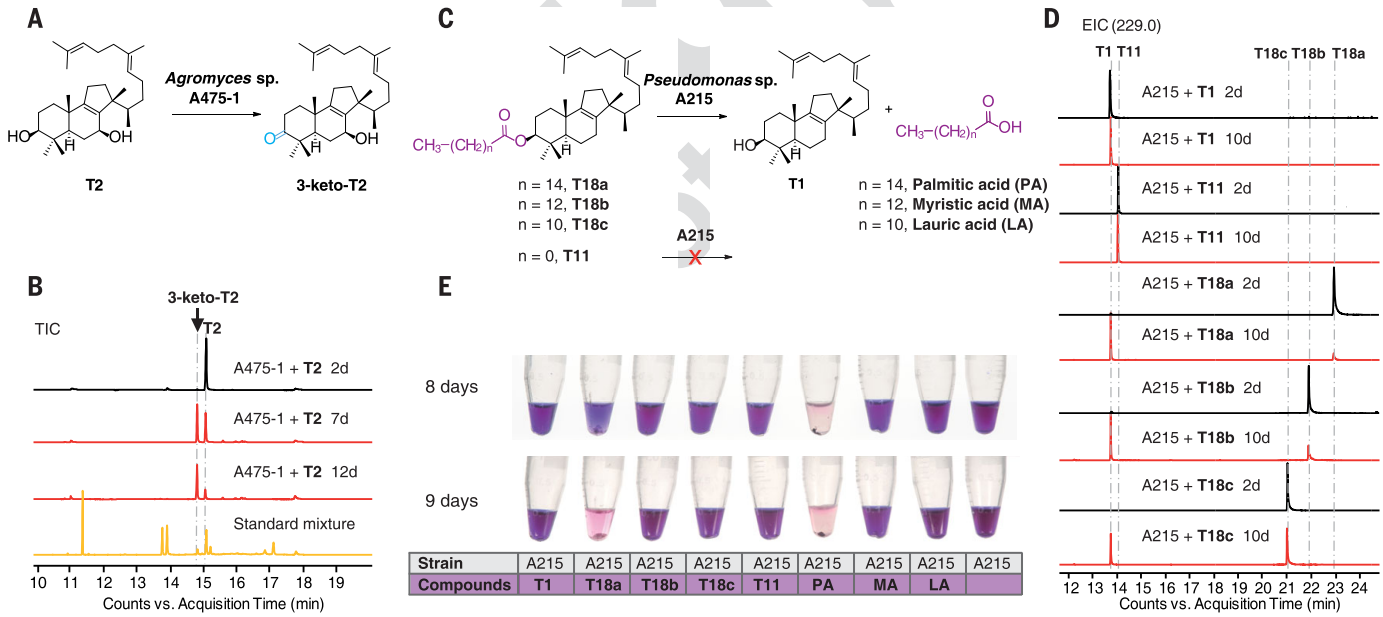


Fig. 6. Bacterium-mediated transformations of triterpenes. (**A**) Conversion of thaliandiol **T2** to **3-keto-T2** by *Agromyces* sp. strain A475-1 as shown by means of (**B**) comparative GC-MS TICs of EtOAc extracts of bacterial cultures of supplemented with **T2** on the 2nd, 7th and, 10th day, respectively. (**C**) Selective conversion of TFAEs **T18a** to **T1c** but not **T11** to **T1** by *Pseudomonas* sp. strain A215 as shown by means of (**D**) comparative GC-MS EICs [mass/charge ratio (*m/z*) = 229] of EtOAc extracts of bacterial strain A215 cultures supplemented with **T1**, **T11**, and **T18a** to **T19c** on the 2nd and 10th day, respectively. (**E**) *Pseudomonas* sp. strain A215 could use palmitic acid as carbon source, as shown by the proliferation of A215 in minimal salt media supplemented with 0.5 mM (in ethanol) of **T18a** or palmitic acid (**PA**) but not in those with **T1**, **T18b**, **T18c**, **T11**, myristic acid (**MA**), lauric acid (**LA**), and control. The minimal salt media contain 0.3 mg/ml resazurin as indicator of bacterial growth (from purple blue to pink).

MS no: RAaau6389/CF/BOTANY

interaction patterns of root metabolites with taxonomically distinct root microbiota members suggest that the metabolites from this biosynthetic network alter the assembly of *A. thaliana* root microbiota. Taking the results of the root microbiota sequencing and in vitro bioassays together, we conclude that this triterpene biosynthetic network tunes the ecological niche for the assembly and maintenance of *A. thaliana* root microbiota.

Conclusion

The triterpene biosynthetic network described here has a latent capacity for the synthesis of more than 50 root metabolites (Fig. 2). This is a relatively large number considering the total of ~300 nonvolatile root metabolites that we detected in the polar methanolic (~220) and nonpolar ethyl acetate extracts (~85) of *A. thaliana* Col-0 roots under our experimental conditions (tables S64 and 65). This network originates from evolutionary divergent biosynthetic gene clusters coupled with cross-talk that involves enzymes encoded by peripheral genes that service multiple biosynthetic pathways. These root metabolites can selectively regulate the growth of *Arabidopsis* root bacteria from different taxa by acting as antibiotics or proliferating agents. Biosynthesis of these specialized root triterpenes dynamically modulate a large portion [52%, 258 (170+88) out of 498 OTUs] (Fig. 4, C and D) of the *A. thaliana*–specific bacteria, shaping the *A. thaliana*–specific root microbiota. Triterpenes, a large and diverse groups of plant natural products (> 20,000 reported so far) (*12*), likely also sculpt the microbiota of other plant species, tailoring them to generate plant species–specific microbial communities, and may be useful for engineering root microbiota for sustainable agriculture (*37*–*39*). It is tempting to speculate that metabolic diversification in the plant kingdom may provide a basis for communication and recognition that enables the shaping of microbial communities tailored to the needs of the host, and that this may in part explain the existence of plant-specialized metabolism.

Materials and methods

Detailed materials and methods can be found in the supplementary materials.

Plant materials

Arabidopsis thaliana accession Columbia (Col-0) was used as wild type in this study. All *A. thaliana* T-DNA insertion mutants were obtained from Nottingham *Arabidopsis* Stock Centre (NASC) unless otherwise stated. Homozygous mutants were identified by PCR-based genotyping using the primers listed in the table S5A. A 35S promoter-driven overexpression line of thalianol synthase (*thas-oe*) was generated previously (*15*). Overexpression lines for *THAA3* were generated by transforming *A. thaliana* Col-0 plants with *Agrobacterium tumefaciens* strain LBA4404 harboring the expression vector pMDC32 which contains the coding sequence of *THAA3* (*At3g51970*) under the control of the 35S promoter via floral dipping (*40*). Homozygous CRISPR knockout lines of *THAA1* (*At5g47980*) and *THAA2* (*At5g47950*) were generated by transforming *A. thaliana* Col-0 wild type with *A. tumefaciens* C58C1 harboring CRISPR/Cas9 constructs with specific single guide RNAs as previously described (*41*).

Cloning and transient expression

The coding sequences for *THAS* (*At5g48010*), *THAH* (*At5g48000*), *THAO* (*At5g47990*), *THAA1* (*At5g47980*) and *THAA2* (*At5g47950*), *THAR1* (*At3g29250*), *THAR2* (*At1g66800*), *THAA3* (*At3g51970*), *ABDS* (*At4g15340*), *CYP705A1* (*At4g15330*), *At1g14960*, *At5g23840*, *At2g16005*, *At5g38030*, *At1g50560*, *At5g38020*, *MRO* (*At5g42580*), *CYP705A12* (*At5g42590*), *MRN* (*At5g42600*), *At5g12420*, *At5g55380*, *At3g49210*, *At1g54570*, *At3g49190*, *At3g26840* were amplified from a root cDNA library of the *A. thaliana* Col-0 accession (*26*). The coding sequences for *PEN3* (*At5g36150*), *CYP716A1* (*At5g36110*) and the truncated HMG CoA reductase (tHMGR) gene from oat had been cloned previously and were included in this study (*18*, *27*). The coding sequence of *CpFatB2* from *Cuphea palustris* (accession no. U38189) was retrieved from NCBI and synthesized by Integrated DNA Technologies. These sequences were cloned into the pEAQ-HT expression vector for transient expression in *N. benthamiana* leaves as detailed in the supplementary materials.

Metabolite extraction and analysis

N. benthamiana leaves from transient expression experiments were harvested 5 days post infiltration, lyophilized, extracted with EtOAc and analyzed by means of gas chromatography–mass spectrometry (GC-MS). Seedlings of *A. thaliana* Col-0 and mutants, rice and wheat were grown on Murashige-Skoog (¼ MS) agar plates or liquid media at 22°C under short day conditions (8 hours light/16 hours dark) for 10 days. Roots were harvested from 10-day-old seedlings grown on agar plates, lyophilized, extracted with EtOAc and analyzed by GC-MS or extracted with MeOH and analyzed by means of liquid chromatography–MS (LC-MS). Spent media and whole seedlings of 10-day-old *A. thaliana* grown in ¼ MS liquid media were harvested separately, lyophilized, extracted with MeOH and analyzed by LC-MS. Targeted and untargeted metabolomics analysis of GC-MS and LC-MS data were performed using Agilent MassHunter and XC-MS, respectively.

Production and isolation of thalianol and arabidiol derived metabolites

Compounds **T1**-**T10**, ***3*-keto-T2**, **T17**, ***cis*-T10**, ***cis*-T17**, **A1**, **A2**, and **A4** were obtained by large-scale vacuum infiltration of *N. benthamiana* leaves with *A. tumefaciens* LBA4404 strains harboring the corresponding expression constructs followed by extraction and purification. Compounds **3-keto-T1**, **T11**, **T18a-c**, **A3**, and **A5** were chemically synthesized from the corresponding precursors.

Microbiota analysis

A. thaliana plants (Col-0 and triterpene mutants) were grown under controlled short-day conditions in the natural soil from Changping Farm (40°5′49”N, 116°24′44” E, Beijing, China). Two wheat (Xiaoyan54 and Jing411) and rice (IR24 and Nipponbare) varieties were grown in Changping Farm under field conditions in 2016 and 2017, respectively. Roots from *A. thaliana*, wheat and rice were harvested six weeks post plantation in natural soil for 16S rRNA gene profiling. Data analysis was performed using QIIME 1.9.1 (*42*), USEARCH 10.0 (*43*) and in-house scripts. OTUs with differential abundance were identified with a negative binomial generalized linear model in the edgeR package with a foldchange threshold >1.2 (*20*). Venn diagrams were generated using the Venndiagram package (*22*).

In vitro bioassays

Root-associated bacteria were isolated and identified from *A. thaliana* Col-0 plants grown in the aforementioned natural soil from Changping Farm with limiting dilution and barcoded sequencing (*35*). 19 bacterial isolates from diverse taxa that share greater than 97% 16*S* rRNA gene similarity with representative OTU sequences were selected for bioassay. The bacteria were cocultured with different formulated cocktails of triterpene mixtures or individual triterpenes in 1/10 TSB media and their growth (OD_{600}) monitored over 48 to 72h. Bacteria-mediated transformation of triterpenes **T1**, **T11**, and **T18a** to **T18c** was tested against all 19 bacteria, whereas **T2**, **T9**, and **T10** against sensitive bacterial strains A224, A475-1, and A479 only.

REFERENCES AND NOTES

1. L. Chae, T. Kim, R. Nilo-Poyanco, S. Y. Rhee, Genomic signatures of specialized metabolism in plants. *Science* **344**, 510–513 (2014). doi: 10.1126/science.1252076; pmid: 24786077
2. R. L. Berendsen, C. M. J. Pieterse, P. A. H. M. Bakker, The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17**, 478–486 (2012). doi: 10.1016/j.tplants.2012.04.001; pmid: 22564542
3. D. Bulgarelli *et al*., Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**, 91–95 (2012). doi: 10.1038/nature11336; pmid: 22859207
4. D. S. Lundberg *et al*., Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86–90 (2012). doi: 10.1038/nature11237; pmid: 22859206
5. C. Nguyen, Rhizodeposition of organic C by plants: Mechanisms and controls. *Agronomie* **23**, 375–396 (2003). doi: 10.1051/agro:2003011
6. K. Zhalnina *et al*., Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* **3**, 470–480 (2018). doi: 10.1038/s41564-018-0129-3; pmid: 29556109
7. I. A. Stringlis *et al*., MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5213–E5222 (2018). doi: 10.1073/pnas.1722335115; pmid: 29686086
8. S. L. Lebeis *et al*., PLANT MICROBIOME. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860–864 (2015). doi: 10.1126/science.aaa8764; pmid: 26184915
9. K. Papadopoulou, R. E. Melton, M. Leggett, M. J. Daniels, A. E. Osbourn, Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12923–12928 (1999). doi: 10.1073/pnas.96.22.12923; pmid: 10536024
10. A. G. Pacheco, A. F. C. Alcântara, G. M. Corrêa, V. G. C. Abreu, *Relationships Between Chemical Structure and Activity of*

Confidential Embargoed

MS no: RAaau6389/CF/BOTANY

Triterpenes Against Gram-Positive and Gram-Negative Bacteria (INTECH Open Access Publisher, 2012).
11. J. M. Augustin, V. Kuzina, S. B. Andersen, S. Bak, Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry* **72**, 435–457 (2011). doi: 10.1016/j.phytochem.2011.01.015; pmid: 21333312
12. R. Thimmappa, K. Geisler, T. Louveau, P. O'Maille, A. Osbourn, Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* **65**, 225–257 (2014). doi: 10.1146/annurev-arplant-050312-120229; pmid: 24498976
13. A. M. Boutanaev, A. E. Osbourn, Multigenome analysis implicates miniature inverted-repeat transposable elements (MITEs) in metabolic diversification in eudicots. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6650–E6658 (2018). doi: 10.1073/pnas.1721318115; pmid: 29941591
14. H.-W. Nützmann, A. Huang, A. Osbourn, Plant metabolic clusters—From genetics to genomics. *New Phytol.* **211**, 771–789 (2016). doi: 10.1111/nph.13981; pmid: 27112429
15. B. Field, A. E. Osbourn, Metabolic diversification—Independent assembly of operon-like gene clusters in different plants. *Science* **320**, 543–547 (2008). doi: 10.1126/science.1154990; pmid: 18356490
16. B. Field *et al*., Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16116–16121 (2011). doi: 10.1073/pnas.1109273108; pmid: 21876149
17. R. Sohrabi *et al*., In planta variation of volatile biosynthesis: An alternative biosynthetic route to the formation of the pathogen-induced volatile homoterpene DMNT via triterpene degradation in *Arabidopsis* roots. *Plant Cell* **27**, 874–890 (2015). doi: 10.1105/tpc.114.132209; pmid: 25724638
18. A. M. Boutanaev *et al*., Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E81–E88 (2015). doi: 10.1073/pnas.1419547112; pmid: 25502595
19. Q. Xiong, W. K. Wilson, S. P. T. Matsuda, An *Arabidopsis* oxidosqualene cyclase catalyzes iridal skeleton formation by Grob fragmentation. *Angew. Chem. Int. Ed.* **45**, 1285–1288 (2006). doi: 10.1002/anie.200503420; pmid: 16425307
20. G. C. Fazio, R. Xu, S. P. T. Matsuda, Genome mining to identify new plant triterpenoids. *J. Am. Chem. Soc.* **126**, 5678–5679 (2004). doi: 10.1021/ja0318784; pmid: 15125655
21. T. Xiang *et al*., A new triterpene synthase from *Arabidopsis thaliana* produces a tricyclic triterpene with two hydroxyl groups. *Org. Lett.* **8**, 2835–2838 (2006). doi: 10.1021/ol060973p; pmid: 16774269
22. P. Morlacchi *et al*., Product profile of PEN3: The last unexamined oxidosqualene cyclase in Arabidopsis thaliana. *Org. Lett.* **11**, 2627–2630 (2009). doi: 10.1021/ol9005745; pmid: 19445469
23. S. Lodeiro *et al*., An oxidosqualene cyclase makes numerous products by diverse mechanisms: A challenge to prevailing concepts of triterpene biosynthesis. *J. Am. Chem. Soc.* **129**, 11213–11222 (2007). doi: 10.1021/ja073133u; pmid: 17705488
24. D. A. Castillo, M. D. Kolesnikova, S. P. T. Matsuda, An effective strategy for exploring unknown metabolic pathways by genome mining. *J. Am. Chem. Soc.* **135**, 5885–5894 (2013). doi: 10.1021/ja401535g; pmid: 23570231
25. J. C. D'Auria, Acyltransferases in plants: A good time to be BAHD. *Curr. Opin. Plant Biol.* **9**, 331–340 (2006). doi: 10.1016/j.pbi.2006.03.016; pmid: 16616872
26. A. C. Huang *et al*., Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6005–E6014 (2017). doi: 10.1073/pnas.1705567114; pmid: 28673978
27. J. Reed *et al*., A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules. *Metab. Eng.* **42**, 185–193 (2017). doi: 10.1016/j.ymben.2017.06.012; pmid: 28687337
28. A. C. Huang, Y. J. Hong, A. D. Bond, D. J. Tantillo, A. Osbourn, Diverged plant terpene synthases reroute the carbocation cyclization path towards the formation of unprecedented 6/11/5 and 6/6/7/5 sesterterpene scaffolds. *Angew. Chem. Int. Ed.* **57**, 1291–1295 (2018). doi: 10.1002/anie.201711444; pmid: 29194888
29. T. Obayashi, Y. Aoki, S. Tadaka, Y. Kagaya, K. Kinoshita, ATTED-II in 2018: A plant coexpression database based on investigation of the statistical property of the mutual rank index. *Plant Cell Physiol.* **59**, e3–e3 (2018). doi: 10.1093/pcp/pcx191; pmid: 29216398
30. D. Ma *et al*., Crystal structure of a membrane-bound O-acyltransferase. *Nature* **562**, 286–290 (2018). doi: 10.1038/s41586-018-0568-2; pmid: 30283133
31. Q. Chen, L. Steinhauer, J. Hammerlindl, W. Keller, J. Zou, Biosynthesis of phytosterol esters: Identification of a sterol o-acyltransferase in Arabidopsis. *Plant Physiol.* **145**, 974–984 (2007). doi: 10.1104/pp.107.106278; pmid: 17885082
32. K. Dehesh, P. Edwards, T. Hayes, A. M. Cranmer, J. Fillatti, Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. *Plant Physiol.* **110**, 203–210 (1996). doi: 10.1104/pp.110.1.203; pmid: 8587983
33. R. Sohrabi, T. Ali, L. Harinantenaina Rakotondraibe, D. Tholl, Formation and exudation of non-volatile products of the arabidiol triterpenoid degradation pathway in *Arabidopsis* roots. *Plant Signal. Behav.* **12**, e1265722 (2017). doi: 10.1080/15592324.2016.1265722; pmid: 27918234
34. L. Hu *et al*., Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nat. Commun.* **9**, 2738 (2018). doi: 10.1038/s41467-018-05122-7; pmid: 30013066
35. Y. Bai *et al*., Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**, 364–369 (2015). doi: 10.1038/nature16192; pmid: 26633631
36. J. Zhang *et al*., Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Sci. China Life Sci.* **61**, 613–621 (2018). doi: 10.1007/s11427-018-9284-4; pmid: 29582350
37. B. O. Oyserman, M. H. Medema, J. M. Raaijmakers, Road MAPs to engineer host microbiomes. *Curr. Opin. Microbiol.* **43**, 46–54 (2018). doi: 10.1016/j.mib.2017.11.023; pmid: 29207308
38. Y. Dessaux, C. Grandclément, D. Faure, Engineering the rhizosphere. *Trends Plant Sci.* **21**, 266–278 (2016). doi: 10.1016/j.tplants.2016.01.002; pmid: 26818718
39. H. Toju *et al*., Core microbiomes for sustainable agroecosystems. *Nat. Plants* **4**, 247–257 (2018). doi: 10.1038/s41477-018-0139-4; pmid: 29725101
40. S. J. Clough, A. F. Bent, Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998). doi: 10.1046/j.1365-313x.1998.00343.x; pmid: 10069079
41. A. Ritter *et al*., The transcriptional repressor complex FRS7-FRS12 regulates flowering time and growth in *Arabidopsis*. *Nat. Commun.* **8**, 15235 (2017). doi: 10.1038/ncomms15235; pmid: 28492275
42. J. G. Caporaso *et al*., QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010). doi: 10.1038/nmeth.f.303; pmid: 20383131
43. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010). doi: 10.1093/bioinformatics/btq461; pmid: 20709691
44. BIG Data Center Members, Database Resources of the BIG Data Center in 2018. *Nucleic Acids Res.* **46**, D14–D20 (2018). pmid: 29036542
45. D. Winter *et al*., An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLOS ONE* **2**, e718 (2007). doi: 10.1371/journal.pone.0000718; pmid: 17684564

ACKNOWLEDGMENTS

J. Pollier and P. Fernandez-Calvo are acknowledged for their support of Y-C.B.; C. Owen is acknowledged for initial testing of the thalianol cluster genes. **Funding:** This work has been supported by the National Institutes of Health Genome to Natural Products Network award U101GM110699 (A.O. and A.C.H.); the "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDB11020700) (Y.B.); the international cooperation and exchanges NSFC grant 31761143017 (Y.B.); the Centre of Excellence for Plant and Microbial Sciences (CEPAMS), established between the John Innes Centre and the Chinese Academy of Sciences and funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC) and the Chinese Academy of Sciences (A.O. and Y.B.); the Priority Research Program of the Chinese Academy of Sciences (QYZDB-SSW-SMC021) (Y.B.); the European Community's Seventh Framework Program (FP7/2007–2013) under grant agreement 613692 (TriForC) (A.O. and A.G.); the joint Engineering and Physical Sciences Research Council/ BBSRC-funded OpenPlant Synthetic Biology Research Centre grant BB/L014130/1 (H.-W.N., A.O.); and the Research Foundation Flanders with a research project grant to A.G. (G008417N). A.C.H. is supported by a European Commission Marie Skłodowska-Curie Individual Fellowship (H2020-MSCA-IF-EF-ST-702478-TRIGEM). H.-W.N. is currently supported by a Royal Society University Research Fellowship (UF160138). A.O.'s laboratory is funded by the UK BBSRC Institute Strategic Programme Grant "Molecules from Nature" (BB/P012523/1) and the John Innes Foundation. Y.-C.B. is supported by a China Scholarship Council (CSC) Ph.D. scholarship. **Author contributions.** A.C.H., Y. B., and A.O. conceived and designed the project. A.C.H. discovered and characterized the biosynthetic network, performed bacterial growth assay, and coordinated the project; T.J. grew plants in natural soils, harvested roots, prepared the 16S amplicon library for sequencing, isolated *A. thaliana* root bacteria and performed bacterial growth assays; Y.-X.L. performed bioinformatics analysis on microbiota sequencing results; Y.-C.B. generated the homozygous *thaa1-crispr* and *thaa2-crispr* lines; J.R. cloned the thalianol and marneral cluster genes; and B.Q. grew and harvested the wheat samples for microbiota analysis. A.C.H., T.J., Y.-X.L., H.-W.N., A.G., and Y.B. analyzed data; A.C.H., T.J., Y.B., and A.O. wrote the manuscript, with contributions from other authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Raw microbiota sequencing data reported in this paper have been deposited in the Genome Sequence Archive in Beijing Institute of Genomics (BIG) Data Center (*44*), Chinese Academy of Sciences under accession no. PRJCA001296 that are public accessible at http://bigd.big.ac.cn/gsa. Scripts used in the microbiota analyses are available under the following link: https://github.com/microbiota/Huang2019SCIENCE.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/[vol]/[issue]/[page]/suppl/DC1
Materials and Methods
Figs. S1 to S64
Tables S1 to S65
References (*46*–*60*)

3 July 2018; accepted 25 March 2019
10.1126/science.aau6389

Copyedited manuscript

MS no: RAaau6389/CF/BOTANY