Plant Communications
Correspondence

CellPress
Partner Journal

A chromosome-scale genome assembly of *Rauvolfia tetraphylla* facilitates identification of the complete ajmaline biosynthetic pathway

Dear Editor,

Rauvolfia tetraphylla (aka the Devil pepper) (Supplemental Figure 1) is a well-known medicinal plant that produces monoterpenoid indole alkaloids (MIAs). This MIA biosynthesis occurs in several organs, including leaves, stems, fruit, and roots, which accumulate the famous antiarrhythmic ajmaline (Kumar et al. 2016a, 2016b; Kumara et al., 2019). MIAs are natural products notably involved in plant adaptation to the environment and defense against aggressors. This mainly results from their high biological activities, which also explain their pharmacological properties. MIAs display complex structures resulting from long and elaborate biogenesis processes, as mainly illustrated in the Madagascar periwinkle *Catharanthus roseus* (Kulagina et al., 2022). Although ajmaline remains an important drug in the general pharmaceutical market, its biosynthetic pathway is still incomplete, precluding a transfer to heterologous organisms as recently achieved for bioproduction of other valuable MIAs (Zhang et al., 2022). Overall, the biosynthesis of ajmaline requires a 10-step modification of strictosidine, catalyzed by enzymes from the cytochrome P450, alcohol dehydrogenase (ADH), and BAHD acyltransferase families, all but one of which have been identified (Dang et al., 2017) (Supplemental Figure 2). The central part of this pathway involves the conversion of vinorine into 17-*O*-acetyl-norajmaline, which relies on hydroxylation of vinorine into vomilenine by vinorine hydroxylase (VH) (Dang et al., 2017). Next, two ADHs successively ensure the reduction of the vomilenine 19,20-double bond and the reduction of its indolenine ring in the 1,2-position. To date, only the vomilenine reductase (from the medium-chain dehydrogenase/reductase family) that produces 19,20-α(S)-dihydrovomilenine has been characterized and named VR2 (vomilenine reductase 2; Geissler et al., 2016). This makes the remaining ADH the only enzyme missing from the ajmaline biosynthetic pathway.

To identify this enzyme, we first assembled a chromosome-scale version of the *R. tetraphylla* genome by generating 43.8 Gb ONT PromethION reads with an N50 of ∼21.8 kb. Reads were assembled with Flye, and the resulting contigs were corrected twice with ONT reads and polished twice with Illumina reads. Using Hi-C data (Supplemental Figure 3A), 89.7% of the unscaffolded assembly was anchored to 33 pseudo-chromosomes (Figure 1A; supplemental materials and methods) in accordance with the 66 chromosomes counted in *R. tetraphylla* cells (2n = 66; Supplemental Figure 2B), resulting in an assembly of ∼733.6 Mb. About 98.3% of the eudicot Benchmarking Universal Single Copy Orthologs (BUSCOs) were annotated, and the LTR Assembly Index (19.21) was higher than that of *C. roseus* (13.11 [Li et al., 2023]–14.62 [Sun et al., 2023]), indicating the high completeness of our assembly in both genic and non-genic regions (Supplemental Table 1). By integrating *ab initio* prediction and *de novo* transcriptome assembly, we annotated 101 883 high-confidence genes (Figure 1A; Supplemental Tables 1 and 2) with a BUSCO completeness score of 97.7%. Functional annotations were assigned to ∼65.7% of the genes (Supplemental Table 2). Transposable-element annotation revealed that ∼39% of the genome consists of transposable elements (Figure 1A; Supplemental Table 3). An evolutionary analysis indicated that *R. tetraphylla* has undergone a whole-genome triplication, which probably resulted from a double hybridization (2n = 6x = 66; Figures 1A and 1B), and a marked expansion of several orthogroups (Figure 1C and Supplemental Figure 4). This also resulted in a substantial expansion of genes encoding ADHs, notably including 372 medium-chain dehydrogenase/reductases, 317 short-chain ADHs, and 135 aldo-keto reductases (Supplemental Table 4).

Using this new genome, we searched for putative natural product biosynthetic gene clusters (supplemental materials and methods; Supplemental Tables 5 and 6). Among them, we identified 3 genomic regions located on chromosomes 11a, 11b, and 11c that consisted of 10, 9, and 9 successive genes, respectively, all of which encoded ADHs corresponding to cinnamyl-ADH-like genes from the medium-chain dehydrogenase/reductase family (Figure 1D). Besides being collinear because they result from polyploidization (Figures 1A and 1D), these regions also shared a high degree of synteny with a locus enriched in genes encoding ADHs involved in heteroyohimbane synthesis (tetrahydroalstonine synthase [THAS]) found on chromosome 4 of *C. roseus* (Sun et al., 2023), suggesting a putative local duplication of *THAS1* and *THAS3* orthologs in *R. tetraphylla* prior to the whole-genome triplication (Supplemental Figure 5). We determined that 25 of the 28 genes were associated with complete ADH proteins and clustered into 7 ADH identity groups (Figure 1E). A homology search revealed that VR2 was located in the studied genomic regions and corresponded to Rte11bG086277, with two close homologs, Rte11aG083588 and Rte11cG087148 (Supplemental Table 7). The three genes show conserved local synteny (Figure 1D), together with a common phylogenetic clustering (Figure 1E).

Such density of ADHs prompted us to investigate the activity of the genomic neighbors of VR2 that may encode the missing ADH of the ajmaline pathway. On the basis of a high expression level in roots (Supplemental Figure 6), one representative of each ADH identity group was amplified and assayed by transient expression in *Nicotiana benthamiana*, together with VH and VR2

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

Please cite this article in press as: Lezin et al., A chromosome-scale genome assembly of *Rauvolfia tetraphylla* facilitates identification of the complete ajmaline biosynthetic pathway, Plant Communications (2023), https://doi.org/10.1016/j.xplc.2023.100784

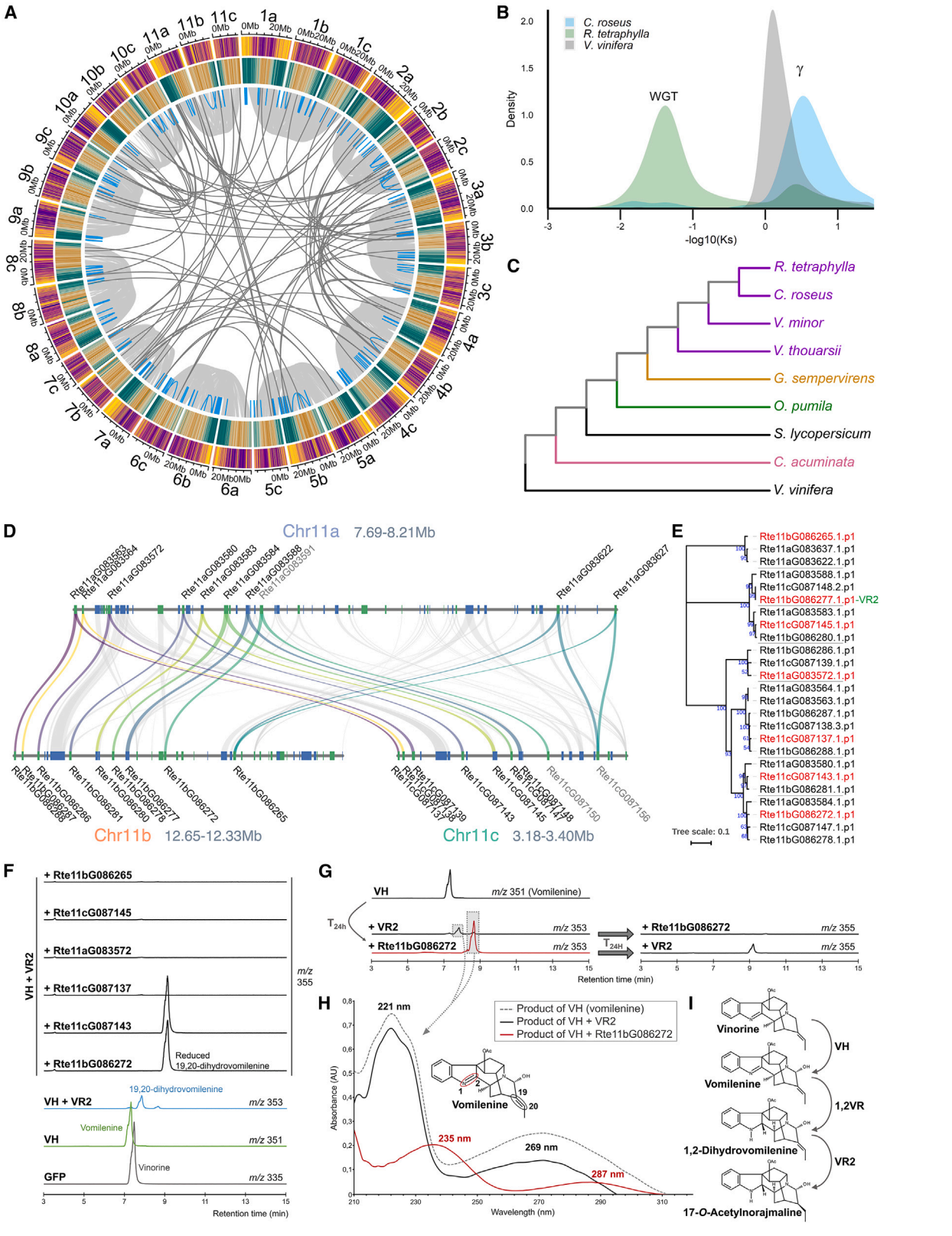

Figure 1. Chromosome-scale genome of *R. tetraphylla* and identification of 1,2-dihydrovomilenine reductase.

(A) Genomic landscape of *R. tetraphylla*. Concentric rings present, from the outside to the inside, pseudo-chromosome name, pseudo-chromosome scale, gene density (purple: low density; yellow: high density), and transposable-element density (blue: low density; brown: high density). Blue central

(legend continued on next page)

(Figure 1F; supplemental materials and methods). Although no modification of the fed vinorine was observed in control leaves expressing only green fluorescent protein (Supplemental Figure 7), overexpression of VH and VH combined with VR2 caused the conversion of vinorine into vomilenine and 19,20-α(S)-dihydrovomilenine, measurable at *m/z* 351 and 353, respectively (Figure 1F). These reactions were fully consistent with the previously reported reaction catalyzed by VH and VR2 (Geissler et al., 2016; Dang et al., 2017). Interestingly, co-expression of Rte11bG086272 and Rte11cG087143 with VH and VR2 led to formation of a new compound whose *m/z* (355) was consistent with an additional reduction of 19,20-α(S)-dihydrovomilenine, potentially yielding 17-*O*-acetyl-norajmaline. By contrast, no such reduction was observed when Rte11cG087137, Rte11cG083572, Rte11cG087145, or Rte11bG086265 was expressed, confirming the specificity of the reaction catalyzed by Rte11bG086272 and Rte11cG087143. In addition, individual co-expression of each of the four aforementioned genes with VH revealed that Rte11cG087145 catalyzed a vomilenine reduction similar to that of VR2 (Supplemental Figure 8).

To gain insight into the identity of the vomilenine derivatives produced in these assays, VR2, Rte11bG086272, and Rte11cG087143 were individually co-expressed with VH (Figure 1G, left, and Supplemental Figure 8). As previously observed for VR2, we noted that Rte11bG086272 and Rte11cG087143 were capable of reducing vomilenine directly, as revealed by formation of an *m/z* 353 product. However, the difference in retention times of the VR2 and Rte11bG086272/Rte11cG087143 products strongly argues for the formation of two distinct compounds. We took advantage of these syntheses to assign the characteristic UV spectrum changes of vomilenine derivatives to the VR2 and Rte11bG086272/Rte11cG087143 products (Figure 1H and Supplemental Figure 8). As described by Geissler et al. (2016), we first observed that both vomilenine and the VR2 product (19,20-α(S)-dihydrovomilenine) displayed similar spectra, reaching two maxima at 221 and 269 nm. Interestingly, the Rte11bG086272/Rte11cG087143 product exhibited a radical spectrum shift, with two maxima at 235 and 287 nm, characteristic of the reduction of the indolenine ring in the 1,2-position found in 1,2-dihydrovomilenine. The identity of this compound was further confirmed by mass fragmentation, which clearly revealed differences in the reductions catalyzed by Rte11bG086272/Rte11cG087143 and VR2 (Supplemental Figure 9). These results indicate that both Rte11bG086272 and Rte11cG087143 encode the missing enzyme of the ajmaline pathway, namely 1,2VR, which catalyzes the 1,2 reduction of vomilenine.

To establish the preferential ADH reaction order, products generated by co-expression of VH and VR2 or VH and Rte11bG086272 or Rte11cG087143 (1,2VR) were further incubated for 24 h with *N. benthamiana* disks expressing Rte11bG086272/Rte11cG087143 (1,2VR) and VR2, respectively (Figure 1G, right, and Supplemental Figure 10). Interestingly, we observed that the enzymes encoded by Rte11bG086272 and Rte11cG087143 (1,2VR) were not capable of reducing the VR2 product (19,20-α(S)-dihydrovomilenine), whereas VR2 efficiently reduced the Rte11bG086272 and Rte11cG087143 products (1,2-dihydrovomilenine). This strongly suggests that 1,2VRs (Rte11bG086272 or Rte11cG087143) catalyze the first vomilenine reduction, and this is followed by the VR2 reaction, in contrast to the previous hypothesis (Geissler et al., 2016).

In conclusion, this chromosome-scale version of the *R. tetraphylla* genome provides valuable insights into MIA biogenesis through identification of the missing enzyme of the ajmaline pathway. The VR2- and 1,2VR-encoding genes were adjacent in the genome and also displayed a distant copy. This identification will undoubtedly pave the way for future bioproduction of ajmaline in a heterologous organism.

DATA AND CODE AVAILABILITY

Raw DNA sequencing data, Hi-C data, RNA sequencing data, and the genome assembly have been deposited under BioProject numbers PRJNA771251 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA771251) and PRJNA1020772 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1020772). The genome annotation, coding sequences, protein sequences, and transcript sequences are available on figshare at https://doi.org/10.6084/m9.figshare.21679628 (private link: https://figshare.com/s/3b080658774cf533f383).

GENE ACCESSION NUMBERS

The gene accession numbers are as follows: Rte11cG087148 (OR571750), Rte11bG086272 (OR571751), Rte11cG087145 (OR571752), and Rte11cG087143 (OR571753).

links represent intra-chromosomal collinearity, light gray central links represent collinearity within a chromosome group, and dark gray central links represent collinearity between chromosome groups.

(B) Synonymous substitution (Ks) rate distribution plot for *R. tetraphylla* (green) orthologs compared with those of *Catharanthus roseus* (blue) and *Vitis vinifera* (gray). WGT, whole genome triplication identified in *R. tetraphylla*; γ, conserved γ whole-genome triplication event that is shared among eudicots.

(C) Phylogenetic tree of *R. tetraphylla* and eight other species, including three Apocynaceae (purple: *C. roseus*, *V. minor*, *V. thouarsii*), one Gelsemiaceae (yellow: *G. sempervirens*), one Rubiaceae (green: *O. pumila*), and one Cornales (pink: *C. acuminata*).

(D) Maps and synteny of ADH-enriched regions on chromosomes 11a, 11b, and 11c. ADH genes are numbered in black and in gray for complete and partial ADH sequences, respectively. Synteny between ADHs is highlighted with different colors.

(E) Neighbor-joining phylogenetic tree (100 bootstrap replications) of predicted ADH protein sequences. Proteins in red were cloned and used for enzymatic assays. The scale bar represents 0.1 substitutions per site.

(F) Screening of ADH activities for reduction of 19,20-α(S)-dihydrovomilenine (VR2 product). ADH candidates were transiently expressed in tobacco leaves, together with VH and VR2, in the presence of vinorine. The reaction products were analyzed by liquid chromatography–mass spectrometry.

(G) Analyses of the enzymatic order of VR2 and Rte11bG086272 after vinorine hydroxylation by VH. Both ADHs were independently transiently overexpressed in tobacco leaves with VH in the presence of vinorine over 24 h. The reaction media of the two enzymatic assays were added, independently, with the other ADH for an additional 24 h, and reaction products were analyzed by liquid chromatography–mass spectrometry.

(H) UV spectra of vomilenine, 19,20-α(S)-dihydrovomilenine (VR2 product), and Rte11bG086272 product (1,2-dihydrovomilenine).

(I) Preferential order of vomilenine reduction successively catalyzed by 1,2VR and VR2.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

FUNDING

This work was supported by the EU Horizon 2020 research and innovation program (MIAMi project-grant agreement N°814645), the ARD CVL Biopharmaceutical program of the Région Centre-Val de Loire (ETOPOCentre project), and the ANR (project MIACYC – ANR-20-CE43-0010).

AUTHOR CONTRIBUTIONS

S.E.O'C., M.K.J., S.B., and V.C. jointly conceived and supervised the work. C.C., M.D., H.J.J., N.C., T.D.d.B., C.S., and R.P.D. performed chromosome counting, sequencing, assembly, and genome annotation and analysis. E.L., I.C., V.V., C.B.W., A.O., J.P., and N.P. performed enzyme characterization. E.L., C.C., M.D., S.B., and V.C. wrote the manuscript and designed the figures with input from all authors. All authors have read and approved the manuscript

ACKNOWLEDGMENTS

R.P.D. and H.J.J. are CEO and CTO, respectively, of Future Genomics Technologies. M.K.J. has a financial interest in Biomia.

Received: September 26, 2023
Revised: December 18, 2023
Accepted: December 21, 2023

Enzo Lezin[1,10], Inês Carqueijeiro[1,10], Clément Cuello[1,10], Mickael Durand[1,10], Hans J. Jansen[2], Valentin Vergès[1], Caroline Birer Williams[1], Audrey Oudin[1], Thomas Dugé de Bernonville[1,9], Julien Petrignet[3], Noémie Celton[4], Benoit St-Pierre[1], Nicolas Papon[5,6], Chao Sun[5,6], Ron P. Dirks[2], Sarah Ellen O'Connor[7], Michael Krogh Jensen[8], Sébastien Besseau[1,] and Vincent Courdavault[1,*]*

[1]Biomolécules et Biotechnologies Végétales, EA2106, Université de Tours, 37200 Tours, France
[2]Future Genomics Technologies, 2333 BE Leiden, the Netherlands
[3]Laboratoire Synthèse et Isolement de Molécules BioActives (SIMBA, EA 7502), Université de Tours, 37200 Tours, France
[4]Laboratoire de Cytogénetique Constitutionnelle, CHRU de Tours - Hôpital Bretonneau, 37044 Tours, France
[5]University Angers, University Brest, IRF, SFR ICAT, 49000 Angers, France
[6]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 2800, China
[7]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany
[8]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 100193 Kgs Lyngby, Denmark
[9]Present address: Limagrain, Center de Recherche, Route d'Ennezat, Chappes, France
[10]These authors contributed equally to this article.
***Correspondence: Sébastien Besseau (sebastien.besseau@univ-tours.fr), Vincent Courdavault (vincent.courdavault@univ-tours.fr)**

https://doi.org/10.1016/j.xplc.2023.100784

Correspondence

REFERENCES

Dang, T.T.T., Franke, J., Tatsis, E., and O'Connor, S.E. (2017). Dual Catalytic Activity of a Cytochrome P450 Controls Bifurcation at a Metabolic Branch Point of Alkaloid Biosynthesis in Rauwolfia serpentina. Angew. Chem. **56**:9440–9444.

Geissler, M., Burghard, M., Volk, J., Staniek, A., and Warzecha, H. (2016). A novel cinnamyl alcohol dehydrogenase (CAD)-like reductase contributes to the structural diversity of monoterpenoid indole alkaloids in Rauvolfia. Planta **243**:813–824.

Kulagina, N., Méteignier, L.V., Papon, N., O'Connor, S.E., and Courdavault, V. (2022). More than a Catharanthus plant: A multicellular and pluri-organelle alkaloid-producing factory. Curr. Opin. Plant Biol. **67**, 102200.

Kumar, S., Singh, A., Bajpai, V., Srivastava, M., Singh, B.P., Ojha, S., and Kumar, B. (2016a). Simultaneous determination of bioactive monoterpene indole alkaloids in ethanolic extract of seven Rauvolfia species using UHPLC with hybrid triple quadrupole linear ion trap mass spectrometry. Phytochem. Anal. **27**:296–303.

Kumar, S., Singh, A., Bajpai, V., Srivastava, M., Singh, B.P., and Kumar, B. (2016b). Structural characterization of monoterpene indole alkaloids in ethanolic extracts of Rauwolfia species by liquid chromatography with quadrupole time-of-flight mass spectrometry. J. Pharm. Anal. **6**:363–373.

Kumara, P.M., Uma Shaanker, R., and Pradeep, T. (2019). UPLC and ESI-MS analysis of metabolites of Rauvolfia tetraphylla L. and their spatial localization using desorption electrospray ionization (DESI) mass spectrometric imaging. Phytochemistry **159**:20–29.

Li, C., Wood, J.C., Vu, A.H., Hamilton, J.P., Rodriguez Lopez, C.E., Payne, R.M.E., Serna Guerrero, D.A., Gase, K., Yamamoto, K., Vaillancourt, B., et al. (2023). Single-cell multi-omics in the medicinal plant Catharanthus roseus. Nat. Chem. Biol. **19**:1031–1041.

Sun, S., Shen, X., Li, Y., Li, Y., Wang, S., Li, R., Zhang, H., Shen, G., Guo, B., Wei, J., et al. (2023). Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartmentation of specialized metabolism. Nat. Plants **9**:179–190.

Zhang, J., Hansen, L.G., Gudich, O., Viehrig, K., Lassen, L.M.M., Schrübbers, L., Adhikari, K.B., Rubaszka, P., Carrasquer-Alvarez, E., Chen, L., et al. (2022). A microbial supply chain for production of the anti-cancer drug vinblastine. Nature **609**:341–347.