

[http://pubs.acs.org/journal/acsodf](http://pubs.acs.org/journal/acsodf?ref=pdf) **Article** Article

# **Dissecting the Herbicidal Mechanism of Microbial Natural Product Lydicamycins Using a Deep Learning-Based Nonlinear Regression Model**

[Xiaoyu](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Xiaoyu+Wang"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) Wang, [Heqian](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Heqian+Zhang"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) Zhang, [Xuanlin](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Xuanlin+Zhan"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) Zhan, Jie [Li,](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Jie+Li"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) [Jiaquan](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Jiaquan+Huang"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) Huang,[\\*](#page-5-0) and [Zhiwei](https://pubs.acs.org/action/doSearch?field1=Contrib&text1="Zhiwei+Qin"&field2=AllField&text2=&publication=&accessType=allContent&Earliest=&ref=pdf) Qin[\\*](#page-5-0)



microbe interactions, but the mechanisms are often complex and nonlinear. Here we show the nonlinear regulatory effects of *Streptomyces ginsengnesis* G7 on *Arabidopsis thaliana* growth. We focused on lydicamycin, a molecule from this bacterium that interferes with auxin polar transport. Using a deep learning approach with a feedforward neural network, we integrated multiomics data to elucidate the mechanism of lydicamycin on plant growth and development. We also examined the impact of flavonol metabolites, particularly isorhamnetin from *A. thaliana*, on the PIN protein family's role in auxin transport. Our findings indicate that lydicamycin regulates auxin transport by inducing flavonol overaccumulation in *A. thaliana*, affecting plant development. This study identifies potential molecular targets for crop enhancement and improved agricultural productivity.

# ■ **INTRODUCTION**

An omics data set stands as a valuable asset in advancing biological research and comprehending complex biological systems. Ideally, a high-quality omics data set demonstrates key characteristics such as comprehensiveness, rigorous documentation, and accessibility. Crucially, it should seamlessly integrate with other omics data sets, promoting a holistic system biology approach. Customized methods are imperative for connecting omics data generated from a single biological sample simultaneously, i.e., at the same spatiotemporal scale. This approach, known as multiomics joint analysis, involves integrating and analyzing data from diverse omics disciplines such as genomics, transcriptomics, proteomics, metabolomics, and epigenomics.

The rhizosphere microbial community plays a crucial role in maintaining plant hormone balance, controlling root development, promoting nutrient absorption, and enhancing plant resistance to both biotic and abiotic stresses.<sup>[1](#page-5-0)</sup> For biocontrol agent development, rhizosphere *Streptomyces*-based biocontrol agents, such as Actinovate and MYCOSTOP, have been widely used to combat agricultural fungal diseases. It is worth noting that different plants exhibit significant variations in their macro- and microscopic responses to the same microorganisms and their metabolites.<sup>[2](#page-5-0)</sup> These differences may be attributed to factors such as plant species, the interaction dynamics between the plants and microbes, and the modes of action of growthpromoting molecules.



Our recent focus is on *Streptomyces ginsengnesis* G7, a species discovered in the Ginseng rhizosphere.<sup>3</sup> Through an integrated approach-encompassing plant-microbe interaction tests, bioinformatics analysis, and antimicrobial assays—we unveiled this strain's dual role as a producer of herbicides and antibiotics. Remarkably, G7 produces type I polyketide alkaloid lydicamycins, which exposes an unprecedented herbicidal mechanism in *Arabidopsis thaliana* (AT) by effectively inhibiting auxin transport. Briefly, lydicamycin can inhibit the growth of both primary and lateral roots in rice weeds, while it has a minor effect on the primary root growth and promotes the emergence of lateral roots of rice. This reveals interspecies differences in plant responses to natural products produced by specific microbes, as well as the potential complexity of the natural products in regulating plant growth.

To further extend our previous work, here we introduce a deep learning-based nonlinear regression model to adeptly analyze the complex, high-dimensional, and interrelated omics data generated during plant−microbe interactions, including the predominant factors induced in AT through G7 treatment.

Received: August 29, 2024 Revised: October 14, 2024 Accepted: October 17, 2024 Published: October 24, 2024





© <sup>2024</sup> The Authors. Published by American Chemical Society **<sup>44778</sup>**

<span id="page-1-0"></span>

Figure 1. Integrated analysis framework and model visualization. (A) Flowchart illustrates the data structure that integrates both upregulated and downregulated genes from transcriptomic, proteomic, and metabolomic data sets. This includes a subset of genes shared between transcriptomic and proteomic data. Each sample is labeled accordingly, with control samples marked as label = 0.5 and treated samples as label = 1. (B) The neural network architecture used in the study comprises four layers of regression coefficients, including one input layer, two hidden layers and one output layer. Tanh, chosen as the activation function, is applied after the first three layers. To prevent overfitting, the model incorporates two dropout layers, Xavier weight initialization, and L2 regularization. Additionally, to augment the data set, three Gaussian random noise samples (mean = 0.0, std = 0.01) are added to each sample in the training set for every LOOCV iteration. (C) The interpretation results showcase the evaluation graph of the 100 bootstrap replications on the evaluation data set; permutation feature importance plotted along the *y*-axis against all omics features on the *x*-axis; and subsequent functional analysis with GO and KEGG pathway annotations.

Our work not only advances data mining methods in complex biological systems but also holds significant practical value for precisely controlling and optimizing interactions between specific plant hosts and growth-promoting microorganisms, as well as replicating these benefits across different agricultural ecosystems.

# ■ **RESULTS**

**Model Structure and Performance Evaluation.** In this part, we aim to enhance the model's predictive ability and demonstrate its generalization ability in handling missing or previously unaccounted data, thereby providing clear insights into various bioprocesses. To this end, we have implemented data preprocessing procedures, which involved the creation of a comprehensive list comprising all three types of up- and down-regulated omics data set (transcriptome, proteome, and metabolome). Furthermore, we identified a subset of genes common to both transcriptomics and proteomic data (Figure 1A), facilitating the analysis of their collective alterations

within the entire system. To accomplish this, we developed a four-layered multilayer perceptron (MLP) deep learning model tailored for fitting of multiomics features, as shown in Figure  $1B<sup>4</sup>$  $1B<sup>4</sup>$  $1B<sup>4</sup>$ 

The raw data for our model was sourced from six AT samples, comprising three from wild-type seedlings and three from G7-treated seedlings. The dimensions of the ultimately chosen omics data are detailed in [Table](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S1. To differentiate between the two groups, target values were intentionally set for regression fitting, deliberately assigning "1" to wild-type samples and "0.5" to G7-treated samples. Note that positive feature values contribute more to wild-type samples, while negative values pinpoint G7-treated samples. Furthermore, the absolute value magnitude reflects their significance in distinguishing between sample types. For optimal utilization of limited-sample data, we designated one sample from each of the two groups for the evaluation set. The remaining four samples were employed as both training and testing sets in the leave-one-out cross validation (LOOCV) deep learning



Figure 2. Multiomics analysis of the interaction process of plant−microbe interactions. (A) A coculturing model demonstrates that *S. ginsengnesis* G7 can effectively inhibit the growth of *A. thaliana* through the production polyketide natural product lydicamycin (an illustrative chemical structure has been shown below). Recent studies suggest a potential mechanism wherein lydicamycins effectively inhibit auxin transport. (B) The diagram indicates the feature importance from different data classifiers (transcriptomic, proteomic, and metabolomic, abbreviated as "trans", "pros", and "meta") to various biological processes categorized by GO terms. The width of each flow represents the magnitude of feature importance, with the color gradient indicating the strength of the contribution, ranging from high (red) to low (green). This integrated approach allows for the visualization of the significant pathways and processes influenced by the omics data, linking molecular interactions to observable phenotypic outcomes.

process.[5](#page-5-0) To prevent information leakage, data scaling was performed after each LOOCV iteration's separation of training and testing sets. Each omics data set, along with the corresponding training and testing sets, underwent independent min−max scaling using MinMaxScaler from the Scikit-learn package.[6](#page-5-0) The model exhibiting the least validation mean absolute error (MAE) across all four iterations in LOOCV was preserved as the best-performing model. Finally, to evaluate the model's consistency with our limited sample size, we conducted 100 replications on the evaluation set. The resulting mean  $R<sup>2</sup>$  score of 99% across all applications underscores the model's robust predictive power ([Figure](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S1), forming a solid foundation for subsequent interpretive analyses [\(Figure](#page-1-0)  $1C$ ).

Next, we employed the local interpretable model-agnostic explanations (LIME) technique to elucidate the model predictions.<sup>[8](#page-6-0)</sup> This method enhances understanding of individual feature contributions by creating perturbed samples around the selected one. Utilizing our fine-tuned model with the normalized data set containing all six samples, we precisely identified the impact of each omics feature, categorizing them based on their positive or negative influences on predictions ([Tables](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S2 and S3).

**G7 Treatment Effects and Biological Implications.** Our previous studies showed that lydicamycin exhibits the inhibition effect on AT seed growth (Figure 2A), but did not fully leverage multiomics to explore the insights into its mechanisms.<sup>[3](#page-5-0)</sup> In this work, using the newly established model, we identified 1151 negative features as potential markers for G7 treatment ([Figure](#page-1-0) 1A), comprising 449 transcripts, 415 proteins, and 287 metabolites. In contrast to prior analyses focusing on differentially expressed genes, proteins, and metabolites, our model's examination of negative features reveals associations with biological processes in GO enrichment analysis, such as stress response, metabolic processes, detoxification and others (Figures 2B, S2, [Table](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S4). First, our comprehensive integration of transcriptome and proteome data highlights a significant involvement of lipid metabolic pathways ([Figure](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S2). In the KEGG pathway analysis of these features, we noted disruptions in glycerolipid and glycerophospholipid metabolism within the transcriptome and proteome, accompanied by decreased levels of hexadecanoic acid and octadecanoic acid in the metabolome ([Figure](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S3). The observed impairment in lipid biosynthesis could adversely affect membrane functionality.<sup>[9](#page-6-0)</sup> Second, previous reports have linked the accumulation of key secondary metabolites to altered root development, with flavonoids identified as negative regulators of auxin transport playing a pivotal role.<sup>[10,11](#page-6-0)</sup>

Note. The spore suspension  $(OD_{600} = 0.6)$  of G7 was smeared at the lower  $1/2$  MS gel area of a  $10 \times 10$  cm square Petri dish, and AT seedlings spread into 3 lines at the upper two-thirds area, displaying a gradient effect on AT growth inhibition where seedlings closer to the G7 area showed more severe growth suppression. According to our previous study, lydicamycin concentrations remained stable after 28 h.

**Unraveling the Interplay between Lydicamycin and Auxin Distribution.** In our earlier investigation, we established that lydicamycin influences the auxin gradient distribution by suppressing the expression of PIN (PINformed) family proteins, yet the underlying mechanism remained unclear.<sup>[3](#page-5-0)</sup> While in this work, the observed upregulation of proteins associated with flavonoid biosynthesis, coupled with substantial compound accumulation supported by metabolomic ([Figure](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S3), suggests that lydicamycin may activate flavonoid biosynthesis. This activation is likely to result in a substantial increase in flavonoid levels, which might consequently exert a negative regulatory effect on auxin



Figure 3. A proposed model in this study showing an overview of the abnormal flavonol accumulation and root development observed upon lydicamycin treatment. Lydicamycin enters root cells through diffusion, initiating an aberrant synthesis of isorhamnetin. The cell diagram illustrates the biosynthetic pathway of flavonoids, although the precise molecular mechanism remains to be elucidated. Following its synthesis, isorhamnetin diffuses and undergoes vesicular trafficking between cells, potentially interacting with PIN proteins. This interaction perturbs the polar transport of auxin, resulting in its accumulation in pericycle cells. Subsequent auxin build up stimulates lateral root growth.

transport through the stabilization of PIN efflux complexes, as indicated in previous studies. $11$  GO annotation of the top 20% of negative values revealed transcripts and proteins annotated for participation in lipid metabolism, root development, and secondary metabolism [\(Table](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S3). The upregulation of biosynthesis and accumulation of flavonoids has been evidenced in plants experiencing various environmental stresses as part of adaptive response, performing as detoxification, antioxidant and signaling molecules, etc.<sup>[12](#page-6-0)</sup> This reinforces the notion that the magnitude of feature values corresponds to their impact on the model's predictions, with larger values indicating involvement in core pathways

([Table](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf) S3). This dual function mediated by flavonoid accumulation highlights the complex interplay between stress tolerance and growth regulation in plants.<sup>[13](#page-6-0)</sup> Future work will delve into these mechanisms in greater detail.

As shown in Figure 3, our model proposes a plausible mechanism for lydicamycin entering root cells and inducing the accumulation of isorhamnetin, which subsequently inhibits PIN protein function. Lydicamycin initially enters cells via diffusion, triggering abnormal isorhamnetin synthesis within them. This compound then diffuses and undergoes vesicular trafficking between cells, potentially interacting with PIN proteins and destabilizing their polymers. This disruption hinders the polar transport of auxin, leading to its accumulation in pericycle cells and thereby stimulating lateral root growth.

## ■ **DISCUSSION**

Plants form complex symbiotic relationships with surrounding microbial communities in natural environments, which can be classified as mutualistic, neutral, or parasitic.<sup>[14](#page-6-0)</sup> Although many growth-promoting microbes have been identified, their varying effects on different plants limit their widespread application. In recent years, the role of microbial-derived natural products as signaling molecules in interspecies communication between plants and microbes has garnered significant attention from the scientific community.<sup>[15](#page-6-0)</sup> When triggering plant immune responses, different plants exhibit substantial variations in both macro- and microscopic responses to the same molecules, though the specific reasons for which remain unclear. Recent studies have shown that plants can selectively attract Streptomycete which produce diverse and complex natural products, and colonize their rhizosphere and internal environments. In turn, Streptomycete forms specific hyphal growth patterns and metabolic networks to promote and regulate their interactions with plant roots.<sup>[16](#page-6-0),[17](#page-6-0)</sup> Notably, Streptomycete can secrete plant growth regulators such as IAA and siderophores to promote plant growth, secrete specific signaling molecules to induce plant systemic resistance, or produce antimicrobial substances to adjust the rhizosphere microbial community, thereby effectively enhancing plant growth and stress tolerance.[17,18](#page-6-0) Moreover, recent studies indicate that Streptomycete can directly activate plant auxin signaling pathways by secreting indole analogs or enhance photosynthesis and regulate stress responses by secreting fervenulin, thereby increasing plant tolerance to drought and salt stress.<sup>[19](#page-6-0)</sup> However, how other small molecules produced by Streptomycete specifically affect plant growth and development remains an open research area. Extensive research in this field will deepen our understanding of plant−microbe interaction mechanisms and provide a scientific basis for discovering new plant growth regulators or developing novel biopesticides.

In this work, an interaction model was previously constructed to study how microbial-derived natural products regulate plant development during plant−microbe interactions. It was found that lydicamycin produced by G7 can silence and disrupt the expression and distribution of the PIN protein family in *Arabidopsis*, thereby inhibiting primary root growth.<sup>20</sup> However, since the structure of lydicamycin is not similar to NPA, it is inferred that lydicamycin does not directly target PIN proteins to affect auxin polar transport. To explain the biological phenomena caused by potential nonlinear regulatory mechanisms, we aim to integrate *Arabidopsis* transcriptomic and metabolomic data by constructing a deep learning framework based on a feedforward neural network. The study revealed an abnormal accumulation of secondary metabolites, including isorhamnetin, which was upregulated 8-fold in the experimental group, while other flavonol metabolites did not show significant changes. Previous studies have shown that flavonols can affect the activity or localization of PIN proteins, thereby inhibiting auxin polar transport.<sup>[11,21](#page-6-0)</sup> Therefore, our future work will focus on determining whether the significant accumulation of isorhamnetin in *Arabidopsis* affects the expression and localization of PIN proteins.

To conclude, we have successfully developed a robust tool for the integrated analysis of multiomics data sets derived from

plant−microbe interactions. As discussed earlier, our findings suggest that there is a wealth of additional insights waiting to be uncovered, particularly because plants and their associated microbes reside in complex, ever-changing, and unpredictable environments.

## ■ **MATERIALS AND METHODS**

**Sample Preparation.** The AT seeds were surface sterilized by immersing them in a 50% ethanol solution for 5 min, followed by a 0.5% NaOCl treatment for another 5 min. The sterilized seeds were then rinsed three times with water. After rinsing, the seeds were sown on 1/2 MS gel and placed in darkness at 4 °C for 2 days. The seeded Petri dishes were subsequently positioned vertically in a growth chamber set to a long-day photoperiod of 16 h of light, with a light intensity of  $300 \ \mu \text{M m}^{-2} \text{ s}^{-1}$ , and maintained at a constant temperature of 22 °C.

The G7 spore suspension  $OD_{600} = 0.6$ ) was spread on a standard  $1/2$  MS gel plate  $(10 \times 10 \text{ cm})$ , covering approximately one-third of the total surface area. After 96 h of fermentation, 2 day-old AT seedlings were vertically transferred to the aforementioned 1/2 MS gel plate with the grown G7. Multiomics analysis was performed on samples from both the G7-treated group and the mock-treated group. All samples were collected from three independent experiments, with multiple plates used within each experiment to ensure robustness and reproducibility.

**Data Preparation.** In our research to estimate the most evoked factors in AT by G7 treatment from multiomics data including transcriptomics, proteomics, and metabolomics, we employed data preprocessing procedures constructing a combined list of all three up/down regulated omics data, plus a group of overlapped genes present both in transcriptomics and proteomics data to analyze their changes as a whole system, a four-layered MLP deep learning model was constructed for multiomics features logistic regression fitting.

Our raw data comes from 6 AT samples, 3 of them are generated from wild-type tissues, and 3 are from G7 treated tissues. The dimensions of the ultimately selected omics data are listed in Table 1. The values of 1 and 0.5 were assigned to

#### Table 1. Dimension for Each Omics Type



wild-type and G7-treated samples, respectively, to serve as the target values for regression fitting task. These values are intentionally set as labels to differentiate the groups for the analysis.

**Deep Learning Model Construction.** The MLP model we have built for regression fitting of all three omics profiles has been depicted in [Figure](#page-1-0) 1. It contains 4 layers with one input layer, two hidden layers, and one output layer of regression coefficients, with Tanh as a better-performed activation function placed after the first 3 layers. Two dropout layers, Xavier weight initialization, and L2 regularization were conducted to prevent overfitting. For data augmentation, we added three Gaussian noise samples (mean = 0, standard deviation  $= 0.01$ ) to each sample in the training set of every LOOCV iteration. Model performance before and after noise is

<span id="page-5-0"></span>evaluated by 100 replications of predictions on the evaluation set.

To utilize small-sample data more effectively, we took each sample from the 2 groups for evaluation set, the rest 4 samples were used as training and testing sets in the LOOCV deep learning process. To avoid information leakage, data scaling is conducted after separating the training and testing sets in each LOOCV iteration. We applied min−max scaling to each omics data set and to the training and testing sets separately, using the MinMaxScaler from the Scikit-learn package. Among the four LOOCV iterations, the model with the lowest validation MAE was saved as the best-performing model.

**Data Interpretation.** We used the LIME technique for explaining predictions made by machine learning models. It provides an understanding of the contribution of each feature by generating a set of perturbed samples around the chosen sample and using the original model to predict the outcomes of these samples. These newly generated data points and their predicted outcomes are used to train a simple model. The coefficients of the simple model can then be interpreted as the contribution of each feature. With LIME assigned importance values to each omics feature, they were separated into two groups with negative/positive influences on the model's predictions.

# ■ **ASSOCIATED CONTENT**

#### **Data Availability Statement**

The authors declare that the data, materials and code supporting the findings reported in this study are available from the authors upon reasonable request. The -omic data is available in GitHub repository: [https://github.com/](https://github.com/Qinlab502/testOmics) [Qinlab502/testOmics.](https://github.com/Qinlab502/testOmics)

#### $\bullet$  Supporting Information

The Supporting Information is available free of charge at [https://pubs.acs.org/doi/10.1021/acsomega.4c07971.](https://pubs.acs.org/doi/10.1021/acsomega.4c07971?goto=supporting-info)

> Comprehensive details regarding the -omic data sets, algorithm development, and the tabulated data supporting the main conclusions in this work [\(PDF](https://pubs.acs.org/doi/suppl/10.1021/acsomega.4c07971/suppl_file/ao4c07971_si_001.pdf))

#### ■ **AUTHOR INFORMATION**

#### **Corresponding Authors**

Jiaquan Huang − *Center for Biological Science and Technology, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China*; Email: [jiaquan\\_terry@bnu.edu.cn](mailto:jiaquan_terry@bnu.edu.cn)

Zhiwei Qin − *Center for Biological Science and Technology, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China;* [orcid.org/0000-0002-1444-3684;](https://orcid.org/0000-0002-1444-3684) Email: [z.qin@](mailto:z.qin@bnu.edu.cn) [bnu.edu.cn](mailto:z.qin@bnu.edu.cn)

## **Authors**

- Xiaoyu Wang − *Center for Biological Science and Technology, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China*
- Heqian Zhang − *Center for Biological Science and Technology, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China*

Xuanlin Zhan − *Center for Biological Science and Technology, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China*

Jie Li − *Department of Biochemistry and Metabolism, John Innes Centre, Norwich NR4 7UH, U.K.*

Complete contact information is available at: [https://pubs.acs.org/10.1021/acsomega.4c07971](https://pubs.acs.org/doi/10.1021/acsomega.4c07971?ref=pdf)

#### **Author Contributions**

X.W. and H.Z. contributed equally to this work. X.W.: Conceptualization, methodology, data curation, writing−review and editing. H.Z.: Conceptualization, methodology, data curation, writing−review and editing, funding acquisition. X.Z.: Data curation, formal analysis. J.L.: Conceptualization, methodology, data curation, writing−review and editing. J.H.: Conceptualization, methodology, data curation, writing−review and editing, funding acquisition. Z.Q.: Conceptualization, methodology, project administration, writing−review and editing, funding acquisition.

#### **Notes**

The authors declare no competing financial interest.

## ■ **ACKNOWLEDGMENTS**

This work was supported by the National Natural Science Foundation of China (32170079 to Z.Q., 32200035 to H.Z., and 32400235 to J.H.), the Natural Science Foundation of Guangdong (2021A1515012026 and 2024A1515012593 to Z.Q., 2023A1515110175 to J.H.), Guangdong Talent Scheme (2021QN020100 to Z.Q.), Guangdong Innovation Research Team for Plant−Microbe Interaction (2021KCXTD037 to Z.Q.), as well as Beijing Normal University via the Youth Talent Strategic Program Project (310432104 to Z.Q.). We are grateful to the Interdisciplinary Intelligence Super Computer Center, Beijing Normal University, for High Performance Computing for access to computational resources.

#### ■ **REFERENCES**

(1) Philippot, L.; Raaijmakers, J. M.; Lemanceau, P.; Van Der Putten, W. H. Going back to the roots: the [microbial](https://doi.org/10.1038/nrmicro3109) ecology of the [rhizosphere.](https://doi.org/10.1038/nrmicro3109) *Nat. Rev. Microbiol.* 2013, *11* (11), 789−799.

(2) (a) Zhalnina, K.; Louie, K. B.; Hao, Z.; Mansoori, N.; da Rocha, U. N.; Shi, S.; Cho, H.; Karaoz, U.; Loque, D.; Bowen, B. P.; et al. Dynamic root exudate chemistry and microbial substrate [preferences](https://doi.org/10.1038/s41564-018-0129-3) drive patterns in [rhizosphere](https://doi.org/10.1038/s41564-018-0129-3) microbial community assembly. *Nat. Microbiol.* 2018, *3* (4), 470−480. (b) Panke-Buisse, K.; Poole, A. C.; Goodrich, J. K.; Ley, R. E.; Kao-Kniffin, J. [Selection](https://doi.org/10.1038/ismej.2014.196) on soil [microbiomes](https://doi.org/10.1038/ismej.2014.196) reveals reproducible impacts on plant function. *ISME J.* 2015, *9* (4), 980−989.

(3) Huang, J.; Li, X.; Zhan, X.; Pan, S.; Pan, C.; Li, J.; Fan, S.; Zhang, L.; Du, K.; Du, Z.; et al. A *[Streptomyces](https://doi.org/10.1093/plphys/kiae006)* species from the ginseng [rhizosphere](https://doi.org/10.1093/plphys/kiae006) exhibits biocontrol potential. *Plant Physiol.* 2024, *194* (4), 2709−2723.

(4) (a) Zhao, L.; Dong, Q.; Luo, C.; Wu, Y.; Bu, D.; Qi, X.; Luo, Y.; Zhao, Y. DeepOmix: A scalable and [interpretable](https://doi.org/10.1016/j.csbj.2021.04.067) multi-omics deep learning framework and [application](https://doi.org/10.1016/j.csbj.2021.04.067) in cancer survival analysis. *Comput. Struct. Biotechnol. J.* 2021, *19*, 2719−2725. (b) Reel, P. S.; Reel, S.; Pearson, E.; Trucco, E.; Jefferson, E. Using [machine](https://doi.org/10.1016/j.biotechadv.2021.107739) learning approaches for [multi-omics](https://doi.org/10.1016/j.biotechadv.2021.107739) data analysis: A review. *Biotechnol. Adv.* 2021, *49*, 107739.

(5) Rivals, I.; Personnaz, L. On cross [validation](https://doi.org/10.1162/089976699300016476) for model selection. *Neural Comput.* 1999, *11* (4), 863−870.

(6) Abd Halim, K. N.; Mohd Jaya, A. S.; Fadzil, A. [Data](https://doi.org/10.35940/ijitee.C8472.019320) Pre-Processing Algorithm for Neural Network Binary [Classification](https://doi.org/10.35940/ijitee.C8472.019320) Model in Bank [Tele-Marketing.](https://doi.org/10.35940/ijitee.C8472.019320) *Int. J. Innovative Technol. Explor. Eng.* 2020, *9*, 272−277.

 $(7)$  Ash, A.; Shwartz, M.  $R^2$ : a useful measure of model [performance](https://doi.org/10.1002/(SICI)1097-0258(19990228)18:4<375::AID-SIM20>3.0.CO;2-J) when predicting a [dichotomous](https://doi.org/10.1002/(SICI)1097-0258(19990228)18:4<375::AID-SIM20>3.0.CO;2-J) outcome. *Stat. Med.* 1999, *18* (4), 375−384.

<span id="page-6-0"></span>(8) Ribeiro, M. T.; Singh, S.; Guestrin, C. J. A. "Why Should I Trust You?": Explaining the Predictions of Any Classifier *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; KDD, 2016.

(9) Ohlrogge, J.; Browse, J. Lipid [biosynthesis.](https://doi.org/10.1105/tpc.7.7.957) *Plant Cell* 1995, *7* (7), 957−970.

(10) Brown, D. E.; Rashotte, A. M.; Murphy, A. S.; Normanly, J.; Tague, B. W.; Peer, W. A.; Taiz, L.; Muday, G. K. [Flavonoids](https://doi.org/10.1104/pp.126.2.524) act as negative regulators of auxin transport in vivo in *[Arabidopsis](https://doi.org/10.1104/pp.126.2.524)*. *Plant Physiol.* 2001, *126* (2), 524−535.

(11) Teale, W. D.; Pasternak, T.; Dal Bosco, C.; Dovzhenko, A.; Kratzat, K.; Bildl, W.; Schworer, M.; Falk, T.; Ruperti, B.; V Schaefer, J.; et al. [Flavonol-mediated](https://doi.org/10.15252/embj.2020104416) stabilization of PIN efflux complexes regulates polar auxin [transport.](https://doi.org/10.15252/embj.2020104416) *EMBO J.* 2021, *40* (1), No. e104416.

(12) Kolb, C. A.; Kaser, M. A.; Kopecky, ́ J.; Zotz, G.; Riederer, M.; Pfundel, E. E. Effects of natural [intensities](https://doi.org/10.1104/pp.010373) of visible and ultraviolet radiation on epidermal ultraviolet screening and [photosynthesis](https://doi.org/10.1104/pp.010373) in grape [leaves.](https://doi.org/10.1104/pp.010373) *Plant Physiol.* 2001, *127* (3), 863−875.

(13) Shomali, A.; Das, S.; Arif, N.; Sarraf, M.; Zahra, N.; Yadav, V.; Aliniaeifard, S.; Chauhan, D. K.; Hasanuzzaman, M. [Diverse](https://doi.org/10.3390/plants11223158) physiological roles of flavonoids in plant [environmental](https://doi.org/10.3390/plants11223158) stress [responses](https://doi.org/10.3390/plants11223158) and tolerance. *Plants* 2022, *11* (22), 3158.

(14) Chagas, F. O.; Pessotti, R. d. C.; Caraballo-Rodríguez, A. M.; Pupo, M. T. [Chemical](https://doi.org/10.1039/C7CS00343A) signaling involved in plant−microbe [interactions.](https://doi.org/10.1039/C7CS00343A) *Chem. Soc. Rev.* 2018, *47* (5), 1652−1704.

(15) Hutchings, M. I.; Truman, A. W.; Wilkinson, B. [Antibiotics:](https://doi.org/10.1016/j.mib.2019.10.008) past, [present](https://doi.org/10.1016/j.mib.2019.10.008) and future. *Curr. Opin. Microbiol.* 2019, *51*, 72−80.

(16) Chater, K. F. Recent advances in [understanding](https://doi.org/10.12688/f1000research.9534.1) *Streptomyces*. *FResearch* 2016, *5*, 2795.

(17) Viaene, T.; Langendries, S.; Beirinckx, S.; Maes, M.; Goormachtig, S. *[Streptomyces](https://doi.org/10.1093/femsec/fiw119)* as a plant's best friend? *FEMS Microbiol. Ecol.* 2016, *92* (8), fiw119.

(18) Santos-Medellín, C.; Liechty, Z.; Edwards, J.; Nguyen, B.; Huang, B.; Weimer, B. C.; Sundaresan, V. [Prolonged](https://doi.org/10.1038/s41477-021-00967-1) drought imparts lasting [compositional](https://doi.org/10.1038/s41477-021-00967-1) changes to the rice root microbiome. *Nat. Plants* 2021, *7* (8), 1065−1077.

(19) (a) Yang, Z.; Qiao, Y.; Konakalla, N. C.; Strøbech, E.; Harris, P.; Peschel, G.; Agler-Rosenbaum, M.; Weber, T.; Andreasson, E.; Ding, L. *[Streptomyces](https://doi.org/10.1038/s41467-023-43177-3)* alleviate abiotic stress in plant by producing [pteridic](https://doi.org/10.1038/s41467-023-43177-3) acids. *Nat. Commun.* 2023, *14* (1), 7398. (b) Lu, L.; Liu, N.; Fan, Z.; Liu, M.; Zhang, X.; Tian, J.; Yu, Y.; Lin, H.; Huang, Y.; Kong, Z. A novel PGPR strain, *[Streptomyces](https://doi.org/10.1111/pce.14847) lasalocidi* JCM 3373T, alleviates salt stress and shapes root [architecture](https://doi.org/10.1111/pce.14847) in soybean by secreting indole-[3-carboxaldehyde.](https://doi.org/10.1111/pce.14847) *Plant, Cell Environ.* 2024, *47* (6), 1941−1956.

(20) (a) Huang, J.; Wu, Y.; Gao, Q.; Li, X.; Zeng, Y.; Guo, Y.; Zhang, H.; Qin, Z. [Metagenomic](https://doi.org/10.1128/aem.02335-23) exploration of the rhizosphere soil microbial community and their [significance](https://doi.org/10.1128/aem.02335-23) in facilitating the development of [wild-simulated](https://doi.org/10.1128/aem.02335-23) ginseng. *Appl. Environ. Microbiol.* 2024, *90* (3), No. e02335. (b) Zhang, H.; Li, X.; Pan, S.; Huang, J.; Qin, Z. Dissecting the [biosynthesis](https://doi.org/10.1039/D3NJ01124K) of the polyketide alkaloid [lydicamycin](https://doi.org/10.1039/D3NJ01124K) using a complex metabolic network. *New J. Chem.* 2023, *47* (26), 12093−12100.

(21) Zhang, X.; Huang, X.; Li, Y.; Tao, F.; Zhao, Q.; Li, W. [Polar](https://doi.org/10.1016/j.phytochem.2021.112702) auxin transport may be [responsive](https://doi.org/10.1016/j.phytochem.2021.112702) to specific features of flavonoid [structure.](https://doi.org/10.1016/j.phytochem.2021.112702) *Phytochemistry* 2021, *185*, 112702.