Identifying Forest Types and Distribution Patterns in Shennongjia National Park

Weiyue Shi [1], Guangfei Zhang [1], Cuili Zhang [1], Haigang Sui [1], Qiming Zhou [1,2], Li Hua [3], Junyi Lui [1]

[1] State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China - (shiweiyue, 2023286190127, zhangcuili)@whu.edu.cn; haigang_sui@263.net; liujunyi_ljy@163.com
[2] Institute of Research and Continuing Education (IRACE), Hong Kong Baptist University, Hong Kong - qiming@hkbu.edu.hk
[3] College of Resources and Environment, Huazhong Agricultural University, Wuhan, China - huali@mail.hzau.edu.cn

Keywords: Forest type identification, Forest distribution pattern, Shennongjia National Park, Multi-source remote sensing.

Abstract

Forests are indispensable ecosystems, providing vital resources and services crucial for human well-being and sustainable development. Remote sensing has emerged as a potent tool for mapping forests across diverse spatial scales. The Shennongjia region stands out globally for its exceptional biodiversity and the presence of rare endangered flora. Despite scholarly attention to the region's forest ecosystem, there remains a gap in detailed and high-resolution assessments of forest type patterns based on remote sensing data, especially following the official establishment of Shennongjia National Park in 2020. This study utilizes multi-resolution and multi-temporal remote sensing data to delineate forest types and distributions within the National Park in 2023. Through the integration of multi-temporal remote sensing data, the Google Earth Engine platform and machine learning techniques, our method achieved an overall accuracy of the six forest types at 86.1%, and the distribution patterns of various forest types generally conform to a natural-law accordance trend with increasing altitude in Shennongjia National Park. We hope that our research results can optimize the workflow for forest type classification, thereby furnishing basic data for national park management.

1. Introduction

Forests play a crucial role in providing essential resources and ecosystem services, which are vital for human well-being and sustainable development (Yu et al., 2023). Acting as natural regulators of climate and environmental characteristics, forests contribute significantly to maintaining ecological balance. They are categorized into different types based on leaf characteristics and phenological traits of tree species, such as needle leaf or broadleaf, evergreen or deciduous. The provision of ecosystem services is intricately linked to forest composition, type, condition, and management practices (Ma et al., 2021), including reserve operations, rotation intervals, and plantation schemes. Therefore, precise monitoring of forests is imperative to effectively manage forest ecosystems for multiple objectives, such as biodiversity conservation, carbon sequestration, and the sustainable supply of timber and non-timber forest products. Enhanced monitoring allows for better assessment and management of these critical ecosystems, ensuring they continue to provide their numerous benefits.

Remote sensing has emerged as a powerful tool for forest mapping across various spatial scales, ranging from global to local levels (Lasota et al., 2019; Li et al., 2014). Recent advancements in satellite sensor technology have facilitated cost-effective and time-efficient opportunities for forest classification. Numerous scholars have conducted research on remote sensing-based forest typing and have developed a range of practical methods based on machine learning techniques (Nink et al., 2019; Yu et al., 2023). The integration of vegetation growth patterns, multi-temporal, and multi-resolution imagery, along with vegetation indices, significantly enhances the accuracy of forest identification. Several open-access forest type products based on medium-resolution remote sensing imagery already exist (Li et al., 2014; Zhang et al., 2021), with spatial resolutions typically ranging from 10 m to 250 m. Moreover, the utilization of high spatial resolution remote sensing images and airborne LiDAR holds promise for specialized regions, further refining forest classification efforts (Li et al., 2021; Lin and Chuang, 2021).

The Shennongjia region is globally renowned for its remarkable biodiversity, rare and endangered flora, and rich biological evolutionary history. This reputation is underscored by its prestigious designations as a "Member of the UNESCO Human and Biosphere Reserve," "World Geopark," and "World Natural Heritage" site. Notably, the region's subtropical forest ecosystem stands as a unique and well-preserved habitat within the middle latitude belt, boasting a gene pool of global significance (Zhang and Li, 2023). Conservation efforts in the area can date back to 1982. While some scholarly endeavors have focused on the forest ecosystem of the region (Dang et al., 2009; Zhao et al., 2005), comprehensive research into detailed and high-resolution forest type classification based on remote sensing data remains limited and outdated. The official establishment of Shennongjia National Park in 2020, alongside the delineation of its boundaries (distinct from the original administrative division of the Shennongjia forest region), underscores the critical need for updated research on forest types and distribution patterns within the park. Such research is essential for guiding spatial allocation and decision-making processes pertaining to forest resource management.

Building upon the groundwork laid by previous studies, this research leverages multi-resolution and multi-temporal remote sensing data to identify forest types and distribution patterns within Shennongjia National Park up to 2023. To fulfill the management needs of the national park administration, this study will analyze forests at the subcompartment scale. Subcompartments are defined and utilized in the Second-class Forest Resources Survey conducted by China. They represent fundamental forest units characterized by consistent internal attributes and distinct differences from neighboring areas. Forests with similar management requirements are grouped together as subcompartments to facilitate efficient forest management and planning. By integrating the capabilities of the Google Earth Engine (GEE) and machine learning models, the study aims to optimize forest type classification workflow and provide foundational data to support national park management and conservation efforts.

2. Study area

Shennongjia National Park, designated as one of the 10 pilot national parks by the National Development and Reform Commission of China, occupies a strategic position in the northwest region of Hubei Province, central China (Figure 1). Spanning longitudinally between 109°42' - 110°40' E and latitudinally between 31°11' - 31°57' N, it encompasses a vast area totaling 2370 km². Positioned within the transitional zone from the plateau in the west to the plain in the east, the park exhibits diverse topography with elevations ranging from 211 to 3105 m. Characterized by a northern subtropical monsoon climate, Shennongjia experiences wet and rainy summers coupled with mild and dry winters, featuring discernible temperature gradients influenced by its mountainous terrain. Precipitation levels increase with elevation due to the orographic effect.

The park stands out for harboring the most representative montane ecosystems and evergreen-deciduous broadleaved mixed forests at this latitude globally. It is the richest area in the world in terms of deciduous woody plants (Zhang and Li, 2023). In the summer and autumn of 2023, our field survey team dedicated two weeks to conduct a comprehensive survey in the national park and established a forest type classification system encompassing needle-leaved forest (NLF), deciduous broadleaved forest (DBF), evergreen broadleaved forest (EBF), mixed broadleaved forest (MBF), mixed leaf forest (MLF), and shrub/grassland (S/G). Subsequently, we collected 1000 samples from the field survey and the Hubei Shennongjia Forest Ecosystem National Field Scientific Observation and Research Station (http://snf.cern.ac.cn/) for validation purposes (Table 1). The distribution of sample points and the survey itinerary are illustrated in Figure 1.

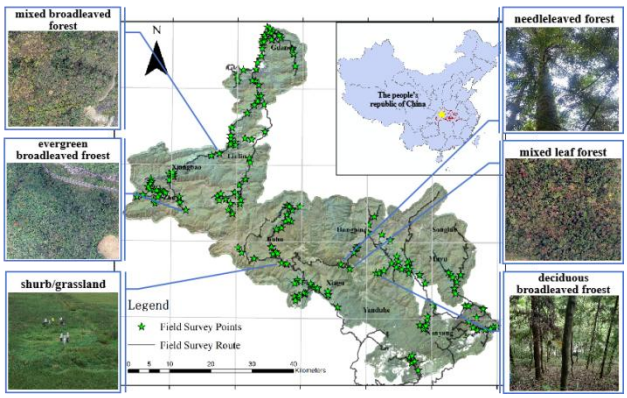

Figure 1. Shennongjia National Park and field survey data

Forest Type		NLF	EBF	DBF	MBF	MLF	S/G
num	station	17	6	19	3	2	0
	survey	51	210	458	48	104	82

Table 1. number of validation points for different forest type

3. Methodology

This study employed a multi-source data approach for forest type classification, utilizing time-series medium resolution Sentinel-2 imagery (10m), high-resolution Planet imagery (3m), and ALOS DEM (12.5m) data. Moreover, a total of 9180 subcompartments were obtained from the Shennongjia National Park Administration within the study region.

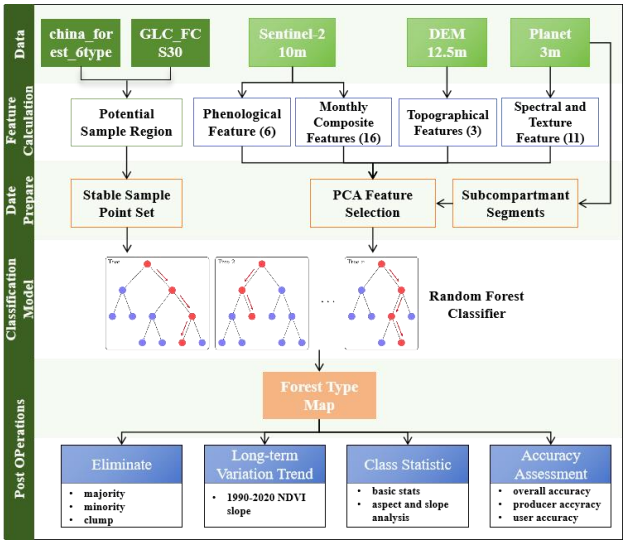

Figure 2. Forest type classification process

3.1 Training Sample Generation

Firstly, we collected multi-source forest vegetation products including china_forest_6type (Li et al., 2014) and GLC_FCS30 (Zhang et al., 2021), and generated a basic forest type sample pool. The china_forest_6type product uses 30m Landsat Thematic Mapper (TM) and 250m MODIS images to produce a 30m forest subclass map of China around 2010. This map includes six forest classes and has an overall accuracy of 72.7%. The GLC_FCS30 product, based on the SpatioTemporal Library (SPECLib), utilizes 30m Landsat data to generate a detailed global land cover map, including five forest classes. Intersection operations were conducted between these products, retaining pixel samples that consistently maintained the same class across multiple products as potential sample areas. Morphological erosion was then applied to the boundaries of potential sample areas.

Subsequently, the standard deviation of the NDVI for a small circular region (with a radius of 7 pixels in this study) centered on each pixel was computed. These standard deviation values were treated as the texture feature of the pixels. Pixels with texture features exceeding a preset threshold were deemed to likely contain mixed land cover types, and thus were masked out and assigned a value of 0. We then exported samples for all the forest types and performed stratified sampling based on geographical distribution to obtain preliminary training samples. This involved dividing the study area into grids, counting the number of pixels of each forest vegetation type within each grid, and determining the number of samples for each type within each grid according to the sampling ratio.

The final training samples were refined and cleaned based on phenological patterns derived from the preliminary training samples. Using field survey samples as constraints and key phenological features derived from time-series imagery, we employed the k-means clustering algorithm to divide the sample set into two clusters. We then calculated the proportion of evergreen and deciduous forests within each cluster. Evergreen forests included evergreen broadleaf and evergreen needleleaf forests, while deciduous forests included deciduous broadleaf and deciduous needleleaf forests. Samples dominated by evergreen forests in the deciduous cluster and samples dominated by deciduous forests in the evergreen cluster were filtered out.

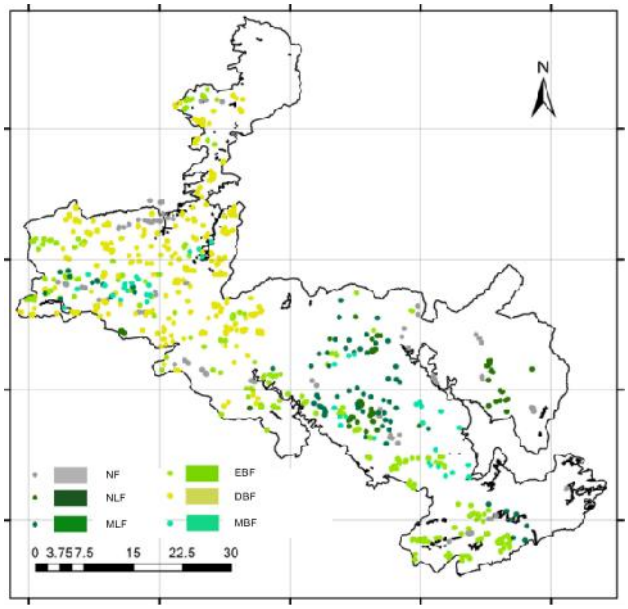

Figure 3. spatial distribution of training samples

3.2 Forest Type Classification

The subsequent step involved calculating the feature set for classification and training the classifier. We constructed a feature stack for classification using medium-resolution time-series imagery and high-resolution imagery from key phenological periods. The classification features encompassed medium-resolution monthly composite spectral features, medium-resolution time-series phenological features, high-resolution spectral features, and topographic features.

Specifically, the medium-resolution monthly composite spectral features include bands B2 to B7, B11, and B12, totaling 10 bands, along with indices NDWI, BSI, NDVI, EVI, RVI, and GCC, totaling 6 indices (Table 2). The medium-resolution time-series phenological features are obtained by calculating the 10th, 20th, 80th, and 90th percentiles of monthly synthesized spectral features, as well as the difference between the 90th and 10th percentiles, and the 80th and 20th percentiles derived from medium-resolution time-series Sentinel-2 imagery in 2023. The high-resolution spectral features are derived from high-resolution remote sensing imagery during key phenological periods (winter 2023 Planet imagery), including bands B2 to B8, totaling 7 bands, and vegetation indices NDVI, EVI, RVI, and GCC, totaling 4 indices. Topographic features, including elevation, slope, and aspect, are extracted from the elevation model data of the ALOS DEM.

The Random Forest algorithm was employed using the above features and samples at the subcompartment scale. The Random Forest algorithm is an ensemble learning method that constructs multiple decision trees during training and outputs the class that is the mode of the classes of the individual trees. The Jeffries-Matusita (JM) distance is used to measure the contribution of each feature to distinguishing sample classes. The JM distance J_{ijk} between classes i and j for feature k can be expressed as:

$$d_{ijk} = \frac{1}{8}(m_{jk} - m_{ik})^2 \frac{2}{v_{ik}^2 + v_{jk}^2} + \frac{1}{2}\ln\frac{\left|\frac{v_{ik} + v_{jk}}{2}\right|}{\sqrt{|v_{ik}||v_{jk}|}}$$

$$J_{ijk} = 2(1 - e^{-d_{ijk}})$$

where i and j represent two different forest vegetation classes, d_{ijk} represents the Bhattacharyya distance between the two classes, m_{ik} and m_{jk} represent the means of feature k for the two classes, and v_{ik} and v_{jk} represent the variances of feature k for the two classes. By adjusting the threshold n_2, features with a JM distance greater than n_2 are selected to train the forest vegetation classification model. This results in a forest vegetation classification Random Forest model and ultimately yields the final subcompartment-level forest type classification results. All steps were conducted on the GEE platform.

Index	Whole Name	Equation
NDVI	Normalized Difference Vegetation Index	(NIR-Red) / (NIR+Red)
EVI	Enhanced Vegetation Index	2.5 ×(NIR -Red) / (NIR+6×Red -7.5×Blue+1)
RVI	Ratio Vegetation Index	NIR /Red
GCC	Green Chromatic Coordinate	Green / (Red+Green+Blue)
NDWI	Normalized Difference Water Index	(Green-NIR) / (Green +NIR)
BSI	Bare Soil Index	((SWIR1+Red) - (NIR+Blue)) / ((SWIR1+Red) + (NIR+Blue))

Table 2. utilized spectral indexes.

4. Results

4.1 Accuracy Assessment

In this study, we use the confusion matrix to assess the classification accuracy of our method. As shown in Table 3, it is an n-row by n-column matrix (where n is the number of classes), with columns representing remote sensing classification data and rows representing the reference ground truth data. The error matrix provides a clear visualization of which classes are subject to misclassification. Specific evaluation metrics include Overall Accuracy (OA), Producer's Accuracy (PA), and User's Accuracy (UA):

$$OA = \frac{TP + TN}{TP + FN + FP + TN}$$

$$PA = \frac{TP}{TP + FN}, UA = \frac{TP}{TP + FP}$$

These accuracy metrics reflect the classification accuracy from different perspectives.

The overall accuracy of the six forest types was recorded at 86.1%. Notably, deciduous broadleaf forest (DBF) exhibited the highest PA of 0.922 and UA of 0.919. This was closely followed by evergreen broadleaf forest (EBF), which demonstrated respective accuracies of 0.819 (PA) and 0.859 (UA). Additionally, needleleaf forest (NLF) achieved an accuracy of approximately 0.8 (0.824 for PA and 0.778 for UA), suggesting satisfactory interpretive performance for these pure stand forest types. However, the accuracies of two mixed forests, mixed leaf forest (MLF) and mixed broadleaf forest (MBF), were relatively lower, with PAs of 0.708 and 0.664 and lower UAs of 0.686 and 0.729, respectively. This diminished accuracy may be attributed to the inconsistency in the definitions of these mixed forest types across different forest products, consequently affecting the quality of the training samples.

		Reference samples						
		NLF	EBF	DBF	MBF	MLF	S/G	UA
predict	NLF	56	10	2	1	2	1	0.778
	EBF	5	177	15	0	9	0	0.859
	DBF	2	15	440	6	14	2	0.919
	MBF	1	7	0	35	4	1	0.729
	MLF	4	5	20	9	75	0	0.664
	S/G	0	2	0	0	2	78	0.951
	PA	0.824	0.819	0.922	0.686	0.708	0.951	

Table 3. accuracy assessment confusion matrix.

4.2 Classification Results

Table 4 illustrates the distribution patterns of various forest types across three altitude levels. Notably, Needle Leaf Forests (NLFs) and Shrub/Grassland (S/G) are predominantly concentrated in high-altitude regions exceeding 2000 meters, with concentrations of 83.5% and 86.7%, respectively. Despite their preference for high altitudes, these forest types occupy relatively limited areas, with NLF covering 24 km^2 and S/G covering 79 km^2. In contrast, Evergreen Broadleaf Forests (EBFs) are primarily found in regions below 2000 meters, encompassing 98.1% of their total area.

Of particular significance is the extensive coverage of Deciduous Broadleaf Forests (DBFs), which are dispersed across a wide range of altitudes and represent the largest forest type within the national park, covering a total area of 1140 km^2. This accounts for nearly half of the park's total area. Mixed Broadleaf Forests (MBFs) are predominantly situated at elevations ranging from 1000 to 2000 meters, while Mixed Leaf Forests (MLFs) are more prevalent in high-elevation regions, exhibiting a distribution pattern similar to Needle Leaf Forests.

Overall, the distribution patterns of various forest types exhibit a natural trend with increasing altitude, indicating that forest types adapt to specific altitude ranges. This distribution aligns with ecological principles, highlighting the influence of altitude on forest composition and structure within the national park. The data underscores the importance of considering altitude in forest management and conservation efforts, as it plays a crucial role in determining the distribution and health of forest ecosystems.

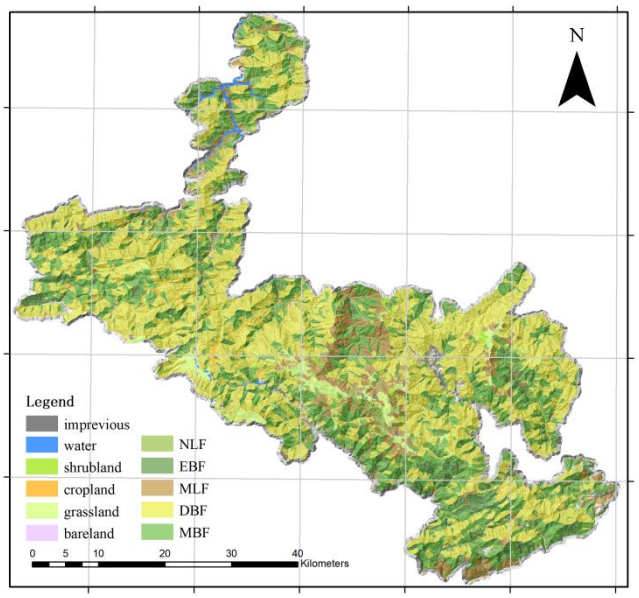

Figure 4. Forest type distribution map in Shennongjia National Park

	L-alti	M-alti	H-alti	L-alt(km^2)	M-alt(km^2)	H-alt(km^2)
NLF	9.5%	6.9%	83.5%	2.308	1.678	20.214
EBF	21.8%	76.3%	1.9%	38.145	133.525	3.296
DBF	2.4%	70.7%	26.9%	27.388	806.143	306.458
MLF	25.1%	20.9%	54.0%	46.096	38.277	98.964
MBF	14.3%	71.7%	14.0%	102.806	514.197	100.458
S/G	0.2%	13.1%	86.7%	0.125	10.291	68.282

Table 4. Forest type distribution at different altitude level.

5. Conclusion

This comprehensive analysis of forest types in Shennongjia National Park provides critical insights into their spatial distribution patterns. By integrating existing datasets and field survey results, we addressed the issue of insufficient sample availability. Leveraging the rich phenological features of medium-resolution imagery and the spatial details of high-resolution imagery, we achieved a detailed and accurate classification of forest types. Our findings indicate that specific forest types, such as Needle Leaf Forests and Shrub/Grassland, predominantly occupy high-altitude regions (>2000 m), while Evergreen Broadleaf Forests are mainly located below 2000 meters. The most extensive coverage is observed in Deciduous Broadleaf Forests, which are distributed across various altitudinal ranges, highlighting their adaptability and significant presence within the park.

The overall accuracy of 86.1% in our classification underscores the robustness of the methodology employed. This high accuracy is crucial for informing management and conservation strategies within Shennongjia National Park. Moreover, the clear altitudinal gradients observed in the distribution of forest types align with ecological principles, emphasizing the role of altitude in shaping forest composition and structure. In conclusion, this study provides a valuable reference for future research and practical applications in forest monitoring, management, and conservation within Shennongjia National Park and similar ecological regions. The integration of remote sensing technologies and machine learning models proves to be an effective approach for detailed forest type classification, contributing to the broader field of ecological research and resource management.

Acknowledgements

This work was funded by the Project of Background Resources Survey in Shennongjia National Park (SNJNP2022001) and the Open Project Fund of Hubei Provincial Key Laboratory for Conservation Biology of Shennongjia Snub-nosed Monkeys (SNJGKL2022001).

References

Dang, H.S., Zhang, K.R., Zhang, Y.J., Tan, S.D., Jiang, M.X., & Zhang, Q.F. 2009. Tree-line dynamics in relation to climate variability in the Shennongjia Mountains, central China. *Canadian Journal of Forest Research, 39*, 1848-1858.

Lasota, J., Blonska, E., Lyszczarz, S., & Sadowy, A. 2019. Forest habitats and forest types on chernozems in south-eastern Poland. *Soil Science Annual, 70*, 234-243.

Li, C.C., Wang, J., Hu, L.Y., Yu, L., Clinton, N., Huang, H.B., Yang, J., & Gong, P. 2014. A Circa 2010 Thirty Meter Resolution Forest Map for China. *Remote Sensing, 6*, 5325-5343.

Li, L.H., Jing, W.P., & Wang, H.H. 2021. Extracting the Forest Type From Remote Sensing Images by Random Forest. *Ieee Sensors Journal, 21*, 17447-17454.

Lin, F.C., Chuang, Y.C. 2021. Interoperability study of data preprocessing for deep learning and high-resolution aerial photographs for forest and vegetation type identification. *Remote Sensing, 13.*

Ma, S., Qiao, Y.P., Wang, L.J., & Zhang, J.C. 2021. Terrain gradient variations in ecosystem services of different vegetation types in mountainous regions: Vegetation resource conservation and sustainable development. *Forest Ecology and Management, 482.*

Nink, S., Hill, J., Stoffels, J., Buddenbaum, H., Frantz, D., & Langshausen, J. 2019. Using Landsat and Sentinel-2 Data for the Generation of Continuously Updated Forest Type Information Layers in a Cross-Border Region. *Remote Sensing, 11.*

Yu, Z.Y., Wang, J.N., Yang, X.K., & Ma, J. 2023. Superpixel-Based Style Transfer Method for Single-Temporal Remote Sensing Image Identification in Forest Type Groups. *Remote Sensing, 15.*

Zhang, B., & Li, L. 2023. Evaluation of ecosystem service value and vulnerability analysis of China national nature reserves: A case study of Shennongjia Forest Region. *Ecological Indicators, 149.*

Zhang, X., Liu, L.Y., Chen, X.D., Gao, Y., Xie, S., & Mi, J. 2021. GLC_FCS30: global land-cover product with fine classification system at 30m using time-series Landsat imagery. *Earth System Science Data, 13*, 2753-2776.

Zhao, C.M., Chen, W.L., Tian, Z.Q., & Xie, Z.Q. 2005. Altitudinal pattern of plant species diversity in Shennongjia Mountains, central China. *Journal of Integrative Plant Biology, 47*, 1431-1449.