Establishment of a Quantitative Evaluation Method of Rice Plant Type Using P-type Fourier Descriptors

Katsuaki Suzuki\(^1\), Zeyu Zheng\(^2\), Yoshiyasu Tamura\(^3\) and Yutaka Hirata\(^1\)

\(^1\)Laboratory of Plant Genetics and Biotechnology, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan; \(^2\)Department of Environmental Information, Tokyo University of Information Sciences, Onaridai, Wakaba-ku, Chiba 265-8501, Japan; \(^3\)Department of Data Science, The Institute of Statistical Mathematics, Tachikawa, Tokyo 190-8562, Japan

Abstract: To evaluate rice plant type precisely at the seedling stage, we established a quantitative evaluation method using an image analysis. P-type Fourier descriptors, which could apply an “open-curve”, were used for the plant type, and the coefficients were summarized as the scores of principal components (PCs) by principal component analysis. At the same time, “conventional plant type traits”, leaf blade length, leaf blade angle and inter-leaf-blade length, were measured by traditional plant type measurement methods. Based on the PC scores, the plant type for each PC was reconstructed by inverse Fourier transformation and the morphological characteristics were evaluated. To examine whether our method could appropriately identify the characteristics of the varieties, we discriminated the varieties using the PC scores by support vector machines. The varieties were also discriminated using the conventional traits of plant type. The results indicated that both traits had equal discrimination efficiency. In addition, the combination of conventional traits and scores had the highest discrimination efficiency. The relationship between the PC scores and conventional traits was examined by multiple regression analysis. The PC scores were not correlated with size-related traits. From these results, our research clarified that a plant type evaluation method using P-type Fourier descriptors could evaluate rice plant type precisely combined with conventional methods.

Key words: Image analysis, Plant type, P-type Fourier descriptors, Quantitative evaluation, Rice.

Attempts have been made to improve the yield by genetically modifying the “plant type” of crops with the corresponding improvement of cultivation methods. In Japanese plant-type breeding, two main definitions of rice plant type have been developed (Tsunoda, 1987). The first, called the “old” plant-type concept by Tsunoda, classifies the plant type as “long-culm and low-tiller type”, “short-culm and high-tiller type”, “panicle weight type”, or “panicle number type” by measuring the plant height, tiller number and panicle number in the mature stage. It is still used for plant type evaluation today because it is simple and relatively effective. However, the perspective of the plant type has limited effectiveness for grasping the correspondence between the morphology–structure and the function of plant aerial parts. The second, called the “new” plant-type concept by Tsunoda (1987) defined the plant type as the morphology and structure of the plant aerial parts involved in the life and the production. Based on this concept, the relationship between the shapes of the plant aerial parts, especially the growth characteristics of leaves, and the photosynthesis of the whole leaf canopy was thought to be important.

Historically, rice plant type developed to the erect type from the prostrate type in the domestication process (Tan et al., 2008). Through green evolution (Khush, 1999), high yielding varieties bred using semi-dwarf gene \(sd1\) (Ashikari et al., 2002; Sasaki et al., 2002) improved the light-receiving characteristics under intensive cultivation and these varieties are cultivated widely. However, on a global scale, there are many areas where sufficient fertilizer and water cannot be applied; therefore, high yielding varieties which have a plant type suited to the conditions of each area need to be developed by breeding (Tsunoda, 1959).

There are several methods for measuring and evaluating plant type using conventional measurements of plant height, length of leaf blade and sheath and leaf angle (Morishima et al., 1967), the stratified clip method (Monsi and Saeki, 1953), the silhouette method (Udagawa et al., 1968), image analysis (Oka and Hinata 1988, 1989), and the 3D digitizer method (Oka and Ogawa, 2005); however, these methods evaluate only part of the plant, the dissected plant body, and are difficult to apply to many populations.
Therefore, it is necessary to establish a quantitative, objective and precise evaluation method using the shape of the plant that is nondestructive and can be applied to many populations.

The image analytical method using elliptic Fourier descriptors and principal component analysis has been used as one of the most improved quantitative evaluation methods for measuring plant morphology (Iwata et al., 1998; Yoshioka et al., 2004). Elliptic Fourier descriptors can only be applied to the analysis of a “closed curve” with the same starting point and endpoint of measurement. “P-type Fourier descriptors”, one type of Fourier descriptors, can be applied to the analysis of an “open curve” with a different starting point and endpoint of measurement (Uesaka, 1984). With the P-type Fourier descriptors used in this study, the target curves for analysis are regarded as polygonal lines consisting of equal length segments, and P-type Fourier descriptors are calculated by Fourier transformation from the angles between the segments. As examples using plant shape, there have been evaluation studies on petal tips in the sacred lotus (Zheng et al., 2005; Zheng and Tamura, 2005) and leaf bending in rice (Zheng et al., 2008).

Based on the new plant-type concept, we aim to predict the plant type from the seedling stage to the mature stage and finally to evaluate yield components from the plant type. From these studies, our method is expected to be used for juvenile selection in breeding. In this study, we established a method of identifying rice plant type at the seedling stage that can quantitatively evaluate and analyze the target shapes using P-type Fourier descriptors with principal component analysis, demonstrating the efficiency of this method for objective and precise evaluation of plant type.

Materials and Methods

1. Plant materials

Two elite indica-type varieties, “IR64” and “MTL250”, an indica-type “Kasalath” utilized in many researches and two elite japonica-type varieties, “Nipponbare” and “Koshihikari”, were used. Forty plants per variety were sown in 7.7 cm diameter pots. Plants were grown in a greenhouse. These pots were located close together and arranged in random order. Twenty plants per variety (one hundred plants in total) were selected and used in this study.

2. Image processing

Images were taken of rice in the fifth leaf stage at the furthestmost positions in a horizontal direction between the culm and each tip of the second, third and fourth leaf blades (Fig. 1). In the second leaf, the contours were extracted from the second leaf blade and the second inter-leaf blade that composed the culm part between the first and second collars. Similarly, the contours of the third and fourth leaves were extracted. The second leaf blade and second inter-leaf blade, third leaf blade and third inter-leaf blade, and fourth leaf blade, and fourth inter-leaf blade were called the second leaf, third leaf and fourth leaf, respectively.
3. Methods

The 64 coefficients of P-type Fourier descriptors were calculated from the plant-type contours in each leaf using 16 harmonics. To summarize the shape information in the coefficients, we performed principal component analysis based on the variance-covariance matrix using the coefficients of these one hundred plants and contribution ratios, and principal component scores were obtained. The calculated principal component scores were applied as the characteristics of the plant-type contours. At the same time, conventional plant type traits were measured using traditional plant-type measurement methods and compared with the method using P-type Fourier descriptors. On the defined second leaf, the second inter-leaf blade length (L2ILL), which is the culm length between the first and second collars, the second leaf blade length (L2LBL) and the second leaf blade angle (L2LBA) which was angle between culm and leaf blade basal point were measured. Similarly, on the third and fourth leaves, L3ILL, L3LBL, L3LBA, L4ILL, L4LBL and L4LBA were measured, respectively.

To detect plant-type variations among varieties, we performed the analysis of variance (ANOVA) using the principal component scores and measured values of the conventional plant type. The plant-type contours of the principal components (PCs) were reconstructed by inverse Fourier transformation to elucidate visually the plant-type traits evaluated by each PC. The reconstructed solid, broken and dotted lines showed that the PC scores were 0 (mean), +1 and -1 times of standard deviation, respectively. From the reconstructed lines, we evaluated the morphological characteristics of each PC.

The varieties were discriminated using a support vector machine (SVM) (Burges 1998; Maeda 2001), which is one of the most powerful learning models for pattern recognition ability to examine whether our method using P-type Fourier descriptors could appropriately grasp the characteristics of the varieties. Misclassification rates were calculated by leave-one-out cross-validation to compare the discrimination efficiencies derived from SVM, and the efficiency of our method was examined on plant-type evaluations. In the leave-one-out cross validation, one plant was omitted from 100 plants and computation of principal component analysis and SVM model building were conducted using the coefficients of P-type Fourier descriptors for the remaining 99 plants. Then, these PCA coefficients and the SVM classifier were used for calculation of the PC scores and discrimination of variety based on the PCA scores, respectively, for the omitted plant. This process was repeated until each observation had been left out and classified exactly once.

We investigated whether the PCs in each leaf corresponded to the traits of the conventional plant type by multiple regression analyses with each ILL, LBL and LBA in each leaf as objective variables, and the PCs in the same leaf position as explanatory variables. To choose the most appropriate multiple regression model, we calculated Akaike’s information criteria (AIC) and selected the multiple regression model with the smallest AIC.

The method of Zheng was used to calculate the P-type Fourier descriptors and inverse Fourier transformation (Zheng et al., 2005; Zheng and Tamura, 2005; Zheng et al., 2008). For principal component analysis, ANOVA, inverse Fourier transformation, SVM, multiple regression analysis and AIC, were performed using statistical package “R” (ver. 2.41). For the calculation using SVM, the linear kernel SVM function included in the e1071 package was applied.

### Results and Discussion

We evaluated rice plant type using the P-type Fourier descriptor method. PC analysis was performed from the coefficients of P-type Fourier descriptors and the contribution ratios of each PC and the PC scores were obtained (Table 1). In the second leaf, the cumulative

---

### Table 1. The contribution rate of each principal component and the result of ANOVA by principal component score in three leaves.

| Component | Second Leaf | Third Leaf | Fourth Leaf |
|-----------|-------------|------------|-------------|
|           | Contr (%)   | F value    | Contr (%)   | F value    | Contr (%)   | F value |
| PC1       | 35.1        | 4.49 **    | 73.8        | 3.37 **    | 74.9        | 26.46 **  |
| PC2       | 17.8        | 19.36 **   | 7.8         | 4.14 **    | 10.8        | 24.32 **  |
| PC3       | 13.7        | 15.29 **   | 7.3         | 2.94       | 6.6         | 1.43      |
| PC4       | 7.1         | 1.57       | 2.8         | 26.74 **   | 1.7         | 0.22      |
| PC5       | 5.8         | 7.28 **    | 1.7         | 4.16 **    | 1.5         | 0.71      |
| ILL       | 20.07 **    | 68.61 **   | 3.40        | 22.08 **   |
| LBL       | 17.78 **    |            | 4.14 **     |
| LBA       | 22.69 **    | 16.93 **   |            |

**, * P <0.01, P <0.05, PC: Principal component, Contr: Contribution rate, ILL: Inter-leaf-blade length, LBL: Leaf blade length, LBA: Leaf blade angle.
contribution ratio of the first five components was 79.5%. In the third and fourth leaves, the cumulative contribution ratios of the first five components were 93.5% and 95.5%, respectively. All three leaves could be evaluated for most plant-type variations using the first five components. In subsequent analysis, we used the first five components of each leaf.

ANOVA among varieties was calculated from the PCs and the conventional plant-type traits (Table 1). In the second leaf, PCs that significantly differed with the variety were found in L2PC1 (the first component of the second leaf), L2PC2, L2PC3 and L2PC5. Significantly different PCs were found in L3PC1, L3PC2, L3PC4 and L3PC5 in the third leaf, and in L4PC1 and L4PC2 in the fourth leaf. For conventional plant-type traits, all ILLs, LBLs and LBAs in the three leaves had less than 1% significance except for L3LBL, which had 5% significance.

The contours estimated by each PC were reconstructed by inverse Fourier transformation and the morphological characteristics in each PC were evaluated (Fig. 2). The first components of all three leaves, for example, seemed to evaluate the degree of leaf bending. L2PC2 seemed to evaluate the relationship between LBA and the ratio of ILL to LBL, and both L3PC2 and L4PC2 seemed to evaluate the angle of leaf blade tip. The other components evaluated the other plant-type traits indicated in Fig. 2.

For conventional plant-type traits, all ILLs, LBLs and LBAs in the three leaves had less than 1% significance except for L3LBL, which had 5% significance.

Table 2. Varietal discrimination of principal components and plant type traits by SVM.

| plant type traits | misclassification |
|-------------------|-------------------|
| Group 1\(^a\)     | 14%               |
| Group 2\(^b\)     | 24%               |
| Group 3\(^c\)     | 17%               |
| Group 1 + Group 3 | 5%                |

\(a\): Group 1 includes the conventional plant type traits. \(b\): Group 2 includes PCs that were significant by ANOVA from PC1 to PC5 in the three leaves. \(c\): Group 3 included the first five PCs in each leaf.

The contours estimated by each PC were reconstructed by inverse Fourier transformation and the morphological characteristics in each PC were evaluated (Fig. 2). The first components of all three leaves, for example, seemed to evaluate the degree of leaf bending. L2PC2 seemed to evaluate the relationship between LBA and the ratio of ILL to LBL, and both L3PC2 and L4PC2 seemed to evaluate the angle of leaf blade tip. The other components evaluated the other plant-type traits indicated in Fig. 2.

The varieties were discriminated by SVM using the PC scores and conventional plant-type traits (Table 2). Compared with the misclassification rates of the traits in the two evaluation methods, the method with a lower rate was better able to grasp the plant-type characteristics of the varieties, i.e., the plant-type evaluation method was more efficient. The misclassification rate of Group 1 using all nine conventional plant-type traits was 14%. In Group 2, using ten PCs that were significant by ANOVA, the misclassification rate was 24%. In Group 3 using fifteen PCs, the misclassification rate was 17%. The discrimination efficiency was lower in Group 2 than in Group 1 or Group 3. The first five PCs (Group 3) could adequately identify the characteristics of the varieties. The first three PCs, however, could not adequately identify the characteristics (data not shown). Group 3 had discrimination efficiency equal to the conventional plant-type traits (Group 1). In plant-type evaluation using P-type Fourier descriptors, it was necessary to use at least the first five PCs. In addition, the combination of the conventional traits (Group 1) and...
the first five PCs (Group 3) had the highest discrimination efficiency; therefore, it was thought to be necessary to combine the P-type Fourier descriptor method and conventional measurements to evaluate the plant type most precisely.

To examine the correspondent relationship between the PCs and the conventional plant-type traits, we performed multiple regression analysis using each conventional plant-type trait in each leaf as the objective variable, and the PCs in the same leaf as explanatory variables (Table 3). It was thought that the closer the score of R² (coefficient of determination adjusted for the degrees of freedom) was to 1.0, the more the PCs corresponded with the conventional plant-type traits, i.e., PCs evaluated conventional plant-type traits. In the second leaf, ILL and LBA had higher R² and corresponded about 60% with the PCs. ILL and LBA of the third leaf had more than 60% correspondence with the PCs. The other conventional traits, however, could not be evaluated with the PCs. This indicated adversely that the PCs evaluated plant-type traits which the conventional traits could not. The traits related to length (ILL and LBL) were difficult to evaluate because P-type Fourier descriptors do not include size information (Zheng et al., 2008). Therefore, it was clear that populations with an extremely similar genetic background could be evaluated by our plant-type evaluation method. This indicated the possibility of applying QTL analysis to segregation generations using the P-type Fourier descriptor method. So far, there have been examples of QTL analyses of such as lemma and palea of rice (Uga et al., 2003), soybean leaflet (Yamanaka et al., 2001) and Drosophila posterior lobe (Zeng et al., 2000) using elliptic Fourier descriptors but not P-type Fourier descriptors. To perform a detailed QTL analysis, we need to evaluate the plant type of many lines without environmental variations. Since the P-type Fourier descriptors could detect the traits (PCs) that were easily affected by the environment, our method could be useful for QTL analysis of rice plant type.

In this study, we established a method of evaluating the plant type of rice seedlings (fifth leaf stage) using P-type Fourier descriptors. When the rice plant body is divided into the main stem and each tiller, the characteristics are not markedly different from those of the body in the seedling stage. Therefore, we can evaluate the plant type of the main culm and tillers as well as in the seedling stage using P-type Fourier descriptors. In particular, the upper three leaves of the main culm in the mature stage are expected to be related to the plant type at the seedling stage. In future studies we will analyze the plant type at the seedling and mature stages to determine the correlation between plant type and yield components.

Acknowledgements

We would like to thank Dr. H. Iwata and Dr. S. Ninomiya of the University of Tokyo for their valuable advice.

References

Ashikari, M., Sasaki, A., Ueguchi-Tanaka, M., Inoh, H., Nishimura, A., Datta, S., Ishiyama, K., Saiito, T., Kohayashi, M., Khush, G. S., Kitano, H. and Matsuoka, M. 2002. Loss-of-function of a rice gibberellin biosynthetic gene, GA20 oxidase (GA20ox-2), led to the
Monsi, M. and Saeki, T. 1953. Über den lichtfaktor in den pflanzengesellschaften und seine bedeutung für die stoffproduktion. *Euphytica*. 102: 143-149.

Khush, G.S. 1999. Green revolution: preparing for the 21st century.

Iwata, H., Niikura, S., Matsuura, S., Takano, Y. and Ukai, Y. 1998. Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*. 102: 676-683*.

Morishima, H., Oka, H. and Chang, T. T. 1967. Analysis of genetic variations in plant type of rice: 1. Estimation of indices showing genetic plant type and their correlations with yielding capacity in a segregating population. *Jpn. J. Breed*. 17: 73-84.

Oka, M. and Hinata, K. 1989. Comparison of plant type between new and old rice cultivars using computer image analysis for characterization of plant type in rice cultivars. *Jpn. J. Breed*. 38: 449-458.

Oka, M. and Hinata, K. 1989. Comparison of plant type between new and old rice cultivars using computer image analysis. *Jpn. J. Crop Sci*. 58: 232-239.

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapana, D., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G. S., Kitano, H. and Matsuoka, M. 2002. Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature*. 416: 701-2.

Suzuki, K., Sarhadi, W.A., Zheng, Z. and Hirata, Y. 2008. A comparison between a new P-type Fourier descriptors analysis and a conventional method to evaluate rice plant type. Proc. World Conference on Agric. Info. and IT [CD-ROM]. Kanagawa, 1041-1046.

Tan, L., Li, X., Liu, F., Sun, X., Li, C., Zhu, Z., Fu, Y., Cai, H., Wang, X., Xie, D. and Sun, C. 2008. Control of a key transition from prostrate to erect growth in rice domestication. *Nat Genet*. 40: 1360-1364.

Tsunoda, S. 1959. A developmental analysis of yielding ability in varieties of field crops: II. The assimilation-system of plants as affected by the form, direction and arrangement of single leaves. *Jpn. J. Breed*. 9: 237-244.

Tsunoda, S. 1987. History and perspectives of plant type investigation. *Rec. Prog. Breed*. 28: 29-39*.

Udagawa, T., Uchijima, Z., Horie, T. and Kobayashi, K. 1968. Studies of energy and gas exchange within crop canopies (3): Canopy structure of corn plants. *Proc. Crop Sci. Soc. Jpn.*, 37: 589-596**.

Usaka, Y. 1984. A new type Fourier descriptor method that is effective also to open contour. *IEICE Trans. Fundamentals*. 67: 166-173*.

Uga, Y., Fukuta, Y., Cai, H. W., Iwata, H., Ohsawa, R., Morishima, H. and Fujimura, T. 2003. Mapping QTLs influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *O. rufipogon* Griff. *Theor. Appl. Genet*. 107: 218-226.

Yamanaka, N., Ninomiya, S., Hoshi, M., Tsubokura, Y., Yano, M., Nagamura, Y., Sasaki, T. and Harada, K. 2001. An informative linkage map of soybean reveals QTLs for flowering time, leaflet morphology and regions of segregation distortion. *DNA Res*. 8: 61-72.

Yoshioka, Y., Iwata, H., Ohsawa, R. and Ninomiya, S. 2004. Analysis of petal shape variation of *Primula sieboldii* E. Morren by elliptic Fourier descriptors and principal component analysis. *Ann. Bot*. 94: 657-664.

Zeng, Z.B., Liu, J., Stam, L.F., Kao, C. H., Mercer, J.M. and Laurie, C. C. 2000. Genetic architecture of a morphological shape difference between two *Drosophila* species. *Genetics*. 154: 299-310.

Zheng, Z., Iwata, H., Ninomiya, S. and Tamura, Y. 2005. Quantitative evaluation of partial shape characteristics of petals of the sacred lotus based on P-type Fourier descriptors. *Breed. Res. 7*: 133-142**.

Zheng, Z. and Tamura, Y. 2005. Cultivar identification of lotus (*Nelumbo nucifera* Gaertn) by P-type fourier descriptors with petal tip contour. *Horticul. Res*. 4: 385-390**.

Zheng, Z., Iwata, H., Hirata, Y. and Tamura, Y. 2008. Quantitative evaluation of the degree of sprout leaf bending of rice cultivars using P-type Fourier descriptors and principal component analysis. *Euphytica*. 163: 259-266.

* In Japanese.

** In Japanese with English abstract.