

PHÂN TÍCH BIẾN DỊ DI TRUYỀN Ở DẠNG VÀ SỐ TRÁI ĐẬU NÀNH BẰNG PHƯƠNG PHÁP ĐƯỜNG VIÊN

Trương Trọng Ngôn¹, Trần Thị Thanh Thủy², Nguyễn Văn Chương³

¹Viện Nghiên cứu & Phát triển Công nghệ Sinh học

²Khoa Nông nghiệp & SHƯĐ

³Viện Khoa học Kỹ thuật Nông nghiệp miền Nam

TÓM TẮT

Số trái và dạng trái đậu nành là hai yếu tố quan trọng có ảnh hưởng đến năng suất sau cùng của hạt. Phương pháp phân tích “Fourier Elliptic” được sử dụng trong nghiên cứu này. Tám mươi giống đậu nành có nguồn gốc khác nhau được gieo từ tháng 1 đến tháng 4/2010 tại tại lô đất thuộc Khu vực V, phường Hưng Thạnh, quận Cái Răng, thành phố Cần Thơ. Bố trí thí nghiệm theo thể thức hoàn toàn ngẫu nhiên, với 3 lần lặp lại. Mỗi giống được gieo trên 2 hàng, mỗi hàng dài 5 m, khoảng cách gieo là 40 x10 cm, mỗi hốc 3 hạt, sau tỉa chừa 2 cây/hốc. Chọn 5 cây ngẫu nhiên để đo các đặc tính nông học như tổng số trái, số trái mang 1, 2 và 3 hạt... 12 trái chín hoàn toàn của mỗi giống được chọn ngẫu nhiên để chụp ảnh bằng máy kỹ thuật số. Các ảnh được phân tích dạng hình bằng phần mềm SHAPE (Iwata, 2002). Kết quả cho thấy nhóm giống Việt Nam có tổng số trái cao nhất (34 trái) so với các nhóm giống khác. Nhưng nhóm giống Trung Quốc có số trái mang 3 hạt nhiều hơn (6 trái). Sự phân tích ở năm thành phần chính giải thích hơn 95% sự biến dị ở dạng trái một hạt, 85% đối với trái 2 hạt, và trên 82% ở trái 3 hạt.

Từ khóa: *Elliptic Fourier, dạng trái, đường viên, phân tích thành phần chính.*

1. ĐẶT VẤN ĐỀ

Số trái và dạng trái đậu là hai yếu tố quan trọng ảnh hưởng đến năng suất hạt ở đậu nành. Phân tích hình dạng của vật thể là một trong các phương pháp quan trọng trong phân tích dạng hình ở các cơ quan thực vật nói chung và ở dạng trái đậu nành nói riêng. Có nhiều phương pháp để phân tích dạng hình như đo chiều dài, đo đường viên (contour) hoặc dùng phương pháp “Fourier elliptics” v.v... trong đó phương pháp Fourier được xem là phổ biến và hiệu quả (Bookstein *et al.*, 1982; Diaz *et al.*, 1989; Ferson *et al.*, 1985), vì nó trực tiếp phân tích trên chính dạng hình của các vật thể hoặc cơ quan nghiên cứu.

Đặc điểm của phương pháp đường viên Fourier được xây dựng bởi các phương trình, các dạng đường viên tiêu biểu với các thông số Fourier. Các thông số Fourier là một trong số các thông số tượng trưng cho các dạng đường viên (Giardina and Kuhl, 1977; Granlund, 1972; Kuhl and Giardina, 1982). Phương pháp Fourier được sử dụng trong phân tích dạng hình bằng cách dùng hai trục X và Y trên đó đường viên sẽ di chuyển chung quanh vật thể nghiên cứu (như dạng trái, dạng lá, dạng thân v.v...) và xoay quanh trên hai trục này ở những góc cạnh khác nhau (Rohlf and Archie, 1984; Ferson *et al.*, 1985; Bierbaum and Ferson, 1986; White *et al.*, 1988; Diaz *et al.*, 1989).

Việc phân tích dạng hình qua ảnh chụp kỹ thuật số đã được áp dụng cho sự phân biệt dạng hạt của các giống (Keefe & Draper, 1986; Neuman *et al.*, 1987; Myers & Edsall, 1989). Gần đây, Furuta và *csv.* (1995) đã thành công trong việc đánh giá dạng lá

chết ở đậu nành bằng các điểm thành phần chính dựa trên thông số Fourier elliptic. Tiếp theo các nghiên cứu này, nhiều nghiên cứu khác được áp dụng trên việc phân tích dạng hạt kiều mạch (buckwheat), Ninomiya và *csv.* (1995), Yoshida và *csv.* (1995) cho thấy phương pháp này rất hiệu quả trong việc xếp nhóm các giống.

Thí nghiệm “Phân tích sự đa dạng di truyền của dạng trái đậu nành dựa trên phương pháp đường viền Fourier” được thực hiện nhằm đánh giá dạng trái của các giống đậu nành có nguồn gốc khác nhau, đồng thời sẽ bổ sung thêm đặc tính này trong việc mô tả giống khi xây dựng bảng tổng kê tập đoàn.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP

2.1 Vật liệu

Tám mươi giống đậu nành có nguồn gốc từ Trung Quốc, Nhật Bản, Hàn Quốc, và Việt Nam được dùng để khảo sát các đặc tính nông học cũng như phân tích dạng trái đậu nành. Số lượng và nguồn gốc của các giống được trình bày ở Bảng 1.

Bảng 1. Số lượng và nguồn gốc của các giống dùng trong thí nghiệm

Nhóm giống	Số lượng	Nguồn gốc	Vĩ tuyến	Kinh tuyến
Trung Quốc*	20	Heilongjiang; Jilin; Liaoning	39°-54°00'N	130°00'-138°00'E
Nhật Bản*	20	Hokkaido; Honshu; Kyushu	30°-46°00'N	130°00'-144°00'E
Hàn Quốc*	20	Gyeonggi; Gangwon; Chungcheong; Cheolla	33°-39°00'N	124°11'-131°52'E
Việt Nam**	20	Miền Bắc, Trung, Nam	8°30'-23°25'N	102°10'-109°30'E

Ghi chú:

* *Nguồn: Ngân hàng gen (Genebank), Bộ Nông Nghiệp, Hàn Quốc.*

** *Nguồn: Viện Nghiên cứu & Phát triển Công Nghệ Sinh học, Đại học Cần Thơ.*

Các giống được gieo từ tháng 1 đến tháng 4 năm 2010 tại phường Hưng Thạnh, quận Cái Răng, Thành phố Cần Thơ. Mỗi giống được gieo trên hàng dài 5 m, với 3 lần lặp lại. Khoảng cách 40 x 10 cm, 2 cây/hốc.

Số trái và các chỉ tiêu nông học được đo ngẫu nhiên trên năm cây. Để chụp hình dạng trái, mỗi giống chọn ngẫu nhiên 12 trái chín hoàn toàn bao gồm các loại trái như trái mang 1 hạt, trái mang 2 hạt và trái mang 3 hạt.

2.2 Phương pháp phân tích dạng trái

Tất cả từng loại trái đều được chụp ảnh bằng máy kỹ thuật số hiệu Nikon Coolpix 5700 (Nikon Japan Incorporation). Bước kế tiếp là chuyển trực tiếp các ảnh vào máy tính để lưu và phân tích. Khoảng cách từ máy chụp ảnh đến trái đậu nành được cố định là 45cm sao cho ảnh được rõ nét và không có bóng viền xung quanh trái. Độ phân giải của các ảnh chụp cố định và kích thước các ảnh là 640 x 480 pixel. Phần mềm thông dụng Photo Impact Bundle 5.0 được dùng để xử lý ảnh như chỉnh độ tương phản, biến đổi màu

nền để nhận được đường viền rõ nhất. Phần mềm chuyên dụng Shape analysis (Iwata, 2002) được dùng để phân tích dạng trái.

2.3 Phương pháp đường viền (Fourier elliptic)

Có 80 giống đậu nành được dùng để phân tích dạng hình dựa trên phương pháp Fourier elliptic. Các hệ số được tính bằng phép biến đổi riêng biệt các đường viền thông qua chuỗi mã hoá dựa trên phương pháp đề nghị của Kuhl and Giardina (1982). Phương pháp này bao gồm các bước chính như sau: đường viền của dạng hình mã số có thể được tượng trưng bởi các chuỗi xoay quanh trục X và Y từ điểm bắt đầu được chọn ngẫu nhiên và di chuyển theo chiều kim đồng hồ. Giả định rằng đường viền giữa hai điểm kề nhau là đường thẳng nội suy, chiều dài giữa hai điểm $(i - 1)^{th}$ và i^{th} là Δt_i , chiều dài đường viền từ điểm bắt đầu tới điểm thứ p là $\Delta t_p = \sum_{i=1}^p \Delta t_i$, chu vi của các đường viền là $T = t_k$, trong đó K là tổng các điểm trên đường viền. Điểm thứ K^{th} tương đương với điểm bắt đầu. Trục X của điểm thứ p^{th} là $x_p = \sum_{i=1}^p \Delta x_i$, trong đó Δx_i là sự dịch chuyển dọc theo trục X của đường viền giữa điểm thứ $(i - 1)^{th}$ và điểm thứ i^{th} . Sự kéo dài của các chuỗi dọc trên trục X được tính bởi phương trình như sau:

$$x_p = x_{cen} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2n\pi t_p}{T} + b_n \sin \frac{2n\pi t_p}{T} \right)$$

trong đó,

$$a_n = \frac{T}{2n^2 \pi^2} \sum_{p=1}^K \frac{\Delta x_p}{\Delta t_p} \left(\cos \frac{2n\pi t_p}{T} - \cos \frac{2n\pi t_{p-1}}{T} \right)$$

và

$$b_n = \frac{T}{2n^2 \pi^2} \sum_{p=1}^K \frac{\Delta x_p}{\Delta t_p} \left(\sin \frac{2n\pi t_p}{T} - \sin \frac{2n\pi t_{p-1}}{T} \right)$$

Trong các phương trình kể trên, X_{cen} được xem là trục điểm ở vị trí trung tâm, và n là các hệ số trong hàm điều hoà (a_n and b_n). Các giá trị hệ số trên trục Y, là c_n và d_n được tìm thấy như cách tính trên trục X.

Các hệ số của phương pháp Fourier elliptic về mặt toán học có thể được chuẩn hoá để không bị biến động về kích cỡ, điểm xoay vòng và điểm bắt đầu của đường viền. Đối với phần mềm SHAPE, các hệ số có thể được chuẩn hoá bằng cách sử dụng hai phương pháp: (i) dựa trên đường cong (ellipse) của chiều dài điều hoà đầu tiên, (ii) dựa trên bán kính dài nhất (Kuhl và Giardina, 1982).

2.4 Phân tích dữ liệu

Phần mềm Excel 2003 được dùng để phân tích các đặc số thống kê của các tính trạng nông học. Phần mềm SHAPE (Iwata, 2002) được dùng để phân tích thành phần chính của dạng trái, và vẽ hình các loại dạng trái đậu nành.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1 Kết quả

Đặc tính nông học

Màu vỏ trái, màu vỏ hạt, và màu tể thường là những tính trạng do đơn gen kiểm soát. Sự đa dạng của những tính trạng này thường biểu hiện như tính trạng chất lượng, nên ít bị ảnh hưởng bởi yếu tố môi trường. Tất cả các giống khảo sát cho thấy đa dạng về dạng trái, màu vỏ hạt.

Trái đậu nành thường chứa từ 1 đến 5 hạt, nhưng phổ biến nhất là dạng trái mang 2 và mang 3 hạt. Tuy nhiên, có một số ít giống có cả trái mang 4 hạt và trái mang 5 hạt.

Số trái là đặc điểm do di truyền và thay đổi từ giống này sang giống khác. Nhóm giống Việt Nam cho tổng số trái cao nhất (34 trái/cây) khi so sánh với các nhóm giống khác (Bảng 2). Tuy nhiên, nhóm giống Hàn Quốc có số trái mang 3 hạt nhiều nhất (6 trái) so với các nhóm còn lại (Bảng 2).

Bảng 2. Trung bình tổng số trái, số trái mang 1 và 2 hạt với số trái mang 3 hạt của các nhóm giống đậu nành

Nhóm giống	Tổng số trái trung bình	Số trái mang 1 hạt và mang 2 hạt trung bình	Số trái mang 3 hạt trung bình
Trung Quốc	29	23 ± 10,2	6 ± 6,5
Nhật Bản	24	21 ± 12,6	3 ± 4,8
Hàn Quốc	21	19 ± 8,3	2 ± 2,5
Việt Nam	34	30 ± 14,5	4 ± 5,9

Phân tích thành phần chính

Dạng hình trái đậu nành có thể được đánh giá chủ yếu bằng mắt, và việc xếp nhóm thường dựa vào ước lượng cảm quan. Sự đa dạng về dạng trái và dạng hạt, do các nhóm giống có nguồn gốc địa lý khác nhau.

Qua đó cho thấy có sự biến dị rõ rệt về dạng trái đậu nành. Dạng hình trái trung bình của mỗi giống được vẽ bằng cách dựa trên các giá trị trung bình của các hệ số chuẩn hoá Fourier. Việc phân tích thành phần chính được tính từ ma trận hiệp phương sai của các hệ số chuẩn hoá Fourier bằng phần mềm SHAPE (Iwata, 2002).

Các hệ số Fourier elliptic được ước lượng bằng cách nghịch đảo cho trường hợp mà điểm tại thành phần chính có giá trị $\pm 2\sigma$ (độ lệch chuẩn của các điểm tại thành phần chính), trong khi đó các điểm cho khác thành phần chính còn lại được cho bằng không (ghi chú giá trị trung bình của các điểm tại mỗi thành phần chính bằng không). Đây là bài toán được giải quyết một cách đồng thời cho các phương trình nếu các giá trị vec tơ riêng (Eigen vectors) nhận được từ việc phân tích thành phần chính.

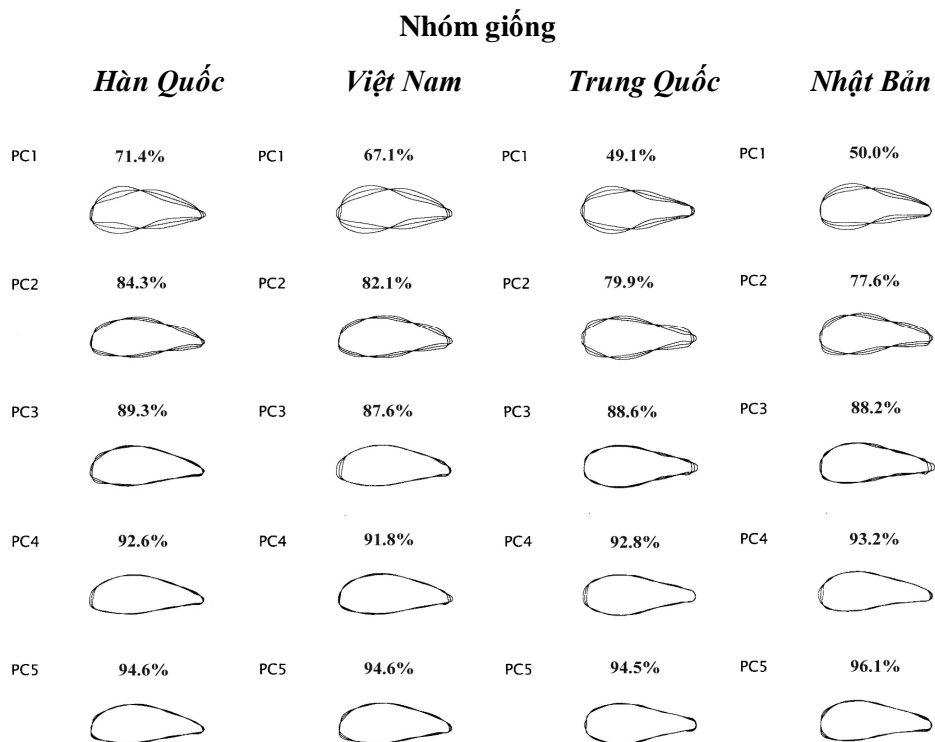
Chúng ta có thể nhận biết bề ngoài mối quan hệ giữa mỗi thành phần chính và dạng hình trái bằng cách vẽ đường viền với các hệ số ước lượng cho thành phần chính tương ứng (Hình 1 - Hình 3). Thành phần chính thứ nhất tượng trưng cho tỉ số phần ngoại vi của dạng hình. Điều này có thể được giải thích một cách đơn giản bởi yếu tố giá

trị vec tơ riêng d_1^{**} lớn đáng kể tại thành phần chính thứ nhất, vì giá trị d_1^{**} tương trưng cho chiều dài tương đối của trục phụ so với trục chính ở đường cong ellip điều hoà thứ nhất, tức là điểm thành phần chính thứ nhất đôi khi tương đương chỉ số dạng hình (tỉ số bề ngoài của các trái).

Điểm thành phần chính thứ hai tương trưng cho vị trí của tâm trái dọc theo đường chính giữa. Điểm thành phần chính thứ ba và thứ tư tương trưng đường thẳng hoặc đường cong của dạng trái. Mỗi thành phần chính được xem là độc lập với các thành phần chính khác. Với giá trị đóng góp tích lũy tại năm thành phần chính đầu tiên đã giải thích trên 95% đối với dạng trái mang 1 hạt, do đó chúng ta có thể nói năm thành phần chính đầu tiên là có ý nghĩa về mặt thống kê trong việc giải thích sự biến động dạng hình của trái đậu nành. Điều này cũng được giải thích tương tự dạng trái mang 2 hạt và mang 3 hạt.

Kết quả phân tích thành phần chính đối với các loại trái như trái mang 1 hạt, trái mang 2 hạt, và trái mang 3 hạt được trình bày qua các Hình 1, Hình 2, và Hình 3. Sự đóng góp ở năm thành phần chính đầu tiên đối với toàn bộ biến dị của dạng trái mang 1 hạt là 95%, 85% đối với toàn bộ biến dị của dạng trái mang 2 hạt, và tương tự 82% cho dạng trái mang 3 hạt.

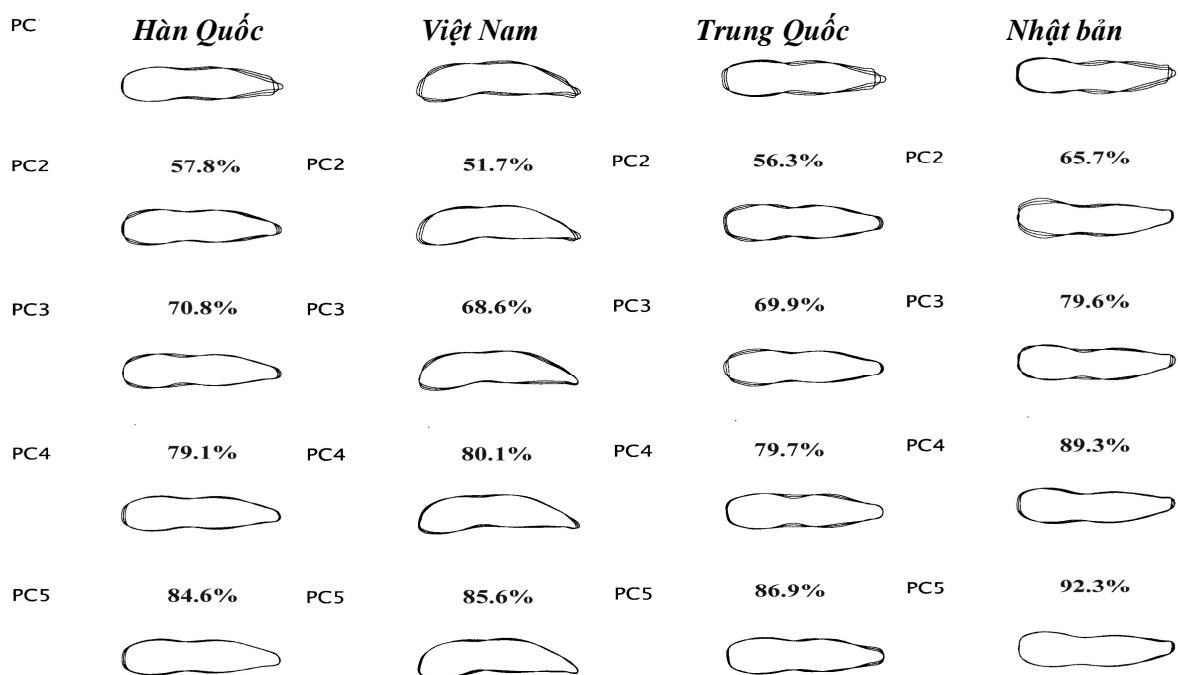
Chúng ta có thể nhận biết rõ mối liên hệ giữa mỗi thành phần chính và dạng hình thông qua hình vẽ đường viền với các hệ số ước lượng cho từng thành phần chính. Từ kết quả này, chúng ta có thể phân biệt các kiểu gen dựa trên dạng hình của trái.



Hình 1. Phân tích dạng hình trái mang 1 hạt của 4 nhóm giống đậu nành

Ghi chú: Mỗi cột tương trưng dạng đường viền của trái được vẽ dựa trên các hệ số Fourier bằng cách ước lượng điểm thành phần chính (+2 σ , -2 σ , và 0)

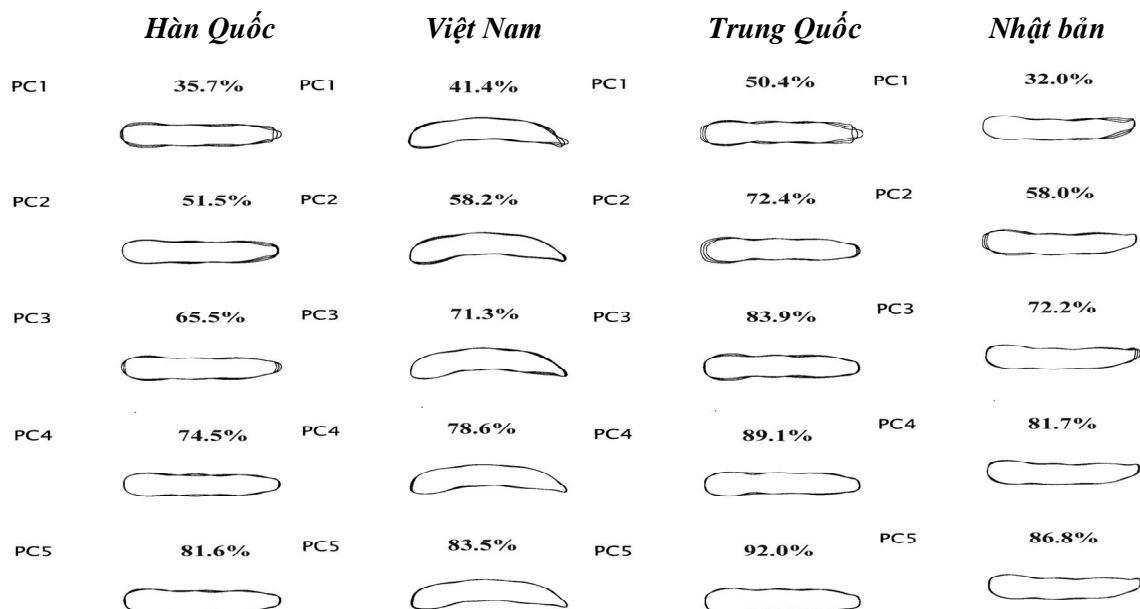
Nhóm giống



Hình 2. Phân tích dạng hình trái mang 2 hạt của 4 nhóm giống đậu nành

Ghi chú: Mỗi cột tượng trưng dạng đường viền của trái được vẽ dựa trên các hệ số Fourier bằng cách ước lượng điểm thành phần chính (+2 σ , -2 σ , và 0)

Nhóm giống



Hình 3. Phân tích dạng hình trái mang 3 hạt của 4 nhóm giống đậu nành

Ghi chú: Mỗi cột tượng trưng dạng đường viền của trái được vẽ dựa trên các hệ số Fourier bằng cách ước lượng điểm thành phần chính (+2 σ , -2 σ , và 0)

4. KẾT LUẬN VÀ ĐỀ NGHỊ

4.1 Kết luận

Nhóm giống Việt nam có tổng số trái đạt nhiều nhất (34 trái) so với các nhóm giống khác, và đồng thời cũng có số trái mang 1 hạt và số trái mang 2 hạt nhiều hơn. Riêng số trái mang 3 hạt thì nhóm giống Trung quốc đạt nhiều nhất (6 trái) so với các nhóm giống còn lại.

Phân tích đường viền với năm thành phần chính thứ nhất đã giải thích 95% biến dị dạng hình của trái mang 1 hạt qua 80 hệ số Fourier, tương tự 85% đối với trái mang 2 hạt, và 82% đối với trái mang 3 hạt.

Việc phân tích phương sai và so sánh đồng thời các thông số ở thành phần chính chỉ ra rằng có sự khác biệt lớn về kiểu gen ở 5 thành phần chính đầu tiên, do các biến dị ở những thành phần này là liên tục và dạng hình khi đó được xem là do đa gen kiểm soát.

4.2 Đề nghị

Phương pháp phân tích đường viền (Fourier elliptic) đã và đang được áp dụng nhiều trong việc phân tích dạng hình vật thể. Thực tế, phương pháp Fourier elliptic là phương pháp hữu ích cho việc mô tả dạng hình sinh học, và có thể được dùng để phân tích dạng hình trên các đối tượng cây trồng khác như dạng trái, dạng lá, dạng thân cụ thể như: dạng lá và dạng trái xoài, dạng lá và củ cà rốt.....

CẢM TẠ

Đề tài được thực hiện với sự hỗ trợ kinh phí của đề tài cấp Bộ; tác giả xin chân thành cảm ơn các em sinh viên Nông học khóa 33, Khoa NÔNG NGHIỆP VÀ SINH HỌC ỨNG DỤNG đã giúp đo đếm các chỉ tiêu nông học, năng suất và các thành phần năng suất.

TÀI LIỆU THAM KHẢO

- Bookstein FL, RE Strauss, JM Humphries, B Cheronoff, RL Elder, and GR Smith. 1982. A comment upon the uses of Fourier methods in systematics. Syst. Zool. 31: 85-92.
- Diaz G, A Zuccarelli, I Pelligra, and A Ghiani. 1989. Elliptic Fourier analysis of cell and nuclear shape. Comp. Biomed. Res. 22: 405-414.
- Ferson S, FJ Rohlf, and RK Koehn. 1985. Measuring shape variation of two dimensional outlines. Syst Zool 34: 59-68.
- Furuta N, S Ninomiya, N Takahashi, H Ohmori, and Y Ukai. 1995. Quantitative evaluation of soybean (*Glycine max* L. Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. Breed. Sci 45: 315-320.
- Giardina, C.R., and F.P. Kuhl. 1977. Accuracy of curve approximation by harmonically related vectors with elliptical loci. Comp. Grap. Ima. Proc. 6: 277-285.
- Granlund GH. 1972. Fourier preprocessing for hand print character recognition. IEEE. Trans. Comput. C-21: 269-281.

- Iwata H, and Y Ukai. 2002. SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *Journal of Hered.* 93: 384-385.
- Iwata H, H Nesumi, S Ninomiya, Y Takano and Y Ukai. 2002. Diallel analysis of leaf shape variations of citrus varieties based on elliptic Fourier descriptors. *Breeding Sci.* 52: 89-94.
- Iwata H, H Nesumi, S Ninomiya, Y Takano and Y Ukai. 2002. The evaluation of genotype x environment interactions of citrus leaf morphology using image analysis and elliptic Fourier descriptors. *Breeding Sci.* 52: 243-251.
- Iwata H, S Niikura, S Matsuura, Y Takano and Y Ukai. 2000. Diallel analysis of root shape of Japanese radish (*Raphanus sativus* L.) based on elliptic Fourier descriptors. *Breeding Sci.* 50: 73-80.
- Keefe PD, and SR Draper. 1986. An automated machine vision system for the morphometry of new cultivars and plant gene bank accessions. *Plant Varieties and Seeds* 1: 1-11.
- Kuhl FP, and CR Giardina. 1982. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing* 18: 236-258.
- Matano T and A Ujihara. 1973. Agroecotypes of *Fagopyrum* II. Classification of common buckwheat (*Fagopyrum esculentum* Moench.) by kernel shape and these geographical distributions in Japan. *Jpn. J. Crop Sci.* 42: 29-30.
- Myers DG, and KJ Edsall. 1989. The application on image processing techniques to the identification of Australian Wheat Varieties. *Plant Varieties and Seeds* 2: 109-116.
- Neuman M, HD Saperstein, E Shwedy and W Bushuk. 1987. Discrimination of wheat class and variety by digital image analysis of whole grain samples. *J. Cereal Sci.* 6: 125-132.
- Ninomiya S, R Ohsawa and M Yoshida. 1995. Evaluation of buckwheat and tartary buckwheat kernel shape by elliptic Fourier method. In: T. Matano & Ujihara (Eds.), *Current Advances in Buckwheat Research*, pp. 389-396. Shinshu Univ. Press.
- Yoshida M, S Ninomiya, M Oide, M Hagiwara, A Ujihara and T Matano. 1995. Geographical variation of kernel shape on tartary buckwheat from the world. In: T. Matano & A. Ujihara (Eds). *Current Advances in Buckwheat Research*, 397-404. Shinshu Univ Press.

Genetic Diversity Analysis of soybean pod shape and pod number by Fourier Elliptic

Truong Trong Ngon, Tran Thi Thanh Thuy

Abstract

Pod number and pod shape are two important factors influencing final yield in soybean. Pod number of eighty genotypes was recorded, and pod shape was evaluated quantitatively by using elliptic Fourier descriptors. Eighty soybean cultivars from different regions were sown from January to April 2010, Hung Thanh

ward, Cai Rang district, Can Tho city. The experiment was carried out in Random Complete Block Design with three replications. Each cultivar was sown on two rows, each row with five metre length. The distance for rows and plants was 40 x 10cm, with 2 plants per hill. Five plants were randomly selected for measuring agronomic traits. Twelve pods per cultivar were used to analyzed pod shape. 80 elliptic Fourier coefficients for each kind of pod were calculated for each contour. Vietnamese cultivar group had highest pod number (about 34 pods) as compared with the other groups, but China cultivar group gave more three-seeded pods than others. The cumulative contribution at the fifth principal component was more than 95% for one-seeded pods, and more than 85% and 82% for two-seeded and three-seed pods, respectively.

Keywords: *Elliptic Fourier descriptors, pod shape, principal component analysis (PCA)*

Phản biện: GSTS. Bùi chí Bửu