

# PHÁT HIỆN MỘT SỐ ĐỘT BIẾN TRÊN VÙNG MÃ HÓA CỦA GEN BGIOSGA024502 (*Ghd7*)

Nguyễn Thị Hồng<sup>1,\*</sup>, Yoshikazu Tanaka<sup>2</sup>, Võ Thị Minh Tuyền<sup>1</sup>, Lê Huy Hàm<sup>1</sup>  
<sup>1</sup>Viện Di truyền Nông nghiệp, đường Phạm Văn Đồng, Bắc Từ Liêm, Hà Nội, Việt Nam  
<sup>2</sup>Trung tâm Nghiên cứu Năng lượng Wakasa-wan, Fukui, Nhật Bản  
\*Email: nguyenhongdhnn@gmail.com

**Tóm tắt:** Nhiều công bố chỉ ra rằng ion beam là tác nhân tạo ra nhiều đột biến điểm có ý nghĩa trong chọn giống cây trồng. Gen BGIOSGA024502 (*Ghd7*) có tác động lớn đến việc quy định các tính trạng như: số hạt trên bông, kích thước hạt, chiều cao cây, ngày trổ bông. Gen BGIOSGA024502 (*Ghd7*) được định vị trên nhiễm sắc thể số 7 (vị trí 9.172.628 đến 9.175.046 reverse strand), với tổng chiều dài 2418 bp, gồm 2 vùng mã hóa (exon1 với chiều dài 444 bp; exon 2 với chiều dài 330 bp) và 1 vùng không mã hóa (intron với chiều dài 1644 bp). Bốn mô hình được thiết kế dựa trên cơ sở dữ liệu nhằm khuếch đại gen BGIOSGA024502 và giải trình tự vùng mã hóa. Tổng số 1548 trình tự mã hóa cho gen BGIOSGA024502 (*Ghd7*) (774 trình tự của dòng đột biến và 774 trình tự của dòng gốc) đã được đọc trình tự theo phương pháp Sanger. Bốn đột biến điểm được xác định bao gồm, hai nucleotit tại vị trí 332 và 336 trên vùng mã hóa 1 (exon 1) và hai nucleotit tại vị trí 72 và 253 trên vùng mã hóa 2 (exon 2). Dựa trên các đột biến này, hai chỉ thị phân tử mới đã được phát triển nhằm nâng cao hiệu quả của chọn giống lúa đột biến.

**Từ khóa:** BGIOSGA024502, *Ghd7*, ion beam, chọn giống đột biến, đột biến điểm, giải trình tự

## 1. MỞ ĐẦU

Bức xạ ion được đánh giá là dạng bức xạ có hệ số truyền năng lượng cao, có khả năng tạo ra nhiều sự xáo trộn trong hệ gen [2, 7, 8]; gây ra nhiều đột biến trên cấu trúc ADN bao gồm cả đột biến lớn và đột biến điểm [5]; có tiềm năng trong việc tạo ra nhiều sự tái tổ hợp khác nhau, là cơ sở để tạo nên một giống cây trồng mới [1]. Gen BGIOSGA024502 (*Ghd7*) quy định các tính trạng quan trọng trên lúa như: số gié/bông, số hạt trên bông, kích thước bông, kích thước hạt, chiều cao cây, ngày trổ bông, phản ứng với môi trường [4, 6, 9]. Một số nghiên cứu chỉ ra rằng, *Ghd7* có năm trạng thái alen [6] và rất nhiều đột biến SNP [3] quy định nên các kiểu hình khác nhau. Trên cơ sở dữ liệu ([www.grammene.org](http://www.grammene.org)), gen BGIOSGA024502 (*Ghd7*) được định vị trên nhiễm sắc thể số 7 (vị trí 9.172.628 đến 9.175.046), với tổng chiều dài 2418bp, gồm 2 vùng mã hóa (exon1 - 444 bp; exon2 - 330 bp) và 1 vùng không mã hóa (intron - 1645 bp).

Dòng đột biến triển vọng mang một số đặc điểm khác biệt nổi bật so với dòng gốc như: kích thước hạt lớn hơn, hạt sáng màu hơn, cao cây hơn và trổ sớm hơn. Vì vậy, sử dụng phương pháp PCR-giải trình tự kết hợp so sánh bằng phương pháp BLAST, chúng tôi tập trung nghiên cứu gen BGIOSGA024502 (*Ghd7*) trên dòng đột biến và giống gốc nhằm tìm hiểu các đột biến phát sinh nếu có.

## 2. NỘI DUNG

**2.1. Vật liệu:** Giống lúa (hai mẫu giống lúa của dòng đột biến và giống gốc) và các hóa chất thí nghiệm sử dụng cho PCR, điện di và giải trình tự.

### 2.2. Phương pháp

#### a. Phương pháp tách chiết ADN

ADN tổng số được tách chiết từ mẫu lá bằng DNeasy Plant Mini Kit (QIAGEN) theo hướng dẫn của nhà sản xuất.

#### b. Phương pháp PCR

Gen BGIOSGA024502 được khuếch đại theo thành phần phản ứng và chu trình nhiệt như sau: Tổng thể tích phản ứng 20  $\mu$ l bao gồm: 1  $\mu$ l ADN tổng số (1ng/ $\mu$ l); 10  $\mu$ l 2X Prime STAR MAX; 0,5  $\mu$ l Môi xuôi (20pmol/ $\mu$ l); 0,5  $\mu$ l Môi ngược (20pmol/ $\mu$ l); 8  $\mu$ l H<sub>2</sub>O. Chu trình phản ứng PCR: 98<sup>0</sup>C - 2 phút; 30 chu kỳ của: 98<sup>0</sup>C - 5 giây, 60<sup>0</sup>C - 5 giây, 72<sup>0</sup>C - 30 giây; 72<sup>0</sup>C - 5 phút; giữ mẫu ở 4<sup>0</sup>C. Sản phẩm PCR được kiểm tra trên gel agarose 1,5% đệm TAE 1X sau đó tinh sạch bằng PCR product purification Kit- QIAGEN

### **b. Phương pháp giải trình tự**

Gen mục tiêu được giải trình tự theo phương pháp Sanger gồm các bước: Bước 1. Phản ứng giải trình tự được tiến hành với tổng thể tích 20  $\mu$ l bao gồm: 1  $\mu$ l ADN gen mục tiêu (1ng/ $\mu$ l); 4  $\mu$ l SeqSaver Sequencing Pre-mix (Sigma); 4  $\mu$ l Môi (20pmol/ $\mu$ l); 11  $\mu$ l H<sub>2</sub>O. Chu trình phản ứng: 94<sup>0</sup>C - 2 phút; 30 chu kỳ của: 98<sup>0</sup>C - 5 giây, 55<sup>0</sup>C - 5 giây, 72<sup>0</sup>C - 15 giây; 72<sup>0</sup>C - 7 phút; giữ mẫu ở 4<sup>0</sup>C. Bước 2. Thực hiện tinh sạch phản ứng theo quy trình DyeEx 2.0 SpinKit (QIAGEN). Bước 3. Đọc trình tự: Sử dụng BigDye Terminator Sequencing Standard Kit (Thermofisher) và đọc kết quả bằng máy ABI PRISM 3100 Genetic Analyzer.

### **c. Phương pháp xử lý số liệu**

Kết quả giải trình tự được xử lý và so sánh bằng phần mềm Bioedit và công cụ BLAST (Basic Local Alignment Search Tool).

## **2.3. Kết quả**

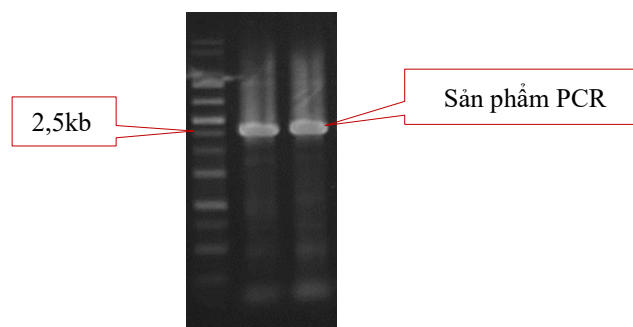
### **a. Khai thác dữ liệu và khuếch đại gen BGIOSGA024502 (*Ghd7*)**

Khai thác dữ liệu về trình tự gen BGIOSGA024502 (*Ghd7*) chúng tôi đã thiết kế một số môi phục vụ nghiên cứu trong bảng 1.

**Bảng 1. Thông tin môi mới thiết kế cho nghiên cứu gen BGIOSGA024502 (*Ghd7*)**

TT	Tên môi	Trình tự (5'-3')	Chiều dài (nu)	GC (%)	T <sub>m</sub> (oC)
1	Ghd7-1F	AGCTCAAGTGACCTCACCTGCTATA	25	48	60,7
2	Ghd7-1R	GATCATGCCGCGCGGATCAGGATTA	25	53	65,4
3	Ghd7-2F	AGGGAGGTTACAATACTGCATTT	25	40	57,2
4	Ghd7-2R	AGTGGTATATACGCACTGTAATTAT	25	32	51,6

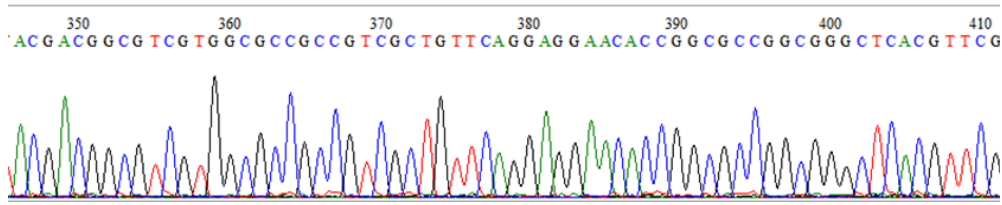
Kết quả khuếch đại toàn bộ chiều dài gen BGIOSGA024502 (*Ghd7*) (Hình 1).



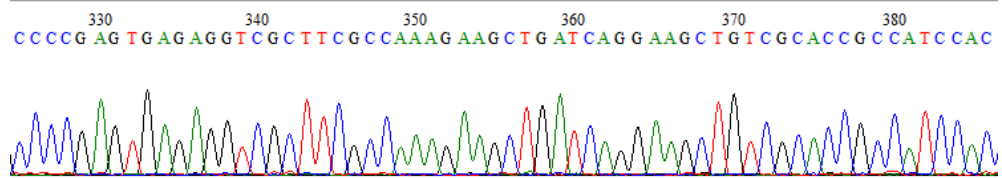
**Hình 1. Sản phẩm khuếch đại gen BGIOSGA024502 (*Ghd7*) trên gel agarose 1,5% (Ghi chú: Làn 1 - 1kb; làn 2 và 3 - sản phẩm PCR)**

**b. Giải trình tự vùng mã hóa của gen BGIOGA024502 (Ghd7)**

Kết quả đọc trình tự được thể hiện trong hình 2.



a. Trình tự vùng mã hóa 1 của dòng đột biến (môi Ghd7-1F)



b. Trình tự vùng mã hóa 2 của dòng đột biến (môi Ghd7-2F)

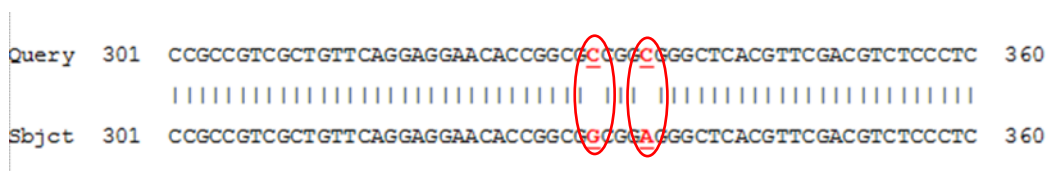
**Hình 2. Kết quả giải trình tự gen BGIOGA024502 (Ghd7)**

**c. Phát hiện đột biến và phát triển chỉ thị ADN**

Tổng hợp kết quả phát hiện đột biến được thể hiện trong bảng 2, hình 3, 4 và 5.

**Bảng 2. Phân tích đột biến trên vùng mã hóa của gen BGIOGA024502 (Ghd7)**

Vùng mã hóa	Số trình tự đọc	Số trình tự sai khác	Tỷ lệ sai khác (%)	Vị trí sai khác
Vùng mã hóa 1	444	2	0,45	362: G -> C 366: A -> C
Vùng mã hóa 2	330	2	0,61	73: A -> G 253: C -> G
<b>Tổng</b>	<b>774</b>	<b>4</b>	<b>0,51</b>	



**Hình 3. Phát hiện đột biến trên vùng mã hóa 1 của gen BGIOGA024502 (Ghd7)**

Ghi chú: Query -trình tự dòng đột biến; Subject = trình tự của dòng gốc



**Hình 4. Phát hiện đột biến trên vùng mã hóa 2 của gen BGIOGA024502 (Ghd7)**

Ghi chú: Query -trình tự dòng đột biến; Subject = trình tự của dòng gốc

Query 241 TTCGCCAAAGAAGCTGATCAGGAAGCTGTCGCACCGCCATCCACCTATGTCGATCCTAGT 300  
 Sbjct 241 TTCGCCAAAGAAGCTGATCAGGAAGCTGTCGCACCGCCATCCACCTATGTCGATCCTAGT 300

**Hình 5. Phát hiện đột biến trên vùng mã hóa 2 của gen BGIOGA024502 (*Ghd7*)**

Ghi chú: Query -trình tự dòng đột biến; Subject = trình tự của dòng gốc

Dựa trên các đột biến điểm, hai chỉ thị ADN mới được phát triển. Thông tin về hai chỉ thị ADN này được thể hiện trong bảng 3.

**Bảng 3. Phát triển chỉ thị ADN mới dựa trên đột biến thu nhận được**

TT	Đặc điểm mỗi				Chiều dài sản phẩm	Mục tiêu nhận biết
	Trình tự (5'-3')	Chiều dài	Tm (oC)	GC (%)		
1	F: CCGGTGCACGAGTTCCAGTTCT R: CGAACGTGAGCCC <u>G</u> CCCG <u>G</u>	22 nu 18 nu	63,8 67,7	59,1 77,8	202 bp	Trình tự <u>C</u> ở vị trí <u>362</u> và trình tự <u>C</u> ở vị trí <u>366</u> trên exon 1
2	F: GAGATGGTGGCCGCCATGGCC <u>G</u> R: GTGCGACAGCTTCCTGATCAG <u>C</u>	22 nu 22 nu	71,1 63,0	72,7 59,1	223 bp	Trình tự <u>G</u> ở vị trí <u>73</u> và trình tự <u>G</u> ở vị trí <u>253</u> trên exon 2

Ghi chú: ký tự đậm, gạch chân - điểm đột biến

## 2.4. Thảo luận

### a. Khai thác dữ liệu, khuếch đại và giải trình tự gen BGIOGA024502 (*Ghd7*)

Khai thác cơ sở dữ liệu, chúng tôi tìm thấy các thông tin liên quan đến gen BGIOGA024502 (*Ghd7*), với trình tự đã được giải mã hoàn chỉnh. Bốn vị trí thiết kế mỗi (hai mỗi xuôi, hai mỗi ngược) được chọn nhằm đảm bảo tối ưu các yêu cầu về thiết kế mỗi (bảng 1). Các mỗi được thiết kế với chiều dài 25 nucleotit, tỷ lệ GC dao động từ 32-53 %, nhiệt độ gắn mỗi từ 51,6 –65,4°C. Mỗi *Ghd7*-2R có tỷ lệ GC thấp nhất (32%) kéo theo nhiệt độ gắn mỗi cũng thấp nhất do phải tránh vùng lặp (AT)<sup>23</sup> và không quá xa vùng mã hóa. Các mỗi còn lại đều đảm bảo các yêu cầu quan trọng được tối ưu hóa như tỷ lệ GC từ 40-60 %; nhiệt độ gắn mỗi từ 55-65°C; chiều dài 18-30 nucleotit....

Toàn bộ chiều dài gen BGIOGA024502 (*Ghd7*) được khuếch đại bằng cặp mỗi *Ghd7*-1F và *Ghd7*-2R, với chiều dài đoạn khuếch đại khoảng 2,5 kb. Từ hình 1 ta có thể thấy, các sản phẩm không mong muốn xuất hiện không đáng kể (băng rất mờ), trong khi đó sản phẩm chính (gen mục tiêu) thì xuất hiện rất rõ với vạch băng dày trên gel agarose ở chiều dài khoảng 2,5 kb. Điều đó cho thấy sản phẩm PCR đủ chất lượng và số lượng để giải trình tự. Sản phẩm PCR khuếch đại gen *Ghd7* của dòng gốc và dòng đối chứng được tinh sạch bằng DyeEx 2.0 Spin Kit của QIAGEN và được đo nồng độ bằng máy Nano Drop. Nồng độ ADN sau khi tinh sạch của gen *Ghd7* đo được ở mẫu gốc là 65,3 ng/μl; và ở mẫu đột biến là 45,4 ng/μl.

Sản phẩm PCR sau khi được tinh sạch đã được sử dụng làm khuôn trong phản ứng giải trình tự bằng BigDye Terminator Sequencing Standard Kit của Thermofisher. Sản phẩm được chạy và đọc bằng máy ABI PRISM 3100 Genetic Analyzer. Kết quả đọc trình tự vùng mã hóa 1 bằng mỗi *Ghd7*-1F và vùng mã hóa 2 bằng mỗi *Ghd7*-2F (hình 2) cho thấy: Các trình tự được đọc rất rõ ràng, không có trình tự nào bị lỗi đọc

(N); các đỉnh (peak) cao, gọn; đường nhiễu (baseline-noise) rất nhỏ, gần như bằng không. Toàn bộ hai vùng mã hóa đã được đọc hoàn chỉnh nhờ hai môi Ghd7-1F và Ghd7-2F nên không cần thiết đọc trình tự bằng hai môi ngược Ghd7-1R và Ghd7-2R.

### **c. Phát hiện đột biến và phát triển chỉ thị ADN**

Trình tự hai vùng mã hóa của gen *Ghd7* giữa dòng gốc và dòng đột biến được so sánh để phát hiện các đột biến thông qua công cụ BLAST. Số liệu trong bảng 2 cho thấy, trong tổng số 774 trình tự mã hóa của gen *Ghd7* thì có 4 trình tự khác nhau giữa dòng gốc và dòng đột biến (chiếm tỷ lệ 0,51%) đã được phát hiện, bao gồm 2 sai khác ở vùng mã hóa 1 và hai sai khác ở vùng mã hóa 2. Các đột biến được thể hiện chi tiết trong hình 4, 5 và 6 bao gồm: trên vùng mã hóa 1, tại vị trí 362 có sự thay thế G thành C và vị trí 366 có sự thay thế A thành C (hình 3); trên vùng mã hóa 2 có sự thay thế nucleotit A thành G tại vị trí 73 (hình 4); và thay thế nucleotit C thành G tại vị trí 253 (hình 5). Dựa trên các đột biến điểm này, hai chỉ thị ADN mới đã được phát triển với mục tiêu nâng cao hiệu quả của chọn giống đột biến liên quan đến gen *Ghd7*. Hai chỉ thị (môi) mới có chiều dài 18-20 nucleotit; nhiệt độ gắn môi trong khoảng 63- 71°C; thành phần GC từ 59,1-77,8 % và sản phẩm khuếch đại mong muốn là 202 bp và 223 bp. Các đột biến được xác định tại những vị trí cố định, vì vậy việc thiết kế chỉ thị dựa trên chúng sẽ khó có thể đảm bảo tối ưu tất cả các yêu cầu đặt ra. Nhiệt độ gắn môi không tối ưu có thể được khắc phục bằng cách điều chỉnh chu trình PCR, hoặc giảm bớt chiều dài môi, tuy nhiên không nên ngắn hơn 18 nucleotit để đảm bảo tính đặc hiệu. Một điều kiện quan trọng khác là các trình tự nhận biết đột biến nên được đặt ở đầu 3' giúp tăng tính chính xác trong quá trình tìm và bắt cặp của môi.

Sự khác nhau đặc trưng nhất giữa dòng đột biến và giống gốc là ở màu sắc lá, màu sắc vỏ trấu và kích thước hạt. Giống gốc có vỏ trấu vàng sẫm, lá xanh đậm, dài hạt thóc 10 mm, dài hạt gạo 7 mm. Trong khi đó, dòng đột biến có lá sáng màu hơn, vỏ trấu vàng sáng, dài hạt thóc 12 mm, dài hạt gạo 8 mm. Có nhiều công trình công bố phát hiện nhiều SNPs cũng như nhiều trạng thái allel khác nhau của gen *Ghd7* [3, 6], nhưng không có đột biến nào giống các đột biến mà nghiên cứu thu được. Kết quả này khẳng định lại hiệu quả tạo đột biến của ion beam trong cải tiến năng suất trên cây lúa. Tuy nhiên, để xác định chính xác các đột biến này quy định nên các tính trạng cụ thể nào thì chúng ta cần thực hiện nhiều nghiên cứu sâu hơn nữa.

## **3. KẾT LUẬN**

Phát hiện bốn đột biến điểm trên vùng mã hóa của gen BGIOSGA024502 (*Ghd7*) bao gồm: đột biến thay thế G thành C tại vị trí 362 và đột biến thay thế A thành C tại vị trí 366 trên vùng mã hóa 1; đột biến thay thế A thành G tại vị trí 73 và đột biến thay thế C thành G tại vị trí 253 trên vùng mã hóa 2.

Dựa trên bốn đột biến điểm được phát hiện ở trên, hai chỉ thị ADN mới đã được phát triển. Tiếp tục đánh giá, thử nghiệm hai chỉ thị mới thiết kế trước khi áp dụng vào chọn giống đột biến liên quan đến gen BGIOSGA024502 (*Ghd7*).

## **LỜI CẢM ƠN**

Nghiên cứu này được thực hiện tại Trung tâm nghiên cứu Năng lượng Wakasan, Fukui, Nhật Bản với sự tài trợ của Chương trình “Fukui International Human Resources Development Center For Atomic Energy (FIHRDC)/The Wakasa Wan Energy Research Center (WERC) FY 2016”.

## TÀI LIỆU THAM KHẢO

- [1] Hirano T., Kazama Y., Ishii K., Ohbu S., Shirakawa Y., Abe T., “Comprehensive identification of mutations induced by heavy-ion beam irradiation in *Arabidopsis thaliana*”, *Plant J.*, 82(1), 93-104, 2015.
- [2] Ishikawa S., Ishimaru Y., Igura M., Kuramata M., Abe T., Senoura T., Hase Y., Arai T., Nishizawa N. K., Nakanishi H., “Ion-beam irradiation, gene identification, and marker-assisted breeding in the development of low-cadmium rice”, *PNAS Early Edition*, ([www.pnas.org/cgi/doi/10.1073/pnas.1211132109](http://www.pnas.org/cgi/doi/10.1073/pnas.1211132109)), 2012.
- [3] Lu L., Yan W., Xue W., Shao D., Xing Y., “Evolution and association analysis of *Ghd7* in rice”, *PLoS ONE* [online], 7(5), e34021, 2012.
- [4] Nemoto Y, Nonoue Y, Yano M, Izawa T., ”*Hdl*, a CONSTANS ortholog in rice, functions as an Ehd1 repressor through interaction with monocot-specific CCT-domain protein *Ghd7*”, *The plant Journal*, 86(3), 221-33, 2016.
- [5] Tanaka A., Shikazono N., Hase Y., “Studies on Biological Effects of Ion Beams on Lethality, Molecular Nature of Mutation, Mutation Rate, and Spectrum of Mutation Phenotype for Mutation Breeding in Higher Plants”, *J.Radiat. Res.*, 51, 223–233, 2010.
- [6] Xue W., Xing Y., Weng X., Zhao Y, Tang W., Wang L., Zhou H., Yu S., Xu C., Li X., Zhang Q., “Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice”, *Nature Genetics*, 40, 761 – 767, 2008.
- [7] Yamaguchi H., Hase Y., Tanaka A., Shikazono N., Degi K., Shimizu A., Morishita T., “Mutagenic effects of ion beam irradiation on rice”, *Breeding Science*, 59(2), 169-177, 2009.
- [8] Yu X., Li Y., Ya H., “The use of microarrays to reveal the probabilistic gene network associated with the response of rice to low-energy ion beam bombardment” *Life Science Journal*, 10(1), 4066-4069, 2013.
- [9] Zhang J., Zhou X., Yan W., Zhang Z., Lu L., Han Z., Zhao H., Liu H., Song P., Hu Y., Shen G., He Q., Guo S., Gao G., Wang G., Xing Y., “Combinations of the *Ghd7*, *Ghd8* and *Hdl* genes largely define the ecogeographical adaptation and yield potential of cultivated rice”, *New Phytologist*, 4, 1056-66, 2015.
- [10] [www.gramene.org](http://www.gramene.org)
- [11] [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

## IDENTIFY SOME POINT MUTATIONS IN CODING REGION OF GENE BGIOSGA024502 (*Ghd7*)

Nguyen Thi Hong<sup>1</sup>, Vo Thi Minh Tuyen<sup>1</sup>, Yoshikazu Tanaka<sup>2</sup>, Le Huy Ham<sup>1</sup>

<sup>a</sup> *Agricultural Genetics Institute, Pham Van Dong, Bac Tu Liem, Hanoi, Vietnam*

<sup>b</sup> *TheWakasa-wan Energy Research Center, Fukui, Japan*

\**Email: nguyenhongdhnn@gmail.com*

**Abstract:** It was reported that ion beams are considered effective for the induction of point mutation valuable for breeding. Gene BGIOSGA024502 (*Ghd7*) has major effects on regulation of main traits in rice, including number of grains per panicle, grain size, plant height and heading date. BGIOSGA024502 (*Ghd7*) was located on chromosome 7 (from 9.172.628 to 9.175.046 on reverse strand), with total length 2418 bp, including two coding regions (exon 1 with length 444 bp; exon 2 with length 330 bp) and one non-coding region (intron with length 1645 bp). Four primers based on database were designed for amplifying and sequencing the target region. A total 1548 coding nucleotides (774 nucleotides of original type and 774 nucleotides of mutant type) was sequenced by Sanger method. Four point mutations were identified including: two nucleotides at points 332 and 336 in exon 1 and two nucleotides at points 72 and 253 in exon 2. Based on these mutations, two new DNA markers were developed for improving efficiency of mutation breeding for rice.

**Keywords:** BGIOSGA024502, *Ghd7*, ion beam, mutation breeding, point mutation, sequencing.