

CONSTRUCTION OF CRISPR/CAS9 VECTOR TO EDIT THE GENE AMYLOSE SYNTHESIS *StGBSS2* OF POTATO VARIETY ATLANTIC

Nguyen Duy Phuong¹, Tran Lan Dai^{2*}

¹Agricultural Genetics Institute, ²Hanoi National University of Education

ARTICLE INFO		ABSTRACT
Received:	03/4/2025	The ratio of resistant starch (amylose) to digestible starch (amylopectin) is one of the key factors determining the quality of potato tuber starch. The <i>StGBSS</i> (granule-bound starch synthase) gene family, which encodes important enzymes involved in amylose biosynthesis, is considered to have a significant impact on potato tuber starch composition. This study aimed to design a CRISPR/Cas9 vector to knock out the <i>StGBSS2</i> gene in the Atlantic potato variety, with the goal of producing potatoes with low amylose content. We successfully isolated and sequenced exon 2 of <i>StGBSS2</i> , designed two specific 22-nucleotide crRNAs, and constructed a transgenic vector pCas9/sgRNA-GBSS2 containing Cas9 expression construct and two sgRNAs. This vector was designed to induce mutations in all four alleles of the target <i>StGBSS2</i> gene. The research results provide a foundation for developing potato varieties with starch composition more suitable for industrial applications and food processing, while also contributing to the functional study of <i>StGBSS2</i> in potato starch synthesis.
Revised:	29/6/2025	
Published:	29/6/2025	
KEYWORDS		
Amylose		
Potato		
CRISPR/Cas9		
Gene editing		
StGBSS2		

THIẾT KẾ HỆ THỐNG CRISPR/CAS9 BẤT HOẠT GENE *StGBSS2* LIÊN QUAN ĐẾN TỔNG HỢP AMYLOSE Ở GIỐNG KHOAI TÂY ATLANTIC

Nguyễn Duy Phương¹, Trần Lan Đài^{2*}

¹Viện Di truyền nông nghiệp, ²Trường Đại học Sư phạm Hà Nội

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	03/4/2025	Tỉ lệ thành phần tinh bột khó tiêu (amylose) và dễ tiêu (amylopectin) là một trong những yếu tố quyết định chất lượng tinh bột củ khoai tây. Nhóm gene <i>StGBSS</i> (<i>granule-bound starch synthase</i>) mã hóa enzyme tham gia vào quá trình sinh tổng hợp amylose, được coi là có ảnh hưởng quan trọng đến thành phần tinh bột củ khoai tây. Nghiên cứu này được thực hiện với mục đích thiết kế vector CRISPR/Cas9 để gây bất hoạt <i>StGBSS2</i> trên giống khoai tây Atlantic, hướng tới tạo ra khoai tây có hàm lượng amylose thấp. Chúng tôi đã phân lập và giải trình tự đoạn exon 2 của <i>StGBSS2</i> , thiết kế hai crRNA đặc hiệu dài 22 nucleotide, và tạo thành công vector chuyển gene pCas9/sgRNA-GBSS2 chứa cấu trúc biểu hiện Cas9 và hai sgRNA. Vector này được thiết kế nhằm hướng đến khả năng gây đột biến trên cả bốn allele của gene đích <i>StGBSS2</i> . Kết quả nghiên cứu là tiền đề trong việc tạo ra giống khoai tây có thành phần tinh bột phù hợp hơn cho các ứng dụng công nghiệp và chế biến thực phẩm, đồng thời góp phần vào nghiên cứu chức năng của <i>StGBSS2</i> trong quá trình tổng hợp tinh bột ở khoai tây.
Ngày hoàn thiện:	29/6/2025	
Ngày đăng:	29/6/2025	
TỪ KHÓA		
Amylose		
Khoai tây		
CRISPR/Cas9		
Chỉnh sửa gene		
StGBSS2		

DOI: <https://doi.org/10.34238/tnu-jst.12480>

* Corresponding author. Email: Daitl@hnue.edu.vn

1. Giới thiệu

Khoai tây (*Solanum tuberosum* L.) được xếp hạng là cây trồng được sản xuất nhiều thứ 4 và được tiêu thụ đứng thứ 3 trên toàn thế giới. Củ khoai tây giàu carbohydrate, vitamin và khoáng chất, đóng vai trò quan trọng cung cấp dinh dưỡng cho con người. Ngoài ra, tinh bột trong củ khoai tây cũng là nguyên liệu cho nhiều ngành công nghiệp bao gồm cả lĩnh vực thực phẩm và phi thực phẩm [1]. Vì vậy, khoai tây luôn được xem là đối tượng quan trọng của các nghiên cứu chọn giống nhằm cải thiện chất lượng dinh dưỡng và các ứng dụng trong công nghiệp.

Tinh bột trong củ khoai tây chiếm 80% chất khô, với thành phần chính là amylose và amylopectin. Tinh bột có hàm lượng amylopectin cao thường được ưa thích trong công nghiệp chế biến thực phẩm và các ứng dụng công nghiệp khác do các đặc tính độc đáo của chúng. Amylopectin dễ dàng gelatin hóa thành dạng sệt, trong, mềm, ổn định và ít bị biến tính hơn khi đông lạnh và rã đông hay bảo quản trong thời gian dài [2], [3]. Bên cạnh đó, loại tinh bột này cũng đã được chứng minh có hiệu quả chuyển hoá thành ethanol cao hơn so với tinh bột chứa tỷ lệ amylopectin thấp, phù hợp cho ngành sản xuất nhiên liệu sinh học [4]. Tổng hợp tinh bột ở khoai tây là một quá trình phức tạp với sự tham gia của nhiều nhóm enzyme khác nhau như AGPase (ADP-glucose pyrophosphorylase), SBE (starch branching enzyme), DBE (starch debranching enzymes), SS (starch synthase), GWD (glucan/water dikinase), PWD (phosphoglucan/water dikinase) [5], [6]. GBSS (granule-bound starch synthase) thuộc nhóm enzyme SS, trong đó GBSS1/waxy được cho là có vai trò quan trọng đối với quá trình sinh tổng hợp amylose [7]. Nhiều nghiên cứu đã chứng minh việc giảm biểu hiện GBSS1 hay mất hoạt tính GBSS1 làm tăng tỉ lệ amylopectin trong thành phần tinh bột củ khoai tây [5], [8], [9]. Trong khi vai trò của StGBSS1 đối với đặc tính tinh bột khoai tây đã được nghiên cứu chi tiết, nghiên cứu về các enzyme khác còn khá hạn chế.

Hệ thống CRISPR/Cas là công cụ chỉnh sửa gene phổ biến nhất hiện nay. Hệ thống này hoạt động dựa trên bộ máy phòng vệ virus CRISPR/Cas của vi khuẩn và bộ máy sửa chữa DNA của sinh vật nhân chuẩn theo cơ chế NHEJ (non-homologous end joining) và HDR (homology-directed repair) để tạo ra các đột biến mong muốn trên gene đích. Nhờ tính đơn giản và hiệu quả, hệ thống chỉnh sửa gene CRISPR/Cas đã được sử dụng cho nhiều mục đích nghiên cứu trên cây trồng, bao gồm cả nghiên cứu cơ bản (nghiên cứu chức năng hệ gene) và nghiên cứu ứng dụng (cải tiến tính trạng) [1], [10], [11].

Trong nghiên cứu này, chúng tôi đã phân tích cơ sở dữ liệu gene và protein khoai tây để xác định gene tương đồng với *StGBSS1* (đặt tên là *StGBSS2*). Để phục vụ tìm hiểu chức năng của *StGBSS2* thông qua bất hoạt gene đích, chúng tôi đã thiết kế hệ thống chỉnh sửa gene CRISPR/Cas9 nhằm vào vị trí exon 2 trên *StGBSS2* của giống khoai tây Atlantic – một giống thương mại rất phổ biến hiện nay. Nghiên cứu làm rõ vai trò của các gene trong nhóm SS sẽ là tiền đề cho việc phát triển giống khoai tây mới có hàm lượng amylopectin cao đáp ứng thị hiếu của thị trường tiêu thụ.

2. Vật liệu và phương pháp nghiên cứu

2.1. Vật liệu nghiên cứu

Giống Khoai tây Atlantic do Trung tâm Nghiên cứu Khoai tây, Rau và Hoa cung cấp. Vi khuẩn *Escherichia coli* DH5 α được mua từ hãng Thermo Scientifics (Mỹ). Vector pENTR4-V5 (mang cấu trúc biểu hiện sgRNA) và pCas9 (mang cấu trúc biểu hiện Cas9) được cung cấp bởi Trung tâm Nghiên cứu vì sự phát triển (Montpellier, Pháp).

Các oligonucleotide sử dụng trong nghiên cứu được thiết kế bằng phần mềm Primer3 và đặt sinh tổng hợp từ Công ty Cổ phần Phù Sa Genomics (Việt Nam).

2.2. Phương pháp nghiên cứu

2.2.1. Phân tích đặc điểm gene in silico

Đặc điểm trình tự và cấu trúc gene được sàng lọc từ cơ sở dữ liệu Ensembl Plant. Quan hệ di truyền giữa các protein được dựng bằng công cụ UniPro. Đặc điểm cấu trúc và chức năng của

protein được sàng lọc từ cơ sở dữ liệu Ensembl Plant, PANTHER, Superfamily, CDD, Pfam, NCBIFAM, TIGRFAM, HAMAP, InterPro.

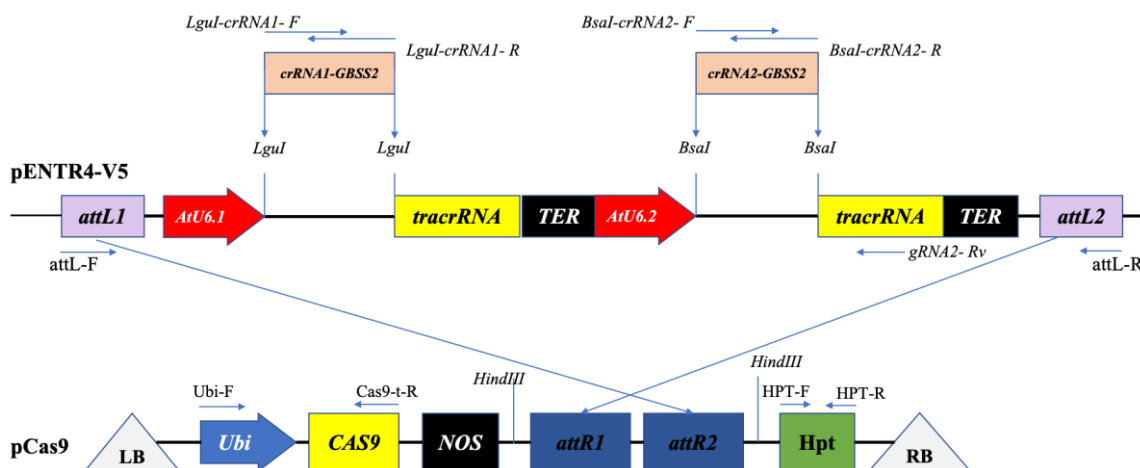
2.2.2. Phân lập và phân tích trình tự *StGBSS2*

DNA tổng số của khoai tây được tách chiết từ lá bằng dung dịch CTAB 2% [12]. Cặp mồi GBSS2-F/GBSS2-R (5'-GGTGGGAGTGGTAGCTTATCA-3'/5'-CCAACTTACGCACTTGTCCC-3') được sử dụng để phân lập đoạn gene *StGBSS2* từ DNA tổng số khoai tây bằng kỹ thuật PCR [13]. Sản phẩm PCR (khoảng 607 bp) được tinh sạch bằng bộ kit *GenJET-TM Gel Extraction* (Thermo Fisher Scientific, Hoa Kỳ) và được giải trình tự nucleotide bởi Công ty Cổ phần Phù Sa Genomics (Việt Nam). Kết quả giải trình tự được xử lý bằng phần mềm BioEdit 4.0 và phân tích bằng công cụ BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

2.2.3. Thiết kế trình tự crRNA chỉnh sửa *StGBSS2*

Trình tự crRNA trên phân tử sgRNA chỉnh sửa *StGBSS2* được xác định bằng phần mềm CRISPR-P v2.0 [14]. Cấu trúc bậc 2 của các sgRNA được phân tích bằng công cụ Mfold 2.3 [15]. Các trình tự DNA trong hệ gene khoai tây tương đồng với crRNA đã thiết kế được xác định bằng công cụ CCTop [16]. Trình tự crRNA được lựa chọn dựa trên chỉ tiêu: hàm lượng GC (30 - 80%), khả năng hình thành và duy trì cấu trúc bậc 2 ổn định (chứa vòng loop RAR, SL2 và SL3, tổng số cặp base trong crRNA < 13, số cặp base liên tục < 8, số cặp base bên trong cấu trúc crRNA < 7) [17], điểm đặc hiệu (on-score) > 0,2, vị trí tạo ra điểm cắt không rơi vào vị trí chứa nucleotide đa hình.

2.2.4. Thiết kế T-DNA biểu hiện cấu trúc sgRNA chỉnh sửa *StGBSS2*



Hình 1. Sơ đồ thiết kế vector biểu hiện pCas9/sgRNA-GBSS2

Ghi chú: (attL1, attL2, attR1, attR2) vị trí nhận biết của LR clonase; (U6.1) promoter *OsU6*; (TER) vùng kết thúc phiên mã; (NOS) vùng kết thúc phiên mã *Nopaline synthase*; (LB, RB) trình tự biên trái, biên phải của T-DNA; (HPT) cấu trúc biểu hiện gene *Hygromycin phosphotransferase*; mũi tên (→) thể hiện vị trí của các mồi.

Các cặp oligonucleotide LguI-crRNA1-F/LguI-crRNA1-R (5'-GGTGGCTTCAGAATGCGCTCCATGG-3'/5'-CCATGGAGCGCATTCTGAAGCCACC-3') và BsaI-crRNA2-F/BsaI-crRNA2-R (5'-CCTCTACTGTAAAGTTCCAAAAGAAC-3') và (5'-GTTCTTTTGGAACTTACAGTAGAGG-3') được liên kết với nhau bằng phản ứng biến tính/phục hồi DNA (95°C - 10 phút, 4°C - ∞) để tạo crRNA sợi đôi hoàn chỉnh (*crRNA1-GBSS2* và *crRNA2-GBSS2*). Hai crRNA này lần lượt được chèn vào vị trí enzyme LguI và BsaI trên vector pENTR4-V5 (Hình 1). Vector tái tổ hợp pEN/sgRNA-GBSS2 được kiểm tra bằng PCR với 2 cặp mồi: LguI-crRNA1-F/BsaI-crRNA2-R (0,45 kb); attL-F/ gRNA2-R (5'-GTACAAAAAAGCAGGCT-3'/5'-GAATTCGATATCAAGCT-3') (0,9 kb). Tiếp theo, đoạn trình tự chứa cấu trúc [*AtU6::sgRNA-GBSS2*] được ghép nối vào pCas9 bằng bộ kit Gateway (Invitrogen). Vector pCas9/sgRNA-GBSS2 được kiểm tra bằng PCR

với các cặp môi *LguI*-crRNA1-F/pEN-R (0,4 kb); Ubi-F/ Cas9-t-R (5'-CCCTGCCTTCATACG CTATT-3')/5'-GCCTCGGCTGTCTCGCCA-3') (0,36 kb); HPT-F/HPT-R (5'-AAGGAGGTG ATCCAGCC-3')/5'-GAGTTTGATCCTGGCTCAG-3') (0,8 kb) và giải trình tự DNA (Công ty DNA sequencing, Việt Nam).

3. Kết quả và bàn luận

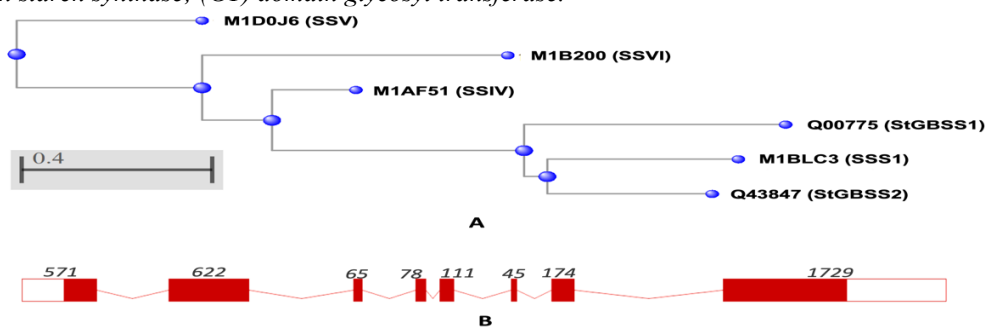
3.1. Phân tích đặc điểm gene *StGBSS2*

StGBSS1/waxy đã được nghiên cứu khá chi tiết về đặc điểm chức năng và vai trò trong quá trình tổng hợp amylose, có vai trò quan trọng đối với thành phần tinh bột củ khoai tây [5], [8], [9]. Sử dụng cơ sở dữ liệu Ensembl Plant và Pfam, chúng tôi đã sàng lọc được trong hệ gene *S. tuberosum* có 6 gene mã hóa các protein có chứa domain SS (Bảng 1), trong đó 4 protein M1AF51 (SS4), Q43847 (SS2), M1BLC3 (SSS1) và Q00775 (*StGBSS1*) cùng mang 2 domain chức năng SS và Glycosyl transferase (GT). Hơn nữa, phân tích quan hệ di truyền giữa các protein mang domain SS của khoai tây cũng chỉ ra rằng Q43847 (SS2) có quan hệ di truyền gần nhất với *StGBSS1* (Hình 2A), gợi ý rằng protein này có chức năng tương tự với *StGBSS1*. Gene mã hóa Q43847 (SS2) bao gồm 8 exon, với vùng mã hóa dài 2301 nucleotide (Hình 2B), mã hóa phân tử protein có kích thước 767 amino acid, khối lượng phân tử 85,2 kDa (Bảng 1).

Bảng 1. Kết quả sàng lọc gene mã hóa protein chứa domain SS

TT	Mã số gene ^a	Vị trí	Mô tả ^a	Mã số protein ^b	Kích thước protein (aa)	Khối lượng protein (kDa)	Vị trí domain chức năng ^{a,b,c}
1	PGSC0003DMG 400008322	2:30142740-30152314	<i>Starch synthase 4</i>	M1AF51	1001	113,7	SS: 505-745 GT: 804-952
2	PGSC0003DMG 400030619	2:42100875-42107550	<i>Starch synthase 5</i>	M1D0J6	390	43,5	SS: 58-297
3	PGSC0003DMG 400001328	2:46189568-46196715	<i>SS2/GBSS2</i>	Q43847	767	85,2	SS: 277-519 GT: 582-697
4	PGSC0003DMG 402018552	3:48221882-48230897	<i>Soluble starch synthase 1</i>	M1BLC3	641	70,6	SS: 132-390 GT: 443-591
5	PGSC0003DMG 402013540	7:40932630-40939876	<i>SS6</i>	M1B200	412	46,3	SS: 126-354
6	PGSC0003DMG 400012111	8:56782186-56785790	<i>GBSS1/WAXY</i>	Q00775	607	66,6	SS: 83-343 GT: 394-524

Ghi chú: ^aTheo cơ sở dữ liệu Ensembl Plant; ^bTheo cơ sở dữ liệu UniPro; ^cTheo cơ sở dữ liệu Pfam; (SS) domain starch synthase; (GT) domain glycosyl transferase.



Hình 2. Đặc điểm gene *StGBSS2*

Ghi chú: (A) Quan hệ di truyền giữa các protein chứa domain SS của khoai tây. Cây phả hệ được dựng bằng công cụ UniPro. Kí tự thể hiện mã số và tên gọi protein (theo cơ sở dữ liệu Ensembl Plant). (B) Cấu trúc gene *StGBSS2*. Hình chữ nhật thể hiện vùng exon: màu trắng thể hiện vùng 5' và 3' không dịch mã, màu đỏ thể hiện vùng mã hóa amino acid; chữ số thể hiện độ dài (nucleotide) của các exon; đường kẻ màu đỏ thể hiện vùng intron.

Việc phân loại các gene/protein thuộc nhóm SS ở khoai tây cho đến nay chưa có sự thống nhất. Dry và cộng sự [18] đã sử dụng kháng huyết thanh kháng GmGBSS1 (đậu tương) để sàng lọc từ thư viện cDNA khoai tây được 2 gene tương đồng với *GmGBSS1* và *GmGBSS2* của đậu tương. Trong khi đó, Nazarian-Firouzabadia và Visser [7] lại xác định nhóm SS ở khoai tây có 4 isoform, đặt tên là StSS1, StSS2, StSS3 và StGBSS1. Gần đây, nhóm SS ở khoai tây được chia thành 2 loại StGBSS và soluble starch synthase (StSSS), trong đó StGBSS bao gồm StGBSS1 và StGBSS2, StSSS gồm 6 loại từ StSS1 – StSS5 [5]. Trong nghiên cứu này, gene mã hóa protein Q43847 (SS2) (được đặt tên là *StGBSS2*) được xác định có sản phẩm protein có mức độ tương đồng cao nhất với StGBSS1 đã được nghiên cứu trước đây. Hơn nữa, 2 protein này cũng được chứng minh có hàm lượng cao nhất trong số các enzyme liên quan tới tổng hợp tinh bột ở khoai tây [5]. Chính vì vậy, chúng tôi đã lựa chọn gene *StGBSS2* này cho nghiên cứu tiếp theo.

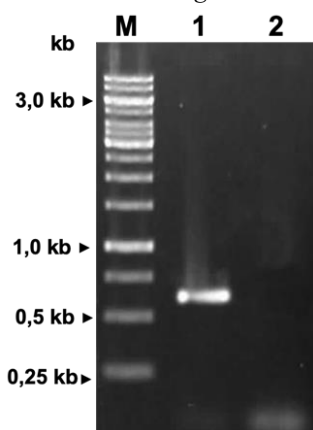
3.2. Phân lập đoạn gene *StGBSS2* của giống khoai tây Atlantic

Đa số các giống khoai tây trong sản xuất có hệ gene tứ bội phức tạp, có sự khác biệt về trình tự nucleotide giữa các allele trong cùng một gene. Tính đa hình nucleotide có thể tạo ra các sự kiện chỉnh sửa gene khác nhau trên mỗi allele khi sử dụng công cụ CRISPR/Cas9 [11]. Do vậy, việc phân tích trình tự đầy đủ của gene đích là cần thiết khi thực hiện nghiên cứu chỉnh sửa gene trên đối tượng cây trồng này.

Bảng 2. Phân tích in silico các domain chức năng của *StGBSS2*

TT	Cơ sở dữ liệu ^a	Vị trí ^b	Mô tả	Vùng exon mã hóa
1	PANTHER (PTHR45825)	26-767	Granule-bound starch synthase 1	2 - 8
2	Superfamily (SSF53756)	276-759	UDP-Glycosyltransferase/ glycogen phosphorylase	2 - 8
3	CDD (cd03791)	277-761	Glycogen synthase	2 - 8
4	Pfam (PF13692)	582-697	Glycosyl transferase	8
5	Pfam (PF08323)	277-519	Starch synthase (domain xúc tác)	2 - 8
6	NCBIFAM	276-760	Glycosyl transferase	2 - 8
7	TIGRFAM (TIGR02095)	276-760	Glycosyl transferase	2 - 8
8	HAMAP (MF_00484)	276-763	Glycogen synthase	2 - 8
9	InterPro (IPR011835)	276-763	Glycogen synthase/ starch synthase	2 - 8
10	InterPro (IPR013534)	277-519	Starch synthase (domain xúc tác)	2 - 8

Ghi chú: ^aKí tự trong ngoặc thể hiện mã số của domain trong cơ sở dữ liệu; ^bTheo cơ sở dữ liệu Ensembl Plant.



Hình 3. Phân lập *StGBSS2* từ giống khoai tây Atlantic

Ghi chú: Nhân bản *StGBSS2* bằng PCR. Giếng M: thang chuẩn DNA 1,0 kb (Thermo Scientific); giếng 1: Sản phẩm PCR của gene *StGBSS2*; giếng 2: đối chứng âm (không có DNA khuôn).

Phân tích cấu trúc gene *StGBSS2* cho thấy tất cả các domain chức năng của *StGBSS2* đều được mã hóa bởi vùng exon 2 - 8 (Bảng 2). Trong nghiên cứu gây đột biến bất hoạt gene chức năng bằng

hệ thống CRISPR/Cas9, vị trí đích của sgRNA càng gần vị trí khởi đầu của domain chức năng thì khả năng thành công càng lớn do có thể tạo đột biến dịch khung làm sai lệch toàn bộ trình tự amino acid của sản phẩm protein. Chính vì vậy, để phục vụ mục đích chỉnh sửa gene *StGBSS2* của giống khoai tây Atlantic, một đoạn trình tự chứa vùng exon 2 của *StGBSS2* đã được lựa chọn để phân lập và giải trình tự nucleotide. Đoạn gene nhân bản được bằng PCR với cặp mồi đặc hiệu từ DNA tổng số có kích thước khoảng 0,6 kb (Hình 3, giống 1). Kết quả giải trình tự sản phẩm PCR cho thấy trình tự phân lập được dài 607 bp, giống >99,8% so với trình tự tương đồng đã được công bố trên GenBank (Mã số NM_001288048.1) (Hình 4). Đáng chú ý, đoạn gene *StGBSS2* của giống Atlantic tồn tại 4 allele có trình tự khác nhau, với 6 vị trí đa hình. Hiện tượng đa hình nucleotide của các allele trên cùng một gene này phù hợp với thực tế thường gặp ở các giống cây trồng đa bội nhân giống vô tính như khoai tây.

Ref	GATGTTGATA	TCCCTGACGT	GGATAAAGAT	TATAATGTTA	CTGTACCTAG	TACTGCTGCT	ACTCCAATTA	CTGATGTCTGA	80
Allele 1:									80
Allele 2:									80
Allele 3:									80
Allele 4:								T	80
Ref	TAAAAATACA	CCGCCTGCTA	TAAGCCAAGA	TTTGTGTGAA	AGTAAAAGAG	AAATCAAAAG	GGACCTGGCC	GATGAAAGGG	160
Allele 1:									160
Allele 2:									160
Allele 3:				A.A					160
Allele 4:									160
Ref	CACCCCACT	GTCGAGATCA	TCTATCAG	CCAGTAGCCA	GATTTCTCT	ACTGTAAGTT	CCAAAAGAAC	GTTGAATGTC	240
Allele 1:									240
Allele 2:		A							240
Allele 3:									240
Allele 4:									240
Ref	CCTCCAGAAA	CTCCGAAGTC	CAGTCAAGAG	ACACTTTTGG	ATGTGAATTC	ACGCAAAAGT	TTAGTAGATG	TTCCTGGAAA	320
Allele 1:									320
Allele 2:						A			320
Allele 3:									320
Allele 4:									320
Ref	GAAGATCCAG	TCTTATATGC	CTTCATTACG	TAAAGAATCC	TCAGCATCCC	ATGTGGAACA	GAGGAATGAA	AATCTTGAAG	400
Allele 1:									400
Allele 2:									400
Allele 3:									400
Allele 4:									400
Ref	GATCAAGTGC	TGAGGCAAAAC	GAAGAGACTG	AAGATCCTGT	GAATATAGAT	GAGAAACCCC	CTCCATTGGC	AGGAACAAAT	480
Allele 1:									480
Allele 2:									480
Allele 3:									480
Allele 4:									480
Ref	GTTATGAACA	TTATTTTGGT	GGCTTCAGAA	TGCGCTCCAAT	GGTCTAAAAC	AG			532
Allele 1:	G								532
Allele 2:									532
Allele 3:									532
Allele 4:									532

Hình 4. Phân tích trình tự *StGBSS2*

Ghi chú: Trình tự 4 allele đoạn gene *StGBSS2* của giống Atlantic (Allele 1 – 4) được so sánh với trình tự tham chiếu trên GenBank (Ref). Dấu chấm thể hiện các nucleotide giống nhau. Vị trí crRNA được đóng khung.

Trong nỗ lực cải tạo chất lượng tinh bột ở khoai tây những năm gần đây, các nhà nghiên cứu tập trung vào việc chuyển gene hoặc chỉnh sửa gene mã hoá các loại enzyme liên quan đến quá trình tổng hợp tinh bột bao gồm *GBSS*, *SS* và *SBE*. Kết quả cho thấy việc bất hoạt (knock out) gene *StGBSS1* làm giảm đáng kể sự tích lũy amylose trong củ khoai tây; trong khi các đột biến trên gene *StSS1*, *StSS2* và *StSBE* làm mất hầu như hoàn toàn sự có mặt của amylopectin [5]. Với mục tiêu tạo ra giống khoai tây không chứa amylose, các nghiên cứu chỉnh sửa gene CRISPR/Cas9 gần đây đã lựa chọn gây đột biến vào các vùng exon trên gene *StGBSS*. Cụ thể, khi biến nạp hệ thống CRISPR/Cas9 vào giống khoai tây tứ bội Sayaka, nhóm nghiên cứu của Kusano đã tạo được các đột biến trên vùng exon 1 của gene *StGBSS1* làm giảm hàm lượng amylose đáng kể so với cây đối chứng [11]. Mới đây, Abeuova và cộng sự [19] cũng thiết kế sgRNA tác động vào 3 vị trí khác nhau trên exon 1 của *StGBSS* các giống khoai tây bản địa Astanalyk, Tokhtar và Aksor và thu được kết quả tương tự. Trong nghiên cứu này, chúng tôi lựa chọn trình tự exon 2 trên gene *StGBSS2* để gây đột biến vì trình tự này có kích thước lớn, nằm gần mã mở đầu ATG, exon này mã hoá amino

acid nằm trong vùng chức năng của StGBSS2 nên có thể bất hoạt chức năng của protein nếu xảy ra đột biến dịch khung.

3.3. Thiết kế trình tự crRNA chỉnh sửa StGBSS2

Với mục tiêu bất hoạt gene ở giống khoai tây Atlantic bằng hệ thống CRISPR/Cas9, trình tự crRNA đã được thiết kế nhằm vào vùng exon 2 trên gene đích. Bốn mươi crRNA (dài 19-22 nucleotide) trên đoạn gene *StGBSS2* đã được xác định bằng công cụ CRISPR-P v2.0, trong đó 8 crRNA có điểm đặc hiệu > 0,2 và không chứa vị trí nucleotide đa hình. Phân tích bằng công cụ Mfold các sgRNA chứa các crRNA này cho thấy 7/8 trình tự có thể tạo ra cấu trúc bậc 2 phù hợp với hệ thống CRISPR/Cas9 (Bảng 3, Hình 5). Để xác định số lượng trình tự off-target (trình tự tương đồng với crRNA trong hệ gene nhưng nằm ngoài gene đích), công cụ CCTop đã được sử dụng; kết quả cho thấy crRNA5 và crRNA6 không có trình tự off-target, trong khi các crRNA khác đều có ít nhất 1 trình tự off-target trong hệ gene tham chiếu. Hai crRNA này (đặt tên lần lượt là *crRNA1-GBSS2* và *crRNA2-GBSS2*) đã được lựa chọn để thiết kế hệ thống CRISPR/Cas9 chỉnh sửa *StGBSS2* (Hình 4).

Bảng 3. Trình tự các crRNA chỉnh sửa *StGBSS2*

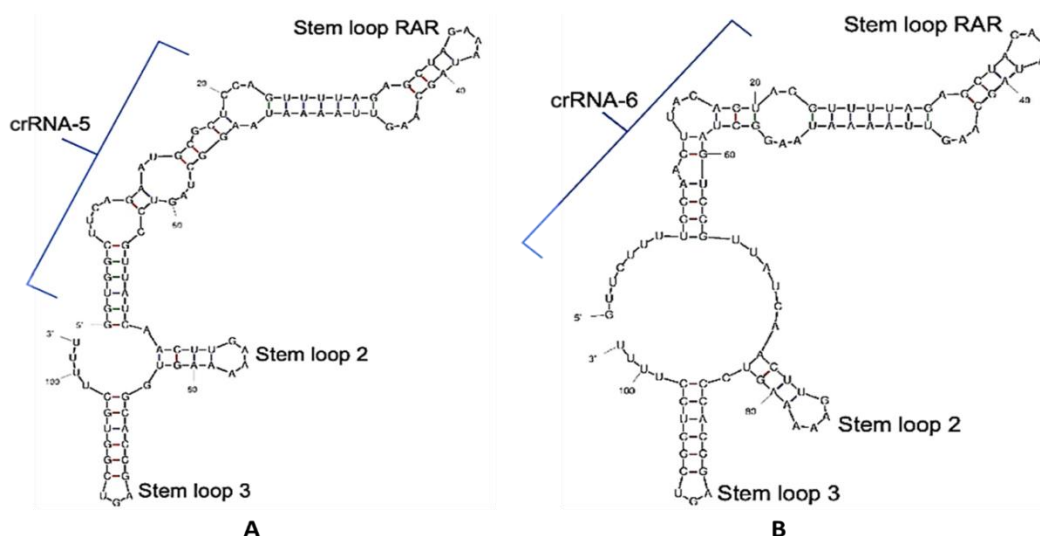
Tên	Trình tự crRNA	Kích thước	Cấu trúc bậc 2 ^a	On score	Off-target ^b
crRNA1	GGACTTCGGAGTTTCTGGA	19	9 TBP; 5 CBP; 4 IBP; 1 GSL	0,6836	19(1)
crRNA2	GCGCTCCATGGTCTAAAAC	19	7 TBP; 4 CBP; 0 IBP; 1 GSL	0,6563	19(0)
crRNA3	GATGAGAAACCCCTCCAT	19	11 TBP; 5 CBP; 0 IBP; 1 GSL	0,3647	13(3)
crRNA4	GTGTCTCTTGACTGGACTT	19	9 TBP; 6 CBP; 0 IBP; 1 GSL	0,3061	19(3)
crRNA5	GGTGGCTTCAGAATGCGCTCCA	22	12 TBP; 6 CBP; 0 IBP; 1 GSL	0,4159	0 (0)
crRNA6	GTTCTTTTGGAACTTACAGTAG	22	9 TBP; 4 CBP; 0 IBP; 1 GSL	0,3217	0(0)
crRNA7	GACTGGACTTCGGAGTTTCTGG	22	12 TBP; 6 CBP; 3 IBP; 1 GSL	0,2295	1(0)

Ghi chú: ^a(TBP) tổng số cặp base; (CBP) số cặp base liên tục; (IBP) số cặp base của trình tự crRNA trong cấu trúc bậc 2 của sgRNA; (GSL) số vòng loop bị mất; On score: Điểm dự đoán khả năng nhận biết đúng vị trí đích.

^bSố trình tự tương đồng nằm trong hệ gene giống khoai tây tham chiếu; chữ số trong ngoặc thể hiện số trình tự tương đồng nằm trên vùng mã hóa của gene chức năng.

Trong các nghiên cứu chỉnh sửa gene bằng CRISPR/Cas9, việc sử dụng nhiều hơn một sgRNA thường giúp tăng hiệu quả của quá trình chỉnh sửa. Điều này càng có ý nghĩa hơn khi chỉnh sửa gene trên các sinh vật đa bội như khoai tây. Villafuerte và cộng sự [20] đã thực hiện chỉnh sửa gene *GBSS* ở giống khoai tây Yukon Gold và thu được dòng khoai tây mang đột biến trên ba allele của gene với kiểu hình hàm lượng amylose không có sự thay đổi đáng kể. Trong nghiên cứu chỉnh sửa gene *StGBSS1* của Kusano và cộng sự [11], việc sử dụng hệ thống chỉnh sửa với ba phân tử sgRNA đã tạo ra hiệu suất đột biến trên cả 4 allele cao hơn rõ rệt so với hệ thống có ít sgRNA hơn (28% so với 7%). Các kết quả nghiên cứu này cho thấy để đảm bảo thu được kiểu hình mong muốn, việc gây đột biến cả 4 allele trong hệ gene khoai tây là rất cần thiết. Trong nghiên cứu này, chúng tôi cũng đã thiết kế 2 sgRNA chỉnh sửa gene *StGBSS2* của giống khoai tây Atlantic để tăng hiệu quả gây đột biến trên cả 4 allele của gene đích.

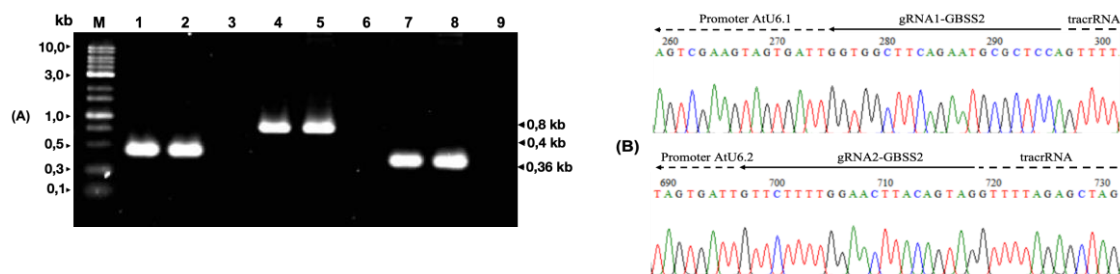
Quan trọng hơn, trong quá trình thiết kế sgRNA, chúng tôi đã phân tích kỹ lưỡng 6 vị trí đa hình đã được xác định trên gene *StGBSS2* của giống Atlantic để đảm bảo chúng không ảnh hưởng đến hiệu quả gây đột biến. Bởi vì các SNP (Single Nucleotide Polymorphism) nằm trong vùng nhận biết của sgRNA hoặc chuỗi PAM (Protospacer adjacent motif) (5'-NGG-3') có thể làm giảm đáng kể hoặc hoàn toàn mất khả năng cắt của phức hợp CRISPR/Cas9 [17]. Hai trình tự crRNA được lựa chọn đều nằm ở vị trí bảo thủ trên tất cả các allele của *StGBSS2*, không chứa SNP trong cả vùng nhận biết và vùng PAM liên kề. Điều này đảm bảo khả năng nhận diện và cắt đồng thời cả 4 allele của gene đích, là yếu tố then chốt để tạo ra kiểu hình mong muốn trong giống khoai tây tứ bội. Đây chính là ưu thế vượt trội của công nghệ chỉnh sửa gene CRISPR/Cas9 so với các phương pháp cải tiến giống truyền thống khi làm việc với các giống khoai tây Atlantic có mức độ dị hợp tử cao và bộ gene tứ bội phức tạp.



Hình 5. Cấu trúc bậc 2 của sgRNA chỉnh sửa *StGBSS2*

Ghi chú: Cấu trúc bậc 2 của sgRNA chứa trình tự crRNA5 (A) và crRNA6 (B) được dự đoán bằng công cụ Mfold v2.3 mang đầy đủ các cấu trúc vòng loop chính (Stem loop 2, 3 và RAR).

3.4. Thiết kế cấu trúc T-DNA chỉnh sửa gene *StGBSS2*



Hình 6. Kiểm tra vector pCas9/sgRNA-GBSS2

Ghi chú: (A) Kiểm tra vector tái tổ hợp bằng PCR. Giếng M: thang chuẩn DNA 1,0 kb (Thermo Scientific); giếng 1, 4, 7: sản phẩm PCR với khuôn là vector tái tổ hợp; giếng 2: đối chứng dương với khuôn là pEN/sgRNA-GBSS2; giếng 5, 8: đối chứng dương với khuôn là pCas9; giếng 3, 6, 9: đối chứng âm (không có DNA khuôn); giếng 1-3: phản ứng sử dụng môi LguI-crRNA1-F/BsaI-crRNA2-R; giếng 4-6: phản ứng sử dụng môi HPT-F/HPT-R; giếng 7-9: phản ứng sử dụng môi Ubi-t-F/Cas9-R. (B) Giải trình tự pCas9/sgRNA-GBSS2. Mũi tên nét đứt thể hiện trình tự promoter AtU6.1 và AtU6.2; mũi tên nét liền thể hiện trình tự crRNA-1 và crRNA-2; đường nét đứt thể hiện trình tự tracrRNA.

Để tạo vector tái tổ hợp mang cấu trúc biểu hiện sgRNA (gọi tắt là pEN/sgRNA-GBSS2), crRNA1-GBSS2 và crRNA2-GBSS2 đã được ghép nối vào cấu trúc biểu hiện sgRNA trên vector pENTR4-V5 (Hình 1). Trình tự chứa 2 cấu trúc biểu hiện sgRNA sau đó được tái tổ hợp vào vector pCas9 nhờ LR clonase. Plasmid tái tổ hợp sau đó được kiểm tra bằng 3 cặp môi đặc hiệu cho gene chọn lọc *HPT*, cấu trúc biểu hiện Cas9 và cấu trúc biểu hiện sgRNA nhắm gene đích *StGBSS2* (Hình 6A). Hình ảnh điện di cho thấy tất cả sản phẩm PCR thu được đều có kích thước phù hợp với kích thước theo tính toán lý thuyết của các đoạn DNA cần nhân bản. Phân tích trình tự vector tái tổ hợp (Hình 6B) cho thấy crRNA1-GBSS2 và crRNA2-GBSS2 đã được chèn đúng vào vị trí giữa promoter *U6* và tracrRNA của cấu trúc biểu hiện sgRNA và không xảy ra hiện tượng đứt gãy tại các vị trí ghép nối. Các kết quả này chứng tỏ vector chuyển gene biểu hiện phức hệ CRISPR/Cas9 nhắm gene đích *StGBSS2* của giống khoai tây Atlantic pCas9/sgRNA-GBSS2 đã được thiết kế thành công. Hiệu quả chỉnh sửa gene của hệ thống sgRNA được điều khiển bởi promoter *AtU6* (phân lập từ *Arabidopsis*) đã được chứng minh bởi các nghiên cứu chỉnh sửa gene khoai tây gần đây [1], [19]. Vì vậy, trong nghiên cứu này, chúng tôi sử dụng promoter *AtU6* để

thiết kế cấu trúc biểu hiện sgRNA chỉnh sửa exon 2 của *StGBSS2* trên giống khoai tây Atlantic. Cấu trúc biểu hiện gene mã hoá protein Cas9 được sử dụng trong nghiên cứu này cũng cho thấy khả năng hoạt động hiệu quả trong nghiên cứu chỉnh sửa gene trên cây lúa trước đây của nhóm chúng tôi [21]. Việc thiết kế vector chỉnh sửa gene trong nghiên cứu này tập trung vào quá trình xây dựng cấu trúc phân tử, bao gồm lựa chọn vị trí đích trên gene *StGBSS2*, thiết kế sgRNA có cấu trúc bậc 2 ổn định và đặc hiệu cao, cũng như lắp ráp thành công các thành phần chức năng vào T-DNA. Mặc dù chưa có kết quả thực nghiệm *in vivo* về hiệu quả gây đột biến trên cây khoai tây, việc xác nhận các thiết kế phân tử này đã thành công (nguyên vẹn trình tự, đúng vị trí ghép nối) là bước quan trọng đầu tiên trong quy trình phát triển công cụ chỉnh sửa gene. Hơn nữa, các thành phần được sử dụng trong vector của chúng tôi như promoter *AtU6*, protein Cas9 và phương pháp thiết kế sgRNA bằng CRISPR-P đã được nhiều nghiên cứu trước đây chứng minh hiệu quả, gợi ý khả năng hoạt động của cấu trúc này [21] - [24].

Để tiến tới ứng dụng thực tế, các nghiên cứu tiếp theo sẽ tập trung vào chuyển cấu trúc vector đã thiết kế vào giống khoai tây Atlantic, đánh giá hiệu suất hoạt động *in vivo*, tối ưu hóa quy trình chuyển gene và sàng lọc dòng khoai tây mang đột biến mong muốn. Những kết quả này sẽ làm cơ sở cho việc tạo giống khoai tây bất hoạt gene *StGBSS2*, từ đó làm sáng tỏ vai trò phức tạp của các enzyme tổng hợp tinh bột, hướng đến mục tiêu cải tiến chất lượng tinh bột cho giống khoai tây thương mại Atlantic. Nghiên cứu hiện tại đã cung cấp công cụ thiết yếu và là tiền đề cho các nghiên cứu tiếp theo.

4. Kết luận

Đoạn trình tự chứa exon 2 gene *StGBSS2* của giống khoai tây Atlantic đã được phân lập và giải trình tự, giống > 99,8% so với trình tự tham chiếu trên GenBank (mã số NM_001288048.1). Hai trình tự sgRNA đặc hiệu cho *StGBSS2* của giống Atlantic dài 22 nucleotide đã được thiết kế bằng công cụ tin sinh, có cấu trúc bậc 2 ổn định và khả năng hoạt động đặc hiệu cao. Vector chuyển gene thực vật biểu hiện phức hệ CRISPR/Cas9 chỉnh sửa *StGBSS2* của giống khoai tây Atlantic đã được tạo thành công, chứa đầy đủ các cấu trúc biểu hiện gene cần thiết. Kết quả nghiên cứu là tiền đề cho nghiên cứu chức năng *StGBSS2*, hướng tới phục vụ chọn tạo giống khoai tây có hàm lượng amylose thấp bằng công nghệ gene.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] A. Tuncel and Y. Qi, "CRISPR/Cas mediated genome editing in potato: Past achievements and future directions," *Plant Sci.*, vol. 325, 2022, Art. no. 111474.
- [2] H. W. Maurer, "Starch in the Paper Industry," in *Starch* (Third Edition), J. BeMiller and R. Whistler (Eds.), Academic Press, San Diego, 2009, pp. 657-713.
- [3] M. Jiao, Y. Gao, and Y. Tian, "Study the relationship between the microstructure and characteristics of quinoa starch by compared with common cereal starches," *Am. J. Biochem. Biotechnol.*, vol. 16, no. 4, pp. 561-567, 2020.
- [4] H. Yangcheng, H. Jiang, M. Blanco, *et al.*, "Characterization of normal and waxy corn starch for bioethanol production," *J. Agric. Food Chem.*, vol. 61, pp. 379-386, 2013.
- [5] C. Tong, Z. Ma, H. Chen, *et al.*, "Toward an understanding of potato starch structure, function, biosynthesis, and applications," *Food Front.*, vol. 4, no. 3, pp. 980-1000, 2023.
- [6] X. Zhao, M. Andersson, and R. Andersson, "Resistant starch and other dietary fiber components in tubers from a high-amylose potato," *Food Chem.*, vol. 251, pp. 58-63, 2018.
- [7] F. Nazarian-Firouzabadi and R. G. F. Visser, "Potato starch synthases: Functions and relationships," *Biochem. Biophys. Rep.*, vol. 10, pp. 7-16, 2017.
- [8] F. Veillet, L. Chauvin, M. P. Kermarrec, *et al.*, "The *Solanum tuberosum* GBSSI gene: A target for assessing gene and base editing in tetraploid potato," *Plant Cell Rep.*, vol. 38, pp. 1065-1080, 2019.
- [9] M. Andersson, H. Turesson, A. Nicolai, *et al.*, "Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts," *Plant Cell Rep.*, vol. 36, pp. 117-128, 2017.

- [10] I. E. Johansen, Y. Liu, B. Jørgensen, *et al.*, “High efficacy full allelic CRISPR/Cas9 gene editing in tetraploid potato,” *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 17715.
- [11] H. Kusano, A. Takeuchi, and H. Shimada, “Efficiency of potato genome editing: Targeted mutation on the genes involved in starch biosynthesis using the CRISPR/dMac3-Cas9 system,” *Plant Biotechnol. (Tokyo)*, vol. 40, no. 3, pp. 201-209, 2023.
- [12] J. J. Doyle and J. L. Doyle, “Isolation of plant DNA from fresh tissue,” *Focus*, vol. 12, no. 1, pp. 13-15, 1990.
- [13] J. Sambrook and D. W. Russel, *Molecular cloning: A laboratory manual*, 3rd edition. Cold Spring Harbour Laboratory, Cold Spring Harbour, NY, 2001.
- [14] H. Liu, Y. Ding, Y. Zhou, W. Jin, *et al.*, “CRISPR-P 2.0: an improved CRISPR/Cas9 tool for genome editing in plants,” *Mol. Plant*, vol. 10, no. 3, pp. 530-532, 2017.
- [15] M. Zuker, “Mfold web server for nucleic acid folding and hybridization prediction,” *Nucleic Acids Res.*, vol. 31, no. 13, pp. 3406-3415, 2003.
- [16] M. Stemmer, T. Thumberger, M. D. Keyer, *et al.*, “CCTop: An intuitive, flexible and reliable CRISPR/Cas9 target prediction tool,” *PLoS One*, vol. 10, no. 4, 2015, Art. no. 0124633.
- [17] G. Liang, H. Zhang, D. Lou, *et al.*, “Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing,” *Scientific Rep.*, vol. 6, 2016, Art. no. 21451.
- [18] I. Dry, A. Smith, A. Edwards, *et al.*, “Characterization of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organs of pea and potato,” *The Plant J.*, vol. 2, no. 2, pp. 193–202, 1992.
- [19] L. Abeuova, B. Kali, D. Tussipkan, *et al.*, “CRISPR/Cas9-mediated multiple guide RNA-targeted mutagenesis in the potato,” *Transgenic Res.*, vol. 32, no. 5, pp. 383–397, 2023.
- [20] S. Toinga-Villafuerte, M. I. Vales, J. M. Awika, *et al.*, “CRISPR/Cas9-Mediated Mutagenesis of the Granule-Bound Starch Synthase Gene in the Potato Variety Yukon Gold to Obtain Amylose-Free Starch in Tubers,” *Int. J. Mol. Sci.*, vol. 23, no. 9, 2022, Art. no. 4640.
- [21] D. P. Nguyen, L. D. Tran, T. H. Pham, *et al.*, “Improved bacterial leaf blight disease resistance in the major elite Vietnamese rice cultivar TBR225 via editing of the *OsSWEET14* promoter,” *PLoS One*, vol. 16, no. 9, 2021, Art. no. e0255470.
- [22] G. F. Schramma, C. W. Schmid-Burgk, and L. F. Gootenberg, “A screening method to identify efficient sgRNAs in Arabidopsis,” *Sci. Rep.*, vol. 9, no. 1, pp. 1-8, 2019.
- [23] J. Li, S. Zhang, R. Zhang, *et al.*, “Efficient multiplex genome editing by CRISPR/Cas9 in common wheat,” *Plant Biotechnol. J.*, vol. 19, no. 3, pp. 427-429, 2021.
- [24] X. Jin, Z. Li, P. Xie, *et al.*, “Genome editing in cotton with the CRISPR/Cas9 system,” *Front. Plant Sci.*, vol. 11, pp. 1-9, 2020.