



KHOA HỌC KỸ THUẬT Thú y

JOURNAL OF VETERINARY SCIENCE AND TECHNOLOGY



ẤT TỴ

ISSN 1859 - 4751

Chúc Mừng
Năm Mới
HAPPY NEW YEAR

Quân
2025

Tập XXXII • Số 1 - 2025

HỘI THÚ Y VIỆT NAM
VIETNAM VETERINARY ASSOCIATION

TỐI ƯU QUÁ TRÌNH NHÂN DÒNG VÀ GIẢI TRÌNH TỰ MỘT SỐ GEN ĐIỀU HOÀ TẠO CƠ Ở BÒ VÀNG VIỆT NAM

*Đào Văn Cường^{1,2}, Từ Trung Kiên¹, Phạm Diệu Thùy¹,
 Dương Thị Hồng Duyên¹, Nguyễn Quang Tính¹, Nguyễn Hữu Hoà¹,
 Nguyễn Thị Ngân¹, Phan Thị Hồng Phúc¹, Lưu Hồng Sơn³,
 Nguyễn Thanh Hải⁴, Nguyễn Thị Thanh Hà⁵, Nguyễn Ngọc Lương^{6*}*
**Tác giả liên hệ email: luongnguyenbio@hueuni.edu.vn*

TÓM TẮT

Bò vàng là một giống bò bản địa tại Việt Nam và hiện là đối tượng được quan tâm bảo tồn nguồn gen. Mặc dù là giống bò lấy thịt, khả năng cho thịt và chất lượng thịt của bò vàng chưa tốt so với một số giống bò nhập ngoại. Nghiên cứu trước đây cho thấy có sự biểu hiện khác biệt của một số gen tham gia quá trình tạo cơ gồm MRF4, MYF5, MYOG, và MYOD1 ở tế bào vệ tinh bò vàng và tế bào vệ tinh bò Hanwoo của Hàn Quốc. Để làm sáng tỏ cơ chế biểu hiện khác biệt cần có được trình tự đầy đủ của các gen nói trên, bao gồm vùng điều hoà promoter và terminator. Nghiên cứu này báo cáo quá trình nhân dòng, giải trình tự và phân tích biến dị của bốn gen nói trên ở tế bào vệ tinh bò vàng Việt Nam nuôi cấy in vitro. Về nhân dòng, khảo sát bốn enzyme DNA Polymerase thương mại cho thấy GXL của Takara cho kết quả ổn định nhất. Về giải trình tự, phương pháp giải trình tự NGS có giá thành thấp hơn và cho kết quả tốt hơn giải trình tự bằng phương pháp Sanger theo kỹ thuật primerwalking. Kết quả này là tiền đề để hướng tới giải trình tự đầy đủ các gen thuộc gen panel tham gia vào quá trình tạo cơ và mỡ ở bò vàng Việt Nam.

Từ khóa: Bò vàng Việt Nam, tế bào vệ tinh, gen điều hoà phát triển cơ, nhân dòng, giải trình tự.

Cloning and sequencing of four genes involving muscle development in muscle satellite cells derived from vietnamese native yellow cow

SUMMARY

Vietnam yellow cattle, a collective term for various indigenous cattle breeds in Vietnam, are primarily raised for meat and are currently an important target for genetic conservation. Despite being raised for meat, their carcass performance and meat quality are considered inferior to some imported breeds. Previous studies showed differential expression of genes involved in muscle formation between muscle satellite cells derived from the yellow cattle and Korean Hanwoo. These genes include MRF4, MYF5, MYOG, and MYOD1. To assist in elucidating the mechanisms of differential expression of these genes, the full sequences of these genes, including promoters and terminators, are required. To this end, we explored methods to clone and sequence these genes. The result showed that Takara's GXL gave the most consistent result among the four commercial DNA Polymerases. For sequencing, Next-Generation-Sequencing technology using the Illumina platform is more economical and more sensitive than

¹. Khoa Chăn nuôi Thú y, trường Đại học Nông Lâm Thái Nguyên

². Văn phòng SPS Việt Nam, Bộ Nông nghiệp và Phát triển Nông thôn

³. Viện Công nghệ sinh học và Công nghệ thực phẩm, trường Đại học Nông Lâm Thái Nguyên

⁴. Khoa Công nghệ sinh học, Học viện Nông nghiệp Việt Nam

⁵. Khoa Thú y, Học viện Nông nghiệp Việt Nam

⁶. Trường Đại học Khoa học, Đại học Huế

traditional Sanger sequencing. The results serve as a baseline for sequencing the other genes in the muscle and fat formation gene panel.

Keywords: Vietnam yellow cow, muscle satellite cells, genes involved muscle formation, cloning, sequencing.

I. ĐẶT VẤN ĐỀ

Theo điều tra trước đây bò vàng Việt Nam là giống bò lai giữa bò vàng Trung Quốc (*Bos taurus*) và bò u Ấn Độ (*Bos indicus*) (Nguyễn Bình Trường và Phạm Ngọc Du, 2017). Giống bò này được nuôi lấy thịt, tuy nhiên năng suất cho thịt cũng như chất lượng thịt chưa cao (Nguyễn Thị Thuỷ và ctv., 2014). Việc người tiêu dùng Việt Nam gần đây được tiếp cận một số sản phẩm thịt bò nhập ngoại như bò Kobe/Wagyu (Nhật) và Hanwoo (Hàn Quốc), sẵn sàng chi trả giá cao cho các loại thịt bò này cho thấy người tiêu dùng ưa chuộng các loại thịt bò mềm và có mỡ giắt. Phân tích so sánh sự biểu hiện của một số gen điều hoà quá trình tạo cơ giữa bò vàng Việt Nam và bò Hanwoo cho thấy MRF4, MYF5, MYOD1 và MYOG biểu hiện mạnh hơn ở tế bào vệ tinh thu nhận từ bò Hanwoo, một giống bò có khả năng cho thịt cao và chất lượng thịt tốt nhờ lượng mỡ giắt lớn (Nguyen *et al.*, 2023). Do đó nhiều khả năng các gen này góp phần tạo nên sự khác biệt giữa chất lượng thịt bò Hanwoo và bò vàng Việt Nam. Một số nghiên cứu trước đây về các gen MRF4, MYF5, MYOD1 và MYOG cho thấy đa hình ở các gen này có thể ảnh hưởng đến chất lượng cũng như khả năng cho thịt của bò. Đặc biệt vị trí đa hình thường nằm ngoài vùng mã hoá, trong các vùng điều hoà như promoter và terminator (Li *et al.*, 2004; Chung & Kim, 2005; Wang *et al.*, 2011; Ujan *et al.*, 2011; Xue *et al.*, 2011; Guo *et al.*, 2013). Để làm sáng tỏ cơ chế tạo nên sự khác biệt về biểu hiện, cần phải phân tích sự sai khác về trình tự của các gen này giữa bò vàng và bò Hanwoo. Điều này đòi hỏi trình tự đầy đủ của bốn gen nói trên, đặc biệt phải bao gồm cả phần trình tự điều hoà gồm promoter và terminator. Mặc dù giải pháp triệt để nhất là giải trình tự toàn bộ bộ gen bò vàng Việt

Nam, chi phí và yêu cầu về kỹ thuật khiến điều này không dễ thực hiện. Một cách tiếp cận khác là tập trung giải trình tự nhóm đa gen (gen panel) liên quan đến một đặc điểm nhất định bằng cách làm giàu các gen này bằng PCR (Weber *et al.*, 2020).

MRF4, MYF5, MYOD1 và MYOG là những gen có kích thước tương đối lớn, mỗi gen có chiều dài (bao gồm cả trình tự điều hoà) khoảng từ 4-5 kb bao gồm cả vùng điều hoà. Nhằm mục đích phân tích biến dị cần thu nhận trình tự của các gen này với số lượng lỗi tạo tác, bao gồm lỗi nhân dòng và giải trình tự, là tối thiểu. Kỹ thuật PCR truyền thống sử dụng các enzyme Taq Polymerase sẽ không phù hợp trong trường hợp này do tỉ lệ lỗi cao. Long-range PCR là những kỹ thuật PCR cho phép tạo dòng các trình tự >4 kb trở lên, sử dụng những enzyme có khả năng khuếch đại trình tự lớn, đồng thời có tỉ lệ lỗi khuếch đại thấp (Jia *et al.*, 2014; Kubo *et al.*, 2015). Trong nghiên cứu này, nhằm xác định điều kiện tối ưu để nhân dòng bốn gen nói trên, chúng tôi khảo sát một số enzyme thương mại có khả năng đọc lỗi và khuếch đại được sản phẩm kích thước lớn. Để xác định trình tự của sản phẩm khuếch đại, chúng tôi khảo sát giải trình tự và phân tích trình tự bốn gen bằng hai phương pháp: phương pháp truyền thống Sanger sử dụng kỹ thuật primer-walking và phương pháp giải trình tự thế hệ mới (NGS) thông qua hệ thống Illumina. Báo cáo này trình bày kết quả khảo sát điều kiện tối ưu để khuếch đại, giải trình tự và phân tích bốn gen MRF4, MYF5, MYOD1 và MYOG. Kết quả sẽ là tiền đề tiếp tục giải trình tự đầy đủ những gen dài hơn thuộc nhóm gen tham gia điều hoà quá trình tạo cơ và mỡ ở bò vàng Việt Nam, đồng thời làm nền tảng cho các nghiên cứu gen panel của các giống vật nuôi quan trọng tại Việt Nam trong tương lai.

II. VẬT LIỆU PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Vật liệu

Là tế bào vệ tinh bò vàng Việt Nam do phòng thí nghiệm chuyên sâu, Viện công nghệ sinh học và Công nghệ thực phẩm, Trường Đại học Nông Lâm, Đại học Thái Nguyên cung cấp.

2.2. Phương pháp nghiên cứu

2.2.1. Tách chiết DNA tổng số

Từ nguyên liệu là chai nuôi T-flask chứa tế bào vệ tinh, tiến hành tách chiết DNA tổng số bằng kit FavorPrep™ Tissue Genomic DNA Extraction kit (Favorgen, cat No FATGK 001-2) theo hướng dẫn của nhà sản xuất. DNA tổng số được định lượng bằng máy đo quang phổ UV/Vis ở bước sóng 260/280, sau đó kiểm tra lại một lần nữa bằng điện di trên gel Agarose 0,8%.

2.2.2. Thiết kế mồi

Trình tự của MRF4 (NC_307332.1), Myf5 (NC_307332.1), MyoD1 (NC_37342.1) và Myog (NC_037343.1) được thu nhận từ cơ sở dữ liệu NCBI Gene. Để đảm bảo trình tự thu nhận chứa vùng promoter và terminator, tiến hành mở rộng về hai phía của trình tự khoảng 1.5-2 kb. Sử dụng phần mềm Primer BLAST (Ye *et al.*, 2012). Để đảm bảo tính đặc hiệu của mồi thiết kế, tiến hành thí nghiệm PCR in silico tại UCSC In-silico PCR (<https://genome.ucsc.edu/cgi-bin/hgPcr>) và chọn các cặp mồi đặc hiệu nhất. Khảo sát từ 2-3 cặp mồi cho mỗi gen.

2.2.3. Nhân dòng bằng kỹ thuật Long-range PCR

Phản ứng Long-range PCR sử dụng các cặp mồi đặc hiệu đã chọn và enzyme DNA Polymerase có khả năng khuếch đại sản phẩm lớn (bao gồm 4 loại enzyme là Phusion™ High Fidelity DNA Polymerase,

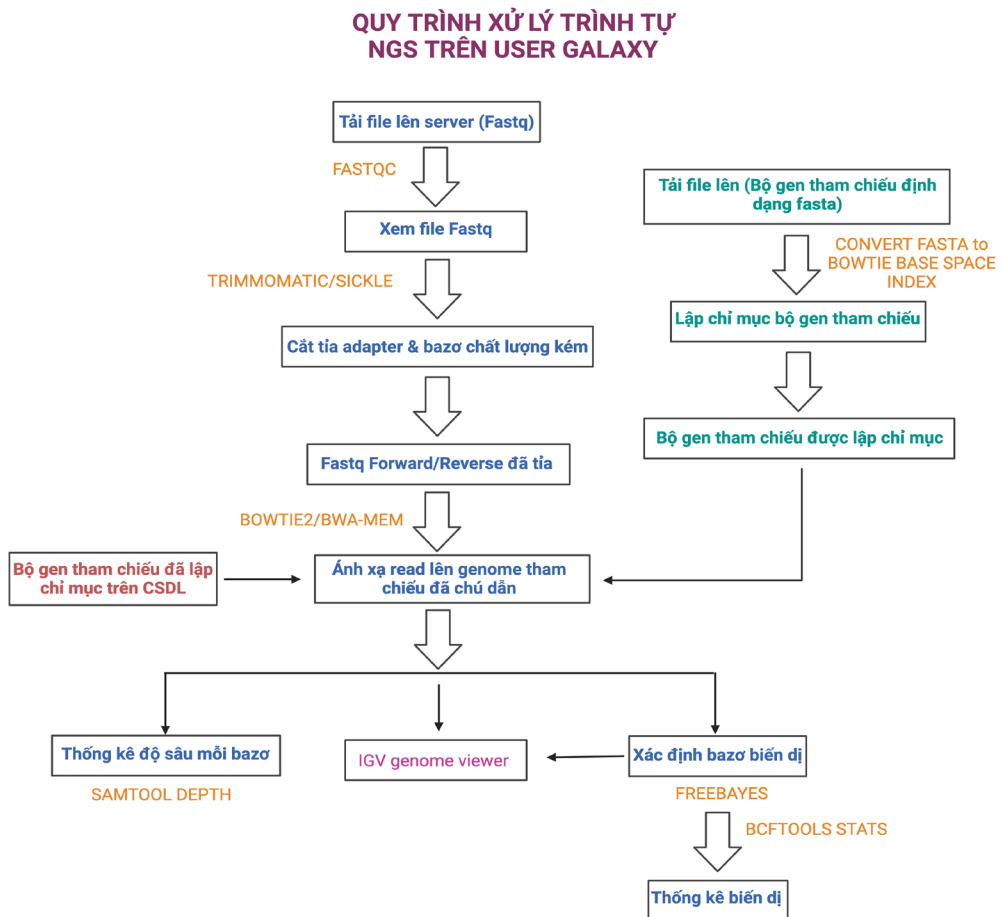
Plantinum™ Superfi™ DNA Polymerase (ThermoFisher scientific, cat No lần lượt là F530S và 12351010), Q5® của NEB (cat No M0491S) và PrimeSTAR GXL của Takara (cat No R050A). Hỗn hợp phản ứng chứa 1 mỗi xuôi và 1 mỗi ngược, khoảng 60-100 ng DNA tổng số làm khuôn mẫu (~2 , đậm phản ứng (5 , dNTP (~2 μ l, **nồng độ 200 μ M**), và 1 đơn vị enzyme DNA-Polymerase. Quy trình phản ứng PCR gồm các bước như sau: 2 phút biến tính ở 98°C, tiếp đến là 30 chu kỳ gồm 98°C trong 30 giây, 58-60°C trong 30 giây, 72°C với 30 giây/1 kb cho ba enzyme Phusion, Superfi, Q5 và 1 phút/1 kb cho GXL, 72°C trong 10 phút và cuối cùng là 4°C trong 10 phút. Nếu không xuất hiện băng đặc hiệu, giảm nhiệt độ gắn mồi xuống khoảng 2°C. Nếu xuất hiện băng tạp nhiễm cùng băng đặc hiệu, tăng nhiệt độ gắn mồi lên khoảng 2°C.

2.2.4. Giải trình tự và phân tích kết quả

Để giải trình tự sản phẩm PCR, chúng tôi thử nghiệm hai phương pháp là giải trình tự Sanger các đoạn chồng lấp để nối chúng lại với nhau (còn gọi là primer walking) và giải trình tự thế hệ mới (NGS) bằng hệ thống Illumina. Trước tiên, tiến hành làm giàu mẫu bằng cách tủa DNA và điều chỉnh nồng độ mỗi gen đến nồng độ bằng nhau (25-39 ng/ml). Đối với giải trình tự Sanger, tiến hành đặt các cặp mồi giải trình tự sao cho các trình tự nằm cạnh nhau chồng lấp tối thiểu 200-300 nucleotide. Giải trình tự cho từng gen riêng rẽ ở công ty FirstBase (Malaysia). Sau khi giải xong, tiến hành tia vùng trình tự xấu và lắp ráp các trình tự thành contig. Việc phân tích biến dị đối với giải trình tự Sanger được tiến hành bằng phần mềm Snapgene phiên bản 7.0 bằng cách giống cột trình tự giải được với trình tự tham chiếu. Khi xác định được sai khác, kiểm tra kết quả bằng mắt thường các vị trí đa hình ở kết quả thô (trace file).

Đối với giải trình tự bằng hệ thống NGS Illumina, trộn 4 sản phẩm PCR của 4 gen với tỉ lệ bằng nhau và gửi đi giải trình tự trên hệ thống Illumina ở công ty K-test (Việt Nam). Phân tích kết quả giải trình tự NGS được tiến hành trên nền tảng Galaxy trực tuyến (<https://usegalaxy.eu/>). Các bước cơ bản gồm: đánh giá chất lượng

gọi bazơ (Q) tại mỗi vị trí bằng FastQ, sau đó tỉa các vị trí có chất lượng thấp bằng Trimmomatic/Sickle. Tiếp đến sử dụng Bowtie2/BWA-MEM để ánh xạ kết quả đọc lên trình tự tham chiếu. Sau khi tiến hành sắp xếp kết quả đọc theo toạ độ, tiến hành gọi biến dị (variance calling). Quy trình phân tích được tóm tắt ở Hình 1.



Hình 1. Sơ đồ tóm tắt kết quả giải trình tự Illumina Miseq 4 gen điều hoà phát triển cơ

Để phân tích biến dị giữa các gen tham gia điều hoà quá trình tạo cơ của bò vàng Việt Nam và bò Hanwoo Hàn Quốc, tiến hành sắp xếp gióng cột trình tự đại diện đã chú giải (annotated consensus sequences) với trình tự gen tương ứng ở bò Hanwoo Hàn Quốc thu nhận từ NCBI Nucleotide.

III. KẾT QUẢ VÀ THẢO LUẬN

3.1. Thiết kế và khảo sát môi

Thông tin về mã truy cập, cấu trúc 4 gen và trình tự các cặp mỗi đặc hiệu đã khảo sát để khuếch đại chúng được cung cấp ở Bảng 1. Chiều dài trong bảng là chiều dài gen. Chiều dài sản phẩm khuếch đại chứa thêm khoảng 1-2kb ở đầu 5' và 3' của gen. Đối với mỗi gen, chọn khoảng 2-3 cặp mỗi đặc hiệu để thử nghiệm với cả bốn loại enzyme.

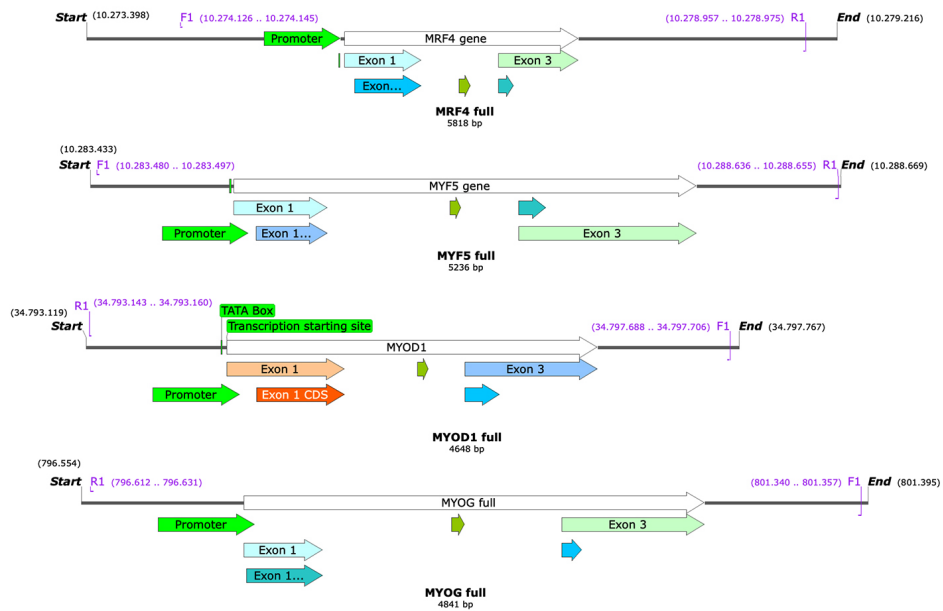
Bảng 1. Thông tin mã truy cập gen, số lượng exon/intron và trình tự mỗi khảo sát

Tên gen	Mã truy cập (NCBI gene)	Chiều dài (bp)	Số lượng exon/intron	Trình tự các cặp mỗi khảo sát (5'→3')
MRF4	281336	1818	3/2	F1: GACATAGCCTTATGCTTTCC R1: TCTCTGCTCCAAAGTATCC F2: CACACGCATGGGATCAGTTAATCG R2: AAGGTTGGGCTAGGGGATGCTG F3: GTTCTAGAGAGCACTGGGTGT R3: GCAGGGGAGTTTGTGTTCT
Myf5	281335	3236	3/2	F1: GTTAGTTTGCCTGAAAGC R1: TATCAAGTACATTTGCAAGC F2: CAAGCCCCTGGGAAGGAGTTCATC R2: TTGCAAGCAACTGGGGTTTGGTAG
Myod1	281938	2648	3/2	F1: CCTCTATGTGTGTGTGTCC R1: TGTTGGTCTGCAAACTCC F2: GCTGTGACGTGCTGTTATGG R2: TTGCAGATGGACTACAGCGG F3: TATGCAGGGGTGCACAAGTC R3: GACGGCTCTCTGCAACTT
Myog	281343	2841	3/2	F1: TCAAGCCTGTTGTGAAGG R1: TGGTGTCTAAATTGGGTAGG F2: TCTTTGCCAGCCAATGAAC R2: AGTGGGCTAGAATGCCCTGT F3: TCTACTCCACAGTCCATGCCT R3: CCATCCAGTACATAGAGCGCC

3.2. Khảo sát enzyme DNA Polymerase

Kết quả PCR sử dụng các cặp mỗi đặc hiệu này thường cho một sản phẩm đặc hiệu và duy nhất. Tuy nhiên đối với 3 loại enzyme DNA Polymerase là Phusion, Plantinum và Q5, luôn cần phải điều chỉnh một trong các thông số và điều kiện PCR, ví dụ điều chỉnh nhiệt độ gắn mỗi, lượng khuôn và thời gian kéo dài. Ngược lại, GXL luôn cho kết quả đặc hiệu và ổn định ở nhiệt độ gắn mỗi 58°C cho cả 4 gen. Tính ổn định và khả năng khuếch đại sản phẩm kích thước lớn được sắp xếp theo thứ tự là: GXL → Plantinum™ → Phusion → Q5. Jia *et al.*, 2014 tiến hành khảo sát 4 loại enzyme thương mại khác nhưng cũng đi đến kết luận tương tự là GXL cho kết quả ổn định nhất. Hình 2 trình bày bản đồ 4 gen mô tả vị trí gắn mỗi của một trong các cặp mỗi đặc hiệu đã chọn. Hình được vẽ bằng phần mềm Snapgene phiên bản 7 dựa trên thông tin vị trí intron, exon ở CSDL NCBI Gene và dự đoán vùng promoter và

terminator từ CSDL promoter sinh vật nhân chuẩn (<https://epd.expasy.org/epd/>) và định nghĩa của terminator. Mỗi được chọn để thể hiện trong hình được bôi đậm ở Bảng 1. Hình 3 trình bày kết quả điện di sản phẩm PCR khi khảo sát 4 enzyme DNA Polymerase. Hình 3A là kết quả PCR cho sản phẩm đặc hiệu đối với cả 4 gen khi sử dụng enzyme Plantinum™ SuperFi™. Hình 3B là hình đại diện kết quả PCR sử dụng enzyme Q5 hoặc Phusion ở điều kiện chưa tối ưu nhiệt độ gắn mỗi và điều kiện phản ứng PCR, dẫn đến không cho sản phẩm hoặc sản phẩm không đặc hiệu. Hình 3C là kết quả PCR sử dụng enzyme GXL cho sản phẩm đặc hiệu đối với cả 4 gen. Đặc biệt không thể hiện trong hình là sản lượng sản phẩm khuếch đại của GXL cao hơn 1,5 lần so với các enzyme DNA Polymerase còn lại. Thứ tự của mẫu trên các gel là: mẫu số 1: MRF4; mẫu số 2: MYF5; mẫu số 3: MYOD1; mẫu số 4: MYOG.



Hình 2. Bản đồ bốn gen tham gia điều hoà phát triển cơ ở bò (vẽ bằng phần mềm Snapgene phiên bản 7 theo chú giải trên CSDL NCBI gene)

3.3. Giải trình tự và phân tích kết quả

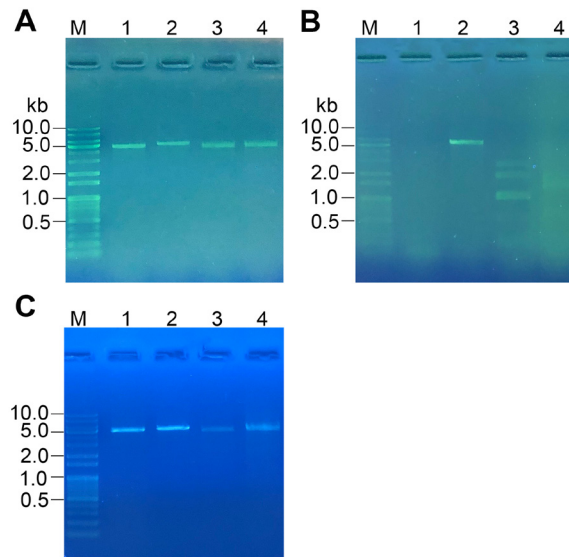
Kết quả giải trình tự Sanger cho 4 gen cho thấy một số vùng của các gen này không thể giải được do giàu GC hoặc các lý do kỹ thuật khác như mỗi bắt cặp không tốt với khuôn. Trình tự lắp ráp của 4 gen giải theo phương pháp Sanger được gửi đi đăng ký ở NCBI Genbank với mã truy cập lần lượt là: MRF4 (OR554977), MYF5 (OR631211), MYOD1 (OR554978), MYOG (OR554979). Chiều dài giải được và số vị trí đa hình của 4 gen giải trình tự theo phương pháp Sanger được thống kê ở Bảng 2. Kết quả giải trình tự NGS cho trình tự dài hơn kết quả giải trình tự Sanger và đồng thời cho nhiều thông tin hơn về biến dị. Bảng 2 tóm tắt một số thông số so sánh kết quả giải trình tự Sanger bằng phương pháp primer walking và NGS Illumina. Trình tự NGS thô được đăng ký ở ngân hàng GenBank SRA với mã truy cập là PRJNA1107566.

Nhằm góp phần làm sáng tỏ cơ sở di truyền học của sự biểu hiện khác biệt giữa các gen điều hoà phát triển cơ ở bò vàng Việt Nam và bò Hanwoo, chúng tôi tiến hành so sánh trình tự 4

gen bò vàng giải được bằng phương pháp NGS với trình tự 4 gen bò Hanwoo công bố trên ngân hàng gen NCBI Nucleotide. Mã truy cập của các nhiễm sắc thể chứa trình tự các gen điều hoà phát triển cơ của bò Hanwoo là CM054874.2 (NST số 5 chứa MRF4 và MYF5), CM054884.2 (NST số 15 chứa MYOD1) và CM054885.2 (NST số 16 chứa MYOG). Kết quả cho thấy có một số sai khác đáng chú ý, xứng đáng được nghiên cứu kỹ hơn để làm rõ mối quan hệ với sự khác biệt về biểu hiện gen cũng như tính trạng cho thịt và chất lượng thịt. Ví dụ một đột biến sai nghĩa ở exon số 1 (183T → C) ở bò Hanwoo khiến một axit amin thay đổi từ cysteine thành arginine ở MRF4; đa hình ở intron số 2 324A → G của MYF5 liên quan đến khả năng cho thịt ở bò Hanwoo đã được Chung và Kim báo cáo trước đây (Chung & Kim, 2005); đột biến sai nghĩa ở exon số 1 567C → T của MYOD1 ở bò Hanwoo khiến arginine trở thành leucine; và đột biến sai nghĩa ở exon số 3 429G → A của MYOG ở bò Hanwoo khiến proline trở thành leucine. Đây là những sai khác về trình tự giữa các gen bò vàng Việt Nam và bò Hanwoo nằm trong vùng mã hoá, hoặc đột biến

được ghi nhận bởi các nghiên cứu trước đây, do đó chúng có nhiều khả năng góp phần tạo nên sự

sai khác về tính trạng giữa bò vàng Việt Nam và bò Hanwoo Hàn Quốc.



Hình 3. Kết quả khảo sát enzyme DNA Polymerase. Hình 3A đại diện cho kết quả khuếch đại 4 gen bằng Plantium™ SuperFi™; 3B đại diện cho kết quả khuếch đại 4 gen bằng Q5 hoặc Phusion™; 3C đại diện cho kết quả khuếch đại 4 gen bằng GXL (xem thêm phần Kết quả 3.2 để biết thêm chi tiết).

Bảng 2. Phân tích một số thông số của phương pháp giải trình tự Sanger và NGS đối với 4 gen điều hoà tạo cơ ở bò.

	SANGER			NGS		
	Chiều dài giải được/ lý thuyết	Tổng số đa hình xác định được	Giá thành (tr. Đồng)	Chiều dài giải được/ lý thuyết	Tổng số đa hình xác định được	Giá thành (tr. Đồng)
MRF4	3035/5001	20	1,8	4745/5001	30	0,6
MYF5	3761/5176	19	2,2	5175/5176	45	0,6
MYOD1	4404/4564	17	2,6	4564/4564	26	0,6
MYOG	3625/4746	10	2,4	4617/4746	49	0,6

*Giá thành mỗi gen được tính bằng giá tổng chia cho số gen, do đó số gen càng lớn giá thành càng giảm.

Bảng 2 cho thấy giải trình tự NGS có ưu thế về giá thành cũng như khả năng phân tích sâu các biến dị so với giải trình tự Sanger. NGS cho phép giải trình tự các vùng giàu GC hoặc vùng lặp, khiến việc bắt cặp mỗi trở nên khó khăn trong giải trình tự Sanger. Việc phân tích kết quả NGS thường khó khăn đối với các nhà Sinh học, tuy nhiên các công cụ trên nền tảng Galaxy trực tuyến có thể giúp đơn giản hoá quá trình phân

tích. Bước đầu phân tích 4 gen nói trên đã phát hiện một số đa hình được báo cáo trước đây có liên quan đến chất lượng thịt cũng như khả năng cho thịt của bò. Cụ thể, đã xác định được 3 vị trí đa hình liên quan đến độ dày mỡ và độ mềm phần thịt lưng (Ujan *et al.*, 2011), trọng lượng bò con mới sinh và tốc độ sinh trưởng khi cai sữa (Li *et al.*, 2004) và chất lượng thịt thăn lưng ngoại (sirloin) (Robakowska-Hyzorek

et al., 2010) đã được báo cáo trước đây. Bên cạnh đó chúng tôi tìm thấy một số lỗi đa hình MRF4 được báo cáo trước đây bởi Du *et al.* 2013 và Wang *et al.*, 2011 nhờ có được trình tự đầy đủ hơn bằng phương pháp NGS. Cụ thể Du *et al.* báo cáo đa hình T186C ở vị trí exon 1 của MRF4 tuy nhiên căn cứ theo thông tin cập mới nhóm tác giả cung cấp, vị trí này nằm ở exon 3 và không chứa đột biến như nhóm tác giả đã báo cáo. Wang *et al.* báo cáo đa hình vị trí cắt hạn chế ở vùng promoter của MRF4, tuy nhiên chúng tôi chỉ phát hiện một kiểu gen duy nhất tại vị trí này thay vì 3 kiểu gen như nhóm tác giả đã báo cáo.

Giải trình tự Sanger trong nghiên cứu này xác định được ít biến dị hơn do chỉ giải được trình tự một phần gen đầy đủ, đặc biệt không thể xác định tốt trình tự nằm ở vùng điều hoà là nơi thường mang nhiều biến dị. Trong khi đó giải trình tự NGS xác định được số lượng biến dị nhiều hơn và đồng thời nhiều biến dị dị hợp tử được xác định hơn so với giải trình tự Sanger. Điều này cho thấy tính ưu việt của NGS trong việc xác định biến dị, đặc biệt khi ta quan tâm đến số lượng biến dị hơn là chất lượng của biến dị. Mặc dù giải trình tự Sanger thường được xem là tiêu chuẩn vàng trong giải trình tự DNA để xác định biến dị, thông thường nó thường được dùng để kiểm tra biến dị biết trước thay vì dùng để tìm các biến dị mới (Pirooznia *et al.*, 2014). Một nghiên cứu gần đây cho thấy các phương pháp giải trình tự NGS gồm Nanopore và PacBio cho kết quả ưu việt hơn Sanger và đặc biệt là không cần sự có mặt của trình tự tham chiếu (Nolan *et al.*, 2024).

IV. KẾT LUẬN

Nghiên cứu đa hình gen để làm sáng tỏ cơ chế phân tử của các tính trạng ở vật nuôi và cây trồng là một kỹ thuật rất thông dụng. Tuy nhiên đối với các gen có kích thước lớn, nhiều rào cản đặt ra với nhà khoa học, bao gồm nhân dòng gen, giải trình tự và phân tích biến dị. Trong báo cáo này chúng tôi trình bày quy trình tối ưu để nhân dòng, giải trình tự và phân tích biến dị của

4 gen MRF4, MYF5, MYOD1 và MYOG ở tế bào vệ tinh bò vàng Việt Nam. Kết quả khảo sát cho thấy để tạo dòng các gen kích thước > 3 kb, enzyme GXL của Takara cho kết quả ổn định nhất, phù hợp với kết luận của một nghiên cứu trước đây. Đối với giải trình tự và phân tích biến dị, kỹ thuật giải trình tự NGS có ưu thế vượt trội so với Sanger về độ phủ (chiều dài trình tự giải được và độ sâu) và giá thành. Bên cạnh đó NGS cũng thể hiện độ nhạy hơn Sanger trong việc phát hiện các biến dị dị hợp tử. Sử dụng nền tảng Galaxy có thể giúp đơn giản hoá quá trình phân tích trình tự NGS.

Lời cảm ơn: Nhóm tác giả trân trọng cảm ơn: i) Bộ Khoa học và Công nghệ Việt Nam (Mã số nhiệm vụ Nghị định thư: NĐT.91.KR/20); ii) Phòng thí nghiệm Công nghệ Sinh học, Trường Đại học Khoa học, Đại học Huế đã tạo điều kiện để thực hiện một số thí nghiệm liên quan đến nghiên cứu này.

TÀI LIỆU THAM KHẢO

1. Nguyễn Bình Trường & Phạm Ngọc Du (2017). Một góc nhìn về bảo tồn giống bò vàng Việt Nam. *Tạp chí Khoa học Kỹ thuật Chăn nuôi*, 216: 95-97.
2. Nguyễn Thị Thuỷ, Đậu Văn Hải, Hoàng Thị Ngân, Phạm Văn Quyến, Nguyễn Thanh Tùng, Huỳnh Văn Thảo (2014). Khảo sát hàm lượng mỡ giắt, dày cơ thần và dày mỡ lưng của một số nhóm bò lai hướng thịt nuôi tại Trà Vinh. *Tạp chí Khoa học Công nghệ chăn nuôi*, 143: 57-65.
3. Nguyen T.U., Dao V.C., Pham D.T., Luu H.S., Nguyen T.N., Nguyen H.Q., Nguyen D.T., & Hwang I.H. (2023). A comparative study on the adipogenic and myogenic capacity of muscle satellite cells, and meat quality characteristics between Hanwoo and Vietnamese yellow steers. *Food Science and Animal Resources*, 18: 11-20.
4. Li C., Basarab J., Snelling W. M., Benkel B., Murdoch B., Hansen C., Moore S. S. (2004) Assessment of positional candidate

- genes myf5 and igf1 for growth on bovine chromosome 5 in commercial lines of *Bos taurus*. *Journal of Animal Science*, 1:1-7.
5. Du X. H., Gan Q. F., Yuan Z. R., Gao X., Zhang L. P., Gao H. J., Li J. Y., Xu S. Z. (2013) Polymorphism of MyoD1 and Myf6 genes and associations with carcass and meat quality traits in beef cattle. *Genetics and Molecular Research*, 12(4):6708-6717.
 6. Wang S., Cai X., Xue K., Chen H. (2011) Polymorphisms of MRF4 and H-FABP genes association with growth traits in Qinchuan cattle and related hybrids. *Molecular Biology Reports*, 38:1013-1020.
 7. Ujan J. A., Zan L. S., Ujan S. A., Adoligbe C., Wang H. B. (2011) Back fat thickness and meat tenderness are associated with a 526 T → A mutation in the exon 1 promoter region of the MyF-5 gene in Chinese *Bos taurus*. *Genetics and Molecular Research*, 10 (7):3070-3079.
 8. Xue M., Zan L. S., Gao L., Wang H. B. (2011) A novel polymorphism of the myogenin gene is associated with body measurement traits in native Chinese breeds. *Genetics and Molecular Research*, 10(4): 2721-2728.
 9. Chung E. R., Kim W. T. (2005) Association of SNP marker in IGF-I and MYF5 candidate genes with growth traits in Korean cattle. *Asian Australian Journal of Animal Science*, 18(8): 1061-1065.
 10. Weber T.M., Buseman B.J., Nasados J.A., Lancaster J.M., Van Buren J.B., Smart J.H., Bass P.D., Murdoch G.K., Insausti K., Colle M.J, 2020. Assessing outcomes of genetic selection panels to predict marbling in crossbred beef cattle. *Translational Animal Science*: 4: 1238 – 1246.
 11. Jia H., Guo Y., Zhao W., Wang K. (2014) Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Scientific Reports*, 4, 5737.
 12. Kubo Y., Nishio H., Saito K. (2015) A new method for SMN1 and hybrid SMN gene analysis in spinal muscular atrophy using long-range PCR followed by sequencing. *Journal of Human Genetics*, 60: 233-239.
 13. Ye J., Coulouris G., Zaretskaya I., Cutcutache I., Rozen S., Madden TL. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134.
 14. Robakowska-Hyzorek D., Oprzadek J., Zelazowska B., Olbromski R., Zwierzchowski L. (2010) Effect of the g.-723G-->T polymorphism in the bovine myogenic factor 5 (Myf5) gene promoter region on gene transcript level in the longissimus dorsi muscle and on meat traits of Polish Holstein-Friesian cattle. *Biochemical Genetics* 48(5-6):450-464.
 15. Pirooznia M., Kramer M., Parla J., Goes S. F., Potash J. B., McCombie W.R, Zandi P. P. (2014) Validation and assessment of variant calling pipelines for next- generation sequencing. *Human Genomics* 8(1):14.
 16. Nolan D. J., DaRoza J., Brody R., Ganta K., Luzuriaga K., Huston C., Rosenthal S., Lamers, Rose R. (2024) Comparing Gold-standard Sanger sequencing with two Next-Generation Sequencing Platforms of HIV-1 gp160 single genome amplicons. *AIDS research and human retroviruses*. Doi: 10.1089/AID.2024.0012.

Ngày nhận: 4-9-2024

Ngày phản biện: 4-9-2024

Ngày đăng: 1-1-2025