



▶ LÊ NỮ ANH THƯ

Lê Nữ Anh Thư tốt nghiệp kỹ sư chuyên ngành Công nghệ sinh học tại trường Đại học Khoa học, Đại học Huế năm 2009. Thạc sĩ Nông nghiệp năm 2014, Đại học Kyoto, Nhật Bản. Tiến sĩ Nông nghiệp năm 2020, Đại học Okayama Nhật Bản. Giáo sư trợ giảng (Special Assistant Professor) năm 2021 – 2022, Khoa Thú y, trường Đại học Khoa học Okayama, Nhật Bản. Từ năm 2009 – hiện tại là giảng viên Khoa Chăn nuôi Thú y, Đại học Nông Lâm, Đại học Huế.

ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ THỂ HỆ MỚI KHÁM PHÁ CÁC ĐỘT BIẾN DI TRUYỀN HIẾM LIÊN QUAN ĐẾN NĂNG SUẤT SINH SẢN Ở BÒ ĐEN NHẬT BẢN

Lê Nữ Anh Thư¹

¹ Giảng viên, Trường Đại học Nông Lâm, Đại học Huế
Email: lenuanhthu@huaf.edu.vn

TÓM TẮT:

Bò Đen là giống bò bản địa được nuôi phổ biến ở Nhật Bản và có giá trị kinh tế cao do sở hữu hàm lượng mỡ giết cao. Tuy nhiên, gần đây do năng suất sinh sản của bò Đen ngày càng giảm nên số lượng bê sinh ra giảm và giá bê gia tăng. Do vậy, cần có giải pháp để cải thiện năng suất sinh sản của bò Đen Nhật Bản vì nó là một trong những tính trạng kinh tế quan trọng nhất trong sản xuất chăn nuôi. Giải pháp hiệu quả cho vấn đề này là cải thiện các đặc điểm di truyền bằng phương thức xác định các alen nguy cơ ảnh hưởng đáng kể đến năng suất sinh sản. Vì vậy, mục tiêu của bài báo là giới thiệu những tiến bộ gần đây về ứng dụng công cụ di truyền phân tử để khám phá ra các đột biến hiếm liên quan đến tính trạng sinh sản ở bò đen Nhật Bản. Các kết quả nghiên cứu đã chỉ ra rằng ứng dụng GWAS và công nghệ giải trình tự thể hệ mới đã cho phép phát hiện ra nhiều đa hình nucleotide đơn (SNPs) và đột biến chèn/xóa (Indels) có tác động lớn đến năng suất sinh sản. Những biến dị di truyền hiếm này được đề nghị sử dụng làm chỉ thị phân tử để sàng lọc và cải thiện khả năng sinh sản ở bò đen Nhật Bản..

Từ khóa: Bò đen Nhật Bản, biến dị di truyền, giải trình tự thể hệ mới (NGS), năng suất sinh sản, SNPs/Indels.

1. GIỚI THIỆU

Sinh sản là một trong những tính trạng kinh tế quan trọng nhất quyết định sự thành công của sản xuất chăn nuôi [1]. Trong đó, tỷ lệ thụ thai là một trong những chỉ tiêu quan trọng đo lường năng suất sinh sản của bò [2]. Các yếu tố về sinh lý sinh sản và môi trường khác nhau có thể ảnh hưởng đến tỷ lệ thụ thai của bò. Thêm vào đó, quá trình hình thành giao tử và sự chết phôi là những yếu tố quan trọng ảnh hưởng đến tỷ lệ thụ thai. Diskin và cs (2012) ước tính tỷ lệ chết phôi ở bò sữa ~40% và ảnh hưởng trực tiếp của sự chết phôi được phản ánh qua tỷ lệ thụ thai thấp [3]. Trong những thập kỷ gần đây, tỷ lệ thụ thai ở bò đen Nhật Bản ngày càng giảm dần dẫn đến hiệu quả sản xuất bê thấp và giá bê tăng dần [4], [5] Vì vậy, việc cải thiện năng suất sinh sản bằng cách giảm tỷ lệ chết phôi là vấn đề quan trọng trong chăn nuôi bò đen Nhật Bản. Ở người, nhiều nghiên cứu cho thấy rằng tỷ lệ chết phôi cao ở thời kỳ đầu của thai kỳ, xấp xỉ 70% tổng số ca thụ thai. Đáng lưu ý, nguyên nhân gây tỷ lệ chết phôi thai ở người do các bất thường trong di truyền nhiễm sắc thể (NST). Yahaya và cs (2021) báo cáo rằng tỷ lệ 2-15% vô sinh ở nam và >10% vô sinh ở nữ đến từ nguyên nhân này [6]. Lee và Kieesling (2016) chỉ ra rằng khoảng 25% NST bất thường

về số lượng do những rối loạn trong quá trình phát sinh giao tử gây ra [7]. Có nhiều gen mã hóa các protein tham gia vào quá trình phát sinh giao tử bao gồm phân chia nguyên nhiễm và giảm nhiễm để tạo ra các giao tử bình thường và một số gen này đã được chứng minh là có liên quan đến khả năng sinh sản ở chuột và người [8]–[10]. Vì vậy, giải pháp hiệu quả để cải thiện khả năng sinh sản là khám phá các biến dị di truyền hiếm trong các gen liên quan đến năng suất sinh sản. Phương pháp tiếp cận gần đây ứng dụng công nghệ giải trình tự thế hệ mới (NGS) đã khám phá một số lượng lớn các biến dị di truyền bao gồm các đột biến đa hình nucleotide đơn (SNPs), các đột biến chèn/xóa đoạn nhỏ (Indels), và các đột biến chèn/xóa kích thước lớn (CNV) [11]. Nhờ tiến bộ của công nghệ NGS kết hợp với GWAS (nghiên cứu liên kết trên toàn bộ gen) đã góp phần phát hiện các đột biến hiếm gây ra rối loạn sinh sản và làm thay đổi những hiểu biết của chúng ta về đặc điểm của tính trạng phức tạp này. Vì vậy, trong bài báo này, tình hình về năng suất sinh sản của bò đen Nhật Bản, lợi ích của ứng dụng công nghệ NGS để phát hiện các biến dị di truyền hiếm liên quan đến năng suất sinh sản ở bò đen Nhật Bản sẽ được thảo luận nhằm làm sáng tỏ tác động của các biến dị này đến khả năng sinh sản ở bò, đồng thời thảo luận về những triển vọng của ứng dụng công nghệ này ở Việt Nam.

2. NĂNG SUẤT SINH SẢN CỦA BÒ ĐEN NHẬT BẢN

Bò Đen là giống bò được nuôi phổ biến nhất ở Nhật Bản, chiếm đến 97% tổng đàn bò nội địa. Thịt bò Đen nổi tiếng với hương vị thơm ngon đặc trưng và hàm lượng mỡ giết cao (>50%) [12]. Tuy nhiên năng suất sinh sản của bò Đen, một trong những tính trạng kinh tế quan trọng nhất trong sản xuất bò thịt, đã giảm dần trong những thập kỷ gần đây mặc dù các tính trạng kinh tế quan trọng khác như sinh trưởng và chất lượng thịt đã được cải thiện đáng kể [13]. Hiện trạng về năng suất sinh sản như tuổi phối giống lần đầu, số lần phối giống thành công, tỷ lệ phối giống thành công (tỷ lệ thụ thai) ở bò Đen Nhật Bản được thể hiện chi tiết qua bảng 1. Trong đó, tỷ lệ thụ thai bị ảnh hưởng bởi các yếu tố như lứa đẻ, thời gian phối lại giống thành công sau khi đẻ, số lần phối giống.

Bảng 1. Năng suất sinh sản của bò Đen Nhật Bản [14]

Chi tiêu	Bò mẹ M ± SEM
Lứa đẻ	4,8 ± 0,1
Tuổi phối giống lần đầu, ngày	-
Thời gian phối giống lại sau khi đẻ, ngày	77,0 ± 0,2
Tỷ lệ thụ thai, %	47,8 ± 0,2
Số lần phối giống	1,7 ± 0,01

(M: giá trị trung bình, SEM: sai số chuẩn của giá trị trung bình)



Hình 1. Bò đen Nhật Bản

Irikura và cs (2018) [5] đã báo cáo rằng tỷ lệ thụ thai của bò Đen giảm từ 55,1% xuống 34,4% khi lứa đẻ gia tăng. Thêm vào đó, khi thời gian phối lại giống lần đầu >90 ngày, tỷ lệ thụ thai ở lần phối giống thứ 4 (51,3%) giảm đáng kể so với tỷ lệ thụ thai ở lần phối giống lần đầu (23,8%). Sasaki và cs (2015) đã báo cáo rằng thời gian phối lại giống thành công sau khi đẻ của bò Đen là 115 ngày. Tuy nhiên, trong chăn nuôi bò sinh sản, chỉ tiêu này không được vượt quá 80 – 85 ngày [13]. Gần đây, Setiaji và cs (2020) cũng đã báo cáo rằng thời gian phối lại giống lần đầu sau khi đẻ là 73,81 ngày và thời gian phối lại giống thành công là 102,63 ngày ở bò đen Nhật Bản [15]. Kết quả tỷ lệ thụ thai thấp đã dẫn đến số bê sinh ra giảm và giá bê trở nên tăng dần. Giá bê trung bình năm 2013 cao hơn 20% so với giá bê năm 2012. Việc thực hiện phối giống (AI) lần thứ 2 trở đi cũng làm tăng chi phí liên quan đến tinh dịch, điều trị nội tiết tố, phí kỹ thuật viên AI cũng như khâu chăm sóc bò cho đến lần phối giống tiếp theo. Vì vậy, việc cải thiện tỷ lệ thụ thai là vấn đề cấp thiết để cải thiện hiệu quả sản xuất chăn nuôi bò sinh sản. Có nhiều yếu tố tác động đến sự thành công của chăn nuôi bò sinh sản bao gồm quản lý, chế độ chăm sóc dinh dưỡng, và yếu tố di truyền. Tuy nhiên, tỷ lệ thụ thai được báo cáo có hệ số di truyền thấp và do đó cải thiện di truyền tính trạng sinh sản bằng phương pháp nhân giống truyền thống có thể không hiệu quả. Do vậy, cách thức tiếp cận được áp dụng phổ biến gần đây là ứng dụng chi thị phân tử trong chọn lọc giống nhằm chọn lọc giống chính xác và có thể cải thiện được các đặc điểm di truyền giống.

3. GIẢI TRÌNH TỰ THẾ HỆ MỚI

Giải trình tự thế hệ mới (next generation sequencing) hay còn gọi là công nghệ giải trình tự thông lượng cao (high-throughput sequencing) đã phát triển nhanh chóng và tác động đáng kể đến cách tiến hành nghiên cứu di truyền. Việc quét đồng thời tất cả các gen ứng dụng NGS đã giúp xác định được một số lượng lớn các biến dị di truyền bao gồm những đột biến hiếm, đột biến nguy cơ trong các gen chức năng liên quan đến bệnh và rối loạn di truyền, từ đó đã làm thay đổi những chiến lược cải tiến nông nghiệp. Bước đột phá của NGS trong nghiên cứu bộ gen là sự ra đời của các phương pháp giải trình tự mục tiêu cho phép giải trình tự có chọn lọc các vùng quan tâm

Bảng 2. Các hệ thống giải trình tự phổ biến hiện nay

Công ty sản xuất	Illumina			Pacific Bioscience		Oxford Nanopore	
Hệ thống giải trình tự	Miseq	Nextseq2000	Novaseq 6000	Sequel	Sequel II	MinION	PromethION
Nguyên lý giải trình tự	Giải trình tự bằng phương pháp tổng hợp			Giải trình tự đơn phân tử Pacbio		Giải trình tự đơn phân tử Nanopore	
Hệ thống phát hiện	Huỳnh quang			Tính dẫn điện			
Độ dài trình tự	2 x 300 bp	2 x 300 bp	2 x 250 bp	300 kb		4 Mb	
Flow cell / thiết bị	1	1	2	12		1	24 hoặc 48
Ưu điểm	Độ chính xác cao			Đầu đọc dài, không cần khuếch đại DNA		Đầu đọc dài, giải trình tự nhanh, hệ thống máy nhỏ gọn	
Nhược điểm	Giải trình tự lâu			Tỷ lệ lỗi cao			

(Flow cell: một thiết bị nhỏ chứa thư viện các phân đoạn DNA, nơi xảy ra giải trình tự)

và từ đó làm giảm đáng kể số lượng trình tự cần được tạo ra. Giải trình tự exome (whole exome sequencing - WES) cho phép giải trình tự tất cả các vùng mã hóa protein và nhanh chóng đã trở thành phương pháp được sử dụng rộng rãi nhất, đặc biệt được sử dụng để khám phá những rối loạn di truyền Mendel. Phương pháp này cũng trở nên phổ biến để khám phá các đột biến trong các gen liên quan đến chức năng sinh sản ở cả người và gia súc. Ở người, các alen nguy cơ thường nằm trong vùng mã hóa protein và vùng mã hóa này chỉ chiếm khoảng 2% của bộ gen. Cách tiếp cận này cũng cho phép phát hiện các biến dị tại các vị trí ghép nối RNA (splicing donor/splicing acceptor sites) có tác động lớn đến quá trình dịch mã protein. Ở người, hệ exome chứa khoảng 85% đột biến có ảnh hưởng lớn đến các đặc điểm liên quan đến bệnh. Trong đó các đột biến nhằm nghĩa và vô nghĩa chiếm khoảng 60% các đột biến gây bệnh [16]. Việc sử dụng WES trong sàng lọc biến dị di truyền đã dẫn đến sự gia tăng đáng kể trong việc khám phá ra các biến dị hiếm gây ra những rối loạn di truyền Mendel [16]–[18]. Điều này được phản ánh qua dữ liệu di truyền trong OMIM (Online Mendelian Inheritance in Man) [19] và OMIA (Online Mendelian Inheritance in Animals) [20]. Thêm vào đó, cách thức tiếp cận khác trong nghiên cứu di truyền ứng dụng NGS là giải trình tự bộ gen (Whole genome sequencing - WGS). Ưu điểm của công nghệ này so với WES là độ bao phủ cao, giúp xác định được một lượng lớn SNPs và Indels trong bộ gen, nhiều biến dị tiềm năng còn lại khác ngoài vùng exon cũng sẽ được phát hiện. Điển hình như, hơn 2 triệu SNPs đã được khám phá ở bò đực Fleckvieh nhờ ứng dụng WGS [21]. Kawahara-Miki và cs (2011) đã xác định được 6,3 triệu SNPs, trong đó có hơn 5,5 triệu SNPs tiềm năng (chiếm 87% của bộ gen) được phát hiện ở bò Kuchinoshima

(bò bản địa Nhật) nhờ ứng dụng WGS [22]. Choi và cs (2014) đã báo cáo rằng bò Hanwoo Hàn Quốc, bò Heugu Jeju, và bò Hà Lan sở hữu xấp xỉ 10,4 triệu SNPs, trong đó phát hiện 54,12% là SNPs tiềm năng và 1.063.267 Indels trong những bộ gen này [2]. Như vậy, những kết quả này đã chỉ ra rằng WES/WGS là chìa khóa để khám phá các đột biến nguy cơ/tiềm năng liên quan đến những tình trạng kinh tế quan trọng.

4. CÁC HỆ THỐNG VÀ CÔNG CỤ PHÂN TÍCH DỮ LIỆU GIẢI TRÌNH TỰ THỂ HỆ MỚI

4.1. Các hệ thống giải trình tự thế hệ mới

Hiện tại có nhiều hệ thống giải trình tự thế hệ mới bao gồm giải trình tự thế hệ mới các đầu đọc ngắn (short read NGS) và đầu đọc dài (long read NGS). Trong đó, giải trình tự đầu đọc ngắn đang thống trị thị trường giải trình tự thế hệ mới. So với giải trình tự đầu đọc dài, giải trình tự đầu đọc ngắn có độ chính xác cao hơn, nhiều công cụ phân tích dữ liệu được thiết kế và dành riêng cho dữ liệu đầu đọc ngắn. Tuy nhiên, thời gian chạy lâu hơn và gặp khó khăn hơn khi lắp ráp *de novo*, xác định các dạng đồng phân phiên mã và cấu trúc của các biến thể. Một số hệ thống giải trình tự được mô tả như ở bảng 2.

4.2. Công cụ phân tích dữ liệu giải trình tự thế hệ mới

Ứng dụng các công cụ tin sinh học, quy trình phân tích dữ liệu giải trình tự thế hệ mới bao gồm các bước phân tích sơ cấp và thứ cấp. Phân tích sơ cấp được thực hiện bằng cách sử dụng các phần mềm như FastQC, Trimmomatic để kiểm tra chất lượng sau khi giải trình tự, lọc và cắt xén đầu đọc. Phân tích thứ cấp bao gồm sắp xếp và căn chỉnh trình tự đọc sử dụng công cụ BWA (Burrows-Wheeler Aligners) và phân loại biến thể (SNPs,



Bảng 3. Các công cụ tin sinh học được ứng dụng để phân tích dữ liệu NGS

Phân tích	Công cụ / Phần mềm được sử dụng phổ biến
Kiểm tra chất lượng sau khi giải trình tự	FastQC, MultiQC, FASTX-toolkit
Cắt xén đầu đọc và các trình tự chất lượng thấp	Trimmomatic, Cutadapt, Fastp
Lọc và loại bỏ các đầu đọc lặp lại	Picard
Sắp xếp và căn chỉnh trình tự đọc	BWA, Bowtie
Phân loại biến thể	GATK, freeBayes
Lọc và hợp nhất các biến thể	Bcftools
Phân tích biến thể theo chức năng	ANNOVAR, SnpEff, EnsemblVEP
Hiển thị kết quả	IGV, Genome Browser, Varsseq, Geneious

Bảng 4. Phân loại biến dị di truyền theo tác động sử dụng SNPEff

Tác động	Biến dị di truyền
Cao	Tạo bộ ba kết thúc, thay đổi chức năng của bộ ba kết thúc, thay đổi chức năng của bộ ba khởi đầu, xóa vùng mã hóa, lặp vùng mã hóa, thay đổi vùng ghép nối ở hai đầu exon
Trung bình	Đột biến nhầm nghĩa, đột biến ở trình tự mã hóa, chèn/xóa một bộ ba
Thấp	Đột biến đồng nghĩa
Sửa đổi	Đột biến xảy ra tại vùng không mã hóa, vùng không dịch mã ở đầu 5' và 3'

Indel, CNV) sử dụng công cụ GATK. Kết quả sắp xếp trình tự đọc và định dạng biến thể sẽ được lưu lại ở dạng BAM và VCF. Để mở các định dạng này có thể sử dụng các phần mềm miễn phí như IGV (Inter Genome Viewer) hoặc Genome Browser. Sau các bước phân tích sơ cấp và thứ cấp, công cụ ANNOVAR và VEP được sử dụng phổ biến để phân tích chức năng sinh học và bệnh lý của các biến thể. Bảng 3 trình bày tổng hợp các công cụ và các bước để phân tích dữ liệu NGS

4.3. Một số công ty và trường học tại Nhật phân tích giải trình tự thế hệ mới

Trong những năm gần đây, công nghệ giải trình tự thế hệ mới đã được ứng dụng phổ biến trong nghiên cứu để phát hiện các biến thể chức năng liên quan đến các tính trạng năng suất, thích nghi, và bệnh tật trên động vật bản địa. Một số trường đại học (Khoa Thú Y, trường Đại học Khoa học Okayama; Phòng thí nghiệm di truyền động vật ứng dụng, Trường sau đại học về Khoa học Tự Nhiên, Sự sống, Môi trường, và Công nghệ, Đại học Okayama) có các hệ thống giải trình tự đã đưa vào giảng dạy và nghiên cứu, bên cạnh đó, nhiều nhà nghiên cứu cũng chọn lựa các công ty giải trình tự để gửi mẫu nhằm đạt được kết quả nhanh, chính xác, và tiết kiệm được kinh phí nếu số lượng mẫu phân tích ít.

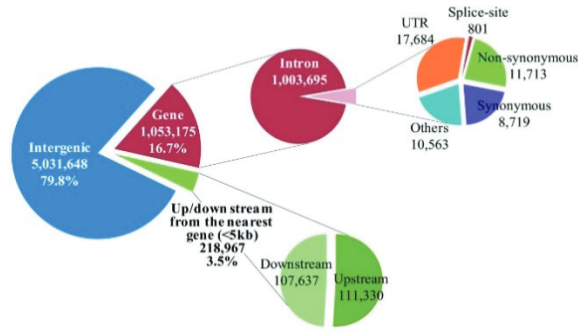
5. CÁC ĐỘT BIẾN NGUY CƠ LIÊN QUAN ĐẾN NĂNG SUẤT SINH SẢN ĐÃ ĐƯỢC PHÁT HIỆN Ở BÒ ĐEN NHẬT BẢN NHỜ ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ THẾ HỆ MỚI

5.1. Phân loại các biến dị di truyền theo tác động đến chức năng sinh học của gen

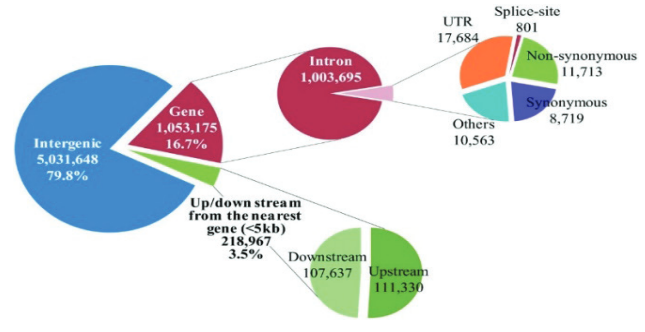
Ứng dụng NGS để khám phá dữ liệu nguồn gen của cá thể sẽ trích xuất dữ liệu trình tự bộ gen dưới định dạng FastQ. Sau đó, bộ dữ liệu trình tự này sẽ được lắp ráp và căn chỉnh theo trình tự tham chiếu. Để phát hiện và phân loại các biến dị di truyền từ bộ dữ liệu đã xử lý này, một số công cụ tin sinh học như VEP (Variant Effect Predictor) [23] và SNPEff [24] sẽ được áp dụng. Các biến dị di truyền được phát hiện sẽ được phân loại theo tác động bao gồm các biến dị tác động cao, tác động trung bình, tác động thấp, và tác động sửa đổi (Bảng 4).

Các SNPs/Indels xảy ra ở vùng mã hóa mà có thể tạo ra bộ ba kết thúc dẫn đến quá trình dịch mã kết thúc sớm hơn, thay đổi bộ ba khởi đầu dẫn đến quá trình dịch mã không được bắt đầu, thay đổi chuỗi trình tự axit amin, thay đổi nucleotide ở vùng ghép nối RNA ảnh hưởng đến quá trình phiên mã và dịch mã được xếp loại là các đột biến tác động cao. Những đột biến tác động cao được cho rằng có thể làm mất chức năng sinh học của gen [25], [26]. Ứng dụng WES/WGS cho phép xác định được sự có mặt hay không của các đột biến tác động cao này trong bộ gen của cá thể/quần thể. Kawahara-Miki và cs (2011) đã ứng dụng WGS để khám phá các biến dị di truyền ở bò Kuchinoshima đã phát hiện 801 SNPs nằm trong vùng ghép nối, 8.719 đột biến đồng nghĩa, và 11.713 đột biến nhầm nghĩa. Trong số các Indels, có 2.942 Indels trong vùng exon và 138 trong vùng ghép nối (Hình 1). Dựa vào kết quả phân tích này, ứng dụng GWAS (nghiên cứu

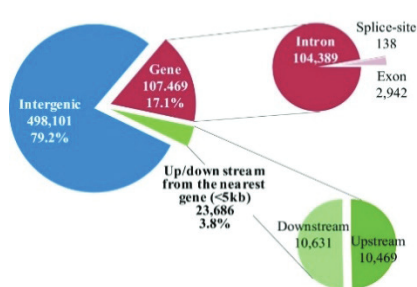
A. SNP



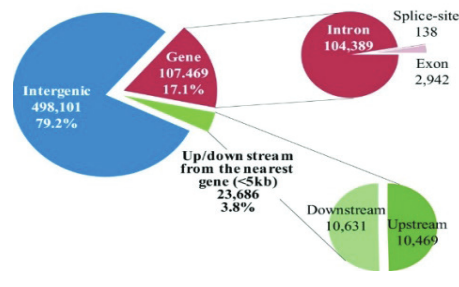
A. SNP



B. Indel



B. Indel



Hình 2. Tỷ lệ các biến dị di truyền SNPs/Indels ở bộ gen của bò Kuchinoshima Nhật Bản [22]

liên kết trên bộ gen giữa cá thể đối chứng và cá thể mang bệnh) có thể khám phá được những biến dị nguy cơ liên quan đến khả năng sinh sản thấp ở bò đen Nhật Bản. Từ đó, có thể có những chiến lược quản lý giống phù hợp.

5.2. Các đột biến/nguy cơ đã được phát hiện liên quan đến năng suất sinh sản thấp ở bò đen Nhật Bản

Ứng dụng tiến bộ của di truyền phân tử trong chăn

nuôi, nhiều SNP/Indels đã được phát hiện và nhanh chóng được lựa chọn làm chỉ thị phân tử trong chọn lọc giống chính xác, có thể loại bỏ được nguy cơ sự lan rộng của các alen lặn gây chết hay ảnh hưởng nghiêm trọng đến chức năng sinh học của gen. Để khám phá được những SNPs/Indels nguy cơ đó cũng như tần suất của chúng trong quần thể, cách tiếp cận gần đây được ứng dụng rộng rãi là sự kết hợp giữa GWAS và WES/WGS

Bảng 5. Một số SNPs/Indels đã được khám phá liên quan đến năng suất sinh sản ở bò đen Nhật Bản

Gen	*Vị trí	Loại đột biến	Phenotypes
<i>CDC45</i>	7474351	SNPs	Chết phôi [27]
<i>PID1</i>	117,836,406 - 118,104,185	SNPs	Tuổi đẻ lứa đầu [29]
<i>DNER</i>	118,193,385 - 118,581,892	SNPs	
<i>FBXO36</i>	118,773,130 - 118,879,541	SNPs	
<i>SLC16A14</i>	118,886,691 - 118,909,613	SNPs	
<i>SP140</i>	118,927,758 - 118,962,090	SNPs	
<i>SP110</i>	119,055,328 - 119,059,88	SNPs	
<i>TRIP12</i>	118,622,220 - 118,729,375	SNPs	Chết phôi [14]
<i>ANXA10</i>	NST8: 378,127 to 412,061	CNV	
<i>ACVR2A</i>	48476925	SNPs	Tỷ lệ thụ thai [30]
<i>ACVR2A</i>	48476943_48476946	Indels	Tỷ lệ thụ thai [30]
<i>PTH2R</i>	97,359,632 97,361,855 97,363,758 97,364,399 97,365,727 97,368,169	SNPs	Thời gian phối giống lại thành công [13]

*Vị trí của đột biến được sắp xếp theo bộ gen tham chiếu UMD-3.1.1



cho phép phát hiện nhiều SNPs/Indels quan trọng. Bước đầu tiên trong cách tiếp cận này là thu thập các thông tin về kiểu hình, được đề cập chủ yếu trong bài này ở đây là năng suất sinh sản của bò đen Nhật Bản bao gồm các chỉ tiêu như số lần giao phối của cá thể qua mỗi lứa đẻ, tỷ lệ chết phôi sớm/muộn, tỷ lệ thụ thai, sau đó ứng dụng WES hay WGS để xác định các SNPs/Indel tác động cao ở các cá thể trong quần thể có hiệu suất sinh sản thấp và cao để từ đó tìm ra mối liên quan giữa kiểu gen và kiểu hình. Sử dụng WES/WGS cho phép xác định được những SNPs/Indels đã được phát hiện trước đây trong quần thể và cả những đột biến mới có trong quần thể bò khảo sát. Sasaki và cs (2021) [27] đã ứng dụng công nghệ GWAS và WES nhằm xác định tần suất của alen nguy cơ ở quần thể bò đen Nhật Bản và đã khám phá được một đột biến nguy cơ trong vùng ghép nối với khoảng cách 5 bp từ exon 14 của gen *CDC15* liên quan đến tỷ lệ chết phôi ở bò đen Nhật Bản. *CDC45* mã hóa phức hợp enzyme khởi đầu helicase và preIC trong quá trình tổng hợp và sao chép DNA ở pha S. Yoshida và cs (2001) đã báo cáo rằng sự chết phôi đã được phát hiện ở chuột mang đột biến gen *Cdc45^{-/-}* [28]. Sasaki và cs cũng đề nghị rằng sự thay thế nucleotide G>T ở vị trí g.74743512 trong gen *CDC15* ảnh hưởng đáng kể đến sự ghép nối trong giai đoạn tạo RNA trưởng thành và tính ổn định của RNA trưởng thành. Tại thời điểm khám phá ra đột biến này, tần suất alen nguy cơ này đã có thể chiếm một tần suất khá cao trong quần thể bò đen Nhật Bản. Tần suất alen T được khảo sát ở một trung tâm giết mổ gia súc là 0,066. Kết quả này chỉ ra rằng alen nguy cơ này đã sẵn sàng lan rộng trong quần thể bò đen Nhật Bản. Mặc dù hiện tại không có cá thể mang kiểu gen đồng hợp lặn TT, nhưng tần suất mong đợi cá thể mang đồng hợp lặn TT được ước tính là 8,02 cá thể trong 1,137 cá thể dựa trên tần suất alen T khảo sát [27]. Cách tiếp cận nghiên cứu này có thể quản lý hiệu quả sự di truyền của các alen nguy cơ trong quần thể. Ưu điểm khác nữa của cách tiếp cận phương pháp này là xác định được nhiều SNP/Indels khác trong các gen liên quan đến tính trạng sinh sản như tỷ lệ thụ thai, tuổi đẻ lứa đầu, khoảng cách lứa đẻ. Bảng 5 liệt kê một số SNPs/indels đã được phát hiện liên quan đến tính trạng sinh sản ở bò đen Nhật Bản [13], [29]–[31].

Những kết quả này đề nghị rằng các đột biến nguy cơ này có thể được sử dụng như chỉ thị phân tử trong chọn lọc giống sinh sản để cải thiện tính trạng sinh sản ở bò đen Nhật Bản. Như vậy, để cải thiện tính trạng sinh sản, điều quan trọng là phải khai thác hay phân tích bộ gen để thiết lập cơ sở dữ liệu nguồn gen gồm các SNPs/Indels/CNV nhờ ứng dụng công nghệ NGS

6. TIỀM NĂNG ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ THỂ HỆ MỚI TRONG CHĂN NUÔI TẠI VIỆT NAM

Hiện nay, công nghệ NGS đã được ứng dụng khá phổ biến trong lĩnh vực y tế ở Việt Nam. Sử dụng NGS trong các xét nghiệm sàng lọc trước sinh không xâm lấn, sàng

lọc ung thư di truyền, xác định các đột biến hiếm gặp và bệnh lý di truyền đa gen. Đặc biệt, NGS hiện đang là xu hướng và là tiêu chuẩn về công nghệ trong sàng lọc di truyền phôi, góp phần tăng tỷ lệ thành công cho các ca IVF ở người. Theo Đỗ Tân Khang và cs (2019), trên cả nước có trên 25 hệ thống giải trình tự từ các công ty khác nhau đã được trang bị và vận hành tại nhiều viện, trường, trung tâm cũng như bệnh viện. Riêng ở thành phố Hồ Chí Minh và khu vực Đồng bằng sông Cửu Long hiện có khoảng 15 máy NGS. Nhiều máy đã đưa vào sử dụng cho các công tác nghiên cứu và chẩn đoán, nhưng vẫn còn nhiều máy đang trong tình trạng không sử dụng [32]. Tuy nhiên, thông tin này cho thấy tiềm năng ứng dụng NGS ở Việt Nam đang ngày càng được chú trọng và phát triển. Trong chăn nuôi, mặc dù ứng dụng NGS để cải thiện di truyền giống còn khá mới mẻ và đa số các nghiên cứu được thực hiện dựa trên sự hợp tác với các viện, trường đại học tại các quốc gia phát triển bao gồm Nhật Bản, nơi đáp ứng được cơ sở vật chất cũng như kinh phí để thực hiện nghiên cứu, nhưng việc áp dụng được NGS trong cải thiện di truyền vật nuôi bản địa sẽ là mang lại những thách thức và triển vọng mới cho sự phát triển bền vững chăn nuôi ở Việt Nam. Bên cạnh đó, giới hạn và khó khăn chính trong việc áp dụng công cụ này để phát hiện ra các đột biến tiềm năng hay nguy cơ liên quan đến các tính trạng kinh tế quan trọng ở gia súc là thông tin kiểu hình. Hiện tại hệ thống chăn nuôi ở Việt Nam vẫn chưa đồng bộ, các tính trạng sinh trưởng, sinh sản của gia súc được nuôi tại các hộ gia đình vẫn chưa được ghi lại, và chưa có hệ thống định dạng cá thể gia súc. Việc thiếu dữ liệu thông tin kiểu hình dẫn đến khó khăn trong xác định sự liên quan của kiểu gen được phát hiện với kiểu hình của cá thể/quần thể. Do vậy, để có thể áp dụng hiệu quả những công nghệ tiên tiến này nhằm cải thiện di truyền giống để nâng cao hiệu quả chăn nuôi, cần thiết lập định dạng cá thể và ghi lại thông tin kiểu hình của cá thể qua mỗi lứa đẻ.

TÀI LIỆU THAM KHẢO

- [1] T. Pinto De Melo, G. M. F. De Camargo, L. G. De Albuquerque, and R. Carvalheiro, 'Genome-wide association study provides strong evidence of genes affecting the reproductive performance of Nellore beef cows', PLoS One, vol. 12, no. 5, May 2017, doi: 10.1371/journal.pone.0178551.
- [2] J. W. Choi et al., 'Whole-genome analyses of Korean native and Holstein cattle breeds by massively parallel sequencing', PLoS One, vol. 9, no. 7, Jul. 2014, doi: 10.1371/journal.pone.0101127.
- [3] M. G. Diskin, M. H. Parr, and D. G. Morris, 'Embryo death in cattle: An update', Reprod Fertil Dev, vol. 24, no. 1, pp. 244–251, 2012, doi: 10.1071/RD11914.
- [4] K. Sato and T. Fujita, 'Elucidation of genes involved in the conception rates of cows decrease.', Bulletin of the Oita Prefectural Animal Industry Experiment Station, vol. 39, pp. 17–19, 2010.
- [5] N. Irikura, M. Uematsu, G. Kitahara, T. Osawa, and Y. Sasaki, 'Effects of service number on conception rate in Japanese Black cattle', Reproduction in Domestic Animals, vol. 53,

- no. 1, pp. 34–39, Feb. 2018, doi: 10.1111/rda.13049.
- [6] T. O. Yahaya et al., ‘Chromosomal abnormalities predisposing to infertility, testing, and management: a narrative review’, *Bull Natl Res Cent*, vol. 45, no. 1, Dec. 2021, doi: 10.1186/s42269-021-00523-z.
- [7] A. Lee and A. A. Kiessling, ‘Early human embryos are naturally aneuploid—can that be corrected?’, *J Assist Reprod Genet*, vol. 34, no. 1, pp. 15–21, 2017, doi: 10.1007/s10815-016-0845-7.
- [8] S. A. Yatsenko and A. Rajkovic, ‘Genetics of human female infertility’, *Biology of Reproduction*, vol. 101, no. 3. Oxford University Press, pp. 549–566, Sep. 01, 2019. doi: 10.1093/biolre/ioz084.
- [9] M. G. Butler, S. K. Rafi, A. McGuire, and A. M. Manzarado, ‘Currently recognized clinically relevant and known genes for human reproduction and related infertility with representation on high-resolution chromosome ideograms’, *Gene*, vol. 575, no. 1, pp. 149–159, Jan. 2016, doi: 10.1016/j.gene.2015.08.057.
- [10] Q. Sang, P. F. Ray, and L. Wang, ‘Understanding the genetics of human infertility’. [Online]. Available: <https://www.science.org>
- [11] T. Tozaki et al., ‘Rare and common variant discovery by whole-genome sequencing of 101 Thoroughbred racehorses’, *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-95669-1.
- [12] M. Motoyama, K. Sasaki, and A. Watanabe, ‘Wagyu and the factors contributing to its beef quality: A Japanese industry overview’, *Meat Sci*, vol. 120, pp. 10–18, 2016, doi: 10.1016/j.meatsci.2016.04.026.
- [13] S. Sasaki, T. Ibi, T. Kojima, and Y. Sugimoto, ‘A genome-wide association study reveals a quantitative trait locus for days open on chromosome 2 in Japanese Black cattle’, pp. 102–105, 2015, doi: 10.1111/age.12360.
- [14] S. Sasaki, T. Ibi, T. Akiyama, M. Fukushima, and Y. Sugimoto, ‘Loss of maternal ANNEXIN A10 via a 34-kb deleted-type copy number variation is associated with embryonic mortality in Japanese Black cattle’, *BMC Genomics*, vol. 17, no. 1, Nov. 2016, doi: 10.1186/s12864-016-3312-z.
- [15] A. Setiaji and T. Oikaw, ‘Genetics of heifer reproductive traits in Japanese Black cattle’, *Asian-Australas J Anim Sci*, vol. 33, no. 2, pp. 197–202, 2020, doi: 10.5713/ajas.19.0118.
- [16] Y. Zhao et al., ‘Whole genome and exome sequencing reference datasets from a multi-center and cross-platform benchmark study’, *Sci Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1038/s41597-021-01077-5.
- [17] H. Satam et al., ‘Next-Generation Sequencing Technology: Current Trends and Advancements’, *Biology*, vol. 12, no. 7. Multidisciplinary Digital Publishing Institute (MDPI), Jul. 01, 2023. doi: 10.3390/biology12070997.
- [18] B. S. Petersen, B. Fredrich, M. P. Hoepfner, D. Ellinghaus, and A. Franke, ‘Opportunities and challenges of whole-genome and -exome sequencing’, *BMC Genetics*, vol. 18, no. 1. BioMed Central Ltd., Feb. 14, 2017. doi: 10.1186/s12863-017-0479-5.
- [19] J. S. Amberger, C. A. Bocchini, F. Schiettecatte, A. F. Scott, and A. Hamosh, ‘OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an Online catalog of human genes and genetic disorders’, *Nucleic Acids Res*, vol. 43, no. D1, pp. D789–D798, Jan. 2015, doi: 10.1093/nar/gku1205.
- [20] F. Nicholas, I. Tammen, and S. I. Hub, ‘Online Mendelian Inheritance in Animals (OMIA)’. The University of Sydney, 1995. [Online]. Available: <https://omia.org>
- [21] S. H. Eck, A. Benet-Pagès, K. Flisikowski, T. Meitinger, R. Fries, and T. M. Strom, ‘Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery.’, *Genome Biol*, vol. 10, no. 8, 2009, doi: 10.1186/gb-2009-10-8-r82.
- [22] R. Kawahara-Miki et al., ‘Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi’, *BMC Genomics*, vol. 12, Feb. 2011, doi: 10.1186/1471-2164-12-103.
- [23] C. Zhang and I. Ochoa, ‘VEF: A variant filtering tool based on ensemble methods’, *Bioinformatics*, vol. 36, no. 8, pp. 2328–2336, Apr. 2020, doi: 10.1093/bioinformatics/btz952.
- [24] P. Cingolani et al., ‘A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3’, *Fly (Austin)*, vol. 6, no. 2, pp. 80–92, 2012, doi: 10.4161/fly.19695.
- [25] D. G. MacArthur and C. Tyler-Smith, ‘Loss-of-function variants in the genomes of healthy humans’, *Hum Mol Genet*, vol. 19, no. R2, Oct. 2010, doi: 10.1093/hmg/ddq365.
- [26] D. G. MacArthur et al., ‘Guidelines for investigating causality of sequence variants in human disease’, *Nature*, vol. 508, no. 7497. Nature Publishing Group, pp. 469–476, 2014. doi: 10.1038/nature13127.
- [27] S. Sasaki et al., ‘Identification of deleterious recessive haplotypes and candidate deleterious recessive mutations in Japanese Black cattle’, *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-86225-y.
- [28] K. Yoshida, F. Kuo, E. L. George, A. H. Sharpe, and A. Dutta, ‘Requirement of CDC45 for Postimplantation Mouse Development’, *Mol Cell Biol*, vol. 21, no. 14, pp. 4598–4603, Jul. 2001, doi: 10.1128/mcb.21.14.4598-4603.2001.
- [29] S. Sasaki, T. Ibi, S. Ikeda, and Y. Sugimoto, ‘A genome-wide association study reveals a quantitative trait locus for age at first calving in delta/notch-like EGF repeat containing on chromosome 2 in Japanese Black cattle’, *Anim Genet*, vol. 45, no. 2, pp. 285–287, 2014, doi: 10.1111/age.12109.
- [30] S. Sasaki, T. Ibi, T. Matsushashi, K. Takeda, S. Ikeda, and M. Sugimoto, ‘Genetic variants in the upstream region of activin receptor IIA are associated with female fertility in Japanese Black cattle’, *BMC Genet*, vol. 16, no. 123, pp. 1–14, 2015, doi: 10.1186/s12863-015-0282-0.
- [31] T. Arishima et al., ‘Maternal variant in the upstream of FOXP3 gene on the X chromosome is associated with recurrent infertility in Japanese Black cattle’, *BMC Genet*, vol. 18, no. 1, p. 103, Dec. 2017, doi: 10.1186/s12863-017-0573-8.
- [32] Đ. Tấn Khang, T. Thị, T. Khương, N. Phạm, A. Thi, and M. Duyên, ‘TIỀM NĂNG MỞ RỘNG ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ GEN THỂ HỆ MỐI Ở ĐỒNG BẰNG SÔNG CỬU LONG Title: Potential in developing application of next generation sequencing technology in the Mekong Delta’, doi: 10.22144/ctu.jsi.2019.001.