



DOI:10.22144/ctu.jsi.2019.001

TIỀM NĂNG MỞ RỘNG ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ GEN THỂ HỆ MỚI Ở ĐỒNG BẰNG SÔNG CỬU LONG

Đỗ Tấn Khang^{1*}, Trần Thị Thanh Khương¹, Nguyễn Phạm Anh Thi¹ và Trần Thị Mỹ Duyên²

¹Viện Nghiên cứu và Phát triển Công nghệ sinh học, Trường Đại học Cần Thơ

²Khoa Thủy sản, Trường Đại học Cần Thơ

*Người chịu trách nhiệm về bài viết: Đỗ Tấn Khang (email: dtkhang@ctu.edu.vn)

Thông tin chung:

Ngày nhận bài: 13/11/2018

Ngày nhận bài sửa: 04/01/2019

Ngày duyệt đăng: 12/04/2019

Title:

Potential in developing application of next generation sequencing technology in the Mekong Delta

Từ khóa:

Chẩn đoán, di truyền, gen, giải trình tự thể hệ mới (NGS), ung thư

Keywords:

Cancer, diagnostics, genetic, medicine, next generation sequencing

ABSTRACT

Next generation sequencing (NGS) is one of the modern tools strongly supporting for DNA sequencing. Many applications of such innovative technology have been successfully conducted in various research areas worldwide. In Vietnam, especially Mekong delta region, there are several NGS systems installed in some research institutions, universities and hospitals. However, applications have been slowly implemented due to numerous barriers. With the aim of effectively exploring the installed NGS instruments, we are going to demonstrate some feasible applications related to plant classification, aquaculture and medicine. In medicine, the novel and remarkable applications of NGS in early detection of cancers, non-invasive prenatal testing, and common genetic diseases will be revealed. In aquaculture, the practicable applications on genetic breeding and pathology will be illustrated. In plant taxonomy, important gene regions involving in evolution will be described in detail. Additionally, this article also discusses current research achievements and potential applications of NGS in other fields in Vietnam.

TÓM TẮT

Giải trình tự thể hệ mới (NGS) là một trong những công cụ hiện đại hỗ trợ đắc lực trong việc giải trình tự DNA. Trên thế giới, rất nhiều ứng dụng của công nghệ này đã được triển khai trên nhiều lĩnh vực khác nhau. Ở Việt Nam, đặc biệt là vùng Đồng bằng sông Cửu Long nhiều viện, trường, bệnh viện đã đầu tư hệ thống giải trình tự thể hệ mới. Tuy nhiên, việc triển khai ứng dụng còn nhiều bất cập. Nhằm khai thác hiệu quả các hệ thống giải trình tự thể hệ mới đã được lắp đặt, các ứng dụng khả thi ở Đồng bằng sông Cửu Long liên quan đến phân loại thực vật, nuôi trồng thủy sản và y học sẽ được trình bày. Về y học, những ứng dụng mới và nổi bật của NGS trong chẩn đoán chính xác một số bệnh ung thư, chẩn đoán tiền sinh không xâm lấn và một số bệnh lý phổ biến do gen sẽ được thảo luận. Ở lĩnh vực thủy sản, các ứng dụng trong di truyền chọn giống và bệnh học sẽ được đề cập. Còn phần phân loại thực vật, các vùng gen và hệ gen quan trọng liên quan đến phân loại sẽ được mô tả chi tiết. Ngoài ra, bài báo cũng thảo luận các thành tựu nghiên cứu đến thời điểm hiện tại cũng như tiềm năng ứng dụng trong các lĩnh vực khác của NGS ở Việt Nam

Trích dẫn: Đỗ Tấn Khang, Trần Thị Thanh Khương, Nguyễn Phạm Anh Thi và Trần Thị Mỹ Duyên, 2019. Tiềm năng mở rộng ứng dụng công nghệ giải trình tự gen thể hệ mới ở Đồng bằng sông Cửu Long. Tạp chí Khoa học Trường Đại học Cần Thơ. 55(Số chuyên đề: Công nghệ Sinh học)(1): 1-13.

1 GIỚI THIỆU

Công nghệ sinh học trên thế giới đang phát triển với tốc độ khá nhanh, nhiều công nghệ, kỹ thuật mới liên tiếp được công bố, trong đó đáng chú ý là sự bùng nổ của công nghệ giải trình tự thế hệ mới. Công nghệ này cho phép phân tích trình tự song song với khối dữ liệu cực lớn từ nhiều mẫu với chi phí thấp, đồng thời giảm thời gian phân tích. Thời gian cần thiết để thiết bị cho ra các trình tự có kích thước lên đến nhiều gigabase được tối giản từ hơn

mười năm xuống còn vài tiếng đồng hồ với chi phí giảm đáng kể. Một ví dụ điển hình là dự án giải trình tự bộ gen người được thực hiện trong gần 15 năm với chi phí hơn 1 triệu đô la Mỹ khi sử dụng phương pháp Sanger (NIH, 2019). Trong khi đó hệ thống giải trình tự thế hệ mới sử dụng hệ thống 454 Genome Sequencer FLX chỉ mất 2 tháng với chi phí thấp hơn 100 lần (Patrick, 2008). Giá giải trình tự bộ gen vi khuẩn sử dụng công nghệ nanopore chỉ tốn khoảng 1.000 đô la Mỹ (Kulski, 2016) (Bảng 1).

Bảng 1: Đặc điểm về thời lượng và giá của một số thiết bị giải trình tự thế hệ mới (Kulski, 2016)

Thiết bị/Công ty	Thời gian chạy (giờ hoặc ngày)	Chi phí (USD)/1 triệu base	Phần trăm lỗi dữ liệu	Giá thiết bị (USD)
454 GS FLX/Roche	24/48 giờ	10	1	500.000
GS Junior/Roche	18 giờ	9		100.000
Hiseq/Illumina	27/240 giờ	0,1	0,8	750.000
Miseq/Illumina	27 giờ	0,13	0,8	125.000
SOLID/Life Technologies	14 ngày	0,13	0,01	350.000
Ion Proton/Life Technologies	2-5 giờ	1	1,7	215.000
Ion PGM/Life Technologies	2-5 giờ	1	1,7	80.000
SMRT/Pac Bio	1-2 giờ	2	12,9	750.000
Nanopore/Oxford Nanopore Technologies	48/72 giờ	<1	34	1.000

Hiện nay, các ứng dụng của giải trình tự thế hệ mới (NGS) bao gồm: giải trình tự toàn bộ hệ gen sinh vật, giải trình tự hệ exon, giải trình tự RNA, sàng lọc trước sinh không xâm lấn (NIPT), sàng lọc di truyền trước chuyển phôi (PGT), sàng lọc và chẩn đoán ung thư (vòm họng, phổi, máu, gan, dạ dày, đại trực tràng), bệnh di truyền, trong nông nghiệp (xác định kiểu gen của dòng bố mẹ), xác định cộng đồng vi sinh vật, khoa học hình sự (xác định tội phạm, xác định danh tánh hài cốt, xác định quan hệ huyết thống),... (Yadav *et al.*, 2014; Motro and Moran-Gilad, 2017). 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 (ĐBSCL) hiện

có khoảng 15 máy giải trình tự thế hệ mới đã được lắp đặt (Bảng 2). 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. Nếu các máy này không được sử dụng sẽ là một sự lãng phí vô cùng lớn xét về mặt tiền của và cả về mặt công nghệ. Sự phát triển của công nghệ có thể nói từng ngày từng giờ, nhưng việc đưa các công nghệ này vào thực tiễn ở nước ta còn quá chậm do nhiều nguyên nhân khác nhau. Một trong những nguyên nhân quan trọng là người dùng chưa nắm rõ được hết công dụng của thiết bị này. Thứ hai, chi phí cho các bộ kit hóa chất để giải trình tự còn khá cao nếu số lượng mẫu được giải không nhiều. Thứ ba, một số hệ thống có quy trình phức tạp và cần sự hỗ trợ của nhiều thiết bị khác.

Bảng 2: Các đơn vị trang bị hệ thống giải trình tự thế hệ mới ở thành phố Hồ Chí Minh và vùng ĐBSCL

STT	Đơn vị	Ứng dụng
1	Tổ chức Oucru	Giải trình tự virus, vi khuẩn
2	Viện Pasteur	Giải trình tự virus, vi khuẩn Có định hướng triển khai về ung thư
3	Viện Y tế Công cộng	Giải trình tự virus, vi khuẩn
4	Bệnh viện Hùng Vương, TP.HCM	Định hướng phát triển hỗ trợ sinh sản: PGS, NIPT đơn gen...
5	Công ty Gene Solution	Dịch vụ NIPT đơn gen Định hướng phát triển chẩn đoán ung thư
6	Công ty Khoa Thương	Triển khai nghiên cứu và phát triển dịch vụ về ung thư, bệnh liên quan tới tim mạch...
7	Bệnh viện Nhi Đồng, TP.HCM	Định hướng triển khai về các bệnh thần kinh đầu và cổ Định hướng triển khai về chẩn đoán ung thư
8	Trường Đại học Kiên Giang	Ứng dụng giải trình tự cộng đồng vi sinh vật
9	Trường Đại học Y Dược Cần Thơ	Sàng lọc HBV, xác định biến chủng vi sinh vật gây bệnh

Với mong muốn tăng cường sự vận hành của các thiết bị giải trình tự thế hệ mới, đồng thời mở rộng các ứng dụng của công nghệ vượt trội này, một số ứng dụng tiềm năng khả thi ở ĐBSCL nói riêng và cả nước nói chung liên quan đến hệ thống giải trình tự thế hệ mới sẽ được thảo luận. Các lĩnh vực được đề cập là y học chẩn đoán, bệnh học và bộ gen thủy sản và các ứng dụng trong nghiên cứu thực vật.

2 CÁC ỨNG DỤNG CỦA GIẢI TRÌNH TỰ THẾ HỆ MỚI TRONG Y KHOA

Sinh thiết lỏng là một kỹ thuật không xâm lấn được sử dụng để chẩn đoán bệnh và theo dõi tình trạng bệnh. Trong sinh thiết lỏng, thay vì tác động vào mô đích, máu ngoại vi của bệnh nhân được lấy một cách đơn giản, không gây đau đớn để phục vụ cho các xét nghiệm hoặc phân tích (Dahl and Klotten, 2015). Trong huyết thanh bệnh nhân, có một thành phần nhỏ các tế bào khối u tuần hoàn (CTC) và các phân tử DNA tự do (cfDNA) là mục tiêu di truyền cho các xét nghiệm (Sausen *et al.*, 2014). Các DNA tự do trong máu này có nguồn gốc từ khối u hoặc từ thai nhi được phát hiện lần đầu tiên vào năm 1948 bởi Mandel and Matais. Nhưng phải đến những năm gần đây, cfDNA cùng với NGS đã mở ra rất nhiều ứng dụng trong chẩn đoán trên nhiều lĩnh vực an toàn và chính xác cao, nổi bật nhất trong hai lĩnh vực chẩn đoán tiền sinh không xâm lấn và chẩn đoán sớm các bệnh ung thư (Abou and Mahfouz, 2018; Mellis *et al.*, 2018). Bộ đôi NGS và cfDNA đã giúp các nhà nghiên cứu cũng các bác sĩ lâm sàng có những bước tiến vượt bậc trong việc tầm soát, chẩn đoán và điều trị các bệnh nguy hiểm, giảm chi phí điều trị và đặc biệt gia tăng tỉ lệ sống đối với các bệnh nhân ung thư một cách đáng kể (Paolillo *et al.*, 2016).

Chẩn đoán tiền sinh không can thiệp (non-invasive prenatal testing- gọi tắt là NIPT)

Trước đây, khi chưa phát hiện ra tế bào thai nhi tự do lưu hành trong huyết thanh người mẹ (cell – free fetal DNA gọi tắt là cffDNA), người phụ nữ mang thai thường được chỉ định thực hiện những xét nghiệm sàng lọc trước sinh truyền thống như siêu âm, chọc ối, triple test (Beta *et al.*, 2018). Những thủ thuật xâm lấn này có mang một nguy cơ xảy thai nhất định, đặc biệt đối với những thai phụ lớn tuổi hoặc có tiền sử các bệnh nguy hiểm liên quan đến thai kỳ (Niederstrasser *et al.*, 2017). Hơn thế nữa, để chẩn đoán một cách chính xác hơn, các phương pháp truyền thống này thường được thực hiện ở những giai đoạn sau của thai kỳ và đây sẽ là hạn chế lớn nếu như có sự can thiệp trong điều trị bệnh (Liao *et al.*, 2014).

Với sự hiện diện của cffDNA trong huyết tương người mẹ đã mở ra cánh cửa chẩn đoán tiền sinh không xâm lấn. Những đoạn DNA ngắn (100 – 200bp) này cung cấp một phương tiện phân tích di truyền vừa chính xác, đơn giản lại không gây nguy cơ xảy thai. Trong quá trình mang thai, khoảng bốn tuần thai, một lượng nhỏ DNA của thai nhi được giải phóng vào máu người mẹ từ nhau thai, chiếm 5-20% tổng lượng cfDNA có trong huyết thanh của người mẹ (Liao *et al.*, 2014). Những đoạn DNA nhỏ này chứa hầu hết các thông tin di truyền đến từ thai nhi và đây chính là nguồn nguyên liệu di truyền đại diện cho bào thai được dùng cho các xét nghiệm. Nhờ vào cffDNA và các kỹ thuật hiện đại của công nghệ sinh học, đặc biệt là NGS đã cho phép thực hiện lâm sàng những sàng lọc bất thường phổ biến trong thai nhi như rối loạn về số lượng nhiễm sắc thể thường, mất đoạn trên nhiễm sắc thể thường, rối loạn nhiễm sắc thể giới tính hoặc rối loạn đơn gen (Wei *et al.*, 2011; Taneja *et al.*, 2016).

Khoảng 50% trường hợp sảy thai ngẫu nhiên có nguyên nhân do bất thường về số lượng nhiễm sắc thể. Trong đó tam nhiễm 21 là thường gặp nhất. Các trường hợp tam nhiễm 16 gần nhau không tìm thấy ở các trẻ em sinh sống và trường hợp này nếu không xảy thai thì sẽ chết ngay sau khi sinh. Vì vậy xét nghiệm về tam nhiễm hầu như được chỉ định cho những phụ nữ mang thai có độ tuổi từ 35 trở lên. NIPT dựa vào cffDNA và NGS trong xét nghiệm tam nhiễm 16, 18, 21 là phương pháp NIPT đầu tiên được áp dụng lâm sàng với độ chính xác cao. Hiện nay, trên thế giới đã cho ra đời những bộ kit thương mại với quy mô sử dụng ngày một lớn. Nổi bật và đầu tiên là bộ kit MaterniT21 Plus được đưa vào thị trường vào khoảng cuối năm 2011 của Sequenom's Laboratories. Phương pháp này với mục đích xác định những bất thường trên nhiễm sắc thể 21, 18, 13 và trên nhiễm sắc thể giới tính. Ứng dụng công nghệ giải trình tự gen thế hệ mới, toàn bộ gen đã được kiểm tra chính xác. Cùng với đó là áp dụng phân tích sinh tin học cao cấp với hàng triệu điểm dữ liệu từ nhiễm sắc thể trên bộ gen, kết quả cuối cùng là hình ảnh lâm sàng rất chính xác. Cụ thể Palomaki *et al.* (2011) đã khảo sát trên 1.696 bệnh nhân có nguy cơ mang thai bị hội chứng Down. Chỉ khoảng 10 ml máu ngoại vi lấy từ người mẹ mang thai ở tuần thai thứ 10, kit MaterniT21 Plus đã cho kết quả độ chính xác cao > 99,1 % cho tam nhiễm 21, > 99,99 cho tam nhiễm 18 và 91,7% cho tam nhiễm 13 (Palomaki *et al.*, 2011). Hay trong một nghiên cứu khác, tác giả sử dụng NIPT – ILLumina, khảo sát ở 85.000 trường hợp với độ chính xác rất cao >99,9% (Bảng 3) (Taneja *et al.*, 2016).

Bảng 3: Kết quả xét nghiệm tam nhiễm của bộ kit MaterniT21 Plus- Illumina trên 85.000 trường hợp

Nhiễm sắc thể	Độ nhạy quan sát được (Observed sensitivity)	Độ nhạy	Độ đặc hiệu
21	>99,49%	100%	99,76%
18	97,23%	97,37%	99,57%
13	97,98%	87,50%	100%
Nhiễm sắc thể			Sự phù hợp
X			99,90%
XX			99,60%
XY			99,60%
XXX/XXY/YYY			97,40%

Bảng 4 tổng hợp một số kit thương mại nổi bật đang được ứng dụng tại Mỹ (số liệu khảo sát vào năm 2013). Với tiêu chí độ chính xác cao, xét nghiệm đơn giản, lượng máu ít, thời gian xét nghiệm nhanh và đặc biệt là giá thành thấp luôn mở ra một hướng nghiên cứu đặc biệt ứng dụng ở các nước

đang phát triển. Ở các nước phát triển, phương pháp sàng lọc trước sinh rất phổ biến. Phụ nữ mang thai từ 10 tuần thai đã có thể được xét nghiệm NIPT, do đó tình trạng bệnh và dị tật thai ở các nước này được kiểm soát rất chặt chẽ (Norton *et al.*, 2012; Kotsopoulou *et al.*, 2015; Renga, 2018).

Bảng 4: Thông số kỹ thuật một số kit NIPT thương mại tại Mỹ (nguồn Agarwal *et al.*, 2013)

Nơi sản xuất	Sequenin Lab	Illumina	Roche	Natera Inc
Tên bộ kit	Materni T21 Plus	Verifi prenatal	Harmony prenatal	Panorama
Mục tiêu	Tam nhiễm 13, 18, 21, bất thường số lượng NST giới tính và mất đoạn NST	Tam nhiễm 13, 18, 21, bất thường số lượng NST giới tính và xác định giới tính	Tam nhiễm 13, 18, 21, bất thường số lượng NST giới tính và xác định giới tính	Tam nhiễm 13, 18, 21, bất thường số lượng NST giới tính và xác định giới tính, hội chứng mất đoạn 22q11.2 và các mất đoạn khác
Thể tích máu	20ml (2 ống)	7ml (1 ống)	10ml (1 ống)	20 ml (1 ống)
Thai kỳ	10 tuần	10 tuần	10 tuần	7 tuần
Độ chính xác	>99%	100%	>99%	100%
Độ nhạy	92 - 99%	87 - 99%	80 - 99%	92 - 99%
Giá thành (vào năm 2013)	2700\$	1500\$	800\$	1500\$
Thời gian thực hiện	7 ngày	3-6 ngày	8-10 ngày	7-10 ngày

NIPT không chỉ được áp dụng trong tầm soát các bị tật do bất thường NST, mà còn được sử dụng rất nhiều trong các rối loạn di truyền khác ở thai nhi như fetal rhesus D determination (Saramago *et al.*, 2018), bệnh liên quan tới NST giới tính như duchenne (Parks *et al.*, 2016) và các bệnh do rối loạn đơn gen (Lench *et al.*, 2013; Verhoef *et al.*, 2016) .

Thalassemia, bệnh tan máu bẩm sinh là một trong những bệnh di truyền rối loạn đơn gen lặn phổ biến nhất trên thế giới. Theo thống kê của Liên đoàn Thalassaemia và WHO (2007), thalassaemia chiếm khoảng 7% dân số thế giới, với khoảng 330.000 trẻ em sinh ra hàng năm bị mắc bệnh và 1,1 % các cặp vợ chồng sinh có nguy cơ mang bệnh. Tại Việt Nam, tan máu bẩm sinh di truyền rất thường gặp và chiếm tỉ lệ cao so với các bệnh di truyền do rối loạn đơn gen khác. Bệnh tan máu bẩm sinh cũng như các bệnh do rối loạn đơn gen thường được chẩn đoán bằng cfDNA kết hợp với phương pháp PCR như AS-PCR, COLD-PCR hoặc enzyme cắt giới hạn đã cho kết quả nhất định trên nghiên cứu cũng như lâm sàng

(Van *et al.*, 2017). Tuy nhiên, một trong những hạn chế lớn nhất của các kỹ thuật sinh học phân tử này là độ chính xác chỉ phù hợp trong trường hợp đột biến tại một điểm, nhưng không thể phát hiện bệnh do đột biến nhiều điểm như bệnh thanataphoric dysplasia (Chitty *et al.*, 2015). Sự ra đời và cải tiến của hệ thống NGS đã khắc phục được hạn chế của các kỹ thuật PCR khi có nhiều đột biến trên cùng một gen. Chỉ trong một lần chạy, NGS cùng với NIPT đã phát hiện ra nhiều đột biến một cách đơn giản. Cụ thể, Chitty *et al.*, 2015 đã phát triển một hệ gen đột biến (panel of mutation) dành cho bệnh loạn sản xương nguyên nhân là do đột biến ở nhiều vị trí khác nhau trên gen FGFR3 và đã chứng minh độ chính xác vượt trội khi so sánh với phương pháp PCR-RELP trước đó. Trong lần thử nghiệm này, nhóm đã phát hiện ra 29 đột biến chỉ trong một lần chạy. Phương pháp sử dụng NGS với hệ gen đột biến này đã được đưa vào chẩn đoán lâm sàng tại rất nhiều bệnh viện ở Anh từ năm 2014. Những nghiên cứu và áp dụng trong chẩn đoán rối loạn gen lặn khác như ở bệnh xơ nang (Hill *et al.*, 2015) và

thalassemia (Xiong *et al.*, 2015) đang được tiến hành rộng rãi ở các phòng thí nghiệm trên thế giới, và hướng đến phát triển các kit chẩn đoán ở phổ rộng nhiều đột biến và nhiều gen hơn chỉ trong một phản ứng.

Tại Việt Nam, NIPT đã bước đầu được ứng dụng ở các bệnh viện và trung tâm phụ sản tại thành phố Hồ Chí Minh và Hà Nội, qua đó cho thấy tiềm năng ứng dụng rất cao. Như vậy, NGS và cfDNA sẽ là xu thế tương lai trong chẩn đoán tiền sinh không xâm lấn trên thế giới và tại Việt Nam. Với kết quả chính xác, giá cả phù hợp, thời gian tầm soát sớm sẽ phần nào giảm đi lo lắng của các bậc cha mẹ và đồng thời giúp xã hội giảm đi rất nhiều những bệnh nguy hiểm không mong muốn.

Trong lĩnh vực điều trị các bệnh ung thư, những năm gần đây NGS đã đóng góp đặc lực vào các lĩnh vực như tầm soát ung thư di truyền (Price *et al.*, 2018), chẩn đoán ung thư giai đoạn sớm (Paolillo *et al.*, 2016) và phát hiện đột biến kháng thuốc trong quá trình điều trị ung thư (Xue and Wilcox, 2016) đã và đang mang lại những kết quả rất khả quan. Ứng dụng phổ biến và hiệu quả nhất là nay của NGS trong lĩnh vực tầm soát ung thư di truyền. Từ các số liệu thống kê, ung thư di truyền chiếm khoảng 10% các loại bệnh ung thư (Stanislaw *et al.*, 2016). Các nhà nghiên cứu đã tìm ra mối liên hệ chặt chẽ giữa đột biến các gen với trên 35 hội chứng ung thư di truyền. Nghiên cứu được thực hiện giải trình tự các gen hội nhập cao lấy từ dự án bản đồ gen ung thư của viện sức khỏe quốc gia Mỹ (NIH) trên 3.000 trường hợp ung thư. Kết quả cho thấy, trong 12 loại

ung thư phổ biến, ung thư buồng trứng có đột biến di truyền cao nhất, tiếp đến là ung thư dạ dày, ung thư vú, 4% xảy ra ở ung thư thần kinh đệm, ung thư đầu, cổ, ung thư tuyến tiền liệt và ung thư phổi. Trong các gen đột biến di truyền, nổi bật và được nghiên cứu nhiều nhất là bộ đôi gen BRCA1 và BRCA2. Những người bình thường được xét nghiệm mang đột biến 2 gen này có nguy cơ 65% mắc bệnh ung thư vú và 50% mắc bệnh ung thư buồng trứng. Trên nhiều nghiên cứu khác, các bệnh nhân ung thư tuyến tụy, ung thư tuyến tiền liệt cũng như ung thư vú ở nam giới đều được phát hiện có xảy ra hiện tượng đột biến trên bộ đôi gen này. Ngoài ra, đột biến xảy ra trên nhóm gen MLH1, MSH2, MSH6, PMS2 và EPCAM được phát hiện ở những bệnh nhân có hội chứng Lynn, tiền thân của ung thư trực tràng (Rahner *et al.*, 2010, 2013). Với số liệu hiện tại, đã có trên 27 gen có liên quan tới ung thư di truyền. Và số lượng các gen tương tự ngày càng được phát hiện (Gardner *et al.*, 2018). Với công nghệ giải trình tự thế hệ mới, các nhà khoa học đã nhận biết chính xác trình tự đột biến cùng một lúc số lượng rất lớn các gen liên quan tới các bệnh ung thư di truyền. Kỹ thuật này thực hiện một cách đơn giản và tiết kiệm chi phí, làm giảm thời gian xét nghiệm nhiều gen ung thư trên chỉ một lần chạy. Bảng 5 liệt kê một số kit tầm soát ung thư áp dụng NGS đang sử dụng phổ biến và đưa vào lâm sàng ở một số quốc gia trên thế giới, chứng tỏ được hiệu quả và tiềm năng cao của NGS trong tầm soát các bệnh ung thư di truyền.

Bảng 5: Một số bộ kit thương mại tầm soát ung thư ứng dụng kỹ thuật NGS (Paolillo *et al.*, 2016)

Hãng sản xuất	Tên	Số lượng gen
Illumina, Inc (Mỹ)	TruSight Caner	94
	TruSight Tumor 15	15
	TruSight Myeloid	54
	TruSeq Amplicon Cancer Panel	48
Thermo Fisher Scientific (Mỹ)	Ion AmpliSeq Cancer Hotspot Panel v2	50
	Ion AmpliSeq Comprehensive Cancer Panel	400
Qiagen (Đức)	Actionable Inside Tumor Panel	12
	ClearSeq Comprehensive	151
Agilent Technologies (Mỹ)	ClearSeq Cancer	47
	ClearSeq AML	20
	Cancer research Panel	47
Foundation Medicen, Inc (Mỹ)	Foundation One	315
University of Washington (Mỹ)	UW-Oncoplex	234
ParadigmDX (Mỹ)	PCDx	114
ARUP Lab (Mỹ)	Solid Tumor Mutation Panel	48
Caris Life Science (Mỹ)	MI Profiles	46
Kight Diagnostic Lab (Mỹ)	GeneTrails Cancer Gene Panel	38
PahtGroup (Mỹ)	SmartGenomics	35
Life Technologies (Mỹ)	Pervenio Lung NGS assay	25

Với việc tầm soát ung thư di truyền bằng công nghệ giải trình tự, đội ngũ bác sĩ sẽ đánh giá được các nguy cơ mắc bệnh và đưa ra dấu chỉ điểm bệnh lý tiền ung thư và ung thư thông qua các xét nghiệm tại mô đích. Tầm soát sớm ung thư di truyền đang được cho là chìa khóa vàng để phát hiện và đưa ra các phương pháp ngăn chặn ung thư, giảm đáng kể nguy cơ tử vong và chi phí cho người bệnh.

Bên cạnh tầm soát ung thư, những khó khăn trong quá trình điều trị ung thư cũng đang là một bài toán rất khó đối với các nhà nghiên cứu. Một trong những nguyên nhân gây thất bại trong quá trình điều trị chính là việc tế bào ung thư không đáp ứng thuốc điều trị. Nguyên nhân của hiện tượng này được cho là do đột biến các gen liên quan (Adashek *et al.*, 2018). Với công nghệ giải trình tự thế hệ mới, toàn bộ quá trình điều trị của bệnh được kiểm soát một cách chặt chẽ, theo dõi khả năng đáp ứng điều trị và phát hiện ra đột biến kháng thuốc phát sinh trong quá trình điều trị một cách chuẩn xác. Với việc xác định được nguyên nhân do đột biến gen cụ thể, những thuốc điều trị đích sẽ được sử dụng để ức chế gen đột biến, đúng liều và đúng thời điểm, qua đó mang lại hiệu quả điều trị cao nhất (Stanislaw *et al.*, 2016).

Như vậy, với khả năng giải trình tự chính xác và số lượng lớn trong cùng một lúc, việc hoàn thiện quy trình NGS được hứa hẹn sẽ là công cụ đắc lực giúp các nhà khoa học chiến thắng trong cuộc chiến chống ung thư khốc liệt đang diễn ra trên toàn thế giới.

3 CÁC TIỀM NĂNG ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ THẾ HỆ MỚI TRONG THỦY SẢN Ở ĐBSCL

Trong lĩnh vực thủy sản, công nghệ NGS đã được sử dụng rộng rãi trên thế giới, giúp các nhà nghiên cứu hiểu rõ hơn về tổ chức, cấu trúc, chức năng cũng như sự tiến hoá của bộ gen của động vật thủy sản (Yáñez *et al.*, 2015; Yue and Wang, 2017), tìm hiểu cơ chế di truyền của các đặc điểm nổi bật của động vật thủy sản và sự tương tác của nó với môi trường, cung cấp dữ liệu trong việc tìm kiếm các chỉ thị phân tử trong quá trình chọn giống, xác định giới tính hay đánh giá đa dạng di truyền (Kumar and Kocour, 2017; Robledo *et al.*, 2017). Bên cạnh đó, công nghệ NGS còn được ứng dụng trong việc quản lý bệnh truyền nhiễm trên động vật thủy sản, ví dụ như giải trình tự toàn bộ bộ gen của vi khuẩn và vi-rút gây bệnh cho động vật nuôi thủy sản để giảm thiểu sự xuất hiện và lây lan của bệnh (Bayliss *et al.*, 2017; Nkili-Meyong *et al.*, 2017), trong các nghiên cứu về độc chất học trong môi trường nước (Mehinto *et al.*, 2012) hay trong các nghiên cứu về đa dạng sinh học của sinh vật biển

(Huete-Pérez and Quezada, 2013). Trong phần này, tiềm năng ứng dụng của NGS trong quản lý dịch bệnh cũng như trong di truyền và chọn giống sẽ được trình bày.

3.1 Ứng dụng công nghệ NGS trong quản lý dịch bệnh thủy sản

Quản lý dịch bệnh là một thách thức quan trọng trong phát triển NTTS bền vững. Nhằm giảm thiểu sự xuất hiện và lây lan của dịch bệnh, nguy cơ bùng phát và phương thức lây truyền của các tác nhân gây bệnh cần được tìm hiểu. Nguyên nhân bùng phát bệnh có thể do sự tiến hoá của các chủng mới (Côté *et al.*, 2008), sự lan truyền của các chủng giữa các quốc gia thông qua xuất nhập con giống, cá thể trưởng thành, cá thể bố mẹ hoặc do các tác động của biến đổi môi trường (Marcos-López *et al.*, 2010; Walker and Winton, 2010). Ngoài ra, chủng mới xuất hiện có thể do thay đổi vật chủ trong quá trình nuôi ghép hoặc từ những sinh vật hội sinh trong cùng một loài, điều này có thể do đột biến hoặc sự lan truyền theo chiều ngang giữa các gen mang độc lực thông qua sự tái tổ hợp giữa các chủng gây bệnh trước đó (Wijegoonawardane *et al.*, 2009).

Công nghệ NGS, đặc biệt là illumina miseq, cho phép giải trình tự toàn bộ hệ gen của vi khuẩn và vi rút gây bệnh (Nkili-Meyong *et al.*, 2017). Hơn 20 hệ gen tham chiếu của vi khuẩn, vi rút gây bệnh trên động vật thủy sản đã được công bố, cung cấp phần lớn thông tin về trình tự gen của các chủng vi khuẩn, vi rút khác nhau trong cùng một loài, cho thấy sự biến đổi trong hệ gen của loài, từ đó tìm lại nguy cơ bùng phát và phương thức lan truyền thông qua các biến thể gen cũng như cung cấp thông tin dự đoán liên quan đến các kiểu hình quan trọng (Bayliss *et al.*, 2017). Barnes *et al.* (2016) đã giải trình tự 51 hệ gen của vi khuẩn *Yersinia ruckeri*, một tác nhân gây bệnh phổ biến trên cá hồi. Dựa vào các chỉ thị phân tử SNPs, so sánh phát sinh loài của 58 chủng vi khuẩn thu thập từ Úc, New Zealand, Mỹ, Chile, Phần Lan và Trung Quốc cho thấy loài vi khuẩn này có nhiều chủng khác nhau trên toàn cầu cũng như trong ranh giới của nước Úc. Kết quả cho thấy kiểu huyết thanh O1b của chủng vi khuẩn thu ở Úc khác biệt về mặt di truyền so với kiểu huyết thanh của các chủng vi khuẩn thu ở nơi khác, và kiểu huyết thanh này chưa được công bố trước đây. Nghiên cứu cũng lần đầu tiên chứng minh sự hiện diện của kiểu huyết thanh O2 ở Tasmania, kiểu huyết thanh này cũng khác biệt hoàn toàn về mặt di truyền với kiểu huyết thanh O2 của chủng vi khuẩn thu ở Mỹ. Điều này cho thấy 2 chủng vi khuẩn *Yersinia ruckeri* là chủng bản địa.

Khi so sánh trình tự hệ gen của 3 dòng vi rút Cyprinid herpesvirus (CyHV), tác nhân gây bệnh

cho cá chép (*Cyprinus carpio carpio*) và cá chép koi (*Cyprinus carpio koi*), kết quả cho thấy độ tương đồng giữa các chủng là 99%, với phần lớn sự khác biệt nằm trên các vùng lặp lại liên kề khác nhau (variable number of tandem repeats – VNTRs) (Avarre *et al.*, 2011). Dựa vào sự khác biệt của các VNTRs, các nhà khoa học có thể phát triển các kỹ thuật chuẩn đoán mầm bệnh chính xác, nhanh và nhạy như kỹ thuật khuếch đại gen (polymerase chain reaction – PCR), kỹ thuật khuếch đại đẳng nhiệt (loop-mediated isothermal amplification – LAMP).

Công nghệ NGS rất hữu ích khi xác định các gen liên quan đến đáp ứng miễn dịch không đặc hiệu cũng như hệ miễn dịch không đặc hiệu của vật chủ khi chống lại tác nhân gây bệnh. Dựa trên dữ liệu của hệ gen và hệ phiên mã, Ordas *et al.* (2011) đã lập được một bộ tham chiếu chủ thích của các gen đáp ứng, các protein chức năng liên quan đến đáp ứng của hệ miễn dịch khi phơi cá ngựa vằn nhiễm *Salmonella typhimurium*. Ngoài ra, nhóm tác giả đã tìm được nhóm gen đáp ứng và dự đoán những gen có tiềm năng đáp ứng trong hệ miễn dịch không đặc hiệu, những gen này đều hoạt động khi có sự hiện diện hay vắng mặt của hệ miễn dịch đặc hiệu. Trong một nghiên cứu khác, Zhang *et al.* (2011) dựa vào trình tự hệ gen và hệ phiên mã để xác định các gen miễn dịch không đặc hiệu của cá chép (*Cyprinus carpio*).

3.2 Ứng dụng công nghệ NGS trong di truyền và chọn giống động vật thủy sản

Chọn giống là một bước quan trọng trong NTTS. Mục tiêu của nghiên cứu chọn giống và nhân giống là để tăng cường hiệu quả sản xuất NTTS, tăng chất lượng sản phẩm cũng như tăng lợi nhuận của người nuôi. Bên cạnh những kết quả đã đạt được nhờ vào các phương pháp chọn giống truyền thống (lai tạo, chọn lọc dựa trên đặc điểm kiểu hình...), kỹ thuật di truyền phân tử đang dần được áp dụng vào công tác chọn giống trong NTTS. Kỹ thuật di truyền phân tử cho phép so sánh sự đa dạng di truyền giữa các giống động vật nuôi thủy sản dựa vào cấu trúc hệ gen, chức năng gen của động vật thủy sản, các biến thể về kiểu gen và kiểu hình. Cho đến năm 2017, với những tiến bộ của công nghệ NGS, trình tự toàn hệ gen của hơn 24 loài động vật thủy sản đã được công bố (Yue and Wang, 2017) trong đó có cá tuyết Đại Tây Dương (Star *et al.*, 2011), hào Thái Bình Dương (Zhang *et al.*, 2012), cá chép (Xu *et al.*, 2014), cá rô phi (Xia *et al.*, 2014)... Dữ liệu trên hệ gen tham chiếu bao gồm trình tự DNA, chú thích và thông tin khác liên quan đến trình tự DNA, nhiễm sắc thể và các gen chức năng...

Nghiên cứu tương quan trên toàn hệ gen (genome-wide association studies – GWAS), là một

cách tiếp cận tương đối mới trong khoa học thủy sản để xác định các gen liên quan đến các tính trạng quan trọng như tính trạng qui định tăng trưởng, giới tính và kháng bệnh thông qua các chỉ thị phân tử (SNPs). Ví dụ, 4 locus điều khiển tính trạng (quantitative trait loci - QTL) kháng bệnh columnaris trên cá da trơn (*Ictalurus spp.*) (Geng *et al.*, 2015) và 9 QTL kháng bệnh do vi khuẩn *Vibrio anguillarum* trong hệ gen của cá bon Nhật (*Paralichthys olivaceus*) (Shao *et al.*, 2015) đã được xác định. Công nghệ NGS cũng đã được sử dụng để nghiên cứu những thay đổi trong biểu hiện gen để đáp ứng với stress và cơ chế tử vong do stress trong cá com Nhật Bản, *Coilia nasus* (Du *et al.*, 2014). Sự điều hòa các gen chuyển hóa glucose và chuyển hóa lipid tăng, các gen liên quan đến quá trình chết tế bào và các gen chỉ thị tổn thương gan cho thấy stress gây ra tổn thương gan thông qua quá trình chết của tế bào ty thể. Hơn nữa, kết quả phân tích hệ phiên mã đường ruột của hai loài cá nước ngọt về khả năng chịu mặn (cá rô phi Mozambique có khả năng chịu mặn cao, *Oreochromis mossambicus*, và cá rô phi ít chịu mặn, *Oreochromis niloticus*) cho thấy độ tương đồng về biểu hiện gen đáp ứng với sự thay đổi độ mặn giữa 2 loài ở ruột sau nhưng có sự khác biệt ở ruột trước. Trong đó, cá rô phi *Oreochromis mossambicus* có 21 gen ở ruột trước tăng hoạt động với độ mặn trong khi 21 gen này giảm hoạt động ở cá rô phi *Oreochromis niloticus*. Các gen này có chức năng vận chuyển và tham gia vào các hoạt động trong các kênh ion. Điều này cho thấy ruột trước đóng một vai trò quan trọng trong cơ chế điều chỉnh khả năng chịu mặn của cá rô phi (Ronkin *et al.*, 2015).

Chọn lựa giới tính là một trong những yếu tố quan trọng đối với nhiều loài nuôi trồng thủy sản vì kiểu hình giới tính ở một số loài động vật thủy sản liên quan đến tính trạng tăng trưởng. Việc kiểm soát tỷ lệ giới tính giúp đạt được tốc độ tăng trưởng cao hơn, cũng như tránh được sự phân tán kích cỡ. Ví dụ như ở tôm càng xanh, con đực có tốc độ tăng trưởng nhanh hơn và kích thước tối đa to hơn con cái (FAO, 2016). Hiện nay, rất ít gen qui định giới tính được xác định ở các loài động vật thủy sản. Ứng dụng công nghệ NGS giúp nâng cao kiến thức về chọn lựa giới tính trong NTTS, như các nghiên cứu về kiểu gen qui định kiểu hình giới tính trên cá rô phi *Oreochromis niloticus* (Palaiokostas *et al.*, 2013, 2015). Nhóm tác giả đã tìm ra được các vùng trình tự có chứa các locus liên quan đến các đặc điểm kiểu hình giới tính trong hệ gen của cá thể đực hoặc cái, cung cấp thông tin các đa hình nucleotide đơn (single nucleotide polymorphism-SNPs) liên quan đến các locus này. Kết quả này sẽ được sử dụng trong việc chọn giống cũng như sản xuất giống cá rô phi *Oreochromis niloticus* toàn đực nhằm cải thiện

chất lượng di truyền, cung cấp con giống có chất lượng cao cho ngành NTTS.

Ứng dụng của công nghệ NGS trong lĩnh vực thủy sản ở ĐBSCL còn nhiều hạn chế. Một số nghiên cứu phân tích về đa dạng di truyền của một số loài cá như cá bông lau, cá dứa và cá tra bần (Dương Thúy Yên và *ctv.*, 2015), cá rô đồng (Dương Thúy Yên, 2014; Phạm Thị Trang Nhung và Dương Thúy Yên, 2014) và cá hường (Dương Thúy Yên và *ctv.*, 2018; Dương Thúy Yên, 2014; Phạm Thị Trang Nhung và Dương Thúy Yên, 2014) cũng như các nghiên cứu về dịch tễ học, tìm hiểu đặc điểm gen của tác nhân gây bệnh đốm trắng trên tôm cũng như con đường lan truyền của nó ở mức độ phân tử (Hoa *et al.*, 2012a; Hoa *et al.*, 2012b) đã được thực hiện ở ĐBSCL. Bên cạnh đó, nhờ vào kết quả giải trình tự, cơ sở dữ liệu về hệ gen chức năng liên quan đến tính trạng chịu mặn của cá tra (Nguyễn Minh Thành và *ctv.*, 2015), kết quả định danh một số loài *Moina* spp. như *Moina micrura* và *Moina macrocopa* (Lê Văn Hậu và *ctv.*, 2018) hay các tác nhân gây bệnh trên động vật nuôi thủy sản như vi khuẩn *Streptococcus dusalactiae* gây bệnh trên cá bông kèo (Nguyễn Thu Dung và *ctv.*, 2016), vi bào tử trùng (*Microsporidia*) gây bệnh trên cá tra đã được công bố (Nguyễn Thị Thu Hằng và Đặng Thị Hoàng Oanh, 2016). Tuy nhiên, các mẫu giải trình tự đều cần phải gửi đến phòng thí nghiệm tại thành phố Hồ Chí Minh hoặc thực hiện tại nước ngoài. Với các ví dụ về ứng dụng của công nghệ NGS trong lĩnh vực thủy sản trên đây, cùng với ưu thế về giá thành thấp (nếu tính trên dữ liệu xuất ra của một lần chạy), thời gian phân tích nhanh và chính xác cho thấy công nghệ NGS có nhiều triển vọng ứng dụng cho nhiều mục đích nghiên cứu trong lĩnh vực thủy sản ở ĐBSCL.

4 ỨNG DỤNG NGS TRÊN THỰC VẬT

Trong những năm gần đây, giá thành giải trình tự các đoạn gen ngày càng giảm đáng kể thì việc giải trình tự gen, bộ gen động – thực vật ngày càng phổ biến. Đây được xem như là xu hướng tiếp cận nghiên cứu khoa học trong nông nghiệp. Những tiến bộ mang tính cách mạng trong công nghệ giải trình tự DNA là sự ra đời của kỹ thuật giải trình tự thế hệ mới (NGS). Kỹ thuật NGS cho phép giải mã chuỗi trình tự hàng triệu nucleotide chỉ trong 1 phản ứng với chi phí ít hơn rất nhiều so với phương pháp giải trình tự Sanger truyền thống. Chi phí và khả năng những công nghệ này tiếp tục đang được cải thiện, cùng với sự phát triển song song của nền tảng NGS kết hợp với các dữ liệu sinh học đang có hiện nay cho phép chúng ta có thể trả lời những câu hỏi khác nhau mà trước đây không thể tìm được đáp án. Khi các kỹ thuật và bộ xử lý dữ liệu tiếp tục được cải thiện và phát triển, nhà khoa học không chỉ làm việc

trên một loài mẫu (model organism), sức mạnh của NGS còn xa hơn thế. Công nghệ NGS và sự phát triển những ứng dụng trong các lĩnh vực nông nghiệp như phát triển marker phân tử, lai tạo, nghiên cứu phát sinh chủng loài và hệ sinh thái, trình tự phiên mã, di truyền đa bội, biến dị di truyền và ứng dụng trong ngân hàng gen.

4.1 NGS trong giải mã trình tự bộ gen thực vật

Bộ gen thực vật đầu tiên được giải mã và công bố vào năm 2000, đó là bộ gen của một loài cây có hoa nhỏ thuộc họ Cải (*Arabidopsis thaliana*) (*Arabidopsis genome initiative*, 2000). Hai năm sau đó, bộ gen của cây lúa đã được giải mã (Goff *et al.*, 2002; Yu *et al.*, 2002). Tại thời điểm bấy giờ, trên cơ sở dữ liệu sinh học (NCBI) cũng cung cấp thông tin một vài bộ gen thực vật đang được tiến hành giải mã. Sau khi bộ gen đầu tiên của thực vật được công bố đã mở ra một kỉ nguyên mới trong nghiên cứu bộ gen thực vật. Tuy nhiên, vì những giới hạn về kỹ thuật và giá thành lúc đó, việc tiến hành giải mã hoàn chỉnh bộ gen chỉ tập trung vào các loài mẫu (model species) với kích thước bộ gen nhỏ hoặc các loài có giá trị kinh tế cao. Cùng với sự phát triển của công nghệ, những kỹ thuật mới sẽ làm thay đổi cách nhà khoa học tiếp cận với các dự án giải trình tự trong tương lai.

Trên thế giới, hiện có hơn 260.000 loài thực vật được biết đến, vì thế việc cần thận lựa chọn bộ gen nào sẽ được giải mã là cần thiết. Việc lựa chọn ưu tiên này sẽ ảnh hưởng đến mục tiêu nghiên cứu và so sánh các bộ gen trong thực vật. Việc tìm hiểu bộ gen sẽ mang đến lợi ích nhất định trong tương lai, ít nhất là 3 lĩnh vực sau:

(1) Việc so sánh các trình tự của bộ gen sẽ cung cấp cơ hội cho các nhà nghiên cứu hiểu được sự tiến hóa của thực vật ở mức độ phân tử.

(2) Cách tiếp cận các gen mục tiêu, các thành phần chức năng của gen và cung cấp dữ liệu quan trọng về những chủ thích cho các bộ gen đã được giải mã.

(3) Trình tự bộ gen thực vật cung cấp những thông tin quan trọng, hữu ích sẽ được ứng dụng trong việc ly trích những gen mục tiêu ở các loài mới.

Theo thống kê, hiện nay đã có hơn 200 loài thực vật đã được giải mã bộ gen, và con số này sẽ không ngừng tăng lên (Royal Botanical Gardens, 2017). Trong đó, 57,7% bộ gen của cây trồng được giải mã, 17,7 % cây trồng hoang dại, 22,3% nhóm thực vật có bó mạch không thuộc nhóm cây trồng và 2,3% còn lại được giải trình tự bộ gen với các mục đích khác nhau. Kích thước các bộ gen thực vật đã được

giải mã cũng vô cùng phong phú. Bộ gen thực vật nhỏ nhất được công bố có kích thước 82 Mb thuộc về loài thực vật ăn thịt sống dưới nước hoặc những vùng ngập nước (*Utricularia gibba*). Bộ gen có kích thước lớn nhất là 19.600 Mb đã được ghi nhận thuộc về một loài thực vật hạt trần thuộc họ thông có tên

thường gọi là Vân sam Na-uy (Norway Spruce) (*Picea abies*), kích thước này chênh lệch khá nhiều so với bộ gen của cây ngô được cho là lớn thứ hai trong các loài thực vật (2.300 Mb) (Michael and Jackson, 2013).

Bảng 6: Danh sách một số cơ sở dữ liệu bộ gen thực vật

STT	Cơ sở dữ liệu	Trang web
1	Plant GDB	http://www.plantgdb.org/
2	Wheatgenome.info	http://www.wheatgenome.info/
3	Phytozome	http://www.phytozome.net
4	EuroPineDB	http://www.scbi.uma.es/pindb/
5	TreeGenes	http://dendrome.ucdavis.edu/treegenes/
6	Gramene	http://www.gramene.org/
7	Sol Genomics	http://github.com/solgenomics/ ; http://solgenomics.net/
8	Legume Information System	http://legumeinfo.org

Công nghệ giải trình tự gen thế hệ mới và kỹ thuật phân tích các dữ liệu này đã phát triển một cách nhanh chóng trong vài năm qua. Trình tự gen đơn hay trình tự toàn bộ bộ gen đã được giải mã sẽ cung cấp rất nhiều thông tin hữu ích cho các nhà khoa học, giúp rút ngắn thời gian nghiên cứu. Thách thức còn lại là làm thế nào sử dụng triệt để và hiệu quả từ nguồn dữ liệu khổng lồ này mang lại.

ĐBSCL hiện nay cũng không đứng ngoài sự tiến bộ khoa học công nghệ. Nhiều viện – trường đã đầu tư các hệ thống thiết bị hiện đại phù hợp với xu hướng phát triển. Bên cạnh đó nguồn thực vật cần được nghiên cứu vô cùng phong phú từ cây trồng mang lại giá trị kinh tế cao đến các cây dược liệu và một số loài thực vật địa phương đặc thù, quý hiếm. Nguồn vật liệu nghiên cứu này sẽ cung cấp rất nhiều thông tin, dữ liệu khoa học quý báu.

4.2 NGS trong công tác chọn giống

Kỹ thuật giải trình tự thế hệ mới (next-generation sequencing – NGS) được ứng dụng giải mã bộ gen thực vật tạo ra cơ hội cho các nhà nghiên cứu hiểu được các đặc tính, cấu trúc, và thành phần các gen trong bộ gen thực vật. Trình tự toàn bộ bộ gen thực vật, trình tự phiên mã (transcriptome sequencing) và trình tự vùng mã hóa protein của các gen (exome sequencing) sẽ cho phép các nhà khoa học khám phá các gen mục tiêu đóng vai trò quan trọng trong nông học, ví dụ như: gen điều khiển năng suất của cây, gen có khả năng chống chịu stress sinh học và phi sinh học. Những dữ liệu có sẵn của bộ gen và các trình tự phiên mã sẽ giúp ích đồng thời cho việc phát triển các marker di truyền. Hiện nay, các công cụ NGS đủ mạnh để cung cấp phân tích độ phân giải cao của bộ gen thực vật. NGS tạo ra một lượng lớn dữ liệu được sắp xếp một cách hiệu quả về chi phí và cho phép lược tả sự biến đổi nucleotide và khám phá một lượng lớn các marker chức năng.

Những marker này sẽ giúp ích cho công tác chọn giống trong việc chọn ra các đặc tính có giá trị kinh tế cao.

5 KẾT LUẬN

Công nghệ giải trình tự thế hệ mới mở ra nhiều cơ hội để chúng ta khám phá nhiều loài sinh vật mới cũng như các bệnh lý phức tạp đang xuất hiện ngày càng nhiều. Nhiều công ty đã sáng chế ra nhiều hệ thống giải trình tự mới với nhiều tính năng mới cũng như đáp ứng nhu cầu của từng ứng dụng cụ thể với giá cả phù hợp và độ chính xác cao. Do đó, việc ứng dụng NGS ở ĐBSCL trong y học, thủy sản và trong nghiên cứu chọn tạo giống thực vật là hoàn toàn khả thi.

TÀI LIỆU THAM KHẢO

Abou, D.S. and Mahfouz, R., 2018. Circulating tumor DNA, liquid biopsy, and next generation sequencing: A comprehensive technical and clinical applications review. *Meta Gene*. 17: 192–201.

Adashek, J.J., Salgia, M. and Pal, S.K., 2018. Next Generation Sequencing in metastatic castrate-resistant prostate cancer. *Urology Case Reports*. 19: 60–62.

Agarwal, A., Sayres, L.C., Cho, M.K., Cook-Deegan, R., and Chandrasekharan, S., 2013. Commercial landscape of noninvasive prenatal testing in the United States. *Prenatal Diagnosis*. 33: 521–531.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 408796–408815.

Avarre, J.C., Madeira, J.P., Santika, A., et al., 2011. Investigation of cyprinid herpesvirus-3 genetic diversity by a multi-locus variable number of tandem repeats analysis. *Journal of Virological Methods*. 173(2): 320–27.

- Barnes, A.C., Delamare-Deboutteville, J., Gudkovs, N., Brosnahan, C., Morrison, R., and Carson, J., 2016. Whole genome analysis of *Yersinia ruckeri* isolated over 27 years in Australia and New Zealand reveals geographical endemism over multiple lineages and recent evolution under host selection. *Microbial Genomics*. 2(11): 1–17.
- Bayliss, S.C., Verner-Jeffreys, D.W., Bartie, K.L., et al., 2017. The promise of whole genome pathogen sequencing for the molecular epidemiology of emerging aquaculture pathogens. *Frontiers in Microbiology*. 8(2): 1–18.
- Beta, J., Lesmes-Heredia, C., Bedetti, C., and Akolekar, R., 2018. Risk of miscarriage following amniocentesis and chorionic villus sampling: a systematic review of the literature. *Minerva Ginecol*. 70: 215–219.
- Bombarely, A., Menda, N., Tecle, I.Y., et al., 2011. The sol genomics network (solgenomics.net): growing tomatoes using perl. *Nucl. Acids Res*. 39: D1149-1155.
- Côté, I., Navarro, S., Tang, K.F., Noble, J.B., and Lightner, D.V., 2008. Taura syndrome virus from Venezuela is a new genetic variant. *Aquaculture*. 284(1–4): 62–67.
- Chitty, L.S., Mason, S., Barrett, A.N., et al., 2015. Non-invasive prenatal diagnosis of achondroplasia and thanatophoric dysplasia: next-generation sequencing allows for a safer, more accurate, and comprehensive approach. *Prenat. Diagn*. 35: 656–662.
- Dahl, E. and Kloten, V., 2015. Liquid biopsy analysis using cell-free DNA (cfDNA): Opportunities and limitations. *Pathologie*. 36: 572–578.
- Dash S., Campbell, J.D., Cannon, E.K., et al., 2016. Legume information system (LegumeInfo. org): A key component of a set of federated data resources for the legume family. *Nucl. Acids Res*. 44: D1181-D1188.
- Database and portal for wheat genome information, 2012. *Plant Cell Physiol*. 53: e2.
- Du, F., Gangchun, X., Zhijuan, N., Pao, X., and Ruobo, G., 2014. Transcriptome analysis gene expression in the liver of coilia nasus during the stress response. *BMC Genomics*. 15(1): 1–11.
- Dương Thúy Yên, Nguyễn Kiệt, Bùi Sơn Nền, Nguyễn Văn Thương, Nguyễn Bạch Loan và Trần Đắc Định, 2016. DNA mã vạch và đặc điểm hình thái của cá bông lau (*Pangasius krempfi*), cá tra bản (*P. mekongensis*) và cá dứa (*P. elongatus*). *Tạp chí Công nghệ sinh học*. 14(1): 29-37.
- Dương Thúy Yên, Nguyễn Phương Thảo, Tiêu Văn Út và Trần Đắc Định, 2018. Đa dạng di truyền của cá hường (*Helostoma temminckii*) ở đồng bằng sông Cửu Long. *Tạp chí Khoa học Trường Đại học Cần Thơ*. 7B (2018): 86-93.
- Dương Thúy Yên, 2014. So sánh trình tự một số gene mã vạch của cá rô đầu vuông và cá rô đồng tự nhiên (*Anabas testudineus* Bloch, 1972). *Can Tho University Journal of Science*. 30: 29-36.
- FAO, 2016. Cultured aquatic species information programme *Macrobrachium rosenbergii* (De Man, 1879). *FAO*. 455(65): 447–55.
- Fernandez-Pozo, N., Menda, N., Edwards, J.D., et al., 2015. The sol genomics network (SGN)-from genotype to phenotype to breeding. *Nucl. Acids Res*. 43: D1036-D1041.
- Fernández-Pozo, N., Canales, J., Guerrero-Fernández, D., Villalobos, D.P., and Díaz-Moreno, S.M., 2011. EuroPineDB: a high-coverage web database for maritime pine transcriptome. *BMC Genom*. 12:366.
- Gardner, S.A., Weymouth, K.S., Kelly, W.S., et al., 2018. Evaluation of a 27-gene inherited cancer panel across 630 consecutive patients referred for testing in a clinical diagnostic laboratory. *Hereditary Cancer in Clinical Practice*. 16: 1.
- Geng, X., Sha, J., Liu, S., Bao, L., Zhang, J., and Wang, R., 2015. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance. *BMC Genomics*. 16(1): 1–12.
- Goff, S.A., Ricke, D., Lan, T.H., et al., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science*. 296: 92–100.
- Gonzales, M.D., Archuleta, E., Farmer, A., et al., 2005. The Legume Information System (LIS): an integrated information resource for comparative legume biology. *Nucl. Acids Res*. 33: D660-D665.
- Hill, M., Twiss, P., Verhoef, T.I., et al., 2015. Non-invasive prenatal diagnosis for cystic fibrosis: detection of paternal mutations, exploration of patient preferences and cost analysis. *Prenat. Diagn*. 35: 950–958.
- Hoa, T.T.T., Zwart, M.P., Phuong, N.T., De Jong, M.C.M., and Vlask, J.M., 2012a. Low numbers of repeat units in variable number of tandem repeats (vntr) regions of white spot syndrome virus are correlated with disease outbreaks. *Journal of Fish Diseases*. 35(11): 817–26.
- Hoa, T.T.T., Zwart, M.P., Phuong, N.T., De Jong, M.C.M., and Vlask, J.M., 2012b. Indel-II region deletion sizes in the white spot syndrome virus genome correlate with shrimp disease outbreaks in Southern Vietnam. *Diseases of Aquatic Organisms*. 99(2): 153–62.
- Huete-Pérez, J.A. and Quezada, F., 2013. Genomic approaches in marine biodiversity and aquaculture. *Biological Research*. 46(4): 353–61.
- Kotsopoulou, I., Tsoplou, P., Mavrommatis, K., and Kroupis, C., 2015. Non-invasive prenatal testing (NIPT): limitations on the way to become diagnosis. *Diagnosis*. 2: 141–158.

- Kulski, J.K., 2016. Next-Generation Sequencing — An overview of the history, tools, and “omic” applications, Next Generation Sequencing, ed. Jerzy K.K., IntechOpen, DOI: 10.5772/61964.
- Kumar, G. and Kocour, M., 2017. Applications of next-generation sequencing in fisheries research: A Review. *Fisheries Research*. 186: 11–22.
- Lench, N., Barrett, A., Fielding, S., *et al.*, 2013. The clinical implementation of non-invasive prenatal diagnosis for single-gene disorders: challenges and progress made. *Prenat. Diagn.* 33: 555–562.
- Lê Văn Hậu, Lê Lư Phương Hạnh, Ngô Huỳnh Phương Thảo, Nguyễn Phúc Cẩm Tú và Nguyễn Quốc Bình, 2018. Ứng dụng marker phân tử DNA barcode trong định danh các mẫu *Moina* spp. phân lập tại khu vực Đồng bằng sông Cửu Long. *Tạp chí Khoa học Trường Đại học Cần Thơ*. 54 (Số chuyên đề: Thủy sản). (2): 36-44.
- Liao, G.J.W., Gronowski, A.M., and Zhao, Z., 2014. Non-invasive prenatal testing using cell-free fetal DNA in maternal circulation. *Clin. Chim. Acta* 428: 44–50.
- Mandel, P. and Metais, P., 1948. Les acides nucléiques du plasma sanguin chez l’homme. *C. R. Seances Soc. Biol. Fil.* 142: 241–243.
- Marcos-López, M., Gale, P., Oidtmann, B.C., and Peeler, E.J., 2010. Assessing the impact of climate change on disease emergence in freshwater fish in the United Kingdom. *Transboundary and Emerging Diseases*. 57(5): 293–304.
- Mehinto, A.C., Martyniuk, C.J., Spade, D.J., and Denslow, N.D., 2012. Applications of Next-Generation Sequencing in fish ecotoxicogenomics. *Frontiers in Genetics*. 3: 1–10.
- Mellis, R., Chandler, N., and Chitty, L.S., 2018. Next generation sequencing and the impact on prenatal diagnosis. *Expert Review of Molecular Diagnostics*. 18(8): 689-699.
- Michael, T.P. and Jackson S., 2013. The first 50 plant genomes. *Plant Genome*. 6: 1-7.
- Motro, Y. and Moran-Gilad J., 2017. Next-generation sequencing applications in clinical bacteriology. *Biomolecular Detection and Quantification*, 14: 1-6.
- Niederstrasser, S.L., Hammer, K., Möllers, M., *et al.*, 2017. Fetal loss following invasive prenatal testing: a comparison of transabdominal chorionic villus sampling, transcervical chorionic villus sampling and amniocentesis. *J Perinat Med*. 45: 193–198.
- NIH, 2019. All about the human genome project (HGP), truy cập ngày 20/9/2018. <https://www.genome.gov/10001772/all-about-the-human-genome-project-hgp/#al-2>.
- Nkili-Meyong, A.A., Bigarré, L., Labouba, I., Vallaey, T., Avarre, J.C., and Berthet, N., 2017. Contribution of Next-Generation Sequencing to aquatic and fish virology. *Intervirology*. 59(5–6): 285–300.
- Norton, M.E., Brar, H., Weiss, J., *et al.*, 2012. Non-Invasive Chromosomal Evaluation (NICE) Study: results of a multicenter prospective cohort study for detection of fetal trisomy 21 and trisomy 18. *American Journal of Obstetrics and Gynecology*. 207: 137.e1-137.e8.
- Nguyễn Minh Thành, Võ Thị Minh Thư, Hyungtaek Jung và Peter Mather, 2015. Phân tích hệ gen chức năng từ mô thận cá tra (*Pangasianodon hypophthalmus*) nuôi ở điều kiện mặn: lập ráp, chú giải, phân tích chỉ thị SNP. *Tạp chí sinh học*. 37(2): 220-227.
- Nguyễn Phương Thảo và Dương Thúy Yên, 2015. So sánh đặc điểm hình thái và dna mã vạch của hai loài cá bống trần *Butis butis* và *Butis humeralis*.” *Tạp Chí khoa học Trường Đại học Cần Thơ*. 30: 29–36.
- Nguyễn Thị Thu Hằng và Đặng Thị Hoàng Oanh, 2015. Xác định mầm bệnh vi bào tử trùng (Microsporida) nhiễm trong cơ cá tra (*Pangasianodon hypophthalmus*). *Tạp chí Khoa học Trường Đại học Cần Thơ*. 42(2016): 101-110.
- Nguyễn Thu Dung, Lê Thanh Cần và Đặng Thị Hoàng Oanh, 2015. Phát hiện vi khuẩn *Streptococcus dusgalactiae* gây bệnh xuất huyết trên cá bống kèo (*Pseudapocryptes elongatus*) bằng phương pháp PCR. *Tạp chí Khoa học Trường Đại học Cần Thơ*. 42(2016): 111-117.
- Ordas, A., Hegedus, Z., Henkel, C.V., *et al.*, 2011. Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. *Fish and Shellfish Immunology*. 31(5): 716–24.
- Palaiokostas, C., Bekaert, M., Khan, M.G.Q., Taggart, J.B., and Gharbi, K., 2013. Mapping and validation of the major sex-determining region in Nile Tilapia (*Oreochromis niloticus* L.) using RAD sequencing. *PLoS ONE*. 8(7): 1–9.
- Palaiokostas, C., Bekaert, M., Khan, M.G.Q., Taggart, J.B., and Gharbi, K., 2015. A novel sex-determining QTL in Nile Tilapia (*Oreochromis niloticus*). *BMC Genomics*. 16(1): 1–10.
- Palomaki, G.E., Kloza, E.M., Lambert-Messerlian, G.M., *et al.*, 2011. DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. *Genet. Med*. 13: 913–920.
- Paolillo, C., Londin, E. and Fortina, P., 2016. Next generation sequencing in cancer: opportunities and challenges for precision cancer medicine. *Scandinavian Journal of Clinical and Laboratory Investigation*. 76: S84–S91.
- Parks, M., Court, S., Cleary, S., *et al.*, 2016. Non-invasive prenatal diagnosis of Duchenne and Becker muscular dystrophies by relative haplotype dosage. *Prenat. Diagn.* 36: 312–320.

- Patrick, K., 2008. 454 life sciences: illuminating the future of genome sequencing and personalized medicine. The Yale journal of biology and medicine. 80(4): 191-4.
- Price, K.S., Svenson, A., King, E., Ready, K. and Lazarin, G.A., 2018. Inherited cancer in the age of next-generation sequencing. Biological Research For Nursing. 20: 192–204.
- Phạm Thị Trang Nhung và Dương Thúy Yên. 2014. Đánh giá sự đa dạng di truyền của các dòng cá rô đồng (*Anabas testudineus*, bloch 1972) bằng các chỉ thị phân tử RAPD và ISSR. Tạp chí Khoa học Trường Đại học Cần Thơ. 1: 101-108.
- Rahner, N., Steinke, V., Schlegelberger, B., Eisinger, F., Hutter, P. and Olschwang, S., 2013. Clinical utility gene card for: Lynch syndrome (MLH1, MSH2, MSH6, PMS2, EPCAM) - update 2012. European Journal of Human Genetics. 21.
- Rahner, N., Steinke, V., Schlegelberger, B., Olschwang, S., Eisinger, F. and Hutter, P., 2010. Clinical utility gene card for: Lynch syndrome (MLH1, MSH2, MSH6, PMS2). European Journal of Human Genetics. 18: 1071–1071.
- Renga, B., 2018. Non invasive prenatal diagnosis of fetal aneuploidy using cell free fetal DNA. Eur. J. Obstet. Gynecol. Reprod. Biol. 225: 5–8.
- Robledo, D., Palaikostas, C., Bargelloni, L., Martínez, P., and Houston R., 2017. Applications of genotyping by sequencing in aquaculture breeding and genetics. Reviews in Aquaculture. 10(3):670-682.
- Ronkin, D., Seroussi, E., Nitzan, T., Doron-Faigenboim, A., and Cnaani, A., 2015. Intestinal transcriptome analysis revealed differential salinity adaptation between two tilapiine species. Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics. 13: 35–43.
- Royal Botanical Garden, 2017. Overview of the state of the World's plants 2017. Assessed on 8/6/2018. <https://stateoftheworldsplants.org/>
- Saramago, P., Yang, H., Llewellyn, A., *et al.*, 2018. High-throughput non-invasive prenatal testing for fetal rhesus D status in RhD-negative women not known to be sensitised to the RhD antigen: a systematic review and economic evaluation. Health Technol Assess. 22: 1–172.
- Sausen, M., Parpart, S., and Diaz, L.A., 2014. Circulating tumor DNA moves further into the spotlight. Genome Med. 6: 35.
- Shao, C., Niu, Y., Rastas, P., *et al.*, 2015. Genome-Wide SNP identification for the construction of a high-resolution genetic map of Japanese flounder (*Paralichthys olivaceus*): Applications to QTL mapping of vibrio anguillarum disease resistance and comparative genomic analysis. DNA Research. 22(2): 161–70.
- Stanislaw, C., Xue, Y., and Wilcox, W.R., 2016. Genetic evaluation and testing for hereditary forms of cancer in the era of next-generation sequencing. Cancer Biol Med. 13: 55–67.
- Star, B., Nederbragt, A.J., Jentoft, S., *et al.*, 2011. The genome sequence of Atlantic cod reveals a unique immune system. Nature. 477(7363): 207–10.
- Taneja, P.A., Snyder, H.L., de Feo, E., *et al.*, 2016. Noninvasive prenatal testing in the general obstetric population: Clinical performance and counseling considerations in over 85 000 cases: NIPT in the general obstetrics population. Prenatal Diagnosis. 36: 237–243.
- Van, B.E.T., Thong, N.V., Kieu, N.T.T, and Hang, Đ.T.T., 2017. Primary study on application of cell-free fetal DNA (cffDNA) in non-invasive prenatal testing of β -thalassemia. VNU Journal of Science: Natural Sciences and Technology. 33:4546.
- Verhoef, T.I., Hill, M., Drury, S., *et al.*, 2016. Non-invasive prenatal diagnosis (NIPD) for single gene disorders: cost analysis of NIPD and invasive testing pathways. Prenat. Diagn. 36: 636–642.
- Walker, P.J. and Winton, J.R., 2010. Emerging viral diseases of fish and shrimp. Veterinary Research. 41(6): 51.
- Wei, X., Ju, X., Yi, X., *et al.*, 2011. Identification of sequence variants in genetic disease-causing genes using targeted Next-Generation Sequencing. PLoS ONE. 6: e29500.
- Wijegoonawardane, P.K.M., Sittidilokratna, N., Petchampai, N., Cowley, J.A., Gudkovs, N., and Walker, P.J., 2009. Homologous genetic recombination in the yellow head complex of nidoviruses infecting *Penaeus monodon* shrimp. Virology. 390(1): 79–88.
- WHO, 2007. Meeting on Management of Haemoglobin Disorders. Accessed on 12/7/2018. https://www.who.int/genomics/WHO-TIF_genetics_final.pdf
- Xia, J.H., Wan, Y.Z., Zhen, L., *et al.*, 2014. Genome-Wide discovery and in silico mapping of gene-associated SNPs in Nile Tilapia. Aquaculture. 432: 67–73.
- Xiong, L., Barrett, A.N., Hua, R., *et al.*, 2015. Non-invasive prenatal diagnostic testing for β -thalassaemia using cell-free fetal DNA and next generation sequencing. Prenat. Diagn. 35: 258–265.
- Xu, P., Zhang, X., Wang, *et al.*, 2014. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. Nature Genetics. 46(11): 1212–19.
- Xue, Y., and Wilcox, W.R., 2016. Changing paradigm of cancer therapy: precision medicine by next-generation sequencing. Cancer Biol Med. 13: 12–18.
- Yadav, N.K., Shukla, P., Omer, A., Pareek, S., and Singh, R.K., 2014. Next generation sequencing: Potential and Application in Drug discovery. The Scientific World Journal. 2014: 1-7.

- Yáñez, J.M., Newman, S., and Houston, R.D., 2015. Genomics in aquaculture to better understand species biology and accelerate genetic progress. *Frontiers in Genetics*. 6: 1–3.
- Yue, G.H. and Wang, L., 2017. Current status of genome sequencing and its applications in aquaculture. *Aquaculture*. 468: 337–47.
- Zhang, G., Fang, X., Guo, X., *et al.*, 2012. The oyster genome reveals stress adaptation and

complexity of shell formation. *Nature*. 490(7418): 49–54.

- Zhang, Y., Stupka, E., Henkel, C.V., Jansen, H.J., Spalink, H.P., and Verbeek, F.J., 2011. Identification of common carp innate immune genes with whole-genome sequencing and RNA-Seq Data 1 introduction." *Journal of Integrative Bioinformatics*. 8(2): 1–11.