

ỨNG DỤNG PAN-GENOME TRONG SÀNG LỌC CÁC MARKER DÙNG PHÁT TRIỂN VACCIN ĐẢO NGƯỢC PHÒNG BỆNH EDWARDSIELLOSIS TRÊN CÁC LOÀI THỦY HẢI SẢN

Nguyễn Thành Luân

Viện Khoa học ứng dụng HUTECH, Đại học Công Nghệ Tp. Hồ Chí Minh

Sự gia tăng và biến đổi đa dạng kiểu hình của các chủng vi khuẩn gây bệnh thủy sản bao gồm chi *Edwardsiella* đang đặt ra những thách thức lớn cho việc thay đổi liên tục chiến lược sản xuất vaccin phòng dịch bệnh trong nuôi trồng thủy sản. Phương pháp giải trình tự bộ gen và so sánh các bộ gen của vi khuẩn bằng các công cụ pan-genome đang được xem là một công cụ hiệu quả hỗ trợ mở rộng để phân tích các vi sinh vật gây bệnh thủy sinh và cách thức các đặc điểm linh động và sự thích nghi với nhiều loài vật chủ cũng như các hệ sinh thái vi môi trường (niches).

Mục đích chính của việc phân tích pan-genome là so sánh bộ gen của các chủng vi khuẩn khác nhau trong một loài (intra-species) hoặc chi/giống (inter-species) (Snipen và ctv., 2009). Nhìn chung, phân tích pan-genome cho phép hiểu về sự tiến hóa của loài gây bệnh đa kháng thuốc. Phương pháp này cũng cho phép hiểu rõ hơn về yếu tố di truyền mang khả năng gây bệnh ở nhiều loại vi khuẩn bằng cách so sánh với các chủng độc lực và không độc lực khi chỉ sử dụng kết quả phân tích các gen trung tâm. Đặc biệt, sự phân tích các gen/cụm gen (core, dispensable, singleton) kết hợp hệ thống phân tích chức năng như RAST (Rapid Annotation using Subsystem Technology), KEGG, WebMGA (một máy chủ web tùy biến để phân tích các chuỗi trình tự) có thể giúp xác định các yếu tố quan trọng xuất hiện trong khả năng gây độc của chủng độc lực. Ngoài ra, khả năng chuyển hóa đường phức tạp có thể được hiển thị bằng các so sánh sự phân bố của các enzyme sử dụng đường (cấu hình Cazy, carbohydrate-active enzymes) giữa các phân lập. Thông tin Cazy đặc trưng là một lợi thế của quá trình chọn lọc cho phép vi sinh vật thích

nghi với hệ sinh thái của chúng.

Kết quả phân tích pan-genome trước đây của chúng tôi (Nguyen và Kim 2018; Nguyen và ctv., 2018) đã chỉ ra rằng chủng WFLU12 phân lập từ cá biển biểu hiện các đặc tính đặc thù cho việc sản xuất, chuyển đổi năng lượng, sự vận chuyển và chuyển hóa đường bằng cách phân tách các gen đặc trưng cho từng chủng. Bên cạnh đó, dựa vào các gen phân tách từ pan-genome, các gen được chọn hoàn toàn có thể được kiểm tra lại chức năng trên cơ sở dữ liệu các gen độc lực (VFDB), các nhóm gen chức năng (COG), từ điển bách khoa về gen và gen của Kyoto (KEGG) và cơ sở dữ liệu gen kháng kháng sinh (ARGB). Do đó, dữ liệu bộ gen của vi sinh vật mang lại những hiểu biết đáng kể về sự hiểu biết về sự tiến hóa của vi khuẩn, sự thích nghi thích hợp, cấu trúc quần thể và tương tác vật chủ. Những nghiên cứu này cũng có thể được áp dụng cho các vấn đề như xác định gen độc lực, thiết kế vaccin và thuốc (Chaplin và ctv., 2015).

1. Khai thác dữ liệu bộ gen của vi khuẩn gây bệnh thủy sản

1.1. Sự cần thiết của pan-genome trong phát triển vaccin đảo ngược

Trong số các bệnh truyền nhiễm được chẩn đoán ở loài cá tra (*Pangasianodon hypophthalmus*) ở Việt Nam, bệnh gan thận mù gây ra bởi *Edwardsiella ictaluri* là bệnh xảy ra thường xuyên nhất (Ferguson và ctv., 2001). Bệnh xảy ra trong các hệ thống sản xuất cá ở mọi lứa tuổi. Điều này dẫn đến hệ quả thiệt hại kinh tế nghiêm trọng vì cá chết hàng loạt, sản lượng cá giảm, chi phí điều trị tăng (Ferguson et al., 2001). Đối với cá bột ương giống, tỷ lệ

chết có thể lên đến 100%, và từ 30 - 50% đối với cá thịt (Dung và ctv., 2010). Việc phân lập các loài *Edwardsiella* spp. từ các nguồn khác nhau cho thấy chúng có khả năng thích nghi với những điều kiện thay đổi của môi trường (ví dụ: nhiệt độ, kháng sinh, vật chủ) đang đặt ra thách thức cho việc phát triển các phương pháp phòng Edwardsiellosis bằng vaccin an toàn và có khả năng phản ứng chéo.

Sự gia tăng về dữ liệu genomic và proteomic là điều kiện thuận lợi cho quá trình thiết kế và phát triển vaccin. Sử dụng vaccin đảo ngược được coi là một phương pháp kinh tế và hiệu quả để sàng lọc toàn bộ bộ gen của vi sinh vật gây bệnh. Để tìm kiếm các mục tiêu kháng nguyên có phổ tác dụng rộng và phân tích các loài vi khuẩn có sự liên quan chặt chẽ, việc đồng hóa khái niệm pan-genome vào phương pháp vaccin đảo ngược là rất cần thiết và là xu thế của các nghiên cứu hiện đại. Các nhóm gen có thể phân tích từ pan-genome như gen lõi, gen phân tán và gen đơn lẻ được phân tách để xác định các ứng viên vaccin thông qua quy trình tạo vaccin đảo ngược.

1.2. Các gen mục tiêu trong sàng lọc vaccin đảo ngược

So sánh pan-genome của các nhóm vi sinh vật gây bệnh cho phép dự đoán chức năng các gen mang đặc điểm của protein tiếp xúc bề mặt tế bào (cell surface-exposed proteins - SEPs), bao gồm protein màng ngoài (OEP) và protein ngoại bào (EP). Những gen mã hóa protein bề mặt này đang được sử dụng như là ứng viên vaccin trong các mô hình động vật (được gọi là Reversed Vaccinology). Tương tự, trong vi khuẩn gây bệnh thủy sản, các SEP tham gia làm yếu tố độc lực đóng vai trò chính trong khả năng gây bệnh và gây phản ứng miễn dịch lên vật chủ. Ví dụ, sự biểu hiện của gen *esa1* từ *E. tarda*, như một kháng nguyên bề mặt D15, trong mô hình cá bơn Nhật Bản đã kích thích sự biểu hiện của một phổ rộng các gen có thể liên quan đến khả năng đáp ứng miễn dịch tự nhiên và đặc hiệu, với kết quả bảo vệ cá sống sót tỷ lệ cao và có khả năng sản xuất kháng thể đặc hiệu (Sun

và ctv., 2011). Tiêm vaccin bằng SEPs kích thích sự phát triển các yếu tố có tác dụng bảo vệ cá chống lại bệnh gây chết do *Aeromonas hydrophila* (Khushiramani và ctv., 2012), nhiễm *Flavobacterium columnare* (Luo và ctv., 2016), và Edwardsiellosis (Park và ctv., 2011). Nghiên cứu gần đây (Zeng và ctv., 2017) đã thực hiện kết hợp pan-geome để sàng lọc SEP từ 17 chủng *Leptospira interrogans* đại diện bao gồm các serovar dịch tễ học từ khắp nơi trên thế giới, và 118 kháng nguyên ứng viên mới được xác định là các protein màng ngoài và các lipoprotein. Do đó, sự gia tăng nhanh chóng về số lượng trình tự bộ gen của vi khuẩn gây bệnh thủy sản sẽ là một tiềm năng lớn cho phép xây dựng một quy trình kiểm soát bệnh lây nhiễm cũng như là một xu hướng tiềm năng để nghiên cứu vi khuẩn gây bệnh dưới nước để cải thiện hiệu quả của các loại vaccin ở cá nuôi và ngăn chặn sự bùng phát bệnh khi thực hiện cùng phân tích pan-genome.

2. Ứng dụng phân tích *Edwardsiella* pan-genome trong tìm các ứng viên vaccin

Phân tích pan-genome của chủng *Edwardsiella* là phương pháp để phân tách các gen mã hóa chung cho các protein bề mặt tế bào (SEPs), bao gồm các protein màng ngoài và protein ngoại bào. Những gen này sẽ được xem như là ứng viên để sản xuất vaccin đảo ngược trong một mô hình cá tra khi chúng thỏa mãn các đặc tính bám và bao gồm các vị trí kháng nguyên thông qua các kiểm tra bằng công cụ đặc trưng.

2.1. Mô hình thực hiện sàng lọc các gen SEP từ *Edwardsiella* pan-genome

2.1.1. Chọn lọc dữ liệu

Trong nghiên cứu này, tổng cộng có 15 bộ gen hoàn chỉnh của chủng *Edwardsiella* và được phân lập từ các hệ sinh thái khác nhau được thu thập từ Ngân hàng bộ gen vi khuẩn NCBI (<ftp://ftp.ncbi.nih.gov/genomes/>).

2.1.2. Phân tích *Edwardsiella* pan-genome

Pan-genome của 15 chủng *Edwardsiella* được phân tích bằng hệ thống cơ sở dữ liệu của

EDGAR v2.2 (Blom và ctv, 2016). Theo đó, nhóm gen bao gồm số lượng gen thuộc gen lõi, gen phân tán và gen đơn lẻ, được trích xuất theo các thông số mặc định của chương trình.

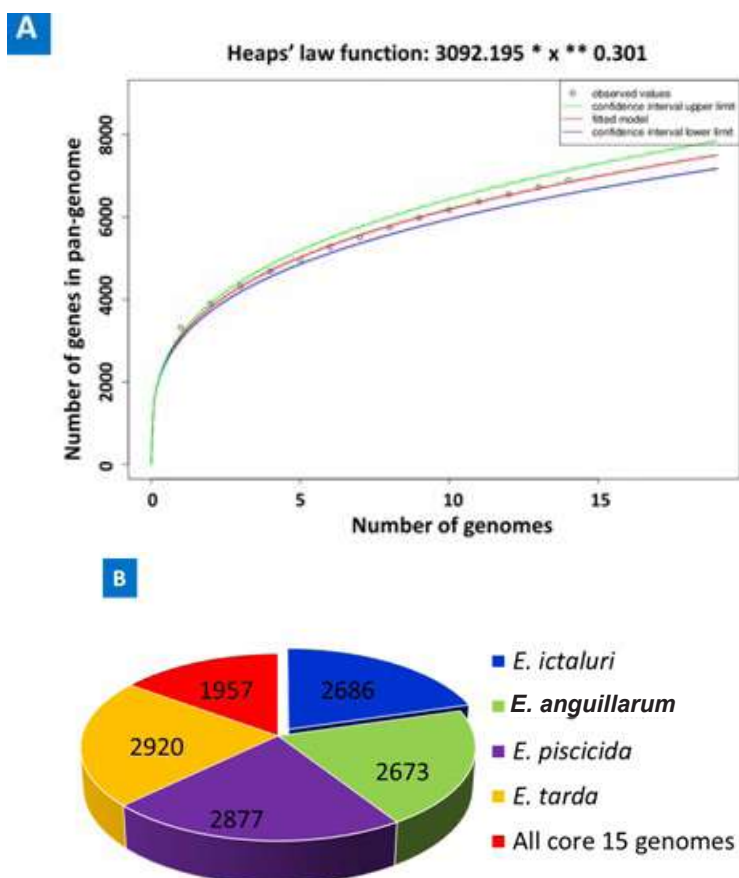
2.1.3. Sàng lọc gen ứng viên vaccin

Để tìm các gen có thể dùng làm ứng viên vaccin, gen lõi của 4 loài xác định từ kết quả phân tích phylogenomics và gen lõi của 15 chủng *Edwardsiella* được phân tách và phân tích lại bằng phần mềm các công cụ khác nhau như PSORTb3.0, CELLO, và SOSUI-GramN để phân tích chức năng gen mã hóa các thành phần tế bào (subcellular localization), đặc tính bám (adhesin probability) bằng công cụ SPAAN software, và dự đoán gen có tính kháng nguyên cao (antigen) bằng VaxiJen server từ các gen lõi

phân tách được. Các gen sàng lọc được tiếp tục khảo sát đặc tính B-cell Linear Epitope bằng IEDB sever (<http://tools.iedb.org/>) và CTL Epitopes bằng CTLPred sever (<http://crdd.osdd.net/>).

2.2. Đặc điểm của các nhóm gen thành phần trong *Edwardsiella* pan-genome

Trong thành phần pan-genome của 15 chủng *Edwardsiella* có 6733 gen mã hóa protein với 29,07% (1957 gen) là gen lõi (core gene), và 70,93% còn lại là gen phân tán (assessor gene) và gen đơn lẻ (singleton gene). Theo Tettelin và ctv., (2005), công thức Heap's Law $n = k * N^{-\alpha}$ với hệ số α trong nghiên cứu này là 0,301 (hình 1A), tương ứng với mô hình pan-genome mở.



Hình 1. Pan-genome của 15 chủng *Edwardsiella*. (A) Số lượng gen trong pan-genome và hệ số α trong công thức Heap's Law. (B) Số lượng gen lõi của 4 loài *E. ictaluri*, *E. anguillarum*, *E. piscicida*, *E. tarda* và của tất cả 15 chủng khảo sát trong nghiên cứu này.

Điều này có nghĩa là việc bổ sung thêm một bộ gen *Edwardsiella* khác sẽ tiếp tục giúp phát hiện các gen mới cho loài này, và cũng cho thấy rằng *Edwardsiella* spp. có thể thích nghi với nhiều điều kiện môi trường khác nhau.

Phân tích sự phát sinh loài từ dữ liệu gen lỗi và hiện diện/vắng mặt của các gen trong dữ liệu gen phân tán cho thấy 15 chủng *Edwardsiella* được chia ra thành 5 nhóm loài như sau: *E. ictaluri* bao gồm chủng 93-146, RUSVM-1 và MS-17-156; *E. anguillarum* bao gồm ET080813, EA181011 và LADL05-105; *E. piscicida* bao gồm C07-087, S11-285, ETW41, ET-1, EIB202 và FL6-60 (với ET-1, EIB202, FL6-60 trước đây được định danh là *E. tarda*); *E. tarda* bao gồm FL95-01 và KC-Pc-HB1; *E. hoshinae* ATCC_35051. Các gen lỗi của 4 loài *E. ictaluri*, *E. anguillarum*, *E. piscicida*, *E. tarda* và 15 chủng khảo sát được phân tách (hình 1B) và sử dụng để sàng lọc các gen ứng viên làm vaccin. Đối với loài *E. hoshinae*, dữ liệu NCBI cho thấy chỉ có 1 bộ gen hoàn chỉnh được công bố tại thời điểm phân tích, vì vậy chúng không

được chọn để phân tách gen lỗi trong nghiên cứu này. Tuy nhiên, *E. hoshinae* cũng được xem là tác nhân có khả năng gây bệnh (Janda và ctv., 1991) nên chúng được đưa vào phân tích với tất cả các chủng còn lại để tìm ra các gen liên quan hướng tới phát triển vaccin đa kháng.

2.3. Kết quả sàng lọc các gen mã hóa SEPs

Trong nghiên cứu này, các gen mã hóa các loại protein bề mặt tế bào (SEPs) bao gồm protein màng ngoài (outer membrane), protein ngoại bào (extracellular) được phân tích từ dữ liệu các gen lỗi. Kết quả cho thấy, 73, 69, 72, 77 và 96 gen liên quan SEPs được tuyển chọn từ 2686, 2673, 2877, 2920, và 1957 gen lỗi tương ứng của *E. ictaluri*, *E. anguillarum*, *E. piscicida*, *E. tarda* và 15 chủng *Edwardsiella* khảo sát. Các gen được tiếp tục sàng lọc với các đặc tính B-cell Linear Epitope và CTL Epitopes gồm 9 gen của *E. ictaluri*, 13 gen của *E. anguillarum*, 9 gen của *E. piscicida*, 12 gen của *E. tarda* và 14 gen từ 15 chủng *Edwardsiella* khảo sát có tiềm năng sử dụng trong nghiên cứu tạo dòng và sản xuất vaccin đảo ngược (bảng 1).

Bảng 1. Một số ứng viên vaccin tiềm năng dùng phát triển vaccin phòng Edwardsiellosis sàng lọc trong nghiên cứu này

STT	Protein Accession	Protein length	BLAST with VFDB_set A	Loài	Chức năng
1	AAW15_11440	859	VFG000456	<i>Edwardsiella_tarda</i>	pilus assembly protein
2	A9798_03900	190	VFG000442	<i>Edwardsiella_hoshinae</i>	virulence protein
3	A9798_10515	367	VFG043487	<i>Edwardsiella_hoshinae</i>	porin
4	A9798_11675	177	VFG000456	<i>Edwardsiella_hoshinae</i>	pilus assembly protein
5	AAW15_03775	190	VFG000442	<i>Edwardsiella_tarda</i>	virulence protein
6	AAW15_10465	386	VFG043487	<i>Edwardsiella_tarda</i>	porin
7	ETEE_0777	190	VFG000442	<i>Edwardsiella_anguillarum</i>	Attachment invasion locus protein precursor
8	ETEE_2462	375	VFG043487	<i>Edwardsiella_anguillarum</i>	Outer membrane protein C precursor
9	ETEE_3185	376	VFG043487	<i>Edwardsiella_anguillarum</i>	Outer membrane protein C precursor

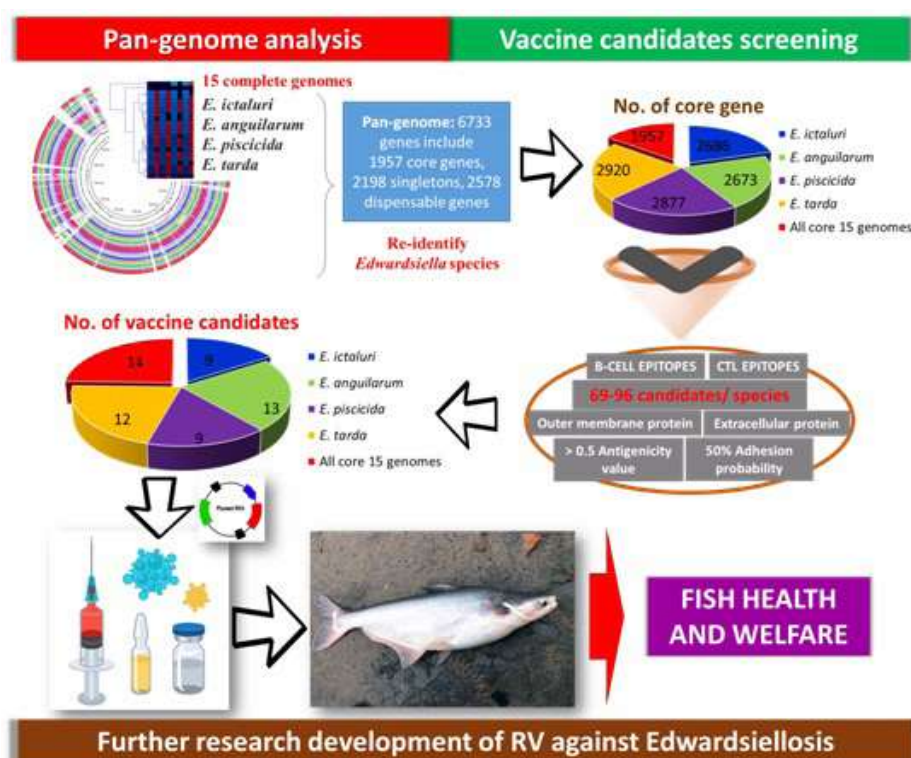
Các SEPs được xem là ứng viên tiềm năng trong mô hình sản xuất vaccin cho động vật và

cá (được gọi là Reversed Vaccinology). Ví dụ, sự biểu hiện của gen *esa1* từ *E. tarda*, một kháng

nguyên bề mặt giống D15, trong mô hình cá bơn Nhật Bản đã tạo ra sự biểu hiện của một phổ rộng các gen có thể liên quan đến khả năng miễn dịch tự nhiên và đặc hiệu, cũng như tăng tỷ lệ sống sót của cá và khả năng sản xuất kháng thể huyết thanh đặc hiệu (Sun và ctv., 2011). Trong nghiên cứu này, ứng dụng pan-genome kết hợp với các công cụ phân tích chức năng và đặc tính của các protein SEPs của chủng *Edwardsiella* bước đầu đã giúp chúng tôi rút ngắn được thời gian để sàng lọc và giảm số lượng các ứng viên tiềm năng nếu chỉ dựa trên trình tự một bộ gen. Trong đó, 9 gen tiềm năng sản xuất vaccin đảo ngược phòng chống nhiễm *E. ictaluri* sẽ đóng vai trò rất quan trọng cho việc phát triển vaccin thử nghiệm phòng bệnh Edwardsiellosis ở cá tra và các loài cá nước ngọt khác ở Việt Nam. Đặc biệt, để nâng cao hiệu quả của một loại vaccin đảo ngược trong phòng bệnh Edwardsiellosis do nhiều loài *Edwardsiella* gây ra, 14 gen tiềm năng trên sẽ là thông tin rất quan trọng cho các nghiên cứu tiếp theo.

4. Tóm tắt

Kết quả của chúng tôi có thể góp phần dự đoán các kháng nguyên phù hợp với đặc tính là ứng cử viên sản xuất vaccin đảo ngược có tiềm năng chống lại nhiều chủng *Edwardsiella* gây bệnh thủy sản và cũng cung cấp những hiểu biết sâu sắc hơn về quy trình chọn lọc các ứng viên (hình 2). Ngoài ra, sự gia tăng nhanh chóng về số lượng genome của vi khuẩn gây bệnh trên thủy sản sẽ cho phép chúng tôi xây dựng một quy trình phản ứng nhanh trong kiểm soát dịch bệnh, và là dữ liệu quý để nghiên cứu cải thiện hiệu quả của các loại vaccin đa giá nhằm ngăn chặn các dịch bệnh có xu hướng biến đổi liên tục trong nuôi trồng thủy sản. Cuối cùng, chiến lược sàng lọc của chúng tôi kết hợp với dữ liệu phân tích pan-genome có thể là một xu hướng nghiên cứu và sản xuất vaccin đảo ngược thông thường được áp dụng cho nhiều tác nhân gây bệnh khác trên cả người và động vật trên cạn.



Hình 2. Sơ đồ sàng lọc ứng viên vaccin đảo ngược phòng Edwardsiellosis kết hợp phân tích pan-genome

Lời cảm ơn: Quỹ nghiên cứu Khoa học và công nghệ, Trường Đại học Công nghệ, TP. HCM (HUTECH) đã tài trợ cho nghiên cứu này.

TÀI LIỆU THAM KHẢO

1. Blom, J., Kreis, S., Spänig, S., Juhre, T., Bertelli, C., Ernst, C., Goesmann, A., 2016. EDGAR 2.0: an enhanced software platform for comparative gen content analyses. *Nucleic Acids Res.* 44, 22–28.
2. Chaplin, A.V., Efimov, B.A., Smeianov, V.V., Kafarskaia, L.I., Pikina, A.P., Shkoporov, A.N., 2015. Intraspecies Genomic Diversity and Long-Term Persistence of *Bifidobacterium longum*. *PloS one* 10(8), e0135658. doi:10.1371/journal.pone.0135658
3. Dung T.T, Haesebrouck F., *et al.*, 2010. Hiện trạng kháng thuốc kháng sinh trên vi khuẩn *Edwardsiella ictaluri* gây bệnh gan, thận mũ trên cá tra (*Pangasianodon hypophthalmus*) ở ĐBSCL. *Tạp chí Khoa học ĐH Cần Thơ*, 5a 162-171.
4. Ferguson, H.W., Turnbull J.F., *et al.*, 2001. Bacillary necrosis in farmed *Pangasius hypophthalmus* (Sauvage) from the Mekong Delta, Vietnam. *J. Fish Dis.* 24:509-513.
5. Janda, J.M., Abbott, S.L., Kroske-Bystrom, S., Cheung, W. K., Powers, C., Kokka, R. P., & Tamura, K., 1991. Pathogenic properties of *Edwardsiella* species. *Journal of clinical microbiology* 29(9), 1997–2001.
6. Snipen, L., Almøy, T., Ussery, D.W., 2009. Microbial comparative pan-genomics using binomial mixture models. *BMC Genomics.* 10-385.
7. Sun, Y., Liu, C. Sun, L., 2011. Construction and analysis of the immune effect of an *E. tarda* DNA vaccine encoding a D15-like surface antigen. *Fish Shellfish Immunol.* 30 273-279.
8. Tettelin, H. Massignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., Deboy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., *et al.*, 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc. Natl. Acad. Sc. USA.* 102 13950–13955.
9. Sun, Y., Liu, C., Sun, L., 2011. Construction and analysis of the immune effect of an *Edwardsiella tarda* DNA vaccine encoding a D15-like surface antigen, *Fish Shellfish Immunol.* 30, 273-279.
10. Khushiramani, R., Maiti, B., Shekar, M., Girisha, S.K., Akash, N. Deepanjali, A., Karunasagar, I., Karunasagar, I., 2012. Recombinant *Aeromonas hydrophila* outer-membrane protein 48 (Omp48) induces protective immune response against *Aeromonas hydrophila* and *Edwardsiella tarda*. *Res. Microbiol.* 163, 286–291.
11. Luo, Z., Fu, J., Li, N., Liu, Z., Qin, T., Zhang, X., *et al.*, 2016. Immunogenic proteins and their vaccine development potential evaluation in outer membrane proteins (OMPs) of *Flavobacterium columnare*, *Aquacult Fish.* 1 1–8.
12. Park, S.B., Jang, H.B., Nho, S.W., Cha, I.S., Hikima, J., Ohtani, M., Aoki, T., Jung, T.S., 2011. Outer membrane vesicles as a candidate vaccine against *Edwardsiella*, *PLoS One* 6 e17629.
13. Zeng, L., Wang, D., Hu, N., Zhu, Q., Chen, K., Dong, K., Zhang, Y., Yao, Y., Guo, X., Chang, Y.F., Zhu, Y., 2017. A novel pan-genome reverse vaccinology approach employing a negative-selection strategy for screening surface-exposed antigens against leptospirosis. *Frontiers in microbiology* 8 396.