

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE VETERINARIA



TESIS DOCTORAL

**Utilización de herramientas genómicas para impulsar los
programas de selección en acuicultura**

**Applying genomic tools to boost aquaculture selective
breeding programs**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Silvia García-Ballesteros García

Directores

**Beatriz Villanueva Gaviña
Jesús Fernández Martín**

Madrid

© Silvia García-Ballesteros García, 2023

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE VETERINARIA



TESIS DOCTORAL

**Utilización de herramientas genómicas para impulsar
los programas de selección en acuicultura**

**Applying genomic tools to boost aquaculture selective
breeding programs**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Silvia García-Ballesteros García

Directores

Beatriz Villanueva Gaviña

Jesús Fernández Martín

Madrid, 2023

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE VETERINARIA



TESIS DOCTORAL

**Utilización de herramientas genómicas para impulsar
los programas de selección en acuicultura**

**Applying genomic tools to boost aquaculture selective
breeding programs**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Silvia García-Ballesteros García

Directores

Beatriz Villanueva Gaviña

Jesús Fernández Martín

Madrid, 2023

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE VETERINARIA



TESIS DOCTORAL

**Utilización de herramientas genómicas para impulsar
los programas de selección en acuicultura**

**Applying genomic tools to boost aquaculture selective
breeding programs**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Silvia García-Ballesteros García

Directores

Beatriz Villanueva Gaviña

Jesús Fernández Martín

Madrid, 2023

La doctora Beatriz Villanueva Gaviña, Profesora de Investigación y el doctor Jesús Fernández Martín, Investigador Científico, ambos del Departamento de Mejora Genética Animal del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, y la doctora Isabel Cervantes Navarro, Profesora del Departamento de Producción Animal de la Facultad de Veterinaria de la Universidad Complutense de Madrid,

CERTIFICAN:

Que la Tesis Doctoral titulada “**Utilización de herramientas genómicas para impulsar los programas de selección en acuicultura**” presentada por Silvia García-Ballesteros García para optar al grado de Doctor, ha sido realizada bajo su dirección, cumple las condiciones exigidas para obtener dicho título y autorizan su presentación para que sea juzgada por la comisión correspondiente.

Y para que así conste, firman en Madrid, a 17 de noviembre de 2022

Beatriz
Villanueva Gaviña

Jesús
Fernández Martín

Isabel
Cervantes Navarro

Esta Tesis Doctoral se realizó en el Departamento de Mejora Genética Animal del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC). El trabajo expuesto en el capítulo 3 se realizó parcialmente en el Instituto Roslin, de la Universidad de Edimburgo durante una estancia predoctoral.

La autora ha recibido financiación del INIA en el periodo 2017-2021 a través de un contrato predoctoral de investigador FPI-INIA (FPI-SGIT2016-02).

Los trabajos que contiene esta tesis han sido también financiados a través del: ‘European Union’s Seventh Framework Program (KBBE.2013.1.2-10) under Grant Agreement No. 613611 (FISHBOOST project)’; ‘European Commission Horizon 2020 (H2020) Framework Programme through Grant Agreement No. 727315 (MedAID project, Mediterranean Aquaculture Integrated Development)’; Ministerio de Ciencia e Innovación (Proyectos CGL2016-75904-C2-2-P, PID2020-114426GB-C22).

Los capítulos de esta Tesis se corresponden con las siguientes publicaciones científicas:

Capítulo 1. García-Ballesteros, S., Fernández, J., Toro, M.Á., and Villanueva, B. (2021). *Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families*. *Aquaculture*, 543, 1-8.

Capítulo 2. García-Ballesteros, S., Fernández, J., Kauser, A., and Villanueva, B. (2022). *Predicted genetic gain for carcass yield in rainbow trout from indirect and genomic selection*. *Aquaculture*, 554, 1-8.

Capítulo 3. *Impact of selection for disease resistance on the spread of infection in a simulated aquaculture population*. Resultados aún no publicados.

Capítulo 4. García-Ballesteros, S., Villanueva, B., Fernández, J., Gutiérrez, J. P., and Cervantes, I. (2021). *Genetic parameters for uniformity of harvest weight in Pacific white shrimp (*Litopenaeus vannamei*)*. *Genetics Selection Evolution*, 53, 1-9.

A Maya,

*“The answer to the ultimate question of life, the
universe, and everything is 42”*

Douglas Adams

Agradecimientos

En primer lugar, quisiera agradecer a mis directores de tesis Beatriz Villanueva Gaviña y Jesús Fernández Martín por el apoyo durante todos estos años de doctorado. No solo me han guiado a través de esta etapa, sino que han escuchado todas mis ideas (por absurdas que fueran) y he podido decidir e intervenir en todo momento. Además, los dos se complementan y trabajan muy bien juntos lo cual crea un entorno de trabajo ideal. No han escatimado tiempo, esfuerzo, ni dinero en cuanto a mi formación se refiere. Me han hecho sentir valorada y aunque me queda mucho por aprender, termino la tesis con la sensación de que estoy preparada para lo que me depara el futuro.

A Beatriz, qué decir, necesitaría muchísimas páginas para poder agradecerte por todos estos años, pero quedaría feo que los agradecimientos ocupasen más que la tesis. Me has aportado infinidad de cosas, pero quizás lo que más valoro es ese punto de positividad, de resolución, de buscar siempre la perfección y el saber que lo que estamos haciendo cuenta y sirve para algo. Todos podemos aprender genética cuantitativa, programas de mejora en acuicultura... leyendo libros. Tú has sabido poner orden en mi cabeza para que diera lo mejor de mí en cada momento. Siempre has estado ahí en cuanto caía en el agobio o el pesimismo para sacarme del pozo poniendo orden, guiándome para que no me agobiara y siguiera adelante.

A Jesús, quién nos lo iba a decir, ocho años desde nuestro primer programa Fortran y siete años desde nuestro primer artículo juntos. La gente que programa sabe que la mayoría de las veces es más fácil crear tu propio programa que entender y usar el de otro. Sin embargo, en nuestro caso no es así, la mayoría de mis líneas de código serían iguales a las tuyas. Gracias a ti no solo he aprendido muchísimo sobre los programas de mejora en acuicultura, si no que he mejorado muchísimo mi nivel de programación. Tampoco hay que olvidar que fuiste la primera persona que dedicó su tiempo para enseñarme a escribir mi primer artículo, si hoy en día dicen que mis borradores son buenos es gracias a ti. Además, no todo ha sido trabajo, qué sería de la vida sin esos viajes a congresos y reuniones recitando a “Les Luthiers”.

A mi tutora Isabel Cervantes y a mi tutor en la sombra Juan Pablo Gutiérrez, muchas gracias por estar siempre ahí, estos últimos años habéis sido mis consejeros. Todo empezó con vosotros hace 11 años, ni siquiera hubiera hecho el doctorado si no hubiera sido gracias a la formación previa que me disteis en programación y en programas de mejora genética.

Muchas gracias a Andrea Doeschl-Wilson por acogerme y hacerme sentir que Edimburgo también era mi hogar. No solo me enseñaste muchísimas cosas, si no que hiciste que la experiencia fuese increíble y no me quisiese volver ni aunque hubiera una pandemia de por medio. Aunque sé que vamos a seguir trabajando a distancia, espero poder volver algún día y trabajar juntas de nuevo, siento que aún me queda muchísimo por aprender de ti.

A Antti Kause por su conocimiento, ayuda, apoyo y paciencia para que entendiera bien el programa de mejora de la trucha. Gracias por tus rápidas revisiones, comentarios y respuestas incluso en mitad de una pandemia o de vacaciones.

Qué sería de mí sin el diverdespacho de enfrente, sin mis compañeros Alejandro, Marta, Bea, Mónica, Adrián, Edgar y Eli. A mis compañeros de despacho Adrián, Consuelo, Anita, Raúl y Gabriel. También a Diego, María Muñoz, María Saura y Óscar. Gracias a vosotros he tenido vida más allá de la tesis y ni os hacéis una idea de la de veces que habéis evitado que lanzase el ordenador por la ventana. Habéis hecho que los congresos, reuniones y cursos fuesen mucho más divertidos. No solo somos compañeros de trabajo, me llevo muchos amigos de aquí.

A nadie le extrañará que le dedique unas palabras en especial a Eli, “mi compañera de trabajo”, ya saben que somos Pili y Mili. Congeniamos desde el primer día y gracias a eso todo el doctorado ha sido superdivertido, para mí, mis mejores años. Hemos compartido directores de tesis, despacho, habitaciones en congresos y cursos, hemos coincidido en el mismo sitio de estancias, hemos viajado en vacaciones (porque no pasábamos suficiente rato juntas), hemos hecho el camino de Santiago solas. Ninguna tenemos hermanas, pero me imagino que sería una relación parecida a la que tenemos y que no cambiaría por nada. Gracias por estar ahí y por muchos años más.

Gracias al resto de compañeros del INIA (Almudena, Cristina, María Jesús, Rita, Yolanda, Fabián, Clara, Teresa, Malena, Ramón, Luis Alberto...) algunos por sus consejos y participación en diversos puntos de mi tesis, a otros simplemente por estar ahí y dar su apoyo y cariño en las comidas de navidad, cafés, seminarios, reuniones de departamento, etc. Todos ellos han participado en crear un ambiente muy bueno en el INIA.

Por último, pero no menos importante, a mi familia y a Cookie por aguantarme durante todo este proceso, que no siempre he sido fácil de llevar.

Acknowledgments

I would like to thank my supervisors Beatriz Villanueva Gaviña and Jesús Fernández Martín for their support during all these years. They have guided me in this stage and have listened to all my ideas (no matter how absurd they were) and I have been able to decide and intervene at all times. The two complement each other and work very well together, creating an ideal work environment. They have spared no time, effort, or money when it comes to my training. They have made me feel valued and although I still have a lot to learn, I finish the thesis with the feeling that I am prepared for what the future holds.

To Beatriz, what can I say, I would need many pages to be able to thank you for all these years, but then the gratitude would take up more than the thesis. You have given me countless things, but perhaps what I value most is that point of positivity, of resolution, of always seeking perfection and knowing that what we are doing is useful for something. We can all learn about quantitative genetics, breeding programs in aquaculture... by reading books. You have known how to put order in my head so that I always give my best. You have always been there when I fell into overwhelm or pessimism, guiding me to keep going and not give up.

To Jesus, who was going to tell us, eight years since our first Fortran program and seven years since our first article together. People who program know that most of the time it is easier to create your own program than to understand and use someone else's. However, this is not our case, most of my lines of code would be the same as yours. Thanks to you, I have not only learned a lot about breeding programs in aquaculture, but I have also greatly improved my level of programming. We must also not forget that you were the first person who dedicated his time to teach me how to write my first article, if today they say that my drafts are good it is thanks to you. But not everything has been work, what would life be without those trips to congresses and meetings reciting "Les Luthiers".

To my tutor, Isabel Cervantes and my shadow tutor Juan Pablo Gutiérrez, thank you very much for always being there, these last years you have been my advisers. It all started with you 11 years ago, I wouldn't even have done my doctorate if it hadn't been thanks to the previous training you gave me in programming and in selective breeding programs.

Many thanks to Andrea Doeschl-Wilson for taking me in and making Edinburgh feel like home too. Not only did you teach me many things, but you made the experience incredible and that I didn't want to go back even when there was a pandemic involved. Although I know that we will continue to work remotely, I hope to be able to return one day and work together again, I feel that I still have a lot to learn from you.

To Antti Kause for his knowledge, help, support and patience in understanding the trout breeding program. Thank you for your quick reviews, comments, and responses even in the middle of the pandemic or on vacation.

What would become of me without the fun-office, without my colleagues Alejandro, Marta, Bea, Mónica, Adrián, Edgar, and Eli. Also, to my colleagues Adrián, Consuelo, Anita, Raúl, Gabriel, Diego, María Muñoz, María Saura, and Óscar. Thanks to you I had a life beyond my thesis, and you have no idea how many times you have prevented me from throwing the computer out the window. You have made the congresses, meetings, and courses much more fun. We are not only co-workers, but I take many friends from here.

No one will be surprised that I dedicate a few words especially to Eli, "my co-worker", they already know that we are 'Pili and Mili'. We hit it off from the first day and thanks to that the whole doctorate has been super fun, for me, my best years. We have shared thesis supervisors, offices, rooms in congresses and courses, we have coincided in the same place of stays, we have traveled on vacations (because we did not spend enough time together), and we have walked the Camino de Santiago alone. None of us have sisters, but I imagine it would be a relationship similar to the one we have, and I wouldn't change it for anything. Thank you for being there and for many more years.

Thanks to the rest of my colleagues at INIA (Almudena, Cristina, María Jesús, Rita, Yolanda, Fabián, Clara, Teresa, Malena, Luis Alberto...) some for their advice and participation in various points of my thesis, others simply for being there and giving your support and affection at Christmas meals, coffees, seminars, department meetings, etc. All of them have participated in creating a very good atmosphere at INIA.

Finally, to my family and Cookie for putting up with me throughout this process, which hasn't always been easy.

INDEX

INDEX OF FIGURES.....	23
INDEX OF TABLES.....	25
RESUMEN.....	27
SUMMARY.....	31
GENERAL INTRODUCTION.....	35
Aquaculture: definition and importance.....	37
Current status of the aquaculture sector.....	38
Aquaculture selective breeding programs.....	40
Traits to be included in complex breeding goals.....	47
Genomic selection.....	51
OBJECTIVES.....	57
Specific objectives.....	60
SOFTWARE DEVELOPED.....	61
CHAPTER 1.....	65
Introduction.....	67
Material and methods.....	68
Results.....	69
Discussion.....	70
Conclusions.....	73
CHAPTER 2.....	75
Introduction.....	77
Material and methods.....	78
Results.....	79
Discussion.....	81
Conclusions.....	82

CHAPTER 3.....	85
Introduction.....	88
Material and methods.....	89
Results.....	91
Discussion.....	92
Conclusions.....	93
CHAPTER 4.....	95
Introduction.....	97
Material and methods.....	98
Results.....	101
Discussion.....	102
Conclusions.....	103
GENERAL DISCUSSION.....	107
Recapitulation.....	109
Real data analyses <i>versus</i> computer simulations.....	109
Factors that affect the benefit obtained from genomic selection.....	111
Genomic selection for candidate traits.....	114
Genomic selection for sib traits.....	116
Inbreeding.....	121
Boosting aquaculture selective breeding programs.....	122
GENERAL CONCLUSIONS.....	125
REFERENCES.....	129
ANNEX I: Manual GAINFISH.....	153
ANNEX II: Supplementary material of Chapter 2.....	171
ANNEX III: Communications at congresses related to this thesis.....	175

INDEX OF FIGURES

GENERAL INTRODUCTION

Figure 1. Global aquaculture and capture fisheries production per year.....	38
Figure 2. European aquaculture production by country in 2018 (in tonnes of live weight).....	39

INDEX OF TABLES

GENERAL INTRODUCTION

Table 1. Genetic gain per generation (in %) for body weight in different aquaculture species.....	44
Table 2. Traits included in the breeding goal of selective breeding programs in different aquaculture species.....	45
Table 3. Commercial SNP arrays developed for different aquaculture species.....	54

GENERAL DISCUSSION

Table 4. Estimates of linkage disequilibrium (measured as r^2 , the squared correlation between pairs of loci) for different aquaculture species.....	113
Table 5. Benefit of genomic evaluations (in terms of accuracy of selection) when compared to traditional BLUP evaluations for candidate traits found in different aquaculture studies.....	115
Table 6. Estimates of residual heritability (h_v^2), genetic coefficient of variation of residual variance (GCV_E) and genetic correlation between the mean and the residual variance (ρ) for different aquaculture species.....	116
Table 7. Benefit of genomic evaluations (in terms of accuracy of selection) when compared to traditional BLUP evaluations for sib traits found in different aquaculture studies.....	119

RESUMEN

La acuicultura es el cultivo de organismos acuáticos que implica algún tipo de intervención en el proceso de crianza. A nivel mundial, representa alrededor del 50% de la producción acuática. El aumento de la población mundial y la creciente demanda de productos acuáticos han llevado a un continuo aumento de la producción acuícola durante las últimas tres décadas. Sin embargo, el desarrollo sostenible de la acuicultura se enfrenta a numerosos desafíos. Por ello son de vital importancia iniciativas dirigidas a mejorar la eficiencia y la competitividad de la industria. Un factor clave para conseguir dicho desarrollo es la aplicación de programas de mejora genética que incluyan la selección de reproductores con un mérito genético alto. Sin embargo, se ha estimado que actualmente solo alrededor del 10% de la producción acuícola mundial se basa en poblaciones mejoradas genéticamente. De hecho, la complejidad de los programas de mejora varía mucho entre diferentes especies y entre diferentes poblaciones de una misma especie. Mientras que algunos programas de mejora solo incluyen componentes básicos (p. ej. selección exclusivamente para crecimiento sin un control sobre el aumento en consanguinidad), otros están en un nivel mucho más avanzado (p. ej. incluyen varios caracteres en el objetivo de mejora y utilizan información de marcadores moleculares).

La investigación desarrollada en esta tesis tiene como objetivo final impulsar los programas de selección existentes para hacerlos más competitivos. En particular, en esta tesis se han investigado las ganancias genéticas que se pueden obtener al i) incluir nuevos caracteres de gran relevancia económica (como son rendimiento a la canal, resistencia a enfermedades y uniformidad de peso) en los objetivos de mejora, y ii) aplicar la selección genómica. El trabajo presentado en esta tesis se ha llevado a cabo mediante análisis de datos reales o mediante simulaciones por ordenador. En la fase inicial de la tesis, se desarrolló un programa de simulación en Fortran 90 (el software **GAINFISH**) para obtener predicciones de la ganancia genética y la consanguinidad en una amplia gama de escenarios. Este software se ha utilizado en los **Capítulos 1 - 3**.

En el **Capítulo 1** se ha investigado, mediante simulaciones, el beneficio de incluir la evaluación genómica (en términos de ganancia genética y consanguinidad) en escenarios donde existe un efecto ambiental común a los miembros de una misma familia. Esta investigación aborda un tipo de problema específico en los programas de selección en acuicultura ya que, para obtener el pedigrí a lo largo de las generaciones, un tipo

habitual de manejo es mantener a las familias separadas en diferentes tanques hasta que los individuos sean lo suficientemente grandes como para poder ser identificados individualmente. El problema con este tipo de manejo es que introduce un efecto ambiental común a los miembros de la misma familia, que puede ser difícil de separar del efecto genético cuando se utiliza la evaluación tradicional mediante BLUP. Este problema podría reducirse utilizando la evaluación genómica, dado que los marcadores pueden capturar mejor las relaciones genéticas entre individuos criados en diferentes tanques y entre hermanos criados en el mismo tanque. Por lo tanto, se han simulado y comparado escenarios que utilizan evaluación genómica o BLUP que, a su vez, incluyen o excluyen los efectos ambientales comunes en sus modelos. Se asume un carácter seleccionado con una heredabilidad de 0,4 y tres proporciones diferentes de la varianza fenotípica debida al efecto ambiental común (0,0, 0,1 y 0,3). Los resultados revelan que el beneficio de la evaluación genómica frente a la evaluación BLUP en términos de ganancia genética es mayor en presencia de efectos ambientales comunes que en su ausencia. Esto se debe a que la presencia de estos efectos no afecta la precisión de la evaluación genómica pero sí disminuye la precisión de la evaluación BLUP. Aunque la evaluación genómica siempre conlleva una tasa de consanguinidad por generación más baja que la evaluación BLUP, no hay una reducción adicional en presencia de efectos ambientales comunes. Además, en este capítulo se muestra que ignorar los efectos ambientales comunes en el modelo de evaluación cuando estos efectos existen, no solo conduce a una reducción en la ganancia genética (como se esperaba) sino a un gran aumento en la consanguinidad, incluso cuando estos efectos ambientales son pequeños.

El rendimiento a la canal y el rendimiento del filete son dos caracteres de gran importancia económica en acuicultura. En particular, son muy importantes en las especies que se venden procesadas como es el caso de la trucha arcoíris. La problemática con estos caracteres es que su medición es laboriosa y está asociada a grandes errores, además implica sacrificar a los peces lo que impide obtener fenotipos en los candidatos a la selección. Investigaciones recientes han descrito un carácter indicador no letal (E8/E23, que es la relación entre el grosor de la pared abdominal y la profundidad de la cavidad peritoneal) que tiene una correlación genética alta tanto con el rendimiento a la canal como con el rendimiento del filete. Sin embargo, se desconocen los beneficios de seleccionar para el carácter indicador cuando se usan fenotipos de los candidatos, en lugar de seleccionar directamente para el carácter canal y sin cabeza cuando se usan fenotipos

de los hermanos. Ambas alternativas podrían ser eficientes para mejorar los rendimientos a la canal y del filete y, por ello, se han comparado para una población de trucha arcoíris en el **Capítulo 2**. Se han simulado cuatro estrategias de selección: i) selección mediante BLUP para el carácter canal sin cabeza, ii) selección indirecta mediante BLUP para el carácter indicador, iii) selección genómica para el carácter canal sin cabeza, y iv) selección genómica para el carácter indicador. Los diferentes escenarios se han comparado para la misma intensidad de selección y número de fenotipos, tanto para el carácter canal sin cabeza (medido en los hermanos) como para el carácter indicador (medido en los candidatos). Como era de esperar, se obtienen mayores ganancias genéticas en los escenarios con selección genómica que en los escenarios con selección BLUP. Las diferencias encontradas en términos de ganancia genética entre la selección para el carácter canal sin cabeza y la selección para el carácter indicador son pequeñas. Por lo tanto, la selección genómica para el indicador parece ser la opción más rentable dado que conlleva a menores gastos asociados con el mantenimiento y el genotipado (al no requerir a los hermanos). Esta misma conclusión se alcanza cuando el objetivo de selección incluye el peso corporal además del indicador. Por lo tanto, se concluye que el enfoque óptimo para mejorar el rendimiento a la canal es una combinación de selección genómica e indirecta en escenarios de selección uni- y multi-carácter.

Otro carácter de gran importancia en acuicultura es la resistencia a enfermedades, dadas las enormes pérdidas económicas que éstas conllevan. De hecho, los programas de selección avanzados como aquellos implementados en el salmón del Atlántico, ya incluyen la resistencia a diversas enfermedades en los objetivos de mejora. La selección para el carácter resistencia a la enfermedad (generalmente medido en acuicultura como fecha de la muerte) tiene como objetivo fortalecer la capacidad del individuo para hacer frente a la infección (es decir, reducir la probabilidad de que los peces mueran cuando se exponen al patógeno). Sin embargo, se desconoce si la selección para el carácter resistencia a la enfermedad reduce no solo la mortalidad sino también la transmisión del patógeno y, por lo tanto, la incidencia y gravedad de los brotes epidémicos. Una complicación adicional con la resistencia a enfermedades es que el fenotipo principal (fecha de la muerte) y la dinámica de propagación de la enfermedad se ven afectados por tres caracteres epidemiológicos subyacentes (no observados): susceptibilidad, infectividad y mortalidad inducida por infección. En el **Capítulo 3**, mediante simulación, se ha investigado el impacto de la selección genómica para la resistencia a enfermedades

sobre los tres caracteres subyacentes y, por lo tanto, sobre la transmisión de la enfermedad (medida a través del número reproductivo básico) en un programa de selección del salmón del Atlántico, mediante la combinación de modelos genéticos y epidemiológicos. Los fenotipos de la fecha de la muerte se obtienen simulando una transmisión rápida del virus de la anemia infecciosa del salmón, en un experimento de desafío realizado para diferentes lapsos de tiempo. Se asume que los tres caracteres subyacentes no están correlacionados. Los resultados muestran que la selección para aumentar la resistencia a la enfermedad no solo reduce la mortalidad sino también la media genética de los tres caracteres subyacentes y, por lo tanto, el número reproductivo básico, lo que contribuye a disminuir la transmisión de la enfermedad en un brote severo.

En el **Capítulo 4**, utilizando datos reales de una población comercial de langostino blanco del Pacífico, se han estimado los parámetros genéticos de uniformidad del peso (otro carácter con gran importancia económica), con el objetivo final de determinar el potencial de incluir este carácter en los programas de selección existentes. Para mejorar genéticamente la uniformidad del peso a través de la selección, es necesaria la existencia de variación genética para la varianza ambiental del peso. Aunque varios estudios han detectado dicha variación en algunas especies acuícolas, no hay estimaciones disponibles para los langostinos. Los resultados de este capítulo muestran la existencia de variabilidad genética de la uniformidad del peso en la población analizada. Las estimaciones de heredabilidad y del coeficiente de variación residual genética son lo suficientemente grandes como para que se pueda mejorar la uniformidad del peso a través de la selección. Además, la correlación genética estimada entre la uniformidad del peso y el peso corporal no es estadísticamente diferente de cero, por lo que la selección para aumentar la uniformidad del peso no conduciría a cambios en el peso.

Los resultados presentados en esta tesis podrían ser muy útiles para ayudar en la toma de decisiones sobre caracteres adicionales que se quieran incluir en el objetivo de mejora y en la aplicación de la selección genómica. Estos resultados indican las posibles modificaciones en los programas de selección existentes en diferentes especies acuícolas con el objetivo de mejorar su competitividad. Además, vale la pena destacar que el software **GAINFISH** desarrollado en esta tesis será de gran utilidad para la industria, ya que permite predecir y comparar las ganancias genéticas y la consanguinidad esperadas al aplicar diferentes estrategias y poder decidir así siguiente paso a dar en términos de complejidad del programa de mejora.

SUMMARY

Aquaculture is the farming of aquatic organisms that implies some form of intervention in the rearing process. Worldwide, it represents about 50% of the aquatic production. The increase of the world population and the growing demand for aquatic products have led to a continuous increase in aquaculture production during the last three decades. However, the sustainable development of aquaculture faces many challenges. Therefore, initiatives aimed at improving the efficiency and competitiveness of the industry are of vital importance. A key factor to obtain such profitable development is the application of genetic improvement programs that include selective breeding. However, it has been estimated that currently only about 10% of the world's aquaculture production is based on genetically improved stocks. In fact, the complexity of breeding programs in aquaculture varies greatly between and within species. While some of them only include basic components of a selection program (e.g., selection exclusively for growth without inbreeding control), others are at a much more advanced level (e.g., they include several traits in the breeding objective and use molecular marker information).

The research included in this thesis has been performed with the final aim of boosting existing selective breeding programs to make them more competitive. In particular, the thesis has investigated the potential genetic gains that can be expected by i) including new economically important traits (such as carcass yield, disease resistance, and weight uniformity) as breeding goals; and ii) applying genomic selection. The work presented in this thesis has been carried out through analyses of real data or through computer simulations. In the initial phase of the thesis, a computer simulation program was developed in Fortran 90 (the software **GAINFISH**) to predict the expected genetic gain and inbreeding for a wide range of aquaculture breeding scenarios. This software has been used for the simulations performed in **Chapters 1 - 3**.

In **Chapter 1**, the benefit of including genomic evaluation (in terms of genetic gain and inbreeding) when common environmental (tank) effects are present has been investigated through computer simulations. This research deals with the fact that in aquaculture selective breeding programs, a common type of rearing that allows recording pedigrees across generations is to keep families separately in different tanks until fish are large enough to be individually tagged. A particular problem with this type of rearing is that it introduces an environmental effect common to the members of the same family.

This effect can be difficult to be disentangled from genetic effects when traditional BLUP evaluation is used. The problem could at least be reduced when using genomic evaluation, given that markers can capture better the genetic relationships between fish raised in different tanks, and between full-sibs raised in the same tank. Thus, scenarios using genomic and BLUP evaluations fitting models that included or excluded common environmental effects are simulated and compared. It is assumed a selected trait with a heritability of 0.4 and three different proportions of phenotypic variance due to common environmental effect (0.0, 0.1 and 0.3). The results reveal that the benefit of genomic selection over BLUP in terms of genetic gain is greater in the presence of common environmental effects than in their absence. This is because the presence of these effects does not affect the accuracy of genomic selection but it does decrease the accuracy of BLUP selection. Although genomic selection always leads to lower inbreeding rate per generation than BLUP, there are no extra reductions in the presence of common environmental effects. In addition, in this chapter, it is shown that excluding the common environmental effects in the evaluation model leads not only to some reduction in genetic gain (as expected) but to high increases in inbreeding, even when these effects are small.

Carcass and fillet yields are traits of great economic importance in aquaculture species. In particular, carcass and fillet yield are very important for species sold processed such as rainbow trout. The problem with these traits is that recording is laborious and associated to large measurement errors, and also implies slaughtering the fish, which prevents obtaining phenotypes for selection candidates. Recent research has described a non-lethal indicator trait (the ratio of abdominal wall thickness to depth of the peritoneal cavity, E8/E23) highly genetically correlated with carcass yield and with fillet yield. However, the potential benefits of selecting on the indicator trait using candidates' phenotypes rather than selecting directly on headless gutted carcass using sibs' phenotypes are unknown. Both strategies could be efficient to improve carcass and fillet yields and they have been compared by simulating a rainbow trout population in **Chapter 2**. Four alternative selection strategies are simulated: i) sib BLUP selection for headless carcass, ii) indirect BLUP selection for the indicator, iii) genomic selection for headless carcass, and iv) genomic selection for the indicator. The different scenarios are compared for the same selection intensity and number of records for headless carcass (recorded on sibs) and for the indicator (recorded on selection candidates). As expected, genomic selection scenarios lead to higher genetic gains than BLUP scenarios. The differences

found in terms of genetic gain for headless carcass between selection for the indicator and selection for headless carcass are small. Thus, given that genomic selection for headless carcass implies extra costs associated with maintaining and genotyping sibs, selection for the indicator appears to be a more cost-effective option. The same conclusions are drawn when the yield trait is selected simultaneously with body weight. It is therefore concluded that the optimum approach to improve carcass yield is a combination of genomic and indirect selection in both single and multi-trait selection scenarios.

Given the enormous economic losses caused by disease outbreaks, another trait of great economic importance for the aquaculture industries is resistance to specific diseases. In fact, advanced selective breeding programs such as those implemented in Atlantic salmon already include resistance to various diseases in the breeding goal. Selection on disease resistance (usually measured as time of death) is aimed at strengthening an animal's ability to cope with infection (i.e., fish are less likely to die when exposed to the pathogen). However, it is unknown whether selecting directly on the time of death reduces not only mortality but also pathogen transmission and, thus, the incidence and severity of outbreaks. An extra complication with disease resistance is that the main recorded phenotype (i.e., time of death) and the dynamics of the spread of the disease are affected by three underlying (unobserved) epidemiological host traits (susceptibility, infectivity, and infection-induced mortality). In **Chapter 3**, the impact of genomic selection for disease resistance on the three underlying traits and, therefore, on the spread of the disease (measured as the basic reproductive ratio) is investigated in a simulated Atlantic salmon selective breeding program, by combining genetic and epidemiological models. Phenotypes for time of death are obtained by simulating a fast transmission of the infectious salmon anemia virus in disease challenge tests varying in the point where the test is stopped. The three underlying traits are assumed to be uncorrelated. The results show that selection for disease resistance not only reduces mortality but also the genetic mean of the three underlying traits and, thus, the basic reproductive ratio. This contributes to reducing disease transmission in a severe outbreak.

Finally, in **Chapter 4**, using real data of a farmed population of Pacific white shrimp, genetic parameters for weight uniformity (another trait with high economic relevance) are estimated, with the final aim of determining the potential of including this trait in the existing selection programs. In order to improve weight uniformity through selective breeding, it is necessary to demonstrate the existence of genetic variation for the

environmental variance of body weight. Although several studies have reported such variance in other aquaculture species, to our knowledge, there are no estimates available for shrimp. The results of this chapter show that genetic variability for weight uniformity exists in the population analyzed, and that the heritability and the genetic residual variation coefficient estimates are large enough to improve weight uniformity through selection. In addition, the estimated genetic correlation between weight uniformity and body weight is not statistically different from zero, and thus selection for weight uniformity would not lead to changes in body weight.

The results presented in this thesis could be very useful in helping to take decisions on extra traits to be included in the breeding goal and on applying genomic selection. This could lead to modifications in the selective breeding programs of different aquaculture species with the aim of improving their competitiveness. In addition, it is worth highlighting the fact that the software **GAINFISH** developed in this thesis will be very useful for the breeding industry, since it allows predicting and comparing the expected genetic gains and inbreeding when applying different strategies, and thus being able to decide the next step to take in terms of complexity of the breeding program.

GENERAL INTRODUCTION





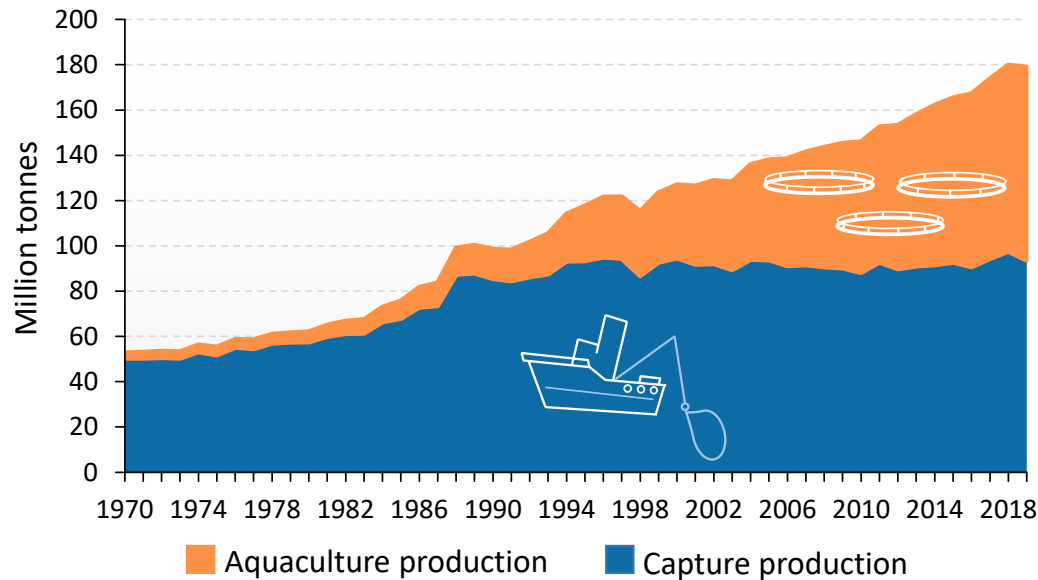
Aquaculture: definition and importance

Aquaculture is the farming of aquatic organisms that implies some form of intervention in the rearing process (e.g., regular stocking, feeding, and protection from predators) to enhance production (FAO, 2020). Fish production under controlled conditions began 3000 - 4000 years ago in China with the farming of common carp (*Cyprinus carpio*) in earthen ponds (Gjedrem and Baranski, 2009; Nakajima et al., 2019) and independently, in medieval Europe in the 12th to 14th centuries (Harland, 2019). The oldest book describing fish farming was written by Fan Lee in 475 BC. Despite this, the domestication of most aquatic species has been complex, and the aquaculture industry has not developed in most species until the last quarter of the 20th century (Teletchea and Fontaine, 2014). For example, the domestication of Atlantic salmon (*Salmo salar*), a species with one of the highest production levels today, began in Scotland in the early nineteenth century (Teletchea, 2019) but was only at the end of the 1960s that the farming in sea cages was used for the first time in Norway (Gjedrem, 2010).

One of the greatest challenges facing humanity is to increase food production to feed a continuously growing world population with limited natural resources while respecting ecosystems. Aquatic foods contribute to healthy diets for billions of people and are one of the most important sources of animal protein (Golden et al., 2021). Aquatic products contain not only high-quality, easily digestible protein but are also rich in vitamins (A, B, and D), amino acids and essential fatty acids (omega 3), offering multiple beneficial effects for health, especially in child development. In addition, when compared with terrestrial farm animals, fish are more efficient feed-to-meat converters (i.e., they require fewer resources to produce a kilogram of food). Thus, it is not surprising that the demand for aquatic products is increasing year after year (FAO, 2020). The world per capita consumption of aquatic products (excluding aquatic mammals, crocodiles, alligators and caimans, seaweeds, and other aquatic plants) increased from 13.5 kg in 1990 to 20.5 kg in 2018. In fact, the world average annual rate of increase in fish consumption is almost twice the corresponding population growth rate, and higher than that of other animal protein foods (FAO, 2020). This increase in consumption has been possible mainly thanks to the increase in aquaculture production (growing an average of 7.7% per year), while capture fisheries production has remained more or less constant since the 1990s due to the limitations and restrictions imposed to avoid overfishing

(Figure 1). It is worth mentioning that aquaculture is the fastest-growing food production sector (Naylor et al., 2021).

Figure 1. Global aquaculture and capture fisheries production per year.



(From FAO's data available at https://www.fao.org/fishery/statistics-query/en/aquaculture/aquaculture_quantity).

Current status of the aquaculture sector

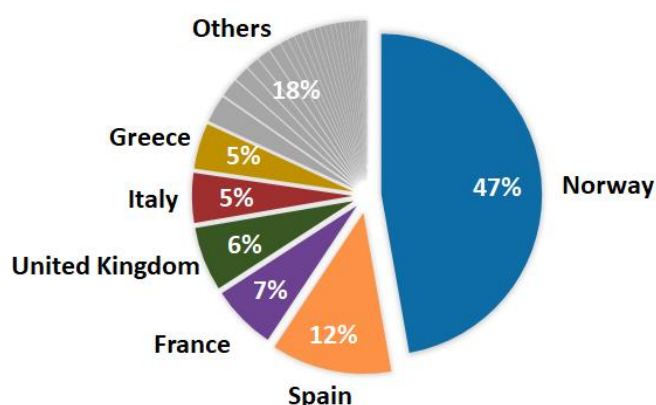
Currently, aquaculture production represents almost 50% of the world's aquatic production (aquaculture and capture fisheries) and is an activity in continuous expansion (FAO, 2020). World aquaculture production is balanced between marine (56.6%) and inland (43.4%) waters. Most of the world's aquaculture production (in tonnes of live weight) comes from fish (47.4%), followed by algae (28.3%), mollusks (15.3%), crustaceans (8.2%) and finally invertebrates (0.4%) and amphibians and reptiles (0.4%). In the last decades, there has been a substantial increase in the number of new species that have been farmed in captivity. There are about 500 aquatic species currently farmed in the world, representing an enormous wealth of genetic diversity (Houston et al., 2020). However, most of the world's aquaculture production comes from a limited number of species. For instance, only 15 species account for 85% of the world's finfish production and most finfish production (84%) comes from freshwater (Teletchea and Fontaine, 2014). The main species produced through aquaculture in the world in 2018, in relation to their production value (in millions of euros), were white shrimp (*Litopenaeus*

vannamei) and Atlantic salmon, followed by marsh crab (*Procambarus clarkii*), Chinese carp (*Ctenopharyngodon idella*) and silver carp (*Hypophthalmichthys molitrix*).

At the European level, aquaculture production has been increasing to currently represent 20% of the total volume of aquatic production, while capture fisheries production has been decreasing in the last decade (FAO, 2020). Most European aquaculture production takes place in marine waters (79.5%) and its main products (in tonnes of live weight) are fish (51%) and mollusks (49%). The main fish species produced by aquaculture in Europe in 2018 in relation to their production value (in millions of euros) were Atlantic salmon, rainbow trout (*Oncorhynchus mykiss*), sea bass (*Dicentrarchus labrax*), sea bream (*Sparus aurata*), common carp, Atlantic bluefin tuna (*Thunnus thynnus*) and turbot (*Scophthalmus maximus*). The main mollusks produced were the Japanese oyster (*Crassostrea gigas*), the Japanese clam (*Ruditapes philippinarum*) and several species of mussels (*Mytilus spp*).

In Spain, the aquaculture production in 2018 accounted for 12% of the European production (Figure 2). The main aquaculture products (in tonnes of live weight) were mollusks (76%) and fish (24%). Within Europe, Spain ranks first in mollusk and third in fish production (FAO, 2020). The main fish species produced through aquaculture about their production value (in millions of euros) in Spain were sea bass, rainbow trout, and sea bream and the main mollusk species were mussels.

Figure 2. European aquaculture production by country in 2018 (in tonnes of live weight).



(From FAO's data available at https://www.fao.org/fishery/statistics-query/en/aquaculture/aquaculture_quantity).

Aquaculture selective breeding programs

In general, the sustainable development of aquaculture faces many challenges given growing concerns about environmental impact, competition for water resources, and increased demand for aquaculture products due to global population growth and declining of wild catches (Gjedrem and Baranski, 2009). Initiatives aimed at improving production efficiency and competitiveness are thus of vital importance to successfully face these challenges. One of the key factors to increase the efficiency and competitiveness of aquaculture industries is the application of genetic improvement programs that include selective breeding. Selective breeding programs exploit the genetic variation present for traits of economic importance in order to improve them. It needs to be highlighted that the improvement for the traits of interest obtained with these programs is permanent (unlike that obtained by improving environmental or management conditions) and cumulative (improvements obtained in one generation are added to those obtained in previous generations).

A basic selective breeding program consists of selecting the elite sires and dams for a trait of interest and mating them to produce the next improved generation. The first step in any selective breeding program is choosing the breeding goal (i.e., which traits we want to improve). The next step is to choose the appropriate selection criteria (i.e., which traits will be measured and selected on). In some cases, the selection criteria may be the same as the breeding goal; in others, some indirect measurements are needed (e.g., when a particular goal trait can only be recorded after slaughter). Once the selection criteria have been decided, the next step is to determine which population and individuals within the population will be recorded. This, in turn, will depend on the trait and the species production system. Two types of traits can be distinguished: i) ‘candidate’ traits, defined as traits that can be recorded on the selection candidates; and ii) ‘sib’ traits, defined as traits that cannot be recorded on the candidates (e.g., traits for which recording implies the sacrifice of individuals) and are recorded on their sibs. The genetic merit of the candidates can be determined by their estimated breeding values (EBVs) which are traditionally obtained using BLUP (Best Linear Unbiased Prediction), a statistical technique that separates, in an optimal way, the genetic and the environmental effects using phenotypic and pedigree records. Recently, more accurate genomic evaluations using thousands of genetic markers (Meuwissen et al., 2001) have started to be performed

in aquaculture selection programs. Once the EBVs for the candidates have been obtained, fish with the highest EBVs are selected and mated to produce the next generation.

When compared with crop and terrestrial livestock industries, the aquaculture industry is young and, thus, domestication and selective breeding are more recent and limited to relatively few species. It was estimated that only about 10% of world aquaculture production was based on stocks coming from selective breeding programs (Gjedrem et al., 2012; Janssen et al., 2017). However, this proportion varies greatly across continents. For instance, in Europe, about 80% of the aquaculture production is based on genetically improved stocks (Janssen et al., 2017). Also, this proportion increases for species with high production volume and economic value. In particular, the Atlantic salmon industry began its first family-based breeding programs in the 1970s, and in 2010 it was estimated that more than 95% of the world production for this species came from genetically improved stocks (Gjedrem et al., 2012). This percentage is reduced to 22% for rainbow trout for which the first breeding programs also began in the early 1970s (Gjedrem and Baranski, 2009). Breeding programs for other species of lower economic value such as tilapia began in the 1990s and only 9% of the production is based on improved stocks (Teletchea and Fontaine, 2014). In the case of European sea bass and gilthead sea bream, breeding programs were initiated in the mid-1990s and in the early 2000s, respectively (Janssen et al., 2017).

Although the development of new aquaculture programs has been slow, its potential is great as very large genetic improvements have been shown (Gjedrem and Baranski, 2009). The great potential of aquaculture selective breeding programs is in part due to the high fecundity of fish species that allows intense selection (large numbers of candidates and low number of selected parents) and therefore, fast genetic gains. However, high intensities of selection can also lead to large reductions of genetic variability and high rates of inbreeding, with the consequent increased risk of inbreeding depression (i.e., reduced fish performance due to inbreeding). Therefore, controlling rates of inbreeding and loss of genetic variation is especially important in aquaculture. Computing inbreeding and imposing restrictions on it can be done using pedigree data (e.g., Gjerde et al., 1996; Villanueva et al., 1996). Also, tracking pedigrees over generations may help to increase the accuracy of selection, and therefore selection responses, by incorporating information not only on the selection candidates but also on their relatives into the genetic evaluation.

In animal breeding programs, ideally, individuals are identified when they are born and then the pedigrees of individual animals can be tracked across generations. However, this represents a problem in aquaculture given that newborn fish are too small to be tagged physically. Thus, selective programs making use of family information need to keep families separately until fish is large enough to be individually tagged or mix families from the start and afterwards assign the offspring to parents using molecular markers (e.g., microsatellites or SNPs). The second type of rearing is the standard in species such as sea bass and sea bream where production usually implies mass spawning due to the particular physiology of the species (Janssen et al., 2017). The problem with mixed rearing is that it can lead to unequal family contributions and hence to high risk of inbreeding (Fernandes et al., 2017).

Separate family rearing is commonly applied in salmonids, shrimp, and Nile tilapia breeding programs (Gjedrem, 2010; Castillo-Juárez et al., 2015). A particular problem with this type of rearing is that it introduces an environmental effect common to the members of the same family (usually known as the ‘tank effect’) that is difficult to disentangle from genetic effects and, thus, reducing the expected genetic gain (Gjerde et al., 1996; Gjedrem and Baranski, 2009). A possible solution would be having a family replicated in different tanks. However, in the aquaculture industry, space is a limiting factor that will determine the number of tanks available for breeding. Therefore, if a family is replicated in different tanks, the number of families will be reduced. For example, the same resources would be needed with 70 tanks and 70 different families and with 70 tanks with 35 different families, each raised in two tanks. Thus, given that replicates would imply a smaller number of families (and therefore a lower genetic response and a higher risk of inbreeding) in general, each family is raised in a single tank so all members of the same family share the same environment. When standard BLUP is used in the genetic evaluation, it is very difficult to disentangle the common environmental effect from the genetic effect (Pante et al., 2002; Montaldo et al., 2012). Consequently, since the estimates of the former effect are usually of low magnitude (Pante et al., 2002; Gjerde et al., 2004; Kjøglum et al., 2005; Kause et al., 2007), common environmental effect is usually ignored in the models assumed in genetic evaluations (Ødegård et al., 2014; Khaw et al., 2016).

Growth rate is the main breeding goal trait in aquaculture selective breeding programs (Gjedrem et al., 2012; Janssen et al., 2017; Gjedrem and Rye, 2018; Houston et al., 2020). It is an economically important trait that is generally measured as overall body weight at harvest and can be easily recorded on selection candidates. As growth rate increases, production time will automatically be reduced along with the maintenance costs. Also, growth has a favorable (negative) genetic correlation with feed conversion ratio (e.g., Besson et al., 2020) and thus, selecting on growth will increase farm profit and decrease environmental impacts. The genetic gains for body weight obtained per generation for different aquaculture species are given in Table 1. These gains in aquaculture species are 4–5 times greater than those in livestock terrestrial species (Gjedrem and Rye, 2018; Zenger et al., 2019) mainly due to two reasons. Firstly, given that selective breeding programs for aquaculture species are in the early stages of development, the genetic diversity in these species has scarcely been exploited. Secondly, as already stated, aquaculture species have a high fecundity which makes possible to apply a high intensity of selection.

However, as growth rate increases, expanding the breeding goals of aquaculture breeding programs is essential to maximize productivity and profit for aquaculture companies. These expanded breeding goals should include economically important traits such as disease resistance, feed efficiency, fillet yield, flesh quality, and adaptability to plant-based feeds (Gjedrem and Rye, 2018; Houston et al., 2020). In fact, the complexity of breeding programs varies greatly between aquaculture species and companies according to the traits included into their breeding goals (Table 2). Note that species with advanced breeding programs (e.g., Atlantic salmon programs) already include a large number of traits.

Table 1. Genetic gain per generation (in %) for body weight in different aquaculture species.

Species	Genetic gain	Reference
Atlantic salmon (<i>Salmo salar</i>)	7 – 15 12 11 – 14	Lhorente et al., 2019 Janssen et al., 2017 Gjedrem, 2000
Coho salmon (<i>Oncorhynchus kisutch</i>)	6 – 13 6 – 13 9 – 10	Lhorente et al., 2019 Dufflocq et al., 2017 Neira et al., 2004
Gilthead sea bream (<i>Sparus aurata</i>)	13 29 5 – 10	Thorland et al., 2015 Brown, 2003 Knibb, 2000
European sea bass (<i>Dicentrarchus labrax</i>)	26 25 – 30 23 16 – 25	Le Boucher et al., 2013 Chatain and Chavanne, 2009 Vandeputte et al., 2009 Duppont-Nivet et al., 2008
Nile tilapia (<i>Oreochromis niloticus</i>)	3 13 – 18 7 12 – 17	Yoshida et al., 2022 Bentsen et al., 2017 Khaw et al., 2008 Eknath et al., 1998
Rainbow trout (<i>Oncorhynchus mykiss</i>)	7 – 13 7 – 13 6 5 – 8 13	Lhorente et al., 2019 Janssen et al., 2017 Janhunen et al., 2012 Kause et al., 2005 Gjedrem, 2000
Turbot (<i>Scophthalmus maximus</i>)	10 – 15	Janssen et al., 2017
Shrimp (<i>Litopenaeus vannamei</i>)	11 4	Andriantahina et al., 2012 Gitterle et al., 2007

Table 2. Traits included in the breeding goal of selective breeding programs in different aquaculture species.

Species	Trait	Reference
Atlantic salmon (<i>Salmo salar</i>)	Body weight	Lhorente et al., 2019 Solberg et al., 2013
	Fat content	Horn et al., 2020
	Survival	Regan et al., 2021
	Flesh color	Lhorente et al., 2019
	Pigment content	Lhorente et al., 2019
	Fillet color	Regan et al., 2021
	Sexual maturation	Regan et al., 2021
	Resistance to furunculosis	Kjøglum et al., 2008
	Resistance to infectious salmon anemia	Kjøglum et al., 2008
	Resistance to pancreas disease	Regan et al., 2021
	Resistance to amoebic gill disease	Boudry et al., 2021
	Resistance to sea lice	Regan et al., 2021 Houston et al., 2020 Houston et al., 2017
	Resistance to salmonid rickettsial septicemia	Boudry et al., 2021 Lhorente et al., 2019
Coho salmon (<i>Oncorhynchus kisutch</i>)	Body weight	Lhorente et al., 2019 Dufflocq et al., 2017 Gallardo et al., 2010
	Fat content	Lhorente et al., 2019
	Flesh color	Lhorente et al., 2019
	Resistance to salmonid rickettsial septicemia	Lhorente et al., 2019
Gilthead sea bream (<i>Sparus aurata</i>)	Body weight	Boudry et al., 2021 Fernandes et al., 2017 Thorland et al., 2007
	Deformities	Boudry et al., 2021
	Survival to pasteurellosis	Boudry et al., 2021 Antonello et al., 2009

European sea bass (<i>Dicentrarchus labrax</i>)	Body weight	Boudry et al., 2021 Gjedrem and Rye, 2018 Dupont-Nivet et al., 2008
	Deformities	Boudry et al., 2021
Nile tilapia (<i>Oreochromis niloticus</i>)	Body weight	Yoshida et al., 2022 Joshi et al., 2020 Gjedrem and Rye, 2018
	Fillet yield	Joshi et al., 2020 Gjedrem and Rye, 2018
Rainbow trout (<i>Oncorhynchus mykiss</i>)	Body weight	Lhorente et al., 2019 Janhunen et al., 2012 Kause et al., 2007
	Fat content	Gjedrem and Rye, 2018
	Flesh color	Lhorente et al., 2019
	Feed conversion ratios	Regan et al., 2021
	Carcass yield	Boudry et al., 2021
	Resistance to bacterial coldwater disease	Regan et al., 2021 Liu et al., 2018
	Resistance to infectious pancreatic necrosis	Lhorente et al., 2019
	Resistance to salmonid rickettsial septicemia	Lhorente et al., 2019
Shrimp (<i>Litopenaeus vannamei</i>)	Body weight	Castillo-Juárez et al., 2015 Gitterle et al., 2005
	Survival	Gjedrem and Rye, 2018 Castillo-Juárez et al., 2015 Gitterle et al., 2005
	Resistance to white spot	Gjedrem and Rye, 2018 Castillo-Juárez et al., 2015

Traits to be included in complex breeding goals

Fillet and carcass yield

Fillet yield is an important trait that has gained relevance in the aquaculture industry in recent years due to its relationship with production efficiency and environmental impact, especially for species sold processed. Fillet yield is defined as the edible part of the body expressed as a percentage (i.e., the percentage of fillet weight out of wet body weight). Therefore, an increase in fillet yield translates into more edible meat (increased profitability) and less waste (decreased environmental impact) for a given amount of resources (facilities, feed, labor and energy). This trait has been included in breeding programs for some species such as Nile tilapia (Gjerde et al., 2012). However, improving fillet yield is not an easy task as recording implies slaughtering the fish and, thus, it cannot be done on selection candidates. Also, recording of fillet yield is laborious and measurement errors are large (Rutten et al., 2005; Powell et al., 2008; Haffray et al., 2012). This could be the reason why estimates of heritability for fillet yield and estimates of the genetic correlation between this trait and body weight vary greatly between and within species (Neira et al., 2004; Rutten et al., 2005; Kause et al., 2007, 2011; Navarro et al., 2009; Saillant et al., 2009; Nguyen et al., 2010; Gjerde et al., 2012).

Haffray et al. (2012, 2013) showed that selection for headless gutted carcass could be more efficient than selection for fillet yield itself in rainbow trout. However, recording headless gutted carcass percentage also requires slaughtering the fish, and therefore, it has to be done on sibs of selection candidates. When selection is based on sib's performance, the within-family component of the genetic variance cannot be exploited with traditional pedigree-based BLUP selection given that all candidates of a given family will have the same EBV. Consequently, the accuracy of breeding values is lower and inbreeding is higher than when selection is based on the performance of selection candidates themselves. All of this has led to a large amount of research in recent years devoted to find appropriate indicator traits that can be recorded on live fish and that are genetically correlated with headless carcass percentage in common carp (Prchal et al., 2018, 2020, 2021; Vandeputte et al., 2020), European sea bass (Vandeputte et al., 2017, 2020; Fraslin et al., 2018), gilthead sea bream (Fraslin et al., 2018; Vandeputte et al., 2020) and rainbow trout (Fraslin et al., 2018; Vandeputte et al., 2019, 2020).

This research aimed at finding potential indicator traits has led to the development of useful indicators of headless gutted carcass percentage such as the ratio of abdominal wall thickness to depth of the peritoneal cavity (E8/E23), measured on live fish by ultrasound tomography. The Echo8/Echo23 ratio between belly ultrasound measures is the best predictor of carcass and fillet yield in rainbow trout. This trait has a moderate heritability (about 0.3) and is highly genetically correlated with carcass yield and with headless gutted carcass yield (Haffray et al., 2013; Vandeputte et al., 2019). Similar genetic parameters have been developed for an equivalent ratio indicator in common carp (Prchal et al., 2018, 2020). These indicators can easily be measured on live breeding candidates and, therefore, selection on them allows exploiting not only the between family component of the genetic variation but also the within family component. However, the potential benefits of selecting on the indicator trait using candidates' phenotypes rather than selecting directly on headless gutted carcass using sibs' phenotypes are still unknown, particularly when applying genomic evaluations.

Disease resistance

Another trait of great economic importance for the aquaculture industries is resistance to specific diseases given that diseases represent a major threat to their economic and environmental sustainability (Houston et al., 2020). For instance, the cost of an outbreak of pancreas disease (one of the most important diseases in European and especially, in Norwegian Atlantic salmon production) at a site with one million smolts (young salmon that has undergone smoltification), has been estimated at approximately 7 million € when including direct losses and production costs (Pettersen et al., 2015). It is worth noting that the specific conditions of aquaculture production, such as the sharing of a common environment in tanks or cages and the water flow, facilitate the transmission of infectious diseases and increase the difficulty of controlling them even more than in terrestrial species.

Selective breeding for disease resistance could reduce the financial and environmental costs of disease outbreaks and increase animal welfare (Bishop and Woolliams, 2014; Colditz and Hine, 2016; Houston et al., 2020). Disease resistance, defined as the ability to survive when exposed to the pathogen, is included as a breeding objective in species with advanced selective breeding programs such as Atlantic salmon.

In fact, most selective breeding programs for this species have already included resistance to various diseases in the breeding goal (Houston, 2017). Different studies (particularly in salmonids and shrimp) have shown that resistance against specific diseases has a moderate to high heritability and thus, there is a clear potential for genetic improvement (see reviews by Ødegård et al., 2011; Gjedrem and Rye, 2018). Selection for infectious pancreatic necrosis and pancreatic disease in salmonids has been shown to be very efficient, reducing the incidence of outbreaks which in the case of infectious pancreatic necrosis has become almost zero (Gjedrem, 2015; Norris, 2017; Houston et al., 2020). In recent years, there has been an increase in studies estimating genetic parameters for resistance to important diseases in other species. Some examples are estimates for koi herpesvirus disease in carp (Zao et al., 2020), viral nervous necrosis in sea bass (Palaïokostas et al., 2018), photobacteriosis and parasitosis by *Sparicotyle chrysophrii* in sea bream (Aslam et al., 2018, 2020) and scuticociliatosis in turbot (Saura et al., 2019).

The main phenotype considered for disease resistance is individual mortality recorded as a binary trait or as survival time (Ødegård et al., 2011; Doeschl-Wilson et al., 2021). This implies exposing or inoculating the pathogen causing a specific disease to the fish in challenge tests. For safety reasons, the challenge tests are performed in test stations, separated from the breeding nucleus. Thus, recording cannot be done on selection candidates and it is done on their sibs and as in the case of headless carcass, the within-family component of the genetic variance cannot be exploited with traditional BLUP selection.

Selection for disease resistance is aimed at improving individual resilience to infection that is, strengthening an animal's ability to cope with infection (i.e., fish are less likely to die when exposed to the pathogen). However, it is unknown whether selecting directly on time of death reduces not only mortality but also pathogen transmission and, thus, the incidence and severity of outbreaks. Disease survivors under field conditions often cannot be sold due to secondary infections, parasitism, and poor quality, even when the mortality rate is low or moderate. Thus, reducing the incidence and prevalence of an infectious disease, and not only the mortality (by improving disease resistance), is an important breeding goal in commercial aquaculture.

As already stated, recording mortality may not be the optimal approach to deal with disease outbreaks as other important components of a population's ability to deal

with pathogens are not taken into account. Diverse studies have shown that animals vary not only in their individual disease resistance but also in those epidemiological traits relating to environmental pathogen load (Doeschl-Wilson et al., 2021). There are three epidemiological host traits related to pathogen transmission that are key to the sustainable disease control but that cannot be directly observed or measured: i) susceptibility, defined as the propensity of an uninfected individual to become infected when exposed to infectious material; ii) infectivity, defined as the ability of an individual, once infected, to transmit the infection; and iii) infection-induced mortality, defined as the propensity of an infected individual to die (Tsairidou et al., 2019; Anacleto et al., 2019; Saura et al., 2019; Chase-Topping et al., 2021). Although these underlying traits play an important role in disease transmission, it is unknown the impact it will have on them selecting on mortality.

Uniformity

Another trait with growing relevance in both terrestrial and aquatic farm species is uniformity of the final product. The increase in product uniformity can be achieved by selection aimed at decreasing the variability between individuals for a given trait. This is of great importance due to the relationship of uniformity with economic efficiency. For example, large variation in body size can cause competition between individuals and increases the need for management practices such as grading (i.e., classifying and separating fish according to size). It has been shown that body weight uniformity depends on the sensitivity of an individual to macro-environmental measurable factors (e.g., temperature and diet) and micro-environmental non-measurable animal-specific factors (Sae-Lim et al., 2016; de Souza et al., 2019). The existence of genetic variation in micro-environmental sensitivity among individuals is needed to improve body weight uniformity and indicates that the phenotypes of individuals with different genotypes will be differentially affected by the environment (Waddington, 1960; Scheiner et al., 1991; Sancristobal-Gaudy et al., 1998; Ibáñez-Escriche et al., 2008b). That means that individuals with less environmental variance will be less sensitive to environmental disturbances. Therefore, another indirect benefit of improving uniformity is its potential to improve resilience, which is defined as the ability of an animal to maintain performance in spite of environmental perturbations (de Souza et al., 2019). The existence of genetic variation in micro-environmental sensitivity has been demonstrated for traits such as

weight and litter size in rabbits and mice (Ibáñez-Escriche et al., 2008a; Hill and Mulder, 2010). In addition, selecting to reduce the environmental variance of birth weight in mice (Formoso-Rafferty et al., 2016, 2020) has been successful.

Despite the great interest in improving product uniformity, studies in aquaculture species are very scarce (Sae-Lim et al., 2016; de Souza et al., 2019). Some recent studies in salmonids and in Nile tilapia have shown the existence of genetic variation for weight uniformity (see review by de Souza et al., 2019) but to our knowledge, there are no estimates of genetic variability for weight uniformity in crustaceans. In particular, there is a real interest in understanding the genetic control of uniformity in shrimp weight at slaughter given that the final product for the industry of this species, is the individual, and thus, uniformity has great economic importance. Shrimp are graded according to their size and the number of individuals per unit of weight. There are different categories according to the number of individuals included per weight unit (kg in Europe) and these vary enormously in terms of the final economic benefit of the industry. A larger number of shrimps per weight unit results in a price reduction. Therefore, increasing the consistency of size within a specific count range can increase profit margins of the industry.

Genomic selection

The rapid advances in genomics in the last decades have opened up new opportunities for boosting selective breeding programs. Genomics has the potential to increase the accuracy of EBVs, particularly for lowly-heritable traits (such as uniformity) and for traits that are difficult or expensive to record, for which pedigree-based selection is generally not very effective (Wientjes et al., 2022). In aquaculture, genomic information is especially useful for sib traits such as fillet and carcass yield and disease resistance for which recording is performed on sibs of the selection candidates (Gjedrem et al., 2010). For these traits, BLUP selection (i.e., selection based on BLUP-EBVs) only utilizes the between-family component of the genetic variance and ignores the within-family component (Nielsen et al., 2009; Sonesson and Meuwissen, 2009). The use of genomic information in selection decisions either through marker-assisted selection (MAS) (Fernando and Grossman, 1982; Sonesson, 2007) or through genomic selection

(Meuwissen et al., 2001; Sonesson and Meuwissen, 2009) allows exploiting the within-family component increasing thus the efficiency of selection.

MAS is based on the use of targeted markers linked to major quantitative traits loci (QTL) which are regions of the genome that explain variation for quantitative traits. This requires that QTLs are detected through GWAS (Genome-Wide Association Studies) and that carriers of the favorable alleles are identified through the markers. Houston et al. (2008) and Moen et al. (2009) found a single QTL explaining 83% of genetic variation in infectious pancreatic necrosis (IPN) virus resistance in Atlantic salmon. MAS based on this major QTL reduced the incidence of outbreaks to almost zero in few generations (Norris 2017). Also, a QTL for pancreatic disease has been found to explain a considerable part of the variation in resistance to this disease (Gonen et al., 2015).

However, in aquaculture (and also in terrestrial livestock), most economically important traits are controlled by a very large (theoretically infinite) number of genes, each with a very small effect (theoretically infinitesimal). The genes of this nature are called polygenes. For these traits, the global impact of MAS has been very limited in animal breeding programs. Firstly, since there are usually many genes affecting a given trait, the proportion of the genetic variance explained by just one of them is small. Secondly, the effects of polygenes are difficult to estimate precisely since they are generally small effects. Thus, it is very difficult to detect a large enough number of QTLs to explain a large proportion of the genetic variance for most traits of interest (Goddard, 2009; Zenger et al., 2019; Houston et al., 2020).

A different approach using markers is to apply genomic selection (Meuwissen et al., 2001). As initially developed, genomic selection was based on the estimation of the effects of thousands of markers covering the entire genome. The high density of markers would ensure that each QTL is in linkage disequilibrium with at least one marker. Genomic selection (selection on genomic EBVs) is therefore large-scale MAS and does not require prior detection of QTL (Meuwissen et al., 2001; Meuwissen et al., 2016). For sib traits, genomic selection involves two steps when obtaining genomic EBVs for fish with no phenotypes. In a first step, estimates of the effects of markers are obtained in a test population composed by sibs of the candidates that are genotyped and phenotyped for the sib trait. In a second step, these estimates are used to obtain genomic EBVs for the

selection candidates that are genotyped, but not phenotyped. Consequently, with genomic selection the candidates from the same family will differ in their genomic EBVs, and thus both the between- and the within-family genetic variation can be exploited for sib traits. The closer the relationship between the test population and the selection candidates the higher the selection accuracy, even at low marker densities. Also, given that this approach has the potential of increasing the accuracies of estimated genetic effects, it can help in disentangling them from tank effects.

There are currently numerous available genomic selection methodologies (e.g., Meuwissen et al., 2016; Wang et al., 2018; Kjetsa et al., 2020). Most studies in aquaculture species use the GBLUP (Genomic Best Linear Unbiased Prediction) method (Zenger et al., 2019). In GBLUP, the pedigree-based numerator relationship matrix (used in standard BLUP) is replaced with the realized genomic relationship matrix computed with markers (Nejati-Javaremi et al., 1997; Villanueva et al., 2005; Hayes et al., 2009). GBLUP is equivalent (i.e., it yields identical GEBVs) to the SNP-BLUP method (Habier et al., 2007; VanRaden, 2008; Goddard, 2009). The prior information in these methods assumes that SNP effects are normally distributed with the same variance for every SNP. For this reason, some approaches have been developed that allow the incorporation of a priori information for SNP effects (e.g., Bayes A, Bayes B, Bayes C, Bayes D, and Bayesian Lasso; see Gianola, 2013). However, when they have been tested in several aquaculture species, no consistent differences were found in the prediction accuracy when compared with the simpler GBLUP approach (Dagnachew and Meuwissen, 2019; Houston et al., 2020; Kjetsa et al., 2020).

The application of genomic selection can significantly increase the genetic response (Georges et al., 2019; Houston et al., 2020) and decrease inbreeding (Daetwyler et al., 2007) in aquaculture selective breeding programs. The benefits of genomic selection (when compared with standard BLUP prediction) in terms of genetic gain and inbreeding have been shown through computer simulations (Nielsen et al., 2009, 2011; Sonesson and Meuwissen, 2009; Villanueva et al., 2011; Lillehammer et al., 2013; Sonesson and Ødegård, 2016) and analyses of empirical data for different traits and aquaculture species (reviewed by Houston et al., 2020). However, the application of genomic selection implies a great economic inversion due to genotyping costs. Only the most advanced aquaculture species such as salmonids currently use genomic selection in their breeding programs, while for the rest of the aquaculture species molecular genetic

tools are mostly used for pedigree reconstruction (Chavanne et al., 2016; Houston et al., 2020; Regan et al., 2021). Further research to predict the potential benefits of genomic selection on other species of lower economic value and/or other traits is necessary prior the investment and application in the aquaculture industry.

In aquaculture species, genomic selection was first tested in Atlantic salmon using high-density SNP arrays (Houston et al., 2014; Ødegård et al., 2014), and its utility in accurately predicting breeding values has been shown in a typical salmon breeding program setting (Ødegård et al., 2014; Yáñez et al., 2016). SNP arrays of different densities have been also developed for other aquaculture species (Table 3). Some arrays combine SNPs for different aquaculture species in order to reduce costs. This is the case for the combined SNP array for two oyster species (Gutiérrez et al., 2017) and for sea bream and sea bass (Peñazola et al., 2021) (Table 3). More SNP arrays are expected to be developed in the coming years, and this, along with the development of low-cost genotyping technologies (associated with the new advances in next-generation sequencing), make genomic evaluations more affordable for aquaculture species opening a new door to the genomic era (Joshi et al., 2020; Houston et al., 2020).

Table 3. Commercial SNP arrays developed for different aquaculture species.

Species	Number of SNPs in the array	Reference
Atlantic cod (<i>Gadus morhua</i>)	~10K	Pocwierz-Kotus et al., 2015
Atlantic salmon (<i>Salmo salar</i>)	~200K ~130K	Yáñez et al., 2016 Houston et al., 2014
Coho salmon (<i>Oncorhynchus kisutch</i>)	~200K	Barría et al., 2019
Common carp (<i>Cyprinus carpio</i>)	~250K ~12K	Xu et al., 2014

European flat oyster (<i>Ostrea edulis</i>)	~11K [†]	Gutiérrez et al., 2017
European sea bass (<i>Dicentrarchus labrax</i>)	~30K [‡] ~57K ~57K ~3K	Peñazola et al., 2021 Griot et al., 2021a Vandeputte et al., 2019 Faggion et al., 2019
Gilthead sea bream (<i>Sparus aurata</i>)	~30K [‡] ~12K	Peñazola et al., 2021 Palaikostas et al., 2016
Nile Tilapia (<i>Oreochromis niloticus</i>)	~65K ~50K ~58K	Peñazola et al., 2020 Yáñez et al., 2020 Joshi et al., 2018
Pacific oyster (<i>Crassostrea gigas</i>)	~27K [†]	Gutiérrez et al., 2017 Qi et al., 2017
Rainbow trout (<i>Oncorhynchus mykiss</i>)	~50K ~57K	Salem et al., 2018 Palti et al., 2015
Whiteleg shrimp (<i>Litopenaeus vannamei</i>)	~19K ~9K	Lillehammer et al., 2020 Jones et al., 2017

[†]Combined ~41K SNP array, with approximately 11K for European flat oyster and 27K Pacific oyster.

[‡]Combined ~60K SNP array, with approximately 30K for European sea bass and 30K for gilthead sea bream.



OBJECTIVES





The general aim of the thesis was to advance aquaculture selective breeding programs by including new traits of economic importance in the breeding goal and implementing genomic tools such as genomic selection.

As already mentioned, a general problem in aquaculture breeding programs raising families separately during the first stages is the problem of having a common environmental effect that leads to inaccurate genetic evaluations and hinders genetic progress. This problem could at least be reduced if genomic rather than traditional BLUP evaluations are performed and this is tested in **Chapter 1**. Also, in general, the main breeding goal trait in aquaculture selection programs is growth rate although there are other economically important traits that could be included in the breeding objective. Some of these traits cannot be recorded on live fish and they are challenging to be incorporated in these programs. This is the case of carcass yield and fillet yield, two economically very important traits, especially for species sold processed as gutted carcass or fillets such as rainbow trout. A potential solution would be to select for an indicator trait that can be recorded on live fish and that is genetically correlated with the goal traits or alternatively, to apply genomic selection. Both alternatives could be efficient to improve carcass yield and fillet yield and they are compared in **Chapter 2**. Another very important trait in aquaculture breeding programs that cannot be measured on live selection candidates is resistance to infectious diseases. An extra complication with these traits is that the main recorded phenotype (i.e., mortality) and the dynamics of the spread of the disease are affected by three underlying (unobserved) epidemiological host traits (susceptibility, infectivity, and infection-induced mortality). Thus, in order to reduce the incidence and prevalence of infectious diseases, genetic models need to be combined with epidemiological models. In **Chapter 3**, the impact of selection for disease resistance measured as time to death on the three underlying traits and, therefore, on the spread of the disease, is investigated by simulating a selective breeding program for Atlantic salmon. Finally, another important trait in aquaculture breeding programs is weight uniformity, particularly for species such as shrimp. However, a necessary condition to improve this trait through selective breeding is the existence of genetic variability for the environmental variance of body weight. In **Chapter 4** the existence of such variation and its genetic correlation with body weight in a population of Pacific white shrimp (*Litopenaeus vannamei*) are estimated using industry data, to investigate the viability to include weight uniformity in the breeding goal.

Specific objectives of this thesis

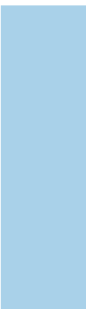
1. Predict, through computer simulations, the extra benefits obtained from genomic evaluation in aquaculture selective breeding programs where families are kept separately and common environmental effects are present.
2. Evaluate the potential of indirect and genomic selection to improve carcass yield in a simulated rainbow trout selective breeding program.
3. Predict the impact of selection for disease resistance on the spread of infection in a simulated salmon population.
4. Estimate genetic parameters for weight uniformity in a farm population of Pacific white shrimp to determine the potential of including this trait in the selection program.

SOFTWARE DEVELOPED

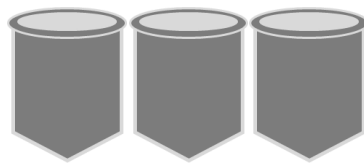




The challenges described above under **Objectives 1 - 3** were tackled through the use of computer simulations. Thus, in the initial steps of this thesis, a software was developed in Fortran 90 within the European project FISHBOOST (<http://cordis.europa.eu/project/id/613611>) for predicting genetic gain and inbreeding expected for a wide range of aquaculture breeding scenarios. The name of the program is **GAINFISH** where GA comes from gain and IN comes from inbreeding. **GAINFISH** was thoroughly tested by running a very large set of scenarios varying in the species characteristics, traits, evaluation, selection, and mating approaches. The **GAINFISH** tool is a suite of programs for windows or linux environments. It is very flexible as it considers a wide range of scenarios including different i) breeding objectives; ii) types of family rearing; iii) genetic models for the different traits; iv) evaluation strategies; v) selection strategies; and vi) mating strategies. The manual of **GAINFISH** is included in Annex I.



CHAPTER 1



García-Ballesteros, S., Fernández, J., Toro, M.Á., and Villanueva, B. (2021). **Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families.** *Aquaculture*, 543, 1-8.

<https://doi.org/10.1016/j.aquaculture.2021.737004>

2021 Journal impact factor 5.135, JCR Rank 5/54 in the *Fisheries* category and 8/113 in the *Marine & Freshwater Biology* category.



Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families

Silvia García-Ballesteros^{a,*}, Jesús Fernández^a, Miguel Ángel Toro^b, Beatriz Villanueva^a

^a Departamento de Mejora Genética Animal, INIA (CSIC), Madrid 28040, Spain

^b Departamento de Producción Agraria, ETSI Agronómica, Alimentaria y de Biosistemas, UPM, Madrid 28040, Spain

ARTICLE INFO

Keywords:

GBLUP
Common environmental effect
Fish
Aquaculture breeding programs

ABSTRACT

A common type of rearing in aquaculture selective breeding programs is to keep families in separate tanks until fish is large enough to be individually tagged. This type of rearing introduces an environmental effect common to the members of the same family which can be difficult to be disentangled from genetic effects when traditional BLUP evaluation is used. The use of genomic evaluation could lead to a better prediction of these effects as markers can capture more accurately genetic relationships between fish raised in different tanks and between full-sibs raised in the same tank. The objective of this simulation study was to predict the benefit of genomic evaluation (GBLUP) when compared to BLUP in the presence of common environmental effects. Models that included or excluded these effects when they were present were compared. Populations with two different levels of linkage disequilibrium were generated. The heritability of the selected trait was 0.4 and three different proportions of phenotypic variance due to common environmental effect (0.0, 0.1 and 0.3) were considered. Nested mating designs where each male was mated to two females were simulated to create 200 full-sib families with 10 or 50 fish per family. GBLUP was less affected than BLUP by the presence of common environmental effects. The loss in genetic gain ranged from 0% to 4% with GBLUP and from 2% to 7% with BLUP. As a consequence, the benefit of GBLUP increased in the presence of common environmental effects, and the extra benefit increased with increasing linkage disequilibrium and family size. In the absence of these effects, genetic gain from GBLUP was up 20% higher than from BLUP and in their presence, genetic gain from GBLUP was 27% higher than from BLUP. Although GBLUP always led to lower inbreeding than BLUP, there were no extra benefits in terms of extra reductions in the presence of common environmental effects. Finally, when these effects were present but ignored in the evaluation model, the loss in genetic gain was moderate but inbreeding increased dramatically, particularly in BLUP scenarios.

1. Introduction

In selective breeding programs, tracking pedigrees across generations is essential for increasing the accuracy of selection by making use of information not only from the candidates for selection, but also from their relatives. Tracking pedigrees is also essential for controlling the increase in inbreeding. In aquaculture programs, given that fish at hatching are too small to be tagged physically, two different types of rearing at early stages of development are frequently used to record pedigrees: i) keep families separately until fish is large enough to be individually tagged (e.g. as in the Norwegian salmon program; Gjedrem, 2010); or ii) keep families together from the start and genotype subsequently with molecular markers (e.g. microsatellites) to assign offspring

to parents (e.g. as in the Aquicultura Balear SAU gilthead sea bream program; Fernandes et al., 2017).

There are potential problems with both types of rearing. Rearing families together can lead to unequal family contributions and hence to high risk of inbreeding (Chatziplis et al., 2007; Herlin et al., 2008; Borrell et al., 2011; Fernandes et al., 2017). Rearing families separately can avoid this problem. However, keeping families in different tanks during the initial phases of development can introduce an environmental effect common to the members of the same family (usually known as the tank effect) that is difficult to disentangle from genetic effects reducing thus the accuracy of estimated breeding values of candidates to selection and the expected genetic gain (Gjerde et al., 1996; Gjedrem and Baranski, 2010).

* Corresponding author.

E-mail address: silvia.garciab@inia.es (S. García-Ballesteros).

Many studies focusing on growth traits have shown that the proportion of phenotypic variance due to common environmental effects is significantly different from 0 (even after more than a year of communal rearing) in Atlantic (*Salmo salar*; Rye and Mao, 1998), chinook (*Oncorhynchus tshawytscha*; Winkelman and Peterson, 1994) and coho (*Oncorhynchus kisutch*; Gallardo et al., 2010) salmon, rainbow trout (*Oncorhynchus mykiss*; Pante et al., 2002; Kause et al., 2005, 2006), Nile tilapia (*Oreochromis niloticus*; Agha et al., 2018) and Atlantic cod (*Gadus morhua*; Tosh et al., 2010). Common environmental effects have also been shown to be significant for disease resistance traits in salmonid species (Kjøglum et al., 2005, 2008). These results indicate that these effects should be included in genetic evaluation models when families are reared separately.

Computer simulations (Nielsen et al., 2009, 2011; Sonesson and Meuwissen, 2009; Villanueva et al., 2011; Lillehammer et al., 2013; Sonesson and Ødegård, 2016) and analyses of empirical data of different aquaculture species (see review by Houston et al., 2020) have shown important benefits of genomic evaluation in aquaculture breeding programs when compared to traditional BLUP evaluation. While empirical studies have focused on benefits of genomic evaluation in terms of prediction accuracy (and consequently genetic gain), simulation studies have also shown benefits in terms of accumulated inbreeding across generations. However, these benefits have been assessed in the absence of common environmental effects. In their presence, the benefit of genomic selection could be even greater given that markers can capture more accurately genetic relationships between genotyped individuals (Habier et al., 2007, 2013). In particular, genetic relationships between fish raised in different tanks that are pedigree unrelated, and between full-sibs raised in the same tank could be determined from the markers. This could increase the accuracy with which both common environmental and genetic effects are estimated.

The objective of this study was to predict, through computer simulations, the benefit of genomic evaluation in terms of genetic gain and inbreeding when compared to BLUP evaluation in aquaculture selective breeding schemes where families are kept separately and common environmental effects are present.

2. Material and methods

Scenarios simulated involved selection for several discrete generations on a trait with a heritability of 0.4 that was measured in all individuals. Nested mating designs where each male was mated with two females were considered. Full-sib families were simulated as being raised separately in different tanks and each tank had a different effect on the trait phenotype. Genetic evaluations were based on BLUP or GBLUP models. The selection program started from different base populations that were created in 2 steps. Firstly, a large population at mutation-drift equilibrium was generated. Secondly, from this population at equilibrium, base populations with family structure were generated. The different base populations varied in the amount of linkage disequilibrium (LD) between loci. The simulations were carried out with our own Fortran 90 codes.

2.1. Generation of the population at mutation-drift equilibrium

The genome consisted of 30 chromosomes and had a total size of 20 Morgans. Each chromosome carried 10,000 evenly spaced biallelic loci that included 2000 SNP markers to be used in genomic evaluations (i.e. a density of 3000 SNP/Morgan). SNP markers were defined at the start of the process and they were also evenly spaced within each chromosome. The population in mutation-drift equilibrium was generated by randomly selecting 500 males and 500 females across 4000 discrete generations. The mutation rate per locus and generation was 2.5×10^{-3} . Mutations were randomly distributed across individuals, chromosomes and loci, switching allele 0 to allele 1 and *viceversa*. Initial frequencies were 0.5 for all loci. When generating the gametes, the number of

crossovers per chromosome was drawn from a Poisson distribution with mean equal to the genome length. Crossovers were randomly distributed without interference. At the end of this process, the expected heterozygosity of the population had already reached an equilibrium value (mutation-drift equilibrium). At this point, all 300,000 loci simulated (including SNPs) were still segregating and 1000 of them were randomly sampled as QTLs affecting the trait selected.

2.2. Generation of the base population

After the QTLs were simulated, two different base populations ($t = 0$) that varied in the amount of LD between SNPs and QTLs were considered. The LD levels simulated were based on those found in the literature for aquaculture species. These levels vary substantially across populations (Kijas et al., 2016; Glover et al., 2017; Barría et al., 2018, 2019; D'Ambrosio et al., 2019).

- A. Base populations with low LD. In these scenarios, 100 males and 200 females were randomly sampled from the population at mutation-drift equilibrium for each replicate. Then, in order to create a family structure, each male was mated randomly with two females, following a nested design and the number of offspring per family (n) was 10 or 50. Thus, in these scenarios the base population consisted of 2000 ($n = 10$) or 10,000 ($n = 50$) fish belonging to 200 full-sib families (and 100 half-sib families). The average LD, measured by r^2 (Hill and Robertson, 1968), between a particular QTL and the closest SNP at $t = 0$ was 0.03 (SD = 0.03).
- B. Base populations with high LD. In these scenarios, as before, 100 males and 200 females were randomly sampled each replicate from the population at mutation-drift equilibrium. Then, the 10 males and 10 females with the highest phenotypic values (see simulation of phenotypes below) were selected and mated randomly. Each male was mated with a single female and each female was mated with a single male and each mating produced 30 offspring (10 males and 20 females). This artificial phenotypic selection was applied for 10 generations. After that, in order to create the base population, each male was mated randomly with two females and the number of offspring per family (n) was 10 or 50. Thus, in these scenarios, the base population also consisted of 2000 ($n = 10$) or 10,000 ($n = 50$) fish belonging to 200 full-sib families (and 100 half-sib families). The average LD between a particular QTL and the closest SNP at $t = 0$ was 0.19 (SD = 0.19).

2.3. Generation of breeding and phenotypic values

The additive effect for QTL i (a_i) was obtained by multiplying a random number sampled from a normal distribution with mean 0 and variance 1 by the factor $\sqrt{\{V_A/[2p(1-p)n_{qtl}]\}}$, where V_A is the additive variance at $t = 0$ in scenarios with low LD and at $t = -10$ in scenarios with high LD, p is the average frequency across QTLs (at $t = 0$ in scenarios with low LD and at $t = -10$ in scenarios with high LD) and n_{qtl} is the number of QTLs (i.e. 1000). Note that, in this way, the expected additive variance summed over all loci equals V_A , assuming that covariances between loci generated by LD are negligible. The true breeding values were obtained as the sum across QTLs of a_i , 0 or $-a_i$ for homozygotes 11, heterozygotes, and homozygotes 00, respectively. In order to have the same value for the additive variance in the base population ($t = 0$) in all scenarios simulated, V_A was set to 0.4 in scenarios with low LD and to 0.5 in scenarios with high LD. Similarly, in order to have the same value for the phenotypic variance at $t = 0$, V_P was set to 1 in scenarios with low LD and to 1.25 in scenarios with high LD. In this way, for all scenarios at $t = 0$ the additive and phenotypic variances were 0.4 and 1, respectively, and the heritability was 0.4. At $t = 0$ a common environmental effect was sampled for each full-sib family from a normal distribution with mean 0 and variance c^2 , where c^2 is the proportion of V_P due to common environmental effects. Three different

values for c^2 (0.0, 0.1 and 0.3) were considered and they remained constant across generations. Finally, phenotypic values were obtained by adding a normally distributed individual environmental effect with mean 0 and variance $V_E = V_P - V_A - c^2$ to the sum of true breeding value and common environmental effect.

2.4. Genetic evaluation methods

The evaluation methods considered for obtaining estimates of breeding values were standard BLUP and genomic evaluation. In both cases, models taking into account or ignoring the common environmental effects when they are present were considered. When the common environmental effects were taken into account, the model assumed for BLUP was

$$y = 1\mu + Zu + Wc + e$$

where y is the vector of observations, μ is the mean, u is the vector of random additive genetic effects, c is the vector of random common environmental effects, e is the vector of random residuals, 1 is a vector of ones and Z and W are the corresponding incidence matrices. The variance-covariance matrices of the random effects were assumed to be $V(u) = A\sigma_u^2$, $V(c) = I_m\sigma_c^2$ and $V(e) = I_r\sigma_e^2$, where A is the pedigree-based numerator relationship matrix, I_m and I_r are identity matrices of order m (number of full-sib families multiplied by the generation number) and r (number of records), respectively, and σ_u^2 , σ_c^2 and σ_e^2 are the additive genetic, common environmental and residual variances, respectively. These 3 variances were assumed to be known. When the common environmental effect was ignored, the term Wc was excluded from the model and $V(e)$ was assumed to be $I_r\sigma_e^2 + I_m\sigma_c^2$.

The genomic evaluations were performed using the GBLUP method (Nejati-Javaremi et al., 1997; Villanueva et al., 2005; Hayes et al., 2009). The assumed models (taking into account or ignoring the common environmental effects) were the same as in BLUP but here the pedigree-based numerator relationship matrix (A) was replaced by the genomic relationship matrix (G) of VanRaden (2008) method 1; i.e., $V(u)$ was assumed to be $G\sigma_u^2$.

For a particular generation t , the information included in both BLUP (phenotypes and pedigree) and GBLUP (phenotypes and genotypes) was that from the selection candidates born at t and from all fish born in previous generations. The BLUPf90 software (Misztal et al., 2015) was used for carrying out both BLUP and GBLUP evaluations.

Results from BLUP scenarios are not affected by the amount of LD present at $t = 0$. However, results from GBLUP scenarios would depend on the amount of LD between SNPs and QTLs. GBLUP scenarios originating from base populations with low LD will be referred to as GBLUP_L and GBLUP scenarios originating from base populations with high LD will be referred to as GBLUP_H.

2.5. Selection and mating

All fish born at a given generation were considered to be selection candidates and the phenotype for the selected trait was available for all of them. Each generation the 100 males and 200 females with the highest estimated breeding values were selected and randomly mated. Nested mating designs where each male was randomly mated with two females were applied and each mating produced n offspring ($n/2$ males and $n/2$ females). Schemes with different family sizes ($n = 10$ or 50) were simulated. Selection for schemes with $n = 10$ lasted for 10 generations, and selection for schemes with $n = 50$ lasted for 4 generations. In all cases, generations were discrete.

2.6. Comparison of scenarios

Scenarios were compared in terms of genetic gain and inbreeding. Each generation t , the cumulative genetic gain and inbreeding were

computed, respectively, as the true breeding value and the genealogical inbreeding coefficient averaged across all fish born in t . Also, the accuracy of evaluation was calculated as the correlation between true and estimated breeding values. Each scenario was replicated 50 times and the results presented are averages over replicates.

3. Results

3.1. Comparison of BLUP and GBLUP scenarios in terms of genetic gain and inbreeding

As expected, in all scenarios, GBLUP led to higher cumulative gains than BLUP (even with low LD) and GBLUP_H led to higher gains than GBLUP_L. The benefit of GBLUP over BLUP increased in the presence of common environmental effects (Tables 1 and 2). For $n = 10$ (Table 1), the benefit of GBLUP_L when compared to BLUP at $t = 10$ increased from 9% in the absence of these effects (i.e. $c^2 = 0.0$) to 11% in their presence (with both $c^2 = 0.1$ and $c^2 = 0.3$). Corresponding figures for GBLUP_H were 15% and 17%. The benefit from GBLUP relative to BLUP also increased with increasing family size (Tables 1 and 2). For $n = 10$ at $t = 4$, this benefit was 15% ($c^2 = 0.0$) and 18% ($c^2 > 0.0$) for GBLUP_H scenarios. Equivalent figures for $n = 50$ were 20% ($c^2 = 0.0$) and 28% ($c^2 > 0.0$). The differences in genetic gain across the different scenarios was mainly a reflection of the differences in the accuracies of breeding values (Fig. 1).

In general, the presence of common environmental effects ($c^2 > 0.0$) led to a loss in genetic gain but this loss was smaller with GBLUP than with BLUP (Tables 1 and 2). This explains the higher benefit of GBLUP when compared with BLUP when $c^2 > 0.0$. For instance, for $n = 10$ the loss in gain at $t = 10$ was 4% ($c^2 = 0.1$) and 2% ($c^2 = 0.3$) in BLUP scenarios and 2% ($c^2 = 0.1$) and 1% ($c^2 = 0.3$) in GBLUP_L scenarios. Also, in general, the loss in gain in the presence of common environmental effects was higher with a larger family size. For $n = 50$ (Table 2), the loss in gain at $t = 4$ was 7% ($c^2 = 0.1$) and 6% ($c^2 = 0.3$) in BLUP scenarios and 4% ($c^2 = 0.1$) and 3% ($c^2 = 0.3$) in GBLUP_L scenarios. However, for $n = 10$, the loss in gain when $c^2 > 0$ at $t = 4$ was similar to that described for $t = 10$ (i.e., up to 4%). Note that GBLUP_H with $c^2 = 0.0$ and $c^2 = 0.3$ led to similar genetic gains for both n .

The benefit of GBLUP_L relative to BLUP in terms of selection response increased with increasing generation number for any c^2 (Tables 1 and 2). This was not the case with GBLUP_H for $c^2 > 0.0$. In this scenario, the maximum benefit was reached in the initial generations of selection and remained practically constant afterwards. Note that the benefit (in %)

was computed as: $100 \left(\frac{\overline{TBV}_{GBLUP} - \overline{TBV}_{BLUP}}{\overline{TBV}_{BLUP}} \right)$, where

$\overline{TBV}_{GBLUP}$ and $\overline{TBV}_{BLUP}$ are the average true breeding values from GBLUP and BLUP schemes, respectively, at a given generation t . For $c^2 = 0.0$ and $n = 10$, the benefit of GBLUP_L (GBLUP_H) relative to BLUP was 1% (11%) at $t = 1$ and 9% (15%) at $t = 10$. For $c^2 > 0.0$ (e.g., 0.3) and $n = 10$, the benefit of GBLUP_L (GBLUP_H) relative to BLUP was 5% (28%) at $t = 1$ and 10% (17%) at $t = 10$.

The accuracy of estimated common environmental effects for BLUP and GBLUP scenarios with $c^2 > 0.0$ and $n = 10$ is shown in Fig. 2. In line with the results already described for genetic gain, in the presence of common environmental effects GBLUP led to higher accuracies of estimated common environmental effects than BLUP both for $c^2 = 0.1$ and $c^2 = 0.3$, particularly in scenarios with high LD (GBLUP_H). As expected, the accuracy for common environmental effects was substantially higher for a higher c^2 (0.3) than for a lower c^2 (0.1) in all scenarios (GBLUP and BLUP). These findings were also observed for scenarios with $n = 50$ (data not shown).

GBLUP scenarios did not only lead to higher gains than BLUP but also to lower inbreeding (Tables 1 and 2) both in the absence and the presence of common environmental effects. Also, GBLUP_H led to lower inbreeding than GBLUP_L for all c^2 values. In addition, lower inbreeding

Table 1

Cumulative genetic gain and inbreeding ($\times 10^2$) across generations (t) from BLUP and GBLUP scenarios with different proportions of the phenotypic variance due to common environmental effects (c^2) and a family size of 10 fish. GBLUP scenarios with low (GBLUP_L) and high (GBLUP_H) linkage disequilibrium were considered.

t	$c^2 = 0.0$			$c^2 = 0.1$			$c^2 = 0.3$		
	BLUP	GBLUP _L	GBLUP _H	BLUP	GBLUP _L	GBLUP _H	BLUP	GBLUP _L	GBLUP _H
Genetic gain									
1	0.73	0.74	0.81	0.66	0.68	0.79	0.63	0.66	0.81
2	1.31	1.34	1.48	1.23	1.28	1.45	1.23	1.27	1.47
3	1.86	1.93	2.11	1.77	1.86	2.08	1.78	1.86	2.11
4	2.39	2.50	2.74	2.29	2.44	2.69	2.32	2.44	2.74
5	2.92	3.08	3.35	2.81	3.02	3.29	2.83	3.02	3.37
10	5.45	5.94	6.25	5.26	5.82	6.14	5.33	5.87	6.26
Inbreeding									
1	0.61	0.60	0.50	0.55	0.52	0.45	0.41	0.43	0.44
2	1.43	1.30	1.00	1.33	1.22	0.96	1.01	1.00	0.85
3	2.20	1.98	1.37	2.12	1.91	1.37	1.61	1.46	1.26
4	2.81	2.51	1.85	2.73	2.51	1.74	2.08	2.00	1.63
5	3.43	3.00	2.18	3.36	3.06	2.11	2.57	2.33	1.99
10	6.43	5.05	3.86	6.23	5.12	3.71	4.72	4.05	3.39

Standard errors were equal to 0.05 for GBLUP_H, 0.01 for both GBLUP_L and BLUP scenarios for genetic gain and 0.01 for inbreeding (in all scenarios).

Table 2

Cumulative genetic gain and pedigree inbreeding ($\times 10^2$) across generations (t) from BLUP and GBLUP scenarios with different proportions of the phenotypic variance due to common environmental effects (c^2) and a family size of 50 fish. GBLUP scenarios with low (GBLUP_L) and high (GBLUP_H) linkage disequilibrium were considered.

t	$c^2 = 0.0$			$c^2 = 0.1$			$c^2 = 0.3$		
	BLUP	GBLUP _L	GBLUP _H	BLUP	GBLUP _L	GBLUP _H	BLUP	GBLUP _L	GBLUP _H
Genetic gain									
1	1.09	1.15	1.29	0.96	1.05	1.25	0.91	1.03	1.29
2	1.94	2.12	2.29	1.76	2.01	2.25	1.76	2.02	2.29
3	2.72	3.05	3.24	2.50	2.92	3.18	2.53	2.94	3.23
4	3.47	3.97	4.15	3.23	3.81	4.09	3.25	3.84	4.15
Inbreeding									
1	2.07	1.57	1.17	1.66	1.31	1.05	1.01	0.93	0.97
2	4.71	3.37	2.30	4.07	3.01	2.03	2.77	2.39	1.90
3	6.63	4.49	2.83	5.96	4.17	2.52	4.19	3.51	2.37
4	8.15	5.33	3.26	7.76	4.85	3.10	5.31	4.08	2.77

Standard errors were equal to 0.05 for GBLUP_H, 0.01 for both GBLUP_L and BLUP scenarios for genetic gain and 0.01 for inbreeding (in all scenarios).

was observed in the presence of common environmental effects (i.e. $c^2 > 0.0$) than in their absence (i.e. $c^2 = 0.0$) and, as in the case of genetic gain, BLUP scenarios shown greater differences between $c^2 = 0.0$ and $c^2 > 0.0$ than GBLUP. For instance, in BLUP scenarios with $n = 10$, inbreeding was up to 27% less at $t = 10$ for $c^2 = 0.3$ than for $c^2 = 0.0$. Equivalent figures for GBLUP_L and GBLUP_H were 20% and 12%, respectively. The decrease in cumulative inbreeding with $c^2 > 0.0$ was due to an increase in the number of families contributing to the next generation (not shown). For instance, in BLUP scenarios with $n = 10$, the number of families that contribute to generate $t = 1$ were 89 ($c^2 = 0.1$) and 108 ($c^2 = 0.3$) of a total of 200 families in the presence of common environmental effects and 79 in their absence. Corresponding figures for GBLUP_H were 100 ($c^2 = 0.1$), 107 ($c^2 = 0.3$) and 95.

3.2. Comparison of scenarios including or excluding common environmental effects in the evaluation model

As expected, excluding common environmental effects (when they existed) in the model led to a lower accuracy of breeding values in all scenarios (Fig. 1). For instance, for $n = 50$ and $c^2 = 0.3$, this loss in accuracy when ignoring the effects was 16% for both GBLUP_L and GBLUP_H and 19% for BLUP at $t = 4$.

Also, the exclusion of common environmental effects in the model implied an important increase in cumulative inbreeding (Fig. 3). For $n = 10$ and $c^2 = 0.3$, inbreeding increased at $t = 10$ by 60% for GBLUP_L, 44% for GBLUP_H and 70% for BLUP when compared to models including these effects. Equivalent figures for a lower c^2 (i.e. $c^2 = 0.1$) were smaller

(11% for GBLUP_L, 16% for GBLUP_H and 23% for BLUP). For $n = 10$, the number of families that contributed to generate $t = 1$ in BLUP scenarios were 61 ($c^2 = 0.1$) and 70 ($c^2 = 0.3$) when common environmental effects were not included in the model and 79 when these effects were absent. Corresponding figures for GBLUP_H were 87 ($c^2 = 0.1$), 93 ($c^2 = 0.3$) and 95 ($c^2 = 0.0$). The same trends were observed in scenarios with $n = 50$.

4. Discussion

This study shows that benefits obtained from GBLUP when compared to BLUP in terms of genetic gain can be higher in the presence of a common environmental effect than in their absence. This extra benefit in the presence of common environmental effects from GBLUP increased with increasing family size and LD. However, there were no extra benefits in terms of inbreeding in the presence of common environmental effects.

The higher accuracy and genetic gain obtained with genomic evaluations have been previously shown in aquaculture schemes (Nielsen et al., 2009, 2011; Sonesson and Meuwissen, 2009; Villanueva et al., 2011; Sonesson and Ødegård, 2016; Houston et al., 2020). The accuracies of breeding values obtained in our study in the absence of common environmental effects for GBLUP and BLUP were in the range of those found in previous simulations (Meuwissen et al., 2001; Nielsen et al., 2011) and empirical studies (Tsai et al., 2015; Palaiokostas et al., 2018b; Nguyen et al., 2020). However, to our knowledge this is the first study to explore the benefit of GBLUP from BLUP in the presence of

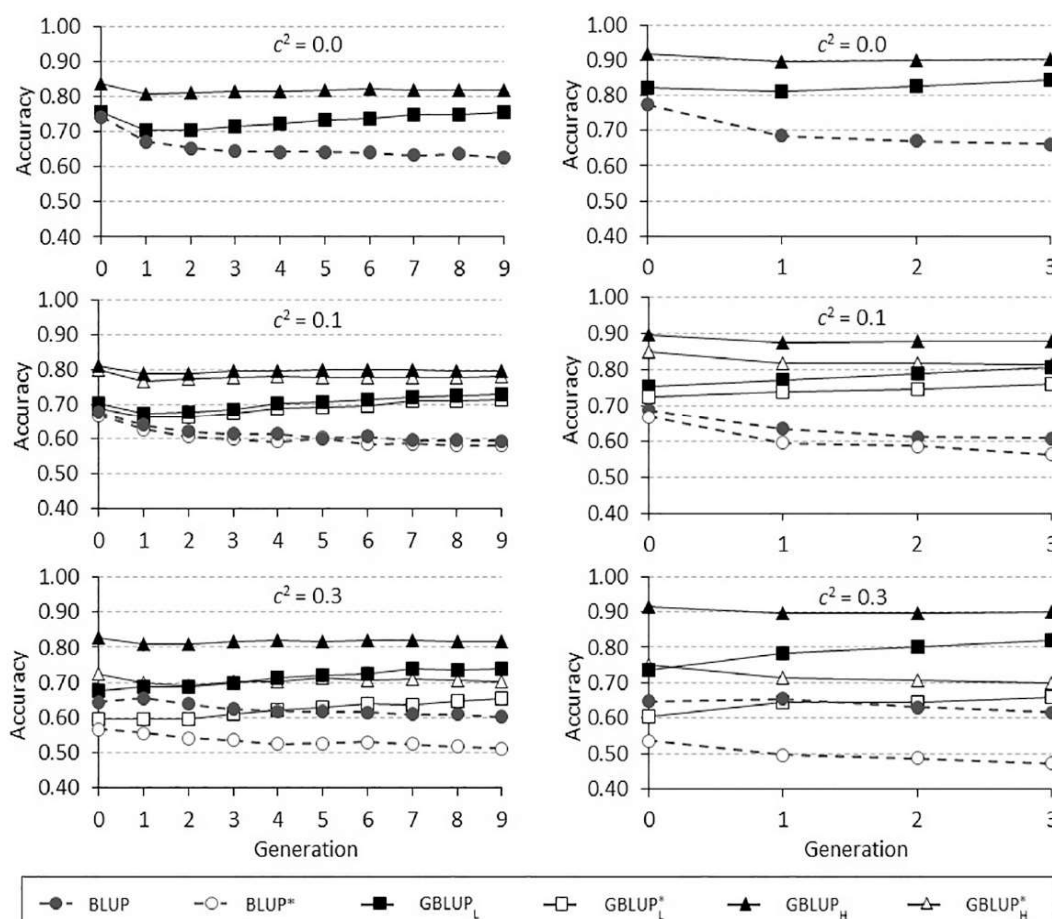


Fig. 1. Accuracy of breeding values across generations for BLUP and GBLUP scenarios with different proportions of the phenotypic variance due to common environmental effects (c^2) and different family sizes ($n = 10$ in left panels and $n = 50$ in right panels). For $c^2 > 0.0$, accuracies for scenarios that included or excluded (superscript *) the common environmental effects in the model are shown. GBLUP scenarios with low (GBLUP_L) and high (GBLUP_H) linkage disequilibrium were considered. Standard errors ranged from 3×10^{-4} to 1.5×10^{-3} .

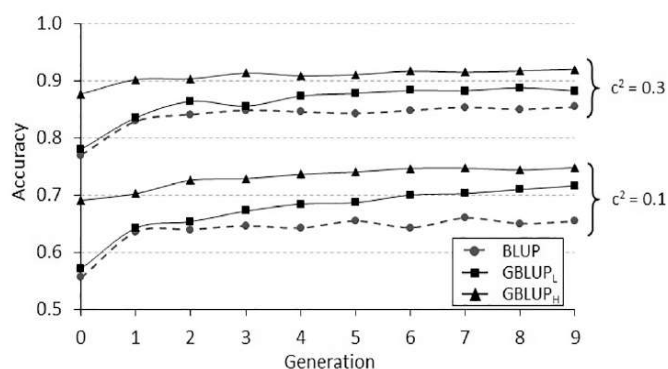


Fig. 2. Accuracy of common environmental effects across generations for BLUP and GBLUP scenarios with different proportions of the phenotypic variance due to common environmental effects (c^2) and a family size of 10 fish. GBLUP scenarios with low (GBLUP_L) and high (GBLUP_H) linkage disequilibrium were considered. Standard errors ranged from 2×10^{-4} to 2.8×10^{-3} .

common environmental effects and we have shown that under this scenario the benefit of GBLUP in terms of genetic gain increases even when LD between QTLs and SNP markers is low (Tables 1 and 2). In general, the presence of a common environmental effect led to decreases in the accuracy of breeding values, genetic gain and inbreeding in both BLUP and GBLUP scenarios because in nested designs it is difficult to separate common environmental from genetic effects. However, these decreases were higher in BLUP scenarios even with the lowest c^2 . GBLUP captures better the genetic relationships between fish raised in different

tanks and between fish raised in the same tank and thus is able to separate better common environmental from genetic effects leading to extra benefits.

Daetwyler et al. (2007) shown that GBLUP leads not only to higher gains but also to lower rates of inbreeding per generation than BLUP, at least when the inbreeding is calculated from pedigree data. This is because genomic information increases differentiation between sibs and reduces their co-selection. In our study, as expected, GBLUP led to lower inbreeding than BLUP, and GBLUP_H to lower inbreeding than GBLUP_L both in the absence and in the presence of common environmental effects. Also, lower inbreeding was observed for $c^2 > 0.0$ than for $c^2 = 0.0$ in both BLUP and GBLUP. In the presence of common environmental effects more families are selected to contribute to the next generation, (up to 37% more families in the case of BLUP) leading to a reduction in inbreeding. Given the design assumed here where each family is raised in a different tank, the estimated common environmental values could include part of the family genetic value and this would decrease the differences between families leading to a higher number of families selected. It is worth noting that if the most efficient method for restricting the rate at which inbreeding increases (i.e., the Optimal Contributions method) is applied, higher extra genetic gains would be expected from GBLUP at the same rate of inbreeding. This method optimizes the contributions of candidates to the next generation to maximize genetic gain while restricting the rate of inbreeding to a specific value (Meuwissen, 1997; Grundy et al., 1998; Sonesson et al., 2012). Thus, if schemes are compared at the same rate of inbreeding, the extra benefit from GBLUP would be even higher given that without any restrictions on inbreeding, BLUP leads to a higher inbreeding.

In nested designs where families are raised in different tanks until

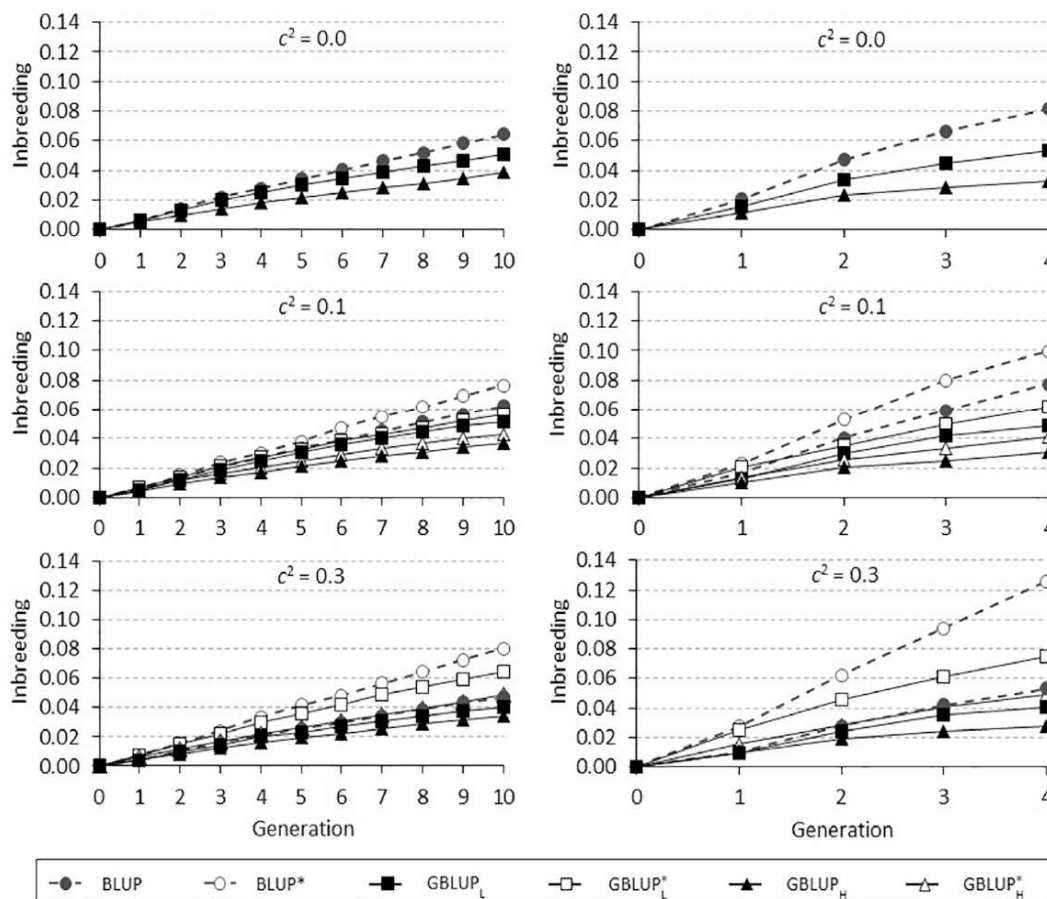


Fig. 3. Genealogical inbreeding across generations for BLUP and GBLUP scenarios with different proportions of the phenotypic variance due to common environmental effects (c^2) and different family sizes ($n = 10$ in left panels and $n = 50$ in right panels). For $c^2 > 0.0$, inbreeding for scenarios that included or excluded (superscript *) the common environmental effects in the model are shown. GBLUP scenarios with low (GBLUP_L) and high (GBLUP_H) linkage disequilibrium were considered. Standard errors ranged from 1×10^{-4} to 1.5×10^{-3} .

tagging and families are not replicated in several tanks, all members of the same family will share the same environment and, as previously mentioned, disentangling common environmental effects from family genetic effects will be difficult (Pante et al., 2002; Kause et al., 2005; Montaldo et al., 2012). Previous studies have shown the need of factorial designs (Kause et al., 2005) or tank replicates for each family (Montaldo et al., 2012) for obtaining high accuracies in BLUP evaluations. If the number of families (and selection intensity) wants to be maintained, replicating tanks per family would imply an increase in the number of tanks and this is not always feasible due to space and logistic restrictions. GBLUP is a good alternative. For instance, in this study GBLUP_H scenarios led to the same accuracy of breeding values in the absence and in the presence of common environmental effects when these were important ($c^2 = 0.3$).

In this study, variance components were assumed to be known without error, and still genetic gain was reduced (up to 7% in BLUP scenarios and up to 3% in GBLUP scenarios) when common environmental effects were present. In practice, variance components will need to be estimated and depending on the data structure estimates may be biased, leading to further reductions in genetic gain. Given this and the fact that the estimated common environmental effects are usually of small magnitude (≤ 0.1), several studies ignored them in the final models of analyses (Ødegård et al., 2014; Khaw et al., 2016). Here, we have shown that excluding these effects in the model leads to a moderate reduction in genetic gain. However, the level of inbreeding was greatly increased especially in BLUP scenarios as less families are selected.

The values considered here for the proportion of the phenotypic variance due to common environmental effects were based on those given in the literature. The lowest value ($c^2 = 0.1$) has been found in

growth traits recorded close to the slaughter time (Pante et al., 2002; Gjerde et al., 2004; Kause et al., 2007) whereas the highest value ($c^2 = 0.3$) has been found in growth traits recorded early in life (Winkelman and Peterson, 1994; Kause et al., 2006). As expected, higher accuracies of common environmental effects were obtained with the highest c^2 (Fig. 2) but the difference between $c^2 = 0.1$ and $c^2 = 0.3$ scenarios in terms of genetic gains was very small. Thus, although in most practical situations c^2 would be low (in most programs the breeding goal would be growth at a time close to slaughter), still there will be an extra benefit of using GBLUP when compared with BLUP.

The simulated trait was a candidate trait (i.e. selection candidates are recorded) with a moderate heritability (0.4), as most studies found in the literature dealing with common environmental effects focus on growth traits. However, at least in the absence of a common environmental effects, most benefit from GBLUP is expected for traits with low heritability and for those that cannot be recorded on the candidates (Sonesson and Meuwissen, 2009; Nielsen et al., 2011; Palaiokostas et al., 2016, 2018a; Sonesson and Ødegård, 2016; Vallejo et al., 2017; Saura et al., 2019; Yoshida et al., 2019). Although few, there are also studies showing the existence of a small but significant proportion of the phenotypic variance due to the common environmental effect for traits with low heritability and recorded on sibs of candidates (Kjøglum et al., 2005, 2008). In our study, the benefit of GBLUP from BLUP was mainly due to the fact that the accuracies of estimated breeding values from BLUP were much more affected when common environmental effects were present (even when they were of low magnitude, $c^2 = 0.1$) than the accuracies from GBLUP. We would expect that for traits with low heritabilities and/or for sib traits, the accuracies from BLUP would be further affected and the benefit from GBLUP could be even higher. It

should be noticed that in practice, aquaculture selective breeding programs focus on candidate traits (i.e. growth traits) and that the benefit of GBLUP when including also a sib trait into the breeding objective (and the selection index) will depend also on the genetic correlations between candidate and sib traits and on their economic values.

5. Conclusions

The benefit of genomic evaluations when compared with traditional BLUP in terms of genetic gain is higher in the presence of common environmental effects than in their absence for schemes applying nested mating designs. Contrarily to what happens with BLUP, the genetic gain from genomic evaluations was only slightly affected by the proportion of the phenotypic variance due to the common environmental effect. The extra benefit of using genomic information was especially clear in schemes with large family size and high LD. However, there were no extra benefits in terms of inbreeding in the presence of common environmental effects. Ignoring the common environmental effect in the evaluation model led to a moderate decrease in genetic gain but to a high increase in inbreeding.

Funding

The research leading to these results has received funding from the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (FPI-SGIT2016-02), the European Union's Seventh Framework Program (KBBE.2013.1.2-10) under grant agreement n° 613611, the Ministerio de Ciencia e Innovación, Spain (grant CGL2016-75904-C2-2-P and PID2020-114426GB-C22) and from the European Commission Horizon 2020 (H2020) Framework Programme through grant agreement no 727315 MedAID project (Mediterranean Aquaculture Integrated Development).

Authors' contributions

Conceptualization SG-B, BV, JFM and MAT; Methodology SG-B, BV, JFM and MAT; Software SG-B and JFM; Writing - original draft SG-B and BV; Writing - review & editing SG-B, BV, JFM and MAT.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Ignazy Misztal for computational help.

References

- Agha, S., Mekki, W., Ibanez-Escribano, N., Lind, C.E., Kumar, J., Mandal, A., et al., 2018. Breeding for robustness: investigating the genotype-by-environment interaction and micro-environmental sensitivity of genetically improved farmed Tilapia (*Oreochromis niloticus*). *Anim. Genet.* 49, 421–427. <https://doi.org/10.1111/age.12680>.
- Barria, A., López, M.E., Yoshida, G., Carvalheiro, R., Lhorente, J.P., Yáñez, J.M., 2018. Population genomic structure and genome-wide linkage disequilibrium in farmed Atlantic salmon (*Salmo salar* L.) using dense SNP genotypes. *Front. Genet.* 9, 649. <https://doi.org/10.3389/fgene.2018.00649>.
- Barria, A., Christensen, K.A., Yoshida, G., Jedlicki, A., Leong, J.S., Rondeau, E.B., et al., 2019. Whole genome linkage disequilibrium and effective population size in a coho salmon (*Oncorhynchus kisutch*) breeding population using a high-density snp array. *Front. Genet.* 10, 498. <https://doi.org/10.3389/fgene.2019.00498>.
- Borrell, Y.J., Gallego, V., García-Fernández, C., Mazzeo, I., Pérez, L., Asturiano, J.F., et al., 2011. Assessment of parental contributions to fast- and slow-growing progenies in the sea bream *Sparus aurata* L. using a new multiplex PCR. *Aquaculture* 314, 58–65. <https://doi.org/10.1016/j.aquaculture.2011.01.028>.
- Chatzpiplis, D., Batargias, C., Tsigonopoulos, C.S., Magoulas, A., Kollias, S., Kotoulas, G., et al., 2007. Mapping quantitative trait loci in European sea bass (*Dicentrarchus labrax*): the BASSMAP pilot study. *Aquaculture* 272, 172–182. <https://doi.org/10.1016/j.aquaculture.2007.08.022>.
- Daetwyler, H.D., Villanueva, B., Bijma, P., Woolliams, J.A., 2007. Inbreeding in genome-wide selection. *J. Anim. Breed. Genet.* 124, 369–376. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>.
- D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., et al., 2019. Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. *Genet. Sel. Evol.* 51, 26. <https://doi.org/10.1186/s12711-019-0468-4>.
- Fernandes, T., Herlin, M., López Belluga, M.D., Ballón, G., Martínez, P., Toro, M.Á., Fernández, J., 2017. Estimation of genetic parameters for growth traits in a hatchery population of gilthead sea bream (*Sparus aurata* L.). *Aquac. Int.* 25, 499–514. <https://doi.org/10.1007/s10499-016-0046-5>.
- Gallardo, J.A., Lhorente, J.P., Neira, R., 2010. The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*). *Genet. Sel. Evol.* 42, 1–8. <https://doi.org/10.1186/1297-9686-42-19>.
- Gjedrem, T., 2010. The first family-based breeding program in aquaculture. *Rev. Aquac.* 2, 2–15. <https://doi.org/10.1111/j.1753-5131.2010.01011.x>.
- Gjedrem, T., Baranski, M., 2010. Selective Breeding in Aquaculture: An Introduction. Springer Science & Business Media, Springer Netherlands, Dordrecht. <https://doi.org/10.1007/978-90-481-2773-3>.
- Gjerde, B., Gjøen, H.M., Villanueva, B., 1996. Optimum designs for fish breeding programmes with constrained inbreeding mass selection for a normally distributed trait. *Livest. Prod. Sci.* 47, 59–72. [https://doi.org/10.1016/S0301-6226\(96\)01000-7](https://doi.org/10.1016/S0301-6226(96)01000-7).
- Gjerde, B., Terjesen, B.F., Barr, Y., Lein, I., Thorland, I., 2004. Genetic variation for juvenile growth and survival in Atlantic cod (*Gadus morhua*). *Aquaculture* 236, 167–177. <https://doi.org/10.1016/j.aquaculture.2004.03.004>.
- Glover, K.A., Solberg, M.F., McGinnity, P., Hindar, K., Verspoor, E., Coulson, M.W., et al., 2017. Half a century of genetic interaction between farmed and wild Atlantic salmon: status of knowledge and unanswered questions. *Fish Fish.* 18, 890–927. <https://doi.org/10.1111/faf.12214>.
- Grundy, B., Villanueva, B., Woolliams, J.A., 1998. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genet. Res.* 72, 159–168. <https://doi.org/10.1017/S0016672398003474>.
- Habier, D., Fernando, R.L., Dekkers, J.C.M., 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177, 2389–2397. <https://doi.org/10.1534/genetics.107.081190>.
- Habier, D., Fernando, R.L., Garrick, D.J., 2013. Genomic BLUP decoded: a look into the black box of genomic prediction. *Genetics* 194, 597–607. <https://doi.org/10.1534/genetics.113.152207>.
- Hayes, B.J., Visscher, P.M., Goddard, M.E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91, 47–60. <https://doi.org/10.1017/S0016672308009981>.
- Herlin, M., Delghandi, M., Wesmajervi, M., Taggart, J.B., McAndrew, B.J., Penman, D.J., 2008. Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. *Aquaculture* 274, 218–224. <https://doi.org/10.1016/j.aquaculture.2007.11.034>.
- Hill, W.G., Robertson, A., 1968. The effects of inbreeding at loci with heterozygote advantage. *Genetics* 60, 615–628.
- Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., et al., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. *Nat. Rev. Genet.* 21, 389–409. <https://doi.org/10.1038/s41576-020-0227-y>.
- Kause, A., Ritola, O., Paananen, T., Wahlroos, H., Mäntysaari, E.A., 2005. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 247, 177–187. <https://doi.org/10.1016/j.aquaculture.2005.02.023>.
- Kause, A., Tobin, D., Houlihan, D.F., Martin, S.A.M., Mäntysaari, E.A., Ritola, O., Ruohonen, K., 2006. Feed efficiency of rainbow trout can be improved through selection: different genetic potential on alternative diets. *J. Anim. Sci.* 84, 807–817. <https://doi.org/10.2527/2006.844807x>.
- Kause, A., Paananen, T., Ritola, O., Koskinen, H., 2007. Direct and indirect selection of visceral lipid weight, fillet weight, and fillet percentage in a rainbow trout breeding program. *J. Anim. Sci.* 85, 3218–3227. <https://doi.org/10.2527/jas.2007-0332>.
- Khaw, H.L., Ponzoni, R.W., Yee, H.Y., bin Aziz, M.A., Bijma, P., 2016. Genetic and non-genetic indirect effects for harvest weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 450, 154–161. <https://doi.org/10.1016/j.aquaculture.2015.07.033>.
- Kijas, J., Elliot, N., Kube, P., Evans, B., Botwright, N., King, H., et al., 2016. Diversity and linkage disequilibrium in farmed Tasmanian Atlantic salmon. *Anim. Genet.* 48, 237–241. <https://doi.org/10.1111/age.12513>.
- Kjøglum, S., Grimholt, U., Larsen, S., 2005. Non-MHC genetic and tank effects influence disease challenge tests in Atlantic salmon (*Salmo salar*). *Aquaculture* 250, 102–109. <https://doi.org/10.1016/j.aquaculture.2005.03.021>.
- Kjøglum, S., Henryon, M., Aasmundstad, T., Korsgaard, I., 2008. Selective breeding can increase resistance of Atlantic salmon to furunculosis, infectious salmon anaemia and infectious pancreatic necrosis. *Aquac. Res.* 39, 498–505. <https://doi.org/10.1111/j.1365-2109.2008.01904.x>.
- Lillehammer, M., Meuwissen, T.H.E., Sonesson, A.K., 2013. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. *Genet. Sel. Evol.* 45, 1–8. <https://doi.org/10.1186/1297-9686-45-39>.
- Meuwissen, T.H.E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.* 75, 934–940. <https://doi.org/10.2527/1997.754934x>.

- Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. <https://doi.org/10.1093/genet/157.6.1819>.
- Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., Vitezica, Z., 2015. Manual for BLUPF90 Family of Programs. University of Georgia, Athens. Available online at: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf (Accessed March 1, 2021).
- Montaldo, H.H., Castillo-Juárez, H., Campos-Montes, G., Pérez-Enciso, M., 2012. Effect of the data family structure, tank replication and the statistical model, on the estimation of genetic parameters for body weight at 28 days of age in the Pacific white shrimp (*Penaeus (Litopenaeus) vannamei* Boone, 1931). *Aquac. Res.* 44, 1–9. <https://doi.org/10.1111/j.1365-2109.2012.03176.x>.
- Nejati-Javaremi, A., Smith, C., Gibson, J.P., 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. *J. Anim. Sci.* 75, 1738–1745. <https://doi.org/10.2527/1997.7571738x>.
- Nguyen, N.H., Phuthaworn, C., Knibb, W., 2020. Genomic prediction for disease resistance to Hepatopancreatic parvovirus and growth, carcass and quality traits in Banana shrimp *Fenneropenaeus merguensis*. *Genomics* 112, 2021–2027. <https://doi.org/10.1016/j.ygeno.2019.11.014>.
- Nielsen, H.M., Sonesson, A.K., Yazdi, H., Meuwissen, T.H.E., 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. *Aquaculture* 289, 259–264. <https://doi.org/10.1016/j.aquaculture.2009.01.027>.
- Nielsen, H.M., Sonesson, A.K., Meuwissen, T.H.E., 2011. Optimum contribution selection using traditional best linear unbiased prediction and genomic breeding values in aquaculture breeding schemes. *J. Anim. Sci.* 89, 630–638. <https://doi.org/10.2527/jas.2009-2731>.
- Ødegård, J., Moen, T., Santi, N., Korsvoll, S.A., Kjøglum, S., Meuwissen, T.H.E., 2014. Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*). *Front. Genet.* 5, 402. <https://doi.org/10.3389/fgene.2014.00402>.
- Palaiokostas, C., Ferrareso, S., Franch, R., Houston, R.D., Bargelloni, L., 2016. Genomic prediction of resistance to pasteurellosis in gilthead sea bream (*Sparus aurata*) using 2B-RAD sequencing. *G3 Genes Genomes Genet.* 6, 3693–3700. <https://doi.org/10.1534/g3.116.035220>.
- Palaiokostas, C., Cariou, S., Bestin, A., Bruant, J.S., Haffray, P., Morin, T., et al., 2018a. Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet. Sel. Evol.* 50, 1–11. <https://doi.org/10.1186/s12711-018-0401-2>.
- Palaiokostas, C., Kocour, M., Prchal, M., Houston, R.D., 2018b. Accuracy of genomic evaluations of juvenile growth rate in common carp (*Cyprinus carpio*) using genotyping by sequencing. *Front. Genet.* 9, 82. <https://doi.org/10.3389/fgene.2018.00082>.
- Pante, M.J.R., Gjerde, B., McMillan, I., Misztal, I., 2002. Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 204, 383–392. [https://doi.org/10.1016/S0044-8486\(01\)00825-0](https://doi.org/10.1016/S0044-8486(01)00825-0).
- Rye, M., Mao, L.L., 1998. Nonadditive genetic effects and inbreeding depression for body weight in Atlantic salmon (*Salmo salar* L.). *Livest. Prod. Sci.* 57, 15–22. [https://doi.org/10.1016/S0301-6226\(98\)00165-1](https://doi.org/10.1016/S0301-6226(98)00165-1).
- Saura, M., Carabaño, M.J., Fernández, A., Cabaleiro, S., Doeschl-Wilson, A.B., Anacleto, O., et al., 2019. Disentangling genetic variation for resistance and endurance to scuticociliatosis in turbot using pedigree and genomic information. *Front. Genet.* 10, 539. <https://doi.org/10.3389/fgene.2019.00539>.
- Sonesson, A.K., Meuwissen, T.H.E., 2009. Testing strategies for genomic selection in aquaculture breeding programs. *Genet. Sel. Evol.* 41, 1–9. <https://doi.org/10.1186/1297-9686-41-37>.
- Sonesson, A.K., Ødegård, J., 2016. Mating structures for genomic selection breeding programs in aquaculture. *Genet. Sel. Evol.* 48, 1–7. <https://doi.org/10.1186/s12711-016-0224-y>.
- Sonesson, A.K., Woolliams, J.A., Meuwissen, T.H.E., 2012. Genomic selection requires genomic control of inbreeding. *Genet. Sel. Evol.* 44, 1–10. <https://doi.org/10.1186/1297-9686-44-27>.
- Tosh, J.J., Garber, A.F., Trippel, E.A., Robinson, J.A.B., 2010. Genetic, maternal, and environmental variance components for body weight and length of Atlantic cod at 2 points in life. *J. Anim. Sci.* 88, 3513–3521. <https://doi.org/10.2527/jas.2009-2676>.
- Tsai, H.Y., Hamilton, A., Tinch, A.E., Guy, D.R., Gharbi, K., Stear, M.J., et al., 2015. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high-density SNP array. *BMC Genomics* 16, 1–9. <https://doi.org/10.1186/s12864-015-2117-9>.
- Vallejo, R.L., Leeds, T.D., Gao, G., Parsons, J.E., Martin, K.E., Evenhuis, J.P., et al., 2017. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genet. Sel. Evol.* 49, 1–13. <https://doi.org/10.1186/s12711-017-0293-6>.
- VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- Villanueva, B., Pong-Wong, R., Fernández, J., Toro, M.A., 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. *J. Anim. Sci.* 83, 1747–1752. <https://doi.org/10.2527/2005.8381747x>.
- Villanueva, B., Fernández, J., García-Cortés, L.A., Varona, L., Daetwyler, H.D., Toro, M.A., 2011. Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs. *J. Anim. Sci.* 89, 3433–3442. <https://doi.org/10.2527/jas.2010-3814>.
- Winkelman, A.M., Peterson, R.G., 1994. Heritabilities, dominance variation, common environmental effects and genotype by environment interactions for weight and length in Chinook salmon. *Aquaculture* 125, 17–30. [https://doi.org/10.1016/0044-8486\(94\)90278-X](https://doi.org/10.1016/0044-8486(94)90278-X).
- Yoshida, G.M., Lhorente, J.P., Correa, K., Soto, J., Salas, D., Yáñez, J.M., 2019. Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in Nile Tilapia (*Oreochromis niloticus*). *G3 Genes Genomes Genet.* 9, 2597–2607. <https://doi.org/10.1534/g3.119.4001164>.

CHAPTER 2





García-Ballesteros, S., Fernández, J., Kause, A., and Villanueva, B. (2022). **Predicted genetic gain for carcass yield in rainbow trout from indirect and genomic selection.** *Aquaculture*, 554, 1-8. Supplementary material in Annex II.
<https://doi.org/10.1016/j.aquaculture.2022.738119>

2021 Journal impact factor 5.135, JCR Rank 5/54 in the *Fisheries* category and 8/113 in the *Marine & Freshwater Biology* category.



Predicted genetic gain for carcass yield in rainbow trout from indirect and genomic selection

Silvia García-Ballesteros^{a,*}, Jesús Fernández^a, Antti Kause^b, Beatriz Villanueva^a

^a Departamento de Mejora Genética Animal, INIA (CSIC), 28040 Madrid, Spain

^b Natural Resources Institute Finland (Luke), Biometrical Genetics, FI-31600 Jokioinen, Finland

ARTICLE INFO

Keywords:

Carcass yield
Fillet yield
Sib selection
Indirect selection
Genomic selection

ABSTRACT

Carcass and fillet yields are traits of great economic importance in aquaculture species, including rainbow trout. Headless gutted carcass percentage (HC) is a convenient selection criterion to improve carcass yield given that it is highly genetically correlated with the latter and also with fillet yield. However, HC is a sib trait that cannot be recorded on selection candidates. Consequently, the within-family component of the genetic variance cannot be exploited with traditional pedigree-based BLUP selection. Two alternatives to exploit this component would be to select directly on an indicator trait genetically correlated with HC that can be recorded on live candidates or to apply genomic selection. The objective of this simulation study was to predict the phenotypic gains for HC in rainbow trout breeding programs when four alternative selection strategies are used: i) sib selection for HC; ii) indirect selection for a morphological indicator recorded in vivo; iii) genomic selection for HC; and iv) genomic selection for the indicator. Also, the four strategies were compared in a multitrait selection scenario where body weight was also included in the breeding objective. The different scenarios were compared at the same selection intensity and number of records (2000) for HC (on sibs) and IHC (on candidates). Two different heritabilities for HC (0.55 and 0.25) were considered. For the highest heritability, the phenotypic gain for HC was higher with sib than with indirect selection for both BLUP (1.03 versus 0.98) and genomic selection (1.22 versus 1.04). However, for the lowest heritability, the phenotypic gain for HC was lower with sib than with indirect selection for both BLUP (0.60 versus 0.64) and genomic selection (0.70 versus 0.71). In any case, the differences in phenotypic gains for HC between sib and indirect selection were not large. Therefore, given that sib selection implies extra costs associated with maintaining and genotyping sibs and indirect selection allows higher selection intensity for a given number of fish, indirect selection appears to be a more cost-effective option. The patterns found for single trait selection were maintained when the yield trait (HC or indicator) was selected simultaneously with body weight. We conclude that the optimum approach to improve carcass yield is a combination of genomic and indirect selection in both single and multitrait selection scenarios.

1. Introduction

One of the main breeding goal traits in aquaculture selection programs is growth rate, a trait that can be easily recorded and improved (Gjedrem et al., 2012). However, other economically important traits cannot be recorded on live animals and they are challenging to be incorporated in these programs. This is the case of carcass yield (i.e., the percentage of carcass weight out of wet body weight) and fillet yield (i.e., the percentage of fillet weight out of wet body weight), two economically very important traits, especially for species sold processed as gutted carcass or fillets (Kankainen et al., 2016; Fraslin et al., 2018;

Prchal et al., 2018).

Direct selection for fillet yield is not an easy task as fillet recording is laborious and is associated to large measurement errors, which lead to a relatively low heritability (Rutten et al., 2005; Powell et al., 2008; Haffray et al., 2012). Carcass yield is easier to record and several studies in rainbow trout have demonstrated that selection for this trait would result in a correlated response in fillet yield (Kause et al., 2007; Haffray et al., 2012; Vandeputte et al., 2019).

A more suitable selection criterion to improve carcass and fillet yield is headless gutted carcass percentage (HC). The heritability of HC is similar to that of carcass yield and higher than that of fillet yield, and the

* Corresponding author.

E-mail address: silvia.garciab@inia.es (S. García-Ballesteros).

genetic correlation between HC and fillet yield is higher than the correlation between carcass yield and fillet yield (Haffray et al., 2012, 2013; Vandeputte et al., 2017, 2019; Prchal et al., 2018). In particular, in rainbow trout, Haffray et al. (2012, 2013) showed that selection for HC could be more efficient than selection for fillet yield itself.

However, recording HC also requires sacrificing the fish, and thus, it is impossible to do it on selection candidates. For this kind of traits, selection is generally based on sibs' performance and consequently, only the between family component of the genetic variance can be exploited when applying traditional pedigree-based BLUP selection (Nielsen et al., 2009; Sonesson and Meuwissen, 2009). An alternative to improve HC would be to select for an indicator trait genetically correlated with HC that can be recorded on live fish (i.e., on the candidates themselves). In fact, HC predictors based on morphological records that permit non-lethal trait recording have been recently developed in different aquaculture species (Haffray et al., 2013; Vandeputte et al., 2017; Prchal et al., 2018, 2021).

Another alternative that allows exploiting the within-family component in the selection decisions would be to apply genomic selection (Meuwissen et al., 2001). Computer simulations (Nielsen et al., 2009, 2011; Sonesson and Meuwissen, 2009; Villanueva et al., 2011; Lillehammer et al., 2013; Sonesson and Ødegård, 2016) and analyses of empirical data of different aquaculture species (see reviews by Zenger et al., 2019; Houston et al., 2020 and D'Agaro et al., 2021) have shown important benefits of this approach in terms of genetic gain compared to traditional BLUP selection. These benefits are expected when selecting directly or indirectly for HC.

The objective of this study was to predict, through computer simulations, the potential of indirect and genomic selection to improve carcass yield in selective breeding programs for rainbow trout. Different selection strategies were evaluated both in a single trait setting, where selection was only for a yield trait and in a multi-trait setting, where body weight was also selected.

2. Material and methods

The simulated selection program started from a base population with a family structure (i.e., composed by full- and half-sib families). Three correlated traits were simulated in all scenarios, including body weight (BW, in g), percentage of headless carcass (HC), and a morphological indicator of HC measured by ultrasound tomography (IHC). Specifically, the indicator simulated was the ratio of abdominal wall thickness (Echo8) to depth of the peritoneal cavity (Echo23) as described in Haffray et al. (2013) and Vandeputte et al. (2019). Genetic parameters assumed for the three traits were taken from the literature (Kause et al., 2007; Haffray et al., 2012, 2013) and are shown in Table 1. These parameters correspond to traits recorded at ~510 days post fertilization (Haffray et al., 2013). The estimate of the heritability for HC (h_{HC}^2) found in previous studies was 0.55, but an alternative value (0.25) was also simulated to investigate the scenario where HC and IHC have similar heritabilities. Phenotypic gains for HC obtained from four different selection strategies were compared. These strategies included i) sib selection for HC based on sibs' phenotypes; ii) indirect selection for IHC based on candidates' phenotypes; iii) genomic selection for HC based on sibs' phenotypes; and iv) genomic selection for IHC based on

candidates' phenotypes. The four strategies were compared at the same selection intensity. A total of 100 males and 200 females were selected and mated following a nested design to generate 200 full-sib families, and each mating produced 20 offspring (10 males and 10 females). The simulations were carried out with our own Fortran 90 codes.

2.1. Genome structure

The genome simulated was composed of 30 chromosomes and had a total size of 20 Morgans, mimicking the trout genome. A total of 3,600,000 evenly spaced biallelic loci were simulated throughout the genome (120,000 loci per chromosome), of which 600,000 (20,000 per chromosome) were considered to be potential SNP (single nucleotide polymorphism) markers to be used in genomic selection. SNP markers were also evenly spaced within each chromosome.

2.2. Generation of the base population

Firstly, a population in mutation-drift equilibrium was simulated by randomly generating 100 males and 100 females across 4000 discrete generations. At each generation, sires and dams were randomly sampled with replacement. Population size was kept constant across generations. The mutation rate per locus and generation was 2.5×10^{-4} . Mutations were randomly distributed across individuals, chromosomes, and loci, switching allele 0 to allele 1 and vice versa. Initial frequencies were 0.5 for all loci. When generating the gametes, the total number of crossovers was drawn from a Poisson distribution with mean equal to the genome length. Crossovers were randomly distributed without interference. At the end of this process, the expected heterozygosity of the population had already reached an asymptote (i.e., mutation-drift equilibrium). At this point, the population was expanded over one generation in order to have enough individuals to sample replicates. One male and one female were randomly sampled with replacement and mated to produce 1 offspring and this process was repeated until a total of 1000 individuals were generated. These 1000 individuals constituted the expanded population from which replicates for the different scenarios were generated.

In order to create a base population ($t = 0$) with a family structure, 100 males and 200 females were randomly sampled from the expanded population for each replicate. Each male was randomly mated with two females following a nested design and 20 offspring per mating were generated. Thus, $t = 0$ was composed of 4000 fish belonging to 200 full-sib families (and 100 half-sib families). At $t = 0$, loci with a minor allele frequency less than 0.1 were discarded as QTLs (loci affecting the traits) or SNPs (loci to be used in genomic selection). Across the genome, a total of 48,000 loci (an average of 1600 loci per chromosome) were then randomly sampled as SNP markers (i.e., a marker density of 2400 SNPs/Morgan). Also, from the total number of loci that remained segregating (1,005,000 loci, including the 48,000 SNPs), a total of 1000 loci were randomly sampled as QTLs that simultaneously affected the three polygenic traits (BW, HC, and IHC). In order to simulate specific pairwise genetic correlations between traits, each QTL had different effects on the three traits (see the specific approach taken below, in section 2.3). At $t = 0$, the average linkage disequilibrium measured as r^2 (the squared correlation between pairs of loci; Hill and Robertson, 1968) between SNPs separated 50 kb and 100 kb was 0.308 and 0.282, respectively. These r^2 values are similar to those found in the literature for rainbow trout (D'Ambrosio et al., 2019).

2.3. Generation of breeding and phenotypic values

The true breeding value of individual j for trait k (TBV_{kj} ; $k = BW, HC, IHC$) was obtained as the sum of the (additive) genetic values across QTLs. For QTL i , these values were a_{ik} , 0 and $-a_{ik}$ for homozygotes 11, heterozygotes 10 and homozygotes 00, respectively. For QTL i affecting trait k , a_{ik} was obtained in two steps (Fernández et al., 2021; García-Ballesteros et al., 2021). Firstly, a random number (a_{ik}^*) was sampled

Table 1

Heritabilities (on diagonal), and genetic (above diagonal) and phenotypic correlations (below diagonal) between the traits considered.

Trait	BW	HC	IHC
BW	0.30	0.15	0.30
HC	0.20	0.55 or 0.25	0.90
IHC	0.40	0.55	0.25

BW: body weight; HC: headless carcass percentage; IHC: indicator of headless carcass percentage.

from the following multivariate normal distribution:

$$N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{a_{BW},IHC} & \rho_{a_{BW},IHC} \\ \text{Symm} & 1 & \rho_{a_{HC},IHC} \\ & \rho_{a_{HC},IHC} & 1 \end{bmatrix} \right),$$

where $\rho_{a_{k,l}}$ is the genetic correlation between traits k and l ($k = BW, HC, IHC; l = BW, HC, IHC; k \neq l$). Secondly, a_{ik} was obtained by multiplying a_{ik}^* by the factor $\sqrt{\frac{\sigma_{a_k}^2}{[2p_k(1-p_k)n_{QTL}]}}$, where $\sigma_{a_k}^2$ is the assumed additive genetic variance for trait k at $t = 0$, p_k is the average frequency across QTLs for trait k at $t = 0$ and n_{QTL} is the number of QTLs (i.e., 1000). Note that in this way the expected additive variance summed over all QTLs equals $\sigma_{a_k}^2$, assuming that covariances between loci generated by linkage disequilibrium are negligible. The phenotypic value of individual j for trait k was obtained by adding an environmental effect to TBV_{kj} . This environmental effect was sampled from the following multivariate distribution:

$$N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e_{BW}}^2 & \rho_{e_{BW},IHC} \sqrt{\sigma_{e_{BW}}^2 \sigma_{e_{HC}}^2} & \rho_{e_{BW},IHC} \sqrt{\sigma_{e_{BW}}^2 \sigma_{e_{IHC}}^2} \\ \text{Symm} & \sigma_{e_{HC}}^2 & \rho_{e_{HC},IHC} \sqrt{\sigma_{e_{HC}}^2 \sigma_{e_{IHC}}^2} \\ & \rho_{e_{HC},IHC} \sqrt{\sigma_{e_{HC}}^2 \sigma_{e_{IHC}}^2} & \sigma_{e_{IHC}}^2 \end{bmatrix} \right),$$

where $\sigma_{e_k}^2$ is the environmental variance for trait k and $\rho_{e_{k,l}}$ is the environmental correlation between traits k and l . The environmental variance for trait k was obtained as $\sigma_{p_k}^2 - \sigma_{a_k}^2$, where $\sigma_{p_k}^2$ is the phenotypic variance for trait k that was set to 1 for the three traits). The environmental correlation was obtained as $\rho_{e_{k,l}} = (\rho_{p_{k,l}} - h_k h_l \rho_{a_{k,l}}) / e_k e_l$, where h_k, h_l are the square roots of the heritabilities for traits k and l , respectively, e_k is $1 - h_k^2$, e_l is $1 - h_l^2$ and $\rho_{p_{k,l}}$ is the phenotypic correlation between both traits (Falconer and Mackay, 1996).

2.4. Genetic evaluation methods

The evaluation methods considered for obtaining estimates of breeding values were standard BLUP and genomic evaluation. The univariate model assumed for BLUP and trait k was.

$$y_k = 1\mu_k + Z\mathbf{a}_k + \mathbf{e}_k,$$

where y_k is the vector of phenotypes, μ_k is the mean, $\mathbf{a}_k$ is the vector of random additive genetic effects, $\mathbf{e}_k$ is the vector of random residuals, $\mathbf{1}$ is a vector of ones and Z is the incidence matrix for genetic effects. The variance-covariance matrices of random effects were assumed to be $V(\mathbf{a}_k) = A\sigma_{a_k}^2$ and $V(\mathbf{e}_k) = I_r\sigma_{e_k}^2$, where A is the pedigree-based numerator relationship matrix and I_r is an identity matrix of order r (number of records). Both $\sigma_{a_k}^2$ and $\sigma_{e_k}^2$ were assumed to be known.

The genomic evaluations were performed using the GBLUP method (Nejati-Javaremi et al., 1997; Villanueva et al., 2005; Hayes et al., 2009). The assumed model was the same as in BLUP but here A was replaced with the genomic relationship matrix (G) computed following VanRaden's method 1 (VanRaden, 2008). Thus, $V(\mathbf{a}_k)$ was assumed to be $G\sigma_{a_k}^2$.

The BLUPf90 software (Miszta et al., 2015) was used for carrying out both BLUP and GBLUP evaluations.

2.5. Trait recording and genotyping

Recording HC implies slaughtering the fish and, thus, it cannot be done on selection candidates. Thus, HC is what is known as a 'sib trait' that was assumed to be recorded only in half of the fish per family (i.e., 10 fish per family). The other 10 fish per family were considered to be the selection candidates. Traits BW and IHC are 'candidate traits' that can be recorded on the selection candidates. To maintain the same number of records available for all traits, IHC and BW were assumed to

be recorded on the 10 selection candidates per family. Thus, in the scenarios where selection was for a single trait (i.e., either for HC, IHC, or BW) the total number of fish recorded for all traits was 2000.

In the scenarios where selection was simultaneous for two traits (BW and HC or BW and IHC), the total number of records for BW was 4000 (2000 candidates and 2000 sibs). This is because when BW and HC are jointly selected, it is necessary to record BW to obtain the HC phenotype and thus, BW records are available not only for the candidates but also for their sibs. This is not the case in the scenarios where BW and IHC are jointly selected but in order to make scenarios comparable, the total number of records for BW was also 4000.

In summary, the total number of records was 2000 for HC (sibs) and IHC (candidates), but the total number of records for BW was 2000 (candidates) in scenarios where selection was only for this trait and 4000 (2000 candidates and 2000 sibs) in scenarios where selection was jointly for BW and a yield trait (HC or IHC).

When performing genomic evaluations, fish were considered to be genotyped for the 48,000 SNPs (on average 1600 SNPs per chromosome). Thus, as indicated above, the marker density was 2400 SNPs/Morgan.

2.6. Selection scenarios

Selection was performed for a single generation, in which the 100 males and 200 females with the highest estimated breeding values were selected and randomly mated. Nested mating designs, where each male was randomly mated with two females, were applied. In all scenarios simulated the number of selection candidates per family was 10 (2000 in total) which implies proportions selected of 10% and 20% for males and females, respectively. Thus, the different scenarios were compared at the same selection intensity.

Different BLUP (scenarios B_X) or GBLUP (scenarios G_X) selection scenarios were considered:

- Selection for a single trait including direct selection for BW (B_W and G_W), sib selection for HC (B_H and G_H), or indirect selection for IHC (B_I and G_I). Scenarios B_H and G_H (sib selection for HC) were also run with an increased number of sibs tested per family (20, 40, 60, 80, and 100).
- Simultaneous selection for both BW and HC (B_{W+H} and G_{W+H}).
- Simultaneous selection for both BW and IHC (B_{W+I} and G_{W+I}).

In scenarios ii) and iii), the final estimated breeding value of an individual i was the sum of the estimated breeding values for each trait (BW and HC or IHC) multiplied by the corresponding relative weight given to each trait. The relative weight for BW was set at 1, and different weights were investigated for the yield trait (HC or IHC) (0.25, 0.50, 0.75, and 1).

Scenarios were compared in terms of the accuracy of evaluation (measured as the correlation between true and estimated breeding values) and phenotypic gain in trait units after one generation of selection. We assumed initial phenotypic means of 1639 g for BW, 76.6% for HC, and 72.4 units for IHC and phenotypic variances of 351 for BW, 2.25 for HC, and 1.30 for IHC (Kause et al., 2007; Haffray et al., 2012, 2013).

Each scenario was replicated 50 times and the results presented are averages over replicates.

3. Results

3.1. Selection for HC or IHC

With BLUP selection and $h_{HC}^2 = 0.55$ (Table 2), the highest response in HC was achieved with sib selection for this trait (1.03 with scenario B_H versus 0.98 with scenario B_I). However, for a lower heritability of HC ($h_{HC}^2 = 0.25$, Table 3) and also with BLUP selection, the highest increase

Table 2

Increase in the phenotypic means of body weight (BW), headless carcass percentage (HC) and the indicator of headless carcass percentage (IHC) after one generation of selection when giving different relative weights to the three traits and when using two different evaluation methods (BLUP, scenarios B_x, or GBLUP, scenarios G_x), for a heritability of HC of 0.55. ^{a, b}

Scenarios ^c	Selection for	Weights			BLUP			GBLUP		
		BW	HC	IHC	BW	HC	IHC	BW	HC	IHC
B _W and G _W	BW	1	0	0	187.50	0.17	0.17	204.21	0.19	0.18
B _H and G _H	HC	0	1	0	25.17	1.03	0.48	36.47	1.22	0.57
B _I and G _I	IHC	0	0	1	54.37	0.98	0.56	61.53	1.04	0.59
B _{W+H.1} and G _{W+H.1}	BW and HC	1	0.25	0	189.72	0.46	0.30	207.65	0.56	0.36
B _{W+H.2} and G _{W+H.2}	BW and HC	1	0.50	0	173.87	0.66	0.38	189.94	0.78	0.45
B _{W+H.3} and G _{W+H.3}	BW and HC	1	0.75	0	154.28	0.78	0.43	166.01	0.93	0.50
B _{W+H.4} and G _{W+H.4}	BW and HC	1	1	0	137.73	0.85	0.45	143.70	1.03	0.54
B _{W+L.1} and G _{W+L.1}	BW and IHC	1	0	0.25	188.84	0.38	0.28	210.69	0.40	0.30
B _{W+L.2} and G _{W+L.2}	BW and IHC	1	0	0.50	180.03	0.50	0.35	199.47	0.54	0.38
B _{W+L.3} and G _{W+L.3}	BW and IHC	1	0	0.75	169.74	0.59	0.39	187.71	0.63	0.43
B _{W+L.4} and G _{W+L.4}	BW and IHC	1	0	1	158.21	0.70	0.44	176.22	0.74	0.49

^a The initial phenotypic means for BW, HC, and IHC were 1639 g, 76.6%, and 72.4 units, respectively.

^b Standard errors were equal to 0.1 for BW scenarios, 0.03 for HC scenarios, and 0.01 for IHC scenarios.

^c Numbers 1 to 4 simply enumerate the scenarios with different weights for each trait to facilitate referring to them in the text.

Table 3

Increase in the phenotypic means of body weight (BW), headless carcass percentage (HC) and the indicator of headless carcass percentage (IHC) after one generation of selection when giving different relative weights to the three traits and when using two different evaluation methods (BLUP, scenarios B_x, or GBLUP, scenarios G_x), for a heritability of HC of 0.25. ^{a, b}

Scenarios ^c	Selection for	Weights			BLUP			GBLUP		
		BW	HC	IHC	BW	HC	IHC	BW	HC	IHC
B _W and G _W	BW	1	0	0	186.80	0.12	0.17	201.12	0.14	0.19
B _H and G _H	HC	0	1	0	24.89	0.60	0.41	29.52	0.70	0.48
B _I and G _I	IHC	0	0	1	54.88	0.64	0.54	58.25	0.71	0.59
B _{W+H.1} and G _{W+H.1}	BW and HC	1	0.25	0	194.77	0.22	0.24	213.83	0.26	0.28
B _{W+H.2} and G _{W+H.2}	BW and HC	1	0.50	0	186.61	0.31	0.30	205.25	0.36	0.34
B _{W+H.3} and G _{W+H.3}	BW and HC	1	0.75	0	177.69	0.37	0.33	194.59	0.43	0.39
B _{W+H.4} and G _{W+H.4}	BW and HC	1	1	0	165.37	0.43	0.36	182.43	0.49	0.41
B _{W+L.1} and G _{W+L.1}	BW and IHC	1	0	0.25	192.93	0.25	0.28	209.04	0.28	0.30
B _{W+L.2} and G _{W+L.2}	BW and IHC	1	0	0.50	183.75	0.35	0.35	202.61	0.37	0.37
B _{W+L.3} and G _{W+L.3}	BW and IHC	1	0	0.75	174.32	0.41	0.40	189.05	0.44	0.42
B _{W+L.4} and G _{W+L.4}	BW and IHC	1	0	1	163.99	0.46	0.43	176.25	0.50	0.47

^a The initial phenotypic means of BW, HC, and IHC were 1639 g, 76.6%, and 72.4 units, respectively.

^b Standard errors were equal to 0.1 for BW scenarios, 0.02 for HC scenarios, and 0.01 for IHC scenarios.

^c Numbers 1 to 4 simply enumerate the scenarios with different weights for each trait to facilitate referring to them in the text.

in HC was achieved with indirect selection (0.60 with scenario B_H versus 0.64 with scenario B_I).

The highest increase in the phenotypic mean of HC was obtained with genomic selection (Tables 2 and 3) both with sib (scenario G_H) and indirect (scenario G_I) selection. As with BLUP, the highest phenotypic response in HC was achieved with sib selection for the highest h_{HC}^2 and with indirect selection for the lowest h_{HC}^2 with genomic selection. The benefit (in terms of phenotypic means for HC) from genomic selection (compared with BLUP) in sib selection scenarios (scenarios G_H versus scenarios B_H) was 18% for $h_{HC}^2 = 0.55$ and 17% for $h_{HC}^2 = 0.25$. In indirect selection scenarios (scenarios G_I versus scenarios B_I), the benefit from genomic selection was lower (6% for $h_{HC}^2 = 0.55$ and 11% for $h_{HC}^2 = 0.25$). Note that this extra gain or benefit (in %) was computed as $100(\overline{\Delta P}_{B_I} - \overline{\Delta P}_{B_H})/\overline{\Delta P}_{B_H}$, where $\overline{\Delta P}_{B_I}$ and $\overline{\Delta P}_{B_H}$ are the increases in the phenotypic mean of HC in scenarios B_I and B_H, respectively.

The accuracy of evaluation for the different traits and scenarios is given in Tables S1 and S2 of supplementary material. In general, when the accuracy for IHC (scenarios B_I and G_I) was higher (lower) than the accuracy for HC (scenarios B_H and G_H), indirect selection led to higher (lower) phenotypic means for HC. For instance, for $h_{HC}^2 = 0.25$, with both BLUP and GBLUP the accuracy for IHC was higher than the

accuracy for HC and indirect selection led to a phenotypic mean of HC than over sib selection (Table 3). Also, for $h_{HC}^2 = 0.55$, with GBLUP the highest accuracy was for HC and this was reflected in the higher phenotypic mean of HC with sib selection (Table 2). However, with BLUP the accuracy for IHC (scenario B_I) was higher than the accuracy for HC (scenario B_H) but the increase in the phenotypic mean of HC was higher with sib selection than with indirect selection.

Table 4 shows the accuracy and the phenotypic mean of HC scenarios B_H and G_H (sib selection scenarios) when the number of sibs tested per family is increased from 10 to up to 100. For $h_{HC}^2 = 0.25$, the accuracy and phenotypic means increased more rapidly with increasing numbers of sibs compared to $h_{HC}^2 = 0.55$. The highest increase in the accuracy was obtained when doubling the number of sibs recorded from 10 to 20 for both h_{HC}^2 and both evaluation method. However, these high increases in accuracy did not translate into high increases in the HC mean. For $h_{HC}^2 = 0.55$, the increase in accuracy was 6% (8%) with BLUP (GBLUP) but the increase in the HC mean was only 0.05% (0.1%) with BLUP (GBLUP). For $h_{HC}^2 = 0.25$, the increase in the accuracy was 13% with both BLUP and GBLUP and this translated into an increase in the HC mean of less than 0.1%. With BLUP, there were no further increases in the accuracy and phenotypic mean when increasing the number of sibs beyond 20 and 40 for $h_{HC}^2 = 0.55$ and $h_{HC}^2 = 0.25$, respectively.

Table 4

Accuracy of evaluation and phenotypic mean of headless carcass percentage (HC) from sib selection for this trait for different number of sibs tested per family (n_{sibs}), evaluation methods (BLUP and GBLUP), and heritabilities for HC (h_{HC}^2).^a

	n_{sibs}	Accuracy		Phenotypic mean	
		BLUP	GBLUP	BLUP	GBLUP
$h_{\text{HC}}^2 = 0.55$	10	0.63	0.76	77.63	77.82
	20	0.67	0.82	77.68	77.92
	40	0.68	0.87	77.69	78.01
	60	0.69	0.90	77.70	78.06
	80	0.70	0.91	77.73	78.08
	100	0.70	0.92	77.73	78.08
$h_{\text{HC}}^2 = 0.25$	10	0.55	0.63	77.20	77.30
	20	0.62	0.71	77.27	77.39
	40	0.65	0.78	77.33	77.46
	60	0.67	0.82	77.36	77.50
	80	0.68	0.84	77.36	77.52
	100	0.68	0.86	77.36	77.54

^a Standard errors were equal to 0.02 for accuracies and 0.03 for phenotypic gains.

However, with GBLUP, the accuracy continued increasing beyond 40, particularly for $h_{\text{HC}}^2 = 0.25$.

Note that for $h_{\text{HC}}^2 = 0.25$, sib selection requires at least 4000 sib records (and a total of 6000 fish to be managed) to achieve a phenotypic mean for HC higher than that obtained with indirect selection using 2000 records (only 2000 fish to be managed) (see also Table 3).

3.2. Simultaneous selection for BW and a yield trait (HC or IHC)

As expected, the increase in the phenotypic mean of HC was lower when BW was also selected, particularly when the weight given to the latter was relatively high (Tables 2 and 3).

The patterns described in the previous section when BLUP single trait selection was applied (scenarios B_{H} and B_{I}) were also observed when BW was also selected (scenarios $B_{\text{W+H}}$ and $B_{\text{W+I}}$). Thus, the highest gain in HC was obtained from sib selection ($B_{\text{W+H}}$) for $h_{\text{HC}}^2 = 0.55$ (Table 2) and from indirect selection ($B_{\text{W+I}}$) for $h_{\text{HC}}^2 = 0.25$ (Table 3), for the whole range of relative weights considered. For $h_{\text{HC}}^2 = 0.55$, the difference between sib and indirect selection in terms of gain in HC was higher when BW was also selected (see gains in scenarios $B_{\text{W+H}}$ and $B_{\text{W+I}}$ for the whole range of relative weights in Table 2).

The benefit of genomic selection, in terms of extra gains in HC, described above for single trait selection scenarios was also observed when BW was also selected. Genomic selection led to increased gains in HC with both sib (scenarios $G_{\text{W+H}}$) and indirect (scenarios $G_{\text{W+I}}$) selection for any combination of relative weights. The highest response in HC was achieved again with genomic sib selection ($G_{\text{W+H}}$) for $h_{\text{HC}}^2 = 0.55$ and with indirect genomic selection ($G_{\text{W+I}}$) for $h_{\text{HC}}^2 = 0.25$. The benefit from genomic selection (compared with BLUP) was higher in sib selection scenarios (up to 22% for $h_{\text{HC}}^2 = 0.55$ and 18% for $h_{\text{HC}}^2 = 0.25$) than in indirect selection scenarios (up to 8% for $h_{\text{HC}}^2 = 0.55$ and 9% for $h_{\text{HC}}^2 = 0.25$).

As expected, when selecting simultaneously for BW and a yield trait, genomic selection led not only to benefits in HC but also in BW. The extra gains in BW from genomic selection were up to 9% (10%) for $h_{\text{HC}}^2 = 0.55$ ($h_{\text{HC}}^2 = 0.25$) with sib selection and up to 12% (10%) for $h_{\text{HC}}^2 = 0.55$ ($h_{\text{HC}}^2 = 0.25$) with indirect selection when compared to the equivalent BLUP scenarios. The scenarios leading to higher increases in the phenotypic mean of BW were those leading to lower increases in the phenotypic mean of HC. Thus, a higher increase in BW was obtained from indirect genomic selection (scenario $G_{\text{W+I}}$) for $h_{\text{HC}}^2 = 0.55$ and from genomic sib selection (scenario $G_{\text{W+H}}$) for $h_{\text{HC}}^2 = 0.25$.

4. Discussion

This study has investigated the potential of indirect and genomic

selection to increase carcass yield in rainbow trout. This is a very important trait due to its relationship with production efficiency as more edible meat (increased profitability) and less waste (decreased environmental impact) for a given amount of resources (such as facilities, feed, labor, and energy) are produced. However, it is difficult to improve. A wide range of selection strategies differing in the traits selected (the carcass yield itself or an indicator trait) and the evaluation method (BLUP or GBLUP) have been compared at the same selection intensity. Similar patterns in the efficiency of the different strategies were found when selection was applied for the yield trait (HC or IHC) and when this trait was simultaneously selected with growth (BW).

In general, when the accuracy of evaluation for the indicator trait was higher (lower) than the accuracy for HC, indirect selection led to higher (lower) phenotypic means for HC than sib selection. However, for $h_{\text{HC}}^2 = 0.55$ higher gains in HC were obtained in the scenarios applying BLUP sib selection than in those applying indirect selection despite the fact that the accuracy of the evaluation for IHC was higher than that for HC. This is because, although the accuracy of the evaluation was lower with sib selection (records on sibs rather than on candidates), the genetic variance for HC was much higher than that for IHC. Previous studies that focused on mass selection where fish are selected solely in accordance to phenotypic values of sibs or candidates also found higher gains in HC with sib selection (Haffray et al., 2013; Vandeputte et al., 2017; Prchal et al., 2018).

Even with a h_{HC}^2 more than twice the heritability of IHC, the differences in phenotypic gains for HC between sib and indirect selection scenarios were not large (77.63% for B_{H} versus 77.58% for B_{I} and 77.82% for G_{H} versus 77.64% for G_{I}). These differences increased in favor of sib selection as we increased the number of sibs tested (up to 20 in the case of BLUP and up to 60 in the case of GBLUP). However, sib selection implies costly lethal recording and extra costs associated with maintaining and genotyping a higher number of fish (sibs). Also, for a given number of fish kept in the program, indirect selection permits a higher intensity of selection as more candidates can be raised, leading to higher responses. Extra simulations were run for indirect selection (and $h_{\text{HC}}^2 = 0.55$) with an increased selection intensity (4000 candidates instead of 2000 candidates and 2000 sibs which implies proportions selected of 5% for males and 10% for females instead of 10% and 20%), and the results (not shown) indicated that indirect selection led to a higher gain in HC than sib selection for both BLUP and GBLUP. In these extra scenarios, gains from BLUP indirect selection and from genomic sib selection were very similar. Taking into account these considerations, it is very likely that indirect selection would be a more profitable option than sib selection with either BLUP or GBLUP.

The benefit of indirect selection (when compared with sib selection) was higher in a more hypothetical scenario with a lower heritability for HC ($h_{\text{HC}}^2 = 0.25$). In this scenario, a higher gain in HC was obtained from indirect selection in both BLUP and GBLUP scenarios. Decreasing h_{HC}^2 to be at the same level as the heritability of the indicator trait increased the efficiency of indirect selection, especially for BLUP selection. In this case, for sib selection to give the same gain as indirect selection, we would need to increase the number of tested sibs to 20 per family.

With indirect selection, the benefit of genomic selection (i.e., GBLUP versus BLUP indirect selection) assuming 2000 candidates, was 6% and 11% for $h_{\text{HC}}^2 = 0.55$ and $h_{\text{HC}}^2 = 0.25$, respectively. Although GBLUP implies extra (genotyping) costs, the benefits for the on-growing farmers are expected to be high due to the high biomass of farmed fish. Small improvements in carcass and fillet yields can result in large profits when farming millions of kilos of fish (Kankainen et al., 2016). Currently, genomic selection is normal practice in the salmon breeding industry, and it can be expected that the same will occur in the near future in other aquaculture species. Industry leaders typically apply the most effective tools and, in the long-term, companies that use the best tools prevail. In recent years, the number of medium- and high-density SNP arrays available for aquaculture species has increased. This together with the

reduction of genotyping costs associated with the new advances in next-generation sequencing, make genomic evaluations more affordable (Joshi et al., 2018; Robledo et al., 2018). Recent research has demonstrated the potential of applying genomic selection in rainbow trout as the prediction accuracies from genomic evaluations are substantially higher than those from pedigree-based BLUP evaluations for reproductive traits (D'Ambrosio et al., 2020) and disease resistance (Vallejo et al., 2017; Silva et al., 2019; Vallejo et al., 2019; Yoshida et al., 2019a; Vallejo et al., 2020; Vallejo et al., 2021). The accuracy of genomic selection could be increased further by including information on allele effects of SNPs affecting the selected traits into the genetic evaluations (e.g., Zhang et al., 2016; Griot et al., 2021). In fact, this application has been investigated to improve yield traits in rainbow trout (Gonzalez-Pena et al., 2016; Salem et al., 2018) and Nile tilapia (Yoshida et al., 2019b).

Interestingly, the accuracy with GBLUP sib selection continued to increase when the number of tested fish per family increased above 60 while the accuracy with BLUP did not increase further when increasing the number above 40. This is because GBLUP captures the genetic variation existing within families, and the benefits from this evaluation method are greater when family sizes are large. These results are in concordance with those found by Dagnachew and Meuwissen (2019) that showed that the accuracy of evaluation was still increasing when 90 sibs per family were tested for a trait with a very low heritability ($h^2 = 0.05$). This means that the structure of traditional pedigree-based breeding programs may need to be changed when genomic selection is implemented.

The SNP density used in GBLUP scenarios in this study was similar to that found in the SNP arrays already developed for rainbow trout (Palti et al., 2015; Salem et al., 2018), and increasing the marker density further is not expected to improve the accuracy of genomic selection. Solberg et al. (2008) showed that the accuracy remained responsive to increasing marker density at least up to $8N_e/\text{Morgan}$, where N_e is the effective population size. Our N_e (estimated as $4N_mN_f/(N_m + N_f)$, where N_m and N_f are the number of breeding males and females, respectively; Falconer and Mackay, 1996) was 267 and thus the accuracy would increase up to a density of about 2136 SNPs/Morgan, a value lower than the density used here (2400 SNPs/Morgan). Recently, Kriaridou et al. (2020) have investigated the value of low- and medium-density SNP arrays to accurately predict breeding values in populations of a variety of aquaculture species (Atlantic salmon, common carp, gilthead sea bream, and Pacific oyster). Despite the differences in genome size, family number, population size, and target trait across the datasets, they found that in general 1000 - 2000 SNPs would be enough to achieve accuracies close to the maximum and explained the consistency across species by their large family sizes (long haplotypes are shared between many individuals in the reference and test population). In particular, recent studies in rainbow trout have shown the benefit of genomic selection when compared to BLUP selection even when using low-density SNP panels and a small training population which facilitates the implementation of genomic selection by reducing the cost of genotyping (Vallejo et al., 2017; Vallejo et al., 2018; Vallejo et al., 2021). This is due to the high LD found in breeding populations of this species (Vallejo et al., 2017; Vallejo et al., 2018).

The growing interest in aquaculture breeding programs to improve traits related to production efficiency that are difficult to quantify and measure (e.g., fillet and carcass yield, fillet and visceral lipid deposition, feed intake or feed conversion ratio), has led to a large amount of research to find appropriate indicator traits (Rutten et al., 2004; Van Sang et al., 2009; Haffray et al., 2012, 2013; Janhunen et al., 2017; Vandeputte et al., 2017, 2019; Knap and Kause, 2018; Prchal et al., 2018, 2020, 2021). These indicators need to be easy to measure and non-lethal so that they can be recorded on candidates thus allowing to exploit the within-family variability. Most indicators for fillet yield developed so far are based on one or a combination of several morphological measurements (Rutten et al., 2004; Van Sang et al., 2009; Haffray et al.,

2012, 2013; Vandeputte et al., 2017, 2019; Prchal et al., 2020, 2021). Here, we have simulated one of these indicators (the ratio of abdominal wall thickness to depth of the peritoneal cavity or E8/E23) that was developed for HC (Haffray et al., 2013; Prchal et al., 2018; Vandeputte et al., 2019) and showed that indirect selection based on it would be a very valuable approach to improve carcass yield, particularly when combined with genomic selection. Although the best scenario would be to record the indicators close to market size (i.e., at harvest), Prchal et al. (2021) have shown recently in common carp that the genetic correlation between the E8/E23 recorded on two-year old fish and at market size is high. Thus, another advantage of applying indirect selection through the indicator could be the possibility of obtaining phenotypes at an earlier age which would facilitate fish manipulation (Prchal et al., 2021).

All selective breeding programs in aquaculture include growth traits in their breeding objectives (Gjedrem et al., 2012). These traits (e.g., body weight and daily gain) are relevant to increase production and reduce the duration of the rearing cycle but have generally a modest direct impact on production efficiency. This is because growth traits have a lower impact on the ratio input/output than yield traits or feed efficiency (Rutten et al., 2004; Kause et al., 2007; Vandeputte et al., 2017; Knap and Kause, 2018; Prchal et al., 2018). Thus, breeding goals should also include other traits such as HC to increase the profitability of the aquaculture industry. Our results show that when the yield trait is simultaneously selected with BW, the selection approach leading to the highest response in HC is the same as when single trait selection (for the yield trait or even only for BW) is performed.

Given the values of the genetic correlations (higher between BW and IHC than between BW and HC), the reduction in the response in BW when selecting also the yield trait was lower when this trait was IHC (scenarios B_{W+I} and G_{W+I}) than when it was HC (scenarios B_{W+H} and G_{W+H}) which reinforces the value of indirect selection. When combining indirect with genomic selection there was an extra response in BW (about 10% extra gain when compared to BLUP indirect selection, for $h_{HC}^2 = 0.55$). It should be noted that, if GBLUP is going to be used for evaluating the yield trait, once the fish are genotyped, other traits including BW can be evaluated using GBLUP at the same genotyping expense. This will increase the benefit-cost ratio of indirect genomic selection.

5. Conclusions

This study has shown that the most profitable option to improve yield traits is a combination of genomic and indirect selection. In other words, indirect genomic selection for the indicator of HC is an interesting cost-benefit option to increase carcass yield. This was the case not only when selection was for a single yield trait but also when a growth trait was simultaneously selected with the yield trait, for a wide range of relative weights given to both traits.

Funding

This work was supported by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (FPI-SGIT2016-02), the European Union's Seventh Framework Program (KBBE.2013.1.2-10) under grant agreement n° 613611 FISHBOOST project, the European Commission Horizon 2020 (H2020) Framework Programme through grant agreement n° 727315 MedAID project (Mediterranean Aquaculture Integrated Development), and MCIN/ AEI /10.13039/501100011033 (Project PID2020-114426GB-C22).

Code availability statement

The Fortran F90 codes are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Silvia García-Ballesteros: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Jesús Fernández:** Conceptualization, Software, Investigation, Writing – review & editing, Visualization, Supervision. **Antti Kause:** Conceptualization, Investigation, Writing – review & editing. **Beatriz Villanueva:** Conceptualization, Investigation, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to Cristina Óvilo for access to her server.

Appendix A. Supplementary data

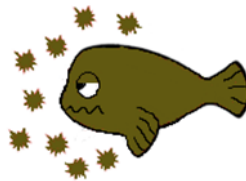
Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aquaculture.2022.738119>.

References

- D'Agaro, E., Favaro, A., Matiussi, S., Gibertoni, P.P., Esposito, S., 2021. Genomic selection in salmonids: new discoveries and future perspectives. *Aquac. Int.* 29, 2259–2289. <https://doi.org/10.1007/s10499-021-00747-w>.
- Dagnachew, B., Meuwissen, T., 2019. Accuracy of within-family multi-trait genomic selection models in a sib-based aquaculture breeding scheme. *Aquaculture* 505, 27–33. <https://doi.org/10.1016/j.aquaculture.2019.02.036>.
- D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., Dupont-Nivet, M., 2019. Genome-wide estimates of genetic diversity, inbreeding, and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. *Genet. Sel. Evol.* 51, 26. <https://doi.org/10.1186/s12711-019-0468-4>.
- D'Ambrosio, J., Morvezen, R., Brard-Fudulea, S., Bestin, A., Acin Perez, A., Guéméné, D., Phocas, F., 2020. Genetic architecture and genomic selection of female reproduction traits in rainbow trout. *BMC Genomics* 21, 58. <https://doi.org/10.1186/s12864-020-06955-7>.
- Falconer, D.S., Mackay, T.F.C., 1996. *Introduction to Quantitative Genetics*, 4th ed. Longman Group Ltd, Edinburgh, England.
- Fernández, J., Villanueva, B., Toro, M.A., 2021. Optimum mating designs for exploiting dominance in genomic selection schemes for aquaculture species. *Genet. Sel. Evol.* 53, 14. <https://doi.org/10.1186/s12711-021-00610-9>.
- Fraslin, C., Dupont-Nivet, M., Haffray, P., Bestin, A., Vandeputte, M., 2018. How to genetically increase fillet yield in fish: new insights from simulations based on field data. *Aquaculture* 486, 175–183. <https://doi.org/10.1016/j.aquaculture.2017.12.012>.
- García-Ballesteros, S., Fernández, J., Toro, M.A., Villanueva, B., 2021. Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families. *Aquaculture* 543, 737004. <https://doi.org/10.1016/j.aquaculture.2021.737004>.
- Gjedrem, T., Robinson, N., Rye, M., 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. *Aquaculture* 350–353, 117–129. <https://doi.org/10.1016/j.aquaculture.2012.04.008>.
- Gonzalez-Pena, D., Gao, G., Baranski, M., Moen, T., Cleveland, B.M., Kenney, P.B., Leeds, T.D., 2016. Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (*Oncorhynchus mykiss*). *Front. Genet.* 7, 203. <https://doi.org/10.3389/fgene.2016.00203>.
- Griot, R., Allal, F., Phocas, F., Brard-Fudulea, S., Morvezen, R., Haffray, P., Vandeputte, M., 2021. Optimization of genomic selection to improve disease resistance in two marine fishes, the European sea bass (*Dicentrarchus labrax*) and the gilthead sea bream (*Sparus aurata*). *Front. Genet.* 12, 665920. <https://doi.org/10.3389/fgene.2021.665920>.
- Haffray, P., Bugeon, J., Pincet, C., Chapuis, H., Mazeiraud, E., Rossignol, M.-N., Dupont-Nivet, M., 2012. Negative genetic correlations between production traits and head or bony tissues in large all-female rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 368–369, 145–152. <https://doi.org/10.1016/j.aquaculture.2012.09.023>.
- Haffray, P., Bugeon, J., Rivard, Q., Quittet, B., Puyo, S., Allamelou, J.M., Dupont-Nivet, M., 2013. Genetic parameters of in-vivo prediction of carcass, head, and fillet yields by internal ultrasound and 2D external imagery in large rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 410–411, 236–244. <https://doi.org/10.1016/j.aquaculture.2013.06.016>.
- Hayes, B.J., Visscher, P.M., Goddard, M.E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91, 47–60. <https://doi.org/10.1017/S0016672308009981>.
- Hill, W.G., Robertson, A., 1968. The effects of inbreeding at loci with heterozygote advantage. *Genetics* 60, 615–628.
- Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Robledo, D., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. *Nat. Rev. Genet.* 21, 389–409. <https://doi.org/10.1038/s41576-020-0227-y>.
- Janhunen, M., Nousiainen, A., Koskinen, H., Vehviläinen, H., Kause, A., 2017. Selection strategies for controlling muscle lipid content recorded with a non-destructive method in European whitefish, *Coregonus lavaretus*. *Aquaculture* 481, 229–238. <https://doi.org/10.1016/j.aquaculture.2017.09.016>.
- Joshi, R., Arnyasi, M., Lien, S., Gjoen, H.M., Alvarez, A.T., Kent, M., 2018. Development and validation of 58K SNP-array and high-density linkage map in Nile tilapia (*O. niloticus*). *Front. Genet.* 9, 472. <https://doi.org/10.3389/fgene.2018.00472>.
- Kankainen, M., Setälä, J., Kause, A., Quinton, C., Airaksinen, S., Koskela, J., 2016. Economic values of supply chain productivity and quality traits calculated for a farmed European whitefish breeding program. *Aquac. Econ. Manag.* 20, 131–164. <https://doi.org/10.1080/13657305.2016.1155961>.
- Kause, A., Paananen, T., Ritola, O., Koskinen, H., 2007. Direct and indirect selection of visceral lipid weight, fillet weight, and fillet percentage in a rainbow trout breeding program. *J. Anim. Sci.* 85, 3218–3227. <https://doi.org/10.2527/jas.2007-0332>.
- Knap, P.W., Kause, A., 2018. Phenotyping for genetic improvement of feed efficiency in fish: lessons from pig breeding. *Front. Genet.* 9, 184. <https://doi.org/10.3389/fgene.2018.00184>.
- Kriaridou, C., Tsairidou, S., Houston, R.D., Robledo, D., 2020. Genomic prediction using low density marker panels in aquaculture: performance across species, traits, and genotyping platforms. *Front. Genet.* 11, 124. <https://doi.org/10.3389/fgene.2020.00124>.
- Lillehammer, M., Meuwissen, T.H.E., Sonesson, A.K., 2013. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. *Genet. Sel. Evol.* 45, 1–8. <https://doi.org/10.1186/1297-9686-45-39>.
- Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., Vitezica, Z., 2015. *Manual for BLUPF90 Family of Programs*. Univ. Georg. Athens, USA.
- Nejati-Javaremi, A., Smith, C., Gibson, J.P., 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. *J. Anim. Sci.* 75, 1738–1745. <https://doi.org/10.2527/1997.7571738x>.
- Nielsen, H.M., Sonesson, A.K., Yazdi, H., Meuwissen, T.H.E., 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. *Aquaculture* 289, 259–264. <https://doi.org/10.1016/j.aquaculture.2009.01.027>.
- Nielsen, H.M., Sonesson, A.K., Meuwissen, T.H.E., 2011. Optimum contribution selection using traditional best linear unbiased prediction and genomic breeding values in aquaculture breeding schemes. *J. Anim. Sci.* 89, 630–638. <https://doi.org/10.2527/jas.2009-2731>.
- Palti, Y., Gao, G., Liu, S., Kent, M.P., Lien, S., Miller, M.R., Moen, T., 2015. The development and characterization of a 57 K single nucleotide polymorphism array for rainbow trout. *Mol. Ecol. Resour.* 15, 662–672. <https://doi.org/10.1111/1755-0998.12337>.
- Powell, J., White, I., Guy, D., Brotherstone, S., 2008. Genetic parameters of production traits in Atlantic salmon (*Salmo salar*). *Aquaculture* 274, 225–231. <https://doi.org/10.1016/j.aquaculture.2007.11.036>.
- Prchal, M., Bugeon, J., Vandeputte, M., Kause, A., Vergnet, A., Zhao, J., Kocour, M., 2018. Potential for genetic improvement of the main slaughter yields in common carp with in vivo morphological predictors. *Front. Genet.* 9, 283. <https://doi.org/10.3389/fgene.2018.00283>.
- Prchal, M., Kocour, M., Vandeputte, M., Kause, A., Vergnet, A., Zhao, J., Bugeon, J., 2020. Morphological predictors of slaughter yields using 3D digitizer and their use in a common carp breeding program. *Aquaculture* 520, 734993. <https://doi.org/10.1016/j.aquaculture.2020.734993>.
- Prchal, M., Zhao, J., Gela, D., Kašpar, J., Lepič, P., Kašpar, V., Kocour, M., 2021. Simplified method for genetic slaughter yields improvement in common carp under European pond conditions. *Aquac. Rep.* 21, 100832. <https://doi.org/10.1016/j.aqrep.2021.100832>.
- Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P., Houston, R., 2018. Applications of genotyping by sequencing in aquaculture breeding and genetics. *Rev. Aquac.* 10, 670–682. <https://doi.org/10.1111/raq.12193>.
- Rutten, M.J.M., Bovenhuis, H., Komen, H., 2004. Modeling fillet traits based on body measurements in three Nile tilapia strains (*Oreochromis niloticus* L.). *Aquaculture* 231, 113–122. <https://doi.org/10.1016/j.aquaculture.2003.11.002>.
- Rutten, M.J.M., Bovenhuis, H., Komen, H., 2005. Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.). *Aquaculture* 246, 125–132. <https://doi.org/10.1016/j.aquaculture.2005.01.006>.
- Salem, M., Al-Tobasei, R., Ali, A., Lourenco, D., Gao, G., Palti, Y., Leeds, T.D., 2018. Genome-wide association analysis with a 50K transcribed gene SNP-chip identifies QTL affecting muscle yield in rainbow trout. *Front. Genet.* 9, 387. <https://doi.org/10.3389/fgene.2018.00387>.
- Silva, R.M.O., Evenhuis, J.P., Vallejo, R.L., Gao, G., Martin, K.E., Leeds, T.D., Lourenco, D.A.L., 2019. Whole-genome mapping of quantitative trait loci and accuracy of genomic predictions for resistance to columnaris disease in two rainbow

- trout breeding populations. *Genet. Sel. Evol.* 51, 42. <https://doi.org/10.1186/s12711-019-0484-4>.
- Solberg, T.R., Sonesson, A.K., Woolliams, J.A., Meuwissen, T.H.E., 2008. Genomic selection using different marker types and densities. *J. Anim. Sci.* 86, 2447–2454.
- Sonesson, A.K., Meuwissen, T.H.E., 2009. Testing strategies for genomic selection in aquaculture breeding programs. *Genet. Sel. Evol.* 41, 37. <https://doi.org/10.1186/1297-9686-41-37>.
- Sonesson, A.K., Ødegård, J., 2016. Mating structures for genomic selection breeding programs in aquaculture. *Genet. Sel. Evol.* 48, 46. <https://doi.org/10.1186/s12711-016-0224-y>.
- Vallejo, R.L., Leeds, T.D., Gao, G., Parsons, J.E., Martin, K.E., Evenhuis, J.P., Palti, Y., 2017. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genet. Sel. Evol.* 49, 17. <https://doi.org/10.1186/s12711-017-0293-6>.
- Vallejo, R.L., Silva, R.M., Evenhuis, J.P., Gao, G., Liu, S., Parsons, J.E., Palti, Y., 2018. Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: evidence that long-range LD is a major contributing factor. *J. Anim. Breed. Genet.* 135, 263–274. <https://doi.org/10.1111/jbg.12335>.
- Vallejo, R.L., Cheng, H., Fragomeni, B.O., Shewbridge, K.L., Gao, G., MacMillan, J.R., Palti, Y., 2019. Genome-wide association analysis and accuracy of genome-enabled breeding value predictions for resistance to infectious hematopoietic necrosis virus in a commercial rainbow trout breeding population. *Genet. Sel. Evol.* 51, 47. <https://doi.org/10.1186/s12711-019-0489-z>.
- Vallejo, R.L., Fragomeni, B.O., Cheng, H., Gao, G., Long, R.L., Shewbridge, K.L., Palti, Y., 2020. Assessing accuracy of genomic predictions for resistance to infectious hematopoietic necrosis virus with progeny testing of selection candidates in a commercial rainbow trout breeding population. *Front. Vet. Sci.* 7, 590048. <https://doi.org/10.3389/fvets.2020.590048>.
- Vallejo, R.L., Cheng, H., Fragomeni, B.O., Gao, G., Silva, R.M.O., Martin, K.E., Palti, Y., 2021. The accuracy of genomic predictions for bacterial cold water disease resistance remains higher than the pedigree-based model one generation after model training in a commercial rainbow trout breeding population. *Aquaculture* 545, 737164. <https://doi.org/10.1016/j.aquaculture.2021.737164>.
- Van Sang, N., Thomassen, M., Klemetsdal, G., Gjøs, H.M., 2009. Prediction of fillet weight, fillet yield, and fillet fat for live river catfish (*Pangasianodon hypophthalmus*). *Aquaculture* 288, 166–171. <https://doi.org/10.1016/j.aquaculture.2008.11.030>.
- Vandeputte, M., Puleda, A., Tyran, A.S., Bestin, A., Coulombet, C., Bajek, A., Haffray, P., 2017. Investigation of morphological predictors of fillet and carcass yield in European sea bass (*Dicentrarchus labrax*) for application in selective breeding. *Aquaculture* 470, 40–49. <https://doi.org/10.1016/j.aquaculture.2016.12.014>.
- Vandeputte, M., Bugeon, J., Bestin, A., Desgranges, A., Allamellou, J.-M., Tyran, A.-S., Haffray, P., 2019. First evidence of realized selection response on fillet yield in rainbow trout *Oncorhynchus mykiss*, using sib selection or based on correlated ultrasound measurements. *Front. Genet.* 10, 1225. <https://doi.org/10.3389/fgene.2019.01225>.
- VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- Villanueva, B., Pong-Wong, R., Fernández, J., Toro, M.A., 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. *J. Anim. Sci.* 83, 1747–1752. <https://doi.org/10.2527/2005.8381747x>.
- Villanueva, B., Fernández, J., García-Cortés, L.A., Varona, L., Daetwyler, H.D., Toro, M.A., 2011. Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs. *J. Anim. Sci.* 89, 3433–3442. <https://doi.org/10.2527/jas.2010-3814>.
- Yoshida, G.M., Carvalheiro, R., Rodríguez, F.H., Lhorente, J.P., Yáñez, J.M., 2019a. Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout. *Genomics* 111, 127–132. <https://doi.org/10.1016/j.ygeno.2018.01.008>.
- Yoshida, G.M., Lhorente, J.P., Correa, K., Soto, J., Salas, D., Yáñez, J.M., 2019b. Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in Nile tilapia (*Oreochromis niloticus*). *G3 Genes Genomes Genet.* 9, 2597–2607. <https://doi.org/10.1534/g3.119.400116>.
- Zenger, K.R., Khatkar, M.S., Jones, D.B., Khalilisamani, N., Jerry, D.R., Raadsma, H.W., 2019. Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Front. Genet.* 9, 693. <https://doi.org/10.3389/fgene.2018.00693>.
- Zhang, X., Lourenco, D., Aguilar, I., Legarra, A., Misztal, I., 2016. Weighting strategies for single-step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. *Front. Genet.* 7, 151. <https://doi.org/10.3389/fgene.2016.00151>.

CHAPTER 3





Impact of selection for disease resistance on the spread of infection in a simulated aquaculture population

Results unpublished

Abstract

Infectious disease outbreaks are a major threat to the economic and environmental sustainability of the aquaculture industry. Given this, disease resistance is included in the breeding objective in aquaculture species with advanced selection programs such as Atlantic salmon. In these programs, challenge tests are performed to record phenotypes on sibs of the selection candidates. The main resistance phenotype recorded is time of death and it is unknown whether direct selection for time of death also reduces disease transmission and severity of outbreaks, which depends on three underlying host traits affecting transmission: susceptibility, infectivity, and infection-induced mortality. This simulation study aimed at predicting the impact of genomic selection for time of death on disease transmission in an Atlantic salmon population. Time of death phenotypes were assumed to be obtained from a population challenged with the Infectious Salmon Anemia (ISA) virus. Three scenarios varying in the time point when the challenge test was stopped were simulated. Selection for time of death led to a reduction in the three underlying transmission traits when these were assumed to be genetically uncorrelated. It also reduced the basic reproductive ratio (R_0) in all scenarios, although the magnitude of reduction differed between scenarios. The greatest decrease in R_0 was for the scenario where the challenge test continued until mortality naturally levels off. In this case, R_0 reached values below 1 indicating that the disease cannot be spread. Thus, genomic selection for time of death can simultaneously increase survival and reduce disease transmission.

Introduction

As already mentioned in the General Introduction, disease resistance is a trait of great economic importance for aquaculture industries where infectious disease outbreaks are a major threat to their economic and environmental sustainability (Houston et al., 2020). Given this, disease resistance (defined as the ability of fish to survive when exposed to the pathogen) is often included in the breeding objective of advanced selective breeding programs such those carried out in salmonid species.

In aquaculture programs, disease resistance is often recorded on sibs of selection candidates in challenge tests, and the main phenotype is time of death (Doeschl-Wilson et al., 2021). Selection on this trait (or on survival) aims at improving individual resilience to infection that is, at strengthening a fish's ability to cope with infection. However, it is unknown whether selecting directly on time of death reduces pathogen transmission and, thus, the incidence and severity of outbreaks (Doeschl-Wilson et al., 2021). This is because the spread of the disease (and also time of death) is affected by three underlying epidemiological host traits, which are difficult to record: susceptibility (defined as the propensity of an uninfected individual to become infected when exposed to infectious material), infectivity (that is the ability of an individual, once infected, to transmit the infection) and infection-induced mortality (defined as the propensity of an infected individual to die) (Tsairidou et al., 2019; Anacleto et al., 2019; Saura et al., 2019; Chase-Topping et al., 2021).

The objective of the study was to predict, through computer simulations (using the software **GAINFISH**), the impact of selection for disease resistance measured as time of death on the three underlying traits and, therefore, on the spread of the disease by mimicking a selective breeding program for Atlantic salmon.

Material and Methods

Simulation of the population

The genome simulated for the individuals in the population was composed of 30 chromosomes of 1 Morgan each, approximating the salmon genome. For each chromosome, 31,500 evenly spaced biallelic loci (including 1,500 SNP markers to be used in genomic evaluations) were simulated. The base population, from which the selection program began, was created in a two-step process. In the first step, a large population in mutation-drift equilibrium was generated by randomly selecting 500 males and 500 females across 4,000 discrete generations. Once the equilibrium was reached, the second step was creating a family structure by selecting and randomly mating 100 males with 200 females. This process led to 200 full-sib families and 100 half-sib families. Each mating produced 30 fish (10 selection candidates and 20 sibs to be used in the challenge test).

Simulation of underlying traits

Three uncorrelated traits (susceptibility, infectivity, and infection-induced mortality) were considered. A total of 1,000 loci (randomly sampled from the 31,500 loci) affected simultaneously the three traits for which the initial phenotypic variance and the heritability (h^2) were 1 and 0.5, respectively. These parameter values led to a h^2 for the observed trait (which is simulated from the underlying traits – see next section) of 0.2. The phenotypic values for the underlying traits were obtained by adding to the true breeding values an environmental deviation sampled from a normal distribution with mean 0 and variance 0.5.

Simulation of observed trait

The observed trait was time of death (i.e., the day when an individual dies from the beginning of the challenge test) which depends on when it gets infected and on when it dies once infected. These times were obtained assuming a stochastic compartmental Susceptible-Infected-Removed epidemiological model (e.g. Lipschutz-Powell *et al.* 2012). The algorithm used (based on Anacleto *et al.*, 2015) specifies inter-event times which determine the transition rate of individuals to move between susceptible to infected that is determined by the force of infection (the infection rate of an individual j) as $\lambda_j = \beta \exp(S_j) \sum_i \exp(I_i)$ and the transition rate of individuals to move between infected to dead

as $w_j = (\gamma \exp(M_j))$, where β is a mean *per capita* transmission rate for the disease, γ is a mean *per capita* infection-induced death rate, S_j and M_j represent fractional deviations in the susceptibility and infection-induced mortality rates of individual j compared to a population-wide reference, and I_i represents the fractional deviation in infectivity of individual i (the sum goes over all infected individuals sharing the same contact group as j at any given point in time). Estimates for β and γ were obtained by matching simulated infection and survival profiles to those from published studies of ISA virus challenge experiments in Atlantic salmon with fast transmission (the basic reproductive ratio was assumed to be $R_0 \sim 4$ for $t = 0$) (Ødegård et al., 2007; Chase-Topping et al., 2021).

Simulated challenge tests consisted of 20 groups, each of 200 fish (one from each family). The infection was seeded by randomly choosing five fish per group as initially infected. The rest of the individuals were considered susceptible. Three scenarios varying in the point where the challenge test was stopped were simulated. This point was established when i) the cumulative mortality reached an asymptote, which occurred at ~ 60 days when 80% of the fish were dead (scenario A); ii) 50% of the fish per group were dead, which occurred on day ~ 16 (scenario B); and iii) 25% of the fish per group were dead, which occurred on day ~ 7 (scenario C). The survivors' phenotype was assumed to be the last day of the challenge test plus one (i.e., day 61, 17, and 8 for A, B, and C, respectively).

Genetic evaluation and selection

Standard genomic evaluation (GBLUP) using 1,500 SNPs/M was implemented to obtain estimates of breeding values for the observed trait. The BLUPf90 software (Misztal et al., 2015) was used for carrying out GBLUP evaluation. The 100 males and the 200 females with the highest estimated breeding value for time of death were selected each generation. Each selected male was randomly mated with two females and each mating produced 30 fish (10 selection candidates and 20 sibs to be used in the challenge test). Selection was practiced for 10 discrete generations.

Parameters evaluated

For each generation t , the cumulative genetic gain for each underlying trait was calculated as the true breeding value averaged across all fish born in t . The basic reproductive ratio (R_0) was also evaluated as a measure of the impact of selection for

disease resistance (time to death) on the spread of the disease. Extrapolating from Hyman and Li (2000) the expected theoretical R_0 was estimated for each group by randomly choosing an individual as infected index case (IC) in an otherwise susceptible population:

$$R_0 = (\beta/\gamma) * \left(\sum_{i=1}^n (\exp(S_i)) / n \right) * (\exp(I_{IC}) / \exp(M_{IC}))$$

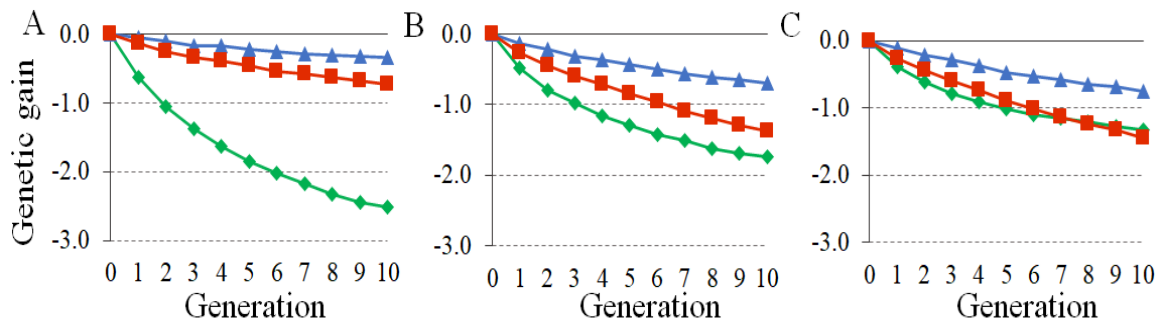
where β and γ (defined above) are 0.8 and 0.4 respectively, n is the total number of susceptible fish per group (excluding the fish chosen as IC), S_i is the susceptibility phenotype of fish i and I_{IC} and M_{IC} are respectively the infectivity and mortality phenotypes of the fish chosen as IC . For each group, R_0 was computed as the average R_0 calculated with the five initially infected fish as IC . Finally, for a given generation, R_0 was calculated as the average R_0 across groups. A total of 25 replicates were simulated and the results presented are averages over replicates.

Results

Cumulative genetic gain for each underlying trait

Selection for time of death led to a reduction in all three underlying traits although the magnitude of the response for each trait differed across the scenarios considered (Figure 1). In scenario A, the genetic gain was mainly for susceptibility. As the duration of the challenge test decreased, the gain of the three traits became more similar, although in the three scenarios the lowest response was obtained for infectivity. At $t = 10$, in scenario A the genetic gain for susceptibility was 70% higher than in scenario B and 53% higher than in scenario C. However, the gain for mortality and infectivity was about half in scenario A than in scenarios B and C.

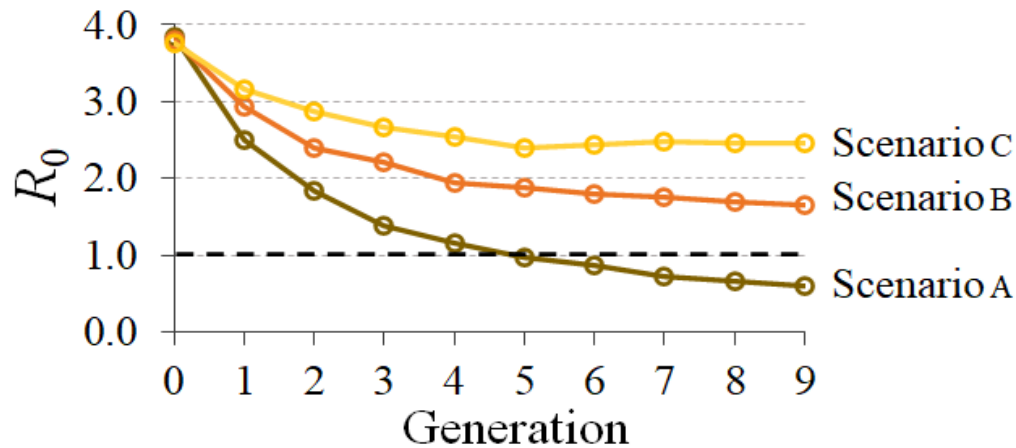
Figure 1. Cumulative genetic gain across generations for each trait (susceptibility in diamonds, infectivity in triangles, and mortality in squares) for scenarios varying in the proportion of cumulative mortality at the end of the challenge (scenarios A, B, and C).



Basic reproductive ratio (R_0)

Selection for time of death decreased R_0 in all scenarios (Figure 2). However, only scenario A led to an R_0 value below 1 and this was reached at $t = 5$. Reducing R_0 below 1 is an important objective, as it implies that the disease cannot spread.

Figure 2. Basic reproductive ratio (R_0) across generations for scenarios varying in the proportion of cumulative mortality at the end of the challenge (scenarios A, B, and C).



Discussion

In this study, we have simulated a disease leading to high mortality within a 60-day time window with mortality profiles similar to those found in some diseases of great importance such as infectious salmon anemia (Ødegård *et al.* 2007, Holborn *et al.* 2020). We show that selection for disease resistance measured as time of death not only increases survival (Houston *et al.* 2020, Doeschl-Wilson *et al.* 2021), but also contributes to reducing disease transmission in a severe outbreak when the underlying traits are assumed to be uncorrelated.

The genetic response for each underlying trait and R_0 depended on when the challenge test was stopped. The closer the end of the challenge test was to the point where cumulative mortality rate reached an asymptote, the greater the response in susceptibility and the greater the decrease in R_0 . When the challenge test was stopped at the point where the asymptote in the cumulative mortality rate was reached, the survivors at the end of the test were mainly fish with low susceptibility (they were not infected or they were infected late and did not die before the end of the test). However, when the challenge test was stopped before, the genetic response in susceptibility and the decrease in R_0 were lower and the response in infectivity and infection-induced mortality was higher. In this

case, the survivors were fish with low susceptibility, but also fish with medium and high susceptibility and medium and low mortality (i.e. fish that were infected but survived with the infection for a long time and had not yet died when the challenge test was stopped) and also fish with high and medium susceptibility that coexist with groups of infected fish with less infectivity (in scenarios B and C, fish with low infectivity and low mortality have less chance/time of infecting).

Conclusion

In summary, our results highlight the fact that, genomic selection for time of death can simultaneously increase survival and reduce disease transmission. In addition, if the selection objective is to control the spread of the disease, the challenge test should continue until mortality naturally levels off.

References

- Anacleto O., Garcia-Cortés L.A., Lipschutz-Powell D., Woolliams J.A., and Doeschl-Wilson A. B. (2015). *A novel statistical model to estimate host genetic effects affecting disease transmission*. Genetics, 201(3), 871-884. <https://doi.org/10.1534/genetics.115.179853>
- Chase-Topping M.E., Pooley C., Moghadam H.K., Hillestad B., Lillehammer M. et al. (2021). *Impact of vaccination and selective breeding on the transmission of infectious salmon anemia virus*. Aquaculture, 535: 736365. <https://doi.org/10.1016/j.aquaculture.2021.736365>
- Doeschl-Wilson A., Knap P.W., Opriessnig T., and More S.J. (2021). *Livestock disease resilience: from individual to herd level*. Animal, 15: 100286. <https://doi.org/10.1016/j.animal.2021.100286>
- Holborn M.K., Ang K.P., Elliott J.A.K., Powell F., and Boulding E.G. (2020). *Genome wide analysis of infectious salmon anemia resistance in commercial Saint John River Atlantic salmon*. Aquaculture, 514: 734514. <https://doi.org/10.1016/j.aquaculture.2019.734514>
- Houston R.D., Bean T.P., Macqueen D.J., Gundappa M.K., Jin Y.H. et al. (2020). *Harnessing genomics to fast-track genetic improvement in aquaculture*. Nature Reviews Genetics, 21(7):389-409. <https://doi.org/10.1038/s41576-020-0227-y>

Hyman J.M., and Li J. (2000). *An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations*. Mathematical Biosciences, 167(1):65-86. [https://doi.org/10.1016/S0025-5564\(00\)00025-0](https://doi.org/10.1016/S0025-5564(00)00025-0)

Misztal I., Tsuruta S., Lourenco D., Aguilar I., Legarra A. et al. (2015). *Manual for BLUPF90 family of programs*. Available online at http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf

Ødegård J., Olesen I., Gjerde B., and Klemetsdal G. (2007). *Evaluation of statistical models for genetic analysis of challenge-test data on ISA resistance in Atlantic salmon (Salmo salar): prediction of progeny survival*. Aquaculture, 266:70-76. <https://doi.org/10.1016/j.aquaculture.2007.02.012>

CHAPTER 4





García-Ballesteros, S., Villanueva, B., Fernández, J., Gutiérrez, J. P., Cervantes, I. (2021).
Genetic parameters for uniformity of harvest weight in Pacific white shrimp (*Litopenaeus vannamei*). Genetics Selection Evolution, 53, 1-9.
<https://doi.org/10.1186/s12711-021-00621-6>2021

2021 Journal impact factor 5.100, JCR Rank 5/62 in the *Agriculture, Dairy & Animal Science* category and 37/175 in the *Genetics & Heredity* category.

RESEARCH ARTICLE

Open Access



Genetic parameters for uniformity of harvest weight in Pacific white shrimp (*Litopenaeus vannamei*)

Silvia García-Ballesteros^{1*} , Beatriz Villanueva¹, Jesús Fernández¹, Juan Pablo Gutiérrez² and Isabel Cervantes²

Abstract

Background: Uniformity of body weight is a trait of great economic importance in the production of white shrimp (*Litopenaeus vannamei*). A necessary condition to improve this trait through selective breeding is the existence of genetic variability for the environmental variance of body weight. Although several studies have reported such variability in other aquaculture species, to our knowledge, no estimates are available for shrimp. Our aim in this study was to estimate the genetic variance for weight uniformity in a farmed population of shrimp to determine the potential of including this trait in the selection program. We also estimated the genetic correlation of weight uniformity between two environments (selection nucleus and commercial population).

Methods: The database contained phenotypic records for body weight on 51,346 individuals from the selection nucleus and 38,297 individuals from the commercial population. A double hierarchical generalized linear model was used to analyse weight uniformity in the two environments. Fixed effects included sex and year for the nucleus data and sex and year-pond combination for the commercial data. Environmental and additive genetic effects were included as random effects.

Results: The estimated genetic variance for weight uniformity was greater than 0 (0.06 ± 0.01) in both the nucleus and commercial populations and the genetic coefficient of variation for the residual variance was 0.25 ± 0.01 . The genetic correlation between weight and weight uniformity was close to zero in both environments. The estimate of the genetic correlation of weight uniformity between the two environments (selection nucleus and commercial population) was 0.64 ± 0.06 .

Conclusions: The existence of genetic variance for weight uniformity suggests that genetic improvement of this trait is possible. Selection for weight uniformity should not decrease weight, given the near zero genetic correlation between these two traits. The strong genetic correlation of weight uniformity between the two environments indicates that response to selection for uniformity in the nucleus will be at least partially transmitted to the commercial population if this trait is included in the breeding goal.

Background

Shrimp production is an important economic activity in the aquaculture industry and ranks third in value, after salmon and trout. The white-leg shrimp (*Litopenaeus*

vannamei) is the crustacean with the highest production level, i.e. thousands of tons of live weight worldwide [1]. In 2016, production of this species accounted for 53% of the total crustacean production and is continuously increasing. In fact, its production increased by 55% from 2010 to 2016 [1]. In addition, due to the growing demand, especially from developed countries, the price of farmed shrimp has increased in recent years.

*Correspondence: silvia.garciab@inia.es

¹ Departamento de Mejora Genética Animal, INIA, 28040 Madrid, Spain
Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Once production management conditions are controlled, the development of genetic improvement programs is one of the key factors for increased production efficiency and competitiveness. Currently, most of the selective breeding programs for shrimp focus on improving growth traits only [2, 3]. However, as growth rate increases and production intensifies [4, 5], other traits related to the quality and uniformity of the final product gain importance for both consumers and producers [6]. Shrimp are graded and classified according to standards that are defined in high-quality marketing evaluations, and are mainly determined by their physical characteristics and uniformity of size [7, 8]. In particular, shrimp are graded according to their size and count per unit of weight. Prices between size categories vary widely, and a larger number of shrimp per weight unit (i.e., of smaller size) results in a price reduction [9, 10]. Therefore, increasing the consistency of size within a specific count range can increase profit margins in the shrimp industry [10]. In addition, large variation in body size can cause competition among shrimp (dominance hierarchies), which negatively affects growth rate, mortality, and feed efficiency, and increases the need for management practices such as size grading [11]. Another indirect benefit of improving uniformity is its potential to improve resilience, which is defined as the ability of an animal to maintain performance in spite of environmental perturbations [11]. For all these reasons, and given that weight is genetically highly correlated with size, uniformity of weight is a clear candidate trait to be included in shrimp breeding programs.

Weight uniformity depends on the sensitivity of an individual to macro- and micro-environmental factors. Macro-environmental factors are measurable factors such as temperature, seasonality, diet and management, whereas micro-environmental factors are non-mensurable animal-specific factors within a given macro-environment. A necessary condition to increase weight uniformity, is the existence of genetic variance for response to such micro-environmental factors [12–14], such that individuals with genotypes that make them less sensitive to environmental disturbances will have more homogeneous offspring and show less environmental within-family variance. Several quantitative genetic models that account for genetic variance of the residual (environmental) variance have been developed [15–17], so-called heterogeneous residual variance models.

The existence of a genetic component for environmental variance has been demonstrated for different species of farmed mammals for traits such as weight and litter size [18–20]. However, in spite of the great interest to improve uniformity of aquaculture

products, studies in aquaculture species are very scarce (see recent review by de Souza et al. [11]), and to our knowledge, there is very little information on weight uniformity in crustaceans [21].

In aquaculture selective breeding programs, the breeding nucleus (in which selection is performed) is usually kept separate from the commercial population that is composed of individuals destined for sale in the market. In aquaculture, the macro-environmental rearing conditions can differ greatly between the nucleus and the commercial population. Thus, if genotype-by-environment interactions exist, genetic improvement achieved in the nucleus may not be fully translated to the commercial population. In fact, in many aquaculture species, growth traits show a significant re-ranking of genotypes in different environments [22]. For body weight uniformity, Sae-Lim et al. [23] showed a moderate re-ranking of families in different environments in rainbow trout.

The aim of this study was to estimate the genetic variance of body weight uniformity in a farmed population of Pacific white shrimp (*Litopenaeus vannamei*), using a heterogeneous residual variance model. In addition, to investigate whether selecting for increased weight uniformity in the breeding nucleus leads to improvement of uniformity in the commercial population, the genetic correlation of weight uniformity between the two environments (selection nucleus and commercial population) was estimated.

Methods

Data

The data used in this study were obtained from the CAMANICA S.A. company, which is based in Nicaragua and is carrying out a breeding program in shrimp with discrete generations. Selection is for body weight and the number of selected individuals per generation is 300 (150 males and 150 females). Each male is mated (through artificial fertilization) with a single female, and each female is mated to a single male. Once shrimp reach the appropriate size, a random sample of individuals per family (between 150 and 200 depending on the year) are tagged, with half of them being individually tagged with eye-rings and assigned to the nucleus (*N*) and the other half being tagged at the family level with elastomers and assigned to the commercial population (*C*). Within the nucleus, all families are reared in the same tank. However, in the commercial population, three to four ponds that are located in different geographical zones are used per generation, with each family equally represented in each pond. Environmental conditions differ greatly between the nucleus and the commercial populations.

Thus, weight in the selection nucleus and weight in the commercial population are considered as two different traits.

The data used here are from three consecutive generations and 425 families. The total number of individuals with phenotypic records for body weight at harvest was 89,643, of which 51,346 belonged to the nucleus and 38,297 belonged to the commercial population. Harvest time was established by estimating the days required to reach an average weight of 15 g in the nucleus. This time was set for both commercial (all ponds) and for the nucleus environments. However, for management reasons, recording the phenotypes of all shrimp can take a few days. Sex, year and pond were also recorded. A description of the number of shrimp in each category and environment is in Table 1. The average number of offspring with harvest body weight recorded per family was 205, with a standard deviation of 43.

Evaluation method

Genetic parameters for weight and weight uniformity (defined as the environmental variability of body weight) in each environment (N and C) were estimated using double hierarchical generalized linear models (DHGLM) [24, 25]. As repeated observations are needed to deal with analysis of individual animal variances, the harvest weight variability was attributed to the family in this study. The family was identified by the mother [26–28] but it also contains a sire effect due to the single pair mating design, resulting in the sire and dam effects to be confounded. Thus, although harvest weight itself was considered as a trait of the individual, its genetic effect was modelled as the sum of the dam plus sire effects. This dam model has been shown to be highly successful to modify environmental variability of birth weight in a mice population under a single pair mating design, after 7 [27] and 17 [28] generations of selection. The analyses were performed separately for N and C using univariate models (including weight and its uniformity), and also jointly using a multivariate model. The same fixed and random effects were used for weight and its uniformity for both N and C . Models were fitted using the ASReml software [29]. The equation for the multivariate DHGLM model was:

Table 1 Description of the data structure in the nucleus (N) and in the commercial population (C)

	N	C	Total
Number of shrimp	51,346	38,297	89,643
Males	25,456	17,327	42,783
Females	25,890	20,970	46,860
N_{11}	16,144	2874; 3436; 3751	26,205
N_{12}	20,050	2870; 4949; 4738	32,607
N_{13}	15,152	3142; 3411; 2536; 6590	30,831
N_{off} (min–max)	118 (23–177)	88 (21–178)	205 (47–299)
A_{wt} (g)	28; 34; 33	19; 20; 23	–
SD_{wt} (g)	4.34; 5.65; 5.04	2.92; 3.03; 3.24	–

N_{ix} : number of shrimp in level x of fixed effects (year in N and year-pond in C);

N_{off} : average number of offspring per family; A_{wt} : average weight per year; SD_{wt} : standard deviation of weight per year

$$\begin{bmatrix} \mathbf{y}_N \\ \mathbf{d}_N \\ \mathbf{y}_C \\ \mathbf{d}_C \end{bmatrix} = \begin{bmatrix} \mathbf{X}_N & 0 & 0 & 0 \\ 0 & \mathbf{X}_N^* & 0 & 0 \\ 0 & 0 & \mathbf{X}_C & 0 \\ 0 & 0 & 0 & \mathbf{X}_C^* \end{bmatrix} \begin{bmatrix} \mathbf{b}_N \\ \mathbf{b}_N^* \\ \mathbf{b}_C \\ \mathbf{b}_C^* \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_N & 0 & 0 & 0 \\ 0 & \mathbf{Z}_N^* & 0 & 0 \\ 0 & 0 & \mathbf{Z}_C & 0 \\ 0 & 0 & 0 & \mathbf{Z}_C^* \end{bmatrix} \begin{bmatrix} \mathbf{u}_{m,N} \\ \mathbf{u}_{m,N}^* \\ \mathbf{u}_{m,C} \\ \mathbf{u}_{m,C}^* \end{bmatrix} + \begin{bmatrix} \mathbf{e}_N \\ \mathbf{e}_N^* \\ \mathbf{e}_C \\ \mathbf{e}_C^* \end{bmatrix},$$

where subscript N refers to the selection nucleus and C to the commercial population, * indicates parameters that are associated with the residual variance (uniformity of body weight), $\mathbf{y}$ is the vector of body weight phenotypes, $\mathbf{d}$ is the vector of residual variances, with elements $d_i = \hat{e}_i^2 / (1 - h_i)$, where $\hat{e}_i^2$ is the squared residual estimate of body weight in each environment (N or C) and h_i is the i th diagonal element in the hat matrix of $\mathbf{y}$ for that trait [24], $\mathbf{b}$ and $\mathbf{b}^*$ are vectors of fixed effects (sex and year for N and sex and the year-pond interaction for C), $\mathbf{X}$ and $\mathbf{X}^*$ are incidence matrices associated with the fixed effects, $\mathbf{u}_m$ and $\mathbf{u}_m^*$ are vectors of additive genetic effects of the mother plus the father, and $\mathbf{Z}$ and $\mathbf{Z}^*$ are incidence matrices associated with the random genetic effects. The distribution of the additive genetic effects was as follows:

$$\begin{bmatrix} \mathbf{u}_{m,N} \\ \mathbf{u}_{m,N}^* \\ \mathbf{u}_{m,C} \\ \mathbf{u}_{m,C}^* \end{bmatrix} \sim N \left(0, \begin{bmatrix} \sigma_{u_{m,N}}^2 & \rho\sigma_{u_{m,N}}\sigma_{u_{m,N}^*} & \rho\sigma_{u_{m,N}}\sigma_{u_{m,C}} & \rho\sigma_{u_{m,N}}\sigma_{u_{m,C}^*} \\ 0 & \sigma_{u_{m,N}^*}^2 & \rho\sigma_{u_{m,N}^*}\sigma_{u_{m,C}} & \rho\sigma_{u_{m,N}^*}\sigma_{u_{m,C}^*} \\ 0 & sym & \sigma_{u_{m,C}}^2 & \rho\sigma_{u_{m,C}}\sigma_{u_{m,C}^*} \\ 0 & & & \sigma_{u_{m,C}^*}^2 \end{bmatrix} \otimes \mathbf{A} \right),$$

where $\mathbf{A}$ is the numerator relationship matrix based on pedigree information available for the mothers (three generations), $\sigma_{u_m}^2$ and $\sigma_{u_m^*}^2$ are the additive genetic variances for body weight and its uniformity, respectively, ρ is the genetic correlation coefficient and $\otimes$ represents the Kronecker product. Residuals of $\mathbf{y}(\mathbf{e})$ and $\mathbf{d}(\mathbf{e}^*)$ were assumed to be independent and normally distributed:

$$\begin{bmatrix} \mathbf{e}_N \\ \mathbf{e}_N^* \\ \mathbf{e}_C \\ \mathbf{e}_C^* \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{W}_N^{-1}\sigma_e^2 & 0 & 0 & 0 \\ 0 & \mathbf{W}_N^{*-1}\sigma_{e^*}^2 & 0 & 0 \\ 0 & 0 & \mathbf{W}_C^{-1}\sigma_e^2 & 0 \\ 0 & 0 & 0 & \mathbf{W}_C^{*-1}\sigma_{e^*}^2 \end{bmatrix} \right),$$

where $\mathbf{W} = \text{diag}(\hat{d}^{-1})$, $\mathbf{W}^* = \text{diag}(\frac{1-h_i}{2})$, and σ_e^2 and $\sigma_{e^*}^2$ are scaling variances (set equal to 1). The multivariate DHGLM used a weighted gamma GLM fitted with response $\mathbf{d}$ and weights $\mathbf{W}^*$ and $\mathbf{W}$ [24]. For the univariate DHGLM, only the parameters for one environment (N or C) were included in the model.

Parameters evaluated

The estimated parameters included the genetic variance for both weight and its uniformity in N and C , heritabilities, genetic correlations between traits (see below), and the genetic coefficient of variation for the residual variance (GCV_e) of weight uniformity. The latter was computed using the approximation $\sqrt{\sigma_{u_m^*}^2}$ because $\sigma_{u_m^*}^2$ is on an exponential scale [17, 20].

Genetic correlations were estimated (i) for weight between the two environments (N and C); (ii) for weight uniformity between the two environments; (iii) between weight and weight uniformity within each environment (N or C); and (iv) between weight in one environment and weight uniformity in the other environment.

Heritability of weight uniformity (h_v^2) was estimated for N and C as the proportion of phenotypic variance in body weight that was due to the genetic variance of the residual variance [17, 20]: $h_v^2 = \sigma_{u^*}^2 / (2\sigma_p^4 + 3\sigma_{u^*}^2)$, where $\sigma_p^2 = 2\sigma_u^2 + \sigma_e^2$. Because $\sigma_{u^*}^2$ is on an exponential scale, it was converted to an additive scale following Mulder et al. [17], i.e., $\sigma_{u_a^*}^2 = \sigma_e^4 \exp(2\sigma_{u^*}^2) - \sigma_e^4$, where σ_e^4 equals $\sigma_e^2 \exp(\frac{1}{2}\sigma_{u^*}^2)$.

Heritability of weight was estimated for N and C . The estimated genetic variance ($\sigma_{u_m}^2$) refers to sire plus dam effects and equals to half of the total genetic variance (σ_u^2). Under heteroscedastic models, heritability (i.e., the usual ratio of additive to phenotypic variance) of the mean (weight) is not unique because it depends on the value of the residual variance, which in turn is conditioned to the levels of the environmental effects ($\mathbf{b}^*$) [19, 26, 30]:

$$h_i^2 = \frac{\sigma_u^2}{\sigma_{u_m}^2 + e^{(\mathbf{Xb}_i^* + 1/2\sigma_{u_m^*}^2)}}.$$

A global heritability was estimated for weight in N and C using a global residual variance obtained by averaging across fixed effect solutions. Estimates of heritability for each combination of level l of a systematic effect s were also obtained, using:

$$\sigma_{e_{sl}}^2 = e^{i=1, \text{systematics}} \left(\sum_{j=1, n_s} \frac{\hat{b}_{ij}}{n_s} \right) + \hat{b}_{sl} + 1/2\sigma_{u_m^*}^2,$$

resulting heritability equal to:

$$h_{sl}^2 = \frac{\sigma_u^2}{\sigma_{u_m}^2 + \sigma_{e_{sl}}^2}.$$

These estimates of heritability were obtained using a Bayesian approach with the GSEVM software [31] and using univariate models with the same systematic and random effects as included when using a frequentist approach. Results from the Bayesian approach were obtained by averaging the results from a Markov chain of Monte Carlo (MCMC) chain of 1,000,000 iterations, sampling one of each 100 iterations, following a burn-in of 100,000. Inferences were based on probabilities obtained from the marginal posterior distributions of

Table 2 Estimates (standard errors) of variance components and genetic parameters for weight and weight uniformity in different environments (nucleus, N , and commercial, C) obtained from multivariate and univariate DHGLM models

	Multivariate		Univariate	
	N	C	N	C
Weight				
$\sigma_{u_m}^2$	4.45 (0.33)	1.60 (0.12)	4.42 (0.33)	1.60 (0.12)
σ_u^2	8.90	3.20	8.85	3.19
σ_e^2	19.06	6.64	18.98	6.60
σ_p^2	23.51	8.23	23.40	8.19
h^2	0.38	0.39	0.38	0.39
Weight uniformity				
$\sigma_{u_m^*}^2$	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)
h_v^2	0.02	0.02	0.02	0.02
GCV_e	0.25 (0.01)	0.24 (0.01)	0.25 (0.01)	0.23 (0.02)

DHGLM: double hierarchical generalized linear model. $\sigma_{u_m}^2, \sigma_{u_m^*}^2$: additive variances of the mother plus father for body weight and its uniformity, respectively; σ_u^2 : additive variance for body weight defined as $2^*\sigma_{u_m}^2$; σ_e^2 : residual variance; σ_p^2 : phenotypic variance; h^2, h_v^2 : heritabilities of weight and its uniformity, respectively; GCV_e : genetic coefficient of residual variation for weight uniformity

the parameters or combinations of parameters and the estimates are the means of these marginal posterior distributions.

Results

Genetic variance components and heritabilities

Estimates of variance components and genetic parameters for weight and its uniformity in the two environments (*N* and *C*) are in Table 2 for the univariate and multivariate models, which led to very similar results. The estimate of the additive genetic variance for body weight was substantially larger in *N* than in *C* but the estimated heritabilities were similar in the two environments. The estimate of the additive genetic variance for weight uniformity was different from 0 for both *N* and *C* (0.06 ± 0.01) and the estimate of heritability was 0.02 for both. Thus, the genetic coefficient of residual variation (weight uniformity) was the same in the two environments (around 0.25), as was the estimate of the global heritability of body weight (around 0.39).

Figure 1 shows the estimates of heritabilities (mean and standard deviation of the marginal posterior distributions obtained using the Bayesian approach) for the different levels of systematic effects. These included sex (male or female) and year (from 1 to 3) for weight in *N* and sex and year-pond (from 1 to 10) for weight in *C*. In *N*, the estimate of heritability for males was about 46% higher than for females (0.38 vs. 0.26). In *N*, the estimate of heritability was highest for the first year and lowest for the second year. The probabilities (obtained from the marginal posterior distributions) that heritability was lower in the third year than in the first year heritability and higher than in the second year were 99.8 and 100%, respectively. Differences in estimates of heritability within levels of systematic effects were also observed for *C*. In *C*, the probability that the heritability was higher

Table 3 Estimates (standard errors) of genetic correlations between weight and weight uniformity within and between environments (nucleus, *N*, and commercial, *C*) obtained from multivariate (above the diagonal) and univariate (below the diagonal) DHGLM models

	<i>N</i>		<i>C</i>	
	Weight	Uniformity	Weight	Uniformity
<i>N</i>				
Weight		0.11 (0.06)	0.61 (0.03)	0.06 (0.07)
Uniformity	0.11 (0.06)		0.01 (0.06)	0.64 (0.06)
<i>C</i>				
Weight	–	–		0.05 (0.07)
Uniformity	–	–	0.07 (0.07)	

DHGLM: double hierarchical generalized linear model

for males than for females was 99.9%, in agreement with the results obtained for *N*. For the year-pond effect in *C*, level 2 shows the lowest estimate of heritability with a probability of 85 to 100% of being lower than for the other year-pond levels.

Genetic correlations

Estimates of the genetic correlation between weight and its uniformity within and between the *N* and *C* environments were also very similar for the univariate and multivariate models (Table 3). In *N*, the estimate of the genetic correlation between weight and its uniformity was positive and relatively low (0.11) but higher than in *C*, for which the estimate was not significantly different from 0.

Estimates of the genetic correlation between weight in *N* and weight in *C* and between weight uniformity in *N* and weight uniformity in *C* were both positive (about 0.60). Estimates of the genetic correlation between weight in *N* and weight uniformity in *C* and between weight in *C*

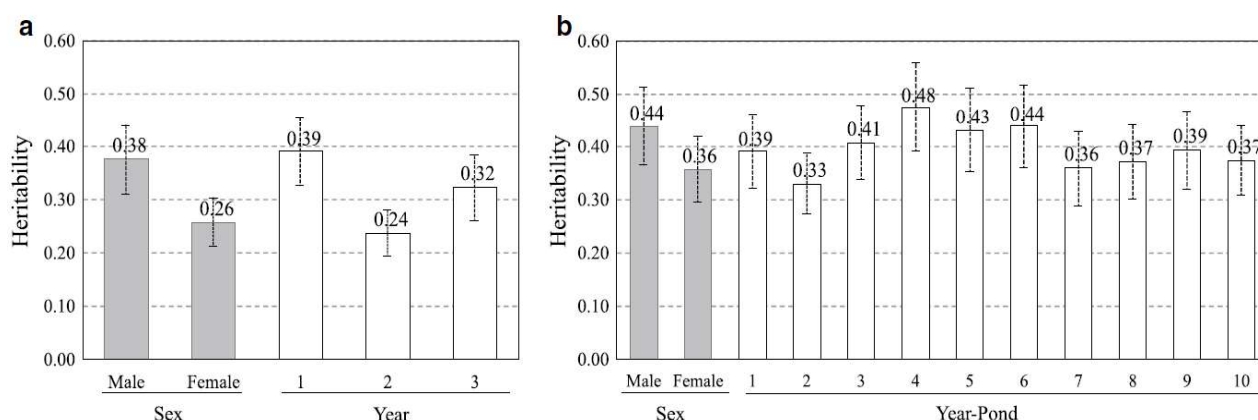


Fig. 1 Estimates of heritability (mean and a posteriori standard deviation) of harvest weight for each level of residual systematic effects. Systematic effects were sex and year for the nucleus (a) and sex and year-pond for the commercial population (b)

and weight uniformity in N were not significantly different from 0.

Ranking of families for weight uniformity in the two environments

Spearman correlations between estimated breeding values for weight uniformity in the nucleus and the commercial population were obtained by assuming a multivariate heterogeneous model. The correlations obtained using all (100%) candidate families (i.e., in N), or families in the top 50, top 20, and top 10% were 0.87, 0.79, 0.77, and 0.76, respectively. 'Top' families in this case were those with the highest estimated breeding value for uniformity (lowest weight variability). As expected, correlations were lower when the number of top families decreased, which indicates that a higher re-ranking between environments N and C for weight uniformity is expected in the best families. Regardless of this, correlations were in general relatively high.

Discussion

Genetic variances and heritabilities

Although weight uniformity is a very relevant trait with the potential of being included in shrimp breeding programs, there is very little information on the existence of genetic variation for this trait [21]. To our knowledge, this is the first study that uses a double hierarchical generalized linear model to estimate genetic variance for body weight uniformity in shrimp, and constitutes a first step to investigate the possibility of including this trait in the breeding goal. This is important since the weight uniformity evaluated here was individual sensitivity to micro-environmental disturbances. Estimates of the additive genetic variance, heritability, and genetic coefficient of residual variation for weight uniformity that were obtained for this Pacific white shrimp population in the nucleus, in which selection takes place, were all different from 0, which indicates that genetic improvement for this trait is possible. In addition, the genetic correlation of weight uniformity between the nucleus and the commercial population was relatively high, which indicates that improvement obtained in the nucleus would be partially transmitted to the commercial population, with the economic benefits that this would entail.

Estimates of the global heritability for body weight at harvest in N and C (0.38 and 0.39) were within the range of those found in the literature for shrimp [4, 5, 21]. More important, is the fact that estimates of the additive genetic variance for uniformity of weight and for the residual heritability were also in the range of those described for shrimp [21], other aquaculture [23, 32–35], and terrestrial [11] species. The estimate of the genetic coefficient of residual variation was lower than that

reported for Nile tilapia [33] but similar to that reported for trout [23], and within the range of that reported for terrestrial animals [11]. This indicates the existence of genetic variation in micro-environmental sensitivity among full sib families (dam plus sire effects), which implies that the phenotypes of offspring of different families will be differentially affected by the environment. Thus, our results show that the potential of genetic selection to improve weight uniformity is similar to that for other species. Selection for improving homogeneity of birth weight in rabbits [36] or environmental variance for birth weight in mice [27, 28] has been shown to be successful, with estimates of genetic parameters similar to those obtained in this study.

Response to selection on a trait depends on the heritability of the trait, among other factors. Under a heteroscedastic model, estimates of heritability can be obtained for each level of the systematic effects [19, 30]. Our results show that estimates of heritability differed between levels of sex and year (or year-pond). Formoso-Rafferty et al. [26] suggested to restrict the information (phenotypes) to the levels of the systematic effects that have the higher heritabilities in order to decrease the residual variance. However, they also recognized that ignoring information from levels that are associated with higher residual variance could have a negative impact on the accuracy of estimated breeding values and thus, on response to selection. In practice, funds to record phenotypes are limited and it might be better to focus on recording phenotypes for traits with the highest heritability.

Genetic correlations

In order to evaluate the potential economic benefit of including weight uniformity in the breeding goal, correlations with other traits that are currently in the breeding goal, such as body weight, must be estimated. The ideal scenario would be the existence of a negative genetic correlation between weight and its variability because it would facilitate selection for higher weight and more uniformity. In aquaculture, estimates of the genetic correlation between weight and its variability vary largely in the literature, i.e. between -0.16 and 0.79 [23, 32–35]. Our estimate was not significantly different from 0 (Table 3), which indicates that it may not be difficult to improve weight and weight uniformity simultaneously through a selection index. This would require the economic value for uniformity to be determined, which is unknown at this point.

It is very important that genetic improvements made in the nucleus are transferred to the commercial population that is composed of individuals for sale in the market. Thus, a high genetic correlation between the nucleus and commercial environments for traits that are selected

for in the nucleus is desirable. This is not always the case because, although conditions are intended to be similar in the two environments, this is not usually feasible. In particular, in aquaculture species, some environmental factors are more important than others in affecting the re-ranking of individuals based on their estimated breeding values. In their review on genotype-by-environment interactions, Sae-Lim et al. [22] found that differences in temperature, rearing system, or stocking density had a greater influence on the re-ranking of individuals for growth traits (estimates of the genetic correlation between environments ranged from 0.36 to 0.56) than differences in farm location (estimates of the genetic correlation ranged from 0.58 to 0.66). In our study, estimates of the genetic correlation between environments *N* and *C* were 0.61 for weight and 0.64 for weight uniformity, which are within the range reported for weight in other aquaculture species [22]. Our estimate of the genetic correlation of weight between environments was lower than that reported for shrimp by Castillo-Juarez et al. [5] but within the range of that reported by Gitterle et al. [4]. Our study provides, for the first time, an estimate of the genetic correlation of weight uniformity between different environments for shrimp, and it is similar to that reported for trout [23].

Statistical models

Estimates and their standard errors of variance components and genetic parameters for weight uniformity and for the genetic correlation of weight with its uniformity obtained using univariate versus multivariate models were very similar (Table 2). In fact, Pearson's correlations of estimated breeding values for weight uniformity between the univariate and multivariate models were high (0.87 in *N* and 0.80 *C*).

To estimate the genetic variation of the residual variance, repeated observations are needed. However, traits such as body weight at harvest are recorded only once. Repeated measurements can be obtained within families by assigning the trait to the mother or the family rather than to the individual. Some studies that fitted the sire and dam effects separately for both the mean (weight) and its uniformity, obtained higher accuracies and less bias for the estimated variance components than when the individual genetic effect was fitted for the mean and the sire and dam effects were fitted for uniformity [23, 32]. In our study, convergence was not obtained when this approach was used because single pair matings were performed and the sire genetic effect could not be distinguished from the dam genetic effect. Thus, to estimate the genetic variance of the environmental variability, the mean and its uniformity were attributed to the mother [26–28]. However, the mother identification

refers also to the father given that they are confounded. Maternal and common environmental effects were not included for the mean in the model because of the data structure, which originates from single pair mating with no replication of tanks per family. The maternal effect is considered to be the part of the environmental effect that affects performance of the offspring that depends on the environment provided by their mother (composition of the eggs among others) [37]. Maternal effects are significant for harvest weight in shrimp [4, 5] and other aquaculture species [34, 38, 39] but of low magnitude. The common environmental family effect originates from rearing families in different tanks during the early stages of growth (until individuals reach a sufficient size for tagging), and thus is common to members of the same family [40]. Many studies have shown that the proportion of phenotypic variance due to common environmental effects, although significantly different from 0 in shrimp [41] and other aquaculture species [35, 40, 42, 43], is of low magnitude. Some studies [42, 44] have suggested that common environmental effects are difficult to separate from family genetic effects, even with data from nested designs. In order to include these random effects in the model and avoid confounding and convergence problems, a possible solution would be to replicate tanks per family or perform factorial matings. The few studies that have included common environmental effects for weight uniformity either did not achieve convergence or did not obtain estimates significantly higher than 0 [32, 33, 35]. Thus, even if maternal and common environmental effects exist, they would probably explain only a small proportion of the phenotypic variance.

Conclusions

Our results show that genetic variability for the environmental variance of weight at harvest exists in shrimp, both in the selection nucleus and in the commercial population. The coefficients of genetic residual variation for these traits (uniformity measured in the nucleus and in the commercial population) were large enough to conclude that response to selection could be obtained if these traits were included in the breeding program. Including weight uniformity should not decrease weight since the genetic correlation between the two traits was not significantly different from zero. Further investigation is necessary to determine what is the best combination of these traits in the selection index to reach the greatest economic benefit. Based on the genetic correlation of weight uniformity between the two environments estimated here, selection in the nucleus will be transmitted to the commercial population.

Acknowledgements

Thanks to CAMANICA S.A. for their contribution to genealogical and phenotypic data. We thank Noelia Ibáñez-Escriche for very useful discussions.

Authors' contributions

JF prepared the database; SG-B, IC and JPG designed the model; SG-B and IC analysed the data; SG-B wrote the first draft of the manuscript. All authors contributed to the discussion of results and the edition of the manuscript revision. All authors read and approved the final manuscript.

Funding

This work was funded by the Ministerio de Ciencia, Innovación y Universidades, Spain (Grant CGL2016-75904-C2-2-P), fondos FEDER and the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (FPI-SGIT2016-02).

Availability of data and materials

The data that support the findings of this study are available from the CAMANICA S.A. company but restrictions apply to the availability of these data, which were used under license for the current study, and thus are not publicly available. However, data are available from the authors upon reasonable request and with permission of CAMANICA S.A.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Departamento de Mejora Genética Animal, INIA, 28040 Madrid, Spain.

² Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain.

Received: 10 July 2020 Accepted: 1 March 2021

Published online: 12 March 2021

References

- FAO. The state of world fisheries and aquaculture 2018. Meeting the sustainable development goals. Rome: FAO and UNEP; 2018.
- Neira R. Breeding in aquaculture species: genetic improvement programs in developing countries. In: Proceedings of the 9th world congress on genetics applied to livestock production, Leipzig, 1–6 August 2010.
- Castillo-Juárez H, Campos-Montes GR, Caballero-Zamora A, Montaldo HH. Genetic improvement of Pacific white shrimp [*Penaeus (Litopenaeus) vannamei*]: perspectives for genomic selection. *Front Genet*. 2015;6:93.
- Gitterle T, Rye M, Salte R, Cock J, Johansen H, Lozano C, et al. Genetic (co) variation in harvest body weight and survival in *Penaeus (Litopenaeus) vannamei* under standard commercial conditions. *Aquaculture*. 2005;243:83–92.
- Castillo-Juárez H, Casares JCQ, Campos-Montes G, Villela CC, Ortega AM, Montaldo HH. Heritability for body weight at harvest size in the Pacific white shrimp, *Penaeus (Litopenaeus) vannamei*, from a multi-environment experiment using univariate and multivariate animal models. *Aquaculture*. 2007;273:42–9.
- Parisenti J, Beirão LH, Tramonte VLGC, Ourique F, da Silveira Brito CC, Moreira CC. Preference ranking of colour in raw and cooked shrimps. *Int J Food Sci Technol*. 2011;46:2558–61.
- Balaban MO, Yeralan S, Bergmann Y. Determination of count and uniformity ratio of shrimp by machine vision. *J Aquat Food Prod Technol*. 1995;3:43–58.
- Zhang D, Lillywhite KD, Lee D-J, Tippetts BJ. Automatic shrimp shape grading using evolution constructed features. *Comput Electron Agric*. 2014;100:116–22.
- Asche F, Benneer LS, Oglend A, Smith MD. US shrimp market integration. *Mar Resour Econ*. 2012;27:181–92.
- Tveteras S. Seafood price indices. In: FAO globefish research programme, vol 78. 2005. p. 7.
- de Souza lung LH, Carvalheiro R, Neves HH, Mulder HA. Genetics and genomics of uniformity and resilience in livestock and aquaculture species: a review. *J Anim Breed Genet*. 2019;137:263–80.
- Waddington CH. Experiments on canalizing selection. *Genet Res*. 1960;1:140–50.
- Gibson JB, Bradley BP. Stabilising selection in constant and fluctuating environments. *Heredity (Edinb)*. 1974;33:293–302.
- Scheiner SM, Caplan RL, Lyman RF. The genetics of phenotypic plasticity. III. Genetic correlations and fluctuating asymmetries. *J Evol Biol*. 1991;4:51–68.
- SanCristobal-Gaudy M, Elsen JM, Bodin L, Chevalet C. Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. *Genet Sel Evol*. 1998;30:423–51.
- Sorensen D, Waagepetersen R. Normal linear models with genetically structured residual variance heterogeneity: a case study. *Genet Res*. 2003;82:207–22.
- Mulder HA, Bijma P, Hill WG. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics*. 2007;175:1895–910.
- Garrick DJ, Pollak EJ, Quaas RL, Van Vleck LD. Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves. *J Anim Sci*. 1989;67:2515–28.
- Ibáñez-Escriche N, Moreno A, Nieto B, Piqueras P, Salgado C, Gutiérrez JP. Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response. *Genet Sel Evol*. 2008;40:279–93.
- Hill WG, Mulder HA. Genetic analysis of environmental variation. *Genet Res*. 2010;92:381–95.
- Castillo-Juárez H, Montaldo HH, Campos-Montes GR. Genetic parameter estimates of the environmental variation for body weight at harvest size in a Pacific white shrimp breeding population. In: Proceedings of the 4th international conference on quantitative genetics, Edinburgh, 17–22 June 2012.
- Sae-Lim P, Gjerde B, Nielsen HM, Mulder H, Kause A. A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. *Rev Aquac*. 2016;8:369–93.
- Sae-Lim P, Kause A, Janhunen M, Vehviläinen H, Koskinen H, Gjerde B, et al. Genetic (co) variance of rainbow trout (*Oncorhynchus mykiss*) body weight and its uniformity across production environments. *Genet Sel Evol*. 2015;47:46.
- Rönnegård L, Felleki M, Fikse F, Mulder HA, Strandberg E. Genetic heterogeneity of residual variance—estimation of variance components using double hierarchical generalized linear models. *Genet Sel Evol*. 2010;42:8.
- Mulder HA, Rönnegård L, Fikse WF, Veerkamp RF, Strandberg E. Estimation of genetic variance for macro- and micro-environmental sensitivity using double hierarchical generalized linear models. *Genet Sel Evol*. 2013;45:23.
- Formoso-Rafferty N, Cervantes I, Ibáñez-Escriche N, Gutiérrez JP. Modulating birth weight heritability in mice. *J Anim Sci*. 2017;95:531–7.
- Formoso-Rafferty N, Cervantes I, Ibáñez-Escriche N, Gutiérrez JP. Genetic control of the environmental variance for birth weight in seven generations of a divergent selection experiment in mice. *J Anim Breed Genet*. 2016;133:227–37.
- Formoso-Rafferty N, Chávez KN, Ojeda C, Cervantes I, Gutiérrez JP. Selection response in a divergent selection experiment for birth weight variability in mice compared with a control line. *Animals*. 2020;10:920.
- Gilmour AR, Gogel BJ, Cullis BR, Welham S, Thompson R. ASReml user guide release 4.1 structural specification. Hemel Hempstead: VSN International Ltd.; 2015.
- Ibáñez-Escriche N, Sorensen D, Waagepetersen R, Blasco A. Selection for environmental variation: a statistical analysis and power calculations to detect response. *Genetics*. 2008;180:2209–26.
- Ibáñez-Escriche N, García M, Sorensen D. GSEVM vol 2: MCMC software to analyze genetically structured environmental variance models. *J Anim Breed Genet*. 2010;127:249–51.
- Sonesson AK, Ødegård J, Rönnegård L. Genetic heterogeneity of within-family variance of body weight in Atlantic salmon (*Salmo salar*). *Genet Sel Evol*. 2013;45:41.

33. Marjanovic J, Mulder HA, Khaw HL, Bijma P. Genetic parameters for uniformity of harvest weight and body size traits in the GIFT strain of Nile tilapia. *Genet Sel Evol*. 2016;48:41.
34. Agha S, Mekawwy W, Ibanez-Escriche N, Lind CE, Kumar J, Mandal A, et al. Breeding for robustness: investigating the genotype-by-environment interaction and micro-environmental sensitivity of genetically improved farmed tilapia (*Oreochromis niloticus*). *Anim Genet*. 2018;49:421–7.
35. Janhunen M, Kauser A, Vehviläinen H, Järvisalo O. Genetics of micro-environmental sensitivity of body weight in rainbow trout (*Oncorhynchus mykiss*) selected for improved growth. *PLoS One*. 2012;7:e38766.
36. Garreau H, Bolet G, Larzul C, Robert-Granié C, Saleil G, SanCristobal M, et al. Results of four generations of a canalising selection for rabbit birth weight. *Livest Sci*. 2008;119:55–62.
37. Green BS. Maternal effects in fish populations. *Adv Mar Biol*. 2008;54:1–105.
38. Martínez V, Neira R, Gall GAE. Estimation of genetic parameters from pedigree populations: lessons from analysis of alevin weight in Coho salmon (*Oncorhynchus kisutch*). *Aquaculture*. 1999;180:223–36.
39. Henryon M, Jokumsen A, Berg P, Lund I, Pedersen PB, Olesen NJ, et al. Genetic variation for growth rate, feed conversion efficiency, and disease resistance exists within a farmed population of rainbow trout. *Aquaculture*. 2002;209:59–76.
40. Herlinger CM, O'Reilly PT, Doyle RW, Wright JM, O'Flynn F. Early growth performance of Atlantic salmon full-sib families reared in single family tanks versus in mixed family tanks. *Aquaculture*. 1999;173:105–16.
41. Montaldo HH, Castillo-Juárez H, Campos-Montes G, Pérez-Enciso M. Effect of the data family structure, tank replication and the statistical model, on the estimation of genetic parameters for body weight at 28 days of age in the Pacific white shrimp (*Penaeus (Litopenaeus) vannamei* Boone, 1931). *Aquac Res*. 2013;44:1715–23.
42. Gallardo JA, Lhorente JP, Neira R. The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*). *Genet Sel Evol*. 2010;42:19.
43. Tosh JJ, Garber AF, Trippel EA, Robinson JAB. Genetic, maternal, and environmental variance components for body weight and length of Atlantic cod at 2 points in life. *J Anim Sci*. 2010;88:3513–21.
44. Kauser A, Ritola O, Paananen T, Wahlroos H, Mäntysaari EA. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*. 2005;247:177–87.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions





GENERAL DISCUSSION





Recapitulation

Aquaculture production faces many challenges to increase the volume and quality of aquatic products while maintaining a sustainable and profitable development. A key factor for achieving this goal is the application of selective breeding programs. In this thesis, the potential genetic gains obtained when including new economically important traits (i.e., carcass yield, disease resistance, and weight uniformity) and when applying genomic selection have been investigated with the final aim of boosting existing programs. Computer simulations or data analyses were used to deal with different questions. Simulations were used to investigate the benefit (in terms of genetic gain and inbreeding) of using genomic evaluation when common environmental (tank) effects are present (**Chapter 1**). Simulations were also used to investigate the advantages of selecting on a non-lethal indicator trait of headless carcass to improve carcass yield in rainbow trout programs, using different selection approaches which include genomic selection (**Chapter 2**). Also, combining genetic and epidemiological models in the simulations, the impact of genomic selection for reducing mortality due to a particular disease on the spread of the disease in an Atlantic salmon breeding program was investigated in **Chapter 3**. Finally, using real data of a farm population, genetic parameters for weight uniformity (another trait with high economic relevance) were estimated for Pacific white shrimp, with the final aim of determining the potential of including this trait in the existing selection programs (**Chapter 4**).

Real data analyses *versus* computer simulations

Animal breeding research uses both real and simulated data to answer specific questions. If the question deals with the genetic background of the organism (e.g., finding the genetic determination of a trait and the amount of heritable variability), analyses of real data are necessary. This has been the case of **Objective 4** where the question was whether there was or not genetic variability for weight uniformity in white shrimp. However, many questions in animal breeding deal with the prediction of genetic gain and inbreeding from alternative breeding programs in the short- and long-term, as are the questions implicit in **Objectives 1 - 3**. In these cases, computer simulation is the approach to be followed.

There are two types of computer simulations used in animal breeding: i) stochastic simulations (also called Monte Carlo simulations); and ii) deterministic simulations. With stochastic simulations, a population of animals is created by generating phenotypes for each ani

mal in the population through random sampling from predefined distributions (which, in turn, are determined by the rules of inheritance, and the environmental effects assumed in the model). Deterministic simulations use equation-based predictions to obtain population average genetic gains and levels of inbreeding. For a given degree of complexity of the breeding program, stochastic simulations are able to mimic populations at the individual animal level and are often relatively easier to develop when compared to deterministic simulations. The result of a particular run of the stochastic simulation is affected by random sampling events, so many repeated runs must be carried out to obtain the expected mean for a particular parameter. This also allows the estimation of the variance of the mean. Given this, the computer simulation program **GAINFISH** was developed for predicting the genetic gain and inbreeding expected for typical scenarios of aquaculture selective breeding programs. This software is very flexible as it considers a wide range of scenarios including: i) breeding nuclei and commercial populations; ii) different breeding objectives; iii) selection for multiple generations on different selection criteria; iv) different genetic models for each trait; v) separate or mixed rearing of families; and vi) different evaluation, selection, and mating strategies. The inputs required to run the software include: i) number and size of chromosomes; ii) breeding objectives and selection criteria; iii) gene action and genetic parameters for breeding objectives and selection criteria; iv) population structure i.e., number of families and number of fish per family; v) evaluation, selection, and mating strategies; vi) number of SNPs used in genomic evaluation; and vii) number of generations of selection. The main outputs of the program are genetic gain and inbreeding per generation. Other outputs include accuracy of evaluation, phenotypic mean, and variance, additive genetic variance, dominance variance, expected heterozygosity, and linkage disequilibrium between SNPs and between SNPs and QTLs. The use of the tool is described in detail in Annex I (**GAINFISH** manual section).

In this thesis, **GAINFISH** has been used to determine the benefit of genomic selection in different scenarios that vary in breeding goals, type of rearing, selection methods, family sizes, and linkage disequilibrium (**Objectives 1 - 3**).

Factors that affect the benefit obtained from genomic selection

Investment in effective aquaculture selective programs may provide high economic returns in future generations. Methodological and technological advances have progressed from mass selection (i.e., individuals are selected according to their own phenotypic value) to pedigree-based BLUP and to genomic selection. The application of genomic selection offers significant advantages over BLUP selection; i.e., increased rates of genetic gain and decreased rates of inbreeding per generation (Meuwissen et al., 2001; Daetwyler et al., 2007). This has been showed in both empirical (Zenger et al., 2019; Houston et al., 2020; Song and Hu, 2021) and simulation (Nielsen et al., 2009, 2011; Sonesson and Meuwissen, 2009; Villanueva et al., 2011; Lillehammer et al., 2013; Sonesson and Ødegård, 2016; **Chapters 1 - 3** of this thesis) aquaculture studies.

Genomic selection exploits the linkage disequilibrium that exists between SNPs and quantitative trait loci (QTL) affecting the trait (Goddard and Hayes, 2009) which depends on marker density. Several studies have aimed at determining the minimum SNP density required to achieve an increase in the accuracy of estimated breeding values when using genomic evaluations rather than BLUP (e.g., Meuwissen et al., 2001; Solberg et al., 2008). Given that the levels of linkage disequilibrium vary across species and across populations within species the marker density required for an efficient genomic selection would also vary. For instance, some studies have shown high levels of linkage disequilibrium for salmonids while low levels have been found for Nile tilapia (Table 4). Thus, the SNP density required for genomic selection being more accurate than BLUP would be lower in salmonids. In fact, Al-Tobasei et al. (2021) and Vallejo et al. (2018) found that, in general, 25 SNPs per Morgan are needed in a rainbow trout population to obtain higher accuracy with genomic selection than with BLUP. The density required in populations of Atlantic salmon ranged from 33 to 100 SNPs/M (Kriaridou et al., 2020; Song and Hu, 2021). However, a higher SNP density (270 to 550 SNPs/M) is needed in Nile tilapia to efficiently implement genomic selection (Yoshida et al., 2019a).

A particular genomic selection approach proposed for aquaculture breeding programs is to apply within-family genomic evaluation (WFGE) (Lillehammer et al., 2013; García-Ballesteros et al., 2018; Dagnachew and Meuwissen, 2019; García-Ballesteros et al., 2019). Under WFGE only the within-family genetic component of the breeding value is estimated using markers. Then the estimated within-family component is combined with an estimate of the between-family breeding value that can be obtained from traditional BLUP. With WFGE,

the marker density required for obtaining high accuracies is much lower than with population-wide genomic evaluation given that high levels of linkage disequilibrium exist within families (e.g., Gutiérrez et al., 2017; Saura et al., 2018). Using computer simulations, Lillehammer et al. (2013) found higher accuracy with WFGE than with BLUP (up to 10%) using only 12 SNPs/M for salmon breeding programs with large family sizes (200 fish per full-sib family). However, as the genomic evaluation is performed only within families, the accuracy of this approach could depend on the family size that is usually much smaller for species other than salmon.

It is clear that in simulation studies aiming at predicting the potential benefits of applying genomic selection, realistic linkage disequilibrium levels should be assumed. However, Zenger et al. (2019) have warned that the levels assumed in such studies have been, in general, much lower than those recently found in different species. In **Chapter 1** of this thesis, low and high levels of disequilibrium were simulated in order to cover all potential situations. In **Chapter 2**, a more species-specific linkage disequilibrium was simulated when investigating the benefit of indirect and genomic selection in a rainbow trout population. In this study, the level of linkage disequilibrium simulated was close to that found in real populations (Vallejo et al., 2018; D'Ambrosio et al., 2019).

Although the level of sophistication of genomic technologies varies greatly across aquaculture species, these technologies are rapidly advancing and becoming more affordable and accessible (Houston et al., 2020). Thus, given the development of commercial SNP arrays for aquaculture species in recent years, the application of genomic selection is already a reality. However, genotyping large numbers of fish with high-density SNP panels is still too costly, particularly given the low value of an individual fish. Therefore, there has been a great interest in developing cost-effective genomic selection strategies for practical application in aquaculture breeding programs. In addition to WFGE, other strategies that have been proposed to reduce the cost of genotyping include reducing the size of the reference population in populations with high linkage disequilibrium (Song and Hu, 2021) and combining low-density genotyping with genotype imputation (Kijas et al., 2017; Tsai et al., 2017; Yoshida et al., 2018; Tsairidou et al., 2020).

Table 4. Estimates of linkage disequilibrium (measured as r^2 , the squared correlation between pairs of loci) for different aquaculture species.

Species	Distance between markers						NoP [†]	Reference
	10kb	50kb	100kb	1Mb	5Mb	10Mb		
Atlantic salmon	0.27 – 0.30					0.07	4	López et al., 2019
			0.35 – 0.62	0.08 – 0.31	0.06 – 0.21	0.06 – 0.19	3	Barría et al., 2018
	0.54	0.31	0.27				1	Kijas et al., 2017
Coho salmon		0.20	0.16	0.14	0.10	0.07	1	Barría et al., 2019
Rainbow trout	0.30 – 0.39	0.22 – 0.30	0.19 – 0.28	0.13 – 0.20	0.09 – 0.12	0.07 – 0.09	6	D'Ambrosio et al., 2019
	0.38			0.25		0.07	1	Vallejo et al., 2018
Nile tilapia	0.19 – 0.20	0.15		0.10 – 0.11	0.04 – 0.05	0.03 – 0.04	3	Yoshida et al., 2019a
Turbot		0.07	0.06	0.05	0.04	0.02	1	Saura et al., 2021
European sea bass			0.07	0.05	0.03	0.03	1	Saura et al., 2021
				0.18 – 0.20	0.16 – 0.18	0.16 – 0.17	3	Griot et al., 2021b
Gilthead sea bream				0.24	0.23	0.21	1	Griot et al., 2021b
Pacific white shrimp		0.22		0.15	0.12	0.07	1	Jones et al., 2017
Pacific oyster		0.09 – 0.17	0.02 – 0.08					Gutiérrez et al., 2017

[†]NoP: number of populations analyzed

Genomic selection for candidate traits

The main traits of interest in the aquaculture industry are related to growth rate because of its direct effect on profitability and because they can be measured directly on the selection candidates (i.e., candidate traits). In **Chapter 1**, simulation results showed a benefit of genomic selection when compared to BLUP selection in terms of accuracy of 11% to 18% (depending on the family size assumed) for a candidate trait with a heritability of 0.4 (similar to that of growth rate). These values are within the range of those found in empirical studies on different aquaculture species, which ranged from 6% to 53% (Table 5). The benefits from genomic selection shown in Table 5 are all calculated in the absence of common environmental (tank) effects. However, in this thesis, it has been shown for the first time that the benefit of genomic selection increases in the presence of common environmental effects (benefits of 20% to 42%). These extra benefits were mainly due to the fact that, with BLUP, the accuracy of EBVs was greatly reduced by the presence of these effects whereas the accuracy of genomic evaluations was hardly affected. This fact and the new advances in the development of low-cost genotyping technologies would make it more appealing to perform genomic evaluations in breeding programs with separate rearing of families rather than BLUP evaluations, as currently done.

Uniformity of body weight

The demand for more uniform products has increased in the latest years in aquaculture species. Table 6 shows the few estimates of genetic parameters for micro-environmental sensitivity available for aquaculture species. In **Chapter 4**, for the first time, genetic parameters for weight uniformity in crustaceans were estimated by using a double hierarchical generalized linear model. The residual heritability and the genetic coefficient of variation of residual variance (or evolvability) found for the shrimp population are in the range of those described for other aquaculture species (Table 6). In general, these estimates are larger in aquaculture than in livestock species (de Souza et al., 2019), indicating that selection for uniformity is even more promising in aquaculture.

There are very few studies on the use of genomic information for estimating genomic parameters for micro-environmental sensitivity, both in livestock and in aquaculture (Mulder et al., 2013; Sell-Kubiak et al., 2015; Sae-Lim et al., 2017). These studies only found a slight improvement in the estimation of genetic parameters for mean and residual variance, using either the genomic relationship matrix (Mulder et al., 2013) or the combined (pedigree and

genomic) relationship matrix (Sae-Lim et al., 2017) instead of the pedigree-based relationship matrix. However, this small increase in accuracy is probably due to the small size of the reference populations. As the costs of genotyping decrease and genomic selection becomes more affordable, genomic selection will open the door to unraveling the quantitative genetics of residual variance.

Table 5. Benefit of genomic evaluations (in terms of accuracy of selection) when compared to traditional BLUP evaluations for candidate traits found in different aquaculture studies.

Species	Trait	Relative increase (%)	Reference
Atlantic salmon (<i>Salmo salar</i>)	Weight	21	Tsai et al., 2015
	Length	18	Tsai et al., 2015
	Weight	11	Yoshida et al., 2018
Banana shrimp (<i>Fenneropenaeus merguensis</i>)	Weight	17	Nguyen et al., 2020
	Length	22	Nguyen et al., 2020
	Width	20	Nguyen et al., 2020
Channel catfish (<i>Ictalurus punctatus</i>)	Weight	23 - 28	Garcia et al., 2018
Common carp (<i>Cyprinus carpio</i>)	Growth	18	Palaiokostas et al., 2018
Nile Tilapia (<i>Oreochromis niloticus</i>)	Weight	25	Yoshida et al., 2019b
	Weight	53	Joshi et al., 2020
Pacific oyster (<i>Magallana gigas</i>)	Wet weight	24	Gutiérrez et al., 2018
	Length	23	Gutiérrez et al., 2018
	Height	28	Gutiérrez et al., 2018
Rainbow trout (<i>Oncorhynchus mykiss</i>)	Weight	6	Gonzalez-Pena et al., 2016
	Female weight	10 - 16	D'Ambrosio et al., 2020

Table 6. Estimates of residual heritability (h_v^2), genetic coefficient of variation of residual variance (GCV_E) and genetic correlation between the mean and the residual variance (ρ) for different aquaculture species.

Species	Trait	h_v^2	GCV_E	ρ	Reference
Atlantic salmon (<i>Salmo salar</i>)	Body weight	0.060	0.34	0.42	Sonesson et al., 2013
	Log (Body weight)		0.34	-0.17	Sonesson et al., 2013
	$\sqrt{\text{Body weight}}$		0.31	0.14	Sonesson et al., 2013
Tilapia (<i>Oreochromis</i>)	Harvest weight	0.029	0.58	0.60	Marjanovic et al., 2016
	Length	0.012	0.39	0.11	Marjanovic et al., 2016
	Depth	0.014	0.42	0.37	Marjanovic et al., 2016
	Width	0.012	0.45	0.20	Marjanovic et al., 2016
	$\sqrt{\text{Weight}}$ Malaysia		0.44	-0.53	Agha et al., 2018
	$\sqrt{\text{Weight}}$ China		0.42	-0.70	Agha et al., 2018
	$\sqrt{\text{Weight}}$ India		0.30	-0.03	Agha et al., 2018
Rainbow trout (<i>Oncorhynchus mykiss</i>)	Weight	0.024	0.38	-0.16	Janhunnen et al., 2012
	Body weight nucleus	0.011	0.21	0.30	Sae-Lim et al., 2015
	Body weight commercial	0.010	0.19	0.79	Sae-Lim et al., 2015

Genomic selection for sib traits

Higher benefits from genomic selection are expected for traits with low heritability and for sib traits. This has been shown in different empirical studies (Tables 5 and 7) as well as in simulation studies (Sonesson, 2007; Sonesson and Meuwissen, 2009; Nielsen et al., 2009, 2011; Sonesson and Ødegård, 2016). As already mentioned above, the main advantage of genomic selection for sib traits is the ability to utilize within-family genetic variation in the selection decisions. In aquaculture, the most important sib traits are yield and disease resistance traits (see Table 7).

Yield traits

For the yield traits studied (i.e., carcass weight, fillet weight, and fillet yield), the benefit from genomic selection varies greatly (from 15% to 323%, Table 7). The highest benefit (> 80%) has been observed in studies using small sample sizes (~1500 fish) and a low number of fish tested per family. Specifically, these studies used ~5 fish tested per family in rainbow trout (Gonzalez-Pena et al., 2016) and ~8 (with a standard deviation of 5) in Nile tilapia (Joshi et al., 2020), and obtained very low accuracies with standard BLUP, and thus very high benefits from genomic selection. In **Chapter 2**, it was shown that the benefit of genomic selection predicted for headless carcass weight under different scenarios varying in the number of sibs tested per family (from 10 to 100) and heritability (0.55 and 0.25), ranged from 15% to 30%. These results are in concordance with those found in other studies for yield traits (García et al., 2018; Horn et al., 2020; Yoshida et al., 2019b; Besson et al., 2022).

Disease resistance

Diverse studies have shown increases in accuracy from genomic selection for resistance to specific diseases in different aquaculture species (Table 7). The benefit from genomic selection when compared to BLUP selection varies greatly (from 5% to 567%) with an average of ~30%. However, as already mentioned in this thesis, selection for disease resistance (measured as mortality) does not necessarily imply a reduction in the incidence and severity of disease outbreaks as those depend on underlying epidemiological traits such as susceptibility, infectivity, and infection-induced mortality (Doeschl-Wilson et al., 2021). In **Chapter 3**, it has been shown that genomic selection for disease resistance contributed to reduce disease transmission in a severe outbreak when these underlying traits were uncorrelated. However, this reduction could vary if the underlying traits are correlated between them and/or with mortality (i.e., observable trait). For infectious pancreatic necrosis in Atlantic salmon, a recent transmission experiment (Doeschl-Wilson et al., 2018) has found an extremely fortunate combination, i.e., a positive correlation between mortality and susceptibility (i.e., fish less likely to die had a lower probability to be infected) and between mortality and infection-induced mortality (i.e., fish die sooner when infected, thus, had shorter infective periods). This has not been the case in another recent study on infectious salmon anemia disease (Chase-Topping et al., 2021) where mortality was uncorrelated with susceptibility. More research is needed for understanding the effect of genomic selection for disease resistance on disease outbreaks for different scenarios varying in how the different traits correlate.

Reference population

In aquaculture programs designed to improve sib traits, the reference population (i.e., individuals actually recorded) is generally constituted by individuals that are full or half-sibs of the selection candidates (Gjedrem et al., 2010; Ødegård et al., 2011). Thus, there is a high degree of relatedness between the reference population and the selection candidates, and therefore, high prediction accuracy can be achieved. However, the accuracy also depends on the sample size of the reference population (i.e., number of sibs tested per family and number of families). The higher the number of sibs tested per family, the higher accuracy of estimated breeding values but there are diminishing returns on selection accuracy relative to increases in number of individuals. Therefore, increasing the number of phenotypes from a certain number of sibs tested would only further increase the cost of maintaining and recording (and genotyping) fish. The optimal (cost-effective) number of sibs tested will depend on the selection approach. In the case of BLUP selection, the accuracy barely increases beyond 20 sibs tested per family (Lillehammer et al., 2013; **Chapter 2**) and stays constant beyond 60 (**Chapter 2**). However, with WFGE selection, the accuracy still increases beyond 40 sibs tested per family for different family sizes (Lillehammer et al., 2013; García-Ballesteros et al., 2019). With population-wide genomic selection, the accuracy barely increases beyond 40 sibs tested per family (Sonesson and Meuwissen, 2009; **Chapter 2**).

A simple way to decrease the costs involved when applying genomic selection is to reduce the number of sibs tested (i.e., to reduce the size of the reference population). Although this leads to a reduction in selection accuracy, this accuracy may still be higher than that obtained from BLUP selection (Sonesson and Meuwissen, 2009; Lillehammer et al., 2013; Dagnachew and Meuwissen, 2019). For a large population (60000 selection candidates), Lillehammer et al. (2013) found 10% higher accuracy from WFGE than from BLUP with only 10 sibs tested per family and 12 SNPs/M. For a smaller population (6000 selection candidates), García-Ballesteros et al. (2019) found that at least 20 sibs tested per family and 50 SNPs/M are needed to obtain a benefit from WFGE. Finally, it has been shown in simulation (Sonesson and Meuwissen, 2009 and **Chapter 2**) and empirical (Table 7) studies that the accuracy of population-wide genomic evaluation was higher than that of BLUP when 10 sibs or more were tested per family for different population sizes.

Table 7. Benefit of genomic evaluations (in terms of accuracy of selection) when compared to traditional BLUP evaluations for sib traits found in different aquaculture studies.

Species	Trait	Relative increase (%)	Reference
Atlantic salmon (<i>Salmo salar</i>)	Fillet pigmentation	22	Ødegård et al., 2014
	Omega-3 fatty acid content	24	Horn et al., 2020
	Resistance to sea lice	7, 25	Tsai et al., 2016
	Resistance to sea lice	22	Correa et al., 2017
	Resistance to sea lice	52	Ødegård et al., 2014
	Resistance to rickettsial syndrome	21 - 30	Bangera et al., 2017
Banana shrimp (<i>Fenneropenaeus merguensis</i>)	Tail weight	17	Nguyen et al., 2020
	Resistance to hepatopancreatic virus	567	Nguyen et al., 2020
Channel catfish (<i>Ictalurus punctatus</i>)	Residual carcass weight	29 - 36	Garcia et al., 2018
Common carp (<i>Cyprinus carpio</i>)	Resistance to koi herpesvirus	8	Palaiokostas et al., 2019
European sea bass (<i>Dicentrarchus labrax</i>)	Feed conversion ratio	10 - 125	Besson et al., 2019
	Resistance to Viral Nervous Necrosis	5 - 89	Griot et al., 2021b
	Resistance to Viral Nervous Necrosis	13	Palaiokostas et al., 2018
	Resistance to vibriosis	10 - 27	Griot et al., 2021b
Gilthead sea bream (<i>Sparus aurata</i>)	Resistance to pasteurellosis	2 - 12	Griot et al., 2021b
	Resistance to pasteurellosis	27 - 53	Palaiokostas et al., 2016
Nile tilapia (<i>Oreochromis niloticus</i>)	Fillet yield	4 - 15	Yoshida et al., 2019b
	Fillet yield	20	Joshi et al., 2020
	Fillet weight	89	Joshi et al., 2020

	Resistance to Francisellosis	57	Joshi et al., 2021
Pacific oyster (<i>Magallana gigas</i>)	Resistance to ostreid herpesvirus	19	Gutiérrez et al., 2020
Rainbow trout (<i>Oncorhynchus mykiss</i>)	Carcass weight	96	Gonzalez-Pena et al., 2016
	Fillet weight	100	Gonzalez-Pena et al., 2016
	Fillet yield	323	Gonzalez-Pena et al., 2016
	Resistance to bacterial cold water disease	12	Vallejo et al., 2016
	Resistance to bacterial cold water disease	83 - 109	Vallejo et al., 2017
	Resistance to bacterial cold water disease	29 - 141	Vallejo et al., 2021
	Resistance to rickettsial syndrome	28	Yoshida et al., 2018
	Resistance to infectious pancreatic necrosis virus	8 - 12	Yoshida et al., 2019c
	Resistance to infectious pancreatic necrosis virus	42 - 154	Vallejo et al., 2019
	Resistance to infectious pancreatic necrosis virus	15	Vallejo et al., 2020
Turbot (<i>Scophthalmus maxima</i>)	Resistance to scuticociliatosis	12	Saura et al., 2019
Whiteleg shrimp (<i>Litopenaeus vannamei</i>)	High salinity tolerance	12	Luo et al., 2022
	Resistance to acute hepatopancreatic necrosis disease	5 - 6	Wang et al., 2019

Inbreeding

Inbreeding is particularly important in aquaculture selective breeding programs. Given the high fecundity of aquaculture species, base populations (from where selection programs start) can be created from a reduced number of broodstock and high selection intensities can be applied once the programs start. This leads to small N_e and, consequently, high rates of inbreeding and losses in genetic variability (Villanueva et al., 1996; Gjerde et al., 1996; Dupont-Nivet et al., 2006; Holtsmark et al., 2008). In general, estimates of N_e in farm aquaculture populations are low (Eknath and Doyle, 1990; Su et al., 1996; Pante et al., 2001; Gallardo et al., 2004; Brown et al., 2005; Borrell et al., 2011; Yáñez et al., 2014; Segvic-Bubic et al., 2017; Barriá et al., 2018; Barriá et al., 2019; D'Ambrosio et al., 2019; Saura et al., 2021; Villanueva et al., 2022), particularly when compared with wild populations (Saura et al., 2021; Villanueva et al., 2022). In fact, many estimates of N_e lay below the critical value of 50 recommended to avoid inbreeding depression and retain fitness in the short-term (Frankham et al., 2002; FAO, 2013). Therefore, the application of selection and mating approaches designed to control the increase in inbreeding and the loss of genetic variability are needed to ensure the long-term sustainability of aquaculture selective programs.

There was a considerable amount of research in the 90s aimed at developing selection and mating strategies for restricting the increase of inbreeding. In aquaculture species where mating can be directed, the traditional mating design has been the nested design in which each sire is mated to N dams, and each dam is mated to only one sire (Gjedrem and Baranski, 2009; Sonesson and Ødegård, 2016). Factorial mating involves more mates per dam (Woolliams, 1989). For instance, in full factorial mating designs each sire mates with each dam, and each dam mates with each sire, producing $s*d$ families, where s and d are the total number of sires and dams, respectively. Therefore, factorial design (compared to nested designs) leads to a proportionally higher number of half-sib and a lower number of full-sib, thus, reducing the risk of selecting many offspring from the same parents (Sorensen et al., 2005). Several studies in different aquaculture species have shown that factorial mating designs (compared to nested designs) result in up to 38% lower inbreeding rates for BLUP (Dupont-Nivet et al., 2002; Busack and Knudsen, 2007) and genomic selection (Nirea et al., 2012; Sonesson and Ødegård, 2016).

When compared with traditional BLUP selection, another advantage of genomic selection (in addition to increasing genetic gain), is that it leads to lower rates of inbreeding per generation (Daetwyler et al., 2007). The decrease in inbreeding with genomic selection is mainly due to the fact that the Mendelian sampling components of breeding values are estimated with a higher accuracy when breeding values are estimated. Consequently, genomic selection allows better differentiation within families and reduces the co-selection of relatives (Daetwyler et al., 2007; Dekkers, 2007; Sonesson and Meuwissen, 2009). This reduction in inbreeding per generation with genomic selection have ranged from 50% to 81% in different previous studies (Dekkers, 2007; Sonesson and Meuwissen, 2009; Sonesson and Ødegård, 2016) and also in **Chapter 1** of this thesis, both in the absence and in the presence of common environmental effects. However, contrarily to what happened with genetic gain (i.e., extra increase in gains in the presence of common environmental effects), there are no extra reductions in inbreeding in the presence of common environmental effects.

Genomic selection for disease resistance was applied in **Chapter 3** for different scenarios varying in the duration of the challenge test. Scenario A, in which the challenge test continued until mortality naturally levels off (with 80% of mortality), led not only to the highest reduction of disease outbreaks (the highest response to selection) but also to the highest inbreeding rate per generation (results not shown). Specifically, after 5 generations of selection, inbreeding in scenario A was double that in the scenario giving the lowest genetic gain (scenario C in which the challenge test stopped when the 25% of mortality was reached). This higher inbreeding in A was due to the fact that more individuals from the same family were selected.

Boosting aquaculture selective breeding programs

As already mentioned, the complexity of aquaculture selective breeding programs varies greatly. For example, selective breeding programs are lacking in common carp, while advanced breeding programs with several traits and routine use of genomic tools are applied in Atlantic salmon. Thus, a major challenge and opportunity for the aquaculture industry is to identify the current level of each selective breeding program and define the next step needed to improve it. This was the general objective of the European project FISHBOOST (<http://cordis.europa.eu/project/id/613611>), which was the framework of this thesis.

A broad categorization of the different levels of complexity in aquaculture programs would be: 1) basic breeding programs with few traits that are measured directly on selection

candidates; 2) advanced breeding programs with several traits and routine sib-testing to improve traits via family selection; and 3) advanced breeding programs with several traits and routine use of genomic tools to improve accuracy on sib-tested traits. Thus, common carp is at level 0 and the next step would be to apply a selective breeding program focused on improving growth while simultaneously controlling inbreeding (i.e., move to level 1). At the other extreme would be the Atlantic salmon, where most breeding programs are at level 2, the next step would be to develop strategies to reduce genotyping costs in order to apply genomic selection routinely (i.e., move to level 3).

This thesis has focused on the advances from level 1 to 2 by evaluating the expected genetic gains when including new traits of interest in selective breeding programs (**Chapters 2, 3, and 4**); and from level 2 to 3, by evaluating the benefit of using genomic information (**Chapters 1, 2, and 3**). Although there are still aspects of the inclusion of new traits in breeding goals and of the implementation of the genomic tools that should be explored in greater depth before they can be used routinely, the results presented in this thesis clearly will help to boost selective breeding programs for different aquaculture species. Also, it is important to highlight the fact that the software developed to perform the simulations will be very useful to the breeding industry to compare, in terms of gains and inbreeding, different strategies in order to decide the next step to be taken to increase competitiveness.



GENERAL CONCLUSIONS



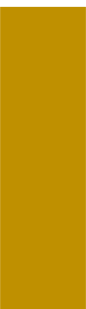


1. In schemes applying nested mating designs, the presence of common environmental effects does not affect the accuracy of genomic selection, but it does decrease the accuracy of traditional BLUP selection. Therefore, genomic selection leads to extra benefits in terms of genetic gain in the presence of common environmental effects.
2. Genomic selection leads to lower inbreeding rates than BLUP in the presence and in the absence of common environmental effects, but this reduction in inbreeding with genomic selection is smaller in the presence of common environmental effects.
3. Even when common environmental effects are small, they should be included in the genetic evaluation model because when they are ignored the increases in inbreeding are very high (although the losses in genetic gains are only moderate).
4. In rainbow trout selective breeding programs, genomic selection for an indicator of headless carcass leads to phenotypic gains similar to those obtained from genomic selection for headless carcass when both traits have the same heritability. However, when the heritability of headless carcass is twice that of the indicator, genomic selection for headless carcass leads to slightly higher gains than genomic selection for the indicator. Nevertheless, given that selection for headless carcass implies extra costs associated with maintaining and genotyping sibs, genomic selection for the indicator appears to be the most cost-effective option.
5. Assuming that the three underlying epidemiological host traits (susceptibility, infectivity, and infection-induced mortality) are genetically uncorrelated, genomic selection for disease resistance measured as time of death not only increases survival but also contributes to reducing disease transmission in a severe outbreak in a simulated Atlantic salmon population. This is because genomic selection for time of death leads to a reduction in the genetic mean of the three underlying traits and in the basic reproductive ratio.
6. The greatest decrease in the basic reproductive ratio occurs when the challenge test continues until mortality naturally levels off. Under this scenario, and assuming uncorrelated underlying epidemiological host traits, the basic reproductive ratio reaches values below one, which means that the disease will not spread in the population.

7. The existence of genetic variance for weight uniformity at harvest in Pacific white shrimp found in the selection nucleus and in the commercial population suggests that genetic improvement of this trait through selection is possible.
8. Weight uniformity and body weight were genetically uncorrelated in Pacific white shrimp and thus, including selection for weight uniformity would not lead to undesirable changes in body weight.
9. The strong positive genetic correlation between weight uniformity in both environments (nucleus and commercial population) indicates that selection for weight uniformity in the nucleus will be transmitted to the commercial population.

REFERENCES





Agha, S., Mekki, W., Ibanez-Escriche, N., Lind, C.E., Kumar, J., et al. (2018). *Breeding for robustness: investigating the genotype-by-environment interaction and micro-environmental sensitivity of Genetically Improved Farmed Tilapia (Oreochromis niloticus)*. *Animal Genetics*, 49(5), 421-427. <https://doi.org/10.1111/age.12680>

Al-Tobasei, R., Ali, A., Garcia, A.L., Lourenco, D., Leeds, T., et al. (2021). *Genomic predictions for fillet yield and firmness in rainbow trout using reduced-density SNP panels*. *BMC Genomics*, 22(1), 1-11. <https://doi.org/10.1186/s12864-021-07404-9>

Anacleto, O., Cabaleiro, S., Villanueva, B., Saura, M., Houston, R.D., et al. (2019). *Genetic differences in host infectivity affect disease spread and survival in epidemics*. *Scientific Reports*, 9(1), 1-12. <https://doi.org/10.1038/s41598-019-40567-w>

Andriantahina, F., Liu, X., Huang, H., and Xiang, J. (2012). *Response to selection, heritability and genetic correlations between body weight and body size in Pacific white shrimp, Litopenaeus vannamei*. *Chinese Journal of Oceanology and Limnology*, 30(2), 200-205. <https://doi.org/10.1007/s00343-012-1066-2>

Antonello, J., Massault, C., Franch, R., Haley, C., Pellizzari, C., et al. (2009). *Estimates of heritability and genetic correlation for body length and resistance to fish pasteurellosis in the gilthead sea bream (Sparus aurata L.)*. *Aquaculture*, 298(1-2), 29-35. <https://doi.org/10.1016/j.aquaculture.2009.10.022>

Aslam, M.L., Carraro, R., Bestin, A., Cariou, S., Sonesson, A.K., et al. (2018). *Genetics of resistance to photobacteriosis in gilthead sea bream (Sparus aurata) using 2b-RAD sequencing*. *BMC Genetics*, 19(1), 1-14. <https://doi.org/10.1186/s12863-018-0631-x>

Aslam, M.L., Carraro, R., Sonesson, A.K., Meuwissen, T., Tsigenopoulos, C.S., et al. (2020). *Genetic variation, GWAS and accuracy of prediction for host resistance to Sparicotyle chrysophrii in farmed Gilthead Sea bream (Sparus aurata)*. *Frontiers in Genetics*, 11, 1652. <https://doi.org/10.3389/fgene.2020.594770>

Bangera, R., Correa, K., Lhorente, J.P., Figueroa, R., and Yáñez, J.M. (2017). *Genomic predictions can accelerate selection for resistance against Piscirickettsia salmonis in Atlantic salmon (Salmo salar)*. *BMC Genomics*, 18(1), 1-12. <https://doi.org/10.1186/s12864-017-3487-y>

Barriá, A., Christensen, K.A., Yoshida, G., Jedlicki, A., Leong, J.S., et al. (2019). *Whole genome linkage disequilibrium and effective population size in a coho salmon (Oncorhynchus kisutch) breeding population using a high-density SNP array*. *Frontiers in Genetics*, 10, 498. <https://doi.org/10.3389/fgene.2019.00498>

Barria, A., López, M.E., Yoshida, G., Carvalheiro, R., Lhorente, J.P., et al. (2018). *Population genomic structure and genome-wide linkage disequilibrium in farmed Atlantic salmon (Salmo*

salar L.) using dense SNP genotypes. *Frontiers in Genetics*, 9, 649. <https://doi.org/10.3389/fgene.2018.00649>

Bentsen, H.B., Gjerde, B., Eknath, A.E., de Vera, M.S.P., Velasco, R.R., et al. (2017). *Genetic improvement of farmed tilapias: Response to five generations of selection for increased body weight at harvest in Oreochromis niloticus and the further impact of the project*. *Aquaculture*, 468, 206-217. <https://doi.org/10.1016/j.aquaculture.2016.10.018>

Besson, M., Allal, F., Chatain, B., Vergnet, A., Clota, F., et al. (2019). *Combining individual phenotypes of feed intake with genomic data to improve feed efficiency in sea bass*. *Frontiers in Genetics*, 10, 219. <https://doi.org/10.3389/fgene.2019.00219>

Besson, M., Komen, H., Rose, G., and Vandeputte, M. (2020). *The genetic correlation between feed conversion ratio and growth rate affects the design of a breeding program for more sustainable fish production*. *Genetics Selection Evolution*, 52(1), 1-10. <https://doi.org/10.1186/s12711-020-0524-0>

Besson, M., Rombout, N., Salou, G., Vergnet, A., Cariou, S., et al. (2022). *Potential for genomic selection on feed efficiency in gilthead sea bream (Sparus aurata), based on individual feed conversion ratio, carcass and lipid traits*. *Aquaculture Reports*, 24, 101132. <https://doi.org/10.1016/j.aqrep.2022.101132>

Bishop, S.C., and Woolliams, J.A. (2014). *Genomics and disease resistance studies in livestock*. *Livestock Science*, 166, 190-198. <https://doi.org/10.1016/j.livsci.2014.04.034>

Borrell, V., Gallego, V., García-Fernández, C., Mazzeo, I., Pérez, L., et al. (2011). *Assessment of parental contributions to fast- and slow-growing progenies in the sea bream Sparus aurata L using a new multiplex PCR*. *Aquaculture*, 314, 58–65. <https://doi.org/10.1016/j.aquaculture.2011.01.028>.

Boudry, P., Allal, F., Aslam, M.L., Bargelloni, L., Bean, T.P., et al. (2021). *Current status and potential of genomic selection to improve selective breeding in the main aquaculture species of International Council for the Exploration of the Sea (ICES) member countries*. *Aquaculture Reports*, 20, 100700. <https://doi.org/10.1016/j.aqrep.2021.100700>

Brown, R.C. (2003). *Genetic Management and Selective Breeding in Farmed Populations of Gilthead Seabream (Sparus aurata)* (Ph.D. Thesis). University of Stirling.

Brown, C., Woolliams, J.A., and McAndrew, B.J. (2005). *Factors influencing effective population size in commercial populations of gilthead seabream, Sparus aurata*. *Aquaculture*, 247, 219–225. <https://doi.org/10.1016/j.aquaculture.2005.02.002>.

Bulmer, M. (1971). *The effect of selection on genetic variability*. *The American Naturalist*, 105(943), 201-211. <https://doi.org/10.1086/282718>

- Busack, C., and Knudsen, C.M. (2007). *Using factorial mating designs to increase the effective number of breeders in fish hatcheries*. Aquaculture, 273(1), 24-32. <https://doi.org/10.1016/j.aquaculture.2007.09.010>
- Caballero, A., Santiago, E., and Toro, M.A. (1996). *Systems of mating to reduce inbreeding in selected populations*. Animal Science, 62(3), 431-442.
- Castillo-Juárez, H., Campos-Montes, G.R., Caballero-Zamora, A., and Montaldo, H.H. (2015). *Genetic improvement of Pacific white shrimp [Penaeus (Litopenaeus) vannamei]: perspectives for genomic selection*. Frontiers in Genetics, 6, 93. <https://doi.org/10.3389/fgene.2015.00093>
- Chase-Topping, M.E., Pooley, C., Moghadam, H.K., Hillestad, B., Lillehammer, M., et al. (2021). *Impact of vaccination and selective breeding on the transmission of Infectious salmon anemia virus*. Aquaculture, 535, 736365. <https://doi.org/10.1016/j.aquaculture.2021.736365>
- Chatain, B. and Chavanne, H. (2009). *La génétique du bar (Dicentrarchus labrax L.)*. Cahiers Agricultures, 18(2-3), 249-255. <https://doi.org/10.1684/agr.2009.0296>
- Chavanne, H., Janssen, K., Hofherr, J., Contini, F., Haffray, P., et al. (2016). *A comprehensive survey on selective breeding programs and seed market in the European aquaculture fish industry*. Aquaculture International, 24(5), 1287-1307. <https://doi.org/10.1007/s10499-016-9985-0>
- Colditz, I.G., and Hine, B.C. (2016). *Resilience in farm animals: biology, management, breeding and implications for animal welfare*. Animal Production Science, 56(12), 1961-1983. <https://doi.org/10.1071/AN15297>
- Correa, K., Bangera, R., Figueroa, R., Lhorente, J.P., and Yáñez, J.M. (2017). *The use of genomic information increases the accuracy of breeding value predictions for sea louse (Caligus rogercresseyi) resistance in Atlantic salmon (Salmo salar)*. Genetics Selection Evolution, 49(1), 1-5. <https://doi.org/10.1186/s12711-017-0291-8>
- Daetwyler, H.D., Villanueva, B., Bijma, P., and Woolliams, J.A. (2007). *Inbreeding in genome-wide selection*. Journal of Animal Breeding and Genetics, 124(6), 369-376. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>
- Dagnachew, B., and Meuwissen, T. (2019). *Accuracy of within-family multi-trait genomic selection models in a sib-based aquaculture breeding scheme*. Aquaculture, 505, 27-33. <https://doi.org/10.1016/j.aquaculture.2019.02.036>
- D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., et al. (2019). *Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding*. Genetics Selection Evolution, 51(1), 1-15. <https://doi.org/10.1186/s12711-019-0468-4>

Dekkers, J.C.M. (2007). *Prediction of response to marker-assisted and genomic selection using selection index theory*. Journal of Animal Breeding and Genetics, 124(6), 331-341. <https://doi.org/10.1111/j.1439-0388.2007.00701.x>

de Souza, L.H., Carvalheiro, R., Neves, H.H., and Mulder, H.A. (2019). Genetics and genomics of uniformity and resilience in livestock and aquaculture species: A review. Journal Animal Breeding Genetics, 137, 263-280. <https://doi.org/10.1111/jbg.12454>

Doeschl-Wilson, A., Anacleto, O., Nielsen, H.M., Karlsson-Drangsholt, T., Lillehammer, M., et al. (2018). *New opportunities for genetic disease control: beyond disease resistance*. Proceedings of the World Congress on Genetics Applied to Livestock Production, January 2018, Auckland, New Zealand, V Biology-Disease Resistance, 3, p. 462.

Doeschl-Wilson, A., Knap, P.W., Opriessnig, T., and More, S.J. (2021). *Livestock disease resilience: from individual to herd level*. Animal, 15, 100286. <https://doi.org/10.1016/j.animal.2021.100286>

Dufflocq, P., Lhorente, J.P., Bangera, R., Neira, R., Newman, S., et al. (2017). *Correlated response of flesh color to selection for harvest weight in coho salmon (Oncorhynchus kisutch)*. Aquaculture, 472, 38-43. <https://doi.org/10.1016/j.aquaculture.2016.08.037>

Dupont-Nivet, M., Vandeputte, M., and Chevassus, B. (2002). *Optimization of factorial mating designs for inference on heritability in fish species*. Aquaculture, 204(3-4), 361-370. [https://doi.org/10.1016/S0044-8486\(01\)00839-0](https://doi.org/10.1016/S0044-8486(01)00839-0)

Dupont-Nivet, M., Vandeputte, M., Haffray, P., and Chevassus, B. (2006). *Effect of different mating designs on inbreeding, genetic variance and response to selection when applying individual selection in fish breeding programs*. Aquaculture, 252(2-4), 161-170. <https://doi.org/10.1016/j.aquaculture.2005.07.005>

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., et al. (2008). *Heritabilities and Gx E interactions for growth in the European sea bass (Dicentrarchus labrax L.) using a marker-based pedigree*. Aquaculture, 275(1-4), 81-87. <https://doi.org/10.1016/j.aquaculture.2007.12.032>

Ek Nath, A.E., Dey, M.M., Rye, M., Gjerde, B., Abella, T.A., et al. (1998). *Selective breeding of Nile tilapia for Asia*. In 6th World Congress on Genetics Applied to Livestock Production. University of New England Armidale, Australia.

Ek Nath, A., and Doyle, R.W. (1990). *Effective population size and rate of inbreeding in aquaculture of Indian major carps*. Aquaculture, 85, 293-305. [https://doi.org/10.1016/0044-8486\(90\)90028-L](https://doi.org/10.1016/0044-8486(90)90028-L)

Faggion, S., Vandeputte, M., Chatain, B., Gagnaire, P.A., and Allal, F. (2019). *Population-specific variations of the genetic architecture of sex determination in wild European sea bass*

Dicentrarchus labrax L. Heredity, 122(5), 612-621. <https://doi.org/10.1038/s41437-018-0157-z>

Falconer, D.S., and Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics*. 4th ed. Longman Group Ltd, Edinburgh, England.

FAO (2013). *In vivo conservation of animal genetic resources*. FAO Animal Production and Health Guidelines. No. 14. Rome. <https://www.fao.org/3/i3327e/i3327e00.htm>

FAO (2020). *The state of world fisheries and aquaculture 2020*. Sustainability in action. Rome. <https://doi.org/10.4060/ca9229en>

Fernandes, T., Herlin, M., Belluga, M.D.L., Ballón, G., Martinez, P., et al. (2017). *Estimation of genetic parameters for growth traits in a hatchery population of gilthead sea bream (*Sparus aurata* L.)*. Aquaculture International, 25(1), 499-514. <https://doi.org/10.1007/s10499-016-0046-5>

Fernando, R.L., and Grossman, M. (1989). *Marker assisted selection using best linear unbiased prediction*. Genetics Selection Evolution, 21(4), 467-477.

Formoso-Rafferty, N., Cervantes, I., Ibáñez-Escriche, N., and Gutiérrez, J.P. (2016). *Genetic control of the environmental variance for birth weight in seven generations of a divergent selection experiment in mice*. Journal of Animal Breeding and Genetics, 133(3), 227-237. <https://doi.org/10.1111/jbg.12174>

Formoso-Rafferty, N., Chavez, K.N., Ojeda, C., Cervantes, I., and Gutiérrez, J.P. (2020). *Selection response in a divergent selection experiment for birth weight variability in mice compared with a control line*. Animals, 10(6), 920. <https://doi.org/10.3390/ani10060920>

Frankham, R., Ballou, S.E.J.D., Briscoe, D.A., and Ballou, J.D. (2002). *Introduction to conservation genetics*. Cambridge university press.

Fraslin, C., Dupont-Nivet, M., Haffray, P., Bestin, A., and Vandeputte, M. (2018). *How to genetically increase fillet yield in fish: new insights from simulations based on field data*. Aquaculture, 486, 175-183. <https://doi.org/10.1016/j.aquaculture.2017.12.012>

Gallardo, J.A., García, X., Lhorente, J.P., and Neira, R. (2004). *Inbreeding and inbreeding depression of female reproductive traits in two populations of Coho salmon selected using BLUP predictors of breeding values*. Aquaculture, 234, 111-122. <https://doi.org/10.1016/j.aquaculture.2004.01.009>.

Gallardo, J.A., Lhorente, J.P., and Neira, R. (2010). *The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*)*. Genetics Selection Evolution, 42(1), 1-8. <https://doi.org/10.1186/1297-9686-42-19>

García, A.L., Bosworth, B., Waldbieser, G., Misztal, I., Tsuruta, S., et al. (2018). *Development of genomic predictions for harvest and carcass weight in channel catfish*. Genetics Selection Evolution, 50(1), 1-12. <https://doi.org/10.1186/s12711-018-0435-5>

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2018). *Estrategias para reducir costes en evaluación genómica en programas de selección de acuicultura*. XIX Reunión Mejora Genética, León, España.

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2019). *Benefits from within-family genomic evaluation in aquaculture schemes with small family sizes*. XVIII Jornadas sobre Producción Animal, Zaragoza, España, 390-392.

Georges, M., Charlier, C., and Hayes, B. (2019). *Harnessing genomic information for livestock improvement*. Nature Reviews Genetics, 20(3), 135-156. <https://doi.org/10.1038/s41576-018-0082-2>

Gianola, D. (2013). *Priors in whole-genome regression: the Bayesian alphabet returns*. Genetics, 194(3), 573-596. <https://doi.org/10.1534/genetics.113.151753>

Gitterle, T., Johansen, H., Erazo, C., Lozano, C., Cock, J., et al. (2007). *Response to multi-trait selection for harvest weight, overall survival, and resistance to white spot syndrome virus (WSSV) in Penaeus (Litopenaeus) vannamei*. Aquaculture, 272, S262. <https://doi.org/10.1016/j.aquaculture.2007.07.072>

Gitterle, T., Rye, M., Salte, R., Cock, J., Johansen, H., et al. (2005). *Genetic (co) variation in harvest body weight and survival in Penaeus (Litopenaeus) vannamei under standard commercial conditions*. Aquaculture, 243(1-4), 83-92. <https://doi.org/10.1016/j.aquaculture.2004.10.015>

Gjedrem, T. (2000). *Genetic improvement of cold-water fish species*. Aquaculture Research, 31(1), 25-33. <https://doi.org/10.1046/j.1365-2109.2000.00389.x>

Gjedrem, T. (2010). *The first family-based breeding program in aquaculture*. Reviews in Aquaculture, 2(1), 2-15. <https://doi.org/10.1111/j.1753-5131.2010.01011.x>

Gjedrem, T. (2015). *Disease resistant fish and shellfish are within reach: a review*. Journal of Marine Science and Engineering, 3(1), 146-153. <https://doi.org/10.3390/jmse3010146>

Gjedrem, T., and Baranski, M. (2009). *Selective breeding in aquaculture*. Springer, Dordrecht.

Gjedrem, T., Robinson, N., and Rye, M. (2012). *The importance of selective breeding in aquaculture to meet future demands for animal protein: a review*. Aquaculture, 350, 117-129. <https://doi.org/10.1016/j.aquaculture.2012.04.008>

Gjedrem, T., and Rye, M. (2018). *Selection response in fish and shellfish: a review*. Reviews in Aquaculture, 10(1), 168-179. <https://doi.org/10.1111/raq.12154>

- Gjerde, B., Gjøen, H.M., and Villanueva, B. (1996). *Optimum designs for fish breeding programmes with constrained inbreeding Mass selection for a normally distributed trait*. *Livestock Production Science*, 47(1), 59-72. [https://doi.org/10.1016/S0301-6226\(96\)01000-7](https://doi.org/10.1016/S0301-6226(96)01000-7)
- Gjerde, B., Mengistu, S.B., and Ødegard, J. (2012). *Quantitative genetics of body weight, fillet weight and fillet yield in Nile tilapia (Oreochromis niloticus)*. *Aquaculture*, 342, 117–124. <https://doi.org/10.1016/j.aquaculture.2012.02.015>
- Gjerde, B., Terjesen, B.F., Barr, Y., Lein, I., and Thorland, I. (2004). *Genetic variation for juvenile growth and survival in Atlantic cod (Gadus morhua)*. *Aquaculture*, 236(1-4), 167-177. <https://doi.org/10.1016/j.aquaculture.2004.03.004>
- Goddard, M. (2009). *Genomic selection: prediction of accuracy and maximization of long term response*. *Genetica*, 136(2), 245-257. <https://doi.org/10.1007/s10709-008-9308-0>
- Goddard, M.E., and Hayes, B.J. (2009). *Mapping genes for complex traits in domestic animals and their use in breeding programmes*. *Nature Reviews Genetics*, 10(6), 381-391. <https://doi.org/10.1038/nrg2575>
- Golden, C.D., Koehn, J.Z., Shepon, A., Passarelli, S., Free, C.M., et al. (2021). *Aquatic foods to nourish nations*. *Nature*, 598 (7880), 315-320. <https://doi.org/10.1038/s41586-021-03917-1>
- Gonen, S., Baranski, M., Thorland, I., Norris, A., Grove, H., et al. (2015). *Mapping and validation of a major QTL affecting resistance to pancreas disease (salmonid alphavirus) in Atlantic salmon (Salmo salar)*. *Heredity*, 115(5), 405-414. <https://doi.org/10.1038/hdy.2015.37>
- Gonzalez-Pena, D., Gao, G., Baranski, M., Moen, T., Cleveland, B.M., et al. (2016). *Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (Oncorhynchus mykiss)*. *Frontiers in Genetics*, 7, 203. <https://doi.org/10.3389/fgene.2016.00203>
- Griot, R., Allal, F., Phocas, F., Brard-Fudulea, S., Morvezen, R., et al. (2021a). *Genome-wide association studies for resistance to viral nervous necrosis in three populations of European sea bass (Dicentrarchus labrax) using a novel 57k SNP array DlabChip*. *Aquaculture*, 530, 735930. <https://doi.org/10.1016/j.aquaculture.2020.735930>
- Griot, R., Allal, F., Phocas, F., Brard-Fudulea, S., Morvezen, R., et al. (2021b). *Optimization of genomic selection to improve disease resistance in two marine fishes, the European sea bass (Dicentrarchus labrax) and the gilthead sea bream (Sparus aurata)*. *Frontiers in Genetics*, 1294. <https://doi.org/10.3389/fgene.2021.665920>
- Grundy, B., Villanueva, B., and Woolliams, J.A. (1998). *Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development*. *Genetics Research*, 72(2), 159-168. <https://doi.org/10.1017/S0016672398003474>

Gutierrez, A.P., Symonds, J., King, N., Steiner, K., Bean, T.P., et al. (2019). *Potential of genomic selection for improvement of resistance to Ostreid Herpes virus in Pacific oyster (Crassostrea gigas)*. BioRxiv, 754473. <https://doi.org/10.1101/754473>

Gutiérrez, A.P., Turner, F., Gharbi, K., Talbot, R., Lowe, N.R., et al. (2017). *Development of a medium density combined-species SNP array for Pacific and European oysters (Crassostrea gigas and Ostrea edulis)*. G3: Genes, Genomes, Genetics, 7(7), 2209-2218. <https://doi.org/10.1534/g3.117.041780>

Habier, D., Fernando, R.L., and Dekkers, J. (2007). *The impact of genetic relationship information on genome-assisted breeding values*. Genetics, 177(4), 2389-2397. <https://doi.org/10.1534/genetics.107.081190>

Haffray, P., Bugeon, J., Pincent, C., Chapuis, H., Mazeiraud, E., et al. (2012). *Negative genetic correlations between production traits and head or bony tissues in large all-female rainbow trout (Oncorhynchus mykiss)*. Aquaculture, 368, 145-152. <https://doi.org/10.1016/j.aquaculture.2012.09.023>

Haffray, P., Bugeon, J., Rivard, Q., Quittet, B., Puyo, S., et al. (2013). *Genetic parameters of in-vivo prediction of carcass, head and fillet yields by internal ultrasound and 2D external imagery in large rainbow trout (Oncorhynchus mykiss)*. Aquaculture, 410, 236-244. <https://doi.org/10.1016/j.aquaculture.2013.06.016>

Harland, J. (2019). *The origins of aquaculture*. Nature Ecology & evolution, 3(10), 1378-1379. <https://doi.org/10.1038/s41559-019-0966-3>

Hayes, B.J., Visscher, P.M., and Goddard, M.E. (2009). *Increased accuracy of artificial selection by using the realized relationship matrix*. Genetics research, 91(1), 47-60. <https://doi.org/10.1017/S0016672308009981>

Hill, W.G., and Mulder, H.A. (2010). *Genetic analysis of environmental variation*. Genetics Research, 92(5-6), 381-395. <https://doi.org/10.1017/S0016672310000546>

Holtmark, M., Klemetsdal, G., Sonesson, A.K., and Woolliams, J.A. (2008). *Establishing a base population for a breeding program in aquaculture, from multiple subpopulations, differentiated by genetic drift: I. Effects of the number of subpopulations, heritability and mating strategies using optimum contribution selection*. Aquaculture, 274(2-4), 232-240. <https://doi.org/10.1016/j.aquaculture.2007.11.042>

Horn, S.S., Meuwissen, T.H.E., Moghadam, H., Hillestad, B., and Sonesson, A.K. (2020). *Accuracy of selection for omega-3 fatty acid content in Atlantic salmon fillets*. Aquaculture, 519, 734767. <https://doi.org/10.1016/j.aquaculture.2019.734767>

Houston, R.D. (2017). *Future directions in breeding for disease resistance in aquaculture species*. Revista Brasileira de Zootecnia, 46, 545-551. <https://doi.org/10.1590/S1806-92902017000600010>

- Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., et al. (2020). *Harnessing genomics to fast-track genetic improvement in aquaculture*. *Nature Reviews Genetics*, 21(7), 389-409. <https://doi.org/10.1038/s41576-020-0227-y>
- Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Tinch, A.E., et al. (2008). *Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar)*. *Genetics*, 178(2), 1109-1115. <https://doi.org/10.1534/genetics.107.082974>
- Houston, R.D., Taggart, J.B., Cézard, T., Bekaert, M., Lowe, N.R., et al. (2014). *Development and validation of a high density SNP genotyping array for Atlantic salmon (Salmo salar)*. *BMC Genomics*, 15(1), 1-13. <https://doi.org/10.1186/1471-2164-15-90>
- Ibáñez-Escriche, N., Moreno, A., Nieto, B., Piqueras, P., Salgado, C., et al. (2008a). *Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response*. *Genetics Selection Evolution*, 40(3), 279-293. <https://doi.org/10.1051/gse:2008003>
- Ibáñez-Escriche, N., Sorensen, D., Waagepetersen, R., and Blasco, A. (2008b). *Selection for environmental variation: a statistical analysis and power calculations to detect response*. *Genetics*, 180(4), 2209-2226. <https://doi.org/10.1534/genetics.108.091678>
- Janhunen, M., Kause, A., and Järvisalo, O. (2012). *Costs of being extreme—Do body size deviations from population or sire means decrease vitality in rainbow trout?*. *Aquaculture*, 370, 123-129. <https://doi.org/10.1016/j.aquaculture.2012.10.013>
- Janssen, K., Chavanne, H., Berentsen, P., and Komen, H. (2017). *Impact of selective breeding on European aquaculture*. *Aquaculture*, 472, 8-16. <https://doi.org/10.1016/j.aquaculture.2016.03.012>
- Jones, D.B., Jerry, D.R., Khatkar, M.S., Raadsma, H.W., Steen, H.V.D., et al. (2017). *A comparative integrated gene-based linkage and locus ordering by linkage disequilibrium map for the Pacific white shrimp, Litopenaeus vannamei*. *Scientific Reports*, 7(1), 1-16. <https://doi.org/10.1038/s41598-017-10515-7>
- Joshi, R., Almeida, D.B., da Costa, A.R., Skaarud, A., de Pádua Pereira, U., et al. (2021). *Genomic selection for resistance to Francisellosis in commercial Nile tilapia population: genetic and genomic parameters, correlation with growth rate and predictive ability*. *Aquaculture*, 537, 736515. <https://doi.org/10.1016/j.aquaculture.2021.736515>
- Joshi, R., Árnýasi, M., Lien, S., Gjøen, H.M., Alvarez, A.T., et al. (2018). *Development and validation of 58K SNP-array and high-density linkage map in Nile tilapia (O. niloticus)*. *Frontiers in Genetics*, 9, 472. <https://doi.org/10.3389/fgene.2018.00472>
- Joshi, R., Skaarud, A., de Vera, M., Alvarez, A.T., and Ødegård, J. (2020). *Genomic prediction for commercial traits using univariate and multivariate approaches in Nile tilapia*

(*Oreochromis niloticus*). Aquaculture, 516, 734641. <https://doi.org/10.1016/j.aquaculture.2019.734641>

Kause, A., Quinton, C., Airaksinen, S., Ruohonen, K., and Koskela, J. (2011). *Quality and production trait genetics of farmed European whitefish, Coregonus lavaretus*. Journal of Animal Science, 89(4), 959-971. <https://doi.org/10.2527/jas.2007-0332>

Kause, A., Paananen, T., Ritola, O., and Koskinen, H. (2007). *Direct and indirect selection of visceral lipid weight, fillet weight, and fillet percentage in a rainbow trout breeding program*. Journal of Animal Science, 85(12), 3218-3227. <https://doi.org/10.2527/jas.2007-0332>

Kause, A., Ritola, O., Paananen, T., Wahlroos, H., and Mantysaari, E. (2005). *Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (Oncorhynchus mykiss)*. Aquaculture, 247, 177-187. <https://doi.org/10.1016/j.aquaculture.2005.02.023>

Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A.L., et al. (2008). *Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel*. BMC Genomics, 9, 187. <https://doi.org/10.1186/1471-2164-9-187>

Khatkar, M.S., Zenger, K.R., Hobbs, M., Hawken, R.J., Cavanagh, J.A.L., et al. (2007). *A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle*. Genetics, 176, 763-772. <https://doi.org/10.1534/genetics.106.069369>

Khaw, H.L., Ponzoni, R.W., and Danting, M.J.C. (2008). *Estimation of genetic change in the GIFT strain of Nile tilapia (Oreochromis niloticus) by comparing contemporary progeny produced by males born in 1991 or in 2003*. Aquaculture, 275(1-4), 64-69. <https://doi.org/10.1016/j.aquaculture.2008.01.022>

Khaw, H.L., Ponzoni, R.W., Yee, H.Y., bin Aziz, M.A., and Bijma, P. (2016). *Genetic and non-genetic indirect effects for harvest weight in the GIFT strain of Nile tilapia (Oreochromis niloticus)*. Aquaculture, 450, 154-161. <https://doi.org/10.1016/j.aquaculture.2015.07.033>

Kijas, J., Elliot, N., Kube, P., Evans, B., Botwright, N., et al. (2017). *Diversity and linkage disequilibrium in farmed Tasmanian Atlantic salmon*. Animal Genetics, 48(2), 237-241. <https://doi.org/10.1111/age.12513>

Kjetså, M.H., Ødegård, J., and Meuwissen, T.H.E. (2020). *Accuracy of genomic prediction of host resistance to salmon lice in Atlantic salmon (Salmo salar) using imputed high-density genotypes*. Aquaculture, 526, 735415. <https://doi.org/10.1016/j.aquaculture.2020.735415>

Kjøglum, S., Grimholt, U., and Larsen, S. (2005). *Non-MHC genetic and tank effects influence disease challenge tests in Atlantic salmon (Salmo salar)*. Aquaculture, 250(1-2), 102-109. <https://doi.org/10.1016/j.aquaculture.2005.03.021>

- Kjøglum, S., Henryon, M., Aasmundstad, T., and Korsgaard, I. (2008). *Selective breeding can increase resistance of Atlantic salmon to furunculosis, infectious salmon anaemia and infectious pancreatic necrosis*. Aquaculture Research, 39(5), 498-505. <https://doi.org/10.1111/j.1365-2109.2008.01904.x>
- Knibb, W. (2000). *Genetic improvement of marine fish—which method for industry?*. Aquaculture Research, 31(1), 11-23. <https://doi.org/10.1046/j.1365-2109.2000.00393.x>
- Kriaridou, C., Tsairidou, S., Houston, R.D., and Robledo, D. (2020). *Genomic prediction using low density marker panels in aquaculture: performance across species, traits, and genotyping platforms*. Frontiers in Genetics, 11, 124. <https://doi.org/10.3389/fgene.2020.00124>
- Le Boucher, R., Vandeputte, M., Dupont-Nivet, M., Quillet, E., Ruelle, F., et al. (2013). *Genotype by diet interactions in European sea bass (*Dicentrarchus labrax* L.): Nutritional challenge with totally plant-based diets*. Journal of Animal Science, 91(1), 44-56. <https://doi.org/10.2527/jas.2012-5311>
- Lhorente, J.P., Araneda, M., Neira, R., and Yáñez, J.M. (2019). *Advances in genetic improvement for salmon and trout aquaculture: the Chilean situation and prospects*. Reviews in Aquaculture, 11(2), 340-353. <https://doi.org/10.1111/raq.12335>
- Lillehammer, M., Bangera, R., Salazar, M., Vela, S., Erazo, E.C., et al. (2020). *Genomic selection for white spot syndrome virus resistance in whiteleg shrimp boosts survival under an experimental challenge test*. Scientific Reports, 10(1), 1-13. <https://doi.org/10.1038/s41598-020-77580-3>
- Lillehammer, M., Meuwissen, T.H.E., and Sonesson, A.K. (2013). *A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values*. Genetics Selection Evolution, 45(1), 1-8. <https://doi.org/10.1186/1297-9686-45-39>
- Liu, S., Vallejo, R.L., Evenhuis, J.P., Martin, K.E., Hamilton, A., et al. (2018). *Retrospective evaluation of marker-assisted selection for resistance to bacterial cold water disease in three generations of a commercial rainbow trout breeding population*. Frontiers in Genetics, 9, 286. <https://doi.org/10.3389/fgene.2018.00286>
- López, M.E., Linderroth, T., Norris, A., Lhorente, J.P., Neira, R., et al. (2019). *Multiple selection signatures in farmed atlantic salmon adapted to different environments across hemispheres*. Frontiers in Genetics, 10, 901. <https://doi.org/10.3389/fgene.2019.00901>
- Luo, Z., Yu, Y., Bao, Z., Xiang, J., and Li, F. (2022). *Evaluation of genomic selection for high salinity tolerance traits in Pacific white shrimp *Litopenaeus vannamei**. Aquaculture, 557, 738320. <https://doi.org/10.1016/j.aquaculture.2022.738320>
- Marjanovic, J., Mulder, H.A., Khaw, H.L., and Bijma, P. (2016). *Genetic parameters for uniformity of harvest weight and body size traits in the GIFT strain of Nile tilapia*. Genetics Selection Evolution, 48(1), 1-10. <https://doi.org/10.1186/s12711-016-0218-9>

- Meuwissen, T.H.E. (1997). *Maximizing the response of selection with a predefined rate of inbreeding*. Journal of animal science, 75(4), 934-940. <https://doi.org/10.2527/1997.754934x>
- Meuwissen, T.H.E, Hayes, B.J., and Goddard, M. (2001). *Prediction of total genetic value using genome-wide dense marker maps*. Genetics, 157(4), 1819-1829. <https://doi.org/10.1093/genetics/157.4.1819>
- Meuwissen, T., Hayes, B., and Goddard, M. (2016). *Genomic selection: A paradigm shift in animal breeding*. Animal Frontiers, 6(1), 6-14. <https://doi.org/10.2527/af.2016-0002>
- Moen, T., Baranski, M., Sonesson, A.K., and Kjøglum, S. (2009). *Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar): population-level associations between markers and trait*. BMC Genomics, 10(1), 1-14. <https://doi.org/10.1186/1471-2164-10-368>
- Montaldo, H.H., Castillo-Juárez, H., Campos-Montes, G., and Pérez-Enciso, M. (2013). *Effect of the data family structure, tank replication and the statistical model, on the estimation of genetic parameters for body weight at 28 days of age in the Pacific white shrimp (Penaeus (Litopenaeus) vannamei Boone, 1931)*. Aquaculture Research, 44(11), 1715-1723. <https://doi.org/10.1111/j.1365-2109.2012.03176.x>
- Mulder, H.A., Crump, R.E., Calus, M.P.L., and Veerkamp, R.F. (2013). *Unraveling the genetic architecture of environmental variance of somatic cell score using high-density single nucleotide polymorphism and cow data from experimental farms*. Journal of Dairy Science, 96(11), 7306-7317. <https://doi.org/10.3168/jds.2013-6818>
- Nakajima, T., Hudson, M.J., Uchiyama, J., Makibayashi, K., and Zhang, J. (2019). *Common carp aquaculture in Neolithic China dates back 8,000 years*. Nature Ecology and Evolution, 3(10), 1415-1418. <https://doi.org/10.1038/s41559-019-0974-3>
- Navarro, A., Zamorano, M.J., Hildebrandt, S., Ginés, R., Aguilera, C., et al. (2009). *Estimates of heritabilities and genetic correlations for growth and carcass traits in gilthead seabream (Sparus auratus L.), under industrial conditions*. Aquaculture, 289(3-4), 225-230. <https://doi.org/10.1016/j.aquaculture.2008.12.024>
- Naylor, R.L., Hardy, R.W., Buschmann, A.H., Bush, S.R., Cao, L., et al. (2021). *A 20-year retrospective review of global aquaculture*. Nature, 591(7851), 551-563. <https://doi.org/10.1038/s41586-021-03308-6>
- Neira, R., Lhorente, J.P., Araneda, C., Díaz, N., Bustos, E., et al. (2004). *Studies on carcass quality traits in two populations of Coho salmon (Oncorhynchus kisutch): phenotypic and genetic parameters*. Aquaculture, 241(1-4), 117-131. <https://doi.org/10.1016/j.aquaculture.2004.08.009>

Nejati-Javaremi, A., Smith, C., and Gibson, J.P. (1997). *Effect of total allelic relationship on accuracy of evaluation and response to selection*. Journal of Animal Science, 75(7), 1738-1745. <https://doi.org/10.2527/1997.7571738x>

Nguyen, N.H., Phuthaworn, C., and Knibb, W. (2020). *Genomic prediction for disease resistance to Hepatopancreatic parvovirus and growth, carcass and quality traits in Banana shrimp Fenneropenaeus merguensis*. Genomics, 112(2), 2021-2027. <https://doi.org/10.1016/j.ygeno.2019.11.014>

Nguyen, N.H., Ponzoni, R.W., Abu-Bakar, K.R., Hamzah, A., Khaw, H.L., et al. (2010). *Correlated response in fillet weight and yield to selection for increased harvest weight in genetically improved farmed tilapia (GIFT strain), Oreochromis niloticus*. Aquaculture, 305(1-4), 1-5. <https://doi.org/10.1016/j.aquaculture.2010.04.007>

Nielsen, H.M., Sonesson, A.K., and Meuwissen, T.H.E. (2011). *Optimum contribution selection using traditional best linear unbiased prediction and genomic breeding values in aquaculture breeding schemes*. Journal of Animal Science, 89(3), 630-638. <https://doi.org/10.2527/jas.2009-2731>

Nielsen, H.M., Sonesson, A.K., Yazdi, H., and Meuwissen, T.H.E. (2009). *Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes*. Aquaculture, 289(3-4), 259-264. <https://doi.org/10.1016/j.aquaculture.2009.01.027>

Nirea, K.G., Sonesson, A.K., Woolliams, J.A., and Meuwissen, T.H.E. (2012). *Effect of non-random mating on genomic and BLUP selection schemes*. Genetics Selection Evolution, 44(1), 1-7. <https://doi.org/10.1186/1297-9686-44-11>

Norris, A. (2017). *Application of genomics in salmon aquaculture breeding programs by Ashie Norris: Who knows where the genomic revolution will lead us?*. Marine Genomics, 36, 13-15. <https://doi.org/10.1016/j.margen.2017.11.013>

Ødegård, J., Baranski, M., Gjerde, B., and Gjedrem, T. (2011). *Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects*. Aquaculture Research, 42, 103-114. <https://doi.org/10.1111/j.1365-2109.2010.02669.x>

Ødegård, J., Moen, T., Santi, N., Korsvoll, S.A., Kjøglum, S., et al. (2014). *Genomic prediction in an admixed population of Atlantic salmon (Salmo salar)*. Frontiers in Genetics, 5, 402. <https://doi.org/10.3389/fgene.2014.00402>

Palaiokostas, C., Cariou, S., Bestin, A., Bruant, J.S., Haffray, P., et al. (2018). *Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (Dicentrarchus labrax) using RAD sequencing*. Genetics Selection Evolution, 50(1), 1-11. <https://doi.org/10.1186/s12711-018-0401-2>

- Palaïokostas, C., Ferrarese, S., Franch, R., Houston, R.D., and Bargelloni, L. (2016). *Genomic prediction of resistance to pasteurellosis in gilthead sea bream (Sparus aurata) using 2b-RAD sequencing*. G3: Genes, Genomes, Genetics, 6(11), 3693-3700. <https://doi.org/10.1534/g3.116.035220>
- Palaïokostas, C., Vesely, T., Kocour, M., Prchal, M., Pokorova, D., et al. (2019). *Optimizing genomic prediction of host resistance to koi herpesvirus disease in carp*. Frontiers in Genetics, 10, 543. <https://doi.org/10.3389/fgene.2019.00543>
- Palti, Y., Gao, G., Liu, S., Kent, M.P., Lien, S., et al. (2015). *The development and characterization of a 57 K single nucleotide polymorphism array for rainbow trout*. Molecular Ecology Resources, 15(3), 662-672. <https://doi.org/10.1111/1755-0998.12337>
- Pante, M.A., Gjerde, B., and McMillan, I. (2001). *Effect of inbreeding on body weight at harvest in rainbow trout, Oncorhynchus mykiss*. Aquaculture, 192, 201–211. [https://doi.org/10.1016/S0044-8486\(00\)00467-1](https://doi.org/10.1016/S0044-8486(00)00467-1).
- Pante, M.J.R., Gjerde, B., McMillan, I., and Misztal, I. (2002). *Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, Oncorhynchus mykiss*. Aquaculture, 204(3-4), 383-392. [https://doi.org/10.1016/S0044-8486\(01\)00825-0](https://doi.org/10.1016/S0044-8486(01)00825-0)
- Peñaloza, C., Manousaki, T., Franch, R., Tsakogiannis, A., Sonesson, A.K., et al. (2021). *Development and testing of a combined species SNP array for the European seabass (Dicentrarchus labrax) and gilthead seabream (Sparus aurata)*. Genomics, 113(4), 2096-2107. <https://doi.org/10.1016/j.ygeno.2021.04.038>
- Peñaloza, C., Robledo, D., Barría, A., Trinh, T.Q., Mahmuddin, M., et al. (2020). *Development and validation of an open access SNP array for Nile tilapia (Oreochromis niloticus)*. G3: Genes, Genomes, Genetics, 10(8), 2777-2785. <https://doi.org/10.1534/g3.120.401343>
- Petterson, E., Guo, T.C., Evensen, Ø., Haugland, Ø., and Mikalsen, A.B. (2015). *In vitro adaptation of SAV3 in cell culture correlates with reduced in vivo replication capacity and virulence to Atlantic salmon (Salmo salar L.)*. Journal of General Virology, 96(10), 3023-3034. <https://doi.org/10.1099/jgv.0.000242>
- Poćwierz-Kotus, A., Kijewska, A., Petereit, C., Bernaś, R., Więcaszek, B., et al. (2015). *Genetic differentiation of brackish water populations of cod Gadus morhua in the southern Baltic, inferred from genotyping using SNP-arrays*. Marine Genomics, 19, 17-22. <https://doi.org/10.1016/j.margen.2014.05.010>
- Powell, J., White, I., Guy, D., and Brotherstone, S. (2008). *Genetic parameters of production traits in Atlantic salmon (Salmo salar)*. Aquaculture, 274(2-4), 225-231. <https://doi.org/10.1016/j.aquaculture.2007.11.036>

- Prchal, M., Bugeon, J., Vandeputte, M., Kause, A., Vergnet, A., et al. (2018). *Potential for genetic improvement of the main slaughter yields in common carp with in vivo morphological predictors*. *Frontiers in Genetics*, 9, 283. <https://doi.org/10.3389/fgene.2018.00283>
- Prchal, M., Kocour, M., Vandeputte, M., Kause, A., Vergnet, A., et al. (2020). *Morphological predictors of slaughter yields using 3D digitizer and their use in a common carp breeding program*. *Aquaculture*, 520, 734993. <https://doi.org/10.1016/j.aquaculture.2020.734993>
- Prchal, M., Zhao, J., Gela, D., Kašpar, J., Lepič, P., et al. (2021). *Simplified method for genetic slaughter yields improvement in common carp under European pond conditions*. *Aquaculture Reports*, 21, 100832. <https://doi.org/10.1016/j.aqrep.2021.100832>
- Qi, H., Song, K., Li, C., Wang, W., Li, B., et al. (2017). *Construction and evaluation of a high-density SNP array for the Pacific oyster (*Crassostrea gigas*)*. *PLoS One*, 12(3), e0174007. <https://doi.org/10.1371/journal.pone.0174007>
- Regan, T., Bean, T.P., Ellis, T., Davie, A., Carboni, S., et al. (2021). *Genetic improvement technologies to support the sustainable growth of UK aquaculture*. *Reviews in Aquaculture*, 13(4), 1958-1985. <https://doi.org/10.1111/raq.12553>
- Rutten, M.J., Bovenhuis, H., and Komen, H. (2005). *Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.)*. *Aquaculture*, 246(1-4), 125-132. <https://doi.org/10.1016/j.aquaculture.2005.01.006>
- Sae-Lim, P., Gjerde, B., Nielsen, H.M., Mulder, H., and Kause, A. (2016). *A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species*. *Reviews in Aquaculture*, 8(4), 369-393. <https://doi.org/10.1111/raq.12098>
- Sae-Lim, P., Kause, A., Janhunen, M., Vehviläinen, H., Koskinen, H., et al. (2015). *Genetic (co) variance of rainbow trout (*Oncorhynchus mykiss*) body weight and its uniformity across production environments*. *Genetics Selection Evolution*, 47(1), 1-10. <https://doi.org/10.1186/s12711-015-0122-8>
- Sae-Lim, P., Kause, A., Lillehammer, M., and Mulder, H.A. (2017). *Estimation of breeding values for uniformity of growth in Atlantic salmon (*Salmo salar*) using pedigree relationships or single-step genomic evaluation*. *Genetics Selection Evolution*, 49(1), 1-12. <https://doi.org/10.1186/s12711-017-0308-3>
- Saillant, E., Dupont-Nivet, M., Sabourault, M., Haffray, P., Laureau, S., et al. (2009). *Genetic variation for carcass quality traits in cultured sea bass (*Dicentrarchus labrax*)*. *Aquatic Living Resources*, 22(1), 105-112. <https://doi.org/10.1051/alr/2009010>
- Salem, M., Al-Tobasei, R., Ali, A., Lourenco, D., Gao, G., et al. (2018). *Genome-wide association analysis with a 50K transcribed gene SNP-chip identifies QTL affecting muscle yield in rainbow trout*. *Frontiers in Genetics*, 9, 387. <https://doi.org/10.3389/fgene.2018.00387>

SanCristobal-Gaudy, M., Elsen, J.M., Bodin, L., and Chevalet, C. (1998). *Prediction of the response to a selection for canalisation of a continuous trait in animal breeding*. *Genetics Selection Evolution*, 30, 423–451.

Saura, M., Caballero, A., Santiago, E., Fernández, A., Morales-González, E., et al. (2021). *Estimates of recent and historical effective population size in turbot, seabream, seabass and carp selective breeding programmes*. *Genetics Selection Evolution*, 53(1), 1-8. <https://doi.org/10.1186/s12711-021-00680-9>

Saura, M., Carabaño, M.J., Fernández, A., Cabaleiro, S., Doeschl-Wilson, A.B., et al. (2019). *Disentangling genetic variation for resistance and endurance to scuticociliatosis in turbot using pedigree and genomic information*. *Frontiers in Genetics*, 10, 539. <https://doi.org/10.3389/fgene.2019.00539>

Saura, M., Fernández, A., Maroso, F., Hermida, M., Martínez, P., et al. (2018). *Development and applications of new genomic tools for turbot selection programmes*. In *Proceedings of the World Congress on Genetics Applied to Livestock Production* (Vol. 11, p. 482).

Scheiner, S.M., Caplan, R.L., and Lyman, R.F. (1991). *The genetics of phenotypic plasticity. III. Genetic correlations and fluctuating asymmetries*. *Journal of Evolutionary Biology*, 4(1), 51-68. <https://doi.org/10.1046/j.1420-9101.1991.4010051.x>

Šegvić-Bubić, T., Grubišić, L., Trumbić, Ž., Stanić, R., Ljubković, J., et al. (2017). *Genetic characterization of wild and farmed European seabass in the Adriatic sea: assessment of farmed escapees using a Bayesian approach*. *ICES Journal of Marine Science*, 74(1), 369-378. <https://doi.org/10.1093/icesjms/fsw155>.

Sell-Kubiak, E., Wang, S., Knol, E.F., and Mulder, H.A. (2015). *Genetic analysis of within-litter variation in piglets' birth weight using genomic or pedigree relationship matrices*. *Journal of Animal Science*, 93(4), 1471-1480. <https://doi.org/10.2527/jas.2014-8674>

Slatkin, M. (2008). *Linkage disequilibrium—understanding the evolutionary past and mapping the medical future*. *Nature Reviews Genetics*, 9(6), 477-485. <https://doi.org/10.1038/nrg2361>

Solberg, T.R., Sonesson, A.K., Woolliams, J.A., and Meuwissen, T.H.E. (2008). *Genomic selection using different marker types and densities*. *Journal of Animal Science*, 86(10), 2447-2454. <https://doi.org/10.2527/jas.2007-0010>

Solberg, M.F., Zhang, Z., Nilsen, F., and Glover, K.A. (2013). *Growth reaction norms of domesticated, wild and hybrid Atlantic salmon families in response to differing social and physical environments*. *BMC Evolutionary Biology*, 13(1), 1-23. <https://doi.org/10.1186/1471-2148-13-234>

Sonesson, A.K. (2007). *Within-family marker-assisted selection for aquaculture species*. *Genetics Selection Evolution*, 39(3), 1-17. <https://doi.org/10.1186/1297-9686-39-3-301>

Sonesson, A.K., and Meuwissen, T.H.E. (2000). *Mating schemes for optimum contribution selection with constrained rates of inbreeding*. Genetics Selection Evolution, 32(3), 231-248. <https://doi.org/10.1051/gse:2000116>

Sonesson, A.K., and Meuwissen, T.H.E. (2009). *Testing strategies for genomic selection in aquaculture breeding programs*. Genetics Selection Evolution, 41(1), 1-9. <https://doi.org/10.1186/1297-9686-41-37>

Sonesson, A.K., and Ødegård, J. (2016). *Mating structures for genomic selection breeding programs in aquaculture*. Genetics Selection Evolution, 48(1), 1-7. <https://doi.org/10.1186/s12711-016-0224-y>

Sonesson, A. K., Ødegård, J., and Rönnegård, L. (2013). *Genetic heterogeneity of within-family variance of body weight in Atlantic salmon (Salmo salar)*. Genetics Selection Evolution, 45(1), 1-8. <https://doi.org/10.1186/1297-9686-45-41>

Song, H., and Hu, H. (2022). *Strategies to improve the accuracy and reduce costs of genomic prediction in aquaculture species*. Evolutionary Applications, 15(4), 578-590. <https://doi.org/10.1111/eva.13262>

Sørensen, A.C., Berg, P., and Woolliams, J.A. (2005). *The advantage of factorial mating under selection is uncovered by deterministically predicted rates of inbreeding*. Genetics Selection Evolution, 37(1), 57-81. <https://doi.org/10.1051/gse:2004036>

Su, G.-S., Liljedahl, L.-E., and Gall, G.A.E. (1996). *Effects of inbreeding on growth and reproductive traits in rainbow trout (Oncorhynchus mykiss)*. Aquaculture, 142, 139–148. [https://doi.org/10.1016/0044-8486\(96\)01255-0](https://doi.org/10.1016/0044-8486(96)01255-0).

Teletchea, F. (2019). *Animal domestication: A brief overview*. IntechOpen.

Teletchea, F., and Fontaine, P. (2014). *Levels of domestication in fish: implications for the sustainable future of aquaculture*. Fish and Fisheries, 15(2), 181-195. <https://doi.org/10.1111/faf.12006>

Thorland, I., Kottaras, L., Refstie, T., Dimitroglou, A., Papaharisis, L., et al. (2015). *Response to selection for harvest weight in a family based selection program of gilthead seabream (Sparus aurata)*. In XII International Symposium on Genetics in Aquaculture, Santiago de Compostela, Spain.

Thorland, I., Papaioannou, N., Kottaras, L., Refstie, T., Papasolomontos, S., et al. (2007). *Family based selection for production traits in gilthead seabream (Sparus aurata) and European sea bass (Dicentrarchus labrax) in Greece*. Aquaculture, (272), S314. <https://doi.org/10.1016/j.aquaculture.2007.07.193>

Toro, M., and Pérez-Enciso, M. (1990). *Optimization of selection response under restricted inbreeding*. Genetics Selection Evolution, 22(1), 93-107.

- Tsai, H.Y., Hamilton, A., Tinch, A.E., Guy, D.R., Bron, J.E., et al. (2016). *Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations*. *Genetics Selection Evolution*, 48(1), 1-11. <https://doi.org/10.1186/s12711-016-0226-9>
- Tsai, H.Y., Hamilton, A., Tinch, A.E., Guy, D.R., Gharbi, K., et al. (2015). *Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array*. *BMC Genomics*, 16(1), 1-9. <https://doi.org/10.1186/s12864-015-2117-9>
- Tsai, H.Y., Matika, O., Edwards, S.M., Antolín-Sánchez, R., Hamilton, A., et al. (2017). *Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon*. *G3: Genes, Genomes, Genetics*, 7(4), 1377-1383. <https://doi.org/10.1534/g3.117.040717>
- Tsairidou, S., Anacleto, O., Woolliams, J.A., and Doeschl-Wilson, A. (2019). *Enhancing genetic disease control by selecting for lower host infectivity and susceptibility*. *Heredity*, 122(6), 742-758. <https://doi.org/10.1038/s41437-018-0176-9>
- Tsairidou, S., Hamilton, A., Robledo, D., Bron, J.E., and Houston, R.D. (2020). *Optimizing low-cost genotyping and imputation strategies for genomic selection in Atlantic salmon*. *G3: Genes, Genomes, Genetics*, 10(2), 581-590. <https://doi.org/10.1534/g3.119.400800>
- Vallejo, R.L., Cheng, H., Fragomeni, B.O., Gao, G., Silva, R.M., et al. (2021). *The accuracy of genomic predictions for bacterial cold water disease resistance remains higher than the pedigree-based model one generation after model training in a commercial rainbow trout breeding population*. *Aquaculture*, 545, 737164. <https://doi.org/10.1016/j.aquaculture.2021.737164>
- Vallejo, R.L., Fragomeni, B.O., Cheng, H., Gao, G., Long, R.L., et al. (2020). *Assessing accuracy of genomic predictions for resistance to infectious hematopoietic necrosis virus with progeny testing of selection candidates in a commercial rainbow trout breeding population*. *Frontiers in Veterinary Science*, 7, 590048. <https://doi.org/10.3389/fvets.2020.590048>
- Vallejo, R.L., Leeds, T.D., Fragomeni, B.O., Gao, G., Hernandez, A.G., et al. (2016). *Evaluation of genome-enabled selection for bacterial cold water disease resistance using progeny performance data in rainbow trout: insights on genotyping methods and genomic prediction models*. *Frontiers in Genetics*, 7, 96. <https://doi.org/10.3389/fgene.2016.00096>
- Vallejo, R.L., Leeds, T.D., Gao, G., Parsons, J.E., Martin, K.E., et al. (2017). *Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture*. *Genetics Selection Evolution*, 49(1), 1-13. <https://doi.org/10.1186/s12711-017-0293-6>
- Vallejo, R.L., Silva, R.M., Evenhuis, J.P., Gao, G., Liu, S., et al. (2018). *Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels:*

evidence that long-range LD is a major contributing factor. Journal of Animal Breeding and Genetics, 135(4), 263-274. <https://doi.org/10.1111/jbg.12335>

Vandeputte, M., Bugeon, J., Bestin, A., Desgranges, A., Allamellou, J.M., et al. (2019). *First evidence of realized selection response on fillet yield in rainbow trout *Oncorhynchus mykiss*, using sib selection or based on correlated ultrasound measurements.* Frontiers in Genetics, 10, 1225. <https://doi.org/10.3389/fgene.2019.01225>

Vandeputte, M., Dupont-Nivet, M., Haffray, P., Chavanne, H., Cenadelli, S., et al. (2009). *Response to domestication and selection for growth in the European sea bass (*Dicentrarchus labrax*) in separate and mixed tanks.* Aquaculture, 286(1-2), 20-27. <https://doi.org/10.1016/j.aquaculture.2008.09.008>

Vandeputte, M., Fraslin, C., Haffray, P., Bestin, A., Allal, F., et al. (2020). *How to genetically increase fillet yield in fish: Relevant genetic parameters and methods to predict genetic gain.* Aquaculture, 519, 734877. <https://doi.org/10.1016/j.aquaculture.2019.734877>

Vandeputte, M., Puleda, A., Tyran, A.S., Bestin, A., Coulombet, C., et al. (2017). *Investigation of morphological predictors of fillet and carcass yield in European sea bass (*Dicentrarchus labrax*) for application in selective breeding.* Aquaculture, 470, 40-49. <https://doi.org/10.1016/j.aquaculture.2016.12.014>

VanRaden, P.M. (2008). *Efficient estimation to compute genomic predictions.* Journal Dairy Science, 901, 4414-4423. <https://doi.org/10.3168/jds.2007-0980>

Villanueva, B., Fernández, A., Peiró-Pastor, R., Peñaloza, C., Houston, R.D., et al. (2022). *Population structure and genetic variability in wild and farmed Mediterranean populations of gilthead seabream and European seabass inferred from a 60K combined species SNP array.* Aquaculture Reports, 24, 101145. <https://doi.org/10.1016/j.aqrep.2022.101145>

Villanueva, B., Fernández, J., García-Cortés, L.A., Varona, L., Daetwyler, H.D., et al. (2011). *Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs.* Journal of Animal Science, 89(11), 3433-3442. <https://doi.org/10.2527/jas.2010-3814>

Villanueva, B., Pong-Wong, R., Fernandez, J., and Toro, M.A. (2005). *Benefits from marker-assisted selection under an additive polygenic genetic model.* Journal of Animal Science, 83(8), 1747-1752. <https://doi.org/10.2527/2005.8381747x>

Villanueva, B., Woolliams, J.A., and Gjerde, B. (1996). *Optimum designs for breeding programmes under mass selection with an application in fish breeding.* Animal Science, 63(3), 563-576. <https://doi.org/10.1017/S1357729800015459>

Villanueva, B., Woolliams, J.A., and Simm, G. (1994). *Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle.* Genetics Selection Evolution, 26(6), 517-535.

Waddington, C.H. (1960). Experiments on canalizing selection. *Genetics Resources*, 1, 140–150.

Wang, Q., Yu, Y., Zhang, Q., Zhang, X., Huang, H., et al. (2019). *Evaluation on the genomic selection in Litopenaeus vannamei for the resistance against Vibrio parahaemolyticus*. *Aquaculture*, 505, 212-216. <https://doi.org/10.1016/j.aquaculture.2019.02.055>

Wang, J., Zhou, Z., Zhang, Z., Li, H., Liu, D., et al. (2018). *Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits*. *Heredity*, 121(6), 648-662. <https://doi.org/10.1038/s41437-018-0075-0>

Wientjes, Y.C., Bijma, P., Calus, M.P., Zwaan, B.J., Vitezica, Z.G., et al. (2022). *The long-term effects of genomic selection: 1. Response to selection, additive genetic variance, and genetic architecture*. *Genetics Selection Evolution*, 54(1), 1-21. <https://doi.org/10.1186/s12711-022-00709-7>

Woolliams, J.A. (1989). *Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle*. *Animal Science*, 49(1), 1-14.

Wray, N.R., and Hill, W.G. (1989). *Asymptotic rates of response from index selection*. *Animal Science*, 49(2), 217-227. <https://doi.org/10.1017/S0003356100032347>

Xu, J., Zhao, Z., Zhang, X., Zheng, X., Li, J., et al. (2014). *Development and evaluation of the first high-throughput SNP array for common carp (Cyprinus carpio)*. *BMC Genomics*, 15(1), 1-10. <https://doi.org/10.1186/1471-2164-15-307>

Yáñez, J.M., Bassini, L.N., Filp, M., Lhorente, J.P., Ponzoni, R., et al. (2014). *Inbreeding and effective population size in a coho salmon (Oncorhynchus kisutch) breeding nucleus in Chile*. *Aquaculture*, 420–421, S15–S19. <https://doi.org/10.1016/j.aquaculture.2013.05.028>.

Yáñez, J.M., Naswa, S., López, M.E., Bassini, L., Correa, K., et al. (2016). *Genomewide single nucleotide polymorphism discovery in Atlantic salmon (Salmo salar): validation in wild and farmed American and European populations*. *Molecular Ecology Resources*, 16(4), 1002-1011. <https://doi.org/10.1111/1755-0998.12503>

Yáñez, J.M., Yoshida, G., Barria, A., Palma-Véjares, R., Travisany, D., et al. (2020). *High-throughput single nucleotide polymorphism (SNP) discovery and validation through whole-genome resequencing in Nile tilapia (Oreochromis niloticus)*. *Marine Biotechnology*, 22(1), 109-117. <https://doi.org/10.1007/s10126-019-09935-5>

Yoshida, G.M., Carvalheiro, R., Lhorente, J.P., Correa, K., Figueroa, R., et al. (2018). *Accuracy of genotype imputation and genomic predictions in a two-generation farmed Atlantic salmon population using high-density and low-density SNP panels*. *Aquaculture*, 491, 147-154. <https://doi.org/10.1016/j.aquaculture.2018.03.004>

Yoshida, G.M., Carvalheiro, R., Rodríguez, F.H., Lhorente, J.P., and Yáñez, J.M. (2019c). *Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout*. *Genomics*, 111(2), 127-132. <https://doi.org/10.1016/j.ygeno.2018.01.008>

Yoshida, G.M., Barria, A., Correa, K., Cáceres, G., Jedlicki, A., et al. (2019a). *Genome-wide patterns of population structure and linkage disequilibrium in farmed Nile tilapia (*Oreochromis niloticus*)*. *Frontiers in Genetics*, 10, 745. <https://doi.org/10.3389/fgene.2019.00745>

Yoshida, G.M., de Oliveira, C.A.L., Campos, E.C., Todesco, H., Araújo, F.C., et al. (2022). *A breeding program for Nile tilapia in Brazil: Results from nine generations of selection to increase the growth rate in cages*. *Journal of Animal Breeding and Genetics*, 139(2), 127-135. <https://doi.org/10.1111/jbg.12650>

Yoshida, G.M., Lhorente, J.P., Correa, K., Soto, J., Salas, D., et al. (2019b). *Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in Nile tilapia (*Oreochromis niloticus*)*. *G3: Genes, Genomes, Genetics*, 9(8), 2597-2607. <https://doi.org/10.1534/g3.119.400116>

Zhao, J., Prchal, M., Palaiokostas, C., Houston, R.D., Kause, A., et al. (2020). *Genetic relationship between koi herpesvirus disease resistance and production traits inferred from sibling performance in Amur mirror carp*. *Aquaculture*, 520, 734986. <https://doi.org/10.1016/j.aquaculture.2020.734986>

Zenger, K.R., Khatkar, M.S., Jones, D.B., Khalilisamani, N., Jerry, D.R., et al. (2019). *Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters*. *Frontiers in Genetics*, 693. <https://doi.org/10.3389/fgene.2018.00693>



ANNEX I: MANUAL GAINFISH



**Financed by: EU-FP7, grant (KBBE.2013.1.2-10)
under grant agreement nº 613611(FISHBOOST)**



Version 1.0

26th of April 2019

Jesús Fernández and Silvia García-Ballesteros



INDEX

Introduction	3
Installation	3
Programs.....	3
<i>Genera_pop</i>	3
<i>Subdiv_pop</i>	7
<i>Breeding program software</i>	11
Mass_select	14
BLUP	14
GBLUP	14
WF_GBLUP	15
Known limitations	17
References	17
Liability	17

Introduction

This suite of programs allows to simulate different aquaculture breeding programs and compare them in terms of gain and inbreeding. A first tool (*genera_pop*), generates a population in mutation-drift equilibrium with a realistic linkage disequilibrium (LD) distribution. A second one (*subdiv_pop*), creates as many replicates as desired of one or several populations (when appropriate) to be the origin of the base population from which the selection program will start; additionally, at this stage the trait(s) to be simulated are defined. The rest of tools will simulate the breeding program itself under truncation selection on phenotypic records (*mass_selec*), BLUP estimated breeding values (BLUP), population-wide genomic breeding values (GBLUP), or within-family genomic breeding values (WF_GBLUP).

Installation

Just uncompress the distribution file (*gainfish_win.zip* or *gainfish_linux.zip*) into a folder. Besides this manual and the example files, six executable files will appear, corresponding to each of the programs conforming the package. Programs will search for input files in the local folder, so they must be there if you want everything to run properly. Compiled executables are provided for both Windows and Linux operative systems.

Programs

Genera_pop

This program creates a population in mutation-drift equilibrium by simulating a large number of discrete generations under random mating with a specific mutation rate (see for example, Fernández et al., 2014).

The user can define the size of the population (males and females separately), the mutation rate for each locus and the number of generations the simulation will run.

Regarding the genome structure, the user has to provide the number of chromosomes and the total genome length in base pairs. **Hint: in this version the software assumes a rate of 1 cM per Mb and, thus, if you want to simulate a particular number of Morgans you should enter this value multiplied by 10^8 .** Chromosomes can be all of the same length or have different sizes. In the latter situation the user has to provide the particular distribution of lengths by chromosome in an extra line of the input file.

A variable number of biallelic loci can be simulated differentiating between marker loci (SNPs) and non-marker loci. SNP genotypes will be assumed to be known and used for genomic selection while the non-markers will be used for the calculation of different parameters (e.g., heterozygosity). Both types of loci may harbour loci with effect on the trait(s) simulated. The distribution of SNPs across chromosomes can be also defined by the user (adding an extra line with the precise number per chromosome in the input file). In fact, if a map is available for the particular species it can be provided to make the simulation to better mimic the real genome structure. Non-marker loci are always distributed across chromosomes in a proportional pattern (i.e. the longer the chromosome the higher the number of loci). At the beginning all loci will have an allelic frequency of 0.5, unless the user prefer to use particular starting genotypes (for example coming from a previous simulation) which will be entered from another file.

The program will search for *genera_pop_input.txt* (compulsory file) whose format is as follows:

- [*genera_pop_input.txt*]

This file corresponds to the example file provided with the software which generates a genome of 15 M and 24 chromosomes of equal length. A total of 150,000 SNPs plus 50,000 non-marker loci are evenly distributed across the genome.

The text file has data separated by blanks (at least one), tabs or commas; the program reads data in free format.

```
500 500 24 4000 ! No. individuals (males and females), no. chromosomes, no. generations
150000 50000 ! Total no. SNPs and non-marker loci
0.0025 ! Mutation rate
1 ! (1) All chromosomes of the same length (0) Different chromosome lengths. If option (0) is
   selected the file should include an extra line with the length of all chromosomes
15000000000 ! Total genome length in bases (1 cM = 1 Mb)
1 ! (1) proportional distribution (0) given distribution. If option (0) is selected the file should
   include an extra line with the number of SNPs per chromosome
0 ! (0) No map provided - evenly spaced SNPs (1) Map provided
0 ! (0) De novo creation of genotypes (1) Genotypes from file
```

Additional (non compulsory files) that can be used are:

- [*real_map.txt*]

When the precise position (physical map) of the SNPs to be simulated are known (option 'Map provided') they can be introduced in this file. The structure will include as many rows as chromosomes. In each line the position of each SNP will appear:

```
... 13527 13594 14022 14151 ...
... 1077 2055 2130 2485 ...
... 227436 227500 227904 228099 ...
```

- [*initial_genotypes.txt*]

The structure of the file is the following:

⇒ *individual id, chromosome, genotypes for this chromosome (first allele first locus, second allele first locus, first allele second locus, second allele second locus, ...).*

```
1 1 0 1 0 1 0 0 0 0 1 0 ...
1 2 1 1 1 1 0 1 1 0 1 0 ...
1 3 1 0 1 0 0 0 0 0 0 1 ...
⋮
1 24 1 0 1 1 0 0 0 0 0 1 ...
2 1 0 1 1 1 0 0 1 1 1 0 ...
⋮
2 24 1 0 1 1 0 1 0 0 1 1 ...
3 1 1 0 0 0 0 1 1 1 0 0 ...
⋮
```

The program reads data in free format, so there is no need of genotypes to be in columns (although it looks nicer).

Hint: the file should include the genotype for all loci (SNPs and non-markers).

The software will produce two kinds of output. Firstly, each generation the file [*results.txt*] will show the value for a set of parameters (mean allelic frequencies, variance of allelic frequencies, ...) for the user to see how the simulation proceed and if the mutation-drift equilibrium has been already reached (for testing purposes). Values are shown for the whole loci and for SNPs and non-marker loci separately. The file has the following structure:

```

Fr_tot Fr_neu Fr_SNP Va_tot Va_neu Va_SNP ⇒ allelic frequencies and their
variances
0 0.5000 0.4999 0.5000 0.0001 0.0001 0.0001
1 0.5000 0.4999 0.5000 0.0002 0.0002 0.0002
2 0.5000 0.4999 0.5001 0.0004 0.0004 0.0004
3 0.5000 0.4999 0.5000 0.0005 0.0005 0.0005
4 0.5000 0.4998 0.5000 0.0006 0.0006 0.0006
5 0.4999 0.4997 0.5000 0.0007 0.0007 0.0007
He_tot Ho_tot He_neu Ho_neu He_SNP Ho_SNP ⇒ expected and observed
heterozygosities
0 0.4999 0.5000 0.4998 0.5000 0.4999 0.5000
1 0.4998 0.4999 0.4996 0.4998 0.4997 0.5000
2 0.4996 0.4999 0.4993 0.4999 0.4995 0.4999
3 0.4993 0.4994 0.4991 0.4995 0.4992 0.4993
4 0.4991 0.4992 0.4988 0.4993 0.4990 0.4992
5 0.4989 0.4991 0.4986 0.4991 0.4988 0.4992
Va_He_t Va_Ho_t Va_He_n Va_Ho_n Va_He_S Va_Ho_S ⇒ variances of the
heterozygosities
0 -0.0000 0.0002 -0.0000 0.0003 -0.0001 0.0002
1 -0.0001 0.0002 -0.0000 0.0003 -0.0001 0.0002
2 -0.0002 0.0002 -0.0000 0.0003 -0.0001 0.0002
3 -0.0002 0.0002 -0.0000 0.0003 -0.0001 0.0002
4 -0.0002 0.0002 -0.0000 0.0003 -0.0001 0.0002
5 -0.0002 0.0002 -0.0000 0.0003 -0.0001 0.0002
seg_tot seg_neu seg_SNP ⇒ segregating loci
0 200000 50000 150000
1 200000 50000 150000
2 200000 50000 150000
3 200000 50000 150000
4 200000 50000 150000
5 200000 50000 150000

```

A second group of output files will be generated just once and will be inputs for other programs of the GAINFISH collection.

- [*parameters.txt*] will include the basic parameters used to generate the population:

```

1000          24 ⇒ used population size and chromosome number
150000       50000 ⇒ total number of SNPs and non markers
2.499999994E-03 ⇒ mutation rate

```

- [*maps.txt*] will show the genome structure with the position of markers in each chromosome

```

Chromosome 1 Length (pb) No.SNP No.neut ⇒ length and number of loci
Chromosome 2 62500000 6250 2084

```

```
Chromosome 23      62500000    6250    2083
Chromosome 24      62500000    6250    2083
```

```
Chromosome          1
  1          9998 ⇒ position of SNPs
  2         19997
  3         29995
```

```
⋮
  8330      62470004
  8331      62480004
  8332      62490004
```

```
Chromosome          2
  1          9998
  2         19997
  3         29995
```

```
⋮
  8330      62470004
  8331      62480004
  8332      62490004
```

```
Chromosome          3
  1          9998
  2         19997
  3         29995
```

```
⋮
```

- [*genotypes.txt*] will store the genotypes of the individuals of the population at the final generation. The structure of the file is the same as explained before:

⇒ *individual id, chromosome, genotypes for this chromosome (first allele first locus, second allele first locus, first allele second locus, second allele second locus, ...).*

```
1 1 0 1 0 1 0 0 0 0 1 0 ...
1 2 1 1 1 1 0 1 1 0 1 0 ...
1 3 1 0 1 0 0 0 0 0 0 1 ...
⋮
1 24 1 0 1 1 0 0 0 0 0 1 ...
2  1 0 1 1 1 0 0 1 1 1 0 ...
⋮
2 24 1 0 1 1 0 1 0 0 1 1 ...
3  1 1 0 0 0 0 1 1 1 0 0 ...
⋮
```

Hint: this file will be generated only at the very end of the simulation. Therefore, even if you check that equilibrium has been reached you have to wait for the total number of generations to run. Put a not too high number or you will waste computing time. If the population is not at equilibrium after all generations you can use the output as the *initial_genotypes* file to go on running the simulation from the same point.

Subdiv_pop

This is a dual purpose program. It creates one or several subpopulations (strains) from the population at equilibrium. These subpopulations will be the source for generating the base population (generation zero) of our breeding program (as in Fernández et al., 2014). Also, at this stage one or several quantitative traits are defined with a particular genetic model. The program will provide different files for as many replicates as required for the next step (i.e. the simulation of the breeding program).

The strain(s) can be created following three different regimes: i) random selection; ii) artificial selection (to mimic already improved strains); and iii) stabilizing selection (to mimic wild strains with local adaptations in nature). These can be run for several generations. **Hint: At least one generation has to be defined; this can be used to expand/collapse to the number of individuals to be used in the base population (founders).**

Regarding the traits, the software allows to assume different genetic models. A variable number of loci with additive and dominant (typically small) effects can be defined. Additionally, one or several loci of large effects can be included. The parameters for each of the traits (heritabilities, phenotypic and genetic correlations) can also be determined by the user.

Besides the files generated by the application **genera_pop** (*parameters.txt*, *maps.txt* and *genotypes.txt*) **subdiv_pop** requires two input files: one for defining the structure of the populations and the way to create them, and another to define the genetic model of the trait(s).

- [*subdiv_pop_input.txt*] should have the following information:

```
3 20 100 ! No. subpopulations, no. generations, no. replicates
30 30 ! Structure (males and females) of subpopulation 1
30 30 ! Structure (males and females) of subpopulation 2
10 10 ! Structure (males and females) of subpopulation 3
0 30 30 ! Type of selection for subpopulation 1 (random)
2 6 6 ! Type of selection for subpopulation 2 (artificial selection)
1 10 5 ! Type of selection for subpopulation 3 (stabilising selection)
3 ! No. traits to be created
```

The final size of the population(s) is the size determined in the second and following lines, irrespective of the size of the subpopulation along the process (see different regimes of the strains later on). These values should be in concordance with the number of individuals desired for the base population. When different strains have been generated a synthetic base population can be optimally constructed using the software **BASEPOP** (<http://www.fishboost.eu/fishboost-tools.html>).

The subpopulations can evolve under three different processes:

- Random contributions of candidates: Drift is the only source of modification of the genetic composition, depending on the size of the subpopulation. In the input file this is coded by including a line with 0 30 30, where zero is the code for Random and the other two values are the number of males and females across generations.
- Stabilising selection: The probability of survival of an individual with phenotypic value P is modelled using a Gaussian distribution (Turelli, 1984; Bürger et al., 1989):

$$W(P) = \exp\left(-\frac{(P - P_{opt})^2}{2\omega^2}\right)$$

where P_{opt} is the optimum phenotype in a particular environment and ω^2 is an inverse measure of the strength of stabilizing selection. The codification in this situation is 1 10 5 where one is for Stabilising selection, 10 is the optimal phenotype (P_{opt}) and 5 is the selection strength (ω^2).

- Artificial selection: Each generation the males and females with the best phenotypes are used for creating the next generation. Selected numbers are defined in the line for this subpopulation as 2 6 6 (artificial selection code, selected males and selected females).

Hint: In the two scenarios involving selection (directional or stabilising) only the first trait is used (if several are defined).

- [trait_input_mv.txt] harbours all the parameters needed to define the trait(s). The structure of the file is:

```
1000 ! No. selective loci
0 ! Position of selective loci: (2) user provided (1) evenly spaced (0) randomly distributed
.3 .0 0. 1. 1 ! Parameters Trait 1
.55 .0 0. 0.15 1. 0 ! Parameters Trait 2
.25 .0 0. 0.3 0.9 1. 1 ! Parameters Trait 3
1 ! Type of selective loci Trait 1: (2) Bimodal, some with large effect (1) normal distributed (0)
Equal effect
1 ! Type of selective loci Trait 2
1 ! Type of selective loci Trait 3
1. ! Environmental correlations and variances Trait 1
0.247786293 1. ! Environmental correlations and variances Trait 2
0.438663106 0.37227298 1. ! Environmental correlations and variances Trait 3
```

To allow for correct genetic correlations between all traits the same loci have effect on all of them but with different magnitude. **Hint: This is also true in this version for the loci with large effect, which may be not very realistic.**

When selective loci are randomly distributed they may be allocated in SNPs as well as in non-marker loci. If a particular distribution of selective loci is to be used (option 2; user provided) a file [def_selec_loci.txt] should be created with the following structure:

```
16      5232
21      3488
15      5658
19      5709
2       1560
!
```

where the first column indicates the chromosome and the second the position (in terms of the total number of loci in each chromosome).

The next lines in the file *trait_input_mv.txt* introduce the parameters for each of the traits to be simulated. The first number is the heritability of the trait and the second is the dominance part (d^2). The third value is the phenotypic mean. Then the genetic covariances (only the lower

triangle of the matrix) and the phenotypic variance. The last number in each line states if the trait can be measured in all individuals (1) or only in some of them (0), for example for disease resistance or carcass traits.

The following lines establish the distribution of effects of loci for each trait: (0) if all loci are to be simulated with the same effect, (1) if effects are normally distributed and (2) if there are some loci with large effect. In the latter scenario an extra line must be added with two figures: the number of loci with large effect and the percentage of the genetic variance explained by them. **Hint: The program assumes that the effect is equally divided between the loci of large effect.**

The last group of lines determine the structure of environmental (co)variances of traits.

The software **subdiv_pop** will produce two type of files:

- *[select_loci_***.txt]* where *** goes from 1 to the number of replicates we want to run. The first and second column show the chromosome and the position within chromosome of the locus. The next two columns present the additive and dominant effects, respectively. The values on top correspond to the first trait, then the lines for the second trait and so on.

```

13      6539  -2.40382403E-02  0.00000000
17      4361  -3.14062508E-03  0.00000000
12      7071   5.55786937E-02  0.00000000
15      7136   2.69308109E-02  0.00000000
 2      1950   2.68000867E-02  0.00000000
 3      7738  -1.96610559E-02  0.00000000
:
13      6539   2.63059903E-02  0.00000000
17      4361   2.13885903E-02  0.00000000
12      7071  -2.20128801E-02  0.00000000
15      7136  -1.85207520E-02  0.00000000
 2      1950  -5.36477333E-03  0.00000000
 3      7738  -2.09480133E-02  0.00000000
:

```

- *[genotypes_***.txt]* is the file with the final genotypes corresponding with the (sub)populations generated. Structure of the file is as in the previous ones:

⇒ *individual id, chromosome, genotypes for this chromosome (first allele first locus, second allele first locus, first allele second locus, second allele second locus, ...).*

```

1 1 0 1 0 1 0 0 0 0 1 0 ...
1 2 1 1 1 1 0 1 1 0 1 0 ...
1 3 1 0 1 0 0 0 0 0 0 1 ...
:
1 24 1 0 1 1 0 0 0 0 0 1 ...
2  1 0 1 1 1 0 0 1 1 1 0 ...
:
2 24 1 0 1 1 0 1 0 0 1 1 ...
3  1 1 0 0 0 0 1 1 1 0 0 ...
:

```

Hint: If a single population was created or if all individuals of the different strains are to be used as base population the *genotypes_***.txt* file can be used directly as the input for the programs simulating the breeding programs. If a synthetic population is created from the strains, the derived file should be renamed to *genotypes_***.txt* in order to be read by the subsequent software.

Breeding program software

The rest of programs perform the simulation of a breeding program under different evaluation strategies. All of them share the same input file [*selection_input.txt*] for the determination of all the available options for the breeding program. The structure of the file is as follows:

```
3 ! No. traits
1 0 1 ! Availability of measured individuals for selection by trait
1 1 0 ! Weight of traits
300 5 100 ! No. base individuals, no. generations and no. replicates
1 ! Interval (in generations) between evaluations (used only for GBLUP)
200 30 ! No. families and no. offspring per family
100 200 ! No. sire and dams selected
1 ! Individuals available for the commercial population: (0) no commercial population is to be
    created (1) best GEBV (2) 'second best' (3) all individuals (mate selection)
0 0 ! Type of mating for nucleus and commercial: (0) Random (1) mcm (2) Optimise offspring
    value
2 1 ! Max. no. matings per sire and dam
1 ! Mass spawning (0), hierarchical (1) or factorial (2)
1 0.33 1 ! Proportion of measured individuals (e.g. disease challenges) in each trait
1 ! Solver to be used in the optimisation of the mate selection (1) Hungarian algorithm (2)
    simulated annealing
0.05 ! Rate of inbreeding to be obtained
```

The first line is self-explanatory. The second line establishes if the individuals that are measured for a particular trait are available (1) or not (0) for the selection. For example, individuals exposed to a disease challenge are not useful for selection (the same for fillet yield or similar traits). However, the individuals measured for size or feed intake, for example, can be candidates to be the parents of the next generation.

The next line determines the weight given to each of the traits in the index to be used as the selection criterion in any of the selection programs. The selection decision will be taken on a linear combination of the phenotypes or the (G)EBVs of the candidates with the coefficients defined by the user.

In the next line the number of individuals of the base population (founders) is defined. **Hint: this number has to be equal to the sum of the number of males and females selected each generation (see some lines below).** The number of generations to be simulated as well as the number of replicates are also determined in the same line of the file.

The next line is intended only for genomic selection and establishes how often (in generations) the effects of the markers are re-evaluated. For generations where no genomic evaluation is performed, GEBVs are calculated using the genotypes of candidates and the latest estimated effects.

The following two lines determine the population structure during the selection process in terms of number of families, number of offspring generated per family, and the number of sire and dams selected. These figures will define the selection pressure (i.e. proportion of selected individuals with respect the total number of candidates).

The software simulates a breeding program where, besides the nucleus for selection, there is a multiplier population to create the individuals to be sent to the market (commercial population, thereafter). The latter is created each generation *de novo* from the nucleus. The individuals used to generate offspring in the commercial population are (1) the ones with the best (G)EBV, (2) the ones with the best (G)EBV once the individuals generating the nucleus have been removed ('second best') or (3) the ones arising from the optimisation of a 'mate selection' scheme. In the latter strategy the selected individuals and the particular mating design are decided in a single step in order to get the maximum expected phenotype in the offspring (see, for example, Toro and Varona, 2014). This way, dominant effects can be exploited.

The mating scheme can be constructed (0) at random, (1) under a minimum coancestry mating design (*mcm*) or (2) following the mate selection procedure cited before (only in the commercial population). The number of matings in which a sire or dam will be involved can be determined.

The user can also decide if the mating design implies (1) hierarchical schemes (i.e. each male mated to an equal number of –different- females, or vice versa) or (2) (partial) factorial schemes (i.e. all possible pairs performed). In the latter scenario the structure of blocks can be defined assigning the correct values to the number of matings allowed for each sex. **Hint: The number of families for the selection program have to coincide with the total number of matings performed.** Additionally, a mass spawning program can be also simulated (option 0). In this situation there is not a fixed structure but different full-sib (FS) and half-sibs (HS) families are created. Moreover, the contribution of parents (families) is not constant and even some selected individuals may end up yielding no offspring. **Hint: If mass spawning is chosen the only sensible option is random mating, as there is no possibility of directing the pairs.**

The optimisation of the mating scheme and the mate selection process are performed through a 'Hungarian algorithm'. Explanations on the application of this algorithm can be found, for example, in Fernández and Toro, 1999.

The proportion of candidates that are measured for each trait is another parameter to be defined by the user. Finally, the maximum rate of increase in inbreeding can be set for the Optimal Contributions (OC) selection. If no restriction is to be included 1 is the value which should appear in the file.

In order to maintain the mating scheme and the required number of families created each generation, OC selection is implemented with fixed numbers of breeding males and females, as in truncation selection schemes. The problem to be solved can be formulated as:

$$\text{Maximise } \mathbf{c}^T \mathbf{g} \text{ subject to } \mathbf{c}^T \Phi \mathbf{c} \leq C$$

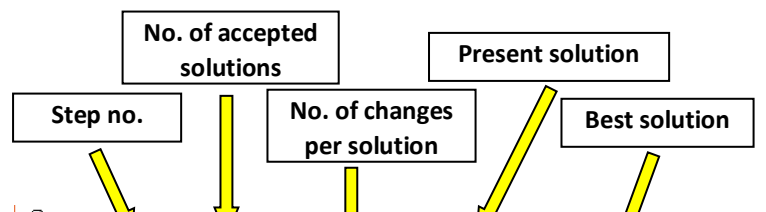
where $\mathbf{c}$ is a vector of solutions (1 if the candidate is selected and 0 otherwise), $\mathbf{g}$ is the vector of BLUP-EBVs of the candidates, Φ is the coancestry matrix and C is the coancestry level (calculated from pedigree or molecular information) that leads to the desired rate of inbreeding (ΔF). The inequality ensures that the constraint on ΔF is met. A couple of extra constraints must be included in order to maintain the number of breeding males and females constant: $\sum_{males} c_i = n_s$ and $\sum_{females} c_i = n_d$, where n_s and n_d are the number of breeding males and females, respectively. Optimisation is solved using a 'simulated annealing' algorithm (Fernández and Toro, 1999). The performance of the algorithm may be improved by modifying the values in the file:

- [anneal_param.txt]

This is a text file including the control parameters of the 'simulated annealing' algorithm. Meaning of each value is:

- 10 $\Rightarrow$ seed for the random number generator
- 200 $\Rightarrow$ maximum no. of allowed steps (temperatures)
- 1000 $\Rightarrow$ no. of solutions tested in each step
- .01 $\Rightarrow$ initial temperature
- .9 $\Rightarrow$ rate of decrease in temperature (i.e. $T_{t+1} = T_t * 0.9$)

During the execution of the 'simulated annealing' algorithm the programme will show some information to check if the process is running correctly and to allow for the 'fine tuning' of the control parameters. Screen will look like this



Level	Accep. sol.	Changes	Present sol.	Best solution
1	98793	10	2.96208132E-02	-0.180914611
2	98675	10	4.64231707E-02	-0.180914611
3	98532	10	9.48169827E-03	-0.188070729
4	98372	10	-2.29285732E-02	-0.188070729
5	98145	10	-5.78846447E-02	-0.204184324
6	97917	10	-4.23245840E-02	-0.204184324
7	97710	10	9.45165381E-03	-0.204184324
8	97510	10	-3.99749540E-02	-0.204184324
9	97145	10	-3.30164321E-02	-0.204184324
10	96945	10	2.77312640E-02	-0.204184324
11	96599	10	0.115985617	-0.204184324
12	96126	10	9.87349264E-03	-0.204184324
13	95743	10	0.126794413	-0.204184324
14	95286	10	-8.02952275E-02	-0.204184324
15	94837	10	9.16208103E-02	-0.204184324
16	94058	10	-4.48284894E-02	-0.204184324
17	93458	10	-2.96605881E-02	-0.212198272
18	92843	10	-1.93363205E-02	-0.212198272
19	91976	10	1.14065120E-02	-0.212198272
20	91043	10	-4.31820471E-03	-0.212198272
21	90225	10	-1.87268928E-02	-0.212198272
22	89020	9	-4.31377301E-03	-0.225963265
23	88665	9	-6.41587302E-02	-0.225963265
24	87099	9	-5.10163531E-02	-0.225963265
25	85885	9	-4.64818068E-03	-0.225963265
26	84413	9	-3.35943736E-02	-0.225963265
27	82257	9	-0.117204830	-0.229584798
28	80533	9	-3.47227901E-02	-0.229584798
29	78358	8	-3.48013602E-02	-0.229584798
30	77521	8	-9.83164981E-02	-0.229584798
31	75063	8	-0.180788472	-0.229584798
32	72401	8	-6.56138584E-02	-0.229717493
33	69400	7	-6.24348782E-02	-0.234246179
34	68304	7	-5.42180389E-02	-0.234246179
35	65118	7	-5.99359125E-02	-0.243717805
36	61189	7	-6.60749301E-02	-0.243717805
37	57461	6	-7.77285621E-02	-0.255200356
38	56007	6	-0.115081243	-0.260297149
39	52071	6	-0.142501980	-0.266750991

For example, too rapid decrease in the number of accepted solutions (and the corresponding number of changes per new solution) may indicate a too low initial temperature, so algorithm is stuck in the starting area. See Kirkpatrick et al. (1983) for thorough explanation of the annealing method and for considerations about the definition of the different parameters.

Mass_select

This program simulates a breeding program under the strategy of mass selection (also called individual selection). In mass selection, individuals are ranked according to their own phenotype and those with the highest phenotype are selected for breeding. No family (i.e. pedigree) information is recorded/used and, thus, there is a significant risk that a large proportion of selected breeders will be close relatives, particularly, when the trait selected has a relatively high heritability.

BLUP

The genetic evaluation method in the breeding program performed by this option is standard BLUP. The model assumed is

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{d} + \mathbf{b}\mathbf{f} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypes, $\mathbf{1}$ is a vector of ones, μ is the population mean, $\mathbf{a}$ and $\mathbf{d}$ are the vectors of random additive and dominance effects, respectively, b is the linear regression of $\mathbf{y}$ on inbreeding coefficients, $\mathbf{f}$ is the vector of inbreeding coefficients, $\mathbf{e}$ is the vector of random residuals and $\mathbf{Z}$ and $\mathbf{W}$ are incidence matrices. The terms $\mathbf{W}\mathbf{d}$ and $\mathbf{b}\mathbf{f}$ are included in the model only if $d^2 \neq 0$. The corresponding variances are $V(\mathbf{a}) = \mathbf{A}\sigma_a^2$, $V(\mathbf{d}) = \mathbf{D}\sigma_d^2$ and $V(\mathbf{e}) = \mathbf{I}\sigma_e^2$, where $\mathbf{A}$ and $\mathbf{D}$ are the additive and dominance, respectively, genetic relationships matrices, $\mathbf{I}$ is the identity matrix and σ_a^2 , σ_d^2 and σ_e^2 are the additive and dominance genetic variances and the error variances which are assumed to be known and taken from the real/simulated values.

GBLUP

In this program, genomic SNP-BLUP evaluation is performed to estimate the breeding values and the dominance deviations assuming the model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{v} + \mathbf{b}\mathbf{h} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypes, $\mathbf{1}$ is a vector of ones, μ is the population mean, $\mathbf{u}$ and $\mathbf{v}$ are the vectors of SNPs additive and dominant effects, respectively, b is the linear regression of $\mathbf{y}$ on observed heterozygosities, $\mathbf{h}$ is the vector of observed heterozygosities, $\mathbf{e}$ is the vector of random residuals and $\mathbf{Z}$ and $\mathbf{W}$ are incidence matrices constructed from individual genotypes. The element of the i th row and j th column in $\mathbf{Z}$ is 2, 0 or 1 if the i th fish is respectively homozygote 11, heterozygote or homozygote 00 for the j th SNP. The corresponding element of $\mathbf{W}$ is 0 if the fish is homozygote and 1 if the fish is heterozygote. The terms $\mathbf{W}\mathbf{v}$ and $\mathbf{b}\mathbf{h}$ are included in the model only if $d^2 \neq 0$. It is assumed that the SNP effects are uncorrelated and that $V(\mathbf{u}) = \mathbf{I}\sigma_u^2$ and $V(\mathbf{v}) = \mathbf{I}\sigma_v^2$, where $\mathbf{I}$ is the identity matrix and σ_u^2 and σ_v^2 are the variances of the additive and dominant marker effects which are assumed to be known. For additive effects, the variance is calculated by dividing the real simulated additive variance by $2\Sigma[p_i(1-p_i)]$, where p_i is the frequency, in the base population, of allele 1 of the i th SNP in the group of individuals evaluated (Gianola et al., 2009). For dominant effects the correction term is $4\Sigma[p_i^2(1-p_i)^2]$. Both σ_u^2 and σ_v^2 are kept constant across generations. The inclusion of the covariate $\mathbf{h}$ makes possible accounting for inbreeding depression in the evaluation when dominant effects exist.

All individuals are genotyped and evaluations are performed using the data from the current generation as well as the genotypes of previous generations.

The estimated breeding value (GEBV) for each individual is calculated from the estimated additive SNPs effects (a_i) by summing across the QTLs $2(1 - p_i)\alpha_i$, $(1 - 2p_i)\alpha_i$ or $-p_i\alpha_i$ for homozygotes 11, heterozygotes and homozygotes 00, respectively, where $\alpha_i = a_i + (1 - 2p_i)d_i$. Similarly, the dominance deviation for each individual is calculated from the estimated dominant SNPs effects (d_i) by summing across the QTLs $-2(1 - p_i)^2d_i$, $2p_i(1 - p_i)d_i$ or $-2p_i^2d_i$ for homozygotes 11, heterozygotes and homozygotes 00, respectively.

When a commercial population is to be created and the option of 'mate selection' has been chosen, all the candidates available for selection enters the optimization. Firstly, the software constructs a matrix with the expected phenotypic value of the offspring of each possible mating pair. This value depends on the SNP genotypes of the involved parents that determine their probability of transmitting allele 1 or 0. Homozygous parents 11 always transmit allele 1, homozygotes 00 never transmit allele 1 and heterozygotes transmit allele 1 to half of their offspring. The proportion of offspring of each possible genotype (P_{11} , P_{10} and P_{00} for genotypes 11, 10 and 00, respectively) can be calculated by multiplying the probabilities of transmitting each allele (assuming no deviations from Mendelian inheritance) in any particular mating pair. Once P_{11} , P_{10} and P_{00} are calculated, the expected phenotype of the offspring of a potential mating pair is estimated as $a_i P_{11} + d_i P_{10} - a_i P_{00}$ for SNP i . The total expected phenotype is the sum of the values for all SNPs plus a term including the mean observed heterozygosity of offspring multiplied by the regression coefficient b calculated in the evaluation, to account for the possible effect of inbreeding depression.

WF_GBLUP

In this scheme, estimated breeding values are combinations of a conventional family estimated breeding value (between family component) and a genomic within-family breeding value. In particular, the EBV for individual i of family l is $EBV_{il} = \frac{1}{2}EBV_{si} + \frac{1}{2}EBV_{di} + w_{il}$, where EBV_{si} and $\frac{1}{2}EBV_{di}$ are the estimated breeding values of the sire and dam, respectively, of individual i (and all its sibs in the corresponding family l) and w_{il} is the within-family breeding value of individual i . Actually, the composite term $(\frac{1}{2}EBV_{si} + \frac{1}{2}EBV_{di})$ is estimated as the average EBV for family l obtained from standard BLUP using the full data set (i.e., all available phenotypes). The within-family breeding value w_{il} was obtained from SNP-BLUP evaluation using only information (genotypes and phenotypes) from family l . The model assumed for both steps are the same as used in BLUP and GBLUP evaluations above.

All evaluations (BLUP, GBLUP and WF_GBLUP) are performed using the software GS3 (Legarra *et al.*, 2011) which is also provided with the suite and has to be located in the same folder as the rest of executables. **Hint: When running several traits, evaluations are performed in a single-trait basis, even if there is some genetic correlation between them.**

In any of the selection scenarios (programs) the output of the simulations will be compiled in the file [results_selection.txt]. This file is created each generation and, thus, user can check if everything is running correctly and results seem sensible without waiting for the whole number of replicates to finish. For all the parameters included, the mean values across replicates and their variances will be shown (for all generations) for the nucleus (first and second column) as well as for the commercial population (third and fourth columns; all numbers set to zero if it does not exist). The appearance of the file will be:

```

Replicate          100

Trait              1

Segregating SNPs
Nucleus
*****

Environmental variance      Inbreeding depression
0.2845      0.0002      0.0000      0.0000

LD between SNPs
      Nucleus      Commercial
0      0.0301      0.0000      0.0000      0.0000
1      0.0636      0.0001      0.0000      0.0000

:
LD between SNPs and QTLs
      Nucleus      Commercial
0      0.0303      0.0000      0.0000      0.0000
1      0.0635      0.0002      0.0000      0.0000

:
Mean phenotype
      Nucleus      Commercial
0      0.0118      0.0664      0.0000      0.0000
1      1.9252      0.0899      0.0000      0.0000

:
Phenotypic variance
:
Mean breeding value
:
Variance of breeding values
:
Additive variance (2pq*alfa**2)
:
Variance of dominant values
:
Expected heterozygosity non-markers
:
Expected heterozygosity SNPs
:
Genealogical inbreeding
:
Genealogical coancestry
:
Accuracy (correlation) of estimates
:
:
Trait              2
:
:
Trait              3
:

```

Known limitations

The limit in the size of the problem is related with the amount of memory available in your computer. No particular limits have been imposed in the genome length (number of loci), number of families, family size, etc.

If you have any particular problem when using the program please, contact with the authors in the e-mail jmj@inia.es. If you discover any bug or you would like any feature to be included in future versions, your comments will be very welcome.

References

- Bürger, R., Wagner, G.P., Stettinger, F. 1989. How much heritable variation can be maintained in finite populations by mutation-selection balance? *Evolution* 43: 1748–1766.
- Fernández, J., Toro, M. A. 1999. The use of mathematical programming to control inbreeding in selection schemes. *Journal of Animal Breeding and Genetics* 116: 447-466.
- Fernández, J., Toro, M. A., Sonesson, A. K., Villanueva, B. 2014. Optimizing the creation of base populations for aquaculture breeding programs using phenotypic and genomic data and its consequences on genetic progress. *Frontiers in Genetics* 5: 414.
- Gianola, D., de los Campos, G., Hill, W. G., Manfredi, E., and Fernando, R. 2009. Additive Genetic Variability and the Bayesian Alphabet. *Genetics* 183: 347-363.
- Legarra A., Ricardi A., Filangi O. 2011. GS3: Genomic Selection, Gibbs Sampling, Gauss-Seidel (and BayesCp). Available at: <http://genoweb.toulouse.inra.fr/~alegarra>.
- Toro, M. A., Varona, L. 2010. A note on mate allocation for dominance handling in genomic selection. *Genetics Selection Evolution* 42: 33.
- Turelli, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical and Population Biology* 25: 138–193.

Liability

The GAINFISH software is developed and owned by INIA. Access and use of the software is free of charge. INIA (and the developers) can not be held responsible for the results coming when using the software.



ANNEX II: SUPPLEMENTARY

MATERIAL OF CHAPTER 2

Table S1. Accuracy of evaluation for body weight (BW), headless carcass percentage (HC) and the indicator of headless carcass percentage (IHC) at generation 0 when giving different relative weights to the three traits and when using two different evaluation methods (BLUP, scenarios B_x, or GBLUP, scenarios G_x), for a heritability of HC of 0.55.^a

Scenarios ^b	Selection for	Weights			BLUP			GBLUP		
		BW	HC	IHC	BW	HC	IHC	BW	HC	IHC
B _W and G _W	BW	1	0	0	0.69	-	-	0.74	-	-
B _H and G _H	HC	0	1	0	-	0.63	-	-	0.76	-
B _I and G _I	IHC	0	0	1	-	-	0.66	-	-	0.71
B _{W+H_1} and G _{W+H_1}	BW and HC	1	0.25	0	0.72	0.63	-	0.79	0.75	-
B _{W+H_2} and G _{W+H_2}	BW and HC	1	0.50	0	0.72	0.63	-	0.79	0.75	-
B _{W+H_3} and G _{W+H_3}	BW and HC	1	0.75	0	0.72	0.63	-	0.79	0.75	-
B _{W+H_4} and G _{W+H_4}	BW and HC	1	1	0	0.72	0.63	-	0.79	0.75	-
B _{W+I_1} and G _{W+I_1}	BW and IHC	1	0	0.25	0.71	-	0.66	0.78	-	0.71
B _{W+I_2} and G _{W+I_2}	BW and IHC	1	0	0.50	0.71	-	0.66	0.78	-	0.71
B _{W+I_3} and G _{W+I_3}	BW and IHC	1	0	0.75	0.71	-	0.66	0.78	-	0.71
B _{W+I_4} and G _{W+I_4}	BW and IHC	1	0	1	0.71	-	0.66	0.78	-	0.71

^a Standard errors were equal to 0.01 in all scenarios.

^b Numbers 1 to 4 simply enumerate the scenarios with different weights for each trait to facilitate referring to them in the text.

Table S2. Accuracy of evaluation for body weight (BW), headless carcass percentage (HC) and the indicator of headless carcass percentage (IHC) at generation 0 when giving different relative weights to the three traits and when using two different evaluation methods (BLUP, scenarios B_x, or GBLUP, scenarios G_x), for a heritability of HC of 0.25.^a

Scenarios ^b	Selection for	Weights			BLUP			GBLUP		
		BW	HC	IHC	BW	HC	IHC	BW	HC	IHC
B _W and G _W	BW	1	0	0	0.68	-	-	0.73	-	-
B _H and G _H	HC	0	1	0	-	0.55	-	-	0.63	-
B _I and G _I	IHC	0	0	1	-	-	0.66	-	-	0.71
B _{W+H_1} and G _{W+H_1}	BW and HC	1	0.25	0	0.71	0.55	-	0.78	0.63	-
B _{W+H_2} and G _{W+H_2}	BW and HC	1	0.50	0	0.71	0.55	-	0.78	0.63	-
B _{W+H_3} and G _{W+H_3}	BW and HC	1	0.75	0	0.71	0.55	-	0.78	0.63	-
B _{W+H_4} and G _{W+H_4}	BW and HC	1	1	0	0.71	0.55	-	0.78	0.63	-
B _{W+I_1} and G _{W+I_1}	BW and IHC	1	0	0.25	0.72	-	0.66	0.78	-	0.71
B _{W+I_2} and G _{W+I_2}	BW and IHC	1	0	0.50	0.72	-	0.66	0.78	-	0.71
B _{W+I_3} and G _{W+I_3}	BW and IHC	1	0	0.75	0.72	-	0.66	0.78	-	0.71
B _{W+I_4} and G _{W+I_4}	BW and IHC	1	0	1	0.72	-	0.66	0.78	-	0.71

^a Standard errors were equal to 0.01 in all scenarios.

^b Numbers 1 to 4 simply enumerate the scenarios with different weights for each trait to facilitate referring to them in the text.



ANNEX III: COMMUNICATIONS AT CONGRESSES RELATED TO THIS THESIS

International congresses

García-Ballesteros, S., Fernández, J., Doeschl-Wilson, A., and Villanueva, B. (2022). Impact of selection for disease resistance on disease transmission in a simulated salmon population, using genomic information. XIV International Symposium on Genetics in Aquaculture (ISGA). Puerto Varas, Chile, November 27 – December 2. Oral presentation (accepted).

García-Ballesteros, S., Fernández, J., Doeschl-Wilson, A., and Villanueva, B. (2022). Impact of selection for disease resistance on the spread of infection in a simulated aquaculture population. World Congress on Genetics Applied to Livestock Production (WCGALP). Rotterdam, The Netherlands, July 3 – 8. Oral presentation.

García-Ballesteros, S., Fernández, J., Toro, M.A., and Villanueva, B. (2021). Benefits of genomic evaluation in aquaculture breeding programs in presence of tank effect. GIA 6th International Symposium on genomics in aquaculture (GIA). Virtual meeting, May 5 – 7. Oral presentation.

García-Ballesteros, S., Villanueva, B., Fernández, J., Gutiérrez, J.P., and Cervantes, I. (2020). Genetic parameters for uniformity of harvest weight in white shrimp (*Litopenaeus vannamei*). 71st Annual Meeting of the European Federation of Animal Science (EAAP). Virtual meeting, December 1 – 4. Oral presentation.

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2019). Benefits from within-family genomic evaluation in aquaculture schemes with small family sizes. European Aquaculture Society (EAS). Berlin, Germany, October 7 – 10. Poster presentation.

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2018). Genetic gain from genomic evaluation in fish breeding programs with separate rearing of families. EAAP 69th Annual Meeting of the European Federation of Animal Science (EAAP). Dubrovnik, Croatia, August 27 – 31. Oral presentation.

Villanueva, B., García-Ballesteros, S., and Fernández, J. (2018). Optimal schemes for advancing selective breeding to the next level for the main species of European aquaculture. European Aquaculture Society (EAS). Montpellier, France, August 25 – 29. Oral presentation.

National congresses

García-Ballesteros, S., Fernández, J., Kause, A., and Villanueva, B. (2022). Combining genomic and indirect selection to improve carcass yield in rainbow trout. XX Reunión Nacional de Mejora Genética Animal. Madrid, Spain, June 1 – 3. Oral presentation.

Toro, M.A., Fernández, J., Saura, M., Fernández, A., García-Ballesteros, S., and Villanueva, B. (2021). Genetic improvement programs in aquaculture. XLII Congreso de la Sociedad Española de Genética (SEG). Online, June 14 – 18. Oral presentation.

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2019). Beneficio de la evaluación genómica dentro de familias en esquemas de acuicultura con tamaños familiares reducidos. XVIII Jornadas sobre Producción Animal de la Asociación Interprofesional para el Desarrollo Agrario (AIDA). Zaragoza, Spain, May 7 – 8. Oral presentation.

García-Ballesteros, S., Fernández, J., and Villanueva, B. (2018). Estrategias para reducir costes en evaluación genómica en programas de selección de acuicultura. XIX Reunión Nacional de Mejora Genética Animal. León, Spain, June 14 – 15. Oral presentation.

