



OPEN

Genomic selection for white spot syndrome virus resistance in whiteleg shrimp boosts survival under an experimental challenge test

Marie Lillehammer¹, Rama Bangera², Marcela Salazar³, Sergio Vela², Edna C. Erazo³, Andres Suarez³, James Cock³, Morten Rye² & Nicholas Andrew Robinson^{1,4}✉

White spot syndrome virus (WSSV) causes major worldwide losses in shrimp aquaculture. The development of resistant shrimp populations is an attractive option for management of the disease. However, heritability for WSSV resistance is generally low and genetic improvement by conventional selection has been slow. This study was designed to determine the power and accuracy of genomic selection to improve WSSV resistance in *Litopenaeus vannamei*. Shrimp were experimentally challenged with WSSV and resistance was evaluated as dead or alive (DOA) 23 days after infestation. All shrimp in the challenge test were genotyped for 18,643 single nucleotide polymorphisms. Breeding candidates (G_0) were ranked on genomic breeding values for WSSV resistance. Two G_1 populations were produced, one from G_0 breeders with high and the other with low estimated breeding values. A third population was produced from “random” mating of parent stock. The average survival was 25% in the low, 38% in the random and 51% in the high-genomic breeding value groups. Genomic heritability for DOA (0.41 in G_1) was high for this type of trait. The realised genetic gain and high heritability clearly demonstrates large potential for further genetic improvement of WSSV resistance in the evaluated *L. vannamei* population using genomic selection.

Disease caused by the White Spot Syndrome Virus (WSSV) is a serious problem for global shrimp aquaculture. It infects all major cultured shrimp species and is highly virulent normally causing death in a few days¹. Cultivation practices, mostly aimed at eliminating the causal agent, may reduce the economic risks associated with white spot disease², but elimination in open pond systems is often impossible³. The main strategies used worldwide to cope with on farm viral infection involve the use of biosecure protocols and fast-growing WSSV pathogen-free but susceptible shrimp in a race to harvest before infection and death occurs^{4–6} or growth of shrimp when water temperatures are higher and the disease does not prosper². Elimination strategies can be effective if used in combination with good management practices⁷, but shrimp often succumb rapidly to disease and provision of WSSV resistant, pathogen-free stocks would greatly benefit the industry.

Shrimp lack an adaptive immune system, relying on the innate immune system to prevent, tolerate and clear infection see review of crustacean immunity⁸. Hence, several conventional approaches used in other live-stock, such as vaccination to boost the immune response and provide protection, have had limited success for shrimp. Although stimulation of the innate immune system through immune priming to prevent and control specific diseases in shrimp shows promise, the efficacy of such methods have not yet been proven under field conditions^{9–11}. Knowledge about the defensive response to infection is building, but the specific mechanisms associated with greater resistance or tolerance to disease in specific populations or individuals have not been conclusively identified. Nevertheless, there is substantial information that can be used to guide breeders in their efforts to produce resistant stocks. Estimates of the heritability of WSSV resistance in *Litopenaeus vannamei* range between 0.01 and 0.31 depending on the batch of shrimp, the trait analysed (e.g. days survival or binary

¹Breeding and Genetics, Nofima, 1430 Ås, Norway. ²Benchmark Genetics Norway AS, 6600 Sunndalsøra, Norway. ³Benchmark Genetics Colombia, Bogota, Colombia. ⁴Sustainable Aquaculture Laboratory-Temperate and Tropical (SALTT), School of BioSciences, The University of Melbourne, Parkville 3010, Australia. ✉email: nick.robinson@nofima.no

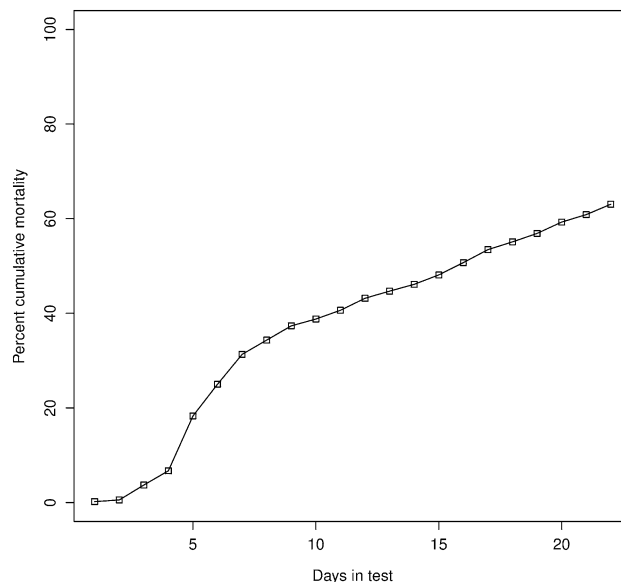


Figure 1. Cumulative mortality observed with WSSV infection among G_0 training animals. The total number of animals evaluated in the challenge test was 1459.

dead or alive), the challenge method applied and statistical models used for genetic parameter estimation^{12–16}. The simultaneous improvement of growth and WSSV resistance with traditional index selection is complicated by a negative genetic correlation (-0.55 to -0.64) between these traits^{12,16}. The heritability for white spot viral copy number (0.18) and pond survival (0.16) in the face of WSSV farm outbreaks is low^{16,17}. The genetic gain after selection depends on the heritability of the trait selected in that population and its phenotypic variance (genetic variance), the accuracy of the breeding value estimation and on the intensity of selection applied. The reported genetic gain per generation for WSSV resistance ranges from 1.7 to 6.5%^{16,18,19}.

If genetic loci with a large effect on WSSV resistance exist, then marker assisted selection (MAS) could be an efficient way to boost resistance. A gene set based association analyses for WSSV resistance (dead/alive) in *L. vannamei* identified a total of 5 SNPs within immune related genes, however, these SNPs were not validated further in a larger population for its application in MAS²⁰. Robinson et al.²¹ identified several linkage groups containing quantitative trait loci associated with time to death after WSSV infection in black tiger shrimp (*Penaeus monodon*). However, as the estimated marker effects on hours survival were relatively small, WSSV resistance is likely to be a typical quantitative trait, affected by the additive effects of many genes of small effect²². Many of the single nucleotide polymorphisms (SNPs) that mapped close to the loci associated with survival occurred near transcripts with homology to putative innate immune genes of interest. For instance, c-type lectin protein which may be involved in immune recognition and microorganism phagocytosis²³, is more highly expressed in resistant than susceptible shrimp exposed to WSSV^{24–27}. This gene was mapped to a QTL region on linkage group 43.

Genomic selection^{28,29} has been proposed as the most effective form of selection available for typical quantitative (polygenic) traits³⁰. Hence, we surmised that genomic selection could be used to improve the genetic gain in WSSV resistance. In this study we experimentally challenged families of white shrimp with WSSV virus and used single nucleotide polymorphism genotype data from challenged and candidate breeders to evaluate the power and accuracy of genomic selection for improving WSSV resistance. Breeding candidates (G_0) were ranked in terms of genomic breeding values for WSSV resistance. The breeding candidates were mated to produce two G_1 populations, one with high and the other with low genomic estimated breeding values. The survival of G_1 populations and offspring from “randomly” mated parent stock was compared in a challenge test.

Results

Filtering of the 465,500 putative SNPs identified from mRNAseq resulted in 10,323 SNPs (Table S1) which were selected for inclusion in the custom 18 K Illumina SNP array (MAF > 0.1, > 51X coverage and eliminating multiple SNPs per contig and SNP alleles fixed in the R- and S-lines). Of the total 19,290 SNPs printed on the chip, post-genotyping quality control left genotypes for 18,643 SNPs that were deemed acceptable for genomic estimated breeding value (gEBV) prediction.

In the G_0 challenge test mortality increased rapidly up to around day five as expected and then, instead of the usual flattening off, a few animals continued to die each day until the test ended (Fig. 1). The test was ended at 22 days according to standard practices, even though some animals were still dying, to facilitate consistent comparisons between G_0 and G_1 .

Estimates of heritability for the dead or alive (DOA) and days survival (DS) traits were similar for the overall G_0 population (0.30 for DOA and 0.53 for DS) to those for the pure R-line animals (DOA 0.22 and DS 0.39) and for crossed resistant by susceptible line (RxS) animals (DOA 0.34 and DS 0.48, Table 1). Similar estimates were obtained both with and without the fixed effect of line included in the analysis. The fixed effect of line revealed

Data	DOA			DS		
	Phenotypic variance	Genetic variance	Heritability	Phenotypic variance	Genetic variance	Heritability
All data	0.23	0.07	0.32 ± 0.05	63.1	34.6	0.55 ± 0.05
Pure RxR	0.26	0.056	0.22 ± 0.06	58.4	22.5	0.39 ± 0.07
RxS cross	0.18	0.061	0.34 ± 0.07	57.6	27.7	0.48 ± 0.08

Table 1. Estimated variance components and heritabilities for both traits from analysing the G_0 -data together or each genetic group (purebred or crossbred) separately.

Number of families			Power	
High-gEBV	Random-gEBV	Low-gEBV	High-gEBV vs Random-gEBV	High-gEBV vs Low-gEBV
40	10	10	0.83	1
30	20	10	0.96	1
40	20	0	0.96	
50	20	0	0.95	

Table 2. Power of experiment to evaluate genomic selection assuming one male is mated with one female to generate each separate family.

0.19 higher survival probability for purebred RxR than crossbred RxS. For the DS trait, the effect of being purebred was 2.71 days longer survival.

Genetic variation within the pure resistant line (R-line) and crossed RxS-line animals was found to be substantial (0.056 for DOA in the pure R-line, 0.061 for DOA in the crossed RxS line, and 22.5 for DS in the pure R-line and 27.7 for DS in the crossed RxS line individuals) so that a good selection response could be expected. This is supported by the accuracy of breeding values, estimated by cross validation of the full dataset, which was 0.69 for DOA and 0.64 for DS.

Power analysis. The power analysis indicated that the minimum number of selected parents in the high resistance group should be 30–40 males and females when there are 20 males and females in the random group (Table 2).

Evaluation of genomic selection on G_1 . Out of the 1885 animals tested, 1 family of 35 siblings (random-gEBV group) was removed due to missing parental data, 3 animals could not be confirmed to be derived from the parents used in the experiment, leaving 1847 records from 59 G_1 families (32 high-, 19 random- and 8 low-gEBV) for analysis.

Cumulative mortality increased at a slower rate and plateaued at a lower level in the challenge test for the high gEBV compared to random and low gEBV groups (Fig. 2). The low gEBV group had the fastest rate of mortality and reached the highest levels of cumulative mortality of the three groups (cumulative mortality at the end of the test of 75% in the low, 63% in the random and 49% in the high gEBV group), giving a 13% improved rate of survival in the high gEBV group. The G_1 random gEBV group reached the same final cumulative mortality during the test as was observed in the G_0 training population (63%, Fig. 1).

Average survival of families in the high gEBV group at 51% was significantly greater than that of the random group (38%) which was, in turn, significantly greater than the 25% survival of the low gEBV group (Fig. 3). The range of survival within families was large, with the greatest variation in the high-gEBV families, which also had the largest range of average family survival. Families with the highest proportion of offspring surviving inherited a greater proportion of the resistant line genes (100% versus 75% or 50%), across the high-, random- and low-groups (Fig. 3). No families from the low group had greater than 50% survival whereas half the high group had greater than 50% survival with four of these families, out of a total of 32, with more than 70% survival.

The heritability of DOA in the G_1 population (Table 3) was greater than those for the G_0 population (Table 1), regardless of whether the effect of selection group was fitted in the model or not. The estimated realised response to selection was greater than predicted (Fig. 4, assuming heritability for DOA of 0.3 as calculated for the G_0 , Table 1), and the phenotypic differences between the groups were larger than the statistically estimated group effect. Tank was found to have a low but significant effect on DOA and was therefore included in the analysis model as a fixed effect.

Discussion

Average survival in the G_1 population increased from 38 to 51% after one generation of genomic selection for high WSSV resistance for the dead or alive binary trait (DOA) relative to average survival in G_1 randomly selected shrimp (Fig. 4). Selection for low WSSV resistance gave a very similar response in the opposite direction. This clearly demonstrates the power of genomic selection as a tool to develop WSSV resistant *L. vannamei* shrimp using this particular “synthetic” population which contains large variation for WSSV resistance. Due to the large

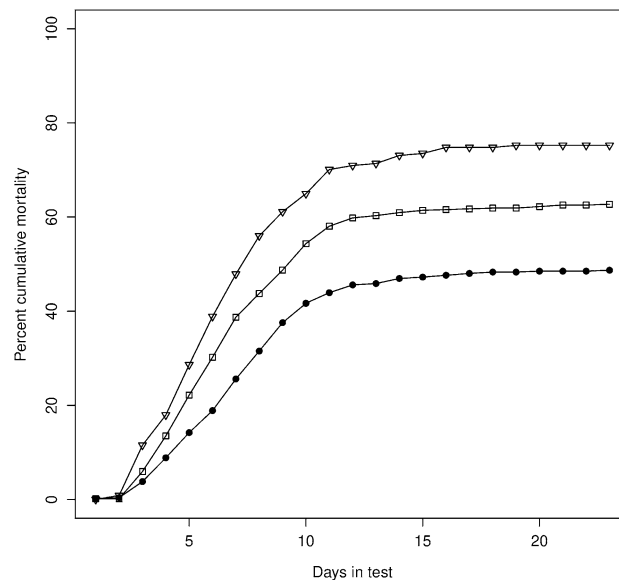


Figure 2. Percent cumulative mortality plotted over days in challenge test for G_1 animals in the high—(closed circles), random—(open squares) and low—(open triangles) gEBV groups. The total number of animals evaluated in the challenge test was 1883 (high 1027, random 622 and low 234).

variation in the DOA trait, and relative high heritability, further selection to produce a G_2 would be expected to produce greater genetic gains than evaluated with production of the G_1 .

Greater genetic gains for the days survival (DS) trait, with genomic heritability (h^2_{DS}) of 0.55 and additive genetic variance ($V_{A,DS}$) of 34.6, would be expected than that for DOA (h^2_{DOA} of 0.32, $V_{A,DOA}$ of 0.07) (Table 1). As a consistently strong positive correlation between DS and DOA is observed in the literature (e.g. $r > 0.8$)³¹ improvement in one trait is likely to result in improvement of the other. This was confirmed in our study, as the shrimp from the group selected for high resistance survived 2 days longer than the randomly selected shrimp, which again survived 2 days longer than the group selected for low resistance (results not shown). Moreover, farmers prefer shrimp populations that survive to those that merely take a few days more to die in the presence of WSSV. Hence, we suggest that genomic selection for WSSV resistance should be based on the DOA trait and not the DS trait despite a potentially faster improvement if DS were to be the trait selected.

Cross-validation, by elimination of individual values and comparison with the predicted value from the remaining data, has been used before to predict the accuracy of genomic selection e.g.³². The accuracy of genomic predictions through cross-validation has been reported in *L. vannamei* for resistance to bacterial disease²⁰, feed efficiency traits³³ and growth rate^{34,35}. Nevertheless, cross-validation only provides an estimate or prediction of the likely genetic gain. The realized genetic gain is a direct measure of the genetic gain and of the effectiveness of a selection method. Hence, we prefer realized genetic gain to cross validation as a means of evaluating distinct selection methods. We have for the first time measured the realized genetic gain from genomic selection for disease resistance in a crustacean species, adding to the earlier evaluation of realized genetic gain from genomic selection for disease resistance in another aquatic species, rainbow trout³⁶.

We used power calculations to design a selection experiment to: (1) give power > 0.95 to detect the selection response we expect from the estimated heritability and selection differential and (2) give as many high resistant families as possible, suitable for further breeding for increased resistance. To achieve this, we aimed to produce 30 high resistance, 20 random and 10 low resistance families. However, not all candidates matured and became available to make the planned crosses. Due to lack of available breeder candidates we produced 32 high, 20 random and 8 low resistance families. The predicted selection response, based on the heritability estimated from the G_0 and the actual selection differential after mating (i.e. the average breeding value of the parents within each selection group) was 0.075 for the selected resistant group and -0.084 for the selected susceptible group. However, the realized response to selection was higher than predicted (Fig. 4), and statistically significant ($p < 0.05$). The actual difference in survival between the groups in G_1 was 0.13 in each direction. The statistical model used to analyse the G_1 data gave an estimated effect of selection of 0.1, i.e. implying that survival should be 0.1 higher or lower in the selection groups than in the random group. However, the estimated response to selection is probably under-estimated because the model aims to distinguish between the effect of genetics and the effect of selection, while these effects are by nature confounded. As this was a controlled challenge test with few other factors to correct for, we suggest the phenotypic differences between the groups are a good estimate of genetic gain.

Heritability estimates differed between G_0 and G_1 and this could explain the discrepancy between predicted and realised selection response, if the predicted response was based on an underestimate of genetic variance. Our experience is that few new mortalities occur after day 15 of the challenge test (as was observed for the challenge of G_1 animals) and the test is normally ended at 22 days. The sustained low level of mortality until the end of test in the G_0 population was abnormal and could be attributed to a lower than normal viral load in the

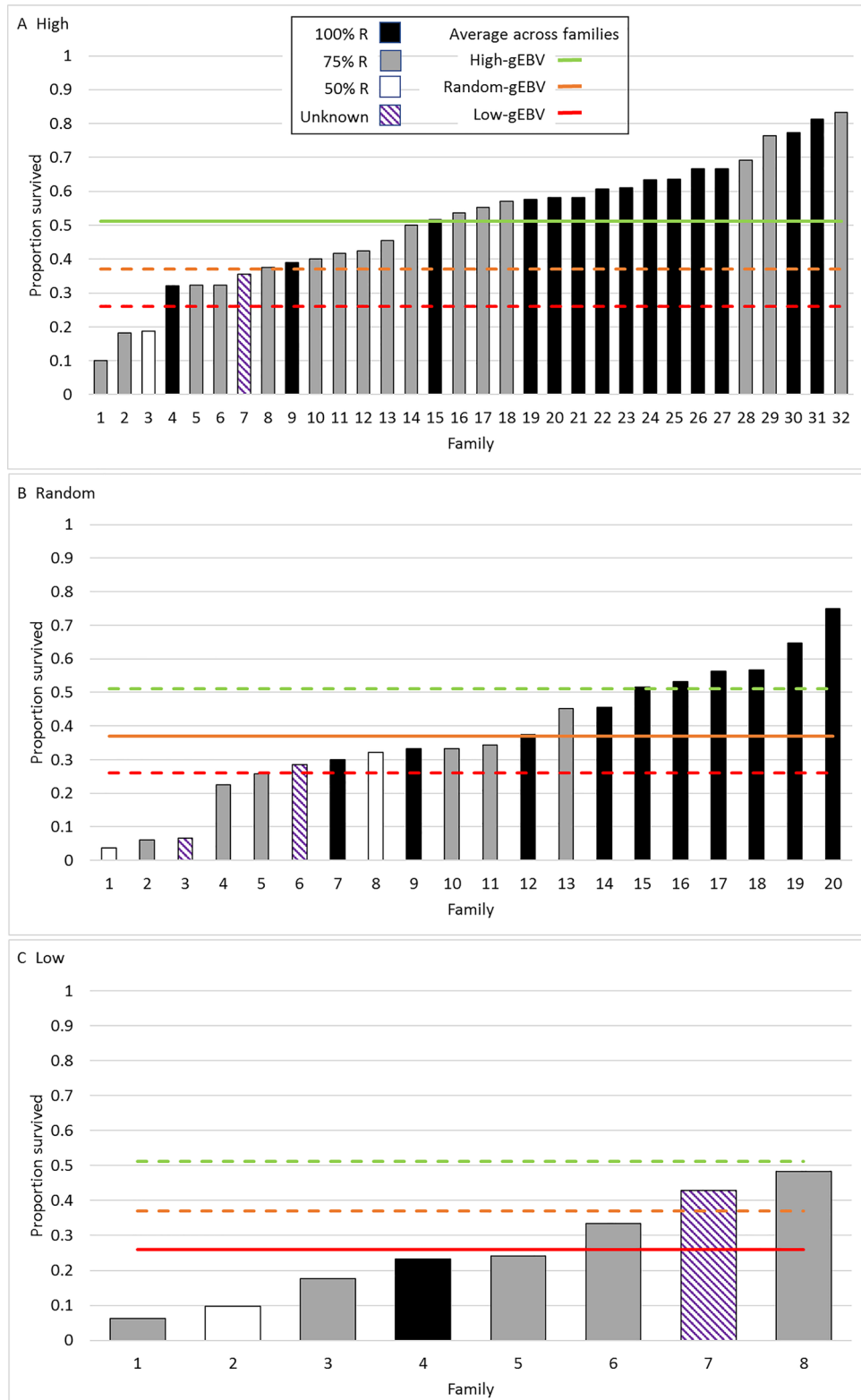


Figure 3. Proportion surviving in each challenged family for high (A) random (B) and low (C) gEBV selected G_1 groups. Shading of the bar plots indicates % of resistant line ancestry in that family (black 100%, grey 75%, white 50% and stripes when the population origin of one parent was unknown). Horizontal lines show the average proportion across families surviving for the high (green), random (orange) and low (red) gEBV G_1 groups. Solid lines relate to the G_1 population that is represented in each plot.

Model	Phenotypic variance	Genetic variance	Heritability
With selection group fitted	0.26	0.11	0.41 ± 0.09
Without selection group fitted	0.27	0.13	0.47 ± 0.09

Table 3. Estimated variance components and heritabilities for “days or alive” from analysing the G_1 -data with or without selection group fitted as a fixed effect.

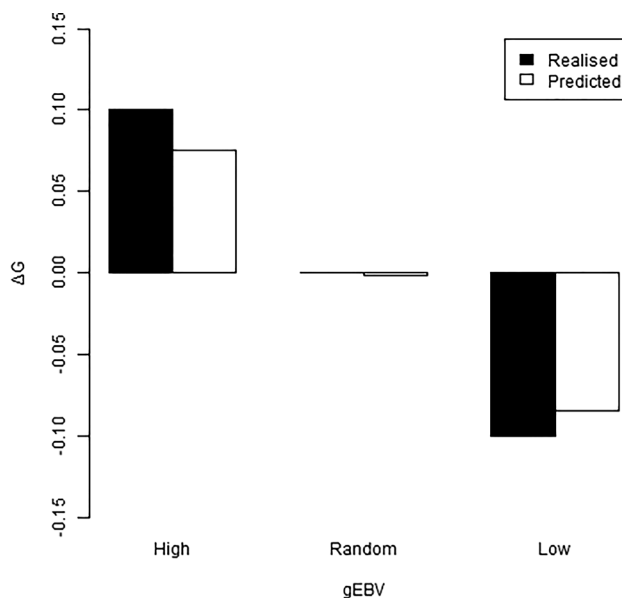


Figure 4. Effect of genomic selection showing realised and predicted genetic gain (ΔG) in the DOA trait for high, random and low gEBV selection groups.

infecting tissue (which was obtained from a population of resistant animals). It is not clear how ending the test before mortalities ceased would affect the genetic parameter estimates for DOA and DS, however this discrepancy could explain the heritability differences between G_0 and G_1 . If we estimate expected response based on the heritability estimate from the G_1 -population, this estimate is similar to the estimated regression coefficient for selection group, but still below the actual realized genetic gain, measured as the phenotypic difference between the selected group and the control group.

In general, our estimates for heritability of WSSV resistance based on genomics were greater than the estimates obtained from previous conventional evaluations^{12–16}. The higher heritability found in G_1 than in G_0 could be due to a difference in true heritability, because of differing test conditions and/or differing genetic composition between the populations. It could however also be due to estimation inaccuracies with one estimate being closer to the true heritability than the other. Possible non-genetic effects common to full-sibs were not fitted, due to confounding with the genetic effect, and could cause over-estimation of the genetic variance. However, as the tested shrimp were young and mixed as soon as they reached a suitable size for tagging, we expect the common effect of full-sib to be small. The test conditions were controlled and standardized and hence should not have caused large differences between the two consecutive generations, although random variation in disease pressure and other environmental factors could still exist. The structures of the two populations tested were however quite different, which could affect the heritability. The G_0 -population consisted of a mix of purebred resistant line animals and crossbred resistant-by-susceptible line animals. The difference in susceptibility between these two genetic groups were fitted as a fixed effect in the analysis model, but the genetic variance was assumed to be the same within these two populations, even though natural selection for resistance in the purebred line might have reduced heritability. The G_1 consisted of animals with various levels of strain composition, and the G_1 also consisted of sub-groups that were selected for high or low resistance. The studied populations would therefore be expected to have unusually high genetic variance for resistance as they constituted a combination of strains with different selection history and due to the application of divergent genomic selection. Statistical corrections made to account for strain in G_0 and selection group in G_1 lead to downward bias in the heritability estimate, as the between-group genetic variance is fitted as a fixed effect. The argument for making this correction is that the genetic variance within the mixed population is not representative of most breeding populations. Analyses of G_1 with and without correction for selection group shows that roughly 10% of the genetic variance exists between the groups, while 90% of the genetic variance is contained within group. This high within group genetic variance, and generally higher heritability with genomics than with conventional methods, suggests that further genomic selection to produce a G_2 should achieve similar levels of genetic progress as documented for the G_1 .

Our finding of similar estimates of heritability for the overall G_0 population, pure R-line animals and crossed RxS-line animals, and our findings of an intermediate level of WSSV resistance in crossed RxS-line animals between that found in the pure R and pure S-lines, indicates additive effects of alleles controlling resistance for both the DOA and DS traits. In this study the role of dominance or heterosis could not be elucidated as the pure S-line could not be compared in the same test (the S-line being highly susceptible). Hence, we could not identify possible non-additive genetic effects from these experiments.

Several factors were identified that make genomic selection for WSSV resistance in *L. vannamei* attractive. First, the methodology based on genomics identified genetic variation in resistance in the initial G_0 population. Our previous trials characterising the starting populations show that hybrid lines contain a level of WSSV resistance somewhere between that of pure R-line and S-line animals, indicating possible additive effects of alleles promoting resistance. Second, the detection of significant differences between the selection groups shows that selection increased WSSV resistance in this first generation of selection. Third, broad variation was detected within groups, particularly in the high-gEBV group. This most likely reflected that the proportion of the R-line in the high-gEBV group which could, in theory, range from 0 to 100% (parents being a mixture of pure R-line and RxS-line individuals). The significant genetic variance remaining in the G_1 , including when selection group was accounted for in the model, coupled with higher heritability in G_1 than in G_0 , indicating no loss of genetic variation for resistance due to selection, indicate that further improvement through selection should be possible with a continued selection response over several generations. Fourth, the selection response was greater than expected from the selection intensity applied, estimated heritability and additive genetic variance. Finally, the 13% improved survival in one generation is greater than with conventional selection methodologies and the genetic gain is commercially interesting.

The survival levels reported here of slightly more than 60% in the best populations may not appear as satisfactory for commercial production. However, when vaccination is used as means of protecting populations against epidemics, the population can be effectively protected without vaccinating every individual due to the herd effect, see for example³⁷. Anche et al.³⁸ suggest that the basic reproduction ratio, R , is the key parameter determining the risk and severity of infectious diseases. Furthermore, they suggest that genetic improvement of R_0 could be used for control of infectious diseases in host populations. Hence, the herd effect of resistant animals could have a similar effect to that of partial vaccination if resistance is associated with a reduced R . In the challenge tests for resistance we used, the animals were not individually infected but rather were maintained in an environment which encouraged reproduction of the disease. Thus, we suggest that not only were we selecting for resistance, but also for a low, desirable, R . Re-analysis of data on WSSV and Taura (TSV) epidemics in shrimp, suggest that the R of Taura was reduced by breeding to a level that contained the disease, whilst in the case of WSSV the reduction was not enough to control the disease³⁹. Furthermore, they suggest that heritability of resistance is not enough on its own to evaluate the economic gain of selection for resistance to a specific disease, because it ignores the fact that resistant animals no longer infect other animals. The inoculum pressure with commercial production is almost certainly much lower than in our challenge tests in which several families now show survival of 70%. Anche et al.³⁸ point out that levels of about 70% survival were sufficient in the case of Taura to maintain the disease under control. Hence, in a similar manner that disease epidemics can be slowed or stopped when a certain level of the population is vaccinated, a certain level of resistant animals in a population, coupled with practices that disfavour the disease, may reduce R to close to one and be sufficient to control WSSV in commercial populations. We have shown that by using genomic selection we can rapidly increase the level of resistance and hopefully will be able to use this tool to offer growers shrimp populations that can survive and produce in the presence of WSSV without having to totally eradicate the causal organism.

In applying genomic selection for disease resistance we also need to consider that an unfavourable genetic correlation between growth rate and WSSV resistance exists^{16,40} and that a faster growth rate is desirable for shrimp aquaculture. Simultaneous improvement of both traits could be achieved by combining a fast-growing S-line with high resistance R-line in a multiplier design for commercial broodstock supply. It may also be possible to use genomic selection in such a way that both traits are positively selected, but this needs further investigation. Another consideration for application is affordability and accessibility. The need for genotyping will add costs to the breeding program, but it may also reduce the need for systematic pedigree recording, which is costly and also induces inefficiencies in the selection process due to separate rearing, as that can be derived from the genotypes. Many SNP markers have already been identified for *L. vannamei*⁴¹ and technologies for ultra-high throughput genotyping of dense markers are becoming more readily available and less costly to apply⁴².

The other major difficulty faced when selecting for WSSV resistance in shrimp is that it is practically impossible to use survivors of challenge tests as breeding candidates as they can vertically transmit the virus and infect the offspring population. Furthermore, the practical difficulties of cleaning survivors to incorporate them into breeding programs which produce WSSV-specific pathogen free broodstock are immense. The practical impossibility of using survivors, and the lack of information on survival performance of breeding candidates and their true genetic relationships with candidates (pedigree relationships have been the standard for previous studies), limits the accuracy of conventional phenotypic selection. These problems also limit the selection intensity and the genetic variance that can be utilized. Greater accuracy for the prediction of breeding values can be gained by obtaining genomic relationships from the dense SNP genotype data. Genome-wide genetic markers make it possible to estimate the degree of genomic relationship between test-animals and individual candidate breeding animals within families. Whereas conventional pedigree relationships estimate the average relationship across the genome, dense SNP data enables estimation of genetic relationships at each position in the genome. The genomic relationship between every animal, and performance data from challenge tested animals, is used to estimate individual genomic breeding values: thus, combined family and individual selection is possible with genomic selection. We contend that the ability to evaluate the individual breeding value through genomic selection is the main reason why a higher rate of genetic gain was obtained with genomic selection (13% improved survival per

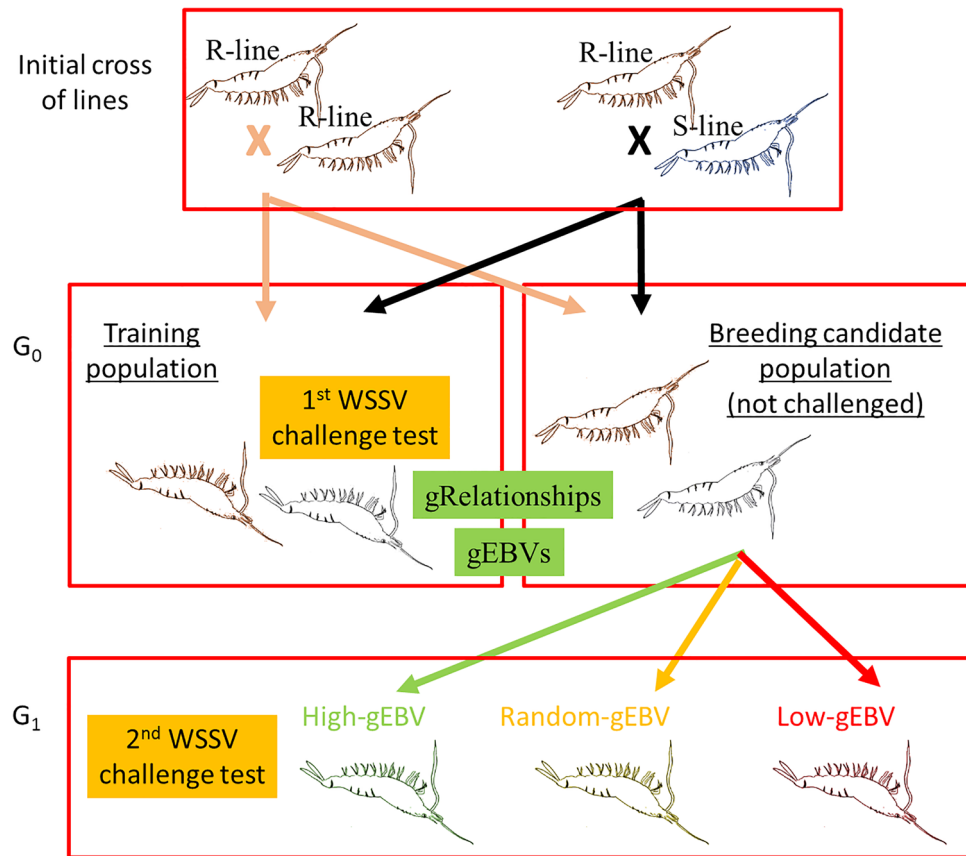


Figure 5. Origin of shrimp for the experiment. The animals were split into training and candidate breeding populations for estimation of white spot syndrome virus (WSSV) resistance genomic estimated breeding values (gEBV) at generation G₀. G₀ breeding candidates were selected and mated to produce high, random and low gEBV G₁ groups whose performance (survival after an experimental challenge test) was compared for the final evaluation of the power of genomic selection.

generation) than that achieved with conventional phenotypic selection of WSSV resistance in *L. vannamei* 1.7% DOA, 6.3% survival and 6.5% survival selection responses per generation^{16,18,19}.

Conclusion

We have shown that significant useful genetic improvement for WSSV resistance can be achieved in a breeding program for *L. vannamei* by applying genomic selection. Compared with conventional methods the use of genomic data resulted in higher heritability estimates and improved the accuracy of selection to levels that are commercially relevant. From a genomic selection experiment, and from cross-validation accuracy predictions, we demonstrated that the moderate marker density of ~18 K used was sufficient for accurate genomic breeding value prediction. This is the first study in *L. vannamei* to demonstrate realized genetic gain from genomic selection. Genomic selection is particularly promising for selection programs that challenge live animals with infectious agents when survivors cannot later be used as breeders.

Materials and methods

Animals used for the study. Two populations of whiteleg shrimp (*Litopenaeus vannamei*) that had been selectively bred for several generations by Benchmark Genetics (formerly Ceniagua) in Colombia were utilised for the study (Fig. 5). The first population was derived from a breeding operation on the Atlantic Coast of Colombia that commenced in 1997. This population has not been exposed to WSSV and is herein referred to as the susceptible line (S-line). Combined family and within-family selection of this population focussed on fast growth, resistance to Taura Syndrome Virus, general pond survival and robustness for 16 generations under strict biosecurity protocols. The second population, herein referred to as the resistant line (R-line), derived from a Pacific Ocean population in 2008, was mass selected with high intensity (1×10^{-4}) for survival in the presence of WSSV for 7 generations. The Atlantic and Pacific populations were maintained in isolation on their respective coasts of Colombia. All animals entering the study were considered as naïve (uninfected with the virus). Animals from the S-line were PCR tested every 3 months for all OIE listed pathogens and were found to be negative for all listed pathogens for more than 4 years. Before the challenge test a representative sample of the animals from the R-line was PCR tested for WSSV⁴³ and found to be negative.

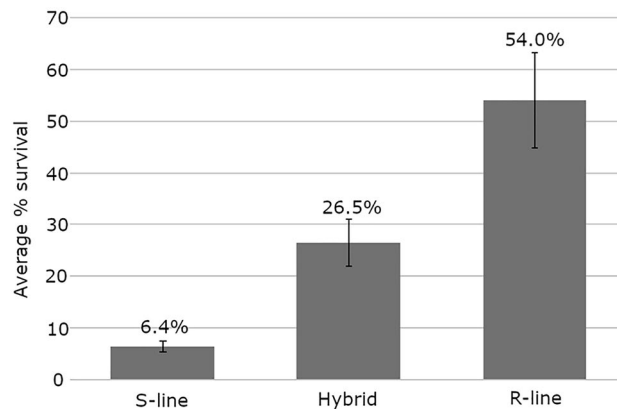


Figure 6. Comparison of average survival in the WSSV susceptible, resistant and hybrid starting populations.

The overall response of the Atlantic and Pacific starting populations to WSSV infection was known from previous trials. Past challenge test trials indicated that the S-line contains little phenotypic variation for resistance (with family survival rates in the range of 0–6%). Deaths occur very rapidly after infection and no animals survive the challenge longer than 4 days. In contrast, variation in post-challenge survival in the R-line ranges from 20 to 54% 15–20 days after infection. The WSSV resistance of Hybrid F1 offspring, produced by crossing R- and S-lines, is intermediate between that of pure lines (Fig. 6).

The first challenge test used 751 pure R-line and 696 crossbred R- and S-line shrimp. These animals were used for training purposes to estimate genomic breeding values. Breeding candidates (not challenged with WSSV) consisted of offspring derived from 34 RxR-line and 24 RxS-line parental crosses (the pure R-line and RxS-line crosses being siblings of the challenged training animals). For the purposes of this experiment this starting population of animals, including test animals and breeding candidates, is referred to as the G_0 population (Fig. 5).

G_0 WSSV challenge tests. Juvenile G_0 shrimp of average weight 3 g were randomly divided into two tanks filled with 4 tons of artificial seawater at 30 ppt salinity at 26 °C temperature and infected with WSSV in a challenge test initiated on April 18th 2018, using minced infected tissue. The tissue used (with a viral load of 1×10^6 copies/ μ g of DNA) was weighed and administered to the tanks at a ratio of 3% of the total shrimp biomass over two consecutive days. Mortalities were recorded and removed hourly until May 10th (22 days) to ensure that no animals were cannibalised and hence not registered as dead. Twelve days post infection, the shrimp were pooled into one single tank to avoid loss of infection pressure due to reduced shrimp density. Infection was confirmed by histopathology⁴⁴ and nested PCR from pleopod samples⁴⁵. Date and hour of mortality was recorded to provide data on two traits, dead or alive after 22 days (DOA) which is a binary trait (assessed 22 days post-infection) and days survival (DS). For the purpose of analysing DS it was assumed that survivors died the day after the trial ended.

Collection of tissue samples and genotyping. Shrimp tissue samples, pleopod or muscle, were collected from all dead animals and survivors in April/May 2018 and November 2018, immersed in 95% ethanol at 4 °C and stored in 1.5 ml centrifuge tubes. SNP genotyping was performed by Neogen (USA) in Nov 2018. DNA was extracted using a BioKits DNA Extraction Kit (Neogen, USA, which utilises proprietary magnetic beads with a high affinity for DNA) and samples were genotyped using a custom 18 K Illumina SNP array designed by Neogen USA (described below).

Single nucleotide polymorphisms. All shrimp were genotyped for 18,643 SNPs. A *L. vannamei* Illumina custom array was developed for SNP genotyping (Neogen USA). A total of 8,967 SNPs were previously identified and published⁴¹ while another 10,323 SNPs were identified using mRNAseq re-sequencing of pooled hepatopancreas, gill and pleopod tissue from 20 naïve generation-3 R-line and 20 naïve generation-3 S-line individuals. Shrimp were dissected under sterile conditions and samples completely immersed in RNAlater (Qiagen Germany and stored at –80 °C to preserve RNA until it could be purified and prepared for RNA-seq (previously described for *P. monodon*, Baranski et al.⁴⁶). No RNA normalization was performed before the samples were pooled (creating one R-line and one S-line pool) and processed with 72 cycles of paired-end sequencing in two separate lanes on a Genome Analyser II (Illumina USA) according to manufacturer's instructions. The CLC Assembly Cell “find variations” program was used to detect putative SNPs in the *L. vannamei* assembly as described for *P. monodon*⁴⁶. Putative SNPs that were identified in homologous *L. vannamei* contigs were filtered to reject those with total sequence depth less than 51, minor allele frequency less than 0.1 and those lacking variation within the R-line or within the S-line (ie SNPs showing fixed allele differences between the two lines were not accepted). In cases where multiple SNPs per contig remained after filtering, the first SNP on the list was the only one accepted. Finally, the basic local alignment search tool (BLAST, National Centre for Biotechnology Information, USA) was used to check all flanking SNP sequences for matches against themselves and against the 8,967 SNPs previously published in order to filter out any repeated SNPs. Both the flanking sequence and the

published SNPs were made into searchable BLAST databases using the *makeblastdb* command and matches were identified using the *blastn* command with e-value 10^{-11} .

Post-genotyping quality control removed SNPs that were not in Hardy Weinberg equilibrium proportions, with minor allele frequency less than 0.05 or call rate less than 95%.

Genomic breeding value and parameter estimation. A genomic relationship matrix (G) between all the animals in G_0 was estimated as described by VanRaden⁴⁷. Genetic evaluation was performed by running an animal model in a multivariate mixed model package DMU⁴⁸. All animals challenged with the virus and all breeding candidates were included in the estimation. The traits DOA and DS were analysed separately, using the animal model:

$$y = Xb + Zu + e$$

where y is a vector of phenotypes (DOA or DS), b is a vector of estimated fixed effects of cross, fitted with two levels: purebred R or R x S cross, u is a vector of estimated breeding values (gEBV) for test-animals and candidates, $u \sim N(0, G\sigma_a^2)$, where G is the genomic relationship matrix and σ_a^2 is the additive genetic variance of y , and e is a vector of random residuals, assumed to be normally distributed with variance σ_e^2 . X and Z are design matrices to link each observation to the correct fixed effect category and animal. Variance components for both traits were estimated separately in single trait models. Both traits were analysed as continuous traits, i.e. analysed on the observed scale only.

Accuracy of genomic EBV prediction. From the G_0 data we estimated the accuracy of genomic EBV prediction using cross validation. Each animal was allocated a number (1–10, creating 10 randomly selected non-overlapping validation subsets), and the analysis was repeated 10 times, each time masking the phenotypes for all traits for one of the validation subsets. Accuracy (Acc) was estimated for each trait as $Acc = \frac{R_{gEBV,y}}{h}$, where $R_{gEBV,y}$ is the correlation between the phenotype and gEBV, using, for each animal, gEBVs from the run in which its phenotype was masked, and h was the square root of the heritability, estimated from the full dataset.

Power analysis. The response to selection can be predicted based on the estimated genetic parameters, given a certain selection intensity. A power analysis was applied to design a selection experiment suitable to document genetic gain and validate the genetic parameters within the limitations of available infrastructure. Predictions of the power for evaluating genomic selection were performed for the DOA trait, because this trait was thought to have the highest economic relevance. The maximum total number of full-sibling families was limited to 60. The number of offspring tested from each family was set at 25 to make sure that the total number of test animals did not exceed the capacity of the test facility. Different options for production of selection groups were compared. High gEBV groups were produced by crossing G_0 individuals with the highest gEBV values. Similarly, low gEBV groups were made from crosses between individuals with the lowest gEBVs. The “random” family group was created by crossing remaining individuals at random. Expected differences between comparisons of the random- and high-gEBV groups, and between the low- and high-gEBV groups (as the trait was selected in both directions) were based on estimated genetic parameters and the selection intensity used. The power analysis assumed that observations were independent (ie. family structure was ignored) and power was estimated using the formula $t_\beta = t_{0.05} - \Delta G / se(\Delta G)$, where $t_{0.05}$ was the critical value needed to reject the null hypothesis of no difference between the selection groups (assuming a standard normal distribution) and ΔG was the expected difference between 2 selection groups. ΔG was estimated using a univariate breeder's equation $\Delta G = h^2 * S$, where h^2 is the heritability estimated from the G_0 -test animals and S is the phenotypic selection differential applied when producing the families. $se(\Delta G)$ was the standard error of the contrast between two selection groups, estimated as $(\frac{1}{n1} + \frac{1}{n2})\sigma^2$, where $n1$ and $n2$ are the number of animals in each group and σ^2 is the phenotypic variance estimated from the challenge test of G_0 . The power was estimated as the area under the cumulative standard normal distribution from $-\infty$ to $1 - t_\beta$.

Selection experiment. Breeders were selected from the G_0 breeding candidate pool based on DOA gEBVs estimated from WSSV challenge tests run in 2018. Males and females, respectively, were ranked on gEBV, and both tails of the ranked list were selected. The top 60 males and 60 females were selected as possible high resistant breeding candidates and the bottom 20 males and 20 females were selected as possible low resistant breeding candidates. Animals not selected for either of these lists, and untested animals from the G_0 , were available as possible “random selection” parents. Sixty families were produced, 32 from matings between high resistant animals, 8 from matings between low resistant males and females and 20 from matings between randomly selected animals (Table 4, close to the 30 high, 20 random and 10 low power test with a value of 0.96 on a 0–1 scale). The average genomic breeding value of the parents were 0.25 for the resistant, -0.005 random and -0.28 for the susceptible group. Mating between likely full-or half-siblings (pairs having close genomic relationship, >0.1) was avoided. The offspring are for the purposes of this experiment referred to as the G_1 population.

Mate lists were generated, and animals reproduced by artificial insemination in March 2019 at Ceniaca's hatchery in Tumaco, Colombia. The artificial insemination of select broodstock, involving extraction of spermatozoa from male candidates and transfer of the spermatid mass to artificially inseminate female candidates, is described by¹². Females were placed in individual hatching tanks. After hatching a random sample of about 5,000 nauplii from each family were stocked in separate 50L-larvae-culture tanks. Larvae were fed a mixed diet of *Chaetoceros* sp. *Artemia* sp. and micro pellets. At the post-larvae 10 stage (PL10) a random sample of 200 post-larvae per family was transferred to separate tanks at a 100/m² density for on-growing until they reached

Group	Female gEBV		Male gEBV		gRelat	
	Average	Range	Average	Range	Average	Range
Low	0.30	0.28:0.33	0.30	0.28:0.33	0.00	-0.06:0.07
Random	-0.04	-0.14 :0.05	0.00	-0.18:0.18	-0.03	-0.12:0.05
High	-0.28	-0.48:-0.21	-0.30	-0.58:-0.20	0.01	-0.12:0.11

Table 4. Choice of mate pairs used to produce the low, random and high gEBV groups showing average and range in genomic estimated breeding values (gEBV) and genomic relationship values (gRelat) for selected G_0 individuals. Two females in the low gEBV group and two males in the high gEBV group were mated twice to produce half-siblings.

a body size suitable for elastomer-tagging (approximately 1 g). Tagged animals were then sent to the challenge test facilities in Bogotá Colombia in May 2019.

The 1885 G_1 -animals were evaluated in two tanks. Family size ranged from 12 to 53 records (average 31). The challenge test was continued for 23 days post-infection. The data from the challenge test was used to estimate genetic gain and to re-estimate genetic parameters, to validate the results from G_0 . The G_1 -data was analysed with a sire-dam-model in Asreml⁴⁹ where,

$$DOA = X\beta + Zu + e$$

where DOA was a vector of the binary dead or alive phenotype. β was the fixed effects of tank and selection group (high, random or low resistance) and X was the design matrix to link observations to the fixed effect categories. u was a matrix of random effects of sire and dam, $u \sim G * \frac{1}{2}\sigma_a^2$ where σ_a^2 was the additive genetic variance, G was the genomic relationship matrix between the parents (the G_1 test shrimp were not genotyped), and Z was the design matrix assigning DOA records to the correct sire and dam. Selection group was fitted as a fixed regression rather than categorical because despite the different number of parents selected for mating to produce the high- and low-gEBV selection groups, selection intensity was of a similar strength in both directions. Expected genetic gain from the realized selection intensity and heritability in G_0 was estimated by average breeding values of the parents multiplied by estimated heritability. Realized and predicted genetic gain were compared.

Data availability

All raw and processed sequence data generated in this study have been deposited in the NCBI Short Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under BioProject accession PRJNA625155.

Received: 9 June 2020; Accepted: 12 November 2020

Published online: 25 November 2020

References

- Flegel, T. W. Detection of major penaeid shrimp viruses in Asia, a historical perspective with emphasis on Thailand. *Aquaculture* **258**, 1–33. <https://doi.org/10.1016/j.aquaculture.2006.05.013> (2006).
- Hernandez-Llamas, A., Ruiz-Velazco, J. M. J. & Gomez-Munoz, V. M. Economic risk associated with white spot disease and stochastic variability in economic, zootechnical and water quality parameters for intensive production of *Litopenaeus vannamei*. *Rev. Aquac.* **5**, 121–131. <https://doi.org/10.1111/raq.12008> (2013).
- Cock, J., Salazar, M. & Rye, M. Strategies for managing diseases in non-native shrimp populations. *Rev. Aquac.* **9**, 211–226. <https://doi.org/10.1111/raq.12132> (2017).
- Lightner, D. V. Virus diseases of farmed shrimp in the Western Hemisphere (the Americas): A review. *J. Invertebr. Pathol.* **106**, 110–130. <https://doi.org/10.1016/j.jip.2010.09.012> (2011).
- Lightner, D. V. *Exclusion of Specific Pathogens for Disease Prevention in a Penaeid Shrimp Biosecurity Program* (World Aquaculture Society, New York, 2003).
- Lotz, J. M. Viruses, biosecurity and specific pathogen-free stocks in shrimp aquaculture. *World J. Microbiol. Biotechnol.* **13**, 405–413. <https://doi.org/10.1023/a:1018572132529> (1997).
- Rahman, M. M. *et al.* Benefits of stocking white spot syndrome virus infection free shrimp (*Penaeus monodon*) post larvae in extensive ghers of Bangladesh. *Aquaculture* **486**, 210–216. <https://doi.org/10.1016/j.aquaculture.2017.12.024> (2018).
- Vazquez, L., Alpuche, J. G., Agundis, C., Pereyra-Morales, A. & Zenteno, E. Immunity mechanisms in crustaceans. A review. *Innate Immun.* **15**, 179–188 (2009).
- Alvarez-Lee, A., Martínez-Díaz, S. F., Gutiérrez-Rivera, J. N. & Lanz-Mendoza, H. Induction of innate immune response in whiteleg shrimp (*Litopenaeus vannamei*) embryos. *Dev. Comp. Immunol.* **105**, 103577. <https://doi.org/10.1016/j.dci.2019.103577> (2020).
- Chang, Y. H., Kumar, R., Ng, T. H. & Wang, H. C. What vaccination studies tell us about immunological memory within the innate immune system of cultured shrimp and crayfish. *Dev. Comp. Immunol.* **80**, 53–66. <https://doi.org/10.1016/j.dci.2017.03.003> (2018).
- Phuoc, L. H. *et al.* Priming the immune system of Penaeid shrimp by bacterial HSP70 (DnaK). *J. Fish Dis.* **39**, 555–564. <https://doi.org/10.1111/jfd.12388> (2016).
- Gitterle, T. *et al.* Genetic (co)variation in harvest body weight and survival in *Penaeus (Litopenaeus) vannamei* under standard commercial conditions. *Aquaculture* **243**, 83–92 (2005).
- Gitterle, T. *et al.* Optimization of experimental infection protocols for the estimation of genetic parameters of resistance to White Spot Syndrome Virus (WSSV) in *Penaeus (Litopenaeus) vannamei*. *Aquaculture* **261**, 501–509 (2006).
- Gitterle, T., Odegard, J., Gjerde, B., Rye, M. & Salte, R. Genetic parameters and accuracy of selection for resistance to White Spot Syndrome Virus (WSSV) in *Penaeus (Litopenaeus) vannamei* using different statistical models. *Aquaculture* **251**, 210–218. <https://doi.org/10.1016/j.aquaculture.2005.06.008> (2006).
- Trang, T. T., Hung, N. H., Ninh, N. H., Knibb, W. & Nguyen, H. Genetic variation in disease resistance against white spot syndrome virus (WSSV) in *Litopenaeus vannamei*. *Front. Genet.* <https://doi.org/10.3389/fgene.2019.00264> (2019).

16. Trang, T. T., Hung, N. H., Ninh, N. H. & Nguyena, N. H. Selection for improved white spot syndrome virus resistance increased larval survival and growth rate of Pacific Whiteleg shrimp *Litopenaeus vannamei*. *J. Invertebr. Pathol.* <https://doi.org/10.1016/j.jip.2019.107219> (2019).
17. Caballero-Zamora, A. *et al.* Genetic parameters for body weight and survival in the Pacific white shrimp (*Litopenaeus vannamei*) affected by a White Syndrome Virus (WSSV) natural outbreak. *Aquaculture* **447**, 102–107 (2015).
18. Gitterle, T. *et al.* Response to multi trait selection for harvest weight, overall survival, and resistance to white spot syndrome virus (WSSV) in *Penaeus* (*Litopenaeus*). *Aquaculture* **272**, S262 (2007).
19. Huang, Y., Yin, Z., Weng, S., He, J. & Li, S. Selective breeding and preliminary commercial performance of *Penaeus vannamei* for resistance to white spot syndrome virus (WSSV). *Aquaculture* **364–365**, 111–117 (2012).
20. Yu, Y. *et al.* Gene set based association analyses for the WSSV resistance of Pacific white shrimp *Litopenaeus vannamei*. *Sci. Rep.* **7**, 66. <https://doi.org/10.1038/srep40549> (2017).
21. Robinson, N. A. *et al.* QTL for white spot syndrome virus resistance and the sex-determining locus in the Indian black tiger shrimp (*Penaeus monodon*). *BMC Genomics* **15**, 731. <https://doi.org/10.1186/1471-2164-15-731> (2014).
22. Hayes, B. & Goddard, M. E. The distribution of the effects of genes affecting quantitative traits in livestock. *Genet. Sel. Evol.* **33**, 209–229 (2001).
23. Marques, M. R. F. & Barracco, M. A. Lectins, as non-self-recognition factors, in crustaceans. *Aquaculture* **191**, 23–44. [https://doi.org/10.1016/s0044-8486\(00\)00417-8](https://doi.org/10.1016/s0044-8486(00)00417-8) (2000).
24. Zhao, Z.-Y. *et al.* Profiling of differentially expressed genes in hepatopancreas of white spot syndrome virus-resistant shrimp (*Litopenaeus vannamei*) by suppression subtractive hybridisation. *Fish Shellfish Immunol.* **22**, 520–534 (2007).
25. He, N. H., Qin, Q. W. & Xu, X. Differential profile of genes expressed in hemocytes of white spot syndrome virus-resistant shrimp (*Penaeus japonicus*) by combining suppression subtractive hybridization and differential hybridization. *Antiviral Res.* **66**, 39–45. <https://doi.org/10.1016/j.antiviral.2004.12.010> (2005).
26. Pan, D., He, N. H., Yang, Z. Y., Liu, H. P. & Xu, X. Differential gene expression profile in hepatopancreas of WSSV-resistant shrimp (*Penaeus japonicus*) by suppression subtractive hybridization. *Dev. Comp. Immunol.* **29**, 103–112. <https://doi.org/10.1016/j.dci.2004.07.001> (2005).
27. García, J. C., Reyes, A., Salazar, M. & Granja, C. B. Differential gene expression in white spot syndrome virus (WSSV)-infected naïve and previously challenged Pacific white shrimp *Penaeus* (*Litopenaeus*) *vannamei*. *Aquaculture* **289**, 253–258 (2009).
28. Meuwissen, T. Genomic selection : marker assisted selection on a genome wide scale. *J. Anim. Breed. Genet.* **124**, 321–322 (2007).
29. Meuwissen, T., Hayes, B. & Goddard, M. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829 (2001).
30. Meuwissen, T., Hayes, B. & Goddard, M. Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* **6**, 6–14. <https://doi.org/10.2527/af.2016-0002> (2016).
31. Gitterle, T., Ødegård, J., Gjerde, B., Rye, M. & Salte, R. Genetic parameters and accuracy of selection for resistance to White Spot Syndrome Virus (WSSV) in *Penaeus* (*Litopenaeus*) *vannamei* using different statistical models. *Aquaculture* **251**, 210–218 (2006).
32. Yoshida, G. M., Carvalheiro, R., Rodriguez, F. H., Lhorente, J. P. & Yanez, J. M. 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> (2019).
33. Dai, P. *et al.* Evaluation of the utility of genomic information to improve genetic evaluation of feed efficiency traits of the Pacific white shrimp *Litopenaeus vannamei*. *Aquaculture* **527**, 735421. <https://doi.org/10.1016/j.aquaculture.2020.735421> (2020).
34. Wang, Q. C. *et al.* 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> (2019).
35. Wang, Q. C. *et al.* Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp *Litopenaeus vannamei*. *BMC Genet.* **18**, 66. <https://doi.org/10.1186/s12863-017-0507-5> (2017).
36. Vallejo, R. L. *et al.* 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> (2018).
37. John, T. J. & Samuel, R. Herd immunity and herd effect: New insights and definitions. *Eur. J. Epidemiol.* **16**, 601–606. <https://doi.org/10.1023/a:1007626510002> (2000).
38. Anche, M. T., de Jong, M. C. & Bijma, P. On the definition and utilization of heritable variation among hosts in reproduction ratio R-0 for infectious diseases. *Heredity* **113**, 364–374. <https://doi.org/10.1038/hdy.2014.38> (2014).
39. Janssen, K. & Bijma, P. The economic value of R0 for selective breeding against microparasitic diseases. *Genet. Sel. Evol.* **52**, 3. <https://doi.org/10.1186/s12711-020-0526-y> (2020).
40. Gitterle, T. *Genetic analyses for resistance to White Spot Syndrome Virus (WSSV), harvest body weight and pond survival in the Pacific white shrimp Penaeus* (*Litopenaeus*) *vannamei*. (Dr. Scientiarum Thesis 2005:9), Agricultural University of Norway, Department of Animal and Aquacultural Sciences, Aas Norway, (2005).
41. Jones, D. B. *et al.* A comparative integrated gene-based linkage and locus ordering by linkage disequilibrium map for the Pacific white shrimp *Litopenaeus vannamei*. *Sci. Rep.* **7**, 66. <https://doi.org/10.1038/s41598-017-10515-7> (2017).
42. Guppy, J. L. *et al.* The state of “Omics” research for farmed penaeids: Advances in research and impediments to industry utilization. *Front. Genet.* **9**, 27. <https://doi.org/10.3389/fgene.2018.00282> (2018).
43. Lo, C. F. *et al.* Detection of baculovirus associated with white spot syndrome (WSBV) in penaeid shrimps using polymerase chain reaction. *Dis. Aquat. Org.* **25**, 133–141. <https://doi.org/10.3354/dao025133> (1996).
44. Bell, T. A. & Lightner, D. V. *A Handbook of Normal Penaeid Shrimp Histology* (World Aquaculture Society, New York, 1988).
45. Lo, C. F. *et al.* Detection and tissue tropism of white spot syndrome baculovirus (WSBV) in captured brooders of *Penaeus monodon* with a special emphasis on reproductive organs. *Dis. Aquat. Org.* **30**, 53–72. <https://doi.org/10.3354/dao030053> (1997).
46. Baranski, M. *et al.* The development of a high density linkage map for black tiger shrimp (*Penaeus monodon*) based on cSNPs. *PLoS ONE* **9**, e85413 (2014).
47. VanRaden, P. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
48. DMU: a user's guide. A package for analysing multivariate mixed models. Version 6, release 5.0 (University of Aarhus, Faculty of Agricultural Sciences (DJF). (Department of Genetics and Biotechnology, Tjele, 2010).
49. Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. & Thompson, R. (VSN International Ltd, Hemel Hempstead, 2015).

Acknowledgements

The research was funded by the Research Council of Norway (project number 244131) and supported by the Norwegian Institute of Food, Fisheries and Aquaculture (Nofima) and Benchmark Genetics. High throughput sequencing was performed by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and Oslo University Hospital and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the South-eastern Regional Health Authorities.

Author contributions

M.L., N.R., J.C., M.S., R.B. and M.R. wrote the paper and prepared the figures. M.L., R.B., M.S., S.V., E.E., A.S., J.C., M.R. and N.R. designed the research. M.S., E.E. and A.S. performed the research. M.L., R.B., S.V., M.S. and N.R. analysed the data. All authors reviewed and provided input for the manuscript.

Competing interests

M.L. and N.R. declare no competing interests. R.B., M.S., S.V., E.E., A.S., J.C. and M.R. are employed by Benchmark who own the selected whiteleg shrimp.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-77580-3>.

Correspondence and requests for materials should be addressed to N.A.R.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access 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 Author(s) 2020