

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/316939971>

Mate selection in aquaculture breeding using differential evolution algorithm

Article in *Aquaculture Research* · May 2017

DOI: 10.1111/are.13365

CITATIONS

0

READS

134

7 authors, including:



Grazyella Massako Yoshida
São Paulo State University

34 PUBLICATIONS 33 CITATIONS

[SEE PROFILE](#)



José Manuel Yáñez
University of Chile

53 PUBLICATIONS 348 CITATIONS

[SEE PROFILE](#)



Carlos A. L. Oliveira
Universidade Estadual de Maringá

100 PUBLICATIONS 243 CITATIONS

[SEE PROFILE](#)



Ricardo Pereira Ribeiro
Universidade Estadual de Maringá

219 PUBLICATIONS 989 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:




Seminal quality of *Colossoma macropomum* [View project](#)



Development of a SNP panel to improve the genetic resistance to Infectious Pancreatic Necrosis virus (IPNV) in salmon and trout [View project](#)

Mate selection in aquaculture breeding using differential evolution algorithm

Grazyella Massako Yoshida^{1,2}  | José Manuel Yáñez² | Carlos Antonio Lopes de Oliveira^{3,4} | Ricardo Pereira Ribeiro^{3,4} | Jean Paul Lhorente⁵ | Sandra Aidar de Queiroz^{1,4} | Roberto Carneiro^{1,4}

¹Animal Science Department, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Faculdade de Ciências Agrárias e Veterinárias (FCAV), Campus Jaboticabal, Jaboticabal, Brazil.

²Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile.

³Animal Science Department, Universidade Estadual de Maringá, Maringá, Brazil.

⁴Conselho Nacional de Desenvolvimento Científico e Tecnológico's (CNPq) researcher, Brazil

⁵Aquainnovo, Puerto Montt, Chile

Correspondence

Roberto Carneiro, Animal Science Department, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Faculdade de Ciências Agrárias e Veterinárias (FCAV), Campus Jaboticabal, Jaboticabal, Brazil.
Email: rcar@fcav.unesp.br

Funding information

Fundação de Amparo à Pesquisa do Estado de São Paulo; Conselho Nacional de Desenvolvimento Científico e Tecnológico

Abstract

An algorithm to perform mate selection in aquaculture breeding using a computational optimization procedure called "differential evolution" (DE) was applied under optimum contribution selection and mate selection scenarios, to assess its efficiency in maximizing the genetic merit while controlling inbreeding. Real aquaculture data sets with 8,782 Nile tilapias from five generations and 79,144 coho salmon from eight generations were used to optimize objective functions accounting for coancestry of parents and expected genetic merit and inbreeding of the future progeny. The mate selection results were compared with those from the realized scenario (real mates), truncation selection and optimum contribution selection method. Mate selection allowed reducing inbreeding up to 73% for Nile tilapia, compared with truncation selection, and up to 20% for coho salmon, compared with realized scenario. There was evidence that mate selection outperformed optimum contribution selection followed by minimum inbreeding mating in controlling inbreeding under the same expected genetic gain. The developed algorithm was computationally efficient in maximizing the objective functions and flexible for practical application in aquaculture breeding.

KEYWORDS

coancestry, evolutionary algorithms, inbreeding, *Oncorhynchus kisutch*, optimum contribution, *Oreochromis niloticus*

1 | INTRODUCTION

At nucleus level, the main goal of aquaculture breeding programmes is to maximize the genetic gain for a specific trait or a combination of traits, which needs to be coupled with some strategy to control inbreeding for avoiding inbreeding depression and for maintaining genetic variability and selection response in the long term. Inbreeding is a critical issue in breeding programmes, especially for aquaculture species, which usually present high fecundity capacity allowing high selection intensities and, as a consequence, a small number of parents may produce all the progeny of each generation. The reduced number of families poses a challenge to control inbreeding rate.

Different strategies and methods have been proposed to control inbreeding rate in breeding programmes, as the optimum contribution (OC) selection (Meuwissen, 1997; Woolliams & Thompson, 1994) that aims to maximize the genetic level under constrained rate of inbreeding. In OC selection, the intensity of use of each selection candidate is optimized according to an index that basically contains the expected genetic merit of the future progeny and the coancestry among selected parents. Several studies have shown the benefits of implementing OC in aquaculture breeding programmes, in terms of maximizing genetic response with controlled rate of inbreeding (D'agaro, Woolliams, Haley & Lanari, 2010; Hinrichs & Wetten, 2006; Holtsmark, Klemetsdal, Sonesson & Woolliams, 2008; Kause,

Ritola, Paananen, Wahlroos & Mäntysaari, 2005; Liu, Meuwissen, Sørensen & Berg, 2015; Nielsen, Sonesson & Meuwissen, 2011; Skaarud, Woolliams & Gjøen, 2011, 2014; Sonesson, 2005).

Another alternative to control inbreeding is the mate selection (MS) strategy (Kinghorn & Shepherd, 1999; Kinghorn, Shepherd & Woolliams, 1999). In MS, selection and mating decisions are performed simultaneously. This single-step approach is suitable to accommodate different key issues faced by animal breeders and to find a global optimum solution (Kinghorn, 2011). To implement MS, an objective function (OF) needs to be defined and maximized. For instance, Kinghorn et al. (1999) applied MS using an OF accounting for expected merit and inbreeding of the future progeny and coancestry among selected parents. In this case, inbreeding of the future progeny would also drive selection besides defining the mates, what is expected to be advantageous as the best animals to be selected can be dependent on the allocation of the mates and vice versa (Kinghorn et al., 1999). Studies conducted by Kinghorn and Shepherd (1999), Shepherd and Kinghorn (1999), Weigel and Lin (2000), Fernández, Toro and Caballero (2001), Carvalheiro, Kinghorn and Queiroz (2009, 2010), Kremer, Newman, Wilson and Kinghorn (2010) and Kinghorn (2011) showed that MS is effective in controlling inbreeding and maximizing genetic gain. To our knowledge, no study has investigated the use of MS in aquaculture breeding programmes. Furthermore, it is not clear in the literature if there is advantage of using MS in comparison with OC selection.

This study was carried out aiming to apply MS in aquaculture breeding, in real data sets under different scenarios, to assess its efficiency in maximizing the genetic merit while controlling inbreeding, as well as compare MS against truncation and OC selection strategies.

2 | MATERIAL AND METHODS

2.1 | Nile tilapia data set

PeixeGen Research Group from Universidade Estadual de Maringá, Maringá, PR, Brazil, provided the Nile tilapia (*Oreochromis niloticus*) data set used in this study. The PeixeGen Research Group run a breeding programme established in 2005 with the importation of 600 animals (~30 families) from Malaysia.

The data set contained pedigree information and standardized estimated breeding values (sEBV) for harvest weight of 8,782 Nile tilapias from five generations. The numbers of sires and dams per generation are presented in Table 1.

The animals evaluated in each generation were obtained from a hierarchical mating design (two dams per sire), producing full and half-sib families. The families were produced through natural mating in separate breeding hapas (1 m³ volume) at the Fish Farming Experimental Station of Universidade Estadual de Maringá (UEM – CODA-PAR), district of Floriano, Maringá, PR, Brazil. Inspection of the presence of spawning was performed two times per week, when the sire was removed and placed together with another dam in a different hapa. The larvae were kept with their dams until the end of the

breeding season (up to 3 months). This common environment effect was adjusted to calculate the EBVs.

At the end of the reproduction period, 100 fingerlings from each family were divided into two equal groups, which were transferred to a nursery structure and kept in hapas of 1 m³ in the earthen pond. When fingerlings reached 10 g, 50 animals of each full-sib family were individually identified by passive integrated transponder (PIT) tags, implanted in the visceral cavity.

Shortly after tagging, the animals were transferred to the grow-out system in cages located in Diamante do Norte, PR, Brazil. The volume of each cage was 6 m³ (2 × 2 × 1.5 m), and the density of fish per cage was approximately 150 fishes/m³. All grow-out cages presented almost the same number of individuals per family. The growing period was approximately equal to 7 months (from March until October). At harvest time, body weight and sex information were recorded on all fishes. EBVs for body weight were calculated using best linear unbiased predictions (BLUP) (Henderson, 1984), under an animal model, including larval and fingerling hapas as random common environment effects, and sex, cage nested to year and age at harvest as fixed effects.

2.2 | Coho salmon data set

Aquainnovo S.A provided the coho salmon (*Oncorhynchus kisutch*) data set of AquaChile Breeding Program in Chile. The breeding programme was established in 1997 and 1998 with two independent populations (odd and even), managed in 2-year reproductive cycle (Dufflocq et al., 2016). The data set used in this study contained pedigree information and standardized economic index of 79,144 coho salmon from the even population, comprising eight generations (Table 2).

The spawning was induced using hormone and all families were generated within one or 2 weeks for minimizing the effect of date of spawning in weight at harvest. A nested design with three to five females per male was adopted and an average of 103 families per year.

As described by Yáñez et al. (2014a), the eggs of each full-sib family were incubated separately, and at eyed stage 2,000 eggs of each selected family were moved to individual tanks (400 l each). Then, the progeny was individually identified using PIT tags, between November and December of each year when the fishes weigh about 5–7 g. At this stage, the fishes were transferred to smoltification sites in freshwater conditions where each full-sib family was randomly stocked in equal numbers (60–80) into two or three rearing cages. Smoltification occurred naturally at 8 months post spawning, and weight at harvest time (~3 kg) was recorded at 20–21 months of age. In the latter generation, resistance against *Piscirickettsia salmonis* was included into the breeding goal. The phenotypes were measured on sibs of the selection candidates by means of a challenge test performed as described by Yáñez, Bangera, Lhorente, Oyarzún and Neira (2013), and the trait was defined as day of death (Yáñez et al., 2014b). The genetic evaluation was carried out using a bivariate animal model, including the contemporary group of

TABLE 1 General information, inbreeding coefficient, standardized estimated breeding value and standard deviation for Nile tilapia data set, per generation

Gen	Number				Inbreeding			sEBV ± SD
	Sires	Dams	Families	Progeny	Mean	Min.	Max.	
1	24	33	33	1,731	0	0	0	-0.01 ± 1.11
2	40	57	58	1,717	0	0	0	0.07 ± 0.55
3	52	79	79	2,695	0	0	0	0.39 ± 0.71
4	39	44	50	1,127	0.00319	0	0.06300	0.83 ± 1.01
5	29	42	51	1,455	0.00016	0	0.00800	0.97 ± 1.28

Gen, generation; Max, maximum; Min, minimum; SD, standard deviation; sEBV = EBV/83.036; sEBV, standardized estimated breeding value.

TABLE 2 General information, inbreeding coefficient, standardized economic index and standard deviation for coho salmon data set, per generation

Gen	Number				Inbreeding			sIndex ± SD
	Sires	Dams	Families	Progeny	Mean	Min	Max	
1998	42	81	81	8,619	0.000	0.000	0.000	-0.05 ± 0.23
2000	36	73	73	6,557	0.002	0.000	0.125	0.37 ± 0.30
2002	59	114	114	9,120	0.025	0.000	0.125	0.63 ± 0.38
2004	49	137	137	10,761	0.047	0.000	0.172	1.32 ± 0.31
2006	37	102	102	10,523	0.062	0.016	0.203	1.65 ± 0.37
2008	34	98	98	8,821	0.063	0.027	0.191	1.97 ± 0.33
2010	45	110	110	8,798	0.070	0.035	0.126	2.37 ± 0.28
2012	61	112	112	15,945	0.081	0.056	0.162	2.74 ± 0.36

Gen, generation; Min, minimum; Max, maximum; SD, standard deviation; sIndex, standardized economic index; sIndex = Index/6.79329.

sex:age:year and age at harvest, and tank and weight at the end of test as fixed effects for weight at harvest and resistance to *P. salmonis* respectively (Yáñez et al., 2016). Thus, the two-trait selection index included EBVs for weight at harvest and resistance to *P. salmonis* weighted according to their relative economic values.

2.3 | Objective function

The MS was defined with the optimization of the following objective function (OF):

$$OF = w_1x'EBV + w_2x'Ax + w_3\bar{F}$$

where $x'EBV$ is the expected merit of the future progeny, $x'Ax$ is the weighted mean coancestry of selected parents, \bar{F} is the expected average inbreeding coefficient of the future progeny, w_1 to w_3 are the corresponding weighting factors and x is the vector to be optimized of genetic contributions for each candidate (the symbol ' denotes a transposed vector).

Although not explicitly described in the OF, the mate allocations were determined by \bar{F} , following the problem representation suggested by Gondro and Kinghorn (2008). In this representation, an auxiliary vector is used internally in the mate selection algorithm with the number of elements equal to the number of mates. Each element of the auxiliary vector is a real number used to indirectly determine the mates. These real numbers are ranked,

and the resultant rankings ultimately define the mates to be performed. As an illustrative example, suppose that five mates need to be performed and the optimized x vector determines the contribution of three candidate sires as 3, 0 and 2 respectively. Let us also suppose that the ranking of the five elements of the auxiliary vector are (in order) equal to second, fifth, first, fourth and third, which determines that the first candidate sire would be mated with the second, fifth and first available dams, and the third candidate sire would be mated with the fourth and third available dams. More information about this mate representation and its reasoning can be obtained in Gondro and Kinghorn (2008).

The optimization of the OF assuming w_2 and w_3 equal to zero corresponds to truncation selection. In turn, the optimization of the OF ignoring \bar{F} ($w_3 = 0$) corresponds to the application of OC selection. Using a relatively very low value for w_3 corresponds to the application of OC selection followed by mating minimizing inbreeding as, in this case, the value of OF would primarily be determined by the expected merit and coancestry and secondary by \bar{F} ; that is, \bar{F} would only drive mate and not selection.

2.4 | Mate selection algorithm

A mate selection algorithm based on differential evolution (DE – Storn & Price, 1995) was developed in FORTRAN language, performing

proper changes in a previous developed DE algorithm to perform OC selection (Carvalho et al., 2010), which loosely mimics a biological process evolving towards an optimal solution.

The terms “generation”, “population”, “chromosome”, “loci”, “allele”, “mutation”, “crossover” and “fitness” will be used to help illustrate the MS method, and these should not be confused with similar terms used for the animal breeding application itself.

The DE algorithm applied consisted in randomly generating an initial population of possible solutions, composed by the number of mates for each candidate (the elements of vector x) and by the real numbers of the auxiliary vector, which indirectly defines the mates. Each individual of the population is considered a chromosome with n loci, where n is the number of candidates for selection (order of vector x) plus the number of mates (order of the auxiliary vector). The alleles are either random values ranging from 0 to the maximum allowed number of mates for each candidate (vector x) or real values (auxiliary vector). The fitness of each individual, determined by the value of the OF described above, is calculated according to the value (allele) of each locus of vector x , and to the expected inbreeding of the future progeny resultant from the mates indirectly determined by the auxiliary vector. Once the initial population is established, several generations are simulated. In each generation, a challenger is constructed for each population member. If this challenger has superior fitness, it will replace the population member in the next generation. To build this challenger, three other individuals are chosen at random. We can label these as A, B and C. Each allele is then addressed in turn. With a probability equal to the crossover rate, the allele is simply adopted from the population member that the challenger is challenging. Otherwise, a new allele value is constructed as the value for member A plus the mutation factor times the difference between the values for B and C. Successful challengers replace their respective population members and, together with surviving members, constitute a new generation with higher mean fitness. The process continues over enough generations to achieve convergence close to an optimal solution, with the fittest solution being chosen.

The operational parameters of the DE algorithm to optimize the OF where population size = 2 times the number of candidates; crossover rate = 0.5; mutation factor = 0.2 (or 0.9 every four generations); and maximum number of generations of the evolutionary process (maxgen) = 1,000,000. Convergence was assumed when the range and the mean absolute deviation of the OF, considering all the possible solutions per generation, were lower than 1×10^{-6} . The best solution from the maxgen generation was considered as the optimum solution when the convergence criterion was not attained.

The approach proposed by Lampinen and Zelinka (1999) was adopted to provide integer solutions for the number of mates per candidate. To increase computational efficiency, Colleau (2002) indirect approach was adopted to calculate coancestry, and linked lists (Knuth, 1969) were used for the storage and calculations involving sparse matrices. The mate selection algorithm is freely available for research purpose and can be obtained under request to the corresponding author.

2.5 | Nile tilapia scenarios

For all scenarios in Nile tilapia, it was considered as selection and mate candidates the best four females (179 animals) and six males (281 animals) per family, excluding those with negative EBVs, from the 51 families of the fifth generation. The males were allowed to be mated with a maximum of four females. Each female was mated once; that is, the contribution of each female was not optimized.

A total of 13 OFs were tested for Nile tilapia, under three different scenarios (Table 3). In the first scenario (parametric space explore), the expected merit, coancestry and inbreeding coefficient were optimized independently in OF1 to OF3, respectively, aiming to explore the parametric space of the different components of the OF. The Scenario 2 (OF4 to OF11) corresponded to OC selection. Firstly, different relative weights for w_1 and w_2 were considered and \bar{F} was ignored (OF4–OF10). After defining proper set of values for w_1 and w_2 , a relatively small weight for \bar{F} was used, corresponding to the application of OC selection followed by minimum inbreeding mating (OF11). In Scenario 3 (OF12 and OF13), MS method was applied using the same weights for merit and coancestry as in OF11, but defining the weight for inbreeding in a way that selection and mating were performed simultaneously.

2.6 | Coho salmon scenarios

The real 2012 coho salmon progeny (eighth generation) was formed by the mate between 61 sires and 112 dams from the year 2010 (seventh generation). This was considered as the real scenario (realized mates) and it was compared to the result of the “phantom” progeny (resultant from the mates recommended by the algorithm) from three different scenarios, to check the efficiency of the mate selection algorithm in more than one data set.

A total of 10 OFs were tested for coho salmon, under three different scenarios (Table 4). For the Scenario 1 (maximize and minimize inbreeding), the sires, dams and their intensity of use were kept equal to the real scenario. The selection was intentionally not optimized and just the mates were optimized, aiming to minimize (OF1) or maximize (OF2) inbreeding. In Scenario 2 (sire use optimization – OF3 to OF6), the sires and dams were kept the same as in the real scenario, but the 61 sires were allowed to be mated with zero to four dams. In Scenario 3 (alternative mates – OF7 to OF10), the dams were kept the same as in the real scenario. The candidate sires were the three best males per family from the year 2010, resulting in 330 candidate sires, which presented similar genetic merit as the 61 original candidates. This scenario provides more opportunity for the algorithm to find alternative mates to optimize the objective function, compared with the previous scenarios.

3 | RESULTS

3.1 | Nile tilapia

The results from the optimization of the different objective functions applied in Nile tilapia are presented in Table 3. OF1 represents

TABLE 3 Genetic merit, coancestry and inbreeding results from the optimization of different objective functions in Nile tilapia

OF	w ₁	w ₂	w ₃	nSire	x'EBV	x'Ax	\bar{F}	dif_x'EBV (%)	dif_x'Ax (%)	dif_ \bar{F} (%)
Scenario 1: Parametric space explore										
1 ^a	1	0	0	60	2.49816	0.03323	0.01685	0.00	0.00	0.00
2	0	1	0	60	1.98516	0.05655	0.02292	-20.54	70.18	36.00
3	0	0	1	135	1.84122	0.02286	0.03787	-26.30	-31.08	124.86
Scenario 2: Optimum contribution selection										
4	1	-1	0	60	2.49758	0.03249	0.01421	-0.02	-2.20	-16.00
5	1	-10	0	65	2.47618	0.02812	0.01384	-0.88	-15.36	-17.87
6	1	-20	0	71	2.43772	0.02560	0.01428	-2.42	-22.95	-15.28
7	1	-30	0	73	2.36095	0.02253	0.01516	-5.49	-32.19	-10.08
8	1	-40	0	79	2.29398	0.02067	0.01334	-8.17	-38.07	-20.85
9	1	-50	0	80	2.22541	0.01904	0.01038	-10.92	-42.70	-38.38
10	1	-100	0	91	1.83347	0.01372	0.01370	-26.61	-58.72	-18.73
11 ^b	1	-20	-0.00001	71	2.43772	0.02560	0.01096	-2.42	-22.95	-34.97
Scenario 3: Mate selection										
12	1	-20	-0.01	70	2.43847	0.02564	0.00580	-2.39	-22.84	-65.56
13	1	-20	-1	71	2.43364	0.02542	0.00459	-2.58	-23.51	-72.75

dif_x'EBV, dif_x'Ax and dif_ \bar{F} are compared to the respective results from the OF1.

OF, objective function; w₁, w₂ and w₃, weights for genetic merit, coancestry and inbreeding, respectively; nSire, number of selected sires; x'EBV, genetic merit; x'Ax, coancestry; \bar{F} , inbreeding; dif_x'EBV, difference of expected genetic merit; dif_x'Ax, difference of coancestry; dif_ \bar{F} , difference of inbreeding.

^aCorresponding to a truncation selection.

^bCorresponding to the application of optimum contribution selection followed by minimum inbreeding mating.

TABLE 4 Genetic merit, coancestry and inbreeding results from the optimization of different objective functions in coho salmon

OF	w ₁	w ₂	w ₃	nSire	x'EBV	x'Ax	\bar{F}	dif_Index	dif_x'Ax	dif_ \bar{F}
Real	-	-	-	61	2.74025	0.17639	0.08130	0.0	0.0	0.0
Scenario 1: Maximize and minimize inbreeding ^a										
1	0	0	-1	61	2.74025	0.17639	0.06894	0.0	0.0	-15.25
2	0	0	1	61	2.74025	0.17638	0.12667	0.0	0.0	55.84
Scenario 2: Sire use optimization ^b										
3	1	0	0	28	2.89665	0.17798	0.08378	5.71	0.91	3.08
4	0	-100	-10	47	2.73764	0.17078	0.06811	-0.10	-3.17	-16.24
5	1	-50	-1	42	2.80045	0.17146	0.06674	2.19	-2.78	-17.96
6	1	-20	-1	37	2.85372	0.17335	0.06711	4.14	-1.70	-17.47
Scenario 3: Alternative mates ^c										
7	1	0	0	28	2.94849	0.18981	0.08737	7.60	7.60	7.50
8	0	-100	-10	89	2.62786	0.16045	0.06523	-4.10	-9.07	-19.80
9	1	-50	-1	84	2.70951	0.16098	0.06500	-1.12	-8.73	-20.05
10	1	-20	-1	80	2.78876	0.16360	0.06636	1.77	-7.26	-18.33

OF, objective function; w₁, w₂ and w₃, weights for genetic merit, coancestry and inbreeding, respectively; nSire, number of selected sires; x'EBV, genetic merit; x'Ax, coancestry; \bar{F} , inbreeding; dif_Index, difference of expected genetic merit (index); dif_x'Ax, difference of coancestry; dif_ \bar{F} , difference of inbreeding.

dif_Index, dif_x'Ax and dif_ \bar{F} are compared to the respective results from the real scenario.

^aThe sires, dams and their intensity of use were kept equal to the real scenario

^bSires and dams were kept the same as in the real scenario, but the sires were allowed to be mated with zero to four dams.

^cThe dams were the same as in the real scenario, but candidate sires were the three best males per family from the year 2010 (total of 330 candidate sires).

truncation selection and provided the highest expected genetic merit of the future progeny (2.4982). Coancestry and inbreeding were not optimized in OF1, and its results were used as a reference for

comparison with those from the other objective functions. In OF2 and OF3, coancestry and inbreeding were maximized, respectively, to explore their parametric space. Compared with OF1, coancestry

was increased by 70.18% in OF2, and inbreeding was increased by 124.86% in OF3 (Table 3).

The results from Scenario 2 (Table 3) correspond to OC selection. In the first seven objective functions from this scenario (OF4 to OF10), inbreeding was ignored and the weight for coancestry varied from -1 to -100. As more importance was given to coancestry (from OF4 to OF10), the number of selected sires increased and the expected genetic merit and coancestry decreased. For instance, compared with OF1, OF10 resulted in 51.67 more sires being selected (60 vs 91 sires) and a reduction of 26.61 and 58.72% in genetic merit and coancestry respectively. The different OF differed not only on the number of selected sires but also on the contributions of the common selected sires. As an example, from the 60 sires selected by OF1, only 30 were in common with OF10. These common sires contributed for 89 mates in OF1 and 61 mates in OF10.

Compared with OF1, results from OF6 revealed that the weights used to optimize this objective function allowed a substantial reduction in coancestry (22.95%) with a small reduction in the genetic merit (2.42%). OF11 corresponded to the application of OC selection followed by minimum inbreeding mating. It provided the same result as OF6 for genetic merit and coancestry, but allowed a greater reduction in inbreeding (34.97% vs 15.28%).

In Scenario 3 (OF12 and OF13), selection and mating decisions were performed simultaneously (MS). The results from OF12 and OF13 (Table 3) showed that MS outperformed OC selection followed by minimum inbreeding mating (OF11) in controlling inbreeding, under similar levels of genetic merit and coancestry. Compared with OF1, the reduction in the expected inbreeding of the future progeny was equal to 65.56% for OF12 and 72.75% for OF13, whereas it was equal to 34.97% for OF11.

3.2 | Coho salmon

The results of the different objective functions applied to coho salmon using the same sires and dams, and their intensity of use, as in the real scenario, showed the possibility to reduce inbreeding by 15.25% (OF1) or to increase inbreeding by 55.84% (OF2), depending on the allocation of the mates, determined by minimizing or maximizing inbreeding of the future progeny respectively (Table 4).

When the 61 sires from the real scenario were allowed to be mated with zero to four dams (Scenario 2; Table 4), the maximum expected index of the future progeny was obtained with OF3. In this case, the top 28 sires would be mated with four dams each, and the remaining 33 sires would not be mated. It is important to emphasize that this strategy would not be recommended in practice otherwise inbreeding rate would rapidly increase. Results from Scenario 2 (Table 4) also showed that, giving the same candidates as in the real scenario, it would be possible to reduce coancestry and inbreeding by 3.17% and 16.24%, respectively, without compromising the expected merit of the future progeny (OF4) or, alternatively, it would be possible to increase the expected merit of the future progeny and, concomitantly, reduce coancestry and inbreeding (OF5 and OF6).

In Scenario 3 (Table 4), the dams were kept the same as in the real scenario and 330 candidate sires were considered. As for OF3, the genetic merit was maximized when the top 28 sires were mated with four dams each, and the remaining (302) sires were not used (OF7). The expected merit of the future progeny was higher for OF7 than for OF3, because more candidate sires were available, providing more opportunity for the algorithm to find alternative mates to optimize the objective function. As for OF5 and OF6, results from OF10 showed that, compared with the real scenario, the algorithm for MS allowed to reduce coancestry and inbreeding with a concomitant increase in the expected merit of the future progeny.

4 | DISCUSSION

To perform MS in aquaculture breeding, a DE algorithm (Storn & Price, 1995) was applied under different scenarios, in two real data sets from Nile tilapia and coho salmon breeding populations. The algorithm showed to be computationally efficient in terms of processing speed and memory requirement (data not shown), making it suitable for practical applications. The computational efficiency came mainly from the use of Colleau (2002) indirect approach to compute coancestry and linked lists (Knuth, 1969) to store and operate the sparse matrices. For instance, the analyses OF4 to OF6 of the coho salmon data set, which had 573 animals (candidates and their ancestors) to compute coancestry, run in approximately 10 minutes in a PC with an Intel® Core™ i7 2.2 GHz processor and 8GB RAM. Simpler OF resulted in faster analyses, especially if inbreeding of the future progeny was not accounted for ($w_3 = 0$). For example, OF3 of the coho salmon data set, which accounted just for the genetic merit, was optimized in <2 seconds.

Besides being computationally efficient, the MS algorithm showed to be flexible regarding the components to be considered in the OF to be optimized. The definition of the proper weight for each component depends on several factors, including number of families (effective population size); current level of inbreeding and, more importantly, current level of coancestry among candidates; the allowed reduction in the genetic progress in the short term; the maximum desired rate of inbreeding; and the time horizon being considered. In the applied MS algorithm, the weights for the different components of the OF need to be determined empirically. In contrast, the analytical derivation of OC selection using LaGrangian multipliers (Meuwissen, 1997) automatically provides the optimum genetic contribution, under constrained rate of inbreeding, without the necessity to define the weights for genetic merit and coancestry *a priori*. This disadvantage of the applied MS algorithm can be compensated by running several analyses varying the weights for the different components of the OF what, in turn, allows to explore the potential outcomes, enabling to better balance the different components and to make the selection and mating decisions in a more dynamic and tactical way (Kinghorn & Shepherd, 1999). In Scenario 2 of Nile tilapia, for example, different weights for coancestry were tested allowing assessing which set of weights resulted in a

significant reduction in coancestry without expressively reducing the genetic merit.

Another aspect of the MS algorithm is that, for being based on an evolutionary algorithm, it is not guaranteed to provide the optimum solution. Because of the stochastic nature of evolutionary algorithms and depending on the complexity of the function to be optimized, a local optimum can be provided as the solution of the optimization process instead of the optimum solution. However, the DE algorithm (Storn & Price, 1995) used in the present study is advocated to be very powerful to optimize diverse objective functions studied in the literature (Carvalho et al., 2010; Kinghorn, 2011; Price, Storn & Lampinen, 2006). OC selection can also be applied using semidefinite programming (Hely, Amer, Walker & Symonds, 2012; Pong-Wong & Woolliams, 2007) that, according to Pong-Wong and Woolliams (2007), guarantees the finding of the optimum solution. The comparison between different optimization methods was outside the scope of the present study.

Although our results focused just on the expected genetic merit and inbreeding of the next generation, they are in accordance with the results from the literature showing the superiority of OC selection over selection based exclusively on genetic merit, in terms of maximizing the genetic gain under controlled rate of inbreeding (Meuwissen, 1997; Meuwissen & Sonesson, 1998). Also, compared with the selection based exclusively on genetic merit, MS results showed the possibility to reduce coancestry and inbreeding without compromising the expected merit of the future progeny or, alternatively, to increase the expected merit of the future progeny and, concomitantly, reduce coancestry and inbreeding. This result highlights the great flexibility of the MS algorithm. Evidence was found that MS outperformed OC selection followed by minimum inbreeding mating in controlling inbreeding, under similar levels of genetic merit and coancestry. A simulation study applying both strategies over subsequent generations is recommended for a better comparison.

Compared with truncation selection or to the real mates, MS allowed a greater reduction in coancestry and inbreeding for the Nile tilapia (Table 3: OF12 and OF13) than for the coho salmon population (Table 4: OF9 and OF10). This result can be associated with the smaller number of controlled generations and lower level of coancestry and inbreeding of the Nile tilapia population, providing more opportunity for the algorithm to find alternative candidates and mates to optimize the OF and to attain better outcomes for the components of the OF. This highlights the importance of adopting strategies to control inbreeding since the establishment of the breeding programmes. It is important to emphasize that the empirical commonly used strategies for controlling inbreeding in aquaculture, as restricting the number of selected animals per family and not allowing full and half-sib mates, can be useful in the beginning of the breeding programme but are not effective in the long term (Skaarud et al., 2011; Yáñez et al., 2014a). In the case where some of the broodstock are used for production or multiplication, coancestry and inbreeding could be just ignored, e.g. OF1 (Table 3), thus maximizing the genetic gain. The algorithm also allow to set specific

values for coancestry (w_2) and inbreeding (w_3), in the case the broodstock manager needs to account for these two parameters to generate fish for production. Besides advocating the use of OC selection and planned mates to better control the rate of inbreeding and the effective population size, Yáñez et al. (2014a) also recommended the use of genetic material to connect different genetically improved populations using, for example, cryopreserved sperm from selected males. Unfortunately, the use of external genetic material is not always possible in aquaculture breeding programmes due to commercial and sanitary barriers.

The mate selection algorithm was computationally efficient and flexible for practical applications in aquaculture breeding. The expected consequence of using the algorithm, in contrast with empirical procedures for controlling inbreeding, is to control inbreeding more effectively and promote higher genetic progress in the long term. Evidence was found that inbreeding can be better controlled by mate selection than by optimum contribution selection followed by minimum inbreeding mating.

ACKNOWLEDGMENTS

The authors are grateful to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process number 2014/20626-4 and 2015/25232-7) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support, and Peixe-Gen research group in Brazil and Aquainnovo in Chile for providing the data sets.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- Carvalho, R., Kinghorn, B. P., & Queiroz, S. A. (2009). Seleção de Acasalamentos Considerando Diferentes Funções Objetivo. In 46^a Reunião Anual da Sociedade Brasileira de Zootecnia: 14–17 junho 2009, Maringá-PR.
- Carvalho, R., Kinghorn, B. P., & Queiroz, S. A. (2010). Mate selection accounting for connectedness. In: Proceedings of the 9th World Congress on Genetics Applied to Livestock Production: 1–6 August 2010, Leipzig.
- Colleau, J. J. (2002). An indirect approach to the extensive calculation of relationship coefficients. *Genetic Selection Evolution*, 34, 409–421.
- D'agaro, E., Woolliams, J. A., Haley, C. S., & Lanari, D. (2010). Optimizing mating schemes in fish breeding. *Italian Journal of Animal Science*, 6, 795–796.
- Dufflocq, P., Lhorente, J. P., Banger, R., Neira, R., Newman, S., & Yáñez, J. M. (2016). Correlated response of flesh color to selection for harvest weight in coho salmon (*Oncorhynchus kisutch*). *Aquaculture*, doi:10.1016/j.aquaculture.2016.08.037.
- Fernández, J., Toro, M. A., & Caballero, A. (2001). Practical implementation of optimal management strategies in conservation programmes: A mate selection method. *Animal Biodiversity and Conservation*, 24, 17–24.
- Gondro, C., & Kinghorn, B. (2008). *Application of evolutionary algorithms to solve complex problems in quantitative genetics and bioinformatics* (pp. 96). Guelph: University of Guelph.

- Hely, F. S., Amer, P. R., Walker, S. P., & Symonds, J. E. (2012). Optimised parent selection and minimum inbreeding mating in small aquaculture breeding schemes: A simulation study. *Animal*, 7, 1–10.
- Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. Guelph, Canada: Guelph Univ. Press.
- Hinrichs, D., & Wetten, M. (2006). An algorithm to compute optimal genetic contributions in selection programs with large numbers of candidates. *Journal of Animal Science*, 84, 3212–3218.
- Holtsmark, M., Klemetsdal, G., Sonesson, A. K., & 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, 232–240.
- 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.
- Kinghorn, B. P. (2011). An algorithm for efficient constrained mate selection. *Genetic Selection Evolution*, 43, 1.
- Kinghorn, B. P., & Shepherd, R. K. (1999). Mate selection for the tactical implementation of breeding programs. *Association For The Advancement Of Animal Breeding And Genetics*, 13, 130–133. 1., 1999, Mandurah. Proceedings Mandurah: 1999.
- Kinghorn, B. P., Shepherd, R. K., & Woolliams, J. A. (1999). An index of estimated breeding value, parental coancestry and progeny inbreeding to help maximize genetic gains. *Association For The Advancement Of Animal Breeding And Genetics*, 13, 412–415. 1., 1999, Mandurah. Proceedings. Mandurah: 1999.
- Knuth, D. E. (1969). *The art of computer programming* (2nd ed., Vol 2: Seminumerical algorithms, pp. 624). Reading, Massachusetts: Addison Wesley.
- Kremer, V. D., Newman, S., Wilson, E. R., & Kinghorn, B. (2010). Mate Selection for sustained genetic improvement in small populations. In Proceedings of the 9th World Congress on Genetics Applied to Livestock Production. Leipzig, Germany.
- Lampinen, J., & Zelinka, I. (1999). Mixed variable non-linear optimization by differential evolution. *Proceedings of Nostradamus*, 99(2), 7–8.
- Liu, H., Meuwissen, T. H., Sørensen, A. C., & Berg, P. (2015). Upweighting rare favourable alleles increases long-term genetic gain in genomic selection programs. *Genetic Selection Evolution*, 47, 19.
- Meuwissen, T. H. E. (1997). Maximizing the response of selection with a predefined rate of inbreeding. *Journal of Animal Science*, 75, 934–940.
- Meuwissen, T. H. E., & Sonesson, A. K. (1998). Maximizing the response of selection with a predefined rate of inbreeding: Overlapping generations. *Journal of Animal Science*, 76, 2575–2583.
- 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. *Journal of Animal Science*, 89, 630–638.
- Pong-Wong, R., & Woolliams, J. A. (2007). Optimisation of contribution of candidate parents to maximise genetic gain and restricting inbreeding using semidefinite programming. *Genetic Selection Evolution*, 39, 1–23.
- Price, K., Storn, R. M., & Lampinen, J. A. (2006). *Differential evolution: A practical approach to global optimization*. Berlin, Germany: Springer Science and Business Media.
- Shepherd, R. K., & Kinghorn, B. P. (1999). A tactical approach to the design of crossbreeding programs. *Proceedings of the Sixth World Congress on Genetics Applied to Livestock Production: 11-16 January, Armidale*, 25, 431–438.
- Skaarud, A., Woolliams, J. A., & Gjøl, H. M. (2011). Strategies for controlling inbreeding in fish breeding programs, an applied approach using optimum contribution (OC) procedures. *Aquaculture*, 311, 110–114.
- Skaarud, A., Woolliams, J. A., & Gjøl, H. M. (2014). Optimising resources and management of genetic variation in fish-breeding schemes with multiple traits. *Aquaculture*, 420, 133–138.
- Sonesson, A. K. (2005). A combination of walk-back and optimum contribution selection in fish: A simulation study. *Genetic Selection Evolution*, 37, 587–600.
- Storn, R., & Price, K. (1995). *Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces*, Vol. 3. Berkeley: ICSI.
- Weigel, K. A., & Lin, S. W. (2000). Use of computerized mate selection programs to control inbreeding of Holstein and Jersey cattle in the next generation. *Journal of Dairy Science*, 83, 822–828.
- Woolliams, J. A., & Thompson, R. (1994). A theory of genetic contributions. In Proceedings of the 5th World Congress on Genetics Applied to Livestock Production (Vol. 19, pp. 127–134).
- Yáñez, J. M., Bangera, R., Lhorente, J. P., Barría, A., Oyarzún, M., Neira, R., & Newman, S. (2016). Negative genetic correlation between resistance against *Piscirickettsia salmonis* and harvest weight in coho salmon (*Oncorhynchus kisutch*). *Aquaculture*, 459, 8–13.
- Yáñez, J. M., Bangera, R., Lhorente, J. P., Oyarzún, M., & Neira, R. (2013). Quantitative genetic variation of resistance against *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*). *Aquaculture*, 414, 155–159.
- Yáñez, J. M., Bassini, L. N., Filp, M., Lhorente, J. P., Ponzoni, R. W., & Neira, R. (2014a). Inbreeding and effective population size in a coho salmon (*Oncorhynchus kisutch*) breeding nucleus in Chile. *Aquaculture*, 420, S15–S19.
- Yáñez, J. M., Lhorente, J. P., Bassini, L. N., Oyarzún, M., Neira, R., & Newman, S. (2014b). Genetic co-variation between resistance against both *Caligus rogercresseyi* and *Piscirickettsia salmonis*, and body weight in Atlantic salmon (*Salmo salar*). *Aquaculture*, 433, 295–298.

How to cite this article: Yoshida GM, Yáñez JM, de Oliveira CAL, et al. Mate selection in aquaculture breeding using differential evolution algorithm. *Aquac Res*. 2017;00: 1–8. <https://doi.org/10.1111/are.13365>