

Importance de l'aquaculture et de la sélection génomique chez les salmonidés dans le monde

1. Importance de l'aquaculture et de la trutticulture dans le monde

La production, issue de la pêche et de l'aquaculture, est en constante augmentation (Figure 1), passant d'environ 20 millions de tonnes en 1950 à 178 millions de tonnes en 2018 (FAO 2020). Depuis les années 1990, c'est avant tout l'aquaculture qui permet cette augmentation. Elle représentait 15% des ces produits dans les années 1990 et 46% en 2018 (soit 73 millions de tonnes) (FAO 2020). La FAO estime que la production aquatique devrait atteindre 204 millions de tonnes en 2030, dont 66% provenant de l'aquaculture. L'augmentation de l'aquaculture d'ici à 2030 concernera tous les continents, même si elle sera plus importante en Afrique (48%), Océanie (41%), Asie (32%), Amérique du Sud (30%) ou du Nord (27%) qu'en Europe (18%) (FAO 2020).

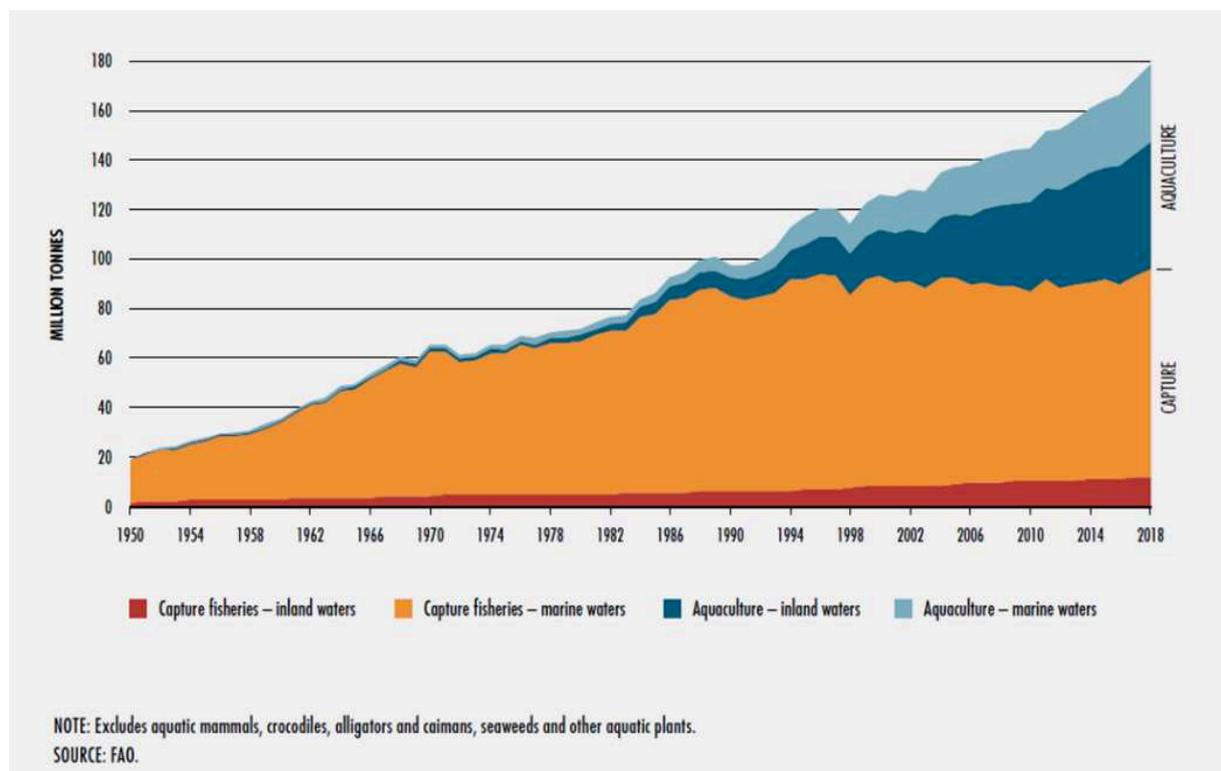


Figure 1 : Évolution de la production de produits aquatiques (pêche et aquaculture) de 1950 à 2018 dans le monde (FAO 2020).

L'aquaculture mondiale est dominée par l'Asie avec 89% de la production (58% par la Chine), suivie par l'Amérique du Sud avec 3,8% (1,5% par le Chili) et l'Europe avec 3,7% (1,6% par la Norvège).

Parmi les cinq espèces de poissons les plus élevées dans le monde, nous retrouvons 4 espèces de carpes représentant 32,8% de la production mondiale (10,5% carpe herbivore, *Ctenopharyngodon idellus* ; 8,8% carpe argentée, *Hypophthalmichthys molitrix* ; 7,7% carpe commune, *Cyprinus carpio* ; 5,8% carpe à grosse tête, *Hypophthalmichthys nobilis*). Le tilapia du Nil (*Oreochromis niloticus*) est la 3^{ème} espèce la plus élevée dans le monde (8,3% de la production totale). Les principales espèces de salmonidés élevées sont le saumon Atlantique (*Salmo salar*) et la truite arc-en-ciel (*Oncorhynchus mykiss*) respectivement en 9^{ème} position et en 15^{ème} position dans le classement des espèces produites et représentant à elles deux 6,1% de la production mondiale respectivement (FAO 2020). Ainsi, 2 436 millions de tonnes de saumon Atlantique et 848 000 tonnes de truite arc-en-ciel ont été produites dans le monde en 2018.

Dans l'Union Européenne, les productions aquacoles principales sont en volume le saumon Atlantique (25%), puis la truite arc-en-ciel (14%), suivies par l'huître, le bar européen, la dorade et les moules soit environ 10% de la production aquacole Européenne représentait par chacune de ces espèces (Figure 2). En Europe, l'essentiel de la production de salmonidés est réalisée en Norvège avec 1,282 million de tonnes de saumon Atlantique (plus de la moitié de la production mondiale) et 68 000 tonnes de truite arc-en-ciel élevées en mer. La France est le 3^{ème} pays en Europe en production de truite arc-en-ciel avec 33 119 de tonnes en 2018 dernière la Russie dont la production s'élève à 35 204 de tonnes en 2018 (AGRESTE 2020 et base de données FAO).

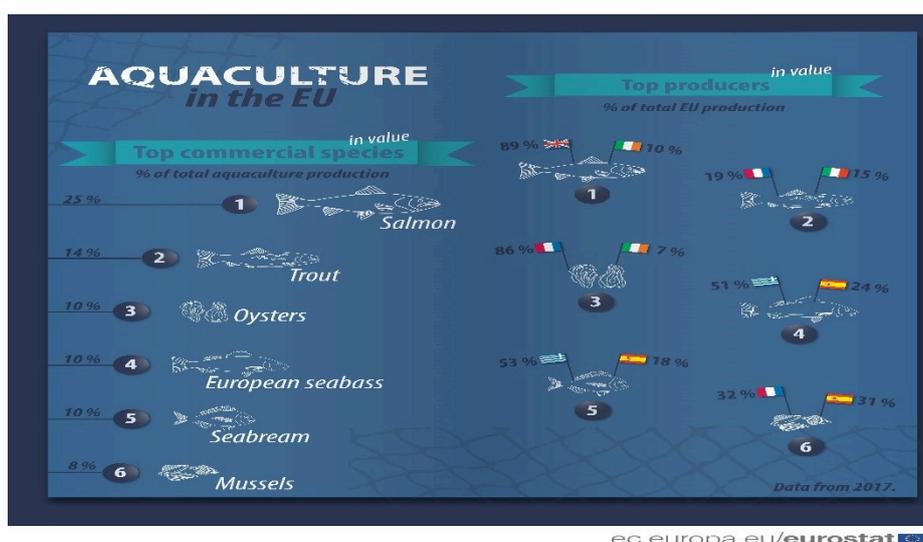


Figure 2 : principales productions aquacoles dans l'Union européenne (2017).

1.1. Quelques chiffres sur la filière française de salmonidés

La filière française aquacole est composée de 3 006 entreprises pour une production de 188 484 tonnes et un chiffre d'affaires annuel de 715 622 k€ (AGRESTE 2020). Les coquillages et crustacés représentent 77% de la production et 70% du chiffre d'affaires. Les poissons représentent 23% de la production et 30% du chiffre d'affaires. La principale espèce de poisson élevée en France est la truite arc-en-ciel.

La production de truite arc-en-ciel représente 80% de la production piscicole en France et 95% de celle des salmonidés (incluant aussi truite Fario, saumon de fontaine, ombre et omble chevalier) en 2018 (AGRESTE 2020). La production en France de truite arc-en-ciel se fait pour sa quasi-totalité en eau douce. L'élevage en mer est très minoritaire (125 tonnes par an) (AGRESTE 2020).

Outre la production de poissons, celle d'œufs de salmonidés pour la consommation est estimée à 111 tonnes par an, soit près du triple de la production française de caviar d'esturgeon (40 tonnes par an). La production d'œufs de truite représente 2% du chiffre d'affaires de la filière française de salmonidés (AGRESTE 2020).

Sur les 456 entreprises salmonicoles recensées pour l'année 2007, 25 entreprises étaient des écloséries, 126 faisaient uniquement du grossissement et 305 conduisaient les deux activités à la fois. En 2008, les 10 entreprises les plus importantes réalisaient 42% du chiffre d'affaires (AGRESTE 2011).

1.2. La France, un pays leader de la sélection de la truite arc-en-ciel

Les premiers essais de sélection expérimentale en aquaculture sont rapportés dès les années 1920 et les premiers programmes de sélection commerciale commencent dans les années 70 (e.g., Haffray 2019). Chavanne et al. (2016) dénombrent 10 programmes de sélection chez la truite arc-en-ciel et sept chez le saumon Atlantique en Europe.

La France a des entreprises leaders au niveau mondial en génétique de la truite arc-en-ciel avec une production totale, incluant les filiales implantées à l'étranger, avoisinant les 850 millions d'œufs embryonnés (communication personnelle SYSAAF). Les quatre entreprises françaises de sélection trutticole sont :

Chapitre 1: Importance de l'aquaculture et de la trutticulture dans le monde

- Les Sources de l'Avance, filiale du Groupe Aqualande (<https://www.groupeaqualande.com>) créée en 1981, qui opère dans 41 piscicultures avec une production de l'ordre 15 000 tonnes de truites par an en France et en Espagne. Le groupe possède une centre de sélection et quatre écloseries de multiplication avec 650 millions d'œufs embryonnés commercialisés chaque année dans le monde à partir des sites de multiplication localisés en France et en Espagne.
- Milin Nevez, filiale de Bretagne Truite (<http://www.bretagne-truite.fr/>) commence son activité en 1996. En 2016, l'entreprise compte 60 sites d'élevage et une production de 7 000 tonnes de salmonidés. Elle vend des œufs embryonnés aux piscicultures adhérentes, mais aussi à des clients extérieurs en France et à l'international. Le programme de sélection a commencé en 2000.
- Viviers de Sarrance (<http://www.oeufsdetruite.fr/>) a été créée en 1975 et a mis en place un programme de sélection en 1995. Aujourd'hui, l'entreprise vend plus de 50 millions d'œufs embryonnés partout dans le monde.
- Charles Murgat (<https://www.charlesmurgat.com>) est la plus ancienne pisciculture familiale de France créée en 1898. Le programme de sélection chez la truite arc-en-ciel a commencé en 1994. Ils produisent quatre espèces de salmonidés (truite fario, omble chevalier, saumon de fontaine et truite arc-en-ciel).

Voici quelques exemples d'entreprises internationales avec des programmes de sélection bien avancés :

- Troutlodge (<https://www.troutlodge.com/en/>) est une entreprise des USA spécialisée en truite arc-en-ciel (rachetée par Hendrix Genetics en 2016 après le rachat de l'entreprise de sélection de saumon Atlantique Landcatch (GB) avec une production d'environ 500 millions d'œufs embryonnés de truite commercialisés dans plus de 70 pays. Dès 2015, Troutlodge a mis en place la sélection génomique pour améliorer la résistance à la flavobactérie, puis, pour la survie, la croissance, la qualité de la chair et l'efficacité alimentaire.
- AquaGen est une entreprise norvégienne spécialisée en sélection du saumon Atlantique et de la truite arc-en-ciel (<https://aquagen.no/en/>), filiale du leader mondial de la sélection de la poule pondeuse Wesjohann Group (Allemagne) aussi propriétaire des entreprises de sélection de tilapia aux Philippines (Genomar) et au Brésil (Aqualabel et AquaAmerica / AquaPorto). Son programme de sélection de la truite arc-en-ciel,

dérivant du programme conduit par Akvaforsk en 1972 a commencé en sélection familiale sur la croissance en mer et concerne maintenant une dizaine de caractères. Depuis 2015, ils ont mis en place une sélection assistée par marqueurs pour des QTLs de la résistance à la nécrose pancréatique infectieuse (IPN) (1 QTL) et/ou à la flavobactériose (2 QTLs). Depuis 2016, AquaGen a mis en place la sélection génomique pour l'une de ses lignées de saumon Atlantique et les caractères de croissance, résistance au pou de mer et résistance à une maladie des branchies (AGD : amoebic gill disease). Ils proposent des œufs embryonnés « estampillés » pour des aptitudes différentes à partir des QTLs identifiés (IPN, flavobactériose, pou de mer, maladie pancréatique du saumon causé par un alphavirus, syndrome de myopathie cardiaque, maturation sexuelle chez le saumon mâle, couleur du filet).

- Benchmark Genetics (<https://salmobreed.no/en/>) est un fond d'investissement anglais impliqué en aquaculture notamment en ayant racheté l'entreprise INVE spécialisée dans la production des aliments en éclosion mais aussi différentes entreprises de sélection aquacoles du saumon Atlantique Stofnfiskdur (Islande), SalmoBreed (Norvège), de tilapia (Honduras), de crevette blanche (Colombie) ou entrée récemment au capital du groupe grec Nireus pour la sélection du bar et de la daurade. Ils produisent environ 550 millions d'œufs embryonnés de saumon Atlantique par an. Pour le saumon, la sélection familiale a commencé en 1991 et la sélection génomique à partir de 2017. Trois lignées de saumon sont sélectionnées sur QTL pour IPN et en sélection génomique pour la résistance au pou de mer. Deux lignées sont sélectionnées pour la croissance en mer, l'une étant également sélectionnée pour la résistance à la maladie du pancréas du saumon et l'autre pour la résistance à l'anémie infectieuse du saumon (ISA). Cette entreprise propose aussi des œufs « estampillés » pour certaines aptitudes spécifiques.

On peut ainsi voir que les principales entreprises concurrentes ont déjà mis en œuvre la sélection génomique et/ou une sélection assistée par marqueurs pour des QTLs à effets forts entre 2015 et 2017. Elles proposent aux multiplicateurs et grossisseurs des œufs « estampillés » pour certains QTLs identifiés à une ou plusieurs résistances aux maladies.

2. Les principes de la sélection génomique

Le génotypage par puce à ADN contenant plusieurs milliers à centaines de milliers de marqueurs SNP (Single Nucleotide Polymorphism) permet, depuis la fin des années 2000, un accès et une lecture facile du patrimoine génétique des animaux, modifiant de ce fait considérablement les schémas de sélection dans toutes les espèces d'élevage (Le Roy et al. 2019). Les puces à ADN permettent la mise en œuvre de la SG qui consiste à prédire la valeur génétique d'un animal à partir de la lecture de son génotype. En effet, Meuwissen et al. (2001) démontrent la possibilité d'utiliser des milliers de marqueurs (SNPs) couvrant l'ensemble du génome pour estimer la valeur génétique des individus au travers des effets estimés pour chaque marqueur. Pour cela, les équations de prédiction des valeurs génétiques (EBV : estimate breeding value) sont établies à l'aide d'une population de référence constituée d'individus génotypés et phénotypés qui permettent d'estimer les effets des marqueurs. Dans un second temps, les EBV des candidats à la sélection génotypés mais sans performance connue, sont ensuite prédites grâce à la somme des effets estimés des marqueurs.

2.1. Les modèles de prédiction génomique

Les premières approches par régression ou stratégies bayésiennes de sélection de variables pour estimer les effets de SNP à des fins de prédiction génomique furent formalisées par Meuwissen et al. (2001). Une large variété de méthodes statistiques a depuis été proposée (Habier et al. 2011). Parmi celles-ci, j'ai utilisé deux modélisations différentes au cours de la thèse que je vous présente.

2.1.1. Modèle animal et évaluation BLUP sur pedigree et/ou génomique

Depuis les années 1990-2000, les valeurs génétiques des animaux d'élevage sont estimées à partir d'un modèle de référence, appelé modèle animal qui est un modèle linéaire mixte décrivant les performances des individus par l'équation suivante :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad (1)$$

Où \mathbf{y} est le vecteur de performances, \mathbf{X} la matrice d'incidence pour le vecteur des effets fixes $\boldsymbol{\beta}$, \mathbf{Z} la matrice d'incidence pour le vecteur \mathbf{u} des valeurs génétiques additives des animaux et \mathbf{e} le vecteur des résidus du modèle.

Les effets environnementaux identifiés (par exemple, les effets de la cohorte de naissance) jouant sur la performance sont inclus dans le modèle en effets fixes afin de discriminer au mieux les individus sur leurs valeurs génétiques. Les effets de milieu non identifiés sont englobés dans les résidus du modèle. Les EBV sont prédites par la méthode Best Linear Unbiased Prediction (BLUP) (Henderson 1975), permettant l'estimation conjointe des effets fixes et des effets aléatoires par la résolution du système d'équations suivant :

$$\begin{bmatrix} \mathbf{X}\mathbf{X}' & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\sigma_e^2}{\sigma_u^2} \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

\mathbf{A}^{-1} : l'inverse de la matrice de parenté des individus basé sur le pedigree (\mathbf{A}), σ_e^2 la variance résiduelle, σ_u^2 la variance génétique.

Le GBLUP (ou BLUP génomique) proposé par Habier et al. (2007) est très similaire à un BLUP classique, mais la matrice de parenté sur pedigree (\mathbf{A}) est remplacée par une matrice de parenté génomique (\mathbf{G}) calculée à l'aide des marqueurs selon la formule suivante (VanRaden 2008) :

$$\mathbf{G} = \frac{\mathbf{z}\mathbf{z}'}{2 * \sum_{i=1}^m (p_i * (1 - p_i))}$$

où m est le nombre de SNP et p_i la fréquence de l'allèle majeur du SNP i .

Le GBLUP permet de ne prédire des EBVs que pour les individus génotypés. Cela a ensuite été étendu afin d'intégrer l'information des animaux génotypés et non-génotypés dans un modèle dit « Single step GBLUP » (Misztal et al. 2009) afin de prendre en compte en une seule étape l'ensemble des informations de phénotypes et génotypes disponibles, et de fournir simultanément une valeur génétique à tous les animaux qu'ils soient génotypés ou non.

Legarra et al. (2009) proposent pour cela de remplacer la matrice \mathbf{G} par une matrice \mathbf{H} construite ainsi :

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_{11} + \mathbf{G}_{12} + \mathbf{A}_{22}^{-1}(\mathbf{G} - \mathbf{A}_{22})\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{G} \\ \mathbf{G}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{G} \end{bmatrix}$$

\mathbf{A} correspondant à la matrice de parenté basée sur le pedigree, \mathbf{G} la matrice de parenté basée sur la génomique, les indices 1 et 2 représentent les individus non génotypés et génotypés, respectivement. L'inverse de la matrice \mathbf{H} pour les équations du modèle mixte est :

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

avec respectivement \mathbf{A}^{-1} et \mathbf{G}^{-1} l'inverse des matrices de parenté pedigree (\mathbf{A}) et de parenté génomique (\mathbf{G}), et \mathbf{A}_{22}^{-1} l'inverse de la matrice de parenté \mathbf{A} uniquement pour les individus génotypés.

Les effets des marqueurs peuvent ensuite être prédits par la résolution de l'équation suivante (Wang et al. 2012):

$$\hat{\mathbf{a}} = \mathbf{DZ}'[\mathbf{ZDZ}']^{-1}\hat{\mathbf{u}}_g \quad (2)$$

où \mathbf{D} est la matrice diagonale des poids associés aux effets de SNPs ($\mathbf{D}=\mathbf{I}$ en GBLUP), \mathbf{Z} la matrice d'incidence qui relie les génotypes aux SNPs des individus génotypés au vecteur $\hat{\mathbf{a}}$ des effets estimés des SNPs et $\hat{\mathbf{u}}_g$ le vecteur des GEBV des individus génotypés ($\hat{\mathbf{u}}_g = \mathbf{Z}\hat{\mathbf{a}}$).

2.1.2. Modèle BayesC π et sélection de variables

Le GBLUP fait l'hypothèse que tous les marqueurs ont un effet sur le caractère, les effets étant distribués selon une loi normale de très faible variance (variance génétique divisée par le nombre de marqueurs). Ce modèle est cohérent avec l'hypothèse du modèle génétique infinitésimal de Fisher (1918) supposant qu'un caractère quantitatif est expliqué par un très grand nombre de gènes ayant tous de très faibles effets additifs. Toutefois, nous ne connaissons pas *a priori* le degré de polygénie des caractères. Certains caractères vont avoir une structure oligogénique avec un ou quelques QTLs majeurs. Par exemple, pour la résistance à l'IPN dans une lignée japonaise de truite arc-en-ciel quelques QTL expliquent 27 à 30% de la variance phénotypique (Ozaki et al. 2001), tandis que d'autres caractères se révèlent très polygéniques comme les caractères de production chez la truite arc-en-ciel (Gonzalez-Pena et al. 2016; Neto et al. 2019; Ali et al. 2020). Le GBLUP va sous-estimer les effets des marqueurs réellement importants et sur-estimer les effets des marqueurs qui ne jouent pas sur l'expression des caractères. Des méthodes de sélection de variables, comme des méthodes bayésiennes, peuvent résoudre cela. De nombreuses méthodes bayésiennes de sélection de variables permettent de s'approcher plus de la réalité biologique et d'un nombre fini de gènes jouant sur les caractères (Meuwissen et al. 2001; Habier et al. 2011).

Pour mieux estimer les effets des SNPs ayant un réel effet sur le caractère, la méthode BayesC π (Habier et al. 2011) a été considérée dans la thèse. Le modèle considéré est un modèle de

mélange avec une proportion π de SNPs ayant des effets non nuls (distribués selon une loi normale de moyenne nulle et variance suivant une loi de Chi2-inverse) et une proportion $(1-\pi)$ de SNPs sans effet par méthode de Monte-Carlo par chaîne de Markov (avec algorithme de Metropolis-Hastings). Celui-ci permet de partir sans *a priori* sur les effets des SNPs. À chaque cycle de tirage, une proportion π de SNPs auront un effet et une proportion $1-\pi$ n'aura aucun effet. Un grand nombre d'itérations (plusieurs centaines de milliers) est nécessaire et les cycles retenus pour les estimations doivent être assez indépendants les uns des autres, ce qui demande un temps de calcul très important (plusieurs heures à quelques jours de CPU pour un caractère et 1300 individus). Le modèle d'analyse des performances est le suivant :

$$y_i = \mu + X\beta_i + \sum_{j=1}^n \delta_j z_{ij} a_j + e_i \quad (4),$$

avec y_i la performance de l'animal i , μ la moyenne du caractère, X la matrice d'incidence pour les effets fixes β d'une performance, n le nombre de SNP, z_{ij} le génotype (codé 0 et 2 pour les homozygotes et 1 pour l'hétérozygote) au locus j pour animal i , a_j l'effet du SNP j , e_i le résidu, δ_j est une indicatrice : à l'itération k , $\delta_j=1$ si le SNP j est inclus dans le modèle (probabilité π) et $\delta_j=0$ s'il n'est pas inclus (probabilité $1-\pi$).

Cette approche a été utilisée dans la thèse pour les études d'associations pangénomiques (GWAS) car elle permet non seulement de mieux estimer les effets des principaux QTLs, mais aussi de préciser leur localisation le long du génome par rapport à une analyse en GBLUP.

La SG a commencé par révolutionner les schémas de sélection des races bovines laitières dès 2008 grâce à l'élaboration de la première puce de génotypage 54K utilisée pour les espèces d'élevage. La SG a été mise en place dans les principales races bovines laitières à partir de 2009 (Guillaume et al. 2011). Il était attendu que le progrès génétique annuel puisse être doublé par la SG en raison de la forte réduction de l'intervalle de génération liée à l'abandon du testage sur descendance des taureaux (Boichard et al. 2016). L'augmentation du progrès génétique annuel réalisé en France grâce à la SG n'a été *in fine* que de 33, 50 et 71% pour les index de synthèse, respectivement pour les races bovines Holstein, Montbéliarde et Normande. Ces écarts entre progrès attendus et observés peuvent être expliqués par une utilisation de reproducteurs peu en cohérence avec les objectifs de sélection définis par les races et, pour la race Holstein, une nette augmentation de la consanguinité engendrant une perte de variabilité génétique (Doublet et al. 2019).

2.2. L'intérêt de la sélection génomique

La SG permet potentiellement une augmentation du progrès génétique annuel (ΔG) en intervenant sur plusieurs composantes de l'équation du sélectionneur décrite ci-dessous.

$$\Delta G = \frac{ir\sigma_g}{T},$$

- i : intensité de sélection ; c'est la différence de performances moyennes exprimée en unité d'écart-type entre les parents sélectionnés et les candidats à la sélection pour le caractère soumis à la sélection.
- r : la précision de l'évaluation génétique ou génomique.
- σ_g : l'écart type génétique du caractère.
- T : intervalle de génération ; c'est l'âge moyen des parents à la naissance de leurs descendants (qui seront eux-mêmes potentiellement reproducteurs), celui-ci est fortement dépendant de la biologie de l'espèce et des pratiques zootechniques.

La SG permet une évaluation précoce des candidats à la sélection, ce qui peut permettre de réduire l'intervalle de génération quand le phénotypage est tardif comme pour les caractères laitiers mesurés chez les filles d'un taureau. Cet aspect n'a que peu d'intérêt *a priori* en sélection piscicole où les caractères sont presque tous évalués avant la maturité sexuelle. Seuls les caractères de reproduction pourraient bénéficier de cet avantage.

En revanche, la SG peut sensiblement améliorer la précision de la sélection, en particulier pour les caractères non mesurables sur les candidats à la sélection comme les caractères létaux intéressants en aquaculture (rendement de découpe, résistance aux maladies) qui sont mesurés sur des collatéraux. En évaluation sur pedigree, sans information d'apparentement génomique, tous les individus d'une même famille partagent alors la même valeur génétique estimée. Or, l'aléa de méiose qui distingue deux pleins-frères ne peut être capturé que par la prise en compte de leurs ressemblances génomiques. La SG permet seule d'estimer une valeur génétique différente pour chaque individu au sein d'une même famille, améliorant ainsi la précision de l'évaluation et l'efficacité de la sélection à court terme, ainsi qu'à plus long terme en évitant de surreprésenter des familles parmi les reproducteurs des nouvelles générations et donc en limitant l'augmentation de la consanguinité (Daetwyler et al. 2007; Meuwissen et al. 2020).

La SG peut aussi permettre d'appliquer des intensités de sélection pour des caractères coûteux à phénotyper sur les candidats à la sélection en permettant d'augmenter le nombre d'individus

génotypés (sans phénotypage) ou en faisant une présélection sur la base des données génomiques avant le phénotypage.

2.3. Facteurs impactant la précision de l'évaluation génomique

Daetwyler et al (2008) et Goddard (2009) ont proposé des formules déterministes pour prédire l'efficacité de la précision de la SG en fonction de la taille de la population de référence, de l'héritabilité du caractère et du nombre de segments chromosomiques indépendants du génome (M_e), une entité théorique qui représente le nombre d'haplotypes indépendants dans la population de référence dont on doit estimer les effets grâce aux marqueurs. M_e est d'autant plus important que la diversité génétique de la population est grande et peut donc être prédite en fonction de l'effectif efficace (N_e) de la population qui représente le nombre théorique de reproducteurs nécessaires pour garder la même diversité génétique à la génération suivante sous l'hypothèse d'une population idéale de Wright-Fisher (taille constante, croisement panmictique, contributions équivalentes des reproducteurs à la génération suivante, ni mutation, migration ou sélection). Diverses formules proposent donc une approximation de M_e en fonction de N_e , mais toutes aboutissent à des prédictions de l'efficacité de la sélection bien souvent éloignées de la réalité (Brard and Ricard 2015).

La [Figure 3](#) illustre comment la précision de la SG varie selon la formule de Goddard (2009) en fonction de N_e , la taille de la population de référence et l'héritabilité du caractère. Nous pouvons voir qu'à taille de population de référence et héritabilité constantes, la précision est d'autant plus élevée que l'effectif efficace N_e est faible. En effet, plus l'effectif efficace est faible, moins le nombre de segments indépendants dont il faut estimer précisément les effets est important : une population de référence de faible taille peut alors être suffisante pour obtenir une bonne précision de la SG.

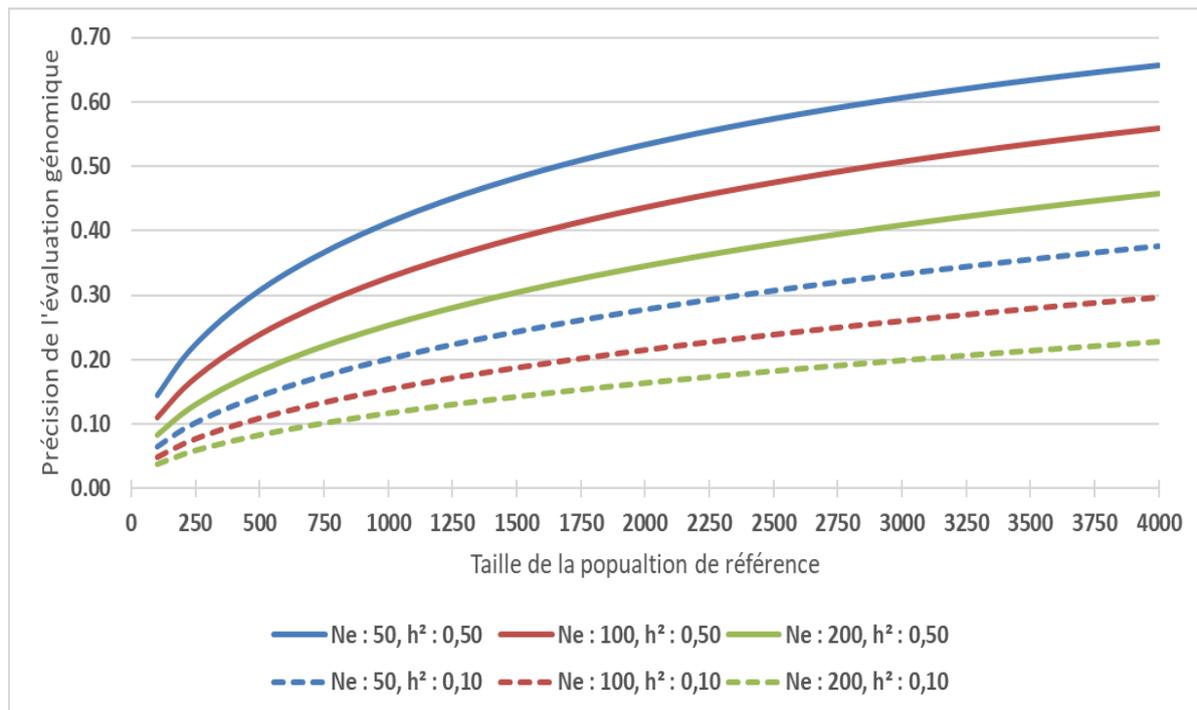


Figure 3 : Précision attendue d'après (Goddard 2009) de l'évaluation génomique en fonction de la taille de la population de référence, de l'héritabilité (h^2) du caractère et de l'effectif efficace de la population (N_e).

Par ailleurs, plus l'effectif efficace est faible, moins nombreux sont les marqueurs nécessaires pour estimer les segments chromosomiques indépendants (M_e). En effet, plus une population a de segments indépendants, plus grand est le nombre de marqueurs nécessaire pour couvrir l'ensemble des segments indépendants. Une étude de simulation d'un schéma de sélection basé sur 3 000 collatéraux de 100 familles à partir de 100 pères et 100 mères (1 père par mère soit environ 30 descendants par famille) avec un caractère d'héritabilité 0,4 montre peu d'évolution de la précision entre une densité de marqueurs de 0,5 et de 1 N_e par Morgan (Sonesson et Meuwissen 2009). Une autre étude avec un schéma d'accouplement de 50 pères et 50 mères où chaque père est accouplé à 20 mères avec 1 descendant par famille montre une augmentation importante de la précision jusqu'à une densité de 4 N_e par Morgan, puis une légère augmentation jusqu'à 8 N_e par Morgan (Solberg et al. 2008). Le N_e est donc un paramètre important à estimer pour évaluer les potentialités de l'utilisation de puces basses densités.

On peut aussi observer sur la [Figure 3](#) que la précision théorique de l'évaluation génomique augmente quand la population de référence augmente (à N_e constant) car l'augmentation de la population de référence permet d'augmenter la quantité d'information permettant une meilleure estimation des effets des M_e . On remarque une augmentation très rapide de la précision avec

l'augmentation de la population de référence mais cette augmentation plafonne lorsque la population de référence devient suffisamment grande. On note aussi que plus un caractère est héritable, plus la précision de la SG est importante à taille efficace et population de référence constantes. En effet, si un caractère est peu héritable, il est nécessaire d'avoir plus individus pour bien estimer les effets des Me dont la variabilité génétique est bruitée par la variabilité environnementale.

3. La sélection des salmonidés

3.1. Schémas de sélection historiques en salmonidés

3.1.1. Sélection massale

La sélection massale individuelle est simple et peu coûteuse à mettre en place et relativement efficace pour améliorer un caractère bien héritable et mesurable chez les candidats à la sélection. Ainsi, la croissance, la morphologie, la date de ponte par exemple peuvent être améliorées par sélection massale. Mais, pour les caractères létaux, la sélection massale est aussi possible en utilisant des critères indirects de mesure qui eux peuvent être évalués sur les animaux vivants. Par exemple, la teneur en lipides des filets peut être estimée par micro-ondes (Douirin et al. 1998) ou le rendement à l'éviscération estimé par des mesures échographiques (ultrasons) (Haffray et al. 2018). Cette méthode de sélection est particulièrement efficace en aquaculture, car la pression de sélection peut être forte grâce à la fécondité importante des espèces aquacoles. Mais sans la connaissance du pedigree des individus, un risque d'augmentation de la consanguinité existe pouvant limiter le progrès génétique à long terme (Felsenstein 1965). Pour maîtriser le risque d'augmentation de consanguinité, il est possible de conduire une sélection par « walk-back » par l'assignation de parenté d'un effectif adapté de candidats grâce à la reconstruction du pedigree par l'assignation de parenté avec des marqueurs génétiques de type microsatellites ou SNPs (Doyle et Herbinger ; 1994) dans le but d'équilibrer les représentations de l'ensemble des parents fondateurs de la génération en cours de sélection (Vandeputte et Haffray 2014).

3.1.2. Schéma de sélection de type « norvégien » en salmonidés

La sélection généalogique en aquaculture a débuté en 1972 (Gjedrem 2010) en Norvège pour le saumon Atlantique et la truite arc-en-ciel, avant d'être étendue aux principales espèces aquacoles (e.g., Vandeputte and Haffray 2014). La sélection généalogique permet i) d'accroître la précision de l'évaluation des caractères mesurés sur le candidat en intégrant l'information familiale ; ii) de sélectionner sur des caractères létaux mesurés sur des collatéraux des candidats ; et iii) de maîtriser l'évolution de la consanguinité. Les alevins à l'éclosion étant trop petits pour être marqués physiquement et ainsi identifiables, les descendants d'une famille (descendants d'un père et d'une mère) sont élevés dans un même bassin jusqu'à ce que leur marquage soit possible à l'aide d'un transpondeur électronique. Pour réaliser ces élevages en familles séparées, il est nécessaire d'investir dans des structures d'élevage très coûteuses avant la mise en commun, après marquage, des différentes familles dans un unique bassin (jusqu'à 800 familles pour le saumon chez AquaGen). Cependant, l'élevage en bassins par famille peut amener à surestimer la composante génétique en ne permettant pas de dissocier les effets génétiques additifs des effets environnementaux et maternels. Ce dispositif est très coûteux à mettre en place pour de petites et moyennes entreprises. Dans les années 1990, l'INRA a évalué chez la truite fario *Salmo trutta* l'intérêt d'une procédure de sélection massale efficace sur la croissance en maîtrisant les effets maternels non-génétiques de différences de tailles des œufs (Chevassus et al. 2004) et les effets de compétition entre les animaux (plusieurs remises en compétition). Cette PROCÉDURE de Sélection Par Epreuves Répétées (PROSPER) a montré une réponse réalisée de +130% du poids à 1 an après quatre générations (Chevassus et al. 2004).

3.1.3. Schéma de sélection truticole français

Suite à l'estimation de la réponse à une première génération de sélection de type PROSPER par l'INRA, le Syndicat des Sélectionneurs Avicoles Français (SYSAF) a été missionné en 1990 par le Ministère de l'Agriculture pour évaluer et proposer les modalités de transfert de la procédure de sélection PROSPER. En 1991, le SYSAF est devenu le Syndicat des Sélectionneurs Avicoles et Aquacoles Français (SYSAAF) en étendant son appui technique aux salmonidés pour conseiller techniquement les entreprises souhaitant transférer la procédure de sélection de type PROSPER. Initialement conduite sans connaissance des généalogies, la

croissance est le premier caractère sélectionné. Puis des critères indirects ont été introduits pour sélectionner des caractères létaux : teneurs en lipides estimée avec le Fatmeter (micro-ondes, Douirin et al. 1998) et les rendements à l'éviscération et au filetage estimés avec des mesures échographiques et une notation morphologique des candidats (Haffray et al. 2018).

Suite au projet de recherche Microsat-Truite (Haffray et al. 2001), les entreprises françaises ont introduit dans leurs programmes de sélection massale de type PROSPER une gestion des appariements assistée par empreintes génétiques (avec des marqueurs microsatellites). La validation répétée et rapide en situation de sélection commerciale de la faisabilité de l'utilisation des empreintes génétiques a justifié une deuxième phase d'expérimentation afin d'évaluer la faisabilité de l'introduction d'une indexation BLUP à partir de performances mesurées sur apparentés phénotypés pour des caractères de rendements de découpe (éviscération, filetage) ou de qualité de chair (couleur, adiposité) dans le cadre des projets PROSPER⁺ et Quality-Truite (Haffray et al, 2007a ; Haffray et al. 2007b).

Depuis 2004, l'assignation de parenté assistée par empreintes génétiques est utilisée en routine pour une indexation BLUP de caractères continus ou discontinus (Vandeputte et Haffray 2014). Cela permet des évaluations génétiques utilisant les informations pedigree qui améliorent la précision de la sélection en combinant phénotypes individuels et phénotypes des apparentés, en particulier pour les caractères non mesurables sur les candidats à la sélection comme les caractères létaux mesurés chez des collatéraux (rendement de découpe, qualité de la chair, caractères de reproduction ou liés au sexe). Il serait aussi possible de prendre en compte les interactions entre génotypes et environnements en considérant des collatéraux élevés dans divers milieux. Suite à des travaux de simulation, ce principe de sélection successivement massale puis généalogique a été étendu à des caractères à seuil comme la résistance à des pathologies (Chapuis et al. 2010).

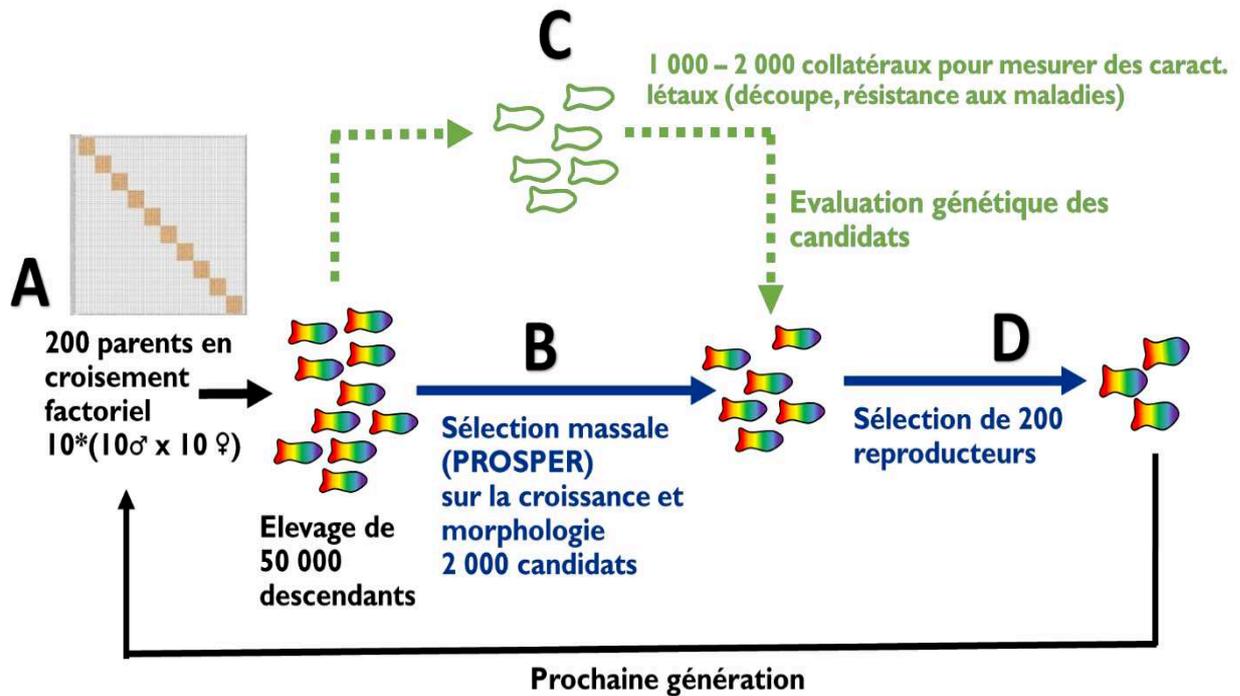


Figure 4 : Description du schéma de sélection actuellement appliqué par les entreprises de sélection trutticole françaises (d'après Haffray, 2018)

Les schémas actuels de sélection chez la truite arc-en-ciel en France peuvent être décrits en 4 étapes :

A – Création de nombreuses familles selon plusieurs petits factoriels $10^*(10♂ \times 10♀)$ pour créer jusqu'à un millier de familles simultanément par fécondation artificielle à partir d'un plan de croisement limitant le risque de croisement entre individus pleins-germains, et équi-représentation des effectifs de descendants entre les familles. Regroupement des pontes pondues le même jour par groupes de taille moyenne d'œufs similaire (Figure 5). Les différents groupes sont ensuite alimentés en fonction de la taille des œufs ou du jour de ponte et de leurs croissances relatives afin de faire converger les moyennes de poids des différents groupes vers un même objectif de poids avant les tris de sélection (Figure 5) et de limiter les biais liés aux effets maternels initiaux.

B – Sélection massale selon la Procédure de Sélection Par Epreuves Répétées (PROSPER) avec 3 à 4 tris successifs sur la longueur avec remise en compétition des candidats entre groupes (Figure 5), permettant de limiter les effets des interactions sociales potentiellement induits par des effets non-génétiques (Chevassus et al. 2004). Cette sélection sur la croissance est suivie par une sélection massale sur les rendements

(ultrasons et morphologie externe) et/ou sur la teneur en lipides (micro-ondes) (Dourin et al. 1998 ; Haffray et al. 2018).

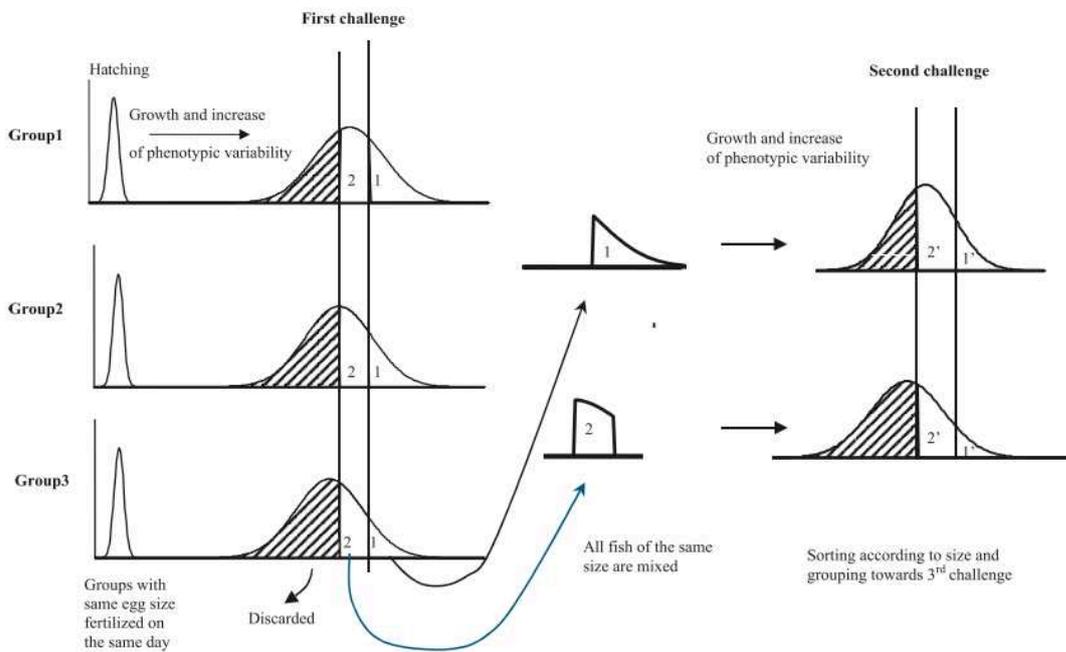


Figure 5 : Illustration de la méthode de sélection massive intra-groupe PROSPER (Chevassus et al. 2004)

C – Assignation de parenté des candidats présélectionnés de façon massive et indexation sur des caractères létaux (résistance à une pathologie, caractères de découpe) mesurés sur des collatéraux dérivés avant le 1^{er} tri de sélection massive des candidats.

D - Sélection des reproducteurs et proposition d'accouplement accouplements pour une espérance de progrès génétique fixée en minimisant l'augmentation de consanguinité par une procédure d'optimisation des contributions (OCS : optimum contribution selection) appelée « Sélection à Parenté Minimum » (SPM). Cette méthode utilise un algorithme de recuit simulé pour résoudre le problème d'optimisation sous contraintes (Colleau et al. 2009; Chapuis et al. 2016)

Ce mode de sélection combinant successivement sélection massive et sélection généalogique de type BLUP sur apparentés a été conduit à partir de 2004 par les entreprises.

3.2. Éléments sur les intérêts potentiels de la sélection génomique chez les salmonidés

Le principal intérêt de la SG dans les schémas de sélection aquacole pourrait être l'augmentation de la précision de l'évaluation des valeurs génétiques pour les caractères mesurés chez des collatéraux des candidats à la sélection. Ces caractères, rendement de découpe, qualité de la chair, sont des composantes importantes des objectifs de sélection actuels.

La SG pourrait aussi permettre d'envisager une sélection sur de nouveaux caractères difficiles et/ou coûteux à mesurer. Ainsi l'un des défis majeurs pour la durabilité des filières aquacoles est l'amélioration des caractères de résistance aux maladies, pour lesquels le potentiel de sélection génétique est important (Gjedrem 2015) et dont la sélection génomique peut apporter un gain majeur d'efficacité, ces caractères étant difficiles à sélectionner dans les schémas généalogiques (Ødegård et al. 2014). Les premières études visant à évaluer l'efficacité de la SG en salmonidés ont été publiées dès 2015 pour le saumon Atlantique et 2016 pour la truite arc-en-ciel. Ces études portent principalement sur les caractères de production et de résistance aux maladies (Tableau 1). Les gains de précision estimés pour la SG par rapport à la sélection sur pedigree varient selon les caractères, de 8% (survie à l'IPN) à plus de 300% (rendement en filet) en fonction des travaux (Tableau 1). Les gains sont d'autant plus importants que les précisions sont faibles en sélection généalogique classique.

Les premières puces à ADN avec plusieurs milliers de marqueurs SNPs datent de 2014 en aquaculture. Elles concernent le saumon Atlantique pour des puces à haute densité (HD) (112K (Houston et al. 2014), 200K (Yáñez et al. 2014)) et 2015 chez la truite arc-en-ciel avec une puce moyenne densité 57K Axiom® Trout Genotyping Array de ThermoFisher (Palti et al. 2015).

Études	Esp.	Caractères étudiés	h ²	N° SNP	PR	BLUP	SG	Gain (%)
Tsai et al. 2015	SA	Poids (1an)	0,48	112K	425	0,56	0,66	18
Tsai et al. 2015	SA	Poids (1an)	0,48	10K	425	0,56	0,66	18
Tsai et al. 2015	SA	Poids (1an)	0,48	5K	425	0,56	0,66	18
Tsai et al. 2015	SA	Longueur (1 an)	0,51	112K	425	0,58	0,70	21
Tsai et al. 2015	SA	Longueur (1 an)	0,51	10K	425	0,58	0,70	21
Tsai et al. 2015	SA	Longueur (1 an)	0,51	5K	425	0,58	0,69	19
Bangera et al. 2017b	SA	Rendement éviscéré	0,46	47K	2031	0,33	0,39	18
Bangera et al. 2017b	SA	Couleur filet	0,31	47K	2031	0,47	0,66	40
Bangera et al. 2017b	SA	Filet (%)	0,02	47K	2031	0,32	0,35	9
Bangera et al. 2017b	SA	Teneur en lipides dans le filet (%)	0,28	47K	2031	0,20	0,22	10
Tsai et al. 2017	SA	Poids (1 an)	-	25K	500	0,58	0,69	19
Yoshida et al. 2018	SA	Poids (25 mois)	-	37K	771	0,66	0,73	11
Gonzalez-Pena et al. 2016	TA	Poids (10 mois, mesure sur candidat)	0,37	42K	2467	0,66	0,70	6
Gonzalez-Pena et al. 2016	TA	Poids de la carcasse	0,65	42K	2467	0,28	0,55	96
Gonzalez-Pena et al. 2016	TA	Poids du filet	0,44	42K	2467	0,25	0,50	100
Gonzalez-Pena et al. 2016	TA	Rendement filets (%)	0,36	42K	2467	0,13	0,55	323
Robledo et al. 2018	SA	Score de résistance à l'Amoebic Gill Disease à <i>Neoparamoeba perurans</i> (AGD)	0,25	7K	1144	0,51	0,62	22
Robledo et al. 2018	SA	Charge qPCR à l'AGD	0,36	7K	1144	0,60	0,70	17
Bangera et al. 2017a	SA	Survie à <i>Piscirickettsia salmonis</i>	0,19	49K	2080	0,58	0,64	10
Bangera et al. 2017a	SA	Statut mort/vivant à <i>P. salmonis</i>	0,26	49K	2080	0,45	0,51	13
Tsai et al. 2017	SA	Résistance à <i>Lepeophtheirus salmonis</i>	-	25K	500	0,48	0,60	25
Tsai et al. 2016	SA	Résistance à <i>L. salmonis</i> (pop 1)	0,33	33K	425	0,48	0,60	25
Tsai et al. 2016	SA	Résistance à <i>L. salmonis</i> (pop 2)	0,22	33K	470	0,41	0,45	10
Vallejo et al. 2017	TA	Résistance <i>Flavobacterium psychrophilum</i> (survie)	0,37	35K	1473 (7900 p.)*	0,34	0,63	85
Vallejo et al. 2017	TA	Résistance <i>F. psychrophilum</i> (statut mort/vivant)	0,35	35K	1473 (7900 p.)*	0,36	0,66	83
Vallejo et al. 2018	TA	Résistance <i>F. psychrophilum</i> (statut mort/vivant)	0,35	9K	1473 (7900 p.)*	0,36	0,64	78
Vallejo et al. 2018	TA	Résistance <i>F. psychrophilum</i> (statut mort/vivant)	0,35	3K	1473 (7900 p.)*	0,36	0,61	69
Yoshida et al. 2018	TA	Résistance à <i>P. salmonis</i> (statut mort/vivant)	0,40	27K	1475	0,78	0,89	13
Yoshida et al. 2018	TA	Résistance à <i>P. salmonis</i> (statut mort/vivant)	0,56	27K	1475	0,69	0,77	13
Yoshida et al. 2019	TA	Survie au virus de la Nécrose Pancréatique Infectieuse, IPN)	0,40	38K	577 (2154 p.)*	0,49	0,53	8
Yoshida et al. 2019	TA	Statut mort/vivant à la IPN	0,35	38K	577 (2154 p.)*	0,50	0,56	12

Tableau 1 : Liste des différentes études sur l'efficacité de la sélection génomique par rapport à une sélection sur pedigree en salmonidés.

Esp : espèces (TA : truite arc-en-ciel, SA : saumon Atlantique) ; h² : héritabilité du caractère ; N° SNP : nombre de marqueurs utilisés dans l'étude ; PR : nombre d'individus dans la population de référence (*, la population est constituée d'individu génotype et phénotypé + p. individus seulement phénotypés); BLUP : précision d'une évaluation sur pedigree, SG : précision de l'évaluation génomique ; gain : gain de précision en pourcentage de la SG par rapport à la sélection BLUP.

Des travaux de recherche internationaux (voir ci-dessous), conduits en partenariat avec le SYSAAF, dans le cadre du projet européen FishBoost (<http://www.fishboost.eu>), sur le bar et la daurade (Aslam et al. 2018; Palaiokostas et al. 2018) et le projet 57K chez la truite arc-en-ciel (Fraslin et al. 2019), ont permis d'évaluer l'intérêt de l'utilisation de la sélection génomique. Sur la base de l'ensemble de ces travaux, le projet de thèse sur l'évaluation de la faisabilité de la sélection génomique chez la truite arc-en-ciel est né.

4. Les objectifs de la thèse

Ma thèse CIFRE (Conventions Industrielles de Formation par la Recherche) s'inscrit dans le cadre d'une collaboration de longue date entre le SYSAAF et l'INRAE pour améliorer l'efficacité et la durabilité des schémas de sélection piscicoles. Elle s'appuie sur les données recueillies au sein des projets « 57K-Truite » financé par FranceAgrimer (n°2015-0638) et « SG-Truite » financé par le Fonds Européen pour les Affaires Maritimes et de la Pêche (FEAMP) (RFEA47 0016FA 1000016), projets en partenariat fort entre le SYSAAF, ses adhérents et l'INRAE.

Dans ce contexte, l'objectif de la thèse est d'évaluer la faisabilité et l'intérêt de la sélection génomique par rapport à la sélection développée en France combinant successivement sélection massale et sélection sur apparentés (Haffray, 2018). En étudiant les principaux caractères d'intérêt aquacole et des objectifs de sélection prioritaires (reproduction, production et résistance à une maladie) pour chacune des entreprises de sélection françaises, il s'agit d'évaluer les méthodes et outils permettant de mettre en place une SG techniquement et économiquement viable.

Le premier chapitre décrivant les travaux de thèse (chapitre 2) concerne l'analyse des premières données obtenues avec la puce 57K dans le projet 57K-Truite. Les objectifs de cette partie ont été de : (i) vérifier la qualité des marqueurs de la puce 57K dans les populations françaises (ii) estimer le déséquilibre de liaison entre marqueurs dans les populations françaises pour estimer si une SG pouvait potentiellement être efficace avec cet outil (iii) étudier la diversité génétique des populations françaises de truites arc-en-ciel en sélection. Les principaux résultats sont présentés dans un article publié dans le journal *Genetics Selection Evolution*. L'étude est prolongée par une analyse de la diversité génétique des données obtenues dans SG-Truite qui

permet de confirmer les premiers résultats obtenus et d'étudier dans deux populations l'évolution de la consanguinité après 1 à 2 génération(s) supplémentaire(s) de sélection.

Les chapitres 3 à 5 traitent séparément les analyses génétiques et génomiques conduites pour chacune des trois populations des sélectionneurs du projet SG-Truite. Le chapitre 3 traite des caractères de reproduction femelle évalués dans la lignée sélectionnée par Viviers de Sarrance, le chapitre 4 des caractères de découpe et qualité de la chair mesurés dans une des lignées d'Aqualande et le chapitre 5 de la résistance à l'IPN évaluée dans la lignée en sélection de Bretagne-Truite. Chaque chapitre fait l'objet d'une étude de l'architecture des caractères en estimant les héritabilités et les corrélations génétiques entre les caractères étudiés. Des études d'association pangénomique (GWAS) sont rapportées d'une part pour les caractères de reproduction (chapitre 3) pour lesquels aucune publication n'a été rapportée à ce jour dans la littérature et d'autre part pour la résistance à IPN où deux QTLs expliquant 27 et 30% de la variance phénotypique chez la truite arc-en-ciel.

Pour les chapitres 3 à 5, l'efficacité de la sélection génomique est estimée et comparée à l'efficacité de la sélection sur pedigree. Dans les chapitres 3 et 4, nous nous sommes intéressés à l'optimisation technico-économique de la SG en identifiant des voies visant à limiter les coûts de la SG comme l'utilisation de puces LD ou le cumul d'information de cohortes différentes en population de référence.

Enfin, le chapitre 6 résume et discute les principaux résultats de la thèse afin de proposer des éléments d'optimisation pour mettre en œuvre des schémas de sélection revisités en introduisant la sélection génomique dans le but d'être techniquement plus efficaces, mais aussi économiquement viables.

5. Références

AGRESTE. Enquête Aquaculture 2018 - Recensement Pisciculture, culture d'algues et de cyanobactéries 2018 [internet : <https://agreste.agriculture.gouv.fr/agreste-web/disaron/ChdAgri2003/detail/>]. Vol. 3, Chiffres et données. 2020 May.

AGRESTE. Recensements 2008 de la salmoniculture et de la pisciculture marine et des élevages d'esturgeons [internet : <https://agreste.agriculture.gouv.fr/agreste-web/disaron/Dos11/detail/>]. Vol. 11, Agreste Les Dossiers. 2011.

Ali A, Al-Tobasei R, Lourenco D, Leeds T, Kenney B, Salem M. Genome-wide identification

- of loci associated with growth in rainbow trout. *BMC Genomics*. 2020;21(1):209. <https://doi.org/10.1186/s12864-020-6617-x>
- Aslam ML, Carraro R, Bestin A, Cariou S, Sonesson AK, Bruant JS, et al. Genetics of resistance to photobacteriosis in gilthead sea bream (*Sparus aurata*) using 2b-RAD sequencing. *BMC Genet*. 2018;19(1):1–14. <https://doi.org/10.1186/s12863-018-0631-x>
- Bangera R, Correa K, Lhorente JP, Figueroa R, Yáñez JM. Genomic predictions can accelerate selection for resistance against *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*). *BMC Genomics*. 2017a;18(1):121. <https://doi.org/10.1186/s12864-017-3487-y>
- Bangera, Thorland, Hillestad, Vela, Johansen, Refstie, et al. Accuracy of genomic predictions for yield traits , flesh colour and fillet fat in Atlantic salmon (*Salmo salar*). In: Proceedings of the World Congresson Genetics Applied to Livestock Production. 2017b. p. 11.627.
- Boichard D, Ducrocq V, Croiseau P, Fritz S. Genomic selection in domestic animals: Principles, applications and perspectives. *Comptes Rendus - Biol*. 2016;339(7–8):274–7. <https://doi.org/10.1016/j.crvi.2016.04.007>
- Brard S, Ricard A. Is the use of formulae a reliable way to predict the accuracy of genomic selection? *J Anim Breed Genet*. 2015;132(3):207–17. <https://doi.org/10.1111/jbg.12123>
- Chapuis H, Pincent C, Colleau JJ. Optimizing selection with several constraints in poultry breeding. *J Anim Breed Genet*. 2016;133(1):3–12. <https://doi.org/10.1111/jbg.12178>
- Chapuis H, Vandeputte M, Dupont-Nivet, M. Haffray P, Quillet E. Selection for an improved disease resistance using factorial mating designs and molecular based pedigrees in fish: a simulation study. In: 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. 2010. p. 1–6.
- Chavanne H, Janssen K, Hofherr J, Contini F, Haffray P, Aquatrace Consortium, et al. A comprehensive survey on selective breeding programs and seed market in the European aquaculture fish industry. *Aquac Int*. 2016;24(5):1287–307. <https://doi.org/10.1007/s10499-016-9985-0>
- Chevassus B, Quillet E, Krieg F, Hollebecq M-G, Mambrini M, Fauré A, et al. Enhanced individual selection for selecting fast growing fish: the “PROSPER” method, with application on brown trout (*Salmo trutta fario*). *Genet Sel Evol*. 2004;36(6):643–61. <https://doi.org/10.1051/gse:2004022>
- Colleau JJ, Tual K, De Preaumont H, Regaldo D. A mating method accounting for inbreeding and multi-trait selection in dairy cattle populations. *Genet Sel Evol*. 2009;41(7):1–10. <https://doi.org/10.1186/1297-9686-41-7>
- Daetwyler HD, Villanueva B, Bijma P, Woolliams JA. Inbreeding in genome-wide selection. *J*

- Anim Breed Genet. 2007;124(6):369–76. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>
- Daetwyler HD, Villanueva B, Woolliams JA. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS One. 2008;3(10). <https://doi.org/10.1371/journal.pone.0003395>
- Doublet AC, Croiseau P, Fritz S, Michenet A, Hozé C, Danchin-Burge C, et al. The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. Genet Sel Evol. 2019;51(1):1–13. <https://doi.org/10.1186/s12711-019-0495-1>
- Douirin C, Haffray P, Vallet JL, Fauconneau B. Détermination de la teneur en lipides des filets de truite arc-en-ciel *Oncorhynchus mykiss* par le Torry Fish Fat Meter. Sci Aliments. 1998;18:527–35.
- Doyle, R. W., Herbinger, C., M., 1994. Walk-back: The use of DNA fingerprinting for high-intensity, within-family selection in fish breeding. Proceedings 5th World Congress, Genetics Applied to Livestock Production. Volume: 19
- FAO. The State of World Fisheries and Aquaculture 2020. FAO; 2020. <https://doi.org/10.4060/ca9229en>
- Felsenstein J. The effect of linkage on directional selection. Genetics. 1965;52(2):349–63.
- Fraslin C, Brard-Fudulea S, D’Ambrosio J, Bestin A, Charles M, Haffray P, et al. Rainbow trout resistance to bacterial cold water disease: two new quantitative trait loci identified after a natural disease outbreak on a French farm. Anim Genet. 2019;50(3):293–7. <https://doi.org/10.1111/age.12777>
- Gjedrem T. The first family-based breeding program in aquaculture. Rev Aquac. 2010;2(1):2–15. <https://doi.org/10.1111/j.1753-5131.2010.01011.x>
- Gjedrem T. Disease resistant fish and shellfish are within reach: A review. J Mar Sci Eng. 2015;3(1):146–53. <https://doi.org/10.3390/jmse3010146>
- Goddard M. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica. 2009;136(2):245–57. <https://doi.org/10.1007/s10709-008-9308-0>
- Gonzalez-Pena D, Gao G, Baranski M, Moen T, Cleveland BM, Kenney PB, et al. Genome-Wide Association Study for Identifying Loci that Affect Fillet Yield, Carcass, and Body Weight Traits in Rainbow Trout (*Oncorhynchus mykiss*). Front Genet. 2016;7(NOV). <https://doi.org/10.3389/fgene.2016.00203>
- Guillaume F, Boichard D, Ducrocq V, Fritz S. Utilisation de la sélection génomique chez les bovins laitiers. INRAE Prod Anim. 2011;24(4):363–8. <https://doi.org/10.20870/productions-animales.2011.24.4.3269>

- Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. 2007;177(4):2389–97. <https://doi.org/10.1534/genetics.107.081190>
- Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*. 2011;12. <https://doi.org/10.1186/1471-2105-12-186>
- Haffray P., Dupont-Nivet M., Amigues Y., Vandeputte M., Pincet C., Deshayes C. and Chevassus B., 2001. Amélioration de la gestion des reproducteurs en trutticulture : Validation de l'intérêt des empreintes génétiques. Compte rendu final de la convention de recherche 033/00/C-OFIMER/CIPA de l'OFIMER.
- Haffray P., Chapuis H., Pincet C., Jégaden M., Coudurier B., Doerflinger J., Petit V., Vandeputte M. et Dupont-Nivet M., 2007a. PROSPER+. Evaluation et amélioration de la procédure de sélection PROSPER : optimisation des méthodes actuelles et introduction d'indexations sur collatéraux. Rapport final Ministère de l'Agriculture et de la Pêche, Génétique animale. Action innovante n° 2005 PA01 et IFOP n°2005/159.
- Haffray, P., Chapuis, H., Coudurier, B., Pincet, C., Petit, V., Mazeiraud, E., Doerflinger, J., Pantano, T., Chatain, B., Choubert, G., Bugeon, J., Vandeputte, M., Dupont-Nivet, M., 2007b. QUALITYTRUITE : Introduction d'une évaluation généalogique qualitative en sélection de type PROSPER. Rapport final des conventions OFIMER n°037/05/C et IFOP n°206/2005.
- Haffray P, Enez F, Bugeon J, Chapuis H, Dupont-Nivet M, Chatain B, et al. Accuracy of BLUP breeding values in a factorial mating design with mixed families and marker-based parentage assignment in rainbow trout *Oncorhynchus mykiss*. *Aquaculture*. 2018;490(March):350–4. <https://doi.org/10.1016/j.aquaculture.2018.03.003>
- Haffray P. Amélioration d'un programme de sélection massale sur la croissance chez la truite arc-en-ciel par introduction d'une sélection BLUP pour des caractères de qualités grâce aux empreintes génétiques. *Génétique animale*. Agrocampus Ouest. 2019.
- Henderson CR. Best linear unbiased estimation and prediction under a selection model published by : international biometric society stable. *Biometrics*. 1975;31(2):423–47.
- Houston RD, Taggart JB, Cézard T, Bekaert M, Lowe NR, Downing A, et al. Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics*. 2014;15(1):1–13. <https://doi.org/10.1186/1471-2164-15-90>
- Legarra A, Aguilar I, Misztal I. A relationship matrix including full pedigree and genomic information. *J Dairy Sci*. 2009;92(9):4656–63. <https://doi.org/10.3168/jds.2009-2061>

- Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001;157(4):1819–29. <https://doi.org/11290733>
- Meuwissen THE, Sonesson AK, Gebregiorgis G, Woolliams JA. Management of Genetic Diversity in the Era of Genomics. *Front Genet*. 2020;11(August):1–16. <https://doi.org/10.3389/fgene.2020.00880>
- Neto RVR, Yoshida GM, Lhorente JP, Yáñez JM. Genome-wide association analysis for body weight identifies candidate genes related to development and metabolism in rainbow trout (*Oncorhynchus mykiss*). *Mol Genet Genomics*. 2019;294(3):563–71. <https://doi.org/10.1007/s00438-018-1518-2>
- Ødegård J, Moen T, Santi N, Korsvoll SA, Kjøglum S, Meuwisse THE. Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*). *Front Genet*. 2014;5(NOV):1–8. <https://doi.org/10.3389/fgene.2014.00402>
- Ozaki A, Sakamoto T, Khoo S, Nakamura K, Coimbra MRM, Akutsu T, et al. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Mol Gen Genet*. 2001;265(1):23–31. <https://doi.org/10.1007/s004380000392>
- Palaiokostas C, Cariou S, Bestin A, Bruant JS, Haffray P, Morin T, et al. Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet Sel Evol*. 2018;50(1):1–11. <https://doi.org/10.1186/s12711-018-0401-2>
- Palti Y, Gao G, Liu S, Kent MP, Lien S, Miller MR, et al. The development and characterization of a 57K SNP array for rainbow trout. *Mol Ecol Resour*. 2015;15(3):662–72. <https://doi.org/10.1111/1755-0998.12337>
- Robledo D, Matika O, Hamilton A, Houston RD. Genome-wide association and genomic selection for resistance to amoebic gill disease in Atlantic salmon. *G3 Genes, Genomes, Genet*. 2018;8(4):1195–203. <https://doi.org/10.1534/g3.118.200075>
- Solberg TR, Sonesson AK, Woolliams JA, Meuwissen THE. Genomic selection using different marker types and densities. *J Anim Sci*. 2008;86(10):2447–54. <https://doi.org/10.2527/jas.2007-0010>
- Sonesson AK, Meuwissen TH. Testing strategies for genomic selection in aquaculture breeding programs. *Genet Sel Evol*. 2009;41(1):1–9. <https://doi.org/10.1186/1297-9686-41-37>
- Tsai HY, Hamilton A, Tinch AE, Guy DR, Bron JE, Taggart JB, et al. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet Sel Evol*. 2016;48(1):1–11. <https://doi.org/10.1186/s12711-016-0226-9>

- Tsai HY, Hamilton A, Tinch AE, Guy DR, Gharbi K, Stear MJ, et al. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genomics*. 2015;16(1):1–9. <https://doi.org/10.1186/s12864-015-2117-9>
- Tsai HY, Matika O, Edwards SMK, Antolín-Sánchez R, Hamilton A, Guy DR, et al. Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. *G3 Genes, Genomes, Genet*. 2017;7(4):1377–83. <https://doi.org/10.1534/g3.117.040717>
- Vallejo RL, Leeds TD, Gao G, Parsons JE, Martin KE, Evenhuis JP, et al. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genet Sel Evol*. 2017;49(1):1–13. <https://doi.org/10.1186/s12711-017-0293-6>
- Vallejo RL, Silva RMO, Evenhuis JP, Gao G, Sixin L, Parsons JE, 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*. 2018;(February):1–12. <https://doi.org/10.1111/jbg.12335>
- Vandeputte M, Haffray P. Parentage assignment with genomic markers: A major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. *Front Genet*. 2014;5(DEC):1–8. <https://doi.org/10.3389/fgene.2014.00432>
- VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91(11):4414–23. <https://doi.org/10.3168/jds.2007-0980>
- Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet Res (Camb)*. 2012;94(2):73–83. <https://doi.org/10.1017/S0016672312000274>
- Yáñez JM, Naswa S, López ME, Bassini L, Cabrejos ME, Gilbey J, et al. Development of a 200K SNP Array for Atlantic Salmon: Exploiting Across Continents Genetic Variation. *Proc World Congr Genet Appl to Livest Prod*. 2014;Species Br:263.
- Yoshida GM, Carvalheiro R, Lhorente JP, Correa K, Figueroa R, Houston RD, et al. Accuracy of genotype imputation and genomic predictions in a two-generation farmed Atlantic salmon population using high-density and low-density SNP panels. *Aquaculture*. 2018;491(March 2017):147–54. <https://doi.org/10.1016/j.aquaculture.2018.03.004>
- Yoshida GM, Carvalheiro R, Rodríguez FH, Lhorente JP, Yáñez JM. Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout. *Genomics*. 2019;111(2):127–32. <https://doi.org/10.1016/j.ygeno.2018.01.008>