

L'essor de l'aquaculture mondiale

Il y a 12 000 ans, l'homme commençait à domestiquer les espèces animales et végétales qu'il consommait. Ce n'est que 8 000 ans plus tard, environ 2 000 ans avant J.-C, qu'il débutait la domestication d'espèces aquacoles. Les premières traces d'élevage d'espèces aquacoles ont été trouvées en Chine où les hommes élevaient des carpes communes (*Cyprinus carpio*) dans un but principalement alimentaire vers 2 000 avant J.-C (Nash, 2011). A cette époque, l'aquaculture consistait essentiellement en un nourrissage de poissons capturés. Il a fallu attendre le V^{ème} siècle avant notre ère et le traité de la culture des poissons de Fan Li pour voir apparaître les prémices de l'aquaculture moderne (création des étangs d'élevage, sélection des géniteurs, gestion de la reproduction et maîtrise de l'alimentation et des conditions d'élevage). Cependant, contrairement à l'élevage, l'aquaculture ne s'est pas développée aussi rapidement et la pêche est restée, jusque très récemment, la principale source de produits aquatiques, aussi bien animales que végétales. A partir des années 1990 et malgré un effort de pêche qui s'intensifiait, les captures ont commencé à stagner (Bell et al., 2017) (Figure 1.1). La pêche souffre aujourd'hui d'une mauvaise image auprès des consommateurs, en lien avec la chute des stocks de poissons et son impact sur la biodiversité. Entre 2017, seulement 65,8% des stocks étaient considérés comme exploités de façon durable, contre 90% en 1990 (FAO, 2020).

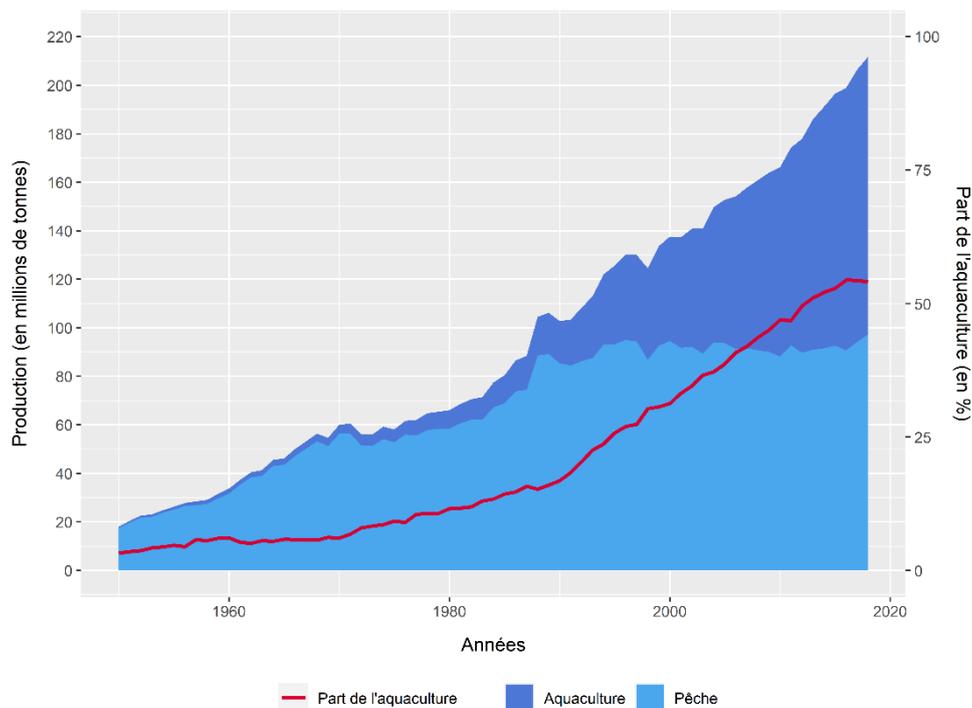


Figure 1.1 : Production de produits aquatiques via l'aquaculture (en bleu foncé) ou la pêche (en bleu clair) dans le monde entre 1950 et 2018. La courbe en rouge représente la part de l'aquaculture dans la production mondiale.

Parallèlement à cette stagnation des captures, la demande mondiale de produits aquatiques n'a cessé de croître, conséquence d'une augmentation de la population, des revenus et de l'urbanisation. La consommation mondiale de poissons est passée de 9,9 kg/an/habitant en 1960 à plus de 20 kg/an/habitant actuellement (FAO, 2020). Dans ce contexte, l'aquaculture s'est rapidement développée depuis les années 1970 pour devenir le secteur agricole à la plus forte croissance à ce jour, avec une croissance de 8,2% par an depuis 1950 (Figure 1.1). En tonnage, l'aquaculture a dépassé la production de viande ovine en 1986 et est aujourd'hui à un niveau comparable à celui de la production bovine (FAO, 1997). Cette croissance n'est cependant pas homogène, tant en localisation qu'en

espèces produites. La majorité de la production mondiale vient d'Asie, en particulier de Chine, qui produit environ 70 millions de tonnes, soit 87% de la production mondiale. C'est également en Chine que la croissance de la production est la plus forte. En Europe, où la croissance de la production est bien plus faible (environ 2,5% par an), la moitié de la production aquacole vient de l'élevage de poissons, qui se concentre autour de 5 espèces majeures : le saumon Atlantique (*Salmo salar*), la truite arc-en-ciel (*Oncorhynchus mykiss*), la carpe (*Cyprinus carpio*), le bar (*Dicentrarchus labrax*) et la daurade royale (*Sparus aurata*) (FAO, 2020). Parmi ces espèces, le saumon a connu une très forte croissance, surpassant de très loin les autres. Seuls le bar et la daurade connaissent aussi une croissance alors que les productions de carpe et de truite arc-en-ciel stagnent, voire diminuent.

Comme toutes les productions agricoles industrialisées, la pisciculture doit faire face à de nombreux enjeux (Ahmed et al., 2019). Le premier est l'utilisation de petits poissons pélagiques issus de captures de pêche pour la production d'huiles et de farines de poissons qui rentrent dans la composition de l'alimentation des animaux omnivores et carnivores. Leur usage pour l'alimentation animale rentre en compétition avec leur consommation par les humains et questionne sur la durabilité des stocks sauvages (Tacon and Metian, 2009). Cette production a augmenté entre 1976 et 1994 pour diminuer depuis (Tacon and Metian, 2009). Cette diminution est principalement due aux innovations dans la composition des aliments destinés aux poissons d'élevage, en particulier la substitution des farines et huiles de poissons par des farines et huiles d'origine végétale. Ainsi, l'alimentation des animaux d'aquaculture n'est plus une menace pour la durabilité de l'aquaculture (Tacon and Metian, 2008). Ensuite, à l'instar des autres productions animales, les questions de santé et de bien-être animal, de sécurité sanitaire et de qualité des produits sont centrales pour la durabilité et le développement de l'aquaculture. Dans le contexte de changement climatique actuel, la hausse des températures, l'acidification des océans ainsi que la diminution du taux d'oxygène dans l'eau sont autant d'enjeux pour l'avenir de l'aquaculture. Ces éléments joueront en particulier sur la survenue et la gravité des épidémies dans les élevages, qu'elles soient favorisées par l'augmentation de la température ou par l'émergence de nouveaux agents pathogènes (Reverter et al., 2020).

1.2. L'amélioration génétique comme réponse aux enjeux

La rapide croissance de la pisciculture a été permise par le perfectionnement des méthodes d'élevage et de la nutrition, ainsi que par la maîtrise de la reproduction et par la sélection génétique. Chez les autres espèces d'élevage, la sélection a permis une augmentation importante de la productivité. Avant même la théorisation de la sélection, les caractères de production comme la croissance, la quantité de lait ou d'œufs ont été les premiers à être sélectionnés. Chez le poulet de chair (*Gallus gallus domesticus*), le taux de croissance a été multiplié par 4,9 entre 1957 et 2001 (Havenstein et al., 2003) tandis que chez le porc (*Sus scrofa domesticus*), entre 1960 et 1996, la croissance a augmenté de 52% (Rauw et al., 1998).

En plus des caractères de productivité, la sélection a permis une amélioration de l'utilisation des ressources, en particulier en ce qui concerne l'alimentation des animaux mais aussi une diminution de l'impact environnemental. Ces deux aspects peuvent être combinés dans l'amélioration de l'efficacité alimentaire (Arthur and Herd, 2005; Besson et al., 2016). Chez le poulet de chair, l'efficacité alimentaire, mesurée par l'indice de conversion, est passée de 2,5 à 1,83 et chez le porc, de 3,24 à 2,26 entre 1960 et 1996 (Rauw et al., 1998).

L'amélioration de la résistance aux agents pathogènes, quant à elle permet à la fois une augmentation de la productivité par une baisse de la mortalité et aussi une diminution de l'utilisation de traitements, que ce soit des antibiotiques ou des vaccins. En aquaculture, la première mise en place documentée d'une sélection génétique a eu lieu en 1919 (Embrey and Hayford, 1925). Elle a été réalisée par sélection massale pour la résistance à la furunculose chez l'omble de fontaine (*Salvelinus fontinalis*). Après quatre générations de sélection, la mortalité avait chuté de 67%.

Cependant, la rationalisation de la sélection ne s'est vraiment mise en place que dans les années 1970 et s'est développée d'abord sur le saumon avec la mise en place de la sélection familiale et sur apparentés, bien plus efficace que la sélection massale (Gjedrem, 1985). Elle coïncide avec l'essor de l'aquaculture dans le monde. En 2010, il est estimé qu'environ 8,2% de la production aquacole mondiale provient de cheptels sélectionnés (Gjedrem et al., 2012), contre plus de 80% en Europe en 2016 (Janssen et al., 2017). Ceci montre à la fois une place particulière de l'Europe dans le monde de la sélection aquacole, et une forte dynamique sur son développement.

La sélection génétique peut permettre d'augmenter le chiffre d'affaire des fermes aquacoles par la réduction du temps d'élevage ou bien par la création de nouveaux marchés, comme cela été le cas avec la vente de truite fumée concurrençant le saumon fumé grâce à l'augmentation de la taille des truites arc-en-ciel obtenue par sélection. Elle peut également améliorer la rentabilité des fermes par la réduction des intrants, en particulier l'alimentation ou les produits de traitements, grâce à l'amélioration de l'efficacité alimentaire et de la résistance aux maladies. La sélection peut également jouer un rôle dans l'acceptation sociétale de l'aquaculture, car elle diminue l'impact écologique par la réduction des émissions de gaz à effet de serre, la réduction des rejets comme le phosphate et le nitrate provenant des déjections ou de l'aliment non ingéré (Besson et al., 2016), mais aussi par la réduction de l'usage des antibiotiques.

1.3. Le bar et la daurade, clés de l'aquaculture méditerranéenne

Le bar et la daurade sont les deux espèces emblématiques de l'aquaculture méditerranéenne. Elles représentaient 91% de la production en 2016, 48% provenant de la production de bar et 43% de celle de daurade, avec une production respective de 200 000 tonnes et de 160 000 tonnes en 2018. Ces deux espèces sont majoritairement élevées en Turquie, en Grèce, en Egypte et en Espagne, qui produisent à elles seules environ 90% de la production (FAOSTAT, 2020). Chez ces deux espèces, l'aquaculture représente une large majorité de la production : depuis l'industrialisation de la production dans les années 1990, elle a rapidement pris le pas sur les captures de pêches jusqu'à représenter 97% de la production totale (Figure 1.2).

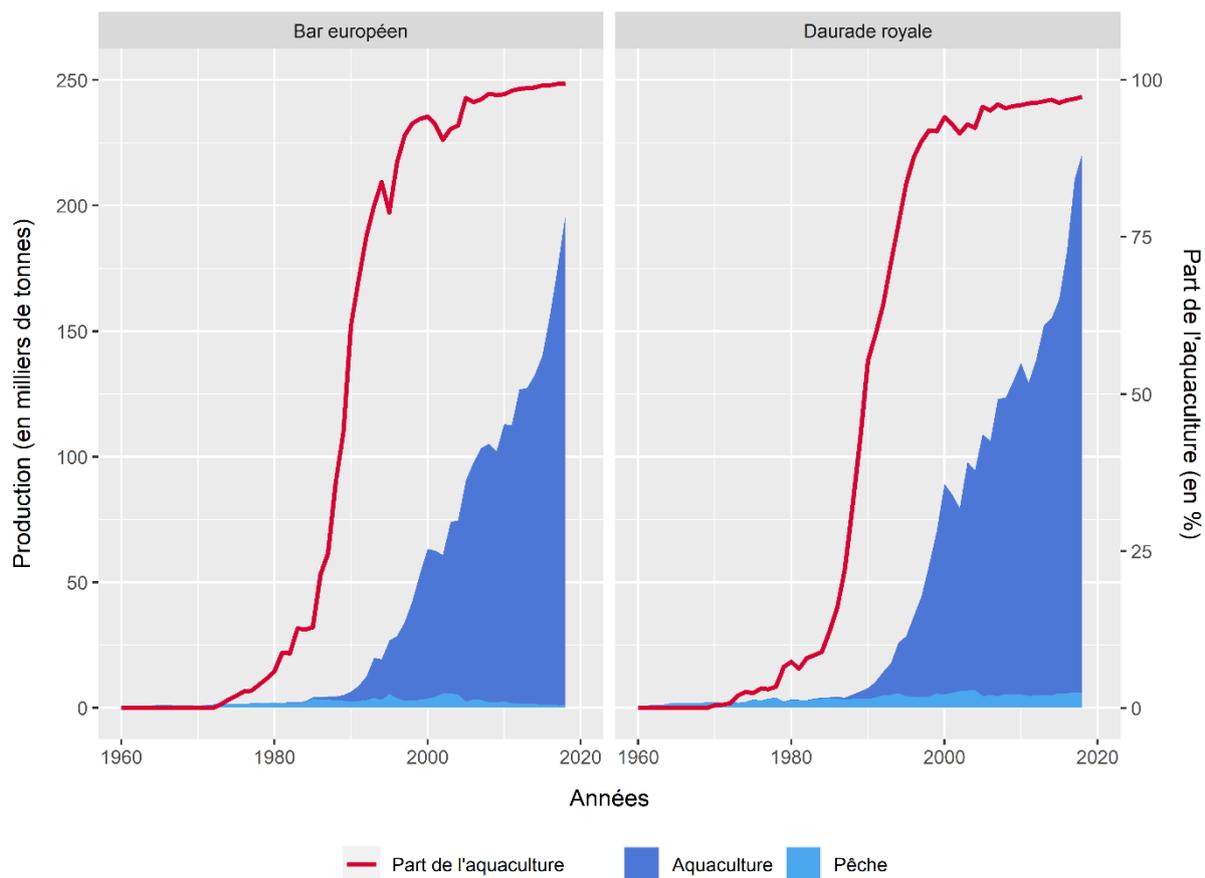


Figure 1.2 : Production de bar européen (panneau de gauche) et daurade royale (panneau de droite) issue de l'aquaculture (en bleu foncé) et de la pêche (en bleu clair) entre 1960 et 2018. La courbe en rouge représente la part de l'aquaculture dans la production.

Dans la nature, ces deux espèces se retrouvent principalement le long des côtes méditerranéennes et atlantiques, mais également en mer du Nord (Figure 1.3). Elles sont euryhalines et peuvent s'adapter aussi bien aux étangs qu'aux estuaires. Leur cycle de reproduction est proche. Le bar se reproduit de décembre à mars en mer Méditerranée et de mars à juin dans l'océan Atlantique (Vandeputte et al., 2019). La daurade se reproduit quant à elle d'octobre à décembre. Elles sont particulièrement prolifiques, produisant environ 200 000 œufs/kg/ponte pour le bar (Vandeputte et al., 2019) et entre 20 000 et 30 000 œufs/kg/jour pour la daurade dont la ponte s'étale sur plusieurs jours (Pousao-Ferreira et al., 1999).

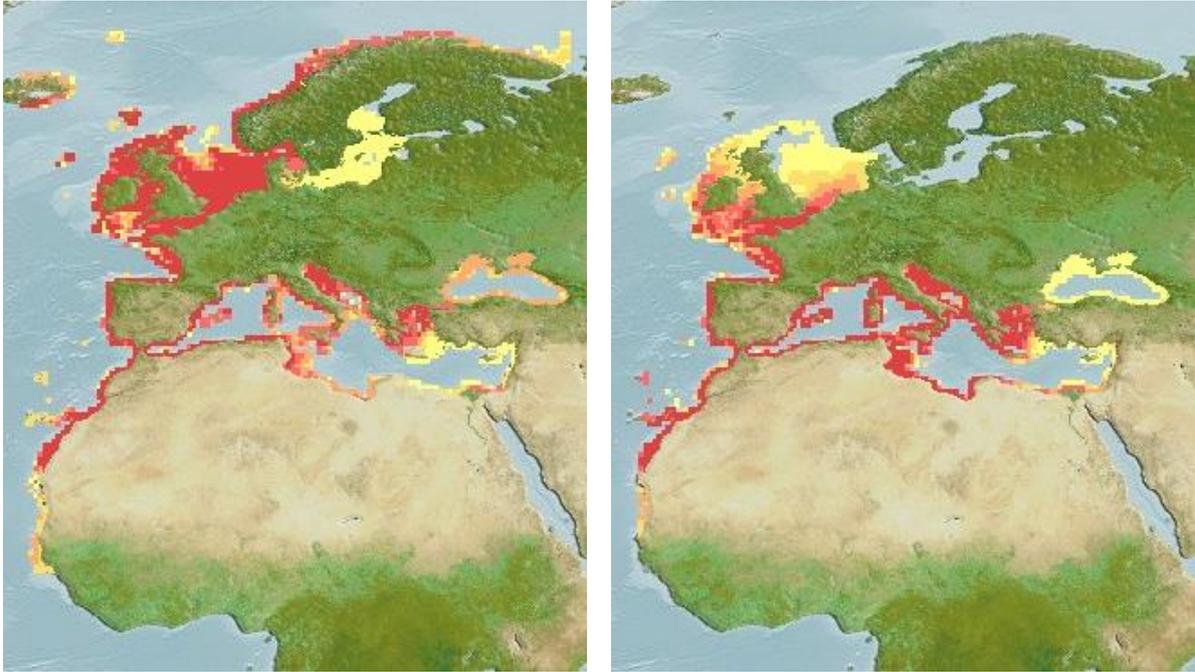


Figure 1.3 : Carte de répartition des populations sauvages de bar européen (panneau de gauche) et de daurade royale (panneau de droite). Source : aquamaps

Comme pour tout élevage, la maîtrise du cycle de vie est un impératif, en particulier celle de la reproduction en captivité, qui permet de choisir des géniteurs, et ainsi de domestiquer les espèces. Chez le bar et la daurade, les élevages ont commencé par la capture d'alevins sauvages dans les années 1970. Les poissons étaient ensuite élevés majoritairement dans des systèmes extensifs jusqu'à taille commerciale puis vendus. Dans les années 1980, chez le bar comme chez la daurade, la reproduction artificielle ainsi que l'élevage des larves et alevins a été maîtrisé, permettant une intensification de la production et le début de l'amélioration génétique par la domestication (Bagni, 2005; Colloca and Cerasi, 2005).

Malgré une part négligeable dans la production de ces deux espèces, la France est un leader dans la sélection génétique et la production d'alevins, qui est exportée à l'étranger à 90%. Deux entreprises réalisent la majeure partie de la production d'alevins en France : le groupe Aqualande par le biais de ses filiales Ferme Marine du Douhet et les Poissons du Soleil et le groupe Gloria Maris avec sa filiale Ecloserie Marine de Gravelines-Ichtus. Ces deux entreprises produisent respectivement 20 millions et 30 millions d'alevins sur les 60 millions produits en France. Elles sont également des leaders au niveau européen. Bien que concurrentes, ces deux entreprises sont adhérentes du SYSAAF (SYndicat des Sélectionneurs Avicoles et Aquacoles Français). Le SYSAAF offre un appui technique à la sélection, en optimisant les plans de croisement selon les objectifs du sélectionneur, en réalisant les estimations de valeurs génétiques sur plusieurs caractères sélectionnés et en favorisant le développement R&D des entreprises par la mutualisation des ressources, en particulier génétiques (performances, pedigree, génotypes).

1.4. Une production sensible aux pathogènes

Comme toute production agricole, l'aquaculture est touchée par de nombreux pathogènes. Ces épidémies sont principalement dues aux fortes densités d'individus dans les systèmes d'élevage (Reno, 1998). D'autres facteurs environnementaux comme la température, la qualité de l'eau et de l'alimentation jouent aussi un rôle.

Le bar et la daurade sont touchés par différentes maladies virales, bactériennes et parasitaires. Parmi ces nombreuses maladies, la nécrose nerveuse virale (NNV), la vibriose et la pasteurellose sont considérées comme ayant le plus d'impact sur la production, autant au stade alevin qu'à taille commerciale (Vendramin et al., 2016).

1.4.1. La nodavirose

La nodavirose est une maladie d'origine virale causée par le virus de la nécrose virale (VNN), appelé aussi virus de l'encéphalopathie et rétinopathie (VER). C'est un virus à ARN appartenant à la famille des betanodavirus qui provoque des nécroses des tissus nerveux (Mori et al., 1992). Il a été rapporté pour la première fois en 1985 chez le poisson perroquet au Japon (Yoshikoshi and Inoue, 1990) et ce n'est qu'en 1988 que les premières épidémies sont apparues en Méditerranée (Breuil et al., 1991). Actuellement, il se retrouve dans toutes les mers du globe, à l'exception de l'Amérique du sud où aucun cas n'a été rapporté à ce jour. Il est capable d'infecter plus de 70 espèces aquatiques, aussi bien d'eau douce que marines. En aquaculture, il touche principalement le bar, la daurade et le turbot en Europe mais aussi le barramundi, l'ombrine et le cabillaud. En Europe, les productions méditerranéennes sont particulièrement affectées car le VNN prolifère dans des eaux dont la température est supérieure à 20°C (Doan et al., 2017b).

Le virus se transmet par l'eau, dans laquelle il peut survivre un long moment sans hôte, mais également par contact d'un individu sain avec un individu porteur, qu'il ait des symptômes ou non. Les symptômes se caractérisent par une nage en rond ou en spirale et une dépigmentation de la peau (Miccoli et al., 2019). Ce comportement erratique est causé par la nécrose des tissus nerveux qui, si elle devient trop importante, provoque la mort de l'individu (Doan et al., 2017b). Dans la majorité des cas, une détection par PCR ou anticorps spécifique ainsi qu'un prélèvement de tissu nerveux pour vérifier la présence de cellules nécrosées permettent de confirmer le diagnostic. En plus de la transmission horizontale, le VNN peut se transmettre par voie verticale, d'un individu à sa descendance par le biais des œufs après infection des gonades (Costa and Thompson, 2016).

Face à ce virus, aucun mode de traitement n'est efficace à ce jour. Seuls des moyens préventifs permettent de réduire la fréquence des épidémies dans les fermes. Les principales méthodes résident dans le fait de ne pas importer le virus à l'intérieur de la structure, en procédant à des contrôles antigéniques ou PCR des poissons provenant de l'extérieur et à une mise en quarantaine. Dans le cas d'œufs, une désinfection par ultra-violet ou ozonation est le moyen le plus efficace (Doan et al., 2017b). A ce jour, un seul vaccin est disponible et offre une protection contre le virus jusqu'à 12 mois (Miccoli et al., 2019).

1.4.2. La vibriose

La vibriose est une maladie causée par les bactéries du genre *Vibrio*. Parmi les 123 espèces décrites appartenant à ce genre, les plus répandues en pisciculture sont *Vibrio parahaemolyticus*, *Vibrio alginolyticus*, *Vibrio harveyi*, *Vibrio owensii* et *Vibrio campbellii*. Principalement répandues dans les estuaires et les zones côtières (Thompson et al., 2004), elles jouent un rôle dans la dégradation de la matière organique (Damir et al., 2013). La première épidémie de vibriose a été rapportée en 1973 chez l'anguille en Norvège et au Royaume-Uni (McCarthy, 1976). Du fait de la grande diversité d'espèces bactériennes provoquant la vibriose, elle est capable de toucher aussi bien les poissons que les mollusques et les crustacés (Ina-Salwany et al., 2019).

La vibriose est une maladie opportuniste, souvent déclenchée par un stress ou lorsque les défenses immunitaires de l'hôte sont amoindries. Elle se développe préférentiellement à une température supérieure à 17°C et se transmet aussi bien par l'eau que par contact. Les individus contaminés présentent une extrême maigreur, une décoloration de la peau et un état léthargique. Dans un second temps, des ulcères et nécroses apparaissent sur la peau. Sans traitement, l'infection se généralise et la septicémie est responsable de la mort. La vibriose est également capable de se transmettre à la descendance en contaminant les œufs (Marco-Noales et al., 2001).

Au sein des fermes piscicoles, une bonne qualité de l'eau, en particulier une faible salinité est une mesure préventive pour éviter les épidémies de vibriose (Prayitno and Latchford, 1995). La réduction du stress, par la réduction des manipulations et de la densité peut également aider dans la prévention. Une dernière mesure préventive est la vaccination du cheptel grâce à des vaccins développés à partir de bactéries inactivées, qui offrent une protection allant de six semaines à une année (Miccoli et al., 2019). Une fois les poissons infectés, un traitement par antibiotique est la solution la plus couramment employée (Ina-Salwany et al., 2019).

1.4.3. La pasteurellose

La pasteurellose, ou photobactériose, est une maladie bactérienne causée par *Photobacterium damsela subsp. piscicida*. Appartenant à la famille des Vibrionacées, la pasteurellose est une forme de vibriose. A cause de sa large distribution géographique, de ses nombreuses antibiorésistances et de sa capacité à infecter différentes espèces marines, elle est vue comme une menace émergente pour la pisciculture méditerranéenne (Andreoni and Magnani, 2014). Elle a été pour la première fois rapportée dans des populations sauvages de perches blanches aux Etats-Unis en 1963 (Snieszko et al., 1964). C'est en 1970 au Japon que les premières espèces d'aquaculture, en l'occurrence des juvéniles de sérioles (*Seriola quinqueradiata*), sont touchées par la pasteurellose (Kubota et al., 1970). Il faudra attendre 1990 pour observer la première épidémie en Europe chez les juvéniles de daurade en Espagne (Toranzo et al., 1991).

Tout comme la vibriose, la pasteurellose se développe préférentiellement à une température supérieure de 15°C avec un optimum aux alentours des 22°C (Romalde, 2002). Les symptômes sont très proches de ceux de la vibriose (dépigmentation de la peau, nécrose - Toranzo et al., 1991).

Les méthodes de prévention contre la vibriose s'appliquent à la pasteurellose. Des vaccins ont été élaborés de la même manière, et consistent en l'injection de bactéries inactivées (Miccoli et al., 2019). Les antibiotiques s'avèrent n'être que très peu efficaces à cause des antibiorésistances de la bactérie, mais restent l'unique méthode de traitement à ce jour.

1.4.4. Vue globale

Bien qu'appartenant à des genres et même des règnes différents, ces pathogènes ont de nombreux points communs. Tout d'abord, ils sont tous capables d'infecter aussi bien le bar que la daurade et causent de fortes mortalités dans les élevages aquacoles (Doan et al., 2017b; Ina-Salwany et al., 2019; Romalde, 2002). Ils sont également plus virulents lorsque la température de l'eau est élevée et sont dits saisonniers puisqu'ils s'observent plus fréquemment pendant l'été. La fréquence d'apparition risque ainsi d'augmenter dans le contexte de changement climatique. Enfin ces pathogènes touchent particulièrement les juvéniles, causant des dégâts importants dans les éclosiers, et sont donc une menace importante pour les sélectionneurs (Doan et al., 2017b; Toranzo et al., 1991). L'ensemble des caractéristiques des agents pathogènes décrits précédemment sont résumés dans le Tableau 1.1.

Tableau 1.1 : Tableau récapitulatif des trois maladies décrites en 1.4.

	NODAVIROSE	VIBRIOSE	PASTEURELLOSE
Organisme pathogène	Virus de l'encéphalopathie et de la rétinopathie (VER) / Virus de la Nécrose Nerveuse (VNN)	<i>Vibrio parahaemolyticus</i> , <i>Vibrio alginolyticus</i> , <i>Vibrio harveyi</i> , <i>Vibrio owensii</i> et <i>Vibrio campbelli</i>	<i>Photobacterium damsela</i> subsp. <i>piscicida</i>
Règne	Virus à ARN	Bactérie gram -	Bactérie gram -
Famille	<i>Nodaviridae</i>	<i>Vibrionaceae</i>	<i>Vibrionaceae</i>
Genre	<i>Betanodavirus</i>	<i>Vibrio</i>	<i>Photobacterium</i>
Distribution géographique	Monde sauf Amérique du sud	Mondiale	Mondiale
Première découverte	Japon en 1985 chez le poisson perroquet	Norvège en 1973 chez l'anguille	USA en 1963 chez la perche blanche et le bar rayé
Symptômes visibles	Nage anormale/erratique (nage en spirale ou en rond), léthargie et assombrissement de la peau	Ulcères et nécroses cutanés, léthargie, maigreur excessive, nage erratique, assombrissement de la peau	Ulcères et nécroses cutanés, maigreur excessive, assombrissement de la peau
Conditions d'apparition	Eau > 15 - 20°C	Eau > 17°C et salinité entre 30 et 35 ppt, stress	Eau > 18 - 20°C
Transmission	<ul style="list-style-type: none"> • Horizontale : contact et eau • Verticale : gonades et œufs 	<ul style="list-style-type: none"> • Horizontale : contact et eau • Verticale : œufs 	<ul style="list-style-type: none"> • Horizontale : contact et eau • Verticale : œufs
Moyens de prévention	Quarantaine, traitement des œufs, contrôle de la qualité de l'eau, vaccination	Quarantaine, contrôle de la qualité de l'eau, vaccination	Quarantaine, contrôle de la qualité de l'eau, vaccination
Moyens de traitement	Aucun	Antibiotiques	Antibiotiques (mais nombreuses antibio-résistances)

1.5. L'apport des outils génomiques en sélection pour améliorer la résistance aux maladies

1.5.1. La sélection sans génomique

L'intérêt pour l'amélioration de la résistance aux maladies par sélection est apparu dès la mise en place de programmes de sélection en aquaculture (Embry and Hayford, 1925). A l'origine, c'est une sélection massale qui a été appliquée pour sélectionner les individus les plus résistants. Avec la sélection massale, les candidats à la sélection sont infectés par un pathogène, volontairement ou non. Les survivants, qui sont qualifiés de résistants, poursuivent les étapes de sélection avant de devenir des géniteurs et de produire la génération future. La sélection massale, bien que très simple à mettre en place, offre de nombreux inconvénients. Premièrement, les candidats à la sélection sont directement infectés par l'agent pathogène, ce qui peut provoquer des épidémies qui se propagent à l'ensemble l'écloserie, aussi bien à cause d'une transmission horizontale entre les différents cheptels présents, mais aussi à cause d'une transmission verticale qui se propagera dans la descendance. Ensuite, la pression de sélection ne peut pas être contrôlée puisqu'elle est égale au taux de survie, qui dépend donc de la sévérité de l'épisode de mortalité.

La sélection sur apparentés est l'alternative la plus couramment employée. Dans ce schéma, les candidats à la sélection sont séparés en deux groupes (Ødegård et al., 2011). L'un des groupes sera volontairement infecté par le pathogène pour lequel le sélectionneur souhaite améliorer la résistance de son cheptel. Cette contamination volontaire en milieu contrôlé est appelée challenge. Lors du challenge, la mortalité est enregistrée tous les jours. A la fin du challenge, les candidats à la sélection qui n'ont pas été challengés peuvent être indexés. Pour cela, leurs valeurs génétiques sont estimées grâce à un modèle BLUP (Best Linear Unbiased Prediction) à partir des phénotypes enregistrés dans le groupe challengé et des relations de parenté entre les individus des deux groupes qui sont enregistrées dans un pedigree. Ainsi, les candidats à la sélection, qui n'ont jamais été infectés par le pathogène, ont une valeur génétique qui retranscrit leur aptitude à résister à ce pathogène, ce qui permet de les classer et de les sélectionner pour produire de futurs géniteurs (Ødegård et al., 2011). La sélection sur apparentés limite le risque d'épidémie au sein de l'écloserie. Elle permet également de contrôler la pression de sélection, et donc de maîtriser l'évolution du gain génétique et de la consanguinité dans les générations futures. Cependant, la connaissance exacte du pedigree est une condition nécessaire à l'obtention de valeurs génétiques par cette méthode.

Pour cela, l'information familiale doit être conservée tout au long de la vie de l'individu. Chez les espèces d'élevage terrestre, cette information est gardée grâce à l'identification de l'individu à sa naissance par un système de marquage physique individuel. Grâce à la traçabilité de l'origine du père et de la mère, le pedigree d'un individu peut facilement être connu. Dans l'élevage de poissons, l'identification individuelle se fait généralement par l'injection d'une puce RFID. Or, à la naissance, il est impossible d'injecter de telles puces car les larves sont trop petites. En attendant une taille suffisante pour les identifier, il faut tout de même conserver l'information familiale. Pour cela, la première méthode a été d'élever chaque famille séparément. Cette méthode a l'inconvénient majeur de produire une confusion entre l'effet de l'environnement d'élevage (le bac) et l'effet de la famille (Vandeputte et Haffray, 2014). Le moyen d'éviter cette confusion est de mélanger plusieurs familles dans un même bac, mais alors l'information familiale est perdue. Cependant, grâce aux marqueurs génétiques, elle peut maintenant être retracée *a posteriori* (Vandeputte et Haffray, 2014). L'apport de l'assignation de parenté par marqueurs génétiques a été considérable dans les programmes de sélection aquacole. En plus de la confusion de l'effet de l'environnement d'élevage avec l'effet génétique, l'élevage en familles séparées ajoute une contrainte sur le nombre de familles à élever, qui

est limité par le nombre de bassins disponibles, et nécessite des infrastructures conséquentes (Vandeputte et al., 2009). Ainsi, l'assignation de parenté *a posteriori* par marqueurs génétiques a permis de lever ces contraintes, permettant un élevage des larves en familles mélangées et une augmentation du nombre de familles. Elle a donc été largement adoptée par les sélectionneurs en aquaculture.

1.5.2. Le développement de marqueurs génétiques

Les premiers marqueurs génétiques à être utilisés en aquaculture ont été des allozymes dans les années 1970 (Brody et al., 1981). Puis dans les années 1990, avec le développement des marqueurs microsatellites, l'utilisation des marqueurs dans les schémas de sélection s'est répandue (Herbinger et al., 1995). En effet, les microsatellites sont bien plus informatifs (ils présentent de nombreux allèles, jusqu'à plusieurs dizaines) et plus fréquents que les allozymes puisque plus nombreux dans le génome. Par leur efficacité en assignation de parenté, permettant de retracer le pedigree d'individus issus de croisement complexes, les marqueurs microsatellites ont permis la mise en place de l'élevage en familles mélangées (Marc Vandeputte and Haffray, 2014). Ensuite, ce sont les marqueurs SNPs (Single Nucleotide Polymorphism) qui ont été utilisés. Bien que ne présentant que deux allèles, ils sont encore plus nombreux que les marqueurs microsatellites et aussi uniformément repartis sur le génome. Chez la plupart des espèces, des millions de marqueurs SNPs peuvent être identifiés.

Dans les années 2000 avec la publication du génome humain (International Human Genome Sequencing Consortium, 2004), la possibilité du séquençage complet de l'ADN a été démontrée. Dans les années qui suivirent, les technologies n'ont cessé de s'améliorer et le coût du séquençage n'a cessé de baisser. De nos jours, le séquençage du génome d'une espèce ne coûte que quelques milliers d'euros contre cent millions de dollars pour celui du premier séquençage de l'humain. Grâce à l'essor des techniques de séquençage haut débit, le développement de marqueurs génomiques chez toutes les espèces, y compris aquacoles, s'est trouvé grandement facilité.

La séquence du génome du bar a été publiée en 2014 (Tine et al., 2014) et celle du génome de la daurade en 2018 (Pauletto et al., 2018). Grâce aux efforts fournis et à la qualité de l'assemblage, ces deux génomes font figures de référence. A partir de ces dates, l'identification de marqueurs génomiques a été beaucoup plus aisée. Les marqueurs SNPs peuvent être génotypés de nombreuses façons. Parmi elles, le RAD-seq fait partie d'une famille de méthodes appelée « génotypage par séquençage », et consiste en un séquençage de certaines régions du génome obtenues après digestion de la molécule d'ADN par des enzymes de restriction (Miller et al., 2007). Une fois la séquence des dites régions obtenue, elle est comparée à celle d'autres individus afin de détecter les SNPs. Le RAD-seq permet dans le même temps d'identifier et de génotyper des marqueurs SNPs, et cet avantage lui a valu d'être très utilisé chez des espèces avec peu ou pas de ressources génomiques disponibles. Un de ses inconvénients est que les marqueurs sont plus ou moins spécifiques de chaque expérimentation, et que de fait leurs répétabilité est faible entre expérimentations (Robledo et al., 2018).

Une seconde méthode est l'utilisation d'une puce à SNPs. Bien que des marqueurs issus de RAD-seq puissent être compilés sur une puce à SNPs, l'obtention des marqueurs se fait généralement autrement. Tout d'abord, une population de référence est séquencée. Cette population doit être suffisamment diversifiée pour refléter la diversité existante dans les autres populations qui seront génotypées avec la puce (population d'un institut de recherche ou d'entreprise, d'un consortium regroupant plusieurs entreprises et/ou instituts, espèce entière). Une fois les séquences obtenues, elles sont alignées sur le génome de référence puis les SNPs sont détectés. Parmi les millions de

marqueurs identifiés à ce stade, une étape de tri est nécessaire afin de sélectionner les marqueurs à figurer sur la puce. Les critères de choix dépendent de la stratégie de design de la puce à SNPs, mais se basent en général sur la qualité (Phred quality score) et la profondeur du séquençage, sur la variabilité de ces marqueurs dans la population de référence (minor allele frequency), sur la présence d'autres variants dans les régions attenantes proches, ainsi que la composition en bases des séquences flanquantes, positionnées en amont et en aval du polymorphisme (LaFramboise, 2009). Pour chaque marqueur sélectionné, une amorce complémentaire à la séquence flanquante du polymorphisme est générée sur la puce. En routine, la molécule d'ADN d'un individu va s'hybrider aux amorces complémentaires correspondantes et une amplification spécifique de la base polymorphe suivante permet de révéler le génotype aux différents marqueurs présents sur la puce. Chez le bar et la daurade, les premières puces ont été développées en 2017 pour le bar et 2018 pour la daurade dans le cadre du projet FEAMP Gènesea (Fond Européen pour les Affaires Maritimes et la Pêche). La puce à SNPs du bar est nommée ThermoFisher Axiom™ DlabChip 57K et le procédé de sélection des marqueurs a été décrit dans Griot et al. (2021). Le design de la puce à SNPs chez la daurade, nommée ThermoFisher Axiom™ SaurChip 60K, n'est quant à lui pas encore publiquement disponible. Ces deux puces, d'environ 57 000 et 60 000 marqueurs SNPs respectivement, ont pour objectif d'être utilisées aussi bien par des entreprises de sélection que par des instituts de recherche. Contrairement au RAD-seq qui nécessite une étape d'identification des marqueurs avant l'étape d'obtention des génotypes, la puce à SNPs permet de produire directement les génotypes, l'identification et la sélection des marqueurs ayant été réalisées lors de la phase de développement, ce qui facilite son utilisation en routine (Robledo et al., 2018).

1.5.3. L'utilisation des marqueurs dans l'estimation des valeurs génétiques : la sélection génomique

L'utilisation de l'information de l'ADN en amélioration génétique a été une révolution dans le domaine. En effet, on ne peut imaginer mieux que l'ADN pour différencier génétiquement des individus entre eux et ainsi mieux comprendre les mécanismes génétiques amenant à des performances différentes. C'est en 2001 que l'utilisation de l'information génomique en sélection a été théorisée sous le terme « sélection génomique ». Dans Meuwissen et al. (2001), les auteurs ont montré que l'utilisation d'un grand nombre de marqueurs répartis tout le long du génome augmente la précision de sélection et donc, le gain génétique par génération. Ils ont montré qu'il était possible de prédire la valeur génétique d'un individu en ne connaissant que son génotype, si les associations génotype-phénotype avaient été documentées chez d'autres individus. Chez toutes les espèces où elle a été mise en place, la sélection génomique a démontré un avantage net par rapport aux méthodes de sélection qui étaient utilisées jusque-là (Lin et al., 2014; Meuwissen et al., 2013; Weller et al., 2017). En aquaculture, sa mise en application a été proposée depuis plus de 10 ans (Sonesson et Meuwissen, 2009a) mais son utilisation par des entreprises privées date de 2016 avec la première commercialisation de saumons atlantiques améliorés par sélection génomique par la société Aquagen. En France, la mise en place de la sélection génomique dans les programmes de sélection aquacole est en cours de développement chez la truite arc-en-ciel, le bar, la daurade et l'huitre creuse grâce à divers projets FEAMP (SG-Truite, Gènesea et Quality Huitre) qui réunissent des entreprises privées, le SYSAAF et des instituts de recherche publique (Ifremer, INRAE, CNRS).

La mise en pratique de la sélection génomique se divise en plusieurs étapes (Taylor, 2014). La première est la création d'une population dite de calibration ou d'entraînement. Cette population sera phénotypée et génotypée et permettra de calibrer le modèle de prédiction génomique. Cette

population doit avoir certaines caractéristiques, comme être proche génétiquement des individus qui seront prédits, et être de taille conséquente. Dans le cas de la sélection pour la résistance aux maladies en aquaculture, la population d'entraînement est constituée des individus qui seront challengés contre l'agent pathogène.

Ensuite, à partir de la population d'entraînement, le modèle de prédiction, comme tout modèle statistique, est calibré. En général, un modèle de prédiction génomique va estimer les effets des marqueurs qui ont été génotypés. La somme de ces effets donne la GEBV (Genomic Estimated Breeding Value) qui est la valeur génétique estimée à partir de l'information génomique. Ce type de modèle est appelé SNP-BLUP lorsque les effets des marqueurs suivent une loi normale (Meuwissen et al., 2001; Whittaker et al., 2000) ou fait partie de la famille des modèles bayésiens (BayesA, BayesB, BayesC, ...) lorsque la distribution des effets des marqueurs suit une autre loi (Habier et al., 2011). Un autre modèle, appelé GBLUP, utilise une autre propriété de l'information génomique. Dans le modèle BLUP (Best Linear Unbiased Prediction), les valeurs génétiques des individus sont estimées grâce aux relations d'apparentement entre les individus. Dans sa variante utilisant l'information contenue dans le pedigree, classiquement utilisée en sélection sur apparentés, les valeurs d'apparentements sont calculées à partir du pedigree et sont donc théoriques, identiques pour tous les animaux d'une même famille. Grâce à l'apport de l'information génomique, il est possible de calculer les valeurs d'apparentements réalisées pour chaque individu, qui sont plus précises que celles calculées à partir du pedigree (VanRaden, 2008), et donc d'obtenir des valeurs génétiques plus précises.

La dernière étape est la prédiction des valeurs génétiques des individus appartenant à la population des candidats à la sélection, qui ne sont que génotypés. Grâce à la calibration du modèle réalisée lors de l'étape précédente, les effets des marqueurs (dans le cas du SNP-BLUP ou des modèles bayésiens) ou des relations d'apparentement réalisées (dans le cas du GBLUP) sont estimés. Puis, les valeurs génétiques de la population des candidats à la sélection sont prédites, connaissant leur génotype, et donc soit en appliquant les effets des marqueurs, soit en calculant les relations d'apparentement réalisées entre la population d'entraînement et celle des candidats à la sélection. A partir des GEBV prédites, les candidats peuvent être classés et ensuite poursuivre le processus de sélection pour devenir de futurs géniteurs.

La mise en place de la sélection génomique dans les schémas de sélection aquacole en France aurait l'avantage, en plus de l'amélioration théorique du gain génétique, de ne nécessiter que peu de modifications techniques et logistiques. En effet, comme le génotypage des candidats et de leurs apparentés est déjà mis en pratique pour l'assignation de parenté dans les schémas de sélection sur apparentés, le génotypage sur une puce à SNPs ne représenterait qu'un changement d'outil de génotypage, et ne constitue donc pas une modification en profondeur de la gestion sur le terrain du schéma de sélection. Il impliquerait néanmoins, au niveau du traitement de l'information, une modification des méthodes de contrôle qualité des génotypages et d'évaluation des valeurs génétiques. Compte tenu du coût supérieur du génotypage de dizaine de milliers de marqueurs SNPs par rapport au coût actuel de génotypage d'une dizaine de marqueurs microsatellites pour assignation de parenté, il faut également évaluer le surcroît de performance que permettrait la sélection génomique dans les programmes de sélection en place.

1.5.4. L'utilisation des marqueurs dans l'étude de l'architecture génétique : l'identification de QTL

Les marqueurs génomiques ont permis de mieux comprendre l'architecture génétique des caractères, en particulier celle des résistances aux maladies (Fraslin et al., 2020b; Hocking, 2005; Rothschild et al., 2007). Alors que la génétique quantitative classique, se basant sur la connaissance du pedigree, permettait d'obtenir une estimation de l'héritabilité d'un caractère ainsi que des corrélations génétiques entre différents caractères, la génomique permet d'aller jusqu'à l'identification de régions génomiques impliquées dans la variation du caractère, appelées QTL (Quantitative Trait Locus). Ces QTLs peuvent donc être dénombrés, positionnés et leurs effets sur le caractère peuvent être quantifiés. Quelle que soit la méthode employée, la détection de QTL repose sur l'association statistique entre la fréquence allélique observée pour un marqueur donné et le phénotype. Un niveau de significativité est donné par une p-value ou un LOD score. Le Logarithm of Odds (LOD) est une mesure du degré de liaison génétique entre deux loci au sein d'une famille. Le LOD est plus précisément le logarithme du rapport de vraisemblance entre un modèle localisant le QTL aux abords d'un marqueur particulier et un autre modèle supposant ce gène non lié. Cette valeur est ensuite comparée à un seuil de significativité choisi, souvent de 5%, pour déclarer la présence ou non d'un QTL.

En pratique, pour que des QTLs soient détectés, il faut qu'ils ségrégent dans la population étudiée. A l'origine, ce sont des populations expérimentales qui ont été utilisées pour détecter des QTLs, en particulier des populations F2, de back-cross ou recombinantes (RIL). Dans ce type de croisement, où les individus sont fortement apparentés, les méthodes d'interval mapping ont été largement employées. Elles réalisent, à intervalle régulier le long du génome, un test statistique visant à déterminer la présence d'un QTL à la position testée (Haley et Knott, 1992). Ensuite, avec l'augmentation du nombre de marqueurs, il a été possible de tirer profit du déséquilibre de liaison au sein de population diversifiées, où les relations d'apparentement sont plus distantes. Ainsi, des méthodes de GWAS (Genome-Wide Association Study) ont été développées (Visscher et al., 2012). En GWAS, les fréquences alléliques pour chacun des marqueurs génotypés au sein d'une population sont mis en relation avec le phénotype d'intérêt. En d'autres termes, on teste si la fréquence allélique explique les différences phénotypiques observées dans la population. Ce test donne une valeur de significativité, exprimée en p-value le plus souvent, pour chaque marqueur génotypé. Il existe plusieurs méthodes pour effectuer ces tests. Le plus couramment utilisé est le modèle GBLUP, capable de tester simultanément tous les marqueurs (Aguilar et al., 2019). Les modèles bayésiens sont également largement utilisés car capables de tester des hypothèses d'architectures génétiques différentes (Legarra et al., 2015). Ces méthodes peuvent également être appliquées à des données de séquence.

En pisciculture, il existe deux exemples majeurs de découverte de QTL ayant un effet fort sur la résistance à une maladie (Fraslin et al., 2020b). Le premier est celui de la découverte d'un QTL impliqué dans la résistance à la NPI (Nécrose Pancréatique Infectieuse) chez le saumon Atlantique qui explique 83% de la variance génétique (Moen et al., 2009). En pratique, les candidats à la sélection sont génotypés au marqueur associé au QTL et, selon leur génotype, peuvent être déclarés comme résistants ou sensibles (Moen et al., 2015). La détection de ce QTL a également permis une approche fonctionnelle, allant jusqu'à la découverte de la mutation dans le gène impliqué dans la résistance (Moen et al., 2015). L'utilisation de ce QTL en sélection a permis de diminuer de 75% le nombre d'épidémies dans la production de saumon (Moen et al., 2015). Le deuxième est celui de la détection d'un QTL expliquant 65% de la variance phénotypique, impliqué dans la résistance à la SHV (Septicémie Hémorragique Virale) chez la truite arc-en-ciel, qui n'est pas utilisé en sélection à ce jour (Verrier et al., 2013).

1.6. Objectifs de la thèse

Cette thèse, inscrite dans le cadre du projet Gènesea, financé par le FEAMP (Fonds Européen pour les Affaires Maritimes et la Pêche) a pour objectif de développer des outils et des méthodes visant à la mise en place de la sélection génomique pour l'amélioration de la résistance aux maladies dans les schémas de sélection chez le bar et la daurade.

Dans ce but, nous avons tout d'abord étudié l'utilisation des données de génotypage, produites par les puces à SNPs de moyennes densités développées dans le cadre du projet, dans l'assignation de parenté. Ensuite, nous avons étudié l'architecture génétique de la résistance à la nodaviriose et à la vibriose chez le bar et à la pasteurellose chez la daurade, dans l'objectif de détecter des QTLs impliqués dans cette résistance. Enfin, nous avons exploré les effets de deux paramètres majeurs influençant la précision de la sélection dans le contexte de la sélection génomique : la densité de marqueurs et la taille de la population d'entraînement.

Pour réaliser ce travail, nous nous sommes principalement appuyés sur trois cohortes de bar et une cohorte de daurade, issues des croisements de sélectionneurs français : la Ferme Marine du Douhet (FMD) et l'Eclosérie Marine de Gravelines-Ichtus (EMG). Dans le cadre de l'étude de l'architecture génétique de la résistance à la nodaviriose chez le bar, une population expérimentale, composée de quatre familles de plein-frères d'origine différente, a également été créée par l'Ifremer. Ces cohortes ont été challengées par différents pathogènes (deux à la nodaviriose, une à la vibriose et une à la pasteurellose) au sein de la plateforme Fortior Genetics, issue de la collaboration entre le SYSAAF et l'ANSES. Les données de génotypage ont été obtenues à partir des puces à SNPs développées dans le cadre du projet Gènesea en collaboration avec le SYSAAF, FMD, EMG, l'Ifremer, INRAE et ThermoFisher. Ces puces, nommées DlabChip 57K pour la puce à SNPs du bar et SaurChip 60K pour celle de la daurade comportent respectivement environ 57 000 et 60 000 marqueurs SNPs. Ces puces ont toutes été traitées par la plateforme Gentyane de l'INRAE de Clermont-Ferrand.

Cette thèse se place dans le cadre du SYSAAF, via le financement d'une bourse CIFRE de l'Association Nationale Recherche Technologie (ANRT) (n° 2017/0731). Les expérimentations et les données utilisées proviennent du programme collaboratif Gènesea (n° R FEA 4700 16 FA 100 0005), financé par le FEAMP (Fonds Européen pour les Affaires Maritimes et la Pêche). Dans ce projet, les sélectionneurs FMD et EMG-Ichtus et le SYSAAF collaborent avec l'Ifremer, INRAE et le CNRS pour avancer vers la mise au point d'outils et méthodes de sélection génomique chez le bar et la daurade



