



**HAL**  
open science

# Estimation du nombre de reproducteurs efficaces : un nouvel indicateur de l'état des populations

Olivier Lepais

► **To cite this version:**

Olivier Lepais. Estimation du nombre de reproducteurs efficaces : un nouvel indicateur de l'état des populations. [Contrat] Inconnu. 2017, 42 p. hal-02187476

**HAL Id: hal-02187476**

**<https://hal.science/hal-02187476>**

Submitted on 17 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Avec le soutien financier de

**AGENCE FRANÇAISE  
POUR LA BIODIVERSITÉ**  
ÉTABLISSEMENT PUBLIC DE L'ÉTAT



## Estimation du nombre de reproducteurs efficaces : un nouvel indicateur de l'état des populations.

### Rapport final

incluant un **Guide méthodologique** d'échantillonnage ainsi qu'un **Protocole** d'analyse génétique à destination des utilisateurs

**Olivier LEPAIS (INRA ECOBIOP, pôle AFB-INRA Gest'Aqua)**

**Décembre 2017**

Document élaboré dans le cadre de :  
la convention AFB-INRA 2016-2018.

En partenariat avec :

INRA U3E

INRA ESE



- **AUTEUR**

**Olivier LEPAIS**, Chargé de recherche (INRA ECOBIOP), olivier.lepais@inra.fr

- **CONTRIBUTEURS**

**Cécile BACLES**, Chercheuse contractuelle (INRA ECOBIOP), cecile.bacles@inra.fr

**Frédéric LANGE**, Technicien (INRA ECOBIOP), frederic.lange@inra.fr

**Aurélié MANICKI**, Assistant ingénieure (INRA ECOBIOP), aurelie.manicki@inra.fr

**Colin BOUCHARD**, Doctorant (INRA ECOBIOP), colin.bouchard@inra.fr

**Cédric TENTELIER**, Maître de conférence (UPPA ECOBIOP), cedric.tentelier@univ-pau.fr

**Etienne PREVOST**, Directeur de recherche (INRA ECOBIOP), etienne.prevost@inra.fr

**Guillaume EVANNO**, Chargé de recherche (INRA ESE), guillaume.evanno@inra.fr

**Lucie MONTORIO**, Doctorante (INRA ESE), lucie.montorio@inra.fr

**Marie NEVOUX**, Chargée de recherche (INRA ESE), marie.nevoux@inra.fr

**Nicolas JEANNOT**, Technicien (INRA U3E), nicols.jeannot@inra.fr

**Frédéric MARCHAND**, Ingénieur d'étude (INRA U3E), frederic.marchand@inra.fr

- **CORRESPONDANTS**

**Partenaire : Olivier LEPAIS**, Chargé de recherche (INRA ECOBIOP, pôle AFB-INRA Gest'Aqua), olivier.lepais@inra.fr

**Agence française pour la biodiversité : Laurent BEAULATON**, Chef du pôle AFB-INRA Gest'Aqua, laurent.beaulaton@afbiodiversite.fr

**Droits d'usage** : accès libre.

**Niveau géographique** : national.

**Couverture géographique** : France ; Nouvelle Aquitaine, Pyrénées Atlantiques, Nouvelle-Aquitaine ; Bretagne, Morbihan et Côtes-d'Armor et Finistère, Scorff ; Normandie, Manche, Oir et La Roche.

**Niveau de lecture** : professionnels et experts.

- **ESTIMATION DU NOMBRE DE REPRODUCTEURS EFFICACES : UN NOUVEL INDICATEUR DE L'ETAT DES POPULATIONS, OLIVIER LEPAIS.**

- **RESUME**

Cette étude évalue si le nombre de reproducteurs efficaces ( $N_b$ ) peut être estimé en utilisant un protocole peu coûteux et rapide basé sur la reconstruction génétique des fratries au sein d'un échantillon unique de juvéniles de l'année dans un contexte de suivi des populations de saumon Atlantique (*Salmo salar*).

La population de saumon de la Nivelle a été utilisée pour valider la méthode, étudier les possibles effets de divers facteurs liés à l'échantillonnage (par des méthodes de sous-échantillonnage et de simulations), et enfin émettre des recommandations précises. Ces recommandations ont été ensuite appliquées à une population de plus grande taille (le Scorff) avec succès permettant de généraliser l'utilité de la méthode dans le cas de grandes populations. Des données disponibles sur La Roche, un affluent de l'Oir, ont permis d'illustrer les limites de l'estimation de  $N_b$  dans le cas d'une population ouverte. Dans ce cas, un grand nombre de juvéniles doivent être échantillonnés pour obtenir une estimation fiable, et l'interprétation de l'estimateur est rendue difficile par les facteurs confondants liés aux mouvements des reproducteurs et des juvéniles entre la zone étudiée et l'ensemble du système.

Enfin,  $N_b$  a été estimé au cours de 10 saisons de reproductions consécutives sur la Nivelle afin d'illustrer les informations fournies par ce paramètre, seul ou en complément d'autres paramètres démographiques disponibles sur cette population.  $N_b$  estimé à l'automne était faible et variable au cours des années, allant de 23 (intervalle de confiance à 95% : 14-41) à 75 (53-101) reproducteurs efficaces. Il ne montrait aucune corrélation avec le nombre d'adultes anadromes estimé, mais était faiblement corrélé positivement avec le nombre de juvéniles de l'année estimé à l'automne. Les facteurs environnementaux tels que la disponibilité et la qualité des zones de frayères et des habitats pour le développement des juvéniles semble prédominer sur les facteurs démographiques (nombre d'adultes anadromes) dans l'évolution de la viabilité à long-terme de la population de saumon de la Nivelle.

Cette étude illustre les conditions d'utilisation et l'utilité de l'estimation de  $N_b$  basée sur la reconstruction génétique des fratries à partir d'un simple échantillonnage de juvéniles de l'année. Ce paramètre intègre des informations sur les facteurs écologiques et démographiques sur un temps court (une saison de reproduction) et peut être estimé sur n'importe quelles populations même celles ne possédant pas d'infrastructures de suivi. Il est pertinent pour évaluer à la fois l'impact à court terme des actions de gestions et le statut conservatoire des populations à plus long terme.

- **MOTS CLES (THEMATIQUE ET GEOGRAPHIQUE)**

Suivi génétique, population de petite taille, Colony, reconstruction de fratries, juvéniles de l'année,  $N_b$ , nombre de reproducteurs efficaces, taille efficace, démographie, saumon Atlantique, *Salmo salar*, ORE DiaFC, Nivelle, Scorff, Oir.

- **ESTIMATION DU NOMBRE DE REPRODUCTEURS EFFICACES : UN NOUVEL INDICATEUR DE L'ETAT DES POPULATIONS, OLIVIER LEPAIS.**
- **SYNTHESE POUR L'ACTION OPERATIONNELLE**

### Contexte général

Le suivi des tendances démographiques au cours du temps est essentiel pour évaluer le statut des populations d'importance conservatoire car la taille des populations a un très fort impact sur leur viabilité. Ceci est particulièrement vrai pour les petites populations pour lesquelles une diminution de taille peut exacerber les risques liés à l'augmentation de la stochasticité démographique, génétique et environnemental. Alors que les principaux corridors de migration bénéficient de ressources financières et humaines pour collecter des données démographiques sur le long terme, de nombreuses petits fleuves ou rivières hébergeant des populations de plus petite taille, potentiellement plus vulnérables, souffrent parfois d'un manque de suivis démographiques, rendant alors impossible l'évaluation de l'évolution de leur état de santé et leur statut conservation.

Parmi les indices traduisant la viabilité des populations sur le long terme, la taille efficace génétique des populations ( $N_e$ ) est un paramètre clé à considérer car il affecte de nombreux processus relatifs à la conservation des espèces. Le paramètre  $N_e$  représente la taille d'une population idéale qui perdrait de la diversité génétique à cause des effets de la dérive et de la consanguinité à la même vitesse que la population étudiée. Dans les populations naturelles,  $N_e$  est généralement inférieur à la taille de la population à cause de facteurs liés aux caractéristiques des systèmes de reproduction, de la variance de la fécondité et de la survie résultant en une variabilité élevée du succès reproducteur individuel. Ainsi, deux populations ayant un même nombre d'individus adultes ( $N_c$ ) auront des tailles efficaces ( $N_e$ ) contrastées si elles diffèrent par le nombre d'individus qui produisent des descendants ou si la variance du succès reproducteur individuel diffère entre elles. En conséquence,  $N_e$  est un indicateur plus robuste que  $N_c$  pour le suivi de populations stables ou en déclin. De plus, le potentiel adaptatif des populations est directement lié à  $N_e$  car cet indice prédit l'efficacité de la sélection, la probabilité de fixation d'allèles délétères et la perte de diversité génétique par le processus de dérive. Par conséquent,  $N_e$  fournit davantage d'information à long terme sur les processus évolutifs à l'œuvre dans les populations d'importance conservatoire, en particulier dans le contexte de suivi de petites populations.

Le manque d'exemple d'utilisation de  $N_e$  et de protocole pour estimer  $N_e$  dans le contexte de gestion des populations naturelles d'espèces à long temps de génération et à génération chevauchantes a probablement limité son utilisation pour le suivi des populations. Cependant, des développements statistiques récents permettant d'utiliser la variabilité de marqueurs moléculaires mesurée à parti d'un seul échantillonnage, par exemple de juvéniles issus d'un événement de reproduction, pour estimer  $N_e$  devrait rendre ces méthodes innovantes plus accessible pour le suivi des populations. Parmi les méthodes récemment développées, celle basée sur l'affectation de fratrie utilise la fréquence des liens d'apparentement au sein d'une cohorte pour estimer  $N_e$ . Cette approche est particulièrement attractive pour le suivi des poissons anadromes car les échantillonnages de juvéniles (truite ou saumon par exemple) qui sont capturés en routine lors des campagnes de pêche à l'électricité dans les populations suivies contiennent des individus apparentés (partageant la même mère et/ou le même père). Cependant, la sensibilité aux effets d'échantillonnages de cette méthode n'a pas été évalué jusqu'à présent. L'objectif principal de cette étude est donc d'explorer la robustesse de la méthode d'estimation de  $N_e$  basée sur la reconstruction des fratries dans le cas d'un protocole d'échantillonnage typique de ce qui est déjà réalisé lors des pêches à l'électricité de juvéniles en conditions réelles de suivi.

Ce projet s'intéresse au saumon Atlantique (*Salmo salar* L. 1758), une espèce de poisson anadrome prisée par la pêche professionnelle et de loisir dont le déclin a débuté dès les années 1970 dans toute son aire de distribution. Les principales populations de saumon Atlantique sont suivies afin d'évaluer leur statut conservatoire et les actions de gestion. Dans ce contexte, le suivi génétique de la taille efficace des populations a été reconnu comme un outil complémentaire pour la gestion. Cependant, l'application en routine de cet outils n'a pas été suivi d'effet, probablement dû au fait que ces méthodes se sont raffinées de façon continue depuis une dizaine d'années et qu'un nombre limité d'études illustre leurs applications pratiques dans un contexte de gestion. Le second objectif

de cette étude est donc de combler le fossé entre la recherche académique et les applications pratiques pour les gestionnaires.

La saumon atlantique est une espèce à générations chevauchantes car les individus peuvent se reproduire potentiellement plusieurs fois à différents âges, et les populations présentent une grande variabilité pour l'âge à la première maturité sexuelle. Des individus de plusieurs cohortes participent donc à la reproduction une saison donnée. De ce fait, l'échantillonnage d'une seule cohorte, par exemple les juvéniles de l'année (juvéniles 0+), permet d'estimer le nombre de reproducteurs efficaces ( $N_b$ ) de cette saison de reproduction ayant produit les juvéniles de l'année, et non la taille efficace de la population ( $N_e$ ) qui est une estimation à l'échelle d'une génération (intégrant le succès reproducteur des individus sur l'ensemble de leur vie et d'autres facteurs). La relation entre  $N_e$  et  $N_b$  est complexe pour les espèces à générations chevauchantes. Même si  $N_e$  fournit des informations à plus long terme sur le potentiel adaptatif des populations, le paramètre  $N_b$  est tout de même pertinent car il permet de prédire les changements génétiques dans une cohorte liés à la contribution reproductive des adultes et au sex ratio et apporte de l'information sur le risque de dépression de consanguinité et la dérive génétique dans un contexte de conservation.

## Objectifs et

### > principaux résultats

L'objectif principal de cette étude est de développer un protocole pour le suivi des petites populations à des fins de gestion et de conservation basé sur l'estimation du nombre efficace de reproducteurs ( $N_b$ ) utilisant la reconstruction des fratries à partir d'un simple échantillonnage de juvéniles de l'année. Afin de garantir une méthode rapide et facile à transférer, les marqueurs génétiques utilisés, de type microsatellites, sont amplifiés grâce à un protocole à faible coût consistant en une seule réaction d'amplification de nombreux marqueurs simultanément, produisant une information génétique suffisamment résolutive pour différencier les individus et reconstruire les fratries.

Dans un premier temps, nous avons déterminé les conditions d'échantillonnage nécessaires à l'obtention d'une estimation fiable et biologiquement pertinente de  $N_b$  pour différents types de configurations pouvant être rencontrées dans différents scénarii de gestion. Pour ce faire, nous avons tiré parti des collections d'échantillons de juvéniles de l'année collectés lors des pêches indice d'abondance d'automne dans le cadre des suivis de l'Observatoire de Recherche en Environnement Diadrome dans les Fleuves Côtiers (ORE DiaFC). Les différentes caractéristiques des fleuves suivis (Nivelle, Oir et Scorff), ainsi que la taille variée des populations de saumon qu'elles hébergent et la disponibilité d'échantillons d'adultes pour certaines, nous ont permis d'identifier les conditions favorables pour une estimation fiable de  $N_b$  et de renseigner les limites de la méthode, grâce à des analyses de sous-échantillonnages et de simulations.

> Nous avons validé le principe qu'un seul échantillon de juvéniles de l'année permettait d'obtenir une estimation fiable de  $N_b$ , sans qu'il soit nécessaire d'échantillonner les parents potentiels. L'information génétique des juvéniles est suffisante pour reconstituer les fratries et estimer  $N_b$  et la méthode peut donc être déployée aisément sur n'importe quelle population. Le nombre de juvéniles de l'année à échantillonner pour obtenir des estimations fiables de  $N_b$  dépend de la taille de la population et se situe autour du nombre d'adultes anadromes remontant dans la population pour s'y reproduire.

Dans un second temps, nous avons étudié l'influence de l'étendue spatiale de la zone échantillonnée et de la saison d'échantillonnage sur l'estimation de  $N_b$  afin de mettre en évidence les éventuels biais d'échantillonnage.

> La stratégie d'échantillonnage a une grande influence sur l'estimation de  $N_b$ . Il est primordial que l'échantillonnage par pêche à l'électricité se fasse de façon standardisée avec le même effort de capture sur l'ensemble des sites échantillonnés. D'autre part, différentes estimations de  $N_b$  sur une même population au cours du temps, ou entre populations peuvent être directement comparées si la stratégie d'échantillonnage est identique (effort de capture, saison d'échantillonnage). Enfin, seules les estimations de  $N_b$  obtenues à partir d'échantillonnages couvrant l'ensemble de la population étudiée peuvent être interprétées en termes de fonctionnement de la population. Les études focalisées sur un affluent ou une portion d'un plus grand bassin versant demandent une taille d'échantillons de juvéniles plus importante et  $N_b$  estimé dans ces conditions est

difficilement interprétable du fait des facteurs confondants liés aux mouvements des individus entre la zone étudiée et le reste du système.

En plus de préciser les recommandations en terme de stratégie d'échantillonnage, cet exercice réalisé sur la Nivelle pour laquelle nous disposons d'échantillonnages de juvéniles au printemps et à l'automne collectées selon deux stratégies différentes pour deux années consécutives nous a permis de mieux comprendre l'information apporté par  $N_b$  au-delà du seul nombre d'adultes ayant produit des descendants.

>  $N_b$  est un paramètre intégratif qui inclut de l'information sur la reproduction (structure du réseau d'apparement entre individus, variabilité du succès reproducteur) mais également sur les facteurs affectant les premiers mois de la vie des juvéniles jusqu'à leur échantillonnage (survie, dispersion...). Un échantillonnage plus ou moins précoce dans la saison apportera donc différentes informations sur la population (par exemple, plutôt sur la reproduction si l'échantillonnage est réalisé au printemps).

Dans un troisième temps, nous avons estimé  $N_b$  au cours de dix années consécutives sur la Nivelle et confronté cet indice à différents paramètres démographiques estimés sur la même période afin d'illustrer l'information apportée par  $N_b$ , soit comme seul outil disponible dans le cas de population non suivi, ou en combinaison avec d'autres indicateurs démographiques ou environnementaux dans le cas de populations faisant l'objet d'un suivi.

> Lorsque considéré seul,  $N_b$  peut servir de paramètre permettant de suivre l'évolution de la santé de la population au cours du temps. Dans le cas d'une population suivie et possédant d'autres information démographiques,  $N_b$  apporte des informations complémentaires pouvant servir à identifier les facteurs démographiques ou environnementaux qui conditionnent la viabilité de la population et son adaptabilité à long terme.

Enfin, les développements techniques et méthodologiques qui ont permis d'obtenir les résultats lors de ce projet sont présentés en détail en annexes afin de pouvoir être reproduits dans un contexte de suivi de population ou d'évaluation d'actions de gestion. A cette fin, un Guide méthodologique pour l'échantillonnage résume les enseignements de ce projet en émettant des recommandations précises en matière d'échantillonnage de juvéniles à des fins d'estimation de  $N_b$  dans les populations (Annexe 1). Un Protocole pour l'analyse génétique décrit également de façon détaillée les différentes étapes nécessaires au laboratoire pour le génotypage rapide et à faible coût des juvéniles de saumon, ainsi que l'analyse des données permettant la reconstitution des fratries et l'estimation de  $N_b$  grâce au logiciel Colony (Annexe 2). L'objectif de ces livrables est d'apporter une information précise pour la mise en œuvre des outils développés et testés lors de ce projet, dans le but d'en faciliter l'appropriation par les utilisateurs potentiels (gestionnaires, bureaux d'études, organismes de recherche appliquée).

#### **Pour en savoir plus :**

Bacles CFE, Bouchard C, Lange F, Manicki A, Tentelier C, Lepais O. (2018) Estimating the effective number of breeders from single parr samples for conservation monitoring of wild populations of Atlantic salmon *Salmo salar*. Journal of Fish Biology, sous presse.

Lepais O. Estimation du nombre de reproducteurs efficaces : un nouvel indicateur de l'état des populations. Rapport final incluant un Guide méthodologique d'échantillonnage ainsi qu'un Protocole d'analyse génétique à destination des utilisateurs. AFB INRA, 2017, accessible en ligne.

Pour tout contact : olivier.lepais@inra.fr

## Table des matières

<b>1. Introduction</b> .....	<b>8</b>
1.1. Contexte .....	8
1.2. Objectifs .....	10
<b>2. Matériels et méthodes</b> .....	<b>11</b>
2.1. Echantillonnages et questions associées .....	11
2.1.1. Nivelles .....	11
2.1.2. Le Scorff .....	13
2.1.3. La Roche .....	13
2.2. Méthodes moléculaires .....	13
2.3. Analyse des données .....	14
2.3.1. Evaluation de la qualité des données génotypiques .....	14
2.3.2. Evaluation de l'estimation de $N_b$ par reconstitution de fratries par Colony .....	14
2.3.2.1. Reconstruction des fratries et estimation de $N_b$ par Colony .....	14
2.3.2.2. Evaluation de l'effet du nombre de juvéniles analysés par sous-échantillonnage et simulation .....	15
2.3.2.3. Evaluation de l'effet de la stratégie d'échantillonnage .....	16
2.3.3. Données démographiques confrontées aux estimations de $N_b$ .....	16
<b>3. Résultats &amp; Discussions</b> .....	<b>16</b>
3.1. Caractéristique des données générées par le protocole de génotypage de microsatellites .....	16
3.2. Variabilité génétique des saumons de la population de la Nivelles .....	16
3.3. Facteurs affectant l'estimation de $N_b$ .....	17
3.3.1. Validation de l'estimation de $N_b$ à partir des seuls juvéniles .....	17
3.3.2. Taille d'échantillonnage de juvéniles de l'année optimale .....	18
3.3.3. Effet de l'étendue spatiale de l'échantillonnage .....	20
3.3.4. Effet de la saison d'échantillonnage .....	21
3.4. Application des recommandations dans le cas d'autres populations .....	22
3.4.1. Scorff : généralisation des recommandations à une population de plus grande taille .....	22
3.4.2. La Roche affluent de l'Oir : application à une portion de population .....	23
3.5. Apport du suivi de $N_b$ au cours du temps .....	24
<b>4. Conclusion</b> : .....	<b>27</b>
<b>7. Bibliographie</b> .....	<b>29</b>
<b>8. Table des illustrations</b> .....	<b>32</b>
<b>9. Annexe 1. Guide méthodologique : Recommandation pour l'échantillonnage dans le but d'estimer le nombre de reproducteurs efficaces chez les poissons à partir des pêches à l'électricité de juvéniles de l'année.</b> .....	<b>33</b>
<b>10. Annexe 2. Protocole : Procédure expérimentale pour l'analyse génétique nécessaire à l'estimation du nombre de reproducteurs efficaces dans les populations de poissons.</b> .....	<b>36</b>
<b>11. Remerciements</b> .....	<b>41</b>



# 1. Introduction

## 1.1. Contexte

Le suivi des tendances démographiques au cours du temps est essentiel pour évaluer le statut des populations d'importance conservatoire car la taille des populations a un très fort impact sur leur viabilité (Reed, 2008). Ceci est particulièrement vrai pour les petites populations pour lesquelles une diminution de taille peut exacerber les risques liés à l'augmentation de la stochasticité démographique, génétique et environnemental (Lande, 1993). Les petites populations peuvent jouer un rôle vital en connectant de plus grandes populations et en maintenant la continuité de l'aire de distribution mais sont également difficiles à suivre à cause des ressources techniques et financières finies, qui en période de restriction peuvent conduire à prioriser le suivi des plus grandes populations (Simberloff et al., 1992). Par exemple, les suivis démographiques des populations de poissons anadromes consistent généralement à compter le nombre d'individus capturés à plusieurs occasions, lors de leur migration de montaison ou de dévalaison, afin d'estimer le nombre d'individus présents dans une population. Ces suivis nécessitent de lourdes infrastructures telles que des passes à poisson équipées de pièges ou de cameras. Alors que les ressources tant financières que humaines sont déployées pour équiper les principaux corridors de migration et collecter des données démographiques sur le long terme, de nombreux petits fleuves ou rivières hébergeant des populations de plus petite taille, potentiellement plus vulnérables, peuvent souffrir d'un manque de suivi démographique nécessaire à l'évaluation de l'évolution de leur état de santé et de leur statut conservation.

Parmi les indices traduisant la viabilité des populations sur le long terme, la taille efficace génétique des populations ( $N_e$ ) est un paramètre clé à considérer car il affecte de nombreux processus relatifs à la conservation des espèces (Waples, 2002). Le paramètre  $N_e$  représente la taille d'une population idéale selon le modèle de Wright-Fisher (reproduction aléatoire entre individus, générations non-chevauchantes, sexe ratio équilibré, variation aléatoire du succès reproducteur (Wright, 1931)) qui perdrait de la diversité génétique à cause des effets de la dérive et de la consanguinité à la même vitesse que la population étudiée. Dans les populations naturelles,  $N_e$  est généralement inférieur à la taille de la population (exprimée par le nombre d'individus adultes  $N_c$ ) à cause de facteurs liés aux caractéristiques des systèmes de reproduction, de la variance de la fécondité et de la survie résultant en une variabilité élevée du succès reproducteur individuel (Frankham, 1995; Palstra & Fraser, 2012). Ainsi, deux populations ayant un même nombre d'individus adultes ( $N_c$ ) auront des tailles efficaces ( $N_e$ ) contrastées si elles diffèrent par le nombre d'individus qui produisent des descendants ou si la variance du succès reproducteur individuel diffère entre elles (Figure 1). En conséquence,  $N_e$  est un indicateur plus robuste que  $N_c$  pour le suivi de populations stables ou en déclin (Tallmon et al., 2010). La comparaison de  $N_c$  et  $N_e$  apporte également des informations sur les processus écologiques pertinents dans le contexte de la conservation (Luikart et al., 2010). De plus, le potentiel adaptatif des populations est directement lié à  $N_e$  car cet indice prédit l'efficacité de la sélection, la probabilité de fixation d'allèles délétères et la perte de diversité génétique par le processus de dérive. Par conséquent,  $N_e$  fournit davantage d'informations à long terme sur les processus évolutifs à l'œuvre dans les populations d'importance conservatoire, en particulier dans le contexte de suivis de petites populations.

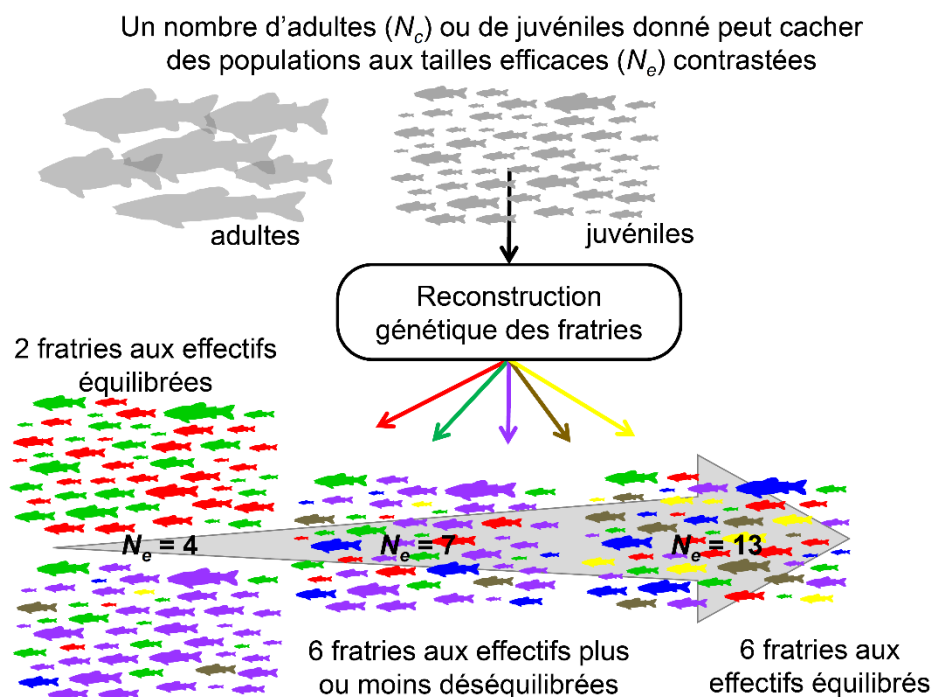


Figure 1. La relation entre le nombre d'adultes, l'effectif de juvéniles et la taille efficace d'une population qui prédit son potentiel adaptatif passe par l'étude de la reproduction. Chaque couleur représente une famille de pleins-frères.

L'absence d'exemple d'utilisation de  $N_e$  et de protocole pour estimer  $N_e$  dans le contexte de gestion des populations naturelles d'espèces à long temps de génération et à génération chevauchantes a probablement limité son utilisation pour le suivi des populations. Cependant, des développements statistiques récents permettant d'utiliser la variabilité de marqueurs moléculaires mesurée à parti d'un seul échantillonnage (Wang et al., 2016), par exemple de juvéniles issus d'un événement de reproduction, pour estimer  $N_e$  devrait rendre ces méthodes innovantes plus accessible pour le suivi des populations. Parmi les méthodes récemment développées, celle basée sur l'affectation de fratrie (*sibship assignment* en anglais, Figure 1) utilise la fréquence des liens d'apparentement au sein d'une cohorte pour estimer  $N_e$  (Wang, 2009, 2016). Cette approche est particulièrement attractive pour le suivi des poissons anadromes car les juvéniles (truite ou saumon par exemple) qui sont capturés en routine lors des campagnes de pêche à l'électricité dans les populations suivies contiennent des individus apparentés (partageant la même mère et/ou le même père, Figure 1). D'autre part, cette méthode est plus robuste que d'autres car elle fait moins d'hypothèses, prend explicitement en compte la reproduction non aléatoire entre individus et n'est pas influencée par la structure des populations, les flux de gènes ou l'erreur de génotypage (Wang, 2016). Néanmoins, elle fait l'hypothèse que la population est échantillonnée de façon aléatoire, et un protocole d'échantillonnage qui aurait pour effet de diminuer ou d'augmenter la proportion d'individus apparentés dans l'échantillon pourrait se traduire par des biais d'estimation de  $N_e$ . Cependant, la sensibilité aux biais d'échantillonnages de cette méthode n'a pas été évaluée jusqu'à présent (Waples, 2016). **L'objectif principal de cette étude est donc d'explorer le comportement de la méthode d'estimation de  $N_e$  basée sur l'affectation des fratries dans le cas d'un protocole d'échantillonnage typique de ce qui est déjà réalisé lors des pêches à l'électricité de juvéniles en conditions réelles de suivi.**

Ce projet s'intéresse au saumon Atlantique (*Salmo salar* L. 1758), une espèce de poisson anadrome prisée par la pêche professionnelle et de loisir dont le déclin a débuté dès les années 1970 dans toute son aire de distribution (Parrish et al., 1998; Klemetsen et al., 2003). Les populations localisées au sud de l'aire de distribution sont particulièrement menacées d'extirpation (Parrish et al., 1998). Les principales populations de saumon Atlantique sont suivies afin d'évaluer leur statut conservatoire (Chaput, 2012) et les actions de gestion (par exemple, Richard et al., 2013; Kennedy et al., 2014; Vøllestad et al., 2014). Dans ce contexte, le suivi génétique de

la taille efficace des populations a été reconnu comme un outils complémentaire pour la gestion (Dodson et al., 1998). Cependant, l'application en routine de cet outils n'a pas été suivi d'effet, probablement dû au fait que ces méthodes se sont raffinées de façon continue depuis une dizaine d'années (Wang et al., 2016) et qu'un nombre limité d'études illustrent leurs applications pratiques dans un contexte de gestion. On trouve cependant des exemples très récents dans un nombre croissant d'articles scientifiques montrant la maturité de ces méthodes et leur potentiel d'application au-delà de la sphère de la recherche académique (Ozerov et al., 2015; Whiteley et al., 2015; Ferchaud et al., 2016; Perrier et al., 2016, par exemple). **Le second objectif de cette étude est donc de combler le fossé entre la recherche académique et les applications pratiques pour les gestionnaires.**

La saumon atlantique est une espèce à générations chevauchantes car les individus peuvent se reproduire plusieurs fois, potentiellement à différents âges, et les populations présentent une grande variabilité pour l'âge à la première maturité sexuelle (Klemetsen et al., 2003). Des individus de plusieurs cohortes participent donc à la reproduction une saison donnée. De ce fait, l'échantillonnage d'une seule cohorte, par exemple les juvéniles de l'année (juvéniles 0+), permet d'estimer le nombre de reproducteurs efficaces ( $N_b$ ) de cette saison de reproduction ayant produit les juvéniles de l'année, et non la taille efficace de la population ( $N_e$ ) qui est une estimation à l'échelle d'une génération (intégrant le succès reproducteur des individus sur l'ensemble de leur vie et d'autres facteurs (Waples, 2005; Wang, 2016)). La relation entre  $N_e$  et  $N_b$  est complexe pour les espèces à générations chevauchantes. Même si  $N_e$  fournit des informations à plus long terme sur le potentiel adaptatif des populations, le paramètre  $N_b$  est tout de même pertinent car il permet de prédire les changements génétiques dans une cohorte liés à la contribution reproductive des adultes et au sexe ratio (Wang, 2009) et apporte de l'information sur le risque de dépression de consanguinité et la dérive génétique dans un contexte de conservation (Ozerov et al., 2015; Whiteley et al., 2015; Ackerman et al., 2016).

## 1.2. Objectifs

L'objectif principal de cette étude est de développer un protocole pour le suivi des petites populations à des fins de gestion et de conservation basé sur l'estimation du nombre efficace de reproducteurs ( $N_b$ ) utilisant la reconstruction des fratries (Wang, 2009) à partir d'un échantillonnage simple de juvéniles de l'année. Afin de garantir une méthode rapide et facile à transférer, les marqueurs génétiques utilisés de type microsatellites sont amplifiés grâce à un protocole à faible coût consistant en une seule réaction d'amplification de nombreux marqueurs simultanément, produisant une information génétique suffisamment résolutive pour différencier les individus et reconstruire les fratries.

Dans un premier temps, nous avons déterminé les conditions d'échantillonnages nécessaires à l'obtention d'une estimation fiable et biologiquement pertinente de  $N_b$  pour différents types de configurations pouvant être rencontrées dans différents scénarios de gestion (l'ensemble d'une petite population, un ruisseau au sein d'un plus grand système, l'ensemble d'une grande population). Pour ce faire, nous avons tiré parti des collections d'échantillons de juvéniles de l'année capturés lors des pêches indice d'abondance d'automne (Lange et al., 2015, 2017; Marchand et al., 2017) dans le cadre des suivis de l'Observatoire de Recherche en Environnement Diadrome dans les Fleuves Côtiers (ORE DiaFC). Les différentes caractéristiques des fleuves suivies (Nivelle, Oir et Scorff) ainsi que la taille variée des populations de saumon qu'elles hébergent et la disponibilité d'échantillons d'adultes pour certaines, nous ont permis d'identifier les conditions favorables pour une estimation fiable de  $N_b$  ainsi que les limites de la méthode, grâce à des analyses de sous-échantillonnages et de simulations.

Dans un second temps, nous avons comparé l'influence de l'étendue spatiale de la zone échantillonnée ainsi que la saison d'échantillonnage sur l'estimation de  $N_b$  afin de mettre en évidence les biais d'échantillonnage pouvant influencer l'estimation de  $N_b$ . Au-delà de préciser les recommandations en terme de stratégie d'échantillonnage, cet exercice réalisé sur la Nivelle pour laquelle nous disposons d'échantillonnages de juvéniles au printemps et à l'automne, collectés selon deux stratégies différentes pour deux années consécutives nous a permis de mieux comprendre l'information apportée par  $N_b$  au-delà du seul nombre d'adultes ayant produit des descendants.

Dans un troisième temps, nous avons estimé  $N_b$  au cours de dix années consécutives sur la Nivelle et confronté cet indice à différents paramètres démographiques estimés sur la même période afin d'illustrer l'information apportée par  $N_b$  et ses limites, soit comme seul outils disponible dans le cas de population non suivi, ou en combinaison avec d'autres indicateurs démographiques ou environnementaux dans le cas de populations faisant l'objet d'un suivi.

Enfin, les développements techniques et méthologiques qui ont permis d'obtenir les résultats présentés dans ce rapport sont présentés en détail en annexes afin de pouvoir être reproduits dans un contexte de suivi de population ou d'évaluation d'actions de gestion. A cette fin, un Guide méthodologique pour l'échantillonnage résume les enseignements de ce projet en émettant des recommandations précises en matière d'échantillonnage de juvéniles à des fins d'estimation de  $N_b$  dans les populations (Annexe 1). Un Protocole pour l'analyse génétique décrit également de façon très détaillée les différentes étapes nécessaires au laboratoire pour le génotypage rapide et à faible coût des juvéniles de saumon, ainsi que l'analyse des données permettant la reconstitution des fratries et l'estimation de  $N_b$  grâce au logiciel Colony (Annexe 2). L'objectif de ces livrables est d'apporter une information précise pour la mise en œuvre des outils développés et testés lors de ce projet, dans le but d'en faciliter l'appropriation par les utilisateurs potentiels (gestionnaires, bureaux d'études, organismes de recherche appliquée).

## 2. Matériels et méthodes

Le but de cette section est de donner un aperçu général des méthodes utilisées afin de pouvoir comprendre les résultats présentés dans ce rapport mais ne cherche pas à atteindre le niveau de détail nécessaire à assurer la reproductibilité. A cette fin, deux annexes sont jointes au rapport : un Guide méthodologique destiné aux gestionnaires (Annexe 1) résume les principales recommandations pratiques en termes d'échantillonnage nécessaires pour l'application de la méthode et un Protocole pour l'analyse génétique (Annexe 2) décrit chacune des étapes avec un niveau de détail nécessaire à reproduire les méthodes d'analyses génétiques développées et optimisées lors de ce projet.

### 2.1. Echantillonnages et questions associées

Ce projet s'est focalisé sur l'étude de collections de tissus biologiques échantillonnées dans le cadre de suivi long-terme des trois populations de saumon Atlantique de l'Observatoire de Recherche en Environnement Diadrome des Fleuves Côtiers (ORE DiaFC; Marchand et al., 2017). Chacune de ces trois populations présente en effet des caractéristiques distinctes en terme de taille de population et de stratégie d'échantillonnage, ce qui permet de tester les méthodes développées dans différents contextes et d'illustrer leurs limites. Les échantillons utilisés et les questions auxquelles ils permettent de répondre sont détaillés par population.

#### 2.1.1. Nivelle

La population de saumon de la Nivelle présente une taille réduite (une centaine d'adultes anadromes remontant chaque année) et se répartie sur un linéaire d'une trentaine de kilomètres accessible par l'espèce (Figure 2). Des pêches d'indice d'abondance (pêches aussi appelés 5 minutes; Prévost & Baglinière, 1995) sont effectuées sur une vingtaine de stations à l'automne chaque année lors desquelles les juvéniles capturés font l'objet d'un prélèvement d'un morceau de nageoire sous anesthésie depuis 2004 (Lange et al., 2017). En parallèle, deux passes à poissons équipées de pièges (Uxondoa et Olha) permettent le suivi des saumons anadromes remontant dans la Nivelle (Figure 2). Ces individus font également l'objet d'un prélèvement d'un morceau de nageoire sous anesthésie ainsi que d'un sexage sur la base de leur morphologie.

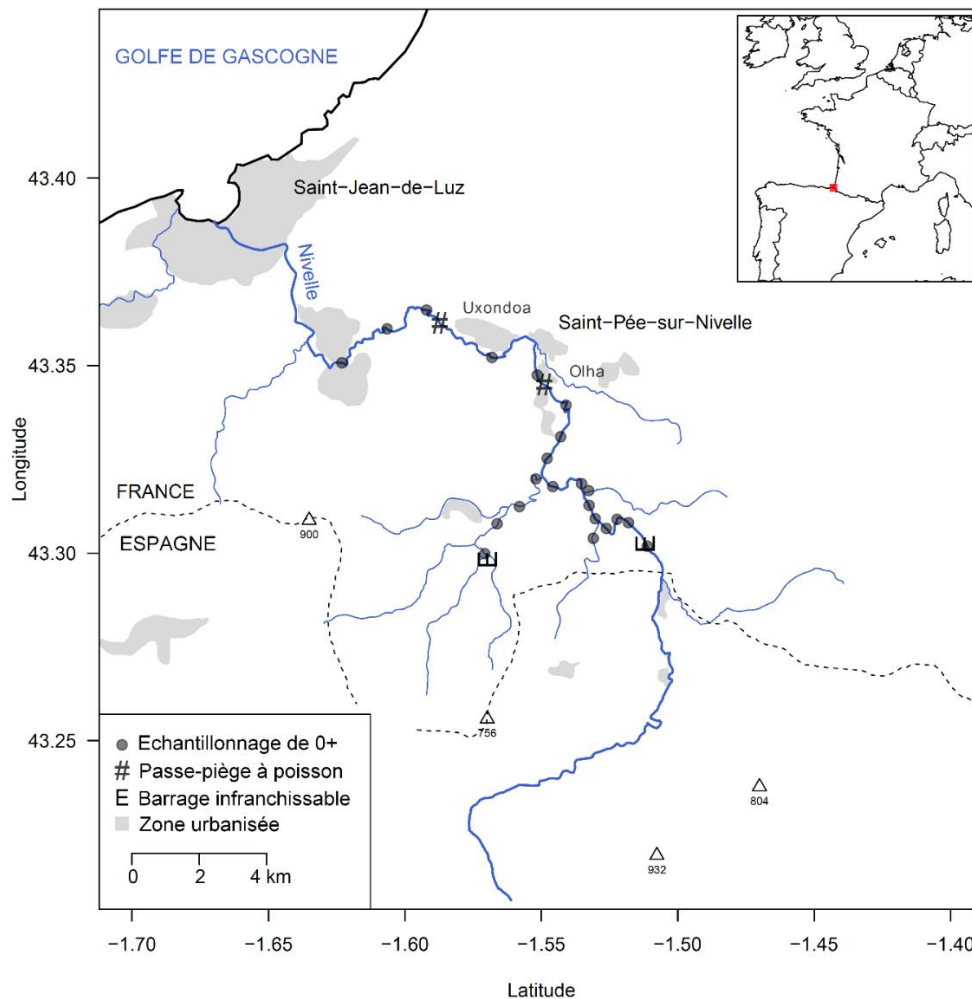


Figure 2. Carte de la Nivelle représentant la localisation des passe-pièges et des points d'échantillonnage de juvéniles (adapté de Bacles et al., 2018).

> **Validation** : Nous avons donc utilisé cette population comme test afin de déterminer si l'analyse des juvéniles seuls donnait des résultats fiables en comparant les résultats obtenus avec ou sans leurs parents potentiels (Tableau 1). En effet, la connaissance des données génétiques des parents potentiels apporte beaucoup d'informations en plus, mais est rarement disponible pour la plupart des populations. Dans un tout premier temps, nous avons donc voulu valider si les seules données génétiques des juvéniles donnaient des informations fiables sur le nombre de reproducteurs efficaces  $N_b$ .

Par le passé, une étude approfondie du comportement de reproduction des saumons a été réalisée lors de deux saisons de reproduction consécutives (2009-2011) pour laquelle un échantillonnage très intense des juvéniles a été effectué au printemps à proximité de l'ensemble des zones de frayères (Tentelier et al., 2016a, 2016b).

> **Recommandations** : La disponibilité de ces données nous a permis de comparer les estimations de  $N_b$  en fonction des stratégies d'échantillonnages (nombreux juvéniles échantillonnés de façon exhaustive ou échantillonnage lâche peu dense (Tableau 2); ensemble de la rivière ou zone de reproduction (Tableau 3), saison d'échantillonnage (Tableau 4)) et de la saison d'échantillonnage (automne ou printemps). Ces comparaisons sur la même rivière et deux années consécutives nous ont permis d'estimer la robustesse de la méthode et de tirer des recommandations visant à limiter les biais inhérents à la stratégie d'échantillonnage.

La collection de tissu issue des pêche « indice d'abondance » des juvéniles (Lange et al., 2017) a été utilisée afin d'effectuer un suivi *a posteriori* de  $N_b$  au cours de dix saisons de reproduction (2004-20014) et de mettre en

relation cet indicateur avec d'autres paramètres démographiques suivis les mêmes années sur cette population (Lange et al., 2015).

> Illustration : L'estimation du nombre de reproducteurs efficaces au cours de 10 années sur une même population permet d'illustrer l'information apportée par cet indicateur seul, dans le cadre d'une population ne faisant l'objet d'aucun suivi, ou en complément avec des paramètres démographiques estimés sur cette population suivie.

### 2.1.2. Le Scorff

Le nombre de saumons anadromes remontant sur Le Scorff est environ dix fois plus important que sur La Nivelle, pour un bassin versant également beaucoup plus étendu (Marchand et al., 2017). Des pêches de juvéniles de type « indice d'abondance » (Prévost & Baglinière, 1995) sont effectuées sur plusieurs dizaine de station chaque année à l'automne au cours desquels des tissus sont prélevés depuis un petite dizaine d'année.

> Recommandations : Cette population de taille nettement plus importante que la Nivelle est un bon site d'étude pour tester l'efficacité de la méthode dans le cas d'une population de grande taille et généraliser les recommandations en matière d'échantillonnage. Ainsi nous avons sous-sélectionné pour génotypage des juvéniles de l'année échantillonnés lors de pêches Indice d'Abondance automnale lors de 3 années présentant des nombres contrastées d'adultes anadromes (Tableau 5) afin d'étudier la sensibilité de l'estimation de  $N_b$  à la taille de la population.

### 2.1.3. La Roche

La Roche est un affluent de l'Oir, lui-même affluent de la Sélune, qui a fait l'objet d'une étude sur le succès reproducteur de différentes stratégies de reproductions chez le saumon (Montorio L, Le cor F, Nevoux M, Evanno G, non publié ; Le cor et al., 2016). Dans ce cadre, les adultes anadromes capturés à diverses occasions lors de leur remontée sur l'Oir ou pêchés à l'électricité dans La Roche, ainsi que plusieurs centaines de juvéniles de l'année échantillonnés sur La Roche, ont été génotypés. L'échantillonnage diffère des pêches de type indice d'abondance car il est plus exhaustif sur l'ensemble de la section de 2 km étudié. Ce système diffère de la Nivelle car, pour un nombre d'adultes anadromes similaire, La Roche représente un linéaire de taille bin inférieure (2.5 km environ contre 35 km pour la Nivelle) et une largeur également beaucoup plus petite. La densité de saumon, qu'il s'agisse des adultes anadromes lors de la reproduction ou des juvéniles produits, est donc beaucoup plus importante sur La Roche. D'autre part, alors que la Nivelle est une population fermée (les juvéniles de l'année sont obligatoirement issus de la population étudiée), le ruisseau de la Roche est un système ouvert où les juvéniles de l'année échantillonnés en automne peuvent être originaires d'une autre partie de l'Oir ou des affluents.

> Recommandations : Il peut être intéressant de suivre l'état d'une population sur un affluent particulier d'un plus grand système, ou sur une échelle spatiale restreinte où la densité de saumon est élevée, au lieu de suivre l'ensemble d'une population de façon plus globale. Ici, la comparaison entre La Roche et la Nivelle va permettre d'étudier l'effet d'un système ouvert et d'une population dense sur l'estimation de  $N_b$ .

## 2.2. Méthodes moléculaires

Le but de l'analyse génétique est de caractériser la variabilité génétique de chaque individu pour un nombre de marqueurs moléculaires suffisants afin d'identifier les individus appartenant à la même fratrie (partageant les deux mêmes parents, on parle de plein-frères ou un seul parent, on parle alors de demi-frères). Le nombre de marqueurs génétiques à analyser dépend du contenu en information des marqueurs, on parle de niveau de polymorphisme. Plus un marqueur génétique présentera de variations (allèles) au sein d'une population, plus il sera puissant pour estimer les relations entre individus. Dans cette étude, nous avons utilisé des marqueurs microsatellites qui ont la particularité d'être très polymorphes (nombreux allèles présents dans les populations) de sorte qu'une dizaine de microsatellites est suffisant pour estimer de façon fiable les relations de fratrie entre individus dans le cas du saumon Atlantique. Nous avons développé un protocole permettant d'analyser

conjointement 15 marqueurs microsatellites en une seule expérience d'amplification et de génotypage réduisant fortement le temps d'analyse ainsi que les coûts (voir Annexe 2 : Protocole d'analyse génétique). L'étape préalable d'extraction d'ADN à partir de morceaux de nageoires ou d'écaillés (dans le cas des adultes) a été réalisé à partir d'un protocole à base de NaCl en format plaque 96 (Gauthey et al., 2015) afin de réaliser rapidement des extraction d'ADN en grande série et à moindre coût (voir Annexe 2 : Protocole d'analyse génétique pour plus de détails).

## 2.3. Analyse des données

### 2.3.1. Evaluation de la qualité des données génotypiques

Afin de valider la qualité des données générées par le protocole de génotypage développé, nous avons analysé un certain nombre d'individus en double pour comparer les génotypes obtenus et estimer le taux d'erreur de génotypage grâce au logiciel Pedant (Johnson & Haydon, 2007).

Nous avons également vérifié que les marqueurs pouvaient être considéré comme indépendants les uns des autres et se comportaient comme attendu selon les hypothèses de Hardy Weinberg (Bacles et al., 2018), prérequis nécessaires pour être utilisés dans la reconstruction de fratrie.

Enfin, le logiciel RClone (Bailleul et al., 2016) a été utilisé pour identifier les possibles individus capturés plusieurs fois. En effet, ces recaptures génèrent des génotypes identiques qui pourraient fausser la reconstruction des fratries.

### 2.3.2. Evaluation de l'estimation de $N_b$ par reconstitution de fratries par Colony

#### 2.3.2.1. Reconstruction des fratries et estimation de $N_b$ par Colony

Le logiciel Colony (Jones & Wang, 2010) a été utilisé pour effectuer la reconstruction des fratries sur laquelle se base l'estimation du nombre de reproducteurs efficaces ( $N_b$ ). Nous avons utilisé la méthode Full-Likelihood (FL) et paramétré le logiciel en considérant que l'espèce est dioïque et diploïde, que les femelles et les mâles sont polygames, qu'il n'y a pas de consanguinité, que nous n'avons pas d'information a priori sur la structure des fratries et qu'il ne faut pas mettre à jour les fréquences alléliques durant l'analyse (pour une reproduction facilitée d'une telle analyse, voir les termes exacts en anglais pour le paramétrage de Colony dans l'annexe 2 : Protocole d'analyse génétique). Nous avons inclu la possibilité d'erreurs de génotypages, en paramétrant un *a priori* de 1% d'erreur à chaque locus pour les deux types d'erreurs modélisées par le logiciel. A partir de ce taux d'erreur *a priori*, Colony estime le taux d'erreur à chaque locus conjointement aux autres paramètres du modèle (Wang, 2017), de sorte que le taux d'erreur a priori n'a pas d'effet sur les résultats pour peu que les données génétiques soient suffisamment informatives et que le taux d'erreur *a priori* ne soit pas trop élevé (ce qui est notre cas). La longueur des analyses a été paramétrée de telle façon à avoir des résultats répétables. Typiquement, une série de trois analyses paramétrées pour une durée moyenne sont lancées et si les résultats concordent, ils sont validés. S'ils diffèrent, trois analyses additionnelles paramétrées sur un temps d'analyse long sont lancées, et ainsi de suite jusqu'à l'obtention de résultats répétables indiquant la convergence des algorithmes. Généralement, des analyses de longueurs moyennes ont été suffisantes pour obtenir une bonne convergence lorsque les juvéniles seuls étaient utilisés pour la reconstruction des fratries, alors que des analyses longues ont été nécessaires lorsque des parents potentiels étaient ajoutés aux analyses. Pour les analyses très nombreuses faisant intervenir des ré-échantillonnages des données (voir ci-dessous), seuls des analyses courtes, mais répétées 10 fois, ont été utilisées afin de rendre le temps de calcul raisonnable.

A partir de la reconstitution des fratries, Colony estime  $N_b$  en utilisant la formule suivante (Wang, 2009):

$$\frac{1}{N_b} = \frac{1 + 3\alpha}{4} (Q_1 + Q_2 + 2Q_3) - \frac{\alpha}{2} \left( \frac{1}{N_1} + \frac{1}{N_2} \right)$$

avec,  $Q_1$ ,  $Q_2$  et  $Q_3$  la probabilité pour une paire de juvéniles de l'année d'être demi-frère paternel, demi-frère maternel et plein frère, respectivement,  $\alpha$  l'écart à la reproduction aléatoire (considérée comme nul ici), et  $N_1$  et  $N_2$  le nombre de parents mâles et femelles présents lors de la reproduction. Comme les populations étudiées



sont conformes aux attendus d'Hardy Weinberg, nous avons considéré que la reproduction entre individus a lieu de façon aléatoire à l'échelle des populations étudiées ( $\alpha = 0$ ). Ceci simplifie la représentation conceptuelle de l'estimation de  $N_b$  qui ne dépend plus que des probabilités des relations de fratrie entre individus. Cependant, le logiciel Colony présente toujours l'estimation de  $N_b$  en faisant ou non l'hypothèse d'une reproduction aléatoire, les paramètres  $\alpha$  ainsi que  $N_1$  et  $N_2$  étant alors directement estimés lors de l'analyse par le logiciel.

#### 2.3.2.2. *Evaluation de l'effet du nombre de juvéniles analysés par sous-échantillonnage et simulation*

Afin d'estimer la taille de l'échantillon nécessaire pour obtenir une estimation fiable de  $N_b$ , un sous-échantillonnage aléatoire d'un nombre croissant de juvéniles de l'année a été réalisé pour certaines années sur les populations de saumon de la Nivelle, du Scorff et de La Roche. L'objectif est de déterminer si le nombre de juvéniles disponibles est suffisant pour une estimation non biaisée de  $N_b$ , ce qui devrait s'observer par la constance de l'estimation de  $N_b$  quel que soit le nombre de juvéniles utilisés dans l'analyse. Un second objectif est de définir une taille d'échantillon minimale à partir de laquelle l'ajout de juvéniles additionnels n'a pas d'impact sur l'estimation de  $N_b$ . Enfin, cette analyse devrait permettre d'identifier des conditions dans lesquelles  $N_b$  ne peut pas être estimé de façon fiable lorsque l'estimateur ne cesse de varier en fonction du nombre de juvéniles utilisés. Dans une telle situation, l'estimation de  $N_b$  est trop dépendante de l'échantillon de telle sorte qu'il n'est pas possible de déterminer des conditions d'utilisation fiable de la méthode.

L'analyse par Colony de nombreux jeux de données obtenus par sous-échantillonnages aléatoires des génotypes de juvéniles a fait appel à des méthodes de programmation en *bash* sous le système d'exploitation Linux permettant d'automatiser le tirage aléatoire des jeux de données, le formatage du fichier d'entrée pour Colony ainsi que la soumission parallèle pour analyse de l'ensemble des jeux de données sur un cluster de calcul du Mésocentre de Calcul Intensif de Bordeaux. Plusieurs centaines de jeux de données peuvent ainsi être générés et analysés de façon automatique et en parallèle en utilisant autant de processeurs et en un temps raisonnable (24h au maximum). Les résultats sont ensuite traités de façon automatique afin d'en extraire les estimations de  $N_b$  qui sont représentés graphiquement en fonction du nombre de juvéniles sous-échantillonnés grâce au logiciel R version 3.2.3 (R Development Core Team, 2016). L'ensemble du script permettant d'automatiser ces tâches est disponible sur demande auprès de l'auteur, l'ensemble étant trop complexe à expliquer pour le non initié et ne faisant pas partie des développements qui ont pour objectif d'être accessible à l'utilisateur final. Typiquement, des jeux de données contenant 15, 25, 50, 75, 100, 125, 150, 175, 200, 250 et plus juvéniles ont été générés avec 10 jeux de données soumis à Colony par nombre de juvéniles. Les paramètres pour l'analyse par Colony sont identiques à ceux décrits précédemment, mais la longueur des analyses a été réglée sur court afin d'obtenir des temps de calcul raisonnables.

L'approche par sous-échantillonnage permet de tester l'effet d'une réduction de la taille d'échantillon sur l'estimation de  $N_b$ , mais montre ses limites pour explorer l'effet d'un échantillon de grande taille. L'approche par simulation est complémentaire car la reproduction entre individus est simulée de façon contrôlée afin de générer un nombre prédéterminé de génotypes de juvéniles dans le cas d'un  $N_b$  connu. Pour effectuer cette simulation, nous avons utilisé le module de simulation du logiciel Colony (Wang, 2013) et nous nous sommes appuyés sur la matrice d'apparement obtenue par analyse de parentés des saumons de la Nivelle lors de l'étude de la saison de reproduction 201011 (Tentelier et al., 2016a). Cette matrice décrit le nombre de descendants assignés à chaque femelle anadrome (en colonne) et mâle anadrome ou tacon précoce (en ligne). Nous avons étendu artificiellement cette matrice pour simuler une population plus grande que celle de la Nivelle tout en maintenant une structure de succès reproducteur typique des saumons de cette population. Un total de 7000 génotypes de juvéniles ont ainsi été simulés, provenant de la reproduction entre 48 femelles et 174 mâles avec un nombre de reproducteurs efficace calculé de 57. Un sous-échantillonnage croissant du nombre de juvéniles soumis à Colony a ensuite été utilisé comme précédemment, avec la possibilité de sous-échantillonner un plus grand nombre de génotypes de juvéniles (jusqu'à 1200 juvéniles ; au-delà, l'analyse de 10 jeux de données prendrait trop de temps de calcul et les résultats n'apporteraient pas grand-chose en plus, voir partie Résultats).



### 2.3.2.3. Evaluation de l'effet de la stratégie d'échantillonnage

Nous avons tiré parti de résultats obtenus dans un précédent projet sur la même population de saumon de la Nivelles lors de deux saisons de reproduction en 2009-10 et 2010-11 (Tentelier et al., 2016a) pour étudier les effets 1) de l'étendue spatiale de l'échantillonnage de juvéniles 0+ (échantillonnage sur l'ensemble de la rivière ou sur la partie amont de la Nivelles), 2) de l'intensité d'échantillonnage (échantillonnage léger dans la présente étude contre intensif autour des zones de reproduction dans l'étude précédente) et 3) de la saison d'échantillonnage (en automne 9 mois après la reproduction dans la présente étude contre au printemps 3 mois après la reproduction dans l'étude précédente) sur l'estimation de  $N_b$ . Cependant pour contrôler les différences de stratégies d'échantillonnage, nous avons effectué différents types de sous-échantillonnage des génotypes de juvéniles 0+ afin que tous les paramètres d'échantillonnage soient comparables sauf le paramètre étudié (Tableau 2, Tableau 3, Tableau 4).

### 2.3.3. Données démographiques confrontées aux estimations de $N_b$

La variation temporelle de  $N_b$  estimé sur la population de saumon de la Nivelles est comparée à d'autres paramètres démographiques estimés sur cette population (Lange et al., 2015; Marchand et al., 2017) comprenant l'estimation du nombre d'adultes anadromes ( $N_c$ ) et du nombre de juvéniles 0+ présents à l'automne dans la rivière (Lange et al., 2015, 2017). Le nombre de frayères observées ainsi que leur localisation a été utilisé pour quantifier l'agrégation spatiale des frayères (Bacles et al., 2018; Bouchard et al., 2018).

## 3. Résultats & Discussions

### 3.1. Caractéristique des données générées par le protocole de génotypage de microsatellites

En ce qui concerne la population de saumon de la Nivelles, sur les 16 marqueurs génotypés (incluant 15 microsatellites et un marqueur sexe spécifique), le locus MHC1 n'a pas été retenu car il présentait un taux de données manquantes et un taux d'erreur élevés. Les autres 14 marqueurs microsatellites présentent entre 3 et 26 allèles et un taux modérément élevé de contenu en information polymorphique (entre 0.681 et 0.689) qui, lorsque combiné entre tous les marqueurs, représente une puissance élevée pour identifier les relations de parenté et l'apparentement entre individus (Bacles et al., 2018).

### 3.2. Variabilité génétique des saumons de la population de la Nivelles

Le niveau de diversité génétique chez les adultes anadromes est similaire entre les années : la diversité génétique (taux d'hétérozygote attendu corrigé pour la taille de l'échantillon  $H_{nb}$ ) varie faiblement entre 0.70 (intervalle de confiance : 0.49-0.81) en 2010 et 0.72 (0.53-0.83) en 2005. Le nombre moyen d'allèle calculé après raréfaction, afin de corriger les biais liés à la taille de l'échantillon, varie entre 9.7 (7.1-12.0) allèles par marqueurs en 2008 et 10.8 (8.1-13.2) allèles par marqueurs en 2009. Le niveau de diversité génétique est très similaire entre années pour les juvéniles, avec une diversité génétique variant entre 0.67 (0.47-0.78) pour 2014 et 0.71 (0.51-0.81) pour 2005 et un nombre moyen d'allèles calculé après raréfaction variant de 8.9 (6.7-10.9) en 2009 à 10.1 (7.5-12.5) en 2011. La diversité génétique n'est pas significativement différente entre les adultes et les juvéniles, que ce soit en comparant les adultes d'une année et leurs descendants ou bien globalement (résultats non montrés).

> Les indices standards de diversité génétique seuls sont très peu sensibles à d'éventuelles variations démographiques sur le court terme. Ces indices sont donc de mauvais indicateurs pour le suivi d'une population chez les saumons, et d'autres espèces en général, car la diminution de la diversité génétique mesurable avec des marqueurs moléculaires couramment utilisés ne sera perceptible qu'après plusieurs générations et dans le cas de variations très fortes de la taille des populations (cas d'effondrement extrêmes, ou de diminution brutale soutenue dans le temps des effectifs). Ce sont donc des indicateurs intéressants pour reconstruire les

événements démographiques relativement anciens mais pas de bon indicateurs pour le suivi d'une population dans le cadre de la gestion ou la conservation qui doivent rendre compte de phénomènes populationnels contemporains. A l'inverse, le nombre de reproducteurs efficaces rend compte de la variabilité du succès reproducteur des individus ainsi que de la survie des juvéniles lors d'une saison de reproduction. S'il peut être estimé correctement dans des conditions d'échantillonnage qui peuvent être mis en œuvre de façon réaliste dans le cadre de suivi de populations, cet indice est plus susceptible d'être en phase avec les objectifs des suivis dans le cadre des plans de gestion des populations.

### **3.3. Facteurs affectant l'estimation de $N_b$**

#### **3.3.1. Validation de l'estimation de $N_b$ à partir des seuls juvéniles**

L'estimation de  $N_b$  utilisant les juvéniles avec ou sans l'aide des parents potentiels donne des résultats similaires avec des intervalles de confiances qui se chevauchent largement (Tableau 1). La probabilité pour une paire d'individus d'être pleins-frères, demi-frère ou non apparentés n'est pas influencée par la connaissance des adultes potentiels et donc n'a pas d'influence sur l'estimation de  $N_b$ . Ceci est lié au niveau de polymorphisme élevé des marqueurs microsatellites utilisés, qui en combinaison, sont suffisamment informatifs pour pouvoir estimer les relations entre juvéniles, et donc  $N_b$ , de façon fiable à partir des seuls génotypes de juvéniles.

> La puissance des données générées par les 15 marqueurs microsatellites utilisés permet de se passer de l'information sur les adultes et d'estimer  $N_b$  de façon fiable à partir du génotypage des seuls juvéniles de l'année.

L'utilisation des génotypes des adultes produit une légère diminution du nombre de fratries identifiées (Tableau 1), car cette information supplémentaire améliore l'estimation et la prise en compte des erreurs de génotypage par Colony : quelques groupes de juvéniles qui seraient classés dans des familles différentes à cause d'une erreur de génotypage non identifiée vont se retrouver groupés dans la même famille grâce à l'information apportée en plus par les génotypes des adultes permettant d'identifier plus efficacement les erreurs de génotypages. Cette tendance à séparer des groupes de juvéniles à tort n'a en général pas d'effet dans l'estimation de  $N_b$  car ce type d'erreur reste marginal dans les cas d'échantillonnages classiques. Mais cet effet peut s'avérer plus problématique dans le cas où de nombreux juvéniles sont analysés (voir ci-après).

Tableau 1. Estimations de  $N_b$  aux cours de 10 années chez les saumons de la Nivelle en utilisant les génotypes des juvéniles 0+ en incluant ou non les génotypes des adultes anadromes.

Saison de reproduction	Génotypes utilisés	Nombre de 0+	Nombre d'adultes anadromes		Nombre de fratries			$N_b$
			mâles	femelles	plein-frère	demi-frère paternelle	demi-frère maternelle	
2004-05	0+	249			161	43	71	72 (53-101)
	0+ & Adultes	249	33	39	161	78	49	76 (56-105)
2005-06	0+	67			38	16	22	23 (14-41)
	0+ & Adultes	67	30	44	38	23	18	24 (15-46)
2006-07	0+	101			69	22	39	43 (28-66)
	0+ & Adultes	101	22	32	67	37	29	44 (29-68)
2007-08	0+	208			73	40	28	33 (21-55)
	0+ & Adultes	208	33	27	74	41	29	32 (20-52)
2008-09	0+	127			73	45	19	31 (19-52)
	0+ & Adultes	127	27	41	71	45	22	30 (19-51)
2009-10	0+	237			120	58	23	39 (26-60)
	0+ & Adultes	237	28	26	118	61	24	37 (24-59)
2010-11	0+	200			103	30	53	52 (36-78)
	0+ & Adultes	200	74	67	99	59	33	54 (38-80)
2011-12	0+	223			119	27	59	51 (35-74)
	0+ & Adultes	223	28	42	117	58	32	46 (31-69)
2012-13	0+	181			75	29	33	31 (19-51)
	0+ & Adultes	181	27	20	75	46	22	32 (20-54)
2013-14	0+	222			129	52	37	35 (23-57)
	0+ & Adultes	222	42	52	124	59	39	40 (27-62)

### 3.3.2. Taille optimale d'échantillonnage de juvéniles de l'année

Dans le cas de la Nivelle, le nombre optimal de juvéniles à échantillonner dépend de la magnitude de  $N_b$  (Figure 3).  $N_b$  peut être estimé de façon fiable à partir d'une taille d'échantillon égale à deux fois  $N_b$  (Figure 3a-c). Une taille d'échantillon faible conduit à une sous-estimation de  $N_b$  et à une augmentation de l'intervalle de confiance.

L'approche par simulation permettant de tester l'effet de grande taille d'échantillon met en évidence une estimation fiable de  $N_b$  pour des échantillons d'effectif compris entre deux fois  $N_b$  et dix fois  $N_b$  (environ 500 juvéniles ; Figure 3d). Utiliser un plus grand nombre de juvéniles conduit à une surestimation de  $N_b$  (Figure 3d). Ce phénomène est lié au fractionnement par erreur des grandes familles de pleins frères (qui sont plus nombreuses au fur et à mesure que la taille de l'échantillonnage augmente) en plusieurs familles à cause de la gestion complexe de l'erreur de génotypage dans le cas d'échantillonnage de grande taille.

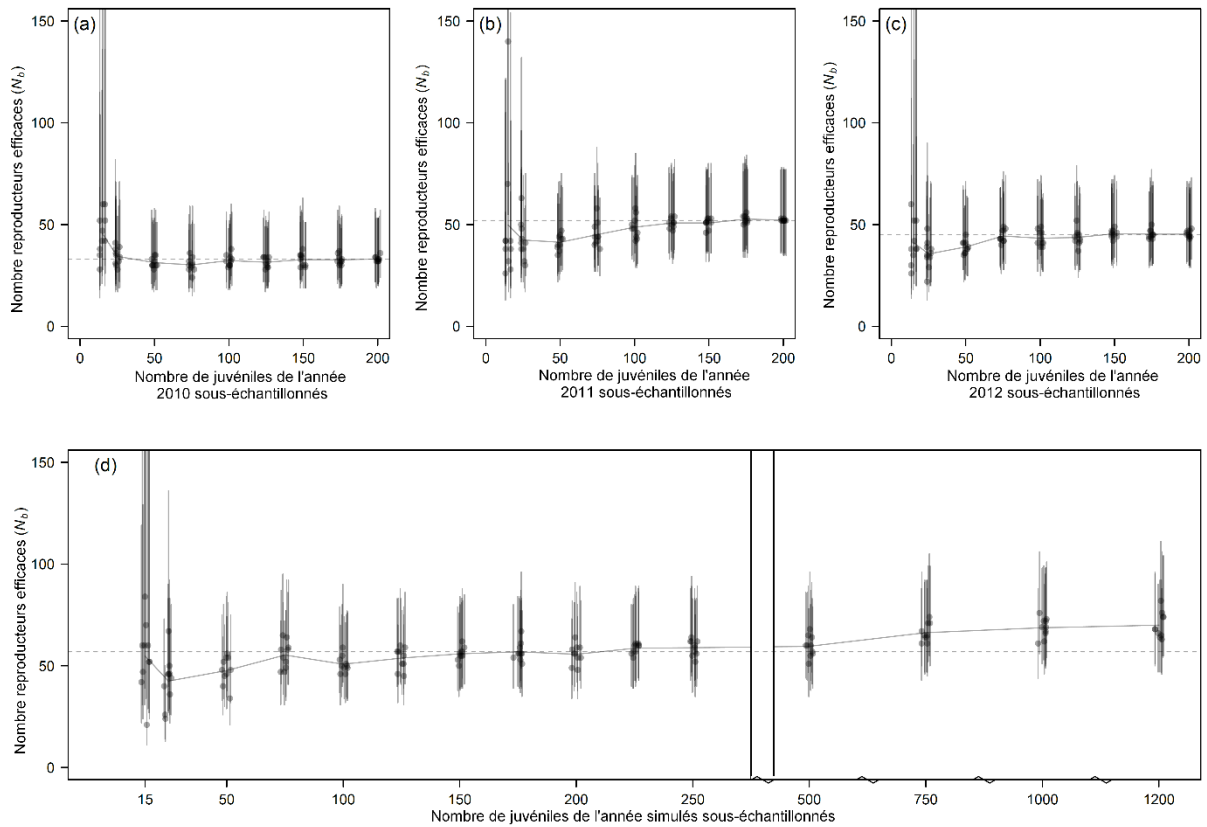


Figure 3. Sous-échantillonnage des génotypes de juvéniles 0+ pour trois années sur la Nivelle (a-c) et des génotypes de juvéniles simulées selon un scénario de 57 reproducteurs efficaces (d). Chaque point représente l'estimation du nombre de reproducteurs efficaces ( $N_b$ ) obtenu par une analyse par Colony pour un certain nombre de juvéniles 0+ tirés aléatoirement (en abscisse) et les traits verticaux délimitent l'intervalle de confiance à 95% de l'estimation. Un léger décalage entre les points proches a été introduit afin de faciliter la distinction entre les points. La ligne horizontale en pointillés indique la valeur de  $N_b$  la plus probable (a-c ; déterminée avec le jeu de donné complet) ou connue (d ; calculée à partir des données simulées). La ligne pleine relie la moyenne des valeurs de  $N_b$  estimée pour l'ensemble des jeux de données présentant le même nombre d'individus afin de souligner la tendance des variations des estimations de  $N_b$  en fonction du nombre d'individus analysés.

Ce phénomène est également observé avec les données réelles correspondant à l'échantillonnage intense des juvéniles de l'année réalisé sur la Nivelle aux printemps 2010 et 2011 (Tableau 2). La grande taille des échantillonnages réalisés au printemps (1677 et 1351 juvéniles 0+ respectivement en 2010 et 2011) conduit à de fortes surestimations de  $N_b$  (121 (95-156) et 95 (73-126), respectivement). Lorsque les génotypes des adultes potentiels sont rajoutés dans l'analyse par Colony, l'erreur de génotypage est mieux prise en compte, les grandes familles de pleins frères sont alors mieux identifiées et l'estimation de  $N_b$  devient alors plus réaliste (19 (11-35 et 23 (14-42)). Si un sous-échantillonnage d'environ 150 à 200 juvéniles est effectué, une estimation réaliste de  $N_b$  est obtenue à partir des seuls juvéniles de l'année (32 (21-52) et 23 (14-42) ; Tableau 2).

Tableau 2. Effet de la taille de l'échantillons de juvéniles 0+ de la Nivelle sur l'estimation de  $N_b$ .

Echantillonnage <sup>a</sup>	Génotypes utilisés	Nombre de 0+	Nombre de fratries			$N_b$
			plein-frère	demi-frère paternelle	demi-frère maternelle	
Printemps 2010	0+	1677	799	57	113	121 (95-156)
Printemps 2010	0+ & Adultes	1677	187	60	16	19 (11-35)
Printemps 2010	0+	189*	104	41	14	32 (21-52)
Printemps 2011	0+	1351	458	115	53	95 (73-126)
Printemps 2011	0+ & Adultes	1351	181	59	28	23 (14-42)
Printemps 2011	0+	147*	62	28	23	23 (14-42)

<sup>a</sup> basé sur une ré-analyse des données issus de Tentelier et al. (2016b) \* sous-échantillonnage aléatoire d'un nombre réduit de juvéniles de l'année.

> Le nombre de juvéniles échantillonnés doit donc être suffisant pour bien caractériser la proportion des différents types d'apparement entre juvéniles (fréquences des relations de fratrie et de non-apparement) et éviter le risque de sous-estimation de  $N_b$ . En même temps, un échantillonnage trop important risque de surestimer  $N_b$  car de trop nombreux juvéniles appartenant à la même famille risque d'être classifiés dans des familles différentes lors de la reconstruction des fratries. Heureusement, une large gamme de taille d'échantillon s'étendant de 2 fois  $N_b$  à 10 fois  $N_b$  (soit de 100 à 500 juvéniles de l'année pour un  $N_b$  de 50) permet une estimation fiable de  $N_b$  avec les seuls génotypes des juvéniles 0+. Il faut noter que pour la Nivelle, 2 fois  $N_b$  correspond grossièrement au nombre d'adultes anadromes estimés dans la population.

### 3.3.3. Effet de l'étendue spatiale de l'échantillonnage

Les estimations de  $N_b$  obtenues à partir d'échantillons provenant d'une partie de la rivière (zone amont, concentrant l'essentiel des zones de reproduction) sont légèrement différentes de celles estimées à partir des juvéniles issus de l'ensemble de la rivière (Tableau 3). D'autre part, les différences observées ne vont pas dans le même sens chaque année : l'estimation de  $N_b$  est plus grande pour un échantillonnage spatialement restreint en 2010 (39 (26-60) sur la rivière entière contre 50 (34-76) sur la seule partie amont en 2010) alors que c'est l'inverse en 2011 (52 (36-76) sur l'ensemble de la rivière contre 40 (27-62) sur la zone amont en 2011). Cette observation démontre que  $N_b$  n'est pas un paramètre linéaire : sa valeur sur la population entière est loin d'être la somme des  $N_b$  estimés localement le long des différentes sections d'une rivière. En fait,  $N_b$  est fortement dépendant de la structure spatiale de la reproduction (structure d'apparement entre individus, localisation des zones de frayères) ainsi que de la variabilité du taux de survie le long du linéaire. De ce fait, des résultats contre-intuitifs peuvent apparaître comme pour l'automne 2010 où  $N_b$  est plus faible lorsque qu'il est estimé sur l'ensemble de la rivière que lorsqu'il est estimé sur une portion de rivière seulement (Tableau 3). Cette année-là, l'essentiel de la reproduction s'est concentré sur la partie amont de la rivière. La prise en compte d'un nombre limité de géniteurs additionnels s'étant reproduit en aval de la rivière lorsque les juvéniles sont échantillonnés sur l'ensemble de la population devrait se traduire par une augmentation de  $N_b$ . Au lieu de cela, la prise en compte de sections favorables (amont) et moins favorable (aval) va augmenter fortement la variation du succès reproducteur entre individus et donc la contribution relative des géniteurs, ce qui aura pour effet de diminuer l'estimation de  $N_b$ . En fait,  $N_b$  estimé en aval en 2010 est faible (14 (8-30)) et donc lorsque ce secteur est pris en compte, il va diminuer le  $N_b$  global, alors que ce même secteur montre un  $N_b$  plus important l'année suivante (24 (15-46)). Il est donc logique que  $N_b$  estimé en 2011 sur l'ensemble de la rivière soit supérieur à celui estimé en amont car, cette année-là, la structure de la reproduction et de la survie des juvéniles est relativement bonne pour la section aval sur la Nivelle.

Tableau 3. Effet de l'étendue spatial de la Nivelles échantillonnée sur l'estimation de  $N_b$ .

Saison d'échantillonnage	Génotypes utilisés	Zone échantillonnée	Nombre de 0+	Nombre de fratries			$N_b$
				plein-frère	demi-frère paternelle	demi-frère maternelle	
Automne 2010	0+	rivière entière	237	118	55	25	39 (26-60)
Automne 2010	0+	amont de Olha*	189	101	36	47	50 (34-76)
Automne 2011	0+	rivière entière	200	103	30	53	52 (36-78)
Automne 2011	0+	amont de Olha*	147	70	29	33	40 (27-62)

\* sous-échantillonnage des juvéniles de l'année capturés à l'amont d'Olha.

> L'étendue spatiale de la zone échantillonnée a un effet important, complexe et imprédictible sur l'estimation de  $N_b$ . Le nombre de reproducteurs efficaces d'une population n'est pas la somme (ni la moyenne) des nombres de reproducteurs efficaces estimés sur différentes sections. Ceci est probablement lié au fait que les sections échantillonnées ne peuvent être considérées comme indépendantes car un géniteur peut éventuellement se reproduire sur plusieurs secteurs ou ses juvéniles peuvent se disperser sur plusieurs secteurs après émergence. L'estimation étant basée sur les probabilités d'apparementement entre paires de juvéniles,  $N_b$  est affecté par de nombreux paramètres liés à l'échantillonnage ou paramètres biologiques sous-jacents. Il est donc important d'obtenir un échantillonnage présentant une bonne couverture spatiale de la population étudiée et de garder la même stratégie d'échantillonnage chaque année afin d'obtenir des estimations qui soient pertinentes biologiquement et comparables. Les estimations de  $N_b$  sur des portions des rivières doivent être interprétées avec la plus grande prudence car même si elles peuvent *a priori* renseigner sur la qualité d'un site pour la reproduction et la survie des juvéniles, elles seront fortement sensibles à d'autres paramètres, en particulier le déplacement des géniteurs lors de la reproduction et la dispersion des juvéniles entre sites. D'autre part, un nombre suffisant de sites doivent être échantillonnés afin d'obtenir un échantillon représentatif de la population. En effet, un sous-échantillonnage de moins de 10 sites échantillonnés sur la Nivelles conduit à une sous-estimation de  $N_b$  (résultats non montrés). Dans le cas de cette population, un point d'échantillonnage tous les 3 km de linéaire du cours d'eau principal (10 sites sur 30 km) semble être un minimum requis.

### 3.3.4. Effet de la saison d'échantillonnage

La comparaison des estimations de  $N_b$  obtenues au printemps et en automne sur deux années montre une augmentation de  $N_b$  au cours de l'année avec des estimations inférieures au printemps (32 (21-52) et 23 (14-42) aux printemps 2010 et 2011 respectivement contre 50 (34-76) et 40 (27-62) à l'automne). Ces résultats peuvent s'expliquer par plusieurs facteurs. D'un part, la comparaison entre saisons n'a pu être effectuée que pour des échantillonnages de l'amont de la Nivelles. Il est donc possible que des juvéniles issus de la partie aval aient eu le temps de se déplacer sur la partie amont entre le printemps et l'automne, augmentant donc  $N_b$  à l'automne alors que  $N_b$  estimé au printemps, avant que les juvéniles n'aient eu l'opportunité de se déplacer loin de leur zone de naissance, traduise plus fidèlement le nombre de géniteurs efficaces de la seule zone amont. Cependant, la zone amont concentre l'essentielle des reproductions sur la Nivelles et d'autres facteurs sont donc à considérer.  $N_b$  estimé à partir des juvéniles échantillonnés à l'automne rend compte à la fois de la reproduction mais également de la migration et de la survie des juvéniles au cours des mois suivant leur émergence. D'un autre côté,  $N_b$  estimé à partir des juvéniles échantillonnés au printemps rend compte plus directement de la reproduction qui est plus proche temporellement. Ainsi, un échantillonnage de printemps est plus susceptible d'échantillonner des familles aux effectifs très déséquilibrés autour des zones de frayères, alors qu'un échantillonnage à l'automne conduira à échantillonner plus de familles (Tableau 4) aux effectifs plus équilibrés liés aux événements de dispersion et de mortalité, ce qui se traduira par des estimations de  $N_b$  plus élevées (Tableau 4).

Tableau 4. Effet de la saison d'échantillonnage sur l'estimation de  $N_b$ .

Saison d'échantillonnage	Génotypes utilisés	Zone échantillonnée	Nombre de 0+	Nombre de fratries			$N_b$
				plein-frère	demi-frère paternelle	demi-frère maternelle	
Printemps 2010	0+	amont de Olha	189	104	41	14	32 (21-52)
Automne 2010	0+	amont de Olha*	189	101	36	47	50 (34-76)
Printemps 2011	0+	amont de Olha*	147	62	28	23	23 (14-42)
Automne 2011	0+	amont de Olha	147	70	29	33	40 (27-62)

\* sous-échantillonnage des juvéniles de l'année capturés à l'amont d'Olha.

> Tout comme pour les effets liés à la spatialisation de l'échantillonnage, les variations de  $N_b$  peuvent être étudiées entre années ou comparées entre rivières à condition qu'ils soient estimés à partir de génotypes de juvéniles échantillonnés à la même saison. Comme  $N_b$  est un paramètre intégratif qui inclut les phénomènes liés à la reproduction mais également à la survie et à la dispersion des juvéniles, la période d'échantillonnage est primordiale pour pouvoir interpréter ce paramètre correctement et en tirer des interprétations biologiques pertinentes. Le suivi des variations de  $N_b$ , ou de façon plus fine, des probabilités d'apparement entre juvéniles, au cours du temps sur un linéaire pourrait fournir des informations sur les paramètres de survie ou de dispersion des juvéniles en lien avec l'espace ou l'environnement fin des cours d'eau. Ce type d'analyses, qui pourraient faire l'objet de recherches ultérieures, dépasse les enjeux liés à ce projet.

### 3.4. Application des recommandations dans le cas d'autres populations

#### 3.4.1. Scorff : généralisation des recommandations à une population de plus grande taille

L'étude d'une population de plus grande taille (le Scorff) permet de généraliser les observations réalisées sur la Nivelle. Ainsi sur les trois années étudiées présentant un nombre d'adultes anadromes contrasté, le dimensionnement de la taille de l'échantillonnage de juvéniles de l'année a été réalisé en tenant compte les résultats obtenus sur la Nivelle :  $2x N_b$  (ce qui en général pour le saumon de la Nivelle équivaut au nombre d'adultes anadromes estimés) juvéniles devraient être en théorie suffisant à estimer convenablement  $N_b$ . Nous avons donc sous-échantillonné les juvéniles de l'année à hauteur de 1.2 fois le nombre d'adultes anadromes estimé (Tableau 5), afin de prévoir un peu de marge et d'anticiper d'éventuelles difficultés techniques liées à de l'ADN dégradé du fait de l'ancienneté des échantillons.

Tableau 5. Application des recommandations de taille d'échantillon à trois années sur le Scorff présentant un nombre d'adultes anadromes contrasté.

Saison de reproduction	Nombre d'adultes anadromes estimés	Nombre de juvéniles sous-échantillonnés <sup>a</sup>	Nombre de génotypes de juvéniles obtenus	$N_b$ <sup>b</sup>	Nombre de juvéniles minimal <sup>c</sup>
2009-2010	376	451	417	140 (110-180)	350
2011-2012	603	724	560*	227 (190-274)	500
2014-2015	868	1042	975	263 (219-310)	800

<sup>a</sup> : déterminé *a priori* à partir des résultats de la Nivelle (cible = 1.2 x le nombre d'adultes anadromes estimés) ; <sup>b</sup> : estimé à partir de l'ensemble des génotypes de juvéniles disponibles ; <sup>c</sup> : estimé par sous-échantillonnages successifs comme le nombre minimal de juvéniles nécessaires pour obtenir une estimation de  $N_b$  similaire à celle obtenue avec l'ensemble des juvéniles disponibles ; \* le faible taux de succès de génotypage (77% cette année-là en comparaison aux plus de 90% obtenus les autres années) correspond à des problèmes liés à la faible qualité de l'ADN extrait à partir des morceaux de nageoire échantillonnés cette année-là.

Les résultats de sous-échantillonnages successifs permettent d'observer que les estimations de  $N_b$  augmentent avec la taille croissante de l'échantillon de juvéniles puis plafonnent au-delà d'un certain seuil (Figure 4), indiquant que la taille des effectifs déterminée *a priori* est suffisante pour des estimations fiables de

$N_b$  même dans le cas d'une population de plus grande taille comme le Scorff. La taille de l'échantillon à partir de laquelle les estimations de  $N_b$  convergent permet de déterminer un nombre minimal mais suffisant de juvéniles pour estimer  $N_b$  (Tableau 5). Pour la saison de reproduction 2011-2012, le nombre de juvéniles 0+ effectivement génotypés est tout juste suffisant pour atteindre le plateau d'estimation de  $N_b$ , soulignant l'importance de prévoir une marge afin de prendre en compte d'éventuelles difficultés d'analyse au laboratoire.

> Les préconisations déterminées sur la population de la Nivelle en terme de nombre de juvéniles à échantillonner pour estimer  $N_b$  se sont montrées pertinentes pour une population de taille plus importante. En fait, la préconisation d'un échantillon de juvéniles 0+ d'effectif égal à 2 fois  $N_b$  peut être difficile à prévoir sans aucune analyse génétique préliminaire sur la population. La taille de l'échantillon nécessaire peut également être estimée par le nombre d'adultes anadromes présents dans la population, une information plus intuitive à prédire *a priori* qu'une valeur de  $N_b$ . Cette préconisation est plus robuste puisque comme on l'observe pour le Scorff,  $N_b$  n'augmente pas proportionnellement au nombre d'adultes anadromes, si bien que le nombre de juvéniles 0+ à échantillonner est plus justement prédit par le nombre d'adultes anadromes que par  $N_b$  (en particulier pour la saison de reproduction 2014-15 présentant de nombreux adultes anadromes ; Tableau 5).

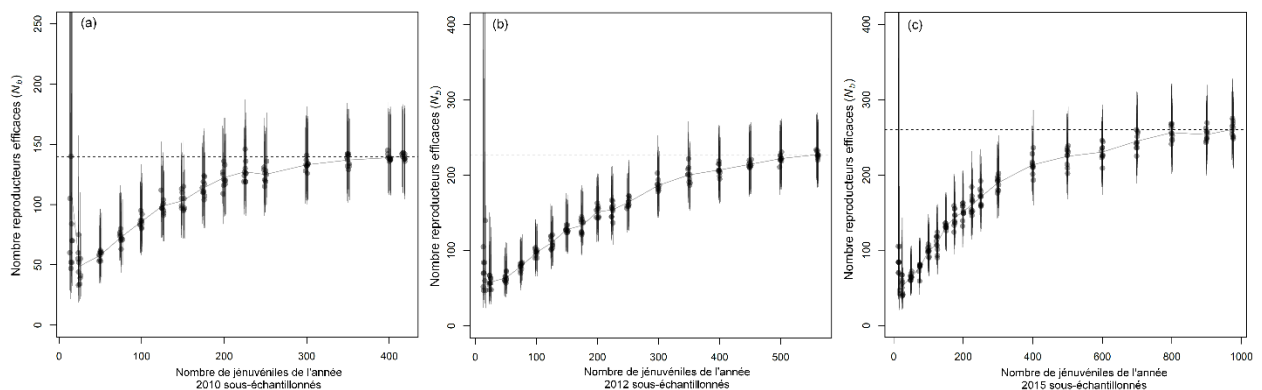


Figure 4. Sous-échantillonnage des génotypes de juvéniles 0+ pour trois saisons de reproduction, (a) 2009-10, (b) 2011-12 et (c) 2014-15, présentant des nombres d'adultes anadromes contrastés sur le Scorff. Chaque point représente l'estimation du nombre de reproducteurs efficaces ( $N_b$ ) obtenu par une analyse par Colony pour un certain nombre de juvéniles 0+ obtenu par tirage aléatoire (en abscisse) et les traits verticaux délimitent l'intervalle de confiance à 95% de l'estimation. Un léger décalage entre les points proches a été introduit afin de faciliter la distinction entre les points. La ligne horizontale en pointillés indique la valeur de  $N_b$  la plus probable (déterminée avec le jeu de données complet). La ligne pleine relie la moyenne des valeurs de  $N_b$  estimé pour l'ensemble des jeux de données présentant le même nombre d'individus afin de souligner la tendance des variations des estimations de  $N_b$  en fonction du nombre d'individus analysés.

### 3.4.2. La Roche affluent de l'Oir : application à une portion de population

L'analyse de nombreux juvéniles 0+ échantillonnés sur La Roche qui est un affluent de l'Oir, lui-même un affluent du fleuve Sélune, montre la difficulté d'estimer convenablement  $N_b$  sur une portion d'un plus grand système (Figure 5). En effet, l'estimation de  $N_b$  ne fait qu'augmenter avec le nombre de juvéniles 0+ utilisés dans l'analyse et aucun plateau ne semble atteint. Aucun échantillon n'est de taille suffisante pour estimer correctement  $N_b$  dans ce cas, alors que la grandeur de  $N_b$  estimé reste modérée (59 (42-85) pour 568 juvéniles 0+ utilisés). L'utilisation des génotypes des adultes anadromes confirme la valeur de  $N_b$  obtenue avec les seuls juvéniles indiquant que les estimations ne sont pas biaisées par des problèmes liés à la détection d'erreurs de génotypages ni à des fractionnements artéfactuels de grandes fratries lors de la reconstruction (résultats non montrés). Clairement, dans ce cas, un effectif en juvéniles 0+ de 2 fois  $N_b$  n'est pas suffisant et il n'est pas possible de prédire avec certitude combien de juvéniles 0+ devraient être échantillonnés pour estimer  $N_b$ , bien qu'une amorce de plafonnement semble se dessiner autour de 450 juvéniles 0+ (Figure 5). Ce résultat peut être lié à la complexification introduite par le fait que l'affluent étudié représente une population ouverte. Dans cette situation, des individus peuvent à la fois se reproduire dans La Roche mais aussi dans d'autres affluents du bassin



versant (l'Oir ou la Sélune) et leurs descendants peuvent s'être déplacés entre ces différents affluents avant l'automne. Ces deux phénomènes complexifient grandement la structure d'apparement observée entre juvéniles 0+ échantillonnés dans La Roche de sorte que l'estimation de  $N_b$  dans ce cas ne reflète plus uniquement la reproduction et la survie des juvéniles produit dans cet affluent, mais intègre également une partie difficile à déterminer du fonctionnement de la population à l'échelle du bassin versant. Comme des juvéniles 0+ échantillonnées sur l'ensemble du bassin versant de la Sélune n'ont pas été génotypés dans cette étude (comme ce fut le cas pour la Nivelles ou du Scorff précédemment), il est très difficile d'interpréter telle quelle l'estimation de  $N_b$  dans ce cas, que ce soit pour caractériser l'affluent La Roche ou bien même la population de saumons de l'Oir ou de la Sélune dans sa globalité.

> L'estimation de  $N_b$  sur une seule portion (affluent) d'un plus grand système est très difficile à interpréter à cause du degré de complexité supérieure introduit par les mouvements des reproducteurs pendant la saison de reproduction et des juvéniles après émergence. Dans ce cas, un nombre très important de juvéniles 0+ doit être échantillonné pour obtenir une estimation valide de  $N_b$  qui lui-même devra faire l'objet de la plus grande attention lors de son interprétation. Cependant, l'utilisation de  $N_b$  comme indice de suivi d'une population localisée dans une portion d'un système plus grand peut tout de même être considérée si la portion du système considéré constitue une population plus ou moins isolée du reste du système (population fermée). Une analyse systématique de la même zone comparée sur plusieurs années peut rendre compte des évolutions de la population sur la zone étudiée (voir par exemple Ozerov et al., 2015 pour une application dans le contexte d'évaluation d'actions de gestion chez la truite commune).

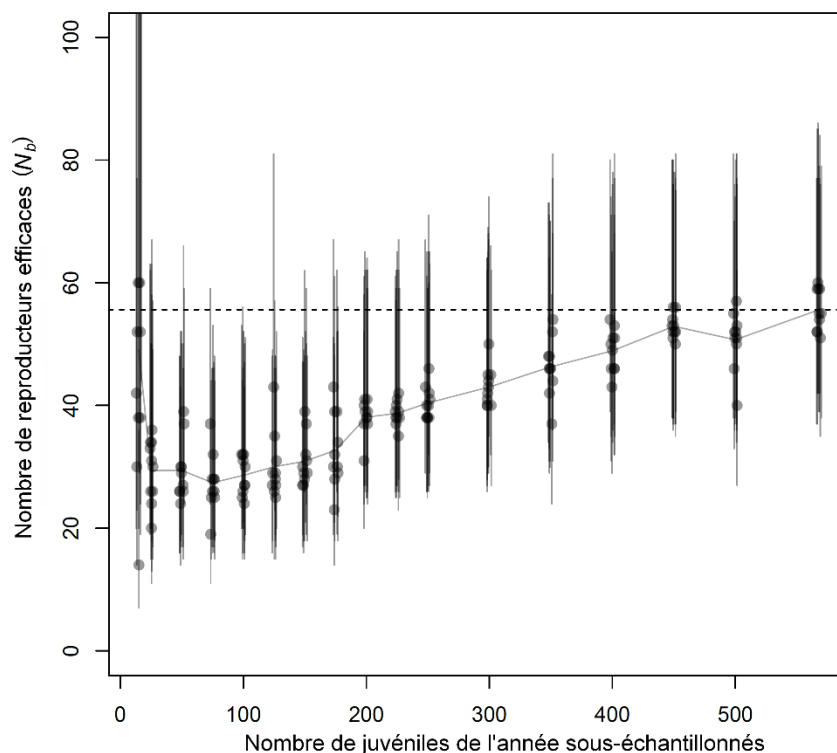


Figure 5. Sous-échantillonnage du nombre de génotypes de juvéniles 0+ utilisé pour estimer  $N_b$  dans le ruisseau de La Roche en 2015 (effectué à partir des données génotypiques fournies par Lucie Montorio, François Le cor, Marie Nevoux et Guillaume Evanno, INRA UMR ESE).

### 3.5. Apport du suivi de $N_b$ au cours du temps

L'estimation de  $N_b$  au cours de 10 saisons de reproduction de la population de saumon de la Nivelles donne l'opportunité d'illustrer les apports de ce paramètre dans le cadre d'un suivi long terme d'une population et sa

complémentarité par rapport aux données d'observations et aux paramètres démographiques classiques accessibles pour cette population (Lange et al., 2015). Le nombre d'adultes anadromes estimé varie entre 72 (65-83) pour la saison de reproduction 2009-10 à 174 (162-191) pour 2010-11 (Lange et al., 2015 ; moyenne globale ( $m$ ) = 99, écart type ( $et$ ) = 31, Figure 6a). Le nombre de frayères observées varie entre 57 pour 2012-13 et 132 pour 2010-11 (Lange et al., 2015 ;  $m$  = 81,  $et$  = 24 ; Figure 6b). Le nombre total de juvéniles 0+ estimé à l'automne varie de 2804 (1847-3934) en 2009 et 15656 (11816-20049) en 2012 (Lange et al., 2015 ;  $m$  = 6263,  $et$  = 3964 ; Figure 6c). Lors de cette étude, l'estimation de  $N_b$  varie de 23 (14-41) pour la saison 2005-06 à 72 (53-101) pour 2004-05 (Figure 6d). Le nombre d'adultes anadromes estimé ( $N_c$ ) n'est pas corrélé avec le nombre de juvéniles estimé (Spearman  $r=0.01$ ,  $p$ -value=0.99,  $N=10$ ), ni avec le nombre de reproducteurs efficaces ( $N_b$  ; Spearman  $r=0.027$ ,  $p$ -value=0.44,  $N=10$ ). Cependant, une corrélation significative a été trouvée entre le nombre de juvéniles estimé et  $N_b$  (Spearman  $r=0.69$ ,  $p$ -value=0.03,  $N=10$ ). L'absence de corrélation entre  $N_c$  et  $N_b$  se traduit par une grande variabilité du ratio  $N_b/N_c$  entre les années (Figure 6e) allant de 0.23 à 0.67 pour les saisons de reproduction 2005-06 et 2004-05 respectivement ( $m=0.43$ ,  $et=0.15$ ). De façon générale, les différents paramètres ne covarient pas le long de la série temporelle étudiée (Figure 6). Une valeur de  $N_b$  relativement élevée (~50) peut résulter de conditions favorables à la reproduction ( $N_c$  et nombre de frayères élevés en 2010-11 par exemple) ou de bon taux de survie des juvéniles (nombreux juvéniles en 2011-12 malgré des valeurs modérées de  $N_c$  et de nombre de frayères). La survie des juvéniles jusqu'à la période d'échantillonnage est un facteur important à considérer pour l'interprétation des variations de  $N_b$  car des années avec des nombres similaires d'adultes anadromes et de frayères (typiquement autour de 100 adultes anadromes et 65 frayères comme pour les saisons 2004-05 et 2005-06) peuvent néanmoins aboutir à des estimations de  $N_b$  très contrastées (73 (53-101) et 23 (14-41) pour ces deux années respectivement), probablement en lien avec des taux de survie de juvéniles contrastés (nombre de juvéniles à l'automne estimés de 9360 (7648-11298) et 2553 (1648-3575), respectivement). De la même façon, des configurations démographiques différentes peuvent conduire à des résultats similaires en termes de nombre de reproducteurs efficaces comme le montre par exemple la comparaison des années 2010-11 et 2011-12 (Figure 6). En 2010-11, un nombre élevé d'adultes anadromes a rencontré des conditions favorables pour la reproduction comme en témoigne le nombre élevé de frayères observées. Cependant cette reproduction s'est traduit par un nombre de juvéniles estimé en automne en dessous de la moyenne ce qui indique une forte mortalité juvénile cette année. A l'opposé, un nombre relativement moyen d'adultes anadromes et de frayères pour la saison 2011-12 a produit un nombre très élevé de juvéniles estimé à l'automne. Dans le premier cas, une mortalité juvénile importante, probablement répartie sur l'ensemble de la rivière, a compensé les conditions favorables à la reproduction alors que l'année suivante, des conditions favorables à la survie des juvéniles ont compensé des conditions moyennes pour la reproduction. Au final, ces deux saisons aux configurations démographiques très différentes résultent en un nombre de reproducteurs efficaces similaire et donc la production de cohortes présentant les mêmes potentialités adaptatives.

Au-delà de la survie des juvéniles, cette grande variabilité du ratio  $N_b/N_c$  qui a été déjà observé chez d'autres populations de saumon Atlantique (Palstra et al., 2009; Ferchaud et al., 2016; Perrier et al., 2016) indique également que la variation du succès reproducteur, en plus du nombre d'individus qui se reproduisent, joue un rôle important dans le nombre de reproducteurs efficaces estimé. Le saumon Atlantique est une espèce polygame : chaque femelle est capable de pondre dans plusieurs frayères, chaque frayère peut contenir des œufs de plusieurs femelles et être fertilisée par plusieurs mâles et chaque mâle peut se reproduire séquentiellement avec plusieurs femelles au cours de la saison de reproduction (de Gaudemar et al., 2000; Garant et al., 2001; Taggart et al., 2001; Grimardias et al., 2010; Weir et al., 2010). Outre ces caractéristiques du système de reproduction, la variabilité du succès reproducteur est également liée au simple fait de se reproduire ou non. En effet, dans une étude précédente sur la Nivelle (Tentelier et al., 2016a, 2016b), il a été montré que la moitié des mâles anadromes et un tiers des femelles anadromes n'avaient aucune descendance détectable parmi les nombreux juvéniles échantillonnés au printemps. La variabilité du succès d'apparement pourrait être expliqué par les conditions environnementales qui peuvent impacter le comportement des individus et expliquer une partie de la variabilité du ratio  $N_b/N_c$  au cours du temps. Par exemple,  $N_b$  est sensible au niveau d'agrégation des frayères : il est maximal les années où les frayères étaient soit très agrégées ou bien au contraire réparties très uniformément le long de la Nivelle et présentait des valeurs faibles pour des niveaux d'agrégation des frayères moyen (Bacles et al., 2018). Le niveau de compétition entre individus pour la reproduction ou la facilité d'accéder

aux différentes frayères spatialement proches et réparties de façon homogène sont des facteurs qui peuvent influencer le comportement reproducteur chez l'espèce et jouer un rôle important sur la variation du succès reproducteur entre individus.

Un autre élément à prendre en compte pour expliquer la variabilité du succès reproducteur et qui n'est pas pris en compte dans les paramètres démographiques classiquement utilisés est la capacité pour certains tacons mâles de mûrir sexuellement et de se reproduire dès l'âge d'un an dans les populations du sud de l'aire de répartition de l'espèce (dont la Nivelle ; Tentelier et al., 2016a). Le fait que ces individus furtifs ne soient pas pris en compte dans l'estimation du nombre de reproducteur dans la population ( $N_c$  n'intégrant que les adultes anadromes), et que le nombre de tacons mâles mûrant de façon précoce puisse fortement varier d'une année à l'autre en fonction des conditions environnementales, peut probablement expliquer une part de la variation du ratio  $N_b/N_c$  observée.  $N_b$  étant estimé à partir des relations d'apparement entre juvéniles et ne nécessitant aucune connaissance ou échantillon concernant les reproducteurs, ce paramètre inclut d'emblée les caractéristiques de la population d'individus reproducteurs, possiblement inobservables, que composent les adultes anadromes si aucun piège n'est disponible à l'entrée du bassin ou les individus furtifs présentant des stratégies de reproduction alternatives (tacon mâles précoces par exemple).

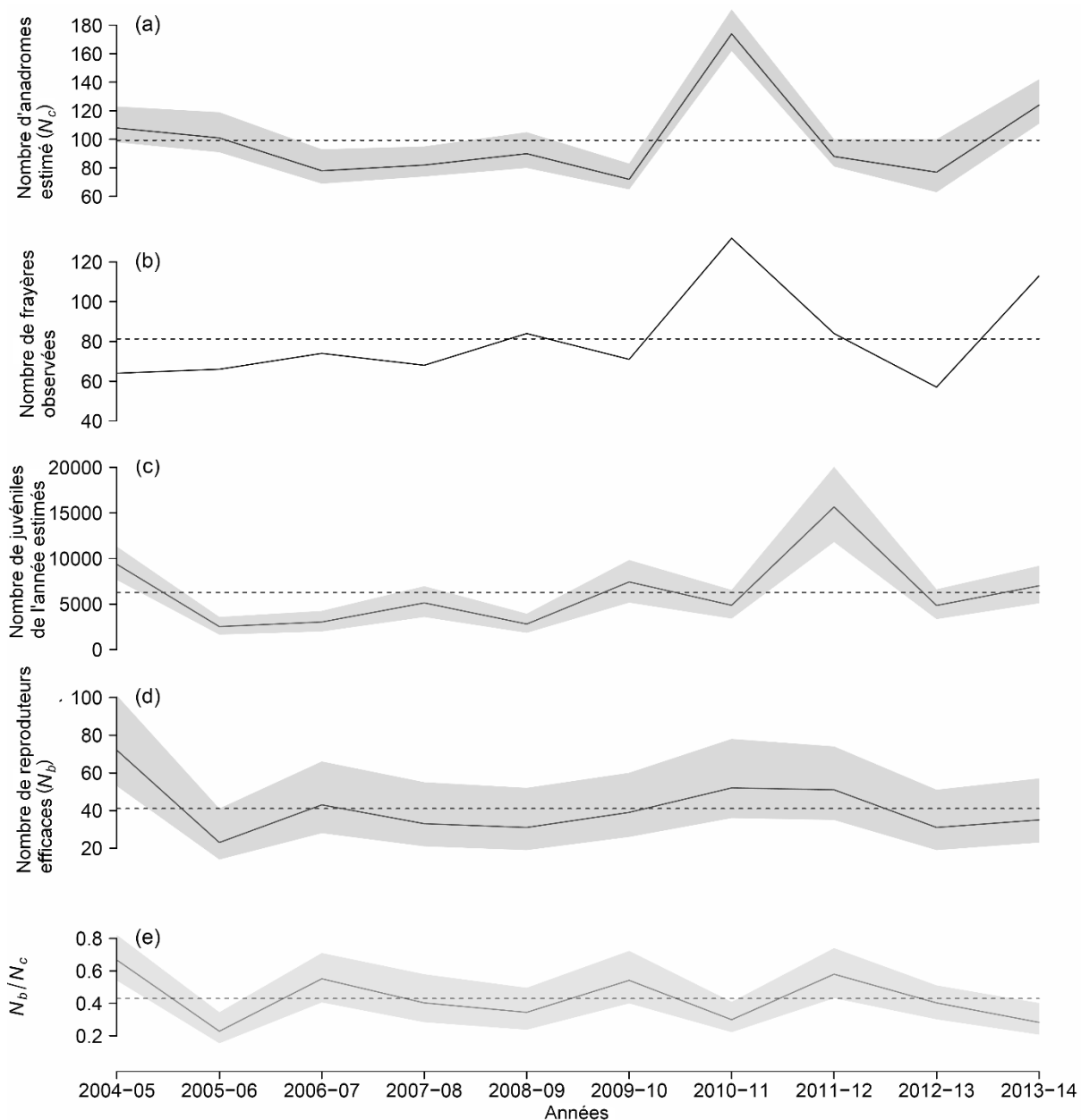


Figure 6. Variation au cours de 10 saisons de reproduction consécutives de la population de saumons de la Nivelle des paramètres populationnelles : (a) le nombre d'adultes anadromes estimés, (b) le nombre de frayères observées, (c) les nombres de juvéniles 0+ estimés en automne, (d) le nombre de reproducteurs efficaces et (e) le ratio nombre de reproducteurs efficaces sur nombre d'adultes anadromes estimés. Les lignes continues indiquent l'évolution des estimateurs entre années et les polygones représentent l'intervalle de crédibilité à 95% (a-c), l'intervalle de confiance à 95% (d) et leur ratio respectif (e). Les lignes horizontales en pointillé représentent les valeurs moyennes des paramètres sur l'ensemble des années.

#### 4. Conclusion :

Cette étude a démontré que les méthodes récentes de la génétique des populations peuvent être utilisées sous certaines conditions comme outil de suivi des populations de saumon. Une estimation précise du nombre de reproducteurs efficaces ( $N_b$ ) peut être obtenue à partir d'un échantillon de taille modéré de juvéniles de l'année collectés lors des pêches à l'électricité effectuées pour suivre la productivité des populations. Pour

certaines systèmes, la possession d'échantillons de tissus biologiques historiques en nombre suffisant permettra d'effectuer un suivi rétrospectif grâce aux méthodes développées lors de cette étude. Les méthodes de génétique des populations mises en œuvre font appel à des techniques robustes, matures et peu coûteuses ainsi qu'à des analyses de données faisant appel à des logiciels libres, bien documentés et accessibles. L'ensemble de ces méthodes est détaillé en Annexes (voir Annexes 1 et 2) afin qu'elles puissent être facilement reproduites et utilisables par les utilisateurs ultérieurs dans la cadre de la gestion des populations.

Ce projet a exploré et quantifié les conditions méthodologiques favorables pour des estimations non biaisées de  $N_b$ . Nous avons mis l'accent sur les facteurs liés à l'échantillonnage qui se sont avérés importants à prendre en compte pour interpréter correctement les résultats obtenus. **Indépendamment de la taille de la population, des juvéniles de l'année, échantillonnés avec une pression de pêche constante en plusieurs points couvrant l'ensemble de l'étendue spatiale d'un bassin versant, permet d'estimer le  $N_b$  de façon fiable pour peu que la taille de l'échantillon soit équivalente au nombre d'adultes anadromes présents dans la population.** Certaines limites à l'estimation de  $N_b$  ont clairement été identifiées, en particulier dans les cas de population ouverte ou d'étude focalisée sur un affluent d'un système plus grand. Dans ces cas, le nombre de juvéniles à échantillonner peut-être plus grand afin d'obtenir des estimations de  $N_b$  fiables et les interprétations des estimations doivent tenir compte des potentiels mouvements des reproducteurs et des juvéniles entre l'affluent et les autres parties du système, complexifiant grandement les interprétations en fonction du degré d'ouverture ou d'isolement de la zone d'étude.

$N_b$  estimé à partir des génotypes des juvéniles de l'année est un paramètre intégré qui inclut de l'information sur le système de reproduction (nombre de reproducteurs, succès reproducteurs et degrés de polygamie) et la survie des juvéniles produits lors d'une saison de reproduction. Les effets liés à des reproducteurs furtifs non observables ou quantifiables (adultes anadromes ou reproducteurs précoces) sont intégrés dans ce paramètre estimable à partir d'un échantillonnage de juvéniles de l'année. L'information apportée par ce paramètre rend donc compte de phénomènes biologiques liés à une saison de reproduction et à la survie lors des premiers mois de développement des juvéniles correspondants. Ces informations sur des pas de temps très immédiat et contemporains de processus écologiques se déroulant à l'échelle populationnel, devrait rencontrer un intérêt auprès des gestionnaires en charge du suivi des populations, par exemple pour évaluer l'impact à court terme d'actions de gestion (amélioration de l'habitat, actions de soutiens au stock...) ou pour suivre le statut de conservation des populations à plus long terme, en particulier lorsque aucune autre information démographique n'est disponible.

Cependant, une compréhension plus poussée des processus écologiques est possible en combinant l'estimation de  $N_b$  à d'autres paramètres populationnels tels que les paramètres démographiques comme nous l'avons illustré sur la population de saumons de la Nivelle. En effet, comme le nombre de reproducteurs efficaces est lié à la taille efficace de la population, ce paramètre complète les paramètres démographiques en apportant de l'information sur les processus éco-évolutifs en œuvre au sein des populations, sur leurs capacités adaptatives et leurs persurances à long terme.

## 7. Bibliographie

Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C. A., Garner, B. A., Mccane, J. & Campbell, M. R. (2016) Effective Number of Breeders from Sibship Reconstruction: Empirical Evaluations Using Hatchery Steelhead. *Evolutionary Applications* **10**, 146–160.

Bacles, C. F. E., Bouchard, C., Lange, F., Manicki, A., Tentelier, C. & Lepais, Olivier. (2018) Estimating the Effective Number of Breeders from Single Parr Samples for Conservation Monitoring of Wild Populations of Atlantic Salmon *Salmo Salar*. *Journal of Fish Biology* **In press**.

Bailleul, D., Stoeckel, S. & Arnaud-Haond, S. (2016) RClone: A Package to Identify MultiLocus Clonal Lineages and Handle Clonal Data Sets in R. *Methods in Ecology and Evolution* **7**, 966–970.

Bouchard, C., Bardonnnet, A., Buoro, M., Prévost, E. & Tentelier, C. (2018) Effects of Spatial Aggregation of Nests on Population Recruitment: The Case of a Small Population of Atlantic Salmon. **In prep**.

Chaput, G. (2012) Overview of the Status of Atlantic Salmon (*Salmo Salar*) in the North Atlantic and Trends in Marine Mortality. *ICES Journal of Marine Science* **69**, 1538–1548.

Le cor, F., Montorio, L., Nevoux, M. & Evanno, G. (2016) Effets Des Stratégies D'histoire de Vie Sur Le Succès Reproducteur Du Saumon Atlantique. *Rapport de Master 2 Ecologie Fonctionnelle Comportementale et Evolutive, Université de Rennes 1, INRA UMR ESE*. 38 p.

Dodson, J. J., Gibson, R. J., Cunjak, R. A., Friedland, K. D., Garcia de Leaniz, C., Gross, M. R., Newbury, R., Nielsen, J. L., Power, M. E. & Roy, S. (1998) Elements in the Development of Conservation Plans for Atlantic Salmon (*Salmo Salar*). *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 312–323.

Ferchaud, A.-L., Perrier, C., April, J., Hernandez, C., Dionne, M. & Bernatchez, L. (2016) Making Sense of the Relationships between  $N_e$ ,  $N_b$  and  $N_c$  towards Defining Conservation Thresholds in Atlantic Salmon (*Salmo Salar*). *Heredity* **117**, 268–278.

Frankham, R. (1995) Effective Population Size/adult Population Size Ratios in Wildlife: A Review. *Genetical Research* **66**, 95.

Garant, D., Dodson, J. J. & Bernatchez, L. (2001) A Genetic Evaluation of Mating System and Determinants of Individual Reproductive Success in Atlantic Salmon (*Salmo Salar* L.). *Journal of Heredity* **92**, 137–145.

de Gaudemar, B., Schroder, S. L. & Beall, E. P. (2000) Nest Placement and Egg Distribution in Atlantic Salmon Redds. *Environmental Biology of Fishes* **57**, 37–47.

Gauthey, Z., Freychet, M., Manicki, A., Herman, A., Lepais, O., Panserat, S., Elozegi, A., Tentelier, C. & Labonne, J. (2015) The Concentration of Plasma Metabolites Varies throughout Reproduction and Affects Offspring Number in Wild Brown Trout (*Salmo Trutta*). *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* **184**, 90–96.

Grimardias, D., Merchermek, N., Manicki, A., Garnier, J., Gaudin, P., Jarry, M. & Beall, E. (2010) Reproductive Success of Atlantic Salmon (*Salmo Salar*) Mature Male Parr in a Small River, the Nivelle: Influence of Shelters. *Ecology of Freshwater Fish* **19**, 510–519.

Johnson, P. C. D. & Haydon, D. T. (2007) Maximum-Likelihood Estimation of Allelic Dropout and False Allele Error Rates from Microsatellite Genotypes in the Absence of Reference Data. *Genetics* **175**, 827–842.

Jones, O. R. & Wang, J. (2010) COLONY: A Program for Parentage and Sibship Inference from Multilocus Genotype Data. *Molecular Ecology Resources* **10**, 551–555.

Kennedy, R. J., Johnston, P. & Allen, M. (2014) Assessment of a Catchmentwide Salmon Habitat Rehabilitation Scheme on a Drained River System in Northern Ireland. *Fisheries Management and Ecology* **21**, 275–287.

Klemetsen, A., Amundsen, P.-A., Dempson, J. B., Jonsson, B., Jonsson, N., O'Connell, M. F. & Mortensen, E. (2003) Atlantic Salmon *Salmo Salar* L., Brown Trout *Salmo Trutta* L. and Arctic Charr *Salvelinus Alpinus* (L.): A

Review of Aspects of Their Life Histories. *Ecology of Freshwater Fish* **12**, 1–59.

Lande, R. (1993) Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. *The American Naturalist* **142**, 911–927.

Lange, F., Carlut, E., Prévost, E. & Servanty, S. (2015) *Les Populations de Saumons, Truites de Mer et Grandes Aloses de La Nivelle En 2014*. Scientific report INRA. Saint Pée sur Nivelle.

Lange, F., Guéraud, F., Huchet, E., Rives, J. & Prévost, E. (2017) *Abundances Indices and Biological Traits of Juvenile Salmon Sampled in the Nivelle River (France). Version 1.7. Institut National de La Recherche Agronomique (INRA). Occurrence Dataset <https://doi.org/10.15468/alsjvy> Accessed via GBIF.org on 2017-12-08.*

Luikart, G., Ryman, N., Tallmon, D. a., Schwartz, M. K. & Allendorf, F. W. (2010) Estimation of Census and Effective Population Sizes: The Increasing Usefulness of DNA-Based Approaches. *Conservation Genetics* **11**, 355–373.

Marchand, F., Beaulaton, L., Prévost, E., Delanoë, R., Destouches, J.-P., Gueraud, F., Guilloux, Y., Jeannot, N., Huchet, E., Lange, F., et al. (2017) Abundance Indices and Biological Traits of Juvenile Salmon (*Salmo Salar*) Sampled in Three Rivers on the Atlantic and Channel Coasts (France). *Biodiversity Data Journal* **5**, e15125.

Ozerov, M., Jürgenstein, T., Aykanat, T. & Vasemägi, A. (2015) Use of Sibling Relationship Reconstruction to Complement Traditional Monitoring in Fisheries Management and Conservation of Brown Trout. *Conservation Biology* **29**, 1164–1175.

Palstra, F. P. & Fraser, D. J. (2012) Effective/census Population Size Ratio Estimation: A Compendium and Appraisal. *Ecology and Evolution* **2**, 2357–2365.

Palstra, F. P., O’Connell, M. F. & Ruzzante, D. E. (2009) Age Structure, Changing Demography and Effective Population Size in Atlantic Salmon (*Salmo Salar*). *Genetics* **182**, 1233–1249.

Parrish, D. L., Behnke, R. J., Gephard, S. R., McCormick, S. D. & Reeves, G. H. (1998) Why Aren’t There More Atlantic Salmon (*Salmo Salar*)? *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 281–287.

Perrier, C., April, J., Cote, G., Bernatchez, L. & Dionne, M. (2016) Effective Number of Breeders in Relation to Census Size as Management Tools for Atlantic Salmon Conservation in a Context of Stocked Populations. *Conservation Genetics* **17**, 31–44.

Prévost, E. & Baglinière, J. L. (1995) Présentation et Premiers Éléments de Mise Au Point D’une Méthode Simple D’évaluation En Eau Courante, Du Recrutement En Juvéniles de Saumon Atlantique (*Salmo Salar*) de L’année. In *Les recherches françaises en évaluation quantitative et modélisation des ressources et des systèmes halieutiques : actes du colloque* (Gascuel, D., Durand, J. L., Fonteneau, A., eds), pp. 39–48 Paris: ORSTOM Editions.

R Development Core Team. (2016) *R: A Language and Environment for Statistical Computing*, R Foundati. Vienna, Austria.

Reed, D. H. (2008) Effects of Population Size on Population Viability: From Mutation to Environmental Catastrophes. In *Conservation biology: evolution in action* (Carroll, S. P., W, F. C., eds), pp. 16–34.

Richard, A., Dionne, M., Wang, J. & Bernatchez, L. (2013) Does Catch and Release Affect the Mating System and Individual Reproductive Success of Wild Atlantic Salmon (*Salmo Salar* L.)? *Molecular Ecology* **22**, 187–200.

Simberloff, D., Farr, J. A., Cox, J. & Mehlman, D. W. (1992) Movement Corridors: Conservation Bargains or Poor Investments? *Conservation Biology* **6**, 493–504.

Taggart, J. B., McLaren, I. S., Hay, D. W., Webb, J. H. & Youngson, A. F. (2001) Spawning Success in Atlantic Salmon (*Salmo Salar* L.): A Long-Term DNA Profiling-Based Study Conducted in a Natural Stream. *Molecular Ecology* **10**, 1047–1060.

Tallmon, D. A., Gregovich, D., Waples, R. S., Scott Baker, C., Jackson, J., Taylor, B. L., Archer, E. R. I. C., Martien, K. K., Allendorf, F. W. & Schwartz, M. K. (2010) When Are Genetic Methods Useful for Estimating Contemporary Abundance and Detecting Population Trends? *Molecular Ecology Resources* **10**, 684–692.

- Tentelier, C., Lepais, O., Larranaga, N., Manicki, A., Lange, F. & Rives, J. (2016a) Sexual Selection Leads to a Tenfold Difference in Reproductive Success of Alternative Reproductive Tactics in Male Atlantic Salmon. *The Science of Nature* **103**, 47.
- Tentelier, C., Larranaga, N., Lepais, O., Manicki, A., Rives, J. & Lange, F. (2016b) Space Use and Its Effects on Reproductive Success of Anadromous Atlantic Salmon. *Canadian Journal of Fisheries and Aquatic Sciences* **73**, 1461–1471.
- Vøllestad, L. A., Skurdal, J. & L'Abée-Lund, J. H. (2014) Evaluation of a New Management Scheme for Norwegian Atlantic Salmon *Salmo Salar*. *Fisheries Management and Ecology* **21**, 133–139.
- Wang, J. (2009) A New Method for Estimating Effective Population Sizes from a Single Sample of Multilocus Genotypes. *Molecular Ecology* **18**, 2148–2164.
- Wang, J. (2013) A Simulation Module in the Computer Program Colony for Sibship and Parentage Analysis. *Molecular Ecology Resources* **13**, 734–739.
- Wang, J. (2016) A Comparison of Single-Sample Estimators of Effective Population Sizes from Genetic Marker Data. *Molecular Ecology* **25**, 4692–4711.
- Wang, J. (2017) Estimating Genotyping Errors from Genotype and Reconstructed Pedigree Data. *Methods in Ecology and Evolution* **Early View**.
- Wang, J., Santiago, E. & Caballero, A. (2016) Prediction and Estimation of Effective Population Size. *Heredity* **117**, 1–14.
- Waples, R. S. (2002) Definition and Estimation of Effective Population Size in the Conservation of Endangered Species. In *Population viability analysis* (Beissinger, S. R., McCullough, D. R., eds), pp. 147–168 Chicago.
- Waples, R. S. (2005) Genetic Estimates of Contemporary Effective Population Size: To What Time Periods Do the Estimates Apply? *Molecular Ecology* **14**, 3335–3352.
- Waples, R. S. (2016) Making Sense of Genetic Estimates of Effective Population Size. *Molecular Ecology* **25**, 4689–4691.
- Weir, L. K., Breau, C., Hutchings, J. A. & Cunjak, R. A. (2010) Multiple Paternity and Variance in Male Fertilization Success within Atlantic Salmon *Salmo Salar* Redds in a Naturally Spawning Population. *Journal of Fish Biology* **77**, 479–493.
- Whiteley, A. R., Coombs, J. A., Cembrola, M., O'Donnell, M. J., Hudy, M., Nislow, K. H. & Letcher, B. H. (2015) Effective Number of Breeders Provides a Link between Interannual Variation in Stream Flow and Individual Reproductive Contribution in a Stream Salmonid. *Molecular Ecology* **24**, 3585–3602.
- Wright, S. (1931) Evolution in Mendelian Populations. *Genetics* **16**, 97–159.



## 8. Table des illustrations

Figure 1. Relation entre effectif de juvéniles, nombre de parents et nombre de reproducteurs efficaces. ....	9
Figure 2. Carte de la Nivelle représentant la localisation des passe-pièges et des points d'échantillonnage de juvéniles (adapté de Bacles et al., 2018). ....	12
Figure 3. Sous-échantillonnage des géotypes de juvéniles 0+ pour trois années sur la Nivelle (a-c) et des géotypes de juvéniles simulées selon un scénario de 57 reproducteurs efficaces (d). ....	19
Figure 4. Sous-échantillonnage du nombre de géotypes de juvéniles 0+ utilisé pour estimer Nb dans le ruisseau de La Roche en 2015 (effectué à partir des données géotypiques fournies par Lucie Montorio, François Le cor, Marie Nevoux et Guillaume Evanno). ....	24
Figure 5. Variation au cours de 10 saisons de reproduction consécutives de la population de saumons de la Nivelle des paramètres populationnelles. ....	27
Tableau 1. Estimations de Nb aux cours de 10 années chez les saumons de la Nivelle en utilisant les géotypes des juvéniles 0+ ou en incluant les géotypes des adultes anadromes. ....	18
Tableau 2. Effet de la taille de l'échantillons de juvéniles 0+ de la Nivelle sur l'estimation de Nb. ....	20
Tableau 3. Effet de l'étendu spatial de la Nivelle échantillonnée sur l'estimation de Nb. ....	21
Tableau 4. Effet de la saison d'échantillonnage sur l'estimation de Nb. ....	22
Tableau 5. Application des recommandations de taille d'échantillon à trois années sur le Scorff présentant un nombre d'adultes anadromes contrasté. ....	22

## 9. Annexe 1. Guide méthodologique : Recommandations pour l'échantillonnage dans le but d'estimer le nombre de reproducteurs efficaces chez les poissons à partir des pêches à l'électricité de juvéniles de l'année.

L'objectif principal de l'échantillonnage de juvéniles de l'année dans le but d'estimer le nombre de reproducteurs efficaces ( $N_b$ ) est d'obtenir suffisamment de juvéniles 0+ sur l'ensemble de la population étudiée dont les relations de fratrie sont représentatives de celles de la population entière. En effet,  $N_b$  est estimé grâce aux probabilités que deux juvéniles de l'échantillon soient demi-frères ou plein-frères ou non apparentés. Il est donc essentiel que ces probabilités d'apparement dans l'échantillon soient les plus représentatives possibles de ces mêmes probabilités à l'échelle de la population entière. Cette étude a permis d'identifier certains points clés, présentés de façon synthétique dans cette Annexe, afin de maximiser la représentativité de l'échantillon et de minimiser les biais liés à l'échantillonnage.

### **Effort de pêche constant**

Un certain nombre de sites le long du cours d'eau principal et ses affluents (au minimum un site tous les 3 km de linéaire) doivent être échantillonnés afin d'avoir une bonne couverture de l'ensemble de la population étudiée. Le facteur le plus important à prendre en compte est l'effort de pêche qui doit être identique sur l'ensemble des sites échantillonnés afin que la taille de l'échantillon soit proportionnelle à l'abondance en chaque point. Le protocole suivi pour estimer l'abondance de juvéniles par la méthode d'indice d'abondance par capture par unité d'effort (CPUE) et qui consiste en une pêche à l'électricité de 5 minutes pratiquée uniquement sur les radiers rapides, habitats préférentiels des juvéniles, est un bon exemple de pêche à effort de capture constant (Prévost & Baglinière 1995) et est tout à fait approprié pour estimer  $N_b$  comme le montre cette étude. Tout autre protocole de pêche standardisé permettant d'assurer un effort de pêche constant peut également convenir. À l'opposé, un protocole visant à échantillonner un nombre prédéfini de juvéniles sur tous les sites de pêche est inapproprié car dans ce cas, l'effort de pêche sera plus important sur les sites moins favorables, augmentant artificiellement le nombre de relations de fratries dans ces sites et créant une distorsion des probabilités d'apparement de l'échantillon par rapport aux probabilités d'apparement dans la population générale.

### **Nombre de juvéniles à échantillonner**

Cette étude a montré que, pour une population fermée, le nombre de juvéniles à échantillonner devait être du même ordre de grandeur que le nombre supposé d'adultes anadromes présents dans la population. Trop peu de juvéniles va conduire à une sous-estimation de  $N_b$ , alors que trop de juvéniles (à partir de 5 fois le nombre d'adultes anadromes environ) peut entraîner une surestimation de  $N_b$ . En général, il est préférable de cibler une gamme de nombre de juvéniles allant du nombre d'adultes anadrome estimé à 5 fois le nombre d'adultes anadromes estimé. Dans cette plage de grandeur, l'estimation de  $N_b$  devrait être fiable tout en optimisant le coût des analyses génétiques.

### **Etendue spatiale des sites échantillonnés**

Il est important de bien répartir les sites échantillonnés sur l'ensemble du réseau hydrographique contenant des habitats favorables aux juvéniles de saumon. Par exemple, pour un petit système tel que la Nivelle (fleuve de 39 km de long et bassin versant de 238 km<sup>2</sup> ; environ 25 km du fleuve étant accessible aux saumon, Figure A1), une vingtaine de sites d'échantillonnage ont donné de bons résultats lors de cette étude.

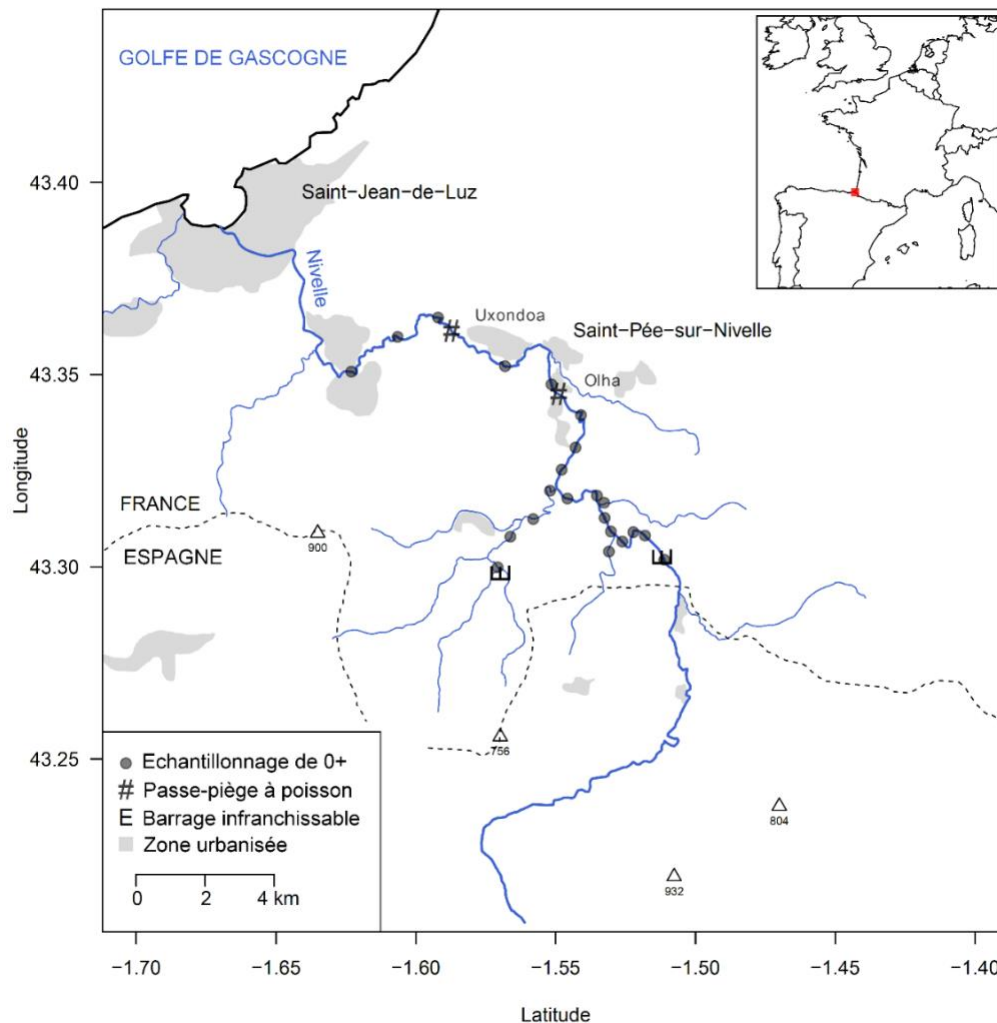


Figure A1. Carte de la Nivelle représentant la localisation des passe-pièges et des points d'échantillonnage de juvéniles (adapté de Bacles et al., 2018).

Pour les populations plus grandes, telle que celle du Scorff par exemple (75 km de long dont 10 km d'estuaire pour un bassin versant de 480 km<sup>2</sup>), une cinquantaine de sites d'échantillonnages dont 39 localisés sur le cours principal ont également donné des résultats satisfaisants lors de cette étude. Par contre, l'étude restreinte à un affluent ou à une partie d'un plus grand système pose d'importantes difficultés d'interprétation liées au fait que les probabilités d'apparement entre individus sont influencées par les mouvements d'entrées et de sorties de la zone étudiée des reproducteurs pendant la reproduction et des juvéniles lors des premiers de mois de vie. Les résultats que nous avons obtenus sur La Roche affluent de l'Oir montrent que même pour une population de petite taille (quelques dizaines d'adultes anadromes sur un linéaire de 2 km), un grand nombre de juvéniles (550) ne suffit pas à estimer convenablement  $N_b$ . En fait, vu les courtes distances en jeu dans cet exemple et le grand nombre d'adultes anadromes présents à proximité de La Roche sur l'ensemble du système hydrographique, le  $N_b$  estimé ne reflète pas seulement le fonctionnement de la population présente dans La Roche mais également une partie de ce qui se passe en dehors, dans le reste du système hydrographique mais de façon incomplètes car l'échantillonnage n'est pas adapté. Il est cependant envisageable de s'intéresser à une zone précise d'un système hydrographique ou à un affluent donné si cette zone est relativement isolée (par une grande distance, ou des zones non favorables à l'espèce) des autres aires de vie de l'espèce sur le reste du bassin versant. Dans ce cas cependant, il faudra prendre des précautions et échantillonner un grand nombre de juvéniles afin d'effectuer une analyse préliminaire servant à dimensionner le nombre de juvéniles nécessaires pour les campagnes de suivis ultérieurs.

### Saison d'échantillonnage

$N_b$  étant estimé à partir de juvéniles, la saison d'échantillonnage a un impact important sur son estimation et son interprétation. Comme nous l'avons montré lors de ce projet,  $N_b$  traduit d'une part les appariements et le succès reproducteur des individus lors d'une saison de reproduction mais également la survie et la dispersion des juvéniles lors des premiers mois jusqu'à l'échantillonnage. Un échantillonnage de juvéniles effectué au printemps, au plus près de la reproduction, renseignera davantage sur la reproduction, alors qu'un échantillonnage à l'automne traduira plus l'état de santé global de la population (reproduction et survie). Par contre, il faut bien veiller à effectuer l'échantillonnage des juvéniles avant les périodes de migration en dehors du système (smoltification) car dans ce cas, une partie de la population n'est plus disponible pour l'analyse et estimer  $N_b$  dans ce cas n'a que peu de sens. Dans les cas de suivis temporels ou de comparaisons entre populations, il faudra veiller à échantillonner les juvéniles à la même saison, voire autour de la même date, afin de pouvoir effectuer des comparaisons qui ont un sens biologique.

### Références

Prévost, E. & Baglinière, J. L. (1995) Présentation et premiers éléments de mise au point d'une méthode simple d'évaluation en eau courante, du recrutement en juvéniles de saumon Atlantique (*Salmo Salar*) de l'année. In Les recherches françaises en évaluation quantitative et modélisation des ressources et des systèmes halieutiques : actes du colloque (Gascuel, D., Durand, J. L., Fonteneau, A., eds), pp. 39–48 Paris: ORSTOM Editions.

## 10. Annexe 2. Protocole : Procédure expérimentale pour l'analyse génétique nécessaire à l'estimation du nombre de reproducteurs efficaces dans les populations de poissons.

Cette annexe présente de façon détaillée et reproductible les différentes étapes d'analyses au laboratoire, de traitement et d'analyse des données nécessaires à l'estimation du nombre de reproducteurs efficaces ( $N_b$ ) par la méthode de reconstruction de fratries à partir d'un échantillonnage de juvéniles de l'année.

### Extraction d'ADN

L'extraction d'ADN a été réalisée à partir d'environ 0.2 cm<sup>2</sup> de morceau de nageoire conservé dans de l'éthanol absolu, mais peut également être réalisée à partir d'une quantité identique de n'importe quel autre tissu biologique conservé dans de l'éthanol absolu, séché ou congelé (écailles ou tout autre tissu).

Nous avons utilisé une version modifiée d'un protocole d'extraction d'ADN basé sur du NaCl et du chloroforme (Müllenbach et al. 1989) mais n'importe quel autre protocole d'extraction d'ADN peut être utilisé (à base de Chelex ou de kits commerciaux standard à base de colonnes échangeuses d'ions). Notre protocole a été adapté pour pouvoir être utilisé en format plaque 96 (microtubes Qiagen num. cat. 19560 et bouchons associés Qiagen num.cat. 19566) permettant d'extraire de l'ADN de haute qualité de 192 individus par jour à faible coût (Gauthey et al. 2015). Le morceau de nageoire est lysé grâce à 200 µL de tampon de lyse (75 nM NaCl, 25 mM EDTA, 1% SDS, pH 8) contenant 10 µL de protéinase K à 20 mg/µL dans un microtube 1.2 mL. Les échantillons sont ainsi incubés pendant toute la nuit à 55°C. 100 µL de NaCl à 5M sont ajoutés et les microtubes sont remués doucement avant que 300 µL de chloroforme soient ajoutés. Les tubes sont remués doucement pendant 10 minutes puis centrifugés à 2000 tpm pendant 10 minutes. La phase supérieure est alors prélevée et déposée dans un nouveau microtube. L'ADN est précipité en utilisant 250 µL d'isopropanol, et après 5 minutes de mélange, une étape de centrifugation à 4100 tpm pendant 5 minutes est effectuée. Le surnageant est enlevé en retournant rapidement le microtube (en veillant à ce que la pelote d'ADN reste collée au fond du microtube), puis une étape de lavage avec 500 µL d'éthanol à 70% est effectuée pendant une heure. Après une étape de centrifugation de 5 minutes à 4100 tpm, l'éthanol est enlevé par retournement du microtube. L'ADN est alors séché à température ambiante et finalement suspendu dans 100 µL de tampon TE 1X.

### Amplification de 15 microsatellite et d'un marqueur de sexe par PCR multiplexée

Les 16 marqueurs étudiés (Figure A2) comprennent SSsp2210, SSsp2215 and SSsp2216 (Patterson et al. 2004), Ssa85, Ssa197 and Ssa171 (O'Reilly et al. 1996), Ssa412UOS and Ssa407UOS (Cairney et al. 2000), SSOSL85 (Slettan et al. 1995), MHC1 (aïssu connu sous le nom de Sasa-UBA, Grimholt et al. 2002), EST19, EST28, EST41, EST107 and EST123 (Vasemägi et al. 2005) and Yanng400, un marqueur spécifique aux mâles (Yanng400F: AGACTGCTCATTGAAA and Yanng400R: TCAGGGAGAAGGGCTGTAG; Guigen Y., communication personnelle ; Yano et al. 2013). L'amplification PCR est réalisée dans un volume finale de 5 µL en utilisant 1X de Type-it Microsatellite PCR Master Mix (Qiagen), des concentrations variables (voir Tableau A1) de chacune des amorces sens et antisens non modifiées (Integrated DNA Technology) spécifiques aux marqueurs (Tableau A1), de la quantité appropriée de chacune des 4 séquence universelles (Blacket et al. 2012) marquées par fluorescence 6-FAM, VIC, NED et PET (Applied Biosystems) (Tableau A1 et A2) et environ 20 ng de matrice d'ADN (correspondant à 0.5 µL d'ADN non diluée). Les cycles d'amplification ont été effectués dans un thermocycleur Applied Biosystems 2720 et consistait en une étape préliminaire de 5 minute à 95°C suivie de 32 cycles comprenant 30 secondes à 95°C, 3 minutes à 59°C et 30 secondes à 72°C, puis une étape d'élongation finale de 30 minutes à 60°C.

## Génotypage et traitement des données électrophorétiques

La taille des produits d'amplification a été déterminée par analyse d'électrophorèse dans le séquenceur capillaire ABI 3730 (Applied Biosystems) de la Plateforme Génome Transcriptome de Bordeaux (PGTB) en utilisant une échelle de taille GeneScan LIZ 500 (Life Technologies). Le logiciel STRand 2.4.110 (Toonen & Hughes, 2001 ; logiciel libre disponible ici : <http://www.vgl.ucdavis.edu/informatics/strand.php>) a été utilisé pour interpréter les profils électrophorétiques manuellement puis le package R MsatAllele v 1.01 (Alberto, 2009) a servi à coder le nom des allèles par le nombre entier le plus proche correspondant à leur taille, en tenant compte du mécanisme de mutation par ajout ou suppression des unités des motifs répétés typique des marqueurs microsatellites.

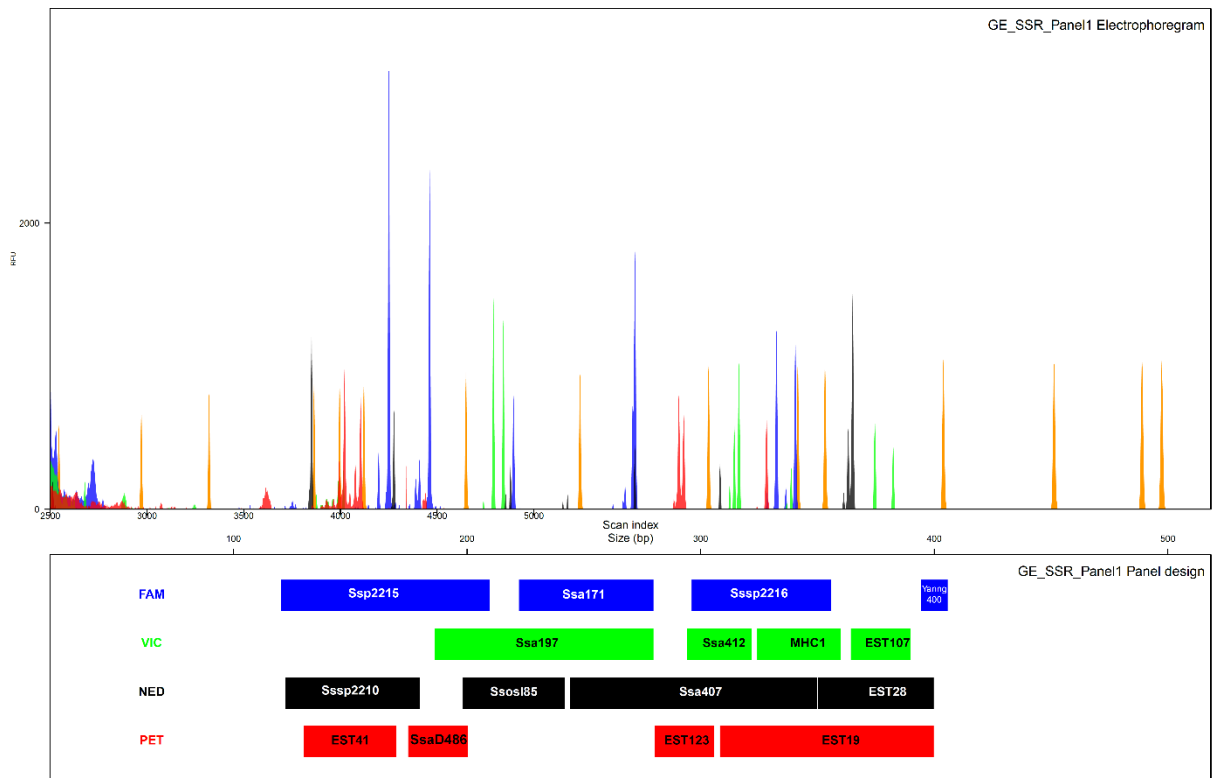


Figure A2 : Exemple de profil d'électrophérogramme obtenu pour un individu (haut) et localisation des tailles d'allèles attendues et des différents marqueurs fluorescents pour les différents marqueurs (bas). La détermination de la taille précise des allèles est effectuée grâce au logiciel STRand en comparant la distance de migration avec des fragments de tailles connues (marqueurs LIZ 500 représenté par les piques oranges).

Tableau A1 : Détail des séquences des différentes amorces utilisées et de leur concentration respective dans le mélange de réactif lors de la PCR.

Marqueur	Fluorescence de l'amorce universelle	Séquence de l'amorce universelle marquée	C. f. [µM]	Séquence de l'amorce sens non marquée	C. f. [µM]	Séquence de l'amorce antisens non marquée	C. f. [µM]	Taille des allèles attendus (pb)
<b>Sssp2215</b>	A-FAM	<b>GCCTTGCCAGCCCGCTCAG</b>	0.360	<b>GCCTTGCCAGCCCGCTCAG</b> GGTCAGTCAGTCACACCATGC	0.010	GTTTGTCACTAGCCAGGTGTCC	0.020	[129-203]
<b>Ssa171</b>	A-FAM	<b>GCCTTGCCAGCCCGCTCAG</b>		<b>GCCTTGCCAGCCCGCTCAG</b> GAGGTCGCTGGGGTTTACTAT	0.025	TTATTATCCAAAGGGGTCAAAA	0.050	[217-280]
<b>Sssp2216</b>	A-FAM	<b>GCCTTGCCAGCCCGCTCAG</b>		<b>GCCTTGCCAGCCCGCTCAG</b> CTCCTCTGGGATTTCTGTCA	0.075	GTTTCTGGAGCAGAGGATTGCTG	0.150	[292-312]
<b>Yanng400</b>	A-FAM	<b>GCCTTGCCAGCCCGCTCAG</b>		<b>GCCTTGCCAGCCCGCTCAG</b> TGCGAGA	0.250	TCAGGGAGAGCTTGTAG	0.500	[395-405]
<b>Ssa197</b>	B-VIC	<b>GCCTCCCTCGCGCCA</b>	0.425	<b>GCCTCCCTCGCGCCA</b> TGGCAGGGATTTGACATAAC	0.050	GTTTGGTTGAGTAGGGAGGCTTG	0.100	[184-279]
<b>Ssa412</b>	B-VIC	<b>GCCTCCCTCGCGCCA</b>		<b>GCCTCCCTCGCGCCA</b> AGTGGAGATACACAGCACTTA	0.075	GTTTCTTGTTAGTACCGGACATG	0.150	[308-314]
<b>MHCI</b>	B-VIC	<b>GCCTCCCTCGCGCCA</b>		<b>GCCTCCCTCGCGCCA</b> AGGAGAGCTGCCAGATGACTT	0.250	GTTTCAATTACCACAAGCCCGCTC	0.500	[325-359]
<b>EST107</b>	B-VIC	<b>GCCTCCCTCGCGCCA</b>		<b>GCCTCCCTCGCGCCA</b> AGCGTTTACGTGAATCCAA	0.050	GTTTCTCATGGAGGGTGGAAAGTGT	0.100	[372-383]
<b>Sssp2210</b>	C-NED	<b>CAGGACCAGGCTACCGTG</b>	0.350	<b>CAGGACCAGGCTACCGTG</b> CCTTTTTCCAATGGGATTCA	0.075	GTTTCATGCACACACATTCACTGC	0.150	[131-176]
<b>Ssosl85</b>	C-NED	<b>CAGGACCAGGCTACCGTG</b>		<b>CAGGACCAGGCTACCGTG</b> TGTGGATTTTGTATTATGTTA	0.125	GTTTATACATTTCTCTCATTGAG	0.250	[196-245]
<b>Ssa407</b>	C-NED	<b>CAGGACCAGGCTACCGTG</b>		<b>CAGGACCAGGCTACCGTG</b> TCGTGACTACTAAGTCTTTGACCA	0.125	GTTTGTGTAGGCAGGTGTGGAC	0.250	[245-348]
<b>EST28</b>	C-NED	<b>CAGGACCAGGCTACCGTG</b>		<b>CAGGACCAGGCTACCGTG</b> CACAGGCACACACTCCTCAT	0.025	GTTTCAGGTGAAGAGCATGACCAA	0.050	[363-377]
<b>EST41</b>	D-PET	<b>CGGAGAGCCGAGAGGTG</b>	0.325	<b>CGGAGAGCCGAGAGGTG</b> TGCAAGTAAAGGCAGGGTTT	0.050	GTTTGTGGTAGGATTGGGGTTCTCT	0.100	[142-178]
<b>SsaD486</b>	D-PET	<b>CGGAGAGCCGAGAGGTG</b>		<b>CGGAGAGCCGAGAGGTG</b> TCGCTGTGTATCAGTATTTGG	0.025	ACTCGGATAAACTCACAGGTC	0.050	[179-194]
<b>EST123</b>	D-PET	<b>CGGAGAGCCGAGAGGTG</b>		<b>CGGAGAGCCGAGAGGTG</b> GCGGCCCTTAGTGAATCAA	0.125	GTTTCTCGCCAGTCACTCTTCAA	0.250	[287-302]
<b>EST19</b>	D-PET	<b>CGGAGAGCCGAGAGGTG</b>		<b>CGGAGAGCCGAGAGGTG</b> CGCTTCTGGACAAAATTA	0.125	GTTTGAGCACACCCATTCTCA	0.250	[314-383]

Seules 4 amorces universelles A, B, C et D chacune marquée par un fluorochrome spécifique sont utilisées afin d'effectuer le marquage par fluorescence de l'ensemble des marqueurs permettant de réduire les coûts des réactifs. Afin que les bons marqueurs soient marqués par les amorces universelles, la séquence universelle voulue a été rajoutée en amont des séquences sens spécifique à chaque marqueur (indiqué en gras coloré). Les amorces sens étant en concentration moindre que les amorces sens, elles vont devenir minoritaires voire s'épuiser pendant la réaction ce qui va conduire en l'incorporation progressive des amorces universelles marquées voulues à chaque marqueur amplifié. C. f. : concentration finale, pb : paire de bases.

## Analyse des données

### Identification des recaptures

Une détection des génotypes multilocus identiques doit être effectuée afin d'identifier d'éventuelles individus capturés plusieurs fois. Cette analyse est réalisée à l'aide du package R Rclone (Bailleul et al., 2016) qui permet de prendre en compte les erreurs de génotypage pour la détection des clones (dans notre cas, individus capturés plusieurs fois). Lorsque qu'un individu a été détecté comme capturé plusieurs fois, seul un génotype multilocus de l'individu correspondant est conservé pour l'analyse ultérieure. Il faut noter que cette détection d'individus recapturés peut également être réalisée en utilisant le logiciel Colony (Jones & Wang 2010).

### Estimation de $N_b$ par Colony

Colony (Jones & Wang, 2010, logiciel libre et disponible ici : <https://www.zsl.org/science/software/colony>) est utilisé pour estimer  $N_b$  par la méthode basée sur la reconstruction des fratries à partir des génotypes de juvéniles de l'année. La méthode d'analyse utilisée doit être la Full-Likelihood (FL) avec les paramètres suivants : Female Polygamy, Male Polygamy, Without Inbreeding, Without Clone, Dioecious, Diploid, Medium Length of Run, Medium Likelihood Precision, No Update of Allele Frequency, Sibship Scaling Yes, Number of Run 1, Random Number Seed : nombre aléatoire à changer à chaque analyse, No Sibship Prior (Figure A3; voir le mode d'emploi du logiciel pour les détails sur la signification de ces paramètres).

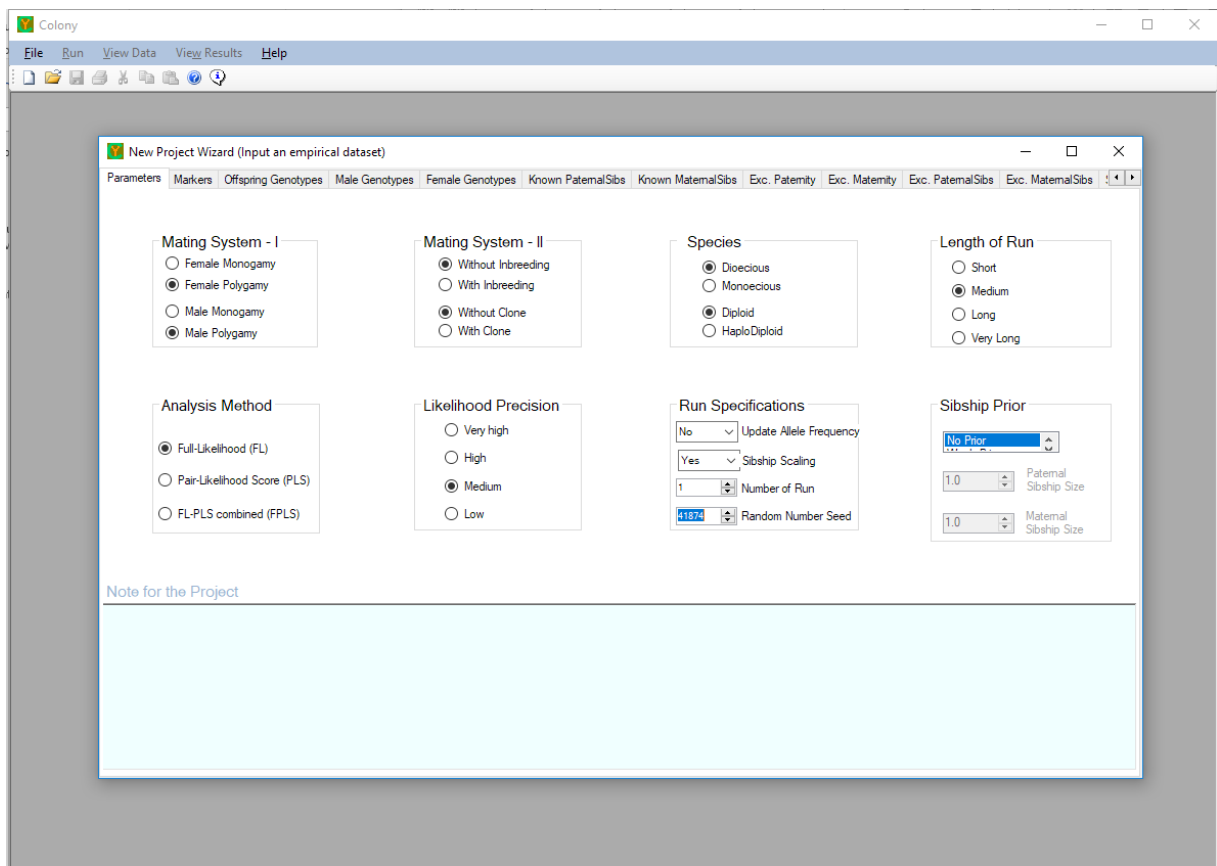


Figure A3 : Capture d'écran des options utilisées pour faire les analyses dans Colony.

À l'issue de l'analyse, les résultats sont présentés sous la forme de nombreux fichiers. Celui qui nous intéresse principalement est le fichier se terminant par « .Ne » qui est un fichier texte ouvrable avec n'importe quel traitement de texte ou bloc note et contenant l'estimation de  $N_e$  (représentant  $N_b$  dans notre cas).

### Validation des estimations de $N_b$

Il est impératif de réaliser trois analyses différentes en changeant seulement le *Random Number Seed* entre les analyses. Il faut en effet s'assurer que les analyses aient convergé vers le résultat le plus probable ce qui se vérifie



par une stabilité des résultats entre les analyses. Si les résultats varient d'une analyse à l'autre, il faut recommencer trois autres analyses plus longues pour laisser plus de temps au modèle de converger (paramètre *Lengh of Run* sur Long).

Afin de s'assurer qu'un nombre suffisant de juvéniles est disponible pour estimer  $N_b$  sans biais, il est possible d'estimer à nouveau  $N_b$  en utilisant un plus petit nombre de juvéniles dans l'analyse. Par exemple, les résultats peuvent être validés si  $N_b$  estimé avec un sous-échantillonnage aléatoire de 80% des juvéniles disponibles est similaire à  $N_b$  estimé avec l'ensemble du jeu de données. L'analyse peut être également effectuée avec un quelques nombres croissants de juvéniles (par exemple 25, 50, 75, 100 juvéniles, sans pour autant faire une analyse aussi exhaustive de celles réalisées lors de ce projet) afin d'observer l'atteinte d'un plateau de valeur de  $N_b$  indiquant que l'échantillonnage est saturant pour l'estimation de  $N_b$ .

## Références :

Alberto F (2009) MsatAllele\_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity* 100:394–397.

Blacket MJ, Robin C, Good RT, Lee SF, Millers AD (2012) Universal primers for fluorescent labelling of PCR fragments - an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources* 12:456-463.

Cairney M, Taggart JB, Høyheim B (2000). Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. *Molecular Ecology* 9:2175-2178.

Gauthey Z, Freychet M, Manicki A, Herman A, Lepais O, Panserat S, Elozegi A, Tentelier C, Labonne J (2015) The concentration of plasma metabolites varies throughout reproduction and affects offspring number in wild brown trout (*Salmo trutta*). *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* 184:90-96.

Grimholt U, Drabløs F, Jørgensen SM, Høyheim B, Stet RJM (2002) The major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics* 54:570-581.

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* 10:551–555.

Müllenbach R, Lagoda PJL, Welter C (1989). An efficient salt-chloroform extraction DNA from blood and tissues. *Trends in Genetics* 5:391.

O'Reilly PT, Hamilton LC, McConnell SK, Wright JM (1996) Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Canadian Journal of Fisheries and Aquatic Sciences* 53:2292-298.

Paterson S, Piertney SB, Knox D, Gilbey J, Verspoor E (2004) Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Molecular Ecology Notes* 4:160-162.

Slettan A, Olsaker I, Lie O (1996) Polymorphic Atlantic salmon, *Salmo salar* L., microsatellites at the SSOSL438, SSOSL439 and SSOSL444 loci. *Animal Genetics* 27:57-58.

Toonen RJ, Hughes S (2001) Increased throughput for fragment analysis on an ABI PRISM 377 automated sequencer using a membrane comb and STRand software. *BioTechniques* 31:1320–1324.

Vasemägi A, Nilsson J, Primmer CR (2005) Seventy-five EST-linked Atlantic salmon (*Salmo salar* L.) microsatellite markers and their cross-amplification in five salmonid species. *Molecular Ecology Notes* 5:282-288.

Yano A, Nicol B, Jouanno E, Quillet E, Fostier A, Guyomard R, Guiguen Y (2013) The sexually dimorphic on the Y-chromosome gene (sdY) is a conserved male-specific Y-chromosome sequence in many salmonids. *Evolutionary Applications* 6 :486-496.

## 11. Remerciements

Je remercie Cécile Bacles pour l'ensemble du travail qu'elle a effectué sur ce projet ; Frédérique Lange pour la mise à disposition des échantillons et des données Nivelles ; Aurélie Manicki pour les extractions d'ADN des échantillons Nivelles et les analyses génétiques du Scorff ; Colin Bouchard et Cédrique Tentelier pour les discussions sur l'interprétation des variations de  $N_b$  en lien avec les comportements reproducteurs ; Etienne Prévost pour avoir partagé des données et rapports non publiés ; Lucie Montorio, Marie Nevoux et Guillaume Evanno pour avoir partagé les géotypes des saumons de La Roche ; Nicolas Jeannot et Frédéric Marchand pour la mise à disposition des échantillons du Scorff ; et Joëlle Chat pour la relecture de ce rapport. Adeline Delcamp a fourni une assistance technique pour les géotypage et Jinling Wang de l'aide pour faire fonctionner Colony, qu'ils soient tous les deux remerciés.

Le géotypage a été effectué sur un ABI 3730 DNA Analyzer de la Plateforme Génome Transcriptome de Bordeaux (financement du Conseil Régional Nouvelle-Aquitaine n°20030304002FA and 20040305003FA, de l'Union Européenne, FEDER n°2003227 et d'un Investissements d'avenir, Convention attributive d'aide N°ANR-10-EQPX-16-01). Les données et collections d'échantillons ont été fournies par l'Installation Expérimentale Ecologie Comportementale des Poissons (ECP) de l'UMR INRA/UPPA 1224 ECOBIOP et par l'Unité Expérimentale U3E de l'INRA. Une partie des calculs nécessaires à cette étude ont été réalisés sur le serveur de calcul du Mésocentre de Calcul Intensif Aquitain (MCIA) de l'Université de Bordeaux et de l'Université de Pau et des Pays de l'Adour. Ce projet a été financé par l'Agence Française pour la Biodiversité, dans le cadre du projet 2016-2018 A13 « NombReEff : Estimation du nombre de reproducteurs efficaces : un nouvel indicateur de l'état des populations ».

Avec le soutien financier de

**AGENCE FRANÇAISE  
POUR LA BIODIVERSITÉ**  
ÉTABLISSEMENT PUBLIC DE L'ÉTAT

[www.agence-francaise-biodiversite.fr](http://www.agence-francaise-biodiversite.fr)



[www.inra.fr](http://www.inra.fr)