

CAMPUS VÉTÉRINAIRE DE LYON

Année 2023 - Thèse n° 162

ÉTUDE DU POTENTIEL ADAPTATIF DE LA STRATÉGIE DE CROISSANCE DU LÉZARD VIVIPARE

THESE

Présentée à l'Université Claude Bernard Lyon 1
(Médecine – Pharmacie)

Et soutenue publiquement le 21 décembre 2023
Pour obtenir le titre de Docteur Vétérinaire

Par

ARAGON Anaïs

CAMPUS VÉTÉRINAIRE DE LYON

Année 2023 - Thèse n° 162

ÉTUDE DU POTENTIEL ADAPTATIF DE LA STRATÉGIE DE CROISSANCE DU LÉZARD VIVIPARE

THESE

Présentée à l'Université Claude Bernard Lyon 1
(Médecine – Pharmacie)

Et soutenue publiquement le 21 décembre 2023
Pour obtenir le titre de Docteur Vétérinaire

Par

ARAGON Anaïs

Liste des enseignants du campus vétérinaire de Lyon

Au 20 mars 2023

Pr	ABITBOL	Marie	Professeur
Dr	ALVES-DE-OLIVEIRA	Laurent	Maître de conférences
Pr	ARCANGIOLI	Marie-Anne	Professeur
Dr	AYRAL	Florence	Maître de conférences
Pr	BECKER	Claire	Professeur
Dr	BELLUCO	Sara	Maître de conférences
Dr	BENAMOU-SMITH	Agnès	Maître de conférences
Pr	BENOIT	Etienne	Professeur
Pr	BERNY	Philippe	Professeur
Pr	BONNET-GARIN	Jeanne-Marie	Professeur
Dr	BOURGOIN	Gilles	Maître de conférences
Dr	BRUTO	Maxime	Maître de conférences
Dr	BRUYERE	Pierre	Maître de conférences
Pr	BUFF	Samuel	Professeur
Pr	BURONFOSSE	Thierry	Professeur
Dr	CACHON	Thibaut	Maître de conférences
Pr	CADORE	Jean-Luc	Professeur
Pr	CALLAIT-CARDINAL	Marie-Pierre	Professeur
Pr	CHABANNE	Luc	Professeur
Pr	CHALVET-MONFRAY	Karine	Professeur
Dr	CHANOIT	Gillaume	Professeur
Dr	CHETOT	Thomas	Maître de conférences
Pr	DE BOYER DES ROCHES	Alice	Professeur
Pr	DELIGNETTE-MULLER	Marie-Laure	Professeur
Pr	DJELOUADJI	Zorée	Professeur
Dr	ESCRIOU	Catherine	Maître de conférences
Dr	FRIKHA	Mohamed-Ridha	Maître de conférences
Dr	GALIA	Wessam	Maître de conférences
Pr	GILOT-FROMONT	Emmanuelle	Professeur
Dr	GONTHIER	Alain	Maître de conférences
Dr	GREZEL	Delphine	Maître de conférences
Dr	HUGONNARD	Marine	Maître de conférences
Dr	JOSSON-SCHRAMME	Anne	Chargé d'enseignement contractuel
Pr	JUNOT	Stéphane	Professeur
Pr	KODJO	Angeli	Professeur
Dr	KRAFFT	Emilie	Maître de conférences
Dr	LAABERKI	Maria-Halima	Maître de conférences
Dr	LAMBERT	Véronique	Maître de conférences
Pr	LE GRAND	Dominique	Professeur
Pr	LEBLOND	Agnès	Professeur
Dr	LEDOUX	Dorothée	Maître de conférences
Dr	LEFEBVRE	Sébastien	Maître de conférences
Dr	LEFRANC-POHL	Anne-Cécile	Maître de conférences
Dr	LEGROS	Vincent	Maître de conférences
Pr	LEPAGE	Olivier	Professeur
Pr	LOUZIER	Vanessa	Professeur
Dr	LURIER	Thibaut	Maître de conférences
Dr	MAGNIN	Mathieu	Maître de conférences
Pr	MARCHAL	Thierry	Professeur
Dr	MOSCA	Marion	Maître de conférences
Pr	MOUNIER	Luc	Professeur

Dr	PEROZ	Carole	Maître de conférences
Pr	PIN	Didier	Professeur
Pr	PONCE	Frédérique	Professeur
Pr	PORTIER	Karine	Professeur
Pr	POUZOT-NEVORET	Céline	Professeur
Pr	PROUILLAC	Caroline	Professeur
Pr	REMY	Denise	Professeur
Dr	RENE MARTELLET	Magalie	Maître de conférences
Pr	ROGER	Thierry	Professeur
Dr	SAWAYA	Serge	Maître de conférences
Pr	SCHRAMME	Michael	Professeur
Pr	SERGEANTET	Delphine	Professeur
Dr	TORTEREAU	Antonin	Maître de conférences
Dr	VICTONI	Tatiana	Maître de conférences
Dr	VIRIEUX-WATRELOT	Dorothée	Chargé d'enseignement contractuel
Pr	ZENNER	Lionel	Professeur

Remerciements

A Monsieur Alain CALENDER

Professeur à la Faculté de Médecine de Lyon

Pour m'avoir fait l'honneur d'accepter la présidence de mon jury de thèse, mes remerciements les plus sincères.

A Monsieur Pierre DE VILLEMEREUIL

Maître de conférence à l'EPHE

Pour m'avoir guidée avec bienveillance tout au long de ce travail, pour votre pédagogie, votre disponibilité et la confiance que vous m'avez accordée, mes remerciements les plus sincères.

A Madame Marie-Laure DELIGNETTE-MULLER

Professeur en statistique à VetAgro Sup

Pour m'avoir co-encadrée pendant ce travail, pour votre engagement et votre réactivité, mes remerciements les plus sincères.

A Madame Marie ABITBOL

Professeur en génétique à VetAgro Sup

Pour avoir accepté de faire partie de mon jury de thèse et pour l'intérêt que vous avez porté à mon travail, mes remerciements les plus sincères.

A Monsieur Alexis RUTSCHMANN

Post-doctorant au CNRS

Pour l'intérêt porté à mon travail et pour vos conseils avisés, mes remerciements les plus sincères.

A Madame Murielle RICHARD

Ingénieur d'étude au CNRS

Pour avoir largement contribué à l'acquisition des données utilisées dans ce travail, mes remerciements les plus sincères.

Table des matières

LISTE DES ANNEXES	11
LISTE DES FIGURES.....	13
LISTE DES TABLEAUX.....	15
LISTE DES ABREVIATIONS ET NOTATIONS MATHÉMATIQUES.....	17
INTRODUCTION	19
PARTIE I - CADRE BIBLIOGRAPHIQUE ET METHODOLOGIQUE	21
I. A BASES THEORIQUES DE LA GENETIQUE QUANTITATIVE.....	21
I. A. 1 Notion de caractère quantitatif.....	21
I. A. 2 Premières études sur la génétique quantitative	21
I. A. 3 Le modèle infinitésimal	22
I. A. 4 Décomposition de la variance phénotypique et variance génétique additive	24
I. A. 5 Héritabilité d'un caractère quantitatif.....	25
I. A. 6 Coefficient de variation additif d'un caractère quantitatif.....	26
I. A. 7 Méthodes d'estimation de l'héritabilité (h^2)	26
I. B LE MODELE ANIMAL.....	29
I. B. 1 Le modèle animal linéaire classique	29
I. B. 2 Le modèle animal hiérarchique non-linéaire.....	33
I. B. 3 Place de l'inférence bayésienne dans l'ajustement des modèles animaux.....	36
I. C TUTORIEL SUR L'AJUSTEMENT DE MODELES ANIMAUX AVEC LE PAQUET BRMS[37] DE R[24].....	38
I. C. 1 Introduction et présentation du paquet brms[37]	38
I. C. 2 Présentation des données	38
I. C. 3 Utilisation de la fonction brm().....	40
I. C. 4 Interprétation de la sortie du modèle	44
I. D ÉTUDE DE LA SELECTION SUR DES CARACTERES QUANTITATIFS	50
I. D. 1 La surface de sélection individuelle.....	50
I. D. 2 Estimation des gradients de sélection directionnelle et stabilisante	51
I. E CONTEXTE GENERAL DE L'ÉTUDE	53
I. E. 1 Éléments sur la biologie du lézard vivipare	53
I. E. 2 Intérêt de l'étude du potentiel adaptatif de la stratégie de croissance du lézard vivipare.....	55
PARTIE II - ÉTUDE EXPERIMENTALE	59
II. A MATERIEL ET METHODES	60
II. A. 1 Suivi de la population de lézards vivipares.....	60
II. A. 2 Analyses génétiques quantitatives sur la stratégie de croissance du lézard vivipare	63
II. A. 3 Estimation des gradients de sélection.....	67
II. B RESULTATS	70

II. B. 1 Reconstitution du pédigrée	70
II. B. 2 Analyses génétiques quantitatives sur la stratégie de croissance.....	72
II. B. 3 Estimation des gradients de sélection	75
II. B. 4 Évolution de la taille des individus	78
II. C DISCUSSION.....	79
II. C. 1 Modélisation des courbes de croissance	79
II. C. 2 Reconstitution du pédigrée de la population	79
II. C. 3 Analyse des composantes de la variance	80
II. C. 4 Étude du potentiel adaptatif de la stratégie de croissance.....	81
II. C. 5 Étude des pressions de sélection au sein de la population étudiée	82
II. C. 6 Mise en perspective des résultats avec le contexte de réchauffement climatique	83
CONCLUSION.....	85
BIBLIOGRAPHIE	87
ANNEXES.....	93

Liste des annexes

Annexe 1 : Évolution de la taille des subadultes au cours de l'été	93
Annexe 2 : Vérification des hypothèses de la régression linéaire entre LMA_{SA} et LMA_N	93
Annexe 3 : Implémentation du modèle de régression linéaire entre LMA_{SA} et LMA_N sur R	94
Annexe 4 : Distributions des <i>a priori</i> des modèles ajustés selon une approche bayésienne dans la partie expérimentale.....	95
Annexe 5 : Implémentation du modèle de croissance logistique sur R	96
Annexe 6 : Implémentation du modèle animal sur la taille des nouveau-nés (LMA_N) sur R ...	98
Annexe 7: Implémentation du modèle animal hiérarchique non-linéaire sur la stratégie de croissance sur R.....	100
Annexe 8 : Implémentation du modèle d'estimation des gradients de sélection sur la taille à la naissance (LMA_N) des femelles sur R	103
Annexe 9 : Implémentation du modèle d'estimation des coefficients de sélection sur les paramètres l_{max} et k des femelles sur R	104

Liste des figures

Figure 1 : Résultat d’une étude de East menée en 1911 sur l’hérédité de la longueur des épis de maïs	22
Figure 2 : Distribution du phénotype pour des caractères mesurables codés par un, deux, cinq ou 50 locus indépendants et équivalents	24
Figure 3 : Lien entre l’héritabilité d’un caractère et son potentiel évolutif	26
Figure 4 : Représentation graphique du modèle de croissance logistique	35
Figure 5 : Extrait du jeu de données utilisé dans la partie expérimentale.	39
Figure 6 : Extrait du pédigrée utilisé dans la partie expérimentale.	39
Figure 7 : Extrait de la Matrice A utilisée dans la partie expérimentale.	40
Figure 8 : Implémentation d’un modèle animal classique avec le paquet brms[37].....	43
Figure 9 : Implémentation d’un modèle animal hiérarchique non-linéaire avec brms[37].....	44
Figure 10 : Densité et trace de quelques paramètres du modèle animal sur LMA _N	46
Figure 11 : Sortie d’un modèle animal ajusté selon une approche bayésienne.....	48
Figure 12 : Estimation de la distribution <i>a posteriori</i> de h^2 et CV _A de LMA _N	49
Figure 13 : Distribution mondiale du lézard vivipare (<i>Zootoca vivipara</i>) selon l’IUCN[51].....	53
Figure 14 : Photographies de lézards vivipares.....	54
Figure 15 : Niveau de vulnérabilité de 130 populations européennes de lézard vivipare selon Bestion et al. 2015.....	56
Figure 16 : Impact de la stratégie de croissance du lézard vivipare sur la dynamique des populations	57
Figure 17 : Chronologie du suivi de la population étudiée	60
Figure 18 : Code utilisé pour le marquage de la population de lézards vivipares	62
Figure 19 : Estimation des gradients de sélection β et γ sur les paramètres k et l_{max} pour les mâles et les femelles	69
Figure 20 : Profondeur du pédigrée entre 2000 et 2019	71
Figure 21 : Représentation graphique des données et des modèles de croissance pour les femelles (à gauche) et les mâles (à droite)	72

Figure 22 : Décomposition de la variance des variables de croissance	74
Figure 23 : Héritabilités et coefficients de variation additifs des variables de croissance	75
Figure 24 : Distributions <i>a posteriori</i> des gradients de sélection sur la taille à la naissance pour les femelles et les mâles	76
Figure 25 : Distributions <i>a posteriori</i> des gradients de sélection sur la taille au stade subadulte pour les femelles et les mâles.....	76
Figure 26 : Surface de sélection individuelle de la taille au stade subadulte chez les mâles ..	77
Figure 27 : Distributions <i>a posteriori</i> des gradients de sélection sur le taux de croissance et la taille asymptotique pour les femelles et les mâles.....	77
Figure 28 : Évolution de LMA_{SA} chez les femelles et les mâles entre 1985 et 2019.	78
Figure 29 : Possibles réponses adaptatives de la stratégie de croissance du lézard vivipare dans un contexte d'élévation des températures.	84

Liste des Tableaux

Tableau I : Valeur prédictive des modèles de croissance.....	73
Tableau II : Décomposition de la variance des variables de croissance et intervalles de crédibilités à 95%.	73

Liste des abréviations et notations mathématiques

β : Gradients de sélection directionnelle

γ : Gradients de sélection stabilisante (ou divergente)

CV_A : Coefficient de Variations Additif

ESS : Taille effective d'échantillon, abréviation de « Effective Sample Size »

h^2 : Héritabilité

k : Taux de croissance

LMA : Distance entre le museau et le cloaque, abréviation de « Longueur Museau-Anus »

LMA_N : Taille des individus à la naissance

LMA_{SA} : Taille des individus au stade subadulte

I_{max} : Taille asymptotique des individus

Matrice A : Matrice d'Apparentement

MCMC : Algorithmes de Monte Carlo par Chaîne de Markov

Rhat : Statistique de Gelman et Rubin

SSI : Surface de Sélection Individuelle

V_A : Variance génétique Additive

V_{AN} : Variance associée à l'effet de l'Année de Naissance

V_E : Variance Environnementale

V_M : Variance associée à l'effet Maternel

V_{NA} : Variance génétique Non-Additive

V_P : Variance Phénotypique

V_{PE} : Variance associée à l'effet de l'Environnement Permanent

V_R : Variance Résiduelle

Introduction

L'un des principaux objectifs de la génétique quantitative est de comprendre le déterminisme génétique des caractères quantitatifs et d'étudier leur transmission entre les générations. Les domaines d'application de cette discipline sont relativement vastes en sciences animales. À l'origine, la génétique quantitative a été développée pour améliorer les programmes de sélection dans le cadre de l'agriculture et de l'élevage, l'objectif étant d'optimiser les croisements pour maximiser la valeur génétique et phénotypique des individus. Ainsi, la génétique quantitative a joué un rôle majeur dans la sélection des lignées de production des différentes filières d'animaux de rente[1–4]. Par ailleurs, le réinvestissement des théories de génétique quantitative en biologie évolutive a suscité un engouement important à la fin du XXe siècle. Les modèles d'analyse de la variance des caractères quantitatifs ont notamment apporté des clés pour évaluer le potentiel évolutif du phénotype dans des populations naturelles[1,4,5]. Aujourd'hui, ces modèles sont largement utilisés pour évaluer le potentiel adaptatif des espèces sauvages face au changement climatique[6–9]. La génétique quantitative a également des applications dans le domaine médical, car elle permet d'étudier le déterminisme génétique de certaines maladies ayant un support polygénique. L'héritabilité de différentes affections a notamment été étudiée chez les carnivores domestiques. Nous pouvons citer la dysplasie de la hanche[10–12], du coude[13], la dermatite atopique[14], la maladie valvulaire dégénérative mitrale[15] ou les malformations vertébrales chez les races chondrodystrophiques[16] par exemple. L'héritabilité des caractères comportementaux a également fait l'objet de nombreuses études chez le chien et le chat[17–20].

Nous présentons ici une analyse des composantes de la variance de plusieurs paramètres décrivant la stratégie de croissance du lézard vivipare (*Zootoca vivipara*). Cette espèce est à la fois un modèle de reptile vivant en milieu tempéré et une espèce sentinelle d'une grande diversité d'écosystèmes frais humides[21]. Plusieurs études récentes ont mis en évidence un déclin rapide du lézard vivipare dans le sud de l'Europe, en lien avec l'élévation des températures[22,23]. Or, chez les ectothermes, la taille des individus joue un rôle clé dans la tolérance thermique. Un fort potentiel évolutif de la stratégie de croissance du lézard vivipare serait donc en faveur d'un fort potentiel adaptatif de l'espèce face au changement climatique.

La première partie de cette thèse a pour objectif d'introduire le cadre théorique et méthodologique nécessaire à la compréhension des analyses réalisées dans la partie

expérimentale. Après avoir rappelé les grands principes de la génétique quantitative, nous présentons des modèles statistiques d'analyse de la variance et nous proposons une méthode d'ajustement de ces modèles à l'aide de packages R[24] selon une approche bayésienne. Nous détaillons également un modèle de régression quadratique permettant d'étudier la pression de sélection sur la stratégie de croissance. Enfin, nous présentons l'espèce étudiée et expliquons l'intérêt de la problématique abordée.

La seconde partie de cette thèse expose les analyses statistiques que nous avons réalisées sur l'évolutivité de la stratégie de croissance du lézard vivipare. Nous nous sommes intéressés à deux variables de croissance mesurées : la longueur museau-cloaque à la naissance et au stade subadulte et deux paramètres estimés par ajustement d'un modèle de croissance logistique : le taux de croissance, et la longueur museau-cloaque asymptotique. Nous avons estimé la variance génétique additive de ces variables puis nous avons estimé leurs gradients de sélection.

Partie I - Cadre bibliographique et méthodologique

I. A Bases théoriques de la génétique quantitative

I. A. 1 Notion de caractère quantitatif

Un caractère quantitatif est une variation phénotypique mesurable qui résulte à la fois d'une structure génétique complexe et de l'hétérogénéité de l'environnement. Cette variation peut concerner des caractères morphologiques, comme la taille ou le poids par exemple, mais également d'autres aspects du phénotype comme la physiologie (pression artérielle, fréquence cardiaque...), le comportement (score évaluant le bien-être, la douleur...) ou même l'histoire de vie (date de ponte dans une colonie d'oiseaux, nombre de petits par portée...). Dans une population, la valeur d'un caractère quantitatif varie selon les individus et est associée à une distribution statistique qui est généralement d'apparence continue[5,25].

I. A. 2 Premières études sur la génétique quantitative

La génétique quantitative est l'étude de la transmission des caractères quantitatifs entre générations au sein des populations. Cette discipline s'est développée tout au long du XXe siècle et a notamment permis d'améliorer les programmes de sélection dans le cadre de l'agriculture et de l'élevage, en optimisant les croisements pour maximiser la valeur génétique et phénotypique des individus[1,4,5]. Les premiers travaux de recherche en génétique quantitative concernaient la morphologie des végétaux[1]. East a par exemple étudié l'hérédité de la longueur des épis de maïs en 1911[26], voir Figure 1. Ces travaux de recherche ont mis en évidence deux observations : (1) le phénotype moyen des descendants était égal à la moyenne du phénotype des parents et (2) si on croisait les descendants entre eux, on observait une explosion de la variance du phénotype de la génération suivante. À terme, on observait une réapparition des phénotypes extrêmes et le point (2) permettait donc une conservation de la diversité phénotypique dans la population[1,26]. À première vue, ces observations n'étaient pas cohérentes avec les lois de la génétique mendélienne sur la transmission des caractères qualitatifs. La question des mécanismes génétiques impliqués dans la transmission des caractères quantitatifs a donc largement été débattue[1]. Le modèle infinitésimal proposé par Fisher en 1918[4,27] a finalement permis de faire le lien entre la transmission des caractères quantitatifs et la génétique mendélienne. Il a remporté ainsi l'adhésion des généticiens et est encore aujourd'hui considéré comme le fondement de la génétique quantitative.

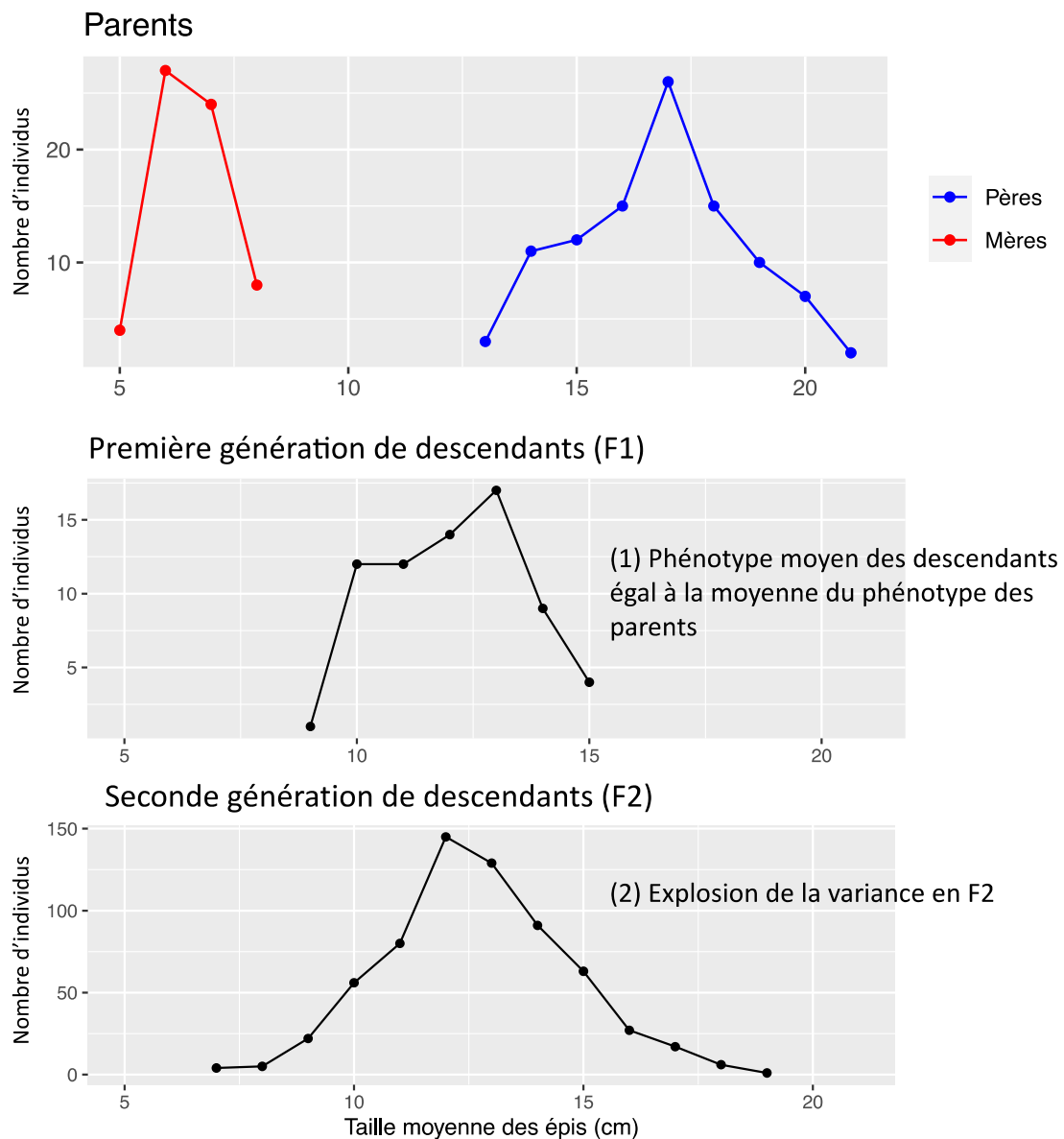


Figure 1 : Résultat d'une étude de East menée en 1911 sur l'hérédité de la longueur des épis de maïs

Distribution de la taille moyenne des épis de maïs chez les parents et les descendants, à la première et seconde génération[26].

I. A. 3 Le modèle infinitésimal

Faisons l'hypothèse que les caractères quantitatifs sont sous la dépendance de locus et d'allèles indépendants et équivalents. Autrement dit, prenons en compte uniquement l'effet additif des locus et des allèles en négligeant les phénomènes de dominance (interactions entre allèles d'un même locus) et l'épistasie (interactions entre différents locus)[4,27]. La Figure 2 illustre les conséquences de ces hypothèses pour des caractères phénotypiques mesurables, codés par un nombre différent de locus, dans des populations théoriques. Le premier encadré montre la distribution théorique d'un caractère phénotypique codé par un unique locus pour

lequel il existerait deux allèles : l'allèle a , associé à une petite valeur du caractère, et l'allèle A , associé à une grande valeur du caractère. Selon les lois de Mendel, si on fait l'hypothèse que les allèles A et a présentent la même fréquence allélique, la moitié des individus de la population devraient être de génotype Aa (ou aA). Puisque l'on considère que les allèles A et a sont équivalents, les individus de génotype Aa (ou aA) présentent un phénotype intermédiaire entre celui associé au génotype aa (petite valeur du caractère) et celui associé au génotype AA (grande valeur du caractère). Ainsi, le phénotype le plus représenté dans la population correspond au phénotype moyen. En appliquant le même raisonnement pour des caractères phénotypiques mesurables codés par deux, cinq ou cinquante locus (voir Figure 2), on constate que la distribution théorique du phénotype tend vers une loi normale lorsque le nombre de locus augmente.

Le modèle infinitésimal fait l'hypothèse que les caractères quantitatifs sont sous la dépendance d'un nombre infini de locus indépendants et équivalents, ayant tous un effet infiniment faible. Sous ces hypothèses, l'application des lois de Mendel permet de prédire une distribution normale des caractères quantitatifs au sein des populations. Ce modèle permet également d'expliquer les observations (1) et (2) sur la transmission des caractères quantitatifs. Malgré les hypothèses du modèle qui constituent une simplification importante de la réalité, le modèle infinitésimal est considéré comme le fondement de la génétique quantitative. Il a permis de développer les outils actuellement utilisés par les généticiens pour comprendre et prédire la réponse des caractères quantitatifs à la sélection[4,5]. Si l'on admet que les caractères quantitatifs sont codés par une infinité de gènes, il est vain de travailler à l'échelle des gènes et des allèles. Ainsi, contrairement à la génétique mendélienne ou à la génétique des populations, la génétique quantitative ne s'intéresse pas aux locus impliqués dans l'expression du caractère, mais repose entièrement sur l'analyse des composantes de la variation du phénotype.

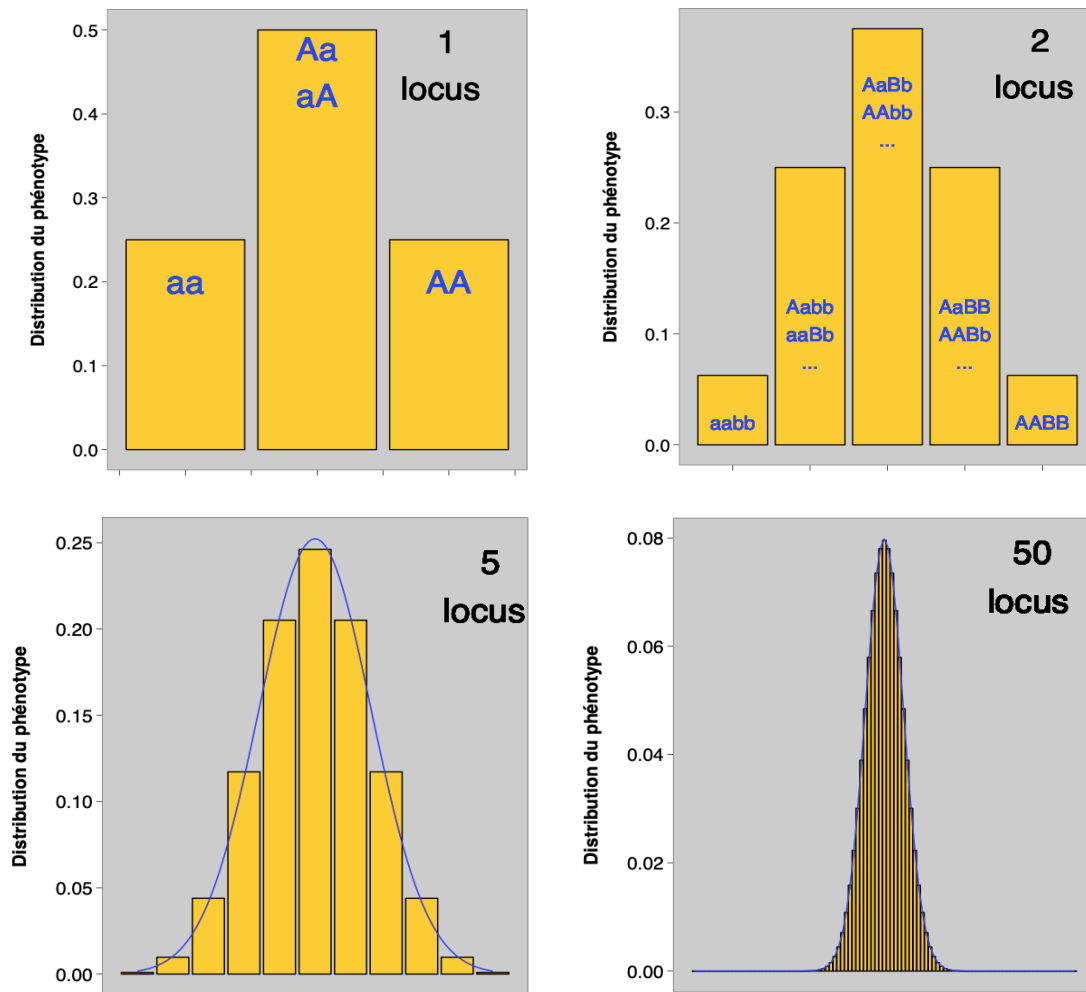


Figure 2 : Distribution du phénotype pour des caractères mesurables codés par un, deux, cinq ou 50 locus indépendants et équivalents

Selon les lois de Mendel, si chaque locus présente 2 allèles ayant la même fréquence au sein de la population, la distribution théorique du phénotype tend vers une loi normale lorsque le nombre de locus augmente[28].

I. A. 4 Décomposition de la variance phénotypique et variance génétique additive

Comme expliqué précédemment, la valeur d'un caractère quantitatif varie entre les individus au sein d'une population. Cette variation est quantifiée par la variance du caractère, appelée variance phénotypique (V_P). V_P résulte de différentes sources de variation qui peuvent être résumées en une composante génétique, notée V_G et une composante environnementale, notée V_E [29].

$$V_P = V_G + V_E$$

V_G correspond à l'effet de tous les allèles des gènes ayant une influence sur le caractère. Cette variance peut être divisée en une composante additive, notée V_A et une composante non-additive, notée V_{NA} [29].

$$V_P = V_A + V_{NA} + V_E$$

V_A correspond à l'effet génétique additif, qui est fondamentalement lié à l'hérédité du phénotype entre les générations. À ce titre, V_A est un paramètre très important en génétique quantitative car il correspond à la part de la variabilité phénotypique qui est transmise à la descendance. V_{NA} correspond à l'effet génétique non-additif, c'est-à-dire l'effet de la dominance et de l'épistasie. V_{NA} n'est pas (ou peu) transmise à la descendance[5,29]. Selon Fisher, les effets génétiques non-additifs sont négligeables à l'échelle des grandes populations, V_{NA} est donc généralement négligée lorsqu'on travaille sur des populations naturelles.⁵

Pour observer une réponse à la sélection, c'est-à-dire un décalage de la valeur moyenne du phénotype sélectionné au cours des générations, il est nécessaire qu'une part de la variabilité phénotypique soit transmissible à la descendance. L'objectif de la génétique quantitative est donc de mesurer la part de variance du phénotype expliquée par la diversité héritable (V_A) et de la distinguer du reste de la variance (notamment V_E)[5,29]. La variation génétique additive d'un caractère quantitatif peut être décrite par différents paramètres dépendants de V_A , dont notamment l'héritabilité et le coefficient de variation additif.

I. A. 5 Héritabilité d'un caractère quantitatif

L'héritabilité d'un caractère quantitatif, notée h^2 , correspond à la proportion de variance génétique additive par rapport à la variance totale du caractère. La valeur de h^2 est donc comprise entre zéro et un[29].

$$h^2 = \frac{V_A}{V_P}$$

Dans un contexte de sélection, qui peut être naturelle ou artificielle, h^2 permet de prédire dans quelle mesure les phénotypes sélectionnés seront transmis à la descendance, voir Figure 3. Ainsi, une héritabilité importante est en faveur d'un fort potentiel évolutif du caractère[29].

Il est intéressant de noter que les sélectionneurs utilisent h^2 pour estimer avec précision le gain génétique, c'est-à-dire la variation de la valeur moyenne du caractère sélectionné qui est attendue entre deux générations. Le gain génétique, noté R , est estimé par le produit de h^2 et d'un différentiel de sélection, noté S , qui correspond à la différence entre le phénotype moyen des individus sélectionnés et celui de la population. Cette formule porte le nom d'équation des sélectionneurs[5,30].

$$R = h^2 \times S$$

Bien que cette relation soit pertinente dans un contexte de sélection artificielle, sur des caractères uniques et dans des conditions contrôlées, il a été montré que l'évolution de la valeur moyenne des caractères étaient difficilement prévisible dans les populations sauvages soumises à la sélection naturelle[5].

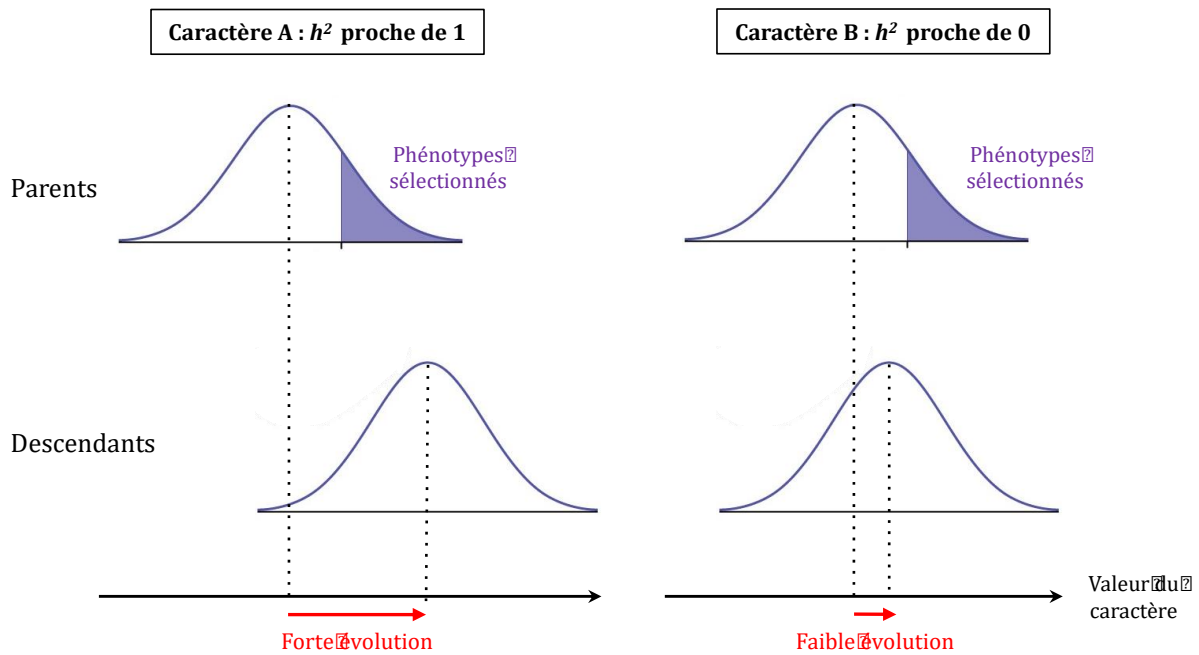


Figure 3 : Lien entre l'héritabilité d'un caractère et son potentiel évolutif

Une forte héritabilité (h^2 proche de 1) est en faveur d'un fort potentiel évolutif du caractère. À l'inverse, une faible héritabilité (h^2 proche de 0) est en faveur d'un faible potentiel évolutif.

I. A. 6 Coefficient de variation additif d'un caractère quantitatif

Le coefficient de variation additif d'un caractère quantitatif, noté CV_A , est obtenu en standardisant la racine carrée de la variance génétique additive par la valeur moyenne du caractère.

$$CV_A = \frac{\sqrt{V_A}}{\mu_{trait}}$$

Contrairement à h^2 , CV_A est une mesure de la variation génétique additive indépendante des autres sources de variation phénotypique. Ce coefficient permet donc d'évaluer le potentiel évolutif d'un caractère quantitatif en termes de proportion par rapport à la valeur moyenne du caractère[31].

I. A. 7 Méthodes d'estimation de l'héritabilité (h^2)

I. A. 7. i Cadre méthodologique commun

L'héritabilité est la mesure de la variation génétique additive la plus largement utilisée en génétique quantitative[5]. Son estimation se base d'une part sur les mesures du phénotype au sein de la population et d'autre part sur les relations de parenté entre individus. Cette approche ne nécessite donc pas de connaissance explicite des locus impliqués dans le déterminisme du

caractère[5,29]. L'idée est que si les individus apparentés (et donc génétiquement similaires) sont phénotypiquement plus proches les uns des autres que les individus qui ne sont pas apparentés (et donc génétiquement différents), on peut déduire que les gènes contribuent fortement aux variations phénotypiques, et donc que h^2 est élevé[32]. Cependant, il est important de souligner que les individus apparentés partagent parfois plus qu'une proximité génétique. Un environnement commun entre individus de la même famille peut être à l'origine d'une forte ressemblance phénotypique et ainsi entraîner une surestimation de V_A au détriment de V_E . Ce facteur doit notamment être pris en compte pour l'étude des populations naturelles vivant dans un milieu hétérogène car les individus de la même famille ont, en général, plus de chances d'occuper un environnement similaire (héritage territorial, faible dispersion, etc.)[5].

L'héritabilité peut être estimée par différentes méthodes plus ou moins complexes. Ces méthodes se distinguent notamment par le degré d'apparentement pris en compte dans les calculs. Nous présentons ici les méthodes les plus classiquement utilisées.

I. A. 7. ii La régression parents-jeunes

La régression parents-jeunes consiste à réaliser une régression du phénotype moyen des parents (ou de l'un des parents) sur le phénotype moyen de la descendance pour estimer l'héritabilité. Il s'agit d'une méthode relativement simple à mettre en place qui n'est pas biaisée par les effets de dominance. Cependant, cette approche ne permet pas de prendre en compte les effets environnementaux. Cette méthode est également sensible aux problèmes liés à la consanguinité. Ainsi, la régression parent-jeune est surtout adaptée aux approches expérimentales qui permettent de contrôler les paramètres environnementaux et le schéma de croisements de la population[5,29].

I. A. 7. iii Le modèle frères-sœurs

Le modèle frères-sœurs prend uniquement en compte les relations de parenté entre frères et sœurs (ou demi-frères et demi-sœurs). L'héritabilité est estimée à l'aide d'un modèle ANOVA (*analysis of variance*) qui compare la somme des carrés intrafamiliaux à la somme des carrés interfamiliaux pour le caractère étudié. Comme la régression parents-jeunes, cette méthode n'est pas adaptée à l'étude des populations sauvages en raison d'un biais par les effets communs de l'environnement. De plus, le protocole expérimental peut être relativement difficile à mettre en œuvre en conditions naturelles[5,29].

I. A. 7. iv Le modèle animal

Tout d'abord, il est important d'insister sur le fait que le terme « modèle animal » désigne ici un modèle statistique de génétique quantitative qui permet d'analyser les composantes de la variance d'un caractère quantitatif (et par extension d'estimer son

héritabilité)[5,29,33]. Bien que cette appellation puisse porter à confusion dans le domaine des sciences vétérinaires, il s'agit d'un terme consacré et insubstituable en génétique quantitative. Dans la suite de ce manuscrit, le terme « modèle animal » fera donc référence à ce modèle statistique de génétique quantitative et ne devra pas être confondu avec l'utilisation d'une espèce animale dans le cadre de la recherche biomédicale.

Le terme « modèle animal » a été choisi par les généticiens quantitatifs travaillant sur la sélection animale car ce modèle est centré sur l'individu, contrairement aux autres méthodes d'estimation de l'héritabilité qui sont habituellement centrées sur un type de relation de parenté (comme les relations parents-descendants ou frères-sœurs). En effet, le modèle animal utilise les coefficients d'apparentement de chaque paire d'individus et prend donc en compte l'intégralité et la diversité des liens de parenté d'un individu avec le reste de la population. Par ailleurs, le modèle animal permet d'inclure divers effets environnementaux qui pourraient biaiser l'estimation de la variation génétique additive[5,29,33]. À ce jour, le modèle animal est considéré comme la méthode la plus performante pour réaliser des analyses génétiques quantitatives sur les populations naturelles qui ont souvent un pédigrée complexe et qui occupent généralement un environnement hétérogène[5].

Le travail présenté dans la partie expérimentale est basé sur l'utilisation de modèles animaux dans une population naturelle de lézards vivipares. Dans le but de faciliter la compréhension de cette étude, nous détaillons les bases mathématiques du modèle animal dans la Partie I.B puis nous présentons un tutoriel sur l'ajustement de ce modèle avec le logiciel R[24] dans la Partie I.C.

I. B Le modèle animal

I. B. 1 Le modèle animal linéaire classique

Le modèle animal classique permet d'analyser les composantes de la variance d'un caractère phénotypique mesuré au sein d'une population. La variance génétique additive (V_A) est estimée à partir de deux éléments : la mesure du phénotype des individus et les relations de parenté au sein de la population, qui sont résumées dans la matrice d'apparentement[29,32].

I. B. 1. i Calcul de la matrice d'apparentement (Matrice A)

Le pédigrée de la population sauvage se présente sous la forme d'un tableau (ou matrice) indiquant l'identité de la mère et du père pour chaque individu. Les individus dont le père et la mère sont inconnus sont appelés individus fondateurs ou population de base : ils correspondent aux individus présents depuis le début du suivi et aux immigrants. Par exemple, le pédigrée d'une population de 10 individus peut être résumé par la matrice P ci-dessous. Les individus sont identifiés par les codes $A1, A2, \dots, A10$ et on note NA les parents inconnus[29].

$$P = \begin{array}{c} \begin{array}{ccc} \text{Individus} & \text{Mères} & \text{Pères} \end{array} \\ \left(\begin{array}{ccc} A1 & NA & NA \\ A2 & NA & NA \\ A3 & NA & NA \\ A4 & NA & NA \\ A5 & A2 & A4 \\ A6 & A1 & A3 \\ A7 & A1 & A3 \\ A8 & A1 & A3 \\ A9 & A2 & A4 \\ A10 & A5 & A6 \end{array} \right) \end{array} \quad \left. \begin{array}{c} \\ \\ \\ \\ \end{array} \right\} \text{Fondateurs}$$

On suppose que les fondateurs sont génétiquement indépendants les uns des autres, c'est-à-dire qu'ils ne partagent aucun lien de parenté. En pratique les fondateurs peuvent être majoritairement des immigrants. La matrice d'apparentement, appelée matrice A , résume l'information disponible dans le pédigrée. Il s'agit d'une matrice carrée dont la dimension correspond au nombre d'individus présents dans le pédigrée (en incluant les fondateurs). Pour construire la matrice A , on calcule le coefficient d'apparentement, généralement noté r , entre chaque paire d'individus[29,32]. Le coefficient d'apparentement est compris entre 0 et 1 et correspond à la proportion de génome qui est identique par descendance entre deux individus. Ainsi, entre individus non apparentés, le coefficient r est de 0 ; entre parents et descendants ou entre frères/soeur, il est de 0,5. Si l'on reprend l'exemple du pédigrée détaillé ci-dessus, la matrice A est la suivante :

$$A = \begin{matrix} & \begin{matrix} A1 & A2 & A3 & A4 & A5 & A6 & A7 & A8 & A9 & A10 \end{matrix} \\ \begin{matrix} A1 \\ A2 \\ A3 \\ A4 \\ A5 \\ A6 \\ A7 \\ A8 \\ A9 \\ A10 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 1/2 & 0 & 1/4 \\ 0 & 1 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 1/4 \\ 0 & 0 & 1 & 0 & 0 & 1/2 & 1/2 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 1/2 & 0 & 0 & 0 & 1/2 & 1/4 \\ 0 & 1/2 & 0 & 1/2 & 1 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 & 1 & 1/2 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 & 1/2 & 1 & 1/2 & 0 & 1/4 \\ 1/2 & 0 & 1/2 & 0 & 0 & 1/2 & 1/2 & 1 & 0 & 1/4 \\ 0 & 1/2 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 1 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 & 1/2 & 1/2 & 1/4 & 1/4 & 1/4 & 1 \end{pmatrix} \end{matrix}$$

I. B. 1. ii Estimation de la Variance génétique additive (V_A)

Soit Y la matrice associée aux mesures du phénotype pour les n individus de la population, notées y_1, \dots, y_n .

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

On définit μ comme la moyenne du phénotype au sein de la population. Le phénotype y_i de l'individu i peut alors être considéré comme une variation autour de μ . Cette variation peut être décomposée en une origine génétique additive transmissible, notée A_i , et le "reste", noté E_i , qui correspond aux effets environnementaux et aux effets génétiques non-additifs[29,32]. On définit ainsi le modèle suivant :

$$y_i = \mu + A_i + E_i$$

Le vecteur des A_i est supposé suivre une loi normale multivariée centrée, ayant pour variance-covariance la matrice $V_A \times A$, avec V_A la variance génétique additive (que l'on souhaite estimer) et A la matrice d'apparentement.

$$\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \sim N(0, V_A \times A)$$

Le vecteur des E_i est supposés suivre une loi normale multivariée centrée ayant pour variance-covariance la matrice $V_R \times I$, avec V_R la variance résiduelle et I la matrice identité. Autrement dit, les E_i sont indépendants et suivent chacun une loi normale centrée en 0 et de variance V_R .

$$\begin{pmatrix} E_1 \\ \vdots \\ E_n \end{pmatrix} \sim N(0, V_R \times I)$$

Ainsi, l'estimation de la matrice de variance-covariance des A_i et des E_i permet d'estimer la variance génétique additive (V_A) et la variance résiduelle (V_R) du modèle. Le modèle peut ainsi estimer V_A et V_R à partir du phénotype Y et de la matrice A [29]. Ce modèle, qui comporte des effets aléatoires est appelé modèle mixte. Pour plus de détail sur ces relations statistiques, le lecteur peut consulter les articles *Quantitative genetic methods depending on the nature of the phenotypic trait*[29] et *An ecologist's guide to the animal model*[32].

L'héritabilité et le coefficient de variation additif de y peuvent être estimés avec les formules suivantes[29] :

$$h^2 = \frac{V_A}{V_A + V_R} \quad \text{et} \quad CV_A = \frac{\sqrt{V_A}}{\mu}$$

I. B. 1. iii Ajout d'effets aléatoires

Comme expliqué précédemment, le phénotype peut être influencé par des variables environnementales. Ainsi, les individus partageant un environnement similaire présentent généralement un phénotype plus proche que les individus occupant des environnements différents. Ce phénomène peut biaiser l'estimation de V_A , notamment car les individus apparentés sont plus susceptibles d'être exposés à un même environnement (héritage territorial, faible dispersion etc.). Pour pallier ce problème, le modèle animal prévoit l'ajout d'effets aléatoires qui permettent de prendre en compte la ressemblance engendrée par des effets communs de l'environnement. Inclure des effets aléatoires sur des variables environnementales connues (notées U_1, \dots, U_k) permet d'estimer la part de variance phénotypique qui peut être expliquée par ces variables. Les facteurs aléatoires, généralement écrits en majuscules, sont intégrés à l'équation du modèle de la manière suivante :

$$y_i = \mu + U_{1i} + \dots + U_{ki} + A_i + E_i$$

L'héritabilité est alors estimée par la formule suivante, avec V_{U_1}, \dots, V_{U_k} la variance associée aux facteurs U_1, \dots, U_k :

$$h^2 = \frac{V_A}{V_A + V_{U_1} + \dots + V_{U_k} + V_R}$$

I. B. 1. iv Ajout d'effets fixes

Le modèle animal permet également la prise en compte d'effets fixes correspondant aux sources de variations biologiques, ou liées à la conception du plan d'expérience, et susceptibles

de biaiser l'estimation de l'héritabilité. Un effet fixe diffère d'un effet aléatoire du fait qu'on ne lui suppose pas de distribution, d'où son nom. Un facteur fixe (f), généralement écrit en minuscule, peut être intégré à l'équation du modèle de la manière suivante :

$$y_i = \mu + f_i + A_i + E_i$$

Par exemple, le sexe des individus est généralement ajouté au modèle en tant qu'effet fixe pour prendre en compte la variance engendrée par le dimorphisme sexuel. Notons que la variance associée aux effets fixes n'est généralement pas incluse dans l'estimation de l'héritabilité.

I. B. 1. v Modèles utilisés dans la partie expérimentale

Des modèles animaux classiques ont été utilisés dans la partie expérimentale pour estimer l'héritabilité de deux variables de croissance (la taille à la naissance, notée LMA_N et la taille au stade subadulte, notée LMA_{SA}). Les variations environnementales ont été prises en compte à travers l'ajout d'effets aléatoires sur l'année de naissance et sur l'effet maternel. Ces effets sont expliqués ci-dessous.

- L'année de naissance

De nombreux paramètres environnementaux peuvent varier en fonction des années. C'est notamment le cas de la température, de la disponibilité en nourriture ou de la pression de prédation par exemple. La population de lézards étudiée dans la partie expérimentale a un cycle biologique circannuel avec une diapause et une unique période de reproduction. Les individus nés la même année (dont les frères et sœurs d'une même fratrie) sont donc plus susceptibles d'être exposés à un environnement similaire durant leur ontogénie. Dans ce contexte, il était intéressant d'inclure l'année de naissance comme effet environnemental.

- L'effet maternel

Les individus issus des mêmes parents peuvent présenter des similitudes qui ne sont pas directement liées aux effets génétiques. Par exemple, un effet maternel célèbre est la qualité du lait chez les mammifères : les mères produisant le meilleur lait seront plus susceptibles d'avoir des descendants en bonne santé qui grandissent rapidement. Prendre en compte l'identité de la mère comme effet aléatoire permet de prendre en compte cette source de ressemblance entre fratries pour ne pas la confondre avec des effets génétiques additifs. Il est important de noter que les effets parentaux n'impliquent aucune ressemblance phénotypique entre parents et descendants, mais prennent uniquement en compte la ressemblance entre frères et sœurs[29]. Bien qu'il n'y ait pas de soins parentaux chez les lézards, la taille et la condition corporelle de la mère durant la gestation peut influencer sur le phénotype des descendants[34]. Il était donc intéressant d'inclure l'effet maternel comme effet environnemental dans nos modèles.

Nous avons également ajouté un effet fixe sur le sexe des individus pour tenir compte des différences systématiques entre mâles et femelles liées au dimorphisme sexuel. Nous avons ainsi pu écrire l'équation des modèles utilisés pour analyser la variance de la taille à la naissance (LMAN) et de la taille au stade subadulte (LMA_{SA}) de la manière suivante :

$$y_i = \mu + MOTHER_i + ANNE_NAISSANCE_i + sexe_i + A_i + E_i$$

Avec y_i , le caractère phénotypique étudié (LMAN ou LMA_{SA}), μ , la moyenne du caractère, A_i , les effets génétiques additifs, $MOTHER_i$ et $ANNE_NAISSANCE_i$, les effets aléatoires, $sexe_i$, l'effet fixe et E_i les résidus du modèle. Ainsi, l'héritabilité et le coefficient de variation additif de LMAN et LMA_{SA} ont été estimés par les formules suivantes dans la partie expérimentale.

$$h^2 = \frac{V_A}{V_A + V_{AN} + V_M + V_R} \quad \text{et} \quad CV_A = \frac{\sqrt{V_A}}{\mu}$$

Avec V_A , la variance génétique additive, V_{AN} la variance associée à l'année de naissance, V_M la variance associée à l'effet maternel, V_R la variance résiduelle et μ la moyenne du caractère.

I. B. 2 Le modèle animal hiérarchique non-linéaire

I. B. 2. i Spécificités du modèle

Le modèle animal hiérarchique non-linéaire permet d'analyser les composantes de la variance de paramètres eux-mêmes estimés par le modèle. Ce type de modèle animal innovant se distingue du modèle animal linéaire classique qui permet simplement d'analyser les composantes de la variance de caractères phénotypiques directement mesurés au sein de la population.

Prenons l'exemple d'un modèle animal hiérarchique non-linéaire qui prend en entrée une variable réponse, notée y , et une variable explicative notée x (ces variables correspondent à des caractères mesurés). Une fonction non linéaire f fait le lien entre x , y et les P paramètres à estimer, notés $\theta_1, \dots, \theta_P$. Ce modèle peut être décrit par les équations (1) et (2) présentées ci-dessous.

$$y_i = f(x_i, \theta_{1_i}, \dots, \theta_{P_i}) + \varepsilon_i \quad (1)$$

Pour un individu i , y_i est un vecteur qui contient les mesures de la variable réponse, x_i est un vecteur qui contient les mesures de la variable explicative, f est une fonction non linéaire de la variable explicative et des paramètres à estimer $\theta_{1_i}, \dots, \theta_{P_i}$ et ε_i est un vecteur qui contient les résidus du modèle pour l'individu i . En supposant les paramètres $\theta_1, \dots, \theta_P$ tous indépendants les uns des autres, on peut écrire l'équation (2) pour chaque paramètre p .

$$\theta_{p_i} = \mu_p + A_{p_i} + E_{p_i} \quad (2)$$

Ici μ_p une constante correspondant à la moyenne du paramètre θ_p au sein de la population, A_{p_i} est la valeur génétique additive du paramètre p pour l'individu i et E_{p_i} est l'effet de l'environnement permanent sur le paramètre p pour l'individu i . Notons que des effets fixes et aléatoires peuvent également être ajoutés à l'équation (2), de la même manière que pour les modèles animaux linéaires classiques. On suppose les distributions suivantes pour les valeurs génétiques des P paramètres :

$$A_p \sim N(0, AV_{A,p})$$

où A est la matrice d'apparentement entre tous les individus et $V_{A,p}$ est la variance génétique additive du paramètre p . De même, on suppose pour E_p que :

$$E_p \sim N(0, IV_{PE,p})$$

où I est la matrice identité et $V_{PE,p}$ est la variance associée à l'environnement permanent pour le paramètre p .

Il est important de noter que le modèle hiérarchique non-linéaire estime une variance résiduelle qui correspond à la variance des résidus ε de l'équation (1). Or, c'est la variance associée à E_p qui doit être incluse dans l'estimation de l'héritabilité (h^2). La variance résiduelle renvoyée par le modèle ne devra donc pas apparaître dans l'estimation de h^2 . La variance associée à E_p est estimée en spécifiant un effet « individu », appelé effet de l'environnement permanent.

1. B. 2. ii Modèle utilisé dans la partie expérimentale

Dans la partie expérimentale, un modèle de croissance logistique a été utilisé pour modéliser les courbes de croissance des individus de la population étudiée. Ce modèle décrivait la taille des individus, qui correspondait ici à la LMA, en fonction de l'âge. Les paramètres du modèle étaient la taille asymptotique (notée l_{max}), la taille à la naissance (noté l_0) et le taux de croissance (noté k). L'équation du modèle de croissance logistique est présentée ci-dessous et une représentation graphique du modèle avec les paramètres fixés à $l_{max} = 68$, $l_0 = 21$ et $k = 1$, est présentée à la Figure 4.

$$LMA = \frac{l_{max}}{1 + \frac{l_{max} - l_0}{l_0} \times e^{-k \times Age}}$$

Un modèle animal hiérarchique non-linéaire a été utilisé pour estimer la variation génétique additive associée aux paramètres k et l_{max} qui étaient eux-mêmes estimés par le modèle. Le modèle peut être décrit par les équations suivantes :

$$LMA = \frac{l_{max}}{1 + \frac{l_{max} - l_0}{l_0} \times e^{-k \times Age}} + E_i \quad (1)$$

$$l_{max_i} = \mu + MOTHER_{1i} + ANNEE_NAISSANCE_{1i} + sexe_{1i} + A_{1i} + PE_{1i} \quad (2)$$

$$k = \mu + MOTHER_{2i} + ANNEE_NAISSANCE_{2i} + sexe_{2i} + A_{2i} + PE_{2i} \quad (3)$$

Comme pour les modèles animaux classiques, l'année de naissance (*ANNEE_NAISSANCE*) et l'effet maternel (*MOTHER*) ont été inclus comme effets aléatoires et le sexe (*sexe*) a été inclus comme effet fixe. Un effet aléatoire de l'environnement permanent (*PE*) a également été ajouté pour estimer la variation de k et l_{max} qui ne peut pas être expliquée par les effets génétiques additifs. L'héritabilité a alors été estimée pour chaque paramètre avec la formule suivante :

$$h^2 = \frac{V_A}{V_A + V_{AN} + V_M + V_{PE}}$$

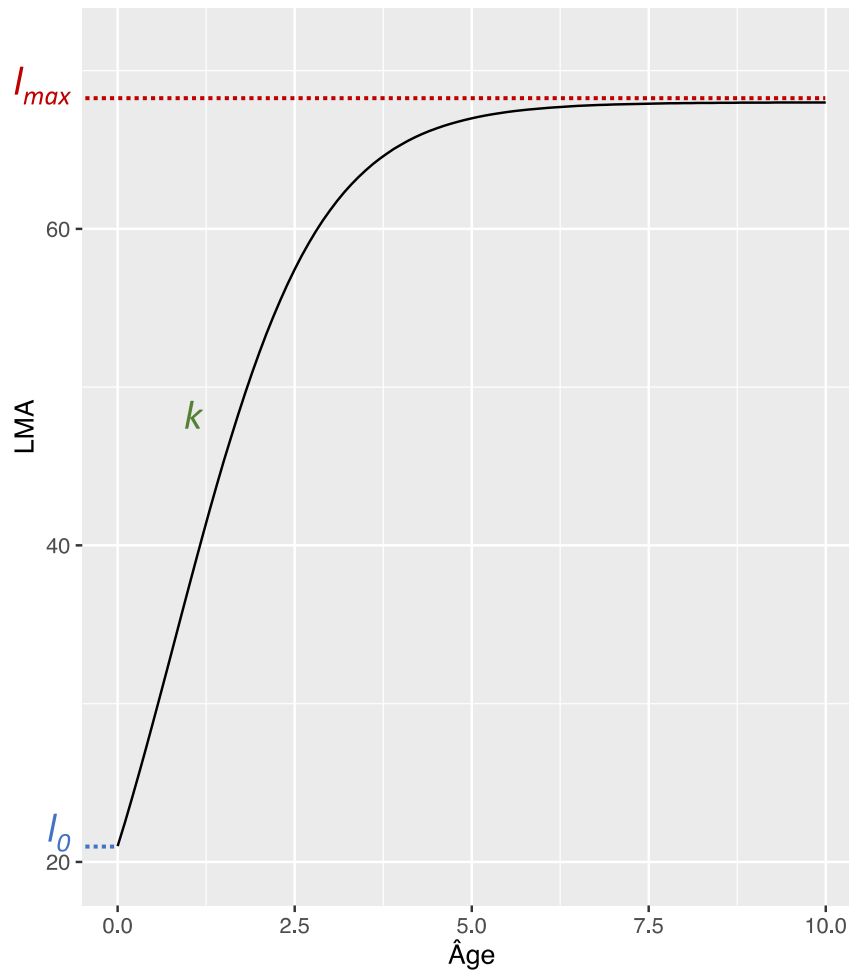


Figure 4 : Représentation graphique du modèle de croissance logistique
Avec représentation de la taille asymptotique (l_{max}), de la taille à la naissance (l_0) et du taux de croissance (k).

I. B. 3 Place de l'inférence bayésienne dans l'ajustement des modèles animaux

I. B. 3. i Cadre général de l'inférence bayésienne

L'inférence bayésienne est une méthode d'inférence statistique qui repose sur l'application du théorème de Bayes. Cette approche s'oppose à la statistique fréquentiste classique ; là où l'approche fréquentiste considère qu'un paramètre correspond à une valeur unique qui peut être estimée à partir des données observées (souvent par la méthode du maximum de vraisemblance), l'approche bayésienne considère les probabilités comme une mesure épistémique de la connaissance de la valeur du paramètre. Ainsi, en statistique bayésienne, un paramètre est estimé par une loi de probabilité qui associe chaque valeur possible du paramètre à un degré de croyance (c'est-à-dire à une probabilité d'être « correcte »). Notons que cette loi de probabilité, également appelée loi *a posteriori*, présente l'avantage de décrire l'incertitude sur l'estimation du paramètre. L'inférence bayésienne peut être appréhendée comme une « révision des croyances » sur la loi de probabilité d'un paramètre à partir des données observées et via une fonction de vraisemblance.

En résumé, la mise en œuvre de l'inférence bayésienne implique une première étape durant laquelle on définit la distribution *a priori* du paramètre à estimer, qui reflète l'état des connaissances sur le paramètre avant l'observation des données. Une seconde étape consiste à utiliser le théorème de Bayes pour actualiser la distribution du paramètre en prenant en compte l'information apportée par les données observées. On estime alors la distribution *a posteriori* qui correspond à la loi de probabilité du paramètre, à interpréter en termes de degré de croyance sur les valeurs du paramètre[5,35].

I. B. 3. ii L'algorithme de Monte Carlo par chaînes de Markov (MCMC)

Les algorithmes de Monte Carlo par Chaîne de Markov (MCMC) sont des méthodes d'inférence bayésienne largement utilisées qui permettent d'approcher numériquement la distribution *a posteriori* des paramètres. Les méthodes MCMC reposent sur l'échantillonnage de l'espace des paramètres grâce à une suite de variables aléatoires, également appelée chaîne. Chaque terme de la chaîne correspond à une itération de l'algorithme MCMC et dépend uniquement de l'itération précédente. Après une phase de convergence, l'algorithme tend à échantillonner dans la distribution *a posteriori* des paramètres. Ainsi, en sauvegardant, après convergence des chaînes, les valeurs d'un paramètre pour un assez grand nombre d'itérations, on obtient une série de valeurs permettant d'estimer la distribution *a posteriori* du paramètre. L'algorithme MCMC est lancé plusieurs fois, de manière à simuler différentes chaînes qui seront combinées pour obtenir une estimation plus fiable de la *distribution a postérieure* des paramètres[29,35]. Pour plus de détails sur le principe de l'inférence bayésienne en écologie animale et sur le fonctionnement des algorithmes MCMC, le lecteur pourra se référer aux ouvrages *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities*[36] et *Handbook of Markov Chain Monte Carlo*[35].

I. B. 3. iii Intérêt de l'inférence bayésienne dans l'ajustement des modèles

L'ajustement de modèles animaux sur des données issues de populations naturelles est souvent un réel défi pour les généticiens. Le pédigrée est généralement complexe et les données, obtenues de manière empirique, sont souvent entachées d'incertitude. Ainsi, le modèle doit pouvoir être ajusté malgré la complexité du jeu de données et retranscrire l'incertitude dans l'estimation des paramètres[5].

L'inférence bayésienne est particulièrement adaptée pour ajuster des modèles statistiques hiérarchiques complexes dans un contexte de forte incertitude. En effet, le fait d'estimer la loi de probabilité des paramètres permet de décrire précisément et complètement l'incertitude associée aux paramètres, là où la moyenne et l'écart-type classiquement utilisés pour résumer l'estimation dans l'approche fréquentiste sont souvent insuffisants. Ainsi, dans le cadre de modèles hiérarchiques où l'estimation des paramètres dépend de variables elles-mêmes estimées avec une certaine incertitude, l'approche bayésienne permet de prendre en compte l'intégralité des sources d'incertitude et de les répercuter avec justesse sur les paramètres estimés. Autrement dit, l'inférence bayésienne permet une propagation naturelle de l'incertitude à travers les niveaux du modèle, de sorte que toutes les informations disponibles soient contenues dans la distribution *a posteriori* des paramètres estimés[5].

I. C Tutoriel sur l’ajustement de modèles animaux avec le paquet brms[37] de R[24]

I. C. 1 Introduction et présentation du paquet brms[37]

Cette section a pour objectif de fournir quelques clés concernant l’aspect pratique de l’ajustement de modèles animaux à l’aide de R[24], afin que le lecteur comprenne les analyses présentées en partie expérimentale. Par ailleurs, ce tutoriel pourra guider le lecteur ayant déjà quelques notions en inférence bayésienne et en programmation R dans la réalisation de ses propres analyses génétiques quantitatives. Nous utilisons ici le paquet d’inférence bayésienne brms[37], disponible sur site du CRAN[38]. Ce paquet permet d’ajuster des modèles statistiques complexes et se distingue des autres paquets d’inférence bayésienne par l’utilisation d’un type d’algorithme MCMC particulier, nommé Monte Carlo hamiltonien, implémenté dans la plateforme Stan[39]. Stan permet généralement une convergence plus rapide d’une grande famille de modèles grâce à un échantillonnage plus efficace de l’espace des paramètres[40]. Le package brms[37] est particulièrement adapté à l’ajustement de modèles hiérarchiques non-linéaires complexes comme les modèles présentés dans la partie expérimentale.

Nous expliquons ici les bases de l’ajustement de modèles animaux (classique et hiérarchique non-linéaire) en nous appuyant sur les données et modèles présentés dans la partie expérimentale. Dans un premier temps, nous présentons le format des jeux de données utilisés, nous donnons quelques clés concernant l’utilisation de la fonction `brm()` et nous indiquons comment définir un modèle animal. Dans un second temps, nous expliquons comment interpréter la sortie du modèle et comment estimer l’héritabilité et le coefficient de variation additif des paramètres.

I. C. 2 Présentation des données

I. C. 2. i Jeu de données sur le suivi de la population

Le jeu de données utilisé dans ce tutoriel regroupe toutes les informations liées au suivi de la population de lézards vivipares étudiée dans la partie expérimentale, les 15 premières lignes sont présentées à la Figure 5. Chaque ligne correspond à l’observation d’un individu, à une date donnée, et chaque colonne correspond à une variable. Dans l’ordre, les variables sont : l’identifiant de l’individu, l’âge (en années), l’année de l’observation, le sexe (prend 0 pour les femelles et 1 pour les mâles), la LMA (en mm), la date d’observation (au format Année-Mois-Jour), une variable qui indique si l’observation est réalisée à la naissance de l’individu, l’année de naissance et l’identifiant de la mère.

##	ID	Age	AN	Sex	LMA	Date	Birth	Annee_nais	Mother
## 1	100012	1	1986	0	33	1986-06-08	FALSE	1985	100012 _ 1
## 2	100012	3	1988	0	47	1988-07-21	FALSE	1985	100012 _ 1
## 3	100025	1	1986	1	42	1986-06-08	FALSE	1985	100025 _ 3
## 4	100026	1	1986	0	32	1986-06-08	FALSE	1985	100026 _ 4
## 5	100031	1	1986	1	32	1986-06-08	FALSE	1985	100031 _ 5
## 6	100031	3	1988	1	53	1988-06-21	FALSE	1985	100031 _ 5
## 7	100035	1	1986	1	27	1986-06-09	FALSE	1985	100035 _ 7
## 8	100035	2	1987	1	48	1987-07-02	FALSE	1985	100035 _ 7
## 9	100036	1	1986	0	29	1986-06-09	FALSE	1985	100036 _ 9
## 10	100036	2	1987	0	47	1987-06-17	FALSE	1985	100036 _ 9
## 11	100037	1	1986	0	30	1986-06-09	FALSE	1985	100037 _ 11
## 12	100039	1	1986	1	27	1986-06-09	FALSE	1985	100039 _ 12
## 13	100039	2	1987	1	47	1987-07-23	FALSE	1985	100039 _ 12
## 14	100039	3	1988	1	51	1988-06-06	FALSE	1985	100039 _ 12
## 15	100047	1	1986	1	33	1986-06-09	FALSE	1985	100047 _ 15

Figure 5 : Extrait du jeu de données utilisé dans la partie expérimentale.

I. C. 2. ii Pédigrée

Le pédigrée de la population se présente sous la forme d'un jeu de données comprenant une ligne par individu et trois colonnes qui contiennent les identifiants de l'individu, la mère et le père, voir Figure 6. Chaque individu de la population doit correspondre à une ligne dans le pédigrée, notamment les individus fondateurs pour lesquels les parents sont inconnus (notés <NA>). La fonction `fixPedigree()` du paquet `Pedantics`[41] peut être utilisée pour rajouter explicitement les individus fondateurs dans le pédigrée.

##	individu	mère	père
##	20636	<NA>	<NA>
##	20783	<NA>	<NA>
##	22037	<NA>	<NA>
##	22318	<NA>	<NA>
##	30424	<NA>	<NA>
...			
##	421028_1999	22037	42512
##	421029_1999	22037	42512
##	421030_1999	22037	42512
##	421031_1999	22037	42512
##	421037_1999	22037	42512
##	421038_1999	22037	42512
##	421409_1999	22318	320285
##	421410_1999	22318	320285
##	421411_1999	22318	320285
##	421417_1999	22318	320285
##	421380_1999	30424	322435
##	421381_1999	30424	322435
##	421407_1999	30424	322080
##	421408_1999	30424	322435
##	421210_1999	32774	<NA>
##	421211_1999	32774	322146
##	421217_1999	32774	322146
...			

Figure 6 : Extrait du pédigrée utilisé dans la partie expérimentale.

I. C. 2. iii Matrice A

La matrice d'apparentement (Matrice A) est une matrice carrée dont la dimension correspond au nombre d'individus présents dans le pédigrée. Cette matrice contient les coefficients de proximité génétique de chaque paire d'individus, voir Figure 7. La matrice A peut être obtenue à partir du pédigrée de la population en utilisant la fonction `pedigreeStats()` du paquet `Pedantics`[41].

##	20636	20783	22037	22318	30424	...	421028_1999	421029_1999	421030_1999
## 20636	1	0	0.0	0.0	0.0	...	0.0	0.0	0.0
## 20783	0	1	0.0	0.0	0.0	...	0.0	0.0	0.0
## 22037	0	0	1.0	0.0	0.0	...	0.5	0.5	0.5
## 22318	0	0	0.0	1.0	0.0	...	0.0	0.0	0.0
## 30424	0	0	0.0	0.0	1.0	...	0.0	0.0	0.0
...
## 421028_1999	0	0	0.5	0.0	0.0	...	1.0	0.5	0.5
## 421029_1999	0	0	0.5	0.0	0.0	...	0.5	1.0	0.5
## 421030_1999	0	0	0.5	0.0	0.0	...	0.5	0.5	1.0
...

Figure 7 : Extrait de la Matrice A utilisée dans la partie expérimentale.

I. C. 3 Utilisation de la fonction `brm()`

La fonction `brm()` est le cœur du paquet `brms`[37] car c'est la fonction qui permet d'ajuster des modèles mathématiques selon une approche bayésienne. Dans un premier temps, nous verrons comment implémenter un modèle animal dans `brms`[37]. Puis, nous détaillerons quelques arguments utiles de `brm()` et nous donnerons des exemples d'appel de la fonction. Enfin, nous parlerons de la distribution *a priori* des paramètres à estimer.

I. C. 3. i Implémentation du modèle

I. C. 3. i. (i) Modèle animal classique

L'argument `formula` de la fonction `brm()` permet de spécifier la formule mathématique du modèle. Les modèles animaux classiques sont des modèles linéaires multiples qui s'écrivent sous la forme $Y \sim 1 + \text{Effet}_1 + \dots + \text{Effet}_k$ [42], avec Y la variable qui correspond au caractère phénotypique étudié (dite variable réponse) et $\text{Effet}_1, \dots, \text{Effet}_k$ les variables correspondent aux effets fixes et aléatoires du modèle (dites variables explicatives). L'effet aléatoire lié à la variation génétique additive s'écrit $(1 \mid \text{gr}(\text{ID}, \text{cov} = \text{Amatrix}))$ avec ID la variable qui correspond à l'identifiant des individus et Amatrix la matrice A[42]. Cette expression permet ainsi d'associer l'identifiant de l'animal à la matrice A. Les effets fixes sont simplement explicités avec le nom de la variable, tandis que les autres effets aléatoires s'écrivent sous la forme $(1 \mid \text{Variable})$.

Prenons l'exemple d'un modèle de la partie expérimentale qui permet d'analyser les composantes de la variance de la taille des individus (LMA) à la naissance, nous appellerons ce modèle « Modèle animal sur LMA_N » dans le reste de ce tutoriel. Ce modèle est ajusté avec un effet fixe sur le sexe des individus et des effets aléatoires sur l'effet génétique additif l'effet de l'année de naissance et l'effet maternel. Ce modèle est donc implémenté avec l'argument suivant dans la fonction `brm()` :

```
formula = LMA ~ 1 +
  Sex + # Effet du sexe
  (1 | gr(ID, cov = Amatrix_N)) + # Effet génétique additif
  (1 | Annee_nais) + # Effet de l'année de naissance
  (1 | Mother), # Effet maternel
```

I. C. 3. i. (ii) [Modèle animal hiérarchique non-linéaire](#)

La fonction `brm()` permet également d'implémenter des modèles animaux hiérarchiques non-linéaires. Comme nous l'avons expliqué dans la section B.2, ce type de modèle est défini par plusieurs équations. Celles-ci doivent être explicitées dans l'argument `formula` de `brm()` en utilisant la fonction `brmsformula()`. Notons que l'argument `n1` doit être défini à `TRUE` (`n1 = TRUE`) pour préciser que le modèle implémenté est non-linéaire. Prenons l'exemple du modèle hiérarchique non-linéaire de la partie expérimentale qui correspond au modèle explicité dans la section B.2.ii. Les équations (2) et (3) de la section B.2.ii peuvent être condensées en un seul terme lors de l'implémentation du modèle avec l'argument `formula` de la fonction `brm()` :

```
formula = brmsformula(
  LMA ~ lmax / (1 + ((lmax - LMA_N) / LMA_N) * exp(- k * Age)),
  # Modèle de croissance logistique
  lmax + k ~ Sex +
    # Effet du sexe
    (1 | Annee_nais) +
    # Effet de l'année de naissance
    (1 | gr(ID_additif, cov = Amatrix)) +
    # Effet génétique additif
    (1 | ID_pe) +
    # Effet de l'environnement permanent
    (1 | Mother),
    # Effet maternel
  n1 = TRUE)
```

I. C. 3. ii Arguments de la fonction `brm()`

Nous détaillons ici les arguments de la fonction `brm()` qui sont utiles dans le cadre de l'implémentation de modèles animaux.

I. C. 3. ii. (i) Arguments sans valeur par défaut

Les arguments présentés ici ne sont pas associés à des valeurs par défaut, ils doivent donc nécessairement être précisés lors de l'utilisation de la fonction `brm()`.

- `formula` : il s'agit du premier argument de `brm()`, il permet de définir le modèle. L'utilisation de cet argument a été détaillée dans la section C.3.i.
- `data` : jeu de données contenant toutes les variables utilisées dans le modèle (variables phénotypiques et environnementales).
- `data2` : argument permettant de préciser la Matrice A.

I. C. 3. ii. (ii) Arguments ayant des valeurs par défaut

Les arguments présentés ici peuvent être spécifiés mais cela n'est pas obligatoire car une valeur par défaut est définie dans la fonction `brm()`.

- `prior` : distribution *a priori* des paramètres du modèle. L'utilisation de cet argument est détaillée dans la section C.3.iv.
- `chains` : nombre de chaînes de Markov calculées par l'algorithme MCMC. Par défaut cet argument est fixé à 4 dans la fonction `brm()`.
- `cores` : nombre de cœurs utilisés par l'ordinateur. Mettre plusieurs cœurs à disposition de l'algorithme permet de paralléliser le calcul des chaînes et ainsi d'obtenir un ajustement plus rapide du modèle. Par défaut cet argument est fixé à 1 dans la fonction `brm()`.
- `iter` : nombre d'itérations total de chaque chaîne (en incluant la phase de chauffe, voir argument `warmup`). Par défaut cet argument est fixé à 2000 dans la fonction `brm()`.
- `warmup` : nombre d'itérations retirées au début de chaque chaîne qui ne sont donc pas incluses dans la distribution *a posteriori* des paramètres. Ces itérations correspondent à la phase de chauffe nécessaire pour que l'algorithme converge et commence théoriquement à échantillonner la distribution *a posteriori* des paramètres. Par défaut, cet argument est fixé à la moitié de l'argument `iter` dans la fonction `brm()`.
- `thin` : argument permettant de définir un intervalle de sauvegarde des itérations. Par exemple, si `thin = 10`, on ne conservera qu'une seule itération sur 10 dans la distribution *a posteriori*. Cette fonctionnalité permet de réduire l'autocorrélation au sein des chaînes et d'économiser de la mémoire. Par défaut, cet argument est fixé à 1 dans la fonction `brm()`.

- `init` : argument permettant de préciser les coordonnées initiales pour l'échantillonnage de l'espace des paramètres (coordonnées initiales des chaînes du MCMC). Par défaut, ces valeurs sont tirées aléatoirement dans l'espace des paramètres.

I. C. 3. iii Appel de la fonction `brm()`

Les exemples présentés aux Figures 8 et 9 montrent comment appeler la fonction `brm()`. La Figure 8 correspond au modèle animal sur LMA_N et la Figure 9 correspond au modèle animal hiérarchique non-linéaire. Notons que la sortie de la fonction `brm()` doit être stockée dans une variable (nommée ici `Model`) pour permettre l'accès aux résultats ultérieurement. Le temps nécessaire pour ajuster un modèle peut être assez long, cela dépend majoritairement des données, des paramètres et bien sûr des capacités de l'ordinateur[43].

```
Model <- brm(
  formula = LMA ~ 1 +
    Sex + # Effet du sexe
    (1 | gr(ID, cov = Amatrix)) + # Effet génétique additif
    (1 | Annee_nais) + # Effet de l'année de naissance
    (1 | Mother), # Effet maternel
  data = donnees_mod_N, # Sous sélection du jeu de données
  data2 = list(Amatrix = Amatrix_N), # Sous sélection de la matrice A
  chains = 4,
  cores = 4,
  iter = 4000,
  warmup = 1000,
  thin = 1
```

Figure 8 : Implémentation d'un modèle animal classique avec le paquet `brms`[37].

Le modèle présenté à la Figure 8 est ajusté sur une sous-sélection du jeu de données et de la matrice A initiales qui comprend uniquement les individus observés à la naissance (respectivement `donnees_mod_N` et `Amatrix_N`). On peut noter que ce modèle est ajusté avec quatre chaînes MCMC de 4000 itérations dont 1000 itérations pour la convergence. La distribution *a posteriori* des paramètres du modèle sera donc estimée avec 12000 itérations (car quatre chaînes de 3000 itérations au final). Le modèle présenté à la Figure 9 est ajusté avec quatre chaînes MCMC de 2000 itérations dont 500 itérations pour la convergence. La distribution *a posteriori* des paramètres du modèle sera donc estimée avec 6000 itérations (car quatre chaînes de 1500 itérations au final). Par ailleurs, une distribution *a priori* et des coordonnées initiales sont précisées.

```
## Récupération des distribution a posteriori des écarts types (sd_)
# et des variances (v_) associés aux effets aléatoires et résiduels

# Effet génétique additif
Model<- brm(
  formula = brmsformula(
    LMA ~ lmax / (1 + ((lmax - LMA_N) / LMA_N) * exp(- k * Age)),
    # Modèle de croissance logistique
    lmax + k ~ Sex +
    # Effet du sexe
    (1 | Annee_nais) +
    # Effet de l'année de naissance
    (1 | gr(ID_additif, cov = Amatrix)) +
    # Effet génétique additif
    (1 | ID_pe) +
    # Effet de l'environnement permanent
    (1 | Mother),
    # Effet maternel
    nl = TRUE),
  data = data,
  data2 = list(Amatrix = Amatrix_mod),
  chains = 4,
  cores = 4,
  iter = 2000,
  warmup = 500,
  prior = Prior,
  init = rep(list(list(b_lmax = array(data = c(68, 0)),
    b_k = array(data = c(1, 0)))), 10))
```

Figure 9 : Implémentation d'un modèle animal hiérarchique non-linéaire avec brms[37].

I. C. 3. iv Quelques éléments sur la distribution *a priori* des paramètres

L'inférence bayésienne nécessite de définir une distribution *a priori* pour chaque paramètre à estimer[35]. Le paquet brms[37] prévoit une distribution *a priori* par défaut qui correspond à une loi de Student centrée en 0. Cette distribution est généralement adaptée s'il n'y a pas de croyance sur les paramètres à estimer avant l'analyse. Elle est utilisée lorsque l'argument `prior` de la fonction `brm()` n'est pas spécifié[44]. Les lois *a priori* utilisées par brms[37] peuvent être obtenues en utilisant la fonction `get_prior()` tandis que la fonction `prior()` permet de définir la distribution *a priori* des paramètres.

I. C. 4 Interprétation de la sortie du modèle

I. C. 4. i Diagnostic de l'algorithme MCMC

Avant d'interpréter les résultats de l'estimation des paramètres, il est important de vérifier le comportement de l'algorithme MCMC pour s'assurer que l'échantillonnage des distributions *a posteriori* ait été correctement réalisé.

I. C. 4. i. (i) Définitions

Deux points importants doivent toujours être contrôlés lors de l'utilisation d'un algorithme MCMC.

- La convergence : suffisamment d'itérations doivent avoir été éliminées au début des chaînes MCMC de manière à ce que les itérations sauvegardées soient bien échantillonnées dans la distribution *a posteriori* des paramètres. Autrement dit, la phase de chauffe doit être suffisamment longue. L'obtention de la convergence peut être vérifiée en contrôlant visuellement les traces des chaînes MCMC : lorsque l'algorithme converge les différentes chaînes du MCMC échantillonnent la même région de l'espace et se mélangent donc complètement[29,35].
- L'autocorrélation : les itérations successives de l'algorithme MCMC ont tendance à être corrélée car le tirage d'une nouvelle valeur dépend de l'itération précédente dans des chaînes de Markov. Or, afin d'obtenir une estimation correcte de la distribution *a posteriori*, il faudrait idéalement obtenir un échantillon assez long des chaînes MCMC pour lequel l'autocorrélation est négligeable. En effet, si l'autocorrélation est trop importante rien ne garantit que l'information obtenue soit suffisante. En pratique, l'autocorrélation est contrôlée lors de l'examen du tracé des chaînes et de la vérification de la taille effective d'échantillon (ESS)[29,35].

I. C. 4. i. (ii) Vérifications

Prenons l'exemple du modèle animal sur LMA_N , la Figure 10 montre la représentation graphique des chaînes de certains paramètres du modèle et la Figure 11 correspond à la sortie du modèle. La première étape du diagnostic consiste à examiner les traces du MCMC. Les traces des chaînes sont représentées à droite des couples de graphiques présentés à la Figure 10 et correspondent à l'évolution des valeurs échantillonnées au fil des itérations des chaînes MCMC. La convergence est vérifiée si on ne voit aucune tendance dans la trace et les chaînes sont mélangées entre elles. L'autocorrélation est faible si les valeurs des chaînes ne suivent pas de chemin traçable[29,35]. Notons que l'estimation de la distribution marginale *a posteriori* de chaque paramètre est représentée, à gauche de ces graphiques, sous forme de densités de probabilités.

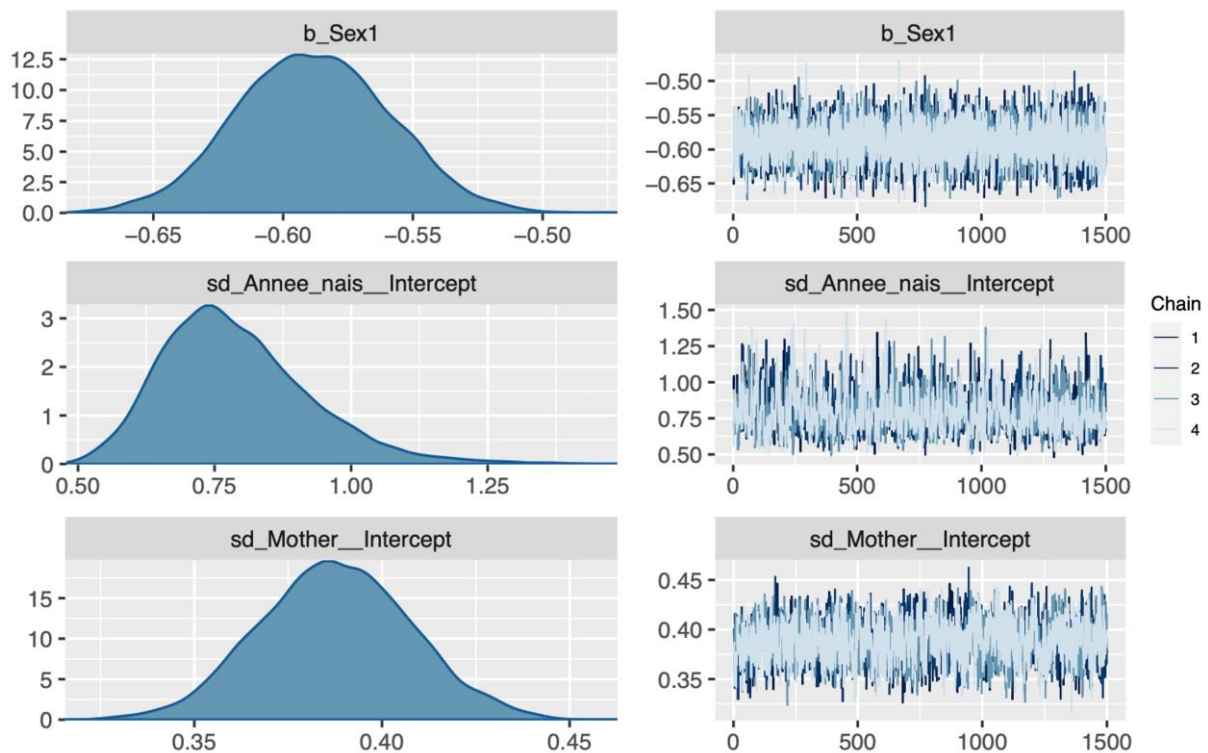


Figure 10 : Densité et trace de quelques paramètres du modèle animal sur LMA_N . Représentation des distributions *a posteriori* (à gauche) et des chaînes (à droite) de l'effet du sexe sur LMA_N (b_sex1), de l'écart-type lié l'année de naissance ($sd_Annee_nais_Intercept$) et de l'écart type lié à l'effet maternel ($sd_Mother_Intercept$).

Pour aller plus loin dans le diagnostic de l'algorithme MCMC, il faut s'intéresser à des statistiques calculées automatiquement par le paquet `brms`[37]. `Brms`[37] estime la taille effective d'échantillon (notée ESS pour « Effective Sample Size ») qui est la taille d'un échantillon non corrélé théorique équivalent à l'échantillon du modèle en termes de quantité d'informations sur la distribution *a posteriori*. L'ESS est estimé dans `brms`[37] sous la forme de 2 paramètres : `Bulk_ESS` et `Tail_ESS`[45] (voir Figure 11). Dans l'idéal, l'ESS devrait faire au moins 1000 itérations, mais cela peut être difficile à obtenir pour des modèles complexes dont l'ajustement est très lent. Par ailleurs, on peut noter que `brms`[37] a tendance à minimiser l'ESS en comparaison d'algorithmes MCMC classiques n'utilisant pas Stan[39]. `Brms`[37] estime également la statistique de Gelman et Rubin (Rhat) qui est un indicateur de la convergence des chaînes (voir Figure 11). Si Rhat est $> 1,05$ il y a de grandes chances qu'il y ait des problèmes de convergence dans l'algorithme MCMC[44]. Dans la partie expérimentale, nous nous sommes assurés que l'intégralité des inférences bayésiennes réalisées avec le paquet `brms` présentent un `Bulk_ESS` minimal supérieur à 200 et un Rhat maximal inférieur à 1,05[46].

I. C. 4. ii Interprétation de l'estimation des paramètres

Après avoir contrôlé la convergence et l'autocorrélation, on peut interpréter l'estimation des paramètres à partir de la sortie du modèle, voir Figure 11. La sortie contient une première partie qui récapitule les caractéristiques du modèle (formule mathématique, données utilisées, nombre de chaînes et d'itérations au sein de chaque chaîne). Les résultats sont ensuite présentés dans trois sections.

- **Group-Level Effects** : contient l'estimation des écarts-types associés aux effets aléatoires. Sur la Figure 11 on peut voir les écarts-types associés à l'année de naissance, l'effet génétique additif (qui correspond à la variable ID) et l'effet maternel.
- **Population-Level Effects**, contient l'estimation des effets fixes. Sur la Figure 11 on peut voir la moyenne du caractère au sein de la population (μ), appelée Intercept, et l'effet du sexe.
- **Family Specific Parameters**, contient l'estimation de l'écart-type résiduel du modèle (sigma).

Les paramètres sont décrits par une colonne **Estimate** qui correspond à la moyenne de l'échantillon MCMC. On peut noter que la médiane constitue parfois une meilleure synthèse statistique que la moyenne, notamment dans le cadre de distributions *a posteriori* très asymétriques. Les colonnes **l-95%** et **u-95%** correspondent aux limites de l'intervalle de crédibilité à 95% (qui est parfois appelé intervalle de confiance bayésien à 95%). Cet intervalle contient 95% de l'échantillon MCMC et, comme l'intervalle de confiance utilisé en statistique fréquentiste, il donne une idée de l'incertitude associée à l'estimation du paramètre[44].

```

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LMA ~ 1 + Sex + (1 | gr(ID, cov = Amatrix_N)) + (1 | Annee_nais) + (1
| Mother)
## Data: donnees_mod_N (Number of observations: 3353)
## Draws: 4 chains, each with iter = 2000; warmup = 500; thin = 1;
## total post-warmup draws = 6000
##
## Group-Level Effects:
## ~Annee_nais (Number of levels: 20)
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)    0.79    0.14    0.57    1.10 1.00    1101    2090
##
## ~ID (Number of levels: 3353)
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)    0.26    0.04    0.17    0.34 1.00     408     701
##
## ~Mother (Number of levels: 1501)
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)    0.39    0.02    0.35    0.43 1.00    1115    2529
##
## Population-Level Effects:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      21.37    0.18    21.03    21.73 1.01     472     848
## Sex1           -0.59    0.03    -0.65    -0.53 1.00    6829    4871
##
## Family Specific Parameters:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma      0.75    0.01    0.72    0.78 1.00     854    1924
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).

```

Figure 11 : Sortie d'un modèle animal ajusté selon une approche bayésienne.

I. C. 4. iii Estimation de l'héritabilité (h^2) et du coefficient de variation addition (CV_A)

Le code permettant d'obtenir la distribution *a posteriori* de h^2 et CV_A de la taille à la naissance (LMA_N) à l'aide du paquet `brms`[37] est présenté à la Figure 12. Dans un premier temps, la fonction `VarCorr()` est utilisée pour récupérer la distribution *a posteriori* des écarts-types associés aux effets aléatoires et aux résidus. La distribution *a posteriori* des variances est ensuite obtenue en utilisant simplement la fonction carré (2). Le CV_A et l' h^2 sont estimés avec les formules détaillées dans la section B.1.v. Notons que la fonction `as.mcmc()` permet de spécifier à R[24] que les distributions *a posteriori* sont des chaînes MCMC (et non pas de simples suites de nombres). On peut ainsi obtenir la représentation graphique des chaînes avec la fonction `plot()`. La fonction `summary()` permet de récupérer la moyenne, la médiane et l'intervalle de crédibilité à 95% des paramètres.

```

## Récupération des distributions a posteriori des écarts types (sd_)
# et des variances (v_) associés aux effets aléatoires et résiduels

# Effet génétique additif
sd_ad <- (VarCorr(m_animal_N, summary = FALSE)$ID$sd)
v_ad <- sd_ad ^ 2

# Effet de l'année de naissance
sd_AN <- (VarCorr(m_animal_N, summary = FALSE)$Annee_nais$sd)
v_AN <- sd_AN ^ 2

# Effet maternel
sd_mother <- (VarCorr(m_animal_N, summary = FALSE)$Mother$sd)
v_mother <- sd_mother ^ 2

# Résidus
sd_r <- (VarCorr(m_animal_N, summary = FALSE)$Residual$sd)
v_r <- sd_r ^ 2

## Estimation de la distribution a posteriori de l'héritabilité (h2)
h2 <- as.mcmc(v_ad / (v_ad + v_AN + v_mother + v_r))

## Estimation de la distribution a posteriori du coefficient de
# variation additif (CVA)
CVA <- as.mcmc(sd_ad / mean(donnees_mod_N$LMA))

```

Figure 12 : Estimation de la distribution *a posteriori* de h^2 et CV_A de LMA_N .

I. D Étude de la sélection sur des caractères quantitatifs

I. D. 1 La surface de sélection individuelle

La surface de sélection individuelle (SSI) d'une population est une fonction qui fait le lien entre la valeur sélective relative des individus (w) et un, ou plusieurs, caractères quantitatifs standardisés. Au sein d'une population, la valeur sélective, également appelée fitness, correspond au nombre moyen de descendants viables issus des individus présentant un même phénotype. Soit z un caractère quantitatif standardisé, la SSI peut être modélisée par la fonction quadratique suivante (1)[47,48].

$$w = \alpha + \beta z + \frac{1}{2} \gamma z^2 \quad (1)$$

Notons que le caractère quantitatif standardisé z correspond à la différence entre la valeur absolue du caractère (Z) et la valeur moyenne du caractère au sein de la population (\bar{Z}). La valeur moyenne du caractère standardisé \bar{z} est donc égale à 0.

$$\begin{aligned} z &= Z - \bar{Z} \\ \bar{z} &= 0 \end{aligned}$$

Par ailleurs, la valeur sélective relative w correspond au quotient de la valeur sélective absolue (W) par la valeur sélective moyenne au sein de la population (\bar{W}).

$$w = \frac{W}{\bar{W}}$$

Le coefficient β de (1) est donc égal à la valeur moyenne de la dérivée partielle de (1) par rapport à z . Autrement dit, β correspond à la pente moyenne de la SSI. La valeur de β permet ainsi d'évaluer si une pression de sélection s'exerce en faveur de l'augmentation ou de la diminution de z au sein de la population étudiée. De ce fait, β est appelé gradient de sélection directionnelle. De la même manière, γ est égal à la valeur moyenne de la dérivée partielle seconde de (1) par rapport à z . Autrement dit, γ correspond à la courbure moyenne de la SSI. La valeur de γ permet donc d'évaluer si une pression de sélection s'exerce en faveur d'une stabilisation ou d'une dispersion des valeurs de z au sein de la population. De ce fait, γ est appelé gradient de sélection stabilisante (ou divergente). On peut noter que le coefficient $\frac{1}{2}$ présent dans l'équation (1) est nécessaire pour que γ corresponde à la dérivée seconde de l'équation[48,49].

Il est important de souligner que les gradients de sélection β et γ permettent d'estimer la pression de sélection qui s'exerce sur des caractères phénotypiques au sein d'une population, mais ne donnent pas d'indication sur l'intensité de la réponse évolutive à cette sélection. En effet, l'évolution d'un caractère dépend de différents facteurs, comme la part de variance génétique additive associée au caractère ou les possibles liens génétiques avec d'autres

caractères par exemple. Ainsi, il est tout à fait possible de mettre en évidence une pression de sélection sur un caractère sans observer de réponse évolutive[49].

La SSI peut être définie pour plusieurs caractères quantitatifs ayant une influence réciproque, comme par exemple le poids et la taille des individus d'une population. Pour deux caractères y et z , la fonction associée à la SSI s'écrit de la manière suivante[48,49] :

$$w = \alpha + \beta_1 y + \beta_2 z + \frac{1}{2} \gamma_1 y^2 + \frac{1}{2} \gamma_2 z^2 \quad (2)$$

La valeur sélective relative w est alors exprimée en fonction des caractères quantitatifs standardisés y et z , de l'intercept α , du gradient de sélection directionnel associé à y , noté β_1 , du gradient de sélection directionnelle associé à z , noté β_2 , du gradient de sélection stabilisante (ou divergente) associé à y , noté γ_1 et du gradient de sélection stabilisante (ou divergente) associé à z , noté γ_2 .

I. D. 2 Estimation des gradients de sélection directionnelle et stabilisante

I. D. 2. i Présentation de la méthode

La méthode présentée par Lande et Arnold en 1983 permet d'estimer les gradients de sélection directionnelle et stabilisante, β et γ , grâce à un modèle de régression quadratique entre la mesure de la valeur sélective relative w^* et le ou les caractères quantitatifs étudiés. La mesure de la valeur sélective correspond au nombre de descendants engendrés par chaque individu au cours de sa vie. Ainsi, la mesure de la valeur sélective relative correspond au rapport entre le nombre de descendants d'un individu et le nombre de descendants moyen, pour la classe mâles ou femelles, au sein de la population.

Pour un caractère étudié de manière indépendante, la mesure de la valeur sélective relative w_i^* de l'individu i est exprimée en fonction du caractère quantitatif standardisé z_i et d'un terme d'erreur ε_i . Les coefficients α , β et γ correspondent à ceux permettant de définir la SSI. L'équation du modèle s'écrit de la manière suivante : [48,49]

$$w_i^* = \alpha + \beta z_i + \frac{1}{2} \gamma z_i^2 + \varepsilon_i$$

De même, pour deux caractères y et z étudiés simultanément, l'équation du modèle s'écrit de la manière suivante :

$$w_i^* = \alpha + \beta_1 z_i + \beta_2 y_i + \frac{1}{2} \gamma_1 z_i^2 + \frac{1}{2} \gamma_2 y_i^2 + \varepsilon_i$$

I. D. 2. ii Ajout de facteurs environnementaux

Certaines variables environnementales ont un impact important sur la valeur sélective des individus. C'est généralement le cas de la disponibilité en nourriture ou de la pression de prédation par exemple. Il peut alors être pertinent d'intégrer ces variables au modèle à travers

l'ajout d'un facteur aléatoire (ENV). Les équations des modèles à un (3) et deux (4) caractères s'écriront alors de la manière suivante :

$$w_i = \alpha + \beta z_i + \frac{1}{2} \gamma z_i^2 + ENV + \varepsilon_i \quad (3)$$

$$w_i = \alpha + \beta_1 z_i + \beta_2 y_i + \frac{1}{2} \gamma_1 z_i^2 + \frac{1}{2} \gamma_2 y_i^2 + ENV + \varepsilon_i \quad (4)$$

I. E Contexte général de l'étude

Cette section a pour but d'expliquer le contexte général et de justifier l'intérêt scientifique des analyses réalisées par la suite. Nous présentons d'abord quelques éléments généraux sur la biologie du lézard vivipare et sur ses traits d'histoire de vie. Puis, nous montrons en quoi l'étude du potentiel adaptatif de la stratégie de croissance du lézard vivipare est pertinente au vu du contexte actuel de réchauffement climatique.

I. E. 1 Éléments sur la biologie du lézard vivipare

Le lézard vivipare (*Zootoca vivipara*, Jacquin 1787) est un squamate de la famille des Lacertidés très largement répandu dans le nord de l'Eurasie, voir Figure 13. Il occupe une grande diversité d'habitats frais et humides comme les landes, les tourbières, les prairies humides, les forêts, les zones côtières et les jardins ruraux. On le retrouve à une altitude allant du niveau de la mer jusqu'à environ 2500 m[50,51]. Il se nourrit majoritairement d'arthropodes, notamment d'hémiptères, d'araignées, de coléoptères et de fourmis[52].



Figure 13 : Distribution mondiale du lézard vivipare (*Zootoca vivipara*) selon l'IUCN[51]

L'espèce présente trois stades de développement : le stade juvénile, entre la naissance et la première hibernation (0-3 mois), le stade subadulte, entre la première et la deuxième hibernation (0-15 mois) et le stade adulte, au-delà de la deuxième hibernation (>22 mois)[53]. La croissance est rapide avant la maturité sexuelle, qui est généralement atteinte à l'âge de deux ans, puis continue théoriquement tout au long de la vie des individus. On observe un dimorphisme sexuel tant sur la longévité que sur la taille. Les femelles peuvent vivre sept à 10 ans et mesurent entre 45 et 75 mm à l'âge adulte, tandis que les mâles dépassent rarement l'âge de six ans et atteignent une taille qui varie entre 40 et 60 mm[54]·[55]. La couleur des écailles présente une variation inter-individuelle considérable en lien avec l'âge et le sexe des individus, voir Figure 14. D'abord presque noirs à la naissance, les jeunes individus s'éclaircissent

rapidement vers des teintes brun foncé au stade subadulte, puis gris clair ou brun jaunâtre à l'âge adulte. La coloration ventrale varie de l'orange au beige clair et est généralement plus vive chez les mâles[53].

L'espèce est majoritairement vivipares, mais il existe quelques populations ovipares inféodées à des climats plus chauds[50,56]. Chez les populations vivipares, les œufs sont retenus dans les oviductes des femelles jusqu'à maturité des fœtus et sont émis sous forme de poches fines et translucides, voire Figure 14. A. Les nouveau-nés s'extraient de ces poches dans les minutes suivant la mise-bas[55].



Figure 14 : Photographies de lézards vivipares

A : Œufs (flèches blanches) et nouveau-né, **B** : Adulte, **C** : Femelle gestante.

Sources : Pierre de Villemereuil (Photos A et B), Thomas Reich (Photo C)

I. E. 2 Intérêt de l'étude du potentiel adaptatif de la stratégie de croissance du lézard vivipare

I. E. 2. i Le lézard vivipare comme espèce modèle et sentinelle

Le suivi de plusieurs populations naturelles de lézards vivipares a été initié dès la fin des années 1980[54] et a fait de l'espèce un des systèmes les plus pertinents pour l'étude écologique des populations animales à l'échelle mondiale. Le lézard vivipare est également un des seuls ectothermes terrestres vivants en milieu tempéré à faire l'objet d'un suivi sur le long terme. Ainsi, l'espèce est aujourd'hui un modèle reconnu en écophysiologie et en écologie évolutive.

Par ailleurs, le lézard vivipare est considéré comme espèce sentinelle d'une grande diversité d'habitats frais et humides menacés par l'élévation des températures. En écologie, les espèces sentinelles sont des espèces particulièrement sensibles qui répondent en premier aux perturbations de l'écosystème, par adaptation ou extinction locale[21]. Ainsi, l'état des populations naturelles de lézards vivipares est un bon indicateur de l'impact du changement climatique sur les écosystèmes. On observe aujourd'hui un déclin rapide du lézard vivipare dans le sud de l'Europe, en lien avec l'élévation des températures et la disparition des zones humides[22,23].

I. E. 2. ii Impact du réchauffement climatique sur les populations de lézard vivipare

Il a été montré que le réchauffement climatique était associé à un déclin important du lézard vivipare au niveau des marges sud de son aire de répartition. La carte présentée à la Figure 15 est issue d'une étude[22] publiée en 2015 et qui a montré le niveau de vulnérabilité de 130 populations européennes de lézards vivipares. Cette étude a estimé que 14 % à 30 % des populations européennes étaient menacées par une augmentation des températures de +2 à 4°C[22]. Bien que l'impact environnemental de l'élévation des températures puisse participer au déclin des populations (diminution des ressources alimentaires, disparition du milieu de vie, etc.), le réchauffement climatique a également des conséquences importantes sur le métabolisme des ectothermes. Chez le lézard vivipare, il a été montré qu'une hausse de 2°C durant la période estivale suffisait à provoquer une accélération globale du métabolisme. Cela se traduisait par une augmentation du taux de croissance des juvéniles, une entrée plus précoce des femelles à la reproduction et une mortalité anticipée des individus de plus d'un an, en particulier des adultes de grande taille[22]-[57]. Ces modifications des traits d'histoire de vie étaient associées à une baisse de l'effectif total de la population.

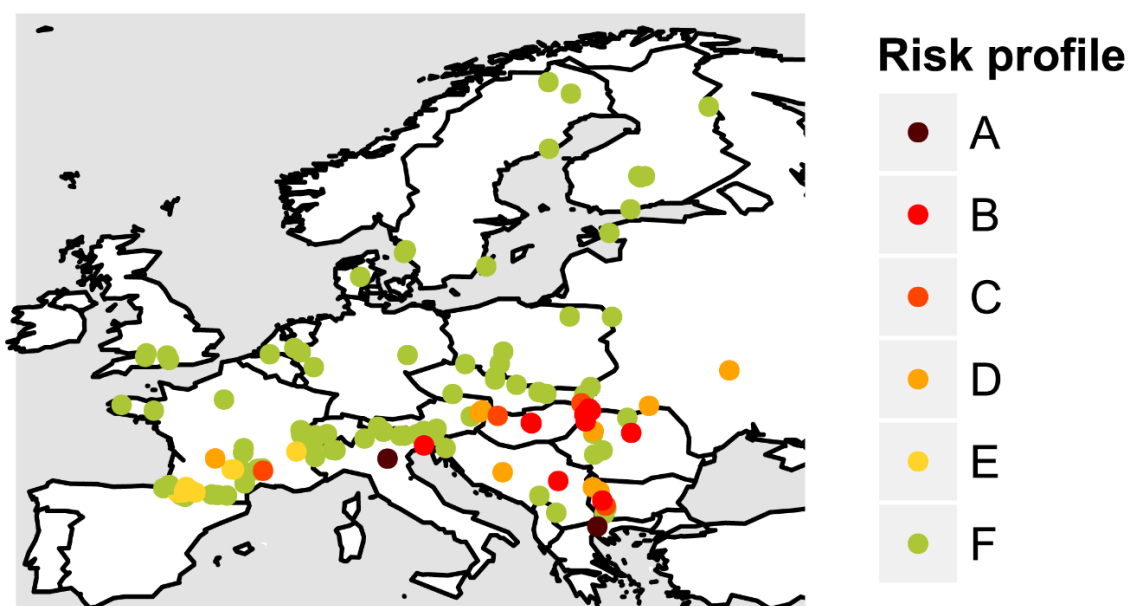


Figure 15 : Niveau de vulnérabilité de 130 populations européennes de lézard vivipare selon Bestion et al. 2015
Populations menacées par une augmentation supérieure à 2°C (A, B, C), 3°C (D) et 4°C (E) [22].

I. E. 2. iii Théorie métabolique de l'écologie

Selon la théorie métabolique de l'écologie[58], le métabolisme basal des ectothermes (I), qui correspond aux dépenses énergétiques strictement nécessaires à la survie d'un individu dans des conditions idéales, est corrélé à la température (T) et à la masse de l'individu (M). Comme on peut le voir dans l'équation suivante, une hausse des températures entraîne une augmentation de l'activité métabolique (et donc des dépenses énergétiques), qui est d'autant plus importante chez les individus de grande taille[22,58,59]. Dans cette équation, i_0 est une constante, E correspond à l'énergie d'activation et K à la constante de Boltzmann.

$$I = i_0 M^{3/4} e^{-E/KT}$$

Une hypothèse avancée pour expliquer la mortalité anticipée des lézards adultes en cas d'élévation de la température est que ces derniers ne pourraient pas compenser l'augmentation du coût métabolique par l'alimentation[22,59]. Les juvéniles, de taille modérée, pourraient plus facilement pallier l'augmentation des dépenses énergétiques en optimisant la prise alimentaire.

I. E. 2. iv Lien entre stratégie de croissance et dynamique des populations

Bien que l'ensemble des mécanismes impliqués dans la tolérance thermique du lézard vivipare ne soit pas encore tout à fait connu, on sait que la taille des individus joue un rôle important dans la dépense métabolique basale[58] d'une part et dans les mécanismes de thermorégulation[60] d'autre part. Par ailleurs, il a été montré que la taille des lézards vivipares conditionnait la fertilité des femelles, le succès reproducteur des mâles et a une influence sur l'âge de la maturité sexuelle[53–55]. La stratégie de croissance, qui est définie ici comme

l'évolution de la taille tout au long de la vie des individus, aurait donc un fort impact sur la dynamique des populations, voire Figure 16. Mieux connaître le potentiel adaptatif de la stratégie de croissance du lézard vivipare permettrait ainsi de donner des éléments en faveur du déclin ou du maintien de l'espèce sur le long terme.

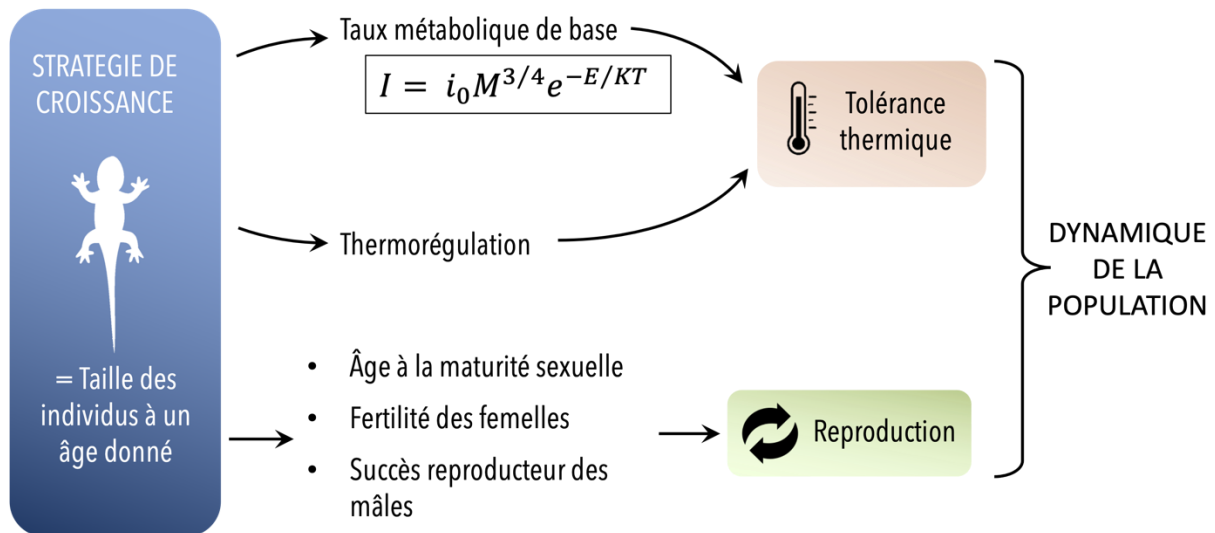


Figure 16 : Impact de la stratégie de croissance du lézard vivipare sur la dynamique des populations

La taille des individus est liée au métabolisme basal, à la thermorégulation et à la dynamique de reproduction de l'espèce. La stratégie de croissance a donc un fort impact sur la dynamique des populations.

Partie II - Étude expérimentale

Comme pour de nombreux reptiles, il a été montré que la stratégie de croissance du lézard vivipare variait beaucoup entre les individus[61][50]. Il a été admis, depuis la fin du XXe siècle, que cette variabilité était associée à une part de plasticité phénotypique et une part de variabilité d'origine génétique[61,62]. Comme expliqué dans la première partie, un caractère quantitatif ayant une part importante de variance génétique additive est plus enclin à répondre aux pressions de sélection. Dans cette étude, nous nous sommes intéressés à l'héritabilité et au coefficient de variation additif de quatre variables de croissance de manière à caractériser le potentiel adaptatif de la stratégie de croissance du lézard vivipare dans sa globalité. Nous avons également étudié les pressions de sélection qui s'exercent sur la stratégie de croissance au sein de la population étudiée.

II. A Matériel et méthodes

II. A. 1 Suivi de la population de lézards vivipares

II. A. 1. i Comité d'éthique

Les méthodes employées pour le suivi de la population correspondaient aux recommandations du guide de référence 2004 du comité IACUC de la société Américaine des Ichtyologues et Herpétologues. Ces protocoles ont été examinés et validés par le comité d'éthique en matière d'expérimentation animale Cuvier (Muséum National d'Histoire naturelle, 57 rue Cuvier, 75005 Paris).

II. A. 1. ii Population étudiée

La population étudiée se situait au niveau du Massif du Mont Lozère (France, latitude : 44,3840, longitude : 3,8778) à environ 1400 m d'altitude et faisait l'objet d'un suivi annuel depuis la fin des années 1980. Les individus étaient actifs d'avril à octobre et entraient en hibernation le reste de l'année. Les femelles ne complétaient qu'un seul cycle reproducteur durant la saison, avec une période d'accouplements en mai et des mises bas durant la seconde moitié du mois de juillet, voir Figure 17. L'effectif total de la population était relativement stable depuis le début du suivi et est estimé à environ 500 individus.

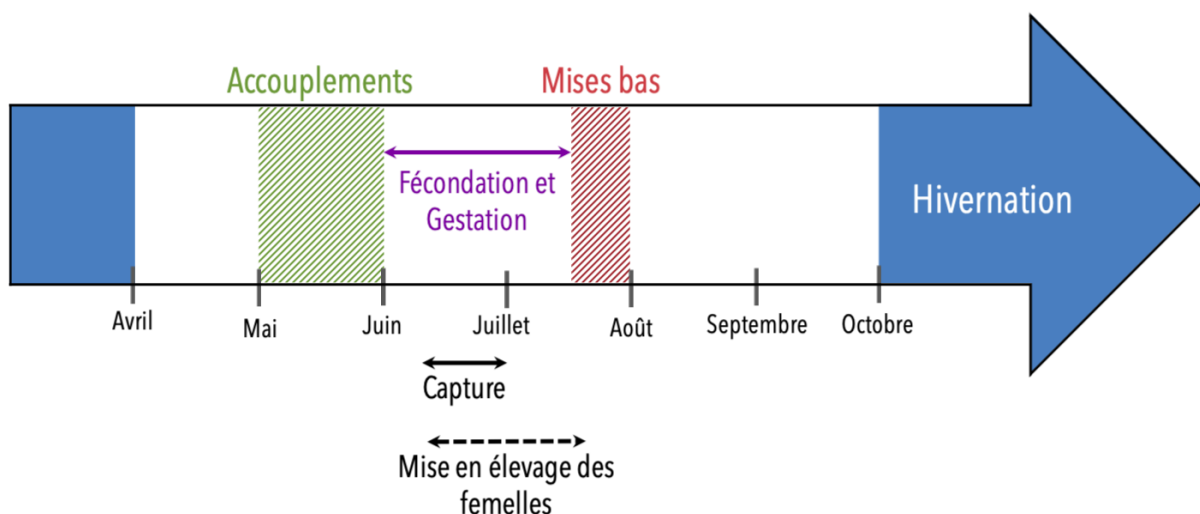


Figure 17 : Chronologie du suivi de la population étudiée

II. A. 1. iii Capture

Les captures ont été effectuées à la main et étaient généralement réalisées durant le mois de juin. Dans un premier temps, le code d'identification des lézards déjà marqués était relevé tandis que les lézards observés pour la première fois étaient marqués avec un code unique. Un prélèvement de tissus caudal de moins d'un cm était également réalisé chez les individus observés pour la première fois. Par la suite, le stade de développement était déterminé et les

individus étaient sexés. La taille était mesurée entre le museau et le cloaque de manière à prévenir les biais engendrés par d'éventuelles pertes de la queue. Enfin, les femelles gestantes étaient transportées au laboratoire et maintenues en élevage jusqu'à la parturition, ce qui permettait de déterminer l'ascendance maternelle des nouveau-nés.

II. A. 1. iv Marquage

Les individus ont été marqués par section de phalanges distales au niveau des quatre membres. Cette technique est bien tolérée chez les espèces de lézards peu mobiles, terrestres et qui perdent régulièrement leur doigt comme le lézard vivipare. Il a notamment été montré que les conséquences biologiques de la section de phalanges (stress, changement d'activité, mortalité) sur les lézards adultes sont moins importantes que celles associées à des méthodes alternatives de marquage permanent comme l'injection de transpondeurs par exemple[63]. Le protocole de marquage par section de phalanges a été optimisé au cours des années et implique une durée de manipulation minimale, des conditions d'asepsie strictes et un temps de repos avant le relâchement des individus dans l'environnement. Les phalanges distales étaient sectionnées à l'aide de ciseaux fins à iridectomie. Le matériel était désinfecté entre chaque utilisation sur des animaux différents avec une solution contenant de la chlorhexidine. Les doigts des juvéniles étant particulièrement fins, la section des phalanges devait parfois être réalisée sous une loupe binoculaire au laboratoire. La section d'un maximum de six phalanges par individu et de deux phalanges par membre était suffisante pour générer environ 3000 combinaisons uniques, permettant l'identification de l'ensemble de la population. Le code utilisé est présenté à la Figure 18. Chaque doigt correspond à un nombre, quand on regarde le lézard avec une incidence dorsale. Il faut alors faire la somme arithmétique des nombres correspondant aux différents doigts marqués pour obtenir le code numérique d'un individu. Notons qu'avec cette méthode de marquage, certaines combinaisons ne doivent pas être réalisées pour éviter une redondance du code.

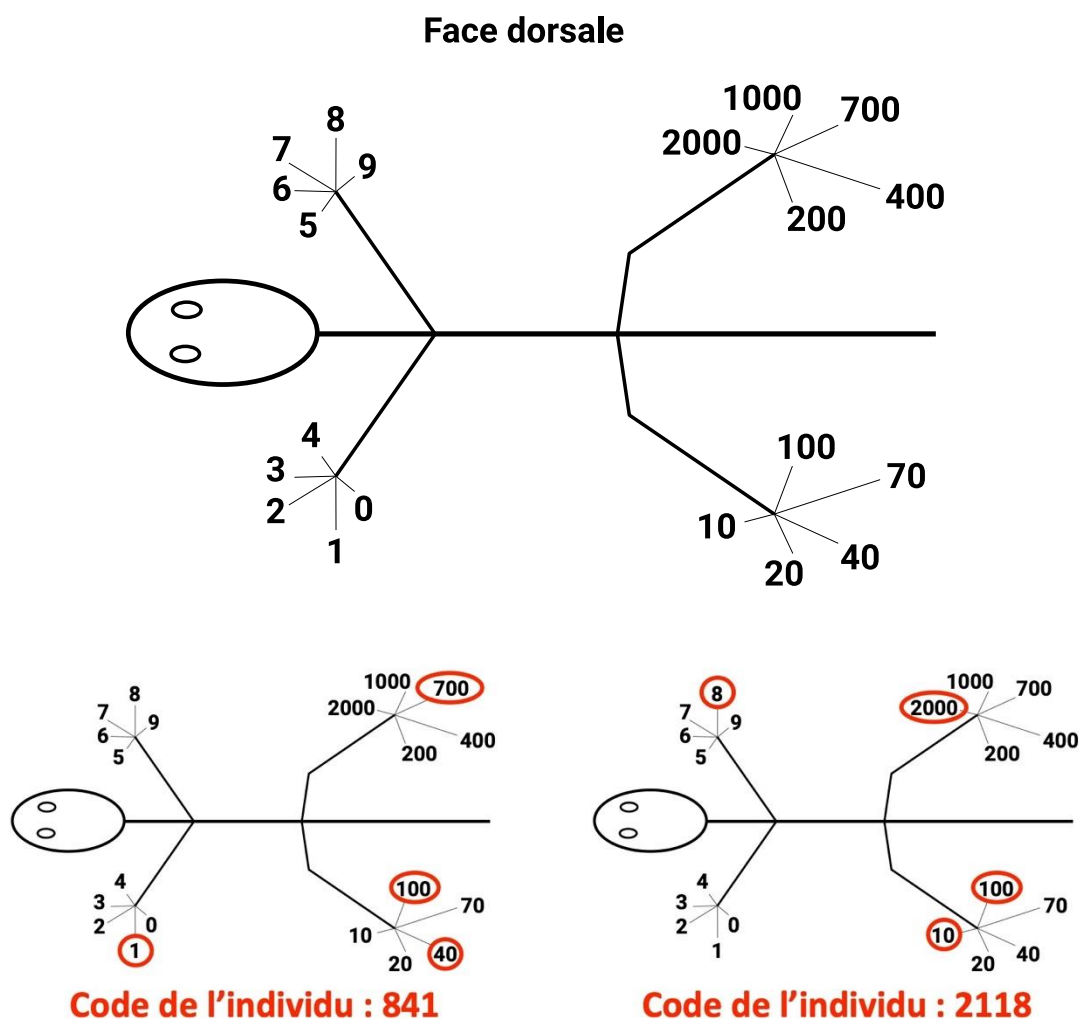


Figure 18 : Code utilisé pour le marquage de la population de lézards vivipares
Chaque doigt correspond à un nombre, quand on regarde le lézard avec une incidence dorsale. Le code numérique d'un individu est obtenu en faisant la somme arithmétique des nombres correspondant aux différents doigts marqués.

II. A. 1. v Élevage

Les femelles gestantes ont été maintenues en élevage entre la saison de capture et la mise-bas. Ce protocole a permis de reconstituer facilement les liens maternels au sein de la population. Dans un premier temps, les femelles ont été examinées et les éventuels parasites externes (tiques et acariens) ont été retirés. Un traitement à base de fipronil (Frontline) a été appliqué sur la face ventrale des individus à l'aide d'une lingette. Par la suite, les femelles ont été installées dans des terrariums individuels (11cm x 17cm x 12cm) contenant de la tourbe du commerce stérilisée à l'étuve pendant 1h à 150°C. La quantité de litière utilisée était suffisante pour permettre l'expression du comportement fouisseur du lézard vivipare. Des filets synthétiques, résistants à la chaleur, assuraient la fermeture des terrariums et permettaient une exposition directe à la chaleur émise par des ampoules à incandescence de 25W qui étaient allumées 6 à 8h par jour. L'éclairage était complété par les lampes UV-B spéciales reptiles et le laboratoire était éclairée entre 8h et 18h. La température et l'hygrométrie étaient

respectivement maintenues à 20-25°C et 50-60 %. Les terrariums étaient équipés d'un abri permettant aux animaux de se dissimuler et de se refroidir. Les lézards étaient nourris avec deux grillons domestiques tous les deux jours et une complémentation avec des larves de Pyralidées était apportée toutes les deux semaines. Les terrariums étaient hydratés trois fois par jour. L'état de santé des femelles était surveillé régulièrement et le poids était contrôlé tous les cinq jours. En cas de détection d'ectoparasites, le traitement à base de fipronil était renouvelé et le terrarium était changé. Ce traitement n'était toutefois pas réalisé si la mise-bas était prévue dans les jours qui suivaient. À la naissance, les jeunes étaient mesurés, marqués et un échantillon d'environ 2 mm de tissus était prélevé au niveau du bout de la queue. La mère et les nouveau-nés étaient ensuite relâchés dans les trois jours suivant la mise-bas, à la position exacte de la capture.

II. A. 1. vi Reconstitution de la parenté

Les échantillons de tissus caudaux ont été utilisés pour génotyper les marqueurs microsatellites de 8 sites distincts du génome. Les génotypes des nouveau-nés, des mères et des mâles ont été analysés par le logiciel CERVUS[64] qui infère la paternité des nouveau-nés à partir d'une modélisation probabiliste de la ségrégation des allèles. Ces analyses ont permis d'assigner l'ascendance paternelle de la majorité des nouveau-nés depuis l'année 2000.

II. A. 2 Analyses génétiques quantitatives sur la stratégie de croissance du lézard vivipare

II. A. 2. i Objectifs

Comme expliqué précédemment, la stratégie de croissance se définit comme l'évolution de la taille tout au long de la vie des individus. Or, il était impossible d'estimer l'héritabilité de la taille pour chaque classe d'âge dans la population étudiée ici car nous ne disposons que de certaines mesures à des âges donnés qui n'étaient pas identiques pour tous les individus. Ainsi, nous avons choisi de nous focaliser sur quatre variables permettant de caractériser la stratégie de croissance du lézard vivipare dans sa globalité. D'une part, nous nous sommes intéressés à la taille à la naissance, notée LM_{AN} et la taille au stade subadulte, notée LM_{SA} , qui étaient directement disponibles dans nos données (ou imputées par un modèle de régression linéaire, voir Partie II. A. 2. ii). L'analyse des composantes de la variance de ces deux premières variables a donc pu être réalisée avec des modèles animaux classiques, comme cela a été expliqué dans la section B.1.v de la première partie. D'autre part, nous nous sommes intéressés à deux paramètres estimés par un modèle de croissance : le taux de croissance, noté k , et la taille asymptotique, notée l_{max} . L'analyse des composantes de la variance de ces paramètres a été réalisée avec un modèle animal hiérarchique non-linéaire, comme cela a été expliqué dans la section B.2.ii de la première partie.

Dans cette sous-partie, nous donnons d'abord quelques éléments sur le formatage des données brutes et la création des jeux de données utilisés pour le reste des analyses. Nous expliquons ensuite le choix du modèle de croissance utilisé puis nous présentons l'analyse des composantes de la variance à l'aide de modèles animaux. Enfin, nous rappelons comment estimer l'héritabilité (h^2) et le coefficient de variation additif (CV_A) des variables de croissance.

II. A. 2. ii Formatage des données en vue de la modélisation

Les données analysées dans cette étude ont été collectées durant des campagnes annuelles de capture-marquage-recapture entre les années 1985 et 2019. Ces données comprenaient la longueur entre le museau et le cloaque (notée LMA pour « longueur museau-anus »), les liens de parenté, le sexe et l'âge des individus.

Dans un premier temps, les incohérences et les erreurs de saisie manuelle ont été recherchées et corrigées. Une redondance des identifiants pour des individus différents a été mise en évidence et rectifiée par la création d'un identifiant unique pour chaque individu. Par ailleurs, les incohérences de sexe au cours de la vie des individus ont été corrigées en prenant comme référence le sexe déterminé à l'âge le plus avancé. De même, l'âge estimé lors de la première observation des individus a servi de référence pour le calcul de l'âge lors des observations suivantes et pour le calcul de l'année de naissance. Une variable associée à l'identité des mères a ensuite été ajoutée et des identifiants maternels uniques ont été créés pour les mères inconnues. Enfin, le nombre d'observations réalisées chaque année pour chaque individu étant très variable (de 0 à 29 observations), nous avons choisi de garder uniquement la première observation réalisée chaque année pour chaque individu.

Dans un second temps, un modèle de régression linéaire multiple a permis d'imputer la taille à la naissance (LMA_N) pour les individus observés au stade subadulte, mais non observés à la naissance. Le modèle a été calibré avec les individus observés à la naissance et au stade subadultes entre 1985 et 2019. LMA_N a été estimée en fonction de la taille au stade subadulte (LMA_{SA}), du sexe et de l'année de naissance. Un terme d'interaction entre la taille au stade subadulte et le mois d'observation des subadultes a été ajouté pour prendre en compte la croissance des subadultes au cours de l'été, voir Annexes 1 et 3. L'année de naissance et le mois de mesure ont été considérés comme des variables catégorielles. Les hypothèses d'applications de la régression linéaire ont été vérifiées (Annexe 2) et le coefficient de détermination du modèle était $R^2 = 0,60$.

Nous avons choisi de ne pas inclure les individus nés avant 2000 dans le jeu de données final car la majorité des liens de parenté était inconnue. La fonction `fixPedigree()` du paquet `pedantics`[41] a été utilisée pour rajouter explicitement les individus fondateurs dans le pédigrée et la matrice d'apparentement A a été calculée à l'aide de la fonction `pedigreeStats()` du même paquet. La profondeur du pédigrée a été évaluée avec la fonction `kindepth()` du paquet

kinship2[65]. Le jeu de données final comprenait 11347 observations réalisées sur 5671 individus.

II. A. 2. iii Choix d'un modèle de croissance

Nous avons comparé trois modèles de croissance classiquement utilisés en biologie animale pour décrire l'évolution de la taille des reptiles en fonction de l'âge[66–72] : les modèles de croissance logistique (1), de Von Bertalanffy (2) et de Gompertz (3). Ces modèles permettent d'estimer la longueur entre le museau et l'anus des individus (LMA) en fonction de l'âge, de la taille à la naissance (l_0), du taux de croissance (k) et de la taille asymptotique (l_{max}). Les équations des modèles sont présentées ci-dessous.

$$LMA = \frac{l_{max}}{1 + \frac{l_{max} - l_0}{l_0} \times e^{-k \times Age}} \quad (1)$$

$$LMA = l_{max} - (l_{max} - l_0) \times e^{-k \times Age} \quad (2)$$

$$LMA = l_{max} \times e^{\log\left(\frac{l_0}{l_{max}}\right) \times e^{-k \times Age}} \quad (3)$$

Ces trois modèles ont été ajustés sur les données de LMAs en utilisant une approche bayésienne. Une sélection préalable des individus observés au moins trois fois a été nécessaire pour permettre la convergence des modèles. Nous nous sommes basés sur les connaissances actuelles[54]·[50]·[73]·[53] pour définir les *a priori* de l_{max} et l_0 et nous avons utilisé un *a priori* très peu informatif pour k , voir Annexe 4. Par ailleurs, nous avons fixé les coordonnées initiales pour l'échantillonnage des paramètres à $l_{max} = 68$, $l_0 = 21$ et $k = 1$, voir Annexe 5.

Les trois modèles de croissance ont été comparés par une approche de validation croisée Leave-One-Out[74]. Cette approche consiste à prédire chacune des observations du jeu de données en ajustant le modèle sans l'observation prédite. Le modèle montrant la valeur prédictive la plus importante[75], c'est-à-dire le modèle faisant le moins d'erreur sur ses prédictions, a été sélectionné pour le reste des analyses.

II. A. 2. iv Analyse des composantes de la variance des variables de croissance

II. A. 2. iv. (i) Analyse des composantes de la variance de LMA_N et LMA_{SA}

Nous avons utilisé deux modèles animaux classiques indépendants pour décomposer la variance phénotypique des variables LMA_N et LMA_{SA} . Dans ces modèles, le sexe des individus a été considéré comme un facteur fixe et l'année de naissance et l'identité maternelle ont été ajoutés comme facteurs aléatoires (voir la section B.1.v de la première partie pour plus de détails). L'équation des modèles est présentée ci-dessous.

$$y_i = \mu + MOTHER_i + ANNE_NAISSANCE_i + sexe_i + A_i + E_i$$

Avec y_i , le caractère phénotypique étudié (soit LMA_N ou LMA_{SA}), μ , la moyenne du caractère, A_i , les effets génétiques additifs, $MOTHER_i$ et $ANNE_NAISSANCE$, les effets aléatoires, $sexe_i$, l'effet fixe et E_i les résidus du modèle.

Les modèles ont été implémentés avec les *a priori* peu informatifs par défaut de brms[37], voir Annexe 4. Les distributions *a posteriori* des modèles ont été estimées avec quatre chaînes MCMC de 3000 itérations et 1000 itérations supplémentaires ont été écartées pour la convergence, voir Annexe 6. Les modèles sur LMA_N et LMA_{SA} ont respectivement été ajustés sur 3353 observations de nouveau-nés et sur 1614 observations de subadultes. La variance phénotypique totale de LMA_N et LMA_{SA} a été calculée directement à partir des observations. Pour chaque variable, les composantes de la variance ont été estimées par les moyennes de leurs distributions *a posteriori*. Nous avons vérifié que la moyenne et la médiane ne présentaient pas de différence majeure pour les différents paramètres et nous avons préféré utiliser la moyenne par soucis de praticité car il s'agit de l'estimateur par défaut des outils statistiques utilisés[37,44].

II. A. 2. iv. (ii) Analyse des composantes de la variance de l_{max} et k

Nous avons utilisé un modèle animal hiérarchique non-linéaire qui reprenait l'équation de croissance logistique (voir Partie I.B.2.ii). Dans un souci de simplification du modèle, le paramètre l_0 a été fixé à la valeur de LMA_N pour chaque individu. Les trois équations du modèle sont présentées ci-dessous.

$$LMA = \frac{l_{max}}{1 + \frac{l_{max} - LMA_N}{LMA_N} \times e^{-k \times Age}} + E_i$$

$$l_{max_i} = \mu + MOTHER_{1i} + ANNE_NAISSANCE_{1i} + sexe_{1i} + A_{1i} + PE_{1i}$$

$$k = \mu + MOTHER_{2i} + ANNE_NAISSANCE_{2i} + sexe_{2i} + A_{2i} + PE_{2i}$$

Comme pour les modèles animaux classiques, l'année de naissance ($ANNE_NAISSANCE$) et l'effet maternel ($MOTHER$) ont été inclus comme effets aléatoires et le sexe ($sexe$) a été inclus comme effet fixe. Un effet aléatoire de l'environnement permanent (PE) a également été ajouté pour estimer la variation de k et l_{max} qui ne peut pas être expliquée par les effets génétiques additifs.

Ce modèle a été ajusté avec les mêmes *a priori* et valeurs initiales que les modèles de croissance, voir Annexe 4. La distribution *a posteriori* du modèle a été estimés avec quatre

chaînes MCMC de 1500 itérations et 500 itérations supplémentaires ont été écartées pour la convergence, voir Annexe 7. Ce modèle a été ajusté sur les individus observés au moins deux fois et pour lesquels $LMAN$ était connue (ou imputée). Cela représentait 4618 observations pour 1848 individus. Pour chaque variable, les composantes de la variance ont été estimées par les moyennes de leurs distributions *a posteriori*. Là encore, nous avons vérifié que la moyenne et la médiane ne présentaient pas de différence majeure pour les différents paramètres du modèle.

II. A. 2. v Estimation de l'héritabilité et des coefficients de variation additifs des variables de croissance

L'héritabilité (h^2) des variables de croissances a été estimée à l'aide des formules suivantes (voir Partie I.B.1.v et B.2.ii pour plus de détails).

$$h^2 = \frac{V_A}{V_A + V_{AN} + V_M + V_R} \quad (1)$$

$$h^2 = \frac{V_A}{V_A + V_{AN} + V_M + V_{EP}} \quad (2)$$

L'équation (1) a été utilisée pour estimer l'héritabilité de $LMAN$ et LMA_{SA} et l'équation (2) pour estimer l'héritabilité de l_{max} et k . On rappelle que la variance génétique additive est notée V_A , la variance associée à l'année de naissance est notée V_{AN} , la variance associée à l'effet maternel est notée V_M , la variance associée à l'effet de l'environnement permanent est notée V_{EP} et la variance résiduelle est notée V_R .

Les coefficients de variation additifs (CV_A) ont été estimés en standardisant la racine carrée de la variance génétique additive (V_A) par la moyenne du caractère (μ_{trait}) pour les variables LMA_{SA} , $LMAN$, l_{max} et k .

$$CV_A = \frac{\sqrt{V_A}}{\mu_{trait}}$$

Contrairement à l'héritabilité, CV_A est une mesure de la variation génétique additive indépendante des autres sources de variance. Ce coefficient permet donc d'évaluer le potentiel évolutif en termes de proportion par rapport à la valeur moyenne du caractère.

II. A. 3 Estimation des gradients de sélection

Pour aller plus loin, nous nous sommes demandé si les variables de croissance $LMAN$, LMA_{SA} , l_{max} et k étaient soumises à des pressions de sélection au sein de la population étudiée. Pour répondre à cette problématique, nous avons estimé les gradients de sélection directionnelle (β) et les gradients de sélection stabilisante (γ) associés à ces quatre variables avec la méthode de Lande et Arnold[49] (voir Partie I, section D pour plus d'informations sur les gradients de sélection).

II. A. 3. i Présentation des modèles

Nous avons utilisé deux modèles indépendants pour étudier la sélection sur les variables LMA_N et LMA_{SA} , voir (1) et (2). Le taux de croissance et la taille asymptotiques des individus étant vraisemblablement liés, nous avons utilisé un modèle à deux caractères pour étudier la sélection sur k et l_{max} , voir (3). Des modèles différents ont été utilisés pour les mâles et les femelles de manière à différencier les gradients de sélection en fonction du sexe. Par ailleurs, l'année de naissance a été ajoutée comme facteur aléatoire dans tous les modèles. Les modèles utilisés pour estimer les gradients de sélection directionnelle (β) et stabilisante (γ) sont présentés ci-dessous.

$$w_i = \alpha + \beta_{sexe} LMA_{Ni} + \frac{1}{2} \gamma_{sexe} LMA_{Ni}^2 + AN + \varepsilon_i \quad (1)$$

$$w_i = \alpha + \beta_{sexe} LMA_{SAi} + \frac{1}{2} \gamma_{sexe} LMA_{SAi}^2 + AN + \varepsilon_i \quad (2)$$

$$w_i = \alpha + \beta_{sexe_1} k + \beta_{sexe_2} l_{max} + \frac{1}{2} \gamma_{sexe_1} k^2 + \frac{1}{2} \gamma_{sexe_2} l_{max}^2 + AN + \varepsilon_i \quad (3)$$

On a noté w_i le nombre de descendants de l'individu i , α l'intercept des modèles régression, AN l'effet associé à l'année de naissance et ε_i un terme d'erreur. Les variables LMA_N , LMA_{SA} , k et l_{max} ont été standardisées avant l'ajustement des modèles.

II. A. 3. ii Ajustement des modèles

Les variables LMA_N et LMA_{SA} ayant été mesurées, nous avons pu directement ajuster les modèles (1) et (2) sur des sous-sélection des jeux de données correspondant aux mâles et aux femelles, voir Annexe 8. Les paramètres k et l_{max} étant estimés par un modèle de croissance ajusté selon une approche bayésienne, nous n'avons pas directement accès aux valeurs exactes de ces paramètres pour chaque individu. Nous avons donc repris le modèle de croissance logistique en utilisant la fonction `save_pars()` du package `brms`[37] qui permet de conserver les paramètres l_{max} et k pour chaque individu, à chaque itération de la distribution *a posteriori* du modèle. Cent itérations ont ensuite été échantillonnées aléatoirement dans cette distribution *a posteriori* et cent jeux de données ont été édités pour chaque sexe à partir des itérations sélectionnées, en incluant les valeurs de l_{max} et k de l'itération pour tous les individus. Le nombre de descendants engendrés, l'année de naissance et les carrés de k et l_{max} ont également été inclus dans ces jeux de données, voir Figure 19 et Annexe 9.1. Le modèle (3) a été ajusté sur les 100 jeux de données mâles et femelles à l'aide de la fonction `brm_multiple()` du paquet `brms`[37], voir Annexe 9.2. Cette fonction permet de combiner les distributions *a posteriori* obtenues pour les gradients de sélection β et γ . L'ajustement des modèles selon une approche bayésienne a permis de prendre en compte de l'incertitude associée à l'inférence des paramètres k et l_{max} dans l'estimation des intervalles de crédibilité des gradients de sélection.

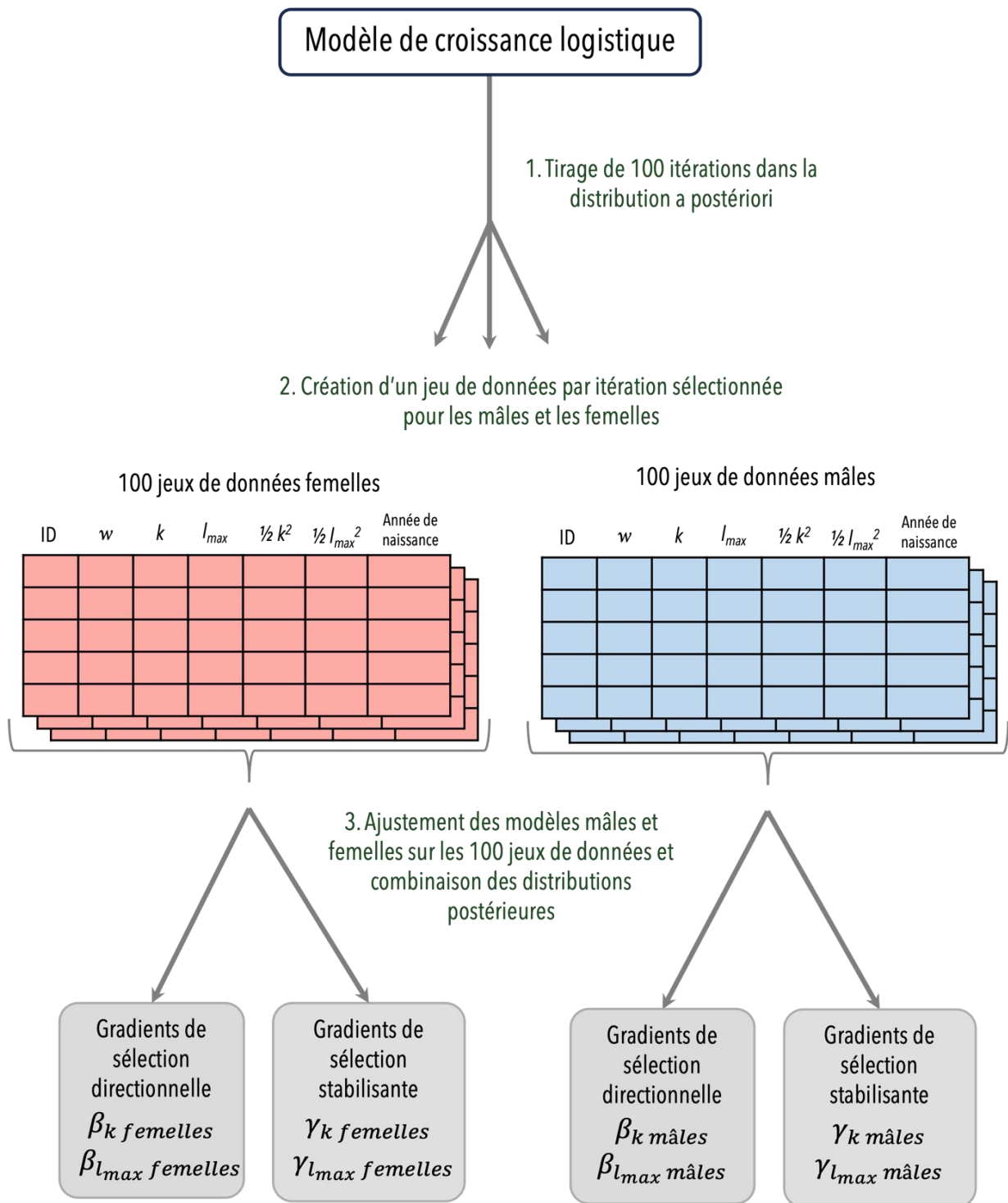


Figure 19 : Estimation des gradients de sélection β et γ sur les paramètres k et l_{max} pour les mâles et les femelles

Cent itérations ont été échantillonnées aléatoirement dans la distribution *a posteriori* du modèle logistique. Pour chaque sexe, cent jeux de données ont été édités à partir des itérations sélectionnées. Ces jeux de données incluaient l'identifiant des individus (ID), le nombre de descendants (w), l'année de naissance, les paramètres k , l_{max} ainsi que leurs carrés. Le modèle a été ajusté sur les 100 jeux de données mâles et femelles et les distributions *a posteriori* ont été combinées pour estimer les gradients de sélection β et γ

II. B Résultats

II. B. 1 Reconstitution du pédigrée

Parmi les 3392 individus utilisés pour l'ajustement des modèles animaux, les mères et les pères étaient connus pour respectivement 2435 et 1927 individus.

La profondeur du pédigrée de la population, qui correspond au nombre de générations entre les nouveau-nés capturés chaque année et leur plus vieil ancêtre identifié, est présentée à la Figure 20. La Figure 20.A représente le nombre absolu de nouveau-nés capturés chaque année et la profondeur de leur pédigrée, la Figure 20.B représente la profondeur du pédigrée en termes de proportion de nouveau-nés capturés chaque année et la Figure 20.C représente l'évolution de la profondeur moyenne du pédigrée des nouveaux nés. La profondeur maximale du pédigrée était de 6 générations entre 2000 et 2019. Nous pouvions noter la présence d'un appauvrissement global du pédigrée en 2010 avec un effondrement de la profondeur moyenne entre 2010 et 2015.



Figure 20 : Profondeur du pédigrée entre 2000 et 2019

La profondeur du pédigrée correspond au nombre de générations entre les nouveau-nés capturés chaque année et leur plus vieil ancêtre identifié.

II. B. 2 Analyses génétiques quantitatives sur la stratégie de croissance

II. B. 2. i Choix du modèle de croissance

Parmi les trois modèles de croissance testés (voir Figure 21), le modèle logistique présentait la meilleure valeur prédictive. Venait ensuite le modèle de Gompertz puis le modèle de Von Bertalanffy, (voir Tableau I). Après ajustement du modèle logistique, l_{max} était supérieur pour les femelles avec une moyenne de 65,48 mm (Intervalle de crédibilité (IC) = [65,13 ; 65,83]) contre seulement 59,13 mm (IC = [58,40 ; 59,87]) pour les mâles. En moyenne, l_0 était également légèrement supérieur pour les femelles avec 21,42 mm (IC = [21,21 ; 21,63]) contre 20,91 mm (IC = [20,53 ; 21,29]) pour les mâles. À l'inverse, le paramètre k était en moyenne supérieur pour les mâles avec $1,36 \text{ an}^{-1}$ (IC = [1,30 ; 1,41]) contre $1,29 \text{ an}^{-1}$ (IC = [1,26 ; 1,32]) pour les femelles.

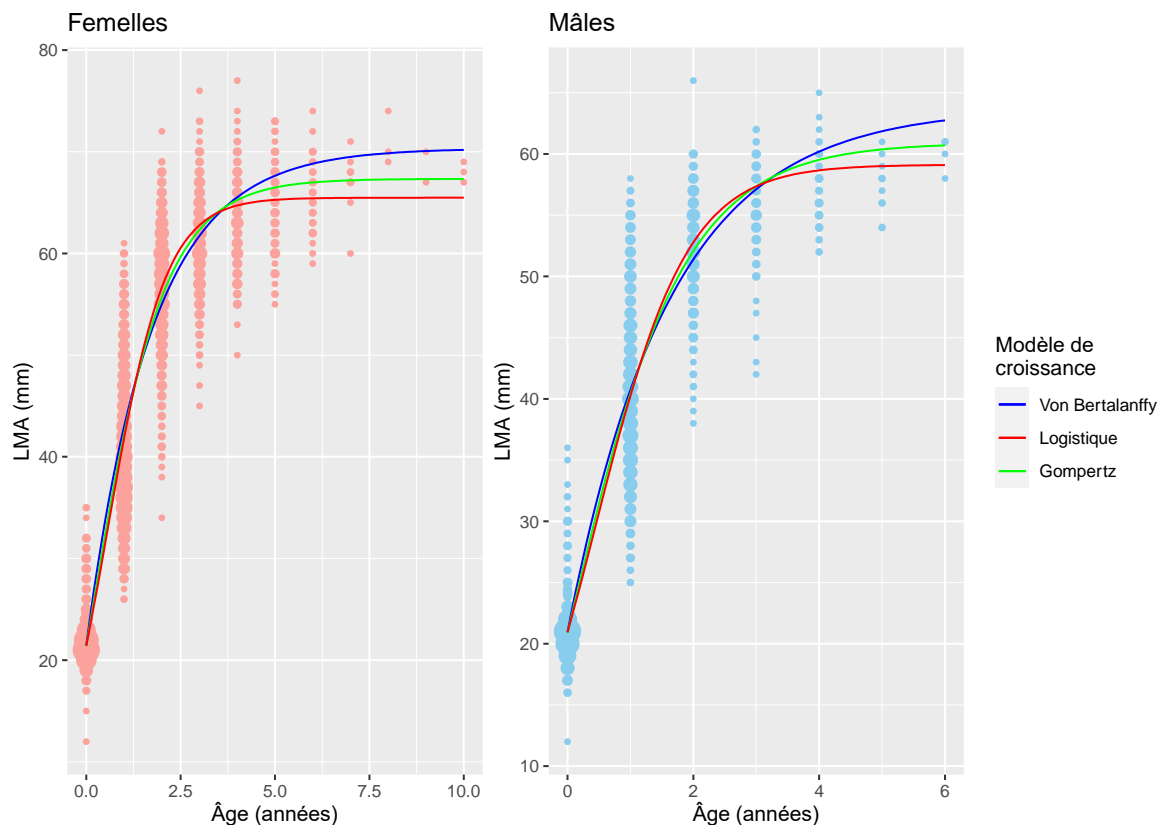


Figure 21 : Représentation graphique des données et des modèles de croissance pour les femelles (à gauche) et les mâles (à droite)

Représentation des données sous forme de cercles (rouges pour les femelles, bleus pour les mâles) et représentation des courbes de croissance des modèles de Von Bertalanffy (bleu), logistique (rouge) et de Gompertz (vert).

Tableau I : Valeur prédictive des modèles de croissance

Modèle	Valeur prédictive (ELPD*)	Différentiel des valeurs prédictive
Croissance logistique	-12869,2	0,0
De Von Bertalanffy	- 13413,7	- 544,5
De Gompertz	- 13125,4	-256,2

* ELPD : expected log pointwise predictive density

II. B. 2. ii Analyse des composantes de la variance

La variance phénotypique totale associée à la LMA était de 1,340 mm² à la naissance puis de 24,581 mm² au stade subadulte. L'année de naissance avait un effet majeur sur les deux mesures de taille et les deux paramètres de croissance. La variance expliquée par les effets génétiques additifs était particulièrement faible à la naissance ($V_A = 0,069$ mm²) mais augmentait substantiellement au stade subadulte ($V_A = 5,330$ mm²) et pour la taille asymptotique l_{max} ($V_A = 7,820$ mm²) (voir Tableau II et Figure 22).

Tableau II : Décomposition de la variance des variables de croissance et intervalles de crédibilités à 95%

	LMA _N	LMA _{SA}	l_{max}	k
Variance totale	1,340	24,581		
Variance génétique additive (V_A)	0,069 [0,028 ; 0,114]	5,330 [2,218; 8,932]	7,820 [0,184 ; 16,779]	0,011 [0,002 ; 0,020]
Année de naissance	0,642 [0,329 ; 1,214]	8,283 [4,086; 16,569]	17,637 [7,835 ; 37,451]	0,049 [0,024 ; 0,099]
Effet maternel	0,151 [0,122 ; 0,183]	1,550 [0,030 ; 3,433]	2,434 [0,011 ; 7,447]	0,004 [0 ; 0,009]
Effet de l'environnement permanent			7,527 [0,346 ; 16,181]	0,024 [0,016 ; 0,032]
Résidus	0,564 [0,523 ; 0,606]	11,061 [8,272; 13,710]		

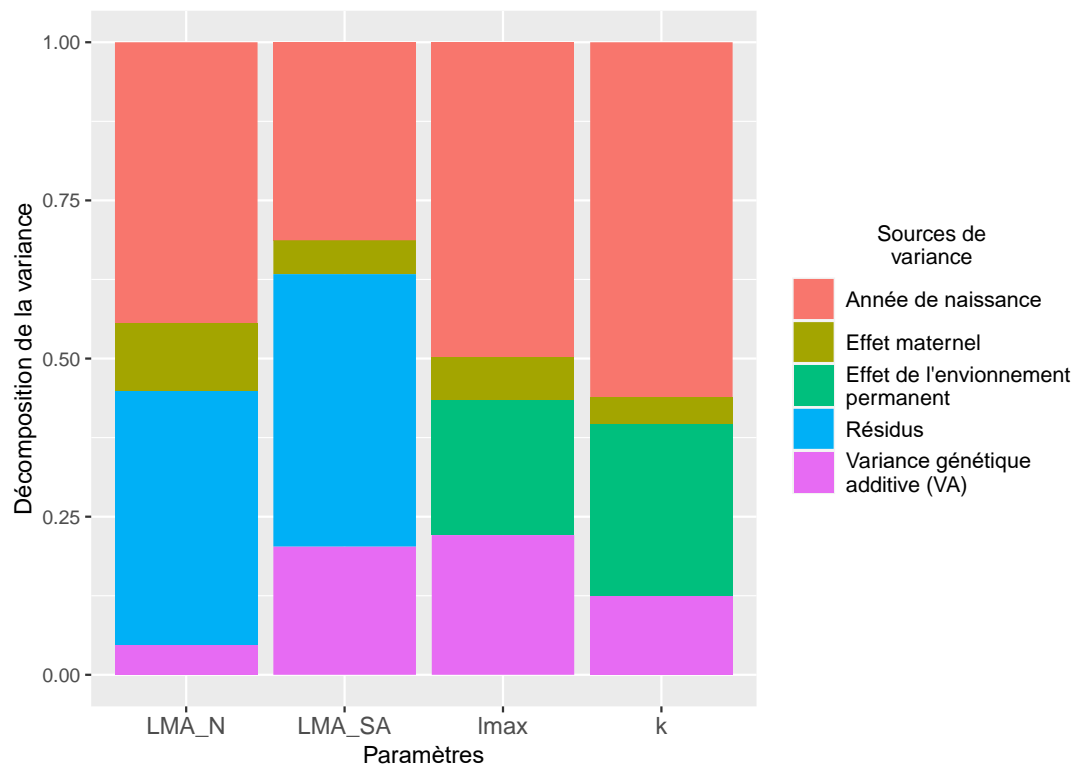


Figure 22 : Décomposition de la variance des variables de croissance
Décomposition de la variance de la taille à la naissance (LMA_N), la taille au stade subadulte (LMA_{SA}), la taille asymptotique (l_{max}) et le taux de croissance (k).

II. B. 2. iii Estimation de l'héritabilité (h^2) et des coefficients de variation additifs (CV_A)

L'héritabilité était minimale pour LMA_N ($h^2 = 0,050$ et $IC = [0,019 ; 0,086]$) et maximale pour l_{max} ($h^2 = 0,228$ et $IC = [0,005 ; 0,507]$). Les héritabilités de LMA_{SA} et k étaient respectivement estimées à $h^2 = 0,206$ ($IC = [0,083 ; 0,340]$) et $h^2 = 0,131$ ($IC = [0,025 ; 0,257]$), voir Figure 22.

Le coefficient de variation additif (CV_A) était maximal pour le paramètre LMA_{SA} avec $CV_A = 0,369$ et minimal pour LMA_N avec $CV_A = 0,013$. Par ailleurs, on a pu noter qu'il était plus important pour k ($CV_A = 0,077$) que pour l_{max} ($CV_A = 0,042$), voir Figure 23.

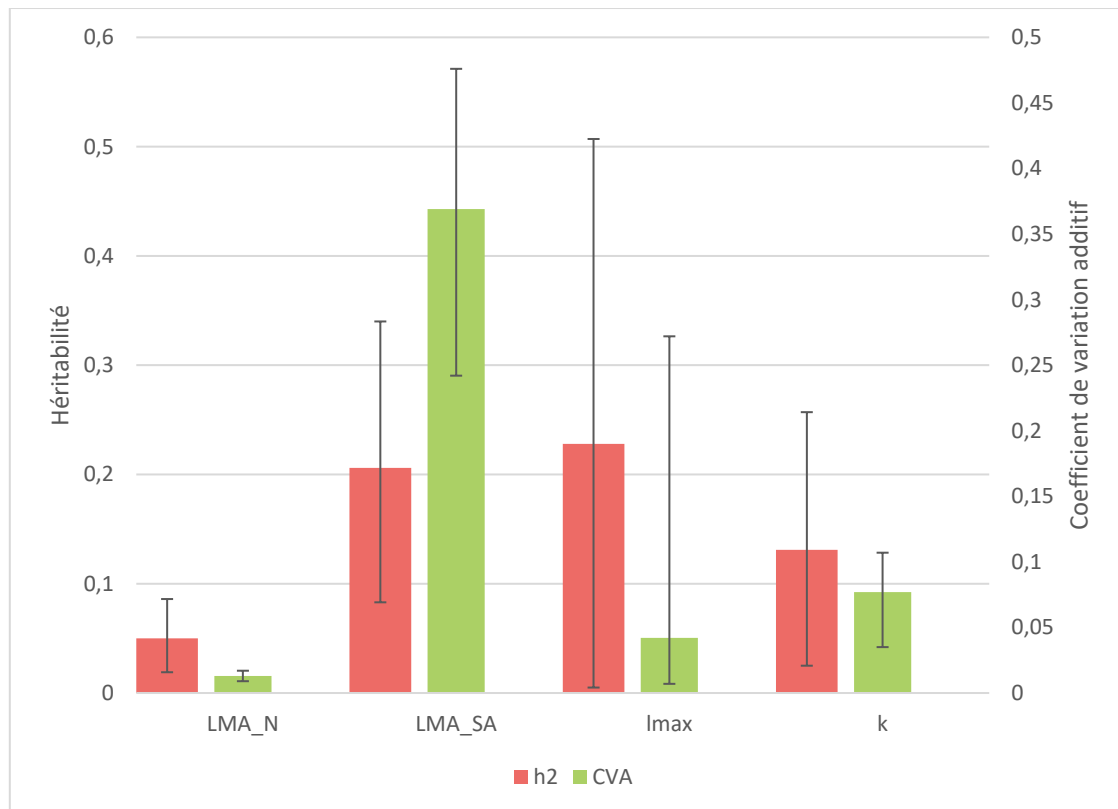


Figure 23 : Héritabilités et coefficients de variation additifs des variables de croissance. Héritabilités (h^2) et coefficients de variation additifs (CV_A) de la taille à la naissance (LMA_N), la taille au stade subadulte (LMA_{SA}), la taille asymptotique (l_{max}) et le taux de croissance (k). Les barres d'erreur représentent les intervalles de crédibilité à 95%.

II. B. 3 Estimation des gradients de sélection

Les graphiques présentés ci-dessous représentent les distributions *a posteriori* des gradients de sélection directionnelle (β) et stabilisatrice (γ) des variables LMA_N (Figure 24), LMA_{SA} (Figure 25), l_{max} et k (Figure 27) pour les mâles et les femelles. Les zones colorées en bleu ou en rouge correspondent aux parts des distributions *a posteriori* situées dans le positif. La moyenne des distributions *a posteriori* et leurs intervalles de crédibilités à 95% sont également indiqués pour chaque gradient de sélection.

Chez les mâles subadultes, le gradient de sélection directionnelle était significativement supérieur à 0 et le gradient de sélection stabilisante était significativement inférieur à 0. Une taille plus importante que la moyenne était donc associée à un meilleur succès reproducteur chez les mâles subadultes. Comme on peut le voir sur la surface de sélection individuelle des mâles subadultes présentée à la Figure 26, il existait cependant un optimum de la valeur sélective autour de + 3 mm par rapport à la taille moyenne. Les autres gradients de sélection n'étaient pas significativement différents de 0.

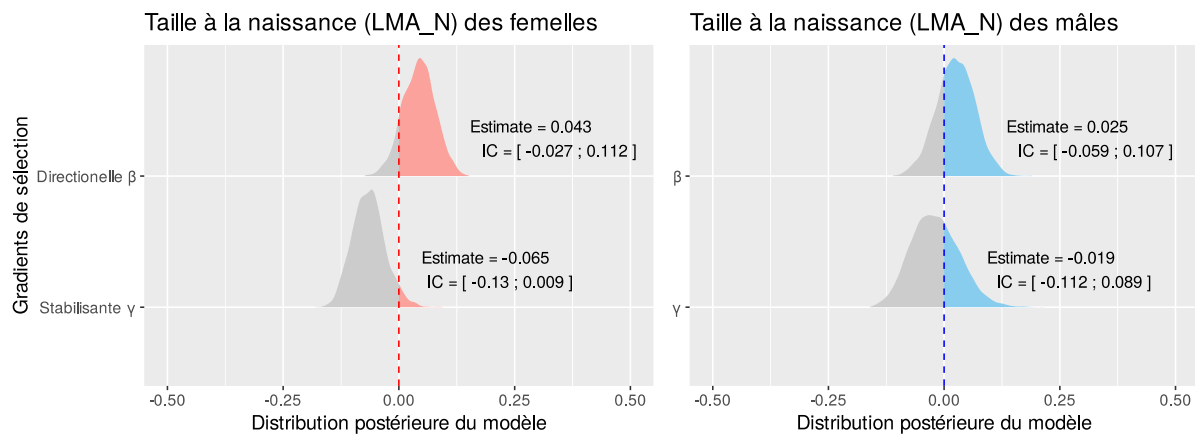


Figure 24 : Distributions *a posteriori* des gradients de sélection sur la taille à la naissance pour les femelles et les mâles
Les étiquettes correspondent aux moyennes (Estimate) et aux intervalles de crédibilités à 95% (IC) des distributions *a posteriori*.

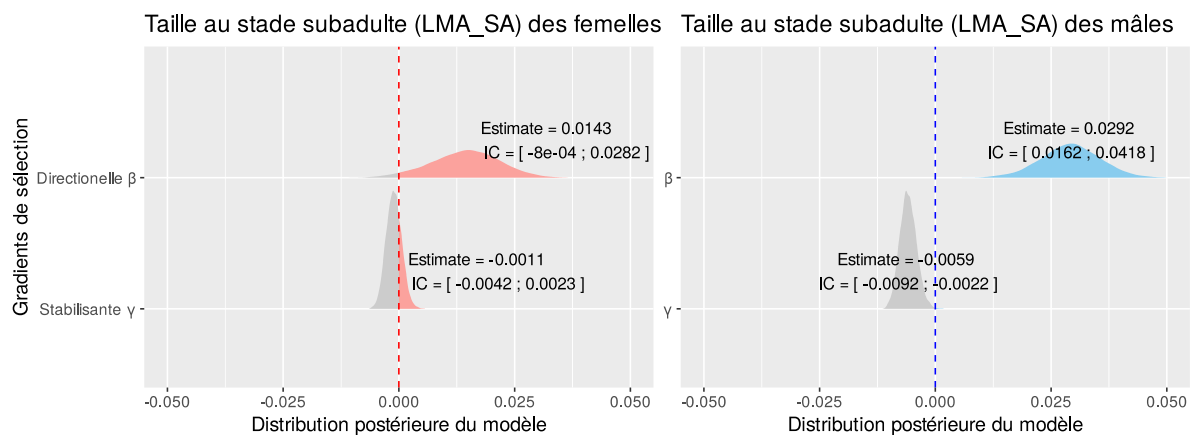


Figure 25 : Distributions *a posteriori* des gradients de sélection sur la taille au stade subadulte pour les femelles et les mâles
Les étiquettes correspondent aux moyennes (Estimate) et aux intervalles de crédibilités à 95% (IC) des distributions *a posteriori*.

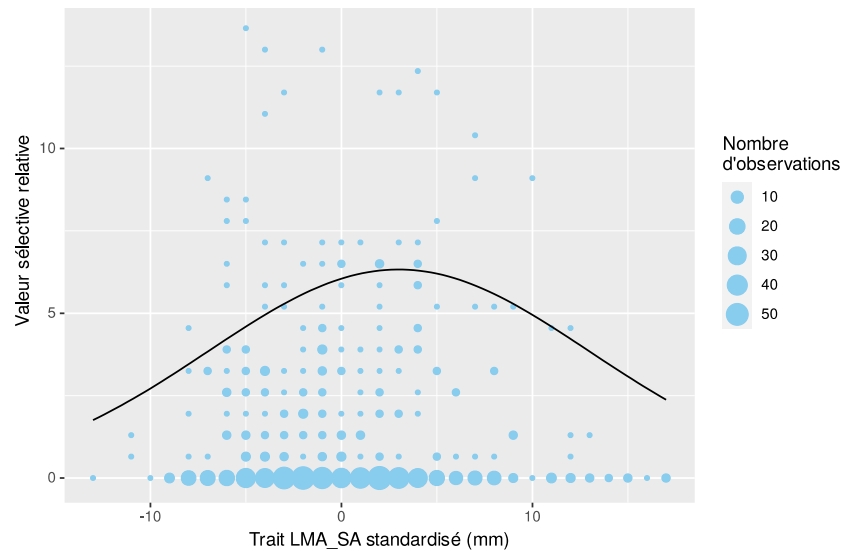


Figure 26 : Surface de sélection individuelle de la taille au stade subadulte chez les mâles
Représentation des données sous forme de cercles bleus et représentation de la surface de sélection individuelle (SSI) en noir.

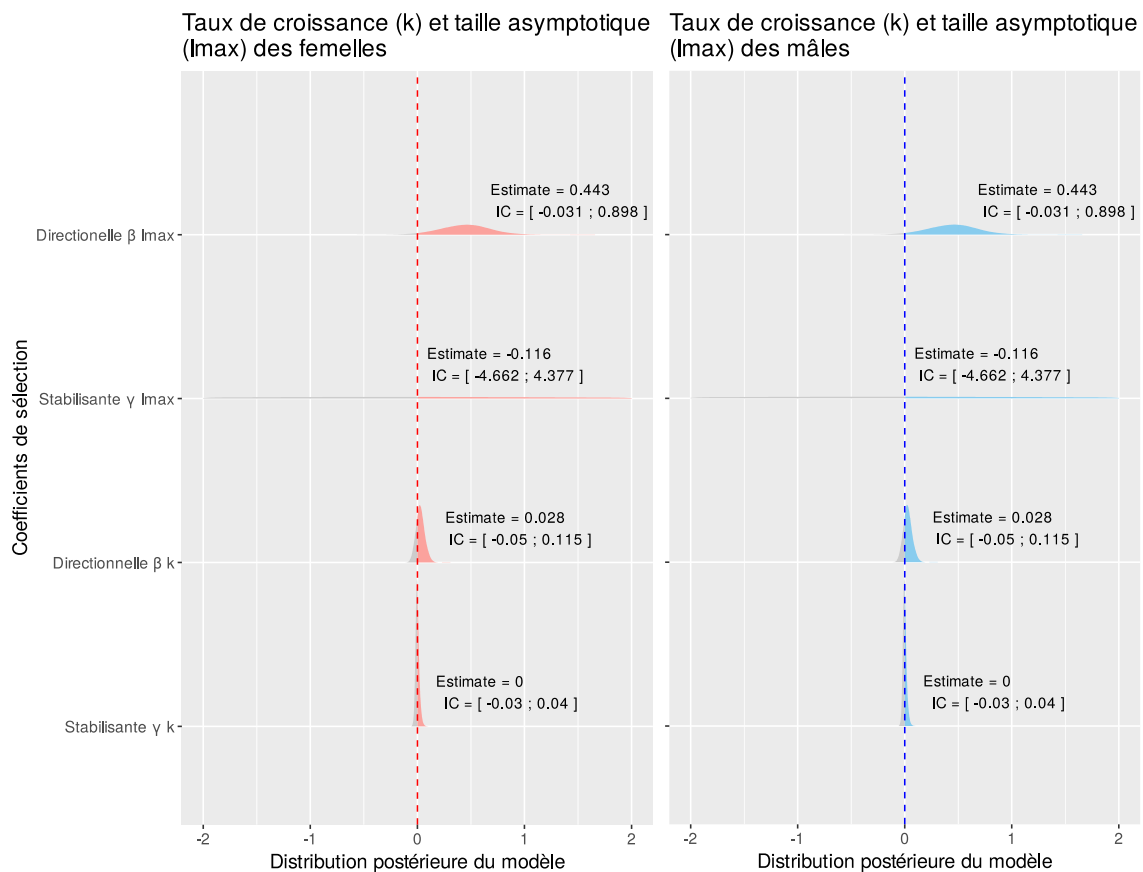


Figure 27 : Distributions *a posteriori* des gradients de sélection sur le taux de croissance et la taille asymptotique pour les femelles et les mâles
Les étiquettes correspondent aux moyennes (Estimate) et aux intervalles de crédibilités à 95% (IC) des distributions *a posteriori*.

II. B. 4 Évolution de la taille des individus

La taille moyenne des individus aux stades naissance et subadulte est présentée à la figure 28. Nous n'avons pas mis en évidence de modification majeure de ces mesures entre les années 1985 et 2019.

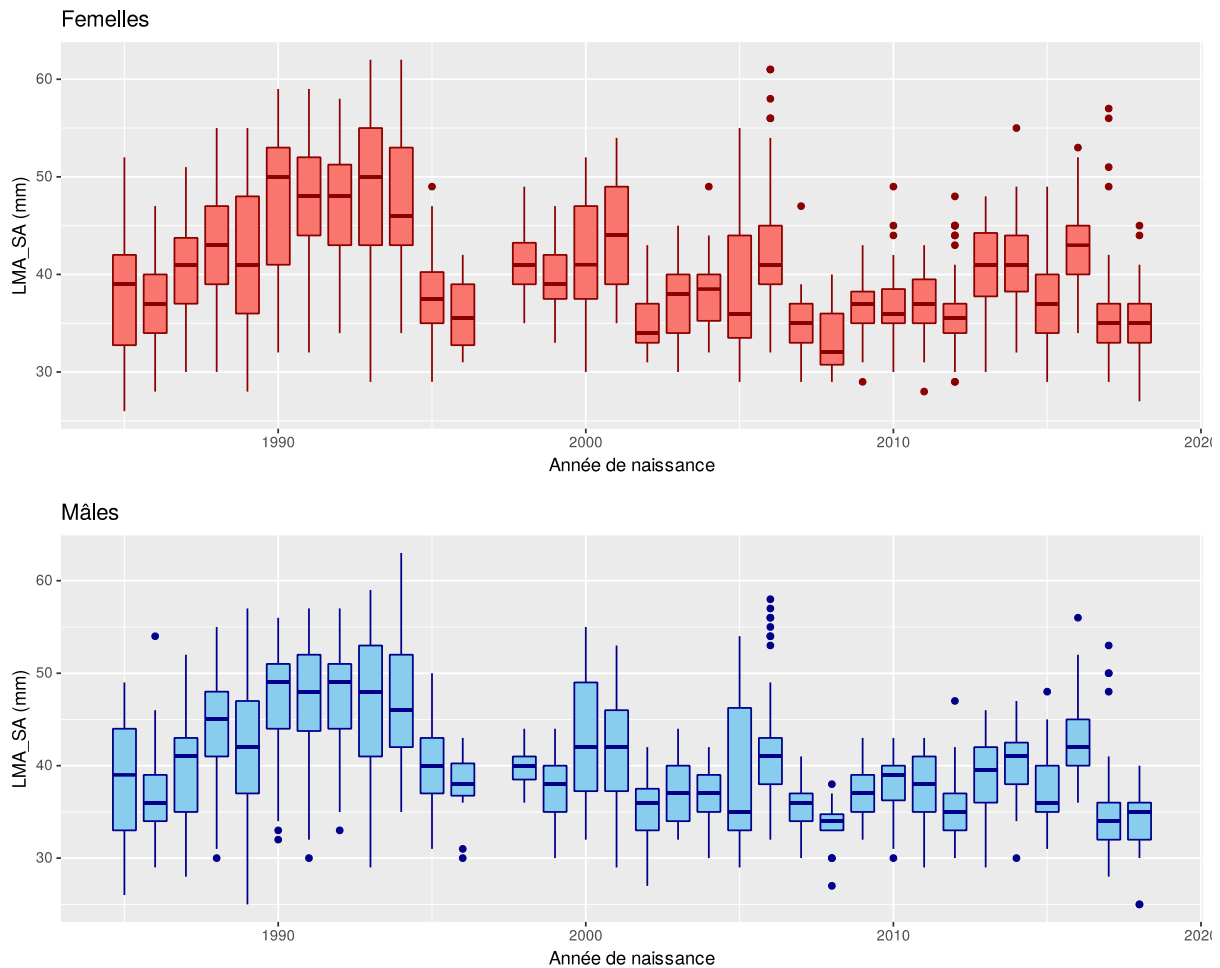


Figure 28 : Évolution de LMA_{SA} chez les femelles et les mâles entre 1985 et 2019.

II. C Discussion

II. C. 1 Modélisation des courbes de croissance

La croissance étant continue tout au long de la vie chez le lézard vivipare, il est impossible de définir un âge à partir duquel les individus atteignent leur taille maximale. Ainsi, les deux intérêts principaux de la modélisation des courbes de croissance sont l'estimation d'une taille asymptotique (l_{max}) pour chaque individu et la possibilité d'aller au-delà de l'étude de la taille à des âges donnés en s'intéressant à la dynamique de croissance dans sa globalité. À ce jour, les modèles de Von Bertalanffy, Gompertz et Logistique sont classiquement utilisés pour décrire des courbes de croissance en biologie animale[69,76–79]. Ces trois modèles ont plusieurs fois été comparés pour décrire la croissance de différentes espèces de lézards. Le modèle le plus adapté variait en fonction de l'espèce, du sexe et du caractère morphologique étudié[72]·[78–80]. Dans cette étude, les modèles de Von Bertalanffy, Gompertz et Logistique ont pour la première fois été ajustés sur la LMA d'une population de lézards vivipares, selon une approche statistique bayésienne.

Nos résultats suggèrent que le modèle de croissance logistique s'ajustait le mieux aux données. La croissance logistique a également été décrite comme le modèle le plus adapté pour modéliser l'évolution de la LMA chez les lézards *Agama impalearis*[78] et *Ctenophorus fionni*[79]. Notre modèle estimait une taille asymptotique l_{max} moyenne supérieure pour les femelles, ce qui est cohérent avec les connaissances actuelles sur la biologie du lézard vivipare[50,53,54]. Par ailleurs le taux de croissance k était supérieur de $0,068 \text{ an}^{-1}$ (IC = [0,014 ; 0,121]) pour les mâles, ce qui signifie que ces derniers atteignaient une taille maximale avant les femelles. Ce phénomène pourrait être en lien avec la compétition intraspécifique, entre les jeunes mâles et les mâles plus âgés, pour l'accès à la reproduction. Cette hypothèse est d'autant plus cohérente que les mâles vivent généralement moins longtemps que les femelles, le succès reproducteur durant les premières années de vie a donc un fort impact sur la valeur sélective des individus.

II. C. 2 Reconstitution du pedigree de la population

Construire un pedigree fiable et complet est l'un des défis majeurs du suivi d'une population naturelle. Dans notre étude, la mise en élevage des femelles reproductrices a permis d'accéder à la majorité des liens maternels. L'intégralité des femelles n'a cependant pas pu être capturée chaque année. La reconstitution de la paternité était plus complexe du fait de l'utilisation de marqueurs génétiques. Par ailleurs, les problèmes de gestion de la campagne de 2010 ont engendré un appauvrissement global du pedigree, avec un effondrement de la profondeur moyenne entre 2010 et 2015 (voir Figure 20). Notons que le faible nombre de nouveau-nés enregistrés en 2018 a traduit une chute des naissances cette année-là, dans un contexte de canicule.

Une connaissance incomplète du pédigrée pourrait vraisemblablement causer une sous-estimation globale des coefficients d'apparement au sein de la population. La variance génétique additive pourrait ainsi être légèrement sous-estimée au profit de la variance résiduelle ou de la variance liée à l'environnement permanent par exemple. Par ailleurs, une meilleure connaissance des liens maternels par rapport aux liens paternels pourrait théoriquement induire une sous-estimation de l'effet génétique additif au profit de l'effet maternel[81]. Cependant, l'effet maternel était ici relativement faible pour les quatre variables de croissances étudiées.

II. C. 3 Analyse des composantes de la variance

Une grande partie de la variance des quatre variables de croissance était expliquée par l'année de naissance. Ces résultats traduisent une forte plasticité phénotypique de la taille du lézard vivipare en réponse aux variations annuelles du milieu (conditions climatiques, nourriture disponible et taux de prédation par exemple). En effet, la population étudiée présentait un cycle biologique circannuel avec une période de reproduction unique, les individus nés la même année étaient donc généralement exposés à des variations environnementales similaires. Ainsi, il est cohérent qu'une grande part de la variance environnementale ait été absorbée par l'année de naissance. Chez des espèces qui se reproduisent tout au long de l'année, ou qui occupent des territoires plus hétérogènes, la variance environnementale peut être absorbée par des effets différents. La variance expliquée par l'année était notamment nettement plus faible dans des populations naturelles de mammifères[82,83]. Chez le suricate par exemple, une part importante de variance environnementale serait absorbée par l'effet portée[83].

Par ailleurs, la proportion de variance associée à l'effet maternel était faible pour les quatre variables de croissance étudiées, avec un maximum de 11,3 % pour LMA_N. En effet, il a été montré que l'effet maternel est généralement maximal durant les premiers stades de vie et diminue au cours du temps[5,84]. La taille de la mère n'étant pas corrélée à la taille des nouveau-nés chez le lézard vivipare[50,55], il n'est pas surprenant d'observer un effet maternel relativement faible dès la naissance. Des études réalisées en milieu contrôlé ont montré des résultats similaires chez le lézard *Lampropholis delicata*¹¹, alors que l'effet maternel expliquait jusqu'à 44,5 % de la variance de la LMA chez le lézard *Eulamprus quoyii*²¹. On notera cependant que les rares analyses issues d'individus vivant en milieux semi-naturels montraient un effondrement de l'effet maternel au profit de l'effet environnemental chez *Eulamprus quoyii*²¹.

II. C. 4 Étude du potentiel adaptatif de la stratégie de croissance

Nous avons évalué le potentiel adaptatif de la stratégie de croissance du lézard vivipare en étudiant la variance génétique additive (V_A) associée à différentes variables de croissance : la taille à la naissance (LMA_N), la taille au stade subadulte (LMA_{SA}), le taux de croissance (k) et la taille asymptotique (l_{max}). La part de V_A a été caractérisée par deux paramètres : l'héritabilité (h^2) qui permet d'évaluer la proportion de V_A par rapport à la variance totale du caractère, et le coefficient de variation additif (CV_A) qui permet de rapporter l'écart type associé à la V_A (donc sa racine carrée) à la moyenne du caractère. La mise en perspective des estimations de h^2 et de CV_A pour les quatre variables de croissance nous a permis d'évaluer précisément le potentiel adaptatif de la dynamique de croissance dans sa globalité.

Une étude de 2006 a évalué l'héritabilité de LMA_N dans des populations de lézards vivipares du massif du mont Lozère[54]. L'héritabilité était largement supérieure au résultat présenté ici, avec une estimation allant de 0,42 à 0,52 en fonction des expériences. Deux sources de biais majeures peuvent expliquer la surestimation de l'héritabilité dans cette étude : d'une part, les analyses concernaient des individus élevés en conditions contrôlées et, d'autre part, l'utilisation d'un modèle frère-sœur ne permettait pas de distinguer précisément les effets maternels et environnementaux des effets génétiques additifs. On peut noter qu'une étude menée sur le lézard *Eulamprus quoyii* a présenté des résultats assez proches de nos estimations concernant l'héritabilité de LMA_N [85]. Étant donné les faibles héritabilités de LMA_N et du paramètre l_0 dans un modèle préliminaire, nous avons choisi de simplifier le modèle animal non-linéaire final sur les courbes de croissance en remplaçant le paramètre l_0 par la mesure de LMA_N disponible dans nos données. L'augmentation de l'héritabilité de la LMA au cours de la vie des individus (LMA_N , LMA_{SA} et l_{max}) était cohérente avec plusieurs analyses génétiques quantitatives ayant mis en évidence une augmentation de l'héritabilité de caractères morphologiques avec l'âge[62,83,86,87]. Dans la population étudiée, l'augmentation de l'héritabilité était liée à une forte augmentation de V_A entre la naissance et les stades de développement plus avancés (voir Figure 9). Des observations similaires ont été réalisées chez le lézard *Lampropholis delicata*¹¹ tandis que, chez d'autres espèces comme le mouton domestique[87] par exemple, l'augmentation de l'héritabilité était liée à une diminution des autres sources de variance. Les différentes études ayant estimé l'héritabilité du taux de croissance chez les lézards présentaient des résultats très hétérogènes[54,85,88]. Il ne faut cependant pas perdre de vue que les définitions et méthodes de mesure du taux de croissance varient fortement selon les études. La comparaison des estimations d'héritabilité est donc difficilement interprétable. À notre connaissance, l'héritabilité des paramètres d'un modèle de croissance n'avait jamais été évaluée chez les reptiles. Les courbes de croissance d'une population naturelle de suricates ont été simulées par un modèle monomoléculaire[83] et l'héritabilité de la taille asymptotique était assez proche des résultats présentés ici, avec une estimation à 0,230. L'héritabilité associée au taux de croissance n'était, en revanche, pas

significativement différente de zéro. L'héritabilité étant maximale pour le paramètre l_{max} , on s'attendait à ce que la taille asymptotique soit le caractère qui réponde le mieux à la sélection dans l'absolu. Cependant, on a noté que le coefficient de variation additif associé à l_{max} était relativement faible. La variable LMA_{SA} présentait à la fois une forte héritabilité et un coefficient de variation additif important, voir Figure 22. Ces résultats suggèrent que la sélection pourrait avoir un effet proportionnellement plus important sur la taille au stade subadulte que sur la taille asymptotique.

II. C. 5 Étude des pressions de sélection au sein de la population étudiée

Tout d'abord, notons que les températures estivales moyennes du mont Lozère n'ont pas augmenté significativement au cours des deux dernières décennies. Nous n'avons pas non plus enregistré de déclin majeur de l'effectif de la population étudiée. De fait, rien ne laisse à penser que les pressions de sélection mises en évidence dans cette étude aient pu être en lien avec le dérèglement climatique.

Comme détaillé plus haut, nous connaissions environ $\frac{2}{3}$ des relations de parenté au sein de la population étudiée. Nous avons ainsi pu obtenir des estimations relativement fiables des gradients de sélection linéaire et non-linéaire. On note cependant que les intervalles de crédibilité des gradients associés à k étaient particulièrement larges. Cela était dû à un manque de puissance du modèle, lié à la prise en compte de l'incertitude sur l'estimation de k et l_{max} . Davantage de données seraient nécessaires pour obtenir des estimations plus précises. L'analyse a montré une sélection directionnelle vers l'augmentation de la taille au stade subadulte pour les mâles. Il est vraisemblable qu'une grande LMA_{SA} soit associée à une taille plus importante à la maturité sexuelle et il a été montré que les mâles de grande taille présentaient un meilleur succès reproducteur [54]. La taille des mâles subadultes était également soumise à une sélection stabilisante, avec un optimum autour de + 3 mm (voir figure 25). Au-delà de la taille optimale, les mâles pourraient être pénalisés par l'augmentation du besoin énergétique sans retirer d'avantages sur le plan reproductif. Des compromis pourraient également exister entre l'énergie allouée à la croissance et l'énergie allouée aux comportements sexuels, à la maturation sexuelle ou à la constitution de ressources en vue de l'hiver par exemple. Des mécanismes sélectifs similaires ont été mis en évidence chez les femelles d'une population naturelle de lézards *Anolis sagrei* [89,90]. Malgré les forces de sélection mises en évidence sur LMA_{SA} , nous n'avons pas constaté de réponse phénotypique de la taille au stade subadulte au cours des 30 dernières années chez les mâles dans la population étudiée (voir Figure 28). Même si une taille importante des mâles au stade subadulte confère un avantage sélectif dans l'absolu, LMA_{SA} est possiblement liée à d'autres variables soumises à sélection, comme la condition corporelle ou le profil d'activité par exemple. Des mécanismes de sélection antagonistes pourraient alors expliquer la conservation de la taille au stade subadultes chez les mâles [5].

II. C. 6 Mise en perspective des résultats avec le contexte de réchauffement climatique

Selon la théorie métabolique de l'écologie[58,91], l'augmentation du coût métabolique lié à l'élévation de la température tend à être compensée par une diminution de la taille chez les ectothermes. Une méta-analyse a notamment montré que le réchauffement climatique entraînait une diminution de la taille moyenne des ectothermes vivant en milieu aquatiques[91]. Ce phénomène s'explique, d'une part, par l'augmentation de la proportion de jeunes au sein des populations et d'autre part par la diminution de la taille moyenne au sein des classes d'âge. Cependant, chez le lézard vivipare, l'élévation des températures est généralement associée à une augmentation de la taille moyenne au sein des classes d'âge juvénile et subadulte[22,23]. L'influence de la température sur la taille à l'âge adulte n'a cependant jamais été évaluée à notre connaissance. Une étude menée sur 15 populations naturelles de lézards vivipares ovipares (*Zootoca vivipara louislantzi*) vivant dans le sud-ouest de la France a récemment évalué le lien entre la température et la LMA des mâles[92]. Comme pour les populations vivipares, la LMA des mâles à un an était plus grande dans les environnements chauds. Cependant, il est intéressant de noter que chez les mâles ayant atteint la maturité sexuelle, la température avait un effet négatif sur la LMA.

Les résultats présentés ici suggèrent que le taux de croissance et la taille à partir du stade subadulte pourraient répondre à des pressions de sélection. On peut ainsi imaginer deux types de réponses adaptatives qui pourraient participer au maintien des populations dans un contexte d'élévation des températures (voir Figure 29). D'une part, l'augmentation de la taille durant les premières années de vie pourrait entraîner une hausse de la fécondité globale de la population. Au-delà de l'augmentation de la taille des portées[50,55], les subadultes de grande taille exposés à des climats chauds et ensoleillés pourraient atteindre la maturité sexuelle dès la première année, comme cela a déjà été rapporté en conditions naturelles et expérimentales[22,53,54]. D'autre part, la sélection des individus présentant une taille modérée à l'âge adulte pourrait permettre de contrer l'augmentation du taux de croissance lié à la plasticité phénotypique dans un contexte d'élévation des températures. Ainsi, la conservation d'une taille modérée à l'âge adulte pourrait limiter la hausse des besoins métaboliques des individus et conférer des avantages sur le plan de la thermorégulation.

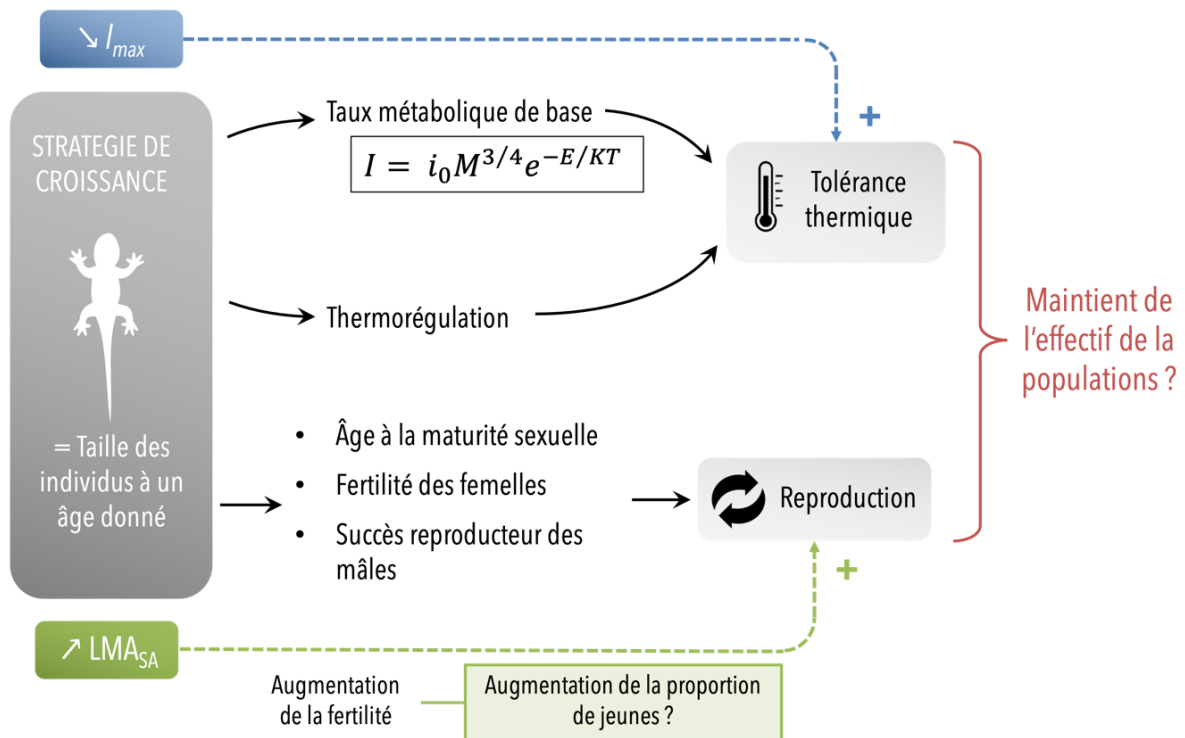


Figure 29 : Possibles réponses adaptatives de la stratégie de croissance du lézard vivipare dans un contexte d'élévation des températures.

D'une part l'augmentation de la taille durant les premières années de vie pourrait entraîner une hausse de la fécondité globale de la population (augmentation de la taille des portées, reproduction dès la première année). D'autre part, la sélection des individus de taille modérée à l'âge adulte pourrait limiter la hausse des besoins métaboliques et conférer des avantages sur le plan de la thermorégulation.

Conclusion

La génétique quantitative permet d'étudier la transmission d'un phénotype complexe entre les générations en estimant la variance génétique additive des caractères quantitatifs. Cette discipline a de nombreux domaines d'application en biologie animale, que ce soit en élevage, en biologie évolutive ou dans le milieu médical.

L'objectif de cette étude était d'étudier le potentiel adaptatif de la stratégie de croissance d'une espèce modèle de lézard vivant en milieu tempéré (*Zootoca vivipara*). Des modèles statistiques de décomposition de la variance ont permis d'estimer l'héritabilité (h^2) et le coefficient de variation additif (CV_A) de quatre variables de croissance d'une population naturelle de lézards vivipares. L'utilisation de l'inférence bayésienne pour ajuster ces modèles mixtes hiérarchiques complexes a permis d'estimer l'incertitude associée à h^2 et CV_A . Nous avons montré que l'héritabilité de la taille du lézard vivipare était très faible à la naissance ($h^2 = 0,050$), augmentait fortement au stade subadulte ($h^2 = 0,206$) et était maximale pour la taille asymptotique ($h^2 = 0,228$). Le taux de croissance présentait une héritabilité intermédiaire ($h^2 = 0,131$). Le coefficient de variation additif était maximal pour la taille au stade subadulte ($CV_A = 0,369$) et minimal pour la taille à la naissance ($CV_A = 0,013$). Il était intermédiaire pour la taille asymptotique ($CV_A = 0,042$) et le taux de croissance ($CV_A = 0,077$). Ces résultats étaient en faveur d'un fort potentiel évolutif de la stratégie de croissance du lézard vivipare à partir du stade subadulte. Nous avons également mis en évidence une pression de sélection vers l'augmentation de la taille des mâles au stade subadulte avec un optimum autour de +3 mm par rapport à la taille moyenne dans la population étudiée.

La taille jouant un rôle important dans la tolérance thermique chez les ectothermes, un fort potentiel évolutif de la stratégie de croissance du lézard vivipare est en faveur d'un fort potentiel adaptatif de l'espèce face à l'élévation des températures. Cependant, une meilleure compréhension des mécanismes impliqués dans le déclin des populations de lézards, constaté dans le sud de l'Europe, est nécessaire pour évaluer quels pourraient être les avantages conférés par une modification de la taille dans un contexte de changement climatique.

Bibliographie

- [1] Lynch M and Walsh B 1998 *Genetics and Analysis of Quantitative Traits*. vol 1 (London: Oxford University Press), 980 p.
- [2] Hill W G and Mackay T F C 2004 D. S. Falconer and Introduction to quantitative genetics. *Genetics* **167** 1529–1536
- [3] Zuidhof M J, Schneider B L, Carney V L, Korver D R and Robinson F E 2014 Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 2005 *Poultry Science* **93** 2970–2982
- [4] Hill W G and Kirkpatrick M 2010 What Animal Breeding Has Taught Us about Evolution *Annual Review of Ecology, Evolution, and Systematics* **41** 1–19
- [5] Charmantier A, Garant D and Kruuk L E B 2014 *Quantitative genetics in the wild* (Oxford: Oxford University Press), 280 p.
- [6] McCairns R J S, Smith S, Sasaki M, Bernatchez L and Beheregaray L B 2016 The adaptive potential of subtropical rainbowfish in the face of climate change: heritability and heritable plasticity for the expression of candidate genes *Evolutionary Applications* **9** 531–545
- [7] Bairos-Novak K R, Hoogenboom M O, van Oppen M J H and Connolly S R 2021 Coral adaptation to climate change: Meta-analysis reveals high heritability across multiple traits *Global Change Biology* **27** 5694–5710
- [8] Boutin S and Lane J E 2014 Climate change and mammals: evolutionary versus plastic responses *Evolutionary Applications* **7** 29–41
- [9] Visser M E 2008 Keeping up with a warming world; assessing the rate of adaptation to climate change *Proceedings of the Royal Society B: Biological Sciences* **275** 649–659
- [10] Bliss S, Todhunter R J, Quaas R, Casella G, Wu R, Lust G, Williams A J, Hamilton S, Dykes N L, Yeager A, Gilbert R O, Burton-Wurster N I and Acland G M 2002 Quantitative genetics of traits associated with hip dysplasia in a canine pedigree constructed by mating dysplastic Labrador Retrievers with unaffected Greyhounds *American Journal of Veterinary Research* **63** 1029–1035
- [11] Low M, Eksell P, Högström K, Olsson U, Audell L and Ohlsson Å 2019 Demography, heritability and genetic correlation of feline hip dysplasia and response to selection in a health screening programme *Sci Rep* **9** 17164
- [12] Hays L, Zhang Z, Mateescu R G, Lust G, Burton-Wurster N I and Todhunter R J 2007 Quantitative genetics of secondary hip joint osteoarthritis in a Labrador Retriever–Greyhound pedigree *American Journal of Veterinary Research* **68** 35–41
- [13] Baers G, Keller G G, Famula T R and Oberbauer A M 2019 Heritability of Unilateral Elbow Dysplasia in the Dog: A Retrospective Study of Sire and Dam Influence *Frontiers in Veterinary Science* **6**
- [14] Shaw S C, Wood J L N, Freeman J, Littlewood J D and Hannant D 2004 Estimation of heritability of atopic dermatitis in Labrador and Golden Retrievers *ajvr* **65** 1014–1020
- [15] Lewis T, Swift S, Woolliams J A and Blott S 2011 Heritability of premature mitral valve disease in Cavalier King Charles spaniels *The Veterinary Journal* **188** 73–76
- [16] Schlensker E and Distl O 2016 Heritability of hemivertebrae in the French bulldog using an animal threshold model *The Veterinary Journal* **207** 188–189
- [17] Salonen M, Vapalahti K, Tiira K, Mäki-Tanila A and Lohi H 2019 Breed differences of heritable behaviour traits in cats *Sci Rep* **9** 7949
- [18] Ruefenacht S, Gebhardt-Henrich S, Miyake T and Gaillard C 2002 A behaviour test on German Shepherd dogs: heritability of seven different traits *Applied Animal Behaviour Science* **79** 113–132
- [19] Hradecká L, Bartoš L, Svobodová I and Sales J 2015 Heritability of behavioural traits

- in domestic dogs: A meta-analysis *Applied Animal Behaviour Science* **170** 1–13
- [20] Villani N A 2011 *Heritability and Characteristics of Catnip Response in Two Domestic Cat Populations* Master of science in Animal Biology (California: University of California), 44 p.
- [21] Mallard F and Couderchet L 2019 Climate Sentinels Research Program: Developing Indicators of the Effects of Climate Change on Biodiversity in the Region of New Aquitaine (South West, France) *Handbook of Climate Change and Biodiversity* ed W Leal Filho, J Barbir and R Preziosi (Cham: Springer International Publishing) pp 223–241
- [22] Bestion E, Teyssier A, Richard M, Clobert J and Cote J 2015 Live Fast, Die Young: Experimental Evidence of Population Extinction Risk due to Climate Change ed G M Mace *PLoS Biol* **13** 1002281
- [23] Dupoué A, Rutschmann A, Le Galliard J F, Clobert J, Angelier F, Marciau C, Ruault S, Miles D and Meylan S 2017 Shorter telomeres precede population extinction in wild lizards *Sci Rep* **7** 16976
- [24] R core Team 2022 R : A language and Environment for Statistical Computation [en ligne], URL : <https://www.R-project.org/> [consulté le 01 juillet 2023]
- [25] Anon 2003 The nature and identification of quantitative trait loci *Nat Rev Genet* **4** 911–916
- [26] East E M 1911 The Genotype Hypothesis and Hybridization *The American Naturalist* **45** 160–174
- [27] Fisher R A 1919 The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans. R. Soc. Edinb.* **52** 399–433
- [28] Szalai C, László V, Pap E, Falus A and Oberfrank F 2014 *Genetics and genomics* (Budapest: Typotex Kiadó), 206 p.
- [29] De Villemereuil P 2018 Quantitative genetic methods depending on the nature of the phenotypic trait: Quantitative genetics for non-classical traits *Ann. N.Y. Acad. Sci.* **1422** 29–47
- [30] Roff D A 2007 A Centennial Celebration For Quantitative Genetics *Evolution* **61** 1017–1032
- [31] Garcia-Gonzalez F, Simmons L W, Tomkins J L, Kotiaho J S and Evans J P 2012 Comparing Evolvabilities : common errors surrounding the calculation and use of coefficients of additive genetic variation *Evolution* **66** 2341–2349
- [32] Wilson A J, Réale D, Clements M N, Morrissey M M, Postma E, Walling C A, Kruuk L E B and Nussey D H 2010 An ecologist’s guide to the animal model *Journal of Animal Ecology* **79** 13–26
- [33] Pemberton J M 2010 Evolution of quantitative traits in the wild: mind the ecology *Phil. Trans. R. Soc. B* **365** 2431–2438
- [34] Westphal M F and Morgan T J 2010 Quantitative Genetics of Pigmentation Development in 2 Populations of the Common Garter Snake, *Thamnophis sirtalis* *Journal of Heredity* **101** 573–580
- [35] Brooks S, Gelman A, Jones G and Meng X-L 2011 *Handbook of Markov Chain Monte Carlo* (Boca Raton: CRC Press), 620 p.
- [36] Royle J A and Dorazio R M 2008 *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities* (London: Elsevier), 463 p.
- [37] Bürkner P-C, Gabry J, Weber S, Johnson A, Modrak M, Badr H S, Weber F, Ben-Shachar M S, Rabel H and Mills S C 2022 Brms: Bayesian Regression Models using “Stan” *Journal of Statistical Software* **80** 1–28
- [38] CRAN team 2023 The Comprehensive R Archive Network [en ligne], URL : <https://cran.r-project.org/> [consulté le 01 juillet 2023]
- [39] Gelman A, Lee D and Guo J 2015 Stan: A Probabilistic Programming Language for

- [40] Hecht M, Weirich S and Zitzmann S 2021 Comparing the MCMC Efficiency of JAGS and Stan for the Multi-Level Intercept-Only Model in the Covariance- and Mean-Based and Classic Parametrization *Psych* **3** 751–779
- [41] Morrissey M B and Wilson A J 2010 pedantics: an r package for pedigree-based genetic simulation and pedigree manipulation, characterization and viewing *Mol Ecol Resour* **10** 711–719
- [42] Martin J and Videlier M Brms | How to fit an animal model [en ligne], URL : https://juliengamartin.github.io/wam_tuto/brms-1.html [consulté le 01 juillet 2023]
- [43] De Villemereuil P 2018 Quantitative genetic methods depending on the nature of the phenotypic trait: Quantitative genetics for non-classical traits *Ann. N.Y. Acad. Sci.* **1422** 29–47
- [44] Bürkner P-C 2017 **brms** : An R Package for Bayesian Multilevel Models Using *Stan J. Stat. Soft.* **80**
- [45] Vehtari A, Gelman A and Gabry J 2017 Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC *Stat Comput* **27** 1413–1432
- [46] Gelman A, Carlin J B, Stern H S, Dunson D B, Vehtari A and Rubin D B 2013 *Bayesian Data Analysis, Third Edition* (CRC Press), 677 p.
- [47] Fear K K and Price T 1998 The Adaptive Surface in Ecology *Oikos* **82** 440
- [48] Stinchcombe J R, Agrawal A F, Hohenlohe P A, Arnold S J and Blows M W 2008 Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution* **62** 2435–2440
- [49] Lande R and Arnold S J 1983 The Measurement of Selection on Correlated Characters *Evolution* **37** 1210–1226
- [50] Horváthová T, Baláž M and Jandzik D 2013 Reproduction and Morphology of the Common Lizard (*Zootoca vivipara*) from Montane Populations in Slovakia *Zoological Science* **30** 92–98
- [51] Shi L, Nazarov R, Terbish K, Munkhbayar K and Zhao W IUCN Red List of Threatened Species: *Zootoca vivipara* [en ligne], URL : <https://www.iucnredlist.org/en> [consulté le 01 juillet 2023]
- [52] Vacheva E and Naumov B 2019 Diet of the Viviparous lizard *Zootoca vivipara* (Lichtenstein, 1823) (Reptilia: Lacertidae) from its southern range *ACA* **2** e46484
- [53] Heulin B 1985 Maturité sexuelle et âge à la première reproduction dans une population de plaine de *Lacerta vivipara* *Can. J. Zool.* **63** 1773–1777
- [54] Le Galliard J-F, Massot M, Landys M M, Meylan S and Clobert J 2006 Ontogenic sources of variation in sexual size dimorphism in a viviparous lizard *J Evolution Biol* **19** 690–704
- [55] Avery R A 1975 Clutch size and reproductive effort in the lizard *Lacerta vivipara* Jacquin *Oecologia* **19** 165–170
- [56] Vercken E, Massot M, Sinervo B and Clobert J 2007 Colour variation and alternative reproductive strategies in females of the common lizard *Lacerta vivipara* *J Evolution Biol* **20** 221–232
- [57] Le Galliard J F, Marquis O and Massot M 2010 Cohort variation, climate effects and population dynamics in a short-lived lizard *Journal of Animal Ecology* **79** 1296–307
- [58] Brown J H, Gillooly J F, Allen A P, Savage V M and West G B 2004 Toward a metabolic theory of ecology *Ecology* **85** 1771–1789
- [59] Huey R B and Kingsolver J G 2019 Climate Warming, Resource Availability, and the Metabolic Meltdown of Ectotherms *The American Naturalist* **194** 140–150
- [60] Nielsen M E and Papaj D R 2015 Effects of developmental change in body size on ectotherm body temperature and behavioral thermoregulation: caterpillars in a heat-stressed

environment *Oecologia* **177** 171–179

[61] Avery R A 1994 Growth in reptiles *Gerontology* **40** 193–199

[62] Kar F, Nakagawa S and Noble D W A 2022 Heritability and developmental plasticity of growth in an oviparous lizard

[63] Paquet M, Galliard J-F L, Perret S and Pantelic Z 2011 Effects of miniature transponders on physiological stress, locomotor activity, growth and survival in small lizards *Amphib Reptilia* **32** 177–183

[64] Kalinowski S T, Taper M L and Marshall T C 2007 Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment *Molecular Ecology* **16** 1099–1106

[65] Sinnwell J, Therneau T, Schaid D, Atkinson E and Mester C 2022 kinship2: Pedigree Functions[en ligne], URL : <https://CRAN.R-project.org/package=kinship2> [consulté le 01 juillet 2023]

[66] Schoener T W and Schoener A 1978 Estimating and Interpreting Body-Size Growth in Some Anolis Lizards *Copeia* **1978** 390

[67] Yang C, Zhao J, Diaz R E and Lyu N 2019 Development of sexual dimorphism in two sympatric skinks with different growth rates *Ecol Evol* **9** 7752–7760

[68] Rotger A, Tenan S, Igual J-M, Bonner S and Tavecchia G 2023 Life span, growth, senescence and island syndrome: Accounting for imperfect detection and continuous growth *Journal of Animal Ecology* **92** 183–194

[69] Hernandez-Llamas A and Ratkowsky D 2004 Growth of fishes, crustaceans and molluscs: -estimation of the von Bertalanffy, Logistic, Gompertz and Richards curves and a new growth model *Mar. Ecol. Prog. Ser.* **282** 237–244

[70] Guarino F M, Di Già I and Sindaco R 2010 Age and growth of the sand lizards (*Lacerta agilis*) from a high Alpine population of north-western Italy *Acta Herpetologica* **23**–29

[71] Frynta D, Frýdlová P, Hnízdo J, Šimková O, Cikánová V and Velenský P 2010 Ontogeny of Sexual Size Dimorphism in Monitor Lizards: Males Grow for a Longer Period, but not at a Faster Rate *Zoological Science* **27** 917–923

[72] Dunham A E 1978 Food Availability as a Proximate Factor Influencing Individual Growth Rates in the Iguanid Lizard *Sceloporus Merriami* *Ecology* **59** 770–778

[73] Clobert J, Massot M, Lecomte J, Sorci G, de Fraipont M and Barbault R 1994 Determinants of Dispersal Behavior: The Common Lizard as a Case Study *Lizard Ecology* ed L J Vitt and E R Pianka (Princeton University Press) pp 183–206

[74] Vehtari A, Gabry J, Magnusson M, Yao Y, Bürkner P-C, Paananen T, Gelman A, Goodrich B, Piironen J and Nicenboim B 2023 loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models *Statistics and Computing* **27** 1413–1432

[75] Vehtari A, Gelman A and Gabry J 2017 Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC *Stat Comput* **27** 1413–1432

[76] Zullinger E M, Ricklefs R E, Redford K H and Mace G M 1984 Fitting Sigmoidal Equations to Mammalian Growth Curves *Journal of Mammalogy* **65** 607–636

[77] Avery R A 1994 Growth in Reptiles *Gerontology* **40** 193–199

[78] El Mouden E H, Znari M and Brown R P 1999 Skeletochronology and mark–recapture assessments of growth in the North African agamid lizard (*Agama impalearis*) *Journal of Zoology* **249** 455–461

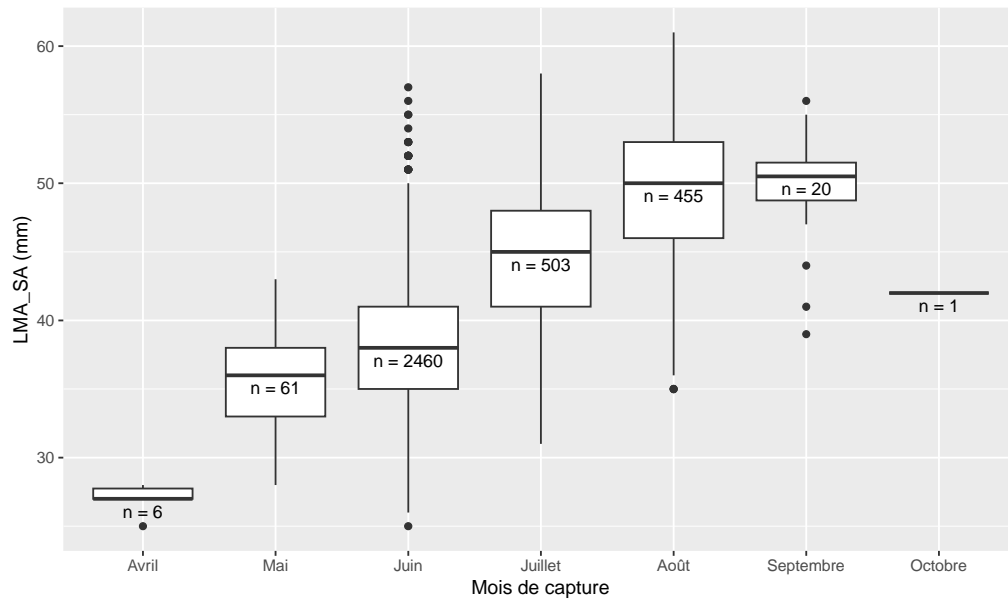
[79] Johnston G 2010 Growth and survivorship as proximate causes of sexual size dimorphism in peninsula dragon lizards *Ctenophorus fionni* *Austral Ecology* **36**

[80] Hernández-Salinas U, Ramírez-Bautista A, Cruz-Elizalde R, Meiri S and Berriozabal-Islas C 2019 Ecology of the growth of *Anolis nebulosus* (Squamata: Dactyloidae) in a seasonal tropical environment in the Chamela region, Jalisco, Mexico *Ecol Evol* **9** 2061–2071

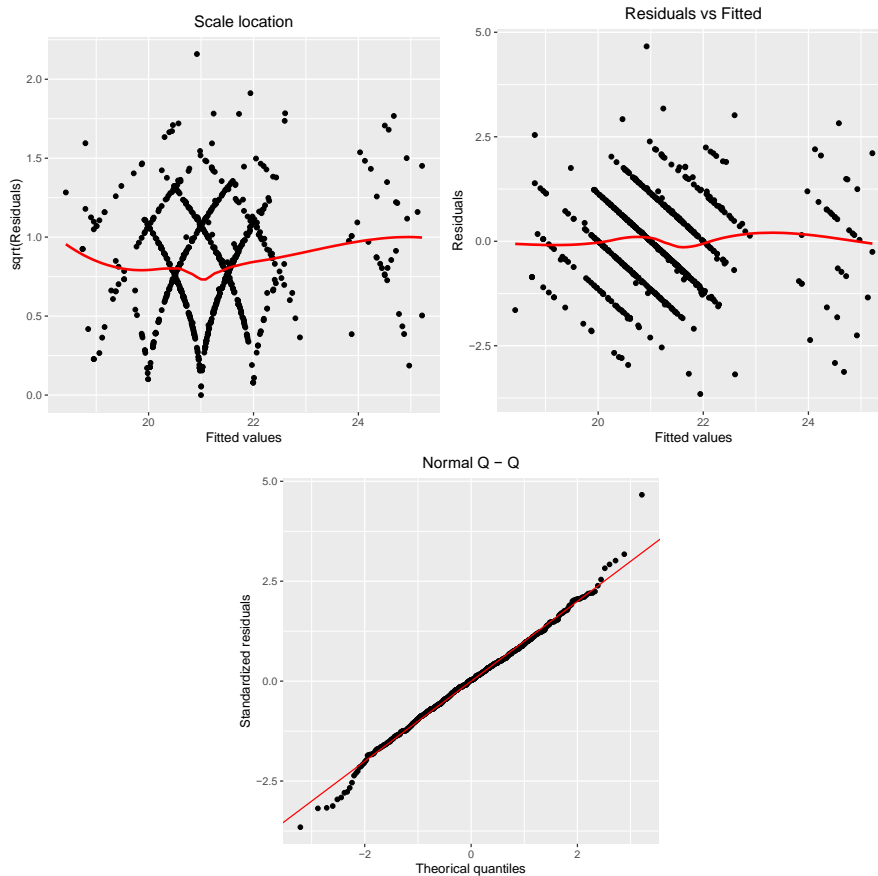
- [81] Kruuk L E B and Hadfield J D 2007 How to separate genetic and environmental causes of similarity between relatives *Journal of Evolutionary Biology* **20** 1890–1903
- [82] Kruuk L E B, Slate J, Pemberton J M, Brotherstone S, Guinness F and Clutton-Brock T 2002 Antler size in red deer: heritability and selection but no evolution *Evolution* **56** 1683–1695
- [83] Huchard E, Charmantier A, English S, Bateman A, Nielsen J F and Clutton-Brock T 2014 Additive genetic variance and developmental plasticity in growth trajectories in a wild cooperative mammal *Journal of Evolutionary Biology* **27** 1893–1904
- [84] Wilson A J and Réale D 2006 Ontogeny of Additive and Maternal Genetic Effects: Lessons from Domestic Mammals. *The American Naturalist* **167** 23–38
- [85] Noble D W A, McFarlane S E, Keogh J S and Whiting M J 2014 Maternal and additive genetic effects contribute to variation in offspring traits in a lizard *Behavioral Ecology* **25** 633–640
- [86] Réale D, Festa-Bianchet M and Jorgenson J T 1999 Heritability of body mass varies with age and season in wild bighorn sheep *Heredity* **83** 526–532
- [87] Wilson A J, Pemberton J M, Pilkington J G, Clutton-Brock T H, Coltman D W and Kruuk L E B 2007 Quantitative genetics of growth and cryptic evolution of body size in an island population *Evol Ecol* **21** 337–344
- [88] Martins F, Kruuk L, Llewelyn J, Moritz C and Phillips B 2019 Heritability of climate-relevant traits in a rainforest skink *Heredity* **122** 41–52
- [89] Duryea M C, Bergeron P, Clare-Salzler Z and Calsbeek R 2016 Field estimates of parentage reveal sexually antagonistic selection on body size in a population of Anolis lizards *Ecology and Evolution* **6** 7024–7031
- [90] Cox R M and Calsbeek R 2010 Sex-specific selection and intraspecific variation in sexual size dimorphism *Evolution* **64** 798–809
- [91] Daufresne M, Lengfellner K and Sommer U 2009 Global warming benefits the small in aquatic ecosystems *Proc. Natl. Acad. Sci. U.S.A.* **106** 12788–12793
- [92] Chabaud C, Berroneau M, Berroneau M, Dupoué A, Guillon M, Viton R, Gavira R S B, Clobert J, Lourdaïs O and Le Galliard J-F 2022 Climate aridity and habitat drive geographical variation in morphology and thermo-hydroregulation strategies of a widespread lizard species *Biological Journal of the Linnean Society* **137** 667–685

Annexes

Annexe 1 : Évolution de la taille des subadultes au cours de l'été



Annexe 2 : Vérification des hypothèses de la régression linéaire entre LMA_{SA} et LMA_N



Annexe 3 : Implémentation du modèle de régression linéaire entre LMA_{SA} et LMA_N sur R

```
# Modélisation de La taille à La naissance à partir des variables explicatives :
#   - Taille au stade subadulte
#   - Sexe
#   - Année de naissance
#   - Terme d'interaction entre la taille des subadultes et le mois de capture
#     des subadultes
model <- lm(LMA_nais ~ LMA_SA + Sex + An_nais + LMA_SA:as.factor(Obs_SA),
            data = donnees_calibrage)

summary(model)

## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9423 -0.5508  0.0255  0.5134  4.0785
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)      19.9012462    0.4438464   44.838 < 2e-16 ***
## LMA_SA          -0.0056471    0.0287193   -0.197  0.844171
## Sex1            -0.4723710    0.0659038   -7.168  1.89e-12 ***
## An_nais1990     -1.2094153    0.6221437   -1.944  0.052289 .
## An_nais1991     -0.1127241    0.4014711   -0.281  0.778961
## An_nais1993     -0.3380813    0.4963240   -0.681  0.495981
## An_nais1994      2.4727601    0.3353235    7.374  4.53e-13 ***
## An_nais1995      0.4952451    0.6937771    0.714  0.475557
## An_nais1998     -0.1105366    0.5337765   -0.207  0.836002
## An_nais1999     -0.2927436    0.5342348   -0.548  0.583883
## An_nais2000      0.1998894    0.4312035    0.464  0.643101
## An_nais2001     -0.0002289    0.3570980   -0.001  0.999489
## An_nais2002      0.7145644    0.3782475    1.889  0.059272 .
## An_nais2003      0.3579074    0.4335056    0.826  0.409296
## An_nais2004     -0.7910694    0.3427213   -2.308  0.021269 *
## An_nais2005     -0.1588719    0.3300754   -0.481  0.630435
## An_nais2006     -0.2660987    0.3337775   -0.797  0.425577
## An_nais2007      0.4817552    0.3577222    1.347  0.178489
## An_nais2008      0.0852045    0.3895572    0.219  0.826929
## An_nais2009      0.0696589    0.3286585    0.212  0.832206
## An_nais2010     -0.7835574    0.3878016   -2.021  0.043698 *
## An_nais2011     -0.8814911    0.3461988   -2.546  0.011096 *
## An_nais2012     -0.7025352    0.3443983   -2.040  0.041724 *
## An_nais2013     -1.3015285    0.3341325   -3.895  0.000107 ***
## An_nais2014     -0.8226666    0.3329056   -2.471  0.013696 *
## An_nais2015     -0.5842523    0.3171556   -1.842  0.065860 .
## An_nais2016     -1.6291048    0.3266231   -4.988  7.66e-07 ***
## An_nais2017     -1.1586828    0.3134958   -3.696  0.000236 ***
## An_nais2018     -2.3378335    0.3392853   -6.890  1.21e-11 ***
## LMA_SA:as.factor(Obs_SA)5  0.0691200    0.0313061    2.208  0.027566 *
## LMA_SA:as.factor(Obs_SA)6  0.0589005    0.0263021    2.239  0.025434 *
## LMA_SA:as.factor(Obs_SA)7  0.0539829    0.0265471    2.033  0.042368 *
## LMA_SA:as.factor(Obs_SA)8  0.0546331    0.0267619    2.041  0.041568 *
## LMA_SA:as.factor(Obs_SA)9  0.0446334    0.0271589    1.643  0.100731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8811 on 724 degrees of freedom
## Multiple R-squared:  0.5993, Adjusted R-squared:  0.581
## F-statistic: 32.81 on 33 and 724 DF, p-value: < 2.2e-16
```


Annexe 4 : Distributions des *a priori* des modèles ajustés selon une approche bayésienne dans la partie expérimentale.

Modèle	Paramètre	Loi mathématique de la distribution <i>a priori</i>
Modèles de croissances (Logistique, de Van Bertalanffy et de Gompertz)	I_{max} (Intercept)	Loi normale $\mathcal{N}(60,10)$
	I_{max} (catégorie Sexe = mâles)	Loi normale $\mathcal{N}(0,10)$
	I_0 (Intercept)	Loi normale $\mathcal{N}(20,2)$
	I_0 (catégorie Sexe = mâles)	Loi normale $\mathcal{N}(0,2)$
	k (Intercept)	Loi uniforme $\mathcal{U}(0,10)$
	k (catégorie Sexe = mâles)	Loi normale $\mathcal{N}(0,5)$
Modèles animaux linéaires classiques (Sur LMA_N et sur LMA_{SA})	Sd _A (Intercept)	Loi de Student $t_3(0,2.5)^*$
	Sd _A (catégorie Sexe = mâles)	Loi de Student $t_3(0,2.5)^*$
	Sd _{AN} (Intercept)	Loi de Student $t_3(0,2.5)^*$
	Sd _{AN} (catégorie Sexe = mâles)	Loi de Student $t_3(0,2.5)^*$
	Sd _M (Intercept)	Loi de Student $t_3(0,2.5)^*$
	Sd _M (catégorie Sexe = mâles)	Loi de Student $t_3(0,2.5)^*$
Modèle animal hiérarchique non linéaire	Sd _R	Loi de Student $t_3(0,2.5)^*$
	Sd _A de k (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _A de k (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _A de I_{max} (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _A de I_{max} (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _{AN} de k (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _{AN} de k (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _{AN} de I_{max} (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _{AN} de I_{max} (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _M de k (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _M de k (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _M de I_{max} (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _M de I_{max} (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _{PE} de k (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _{PE} de k (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _{PE} de I_{max} (Intercept)	Loi de Student $t_3(0,21.1)^*$
	Sd _{PE} de I_{max} (catégorie Sexe = mâles)	Loi de Student $t_3(0,21.1)^*$
	Sd _R	Loi de Student $t_3(0,21.1)^*$

* Lois tronquées en $[0; +\infty[$

Annexe 5 : Implémentation du modèle de croissance logistique sur R

Annexe 5.1 : Ajustement du modèle avec le paquet brms

```
prior <- prior(normal(60, 10), nlpar = 'lmax') +  
  prior(normal(0, 10), nlpar = 'lmax', coef = 'Sex') +  
  prior(normal(20, 2), nlpar = 'l0') +  
  prior(normal(0, 2), nlpar = 'l0', coef = 'Sex') +  
  # Connaissances à priori sur la taille des lézards  
  prior(uniform(0, 10), nlpar = 'k') +  
  prior(normal(0, 5), nlpar = 'k', coef = 'Sex')  
  
model_log <- brm(bf(LMA ~ lmax / (1 + ((lmax - l0) / l0) * exp(- k * Age)),  
  lmax + k + l0 ~ Sex + (1|ID),  
  # l0, lmax et k dépendent de l'individu  
  nl = TRUE),  
  chains = 4,  
  # modèle tourne avec 4 chaînes  
  init = rep(list(list(b_lmax = array(data = c(68, 0)),  
    b_k = array(data = c(1, 0)),  
    b_l0 = array(data = c(21, 0)))), 4),  
  # Coordonnées de départ de l'échantillonnage des chaînes  
  data = donnees_mod,  
  prior = prior)
```

Annexe 5.2 : Sortie du modèle

```
## Family: gaussian  
## Links: mu = identity; sigma = identity  
## Formula: LMA ~ lmax/(1 + ((lmax - l0)/l0) * exp(-k * Age))  
## lmax ~ Sex + (1 | ID)  
## k ~ Sex + (1 | ID)  
## l0 ~ Sex + (1 | ID)  
## Data: donnees_mod (Number of observations: 4584)  
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
## total post-warmup draws = 4000  
##  
## Group-Level Effects:  
## ~ID (Number of levels: 1320)  
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
## sd(lmax_Intercept)	2.27	0.16	1.98	2.59	1.01	936	1762
## sd(k_Intercept)	0.28	0.01	0.26	0.30	1.00	1269	2280
## sd(l0_Intercept)	0.16	0.12	0.01	0.44	1.00	957	1286

```
##  
## Population-Level Effects:  
##
```

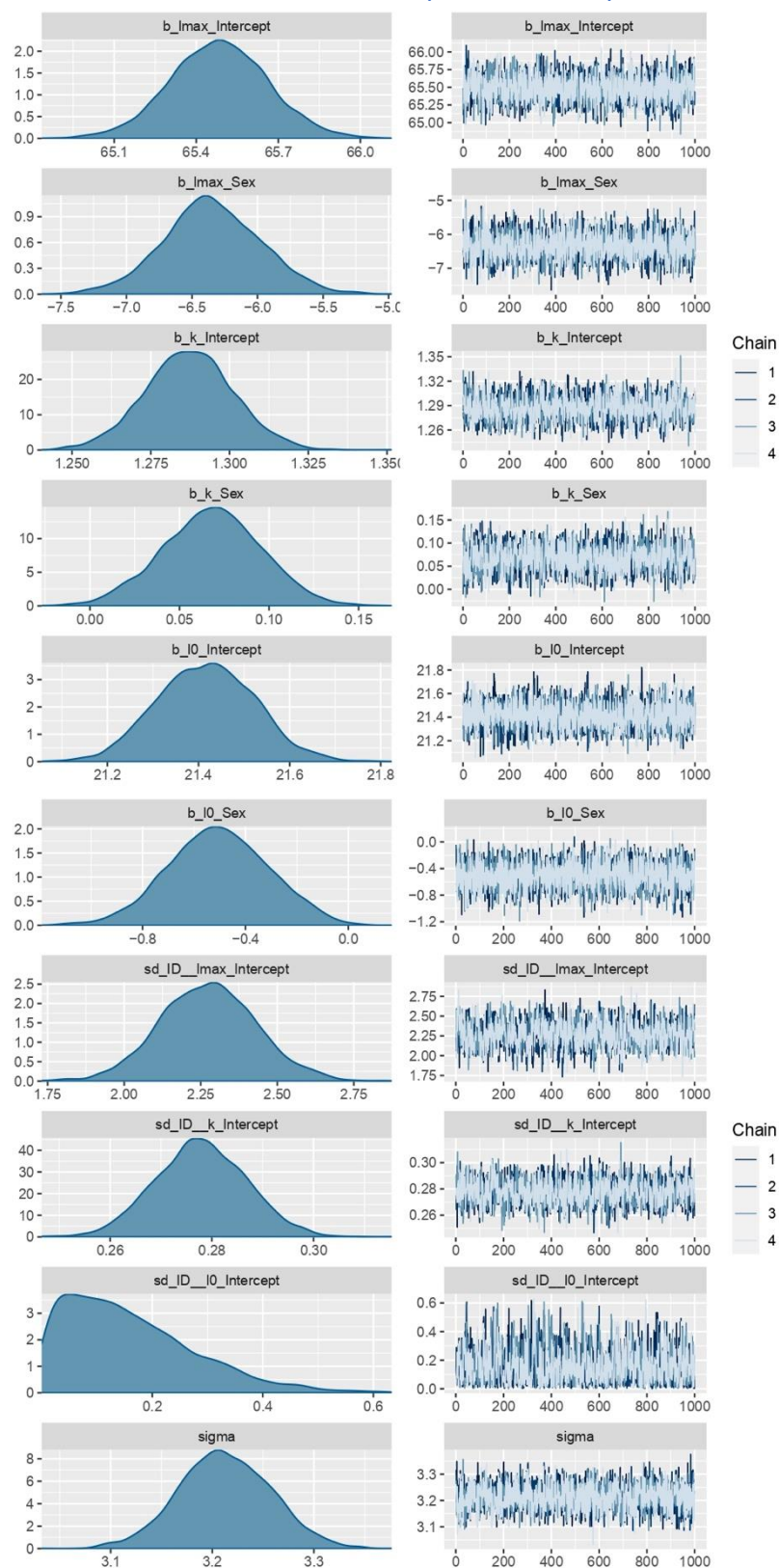
	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
## lmax_Intercept	65.48	0.18	65.13	65.83	1.00	2626	2961
## lmax_Sex	-6.35	0.37	-7.08	-5.61	1.00	2647	2835
## k_Intercept	1.29	0.01	1.26	1.32	1.00	1622	2198
## k_Sex	0.07	0.03	0.01	0.12	1.00	1827	2665
## l0_Intercept	21.42	0.11	21.21	21.63	1.00	3652	2928
## l0_Sex	-0.50	0.19	-0.89	-0.13	1.00	3696	2995

```
##  
## Family Specific Parameters:  
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
## sigma	3.21	0.05	3.12	3.30	1.00	1843	2496

```
##  
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS  
## and Tail_ESS are effective sample size measures, and Rhat is the potential  
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Annexe 5.3 : Tracé des chaînes et distributions *a posteriori* des paramètres du modèle



Annexe 6 : Implémentation du modèle animal sur la taille des nouveau-nés (LMA_N) sur R

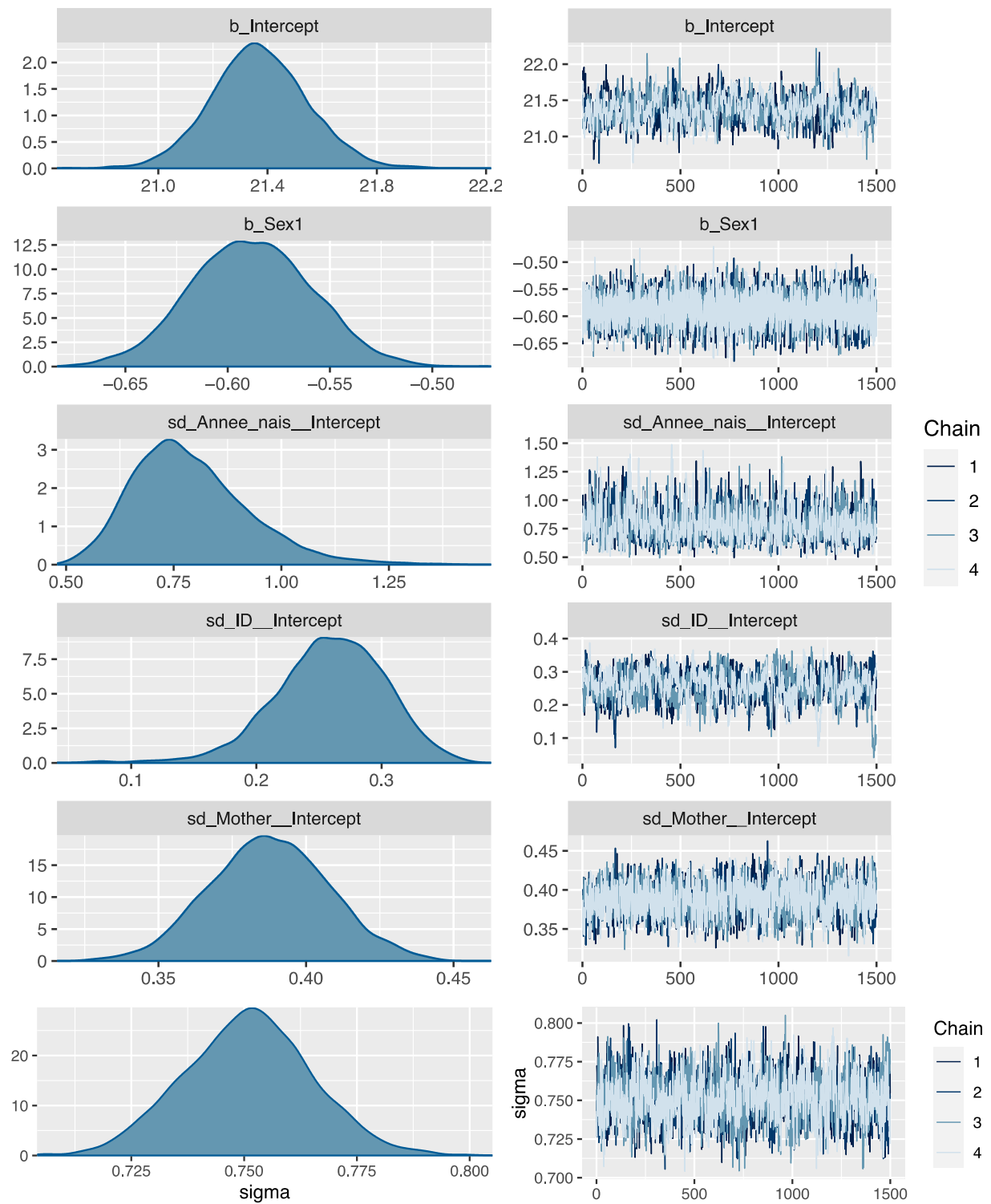
Annexe 6.1 : Ajustement du modèle avec le paquet brms

```
manimal_N <- brm(LMA ~ 1 +  
                  Sex +  
                  (1 | gr(ID, cov = Amatrix_N)) +  
                  (1 | Annee_nais) +  
                  (1 | Mother),  
  data = donnees_mod_N,  
  data2 = list(Amatrix_N = Amatrix_N),  
  family = gaussian(),  
  chains = 4,  
  cores = 4,  
  iter = 4000,  
  warmup = 1000,  
  thin = 1  
)
```

Annexe 6.2 : Sortie du modèle

```
## Family: gaussian  
## Links: mu = identity; sigma = identity  
## Formula: LMA ~ 1 + Sex + (1 | gr(ID, cov = Amatrix_N)) + (1 | Annee_nais) + (1  
| Mother)  
## Data: donnees_mod_N (Number of observations: 3353)  
## Draws: 4 chains, each with iter = 2000; warmup = 500; thin = 1;  
## total post-warmup draws = 6000  
##  
## Group-Level Effects:  
## ~Annee_nais (Number of levels: 20)  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## sd(Intercept) 0.79 0.14 0.57 1.10 1.00 1101 2090  
##  
## ~ID (Number of levels: 3353)  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## sd(Intercept) 0.26 0.04 0.17 0.34 1.00 408 701  
##  
## ~Mother (Number of levels: 1501)  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## sd(Intercept) 0.39 0.02 0.35 0.43 1.00 1115 2529  
##  
## Population-Level Effects:  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## Intercept 21.37 0.18 21.03 21.73 1.01 472 848  
## Sex1 -0.59 0.03 -0.65 -0.53 1.00 6829 4871  
##  
## Family Specific Parameters:  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## sigma 0.75 0.01 0.72 0.78 1.00 854 1924  
##  
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS  
## and Tail_ESS are effective sample size measures, and Rhat is the potential  
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Annexe 6.3 : Tracé des chaînes et distributions *a posteriori* des paramètres du modèle



Annexe 6.4 : Estimation de l'héritabilité et du coefficient de variation additif associé à LMA_N

```
## Récupération des distribution a posteriori des écarts types (sd_)
# et des variances (v_) associés aux effets aléatoires et résiduels
sd_ad <- (VarCorr(m_animal_N, summary = FALSE)$ID$sd) # Effet génétique additif
v_ad <- sd_ad ^ 2
sd_AN <- (VarCorr(m_animal_N, summary = FALSE)$Annee_nais$sd) # Année de naissance
v_AN <- sd_AN ^ 2
sd_mother <- (VarCorr(m_animal_N, summary = FALSE)$Mother$sd) # Effet maternel
v_mother <- sd_mother ^ 2
sd_r <- (VarCorr(m_animal_N, summary = FALSE)$Residual$sd) # Résidus
v_r <- sd_r ^ 2

## Estimation de l'héritabilité (h2) et du coefficient de variation additif (CVA)
h2 <- as.mcmc(v_ad / (v_ad + v_AN + v_mother + v_r))
CVA <- as.mcmc(sd_ad / mean(donnees_mod_N$LMA))
Summary(h2)

## Iterations = 1:6000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 6000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean           SD      Naive SE Time-series SE
##    0.0495888    0.0174652    0.0002255    0.0008021
##
## 2. Quantiles for each variable:
##
##    2.5%    25%    50%    75%    97.5%
## 0.01865 0.03711 0.04866 0.06112 0.08633
```

Annexe 7: Implémentation du modèle animal hiérarchique non-linéaire sur la stratégie de croissance sur R

Annexe 7.1 : Ajustement du modèle avec le paquet brms

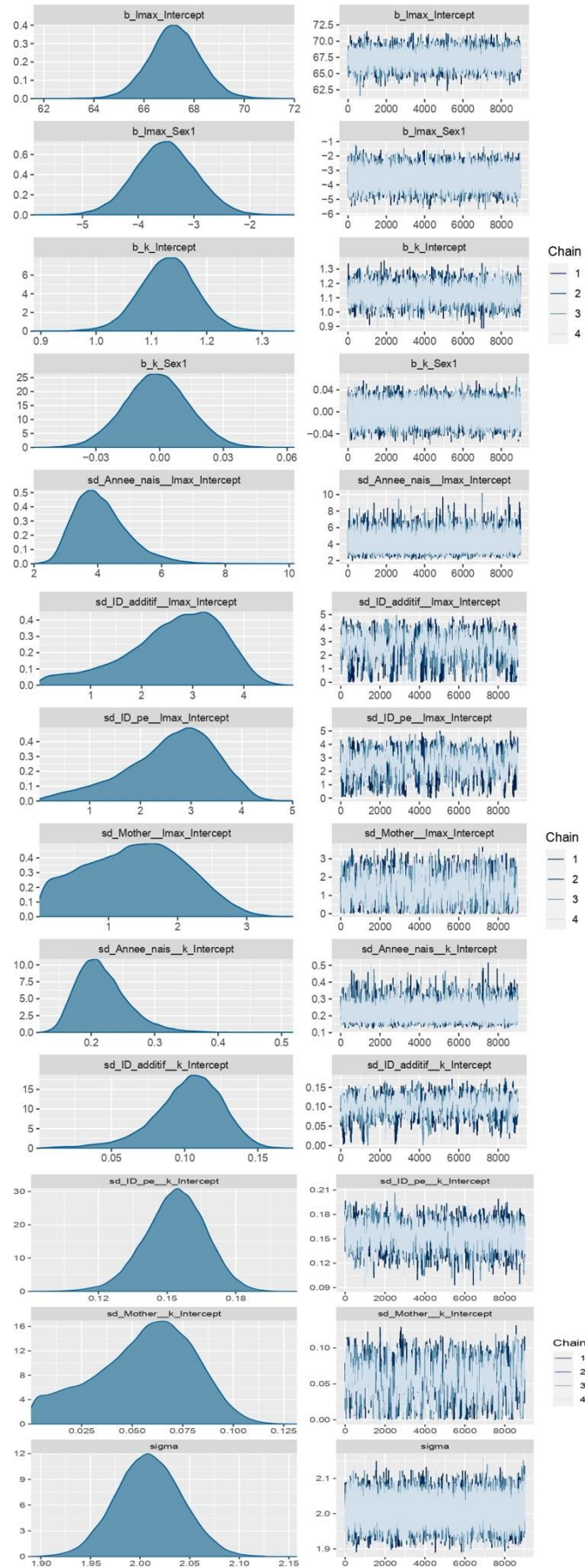
```
model <- brm(bf(LMA ~ lmax / (1 + ((lmax - LMA_N) / LMA_N) * exp(- k * Age)),
# Équation de croissance logistique
  lmax + k ~ Sex +
  (1 | Annee_nais) +
  (1 | gr(ID_additif, cov = Amatrix)) +
  (1 | ID_pe) +
  (1 | Mother),
# Définition des facteurs fixes et aléatoires
  nl = TRUE),
chains = 10,
cores = 10,
iter = 2000,
warmup = 500,
prior = prior,
init = rep(list(list(b_lmax = array(data = c(68, 0)),
                        b_k = array(data = c(1, 0)),
                        b_l0 = array(data = c(21, 0)))), 10),
# Coordonnées de départ

data = donnees_mod,
data2 = list(Amatrix = Amatrix))
```

Annexe 7.2 : Sortie du modèle

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LMA ~ lmax/(1 + ((lmax - LMA_N)/LMA_N) * exp(-k * Age))
##          lmax ~ Sex + (1 | Annee_nais) + (1 | gr(ID_additif, cov = Amatrix)) +
(1 | ID_pe) + (1 | Mother)
##          k ~ Sex + (1 | Annee_nais) + (1 | gr(ID_additif, cov = Amatrix)) + (1
| ID_pe) + (1 | Mother)
## Data: data (Number of observations: 4618)
## Draws: 4 chains, each with iter = 10000; warmup = 1000; thin = 1;
##          total post-warmup draws = 36000
##
## Group-Level Effects:
## ~Annee_nais (Number of levels: 19)
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(lmax_Intercept)      4.11      0.86      2.80      6.12 1.00     11653     17310
## sd(k_Intercept)         0.22      0.04      0.15      0.31 1.00      8913     14451
##
## ~ID_additif (Number of levels: 1848)
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(lmax_Intercept)      2.63      0.94      0.43      4.10 1.01       317       736
## sd(k_Intercept)         0.10      0.02      0.05      0.14 1.01       607       760
##
## ~ID_pe (Number of levels: 1848)
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(lmax_Intercept)      2.60      0.88      0.59      4.02 1.02       278       791
## sd(k_Intercept)         0.15      0.01      0.13      0.18 1.01       859     1990
##
## ~Mother (Number of levels: 1382)
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(lmax_Intercept)      1.39      0.72      0.10      2.73 1.00      1234     3534
## sd(k_Intercept)         0.06      0.02      0.00      0.10 1.01       670     1489
##
## Population-Level Effects:
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## lmax_Intercept      67.24      1.03     65.22     69.27 1.00      6746     12774
## lmax_Sex1           -3.51      0.55     -4.60     -2.42 1.00     19345     24352
## k_Intercept          1.13      0.05      1.03      1.23 1.00      3689      7899
## k_Sex1              -0.00      0.02     -0.03      0.03 1.00     20127     24665
##
## Family Specific Parameters:
##          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma          2.01      0.03      1.94      2.08 1.00      9137     17613
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```


Annexe 7.3 : Tracé des chaînes et distributions *a posteriori* des paramètres du modèle



Annexe 8 : Implémentation du modèle d'estimation des gradients de sélection sur la taille à la naissance (LMA_N) des femelles sur R

Annexe 8.1 : Ajustement du modèle avec le paquet brms

```
model_N_f <- brm(
  Repro ~ LMA_N_stand + quadLMA_N_stand + (1|Annee_nais),
  # Repro : nombre de descendants
  # LMA_N_stand : taille à la naissance standardisée
  # quadLMA_N_stand : (LMA_N_stand ^ 2) / 2
  data = table_N_f,
  # Sous sélection du jeu de donnée contenant uniquement les données des
  # femelles à la naissance
  family = zero_inflated_poisson(link = "log", link_zi = "logit"),
  chains = 4,
  iter = 2000,
  warmup = 1000,
  thin = 1,
  prior = NULL)
```

Annexe 8.2 : Sortie du modèle

```
## Family: zero_inflated_poisson
## Links: mu = log; zi = identity
## Formula: Repro ~ LMA_N_stand + quadLMA_N_stand + (1 | Annee_nais)
## Data: table_N_f (Number of observations: 1555)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
## Group-Level Effects:
## ~Annee_nais (Number of levels: 20)
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept) 1.25 0.31 0.77 1.99 1.00 627 1305
##
## Population-Level Effects:
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept 1.31 0.30 0.66 1.86 1.02 333 671
## LMA_N_stand 0.05 0.04 -0.03 0.12 1.00 2266 2287
## quadLMA_N_stand -0.08 0.04 -0.16 0.01 1.00 2402 2455
##
## Family Specific Parameters:
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## zi 0.74 0.01 0.72 0.77 1.00 2274 2050
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Annexe 9 : Implémentation du modèle d'estimation des coefficients de sélection sur les paramètres l_{max} et k des femelles sur R

Annexe 9.1 : Création des 100 jeux de données par sexe, nécessaires à l'ajustement des modèles

```
# Importation du modèle de croissance logistique
mcrois <- readRDS("Output/Model de croissance logistique_avec_LMAN_apres2000.rds")

# Extraction des itérations de  $l_{max}$  et  $k$  pour chaque individu
draws <- as_draws_df(mcrois)
draws_lmax <- draws |> select(starts_with("r_ID_lmax"))
draws_k <- draws |> select(starts_with("r_ID_k"))
draws_l0 <- draws |> select(starts_with("r_ID_l0"))

# Création jeu de données avec une ligne par individu, LMAN et variable Repro
donnees_mod_kl <- donnees_mod |>
  filter(length(Age) > 2 & TRUE %in% Birth & Annee_nais > 1999) |>
  group_by(ID, Sex, Annee_nais) |>
  summarize(Repro = length(which(donnees_patc$MotherID == ID[1] | donnees_patc$FatherID == ID[1])),
    # Repro : approximation de la valeur sélective absolue
    LMA_N = LMA[which(Birth == TRUE)]) |>
  mutate(Sex = as.character(Sex),
    Annee_nais = as.character(Annee_nais))

# Tirage aléatoire des 100 itérations à utiliser pour faire les 100 jeux de données
random <- as.integer(runif(100, 1, dim(draws_lmax)[1]))

# Fonction permettant d'ajouter les colonnes  $l_{max}$  et  $k$  au tableau, pour un numéro de
# ligne donné (n)
create_dataframe = function(n,
  table = donnees_mod_kl,
  draws_lmax = draws_lmax,
  draws_k = draws_k
){
  ks <- pivot_longer(data = draws_k[n,], everything())
  lmaxs <- pivot_longer(data = draws_lmax[n,], everything())
  table$k = ks$value
  table$lmax = lmaxs$value
  table$quad_k = (table$k**2)/2
  table$quad_lmax = (table$lmax**2)/2
  return(table)
}

# Création de deux listes de 100 dataframes pour les 100 itérations, pour les mâles
# et les femelles avec les colonnes  $l_{max}$  et  $k$ 
list_df_f <- c() #Initialisation liste femelles
list_df_m <- c() #Initialisation liste mâles
for (r in random){
  df <- create_dataframe(n = r,
    table = donnees_mod_kl,
    draws_lmax = draws_lmax,
    draws_k = draws_k)
  list_df_f <- c(list_df_f, list(df |> filter(Sex == 0)))
  list_df_m <- c(list_df_m, list(df |> filter(Sex == 1)))
}
```

Annexe 9.2 : Ajustement du modèle femelles avec brms

```
model_kl_f <- brm_multiple(  
  Repro ~ k + lmax + quad_k + quad_lmax + (1|Annee_nais),  
  # Année de naissance comme effet alléatoire car beaucoup d'années différentes :  
  # simplifie le modèle  
  data = list_df_f,  
  chains = 4,  
  iter = 2000,  
  warmup = 1000,  
  family = zero_inflated_poisson(link = "log", link_zi = "logit"),  
  prior = NULL)
```

Annexes 9.3 : Sortie du modèle femelles

```
##  
## Family: zero_inflated_poisson  
## Links: mu = log; zi = identity  
## Formula: Repro ~ k + lmax + quad_k + quad_lmax + (1 | Annee_nais)  
## Data: list_df_f (Number of observations: 441)  
## Draws: 400 chains, each with iter = 2000; warmup = 1000; thin = 1;  
## total post-warmup draws = 4e+05  
##  
## Group-Level Effects:  
## ~Annee_nais (Number of levels: 18)  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## sd(Intercept) 0.27 0.06 0.17 0.41 1.02 14731 37866  
##  
## Population-Level Effects:  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## Intercept 1.76 0.08 1.59 1.92 1.12 1989 4997  
## k 0.45 0.25 -0.03 0.95 1.49 735 1150  
## lmax 0.03 0.02 -0.01 0.08 1.75 599 870  
## quad_k -0.53 2.41 -5.88 3.73 1.72 609 832  
## quad_lmax -0.00 0.02 -0.04 0.03 1.71 613 1038  
##  
## Family Specific Parameters:  
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS  
## zi 0.33 0.02 0.28 0.37 1.00 667384 292162  
##  
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS  
## and Tail_ESS are effective sample size measures, and Rhat is the potential  
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Ici les Rhats sont très élevés car ils estiment la convergence globale des chaînes lancées à partir des 100 jeux de données. La convergence du modèle a été vérifiée en s'assurant que les Rhats spécifiques à chaque jeu de données étaient inférieurs à 1,05.

ÉTUDE DU POTENTIEL ADAPTATIF DE LA STRATÉGIE DE CROISSANCE DU LÉZARD VIVIPARE

Auteur

Anaïs ARAGON

Résumé

Il a été montré que l'élévation des températures provoquait un déclin rapide du lézard vivipare (*Zootoca vivipara*) dans le sud de l'Europe. Plusieurs arguments laissent à penser qu'une évolution de la stratégie de croissance des individus pourrait participer au maintien de l'espèce. Dans cette étude, nous avons évalué le potentiel adaptatif de la stratégie de croissance du lézard vivipare grâce à des analyses génétiques quantitatives réalisées selon une approche bayésienne. Nous avons ajusté des modèles statistiques d'analyse de la variance sur les données issues d'une population naturelle de lézards suivie depuis la fin du XXe siècle. Nos résultats suggèrent que l'héritabilité de la taille du lézard vivipare est très faible à la naissance ($h^2 = 0,050$), augmente fortement au stade subadulte ($h^2 = 0,206$) et est maximale pour la taille asymptotique ($h^2 = 0,228$). Le taux de croissance présentait une héritabilité intermédiaire ($h^2 = 0,131$). Le coefficient de variation additif était maximal pour la taille au stade subadulte ($CV_A = 0,369$) et minimal pour la taille à la naissance ($CV_A = 0,013$). Il était intermédiaire pour la taille asymptotique ($CV_A = 0,042$) et le taux de croissance ($CV_A = 0,077$). Nous avons également utilisé des modèles de régression quadratique pour étudier la pression de sélection sur la stratégie de croissance au sein de la population étudiée. Nous avons ainsi mis en évidence une sélection directionnelle vers l'augmentation de la taille des mâles au stade subadulte, avec un optimum autour de + 3 mm par rapport à la taille moyenne.

Ces résultats étaient en faveur d'un fort potentiel adaptatif de la stratégie de croissance du lézard vivipare. Une meilleure compréhension des mécanismes impliqués dans le déclin des populations de lézards est néanmoins nécessaire pour mieux comprendre les avantages conférés par une modification de la taille des individus, dans un contexte d'élévation des températures.

Mots-clés

Héritabilité, *Zootoca vivipara*, génétique quantitative, modèle animal, stratégie de croissance, surface de sélection individuelle.

Jury

Président du jury	:	Professeur CALENDER Alain
Directeur de thèse	:	Docteur DE VILLEMEREUIL Pierre
1er assesseur	:	Docteur DELIGNETTE-MULLER Marie-Laure
2ème assesseur	:	Docteur Marie ABITBOL