

Mémoire de fin d'études

présenté pour l'obtention du diplôme d'ingénieur agronome
Option : Amélioration des plantes et ingénierie végétale
méditerranéennes et tropicales (APIMET)

<p>Les pratiques de sélections paysannes de blé dur : Quel regard porter sur les grains ?</p>
--



par Clément BIENVENU

Année de soutenance : 2023

Mémoire préparé sous la direction de : Jean-François MARTIN

Présenté le : 21/09/2023

devant le jury :

Vincent RANWEZ

Anne LAPERCHE

Timothée FLUTRE

Organisme d'accueil : L'Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement (l'Institut Agro) – Chaire partenariale "Semences pour Demain"

Maître de stage : Jacques DAVID

Le délit de plagiat est encadré par les articles L335-2 et suivants du Code de la Propriété Intellectuelle.

Ainsi « toute reproduction, représentation ou diffusion, par quelque moyen que ce soit, d'une œuvre de l'esprit en violation des droits de l'auteur, tels qu'ils sont définis et réglementés par la loi » est un délit de contrefaçon en vertu de l'article L335-3 du Code de la Propriété Intellectuelle.

Vous devez de citer correctement les sources des textes, images, graphiques ou toutes autres informations reprises sur d'autres documents. L'omission de la référence, volontaire ou non, est fautive.

Par ailleurs, en conformité avec la charte éthique de l'Etablissement, l'équipe pédagogique vous rappelle que les travaux écrits que vous livrez, ou déposez sur Ticéa (rapports d'UE ou ECUE, mémoire de fin d'études) doivent s'abstenir de toute reprise textuelle de documents antérieurs. La citation partielle de l'écrit d'un autre auteur reste acceptable, si elle s'avère absolument nécessaire (cas exceptionnel : par ex. première définition d'un concept) ; elle est alors placée entre guillemets et référencée. A contrario, les reprises textuelles des écrits d'autrui sont considérées comme des plagats.

Vos écrits sont soumis à la détection de plagiat via un logiciel. En cas de plagiat caractérisé, les articles 3-5 et suivants du règlement des études en vigueur dans l'Etablissement s'appliqueront.

En cas de manquement à ces consignes, la SEVE/le correcteur se réservent le droit d'exiger la réécriture du document sans préjuger d'éventuelles sanctions disciplinaires.

Engagement :

Je soussigné Clément BIENVENU

Reconnaît avoir lu et m'engage à respecter les consignes de non plagiat.

Reconnaît avoir ~~utilisé~~ ou ne pas utiliser l'Intelligence Artificielle Chat GPT pour la rédaction de mon document. (Rayer la mention inutile)

A Montpellier le 05/09/2023

Signature :



RESUME

Comme toutes les cultures, le blé dur va être impacté par le changement climatique et la demande d'une population croissante. Une transition vers un modèle agricole plus résilient et durable du point de vue environnemental, social et économique est donc nécessaire. Parmi toutes les étapes de la production, la sélection variétale est la plus en amont, et se doit de proposer des innovations pour participer à cette transition. Dans ce contexte, l'utilisation de variétés populations et la pratique de la sélection massale sont intéressantes car elles offrent justement une certaine résilience face aux conditions climatiques rudes, et ont aussi des implications sociales et économiques pour les agriculteurs.

Durant le stage, deux questions ont été abordées. La première est celle de l'efficacité sélection massale sur la taille du grain, un trait important de la qualité chez le blé dur. La deuxième question concerne tout trait mesurable sur le grain et est de savoir s'il est préférable de mesurer ces traits sur des grains individuels ou sur des moyennes de grains du même épi.

Nous avons pu montrer que la sélection massale sur la taille du grain peut être efficace pour améliorer ce trait. Nous avons aussi développé une équation permettant de comparer la sélection sur grain individuel et sur épi. Cette équation nous a permis de montrer que pour la taille du grain, la sélection grains individuels semble avantageuse dans la plupart des situations mais que la sélection sur épi peut devenir plus efficace lorsque les effectifs impliqués sont faibles.

Mots clés

Sélection massale, blé dur, taille des grains, variété population, agroécologie

Pour citer ce document : Bienvenu, Clément, 2023. Les pratiques de sélections paysannes de blé dur : Quel regard porter sur les grains ?. Mémoire d'Ingénieur Agronome, diplôme, Amélioration des plantes et ingénierie végétale méditerranéennes et tropicales, L'Institut Agro Montpellier. 54 pages.

ABSTRACT

Title : Farmer's mass selection of durum wheat : how to look at grains ?

As for many other crops, durum wheat production will have to face climate change and the growing demand of an increasing world population. Thus, a transition toward a more environmentally, socially and economically sustainable agriculture is needed. Breeding is at the beginning of the production line and must provide innovations enabling the transition. In this context, the use of wheat populations and mass selection are interesting because they can be part of more resilient systems and have social and economic implications for farmers.

Two questions were addressed during this internship. First, the efficiency of mass selection on grain size (which is important for grain quality) was investigated. Then, the scope was broadened to any trait of the grain, and it was addressed whether it is better to measure these traits on individual grains or on means of spikes.

It was shown that mass selection on wheat grain size could be efficient and an equation to compare selection on individual grains and on spikes was developed. This equation led us to the conclusion that for grain size, selecting on individual grains is better in most of the situations but selection on spikes can be more efficient when only a few plants can be grown, measured and selected.

Key words

Mass selection, durum wheat, grain size, population, agroecology

REMERCIEMENTS

Je voudrais en premier lieu remercier mes encadrants Jacques David et Martin Ecarnot qui ont fait de ce stage une aventure intellectuelle stimulante et pleine de questionnements. Je remercie particulièrement Jacques pour m'avoir préparé au concours de l'école doctorale, je n'aurais sûrement pas eu ma place sans ses conseils avisés. Je le remercie également pour avoir été flexible et compréhensif dans une période assez éprouvante sur le plan personnel. Je remercie aussi la Chaire partenariale "Semences pour Demain" qui financé ce stage.

Je remercie chaudement ma famille et mes amis qui m'ont soutenu durant cette tranche de vie pétrie de stress et d'incertitudes (concours, orientation, rendez-vous médicaux, maladie, douleurs...). Merci d'avoir été présents lorsque, me trouvant dans un piteux état, tant mentalement que physiquement, je remportais au logis un œil triste et mouillé. J'y ai laissé quelques plumes mais grâce à vous je ne suis pas tombé du nid, merci du fond du cœur. Petite dédicace également à mon kiné, mon chirurgien et ma rhumato pour leur prise en charge de qualité.

Enfin, merci aux techniciens, ingénieurs, chercheurs et thésards du bâtiment ARCAD pour leurs conseils, leur aide, et le partage de leur expérience.

TABLE DES MATIERES

1	Introduction	11
1.1	Le changement climatique	11
1.2	Sélection massale et diversité génétique	11
1.3	L'importance de la qualité des grains chez le blé dur	13
1.4	Les pratiques de sélection massale chez le blé dur	14
2	Matériel et méthodes	16
2.1	Matériel végétal	16
2.2	Dispositifs expérimentaux.....	17
2.2.1	Dispositif de sélection in vivo	17
2.2.2	Dispositif de sélection in silico.....	17
2.3	Phénotypage.....	18
2.4	Mesure de l'effet de la sélection dans le dispositif de sélection in vivo	19
2.5	Mesure de l'effet de la sélection dans le dispositif de sélection in silico.....	19
2.5.1	Simulation de la sélection.....	19
2.5.2	Estimation des héritabilités.....	20
2.5.3	Estimation de l'effet de la sélection.....	21
2.6	Développement analytique pour comparer la sélection sur grain et la sélection sur épi.....	22
2.6.1	Hypothèses du modèle	22
2.6.2	Estimation des paramètres de la variance	24
2.6.3	Validation expérimentale de l'équation	24
2.6.4	Comparaison de la sélection sur grain et de la sélection sur épi.....	25
3	Résultats	26
3.1	Effet de la sélection sur la taille du grain.....	26
3.1.1	Dans le dispositif de sélection in vivo	26
3.1.2	Résultats pour la sélection in silico.....	27
3.2	Comparaison de la sélection sur grain et sur épi.....	29
3.2.1	Rapport théorique entre l'efficacité de la sélection sur grain et l'efficacité de la sélection sur épi (ERS).....	29
3.2.2	Validation du développement analytique	30
3.2.3	Scenarii de sélection	30

4	Discussion	33
4.1	Retour sur les résultats de sélection	33
4.2	Recul sur l'approche théorique	35
4.3	Vers une sélection massale phénotypique sur grains ?	38
4.4	Le potentiel du dispositif de sélection in silico	39

AVANT-PROPOS

Ce mémoire a été rédigé dans le cadre de l'obtention du diplôme d'ingénieur agronome de L'institut Agro – Montpellier, option Amélioration des plantes et ingénierie végétale méditerranéennes et tropicales. Il est le fruit d'un stage de recherche de 6 mois au sein de l'équipe GE2POP de l'INRAe. Il porte sur les pratiques paysannes de sélection massale de blé dur. Ce stage a été financé par la chaire partenariale "Semences pour Demain" qui regroupe l'institut agro et l'interprofession des semences et plants (SEMAE).

Ce thème de recherche a été motivé par la volonté de faire face aux changements climatiques et d'accompagner l'agriculture dans une transition agro-écologique prenant aussi en compte des aspects économiques et sociaux. La sélection massale s'inscrit dans ces objectifs et mérite d'être étudiée pour pouvoir être mise en place.

Cette étude a deux objectifs. D'une part, mesurer les effets et l'efficacité de la sélection massale sur la taille des grains chez le blé dur. D'autre part, pouvoir comparer de façon plus théorique différentes méthodes de sélection massale chez les céréales de sorte à aiguiller les agriculteurs (ou autres personnes souhaitant faire de la sélection massale) dans leurs pratiques de sélection massale.

GLOSSAIRE

Espèce autogame : Espèce se reproduisant majoritairement par autogamie, un mode de reproduction sexuée dans lequel les deux gamètes sont issus du même individu.

Valeur génétique : Pour un trait donné, cela correspond à la somme des effets des gènes que porte un individu, qui est transmise à ses descendants.

Héritabilité : Grandeur se calculant comme la part de variance génétique dans la variance phénotypique. On peut interpréter cette grandeur comme la probabilité pour qu'une caractéristique apparente, manifeste d'un individu soit transmise héréditairement par les facteurs génétiques. On peut aussi interpréter cette grandeur comme la corrélation entre la valeur phénotypique et la valeur génétique d'un individu.

Variété : Population de plantes d'une espèce donnée qui a été sélectionnée pour la production agricole et cultivée. Au sens juridique, une variété est considérée comme telle si elle est distincte (on peut la différencier des autres variétés), homogène (tous les individus de la variété sont identiques génétiquement), stable (il est possible de reproduire la variété à l'identique de génération en génération). Ces conditions sont généralement résumées sous le terme de critère DHS.

Population / Variété population : Une variété population, ou variété de pays ou variété de ferme, est une variété cultivée traditionnelle, hétérogène, constituée d'un ensemble d'individus aux génotypes variés, sélectionnés principalement par les agriculteurs eux-mêmes, dans leurs champs. Ces variétés sont multipliées en pollinisation libre et sélectionnées par sélection massale. Ce ne sont pas des variétés au sens juridique du terme car elles ne répondent pas aux critères DHS (distinction, homogénéité et stabilité).

Sélection variétale : Processus par lequel une variété est créée. La sélection, encore appelée amélioration des plantes, consiste à créer et choisir des plantes qui présentent les caractéristiques souhaitées et répondent à des besoins ciblés (agricoles, alimentaires, médicaux, industriels, technologiques...). La sélection a pour objectif de trouver des variétés combinant plusieurs caractères favorables. Pour cela un grand nombre de plantes sont étudiées et les plus intéressantes sont croisées entre elles. La descendance des croisements est testée dans un grand nombre d'environnements et un grand nombre de caractères sont étudiés afin de sélectionner les meilleures plantes pour en faire des nouvelles variétés.

Sélection massale : Méthode de sélection variétale qui consiste à améliorer une population en sélectionnant dans celle-ci les meilleurs individus. Ce processus aboutit donc à la création de variétés populations.

SIGLES ET ACRONYMES

Vg = Variance génétique

Vinter = Variance inter-épi

Ve = Variance environnementale

Vintra = Variance intra-épi

NGO = Nombre de grains observés

NGE = Nombre de grains par épi

NEO = Nombre d'épis observés

nsel = nombre de grains sélectionnés

PMG = Poids de mille grains

GSV = Variance intra-épi de la taille des grains (Grain Size Variance)

NIRS = Near InfraRed Spectroscopy (spectroscopie proche infrarouge)

ERS = Efficacité Relative de Sélection entre sélection sur grain et sélection sur épi pour un trait du grain

TMG = taille moyenne des grains

PTE = poids total d'épis

NbEP = nombre d'épis par plante

NbE = nombre d'épillets par épi

TPG = taux de protéines des grains

TPF = taux de protéine des feuilles drapeaux à floraison

PRE = précocité

H = hauteur entre la base de la tige et le bout de l'épi en cm

1 Introduction

1.1 Le changement climatique

Le changement climatique impacte déjà l'agriculture, et ses effets vont s'accroître dans le futur. Une transition vers un modèle agricole plus résilient et durable du point de vue environnemental, social et économique est donc nécessaire (Horlings and Marsden, 2011). Pour le blé, certaines simulations prédisent une baisse de rendement de 3 à 10% par degré d'augmentation de la température (Asseng et al., 2015; Wang et al., 2018). Il se peut qu'au cours du siècle, les rendements en France baissent de 14 à 17% (Wang et al., 2018), et les rendements en zone non irriguée pourraient baisser de 9 à 30% (Ishaque et al., 2023). D'autres prédisent des diminutions ou des augmentations du rendement en fonction de la zone de production en insistant sur le caractère incertain et variable des rendements futurs (Wilcox and Makowski, 2014). Des simulations sur l'impact des maladies du blé montrent aussi des effets bénéfiques ou néfastes du changement climatique selon le pathogène et la zone géographique (Juroszek and von Tiedemann, 2013), et il semble également que la qualité des grains et leur valeur nutritionnelle puissent être détériorée (Wang et al., 2018; Zahra et al., 2023). En couplant tout cela à la croissance de la population, il est vraisemblable que le blé devienne une denrée de plus en plus chère et rare, menaçant en priorité la sécurité et la sûreté alimentaire des plus pauvres (Wang et al., 2018).

De plus, pour s'assurer un revenu dans des conditions de productions qui se dégradent à cause du changement climatique, les agriculteurs sont souvent contraints de recourir à des pratiques qui les sauvent à court terme, mais qui empirent le phénomène. Par exemple des événements de fortes pluies fréquents qui lessivent l'azote du sol peuvent contraindre les agriculteurs à utiliser plus d'engrais pour s'assurer un revenu dans un marché compétitif qui force à prioriser de forts rendements (Houser and Stuart, 2020). Cela va en contradiction avec les attentes sociétales d'une partie de la population, souhaitant une agriculture plus respectueuse de l'environnement et de la santé (Andrée et al., 2019).

1.2 Sélection massale et diversité génétique

Face à ce constat, il est nécessaire d'adapter l'agriculture et de trouver des moyens pour continuer à produire dans un climat difficile et incertain, tout en réduisant l'impact de nos productions sur l'environnement et en rémunérant les agriculteurs. À la base des filières de production se trouve la création variétale qui fournit aux agriculteurs les semences devant répondre aux objectifs et contraintes de production (rendement à atteindre, résistance aux maladies, aux stress environnementaux, conformité pour la transformation, etc...). La création variétale a donc un rôle primordial à jouer dans les transitions précédemment évoquées et se doit de proposer des innovations pour les rendre possible, tout en intégrant

les innovations, les besoins et les pratiques en aval dans la filière. Aujourd'hui, la sélection variétale est encore dominée par le paradigme qui s'est imposé durant la révolution verte dans les années 1960. Il s'agit de créer des variétés très performantes dans des conditions de cultures très artificialisées (utilisation d'engrais et de pesticides), et vouées à être utilisées en monoculture pure (Phillips and Wolfe, 2005). Les objectifs de sélection sont souvent orientés vers le rendement et la résistance aux maladies. Cette méthode de sélection, bien qu'efficace, n'est pas réalisable par les agriculteurs car elle nécessite trop d'investissement. Il est souvent déploré que cette façon de produire les variétés rend les agriculteurs dépendants à l'industrie semencière et n'est pas adaptée aux enjeux du changement climatique, aux enjeux sociaux, et aux attentes des agriculteurs soucieux de préserver l'environnement. Certains auteurs et agriculteurs parlent par exemple d'un "vide variétal" pour le blé en agriculture biologique (Chable and Berthelot, 2006). Ce paradigme est donc de plus en plus remis en question, en lien avec des réflexions globales sur les changements de pratiques nécessaires en agriculture (Dawson and Goldringer, 2012; Pingali, 2012). Notamment, la monoculture est critiquée, et les avantages de l'utilisation de la biodiversité intra et inter spécifique est mise en avant par certains chercheurs (Dawson and Goldringer, 2012).

La biodiversité végétale peut être mobilisée de différentes façon par les agriculteurs : au niveau intra-spécifique par les mélanges variétaux, et les variétés populations, au niveau interspécifiques par des cultures associées, ou des plantes compagnes par exemple. Son utilisation offre de nombreux avantages : prolongation de l'efficacité de gènes de résistance, effet tampon contre les stress biotiques et abiotiques (donc bonnes performances avec peu d'intrants), et maintien de la diversité dans l'écosystème et des services écosystémiques associés (Dawson and Goldringer, 2012). L'utilisation de variétés populations permet également la facilitation de l'accès aux ressources génétique, leur conservation in situ, et leur évolution en fonction de l'environnement (Dawson and Goldringer, 2012; Suneson, 1956). De plus, au-delà des considérations écologiques, ce type de matériel végétal est aussi adapté à une sélection participative, plus locale et maîtrisée par les agriculteurs, qui prend en compte la notion de terroir (environnement, pratiques agricoles spécifiques), des aspects sociaux économiques locaux, l'utilisation finale des grains, et s'inscrit dans un changement plus global des modèles agricoles (Dawson and Goldringer, 2012; Desclaux et al., 2012). Néanmoins, l'utilisation de la diversité intra-spécifique n'est pas encore une solution viable pour les agriculteurs car des blocages techniques, économiques et réglementaires existent (Desclaux et al., 2009). Parmi les blocages techniques, la méthodologie et les critères de sélection pour l'utilisation de la biodiversité sont encore à étudier.

La sélection massale est une technique de sélection qui s'inscrit dans la démarche de diversification au niveau intra spécifique. Elle peut permettre de maintenir et gérer une population, et est adaptée à un mode de sélection participatif ou décentralisé. Ce type de sélection est marginal et mené par des

associations la promouvant (comme le Réseau Semences Paysannes) et soutenu par certaines équipes de recherche à travers des programmes de sélection participative (Dawson et al., 2011; Mailhe et al., 2013). L'objectif de ce stage est d'étudier l'efficacité de la sélection massale en prenant comme espèce modèle le blé dur (*Triticum turgidum*).

1.3 L'importance de la qualité des grains chez le blé dur

L'alimentation humaine est le seul débouché de cette espèce qui sert principalement à la fabrication de pâtes et de couscous, mais aussi de pain dans certaines régions du monde. Sa culture en France remonte aux années 1950 dans le Sud-Est et représentait en 2017 300.000 ha répartis sur 24.000 exploitations, ce qui représente 4% de la surface céréalière. Chaque année, 1.8 Mt sont produites ce qui représente 3% de la production céréalière française, faisant du blé dur la quatrième céréale de France. Les principaux bassins de production sont le Sud-Est où il représente 80% des surfaces de céréales, le Sud-Ouest, l'Ouest et le Centre. A l'échelle mondiale, la production Française représente 5% de la production totale, les principaux producteurs étant le Canada et l'Italie. Le blé dur est adapté à des climats plutôt chauds et secs. Il se sème principalement à l'automne, a des graines peu dormantes, des besoins de vernalisations quasiment nuls, et supporte mal le froid hivernal (GIE Blé dur, 2017). Le blé dur est donc une culture qui permet de produire des aliments de base en Europe et dans bassin méditerranéen et des grains de bonne qualité sont nécessaires pour produire ces aliments.

La taille des grains et le poids de mille grains (PMG) sont des traits importants de la qualité. En effet, il sont liés au rendement semoulier, à la texture et la couleur de la semoule et des pâtes (Wang and Fu, 2020). La masse individuelle des grains a aussi une importance sur l'établissement du rendement en conditions de stress thermique post floraison (Sharma et al., 2008), et sur la vigueur germinative, ce qui peut impacter indirectement le rendement (Finch-Savage and Bassel, 2016). La taille des grains est donc un trait intéressant à sélectionner dans un contexte d'agriculture à bas intrants et de sélection massale car c'est une sélection indirecte sur le rendement et la résistance à un stress, et une sélection directe sur l'utilisation finale des grains. Il serait donc intéressant de pouvoir sélectionner des populations avec des grains de grande taille pour la qualité, et qui soient plutôt homogènes pour une bonne germination et de bonnes performances technologiques. Sélectionner sur la taille des grains est *a priori* possible car la taille des grains a doublé entre les premières variétés de blé et les variétés modernes (Donald, 1981 cité dans Sadras, 2007).

1.4 Les pratiques de sélection massale chez le blé dur

Chez le blé, la sélection massale s'opère en conservant les plus beaux épis, ou les épis des plus belles plantes et en semant leurs grains (Dawson et al., 2011; Mailhe et al., 2013). Cette pratique permet de sélectionner sur la partie végétative de la plante (vigueur, aspect sanitaire, productivité) mais elle ne donne pas directement accès à l'aspect des grains à moins d'effectuer un battage épi par épi. L'observation des épis n'est à notre connaissance pas réalisable à haut débit. C'est donc une méthode coûteuse en temps et qui ne permet pas d'observer beaucoup de plantes, ce qui limite les possibilités de sélection pour la qualité des grains. Etant donné que les caractéristiques des grains sont importantes, on pourrait proposer de faire la sélection en observant les grains individuels plutôt que les épis. En effet, l'observation du grain permet de prédire (avec plus ou moins de précision) sa qualité germinative et sa valeur génétique pour les traits liés à sa morphologie. C'est aussi l'organe de la plante qui paraît le plus facile à observer à haut débit en triant sur la taille avec un tamis ou sur le taux de protéine avec un trieur optique équipé d'infrarouge par exemple. Enfin, c'est le produit de la récolte. On peut le sélectionner après la moisson et cela ne nécessite pas d'observations au champ. Par contre, une sélection sur le seul grain implique qu'on ne peut pas sélectionner directement sur des traits végétatifs, par exemple la hauteur de la plante ou la résistance aux maladies. Cependant la morphologie du grain est liée à l'histoire de vie de la plante et il existe des corrélations entre certains de ces traits et ceux du grain, ouvrant des possibilités de sélection indirecte.

Il serait utile de pouvoir comparer l'efficacité d'une sélection sur grain individuel à celle d'une sélection sur la valeur moyenne des grains d'un épi pour optimiser une sélection massale pour la qualité du grain. Sur la taille des grains, on pourrait intuitivement penser que la sélection sur épi est toujours plus efficace que la sélection sur grain. En effet, il existe toujours une forte variabilité de la taille des grains au sein d'un épi (Béral, 2020). Ainsi, les génotypes à gros grains font des épis contenant de très gros grains, des grains moyens et des grains plus petits transmettant tous une grosse taille aux grains de la génération suivante (c'est-à-dire ayant tous une bonne valeur génétique). Cela illustre les résultats historiques de Johannsen, (1903) sur les notions de génotypes et de phénotypes. Les génotypes à plus petits grains produisent dans le même épi des grains se répartissant entre petits, moyens, et gros, qui transmettent tous une plus petite taille aux grains de la génération suivante. Ainsi, en faisant une sélection massale sur les grains pour en augmenter la taille, la majorité des gros grains proviendront de génotype à gros grains mais certains proviendront aussi de génotypes à petits grains, et certains grains de bonne valeur génétique ne seront pas sélectionnés. En faisant une sélection sur épi, on sélectionne sur la moyenne d'une famille apparentée de grains et cette moyenne est bien plus prédictive de la valeur génétique que celle des grains individuels. Cette opposition est classique en théorie de la sélection : une moyenne familiale prédit mieux la valeur des individus que leur valeur individuelle. On peut aussi dire que la

moyenne familiale a une meilleure héritabilité que les valeurs individuelles. A notre connaissance, peu de travaux se sont intéressés à la modélisation théorique de la situation grain vs épi.

A priori la sélection sur épi apparaît comme la meilleure des deux méthodes puisque plus précise (mesure plus héritable). Cependant, en la replaçant dans un contexte où un nombre de grain fixé doit être sélectionné (lorsqu'un agriculteur ou un sélectionneur alloue une partie de sa surface à sa sélection massale), on peut ajouter les arguments suivant :

En sélection sur grain individuel, il est aisé de traiter un très grand nombre de grains, par exemple avec un tamis pour sélectionner sur la taille. A l'extrême, une récolte entière pourrait passer au tamis pour ne récupérer que les grains souhaités. Appelons n_{sel} le nombre de grains sélectionnés et NGO le nombre de grains observés, la proportion de grains sélectionnés dans la population totale de grains est $P\%_{grain} = \frac{n_{sel}}{NGO}$.

En revanche, observer les épis est beaucoup plus fastidieux. Il n'existe pas à notre connaissance d'équipement permettant d'automatiser le battage et la mesure de la moyenne des grains épi par épi sur de très grands effectifs. Sans équipement, la sélection sur épi se ferait plutôt de façon visuelle, et même de cette façon, il paraît infaisable d'observer tous les épis d'un champ pour sélectionner. Il faut donc échantillonner un certain nombre d'épis dans la population pour les observer, et la sélection ne s'opérera que sur cet effectif. Appelons NEO le nombre d'épis observés et NGE le nombre de grains par épi. Pour sélectionner n_{sel} grains, il faut sélectionner $\frac{n_{sel}}{NGE}$ épis (en supposant que chaque épi donne NGE grains). Ainsi, la proportion d'épis sélectionnés dans la population d'épis observés est $P\%_{epi} = \frac{n_{sel}}{NGE \times NEO}$.

En imaginant un exemple réaliste de sélection sur la taille du grain par tamis, mettons qu'un paysan-sélectionneur alloue chaque année une surface constante de 100 m² pour améliorer sa population. Avec une densité de semis de 250 grains par m², il lui faut chaque année sélectionner 25.000 grains pour ressemer les 100 m² ($n_{sel} = 25.000$). Lors de sa récolte, moyennant 1.5 épis par plante et 60 grains par épi ($NGE = 60$), il obtient 2.250.000 grains (soit à peu près 80 kg de grains pour un PMG de 35 grammes, $NGO = 2.250.000$). Ces grains sont issus de 50.000 épis. Imaginons que le paysan sélectionneur observe 1.000 épis en tout ($NEO = 1000$, ce qui est déjà un travail conséquent), on obtient alors $P\%_{grain} = 1.1\%$ et $P\%_{epi} = 35\%$.

Replaçons maintenant ces deux parties de raisonnement dans l'équation du sélectionneur (Lush, 1943): $R = H^2 \cdot i \cdot \sigma_p$ avec R le progrès génétique, i l'intensité de sélection en unité d'écart-type, H^2 l'héritabilité, et σ_p l'écart type phénotypique du trait. Plus la proportion $P\%$ est faible, plus l'intensité de sélection i est forte. Pour les deux méthodes, H^2 , i et σ_p ont des valeurs différentes.

Les mesures de moyenne par épi auront une plus forte héritabilité que les mesures de grains individuels (justification complète en annexe 1) mais l'intensité de sélection sur les grains individuels est bien plus élevée que l'intensité de sélection sur épi. L'équation du sélectionneur fait apparaître entre les deux méthodes de sélection un compromis entre l'héritabilité et l'intensité de sélection. Il est donc difficile de savoir quelle méthode est la plus avantageuse et étudier de manière analytique cette contradiction donnerait des bases pour explorer leurs intérêts respectifs.

Durant le stage, nous avons donc tenté de répondre à deux questions :

- Quel sont les effets directs et indirects d'une sélection massale sur la taille du grain ?
- Dans quelles conditions est-il préférable de sélectionner sur grain ou sur épi ?

Pour cela, deux expérimentations ont été réalisées. La première est l'étude au champ des effets d'une sélection sur la taille individuelle des grains. La seconde est une étude en condition contrôlées où les descendances de plusieurs centaines de grains de génotypes et phénotypes connus ont été étudiées pour simuler l'effet qu'aurait eu une sélection sur la taille du grain. Enfin une approche analytique a été conduite pour formaliser l'efficacité relative de la sélection entre sélection sur grain et sélection sur épi pour un trait du grain (ERS).

2 Matériel et méthodes

Toutes les analyses statistiques décrites dans cette partie ont été réalisées avec le logiciel R (R Core Team, 2022).

2.1 Matériel végétal

Le matériel végétal utilisé durant ce stage est une population de blé dur à base génétique large appelée EPO. Cette population a été constituée en 1997 en croisant des accessions de *T. turgidum* contemporaines avec des accessions sauvages et primitives. Un gène de stérilité mâle ségrége dans la population pour assurer des allofécondations. Chaque année la population est ressemée pour être maintenue. En 2009, 180 lignées ont été extraites de la population alors à sa 17^{ième} génération et ont été fixées par la méthode single seed descent pendant 2 générations (David et al., 2014). Cette population a servi à plusieurs études de génétique et renferme une large diversité génétique pour de nombreux caractères (Ballini et al., 2020; Colombo et al., 2022; Montazeaud et al., 2022). La population ainsi que les lignées qui en sont extraites sont le matériel végétal de ce stage.

2.2 Dispositifs expérimentaux

2.2.1 Dispositif de sélection *in vivo*

Sur l'unité expérimentale INRAe DIASCOPE à Mauguio, un semis a été fait en prenant des grains directement issus de la population composite EPO (en mélange, non fixée et ségrégeant pour l'allèle de stérilité mâle). Le semis s'est fait en plein champ dans des micro parcelles de 1.5 m² le 10/11/2022 chaque parcelle contenant 6 rangs. Avant le semis, les grains ont été tamisés sur des mailles de 34 et 35 mm. Ils ont été séparés entre "gros" (>35 mm), "moyens" (34 < 35), et "petits" (< 34 mm). Une partie des grains n'a pas été tamisée et constitue la population de référence. Les traits morphologiques des grains semés ont été mesurés individuellement (une mesure par grain) par analyse automatique d'image à l'Optomachine¹. L'Optomachine calcule une centaine d'indicateurs, notamment la surface des grains (en mm²). Cette surface est appelée "taille du grain" dans la suite du rapport. L'Optomachine pèse également les grains et calcule le PMG.

Au total 12 micro parcelles, donc 3 répétitions de chaque traitement ont été disposées au hasard. Les micro-parcelles étaient organisées en passages et en planches (qui correspondaient à des coordonnées spatiales). Elles ont été conduites selon un itinéraire technique bio, avec irrigation (aucune mesure précise de la quantité d'eau apportée n'est disponible) et sans traitements phytosanitaires. La récolte a eu lieu de 3 juillet.

2.2.2 Dispositif de sélection *in silico*

Un autre dispositif a été mis en place au bâtiment ARCAD (10 rue Arthur Young, 34000, Montpellier, France). Six bacs de 1 mètre carré ont été semés avec les 180 lignées EPO fixées. Pour chaque lignée, des lots de 12 grains non cassés et non échaudés ont été choisis pour participer au semis. Sur chaque grain, deux spectres NIRS ont été acquis avec un spectromètre ASD², et la taille des grains a été mesurée à l'Optomachine.

Dans chaque bac, 208 grains ont été semés (13 lignes et 16 colonnes) en les répartissant au hasard. Il s'agit d'un plan complet équilibré, chaque génotype est semé au moins une fois dans chaque bac et la position de chaque grain est connue. Pour respecter les distances de semis les grains ont été collés avant semis sur une bande de papier facilement dégradable avec un empois d'amidon fabriqué à partir de farine. Le semis a été fait le 06/01/2023. Un semis de rattrapage a été réalisé le 17/02/2023 pour remplacer les défauts de germination. En moyenne, 7 grains par génotype ont été semés parmi les 12 de

¹ Station de Caractérisation des semences Optoagrimetric développé par la société Optomachines : <https://optomachines.fr/accueil/optomachines-equipements-danalyse-de-semences/optomachines-opto-agrimetric-caracterisation-de-semences/>

² Spectromètre Fieldspec 2500© (Analytical Spectral Devices, Inc. (ASD), Boulder, CO, USA) utilisé en réflexion avec une fibre bifurquée de 3 mm de diamètre

départ. Semer aléatoirement les grains complique l'acquisition des données mais permet de simuler l'effet d'une sélection et d'obtenir les valeurs génétique des génotypes en population, et non en monoculture ou avec un nombre restreint de voisin. Les bacs ont été irrigué une fois par semaine jusque début mars puis deux fois par semaine jusque mi-juin (aucune mesure précise de la quantité d'eau apportée n'est disponible). Enfin, pour simuler un environnement limité en ressources chaque bac n'a reçu que 30 unités d'azote le 20/04. La récolte a eu lieu le 1 juillet.

2.3 Phénotypage

Dans le dispositif de sélection *in vivo*, les deux rangs centraux de chaque micro-parcelle ont été récoltés. Dans cette récolte, 30 tiges ont été choisies aléatoirement en faisant attention à ne pas prendre deux tiges de la même plante et les traits suivants ont été mesurés :

Hauteur du brin entre la base de la tige et la pointe de l'épi en cm (H), taux de protéine des grains en % de la masse (par NIRS des grains) (TPG), PMG en g, nombre d'épillets (NbE), nombre de grains par épi (NGE), taille individuelle de chaque grain en mm² (TIG) par mesure Optomachine, taille moyenne des grains par épi en mm² (TMG), variance de la taille des grains en (mm²)² (GSV).

Pour le dispositif de sélection *in silico*, chaque plante a été mesurée individuellement, et leurs brins maîtres ont été récoltés. Les mesures effectuées sur les brins récoltés sont les même que pour le dispositif de sélection *in vivo*, avec en plus les mesures suivantes :

Taux de protéines dans la feuille drapeau à floraison en % de la masse (par NIRS des feuilles) (TPF), date d'épiaison en degrés jours depuis le semis (PRE), nombre d'épis par plante (NbEP), poids total des épis de la plante (PTE).

Les spectres NIRS des grains ont été acquis avec un spectromètre Perkin-Elmer³ et les spectres des feuilles par un spectromètre ASD.

³ Perkin-Elmer Spectromètre FTIR Frontier © (PerkinElmer), utilisé en réflexion avec le module autosampler, dans des vials de 5 mL

2.4 Mesure de l'effet de la sélection dans le dispositif de sélection in vivo

Les caractéristiques des différentes modalités sélectionnées ont été mesurées directement et ont été comparées à celles de la population de référence. Les changements de valeurs des traits ont été estimé par le modèle suivant :

$$Y_{ijkl} = \mu + R_i + PASSAGE_j + PLANCHE_k + \varepsilon_{ijkl}$$

Avec :

Y_{ijkl} le phénotype de la plante k dans le groupe sélectionné i dans le passage j dans la planche k

μ la moyenne des observations

R_i l'effet de la sélection, effet qualitatif prenant 4 valeurs : "référence", "gros", "moyen", "petit"

$PASSAGE_j$ l'effet du passage j

$PLANCHE_k$ l'effet de la planche k

ε_{ijkl} la résiduelle tel que $\varepsilon_{ijkl} \sim N(0, \sigma)$ iid

L'effet de la sélection a été estimée avec les coefficients R_i en choisissant la modalité "référence" comme référence et la significativité de chaque coefficient a été testée par test de Student. Avec les données morphologiques des grains semés, la différentielle de sélection sur la taille des grains (S) a été calculée pour chaque modalité de sélection. Cela a permis de calculer l'héritabilité réalisée (H^2) pour la taille des grains en divisant les progrès estimés par les différentielles de sélection :

$$S_{modalité} = \mu_{modalité} - \mu_{référence} \quad \text{et} \quad H^2_{modalité} = \frac{R_{modalité}}{S_{modalité}}$$

Modalité étant "gros", "moyen", ou "petit" et μ étant la moyenne de la taille des grains.

2.5 Mesure de l'effet de la sélection dans le dispositif de sélection in silico

2.5.1 Simulation de la sélection

Dans ce dispositif, la sélection sur les valeurs du grain semé et l'effet sur leur descendance ont été simulés. Seules les données des grains issues du premier semis (06/01) ont été utilisées pour éviter de possibles interaction génotype x date de semis. Les simulations ont été réalisées en faisant varier les paramètres nsel, et NEO présentés dans l'introduction pour réaliser la sélection sous différentes

intensités. Les paramètres NGO et NGE n'ont pas été manipulés car imposés par les données. L'effet de la sélection sur grain individuel a été simulée en calculant la valeur des caractères observés sur les descendants de nsel grains semés triés selon leur taille.

La sélection sur épi a été simulée en assimilant la moyenne de la taille des 12 grains des lots de chaque génotype retenus pour le semis à la valeur d'un épi de ce génotype, bien que ces grains soient issus d'une récolte d'une ligne complète de plusieurs plantes. Puis, un nombre NEO de lots a été échantillonné aléatoirement pour constituer la population d'épis observés. Cela permet de simuler l'impossibilité d'observer tous les épis d'une parcelle en sélection sur épi (on pourrait aussi dire que cela simule l'effort de phénotypage réalisé en sélection sur épi). Parmi ces NEO lots échantillonnés aléatoirement, les nsel/NGE lots ayant les plus grandes moyennes ont été sélectionnés de sorte à sélectionner nsel grains. L'effet de la sélection sur épi a été simulé en calculant la valeur des caractères observés sur les descendants de ces grains.

L'effet de la sélection a été estimée en comparant les populations sélectionnées à une population de référence de nsel grains échantillonnés aléatoirement dans les données. Comme des échantillonnages aléatoires interviennent lors de ces simulations, 100 répétitions ont été effectuées pour chaque combinaison nsel x NEO testées.

2.5.2 Estimation des héritabilités

L'héritabilité a été estimée pour chaque trait avec les données utilisées pour simuler la sélection. Le modèle suivant a été utilisé :

$$Y_{ijk} = \mu + G_i + BAC_j + \varepsilon_{ijk}$$

Avec :

Y_{ijk} le phénotype de l'individu l de génotype i dans le bac j

μ la moyenne des observations

$G_i \sim N(0, \sigma_g)$ iid l'effet aléatoire du génotype

BAC_j l'effet fixe du bac j

$\varepsilon_{ijk} \sim N(0, \sigma_r)$ iid la résiduelle

Pour chaque trait l'héritabilité a été calculée comme $H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_r^2}$

2.5.3 Estimation de l'effet de la sélection

Pour chaque simulation réalisée, l'effet de la sélection a été estimé avec le modèle suivant :

$$Y_{ijk} = \mu + BAC_i + SELECTION_j + \varepsilon_{ijk}$$

Avec :

- Y_{ijk} le phénotype de l'individu k dans le bac i avec la modalité de sélection j
- μ la moyenne des observations
- BAC_i l'effet fixe du bac i sur le phénotype
- $SELECTION_j$ l'effet fixe de la sélection prenant trois modalités : "grain", "épi", et "référence"
- $\varepsilon_{ijk} \sim N(0, \sigma)$ iid la résiduelle

L'effet de la sélection a été estimé par les coefficients de la variable SELECTION en prenant la modalité "référence" comme référence. La significativité des effets estimés a été testée par test de Student. Le processus de sélection *in silico* est schématisé en figure 1.

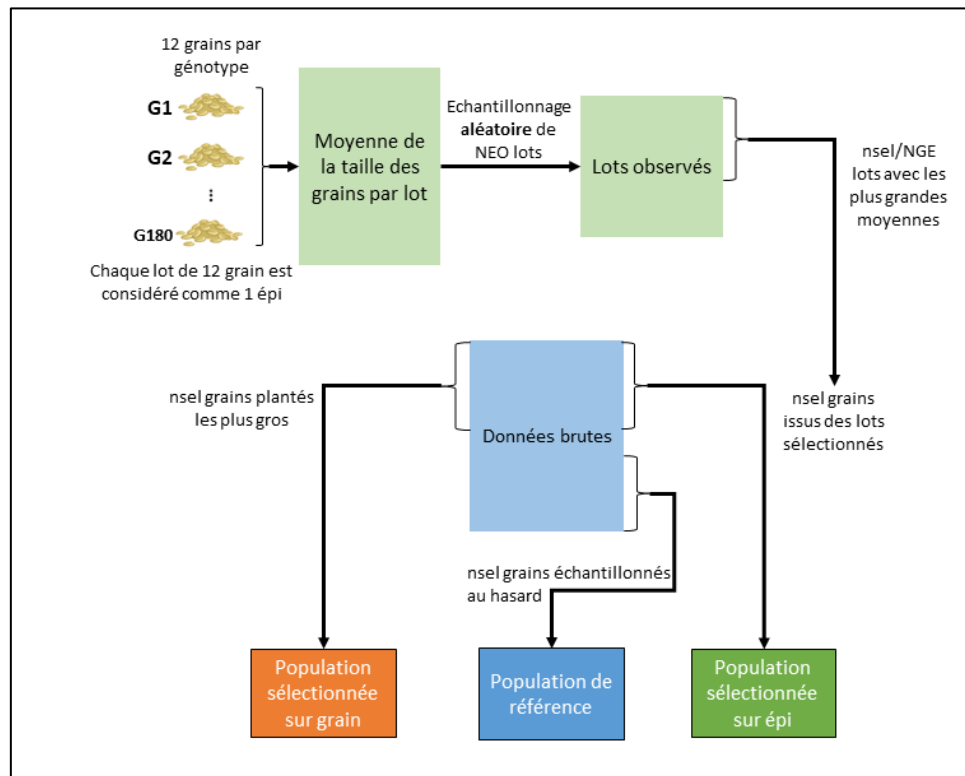


Figure 1 : Schéma récapitulatif de la méthode de sélection *in silico*

Les paramètres nsel et NEO sont choisis pour explorer différents scénarii de sélection. La sélection sur grain se fait en regardant directement la taille des grains plantés à l'échelle individuelle. La sélection sur épi se fait en prenant les grains issus des lots ayant la meilleure moyenne. La population non sélectionnée de référence est tirée aléatoirement dans les données. On utilise un modèle prenant en compte l'effet des bacs et de la sélection pour estimer l'effet de la sélection.

2.6 Développement analytique pour comparer la sélection sur grain et la sélection sur épi

2.6.1 Hypothèses du modèle

On considère le grain et l'épi comme deux unités de sélections différentes. Pour comprendre les différences qui existent entre ces deux unités de sélection (grain individuel vs. moyenne des grains d'un épi), il faut s'intéresser aux paramètres importants en sélection : la variance phénotypique totale du trait sélectionné, la précision avec laquelle une mesure phénotypique permet d'estimer la valeur génétique d'une unité de sélection (l'héritabilité), et l'intensité de sélection appliquée en fonction du nombre d'unités de sélection observées. Commençons par détailler les composantes de la variance phénotypique d'un trait du grain :

- V_g la variance génétique qui existe entre génotypes différents (toute chose égale par ailleurs)
- V_e la variance environnementale qui existe entre grains issus de plantes de même génotype ayant poussé dans des conditions différentes (comme des parcelles différentes dans le même lieu ou des lieux différents la même année, etc.)
- V_{inter} la variance inter-épi qui existe entre épis d'une même plante pour des raisons physiologiques (par exemple entre brin maître et brin secondaire (Acreche and Slafer, 2006))
- V_{intra} la variance intra-épi qui existe entre grains d'un même épi pour des raisons physiologiques également (par exemple grains des fleurs proximales vs fleurs distales (Béral, 2020))

Considérons maintenant une parcelle où la sélection sur grain est appliquée et une parcelle identique où la sélection sur épi est appliquée pour un même trait mesurable sur le grain. Les différences d'intensités de sélections entre les deux types de sélections seront dues à :

- NGO Le nombre de grains observés en sélection sur grain
- NEO le nombre d'épi observés en sélection sur épi
- $nsel$ le nombre de grains à sélectionner pour semer la génération suivante, le même pour la sélection sur épi et la sélection sur grain.
- NGE le nombre de grains par épi

L'efficacité relative d'une sélection sur grain par rapport à la sélection sur épi (ERS) peut être étudié avec le rapport :

$$ERS(NEO, NGO, nsel, NGE, Vg, Ve, Vinter, Vintra) = \frac{R_{grain}}{R_{epi}} = \frac{H_{grain}^2 \times i_{grain} \times \sigma_{grain}}{H_{epi}^2 \times i_{epi} \times \sigma_{epi}}$$

Le développement analytique effectué a consisté à exprimer les écart-types phénotypiques et les héritabilités en fonction des composantes de la variance, et les intensités de sélection en fonction des paramètres NEO, NGO, NGE, nsel. Tous les calculs sont disponibles en annexe 2 et ont été vérifiés avec le logiciel Mathematica (Wolfram Research inc). Le raisonnement et les hypothèses du modèle sont schématisés en figure 2.

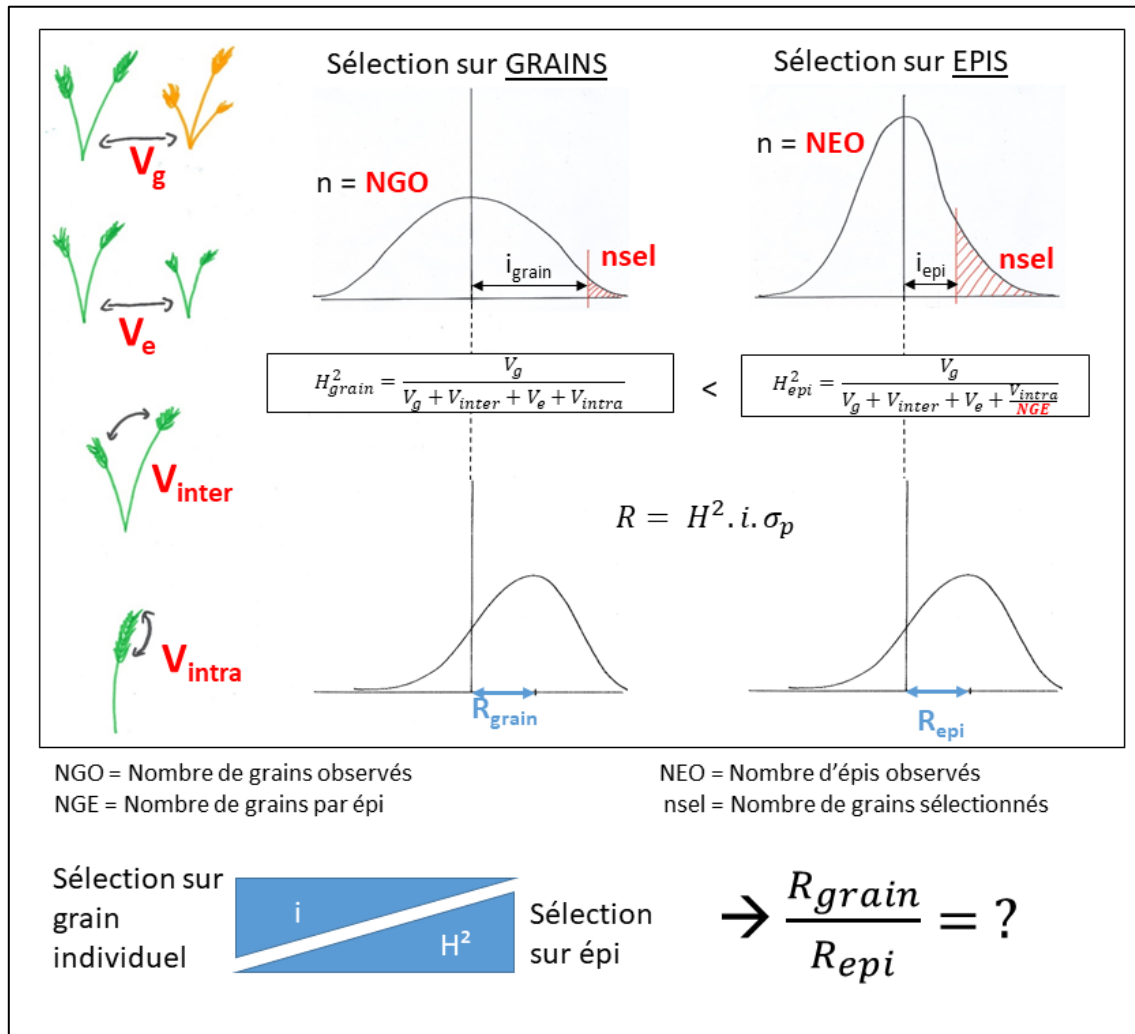


Figure 2 : Schéma récapitulant le raisonnement pour comparer la sélection sur grain et sur épi.

On retrouve les composantes de la variance des traits du grain à gauche. La sélection sur grain se fait en sélectionnant nsel grains parmi une population de NGO grains observés. Selon la valeur de l'héritabilité H^2_{grain} et l'intensité de sélection i_{grain} , on obtient un progrès R_{grain} . La sélection sur épi se fait en sélectionnant un nombre d'épis tel qu'on obtienne nsel grains à la fin, en observant une population de NEO épis. Selon la valeur de l'héritabilité H^2_{epi} et l'intensité de sélection i_{epi} , on réalise un progrès R_{epi} . On rappelle en bas le compromis entre l'intensité de sélection et l'héritabilité qui amène à se demander quelle méthode de sélection est la plus efficace.

2.6.2 Estimation des paramètres de la variance

Le calcul de l'ERS à partir de l'équation développée nécessite d'avoir au préalable une estimation des valeurs des composantes de la variance. Dans le dispositif de sélection *in silico* un échantillonnage supplémentaire a été réalisé pour estimer ces composantes pour la taille des grains. Pour 35 plantes par bac, deux épis ont été récoltés pour lesquels la taille des grains a été mesurée individuellement par Optomachine. L'échantillonnage de deux épis par plante permet d'estimer la variance inter-épi. L'échantillonnage a été fait de sorte que 103 génotypes soient représentés (pour couvrir la variance génétique), et que chaque génotype soit présent dans deux bacs différents (pour estimer la variance environnementale inter-bac). La variance intra-épi a aussi pu être estimée par la mesure individuelle de chaque grain. L'estimation des variances s'est faite avec le modèle suivant :

$$Y_{ijkl} = \mu + G_i + ind_{i(j)} + epi_{ij(k)} + \epsilon_{ijkl}$$

Avec :

Y_{ijkl} la taille du grain l de l'épi k de l'individu j de génotype i

μ la moyenne des observations

$G_i \sim N(0, \sigma_g)$ l'effet aléatoire du génotype i

$ind_{i(j)} \sim N(0, \sigma_e)$ l'effet aléatoire de l'individu j hiérarchisé dans l'effet génotype i (donc l'effet de l'environnement et de l'interaction génotype x environnement car les individus de même génotype se trouvent dans des bacs différents)

$epi_{ij(k)} \sim N(0, \sigma_{inter})$ l'effet aléatoire de l'épi k hiérarchisé dans l'effet individu j et l'effet génotype i

$\epsilon_{ijkl} \sim N(0, \sigma_{intra})$ la résiduelle qui contient la variance intra-épi

2.6.3 Validation expérimentale de l'équation

Pour valider l'équation obtenue, les résultats de sélection *in silico* sur la taille des grains ont été utilisés car on a pu estimer les composantes de la variance pour ce trait. Pour 132 de scenarii de sélection l'ERS a été calculée expérimentalement grâce aux estimations d'effets de la sélection (voir paragraphe 2.5.3), et théoriquement avec l'équation développée. Les valeurs expérimentales et théoriques ont été confrontées avec un modèle linéaire simple $R_{experimental_i} = \alpha + \beta R_{theorique_i} + \epsilon_i$ où $\epsilon_i \sim N(0, \sigma)$ iid.

2.6.4 Comparaison de la sélection sur grain et de la sélection sur épi

L'ERS a été calculée avec l'équation développée pour de nombreux scénarii de sélection. Nous avons fait varier NEO de 500 à 5000, nsel de 10.000 à 100.000, NGO de 10.000 à 5.000.000, et NGE de 40 à 80. Les variances estimées avec le dispositif de sélection *in silico* ont été utilisées pour ces calculs. Les composantes de la variance étant relatives à la taille des grains, l'ERS a été calculée pour ce trait. Ci-dessous (figure 3) un schéma récapitulant les dispositifs expérimentaux :

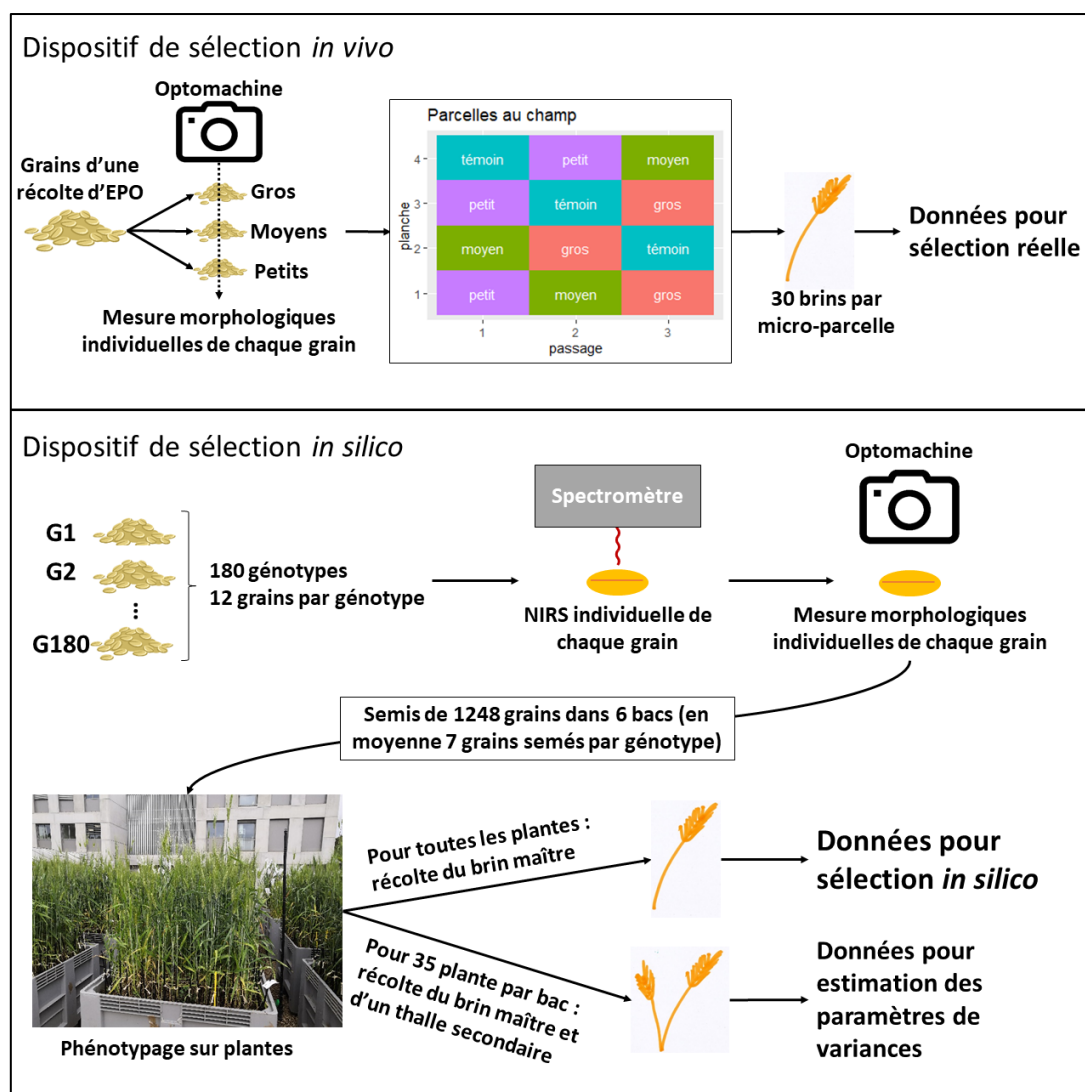


Figure 3 : Schéma récapitulatif des dispositifs expérimentaux.

En haut, le dispositif de sélection *in vivo*. Des grains issus d'une récolte d'EPO ont été tamisés en 3 classes différentes (petits, moyens, gros, + la référence non tamisée) puis semés au champ dans 12 micro parcelles. Les deux rangs centraux de chaque parcelle ont été récoltés pour faire les mesures. En bas, le dispositif de sélection *in silico*. 12 grains des 180 lignées issues de la population EPO ont été passés à la NIRS et à l'Optomachine puis en moyenne 7 grains par génotypes ont été semés dans 6 bacs. Les mesures ont été réalisées sur chaque plante individuellement. Le brin maître de chaque plante et son épi ont été récoltés pour les mesures. Pour 35 plantes par bac, un thalle secondaire et son épi ont également été récolté pour estimer les composantes de la variance.

3 Résultats

3.1 Effet de la sélection sur la taille du grain

3.1.1 Dans le dispositif de sélection *in vivo*

Les effets estimés pour les traits mesurés dans le dispositif de sélection *in vivo* sont donnés en figure 4.

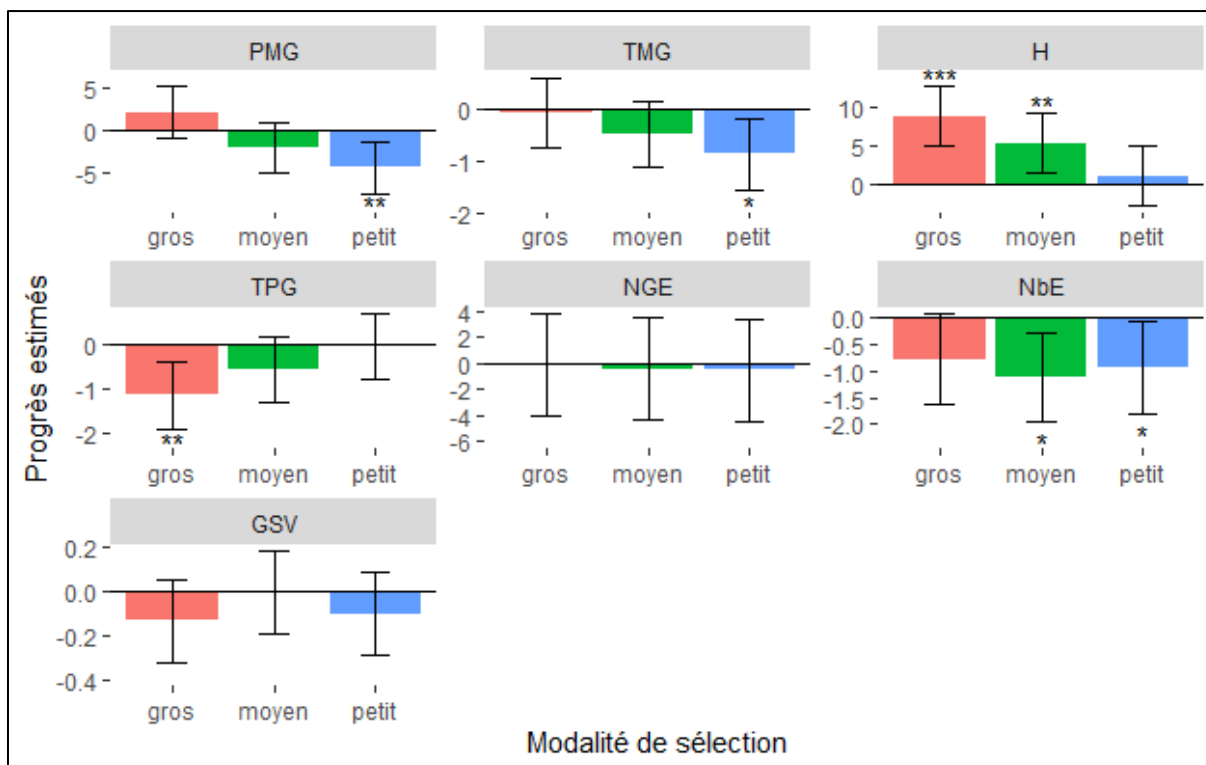


Figure 4 : Effets de la sélection sur la taille des grains dans le dispositif de sélection *in vivo*.

Les effets sont estimés en prenant la population non sélectionnée comme référence. Les barres verticales représentent les intervalles de confiance à 95% des estimations. Les effets ont été estimés pour les trois modalités de sélection : "gros", "moyen", et "petit". Les axes verticaux sont chacun dans l'unité du trait mesuré et ont leurs propres limites. PMG = poids de mille grains en g, TMG = taille moyenne des grains en mm², H = hauteur entre la base de la tige et le bout de l'épi en cm, TPG = taux de protéines des grains en % de la masse, NGE = nombre de grains par épis, NbE = nombre d'épillets, GSV = variance intra-épi de la taille des grains en (mm²)². Les étoiles indiquent si l'effet est significativement différent de 0. * = *p*-value < 0.05, ** = *p*-value < 0.01, *** = *p*-value < 0.001.

La sélection de gros grains n'a pas significativement impacté la taille des grains par rapport à la référence. La sélection de petits grains a en revanche modifié significativement la taille de -0.9 mm². On constate un effet similaire sur le PMG qui a été diminué de -4.4 g par la sélection de petits grains. Pour la hauteur des plantes et le taux de protéine des grains, la sélection des gros grains a eu un effet significatif (hauteur = + 8.7 cm, taux de protéines des grains = -1.14%) mais pas la sélection des petits. Les effets estimés pour le nombre de grain par épi et la variance intra-épi de la taille des grains ne sont pas significativement différents de 0. Les effets estimés pour le nombre d'épillets sont significatifs pour

la sélection des moyens et des petits, mais l'effet de la sélection des gros est presque significatif (p -value = 0.07). Toutes les modalités de sélection ont modifié le trait dans le même sens (diminution du nombre d'épillet).

Les héritabilités réalisées et les intensités de sélection appliquées pour la taille des grains sont présentés dans le tableau 1 :

	PETIT	MOYEN	GROS
i	-1.25	-0.03	1
S	-4.71	-1.17	3.79
R	-0.9	-0.5	-0.08
H ²	0.19	0.43	-0.02

Tableau 1 : Héritabilités, progrès réalisés, et intensités de sélection pour la taille des grains dans le dispositif de sélection in vivo.

Les valeurs sont calculées pour la taille des grains individuels. Tous les grains semés ont été mesurés à l'Optomachine pour avoir les tailles des grains individuels et calculer les différentielles de sélection pour chaque modalité (moyenne de la taille des grains dans la modalité – moyenne de la taille des grains non sélectionnés). S = différentielle de sélection en mm², i = intensité de sélection en unité d'écart-type des grains non sélectionnés, R = progrès estimé en mm², H^2 = héritabilité réalisée calculée comme R/S .

Les intensités de sélection ne permettent pas vraiment d'expliquer les différences de progrès effectués car les intensités de sélection sur petits et gros grains sont proches (en valeur absolue) tandis que l'intensité de sélection sur grains moyens est plus faible. On voit en revanche que l'héritabilité réalisée pour la sélection sur gros grains est nulle. Il semble donc que la sélection ait fonctionné pour réduire la taille du grain mais pas pour l'augmenter.

3.1.2 Résultats pour la sélection in silico

L'héritabilité des traits mesurés à l'échelle de la plante sont en tableau 2. Pour la taille des grains, l'héritabilité individuelle était de 0.13, et l'héritabilité par épi était de 0.2.

PMG	PTE	NbEP	NGE	TPG	TPF	PRE	H	GSV
0.17	0.1	0.1	0.04	0.05	0.07	0.36	0.1	0.13

Tableau 2 : Héritabilités individuelles pour les traits mesurés dans le dispositif de sélection in silico.

PMG = poids de mille grains, PTE = poids total d'épis, NbEP = nombre d'épis par plante, NGE = nombre de grains par épis, TPG = taux de protéines des grains, TPF = taux de protéine des feuilles drapeaux à floraison, PRE = précocité, H = hauteur entre la base de la tige et le bout de l'épi, GSV = variance intra-épi de la taille des grains.

Nous avons choisi de représenter les résultats obtenus pour le nombre d'épi observé le plus grand possible (NEO = 177, ce qui permet la plus grande intensité de sélection possible sur épi) et pour nsel valant 250, qui permet d'avoir une intensité de sélection suffisante pour voir l'effet de la sélection tout en gardant un assez grand nombre d'individus pour l'estimer (figure 5).

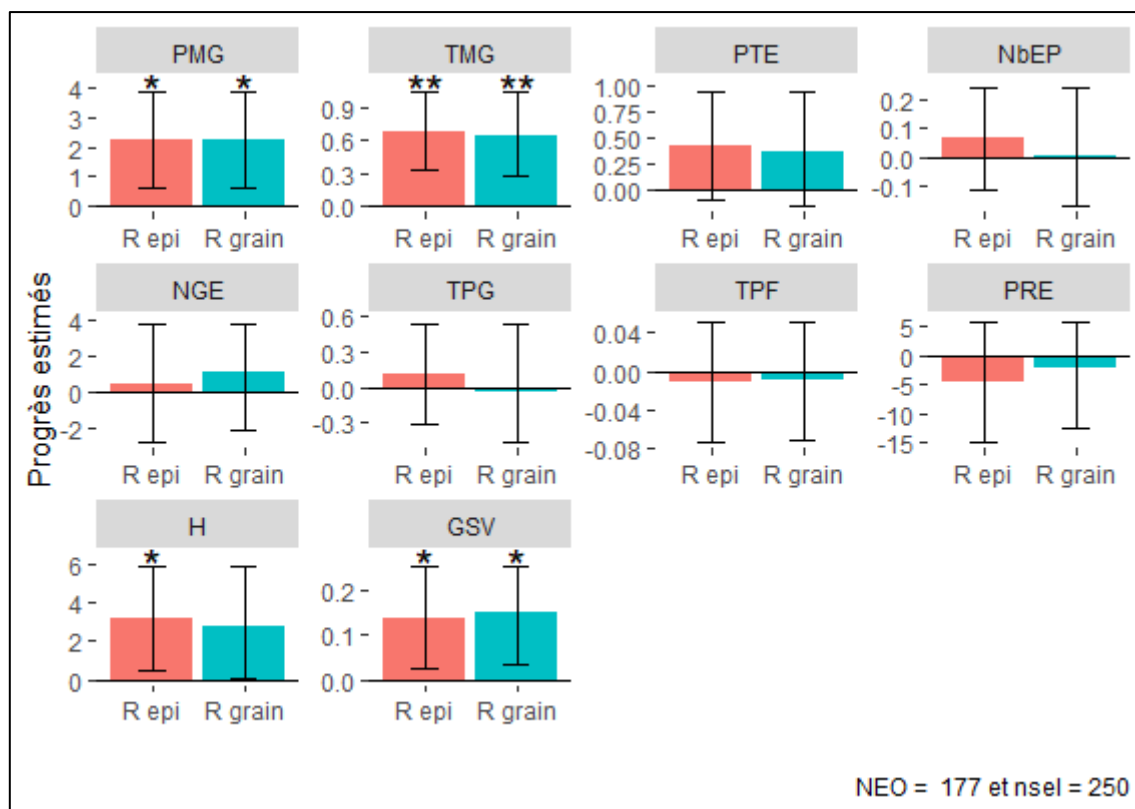


Figure 5 : Effet de la sélection sur la taille du grain dans le dispositif de sélection *in silico*.

Les valeurs présentées sont les moyennes des 100 simulations réalisées. Les barres verticales représentent les intervalles de confiance à 95% des estimations. Les effets ont été estimés pour les deux méthodes de sélection : sélection sur épi (R_epi), et sélection sur grain (R_grain). Les axes verticaux sont chacun dans l'unité du trait mesuré et ont leurs propres limites. PMG = poids de mille grains en g, TMG = taille moyenne des grains en mm², PTE = poids total d'épis en g, NbEP = nombre d'épis par plante, NGE = nombre de grains par épis, TPG = taux de protéines des grains en % de la masse, TPF = taux de protéine des feuilles drapeaux à floraison en % de la masse, PRE = précocité en degré jour depuis le semis, H = hauteur entre la base de la tige et le bout de l'épi en cm, GSV = variance intra-épi de la taille des grains en (mm²)². Les étoiles indiquent si l'estimation est significativement différente de 0. * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001.

Contrairement aux résultats de l'expérience de sélection *in vivo*, la sélection *in silico* de gros grains a permis de réaliser un progrès significatif sur la taille des grains (R_grain = +0.64 mm², R_epi = +0.68 mm²). Et sur le PMG (R_grain = +2.2 g, R_epi = +2.3 g). La variance de la taille des grains au sein d'un épi (GSV) a aussi augmenté avec la sélection (R_grain = +0.15 (mm²)², R_epi = +0.14 (mm²)²). La hauteur des plantes a augmenté avec la sélection (R_grain = +2.8 cm, R_epi = +3.2 cm) et il semble donc qu'une sélection sur gros grains peut faire augmenter indirectement la taille de la population. Ces

valeurs varient en fonction des combinaisons NEO x nsel testée et dans certains cas, aucun progrès n'a été réalisé à cause d'intensités de sélection trop faibles (par exemple lorsque le nombre d'épi observé est proche du nombre d'épi sélectionné, l'intensité de sélection sur épi est très faible, résultat non montré).

Pour les autres traits, aucun changement indirect significatif n'a été estimé quelles que soient les valeurs de NEO et nsel, ce qui est cohérent avec les valeurs d'héritabilité plus faibles de ces traits. La sélection sur épis perd en significativité lorsque NEO baisse, et les deux méthodes de sélections sont de plus en plus significatives à mesure que nsel est faible, ce qui est logique (résultats non montrés).

3.2 Comparaison de la sélection sur grain et sur épi

3.2.1 Rapport théorique entre l'efficacité de la sélection sur grain et l'efficacité de la sélection sur épi (ERS)

Les calculs (disponibles en annexe 2) ont abouti à l'équation suivante (eq. 1) :

$$\frac{R_{grain}}{R_{epi}} = \frac{NGO}{NGE \times NEO} \times \exp \left[\frac{1}{2} \left(\varphi^{-1} \left(0,1 ; 1 - \frac{nsel}{NGE \times NEO} \right)^2 - \varphi^{-1} \left(0,1 ; 1 - \frac{nsel}{NGO} \right)^2 \right) \right] \times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$$

Avec : R_{epi} : le progrès effectué avec la sélection sur épi, R_{grain} : le progrès effectué avec la sélection sur grain, NGE : Nombre de grains par épi, NEO : Nombre d'épis observés, NGO : Nombre de grains observés, $nsel$: Nombre de grains sélectionnés, V_g : variance génétique, V_e : variance environnementale, V_{inter} : variance inter épi, V_{intra} : variance intra épi, $\varphi^{-1}(0,1 ; x)$: complémentaire de la fonction de répartition de la loi normale centrée réduite en x. C'est en fait la fonction de quantile qui associe un seuil de troncation à la proportion d'individus sélectionnés. Cette fonction est disponible sous R et il n'existe pas de forme analytique simple.

Il est important de voir que l'eq. 1 est une multiplication de deux termes qui varient de manière opposée : le rapport des intensités de sélection appliquées sur grain et sur épi (premier terme à gauche de la multiplication) multiplié par la racine carrée du rapport des héritabilités sur grain et sur épi (à droite de la multiplication). Comme de façon générale $\frac{i_{grain}}{i_{epi}} > 1$ et que $\frac{H_{grain}^2}{H_{epi}^2} < 1$ on retrouve bien le compromis entre l'intensité de sélection et la précision de l'estimation de la valeur génétique des grains.

3.2.2 Validation du développement analytique

L'alignement entre prédiction théorique et valeurs expérimentales (figure 6) est très bon ($R^2=0.89$) et apparait non biaisé (non vérifié). L'eq. 1 peut donc donner des résultats réalistes, et nous allons l'utiliser pour explorer des scénarii de sélection au-delà de ceux présentés dans la partie expérimentale dans laquelle l'intensité de sélection était faible.

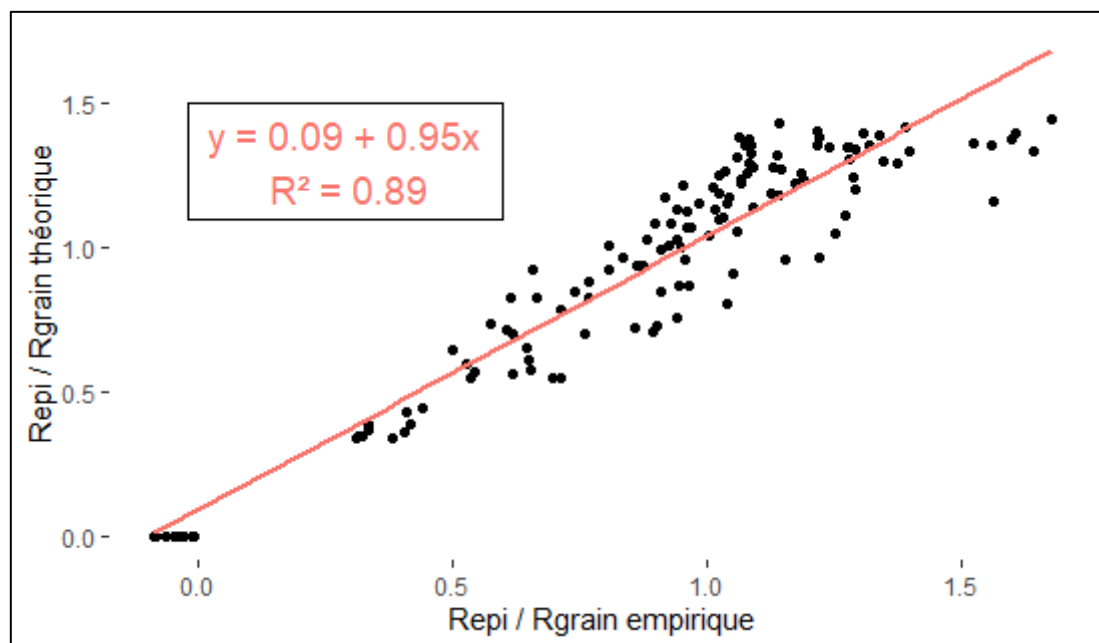


Figure 6 : Validation du développement analytique grâce aux données du dispositif de sélection *in silico*.

Les valeurs d'ERS ont été estimées par sélection *in silico*, puis calculés avec l'eq. 1 pour 132 scénarii de sélection. Chaque point correspond à une valeur calculée théoriquement mise en rapport avec la moyenne des 100 valeurs estimées par sélection *in silico*. R^2 est le carré du coefficient de corrélation de Pearson entre les deux séries de valeurs. L'ordonnée à l'origine vaut 0.09 mais n'est pas significativement différente de 0. La pente vaut 0.95 et est significativement différente de 0.

3.2.3 Scénarii de sélection

L'ERS a été calculée pour la taille du grain, et tous les résultats présentés dans cette partie ne sont valables que pour ce trait. Pour représenter différentes situations paraissant possibles en sélection, la valeur du rapport $R_{\text{grain}}/R_{\text{epi}}$ a été calculée et représenté pour différentes combinaisons de paramètres qui paraissent réalistes (figure 7) : nsel allant de 10.000 à 100.000, NGO allant de 10.000 à 5.000.000, NEO allant de 500 à 5.000, et NGE allant de 40 à 80.

Les estimations expérimentales des composantes de la variance de la taille des grains ont été utilisées pour les calculs ($V_g = 1.68$, $V_{\text{inter}} = 2.02$, $V_e = 1.51$, $V_{\text{intra}} = 7.17$). La variance intra-épi est clairement celle qui contribue le plus à la variance de la taille des grains. Cela correspond à une héritabilité de la

taille des grains individuels de 0.14 (ce qui correspond avec l'héritabilité calculée dans la partie 3.1.2), et une héritabilité de la moyenne de la taille des grains par épi autour de 0.32 pour NGE = 60 (différente de l'héritabilité calculée dans la partie 3.1.2 car on choisit NGE).

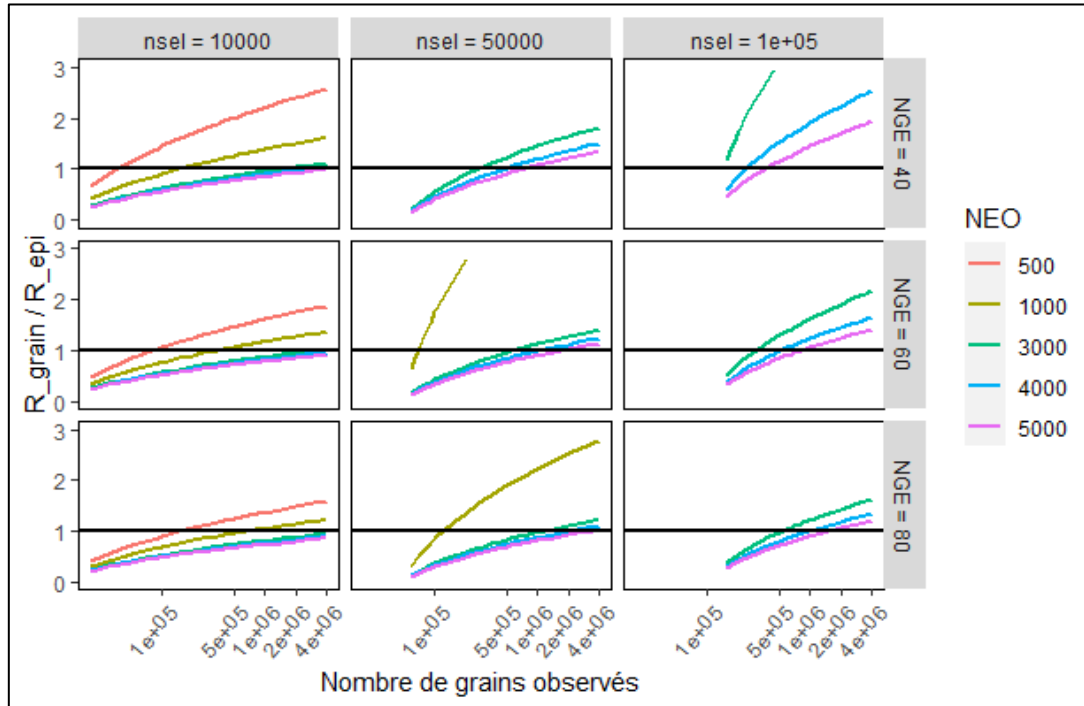


Figure 7 : ERS calculé avec l'eq. 1 pour différents scenarii de sélection.

Les combinaisons de paramètres testées sont : nsel allant de 10.000 à 100.000, NGE allant de 40 à 80, NEO allant de 10.000 à 5.000.000, NEO allant de 500 à 5.000. Les composantes de la variance utilisées sont celles estimées dans le dispositif de sélection in silico. Chaque graph représente le rapport $R_{\text{grain}}/R_{\text{epi}}$ en fonction de NEO pour une combinaison de nsel x NGE. Chaque courbe représentée correspond à une valeur de NEO. Les axes horizontaux sont en échelle logarithmique. Sur chaque graph est représenté par une ligne noire la courbe $R_{\text{grain}}/R_{\text{epi}}=1$.

On peut en premier lieu vérifier que les résultats correspondent bien aux attendus théoriques. Plus le nombre de grain observé est grand, plus la sélection sur grain dépasse celle sur épis car d'une part, l'héritabilité individuelle d'un grain n'est pas négligeable ($H^2 = 0.14$) et d'autre part l'intensité de sélection sur grain s'approche des valeurs les plus extrêmes que l'on peut obtenir. On voit que plus NGE est grand, plus le rapport baisse. C'est logique car un grand nombre de grain par épi augmente l'héritabilité sur épi ce qui fait augmenter R_{epi} si la variance intra épi est importante (ce qui est le cas avec les estimations expérimentales). De plus, comme moins d'épis sont nécessaire, grâce à un taux de multiplication plus important, il est possible d'augmenter la pression de sélection sur les épis et en retenir un plus faible pourcentage. De même, plus NEO est grand, plus le rapport est favorable à la sélection sur épi, car on peut observer plus d'épis et aller plus loin dans l'intensité de sélection sur épi, donc R_{epi} augmente.

Pour bien utiliser ces résultats, il faut comprendre et imaginer à quelle situation réelle peut correspondre chaque jeu de paramètres. Prenons par exemple la colonne $n_{sel} = 10000$. Ce nombre de grain sélectionné correspond à une surface à replanter, de l'ordre de 30 à 40 m² pour une densité autour de 300 grains/m². On peut imaginer une situation où un agriculteur améliore sa/ses populations sur une surface constante d'année en année. Il récolte chaque année les grains issus de ces 40 m², sélectionne dans sa récolte et replante 40 m² avec les grains sélectionnés. Les grains non sélectionnés constituent la population améliorée pour l'année de la récolte et peuvent être semés ou multipliés avant le semis. Dans ce cas, le nombre de grains observé serait de l'ordre de 600.000 grains pour un NGE de 60. Avec ces paramètres l'équation 1 montre que si l'agriculteur ne peut pas observer plus de 1000 épis, alors il aura plutôt intérêt à sélectionner sur grain individuel. A l'inverse, s'il est possible d'observer plus de 1000 épis alors la sélection sur épi permettra un meilleur progrès.

En se plaçant dans une situation où un agriculteur observe plusieurs millions de grains et en sélectionne 10.000 on se retrouve dans un cas où il faudrait observer au minimum 3000 épis pour que la sélection sur épi soit avantageuse. Dans ce cas, augmenter grandement le nombre d'épis observés ne rend pas la sélection sur épi beaucoup plus avantageuse (les courbes $NEO = 3000$ et $NEO = 5000$ sont proches). De plus, le rapport entre le progrès sur épi et le progrès sur grain est proche de 1 dans ce type de situation. Etant donné la complexité de la sélection sur épi il est sûrement plus judicieux de quand même faire de la sélection sur grain ce qui ferait gagner du temps et réduirait probablement les coûts.

Dès lors que l'on se place dans des situations où plus de 50.000 grains doivent être sélectionnés (ce qui correspond par exemple à sélectionner sur une surface constante de 200 m²), on voit que la sélection sur épi ne reste avantageuse que si le nombre d'épi observable est grand (> 3000) et que le nombre de grain observable est relativement faible (< 800.000). Se retrouver dans une telle situation ne paraît pas très réaliste car observer 800.000 grains pour leurs tailles est déjà réalisable aujourd'hui (avec des tamis) alors qu'observer 3000 épis est un gros travail. De plus, la comparaison entre sélection sur épi et sélection sur grain perd de son sens en imaginant une situation où un agriculteur sélectionnerait directement dans sa récolte pour faire le semis de l'année suivante. En effet, en considérant par exemple une surface cultivée de 20 ha, il faudrait observer des milliards de grains et des millions d'épis pour pouvoir sélectionner assez de grains pour replanter la surface. Une telle sélection sur épi paraît infaisable, et la question se pose également pour la sélection sur grain.

Il semble donc d'après ces résultats que la sélection sur épi ne serait avantageuse que lorsque de très petites surfaces sont allouées à la sélection. En effet, cela implique qu'il est possible d'observer une grande partie des épis récoltés ce qui permet d'aller vers de fortes intensités en sélection sur épi sans que beaucoup de grains individuels soient observables. Des représentations graphiques sans échelle

logarithmique et en trois dimensions sont disponibles en annexe 3.

4 Discussion

4.1 Retour sur les résultats de sélection

Mis à part pour la hauteur des plantes, les résultats de la sélection *in vivo* et de la sélection *in silico* ne concordent pas complètement. Il était attendu que la sélection sur grain déplace la hauteur de la population dans la sélection *in vivo* car les grains semés ont été produits dans un contexte de compétition entre les plantes où les grandes plantes ont accès à plus de ressource et font de plus gros grains. Les grains plantés dans le dispositif de sélection *in silico* n'ont pas été produits dans de telles conditions. Il y a donc bien eu une sélection indirecte sur la hauteur des plants lors de la sélection sur la taille des grains. Nous n'avons pas étudié si cela est dû à une corrélation est génétique ou si c'est lié à une meilleure implantation des gros grains : Donald, (1981) (cité dans Sadras, 2007)) a montré qu'à génotype constant, un semis constitué de gros grains avait un meilleur rendement qu'un semis constitué de petits grains, et un effet similaire est peut-être possible sur la hauteur des plants.

Seule la sélection des petits grains a permis de modifier la valeur du caractère en sélection *in vivo* tandis que la sélection de gros grain a fonctionné en sélection *in silico*. On pourrait attribuer cette différence au fait que l'expérimentation de sélection *in vivo* n'était pas d'une grande qualité. Des effets importants liés aux micro-parcelles étaient présents et ce dispositif ne permettait pas une grande puissance statistique. On peut aussi penser que l'intensité de sélection sur grain en sélection *in vivo* n'ait en fait pas été assez forte (la différentielle de sélection était de l'ordre d'un écart-type). Il serait néanmoins intéressant de continuer cette expérimentation et de continuer à sélectionner à partir des populations fondées pour avoir des résultats de meilleure qualité. De plus, en sélection *in silico* nous avons fait l'hypothèse que les plantes se comportaient de la même manière dans une population non sélectionnée (avec toute la diversité génétique) et dans une population sélectionnée (avec moins de diversité). Or, les blés se comportent différemment selon le niveau de diversité de la population dans laquelle ils poussent (Li et al., 2023), et l'effet de la diversité semble être plus marqué en conditions stressantes (Döring et al., 2010).

Les résultats de sélection *in silico* concordent plus avec les résultats trouvés dans la littérature que les résultats de sélection *in vivo*. La littérature sur la sélection massale sur la taille des grains n'est pas très étendue. Les résultats des études que nous avons pu trouver sont résumés dans le tableau 3.

Population	Type de sélection	Nb cycles	PMG	Rdt	Prot	Preco	Grains/epi	Hauteur	Référence
3 pop biparentales	Tamissage grain	2	+2.4%	+30% /cycle					(Derera and Bhatt, 1972)
90 pop biparentales	Visuelle sur épi	3	+		-	-	+	+	(Rivière et al., 2015)
3 pop biparentales	Visuelle sur épi	2	+	+					(Nass, 1987)
	Tamissage grain	2	+	+					(Lang et al., 1989)
6 pop composites	Tamissage grain	3	0						(Blum et al., 1991)
1 pop à 10 parents	Récurrente sur poids du grain	8	+4.5 % /cycle	0		+0.5 jour/cycle	-0.5/cycle	0	(Wiersma et al., 2001)
1 pop à 10 parents	Récurrente sur poids du grain	4	+7% /cycle	0	+1% / cycle	+2% /cycle	0	0	(Busch and Kofoed, 1982)
1 pop à 10 parents	Tamissage grain	3	+	+			0	+	(Sharma et al., 1995)

Tableau 3 : Résultats d'études sur la sélection massale pour la taille du grain dans des populations de blé.

Population = type de population sur laquelle la sélection a été menée, Nb cycles = nombres de cycles de sélection effectués, Rdt = rendement, Prot = taux de protéines des grains, PMG = poids de mille grains, preco = précocité, Grains/epi = nombre de grains par épi. Un "+" indique une augmentation de la valeur du trait après sélection sur la taille du grain, un "-" indique une diminution, et un "0" indique l'absence de changement significatif. Pour la précocité "+" = plus précoce et "-" = plus tardif. Lorsque les changements étaient quantifiés en pourcentage, les chiffres apparaissent après le "+" ou le "-". Une case grisée indique l'absence d'information.

Pour le PMG, et la hauteur des plants, les études montrent des résultats similaires aux nôtres : la sélection sur grain a eu un effet positif sur le PMG et il semble que la hauteur des plantes puisse augmenter avec une sélection de gros grains. Le rendement semble aussi augmenter après une sélection de gros grains. Dans nos résultats, la mesure la plus proche d'un rendement (le poids total d'épi par plante) n'a pas été affecté par la sélection. Les effets ne sont pas clairs sur la précocité, le nombre de grains par épi ou le taux de protéines (les traits ne sont pas modifiés dans le même sens selon les études), qui n'ont pas été affectés par la sélection dans nos résultats. De plus, les études menées par Wiersma et al. (2001) et Busch and Kofoed (1982) portent sur la sélection du PMG et non de la taille du grain, mais nous les avons prises en compte car ce sont des traits corrélés (Béral, 2020). Ces études semblent montrer que la sélection est plus efficace pendant les premiers cycles de sélection. Wiersma et al. (2001) précisent également que la variance génétique pour le PMG a été fortement réduite par la sélection.

De plus, nos résultats montrent que la sélection de gros grains pourrait faire augmenter la GSV. Aucune étude précédente à notre connaissance n'a mesuré l'effet de la sélection sur ce trait. Les résultats obtenus pour ces traits ne peuvent donc pas être comparés à de précédentes expériences de sélection. On peut citer les travaux de Béral, (2020) qui a montré que les différences génotypiques de masses de grains individuels existaient uniquement pour les grains proximaux (les grains issus des fleurs les plus proches

de la tige qui donnent les grains à l'extérieur de l'épillet). Il en découle mécaniquement que les épis portant des grains individuels de taille importante ont également une plus grande variance de la taille des grains. En effet les grains distaux (issus des fleurs éloignées de la tige) d'épis portant de gros grains sont aussi petits les grains distaux d'épis ne portant pas de gros grains. Il a aussi été montré que la GSV avait un déterminisme génétique spécifique et que ce trait pouvait avoir une héritabilité de l'ordre de 0.52 à 0.85 sur le blé tendre (Béral et al., 2020) ce qui est plus élevé que ce que nous avons trouvé dans notre dispositif ($H^2=0.13$). Tout cela est donc cohérent avec une augmentation de la GSV lors de la sélection de gros grains.

Cependant, cette augmentation de la GSV peut être désavantageuse. Premièrement du point de vue de la transformation des grains car cela montre que sélectionner pour des grains de bonne qualité ne permet que d'améliorer la qualité d'une partie des grains (même s'il est *a priori* possible de sélectionner pour une faible GSV indépendamment du reste du fait du déterminisme génétique spécifique de ce trait, ce qui est possible en sélection sur épi). Deuxièmement, Béral, (2020) a aussi montré que plus la GSV est élevée, plus la perte de PMG liée à un stress thermique post-floraison sera importante. Il serait intéressant de voir si une augmentation de la GSV liée à une sélection sur la taille du grain rend le blé plus sensible à ce stress, ce qui serait contre-productif dans un contexte de sélection pour l'adaptation aux changements climatiques.

Enfin, on peut pointer un défaut dans nos résultats : l'absence de mesure directe de rendement. Ce trait n'a été mesuré qu'indirectement par le poids total d'épi dans les bacs. Il manque donc dans l'analyse une information importante. Même si, dans le cadre de sélection massale ou de sélection sur la morphologie du grain, les objectifs de sélection peuvent être différents de ceux travaillés en sélection conventionnelle et ne pas prioriser le rendement, il reste important de savoir comment il est impacté par cette sélection.

4.2 Recul sur l'approche théorique

Dans le cadre du stage, l'eq. 1 a seulement été utilisée pour la taille du grain car c'est le seul trait pour lequel nous avons pu estimer les composantes de la variance nécessaires pour les calculs. Cependant, cette équation peut en théorie être utilisée pour n'importe quel trait qui peut être mesuré à l'échelle du grain individuel, et pour n'importe quelle céréale autogame. De plus, elle pourrait être améliorée. Une des améliorations les plus évidentes est la pondération des progrès par le coût et le temps nécessaire pour chaque méthode. Cela permettrait surtout de prendre en compte la faisabilité d'une sélection sur épi qui n'est aujourd'hui pas possible à grande échelle d'après nos connaissances. Même si dans certaines situations la sélection sur épi devient meilleure à grande échelle que la sélection sur grain, il

se peut que les moyens à investir pour faire ce progrès ne puissent être envisagés que dans le cas d'un groupe d'agriculteur ou d'un réseau animé par une association de semences de fermes et qu'une sélection sur grain soit au final plus efficace relativement à l'investissement nécessaire. De tels investissements pourraient toutefois intéresser les sélectionneurs dans les premières générations de fixation sur bulk ou d'introggression en population de ressources génétiques. En prenant en compte ces aspects plus pratiques, l'équation pourrait donc être utilisée pour aiguiller des agriculteurs ou sélectionneurs souhaitant faire une sélection massale sur les traits du grain.

La validation de l'éq. 1 a été faite sur des données peu réalistes. En effet, les données du dispositif de sélection *in silico* ont contraints le nombre de grains observés à 859 (nombre de plantes du premier semis ayant été récoltées) et le nombre de grain par épi à 5 car parmi les 859 grains utilisés, en moyenne 5 grains de chaque lot étaient présents. Cela correspond donc bien à une sélection sur épi où les épis n'auraient donné que 5 grains. De plus, les variances estimées ne correspondent peut-être pas à des variances en conditions réelles au champ. On ne sait donc pas si l'éq. 1 a le même pouvoir prédictif pour des valeurs de paramètres estimées au champ et il faudrait la valider totalement en sélectionnant réellement sur grain et sur épi au champ. Une telle expérience serait assez lourde à mener étant donné l'effort à faire pour sélectionner sur épi et pour récolter les données permettant d'estimer les paramètres de la variance phénotypique (V_g , V_{inter} , V_{intra} V_e).

Un autre élément important de discussion est justement la nécessité d'estimer les paramètres de la variance pour utiliser l'éq. 1. Etant donné que ces paramètres sont contraints (on ne peut pas les adapter ou les choisir comme NEO, NGO ou nsel), il est nécessaire de les estimer au préalable. L'estimation de ces paramètres requiert un travail conséquent et/ou l'utilisation de technologie avancée (Optomachine ou NIRS) du fait de la nécessité de mesurer des grains individuellement. Dans le cadre d'une sélection massale menée par un groupement d'agriculteur, il est raisonnable de penser que ni le temps ni la technologie ne sont disponibles actuellement. De plus, comme ces paramètres changent en fonction du lieu, de l'année et du trait observé, il faudrait sûrement ré-estimer régulièrement les paramètres de variance pour se servir de l'équation.

L'exploration des domaines de supériorité de la sélection sur grain et sur épi s'est faite en utilisant les variances estimées dans le dispositif de sélection *in silico*. Ces résultats sont donc valables pour ces valeurs de variances mais ne sont pas universels. Nous avons tout de même trouvé une étude qui a comparé la sélection sur grain et la sélection sur épi (Nass, 1987) et nos résultats corroborent sa conclusion (la sélection sur grain est plus efficace dans de grandes populations, mais la sélection sur épi peut être avantageuse avec des effectifs réduits). Il est intéressant de voir que nos résultats ont été

obtenus avec une héritabilité de 0,14 pour les mesures sur grains individuels et de l'ordre de 0.3 pour les mesures sur épi. Il n'était pas trivial que sélectionner avec une faible héritabilité serait avantageux dans un grand nombre de situations grâce à la compensation des fortes intensités de sélection.

De plus, les valeurs de variances utilisées supposent que l'on a un moyen de mesurer précisément des traits de grains individuellement et de faire la moyenne de ces mesures épis par épis. Il n'existe pas de machine ou de méthode permettant de faire cela à haut débit à notre connaissance mais cette démarche va être entreprise dans le projet Bbsocoul (blé bio sous couvert de luzerne) de mon laboratoire d'accueil. On peut donc imaginer en prenant l'exemple de la sélection sur la taille du grain que des tamis permettent une mesure précise du trait à l'échelle du grain individuel mais qu'une sélection sur épi doive se faire de façon visuelle et soit imprécise (à moyenne/grande échelle, seule une observation visuelle des épis permettrait d'en observer assez pour opérer la sélection). L'imprécision d'une sélection visuelle ferait baisser l'héritabilité des observations sur épi et la sélection sur épi serait plus désavantagée par rapport à la sélection sur grain que ce que montrent nos résultats. Toutefois, l'équation 1 pourrait quand même permettre d'explorer ce type de situation en faisant varier les valeurs de H_{epi}^2 tout en gardant une valeur de H_{grain}^2 non négligeable. Ce raisonnement ne s'applique pas nécessairement pour tous les traits. Par exemple sur le taux de protéine, qui se mesure par NIRS, il se peut que les précisions de mesure sur grain individuel ou sur lot de grain soient similaires. Des calibrations sont en cours dans mon laboratoire d'accueil pour prédire le taux de protéines à partir de spectres de grains individuels, et refaire le même type d'analyse que celles réalisées sur la taille des grains.

Enfin, cette équation fonctionne lorsqu'il s'agit de faire de la sélection par troncation. Or, ce n'est pas la seule façon de sélectionner qui existe. L'eq. 1 ne s'applique pas par exemple si l'on veut utiliser un gradient de sélection ou faire de la sélection sur index. En sélection massale particulièrement, un des enjeux de la sélection peut être de conserver de la diversité génétique dans la population. Par exemple, Gorjanc et al., (2018) ont testé par une approche de simulation, une méthode de sélection où les grains sélectionnés sont obtenus en choisissant dans la population les parents à croiser selon leurs caractéristiques et leur niveau de parenté de sorte à conserver un certain niveau de diversité génétique de génération en génération tout en réalisant un progrès. Ils ont montré qu'une telle approche pourrait permettre d'obtenir de meilleurs gains génétiques sur le long terme par rapport à des générations de sélection par troncation. L'eq. 1 n'est donc évidemment pas adaptable à toute situation de sélection massale sur les traits du grain.

4.3 Vers une sélection massale phénotypique sur grains ?

La prédiction phénotypique, un développement récent en génétique quantitative, permet de prédire certains traits d'une plante à partir de spectres infrarouges collectés sur cette plante. De premiers résultats encourageants montrent que la sélection phénotypique a une efficacité souvent comparable à l'efficacité de la prédiction génomique, et parfois même meilleure pour le rendement chez le blé (Rincent et al., 2018). C'est une méthode non destructive et peu coûteuse qui pourrait permettre de traiter les grains individuellement à haut débit. Des trieurs optiques sont déjà utilisés de façon routinière dans certaines entreprises semencières, qui possèdent donc déjà au moins une partie des outils nécessaires à l'implémentation de cette méthode. Pour ces raisons, la sélection phénotypique semble être pertinente dans un cadre de sélection massale, et certains auteurs pensent que cette méthode sera à l'avenir un élément essentiel des programmes de sélection classique (Zhu et al., 2021).

Comme l'éq. 1 peut être utilisée pour n'importe quel trait du grain, elle peut apporter des éléments de discussion intéressants pour la sélection phénotypique. En effet, un spectre (ou la prédiction que l'on fait à partir du spectre) peut être considéré comme un trait du grain, et les compromis entre héritabilité et intensité de sélection s'appliquent (l'héritabilité étant dans ce cas l'accuracy de prédiction du trait). On pourrait donc comparer le progrès atteignable en faisant des modèles prédictifs basés sur le grain individuel ou sur la moyenne de grains de même génotype/épi en et en remplaçant dans l'éq. 1 le rapport des héritabilités par le rapport des accuracy de prédiction des deux méthodes.

La plupart des travaux réalisés en prédiction phénotypique sur des céréales utilisent des mesures sur lots de grains ou sur de la farine de lots de grains (Cuevas et al., 2019; Lane et al., 2020; Rincent et al., 2018; Robert et al., 2022; Weiß et al., 2022; Zhu et al., 2022). On peut considérer le spectre ou le phénotype d'un lot de grain comme une moyenne des spectres/phénotypes individuels des grains du lot. Cette façon de calibrer les modèles prédictifs correspond donc plutôt à une sélection sur épi. Or les trieurs optiques utilisés en sélection sont capables de trier les grains individuellement. Des modèles prédictifs basés sur des moyennes de lots ne sont donc peut-être pas adaptés pour ce type de tri. Il serait donc intéressant de comparer le progrès atteignable à partir de prédiction sur grains individuels et sur moyennes de lots de grains. En effet, si l'on peut montrer qu'une prédiction à partir de grains individuels permet un meilleur progrès malgré la plus faible accuracy de prédiction, cela ouvre des perspectives pour la sélection phénotypique sur grain. On aurait accès à des mesures indirectes pour un ensemble de traits de la plante ce qui faciliterait énormément le travail de sélection. Il serait aussi intéressant de savoir si dans l'absolu (sans comparer des prédictions sur grains individuels et sur lots de grains) une prédiction sur grain individuel serait déjà suffisante pour réaliser un progrès satisfaisant.

Enfin, Gaynor et al. (2017) ont imaginé un schéma de sélection massale utilisant la prédiction génomique pouvant s'appliquer au blé. Ils ont montré par des simulations qu'un tel schéma permettrait

d'obtenir de meilleurs gains génétique sur le long terme que les schémas de sélection actuellement en place sur le blé. L'efficacité de ce schéma de sélection découle de l'utilisation de prédiction génomique pour l'amélioration de la population. On peut donc imaginer remplacer dans ce schéma la prédiction génomique par la prédiction phénomique, et la question de la prédiction sur grain ou sur lots de grains se pose alors encore. Tout cela suppose que l'on puisse prédire des traits d'une plante à partir du spectre du grain dont cette plante est issue, ce qui n'est aujourd'hui pas bien démontré.

4.4 Le potentiel du dispositif de sélection in silico

Dans le dispositif de sélection *in silico*, chaque plante a été mesurée individuellement en gardant trace de sa position dans le bac et de son génotype, puis chaque grain de chaque plante a été mesuré individuellement. Cela nécessite beaucoup d'effort pour ne pas faire d'erreur pendant les mesures, car il ne faut pas se tromper sur la position de la plante, réussir à isoler la bonne plante pour la bonne mesure etc... Cependant, ces efforts permettent d'acquérir une très grande quantité d'informations et ouvrent des perspectives pour un grand nombre d'analyses. En effet, ce dispositif permet d'accéder aux valeurs phénotypiques en population et non en culture pure. La position et le génotype de chaque plante étant connu, il est possible d'étudier en détail les interactions entre les plantes (effets de certains allèles sur les phénotypes voisins, tendance de chaque génotype à collaborer ou à rentrer en compétition, allèles à favoriser dans la population pour améliorer un trait...). Le dispositif permet également d'explorer les questions liées à la phénomique car des spectres de feuilles et des grains de chaque plante ont été acquis. On pourrait par exemple apporter des éléments de réponse aux questionnements évoqués plus haut, ou regarder s'il est possible de prédire l'effet d'un génotype sur ses voisins à partir de son spectre. Au final, les analyses réalisées dans ce stage n'ont permis d'exploiter qu'une partie de toute l'information collectée, et les données peuvent encore donner du grain à moudre pour de futures analyses.

Références bibliographiques

- Acreche, M.M., Slafer, G.A., 2006. Grain weight response to increases in number of grains in wheat in a Mediterranean area. *Field Crops Res.* 98, 52–59. <https://doi.org/10.1016/j.fcr.2005.12.005>
- Andrée, P., Clark, J.K., Levkoe, C.Z., Lowitt, K. (Eds.), 2019. *Civil Society and Social Movements in Food System Governance*. Taylor & Francis.
- Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Müller, C., Naresh Kumar, S., Nendel, C., O’Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. *Nat. Clim. Change* 5, 143–147. <https://doi.org/10.1038/nclimate2470>
- Ballini, E., Tavaud, M., Ducasse, A., Sanchez, D., Paux, E., Kitt, J., Charmet, G., Audigeos, D., Roumet, P., David, J., Morel, J.-B., 2020. Genome wide association mapping for resistance to multiple fungal pathogens in a panel issued from a broad composite cross-population of tetraploid wheat *Triticum turgidum*. *Euphytica* 216, 92. <https://doi.org/10.1007/s10681-020-02631-9>
- Béral, A., 2020. Déterminisme génétique et écophysiologique de la variabilité des masses de grains individuels chez le blé tendre (*Triticum aestivum*). (phdthesis). Université Clermont Auvergne [2017-2020].
- Béral, A., Rincet, R., Gouis, J.L., Girousse, C., Allard, V., 2020. Wheat individual grain-size variance originates from crop development and from specific genetic determinism. *PLOS ONE* 15, e0230689. <https://doi.org/10.1371/journal.pone.0230689>
- Blum, A., Shpiler, L., Golan, G., Mayer, J., Sinmena, B., 1991. Mass selection of wheat for grain filling without transient photosynthesis. *Euphytica* 54, 111–116. <https://doi.org/10.1007/BF00145637>
- Busch, R.H., Kofoed, K., 1982. Recurrent Selection for Kernel Weight in Spring Wheat. *Crop Sci.* 22. <https://doi.org/10.2135/cropsci1982.0011183X002200030032x>
- Chable, V., Berthellot, J.-F., 2006. La sélection participative en France: présentation des expériences en cours pour les agricultures biologiques et paysannes. *Courr. L’environnement L’INRA* 30, 129–138.
- Colombo, M., Roumet, P., Salon, C., Jeudy, C., Lamboeuf, M., Lafarge, S., Dumas, A.-V., Dubreuil, P., Ngo, W., Derepas, B., Beauchêne, K., Allard, V., Le Gouis, J., Rincet, R., 2022. Genetic Analysis of Platform-Phenotyped Root System Architecture of Bread and Durum Wheat in Relation to Agronomic Traits. *Front. Plant Sci.* 13.
- Cuevas, J., Montesinos-López, O., Juliana, P., Guzmán, C., Pérez-Rodríguez, P., González-Bucio, J., Burgueño, J., Montesinos-López, A., Crossa, J., 2019. Deep Kernel for Genomic and Near Infrared Predictions in Multi-environment Breeding Trials. *G3 GenesGenomesGenetics* 9, 2913–2924. <https://doi.org/10.1534/g3.119.400493>
- David, J., Holtz, Y., Ranwez, V., Santoni, S., Sarah, G., Ardisson, M., Poux, G., Choulet, F., Genthon, C., Roumet, P., Tavaud-Pirra, M., 2014. Genotyping by sequencing transcriptomes in an evolutionary pre-breeding durum wheat population. *Mol. Breed.* 34, 1531–1548. <https://doi.org/10.1007/s11032-014-0179-z>
- Dawson, J.C., Goldringer, I., 2012. Breeding for Genetically Diverse Populations: Variety Mixtures and Evolutionary Populations, in: *Organic Crop Breeding*. John Wiley & Sons, Ltd, pp. 77–98. <https://doi.org/10.1002/9781119945932.ch5>
- Dawson, J.C., Rivière, P., Berthellot, J.-F., Mercier, F., De Kochko, P., Galic, N., Pin, S., Serpolay, E., Thomas, M., Giuliano, S., others, 2011. Collaborative plant breeding for organic agricultural systems in developed countries. *Sustainability* 3, 1206–1223.
- Derera, N.F., Bhatt, G.M., 1972. Effectiveness of mechanical mass selection in wheat (*Triticum aestivum* L.). *Aust. J. Agric. Res.* 23, 761–768. <https://doi.org/10.1071/ar9720761>

- Desclaux, D., Ceccarelli, S., Navazio, J., Coley, M., Trouche, G., Aguirre, S., Weltzien, E., Lançon, J., 2012. Centralized or Decentralized Breeding: The Potentials of Participatory Approaches for Low-Input and Organic Agriculture, in: *Organic Crop Breeding*. John Wiley & Sons, Ltd, pp. 99–123. <https://doi.org/10.1002/9781119945932.ch6>
- Desclaux, D., Nolot, J.M., Chiffolleau, Y., 2009. La sélection participative pour élaborer des variétés de blé dur pour l'agriculture biologique. *Innov. Agron.* 7, 65–78.
- Donald, C., 1981. Competitive plants, communal plants, and yield in wheat crops. *Wheat Sci.-Today Tomorrow* 223.
- Döring, T., Wolfe, M., Jones, H., Pearce, H., Zhan, J., others, 2010. Breeding for resilience in wheat—Nature's choice. *Breed. Resil. Strategy Org. Low-Input Farming Syst.* 47–50.
- Finch-Savage, W.E., Bassel, G.W., 2016. Seed vigour and crop establishment: extending performance beyond adaptation. *J. Exp. Bot.* 67, 567–591. <https://doi.org/10.1093/jxb/erv490>
- Gaynor, R.C., Gorjanc, G., Bentley, A.R., Ober, E.S., Howell, P., Jackson, R., Mackay, I.J., Hickey, J.M., 2017. A Two-Part Strategy for Using Genomic Selection to Develop Inbred Lines. *Crop Sci.* 57, 2372–2386. <https://doi.org/10.2135/cropsci2016.09.0742>
- GIE Blé dur, 2017. Description et spécificités - GIE Blé dur Description, spécificités du blé dur. GIE Blé Dur. URL <https://www.gie-bledur.fr/la-filiere-ble-dur/description-du-ble-dur/> (accessed 6.7.23).
- Gorjanc, G., Gaynor, R.C., Hickey, J.M., 2018. Optimal cross selection for long-term genetic gain in two-part programs with rapid recurrent genomic selection. *Theor. Appl. Genet.* 131, 1953–1966. <https://doi.org/10.1007/s00122-018-3125-3>
- Horlings, L.G., Marsden, T.K., 2011. Towards the real green revolution? Exploring the conceptual dimensions of a new ecological modernisation of agriculture that could 'feed the world.' *Glob. Environ. Change, Special Issue on The Politics and Policy of Carbon Capture and Storage* 21, 441–452. <https://doi.org/10.1016/j.gloenvcha.2011.01.004>
- Houser, M., Stuart, D., 2020. An accelerating treadmill and an overlooked contradiction in industrial agriculture: Climate change and nitrogen fertilizer. *J. Agrar. Change* 20, 215–237. <https://doi.org/10.1111/joac.12341>
- Ishaque, W., Osman, R., Hafiza, B.S., Malghani, S., Zhao, B., Xu, M., Ata-Ul-Karim, S.T., 2023. Quantifying the impacts of climate change on wheat phenology, yield, and evapotranspiration under irrigated and rainfed conditions. *Agric. Water Manag.* 275, 108017. <https://doi.org/10.1016/j.agwat.2022.108017>
- Johannsen, W., 1903. Heredity in populations and pure lines. *Class. Pap. Genet.* 20–26.
- Juroszek, P., von Tiedemann, A., 2013. Climate change and potential future risks through wheat diseases: a review. *Eur. J. Plant Pathol.* 136, 21–33. <https://doi.org/10.1007/s10658-012-0144-9>
- Lane, H.M., Murray, S.C., Montesinos-López, O.A., Montesinos-López, A., Crossa, J., Rooney, D.K., Barrero-Farfan, I.D., De La Fuente, G.N., Morgan, C.L.S., 2020. Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *Plant Phenome J.* 3, e20002. <https://doi.org/10.1002/ppj2.20002>
- Lang, L., Balla, L., Bedo, Z., 1989. Machine mass selection of winter wheat hybrid populations. *Novenytermeles Hung.*
- Li, L., Giller, P., Peng, P., Guo, Y., Zhang, S., Hu, Z., Wang, X., Zhao, G., 2023. The genetic diversity-productivity effect in wheat cultivar mixtures at multiple levels. *Eur. J. Agron.* 142, 126676. <https://doi.org/10.1016/j.eja.2022.126676>
- Lush, J.L., 1943. Animal breeding plans. *Anim. Breed. Plans.*
- Mailhe, G., Cazeirgue, F., Gascuel, J., Gasnier, R., Berthelot, J., Baboulène, J., Poilly, C., Lavoyer, R., Hernandez, M., Coulbeaut, J., others, 2013. Mise en place d'une méthodologie de sélection participative sur le blé tendre en France.
- Montazeaud, G., Flutre, T., Ballini, E., Morel, J.-B., David, J., Girodolle, J., Rocher, A., Ducasse, A., Violle, C., Fort, F., Fréville, H., 2022. From cultivar mixtures to allelic mixtures: opposite effects of allelic richness between genotypes and genotype richness in wheat. *New Phytol.* 233, 2573–2584. <https://doi.org/10.1111/nph.17915>

- Nass, H.G., 1987. Selection for grain yield of spring wheat utilizing seed size and other selection criteria. *Can. J. Plant Sci.* 67, 605–610. <https://doi.org/10.4141/cjps87-086>
- Phillips, S., Wolfe, M.S., 2005. Evolutionary plant breeding for low input systems. *J. Agric. Sci.* 143, 245–254. <https://doi.org/10.1017/S0021859605005009>
- Pingali, P.L., 2012. Green Revolution: Impacts, limits, and the path ahead. *Proc. Natl. Acad. Sci.* 109, 12302–12308. <https://doi.org/10.1073/pnas.0912953109>
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rincent, R., Charpentier, J.-P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., Segura, V., 2018. Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. *G3 GenesGenomesGenetics* 8, 3961–3972. <https://doi.org/10.1534/g3.118.200760>
- Rivière, P., Goldringer, I., Berthelot, J.-F., Galic, N., Pin, S., Kochko, P.D., Dawson, J.C., 2015. Response to farmer mass selection in early generation progeny of bread wheat landrace crosses. *Renew. Agric. Food Syst.* 30, 190–201. <https://doi.org/10.1017/S1742170513000343>
- Robert, P., Auzanneau, J., Goudemand, E., Oury, F.-X., Rolland, B., Heumez, E., Bouchet, S., Le Gouis, J., Rincent, R., 2022. Phenomic selection in wheat breeding: identification and optimisation of factors influencing prediction accuracy and comparison to genomic selection. *Theor. Appl. Genet.* 135, 895–914. <https://doi.org/10.1007/s00122-021-04005-8>
- Sadras, V.O., 2007. Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Res.* 100, 125–138. <https://doi.org/10.1016/j.fcr.2006.07.004>
- Sharma, R.C., Tiwary, A.K., Ortiz-Ferrara, G., 2008. Reduction in kernel weight as a potential indirect selection criterion for wheat grain yield under terminal heat stress. *Plant Breed.* 127, 241–248. <https://doi.org/10.1111/j.1439-0523.2007.01460.x>
- Sharma, S.K., Singh, K.P., Singh, I., 1995. SELECTION RESPONSES FOR GRAIN WEIGHT IN SOME MASS SELECTED AND INTERMATED POPULATIONS OF WHEAT (TRITICUM AESTIVUM L.). *INDIAN J. Genet. PLANT Breed.* 55, 365–373.
- Suneson, C.A., 1956. An Evolutionary Plant Breeding Method. *Agron. J.* 48, 188–191. <https://doi.org/10.2134/agronj1956.00021962004800040012x>
- Truncated normal distribution [WWW Document], 2023. . Wikipedia. URL https://en.wikipedia.org/w/index.php?title=Truncated_normal_distribution&oldid=1152823526#cite_note-5 (accessed 6.9.23).
- Wang, J., Vanga, S.K., Saxena, R., Orsat, V., Raghavan, V., 2018. Effect of Climate Change on the Yield of Cereal Crops: A Review. *Climate* 6, 41. <https://doi.org/10.3390/cli6020041>
- Wang, K., Fu, B.X., 2020. Inter-Relationships between Test Weight, Thousand Kernel Weight, Kernel Size Distribution and Their Effects on Durum Wheat Milling, Semolina Composition and Pasta Processing Quality. *Foods* 9, 1308. <https://doi.org/10.3390/foods9091308>
- Weiß, T.M., Zhu, X., Leiser, W.L., Li, D., Liu, W., Schipprack, W., Melchinger, A.E., Hahn, V., Würschum, T., 2022. Unraveling the potential of phenomic selection within and among diverse breeding material of maize (*Zea mays* L.). *G3* 12, jkab445.
- Wiersma, J.J., Busch, R.H., Fulcher, G.G., Hareland, G.A., 2001. Recurrent Selection for Kernel Weight in Spring Wheat. *Crop Sci.* 41, 999–1005. <https://doi.org/10.2135/cropsci2001.414999x>
- Wilcox, J., Makowski, D., 2014. A meta-analysis of the predicted effects of climate change on wheat yields using simulation studies. *Field Crops Res.* 156, 180–190. <https://doi.org/10.1016/j.fcr.2013.11.008>
- Wolfram Research inc, n.d. Mathematica, Version 13.3.
- Zahra, N., Hafeez, M.B., Wahid, A., Al Masruri, M.H., Ullah, A., Siddique, K.H.M., Farooq, M., 2023. Impact of climate change on wheat grain composition and quality. *J. Sci. Food Agric.* 103, 2745–2751. <https://doi.org/10.1002/jsfa.12289>
- Zhu, X., Leiser, W.L., Hahn, V., Würschum, T., 2021. Phenomic selection is competitive with genomic selection for breeding of complex traits. *Plant Phenome J.* 4, e20027. <https://doi.org/10.1002/ppj2.20027>
- Zhu, X., Maurer, H.P., Jenz, M., Hahn, V., Ruckelshausen, A., Leiser, W.L., Würschum, T., 2022. The

performance of phenomic selection depends on the genetic architecture of the target trait. *Theor. Appl. Genet.* 135, 653–665. <https://doi.org/10.1007/s00122-021-03997-7>

Annexe 1 : Variance et héritabilité des traits du grain en mesurant les grains individuels ou la moyenne des grains par épi

Pour un trait du grain, la variance observée entre grains provient de :

- La variance génétique qui existe entre génotypes différents (toute chose égale par ailleurs)
- La variance environnementale qui existe entre grains issus de plantes différentes d'un même génotype ayant poussé dans des conditions différentes (comme des parcelles différentes dans le même lieu ou des lieux différents la même année, etc.)
- La variance inter-épi qui existe entre épis d'une même plante pour des raisons physiologiques, (par exemple entre brin maître et brin secondaire) (Acreche and Slafer, 2006)
- La variance intra-épi qui existe entre des grains d'un même épi de la même plante pour des raisons physiologiques également (par exemple grains des fleurs proximales vs fleurs distales (Béral, 2020))

Ces considération amène à modéliser la valeur phénotypique observée sur un grain individuel de la manière suivante :

$$Y_{ijkl} = \mu + G_i + Env_j + Epi_{ij(k)} + \epsilon_{ijkl}$$

Où :

- Y_{ijkl} est le phénotype d'un grain individuel
- $G_i \sim N(0, \sigma_g)$ l'effet du génotype i
- $Env_j \sim N(0, \sigma_e)$ l'effet de l'environnement j
- $Epi_{ij(k)} \sim N(0, \sigma_{inter})$ l'effet de l'épi k au sein du génotype i dans l'environnement j
- $\epsilon_{ijkl} \sim N(0, \sigma_{intra})$ un effet résiduel non expliquée par les autres effets, et qui correspond à la variance intra-épi.

Donc la variance du trait à l'échelle du grain vaut :

$$\begin{aligned} V(Y_{ijkl}) = & V(\mu) + V(G_i) + V(Env_j) + V(Epi_{ij(k)}) + V(\epsilon_{ijkl}) + 2cov(\mu, G_i) + 2cov(\mu, Env_j) \\ & + 2cov(\mu, Epi_{ij(k)}) + 2cov(\mu, \epsilon_{ijkl}) + 2cov(G_i, Env_j) + 2cov(G_i, Epi_{ij(k)}) \\ & + 2cov(G_i, \epsilon_{ijkl}) + 2cov(Env_j, Epi_{ij(k)}) + 2cov(Env_j, \epsilon_{ijkl}) \end{aligned}$$

On suppose que les effets sont indépendants et donc que toutes les covariances sont nulles, et on sait que $V(\mu) = 0$. On a donc :

$$V(Y_{ijkl}) = V(G_i) + V(Env_j) + V(Epi_{ij(k)}) + V(\epsilon_{ijkl})$$

Et donc :

$$V_{grain} = V_g + V_{inter} + V_e + V_{intra}$$

La moyenne des grains d'un épi est :

$$Y_{ijk\cdot} = \frac{1}{NGE} \sum_{l=1}^{NGE} Y_{ijkl}$$

Avec NGE le nombre de grains par épi.

$$V(Y_{ijk\cdot}) = V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} Y_{ijkl}\right)$$

$$V(Y_{ijk\cdot}) = V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} [\mu + G_i + Env_j + Epi_{ij(k)} + \epsilon_{ijkl}]\right)$$

$$V(Y_{ijk\cdot}) = V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \mu + \frac{1}{NGE} \sum_{l=1}^{NGE} G_i + \frac{1}{NGE} \sum_{l=1}^{NGE} Env_j + \frac{1}{NGE} \sum_{l=1}^{NGE} Epi_{ij(k)} + \frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right)$$

Les variances concernant μ sont toutes nulles. Toujours en supposant que les effets sont indépendants et donc que les covariances sont nulles :

$$V(Y_{ijk\cdot}) = V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} G_i\right) + V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} Env_j\right) + V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} Epi_{ij(k)}\right) + V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right)$$

Comme G , n'est pas indexés sur l , alors :

$$\sum_{l=1}^{NGE} G_i = NGE \times G_i$$

$$\frac{1}{NGE} \sum_{l=1}^{NGE} G_i = G_i$$

$$V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} G_i\right) = V(G_i)$$

Il en est de même pour Env et Epi. En revanche, ε est indexé sur l donc :

$$\begin{aligned}
 V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right) &= \frac{1}{NGE^2} V\left(\sum_{l=1}^{NGE} \epsilon_{ijkl}\right) \\
 V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right) &= \frac{1}{NGE^2} \sum_{l=1}^{NGE} V(\epsilon_{ijkl}) \\
 V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right) &= \frac{1}{NGE^2} \times NGE \times V(\epsilon_{ijkl}) \\
 V\left(\frac{1}{NGE} \sum_{l=1}^{NGE} \epsilon_{ijkl}\right) &= \frac{1}{NGE} \times V(\epsilon_{ijkl})
 \end{aligned}$$

On a donc :

$$V(Y_{ijk.}) = V(G_i) + V(Env_j) + V(Epi_{ij(k)}) + \frac{V(\epsilon_{ijkl})}{NGE}$$

$V_{epi} = V_g + V_{gxe} + V_{inter} + \frac{V_{intra}}{NGE}$

Nécessairement $V_{grain} > V_{epi}$ car $NGE > 1$.

L'héritabilité d'un trait étant le rapport entre sa variance génétique et sa variance phénotypique, il en découle que :

$$H_{grain}^2 = \frac{V_g}{V_g + V_{gxe} + V_{inter} + V_{intra}} \quad \text{et} \quad H_{epi}^2 = \frac{V_g}{V_g + V_{gxe} + V_{inter} + \frac{V_{intra}}{NGE}}$$

Donc nécessairement : $H_{grain}^2 < H_{epi}^2$

Cette modélisation reflète bien un des attendus exposé en introduction.

Annexe 2 : Calculs complets pour obtenir l'équation 1

On cherche à calculer le rapport $\frac{R_{grain}}{R_{epi}}$ où R_{grain} est le progrès effectué en sélectionnant sur un trait du grain individuel et R_{epi} est le progrès effectué en sélectionnant sur un trait moyen des grains d'un épi.

On cherche à calculer ce rapport en fonction des paramètres suivants :

NGE : Nombre de grains par épi

NEO : Nombre d'épis observés

NGO : Nombre de grains observés

n_{sel} : Nombre de grains sélectionnés

V_g : variance génétique du trait

V_e : variance environnementale liée à l'hétérogénéité de la parcelle

V_{inter} : variance inter épi au sein d'une plante liée aux différences physiologiques entre les thalles

V_{intra} : variance intra épi

On pose donc :

$$\frac{R_{grain}}{R_{epi}} = \frac{H_{grain}^2 \cdot i_{grain} \cdot \sigma_{grain}}{H_{epi}^2 \cdot i_{epi} \cdot \sigma_{epi}}$$

Avec :

H^2 l'héritabilité sur grain ou sur épi, i l'intensité de sélection en unité d'écart type sur grain ou sur épi, et σ l'écart type de la distribution du trait des grains individuels (pour la sélection sur grain) ou de la distribution du trait moyen des grains par épi (pour la sélection sur épi).

On sait que

$$\sigma_{grain} = \sqrt{V_g + V_e + V_{inter} + V_{intra}}$$

$$\sigma_{epi} = \sqrt{V_g + V_e + V_{inter} + V_{intra}/NGE}$$

$$H_{grain}^2 = \frac{V_g}{V_g + V_e + V_{inter} + V_{intra}}$$

$$H_{epi}^2 = \frac{V_g}{V_g + V_e + V_{inter} + V_{intra}/NGE}$$

On obtient donc :

$$\frac{R_{grain}}{R_{epi}} = \frac{i_{grain}}{i_{epi}} \times \frac{\frac{V_g}{V_g + V_e + V_{inter} + V_{intra}} \times \sqrt{V_g + V_e + V_{inter} + V_{intra}}}{\frac{V_g}{V_g + V_e + V_{inter} + V_{intra}/NGE} \times \sqrt{V_g + V_e + V_{inter} + V_{intra}/NGE}}$$

Les variances génétiques se simplifient et $\frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}}$ donc :

$$\frac{R_{grain}}{R_{epi}} = \frac{i_{grain}}{i_{epi}} \times \frac{1/\sqrt{V_g + V_e + V_{inter} + V_{intra}}}{1/\sqrt{V_g + V_e + V_{inter} + V_{intra}/NGE}}$$

$$\frac{R_{grain}}{R_{epi}} = \frac{i_{grain}}{i_{epi}} \times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$$

On cherche donc maintenant à exprimer les intensités de sélection en fonction des paramètres voulus. On sait que l'intensité de sélection en unité d'écart-type est liée à la proportion d'individus conservés lors de la sélection par troncation. On appelle cette proportion P%. On a :

$$P\%_{grain} = \frac{nsel}{NGO}$$

$$P\%_{epi} = \frac{nsel}{NGE \times NEO}$$

Car nsel est le nombre de grains sélectionné parmi la population de grain dans laquelle la sélection s'opère, qui contient NGO grains, et $\frac{nsel}{NGE}$ est le nombre d'épis à sélectionner pour obtenir nsel grain à la fin, ce nombre d'épi étant sélectionné parmi la population d'épis observés contenant NEO épis. Il est donc intéressant d'exprimer i en fonction de P% pour relier l'intensité de sélection aux paramètres voulus. On l'obtient de cette façon :

Soit $X \sim N(\mu, \sigma)$ une variable aléatoire suivant une loi normale de moyenne μ et d'écart-type σ .

Soit $\phi(\mu, \sigma; x)$ la fonction de densité de la loi normale de moyenne μ et d'écart-type σ en x.

Soit $\varphi(\mu, \sigma; x)$ la fonction de répartition de la loi normale de moyenne μ et d'écart-type σ en x.

De plus, d'après "Truncated normal distribution," (2023) que :

En posant :

$$\alpha = \frac{a - \mu}{\sigma}$$

$$\beta = \frac{b - \mu}{\sigma}$$

Alors :

$$E(X | a < X < b) = \mu - \sigma \times \frac{\phi(0,1; \beta) - \phi(0,1; \alpha)}{\varphi(0,1; \beta) - \varphi(0,1; \alpha)}$$

C'est l'espérance de X sachant que X se trouve entre a et b.

Opérer une sélection par troncation dans un trait suivant une loi normale revient à dire que $b = +\infty$ et a est égal au seuil de troncation appliqué. En posant S le différentiel de sélection pour un trait suivant une loi normale de moyenne μ et d'écart-type σ , on obtient :

$$\mu + S = E(X | a < X < +\infty)$$

Si $b = +\infty$, alors $\beta = +\infty$. Or :

$$\lim_{\beta \rightarrow +\infty} \phi(0,1; \beta) = 0$$

$$\lim_{\beta \rightarrow +\infty} \varphi(0,1; \beta) = 1$$

$$S = i \times \sigma$$

On obtient donc :

$$\mu + i \times \sigma = \mu - \sigma \times \frac{0 - \phi(0,1; \alpha)}{1 - \varphi(0,1; \alpha)}$$

$$\mu + i \times \sigma = \mu + \sigma \times \frac{\phi(0,1; \alpha)}{1 - \varphi(0,1; \alpha)}$$

$$i = \frac{\phi(0,1; \alpha)}{1 - \varphi(0,1; \alpha)}$$

Or, pour un trait suivant une loi normale de moyenne μ et d'écart-type σ , la proportion P% d'individus conservés après une sélection par troncation vaut :

$$P\% = P(X > a)$$

$$P\% = 1 - P(X \leq a)$$

$$P\% = 1 - P\left(\frac{X - \mu}{\sigma} \leq \frac{a - \mu}{\sigma}\right)$$

$$P\% = 1 - P\left(\frac{X - \mu}{\sigma} \leq \alpha\right)$$

Comme $\frac{X - \mu}{\sigma} \sim N(0,1)$, alors :

$$P\% = 1 - \varphi(0,1 ; \alpha)$$

On obtient donc que :

$$i = \frac{\phi(0,1 ; \alpha)}{P\%}$$

Donc :

$$i = \frac{1}{P\%} \times \frac{1}{\sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \alpha^2\right)$$

Comme la fonction de répartition d'une loi normale a une fonction réciproque (la fonction de quantiles de cette loi normale qui associe un seuil de troncation à la proportion d'individus sélectionnés) :

$$P\% = 1 - \varphi(0,1 ; \alpha)$$

$$1 - P\% = \varphi(0,1 ; \alpha)$$

$$\varphi^{-1}(0,1 ; 1 - P\%) = \alpha$$

Donc :

$$i = \frac{1}{P\% \times \sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \varphi^{-1}(0,1 ; 1 - P\%)^2\right)$$

Ainsi :

$$i_{grain} = \frac{NGO}{n_{sel} \times \sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \varphi^{-1}\left(0,1 ; 1 - \frac{n_{sel}}{NGO}\right)^2\right)$$

$$i_{epi} = \frac{NGE \times NEO}{n_{sel} \times \sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \varphi^{-1}\left(0,1 ; 1 - \frac{n_{sel}}{NGE \times NEO}\right)^2\right)$$

Et donc :

$$\frac{R_{grain}}{R_{epi}} = \frac{i_{grain}}{i_{epi}} \times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$$

$$\frac{R_{grain}}{R_{epi}} = \frac{\frac{NGO}{n_{sel} \times \sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGO}\right)^2\right)}{\frac{NGE \times NEO}{n_{sel} \times \sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGE \times NEO}\right)^2\right)}$$

$$\times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$$

Les 2π et les n_{sel} se simplifient, et comme $\frac{e^a}{e^b} = e^{a-b}$, alors :

$$\frac{R_{grain}}{R_{epi}} = \frac{NGO}{NGE \times NEO}$$

$$\times \exp\left[-\frac{1}{2} \times \varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGO}\right)^2 - \left(-\frac{1}{2} \times \varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGE \times NEO}\right)^2\right)\right]$$

$$\times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$$

Et on obtient finalement :

$\frac{R_{grain}}{R_{epi}} = \frac{NGO}{NGE \times NEO} \times \exp\left[\frac{1}{2}\left(\varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGE \times NEO}\right)^2 - \varphi^{-1}\left(0,1;1 - \frac{n_{sel}}{NGO}\right)^2\right)\right] \times \sqrt{\frac{V_g + V_e + V_{inter} + V_{intra}/NGE}{V_g + V_e + V_{inter} + V_{intra}}}$

Annexe 3 : Représentations graphiques alternatives des scénarii de sélection explorés avec l'équation 1

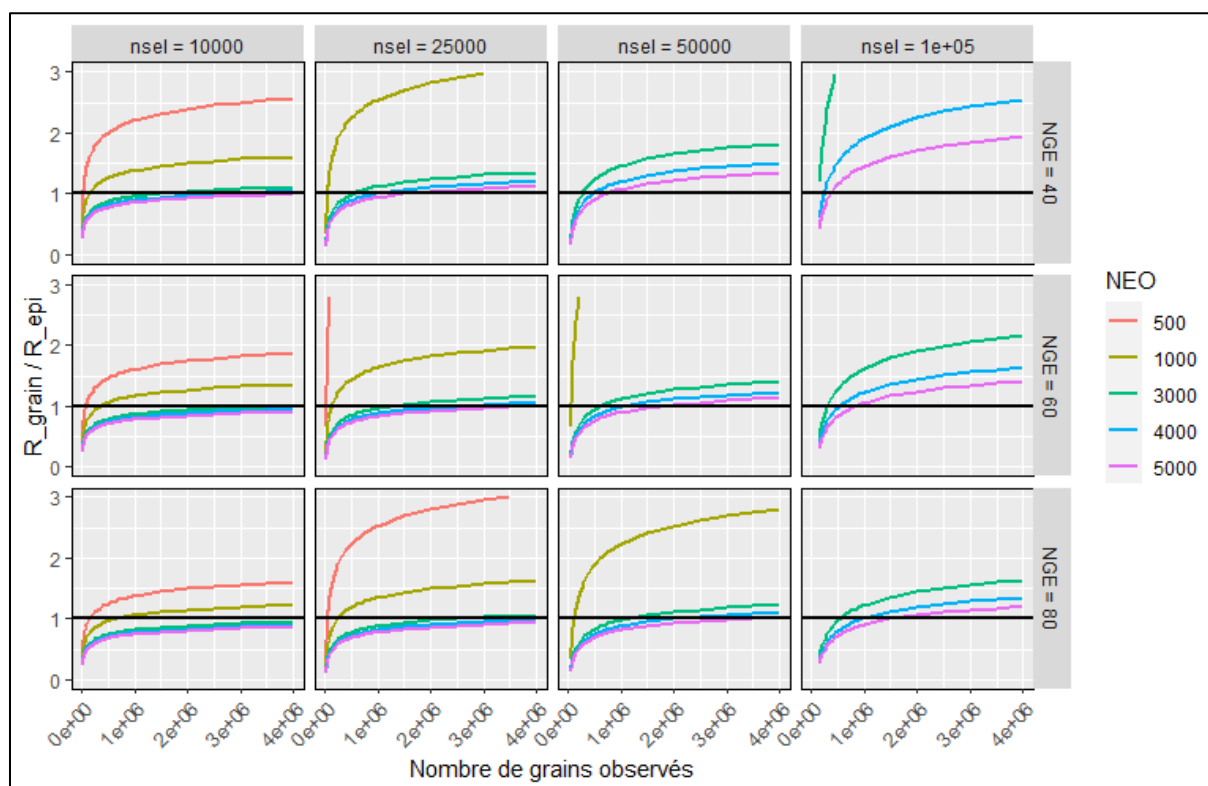


Figure 8 : ERS calculé à partir de l'équation 1 pour différentes combinaisons de paramètres sans échelle logarithmique.

Les combinaisons de paramètres testées sont : n_{sel} allant de 10.000 à 100.000, NGE allant de 40 à 80, NGO allant de 10.000 à 5.000.000, NEO allant de 500 à 5.000. Les composantes de la variance utilisées sont celles estimées dans le dispositif de sélection *in silico*. Chaque graph représente le rapport R_{grain}/R_{epi} en fonction de NGO pour une combinaison de $n_{sel} \times NGE$. Chaque courbe représentée correspond à une valeur de NEO .

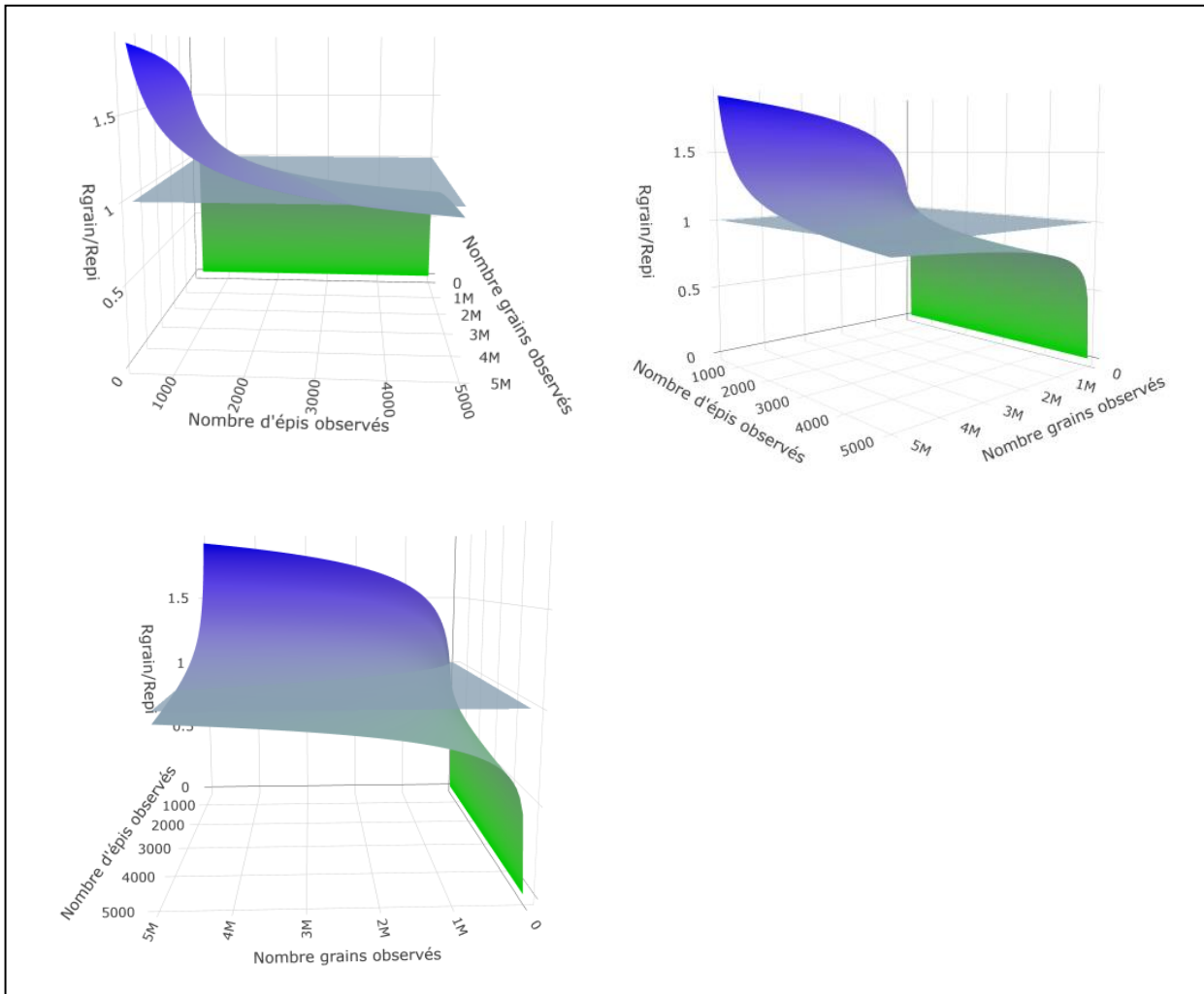


Figure 9 : ERS calculé à partir de l'équation 1 pour différentes combinaisons de paramètres en 3 dimensions.

Les combinaisons de paramètres testées sont : $n_{\text{sel}} = 10.000$, $NGE = 60$, NGO allant de 10.000 à 5.000.000, NEO allant de 500 à 5.000. Les composantes de la variance utilisées sont celles estimées dans le dispositif de sélection in silico. Chaque graph représente le rapport $R_{\text{grain}}/R_{\text{epi}}$ en fonction de NGO et NEO . Sur chaque graph est représenté le plan $R_{\text{grain}}/R_{\text{epi}}=1$.

Résumé

Comme toutes les cultures, le blé dur va être impacté par le changement climatique et la demande d'une population croissante. Une transition vers un modèle agricole plus résilient et durable du point de vue environnemental, social et économique est donc nécessaire. Parmi toutes les étapes de la production, la sélection variétale est la plus en amont, et se doit de proposer des innovations pour participer à cette transition. Dans ce contexte, l'utilisation de variétés populations et la pratique de la sélection massale sont intéressantes car elles offrent justement une certaine résilience face aux conditions climatiques rudes, et ont aussi des implications sociales et économiques pour les agriculteurs.

Durant le stage, deux questions ont été abordées. La première est celle de l'efficacité sélection massale sur la taille du grain, un trait important de la qualité chez le blé dur. La deuxième question concerne tout trait mesurable sur le grain et est de savoir s'il est préférable de mesurer ces traits sur des grains individuels ou sur des moyennes de grains du même épi.

Nous avons pu montrer que la sélection massale sur la taille du grain peut être efficace pour améliorer ce trait. Nous avons aussi développé une équation permettant de comparer la sélection sur grain individuel et sur épi. Cette équation nous a permis de montrer que pour la taille du grain, la sélection grains individuels semble avantageuse dans la plupart des situations mais que la sélection sur épi peut devenir plus efficace lorsque les effectifs impliqués sont faibles.

Mots clés

Sélection massale, blé dur, taille des grains, variété population, agroécologie

Pour citer ce document : Bienvenu, Clément, 2023. Les pratiques de sélections paysannes de blé dur : Quel regard porter sur les grains ?. Mémoire d'Ingénieur Agronome, diplôme, Amélioration des plantes et ingénierie végétale méditerranéennes et tropicales, L'Institut Agro Montpellier. 54 pages.