

**DÉTERMINISME GÉNÉTIQUE DE LA PLASTICITÉ DE LA COMPOSITION
PROTÉIQUE DES GRAINES DE LÉGUMINEUSES VIS-A-VIS DE
L'ENVIRONNEMENT :
RÔLE DU MÉTABOLISME DU SOUFRE**

Thèse de Kévin **CARTELIER**¹

Analysée par Dominique **JOB**²

Co-directrices de la thèse :

Christine **LE SIGNOR**, Ingénieure de recherche INRAE

Karine **GALLARDO**, directrice de recherche INRAE

En France, le pois est le principal facteur de déclin de la production des protéagineux, avec une diminution des surfaces de 700 000 ha à 200 000 ha en 30 ans, principalement en raison de rendements très instables, dus à une forte sensibilité de cette culture aux stress biotiques (champignons : *Aphanomyces*, ascochytose ; insectes : bruches, pucerons) et abiotiques (stress hydrique, thermique, nutritionnels). Ces différents stress peuvent, de plus, impacter la teneur ainsi que la composition protéique des graines, induisant des changements sur la valeur nutritionnelle et donc marchande.

Le regain d'intérêt pour les protéines végétales stimule les recherches visant à développer des marqueurs d'aide à la sélection de variétés de légumineuses mieux adaptées aux besoins nutritionnels de l'homme et des animaux d'élevage. Par ailleurs, de manière générale, les légumineuses sont extrêmement utiles en agriculture dans la mesure où elles sont capables de fixer l'azote atmosphérique grâce à des symbioses racinaires avec des bactéries rhizosphériques fixatrices d'azote. Ainsi, elles sont capables de se développer en l'absence de tout apport de fertilisants de synthèse, préservant des ressources fossiles pour leur fabrication et prévenant, suite à leur épandage, d'importantes pollutions des cours d'eaux, zones littorales et nappes phréatiques.

Les légumineuses sont particulièrement riches en protéines. Parmi les acides aminés essentiels des protéines de réserve majoritaires des graines, la lysine est présente en forte proportion, ce qui en fait un bon complément des céréales, où cet acide aminé est peu représenté. Cependant, la faible proportion en acides aminés soufrés (cystéine, méthionine) (2 %) et en tryptophane (4 %) dans les protéines de réserve des graines de légumineuses est un frein à leur utilisation, ces acides aminés étant essentiels pour l'homme et les animaux d'élevage. Pour promouvoir les protéines issues des graines de légumineuses, il est donc nécessaire, d'améliorer l'équilibre nutritionnel en acides aminés en augmentant la proportion

¹Thèse de doctorat de l'Université Bourgogne Franche-Comté préparée à l'INRAE-UMR 1347 Agroécologie, Pôle GEAPSI – Équipe FILEAS et soutenue le 19 mars 2021.
École Doctorale n°554 « Environnements – Santé », Doctorat de spécialité Sciences agronomiques.

²Directeur de recherche émérite au CNRS, membre de l'Académie d'agriculture de France, section 6, « Sciences de la vie ».

ANALYSE DE THÈSE

d'acides aminés soufrés et de tryptophane, mais aussi de stabiliser cet équilibre face aux stress biotiques et abiotiques. Signalons, dans ce contexte, que des carences en soufre, altérant le rendement des cultures, sont de plus en plus fréquemment observées du fait d'une meilleure maîtrise des pollutions industrielles entraînant des émissions de soufre, jusqu'à présent suffisantes pour la croissance des cultures. Il est donc nécessaire aujourd'hui de mieux pouvoir contrôler la nutrition soufrée des cultures.

Ainsi, outre un enrichissement de nos connaissances sur plusieurs aspects fondamentaux, ces problématiques revêtent un fort intérêt agronomique.

L'objectif de cette thèse était d'identifier, par une approche de génétique d'association (GWAS Génome Wide Association), des gènes candidats pour optimiser ou stabiliser la composition protéique des graines de légumineuses. Le manuscrit de thèse est structuré en quatre chapitres.

Le **Chapitre I** est une synthèse bibliographique particulièrement bien détaillée des connaissances générales utiles à la compréhension des travaux de thèse (définition des plantes légumineuses, dont le pois et *Medicago truncatula*, développement de la graine et accumulation des réserves, composition des graines de légumineuses, voies de biosynthèse des acides aminés soufrés chez les plantes, stratégies visant à modifier le contenu en acides aminés des graines, impact des variations environnementales sur le rendement et la composition protéique des graines de légumineuses, e.g., stress nutritionnel tel que la carence en soufre).

Le **Chapitre II** correspond à l'étude du déterminisme génétique de la plasticité de la composition protéique des graines en réponse à des variations environnementales en utilisant la variabilité naturelle disponible chez la légumineuse modèle *Medicago truncatula*. Les connaissances acquises sur les gènes candidats ainsi identifiés pourront par la suite être transférées à une espèce agronomique (e.g., le pois) et utilisées en sélection afin d'obtenir des variétés améliorées dans la stabilité de leur composition protéique. Ce Chapitre est présenté sous forme d'une publication en anglais, parue dans la revue Plant Journal et dont Kévin Cartelier est premier auteur.

Le **Chapitre III** porte sur la réponse d'une collection d'écotypes de pois à une carence en soufre. Ce troisième chapitre est structuré en deux parties. La première est consacrée à l'étude du déterminisme génétique de la composition protéique des graines et de sa réponse à une carence en soufre. La deuxième traite les composantes du rendement et la teneur en protéines des graines jusqu'à l'identification des régions du génome et des gènes potentiellement impliqués dans leur variabilité en réponse à une carence en soufre.

Enfin, le **Chapitre IV** est une conclusion générale des travaux et apporte des pistes de recherche faisant suite aux résultats du travail de la thèse.

Les principaux résultats de la thèse sont les suivants :

D'une part, l'auteur a analysé par électrophorèse mono-dimensionnelle la composition protéique des graines de 200 génotypes de *Medicago truncatula* cultivés dans quatre environnements contrastés (apport suffisant en eau, stress hydrique, et deux semis décalés). Les indices de plasticité calculés pour chaque bande protéique ont permis d'identifier, par génétique d'association à l'échelle du génome (GWAS), des polymorphismes de séquence associés à des variations de plasticité des protéines de réserve majeures, les globulines 7S et 11S. La liste des gènes candidats montre une forte proportion de gènes liés à la transcription, à la réparation de l'ADN et à la transduction du signal. Fait très intéressant, d'autres gènes surreprésentés sont impliqués dans les métabolismes du soufre et de l'aspartate (précurseur de la lysine) conduisant à la synthèse de méthionine et de lysine (*cf.*

Ravanel S, Gakière B, Job D, Douce R. The specific features of methionine biosynthesis and metabolism in plants. *Proc Natl Acad Sci U S A*. 1998 Jun 23;95(13):7805-712). En exploitant des données d'expression publiquement disponibles chez *Medicago truncatula*, et en remplaçant ces gènes dans les voies métaboliques correspondantes, l'auteur met en évidence l'importance des voies de recyclage de la méthionine dans le contrôle de la plasticité de la composition protéique des graines. Ce recyclage fait intervenir l'enzyme homocystéine S-méthyltransférase 3 impliquée dans la régénération de méthionine à partir de S-méthylméthionine, une forme de transport de méthionine dans le phloème. Ce recyclage est une caractéristique essentielle. En effet, la méthionine étant abondamment utilisée dans une myriade de réactions de méthylation, ce recyclage permet d'« économiser » le soufre nécessaire à la croissance des plantes, en le recyclant *via* le cycle des méthyles (*cf.* Ravanel *et al.* 1998). Cette section de la thèse a fait émerger de nouveaux gènes à cibler pour améliorer et stabiliser le contenu en acides aminés soufrés, ouvrant des perspectives de recherches translationnelles visant à améliorer la qualité des graines de protéagineux comme le pois.

D'autre part, l'auteur a exploré la variabilité de réponse d'une collection d'accessions de pois à une carence en soufre, aux niveaux des composantes de rendement et de la composition protéique des graines, ce qui a permis d'identifier par GWAS des gènes potentiellement impliqués dans cette variabilité de réponse. Deux expériences ont été conduites deux années successives avec respectivement 304 et 198 génotypes développés dans deux conditions : apport suffisant ou carencé en soufre. L'effet année et l'effet de la carence en soufre étant significatifs sur les variables étudiées, des indices de plasticité reflétant l'amplitude de réponse vis-à-vis d'une carence en soufre ont été calculés pour chaque variable. Les résultats obtenus par GWAS ont mis en évidence des gènes candidats pour contrôler la composition protéique des graines dans les environnements pauvres en soufre. Plusieurs sont associés aux mécanismes de transport (ex. d'acides aminés) et de régulation post-traductionnelle, tandis que d'autres jouent un rôle dans le métabolisme du soufre ou sont au carrefour entre les métabolismes azotés et soufrés. Ce travail a permis de mettre en lumière des gènes de pois susceptibles de contrôler le statut azoté et soufré des graines, et ainsi de moduler l'équilibre entre l'accumulation des globulines 7S et 11S dans les graines. La même approche GWAS menée au niveau des composantes de rendement a conduit à proposer des gènes candidats pertinents, en particulier pour le contrôle du poids d'une graine.

L'analyse des composantes du rendement a mis en évidence l'impact négatif de la carence en soufre sur ces variables, répétable sur deux années, et a permis d'identifier trois zones de hotspots communes à plusieurs de ces variables ainsi qu'à leur plasticité pour deux d'entre eux. Ces régions correspondent, dans deux cas sur trois, à l'emplacement de gènes majeurs d'architecture (*Le* pour la longueur des entrenœuds, conditionnant donc la taille des plantes, et *Af* pour la transformation des folioles en vrille, important pour les plantes grimpantes, leur permettant de se dresser au-dessus du sol) ayant des effets pléiotropiques sur le rendement. L'intégration de ces gènes majeurs dans le modèle GWAS a permis à l'auteur d'affiner la détection de gènes candidats en révélant de nouveaux SNP (polymorphisme nucléotidique dans l'ADN) significatifs. De nombreux autres hotspots détectés pour une ou deux variables ont été placés sur la carte génétique du pois. Ils colocalisent souvent avec des régions chromosomiques porteuses de QTLs des composantes de rendement identifiés chez cette espèce.

De façon remarquable, les deux études de plasticité de la composition protéique réalisées sur les deux espèces de légumineuses (*Medicago truncatula* et pois) et dans des gammes d'environnements très différents ont mis en évidence l'implication majeure du métabolisme du soufre. Ces études démontrent que le métabolisme du soufre est un point de contrôle majeur de la plasticité de la composition protéique des graines chez les deux espèces. Ce résultat offre des perspectives intéressantes visant à cibler les voies métaboliques identifiées afin

ANALYSE DE THÈSE

d'augmenter et de stabiliser le contenu en acides aminés soufrés des graines des légumineuses. Cet enjeu est d'importance fondamentale pour promouvoir l'utilisation des graines de pois comme source de protéines en alimentation humaine.

En conclusion, l'ensemble du travail réalisé par Kévin Cartelier et présenté dans ce mémoire, est original et d'excellente qualité. Il est présenté d'une manière claire et rigoureuse. La thèse est fort bien rédigée et illustrée. De manière générale, la présentation est très soignée. Les analyses statistiques et bio-informatiques sont très complètes et leurs interprétations sont appropriées et pertinentes. L'auteur fait preuve d'une grande maîtrise dans l'utilisation des différentes techniques qu'il a utilisées, notamment une combinaison d'approches omiques, biochimiques, moléculaires et génétiques. Cette thèse augmente de manière significative les connaissances sur l'architecture génétique du rendement et de la composition protéique des légumineuses, avec une liste impressionnante de gènes candidats gouvernant la réponse aux contraintes environnementales. L'ensemble des résultats obtenus offrent des perspectives d'amélioration variétale pour stabiliser le rendement et la composition protéique des graines de légumineuses dans des environnements fluctuants, notamment pour la disponibilité en soufre. Cette thèse permet également de comprendre les méthodologies omiques à grande échelle (e.g., GWAS), utiles pour l'analyse de la diversité génétique chez tous les organismes vivants et la mise en évidence de gènes d'intérêt.

Je recommande vivement la lecture de cette thèse qui démontre parfaitement comment les approches omiques permettent d'aborder des questions complexes telles que la composition protéique des graines de légumineuses et l'impact des conditions environnementales sur cette composition. Cet aspect est particulièrement important dans le cadre de la maîtrise du rendement des cultures vivrières face à l'augmentation de la démographie et du changement climatique.

La thèse a fait l'objet de plusieurs publications et présentations, notamment une publication dans *Plant Journal*, une revue de fort facteur d'impact :

- Cartelier K, Aimé D, Ly Vu J, Combes-Soia L, Labas V, Prosperi JM, Buitink J, Gallardo K, Le Signor C. Genetic determinants of seed protein plasticity in response to the environment in *Medicago truncatula*. *Plant J.* 2021 Jun;106(5):1298-1311
- Cartelier K, Aimé D, Ly Vu J, Combes-Soia L, Labas V, Prosperi JM, Buitink J, Gallardo K, Le Signor C. Genetic determinants of seed protein plasticity in response to the environment in K. Cartelier, K. Gallardo, C. Le Signor. « Optimiser et stabiliser la composition protéique des graines de légumineuses ». Journée des doctorants de l'UMR Agroécologie, Dijon, France, 15 mai 2018. Poster
- K. Cartelier, K. Gallardo, C. Le Signor. « Optimiser et stabiliser la composition protéique des graines de légumineuses ». Forum des Jeunes Chercheurs de l'école doctorale Environnement-Santé, Besançon, France, 14 et 15 juin 2018. Poster
- K. Cartelier, K. Gallardo, C. Le Signor. « Optimizing and stabilizing protein composition of legume seeds ». Journée du département INRA de Biologie et Amélioration des Plantes, Paris, France, 20 et 21 septembre 2018. Poster
- K. Cartelier, K. Gallardo, C. Le Signor. « Optimiser et stabiliser la composition protéique des graines de légumineuses ». 2ème rencontres Francophones sur les Légumineuses, Toulouse, France, 17 et 18 octobre 2018. Poster, Prix du meilleur poster
- K. Cartelier, K. Gallardo, C. Le Signor. « Elucidating the genetic determinism of the plasticity of seed proteins in response to the environment using *Medicago truncatula* ». International Conference on Legume Genetics and Genomics 2019, Dijon, France, 13 au 17 mai 2019. Communication orale

ANALYSE DE THÈSE

- K. Cartelier, K. Gallardo, C. Le Signor. « Identification des gènes contrôlant la plasticité de la composition protéique vis-à-vis de l'environnement ». Colloque GRAINES 2019, Angers, France, 21 au 23 mai 2019. Communication orale
- K. Cartelier, K. Gallardo, C. Le Signor. « Identification des déterminismes génétiques de la plasticité de la composition protéique des graines ». Forum des Jeunes Chercheurs de l'école doctorale Environnement-Santé, Dijon, France, 13 et 14 juin 2019. Communication orale.

Les compétences démontrées de Kévin Cartelier lui ont récemment permis d'intégrer l'entreprise Nova Genetics, en tant qu'Assistant coordinateur des programmes de sélection (CDI).

La qualité et les apports des travaux de cette thèse méritent que ce document d'analyse figure sur le site de l'Académie d'agriculture de France, à titre de valorisation.