

# PROSIT : Qualification des protéines d'intérêt pour la qualité brassicole des orges

Julien BILLARD<sup>1</sup>, Sophie SCHWEBEL<sup>1</sup>, Marc SCHMITT<sup>1</sup>, Séverine MAUDEMMAIN<sup>2</sup>, Isabelle CHAILLET<sup>2</sup>, Louis-Marín BOSSUET<sup>3</sup>, Amélie GENTY<sup>4</sup>, Jean-François HERBOMMEZ<sup>4</sup>, Jean-Philippe JELU<sup>5</sup>

1 - IFBM - Institut Français des Boissons de la Brasserie Malterie -

7 rue du Bois de la Champelle - 54500 Vandoeuvre-les-Nncy

2 - ARVALIS - Institut du Végétal - 3 rue Joseph et Marie Hackin - 75116 Paris

3 - GEVES - 25 rue Georges Morel, CS 90024 - 49071 Beaucouze Cedex

4 - UFS - Union Française des Semenciers - 17 rue du Louvre - 75001 Paris

5 - Malteurs de France - 66 rue de la Boétie - 75008 Paris

## 1 Introduction

### ► Contexte

La France est le premier exportateur mondial de malt (80% de la production nationale est exportée) et le second exportateur d'orges brassicoles. Cette position de leader international s'appuie en particulier sur la qualité et la régularité de la production agricole française. Cependant, dans un contexte général de réduction des intrants (dont la fertilisation azotée) et d'aléas climatiques, il est important de savoir comment la filière va pouvoir agir vis-à-vis des contraintes réglementaires et climatiques. Au-delà de la teneur quantitative en protéines des orges de brasserie, se pose la question de leur nature (ou « qualité ») et de leur impact sur le processus de maltage.

Les **attentes qualitatives** de la malterie et de la brasserie pour les orges brassicoles sont nombreuses (*Schmitt, 2013'*). L'orge de brasserie doit répondre à un certain nombre de critères : elle doit être riche en amidon, avoir une bonne énergie germinative (le malt est obtenu par germination en conditions maîtrisées et contrôlées de l'orge), un bon calibrage, et une teneur en protéines totales comprise entre 9,5 et 11,5% du poids sec. Cette situation s'explique schématiquement par un besoin en protéines pour avoir une bonne tenue de mousse et assurer la nutrition azotée de la levure au cours de la fermentation. A l'opposé, un excès de protéines rend le maltage plus difficile et risque de donner du trouble à la bière. Les cahiers des charges des brasseurs sont donc relativement stricts sur ce paramètre.

A ce jour, il y a peu de bibliographie sur la relation entre les classes de protéines et la qualité du malt obtenue contrairement au blé qui a été bien référencé. Dans la filière malterie brasserie, les protéines dans l'orge

sont caractérisées en routine uniquement par leur teneur totale dans le grain. Mais cette donnée ne nous permet pas de différencier les différentes classes de protéines qui la constituent et n'explique pas toujours les différences de qualité de la matière première. L'exemple, présenté en Figure 1, met en lumière la variabilité des paramètres du malt pour une teneur en protéines totales de 10% (en rouge sur le graphique).

L'objectif du projet est d'identifier les fractions protéiques de l'orge qui sont en relation avec sa qualité brassicole, de voir quel est le poids de la variété dans la construction du réseau protéique et également de savoir s'il est possible de les piloter au travers de la fertilisation azotée ou de l'itinéraire de culture.

Ce projet d'intérêt pour la filière réunit tous les acteurs, de la sélection à la transformation (primaire et secondaire) en passant par l'agronomie.

### ► Structure du projet

#### Tâche 1 : Mise au point de l'analyse qualitative des protéines d'intérêt brassicole

Cette première tâche est consacrée au développement et à la mise au point de la méthode d'analyse des protéines dans l'orge. La méthode sera « fixée » lorsqu'elle permettra de lier un ou plusieurs paramètres d'analyse « orge » avec la qualité brassicole (paramètres malt). Elle pourra être éventuellement simplifiée selon les résultats obtenus en Tâche 2.

#### Tâche 2 : Recherche des fractions protéiques d'intérêt

Un premier échantillonnage, produits hors du cadre de PROSIT, présentant une grande variabilité (récolte, itinéraire de culture, qualité brassicole...) a permis de développer la méthode d'extraction et d'analyse des protéines de l'orge pour constituer le début de notre base de données (BDD). Il a ensuite servi à éprouver les modèles de prédiction de la qualité des malts issus des échantillons de la tâche 3.

#### Tâche 3 : Étude des paramètres agronomiques qui impactent la nature et la qualité des protéines de l'orge

Dans ces « essais champ », plusieurs partenaires ont semés des orges de printemps et orges d'hiver sur 2 années en faisant varier l'itinéraire de culture (e.g. fertilisation azotée, apport en soufre...). Une partie de ces échantillons a été sélectionnée par le COPIL pour transformation en micro-malterie pour enrichir la BDD et établir des pistes de liens entre itinéraire agronomique et profil qualitatif des protéines de l'orge.

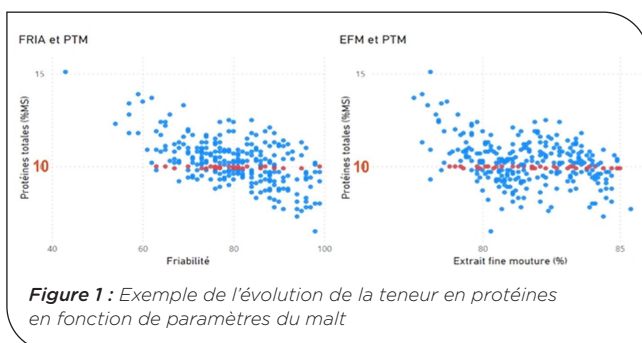


Figure 1 : Exemple de l'évolution de la teneur en protéines en fonction de paramètres du malt

## Tâche 4 : Validation des hypothèses précédentes en brasserie

La dernière étape du projet aurait permis de valider, pour l'ensemble du procédé malterie-brasserie, la nature des protéines de l'orge identifiées comme étant d'intérêt dans les tâches précédentes. Cependant, cela n'a pas pu être fait dans le cadre du projet.

## 2 Matériel et méthode

### ► Production des orges (tâche 2 et 3)

Pour la tâche 2, un appel à échantillons a été lancé pour récupérer des échantillons d'orges de différentes provenances (types d'essai, lieu de culture, variété). Nous avons récupéré les données de 22 essais agronomiques (essais variétaux et essais de fertilisation azotée). Notre choix s'est porté sur 3 essais de fertilisation azotés pour un total de 3 variétés d'orge : KWS IRINA (n=15) et RGT PLANET (n=17) (orges de printemps), ISOCEL (n=13) (orge d'hiver), soit 45 échantillons d'orge. Un ensemble supplémentaire de 56 orges issus du CTPS ont été utilisés au début du projet pour le développement de la méthode d'extraction et d'analyse des protéines de l'orge (résultats non présentés dans cet article).

Dans le cadre de la tâche 3 ; les 9 sélectionneurs et partenaires du projet ont participé aux essais champ (SYNGENTA, SECOBRA, KWS MOMONT, RAGT 2n, LIMAGRAIN, UNISIGMA, FLORIMOND-DESPREZ, ARVALIS et le GEVES).

Dans ces « essais champs » (Figure 2), des orges de printemps (OP) et des orges d'hiver (OH) ont été semées sur 2 années (année 2019 & 2020 ou 2020 & 2021) en faisant varier l'itinéraire de culture. Pour ces essais, il a été décidé de travailler sur la date d'apport plutôt que sur des doses extrêmes d'azote (carences ou surdose). Cependant, des modalités avec plus et moins d'azote (T4 : X+30 et T5 : X-30%) ont quand même été étudiés.

MODALITÉS		T1	T2	T3	T4	T5
ORGE D'HIVER	Dose d'azote (kg N/ha)	X	X	X	X + 30 %	X - 30 %
	Tallage (Z21)	1/3	2/3	0	1/3	1/3
	Epi 1 cm (Z30)	2/3	1/3	2/3	1/3	2/3
	Dernière feuille			1/3	1/3	
MODALITÉS		T1	T2	T3	T4	T5
ORGE DE PRINTEMPS	Dose d'azote (kg N/ha)	X	X	X	X + 30	X - 30
	Au semis	X		X/2	X/2	(X-30) / 2
	Début du tallage		X	X/2	X/2	(X-30) / 2
	Epi 1 cm				+ 30 U N	

Figure 2 : Itinéraire de culture pour les orges d'hiver et les orges de printemps

Sur les 10 variétés (5 OH et 5 OP) semées sur 10 lieux par an, nous avons sélectionné 4 variétés, sur 3 lieux par année et 2 années de récolte (2019 et 2020 pour les OH et 2020 et 2021 pour les OP). Cette sélection s'est faite lors du comité de pilotage annuel du projet. Seuls les lieux les plus répétables avec un rendement satisfaisant, une bonne teneur en protéines et sans souci technique ou sanitaire, ont été envoyés et transformés à l'IFBM.

### ► Extraction des protéines

La méthodologie choisie pour l'extraction des protéines est celle de *Schalk et al. 2017*<sup>2</sup> qui se base sur les travaux d'isolement des protéines selon *Osborne*<sup>3</sup>. Elle consiste à extraire de façon séquentielle les différentes protéines en fonction de leur affinité avec le tampon d'extraction.

La mouture d'orge (<0,2mm) est d'abord extraite avec le premier tampon sous agitation pendant 30 min à température ambiante. Le premier tampon constitué d'eau distillée permet d'extraire les protéines hydrosolubles comme les **albumines**. Une étape de centrifugation permet ensuite de sédimenter les protéines insolubles qui forment un culot. La fraction liquide (contenant les protéines solubles, majoritairement des **albumines**) est récupérée et le deuxième tampon d'extraction, salin (NaCl + KH<sub>2</sub>PO<sub>4</sub> + Na<sub>2</sub>HPO<sub>4</sub>), est ajouté suivi par une étape d'extraction et de centrifugation pour isoler les **globulines**. Les étapes sont ensuite répétées avec le tampon alcoolique pour les **hordéines** (Propanol + eau) et enfin le tampon dénaturant pour les **glutélines** (borate + DTT). Chaque fraction ainsi isolée peut être ensuite analysés.

### ► Analyse des protéines

Après un état des lieux, nous avons identifié un certain nombre de technique utilisés pour la séparation et la quantification des protéines. On peut noter les quantifications par KJELDAHL et DUMAS (*Linko et al. 1989*<sup>4</sup>) pour nous renseigner sur la concentration de chaque fraction protéique après extraction. Des méthodes par chromatographies liquide (ou par électrophorèses SDS-PAGE sur gels (*Molina et al. 2001*<sup>5</sup>, *Marchylo et al. 1986*<sup>6</sup>) permettent de séparer les protéines dans chaque fraction pour quantifier les protéines isolées. Ces techniques ont été mise en œuvre la première année du projet mais n'ont pas été retenues (problème de compatibilité de solvant, temps d'analyse trop long, pas assez discriminant entre les échantillons, etc...)

Le choix de la méthode analytique s'est finalement porté sur l'électrophorèse capillaire, méthode permettant d'analyser un grand nombre d'échantillons, avec un investissement matériel accessible. Cette technique permet d'analyser spécifiquement les protéines en les séparant par leur taille ou masse moléculaire (en kDa) et de les quantifier (*Balazs et al. 2012*<sup>7</sup>).

Les premiers essais réalisés avec le constructeur Perkin Elmer (et les échantillons issus de la tâche 2) ont montré l'intérêt de l'équipement : temps d'analyse très réduit, les 4 fractions protéiques analysables directement, différences observées entre les fractions. Cela a motivé l'intérêt d'analyser la suite des échantillons du projet. De plus cette technique permet de disposer d'informations à de nombreuses échelles : de la quantification globale de chaque fraction jusqu'à une analyse pic par pic.

### ► Maltage des orges

Après récolte, les orges ont été réceptionnées par l'IFBM. Les maltages ont été réalisés dans le service micro-malterie d'IFBM selon le diagramme de maltage de l'année en cours défini dans le cadre du CTPS (Comité Technique Permanent de la Sélection).

Les paramètres technologiques des malts obtenus ont ensuite été caractérisés par IFBM.

### ► Analyse technologique des malts

Les paramètres technologiques étudiés dans le cadre du projet permettent d'évaluer la qualité malticole et brassicole des échantillons. La liste des analyses est la suivante :

(1) Les paramètres de qualité (IQ pour indice qualité), regroupant :

- l'humidité (HUM) : une humidité entre 4 et 5% est importante pour assurer la conservation du grain
- l'extrait fine mouture (EFM) : il définit le potentiel du malt à produire des substances solubles dans le moût (sucres, protéines solubles...)
- le pouvoir diastasique (PD) : c'est un indicateur du potentiel enzymatique du malt nécessaire à la production des sucres fermentescibles lors du brassage.
- la teneur en protéines totales du malt (PT), les protéines solubles du malt (PS), l'indice Kolbach (IK) : Les protéines sont des molécules clés dans le processus de brassage et participent à de nombreux phénomènes : nutrition azotée de la levure lors de la fermentation, activités enzymatiques, trouble, tenue de mousse, couleur...
- la viscosité (VIS), la friabilité (FRIA) et les bêta-glucanes solubles (BGL) : ces trois paramètres sont liés à la dégradation du grain lors du maltage. Une viscosité faible, une friabilité élevée et des bêta-glucanes solubles faibles limiteront les problèmes de filtration au brassage.

Ces paramètres sont utilisés pour l'indice qualité (IQ) dans le cadre du processus d'inscription des variétés par le CTPS.

(2) Les activités enzymatiques : alpha-amylase (AAM), beta-amylase (BAM). Elles sont les composantes du pouvoir diastasique mesuré précédemment et assurent l'hydrolyse de l'amidon en glucides assimilables par la levure lors de la fermentation. Les activités enzymatiques sont d'intérêt pour le procédé de brassage.

(3) Les données récupérées au cours du procédé de maltage : humidité fin de trempage (HFT), humidité fin de germination (HFG) pour les échantillons transformés en malt. Ces données d'humidité en cours de procédé sont un bon indicateur de la prise d'eau du grain, signe de bonne hydratation et transformation de l'orge.

### ► Mise en forme de la base de données

La base de données (BDD) constituée regroupe les informations obtenues dans le cadre du projet pour 220 échantillons d'orge sur 3 années de récoltes (de 2019 à 2021). Les orges sont séparées en 2 espèces (100 orges de printemps et 120 orges d'hiver). Les essais de fertilisation ont été réalisés sur 8 lieux d'essais, 2 espèces d'orge et 8 variétés avec 5 modalités de fertilisation. En plus des informations agronomiques, les 12 paramètres technologiques du malt analysés ont été intégrés. Concernant les fractions protéiques, un ensemble de 53 pics par échantillons d'orge a été intégré après "nettoyage" des spectres obtenus : 18 pour les albumines, 10 pour les globulines, 15 pour les hordéines et 10 pour les glutélines.

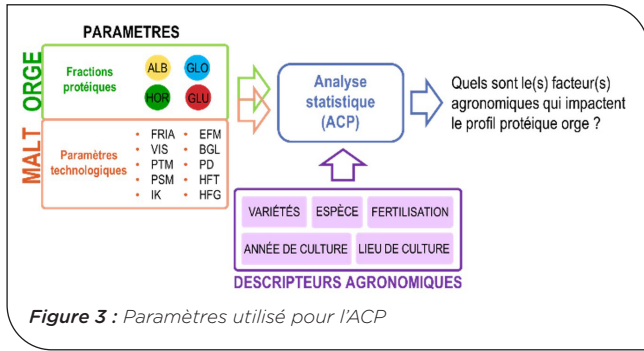
### ► Analyses qualitative et quantitative par chimométrie

Le traitement statistique de la base de données a été réalisé par Arvalis. Le nombre de données ne cesse de croître pour le contrôle de procédé. En effet, le nombre de variables est de plus en plus important. Les informations utiles sont bien souvent difficiles à extraire. Les échantillons tendent aussi à être chimiquement très complexes. Ainsi, il est souvent nécessaire d'utiliser des approches de modélisation multivariée. La chimométrie est une solution parfaite pour de nombreux problèmes physico-chimiques. Massart définit la chimométrie comme une discipline de la chimie qui utilise les mathématiques et les statistiques pour sélectionner les procédures expérimentales optimales et extraire des informations pertinentes (Massart *et al.* 1988<sup>8</sup>). Elle permet notamment d'améliorer la spécificité et la sensibilité de la méthode analytique. Dans le cadre du projet PROSIT, deux méthodes chimométriques ont été employées. L'Analyse en Composantes Principales (ACP) a permis de visualiser les lots de données dans une base vectorielle avec une dimension inférieure à la base des pics du Labchip, de détecter les variables et/ou les échantillons aberrants et d'identifier les facteurs introduisant la variabilité. Après cette phase exploratoire, la régression par la méthode des moindres carrés partiels (PLS) était la méthode de régression linéaire la plus simple et la plus adaptée à la problématique. Celle-ci consiste à construire une équation d'étalonnage à partir des données et des mesures de référence d'une propriété d'intérêt.

### Analyse en composante principale (ACP)

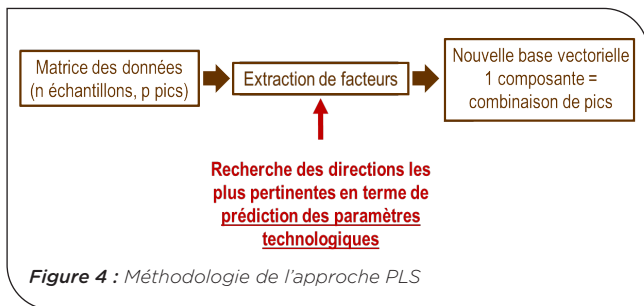
La visualisation des données du LabChip, c'est-à-dire les pics aux différents poids moléculaires, devient difficile lorsque le nombre de variables (les pics) est important. Il est donc nécessaire de réduire le nombre de données en quelques composantes pertinentes. L'ACP est une méthode de condensation des données : elle consiste en une projection dans un espace de représentation de faibles dimensions. En fait, l'ACP sélectionne les projections maximisant la variance entre les points projetés. Chaque axe de projection est appelé « composante principale ». Celle-ci est une combinaison linéaire des pics initiaux et contient ainsi toute l'information utile des données. Le premier axe est choisi dans la direction de la variance maximale de la matrice variance-covariance. De cette façon, la quantité d'information dans la première nouvelle variable est le plus grand possible. Le second axe est choisi orthogonal au premier pour que la seconde composante ne soit pas corrélée à la première.

A partir de là, la réduction du nombre de variables est expliquée par le fait que la variabilité des données du LabChip est causée par les facteurs variant au sein de l'échantillon. Il est à noter que l'information d'un constituant particulier de l'échantillon n'est pas nécessairement contenue dans une seule composante, et l'information utile n'est pas confinée dans la première ou les deux premières composantes même si celles-ci représentent presque toute la variabilité. L'objectif de cette analyse est de voir si les descripteurs agronomiques (Figure 3) peuvent expliquer la représentation des données et s'ils ont un effet sur la qualité des protéines de nos échantillons.

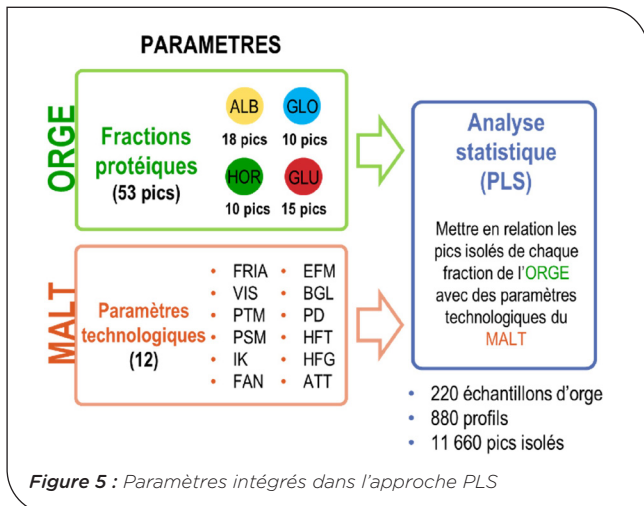


### Régression par la méthode des moindres carrés partiels (PLS)

Comme pour l'ACP, la PLS est basée sur la construction de facteurs à partir des données LabChip initiales. Le but de cette méthode est de réduire la quantité de données et d'éviter ainsi les problèmes de surentraînement sans éliminer les informations utiles. Les composantes ou variables de la nouvelle base vectorielle sont des combinaisons linéaires des variables dans l'ancienne base (Figure 4).



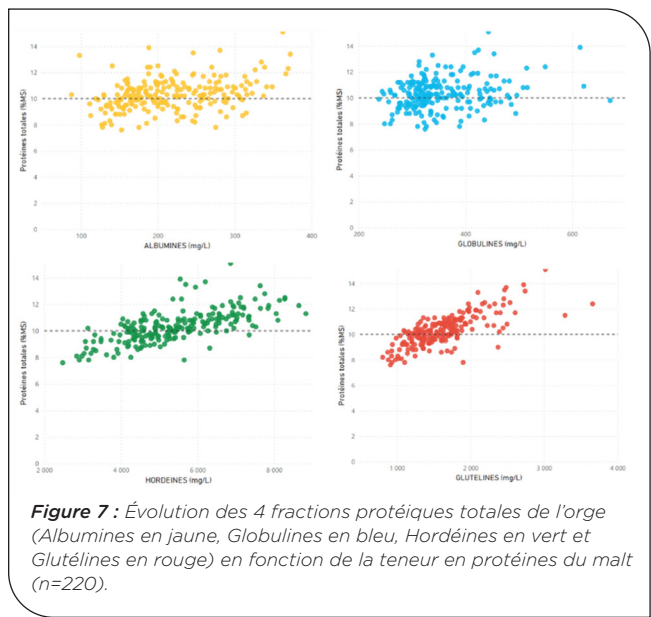
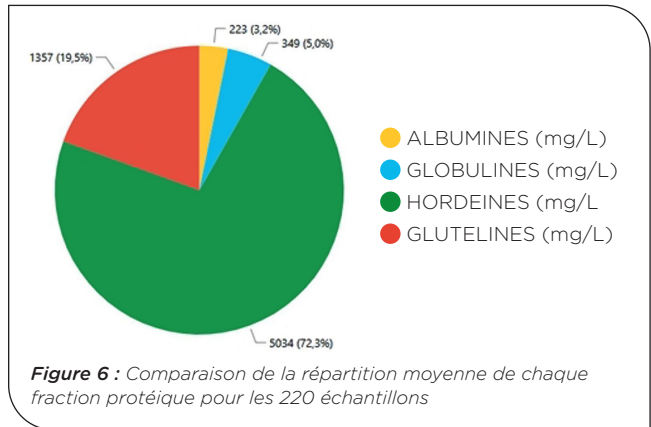
Mais la différence avec l'ACP est que la PLS construit ses facteurs en tenant compte de la corrélation entre les pics aux différents poids moléculaires et les concentrations. La condensation des données se fait donc suivant les directions les plus pertinentes en termes de prédiction des concentrations. La régression sur la matrice des vecteurs propres est utilisée pour construire l'équation de prédiction. La PLS est un processus itératif avec lequel les facteurs sont construits et ajoutés à la régression un à la fois jusqu'à ce que la procédure soit stoppée. Elle permet aussi d'identifier les pics protéiques les plus contributeurs aux modèles de prédiction à partir des paramètres orge et malt intégrés (Figure 5).



## 3 Résultats

### ► Analyse globale des fractions protéiques

La répartition moyenne des 4 fractions protéiques pour les échantillons de la base de données est présentée Figure 6. Les hordéines représentent la grande majorité des protéines extraites du grain d'orge (72%) suivies par les glutélines (21%) et enfin les globulines et albumines avec respectivement 5 et 3% en moyenne (Figure 5).



La Figure 7 montre l'évolution de la quantification de chaque fraction protéique en fonction de la teneur en protéines totales de l'orge. Nous observons que pour une même teneur en protéines (ligne grise pointillée), les concentrations des fractions protéiques évoluent jusqu'à un facteur 4 (entre la plus petite et la plus grande valeur d'une fraction) pour des qualités de malts très différentes. Les fractions protéiques isolées semblent évoluer différemment avec l'augmentation de la teneur en protéines totales. Les hordéines (HOR) et les glutélines (GLU) présentent une relation proportionnelle contrairement aux albumines (ALB) et aux globulines (GLO).

Ces observations justifient la mise en place du projet qui suppose que la teneur en protéines totales de l'orge ne permet pas de définir, à elle seule, sa qualité brassicole, car la « qualité » des protéines présentes compte aussi.



Après cette visualisation générale, une analyse plus poussée en composantes principales a été réalisée par Arvalis pour étudier quel facteur est le plus impactant dans la répartition de ces échantillons.

### ► Impact des facteurs agronomique sur la qualité des protéines de l'orge

Pour rappel, un dispositif agronomique sur 3 ans a été mis en place pour essayer d'identifier des leviers permettant de faire varier la qualité des protéines de l'orge (facteur expliquant la diversité de qualité protéique). Une analyse par ACP de l'effet des différents facteurs (effet de l'année et du lieu de culture, de l'espèce, de la variété et des modalités de culture) sur la qualité des protéines a été réalisée.

#### Analyse en composante principale

L'analyse en composante principale (ACP) a permis dans un premier temps de visualiser la position des échantillons (la différence entre deux échantillons est définie par la distance qui les sépare) et de les regrouper. On visualise 3 groupes (ou clusters) distincts (Figure 6), avec une représentation de plus de 92% de la variabilité existante avec deux dimensions et plus de 98% si on ajoute la composante principale 3 (PC3). Cela veut dire que l'on a au moins 92% de l'information qui est contenue dans 2 dimensions (voir 98 si on intègre PC3) ce qui est robuste.

A présent, l'objectif est de comprendre les causes de ces regroupements et le lien entre fractions protéiques de l'orge et facteurs agronomiques.

#### Effet de l'année de culture

Pour confirmer cela, le facteur « **année de culture** » (2019-2020-2021) a été représenté sur l'ACP en Figure 9.

Les échantillons d'une même année se retrouvent indépendamment dans l'un ou l'autre groupe. Cela veut dire que l'année de culture n'explique pas le regroupement des données.

#### Effet du lieu de culture

De la même manière que l'année de récolte, l'ACP (Figure 10) montre que ce n'est pas le lieu de culture qui permet de définir les clusters.

Nous allons à présent nous intéresser au paramètre « **espèce** » dans la suite de l'analyse et voir son impact sur la définition des clusters.

#### Effet de l'espèce

Dans le cadre de cette étude, il est important prendre en compte l'impact de l'espèce. En effet, les orges d'hiver (OH) et les orges de printemps (OP) ont une physiologie différente. Le cycle de culture est adapté et les qualités finales sont aussi différentes avec une activité enzymatique plus importante pour les orges d'hiver par exemple.

Pour confirmer cette hypothèse le facteur « **espèce** » a été mis en lumière sur l'ACP en Figure 10.

De manière surprenante, les 2 espèces d'orge sont mélangées sauf pour le groupe d'échantillons en haut à droite. Ces résultats expliquent la répartition des échantillons sur les Figure 9 et Figure 10, avec un cluster qui ne regroupe que des OP (et donc que les

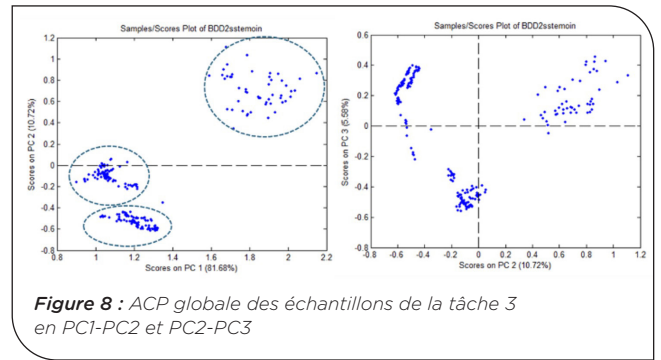


Figure 8 : ACP globale des échantillons de la tâche 3 en PC1-PC2 et PC2-PC3

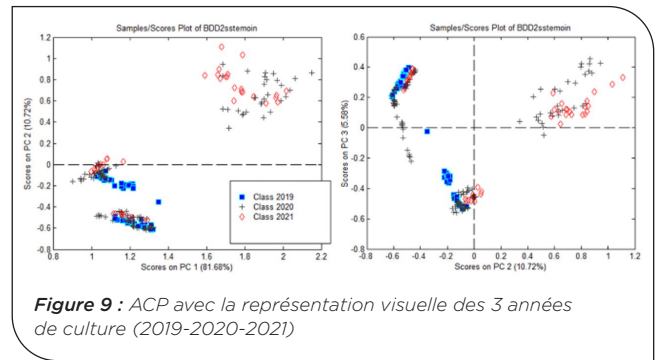


Figure 9 : ACP avec la représentation visuelle des 3 années de culture (2019-2020-2021)

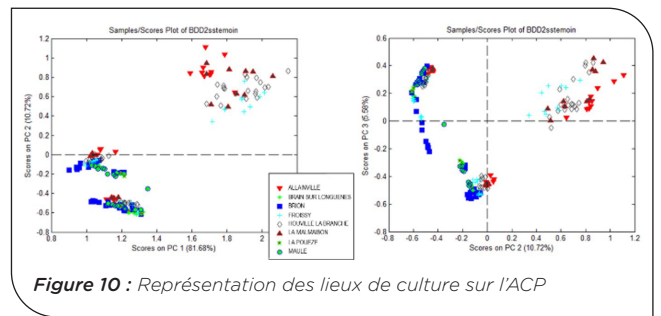


Figure 10 : Représentation des lieux de culture sur l'ACP

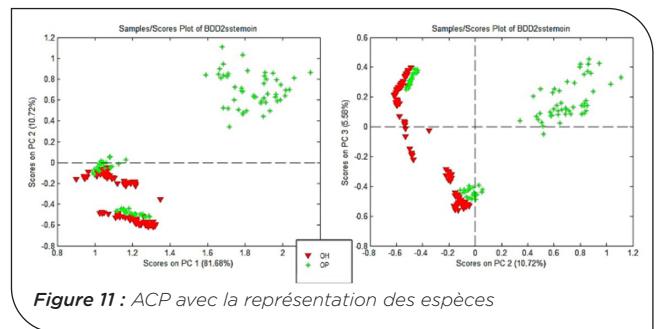


Figure 11 : ACP avec la représentation des espèces

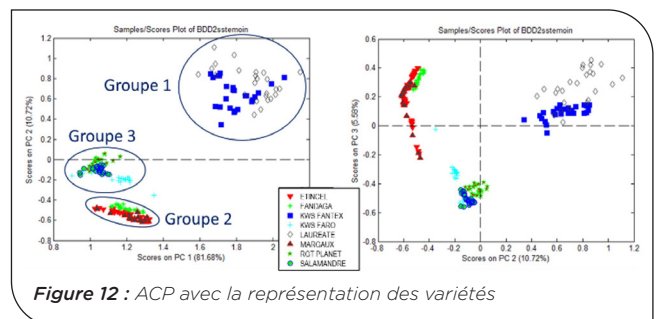


Figure 12 : ACP avec la représentation des variétés

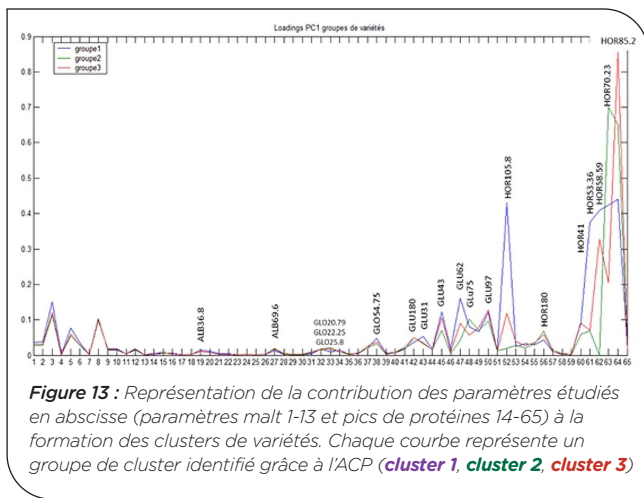
modalités de lieu et d'année de culture correspondant aux OP). Malgré les différences de ratio moyens entre les pics observés (non présenté dans cet article) le facteur « **espèce** » ne semble pas être le facteur le plus explicatif des différences entre échantillons.

### Effet de la variété

L'analyse par ACP (Figure 12) confirme l'effet variétal observé précédemment et il est significatif car tous les échantillons d'une même variété se retrouvent dans un même cluster :

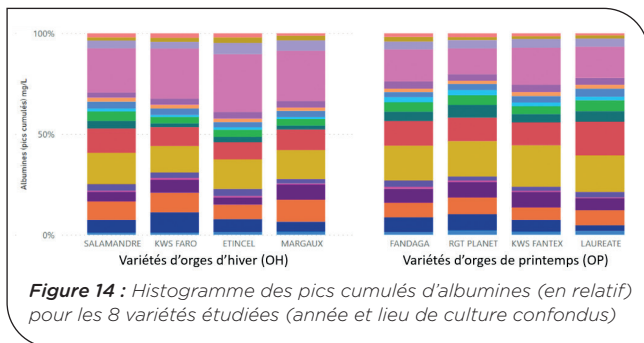
- **Groupe 1** : KWS FANTEX et LAUREATE
- **Groupe 2** : FANDAGA, ETINCEL, MARGAUX
- **Groupe 3** : RGT PLANET, SALAMANDRE, KWS FARO

Pour entrer un peu plus dans les détails, la Figure 13 montre les pics significatifs dans le regroupement des échantillons. Au vu de ces résultats, on confirme que les hordéines mais aussi quelques glutélines sont de forts contributeurs (hauteurs les plus élevées sur la partie droite de la Figure 5). Certains pics d'hordéines sont même présents dans un groupe et absents dans un autre (exemple : HOR 105.8). Les autres paramètres (albumine, globuline, paramètres du malt) ont un impact moindre sur le profil de clustering.



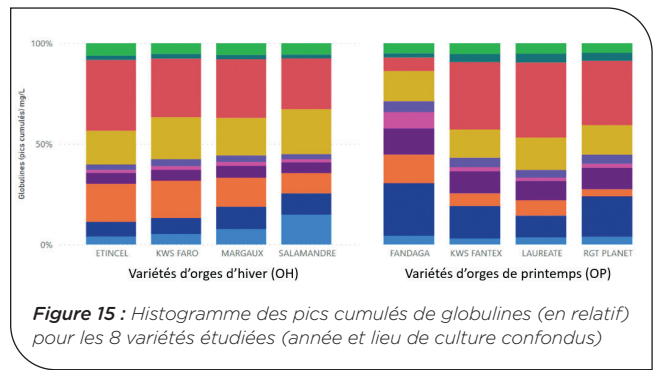
L'analyse précédente s'est focalisée sur l'impact de l'espèce sur la composition protéique de l'orge. Mais il paraît pertinent de descendre à l'échelle de la variété. Les 4 parties suivantes vont s'intéresser à la population de pics entre les variétés des espèces OH et OP (les couleurs peuvent être communes entre les fractions protéiques ALB, GLO, HOR ou GLU mais ils ne s'agit pas des mêmes pics).

### Diversité des albumines



En comparant la répartition des pics entre les espèces OH et les espèces OP (Figure 14), les pics **rose** sont plus importants sur les espèces OH là où les pics **vert foncé, jaune et rouge** sont plus importants chez les espèces OP. Cette classe de protéines semble donc spécifique de l'espèce (OP ou OH) plutôt que de la variété.

### Diversité des globulines

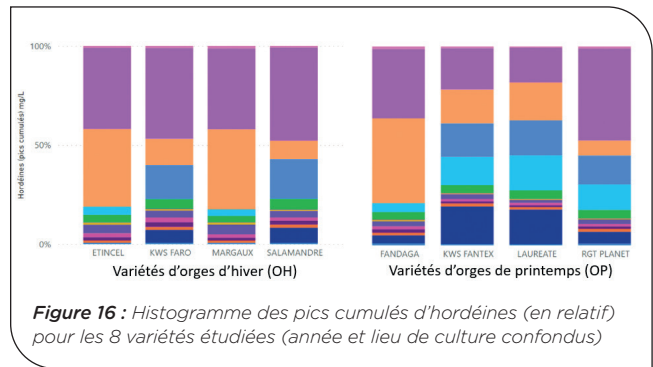


Il y a des différences de profil visibles entre les espèces OH et les espèces OP sur l'ensemble des pics, notamment sur quatre pics (**bleu foncé, bleu clair, orange et violet**) (Figure 15). En regardant les variétés en détail, on se rend compte que FANDAGA (variété OP) présente un profil atypique avec les pics **bleu foncé** et **rose** qui portent une forte contribution protéique. Cette variété est connue pour être plus protéolytique que les autres et ce caractère pourrait être expliqué car la majorité des enzymes sont des globulines.

A l'exception de la variété FANDAGA, il semble donc qu'il existe un profil « globuline » par espèce d'orge, mais qui varie peu entre les variétés d'une même espèce.

### Diversité des hordéines

C'est sur le profil des hordéines que la différence entre les variétés est la plus visible (Figure 16).

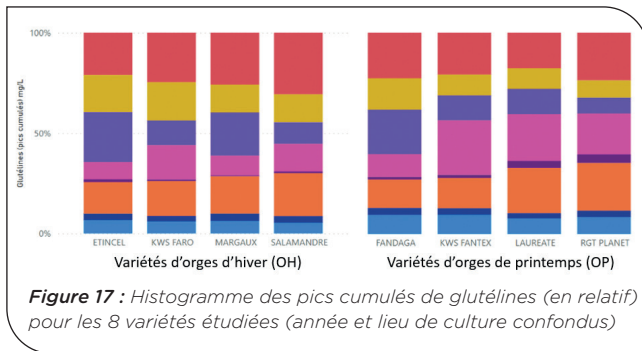


Entre les différentes variétés des espèces OP et OH, les pics **violet, orange, bleu, bleu clair, et bleu foncé** sont significativement différents. En comparant leur profil, les variétés peuvent être regroupés en 3 groupes :

- **Groupe 1** : KWS FANTEX (OP) et LAUREATE (OP)
- **Groupe 2** : FANDAGA (OP), ETINCEL (OH), MARGAUX (OH)
- **Groupe 3** : RGT PLANET (OP), SALAMANDRE (OH), KWS FARO (OH)

On voit ici que des variétés de l'espèce OP (FANDAGA et RGT PLANET) ont des profils proches d'autres variétés de l'espèce OH. L'impact de la variété est donc plus fort que celui de l'espèce sur le profil protéique. D'ailleurs, historiquement, le profil en hordéines est utilisé pour l'identification variétale des orges brassicoles.

## Diversité des glutélines



D'après les résultats en Figure 13, il semble que les glutélines pourraient définir les mêmes groupes que les hordéines avec les pics **jaune**, **rose** et **rouge** qui varient le plus entre les variétés.

## Effet de la fertilisation

L'un des objectifs dans le cadre du projet PROSIT est de savoir si la qualité des fractions protéiques de l'orge peut être pilotée via la fertilisation azotée au champ. En d'autres mots, est-ce qu'une diminution des intrants azotés au champ peut être compatible avec une

préservation de la qualité des protéines à des fins de maltage et brassage. Un itinéraire robuste a été mis en place pour déterminer l'impact de la fertilisation sur les fractions protéiques (T1 à T5, correspondent aux doses présentées dans la Figure 2).

L'effet des modalités de fertilisation est identique pour les OP et les OH avec une augmentation de la somme des fractions protéiques proportionnellement à la teneur en protéines et à la dose d'azote apportée (Figure 16.) Cependant, les différents protocoles de fractionnement de l'azote (T1 à T5) n'ont pas induit de changement dans les ratios moyens des pics.

De manière surprenante, on observe un mélange des modalités dans les différents clusters pour les OH (Figure 19) et les OP (Figure 20), confirmant les résultats précédents (Figure 18). Ainsi, la fertilisation a un impact sur la teneur en protéines totales lié à l'apport d'azote. Cependant elle ne semble pas avoir d'impact significatif sur le profil des protéines du grain. Toutefois, il est possible que les conditions de fertilisation n'aient pas été assez extrémisées pour que nous puissions identifier des différences de compositions notables entre les différentes modalités. De plus, nous n'avons pas pu vérifier la valorisation des apports azotés par des données météorologiques par exemple.



Figure 18 : Histogrammes des fractions protéiques en fonction de la modalité de fertilisation en absolu (à gauche) et en relatif (à droite). Les modalités de fertilisation sont différentes entre les OH et les OP et ont donc été traitées séparément (se référer à la Figure 2). La courbe noire représente l'évolution de la teneur en protéine totale dans le grain.

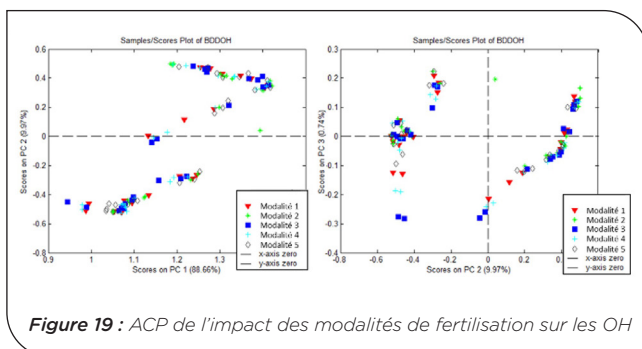


Figure 19 : ACP de l'impact des modalités de fertilisation sur les OH

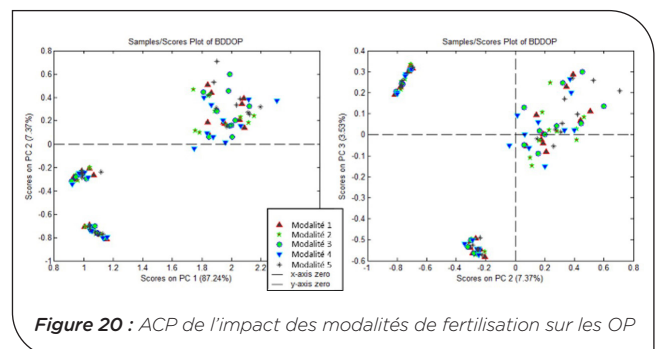


Figure 20 : ACP de l'impact des modalités de fertilisation sur les OP



## ► Prédiction des paramètres technologiques du malt

Un des objectifs principaux de PROSIT est de relier un ou plusieurs pics protéiques de l'orge avec des paramètres technologiques du malt. En d'autres termes, est-ce que certaines protéines sont à privilégier, d'autres à éviter ? Pour ce faire, des outils statistiques (e.g. corrélations simples, PLS) ont été utilisés.

L'objectif de cette analyse est d'obtenir un modèle de prédiction des paramètres technologiques du malt en identifiant les pics protéiques de l'orge les plus contributeurs.

### Prédiction par la méthode des moindres carrés (PLS)

L'analyse par régression PLS a été réalisée sur les 53 pics de l'orge isolés par LabChip et sur 220 échantillons. L'ensemble des résultats est disponible sur la Figure 20 et résumé dans le Tableau 3.

Les graphiques représentent les valeurs mesurées de chaque paramètre malt (en ordonnée) en relation avec les valeurs prédites (abscisse). La droite verte est la droite de régression qui doit s'approcher de l'équation

$$y=x$$

(mesuré=prédit) avec un  $R^2$  le plus proche de 1 (coefficient de corrélation).

La RMSEC représente l'erreur sur la population de développement du modèle. La RMSECV est l'erreur moyenne de sous-ensembles de la population de développement du modèle. C'est un critère statistique utilisé pour sélectionner le nombre de composantes (recombinaison de pics) et avoir une idée de l'erreur sur une population indépendante.

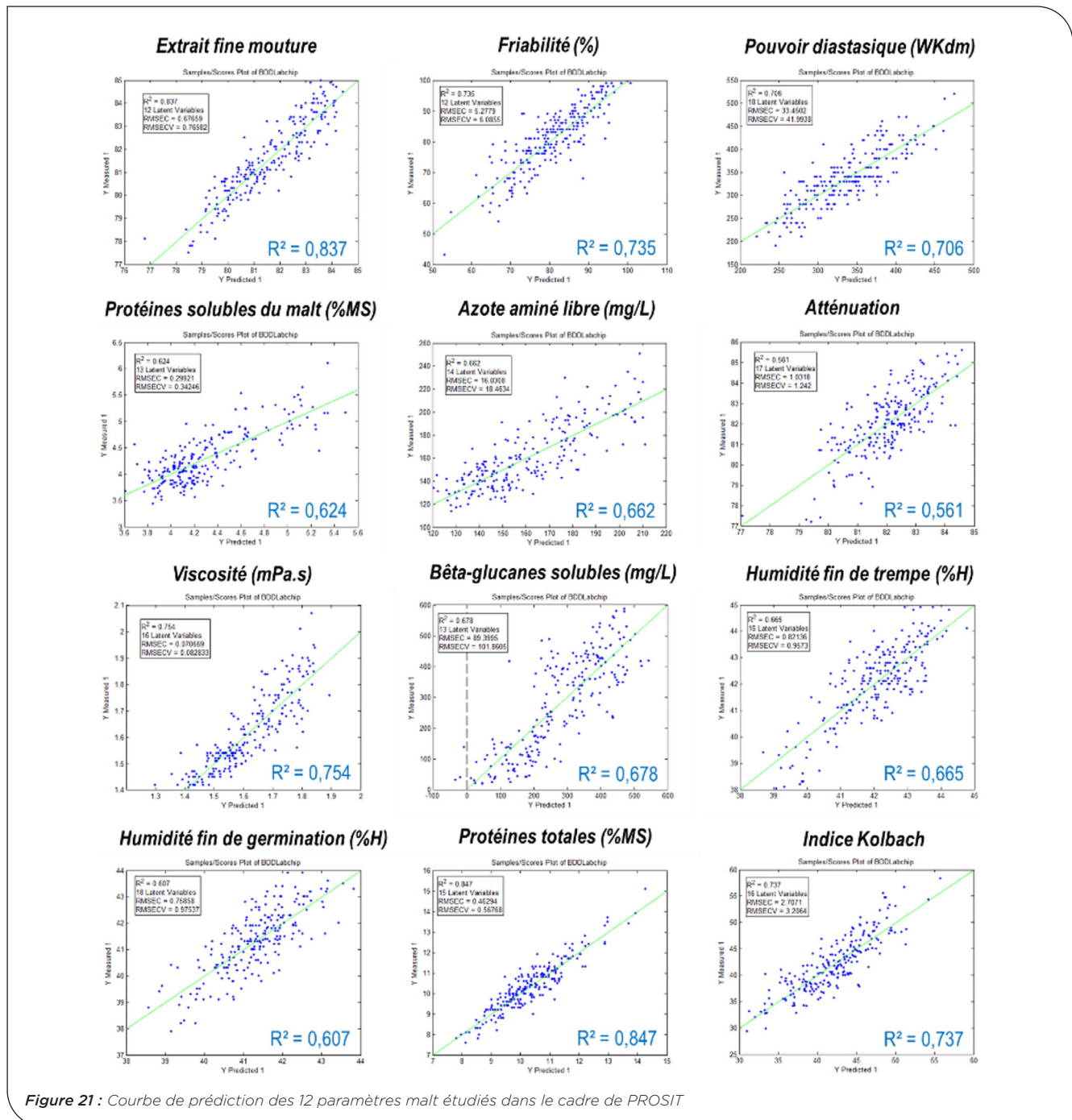


Figure 21 : Courbe de prédiction des 12 paramètres malt étudiés dans le cadre de PROSIT



Paramètres malt	R <sup>2</sup>	Variables	RMSEC	Incertitude brute analytique
Extrait fine mouture	<b>0,837</b>	12	0,68	0,45
Friabilité	<b>0,735</b>	12	5,28	5,47
Pouvoir diastasique	<b>0,706</b>	18	33,45	34,94
Protéines solubles du malt	<b>0,624</b>	13	0,30	0,15
Azote aminé libre	<b>0,662</b>	14	16,03	14,56
Atténuation	<b>0,561</b>	17	1,03	1,43
Viscosité	<b>0,754</b>	16	0,07	0,03
Beta glucans soluble	<b>0,678</b>	13	89,40	39
Humidité fin de trempe	<b>0,665</b>	16	0,82	0,14
Humidité fin de germination	<b>0,607</b>	18	0,76	0,14
Protéines totales malt	<b>0,847</b>	15	0,46	0,29
Indice Kolbach	<b>0,737</b>	16	2,71	-

**Tableau 1 :** Résumé de la force de prédiction des paramètres technologiques avec les pics de protéines

Le Table 1 regroupe les données de prédiction des 12 paramètres malt. La colonne « variables » indique le nombre de pics de protéines contributeurs sélectionnés pour le modèle (entre 12 et 18 variables latentes) pour avoir une prédiction robuste. Cela nous montre que tous les pics ne sont pas pertinents pour la prédiction de chaque paramètre technologique. Les R<sup>2</sup> compris entre 0,561 et 0,847 indique des modèles plus ou moins performants selon les paramètres et perfectibles.

La précision de cette mesure est aussi mise en parallèle aux incertitudes des méthodes de référence utilisées pour mesurer chaque paramètre (RMSEC comparé à l'incertitude analytique). Bien entendu, les performances du modèle prédictif (RMSEC) sont en dessous de l'incertitude, mais reste proches dans certains cas et raisonnable en prédiction.

On peut retenir des résultats intéressants et encourageant pour les trois paramètres : extrait fine mouture (0,837), la

friabilité (0,735) ou le pouvoir diastasique (0,706) avec des erreurs de prédiction raisonnables.

Enfin, les pics les plus contributeurs aux modèles de prédictions ont été listés pour chaque paramètre technologique (Table 2).

Dans ces combinaisons de pics, 20 pics ont pu être identifiés comme fortement contributeurs dans la prédiction des paramètres malt (7 pics d'ALBUMINES, 3 pics de GLOBULINES, 6 pics d'HORDEINES et 4 pics de GLUTELINES). Le poids différent de chaque pic (positif, négatif et grandeur de la variable) en fonction des paramètres technologiques démontre ici toute la complexité de la prédiction.

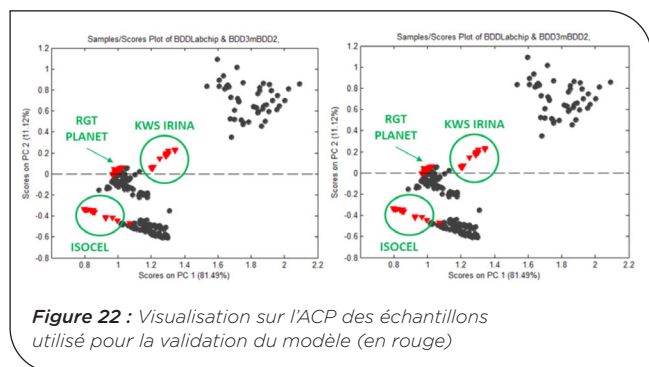
Dans le cas de nos trois paramètres (EFM, FRIA, PD) dont les modèles sont les plus précis, les quatre classes de protéines sont utilisées dans la prédiction. Il n'y a pas un seul type de protéines qui semble significativement plus impliqué dans la qualité du malt qu'un autre.

Pic	ALB110.4	ALB25.8	ALB28.8	ALB36.8	ALB40	ALB69.6	ALB77.3	GLO20.79	GLO25.8	GLO54.75
EFM			-0,0035333		0,0038527	0,0062124	0,0029818	0,0063023		0,0065586
FRIA				0,055939		0,047362		0,09203		0,043588
PD		1,5869	0,89019	0,60109	2,3487	1,4352			0,73322	0,64083
PSM	0,0022646	0,0083859		0,0050576	0,0025424	0,0032811		0,0074488		0,001912
PTM		0,012648	0,0081165		0,012859		-0,0056506		0,022535	0,0076021
FAN		0,32692		0,18629	0,14559	0,13989		0,49414		
ATT	0,028555			0,044975		0,016364	0,010625	0,023746	-0,011314	
BGL		1,3832		0,82815	0,74957		0,74974	-2,2756		0,63251
HFG	0,014414		0,023805	0,03111	0,022006		0,015022		0,058102	
HFT	0,0104		0,015788	0,023434	0,018566	0,0098625				
IK	0,076525	-0,078382	0,028343	0,099954	0,037706		0,036671	0,087862	0,14248	-0,065229
VIS	-0,0013359	0,0017161	0,0025637	0,0031376	0,003295		-0,0016092		0,0030054	
Nbre apports	6	7	7	10	10	7	7	7	6	7
Pic	GLU180	GLU31	GLU35	GLU54.6	HOR12.91	HOR123	HOR23.41	HOR28.71	HOR41	HOR9.13
EFM	-0,0039866	0,0038674		0,0027147		-0,007408	-0,0029213	-0,0031726		0,0084539
FRIA	-0,054493		0,056523	-0,079708	0,067345	-0,073805	-0,094964		0,046571	0,080467
PD			-1,0434				-0,75878	1,374	0,73613	-1,2905
PSM	-0,0046763	0,0034303		-0,0063393	-0,0022737		-0,0037175	0,0021312	0,0035853	
PTM		0,0059351	-0,014169		-0,014972		0,011286	0,0080607	0,0079168	-0,013293
FAN	-0,30763	0,20798		-0,35937			-0,22332	-0,24991		0,16424
ATT		-0,021195		-0,016907	0,018832	-0,010724	-0,032303	-0,0093452	0,012588	0,019669
BGL	0,95125	-1,1606	-0,70101	1,673	-1,1442	1,2465	1,9045		-1,2223	-0,74181
HFG		-0,014874	-0,01974	-0,018855	0,012229	-0,013801		0,029082	0,010273	
HFT		-0,018595		-0,026256	-0,010433			-0,010012	0,010846	0,021651
IK	-0,033756		0,044939	-0,033904				-0,081386		0,055585
VIS				0,0014658	-0,0015803		0,0018108		-0,0010688	-0,0018725
Nbre apports	6	8	6	10	8	6	11	7	10	10

**Tableau 2 :** Pics les plus contributeurs aux modèles de prédiction des paramètres malt

## Validation de la prédiction

Les modèles de prédiction ont été éprouvés avec 45 échantillons de 3 variétés différentes afin de tester leur robustesse. Cette première analyse montre que les nouvelles variétés (KWS IRINA et ISOCEL) ont plus de mal à s'intégrer dans la population initiale alors que RGT Planet est plus souvent intégré. C'est dû au fait que ces 2 variétés n'ont pas été utilisées lors du développement du modèle alors que RGT Planet faisait partie de l'échantillonnage initial.



Cependant, parmi les meilleurs modèles, l'extrait fine mouture et la friabilité semblent bien prédites même pour ces nouvelles variétés. Pour le pouvoir diastasique, la prédiction ne serait valable que pour ISOCEL, et non applicable pour KWS IRINA d'après ces premiers résultats.

En plus de cette première analyse visuelle, nous pouvons comparer avec le Table 3, les deux  $R^2$  obtenues (pour  $n=220$  et  $n=45$ ) et les erreurs relatives pour les 2 populations : RMSEC ( $n=220$ ) et RMSEP ( $n=45$ ).

Paramètres malt	$R^2$ ( $n=220$ )	$R^2$ ( $n=45$ )	RMSEC	RMSEP
Extrait fine mouture	0,837	0,757	0,68	1,51
Friabilité	0,735	0,684	5,28	6,84
Pouvoir diastasique	0,706	0,505	33,45	83,3
Protéines solubles du malt	0,624	0,468	0,30	0,54
Azote aminé libre	0,662	0,482	16,03	31,77
Atténuation	0,561	0,194	1,03	3,29
Viscosité	0,754	0,632	0,07	0,12
Beta glucans soluble	0,678	0,52	89,40	213,58
Humidité fin de trempage	0,665	0,334	0,82	2,97
Humidité fin de germination	0,607	0,549	0,76	1,78
Protéines totales malt	0,847	0,798	0,46	0,84
Indice Kolbach	0,737	0,643	2,71	4,49

**Tableau 3 :** Comparaison des performances des modèles de prédiction pour les 2 populations de développement ( $n=220$ ) et de validation ( $n=45$ )

En appliquant le modèle de prédiction développé sur la première population ( $n=220$ ) on remarque que le modèle perd en puissance. Les performances ( $R^2$  et RMSEP) de prédiction des paramètres sont inférieures, de manière non surprenante, en comparaison à celles de la population d'entraînement.

Les erreurs liées à la prédiction restent modérées pour les paramètres extraits fine mouture et friabilité. En revanche, pour le pouvoir diastasique, la RMSEP est significativement supérieure, notamment par l'éloignement de la population KWS IRINA.

Il semble donc que ce sont les nouvelles variétés qui diminuent la précision des modèles. Cette observation va dans le même sens que l'impact variétal identifié précédemment. A ce stade de l'étude, l'intégration d'un plus grand nombre de variétés dans le modèle permettrait donc de l'améliorer et le rendre plus robuste.

## Conclusion

Des modèles de prédiction des paramètres malt ont été créés à partir des 53 pics sélectionnés. Dans ces combinaisons de pics nous avons identifié 20 pics fortement contributeurs dans la prédiction des paramètres malt présent (7 pics d'ALBUMINES, 3 pics de GLOBULINES, 6 pics d'HORDEINES et 4 pics de GLUTELINES). Pour certains paramètres comme l'extrait fine mouture, la friabilité ou le pouvoir diastasique, la performance de prédiction est encourageante. La validation des modèles sur une population indépendante confirme le potentiel de prédiction de l'extrait fine mouture et la friabilité. Mais l'intégration d'un plus grand nombre de variétés au modèle permettrait d'améliorer la qualité de prédiction.

## 4 Discussion et conclusion

Le projet PROSIT a confirmé l'importance de s'intéresser à la composition fine des protéines des orges brassicoles et a permis de collecter une grande quantité de données et un certain nombre d'informations. Il a notamment permis **(1) la mise au point d'une méthode de quantification des fractions protéiques par électrophorèse capillaire** après un fractionnement séquentiel des types de protéines (albumines, globulines, hordéines et glutélines).

Le projet PROSIT a également confirmé l'intérêt de la démarche car il a montré une grande variabilité de composition pour une même teneur en protéines totales avec **(2) la mise en évidence d'un pool de fractions protéiques d'intérêt en lien avec des paramètres de qualité brassicole par modélisation** (méthode des moindres carrés partiels PLS). Les modèles de prédiction pour l'extrait fine mouture et friabilité sont très encourageants, car validés sur un nouveau pool d'échantillons indépendants. Mais ils restent perfectibles et mériteraient d'être renforcés avec un nombre de variétés plus importants.

**(3) Le projet a mis en évidence que l'impact variétal est significatif sur la composition en protéines des orges** dans le cadre des expérimentations menées. D'après nos résultats, les autres facteurs agronomiques (année de récolte, lieu, espèce d'orge, protocole de fertilisation) n'ont pas d'impact caractéristique.

Ensuite, d'après les résultats collectés sur les 220 échantillons étudiés, on peut retenir deux résultats majeurs :

**Les variétés ne semblent pas être définies par leur profil en albumines et globulines.** Ces deux classes

de protéines semblent plutôt différencier les espèces OP des espèces OH. On peut faire l'hypothèse que ces protéines contribuent à la différence de qualité fonctionnelle (notamment activité enzymatique) entre les OP et les OH. Cependant, la variété FANDAGA fait exception et présente un profil très singulier de globulines, mais intéressant à investiguer pour la caractérisation/prédiction future des variétés protéolytiques.

**Les résultats les plus significatifs de l'effet variétal sont observés sur les hordéines** (et dans une moindre mesure les glutélines) avec trois profils « type » mélangeant les deux espèces d'orges brassicoles confirmés par l'ACP. Ces observations supposent que la sélection variétale des orges brassicoles est et restera le meilleur levier pour contrôler la qualité des protéines.

## 5 Remerciements

Nous tenons à remercier toute l'équipe de micro-maltage de l'IFBM : Catherine Colin, Marjorie Salles et Vanessa Parent, qui ont largement été impliquées dans la transformation de tous les échantillons d'orge. Nous tenons également à remercier Séverine Maudemain (ARVALIS) qui a réalisé tout le traitement statistique des données.

Enfin, merci à tous les partenaires du projet FSOV pour leur contribution significative à ce travail : Arvalis, GEVES, UFS ainsi que tous les sélectionneurs qui ont participé aux essais agronomiques (FLORIMOND-DESPREZ, LIMAGRAIN, KWS MOMONT, RAGT 2n, SECOBRA Recherches, SYNGENTA UNISIGMA). Mais aussi Malteurs de France, ainsi que Brasseurs de France pour leur enthousiasme pour ce sujet.

## Références bibliographiques

<sup>1</sup> **Schmitt, M.** (2013). Attentes qualitatives des Malteurs et des Brasseurs. 15ème Colloque Orge d'Arvalis, 18 avril 2013, Reims.

<sup>2</sup> **Schalk, K., Lexhaller, B., Koehler, P., & Scherf, K. A.** (2017). Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS ONE*, 12(2), 1-20. <https://doi.org/10.1371/journal.pone.0172819>

<sup>3</sup> **OSBORNE, T.B.**, (1907). The proteins of the wheat kernel. Publication 84. Carnegie Institute of Washington, Washington DC

<sup>4</sup> **Linko, R., Lapvetelainen, A., Laakso, P., & Kallio, H.** (1989). Protein Composition of a High-Protein Barley Flour and Barley Grain. *Cereal Chemistry*, 66(6), 478-482.

<sup>5</sup> **Molina-Cano, J. L., Polo, J. P., Romera, E., Arais, J. L., Zarco, J., & Swanston, J. S.** (2001). Relationships between

barley hordeins and malting quality in a mutant of cv. Triumph I. Genotype by environment interaction of hordein content. *Journal of Cereal Science*, 34(3), 285-294.

<sup>6</sup> **Marchylo, B. A., Kruger, J. E., & Hatcher, D.** (1986). High performance liquid chromatographic and electrophoretic analysis of Hordeins during malting for two barley varieties of contrasting malting quality. *Cereal Chemistry*, 63(3), 219-231.

<sup>7</sup> **Balazs G., Tomoskozi S., Harasztos A., Nemeth V., Tamas A., Morgounov A., Belan I., Ma W., Bekes F.** Advantages and Limitation of Lab-on-a-Chip Technique in the Analysis of Wheat Proteins. *Cereal Research Communications*, 2012.

<sup>8</sup> **D.L. Massart, B.G.M.Vandeginste, S.N.Deming, Y. Michotte and L.Kaufman.** Chemometrics: a textbook., Elsevier, Data Handling in Science and Technology, Volume 2, Amsterdam 1988.

# PROSIT : Qualification des protéines d'intérêt pour la qualité brassicole des orges

Julien BILLARD<sup>1</sup>, Sophie SCHWEBEL<sup>1</sup>, Marc SCHMITT<sup>1</sup>, Séverine MAUDEMAIN<sup>2</sup>, Isabelle CHAILLET<sup>2</sup>, Louis-Marin BOSSUET<sup>3</sup>, Amélie GENTY<sup>4</sup>, Jean-François HERBOMMEZ<sup>4</sup>, Jean-Philippe JELU<sup>5</sup>

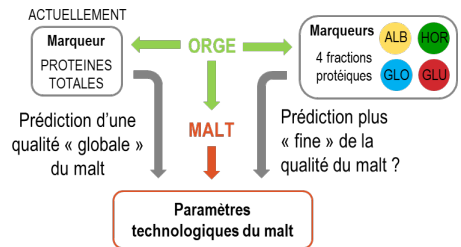
1 - IFBM - Institut Français des Boissons de la Brasserie Malterie  
7 rue du Bois de la Champelle, 54500 Vandoeuvre-les-Nancy

3 - GEVES - 25 rue Georges Morel, CS 90024, 49071 Beaucouze Cedex  
4 - UFS - Union Française des Semenciers - 17 rue du Louvre, 75001 Paris  
5 - Malteurs de France - 66 rue de la Boétie, 75008 Paris

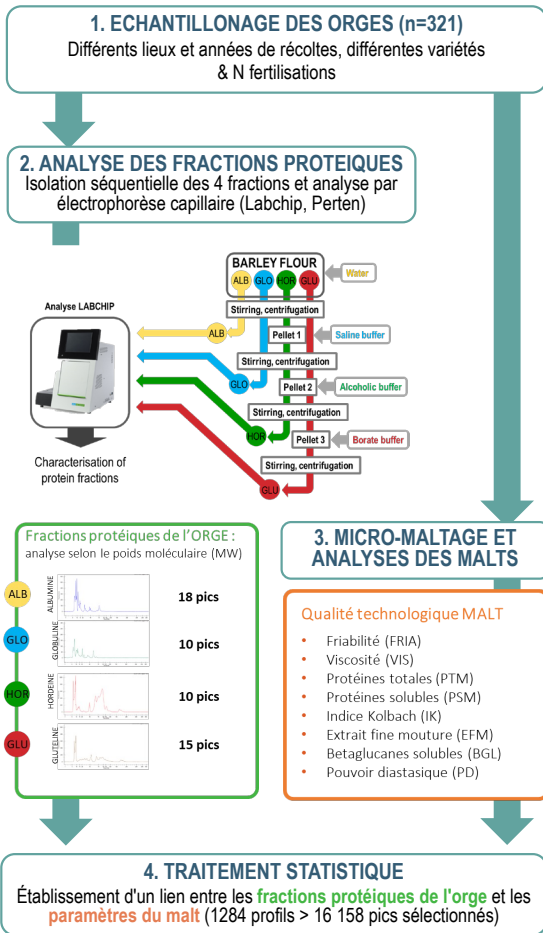
## INTRODUCTION

Les protéines de la matière première (orge et malt) jouent un rôle important dans les processus de maltage et de brassage. Toutefois, les protéines de l'orge sont peu étudiées. Même si les cahiers des charges des malteurs exigent une teneur minimale et maximale en protéines totales de l'orge, la composition détaillée des protéines n'est pas connue et pourrait expliquer les différents niveaux de protéolyse du malt. Dans un contexte de changement climatique et de réduction de la fertilisation azotée, le projet collaboratif PROSIT propose de dresser le profil qualitatif des différentes fractions protéique de l'orge et d'identifier les possibles relations avec les paramètres du malt. L'impact de la variété d'orge et des conditions de fertilisation sur le profil protéique a également été étudié de manière qualitative.

**La teneur en protéines totales de l'orge est une information générale. Peut-on avoir des informations plus fines avec l'étude des classes protéiques ?**

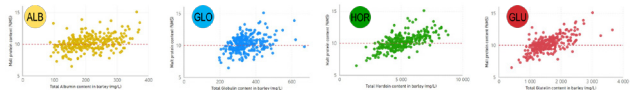


## STRATEGIE DU PROJET

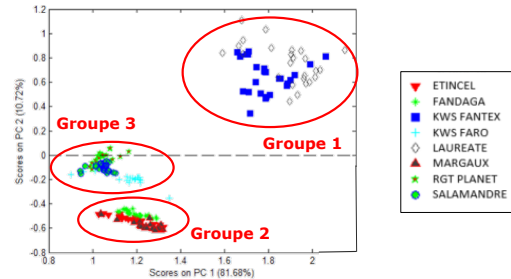


## RESULTATS

Variation très importante de la teneur des fractions pour une même teneur en protéines totales

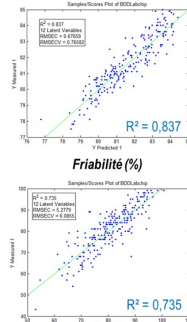


L'analyse ACP des essais de fertilisation azotée (n = 220) montre un effet variétal important sur le profil protéique, mais pas d'effet culture, ni d'effet sur les modalités d'apports.



Des modèles de prédictions des paramètres du malt ont pu être mis au point à partir des pics de protéines identifiés, avec plus ou moins de performances (R<sup>2</sup> entre 0,56 et 0,85) pour les variétés d'orge étudiées.

Extrait fine mouture



Paramètres malt	R <sup>2</sup>	Variables	RMSEC	Incertitude brute analytique
Extrait fine mouture	0,837	12	0,68	0,45
Friabilité	0,735	12	5,28	5,47
Pouvoir diastasique	0,706	18	33,45	34,94
Protéines solubles du malt	0,624	13	0,30	0,15
Azote aminé libre	0,662	14	16,03	14,56
Atténuation	0,561	17	1,03	1,43
Viscosité	0,754	16	0,07	0,03
Beta glucanes soluble	0,678	13	89,40	39
Humidité fin de tremp	0,665	16	0,82	0,14
Humidité fin de germination	0,607	18	0,76	0,14
Protéines totales malt	0,847	15	0,46	0,29
Indice Kolbach	0,737	16	2,71	-

## CONCLUSIONS

Le projet a permis de constituer une base de données caractérisant plus de 320 échantillons dont 220 ont été utilisés pour essayer d'identifier des liens entre fractions protéiques de l'orge et qualité brassicole du malt. Les principales conclusions du projet sont :

- Il existe une très grande diversité dans la composition des protéines de l'orge pour une même teneur totale en azote ;
- Le profil protéique est fortement lié à la variété de l'orge ;
- L'année de culture, le lieu ou le calendrier de fertilisation a peu d'impact sur la composition des protéines, pour les variétés d'orges étudiées ;
- Des modèles de prédictions des paramètres techniques du malt ont été créés à partir de 20 marqueurs, pour les variétés d'orge étudiées ; la diversité génétique doit être renforcée pour rendre plus robuste les modèles.

**Financements :** FSOV, Fondation de la Malterie et de la Brasserie  
**Remerciement à l'ensemble des partenaires du projet :** IFBM, ARVALIS, GEVES, UFS, Malteurs de France & Brasseurs de France

