

LE METABOLOME ET SON UTILISATION EN AMELIORATION DES PLANTES

**Yves GIBON, Sylvain PRIGENT, Thomas DUSSARRAT, Annick MOING,
François TARDIEU et Pierre PETRIACQ**

UMR1332 BFP, Univ. Bordeaux, INRAE, Bordeaux-Nouvelle Aquitaine
LEPSE, INRAE, Université Montpellier, SupAgro, Montpellier, France
yves.gibon@inrae.fr

RESUME

La métabolomique est un outil capable de contribuer significativement au phénotypage des plantes et donc à la sélection de nouvelles variétés mieux adaptées au monde de demain. Grâce à plus de vingt ans de progrès, des nombres toujours plus grands de métabolites peuvent être détectés, de plus en plus rapidement et à des coûts de moins en moins élevés, fournissant une information de plus en plus pertinente, à mi-chemin entre l'expression du génome et le phénotype de la plante. La métabolomique prédictive, qui utilise des approches de modélisation descendante, peut être utilisée pour prévoir la performance des plantes. Des études menées avec des panels de diversité intraspécifique montrent que des prédictions peuvent être obtenues avec des corrélations proches de celles obtenues par la prédiction génomique. Ainsi, des prédictions de rendement dans un essai multisite au champ ont été obtenues à partir du métabolome foliaire mesuré en serre pour un panel d'hybrides de maïs. La métabolomique prédictive, qui s'affranchit des limitations liées aux données de séquence, peut également être appliquée à la diversité interspécifique pour mettre à jour des mécanismes d'adaptation génériques.

Mots-clés : métabolisme, métabolome, métabolomique, modélisation, prédiction, performance

1 – INTRODUCTION

Des premiers agriculteurs à la sélection moderne, le processus de sélection a été en grande partie motivé par l'amélioration de la performance des cultures (Fernandez *et al.*, 2016). Cependant, la définition de la performance a longtemps été limitée au rendement, en particulier pendant la « révolution verte » (Fernandez *et al.*, 2016). Si des améliorations spectaculaires ont contribué à réduire la faim dans le monde, nous sommes aujourd'hui confrontés à une stagnation des rendements des principales cultures malgré des variétés élites pourtant régulièrement remplacées (Long *et al.*, 2015), une situation qui se complique avec les incertitudes liées au changement climatique et à des injonctions parfois contraignantes, comme par exemple l'interdiction de l'usage de pesticides. De plus, en se concentrant sur le rendement, les sélectionneurs ont peut-être négligé la diversité génétique et perdu certaines caractéristiques de performances importantes telles que la résistance aux stress abiotiques et biotiques. Enfin, alors que l'agriculture intensive cherche à maîtriser l'environnement (pesticides, engrais, mécanisation extrême) pour que le potentiel de variétés à haut rendement puisse s'exprimer, l'agriculture de demain devra sans doute prendre davantage en compte les spécificités locales (climat, sol, microflore, etc.), ce qui pourrait se traduire par la génération de variétés adaptées à des scénarios climatiques spécifiques.

Dans ce contexte, la prédiction de la performance des plantes dans n'importe quel scénario de croissance va prendre de l'importance. Outre les promesses de la sélection génomique, la phénotypique végétale a été proposée pour fournir une prédiction de performance basée sur des descripteurs phénotypiques à faible coût et à haut débit plutôt que sur les polymorphismes de l'ADN (Rincint *et al.*, 2018). Cette approche a pour l'instant très largement privilégié l'observation non-destructive des phénotypes en suscitant le développement d'un ensemble de méthodes d'imagerie (Tardieu *et al.*, 2017). A ce jour le "grain" d'observation de ces méthodes se situe à l'échelle de l'organe, de la plante voire de la parcelle. En revanche, l'observation destructive, par laquelle des traits dits cellulaires (transcriptome, protéome, métabolome...) peuvent être observés, a rarement été envisagée car souvent jugée trop coûteuse.

Dans cet article, nous envisageons la métabolomique comme un outil capable de contribuer au phénotypage des plantes et à la sélection de nouvelles variétés mieux adaptées à la diversité des conditions climatiques environnementales associées au changement climatique. Dans un premier temps nous définissons le métabolisme, le métabolome et la métabolomique. Nous considérons ensuite la métabolomique comme un outil potentiellement utile pour des prédictions de performance lorsqu'elle est combinée à des approches de modélisation statistique. Nous illustrons cela avec des résultats obtenus avec un panel de diversité intraspécifique chez le maïs avant de montrer que la métabolomique prédictive, qui s'affranchit des limitations liées aux données génomiques, peut également être utilisée dans un contexte de diversité interspécifique. Enfin, nous proposons d'insérer la métabolomique dans le pipeline de sélection.

2 - METABOLISME, METABOLOME ET METABOLOMIQUE

Le métabolisme est défini comme l'ensemble des réactions chimiques d'un organisme ; il est responsable de la transformation de la matière. Il produit la biomasse et les molécules de défense contre les stress biotiques et abiotiques. Chez les plantes, de nombreux métabolites, en particulier les phytohormones, sont également utilisés comme des signaux régulateurs pour ajuster les programmes de développement, de reproduction et de défense.

Le métabolome est l'ensemble des métabolites d'un échantillon biologique. Suivant les modalités de l'observation, cet échantillon peut être une cellule, un tissu, un organe, un individu voire un écosystème. Il est à noter que le métabolome est particulièrement dynamique et réactif. Ainsi il fluctue au cours du temps, par exemple dans une feuille au cours d'un cycle jour-nuit (Gibon *et al.*, 2006) ou au cours du développement d'un organe (Carrari *et al.*, 2006). Toujours selon les modalités de l'observation, le métabolome peut comprendre les métabolites synthétisés par le système à l'étude, mais aussi tous ceux issus de la nutrition et des interactions avec d'autres systèmes. Plus généralement les métabolites constituent une « empreinte » de l'interaction complexe entre le génome et l'environnement (Viant *et al.*, 2017). Le métabolome contient en effet de l'information reflétant des événements passés (par exemple, des métabolites à renouvellement lent accumulés en réponse à un stress passé), présents (par exemple, des intermédiaires métaboliques à renouvellement élevé) et futurs (par exemple, des précurseurs de la biomasse en construction) (Hall *et al.*, 2022). Les plantes sont de véritables usines chimiques synthétisant un arsenal de molécules aux fonctions multiples. Ainsi, on estime qu'une espèce donnée produit entre 1 000 et 40 000 métabolites sachant que le règne végétal en compterait entre 100 000 et 1 million (Aleekh et Fernie, 2018).

La métabolomique est la tentative de décrire le métabolome (Allwood *et al.*, 2021). D'un côté, on peut chercher à détecter, voire quantifier, tous les métabolites présents dans un échantillon. De l'autre, on peut chercher à expliquer un maximum de la composition de la biomasse. Le fait que quelques dizaines de traits métaboliques expliquent plus de 90% de la composition de la biomasse indique une répartition très inégale des quantités de métabolites. Autrement dit, de très nombreux

métabolites sont présents à des concentrations très faibles. On peut distinguer deux grands types de méthodes utilisées pour détecter et quantifier les métabolites.

Le premier type comprend les méthodes ciblées, par lesquelles un ou plusieurs métabolites sont mesurés. Ces méthodes, généralement quantitatives, utilisent une gamme très variée de principes, se basant sur leurs propriétés chimiques, qui permettent de les quantifier directement ou via des réactions catalysées ou non par des enzymes (le plus souvent par spectrophotométrie ou par fluorescence), mais aussi sur leurs propriétés physiques, en utilisant la spectrométrie de masse, le plus souvent couplée à la chromatographie liquide ou gazeuse. La biochimie analytique a longtemps été plébiscitée, trouvant sans doute son apogée dans les années 70-80 lorsque des centaines d'enzymes étaient disponibles commercialement, offrant la possibilité de doser de nombreux métabolites. La généralisation à partir des années 2000 de l'utilisation de la microplaque, combinée ou non à la robotique, a ouvert de nouvelles perspectives pour la biochimie analytique en augmentant considérablement le débit de l'analyse (le nombre d'échantillons traités par unité de temps). Cependant, cette approche est aujourd'hui limitée à un nombre toujours plus restreint de traits métaboliques en raison de la forte diminution du nombre d'enzymes disponibles commercialement. De plus, les possibilités offertes par la spectrométrie de masse ciblée n'ont cessé de croître en termes de sensibilité, de diversité des métabolites détectés et de débit (Allwood *et al.*, 2021).

Le second type comprend les méthodes non-ciblées qui visent à caractériser le contenu métabolique sans *a priori* à l'aide de la spectrométrie de masse ou de la spectroscopie par résonance nucléaire. Ces méthodes sont le plus souvent semi-quantitatives car produisant des données normalisées à l'aide de standards internes (une ou plusieurs molécules ajoutées aux extraits) ou externes (par exemple un extrait ou mélange d'extraits représentatifs). Elles procurent cependant une information bien plus exhaustive de la diversité métabolique et permettent de découvrir de nouveaux biomarqueurs (ici des signatures métaboliques associées à de la performance). Là encore la spectrométrie de masse semble s'imposer lorsque déclinée en haute résolution elle permet de détecter plus de 20 000 signatures métaboliques dans un échantillon végétal, à partir desquelles jusqu'à plusieurs centaines voire milliers de métabolites peuvent désormais être annotés (Shen *et al.*, 2019).

Un défi majeur sera de parvenir à des méthodes produisant des données reproductibles voire interopérables. En effet, jusqu'à récemment, l'utilisation de la métabolomique en amélioration des plantes a été vue comme commençant par la recherche de biomarqueurs à l'aide de méthodes non-ciblées, puis se poursuivant par l'utilisation de méthodes ciblées moins coûteuses et plus rapides. De récentes avancées technologiques rendent cependant la combinaison haute-densité et haut-débit possible, à savoir la réalisation d'analyses non-ciblées dans de grands nombres d'échantillons en combinant la robotique (pesée des aliquotes, extraction, injection des extraits) et la spectrométrie de masse haute-résolution à acquisition de données ultra-rapide. Des résultats récents montrent que jusqu'à 19 000 signatures métaboliques peuvent être obtenues en moins d'une minute pour un extrait, ce qui ouvre la possibilité de traiter plus de mille échantillons par jour avec un seul spectromètre de masse (Hall *et al.*, 2022). Autrement dit, le même pipeline analytique peut désormais être envisagé à la fois pour la construction et l'utilisation d'un modèle prédictif, d'autant plus que des analyses ciblées et non-ciblées peuvent être conduites dans une même séquence d'analyse.

3 - METABOLOMIQUE PREDICTIVE

La métabolomique prédictive peut se définir comme une approche de modélisation statistique descendante, du phénotype vers les mécanismes et/ou le génotype, visant à associer le métabolisme à des traits d'intérêt. Ainsi, des données métabolomiques peuvent être combinées ou non à d'autres données omiques, à de l'information génétique ou encore à des facteurs

environnementaux pour prédire des caractères complexes tels que la production de biomasse ou le rendement. Un problème qui se pose avec la spectrométrie de masse non-ciblée est que le nombre de variables (jusqu'à plus de 20 000 signatures métaboliques, souvent fortement multi-colinéaires parce que la plupart des métabolites sont décomposés en plusieurs signatures) est nettement plus grand que le nombre d'observations (par exemple le nombre de génotypes utilisés pour construire un modèle). Pour prendre cela en compte, des méthodes de réduction de variables sont souvent utilisées. Ainsi, l'analyse en composantes principales (ACP) permet de rechercher des caractéristiques qui maximisent la variance des données. Cette méthode est beaucoup utilisée en métabolomique pour obtenir facilement une première représentation des données et pour identifier d'éventuelles observations aberrantes. La méthode des moindres carrés partiels (PLS) est une approche supervisée qui permet une discrimination des données en maximisant la covariance avec un trait d'intérêt, augmentant ainsi les capacités de prédiction. Elle permet de rechercher des marqueurs métaboliques *via* le score d'importance des variables correspondantes dans la projection (VIP). Ce score exprime comment chaque variable prédictive contribue à la construction des vecteurs latents du modèle. Cependant la PLS ne permet pas d'écarter des variables fortement corrélées entre elles.

Des variantes de la PLS peuvent être utilisées pour faire des prédictions quantitatives ou qualitatives. Les méthodes basées sur la PLS ont l'avantage d'être relativement peu exigeantes en termes de puissance de calcul. Alors que l'utilisation de la PLS est très répandue en métabolomique végétale, les modèles linéaires multivariés sont encore rarement envisagés. Pourtant, grâce à l'utilisation de méthodes de contraction des coefficients de la régression, la plus connue étant LASSO (*Least Absolute Shrinkage and Selection Operator*), il est possible d'obtenir des modèles n'utilisant qu'une petite fraction des variables, ce qui facilite grandement l'identification de marqueurs et réduit considérablement la demande en puissance de calcul. Une étude récente menée chez le riz avec des données métabolomiques et génomiques suggère que la prédiction linéaire sans biais (*Best Linear Unbiased Prediction*, BLUP) et la régression LASSO sont les méthodes les plus efficaces parmi huit méthodes de prédiction testées, en fonction du trait à prédire et du type de données utilisées pour la prédiction (Wang *et al.*, 2019).

4 - PREDICTION DE LA PERFORMANCE AU SEIN DE PANELS DE DIVERSITE INTRASPECIFIQUE

Les premières tentatives pour prédire la performance des plantes à partir de données métabolomiques remontent à 2007 (Meyer *et al.*, 2007). Dans une population d'introgession d'*Arabidopsis thaliana*, une corrélation médiane de 0,58 a été obtenue pour la prédiction de la biomasse aérienne en utilisant une partie des observations pour le paramétrage d'un modèle d'analyse canonique des corrélations et l'autre pour tester le modèle. Des résultats similaires ont été observés depuis pour d'autres traits dont le rendement mais aussi la résistance à des stress biotiques ou abiotiques et étendus à un nombre croissant d'espèces d'intérêt agronomique comme le maïs (Riedelsheimer *et al.*, 2012), le riz (Wang *et al.*, 2019), le colza (Knoch *et al.*, 2021) ou la tomate (Luna *et al.*, 2020). Les corrélations obtenues entre données prédites et mesurées étaient toujours significatives, avec des coefficients de corrélation (*R*) dépassant souvent 0,8.

Cependant, dans la plupart de ces études, les données prédictives et les données cibles (les traits de performance) ont été obtenues dans les mêmes expériences. De bonnes prédictions de rendement ou d'autres traits complexes ont certes pu être obtenues avec des échantillons de feuilles prélevés à des stades jeunes mais la portée de ces résultats mérite d'être tempérée. En effet, les données métaboliques varient d'une expérience à l'autre parce-que les interactions entre le génotype et l'environnement influencent fortement le métabolisme dans la variabilité naturelle du climat entre années et entre lieux. La question n'est donc pas de savoir si des données métabolomiques obtenues dans un champ permettent de prévoir le rendement dans le même champ, mais jusqu'à quel point la prévision du rendement à partir de la métabolomique a une valeur générique. Ainsi, une démarche

véritablement prometteuse sera d'obtenir des données métabolomiques à partir de plantes cultivées en conditions facilement reproductibles et de préférence à des stades jeunes. Quelques résultats encourageants ont montré qu'une telle démarche est possible. Ainsi chez le blé, des données métabolomiques de plants cultivés en serre ont été associées à l'écart de rendement entre conditions limitées en eau et conditions bien irriguées pour un ensemble de 18 expériences au champ (Yadav *et al.*, 2019). Des résultats obtenus chez le maïs (Figure 1, résultats non publiés) montrent que des coefficients de corrélation dépassant 0,7 peuvent être atteints avec des modèles linéaires multivariés construits à partir de données métabolomiques foliaires de plantes cultivées sur une plateforme de phénomique et des données de rendement obtenues au champ. Des coefficients de corrélation similaires ont été obtenus avec des approches de prédiction génomique et phénomique pour le même panel de diversité (Millet *et al.*, 2019).

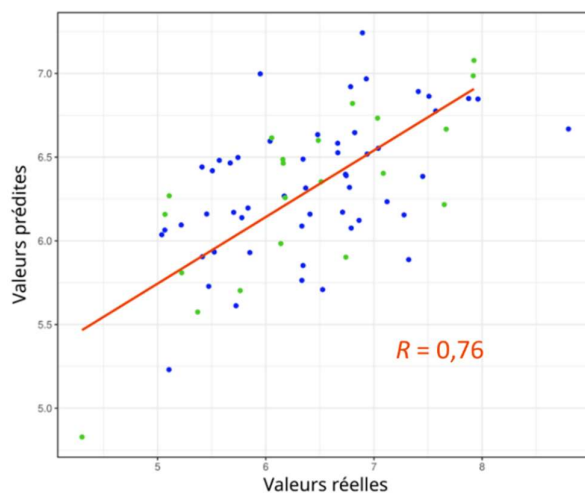


Figure 1. Prédiction du rendement au champ à partir de données métabolomiques foliaires obtenues en serre. Un modèle linéaire multivarié a été construit avec des données métabolomiques obtenues pour des plants cultivés en serre et des données de rendement obtenues au champ pour un panel de 238 hybrides de maïs (données métabolomiques non-publiées et rendements publiés dans Millet *et al.*, 2019). Le modèle a été construit avec 90% des données (bleu) et testé avec les 10% restants (vert). Le coefficient de corrélation a été calculé avec ces dernières.

Une hypothèse souvent invoquée expliquant les bons résultats obtenus avec des données métabolomiques pour prédire des traits de performance est qu'elles intègrent de nombreuses interactions et, qui plus est, à différents niveaux en aval du fonctionnement du génome. Il sera donc important de tester divers scénarios de croissance (conditions optimales, stress, application d'éliciteurs...) dans l'hypothèse où des métabolomes plus prédictifs pourraient être obtenus.

Enfin, il est intéressant de noter que les utilisations de données génomiques, transcriptomiques ou métabolomiques pour prédire la performance de la descendance ont été comparées chez le maïs (Westhues *et al.*, 2017). Pour cela, des mesures ont été effectuées sur des lignées parentales à des stades précoces de développement et leur valeur prédictive évaluée avec des données de performance des hybrides issus de leurs croisements, dont le rendement en matière sèche. Les meilleures prédictions ont été obtenues avec le transcriptome, confirmant sa valeur intégrative. En revanche, les prédictions obtenues avec les données métabolomiques (environ 200 métabolites analysés au total) étaient les moins bonnes quoique significatives. Il sera néanmoins intéressant de réaliser d'autres études de ce type, notamment avec des techniques de spectrométrie de masse plus

performantes en termes de couverture du métabolome. En effet, le coût de l'analyse en spectrométrie de masse haute résolution (moins de 20€ par échantillon) est d'ores et déjà nettement inférieur à celui de la transcriptomique (autour de 100€ par échantillon).

5 - PREDICTION DE LA PERFORMANCE AU SEIN DE PANELS DE DIVERSITE INTERSPECIFIQUE

Nous avons vu précédemment que le règne végétal produirait jusqu'à un million de métabolites. La plupart sont vraisemblablement des métabolites dits « secondaires » ou « spécialisés », souvent considérés comme impliqués dans les interactions avec l'environnement, alors que ceux dits « primaires » sont impliqués dans les fonctions de base du vivant (croissance, développement et reproduction). Il est estimé qu'au moins plusieurs milliers de métabolites sont partagés par les espèces végétales, la quasi-totalité des métabolites primaires et une large proportion des métabolites secondaires, notamment ceux situés en amont des voies de synthèses ou agissant comme régulateurs. Il semble donc pertinent de se demander si les concentrations de ces métabolites « partagés » peuvent être utilisées pour étudier l'adaptation des espèces végétales. En effet, une telle démarche, *a priori* impossible avec des approches basées sur des données de séquence (génomiques, transcriptomiques, protéomiques), a le potentiel de mettre à jour des mécanismes impliqués dans l'adaptation qui pourraient être transférés d'une espèce à l'autre sans qu'une ingénierie complexe de biologie synthétique soit nécessaire. Ainsi, des marqueurs métaboliques associés à un caractère d'intérêt pourraient être utilisés pour sélectionner des polymorphismes impliqués dans la régulation.

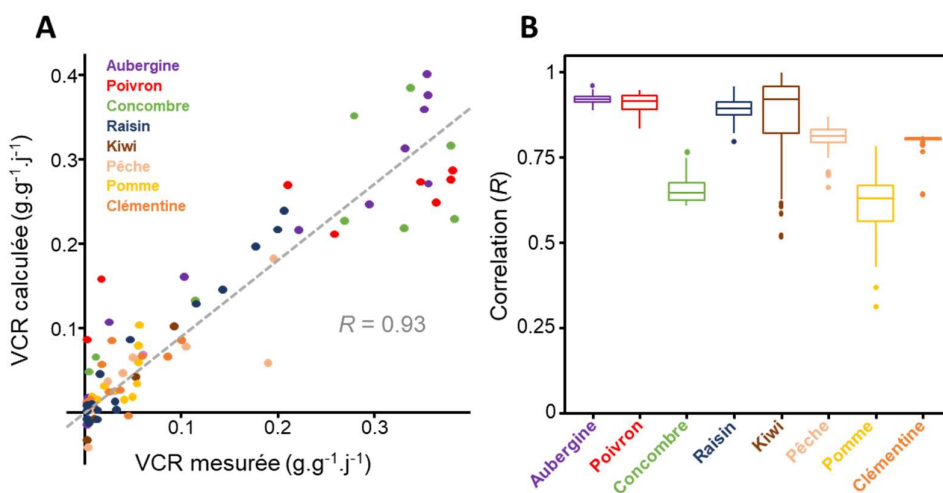


Figure 2. Modélisation linéaire multivariée pour la prédiction de la vitesse de croissance relative (VCR) à partir de la composition de la biomasse de huit espèces de fruits à tous stades de développement. A. Corrélation entre la VCR mesurée et la VCR calculée avec un modèle obtenu pour l'ensemble des fruits. B. Distribution des coefficients de corrélation obtenus pour chaque espèce fruitière en utilisant 100 modèles construits avec les données des sept autres espèces. Adapté de Roch *et al.* (2020).

Une étude menée récemment avec un petit panel de huit espèces de fruits charnus valide l'hypothèse selon laquelle des traits métaboliques peuvent être utilisés pour prédire la performance (Roch *et al.*, 2020). La composition de la biomasse du fruit de trois espèces herbacées (aubergine, poivron, concombre), trois espèces arborescentes (pomme, pêche, clémentine) et deux espèces de vigne (kiwi, raisin) a été analysée tout au long du développement et de la maturation des plantes. Des corrélations significatives ont été trouvées entre la vitesse de croissance relative (VCR) du fruit et

plusieurs constituants de la biomasse, notamment les protéines solubles totales ($R = 0,84$), le stéarate ($R = 0,72$), le palmitate ($R = 0,72$) et le lignocérate ($R=0,68$). Ce lien étroit entre la composition de la biomasse et la VCR a ensuite été confirmé par des modèles linéaires généralisés qui prédisaient la VCR avec des valeurs R parfois supérieures à 0,9 (Figure 2).

Il est étonnant de constater que des prédictions avec des coefficients de corrélation aussi élevés puissent être obtenues avec si peu d'observations (seulement huit espèces et une dizaine de stades de développement du fruit). Cela pourrait être lié au fait que la plage dynamique des données est nettement plus grande avec un panel interspécifique qu'avec un panel intraspécifique. L'exemple suivant, une étude menée avec un panel d'espèces du désert de l'Atacama au Chili semble le confirmer (Dussarrat *et al.*, 2022).

Des échantillons de la partie aérienne de 24 espèces appartenant à 14 familles botaniques et à trois types photosynthétiques (C3, C4 et CAM) ont été prélevés le long d'un gradient altitudinal représentant un dénivelé de plus de 2000 mètres (Figure 3). Des profils métabolomiques avec plusieurs milliers de métabolites ou signatures métaboliques communs à toutes les espèces ont ensuite été obtenus par chromatographie liquide couplée à la spectrométrie de masse haute résolution. Comme indiqué dans la Figure 3 ces espèces sont caractérisées par des aires de répartition plus ou moins grandes en lien avec l'altitude. Considérant que l'altitude intègre les variations d'un certain nombre de facteurs environnementaux comme la température et la disponibilité en eau et en nutriments du sol, un modèle linéaire multivarié a ensuite été construit pour évaluer l'implication du métabolisme dans l'adaptation à ces milieux extrêmes. De très bonnes prédictions de l'altitude à laquelle croissent les plantes ont pu être obtenues à partir du métabolome. Avec seulement trente-neuf composés identifiés ou annotés, un coefficient de corrélation de 0,89 entre altitudes réelles et prédites a été obtenu. Ce modèle a ensuite été validé avec un jeu de données obtenu à partir d'échantillons prélevés une autre année, donnant un coefficient de corrélation de 0,89.

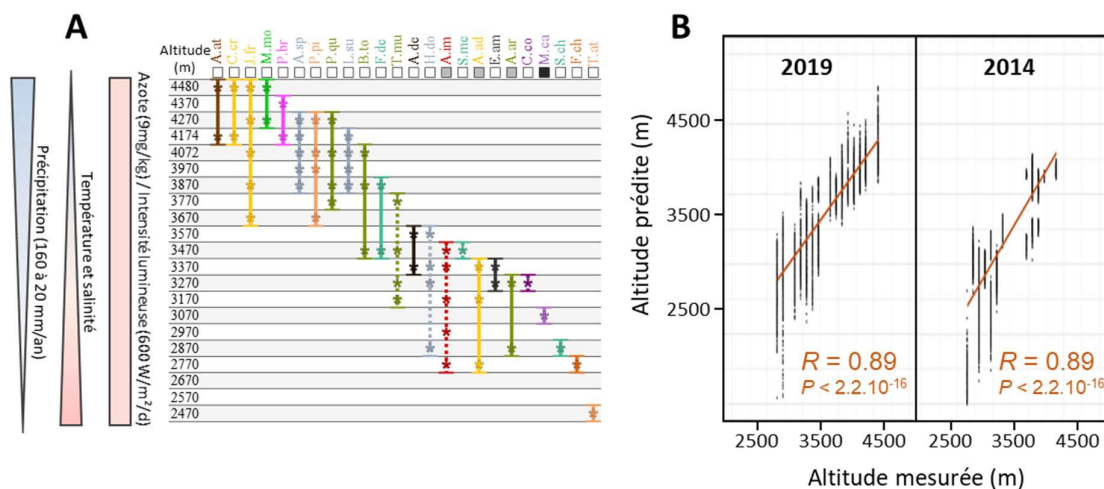


Figure 3. Prédiction de l'altitude à partir du métabolome dans le désert de l'Atacama. A. Description de l'échantillonnage de 24 espèces le long d'un gradient d'altitude dans le désert de l'Atacama. B. Validation du modèle construit à partir de données de 2019 avec des données de 2014. Les carrés blanc, gris et noir indiquent, respectivement, les plantes avec un métabolisme central carboné C3, C4 et CAM. Adapté de Dussarrat *et al.* (2022).

Ces résultats suggèrent que des espèces qui avaient divergé bien avant l'existence du désert de l'Atacama, puisqu'appartenant à des familles botaniques que l'on retrouve sur d'autres continents, ont utilisé les mêmes mécanismes pour s'adapter à des environnements extrêmes. Parmi les 39 meilleurs marqueurs de cette adaptation se trouvent notamment des métabolites déjà connus pour leur implication dans l'adaptation aux milieux extrêmes, comme par exemple le raffinose, la proline ou encore la proline bêtaïne. Les métabolites identifiés, également présents dans des espèces d'intérêt agronomique, constituent autant de pistes pour améliorer les performances de plantes cultivées dans des milieux contraignants.

6 – CONCLUSION ET PERSPECTIVES

Des résultats toujours plus nombreux montrent que la métabolomique peut être utilisée pour prédire des traits d'intérêt agronomique avec une performance proche de celles de la prédiction génomique et de la prédiction phénotypique. Le fait que des données métabolomiques obtenues en conditions contrôlées puissent être utilisées pour prédire la performance au champ semble particulièrement intéressant. Cependant, il existe encore trop peu d'études pour pouvoir juger de la pertinence d'une telle approche pour la sélection. En particulier, il sera important de tester divers scénarios de culture en conditions contrôlées, notamment en imposant des stress, pour vérifier s'il est possible de faire exprimer aux plantes un métabolome encore plus prédictif. L'enjeu est de parvenir à des conditions de culture répétables et peu onéreuses. Parce que la métabolomique prédictive peut également être utilisée avec des panels de diversité interspécifique, des mécanismes génériques impliqués dans l'adaptation des plantes pourront être mis à jour, puis transférés d'une espèce à l'autre. Au-delà de la recherche et du transfert de mécanismes impliqués dans la performance des plantes, des approches complètement nouvelles pourront être envisagées, notamment en lien avec le développement de l'agroécologie.

Ces avancées ont été rendues possibles grâce à des progrès significatifs de l'analyse du métabolome et de la modélisation. La combinaison de la robotique de laboratoire et de la spectrométrie de masse haute résolution permet désormais de réaliser en routine des expériences avec des centaines, voire des milliers d'échantillons, à des coûts raisonnables. Des solutions ont été trouvées pour résoudre les goulots d'étranglement, depuis l'extraction jusqu'au traitement des données. C'est finalement l'échantillonnage qui apparaît comme l'étape la plus contraignante, avec la nécessité de prendre en compte le stade de développement de la plante ou de l'organe, la position sur l'organe si ce dernier n'est pas entièrement échantillonné, l'heure de la journée car le métabolome fluctue et la nécessité d'inactiver les réactions biochimiques (Hall *et al.*, 2022). L'échantillonnage manuel est contraignant et souvent source d'erreur et l'automatisation rarement envisagée. Cependant, des robots capables de prélever des échantillons sur des plateformes de phénotypage pourraient bientôt ouvrir la voie à des observations destructrices. Ces pistes de réflexion sont notamment évaluées par des initiatives nationales comme les infrastructures MetaboHUB 2.0 et PHENOME-EMPHASIS France.

Journée Scientifique ASF du 3 février 2022

« Quoi de neuf sur le phénotypage en amélioration des plantes ? »

Financement : MetaboHUB (ANR-11-INBS-0010), PHENOME (ANR-11-INBS-0012), AMAIZING (ANR-10-BTBR-01).

REFERENCES

- Alseekh S., Fernie A.R. – 2018 - Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant J.* 94: 933-942.
- Allwood J.W., Williams A., Uthe H., van Dam N.M., Mur L.A., Grant M.R., Pétriacq P. – 2021 - Unravelling Plant Responses to Stress - The Importance of Targeted and Untargeted Metabolomics. *Metabolites* 11: 558.
- Carrari F., Baxter C., Usadel B., Urbanczyk-Wochniak E., Zanon M.I., Nunes-Nesi A., Nikiforova V., Centro D., Ratzka A., Pauly M., Sweetlove L.J., Fernie A.R. – 2006 - Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behaviour. *Plant Physiol.* 142: 380-1396. DOI10.1104/pp.106.088534
- Dussarrat T., Prigent S., Latorre C., Bernillon S., Flandin A., Díaz F.P., Cassan C., Van Delft P., Jacob D., Varala K., Joubes J., Gibon Y., Rolin D., Gutiérrez R.A., Pétriacq P. - 2022 - Predictive metabolomics of multiple Atacama plant species unveils a core set of generic metabolites for extreme climate resilience. *New Phytol.* In press.
- Fernandez O., Urrutia M., Bernillon S., Giauffret C., Tardieu F., Le Gouis J., Langlade N., Charcosset A., Moing A., Gibon Y. – 2016 - Fortune telling: metabolic markers of plant performance. *Metabolomics* 12: 158.
- Gibon Y., Usadel B., Blaesing O.E., Kamlage B., Hoehne M., Trethewey R., Stitt M. – 2006 - Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. *Genome Biology* 7: R76.
- Hall R.D., D’Auria J.C., Silva Ferreira A.C., Gibon Y., Kruszka D., Mishra P., van de Zedde R. – 2022 - High-throughput plant phenotyping: a role for metabolomics? *Trends Plant Sci.* In press.
- Knoch D., Werner C.R., Meyer R.C., Riewe D., Abbadi A., Lücke S., Snowdon R.J., Altmann T. – 2021 - Multi-omics-based prediction of hybrid performance in canola. *Theor. Appl. Genet.* 134: 1147-1165.
- Long S.P., Marshall-Colon A., Zhu X.-G. (2015). Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161: 56–66.
- Luna E., Flandin A., Cassan C., Prigent S., Chevanne C., Kadiri C.F., Gibon Y., Petriacq P. – 2020 - Metabolomics to Exploit the Primed Immune System of Tomato Fruit. *Metabolites* 10: 96.
- Meyer R.C., Steinfath M., Lisek J., Becher M., Witucka-Wall H., Torjek O., Fiehn O., Eckardt A., Willmitzer L., Selbig J., Altmann T. – 2007 - The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Nat. Acad. Sci USA* 104: 4759-4764.
- Millet E.J., Kruijer W., Coupel-Ledru A., Prado S.A., Cabrera-Bosquet L., Lacube S., Charcosset A., Welcker C., van Eeuwijk F., Tardieu F. – 2019 - Genomic prediction of maize yield across European environmental conditions. *Nature Gen.* 51: 952-956.
- Riedelsheimer C., Czedik-Eysenberg A., Grieder C., Lisek J., Technow F., Sulpice R., Altmann T., Stitt M., Willmitzer L., Melchinger A.E. – 2021 - Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics* 44: 217-220.
- 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 Genes Genomes Genetics* 8: 3961-3972.
- Shen X., Wang R., Xiong X., Yin Y., Cai Y., Ma Z., Liu N., Zhu Z.-J. – 2019 - Metabolic reaction network-based recursive metabolite annotation for untargeted metabolomics. *Nature Comm.* 10: 1516.
- Tardieu F., Cabrera-Bosquet L., Pridmore T. et Bennett M. – 2017 - Plant phenomics, from sensors to knowledge. *Current Biology* 27: R770–R783.
- Viant M.R., Kurland I.J., Jones M.R., Dunn W.B. – 2017 - How close are we to complete annotation of metabolomes? *Curr. Opinion Chem. Biol.* 36: 64-69.
- Wang S., Wei J., Li R., Qu H., Chater J. M., Ma R., Li Y., Xie W., Jia Z. (2019) Identification of optimal prediction models using multi-omic data for selecting hybrid rice. *Heredity* 123: 395–406.

Westhues M., Schrag T.A., Heuer C., Thaller G., Utz H.F., Schipprack W., Thiemann A., Seifert F., Ehret A., Schlereth A., Stitt M., Nikoloski Z., Willmitzer L. Schoen C.C., Scholten S., Melchinger A.E. – 2017 - Omics-based hybrid prediction in maize. *Theor. Appl. Gen.* 130: 1927-1939.

Yadav, A.K., Carroll A.J., Estavillo G.M., Rebetzke G.J., Pogson B.J. – 2019 - Wheat drought tolerance in the field is predicted by amino acid responses to glasshouse-imposed drought. *J. Exp. Bot.* 70, 4931-4947