

**AFPP – 23<sup>e</sup> CONFÉRENCE DU COLUMA**  
**JOURNÉES INTERNATIONALES SUR LA LUTTE CONTRE LES MAUVAISES HERBES**  
**DIJON – 6, 7 ET 8 DÉCEMBRE 2016**

**IDENTIFICATION DE GENES LIES A LA RESISTANCE AUX HERBICIDES NON LIEE A LA CIBLE CHEZ LES  
IVRAIES: VERS DES TESTS « MOLECULAIRES » DE DIAGNOSTIC**

C. DÉLYE <sup>(1)</sup>, A. DUHOUX <sup>(1)</sup>, S. CARRÈRE <sup>(2)</sup> ET L. BONIN <sup>(3)</sup>

<sup>(1)</sup> INRA, AGROÉCOLOGIE, 17 RUE SULLY, F-21000 DIJON, FRANCE, [delye@dijon.inra.fr](mailto:delye@dijon.inra.fr)

<sup>(2)</sup> INRA, UMR441 LIPM, F-31326 CASTANET-TOLOSAN, FRANCE

<sup>(3)</sup> ARVALIS-INSTITUT DU VÉGÉTAL, F-75116, PARIS, FRANCE

## **RÉSUMÉ**

Le développement de tests moléculaires de détection serait un grand pas en avant dans la prévention et la gestion de la résistance non liée à la cible (RNLC). Les déterminismes génétiques de la RNLC sont cependant complexes, et font largement appel à des différences d'expression d'un nombre inconnu de gènes entre plantes résistantes et plantes sensibles. Leur identification est donc ardue. Dans ce travail, nous décrivons l'identification de 19 gènes dont les profils d'expression sont corrélés à la RNLC à des herbicides inhibiteurs de l'ALS chez les Ivraies (*Lolium* sp.), et nous montrons le potentiel de l'utilisation des données d'expression combinées de plusieurs de ces gènes pour la prédiction de la RNLC.

Mots-clés : ALS, Détection, PCR, Ivraie, Résistance non liée à la cible.

## **ABSTRACT**

### **IDENTIFICATION OF GENES LINKED TO NON-TARGET-SITE RESISTANCE TO HERBICIDES IN RYE-GRASS: TOWARDS "MOLECULAR" DIAGNOSIS ASSAYS**

Setting up molecular diagnosis tools would greatly foster preventing and managing non-target-site-based resistance to herbicides (NTSR). This is currently hampered by the complexity of NTSR genetic determinism, which is largely based upon differences in the expression of an unknown number of genes between resistant and sensitive plants. Identifying NTSR genes is thus a challenging task. Herein, we report the identification of 19 genes which expression patterns are correlated to NTSR to ALS-inhibiting herbicides in rye-grasses (*Lolium* sp.). We demonstrate the potential of using combined expression data from several of these genes to predict NTSR.

Keywords: ALS, Detection, PCR, Rye-grass, Non-target-site-based resistance.

## INTRODUCTION

Les mauvaises herbes sont la première cause biotique de perte de rendement des cultures (Oerke 2006). Leur prolifération doit donc être freinée. En agriculture dite « conventionnelle », ceci est réalisé par l'application d'herbicides. Cependant, les applications d'herbicides répétés sélectionnent les génotypes ayant une sensibilité réduite aux herbicides, et qui peuvent finir par entraîner l'apparition de génotypes résistants à la dose au champ (Délye *et al*, 2013). Trait-clé du succès de la réussite des mauvaises herbes, la résistance aux herbicides est causée par des mutations structurelles entraînant des variations dans les propriétés des protéines et / ou des mutations de régulation modifiant la quantité de protéines (Délye *et al*, 2013). Les mutations affectant les protéines cibles d'herbicides provoquent une résistance dite « de cible », un mécanisme monogénique maintenant bien caractérisé (Beckie et Tardif 2012; Délye *et al*, 2013.). D'autres mutations provoquent une résistance dite « non liée à la cible » (RNLC) qui est en général un caractère quantitatif complexe doté d'une variété de mécanismes provoquant la neutralisation et / ou la compensation de l'action d'un herbicide. La RNLC fait partie des mécanismes de réponse aux herbicides et est supposée être largement due à des différences constitutives ou induites par les herbicides dans la régulation de nombreux gènes entre les plantes résistantes et sensibles (Délye, 2013).

Un diagnostic précoce et précis de la RNLC permettant l'adaptation réactive du programme de lutte contre les mauvaises herbes pour éviter une poursuite de l'évolution de la RNLC dans les populations de mauvaises herbes est essentiel pour un contrôle chimique durable des mauvaises herbes (Délye *et al*, 2013). Les tentatives d'identification de gènes régissant la RNLC afin de développer des tests de diagnostic basés sur les acides nucléiques (ADN, ARN) a longtemps été entravée par la grande diversité des mécanismes impliqués dans la RNLC (Délye 2013). Les nouvelles technologies de séquençage du transcriptome entier ont récemment stimulé la recherche de gènes de RNLC, mais à ce jour, aucun test moléculaire pour le diagnostic de la RNLC n'a émergé, très probablement parce que, comme la RNLC est un caractère complexe, elle ne peut pas être prédite de manière fiable en utilisant seulement un ou quelques gènes. Cela suggère qu'un certain nombre de gènes de RNLC doit être identifié avant que le diagnostic moléculaire ne soit possible. Les gènes pertinents pour la détection de la RNLC sont des gènes dont le niveau d'expression devrait être corrélé à la RNLC et qui devraient également être largement répandus et fréquents dans les populations de mauvaises herbes géographiquement différentes, de sorte qu'un même ensemble de gènes puisse être utilisé pour le diagnostic à grande échelle.

Nous avons éprouvé la véracité de cette hypothèse sur l'ivraie (*Lolium* sp, « ray-grass » adventice.), une des mauvaises herbes mondiales et françaises les plus nuisibles. Les inhibiteurs de l'ALS sont le mode d'action principal utilisé contre les Ivraies en grandes cultures, et les résistances à ce mode d'action sont fréquentes chez cette espèce. Ici, nous avons mis en œuvre une analyse du transcriptome basée sur une ressource transcriptomique décrite précédemment (LOLbase, Duhoux *et al*, 2015) pour identifier des gènes avec un niveau d'expression constitutif corrélé à la RNLC aux inhibiteurs de l'ALS, puis nous avons évalué la pertinence de ces marqueurs pour le diagnostic de la RNLC en utilisant un ensemble divers de populations d'Ivraies du champ.

## MATERIEL ET MÉTHODE

### MATERIEL VEGETAL

Quatorze populations d'ivraie de terrain géographiquement différentes (Figure 1) ont été utilisées pour l'identification et la validation de gènes de RNLC, et 20 populations (incluant 10 des 14 précédentes) pour des essais de prédiction de RNLC sur la base des données d'expression de ces gènes.

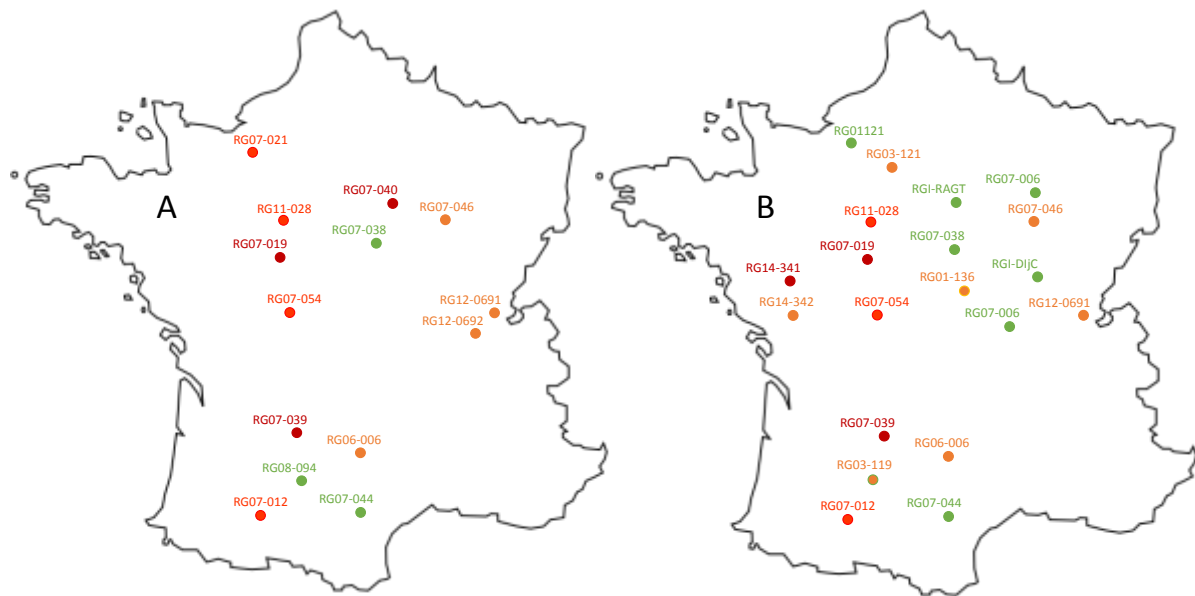


Figure 1. Origines géographiques des populations d'ivraie utilisées pour l'identification et la validation des gènes de RNLG (A) et pour les essais de prédiction de RNLG (B). La couleur des populations indique la fréquence de plantes résistantes à la dose maximale applicable au champ de pyroxsulame et/ou d'iodosulfuron+mésosulfuron déterminée par des tests biologiques de sensibilité (applications en serre sur plantes au stade 3-4 feuilles). Rouge sombre, populations de classe A (fréquence de plantes résistantes > 90% pour les deux herbicides); rouge, populations de classe B (fréquence de plantes résistantes entre 90% et 70% pour les deux herbicides); orange, populations de classe C (fréquence de plantes résistantes entre 70% et 40% pour les deux herbicides); vert, populations de classe D (fréquence de plantes résistantes < 40% pour les deux herbicides).

Figure 1. Geographic localisation of the rye-grass field populations used to identify and validate NTSR genes (A) and for NTSR prediction attempts (B). Dot colour indicates the frequency of plants resistant the maximum allowed field rate of pyroxsulam and/or iodosulfuron+mesosulfuron as determined by herbicide sensitivity assays (herbicide application on 3-4-leaf plants in the greenhouse). Dark red, populations in class A (frequency of resistant plants > 90% for both herbicides); light red, populations in class B (frequency of resistant plants between 90% and 70% for both herbicides); orange, populations in class C (frequency of resistant plants between 70% and 40% for both herbicides); green, populations in class D (frequency of resistant plants < 40% for both herbicides).

#### PRODUCTION DE MATERIEL VEGETAL CARACTERISE

Après vérification de la manière décrite (Délye *et al*, 2009) de l'absence d'allèles mutants ALS conférant une résistance aux herbicides, 299 plantes représentant 14 populations d'ivraie (Figure 1A) ont été soumises à la multiplication végétative comme décrit (Duhoux *et al*, 2015). Ceci a permis de disposer pour chacune des 299 plantes de matériel végétal non traité pour les expérimentations ARN, et de la sensibilité de chaque plante aux deux principales spécialités à base d'inhibiteurs de l'ALS appliquées contre l'ivraie: Abak (pyroxsulame 7,5% p / p; Dow AgroSciences, Saint-Quentin-en-Yvelines, France) et Archipel (iodosulfuron 3% p / p + mésosulfuron 3% p / p, Bayer CropScience, Lyon, France). Chaque herbicide a été appliqué à une ou deux fois la dose maximale recommandée sur deux clones par plante (à savoir 18,75 g pyroxsulame ha<sup>-1</sup> ou 7.5g iodosulfuron + 7.5g mésosulfuron ha<sup>-1</sup>, avec un adjuvant (Actirob B: huile de colza méthylée, Bayer CropScience, 1L ha<sup>-1</sup>). Les plantes ont été affectées à des classes phénotypiques quatre semaines après l'application. Les plantes dont les clones ont été tués aux doses recommandées respectives ont été notés sensibles (S).

Les plantes dont les clones ont été tués à deux fois les doses recommandées respectives ont été affectées à l'une des deux catégories suivantes: résistance moyenne (M) si les clones ont survécu à la dose recommandée d'un herbicide mais ont été tués par l'autre, résistant (R) si les clones ont survécu aux deux doses recommandées respectives. Les plantes dont les clones ont survécu à deux fois les doses recommandées respectives des deux herbicides ont été notés très résistants (RR). Pour chacune des 299 plantes, du matériel végétal non traité a aussi été prélevé pour les extractions d'ARN (la base de la tige et des feuilles de 2 clones).

#### **EXTRACTION DE L'ARN ET SÉQUENÇAGE DE TRANSCRIPTOME (RNA-SEQ)**

L'ARN total a été extrait séparément de chaque clone de chaque plante analysée en utilisant le RNeasy plante kit (Qiagen, Courtaboeuf, France) en suivant les instructions du fabricant. Pour chacune des 14 deux populations, l'ARN de 10 plantes RR a été séquençé en pool et son profil d'expression comparé à l'ARN de 10 plantes S (soient deux comparaisons). Le RNA-Seq a été réalisé comme précédemment (Duhoux *et al*, 2015). En bref, des banques d'ADNc ont été indexées avec des identificateurs uniques 6-base et séquencées dans pistes indépendantes de cellules Illumina pour générer des séquences de 2 × 100 bases appariées. Les 40 millions de séquences obtenues pour chacun des 4 pools ont été appariées contre la base de données de transcriptome LOLbase comme précédemment (Duhoux *et al*, 2015.).

#### **IDENTIFICATION ET VALIDATION DE GENES DE RNLC**

Les gènes-candidats de RNLC ont été identifiés sur la base d'une différence d'expression > 2 fois entre les pools RR et les pools S pour les deux comparaisons, une longueur minimale de 400 nucléotides et une valeur d'expression prédite (RPKM)  $\geq 0.5$  dans au moins un pool en utilisant le package R DESeq. Le niveau d'expression de chaque gène-candidat ainsi identifié a été mesuré individuellement par RT-PCR quantitative selon la procédure décrite précédemment (Duhoux et Délye, 2013) dans chacune des 299 plantes pour valider son lien avec la RNLC. Pour vérifier si les données d'expression de gènes de RNLC combinés pourraient discriminer les plantes résistantes par RNLC de celles sensibles, une analyse en composantes principales (ACP) a été effectuée en utilisant les valeurs logarithmiques des données d'expression obtenues pour les 299 plantes.

#### **PREDICTION DE FREQUENCES DE PLANTES RESISTANTES DANS LES POPULATIONS D'IVRAIES EN UTILISANT L'EXPRESSION DE GENES DE RNLC**

Nous avons testé la faisabilité de prédire directement les fréquences de plantes résistantes dans les populations d'ivraie en utilisant le niveau moyen d'expression des gènes de RNLC mesuré dans des pools de plantes non-clonées et au stade auquel le traitement est recommandé au champ (3-4 feuilles). Les analyses ont été effectuées sur 20 populations de terrain d'ivraies, dont 10 populations utilisées pour la validation des gènes de RNLC et 10 populations supplémentaires (Figure 1B). La sensibilité individuelle des plantes à 3-4 feuilles n'a pas été déterminée. Par contre, la fréquence de plantes résistantes au pyroxsulame et / ou à l'iodosulfuron + mésosulfuron avait été préalablement déterminée dans chacune des 20 populations (Figure 1B). Dans chaque population, quatre pools de cinq plantes ont été utilisés pour la mesure de l'expression de chaque gène, ce qui donne un total de 80 échantillons représentant 400 plantes. Les plantes ont été collectées au hasard au sein de chaque population. L'ARN a été extrait comme précédemment à partir d'une section de 1 cm de base de chaque plante. La RT-qPCR était comme précédemment.

Les corrélations entre le niveau moyen d'expression de gènes de RNLC calculé sur les quatre pools par population et les fréquences des plantes résistantes au pyroxsulame et / ou à l'iodosulfuron + mésosulfuron ont été testées pour chaque gène avec le coefficient de corrélation de Pearson. Des modèles linéaires généralisés quasi-binomiaux destinés à prédire les fréquences de plantes résistantes ont été construites en utilisant la fonction R glm. Un modèle préliminaire utilisant

l'expression moyenne de l'ensemble des 19 gènes de RNLC identifiés a été utilisé pour identifier les gènes qui contribuent de manière significative à la prédiction des fréquences de plantes résistantes. Un modèle final n'utilisant que les données d'expression de ces gènes a ensuite été utilisé pour la prédiction.

## **RESULTATS**

### **IDENTIFICATION DE GENES DE RNLC**

Les 299 plantes analysées incluent 210 plantes résistantes (114 RR, 64 R et 32 M) et 89 plantes sensibles (S). Vingt-sept gènes-candidats, tous surexprimés dans les pools de plantes résistantes, ont été identifiés. Les niveaux d'expression mesurés dans les 299 plantes individuelles représentant 14 populations montraient des variations importantes entre plantes et entre populations pour chaque candidat. La variation était plus élevée entre les plantes qu'entre les populations. Dans l'ensemble, un niveau d'expression significativement plus élevé dans les plantes résistantes a été observé pour 19 gènes (Illustré Figure 2). Ces candidats ont été considérés comme des gènes de RNLC pertinents.

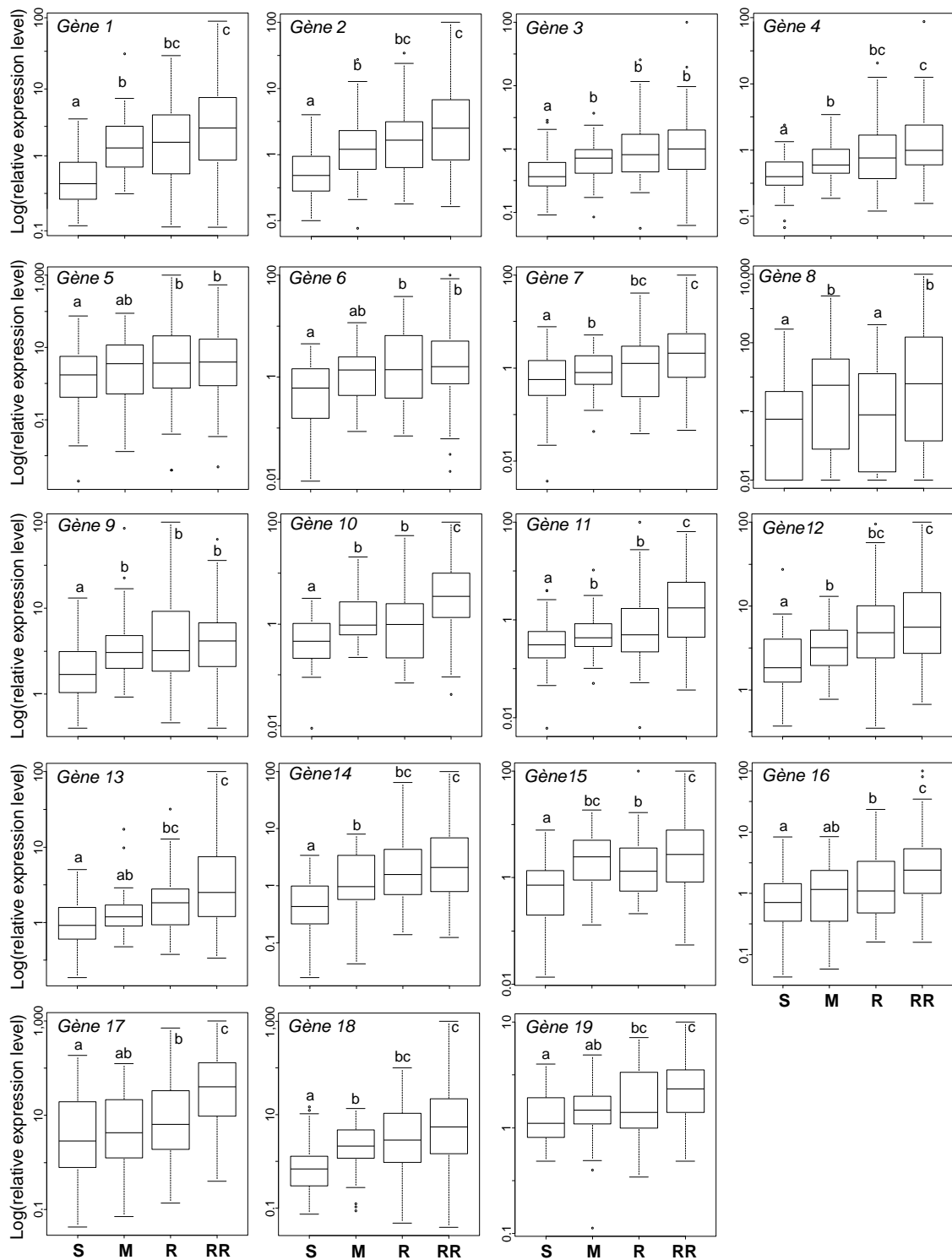


Figure 2. Niveau moyen d'expression des 19 gènes de RNLC retenus dans les plantes notées S, M, R et RR. Des lettres différentes au dessus d'une boîte à moustaches indiquent des niveaux d'expression significativement différents (test de Wilcoxon bilatéral non paramétrique ,  $p < 0,05$ ).

Figure 2. Average Relative expression level of the 19 NTSR genes retained in the plants rated S, M, R and RR. Different letters above a box-plot indicate expression levels significantly different ( $p$  value  $< 0.05$  according to a two-sided Wilcoxon's non-parametric test).

L'Analyse en Composantes Principales (ACP) mis en œuvre sur les données d'expression des 19 gènes pour les 299 plantes a montré une séparation claire de beaucoup, mais pas toutes, les plantes résistantes par rapport aux plantes sensibles (Figure 3).

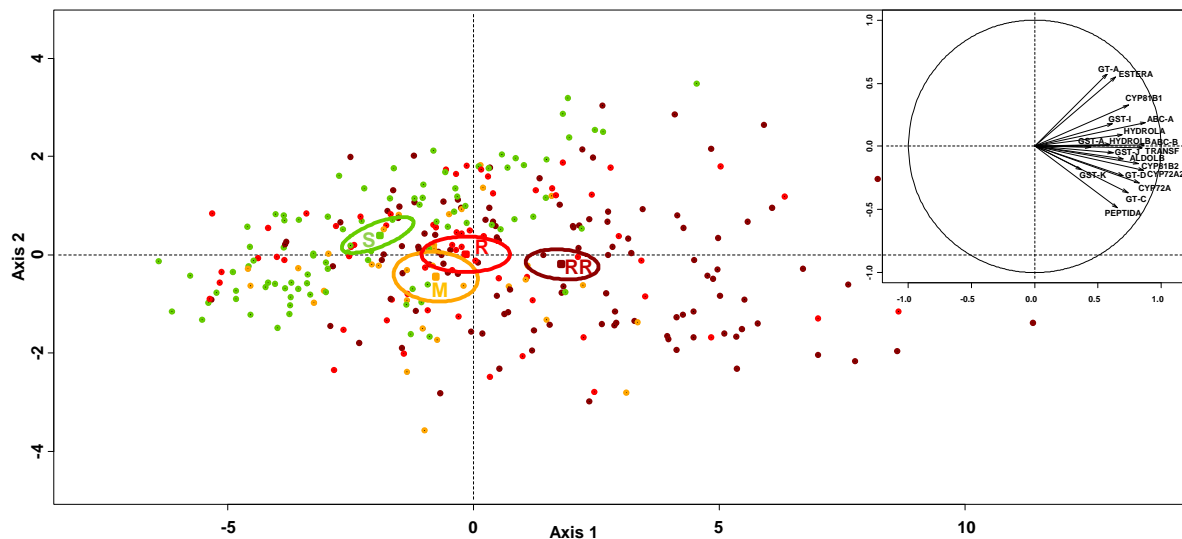


Figure 3. ACP des niveaux d'expression relatifs des 19 gènes dans 299 plantes d'ivraie assignées à quatre classes de phénotypes (S, M, R, RR). Chaque point représente une plante. Les carrés représentent les barycentres de chaque classe et sont inclus dans les ellipses de confiance à 95%. Rouge sombre, RR; rouge, R; orange, M; vert, S. Les axes 1 and 2 portent 48.84% and 7.50% de la variation, respectivement.

Figure 3. PCA of the relative expression levels of the 19 genes in 299 rye-grass plants assigned to four phenotype classes (S, M, R, RR). Dots represent individual plants. Squares represent the barycentre of each phenotype class and are included in the 95% confidence ellipses. Dark red, RR; red, R; orange, M; green, S. Axes 1 and 2 explained 48.84% and 7.50% of the variation, respectively.

#### TENTATIVE DE PREDICTION DES FREQUENCES DE PLANTES RESISTANTES DANS DES POPULATIONS D'IVRAIE A PARTIR DES DONNEES D'EXPRESSION DES 19 GENES MESUREES DANS DES PLANTES PRISES AU HASARD DANS CHAQUE POPULATION

Les 20 populations du champ analysées contenaient des fréquences observées de plantes résistantes allant de 9% à 100% pour l'iodosulfuron + mésosulfuron et de 0% à 100% pour pyroxsulame (Figure 1B). Pour vérifier si les données d'expression de gènes de RNLC pourraient être utilisées pour prédire la fréquence des plantes résistantes dans une population, l'expression des 19 gènes a été mesurée dans quatre pools de cinq plantes pour chacune des 20 populations (c.-à-d 400 plantes dédiées), dont 10 populations non utilisées pour l'identification ou la validation des gènes (Figure 1).

Un premier modèle quasi-binomial a identifié 5 des 19 gènes comme variables explicatives significatives pour la fréquence des plantes résistantes à l'iodosulfuron + mésosulfuron ou au pyroxsulame. L'information apportée par les 14 autres gènes était redondante avec celle fournie par ces 5 gènes. Les fréquences de plantes résistantes prédites pour une population donnée à partir de chaque pool individuel de cinq plantes en utilisant ces 5 gènes étaient peu pertinentes et très variables (effet échantillonnage). En revanche, utiliser les données d'expression moyenne de l'ensemble des 20 plantes de chaque population améliorerait grandement la prédiction (Figure 4,  $R > 0,89$  pour les deux herbicides).

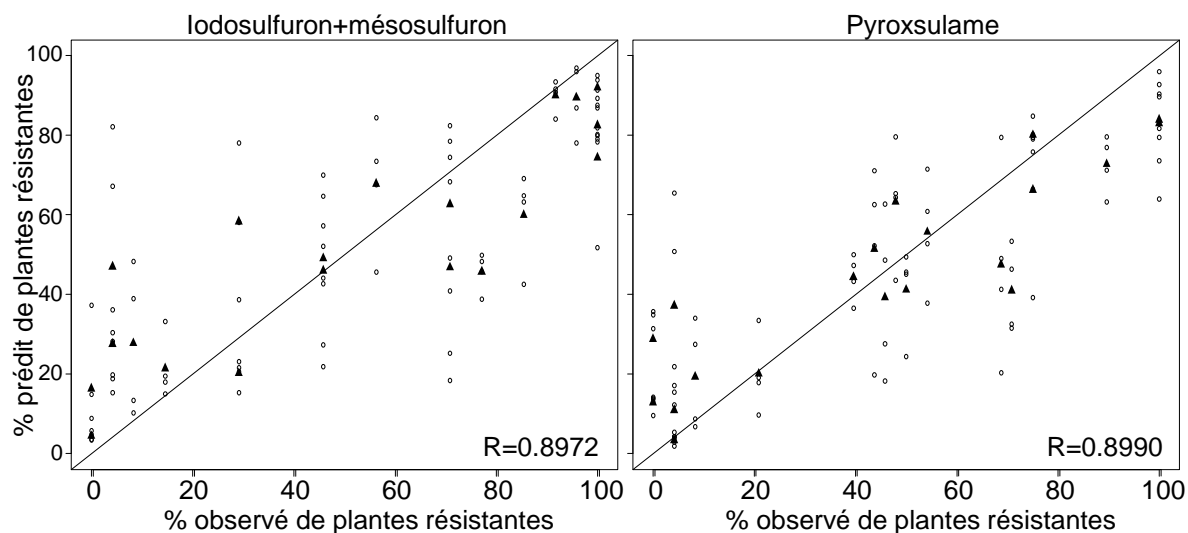


Figure 4. Prédiction des fréquences de plantes résistantes dans 20 populations d'ivraie à partir de données d'expression de gènes corrélés à la RNLC. Les fréquences prédites sont représentées en fonction des fréquences observées. Les points représentent les fréquences prédites à partir des données des pools de 5 plantes (4 pools par population). Les triangles représentent les fréquences prédites à partir des données obtenues sur les 4x20 plantes par population. R, corrélation entre les données prédites et observées.

Figure 4. Prediction of resistant plant frequencies in rye-grass populations using NTSR marker expression data plotted as a function of the observed frequencies in 20 rye-grass populations. Dots represent predicted frequencies using data from individual five-plant pools (4 pools per population), triangles represent predicted frequencies averaged over the 4 pools per population. R, correlation between observed and predicted data.

## DISCUSSION

Ici, à partir de deux populations d'ivraie, nous avons réussi à identifier 19 gènes dont l'expression est significativement régulée à la hausse dans les plantes résistantes par RNLC par rapport aux plantes sensibles (Figure 2). Les gènes identifiés dans ces deux populations ont permis une séparation correcte des plantes résistantes et des plantes sensibles de 14 populations géographiquement diverses sur la base de leurs données d'expression (Figure 3), mais cette séparation n'est pas optimale. La précision de la prédiction du phénotype semble sensiblement augmenter avec le niveau de résistance des plantes (Figure 2, les plantes RR sont mieux prédites que les R et surtout les M).

L'utilisation d'un modèle de base prédit également assez correctement la fréquence des plantes résistantes dans 20 populations d'ivraie en utilisant les données d'expression de seulement 20 plantes prises au hasard dans chaque population (Figure 4). Ceci démontre la faisabilité d'une prédiction de la RNLC en utilisant les données d'expression de gènes corrélés à la RNLC.

Les 19 gènes identifiés ne permettaient pas de distinguer toutes les plantes NTSR de toutes les plantes sensibles. Nos résultats sont en accord avec les données antérieures montrant que les mécanismes de RNLC peuvent varier entre les populations d'ivraie et parmi les plantes au sein des populations (Duhoux *et al*, 2015). De toute évidence, des gènes liés à la RNLC supplémentaires restent à identifier pour augmenter la précision de la prédiction. Ceci peut être réalisé par la

recherche de nouveaux gènes dans les plantes résistantes qui ne pouvaient pas être distinguées des plantes sensibles à l'aide de 19 marqueurs. Comme seulement quelques gènes semblent nécessaires pour la prédiction de la résistance, et que l'information apportée par nombre de gènes semble redondante avec celle apportée par d'autres gènes, les modèles de prévision devront être reconstruits lorsque de nouveaux gènes seront identifiés de manière à ne retenir que les gènes les plus informatifs. Il est très probable que des gènes qui fournissent une information redondante soient co-régulés, et donc appartiennent à la même voie métabolique. Ceci confirme le caractère complexe de la RNLC.

Nos travaux ont également permis de vérifier que les 19 gènes identifiés dans deux populations étaient liés à la RNLC dans une vingtaine de populations différentes (Figure 1). Les gènes identifiés à partir de plantes clonées étaient également pertinents pour étudier des plantes à 3-4 feuilles. Nous avons constaté qu'un modèle simple utilisant des données d'expression de cinq des 19 gènes a permis une prédiction correcte de la fréquence des plantes résistantes dans des populations d'Ivraie (Figure 4). Le nombre optimal de plantes à analyser par population pour améliorer la précision de prédiction reste à déterminer.

## **CONCLUSIONS**

Notre travail a démontré la faisabilité de la prévision de la RNLC basée sur une approche de transcriptomique. Le séquençage quantitatif du transcriptome permet d'identifier des gènes de RNLC potentiels qui peuvent ensuite être validés comme gènes d'intérêt pour la prédiction de RNLC simplement en corrélant leur niveau d'expression au phénotype d'un grand nombre de plantes caractérisées. Cette approche devrait permettre une identification rapide de nombreux gènes supplémentaires et la sélection des plus informatifs.

Nos résultats suggèrent également que les voies de RNLC auxquelles appartiennent les gènes identifiés ont émergé par évolution redondante, puisque les mêmes gènes sont liés à la RNLC dans des populations d'Ivraie géographiquement diverses. Ceci suggère qu'au moins certains des mécanismes qui confèrent la RNLC aux inhibiteurs de l'ALS chez l'Ivraie sont répandus chez cette espèce. Dans cette hypothèse, le diagnostic de la RNLC à grande échelle à partir de gènes liés à la RNLC semble réalisable. Comme l'analyse de l'ADN est plus facile et moins sujette à variation induite par l'échantillonnage que l'analyse de l'ARN telle que réalisée ici, la recherche de mutations de l'ADN liées aux différences d'expression observée est aussi une piste à suivre pour faciliter le diagnostic.

## **REMERCIEMENTS**

Ce travail a été financé par Arvalis-Institut du Végétal

## **BIBLIOGRAPHIE**

Beckie H.J., Tardif F.J., 2012 - Herbicide cross resistance in weeds. *Crop Protection* 35, 15-28.

Délye C., 2013 - Unravelling the genetic bases of non-target-site-based resistance (NTSR) to herbicides: a major challenge for weed science in the forthcoming decade. *Pest Management Science* 69, 176-187.

Délye C., Jasieniuk M., Le Corre V., 2013 - Deciphering the evolution of herbicide resistance in weeds. *Trends in Genetics* 29, 649-658.

Délye C., Boucansaud K., Pernin F., Le Corre V., 2009 - Variation in the gene encoding acetolactate-synthase in *Lolium* species and proactive detection of mutant, herbicide-resistant alleles. *Weed Research* 49, 326–336.

Duhoux A., Délye C., 2013 - Reference genes to study herbicide stress response in *Lolium* sp.: Up-regulation of P450 genes in plants resistant to acetolactate-synthase inhibitors. *Plos One* 8, e63576.

Duhoux A., Carrère S., Gouzy J., Bonin L., Délye C., 2015 - RNA-Seq analysis of rye-grass transcriptomic response to an herbicide inhibiting acetolactate-synthase identified transcripts linked to non-target-site-based resistance. *Plant Molecular Biology* 87, 473-487.

Oerke E.C., 2006 - Crop losses to pests. *Journal of Agriculture Science*, 144, 31-43.